



OPEN ACCESS

EDITED BY

Mingyue Tan,
Shanghai University of Traditional
Chinese Medicine, China

REVIEWED BY

Wei Shenyu,
Zhejiang Chinese Medical University,
China
Jiahang Song,
The First Affiliated Hospital of Nanjing
Medical University, China

*CORRESPONDENCE

Haitao Ma
soochowmht@163.com

†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Molecular and Cellular Oncology,
a section of the journal
Frontiers in Oncology

RECEIVED 30 August 2022

ACCEPTED 28 September 2022

PUBLISHED 17 October 2022

CITATION

Shi Q, Han S, Liu X, Wang S and Ma H
(2022) Integrated single-cell and
transcriptome sequencing analyses
determines a chromatin regulator-
based signature for evaluating
prognosis in lung adenocarcinoma.
Front. Oncol. 12:1031728.
doi: 10.3389/fonc.2022.1031728

COPYRIGHT

© 2022 Shi, Han, Liu, Wang and Ma.
This is an open-access article
distributed under the terms of the
Creative Commons Attribution License
(CC BY). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Integrated single-cell and transcriptome sequencing analyses determines a chromatin regulator-based signature for evaluating prognosis in lung adenocarcinoma

Qingtong Shi^{1,2†}, Song Han^{3†}, Xiong Liu^{2,4}, Saijian Wang^{2,4}
and Haitao Ma^{1*}

¹Department of Thoracic Surgery, The First Affiliated Hospital of Soochow University, Suzhou, China, ²Department of Thoracic Surgery, The Affiliated Hospital of Yangzhou University, Yangzhou, China, ³Department of Thoracic Surgery, Suzhou Science and Technology Town Hospital, Suzhou, China, ⁴Graduate School of Dalian Medical University, Dalian, China

Background: Accumulating evidence has highlighted the significance of chromatin regulator (CR) in pathogenesis and progression of cancer. However, the prognostic role of CRs in LUAD remains obscure. We aim to detect the prognostic value of CRs in LUAD and create favorable signature for assessing prognosis and clinical value of LUAD patients.

Methods: The mRNA sequencing data and clinical information were obtained from TCGA and GEO databases. Gene consensus clustering analysis was utilized to determine the molecular subtype of LUAD. Cox regression methods were employed to set up the CRs-based signature (CRBS) for evaluating survival rate in LUAD. Biological function and signaling pathways were identified by KEGG and GSEA analyses. In addition, we calculated the infiltration level of immunocyte by CIBERSORT algorithm. The expressions of model hub genes were detected in LUAD cell lines by real-time polymerase chain reaction (PCR).

Results: KEGG analysis suggested the CRs were mainly involved in histone modification, nuclear division and DNA modification. Consensus clustering analysis identified a novel CRs-associated subtype which divided the combined LUAD cohort into two clusters (C1 = 217 and C2 = 296). We noticed that a remarkable discrepancy in survival rate among two clusters. Then, a total of 120 differentially expressed CRs were enrolled into stepwise Cox analyses. Four hub CRs (CBX7, HMGA2, NPAS2 and PRC1) were selected to create a risk signature which could accurately forecast patient outcomes and differentiate patient risk. GSEA unearthed that mTORC1 pathway, PI3K/Akt/mTOR and p53 pathway were greatly enriched in CRBS-high cohort. Moreover, the infiltration percentages of macrophage M0, macrophage M2, resting NK cells, memory B cells, dendritic cells and mast cells were statistically significantly different in

the two groups. PCR assay confirmed the differential expression of four model biomarkers.

Conclusions: Altogether, our project developed a robust risk signature based on CRs and offered novel insights into individualized treatment for LUAD cases.

KEYWORDS

lung adenocarcinoma, chromatin regulator, prognosis, risk signature, immune microenvironment

Introduction

Lung cancer (LC) is the major cause of death for men and women with tumor, representing approximately 18% of all cancer deaths worldwide (1). Up to 90% of LC cases are non-small cell lung cancer (NSCLC), including both lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) histological subtypes, with LUAD occurring most frequently (2). Despite recent advances in clinical treatment, the prognosis for LUAD remains dismal, with a 5-year survival rate of only 19%. With the advent of aging and air pollution in developing countries, the incidence of LUAD remains high and early diagnosis of LUAD becomes essential (3). Unfortunately, we still have limited availability of accurate biomarkers for early diagnosis and individualized treatment of LUAD.

Tumor microenvironment (TME) is the internal environment for tumor cell production and survival, and its cellular components include resident stromal cells and recruited immunocytes in addition to tumor cells (4). TME plays an important role in the tumor growth, metastasis, angiogenesis and treatment resistance and has also a crucial impact on prognosis (5). Therefore, systematic exploration of TME is helpful to clarify the mechanism of tumor occurrence and individualized treatment.

Epigenetic modification is a reversible and heritable process of gene expression in the absence of DNA sequence changes. It is one of the critical regulatory mechanisms at the post-transcriptional level of genes by chromatin regulators (CRs), mainly including DNA methylation, histone modifications, chromatin remodeling and RNA regulation (6). CRs mediate epigenetic modification regulates the activation of heterozygous promoters or the activity of repressors and trigger changes in gene transcription levels, resulting in cell differentiation, abnormal proliferation and tumorigenesis (7).

Numerous studies have demonstrated that CRs are tightly bound up to the patient outcomes of LC HMGA1, a chromatin remodeler, has been reported to be involved in DNA transcription, replication and repair. Saed and his colleagues have observed that HMGA1 presented higher expression in lung cancer specimens and overexpressed HMGA1 lead to dismal prognosis of LUAD (8).

Moreover, HMGA1 was proved to facilitate LUAD cell proliferation and migration through GRP75-induced JNK pathway (9). EZH2, belonging to the polycomb-group (PcG) family, has been reported to be greatly overexpressed in lung specimens, and upregulation of EZH2 predicts dismal survival of NSCLC (10). Geng and his colleagues indicated EZH2 enhances the growth and metastasis of lung cell by Akt pathway (10). RAD51 is well known for its important role in homologous recombination. RAD51 has shown to be upregulated in KRAS mutant lung cancer and could regulate cell survival by enhancing DNA damage repair (11). However, the expression patterns and prognostic value of CRs in LUAD remain largely unknown.

In this academic research, we determined CRs with powerful prognostic values in LUAD and created a risk signature for clinical outcome assessment and immune status prediction of LUAD cases.

Methods

Data collection and processing

We obtained the RNA sequencing (RNA-seq) data of 535 LUAD patients and 59 normal controls and their corresponding clinical features from TCGA database (<https://portal.gdc.cancer.gov/>) to construct the prognostic signature. The transcription profiling data was downloaded from GEO dataset and was utilized as the validate set.

Determination of differentially expressed CRs

A total of 870 CRs were retrieved from previous research (6). The gene information of all CRs summarized in [Supplementary Table S1](#). The differentially expressed genes (DEGs) between normal and LUAD tumor tissues were determined using the limma R package with a criteria P value < 0.05 (12). The generated DEGs and CRs gene sets were subsequently intersected to obtain differentially expressed CRs (DECs).

Function and pathways enrichment analyses

GO and KEGG enrichment analysis was conducted to obtain the insight into the biological functions and potential pathways of DECRs. Terms with $p < 0.05$ were listed and visualized using the “clusterProfiler” R packages (13).

Integration of protein–protein interaction (PPI) network

A protein–protein interaction network (PPI) was developed and visualized using the STRING online database (<https://cn.string-db.org/>) and the Cytoscape (<https://cytoscape.org/>), respectively (14, 15). Further, the cubHubba plugin in Cytoscape software was used to filter hub genes of the PPI.

Gene consensus cluster analysis

The consensus cluster analysis was conducted using the “ConsensusClusterPlus” R package, based on the combined LUAD cohort (16). To identify the optimal cluster value, we calculated the Delta area and the cumulative distribution function (CDF). Survival analysis was carried out to compare clinical prognoses between different subtypes using “survival” R package.

Construction of the risk signature

Subsequently, Cox regression analyses were performed to obtain candidate CR with remarkable prognostic value. The formula was set up: Risk score = $\sum_{i=1}^n (\text{coef} \times \text{Exp}_i)$. “Coef” was defined as the corresponding regression coefficient value, and “Exp” was the expression level of genes in the prognostic model. All patients were divided into low- and high-risk groups according to the median score.

Gene set enrichment analysis (GSEA)

We performed GSEA analysis, including GO and KEGG analysis based on CR related DEGs to identify the potential biological and functional differences of different hierarchical clustering (17). A function term with an adjusted p -value < 0.05 and a false discovery rate (FDR) < 0.25 was considered enriched.

Estimate of immune infiltrating status

CIBERSORT tool (<https://cibersortx.stanford.edu/index.php>) was employed to quantify the infiltration status of 22

types of immunocyte fractions in the two LUAD subgroups. $P < 0.05$ was defined as statistically significant.

Single-cell sequencing analysis

We utilized the Seurat clustering to analyze the single-cell data acquired from the GEO databases. The UMAP dimensional reduction and the t-Distributed Stochastic Neighbor Embedding (tSNE) method, were employed to visualize the gene expression and distribution in dataset GSE131907. Next, the cells were re-clustered with the “SingleR” packages to demonstrate the feature genes of different cell types.

Validation of the model CRs

To detect the expression pattern of a model gene at the mRNA level, GEPIA2 tool was applied. Human Protein Atlas (HPA, <https://www.proteinatlas.org/>) database was utilized to confirm the protein level of our model genes between LUAD and normal control (18).

