



OPEN ACCESS

EDITED BY

Anton A. Buzdin,
European Organisation for Research and
Treatment of Cancer, Belgium

REVIEWED BY

Wallax Ferreira,
Evandro Chagas Institute, Brazil
Jing Liu,
Johns Hopkins University, United States

*CORRESPONDENCE

Junling Wang,
✉ wangjl@lzu.edu.cn
Zhitong Bing,
✉ bingzt@impcas.ac.cn

RECEIVED 20 March 2024

ACCEPTED 10 September 2024

PUBLISHED 23 September 2024

CITATION

Liu J, Bing Z and Wang J (2024) Comprehensive
pan-cancer analysis and experiments revealed
R3HDM1 as a novel predictive biomarker for
prognosis and immune therapy response.
Front. Genet. 15:1404348.
doi: 10.3389/fgene.2024.1404348

COPYRIGHT

© 2024 Liu, Bing and Wang. This is an open-
access article distributed under the terms of the
[Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).
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.

Comprehensive pan-cancer analysis and experiments revealed R3HDM1 as a novel predictive biomarker for prognosis and immune therapy response

Jiawei Liu¹, Zhitong Bing^{2,3*} and Junling Wang^{1*}

¹School of Public Health, Lanzhou University, Lanzhou, Gansu, China, ²Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou, Gansu, China, ³Gansu Laboratory of Isotope, Gansu Provincial Laboratory, Lanzhou, Gansu, China

Background: R3HDM1, an RNA binding protein with one R3H domain, remains uncharacterized in terms of its association with tumor progression, malignant cell regulation, and the tumor immune microenvironment. This paper aims to fill this gap by analyzing the potential of R3HDM1 in diagnosis, prognosis, chemotherapy, and immune function across various cancers.

Methods: Data was collected from the Firehost database (<http://gdac.broadinstitute.org>) to obtain the TCGA pan-cancer queue containing tumor and normal samples. Additional data on miRNA, TCPA, mutations, and clinical information were gathered from the UCSC Xena database (<https://xenabrowser.net/datapages/>). The mutation frequency and locus of R3HDM1 in the TCGA database were examined using the cBioPortal. External validation through GEO data was conducted to assess the differential expression of R3HDM1 in different cancers. Protein expression levels were evaluated using the Clinical Proteomics Tumor Analysis Alliance (CPTAC). The differential expression of R3HDM1 was verified in lung adenocarcinoma cell lines and normal lung glandular epithelial cells via RT-qPCR. Cell migration and proliferation experiments were conducted by knocking down the expression of R3HDM1 in two lung adenocarcinoma cell lines using small interfering RNA. The biological role of R3HDM1 in pan-cancer was explored using the GSEA method. Multiple immune infiltration algorithms from the TIMER2.0 database was employed to investigate the correlation between R3HDM1 expression and the tumor immune microenvironment. Validation of transcriptome immune infiltration was based on 140 single-cell datasets from the TISCH database. The study also characterized a pan-cancer survival profile and analyzed the differential expression of R3HDM1 in different molecular subtypes. The relationship between R3HDM1 and drug resistance was investigated using four chemotherapy data sources: CellMiner, GDSC, CTRP and PRISM. The impact of chemicals on the expression of R3HDM1 was explored through the CTD database.

Result: The study revealed differential expression of R3HDM1 in various tumors, indicating its potential as an early diagnostic marker. Changes in somatic copy number (SCNA) and DNA methylation were identified as factors contributing to abnormal expression levels. Additionally, the study found that R3HDM1 expression is associated with clinical features, metabolic pathways, and important pathways related to metastasis and the immune system. High

expression of R3HDM1 was linked to poor prognosis across different tumors and altered drug sensitivity. Furthermore, the expression of R3HDM1 showed significant correlations with immune modulatory molecules and biomarkers of lymphocyte subpopulation infiltration. Finally, the study highlighted four chemicals that could influence the expression of R3HDM1.

Conclusion: Overall, this study proposes that R3HDM1 expression is a promising biomarker for predicting the prognosis of cancer, especially lung adenocarcinoma, and the efficacy of immunotherapy, demonstrating the rationale for further exploration in the development of anti-tumor therapies.

KEYWORDS

R3HDM1, diagnosis, pan-cancer, prognosis, TME, LUAD

1 Introduction

Cancer encompasses a complex array of diseases that currently pose a significant obstacle to global life expectancy. It primarily arises from the progressive development of uncontrolled cell growth and evasion of normal cell death, eventually spreading to adjacent organs and tissues over time. According to the 2020 statistics from the World Health Organization (WHO), cancer causes approximately ten million deaths worldwide, with the number of cancer patients remaining high and continuing to rise (Nassar and Blanpain, 2016). Given that cancer cells stem from normal cells, conventional drug treatments not only target cancer cells but also exhibit significant toxicity to non-cancerous cells in the body. Therefore, there is an urgent need for novel cancer targeted therapies that offer increased efficacy and reduced side effects (Phillips et al., 2006; Hanahan and Weinberg, 2011).

R3HDM1, also known as an RNA-binding protein with an R3H domain, has emerged as a pivotal focus in cancer research (Fukushi et al., 2021). Despite being an uncharacterized protein, the role of R3HDM1 in tumor progression, malignant cell regulation, and the tumor immune microenvironment remains largely unexplored. Understanding the significance of R3HDM1 in the context of cancer could potentially offer new insights into its diagnostic, prognostic, and immunotherapeutic value across a spectrum of cancer types. This study aims to comprehensively investigate the functional implications of R3HDM1 in cancer development and progression. By collecting and analyzing data from various databases and conducting *in vitro* cellular experiments, this research seeks to elucidate the association between R3HDM1 expression levels in tumors and their clinical features, metabolic pathways, and immune modulation. The findings from this study are expected to shed light on how elevated R3HDM1 expression correlates with poor prognosis and altered drug sensitivity in different cancer types. Furthermore, exploring the relationship between R3HDM1 expression and immune modulatory molecules as well as lymphocyte subpopulation infiltration biomarkers could provide valuable insights into the role of R3HDM1 in the tumor immune microenvironment. *In vitro* cellular experiments were conducted to validate the differential expression of R3HDM1 between lung adenocarcinoma cell lines and normal lung glandular epithelial cells. Overall, this investigation aims to establish R3HDM1 as a promising molecular marker for predicting cancer patient outcomes and assessing the efficacy of immunotherapeutic interventions, thereby contributing to the development of more effective cancer treatments and potentially novel anti-tumor therapies.

2 Materials and methods

2.1 R3HDM1 multidisciplinary expression analysis in pan-cancer

First, we detected the differences in R3HDM1 between tumor and normal tissues in TCGA data. Then, we examined the differences in R3HDM1 expression between tumor and normal tissues in paired samples from TCGA cancer subgroups. Finally, we expanded the sample size of normal tissues by combining the data from TCGA and GTEx, and conducted differential analysis using the Wilcoxon test ($*P < 0.05$; $**P < 0.01$; $***P < 0.001$; $****P < 0.0001$). The gene expression distribution in various organs was visualized using the gganatogram package. We used the pROC package to assess the potential role and significance of R3HDM1 in pan-cancer diagnosis and validated protein expression in external gene transcriptome levels in GEO database and CPTAC data. Logistic regression was performed to further validate the accuracy of the Wilcoxon test in four cohorts. It is worth noting that we conducted RT-qPCR experiments to verify the differential expression of R3HDM1 in normal lung epithelial cell line BEAS-2B and lung adenocarcinoma cell lines A549 and H1299.