Somatic mutation and stem cell characteristics analyses

The somatic mutation data were obtained from TCGA Portal and processed to compare the tumor mutation burden (TMB) in two groups. The mRNasi is a quantitative index reflecting cancer cells calculated based on gene profiles; The mRNasi and TMB differences in two subgroups were compared using the independent-samples t-test.

Cell culture

Two human LUAD cell lines (A549 and NCI-H460) and a normal human lung epithelial cell line (BEAS-2B) were purchased from American Type Culture Collection. All cell lines were cultured in RPMI 1640 medium (Sigma) containing 10% fetal bovine serum (Gibco) and 1% antibiotics (100 U/ml penicillin G and 100mg/ml streptomycin) at 37°C in a humidified chamber containing 5% CO₂.

RNA extraction and quantitative real-Time PCR

Total cell RNA was extracted by RNA isolation reagent (Takara), then reversed into cDNA by PrimeScript Mix reagent (Takara). SYBR Green Premix (Vazyme biotech) was utilized for PCR reaction system. The value of individual genes

was standardized to GAPDH expression level. [Supplementary Table S2](#) displays primer sequences of all genes.

Statistical analysis

All statistical data in the present project was analyzed by R version 4.0.5 and GraphPad Prism 9. The Kaplan-Meier (KM) analysis was employed to assess the prognostic value of the signature. Moreover, we plotted the receiver operating characteristic (ROC) curve over time to evaluate the prognostic efficacy of the signature.

Results

Characterization of chromatin regulators in LUAD

We first collected 4846 DEGs between LUAD samples and normal cases. A total of 120 DECRs were obtained by taking the intersection of CRs and DEGs gene lists ([Figure 1A](#)). Then, GO analysis was employed to detect the underlying function of

DECRs. The result disclosed that these genes were mainly enriched in histone modification and DNA modification ([Figure 1B](#)). Next, we generated a PPI network to explore the protein interaction among 120 DECRs ([Figure 1C](#)). Based on the MCC algorithm, the top ten hub genes were selected to set up a hub network, including CHEK1, CDK1, TOP2A, CDC6, UHRF1, AURKB, PBK, BUB1, TTK and RAD54L ([Figure 1D](#)).

Chromatin regulators-based consensus cluster analysis

TCGA-LUAD and GSE14520 were combined into one LUAD cohort (n = 609). We applied consensus cluster analysis to develop a CR-related molecular subtype of LUAD.

The result suggested the entire dataset could be well divided into two subtypes based on the 120 DECRs when k = 2 by increasing the clustering variable (k) from 2 to 9 ([Figures 2A–C](#)). PCA analysis shows that DECRs can clearly distinguish two subgroups for clustering analysis ([Figure 2D](#)). There were remarkable discrepancies in survival rates among the two clusters ([Figure 2E](#)). To evaluate the TME status of two clusters, ESTIMATE algorithm was conducted. As suggested in

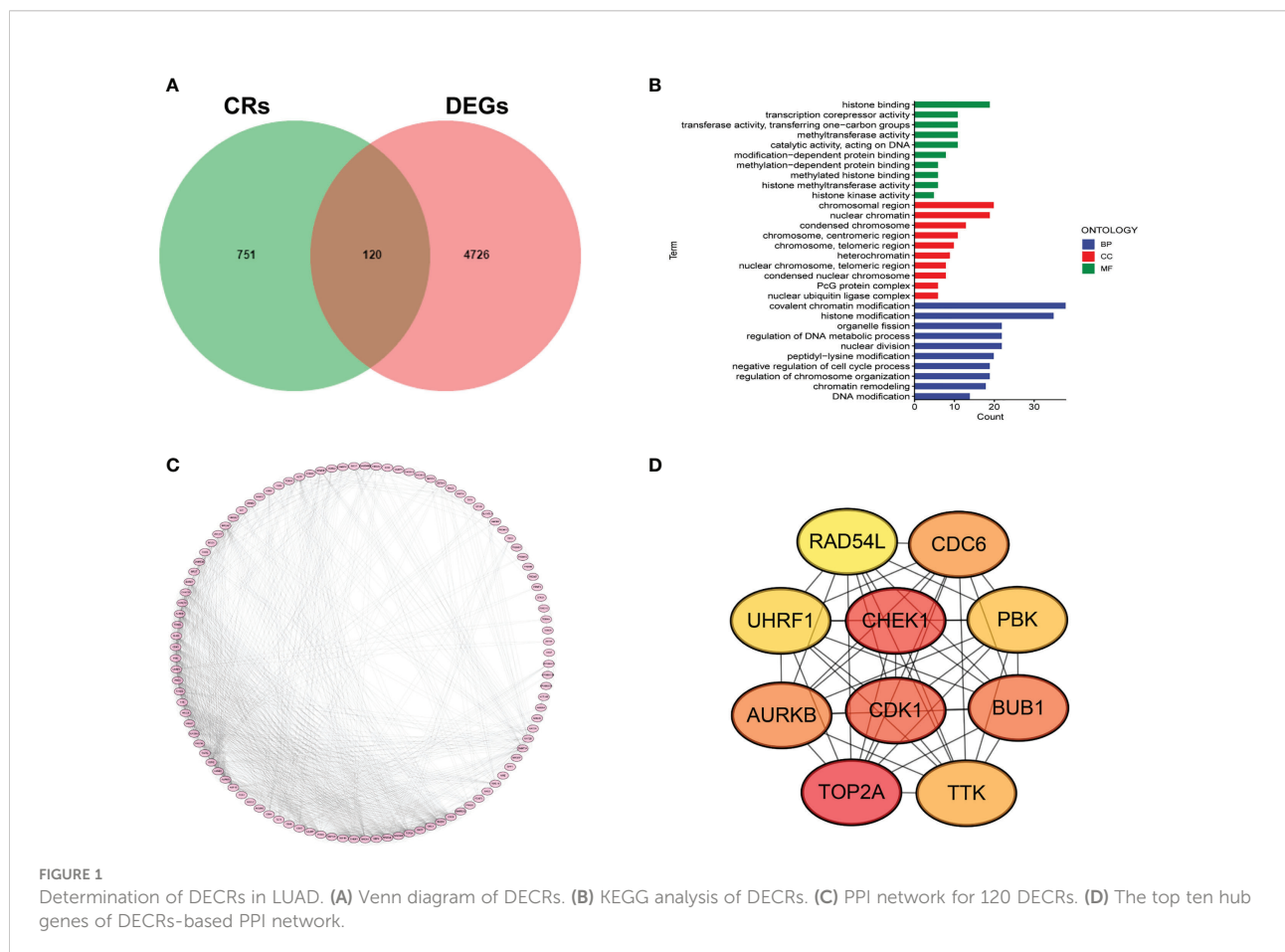
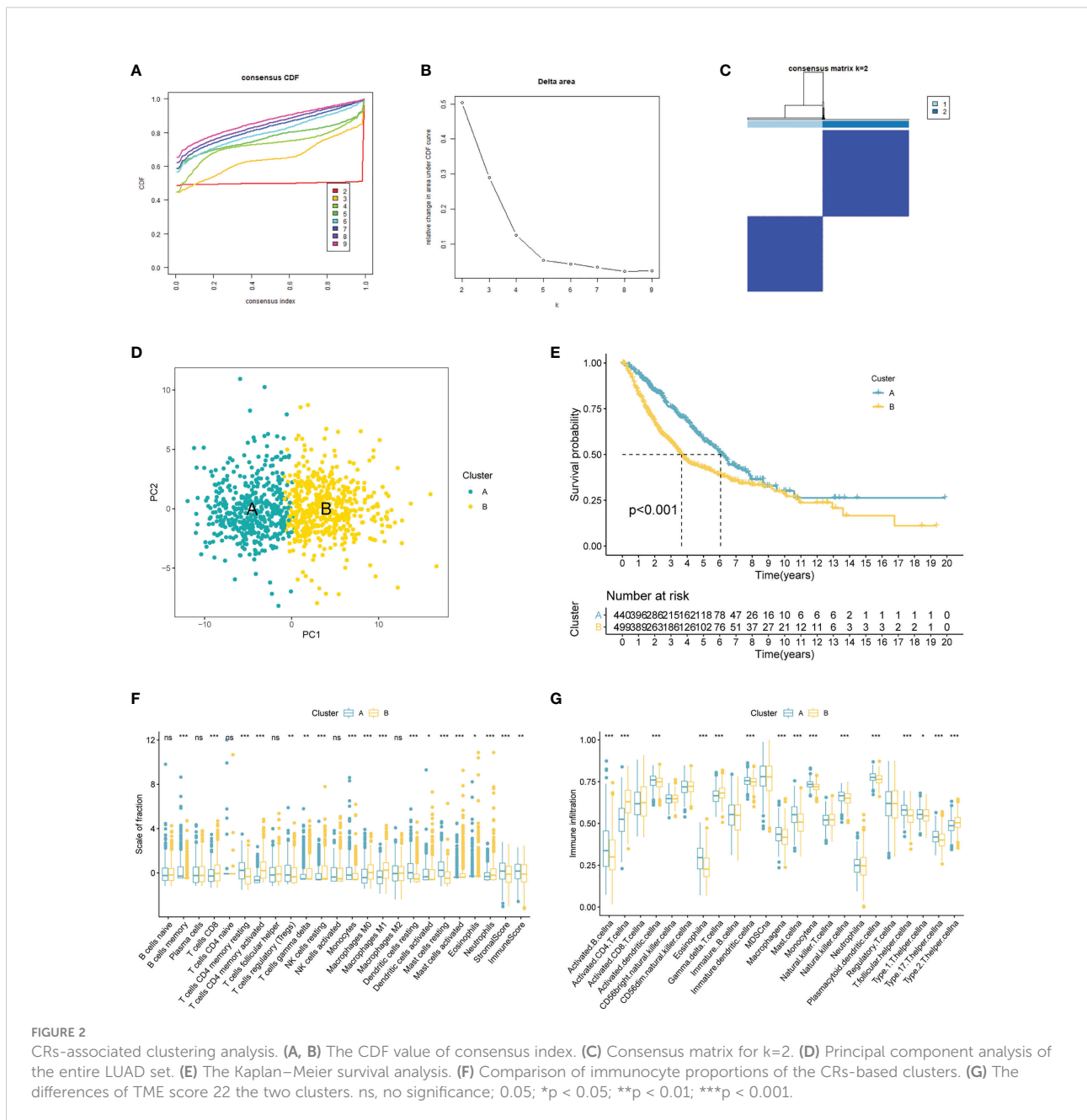


Figure 2F, cluster A presented higher stromal score and immune score than that cluster B. In addition, we observed that B cells, T cells, NK cells, dendritic cells and Macrophages showed the most notable difference among the two clusters (Figure 2G).

Construction of the CRBS

To develop an optimal prognostic signature, TCGA-LUAD cohort was selected as the training set. Univariate Cox regression was first employed to determine possible CRs with significant

prognostic values (Figure 3A). Subsequently, 12 candidate genes were enrolled into multivariate Cox analysis to create a CRBS that included four risk CRs (Figure 3B). The risk formula was shown as follows: $(0.1082 \times HMGA2) + (0.3525 \times NPAS2) + (0.1909 \times PRC1) + (-0.2416 \times CBX7)$. Survival curves illustrated that CBX7 was a potential favorable indicator, and HMGA2, NPAS2 and PRC1 were risky candidate indicators (Figure 3C). Then, we detect the expression differences of four CRs according to TCGA-LUAD dataset. All four CRs were greatly dysregulated between LUAD cases and control samples (Figure 3D). Furthermore, we validated the expression patterns of four



model genes by qRT-PCR in cell lines. Consistent with the above bioinformatics analysis results, we noticed that CBX7 was downregulated in LUAD cell lines (A549 and HCl-H1975), and HMGA2, NPAS2 and PRC1 were overexpressed in LUAD cell lines compared to BEAS-2B (Figure 3E). Consistent with the above results, we detected the expression patterns of four CRs at IHC level based on HPA database (Figure 3F).

Verification of the CRBS

Figure 4A demonstrated that survival rates are lower in CRBS-high group compared to CRBS-low group in the training set. The AUC (area under the curve) values of 1-, 3-, and 5-year survival rates assessed by the CRBS were 0.729, 0.662, and 0.634, respectively (Figure 4B). Figure 4C summarizes the positive correlation between surviving cases and risk score. Moreover, we observed a similar trend of results in the test set, suggesting the favorable prediction ability of the CRBS (Figures 4D–F). To further unearth the independence of our model, univariate and multivariate Cox regression analyses were employed. Univariate analysis indicated that the risk score was an independent indicator for prognosis in both two datasets (Figures 4G, I). The multivariate method disclosed that risk score was independently associated with the dismal outcome of LUAD cases (Figures 4H, J). At the same time, we explore the performance of the CRBS based on a diversity of clinical subgroups. The results revealed that low risk score was correlated with favorable outcomes in different ages, genders, T stage and N stage cohorts (Figures 5A–D). Similarly, the good prediction capability of the CRBS was confirmed in the T stage and N stage subgroups (Figures 5E, F).

Single-cell sequencing analysis

To decipher the single-cell transcriptome dataset GSE131907, Seurat package was performed. The UMAP analysis suggested the distribution of the 22 LUAD samples (N = 11 and T = 11) with no remarkable batch effects (Figure 6A). All the cells were divided into 12 clusters through k-Nearest Neighbor (KNN) clustering algorithm (Figure 6B). After performing cell annotation by different cell surface markers, we obtained eight cell subtypes, including B lymphocytes, endothelial cells, epithelial cells, fibroblasts, mast cells, myeloid cells, NK cells and T lymphocytes (Figure 6C). Next, we investigate the location of four CRs at single-cell transcriptome level. In Figure 6D,

HMGA2 and NPAS2 are mainly located in endothelial cells, and PRC1 and CBX7 are mainly located in NK cells and T lymphocytes. In addition, we noticed that the expression of CBX7 was negatively correlated with endothelial cells, whereas NPAS2 was positively correlated with endothelial cells (Figure 6E).

GSEA determines CRBS-associated pathways

In Figure 7A, the top six cancer hallmarks were remarkably enriched in the CRBS-high group, including glycolysis, hypoxia, mTORC1 pathway, MYC target, PI3K/Akt/mTOR and unfolded protein response. In terms of the KEGG analysis, we observed that CRBS-high group was involved in the cell cycle, p53 pathway and ubiquitination response (Figure 7B).

Immune environment analysis

To depict the immune landscape of LUAD, we evaluated the immunocyte infiltration of each case. Figure 8A summarizes the correlation between the 22 immunocyte types. As suggested in Figure 8B, CBX7 was greatly positively associated with memory B cells and resting mast cells. PRC1 was positively associated with activated memory T cells and negatively correlated with resting mast cells. Moreover, macrophage M0, macrophage M2 and resting NK cells were enriched in the CRBS-high group. Cases in CRBS-low group had greatly higher proportions of memory B cells, dendritic cells and mast cells (Figures 8C–H).

Additionally, some immune functions displayed differences between the two groups, including APC co-stimulation, checkpoint, HLA, MHC class I, T cell co-stimulation, and type II IFN response (Figure 9A). Also, we observed that four immune responses (checkpoint, HLA, MHC class I and type II IFN response) had significant differences in the outcome of patients with LUAD (Figures 9B–E).

Clinical potency analysis of the CRBS

TMB has been demonstrated to be useful as an indicator of the efficacy of immunotherapy. We calculated the TMB of each LUAD sample and found that CRBS-high group had a higher TMB than the CRBS-low group (Figures 10A). Moreover, CRBS-high group presented a high level of mRNasi (Figure 10B). In Figure 10C, most of the immune checkpoint markers were upregulated in the CRBS-high group. The comparison in the expression of m6A markers between the two groups indicated that the expression of ALKBH5, FTO, METTL14, HNRNPC, YTHDF1, YTHDF2, METTL3, RBM15 and WTAP were significant (Figure 10D).

Discussion

LUAD is the most common pathological subtype of lung cancer, which is composed of approximately 40% of lung cancer cases (19). Despite the various efforts in improve, the five-year