2.2 Somatic copy-number alteration (SCNA), mutation analysis and DNA methylation analysis

We use the cBioPortal website for mutation-related analysis and visualization, including the frequency of changes, mutation types, CNA data, and gene mutation locations for all cancers (Cerami et al., 2012). Somatic cell copy number alterations (SCNA) and mutation can enhance the copy number changes of R3HDM1, with over 5% considered high-frequency SCNA. We calculate the Spearman correlation between the expression value and the copy number segment value of R3HDM1 to evaluate the association between SCNA and R3HDM1 expression. We use the Bioconductor R package “IlluminaHumanMmethylation-450kanno.ilmn12. hg19” to annotate methylation probes for R3HDM1 promoter. We test the differential methylation of R3HDM1 in tumor and normal samples using the Wilcoxon rank sum test and use a significance cutoff value of 0.05 to identify genes with significantly low or high methylation. We calculate the Spearman correlation between gene transcription expression and promoter DNA methylation Beta

values, considering it significant if the P -value is < 0.05 (Mei et al., 2017; Zheng et al., 2019). Additionally, we calculate the Spearman correlation between R3HDM1 and 10 genomic feature scores.

2.3 Analysis of immune cell infiltration and assessment of the anti-cancer immune response

Utilizing TCGA tumor data, we employed seven different methods in TIMER2.0 to estimate the relationship between R3HDM1 expression levels and infiltration levels in various cells within the tumor immune microenvironment (Li T. et al., 2020). To investigate the cell types expressing R3HDM1 in tumor tissues, we performed in-depth single-cell analysis on 140 datasets from the TISCH database and visualized the findings (Han et al., 2023). Additionally, UMAP plots displaying the expression patterns of R3HDM1 in NSCLC_GSE143423 and NSCLC_GSE148071 were obtained from the TISCH database. Subsequently, visualizations were conducted to explore the relationship between R3HDM1 expression in different cells in NSCLC_GSE143423 and NSCLC_GSE148071 and functional pathways. Analysis on the impact of R3HDM1 expression levels on anti-cancer immune status in 32 cancer types was then carried out using the Tumor Immune Phenotype (TIP) database (<http://biocc.hrbmu.edu.cn/TIP>), with scoring performed for each step (Xu et al., 2018). Furthermore, the differences in immune activity scores between the high and low R3HDM1 expression groups were calculated. Visualization of the results was achieved using the R package pHeatmap (v1.0.12).

2.4 Prediction of immunotherapy and immune checkpoint inhibitor responses

We used the EaSIeR model and MeTIL scores to forecast the responses to anti-tumor immune checkpoint inhibitors based on RNA-seq data. EaSIeR is a predictive tool for biomarker-guided immunotherapy, relying on a model rooted in cancer-specific immune responses (Lapuente-Santana et al., 2021). MeTIL scores indicate the infiltration levels of T cells, NK cells, B cells, Tregs, and cytotoxic T lymphocyte (CTL) function (Jeschke et al., 2017).

2.5 Analysis of pathways and functional mechanisms of R3HDM1

We calculated the differential genes between the high-expression group and the low-expression group of R3HDM1 in pan-cancer. If a gene shows differential expression in more than five tumors, we defined it as an R3HDM1-related gene and used the clusterProfiler package for KEGG pathway enrichment analysis to identify the conservative function or biological pathways participated by this gene in pan-cancer (Wu et al., 2021). Research on complex diseases such as cancer has shown that the activity levels of proteins, such as expression and modification, have a significant impact on the occurrence and development of diseases. Changes in protein

TABLE 1 List of siRNAs.

Primer	Sequence (5' to 3')
NC	UUCUCGGAACGUGUCACGUTT
GAPDH	TCTTCTCTTGTGCTCTTG
R3HDM1-1	CGCCAGAUUUUAGAGUUAAU
R3HDM1-2	UGGGAAGUCUGUCAUAGUAAA
R3HDM1-3	AGACUUUCAGAAACGUUAAU
R3HDM1-4	GAGAGCCAGAGACCGAAUAAU

levels and structures have been proven to play a key role in tumor development but are not reflected in genetic changes. We used Spearman correlation analysis to identify the correlation between R3HDM1 expression and protein content identified by the RPPA method from the TCPA database (Li et al., 2017). The Sankey diagram visualized the results with correlation coefficients greater than 0.4 in all tumors. CancerSEA redefined 14 functional states (Yuan et al., 2019). In addition, we collected 14 classic tumor-related pathways from the KEGG database. We used the z-score parameter in the GSVA R package to calculate the gene sets of the 14 functional states and conducted Pearson correlation analysis (Lee et al., 2008). To determine the pathways associated with genes, we divided the samples of each tumor type into two groups, including the top 30% and the bottom 30%, and conducted gene set enrichment analysis (GSEA) to explore the gene set activation or inhibition in the high-expression group compared to the low-expression group in different tumors (Subramanian et al., 2005). Finally, to determine the protein-protein interactions associated with genes, we used the ComPPI database to filter out interactions with non-co-localized subcellular localization, introduced localization scores and interaction scores, and ultimately obtained the proteins that interact with genes (Veres et al., 2015).

2.6 Identification of chemical substances interacting with R3HDM1

We utilized the GSCALite database (<http://bioinfo.life.hust.edu.cn/web/GSCALite/>) to analyze gene expression and drug sensitivity (Liu et al., 2023). GSCALite provides 750 small molecule drugs from GDSC and CTRP, and explores valuable small molecule drugs related to gene expression data. In addition, the National Cancer Institute (NCI) has established the cancer cell line platform, which has been widely used for drug screening based on relevant gene expression. NCI-60 is a collection of 60 human cancer cell lines from nine different cancer types. The NCI-60 expression data comes from CellMiner, and we analyzed the relationship between R3HDM1 expression and drug sensitivity z-scores and calculated Spearman correlation coefficients (Reinhold et al., 2023). Furthermore, we identified the differential expression of genes in different cancer types, and collected 150 most upregulated or downregulated genes as gene-related markers. The CMAP_gene_signatures.RData file contains 1,288 compound-related features, downloaded from the database

TABLE 2 List of primer.

Primer	Sequence (5' to 3')
GAPDH-F	GGCTCTCCAGAACATCATC
GAPDH-R	TCTTCCTCTTGTGCTCTTG
R3HDM1-F	GCAGCACAGATTTCAGACA
R3HDM1-R	ACCAGAAGACTCAGAACCT

website (<https://www.pmgenomics.ca/bhklab/sites/default/files/downloads>) and used for calculating matching scores (Malta et al., 2018). The analysis process followed the methods outlined in previous publications, and the results for 32 cancer types were summarized and graphically presented using the R language (Yang et al., 2022).

2.7 Survival and clinical features analysis of R3HDM1 in pan-cancer

Survival data was retrieved from the TCGA database, and the “survival” and “survminer” R packages were used to analyze the correlation between gene expression and prognostic indicators [including overall survival (OS), disease-specific survival (DSS), progression-free interval (PFI), and disease-free interval (DFI)]. We combined two methods, Kaplan-Meier and univariate Cox analysis, to comprehensively determine whether the gene is a risk factor or protective factor, and ultimately created a gene survival map. It is worth noting that when using the Kaplan-Meier method for survival analysis, the best cutoff values for high and low expression queues of R3HDM1 were determined using the R package “survminer,” and the survfit function was used to perform a log rank test to evaluate the significance of high and low expression groups of the gene. In addition, the “forestplot” package was used to visualize the results of Cox analysis of survival data. Furthermore, we used the Wilcoxon rank-sum test and Kruskal-Wallis Rank Sum Test to detect gene expression in different stages and molecular subtypes.

2.8 Construction of cell lines expressing R3HDM1 siRNA

H1299 and A549 cells were acquired from Wuhan Sevilla Biotechnology Company, while BEAS-2B cells were obtained from Hunan Prattze Biotechnology Company. The cells were cultured in a CO₂ incubator maintained at 37°C with a 5% CO₂ atmosphere. Sterile distilled water, sterilized under high-temperature and high-pressure conditions, was regularly replaced in the incubator to maintain appropriate humidity levels. The targeting sequences of R3HDM1 siRNA and the sequence of the negative control siRNA can be found in Table 1. H1299 and A549 cells were seeded into separate culture dishes and transfected when the cell density reached 60%, with mRNA expression levels evaluated after 24 h of standard cultivation.

2.9 RT-PCR and RT-qPCR

RNA was extracted utilizing the TRIzol reagent (Invitrogen Life Technologies; Carlsbad, CA, United States), followed by treating RNA samples with the gDNA remover reagent (GenStar) according to the manufacturer’s protocol to eliminate genomic DNA. Subsequently, cDNA synthesis was performed using the RT-Phusion kit (Thermo Fisher Scientific; Waltham, MA, United States). Gene-specific mRNA levels were quantified through standard and quantitative RT-PCR (RT-qPCR) techniques, employing the $\Delta\Delta C_t$ method. The primer sequences are available in Table 2.

2.10 Detection of cell migration and proliferation

After transfecting A549 and H1299 cells, they were individually seeded into six-well plates. Once reaching approximately 90% confluency, vertical scratch wounds were created using a 200 μ L pipette tip. The scratches were photographed every 12 h, and the migration rate was analyzed using ImageJ software. Suspensions of transfected A549 and H1299 cells were prepared for cell counting purposes. Each well of a 96-well plate was seeded with 1×10^4 cells. Subsequently, after cell adhesion, the absorbance at 450 nm was measured every 24 h upon the addition of CCK-8 reagent to calculate the proliferation rate between the negative control group and the treatment group.

3 Results

3.1 Expression of R3HDM1 in pan-cancer

By combining and mining the resources of TCGA and GTEx databases, we obtained the mRNA expression levels of R3HDM1 from a pan-cancer perspective. Firstly, compared to normal tissues in the TCGA database, we observed significantly higher expression of R3HDM1 in most cancer tissues, including breast cancer (BRCA), Colon adenocarcinoma (COAD), Esophageal carcinoma (ESCA), Head and Neck squamous cell carcinoma (HNSC), Liver hepatocellular carcinoma (LIHC), Lung adenocarcinoma (LUAD), Lung squamous cell carcinoma (LUSC), Pheochromocytoma and Paraganglioma (PCPG), Rectum adenocarcinoma (READ), Stomach adenocarcinoma (STAD), and Uterine Corpus Endometrial Carcinoma (UCEC), among which the differential expression level in lung adenocarcinoma is significantly significant (Figure 1A). Significant differences in expression levels of LUAD were also observed in paired tissues (Figure 1B). Due to the limited number of normal samples in the TCGA data, we combined the GTEx database to further expand the sample size, and obtained consistent results. The significance of the differential expression level in lung adenocarcinoma has not changed (Figure 1C).

Next, we visualize the expression distribution pattern of R3HDM1 using organ graphs (Figure 2A). It is evident that, except for the testis, thyroid, brain, and kidney, cancer tissues show significantly higher expression of R3HDM1 compared to

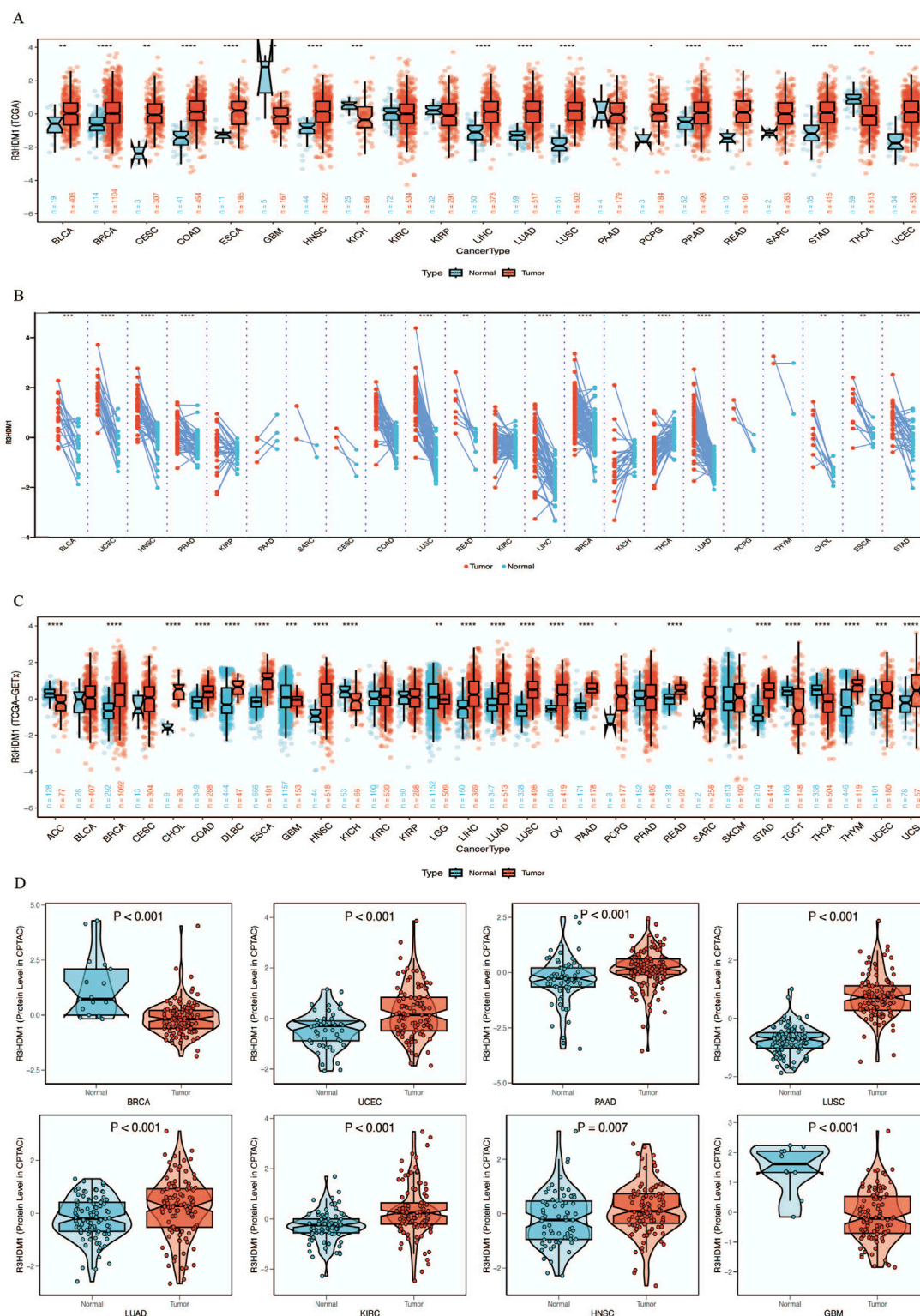


FIGURE 1 mRNA expression levels of R3HDM1 in human normal and tumor tissues. **(A)** mRNA expression levels of R3HDM1 in normal and cancer tissues using data from the TCGA database. **(B)** mRNA expression levels of R3HDM1 in paired tissues using data from the TCGA database. **(C)** The expression level of R3HDM1 in different tumor tissues and corresponding normal tissues from TCGA and GTEx datasets. **(D)** External validation of protein levels was performed using a CPTAC database. The symbols *, **, and *** indicate $P < 0.05$, $P < 0.01$, $P < 0.001$, respectively.