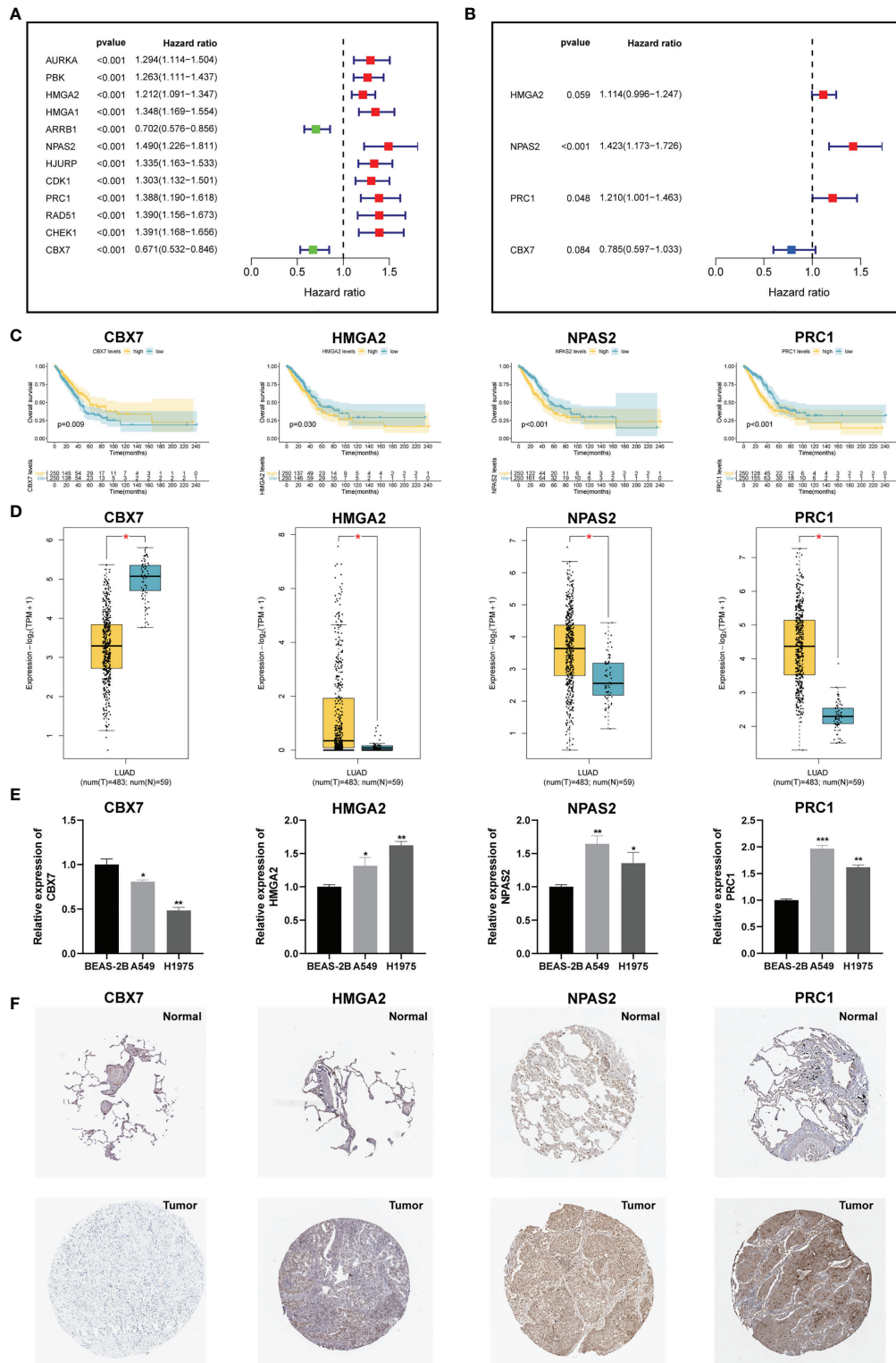
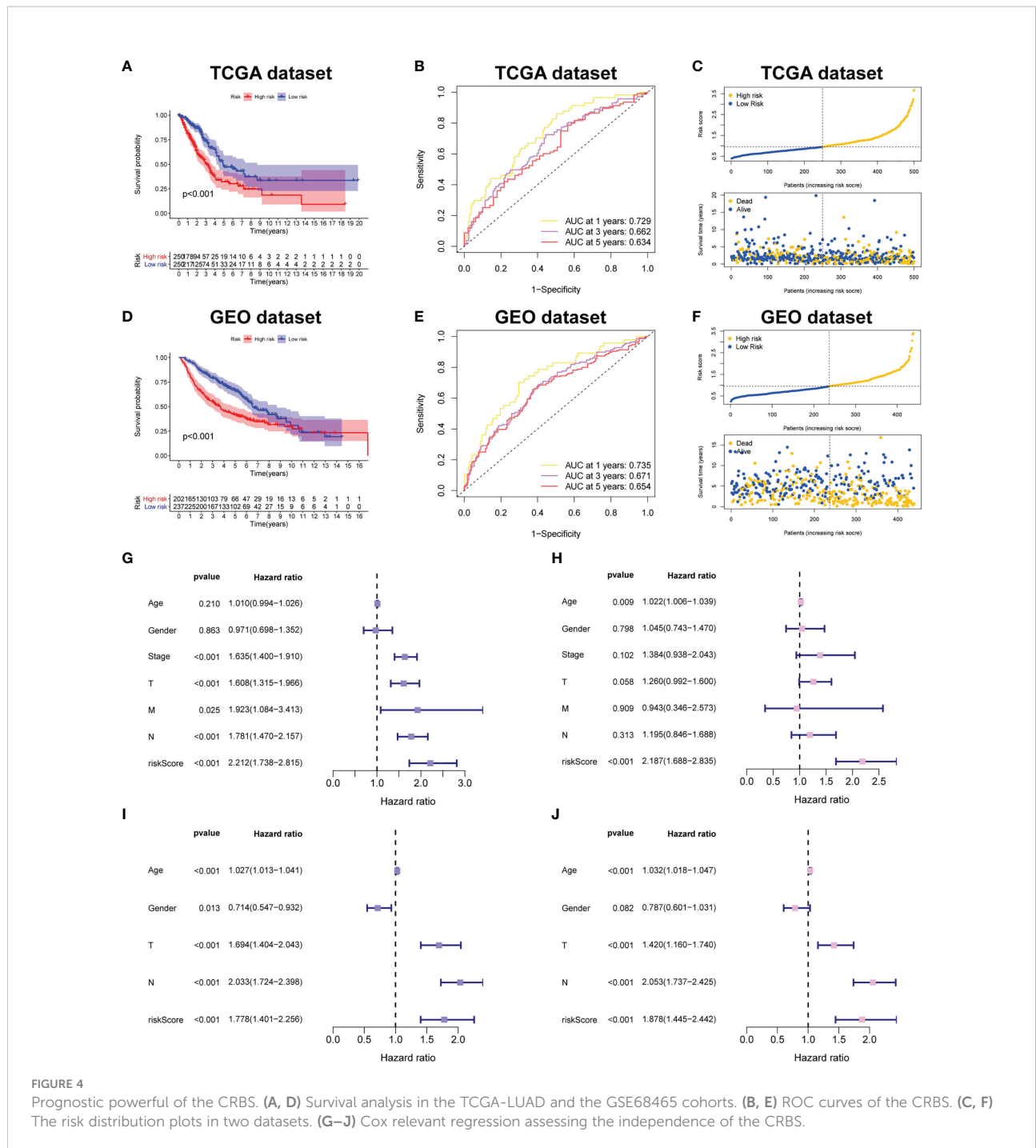


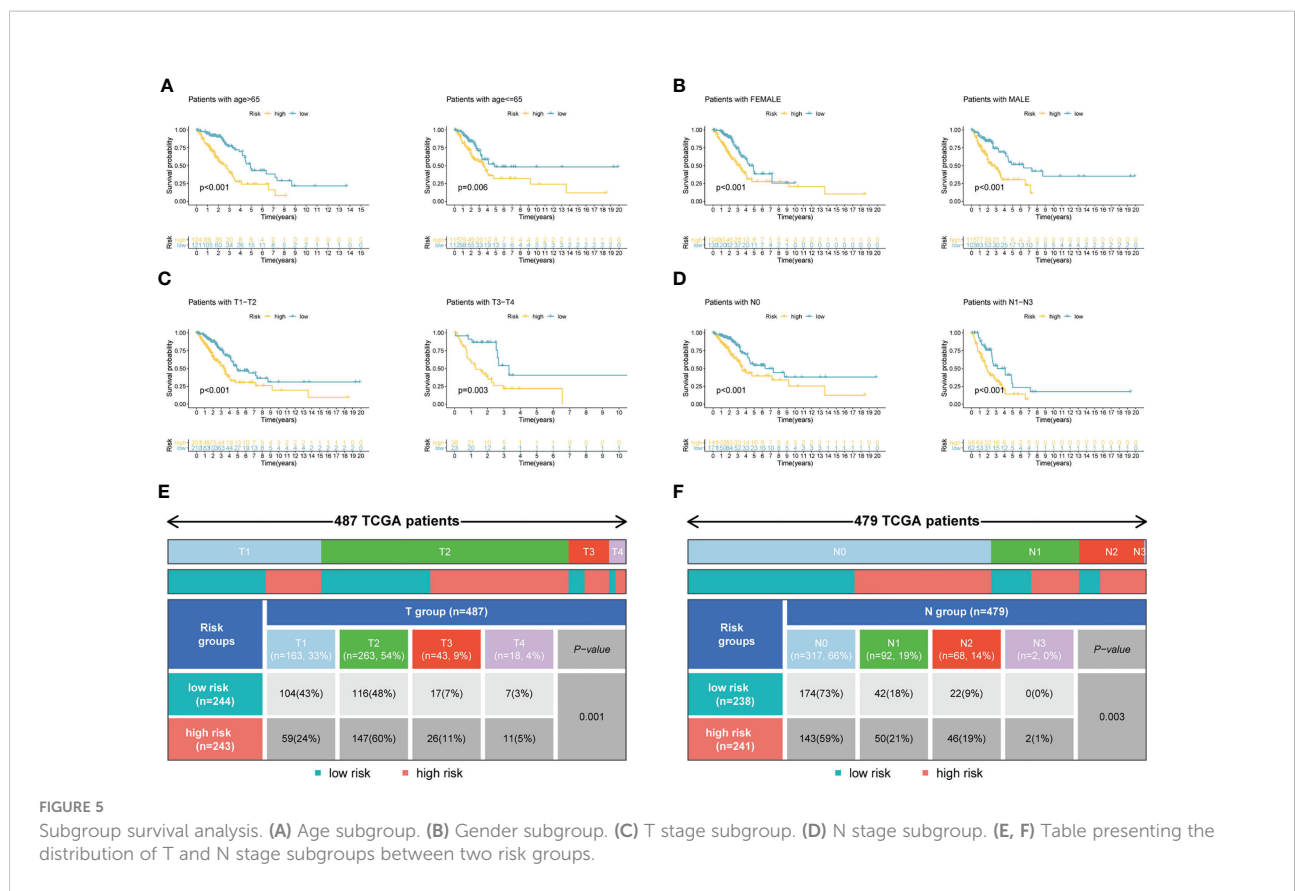
FIGURE 3 Construction of a CRs-based signature (CRBS). **(A, B)** Univariate and multivariate Cox analyses for signature establishment. **(C)** Survival analysis of four model CRs. **(D)** Comparison of differential expression of four model CRs based on GEPIA2 online portal. **(E)** The expression of four model CRs in BEAS-2B, A549 and H1975 cells line. **(F)** Immunohistochemistry of the CGs according to the HPA database. **p* < 0.05; ***p* < 0.01; ****p* < 0.001.



survival rate for LUAD patients remains shabby. Recent studies have suggested that patients with the same histology and TNM stages may have very distinct clinical outcomes, mainly due to their genetic heterogeneities (20). With the rapid development of the next-generation sequencing, a growing number of prognostic signatures based on transcriptome data were established to depict the individual differences, and to forecast the prognosis in various cancers (21–23). Therefore, a more reliable prognostic