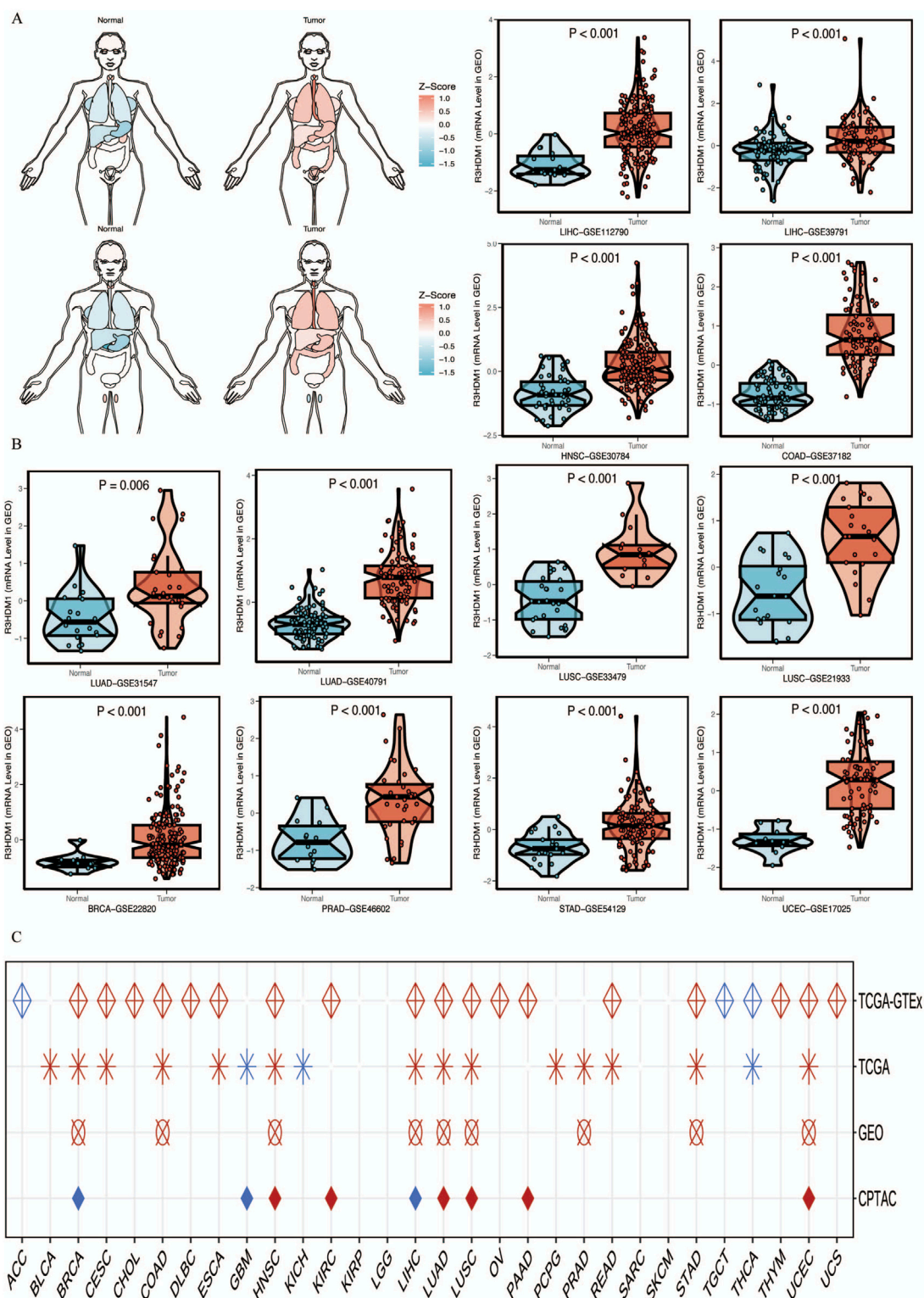


FIGURE 2 mRNA expression levels of R3HDM1. **(A)** Expression and distribution of R3HDM1 in various organs. **(B)** External validation of mRNA levels was performed using GEO database. **(C)** Logistic regression analysis of TCGA, TCGA-GTEX. Red means OR is greater than 1, blue represents an OR value between 0 and 1. A white circle means that there is no significance, and an empty value means that no relevant data set has been collected in the current database. Tumor tissues are represented by red, while normal tissues are represented by blue.

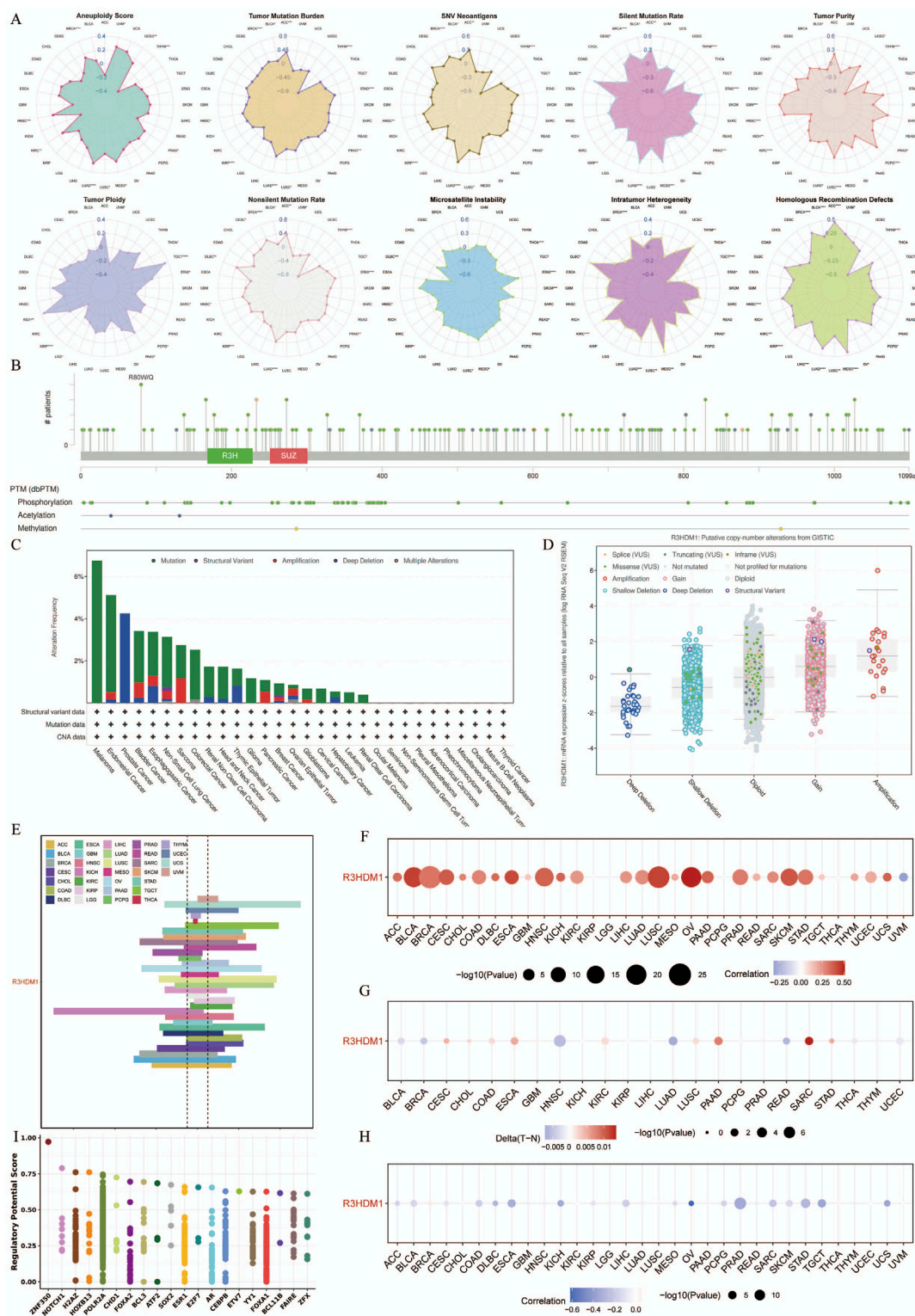


FIGURE 3 Genetic Alterations of R3HDM1 in cancers. **(A)** Radar map visualization of spearman correlation coefficients of R3HDM1 and 10 genomic features. Statistical significance was reported at * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$. **(B)** Sites and number of cases with R3HDM1 genetic alterations in pan-cancer from cBioPortal. **(C)** Frequency of R3HDM1 mutations in different tumor types. **(D)** Relationship between R3HDM1 mRNA expression and genetic alterations. **(E)** Histogram shows the frequency of somatic copy number alterations for R3HDM1 in each cancer type. **(F)** The Spearman's correlation between somatic copy number alterations and the expression of R3HDM1. **(G)** Heatmap shows the differential methylation of R3HDM1 in cancers; hypermethylated and hypomethylated R3HDM1 are marked in red and blue, respectively (Wilcoxon rank-sum test). **(H)** Spearman's correlation of R3HDM1 between transcriptional expression and promoter methylation. Red and blue represent positive and negative correlations, respectively. **(I)** Top 20 factors are showed in this plot. Y-axis represents the RP score. X-axis represents different factors. Dots in a x-axis line means same factor.

normal tissues in other organs, with lung cancer tissues still exhibiting significantly higher expression than normal lung tissue. We found differential expression of R3HDM1 in most types of cancer, and consistent expression patterns across cancer types, with a universal significant upregulation. External validation of mRNA levels was conducted using the GEO database (Figure 2B) and protein levels using the CPTAC database (Figure 1D). Logistic regression analysis based on TCGA, TCGA-GTEX, GEO, and CPTAC datasets fully verified the above results, with good consistency in expression trends across different omics and databases (Figure 2C). The results related to LUAD further confirm the presence of R3HDM1 as a risk factor, warranting further research. In fact, after expanding the sample size of normal tissues by combining the GTEX database, the estimated ROC curve indicates that R3HDM1 demonstrates satisfactory sensitivity and specificity in the diagnosis of 17 cancer types (area under the curve >0.7, Supplementary Figures 1, 2). The reproducibility and consistency of these results have been demonstrated across multiple databases, tumor types, various methods, and omics analyses, suggesting that the dysregulation of R3HDM1 expression may be involved in different cancers and is unlikely to be attributed to technical artifacts, randomness, or biases in sample identification standards within databases.

3.2 Genetic variants of R3HDM1 in cancer

We obtained genomic data of tumors and normal tissues from the TCGA Pan-Cancer dataset, including genetic variations, somatic copy number alterations (SCNA), mRNA expression, and DNA methylation data. It is worth noting that Figure 3A shows the correlation between R3HDM1 and ten genomic features scores, so we further calculated the relationship between R3HDM1 and different mutation scores, especially the importance of SNV Neoantigens, TMB, and MSI as important biomarkers for immunotherapy. The high SNV Neoantigens score of R3HDM1 expression is positively correlated with Adrenocortical carcinoma (ACC), Bladder Urothelial Carcinoma (BLCA), BRCA, HNSC, Kidney renal papillary cell carcinoma (KIRP), LUSC, LUAD, Prostate adenocarcinoma (PRAD), and Thymoma (THYM). Similarly, the expression of R3HDM1 is positively correlated with high TMB scores in ACC, BLCA, BRCA, HNSC, KIRP, LUSC, LUAD, PRAD, THYM, and STAD. Additionally, the high MSI scores in Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC), HNSC, Mesothelioma (MESO), KIRP, LUSC, Thyroid carcinoma (THCA), Rectum adenocarcinoma (READ), Skin Cutaneous Melanoma (SKCM), and STAD are positively correlated with R3HDM1 expression. Simultaneously, R3HDM1 expression is significantly positively correlated with the high score of LUAD in eight scoring methods, indicating that R3HDM1 can be used as a predictive marker for the effectiveness of immunotherapy in LUAD. Figure 3B depicts the 2D structure of the R3HDM1 mutation site. Analysis using the cBioPortal database, we found that R3HDM1 exhibits certain genetic alteration frequencies in most cancers, with amplification and mutation being the most common types of genetic alterations (Figure 3C). The mutation frequencies of R3HDM1 are different in different cancers, with the highest mutation frequencies in melanoma, endometrial cancer, prostate cancer, bladder cancer, esophageal

gastric cancer, and non-small cell lung cancer. The mutation patterns of R3HDM1 vary in different tumors, predominantly mutations in most tumors, but copy number loss is predominant in prostate cancer and thymic epithelial tumors. At the same time, a significant increase in R3HDM1 expression with the progression of SCNA from copy number loss to amplification (Figure 3D). Clearly, SCNA played a crucial role in regulating gene expression in cancer. We next found SCNA occurs at a high frequency (exceeding 5% of all samples) in most cancer types, with very low frequencies only in THCA and THYM (Figure 3E). Next, we evaluated the impact of SCNA on R3HDM1 expression, and gained the Spearman correlation between R3HDM1 expression and copy number across different cancers in TCGA (Figure 3F). The results indicate a significant positive correlation between R3HDM1 expression and SCNA in most tumors. To further study the genetic alterations of R3HDM1 in pan-cancer, we examined the percentage of SCNA. Then, we observed diverse methylation patterns of R3HDM1 in the pan-cancer dataset (Figure 3G), with lower methylation in tumor tissues compared to normal tissues in HNSC, LUAD, and READ, while higher methylation levels were observed in tumor tissues in Sarcoma (SARC), Pancreatic adenocarcinoma (PAAD), Esophageal carcinoma (ESCA), and COAD. After all, we obtained a negative correlation between the expression of R3HDM1 and DNA methylation (Figure 3H). In addition, we showed that the top three transcription factors with the highest RP scores are ZNF350, NOTCH1, and H2AZ (Wang et al., 2013) (Figure 3I).