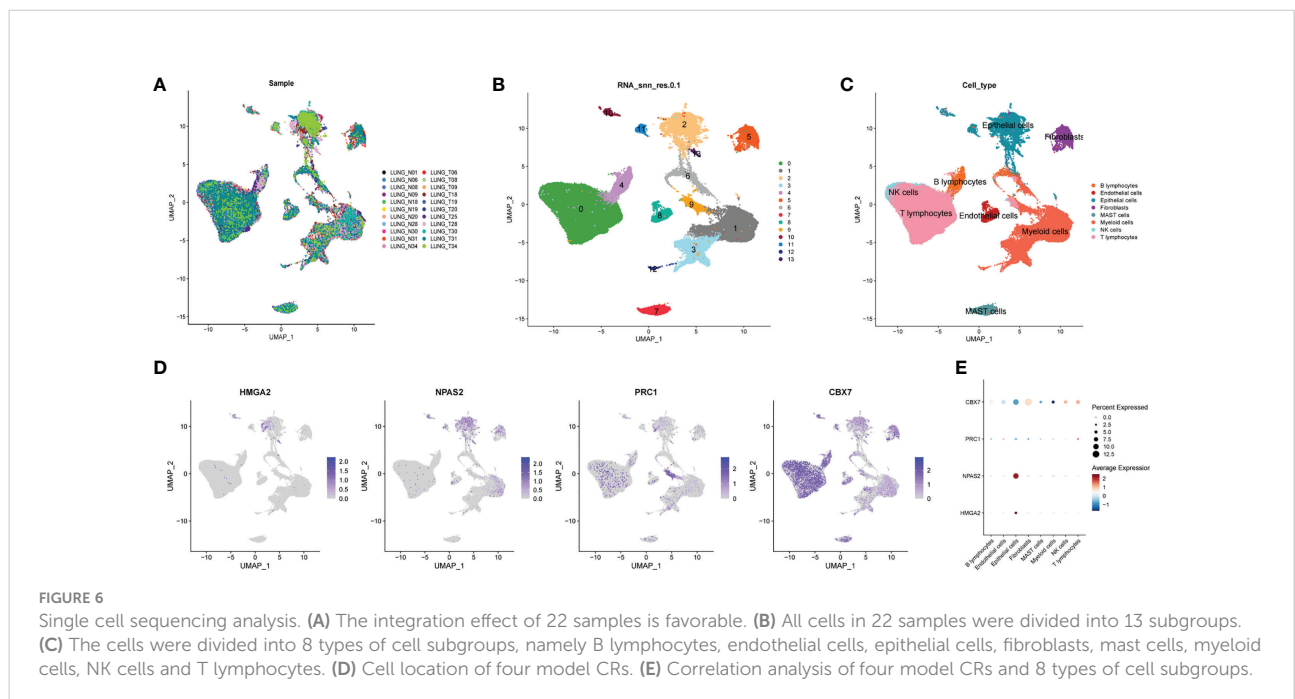
model based on genetic alterations is urgently needed to provide early detection and personalized treatment for LUAD patients.

It is well known that epigenetic alterations play a considerable role in mediating the tumor progression (24). As indispensable regulatory elements of epigenetics, CRs are involved in the onset and development of various cancer types including multiple myeloma, prostate cancer, hepatocellular carcinoma, and LUAD (25–28). In our current work, a total of



four CRs including HMGA2, NPAS2, PRC1, and CBX7, were identified as effective prognostic biomarkers for predicting the prognosis of LUAD. Survival analysis indicated that HMGA2, NPAS2 and PRC1 are potential risky genes since their high

expressions are correlated with dismal outcomes of LUAD samples, whereas CBX7 is candidate protective factor given that its high expression is associated with favorable outcomes of LUAD samples. The pro-tumor role of HMGA2 has been



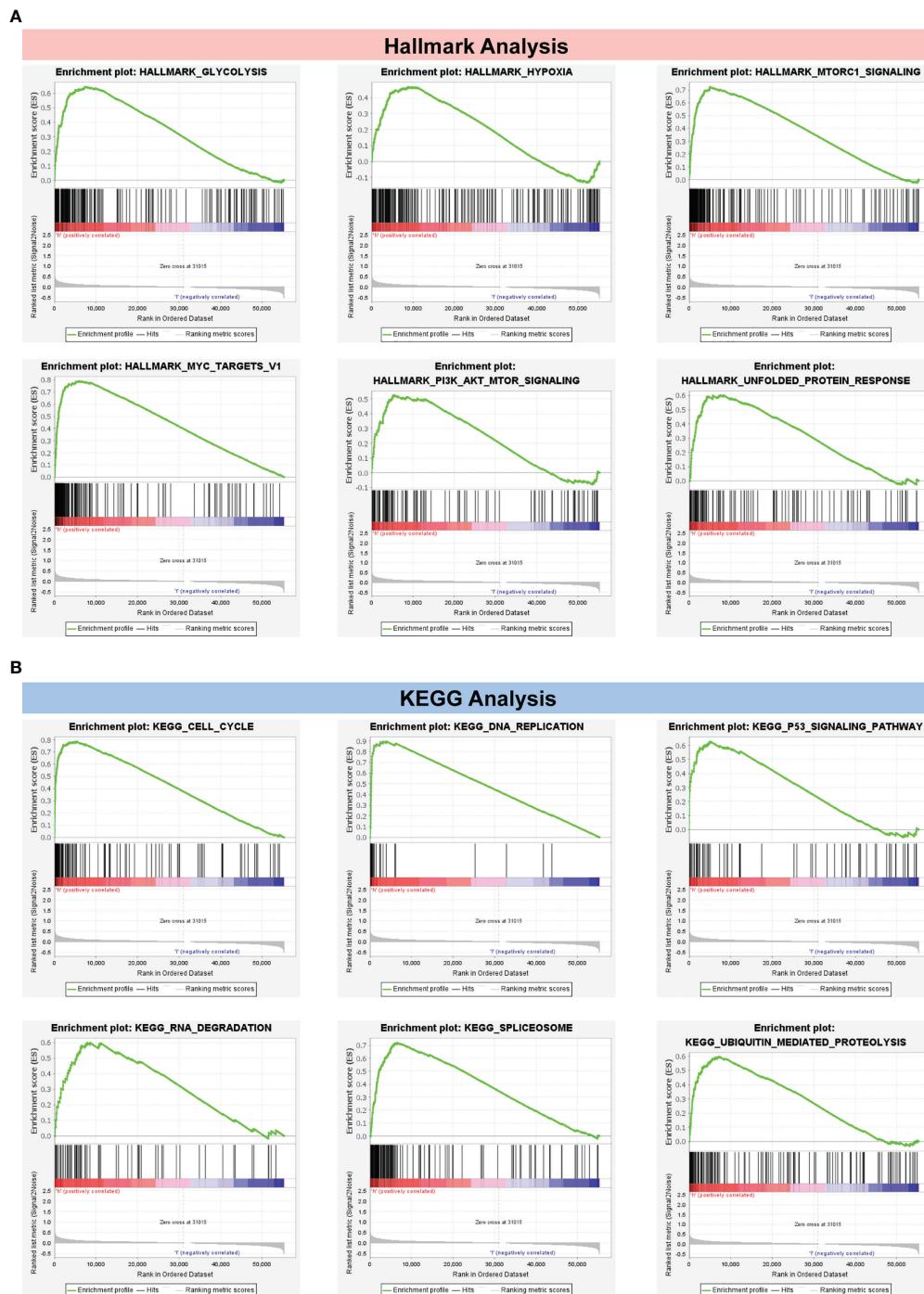
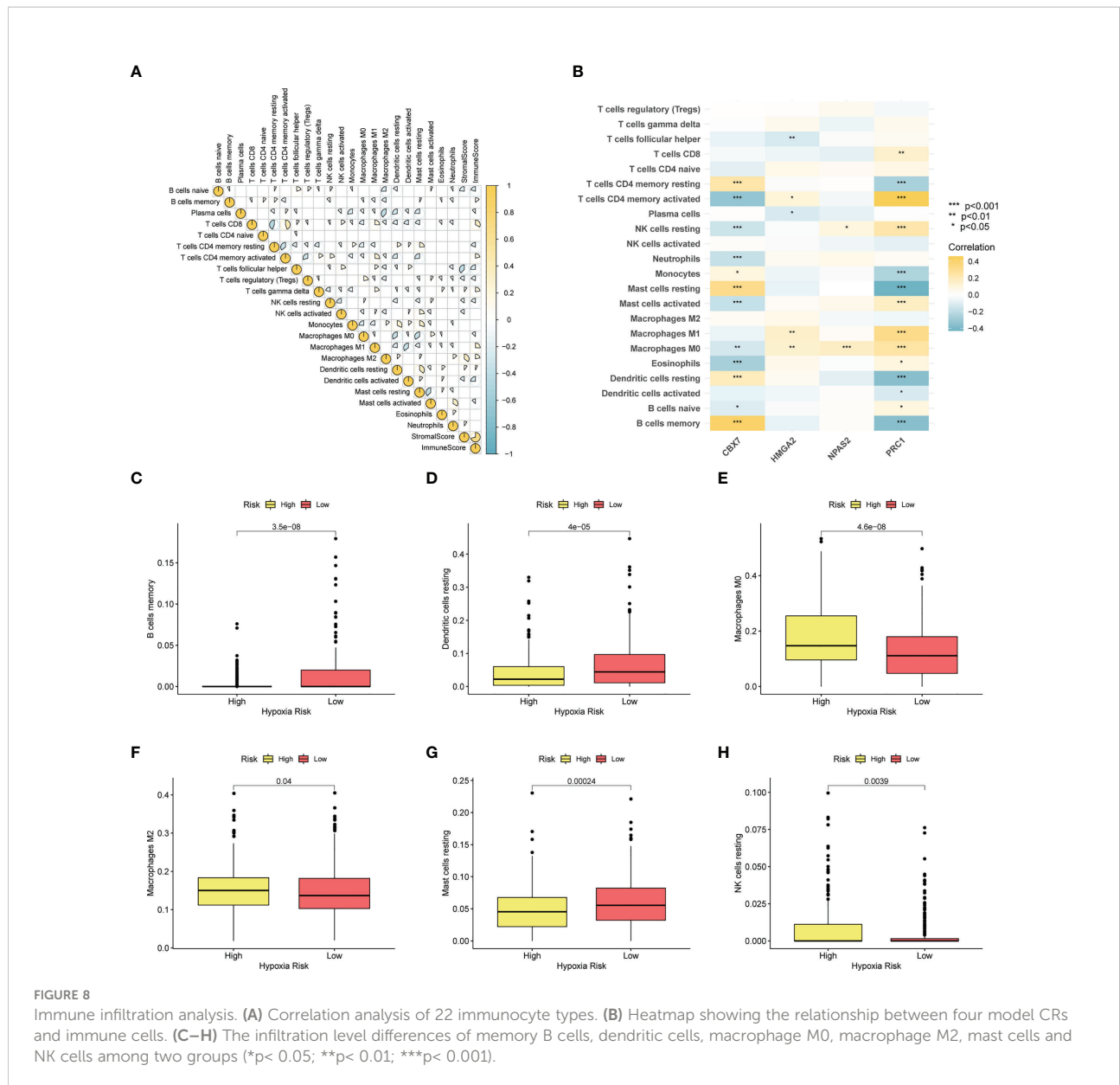