3.3 The correlation between R3HDM1 and immune microenvironment in cancer

The interaction between tumor cells and the tumor microenvironment is crucial for tumor development, progression, metastasis, and treatment response. Analysis of the correlation between R3HDM1 expression and immune infiltration levels in TCGA tumor profiles using seven different algorithms revealed a significant positive association of R3HDM1 expression with B cells, CD4⁺T cells, CD8⁺T cells, and Treg cells in nearly all cancer types. This suggests that R3HDM1 plays a role in immune rejection or immune silence, highlighting its critical involvement in the interactions between the immune system and tumors, especially in immune evasion (Figure 4A). To further study the cell types expressing R3HDM1 in tumor tissues, we analyzed single-cell expression levels of R3HDM1 (Supplementary Figure 7), finding widespread expression in various immune and malignant cells. For instance, UMAP plots and heat maps demonstrate a significant positive correlation of R3HDM1 expression with CD8⁺T cells and proliferative T cells (T-prolif) in non-small cell lung cancer (NSCLC_GSE143423 and NSCLC_GSE148071), indicating its significant role in cellular proliferation, cell death, and mitochondrial energy metabolism pathways (Figures 4B, C). Additionally, R3HDM1 demonstrates a reverse expression pattern between T-prolif and exhausted CD8⁺T cells in multiple cancer tissues, suggesting its potential regulatory role in T cell function (Supplementary Figure 7).

Given the regulatory role of R3HDM1 expression levels in the immune response and immune cell infiltration across various

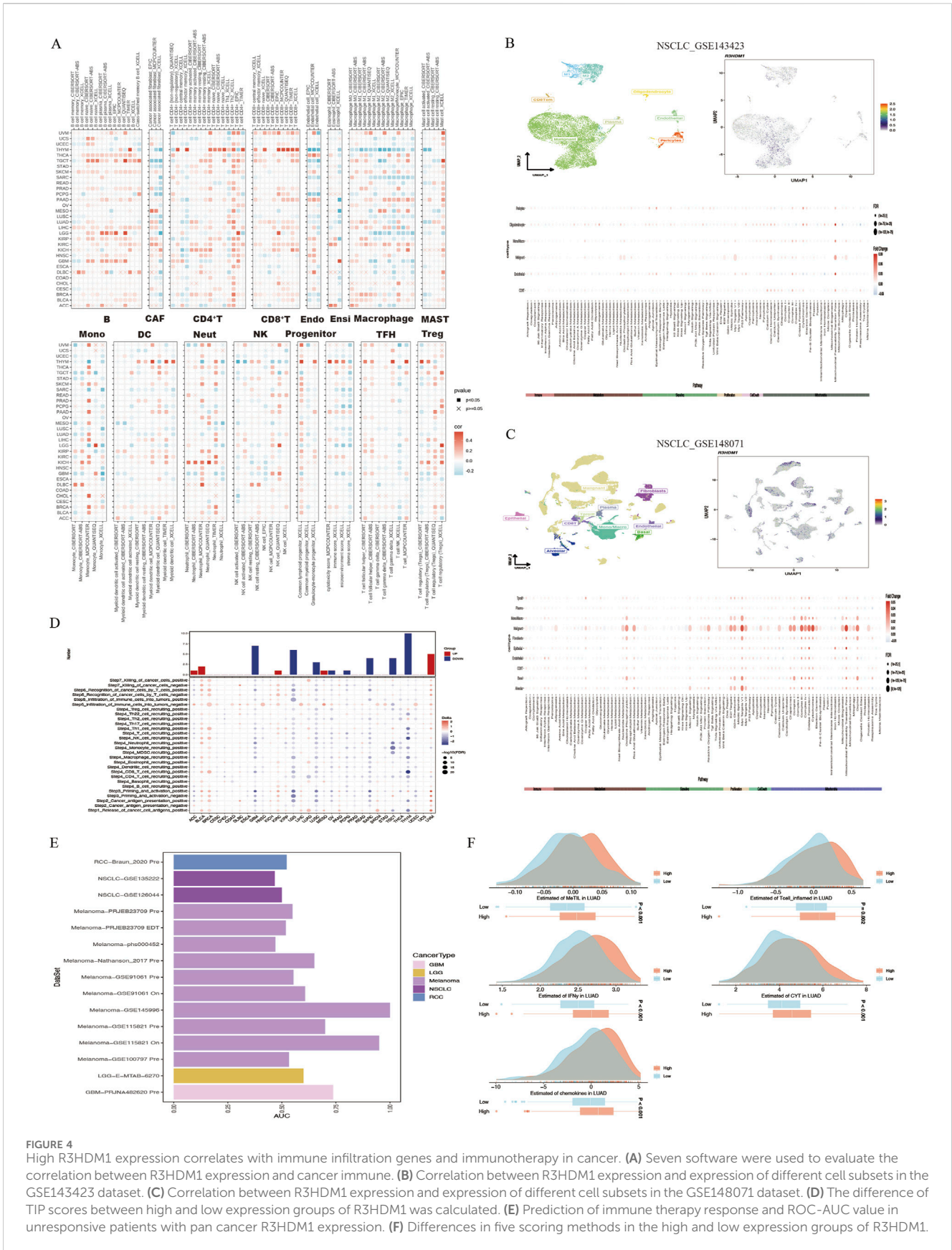


FIGURE 4 High R3HDM1 expression correlates with immune infiltration genes and immunotherapy in cancer. **(A)** Seven software were used to evaluate the correlation between R3HDM1 expression and cancer immune. **(B)** Correlation between R3HDM1 expression and expression of different cell subsets in the GSE143423 dataset. **(C)** Correlation between R3HDM1 expression and expression of different cell subsets in the GSE148071 dataset. **(D)** The difference of TIP scores between high and low expression groups of R3HDM1 was calculated. **(E)** Prediction of immune therapy response and ROC-AUC value in unresponsive patients with pan cancer R3HDM1 expression. **(F)** Differences in five scoring methods in the high and low expression groups of R3HDM1.

human cancers from our previous analyses, we proceeded to investigate the predictive role of R3HDM1 in cancer immunotherapy. To assess this, we downloaded data from the TIP database and evaluated the activity scores of the cancer immunity cycle using the GSVA algorithm, discovering a positive correlation between high R3HDM1 expression and the release of cancer cell antigens (step 1), recruitment of Th1 cells (step 4), and T cell recognition of cancer cells (step 6). However, in most cancer types, a negative correlation was observed between high R3HDM1 levels and the recruitment of CD4⁺T cells (step 4), recruitment of CD8⁺T cells (step 4), and immune cell infiltration into tumors (step 5), further indicating the immunosuppressive role of R3HDM1 in the TME (Figure 4D).

Subsequently, from a pan-cancer perspective on the response to immunotherapy, we visualized the predictive capability of R3HDM1 for predicting responders and non-responders to pan-cancer immunotherapy (Figure 4E). We found varied predictive performance of R3HDM1 across different datasets, with the best performance observed for Melanoma and good performance for RCC, NSCLC, LGG, and GBM. Finally, we used MeTIL scores, CYT scores, IFN- γ scores, Tcell_inflamed scores, and Chemokines scores to predict the response of LUAD patients receiving immune checkpoint inhibitor (ICB) therapy in the LUAD patient dataset. We observed that LUAD patients with high R3HDM1 expression exhibited a significantly favorable response to ICB therapy, indicating the effective predictive ability of R3HDM1 for the response rate of LUAD patients to immune therapy (Figure 4F).

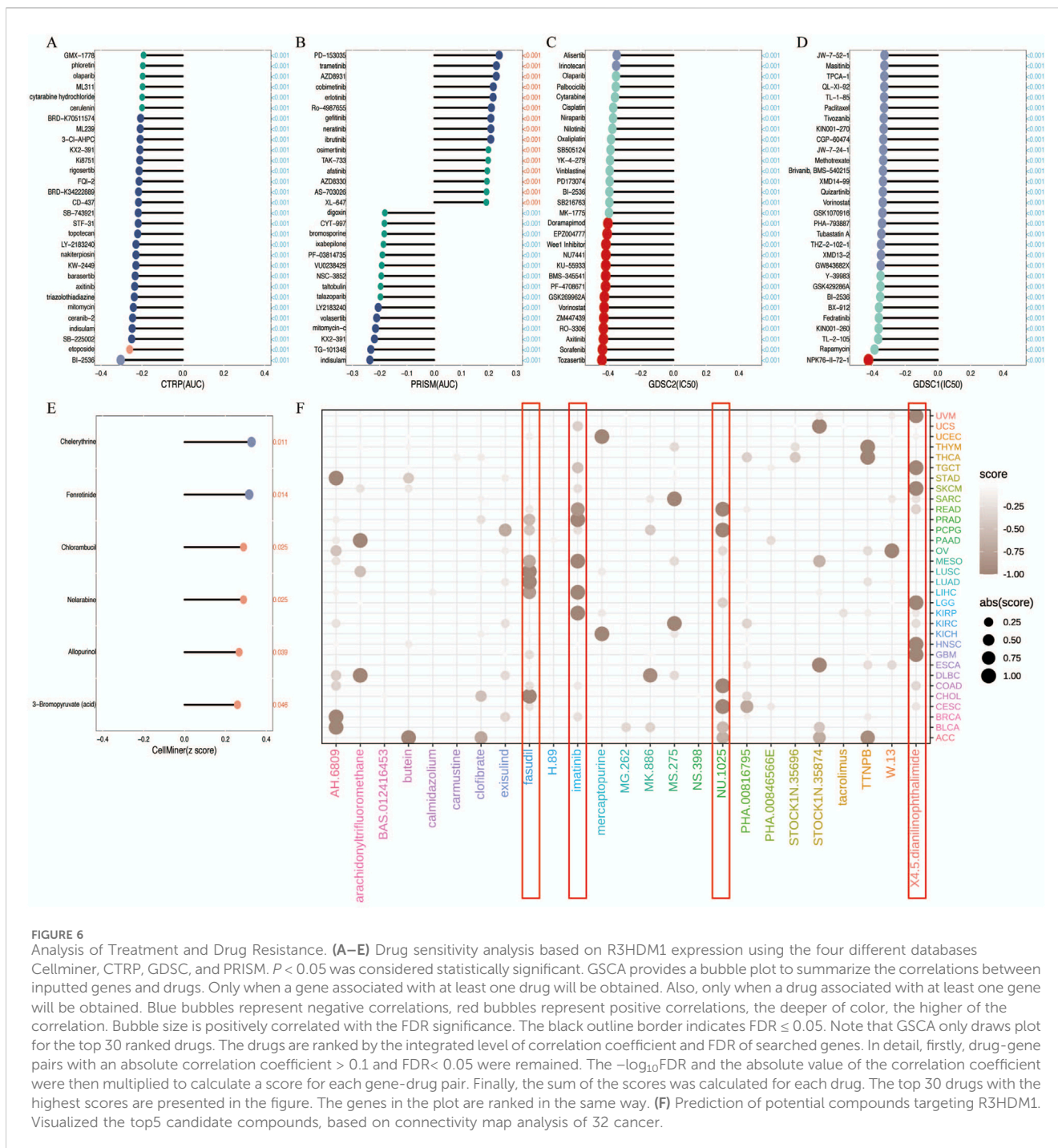
3.4 The relationship between R3HDM1 and pathways

To further understand the impact of R3HDM1 on the prognosis of cancer patients, especially LUAD, we conducted Gene Set Enrichment Analysis (GSEA). Firstly, we examined the differentially expressed genes between high and low R3HDM1 patients in each cancer type. The result shows that in the high R3HDM1 expression group of almost all cancer types, cell proliferation-related signaling pathways (including MYC, mTORC1, spindle, G2M, and E2F pathways) are significantly enriched (Figure 5A). This finding corroborates previous findings of sustained high expression of R3HDM1 in malignant cells, indicating its crucial role in tumor cell proliferation. Furthermore, in BLCA, KICH, KIRC, KIRP, LUAD, SKCM, and MESO patients, high expression of R3HDM1 is positively correlated with epithelial-to-mesenchymal transition (EMT). This correlation may explain why patients with high levels of R3HDM1 in SKCM cancer type are prone to distant metastasis. Interestingly, the enrichment of the EMT pathway is not significant in patients with high levels of R3HDM1 in THCA and TGCT, which is consistent with the analysis of staging status in these two types of cancer, where the expression levels of R3HDM1 in stages I and II are significantly higher than in stages III and IV (Supplementary Figure 5). In addition, the heatmap also highlights differential enrichment of immune-related pathways, such as interferon (IFN)- γ , IFN- α , inflammation, IL-6, IL-12, complement, and allograft rejection pathways (Figure 5A). These pathways are negatively enriched in most cancer patients with high R3HDM1 expression levels, including CESC, DLBC, ESCA, GBM, HNSC, LAML, LGG, LUSC, MESO, OV, PCPG, SARC, and UCS. These findings suggest that

R3HDM1 may be involved in inhibiting the anti-tumor immune response in these cancers. It is worth noting that, for LUAD patients, high levels of R3HDM1 are also associated with immune response, IL-6, IL-12, and complement pathways, albeit to a moderate degree. In summary, these results strongly indicate that the elevated levels of R3HDM1 are closely associated with increased proliferation, EMT, and immune suppression in human cancers. Additionally, we analyzed the relationship between the scores of 14 cancer-related pathways in LUAD and R3HDM1 (Figure 5B). Obviously, the scores related to the cell cycle are higher and positively correlated than other scores, which further confirms the sustained expression of R3HDM1 in malignant cells, indicating its involvement in tumor cell proliferation. Moreover, using comPPI, helped us identify the related genes that may interact with R3HDM1 and represented the interaction sites with different colored lines, suggesting that R3HDM1 plays a wide range of roles through the networks (Figure 5C). At the same time, we found in the high R3HDM1 expression group of LUAD, pathways related to the cell cycle, cell proliferation (including MYC, mTORC1, spindle, G2M, and E2F pathways), and epithelial-mesenchymal transition (EMT) were significantly enriched (Figure 5D). This finding provides additional evidence that R3HDM1 is consistently highly expressed in malignant LUAD cells, and is strongly associated with tumor progression, proliferation, and metastasis. Meanwhile, we investigated proteins in the TCPA database with a correlation of over 0.35 with R3HDM1 and visualized them. We found a positive correlation between the expression levels of R3HDM1 and various functional proteins in LUAD (Figure 5E). Specifically, the expression of R3HDM1 in LUAD is positively correlated with the functional proteins TFRC, MTOR_pS2448, and X4EBP1, which are known to be involved in key cellular processes such as metabolism, apoptosis, signaling, cell cycle regulation, and proliferation.