FIGURE 7
Gene Set Enrichment Analysis. **(A)** Hallmark analysis of the CRBS. **(B)** KEGG analysis of the CRBS.

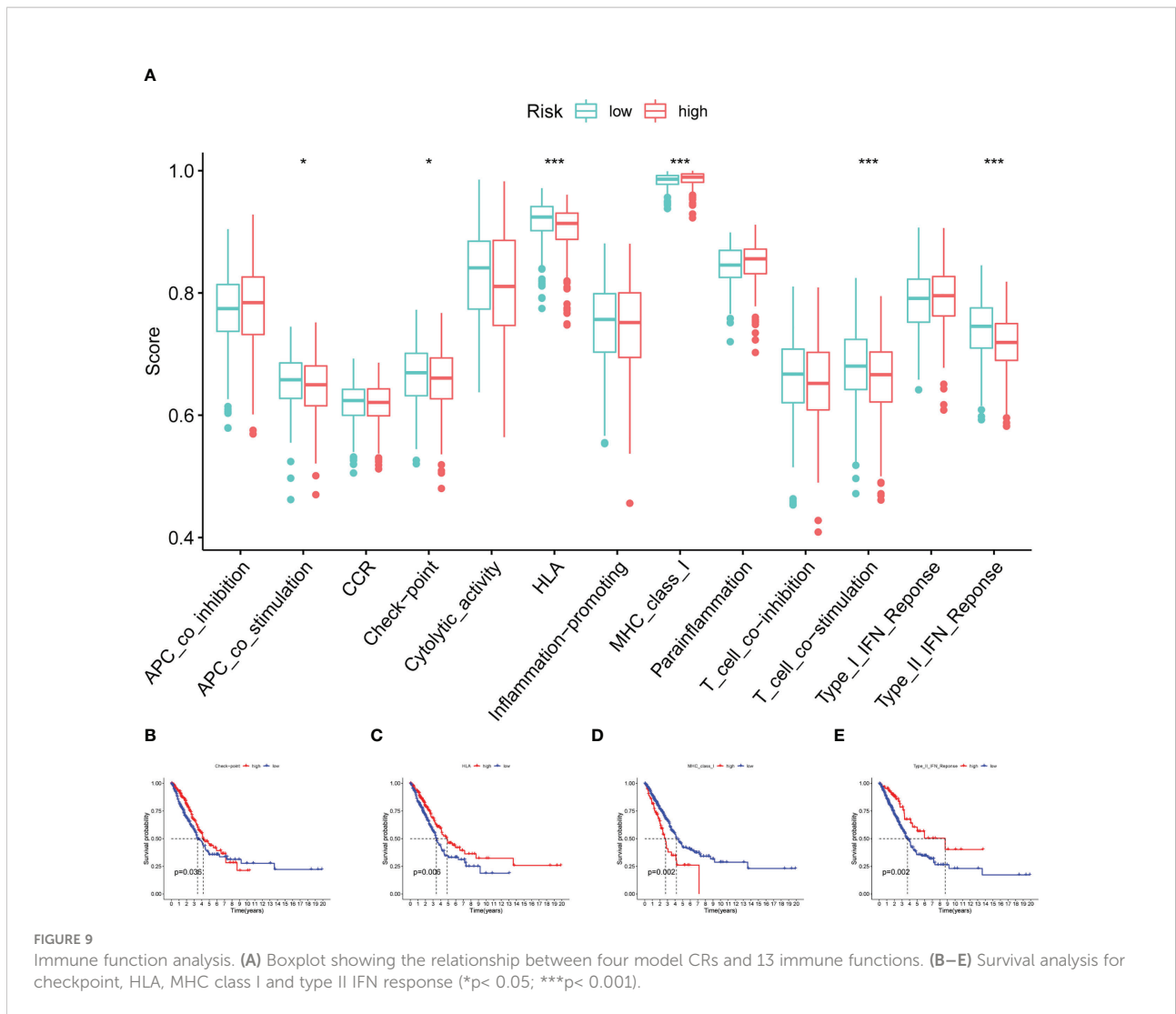
widely reported in multiple cancers including LUAD (29). HMG2 was found highly expressed in the LUAD tissues compared with normal lung tissues, and HMG2 silencing notably reduced the growth and metastasis of LUAD cell lines (30). In addition, a mechanistic study revealed that HMG2

could induce epithelial-mesenchymal transition by activating MAPK/extracellular receptor kinase signaling in LUAD (31). Npas2 has been identified in peripheral tissues, possibly as a modulator of circadian rhythms (32). Qiu et al. once reported that in LUAD, the elevated expression level of NPAS2 is



significantly related to poor prognosis (33). Conversely, it has been indicated that LUAD cases with low NPAS2 expression displayed a favorable clinical outcome by another team (34). Therefore, more basic researches are needed to elucidate the exact role of NPAS2 in LUAD. PRC1 has received widespread attention considering its diverse regulatory roles in a number of diseases, especially tumorigenesis (35). It has been suggested that overexpression of PRC1 triggers the onset of various cancers yet its potential roles in LUAD have not been fully understood (36). An ever-growing series of reports has demonstrated the aberrant expression of PRC1 in a variety of tumors (37). Mechanically, CBX7 may exert its tumor suppressor role by inhibiting the Wnt pathway and subsequently restrain the malignant character in LUAD (38).

GSEA unearthed that CRBS-high group were involved in glycolysis signaling, PI3K/AKT/mTOR signaling, and p53 signaling pathway using GSEA. Suppressed oxidative phosphorylation along with enhanced glycolysis, which is called the Warburg phenotype, is considered as metabolic marker of cancers (39). Vaupel et al. once reported that enhanced glycolysis accelerates lactic acid accumulation to impair the immune functions in TME and finally promote malignant progression (40). The PI3K/AKT/mTOR pathway plays a crucial role in diverse biological behaviors including cell growth, migration, metabolism, and death (41). In LUAD, the aberrant activation in this signaling has been indicated to induce uncontrolled growth, drug resistance, sustained angiogenesis, as well as distant metastasis (42). P53 protein



is a transcription factor known as the “guardian gene” because of its significant role in preserving genomic integrity. The mutation of the p53 gene can be detected in a wide spectrum of human malignancies, including the breast, cervical, lung, and prostate cancer (43). More recently, Vokes and his colleagues provided the evidence that p53 alterations were involved in faster resistance evolution and may cooperate with other genomic events to gain resistance to EGFR tyrosine kinase inhibitors (44).

Immunotherapy that emerged recently has achieved promising results in the treatment of LUAD (45). In our work, a comprehensive analysis of tumor-infiltrating immune cells was further conducted to help to clarify the immune infiltration status between the two different risk groups. As a result, the infiltration level of HLA as well as the type 2 IFN was found downregulated remarkably in CRBS-high group. Also, the expression level of the immune

checkpoint markers was validated to be correlated with the risk score. CD273, also named B3-H7, is overexpressed in various solid malignancies which serve as a potential therapeutic target (46, 47). Yu and his colleagues disclosed CD273 was upregulated in LUAD, and was correlated with lymph node metastasis (48). Likewise, accumulating studies have indicated the close association between the efficacy of immunotherapy and the CD274 expression (49). VTCN1, also named B7-H4, belongs to the co-stimulatory B7 family molecules and is associated with a poor prognosis in multiple cancer types (50–52). As revealed by a recent study, the elevation of VTCN1 expression is associated with LUAD with EGFR-activating mutations, which can ultimately cause resistance to immunotherapy in LUAD patients (53).

In view of the essential effect of m6A methylation modification in LUAD progression, we unearthed the

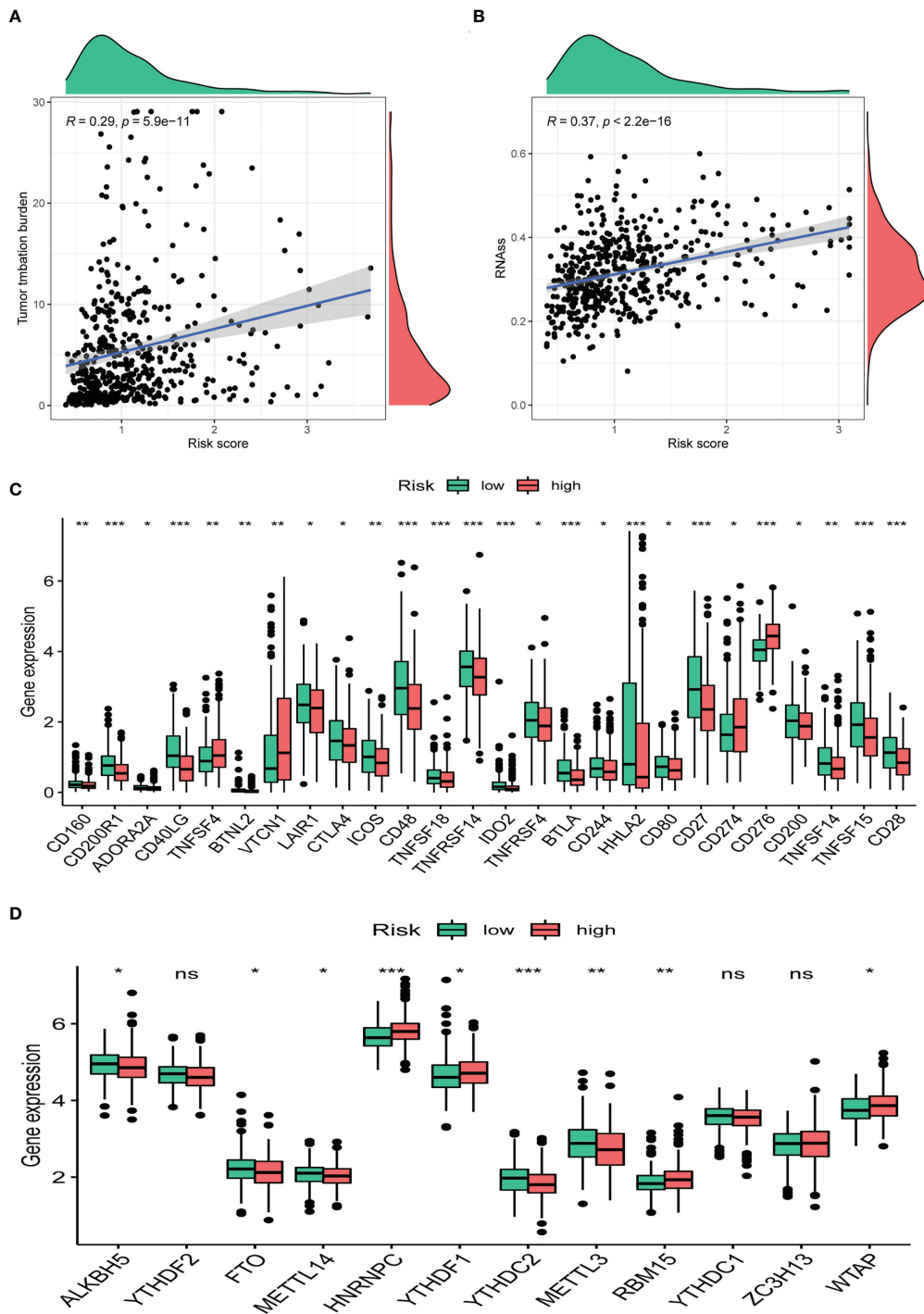


FIGURE 10
 Clinical potency analysis. (A) TMB analysis of the CRBS. (B) Cancer stem cells index analysis the CRBS. Comparison of differential expression of (C) immune checkpoints and (D) m6A markers (* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$). ns, no significance.

expression patterns of m6A regulators between two risk groups. The results indicated that HNRNPC, YTHDF1, RBM15 and WTAP were enriched in the high-risk group. Lou and his colleagues demonstrated that YTHDF1 could facilitate LUAD

growth and survival by enhancing Cyclin B1 translation (54). In addition, YTHDF1 has also been confirmed to have carcinogenic effects in many digestive system tumors including gastric cancer, hepatocellular carcinoma and colorectal cancer (55–57). Cheng

et al. found that overexpression of WTAP correlate with dismal outcome of LC cases. In NSCLC, PCGEM1 could boost cancer cells proliferation by improving WTAP expression (58, 59).

There are still several limitations of the present study that need to be considered. Only expression data in gene level was analyzed to construct the prognostic model, and large-sample clinical data are still needed, as an external cohort, to evaluate the predictive value of our model. Additionally, although we have proven the reliable prognostic capacity of the four CR related genes, fundamental experiments are still needed to validate their precise functions in mediating LUAD progression.

Conclusion

Taken together, our data may help provide opportunities for the development of new therapeutic strategies and elucidate the mechanism of tumor immune escape in LUAD. Our proposed model may usher in novel approaches to predicting prognosis of patients with LUAD.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68:394–424. doi: 10.3322/caac.21492
2. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: Epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc* (2008) 83:584–94. doi: 10.4065/83.5.584
3. Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* (2012) 150:1107–20. doi: 10.1016/j.cell.2012.08.029
4. Bejarano L, Jordão MJC, Joyce JA. Therapeutic targeting of the tumor microenvironment. *Cancer Discov* (2021) 11:933–59. doi: 10.1158/2159-8290.CD-20-1808
5. Jin M-Z, Jin W-L. The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduct Target Ther* (2020) 5:166. doi: 10.1038/s41392-020-00280-x
6. Lu J, Xu J, Li J, Pan T, Bai J, Wang L, et al. FACER: Comprehensive molecular and functional characterization of epigenetic chromatin regulators. *Nucleic Acids Res* (2018) 46:10019–33. doi: 10.1093/nar/gky679
7. Plass C, Pfister SM, Lindroth AM, Bogatyrova O, Claus R, Lichter P. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. *Nat Rev Genet* (2013) 14:765–80. doi: 10.1038/nrg3554
8. Saed L, Jelen A, Mirowski M, Salagacka-Kubiak A. Prognostic significance of HMG1A1 expression in lung cancer based on bioinformatics analysis. *Int J Mol Sci* (2022) 23:6933. doi: 10.3390/ijms23136933

Author contributions

QS and HM visualized the study and took part in the study design. QS, SH, XL, and SW performed the manuscript writing and bioinformatics analysis. All authors read and approved the final manuscript.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2022.1031728/full#supplementary-material>

9. Qiao G-B, Wang R-T, Wang S-N, Tao S-L, Tan Q-Y, Jin H. GRP75-mediated upregulation of HMG1A1 stimulates stage I lung adenocarcinoma progression by activating JNK/c-JUN signaling. *Thorac Cancer* (2021) 12:1558–69. doi: 10.1111/1759-7714.13944
10. Wang X, Zhao H, Lv L, Bao L, Wang X, Han S. Prognostic significance of EZH2 expression in non-small cell lung cancer: A meta-analysis. *Sci Rep* (2016) 6:19239. doi: 10.1038/srep19239
11. Hu J, Zhang Z, Zhao L, Li L, Zuo W, Han L. High expression of RAD51 promotes DNA damage repair and survival in KRAS-mutant lung cancer cells. *BMB Rep* (2019) 52:151–6. doi: 10.5483/BMBRep.2019.52.2.213
12. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43:e47–7. doi: 10.1093/nar/gkv007
13. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *Omics J Integr Biol* (2012) 16:284–7. doi: 10.1089/omi.2011.0118
14. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* (2003) 13:2498–504. doi: 10.1101/gr.1239303
15. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* (2019) 47:D607–13. doi: 10.1093/nar/gky1131
16. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26:1572–3. doi: 10.1093/bioinformatics/btq170

17. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
18. Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome: The human protein atlas. *Protein Sci* (2018) 27:233–44. doi: 10.1002/pro.3307
19. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin* (2018) 68:7–30. doi: 10.3322/caac.21442
20. Benusiglio PR, Fallet V, Sanchis-Borja M, Coulet F, Cadranet J. Lung cancer is also a hereditary disease. *Eur Respir Rev Off J Eur Respir Soc* (2021) 30:210045. doi: 10.1183/16000617.0045-2021
21. Dai D, Li Q, Zhou P, Huang J, Zhuang H, Wu H, et al. Analysis of omics data reveals nucleotide excision repair-related genes signature in highly-grade serous ovarian cancer to predict prognosis. *Front Cell Dev Biol* (2022) 10:874588. doi: 10.3389/fcell.2022.874588
22. Song J, Sun Y, Cao H, Liu Z, Xi L, Dong C, et al. A novel pyroptosis-related lncRNA signature for prognostic prediction in patients with lung adenocarcinoma. *Bioengineered* (2021) 12:5932–49. doi: 10.1080/21655979.2021.1972078
23. Buttarelli M, Ciucci A, Palluzzi F, Raspaglio G, Marchetti C, Perrone E, et al. Identification of a novel gene signature predicting response to first-line chemotherapy in BRCA wild-type high-grade serous ovarian cancer patients. *J Exp Clin Cancer Res* (2022) 41:50. doi: 10.1186/s13046-022-02265-w
24. Wang H, Wang Q, Cai G, Duan Z, Nugent Z, Huang J, et al. Nuclear TIGAR mediates an epigenetic and metabolic autoregulatory loop via NRF2 in cancer therapeutic resistance. *Acta Pharm Sin B* (2022) 12:1871–84. doi: 10.1016/j.apsb.2021.10.015
25. Zhang J, Zhang H, Ding X, Hu J, Li Y, Zhang J, et al. Crosstalk between macrophage-derived PGE2 and tumor UHRF1 drives hepatocellular carcinoma progression. *Theranostics* (2022) 12:3776–93. doi: 10.7150/thno.69494
26. Tirtakusuma R, Szoltysek K, Milne P, Grinev V, Ptasinska A, Chin PS, et al. Epigenetic regulator genes direct lineage switching in MLL/AF4 leukaemia. *Blood* (2022). doi: 10.1182/blood.2021015036
27. Zhang L, Zhang R, Wang J, Chen Y, Qiao C, Shi Q, et al. Identification of clinical implications and potential prognostic models of chromatin regulator mutations in multiple myeloma. *Clin Epigenet* (2022) 14:93. doi: 10.1186/s13148-022-01314-7
28. Wu L, Kou F, Ji Z, Li B, Zhang B, Guo Y, et al. SMYD2 promotes tumorigenesis and metastasis of lung adenocarcinoma through RPS7. *Cell Death Dis* (2021) 12:439. doi: 10.1038/s41419-021-03720-w
29. Li Z, Wu X, Li J, Yu S, Ke X, Yan T, et al. HMGA2-Snai2 axis regulates tumorigenicity and stemness of head and neck squamous cell carcinoma. *Exp Cell Res* (2022) 418:113271. doi: 10.1016/j.yexcr.2022.113271
30. Wang M, Ma M, Yang Y, Li C, Wang Y, Sun X, et al. Overexpression of hsa_circ_0008274 inhibited the progression of lung adenocarcinoma by regulating HMGA2 via sponging miR-578. *Thorac Cancer* (2021) 12:2258–64. doi: 10.1111/1759-7714.14059
31. Wang B, Pan L-Y, Kang N, Shen X-Y. PP4R1 interacts with HMGA2 to promote non-small-cell lung cancer migration and metastasis via activating MAPK/ERK-induced epithelial-mesenchymal transition. *Mol Carcinog* (2020) 59:467–77. doi: 10.1002/mc.23168
32. Zhou YD, Barnard M, Tian H, Li X, Ring HZ, Francke U, et al. Molecular characterization of two mammalian bHLH-PAS domain proteins selectively expressed in the central nervous system. *Proc Natl Acad Sci USA* (1997) 94:713–8. doi: 10.1073/pnas.94.2.713
33. Qiu M, Chen Y-B, Jin S, Fang X-F, He X-X, Xiong Z-F, et al. Research on circadian clock genes in non-small-cell lung carcinoma. *Chronobiol Int* (2019) 36:739–50. doi: 10.1080/07420528.2018.1509080
34. Gao L-W, Wang G-L. Comprehensive bioinformatics analysis identifies several potential diagnostic markers and potential roles of cyclin family members in lung adenocarcinoma. *OncoTargets Ther* (2018) 11:7407–15. doi: 10.2147/OTT.S171705
35. Angrand P-O. Structure and function of the polycomb repressive complexes PRC1 and PRC2. *Int J Mol Sci* (2022) 23:5971. doi: 10.3390/ijms23115971
36. Pallante P, Forzati F, Federico A, Arra C, Fusco A. Polycomb protein family member CBX7 plays a critical role in cancer progression. *Am J Cancer Res* (2015) 5:1594–601.
37. Li R, Yan Q, Tian P, Wang Y, Wang J, Tao N, et al. CBX7 inhibits cell growth and motility and induces apoptosis in cervical cancer cells. *Mol Ther Oncolytics* (2019) 15:108–16. doi: 10.1016/j.omto.2019.09.002
38. Pei Y-F, He Y, Hu L-Z, Zhou B, Xu H-Y, Liu X-Q. The crosstalk between lncRNA-SNHG7/miRNA-181/cbx7 modulates malignant character in lung adenocarcinoma. *Am J Pathol* (2020) 190:1343–54. doi: 10.1016/j.ajpath.2020.02.011
39. Siddiqui A, Ceppi P. A non-proliferative role of pyrimidine metabolism in cancer. *Mol Metab* (2020) 35:100962. doi: 10.1016/j.molmet.2020.02.005
40. Vaupel P, Schmidberger H, Mayer A. The warburg effect: Essential part of metabolic reprogramming and central contributor to cancer progression. *Int J Radiat Biol* (2019) 95:912–9. doi: 10.1080/09553002.2019.1589653
41. Alzahrani AS. PI3K/Akt/mTOR inhibitors in cancer: At the bench and bedside. *Semin Cancer Biol* (2019) 59:125–32. doi: 10.1016/j.semcancer.2019.07.009
42. Sanaei M-J, Razi S, Pourbagheri-Sigaroodi A, Bashash D. The PI3K/Akt/mTOR pathway in lung cancer; oncogenic alterations, therapeutic opportunities, challenges, and a glance at the application of nanoparticles. *Transl Oncol* (2022) 18:101364. doi: 10.1016/j.tranon.2022.101364
43. Azulay EE, Cooks T, Elkabets M. Potential oncogenic roles of mutant-p53-derived exosomes in the tumor-host interaction of head and neck cancers. *Cancer Immunol Immunother CII* (2020) 69:285–92. doi: 10.1007/s00262-019-02450-5
44. Vokes NI, Chambers E, Nguyen T, Coolidge A, Lydon CA, Le X, et al. Concurrent TP53 mutations facilitate resistance evolution in EGFR-mutant lung adenocarcinoma. *J Thorac Oncol Off Publ Int Assoc Study Lung Cancer* (2022) 17:779–92. doi: 10.1016/j.jtho.2022.02.011
45. Kaur J, Elms J, Munn AL, Good D, Wei MQ. Immunotherapy for non-small cell lung cancer (NSCLC), as a stand-alone and in combination therapy. *Crit Rev Oncol Hematol* (2021) 164:103417. doi: 10.1016/j.critrevonc.2021.103417
46. Zhou W-T, Jin W-L. B7-H3/CD276: An emerging cancer immunotherapy. *Front Immunol* (2021) 12:701006. doi: 10.3389/fimmu.2021.701006
47. Wang S, Zhang X, Ning H, Dong S, Wang G, Sun R. B7 homolog 3 induces lung metastasis of breast cancer through Raf/MEK/ERK axis. *Breast Cancer Res Treat* (2022) 193:405–16. doi: 10.1007/s10549-022-06520-8
48. Yu T-T, Zhang T, Lu X, Wang R-Z. B7-H3 promotes metastasis, proliferation, and epithelial-mesenchymal transition in lung adenocarcinoma. *OncoTargets Ther* (2018) 11:4693–700. doi: 10.2147/OTT.S169811
49. Oitabén A, Fonseca P, Villanueva MJ, García-Benito C, López-López A, Garrido-Fernández A, et al. Emerging blood-based biomarkers for predicting immunotherapy response in NSCLC. *Cancers* (2022) 14:2626. doi: 10.3390/cancers14112626
50. Iizuka A, Nonomura C, Ashizawa T, Kondou R, Ohshima K, Sugino T, et al. A T-cell-engaging B7-H4/CD3-bispecific fab-scFv antibody targets human breast cancer. *Clin Cancer Res Off J Am Assoc Cancer Res* (2019) 25:2925–34. doi: 10.1158/1078-0432.CCR-17-3123
51. Liu Z, Jin K, Zeng H, Shao F, Chang Y, Wang Y, et al. B7-H4 correlates with clinical outcome and immunotherapeutic benefit in muscle-invasive bladder cancer. *Eur J Cancer Oxf Engl* 1990 (2022) 171:133–42. doi: 10.1016/j.ejca.2022.05.022
52. Schalper KA, Carvajal-Hausdorf D, McLaughlin J, Altan M, Velcheti V, Gaule P, et al. Differential expression and significance of PD-L1, IDO-1, and B7-H4 in human lung cancer. *Clin Cancer Res Off J Am Assoc Cancer Res* (2017) 23:370–8. doi: 10.1158/1078-0432.CCR-16-0150
53. Lu Y, Wu F, Cao Q, Sun Y, Huang M, Xiao J, et al. B7-H4 is increased in lung adenocarcinoma harboring EGFR-activating mutations and contributes to immunosuppression. *Oncogene* (2022) 41:704–17. doi: 10.1038/s41388-021-02124-6
54. Lou X, Ning J, Liu W, Li K, Qian B, Xu D, et al. YTHDF1 promotes cyclin B1 translation through m6A modulation and contributes to the poor prognosis of lung adenocarcinoma with KRAS/TP53 Co-mutation. *Cells* (2021) 10:1669. doi: 10.3390/cells10071669
55. Wang S, Gao S, Zeng Y, Zhu L, Mo Y, Wong CC, et al. N6-methyladenosine reader YTHDF1 promotes ARHGEF2 translation and RhoA signaling in colorectal cancer. *Gastroenterology* (2022) 162:1183–96. doi: 10.1053/j.gastro.2021.12.269
56. Zhao X, Chen Y, Mao Q, Jiang X, Jiang W, Chen J, et al. Overexpression of YTHDF1 is associated with poor prognosis in patients with hepatocellular carcinoma. *Cancer Biomark Sect Dis Markers* (2018) 21:859–68. doi: 10.3233/CBM-170791
57. Chen X-Y, Liang R, Yi Y-C, Fan H-N, Chen M, Zhang J, et al. The m6A reader YTHDF1 facilitates the tumorigenesis and metastasis of gastric cancer via USP14 translation in an m6A-dependent manner. *Front Cell Dev Biol* (2021) 9:647702. doi: 10.3389/fcell.2021.647702
58. Cheng H, Wang S-J, Li Z, Ma Y, Song Y-R. ING2-WTAP is a potential therapeutic target in non-small cell lung cancer. *Biochem Biophys Res Commun* (2022) 605:31–8. doi: 10.1016/j.bbrc.2022.02.041
59. Weng L, Qiu K, Gao W, Shi C, Shu F. LncRNA PCGEM1 accelerates non-small cell lung cancer progression via sponging miR-433-3p to upregulate WTAP. *BMC Pulm Med* (2020) 20:213. doi: 10.1186/s12890-020-01240-5