3.5 The association between R3HDM1 and chemotherapy

In the chemotherapeutic analysis, we investigated the potential correlation between drug sensitivity and R3HDM1 expression using four different databases (Figures 6A–E). It illustrates a substantial negative correlation between the sensitivity AUC and IC50 of numerous drugs across the CTRP, PRISM, and GDSC databases. This suggests that elevated R3HDM1 expression renders cells highly responsive to a wide array of drugs. Conversely, in the Cellminer database, R3HDM1 showed a significant positive correlation with the z scores of all drugs. Clearly, R3HDM1 emerges as a potential chemosensitive gene. To explore potential therapeutic strategies that can inhibit the tumor-promoting effects mediated by R3HDM1, we conducted CMap analysis. We constructed a set of R3HDM1-related gene signature, comprising the top 150 upregulated and 150 downregulated genes, determined by comparing patients with high and low gene expression within each cancer type. Utilizing the optimal feature matching method XSum (eXtreme Sum), we compared the gene-related features with CMap gene features, resulting in similarity scores for 1,288 compounds. From the result, fasudil, imatinib, NU1025, and 4,5-Dianilinophthalimide exhibited significantly lower scores across most cancer types, suggesting their potential to inhibit R3HDM1-



mediated oncogenic effects (Figure 6F). Surprisingly, previous studies have demonstrated the anti-tumor effects of fasudil, imatinib, and 4,5-Dianilinophthalimide. These findings substantiate the validity of our predictive outcomes, although further research is warranted to elucidate the underlying mechanisms.

3.6 Clinical relevance of R3HDM1

The study explored the clinical relevance of R3HDM1 in cancer by analyzing its role in cancer survival. Survival analysis across various

cancer types revealed a consistent association of R3HDM1 with multiple survival periods, primarily as a risk factor for various cancers, notably in LUAD. While R3HDM1 typically acts as a risk factor at different survival stages in most tumors, exceptions exist where it serves as a protective factor; for instance, higher R3HDM1 expression is linked to better survival rates in LGG, THYM, and similar conditions. These results underscore the diverse role of R3HDM1 in different cancers, highlighting the need for further investigation into its functional significance in cancer survival. Clustering results (Figure 7A) underscored R3HDM1's prognostic value for 11 tumor types, including ACC, BRCA, KICH, KIRC, KIRP, LUAD, MESO, THCA, UCEC, LIHC, and UVM.

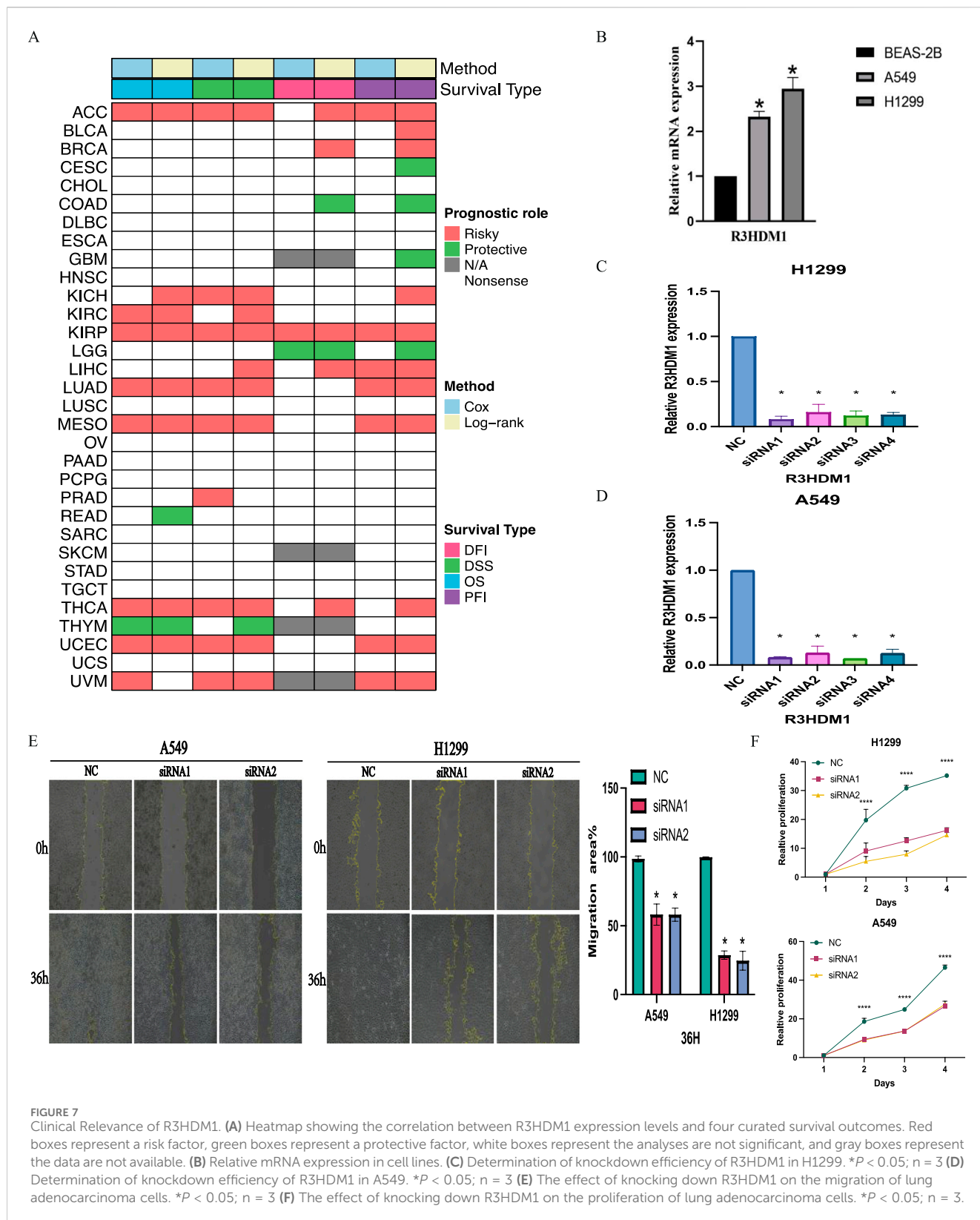


FIGURE 7

Clinical Relevance of R3HDM1. (A) Heatmap showing the correlation between R3HDM1 expression levels and four curated survival outcomes. Red boxes represent a risk factor, green boxes represent a protective factor, white boxes represent the analyses are not significant, and gray boxes represent the data are not available. (B) Relative mRNA expression in cell lines. (C) Determination of knockdown efficiency of R3HDM1 in H1299. * $P < 0.05$; $n = 3$ (D) Determination of knockdown efficiency of R3HDM1 in A549. * $P < 0.05$; $n = 3$ (E) The effect of knocking down R3HDM1 on the migration of lung adenocarcinoma cells. * $P < 0.05$; $n = 3$ (F) The effect of knocking down R3HDM1 on the proliferation of lung adenocarcinoma cells. * $P < 0.05$; $n = 3$.

Moreover, forest plots and Kaplan-Meier curves (Supplementary Figures 3, 4) presented the Cox survival rate analysis and log-rank test outcomes across various cancers. We conducted further investigation into the correlation between R3HDM1 expression and tumor staging, revealing

its association with multiple tumor stages (Supplementary Figure 5), implying a link between R3HDM1 and the progression of these tumors. Notably, R3HDM1 displayed variations across various molecular subtypes in different cancers (Supplementary Figure 6). For instance,

it exhibited the lowest expression in the LumA subtype of BRCA, higher expression in the Basal subtype, and significantly higher expression in Triple-Negative Breast Cancer (TNBC). In LUSC, its expression was lowest in the Basal subtype and highest in the Primitive subtype. The expression across different subtypes in LUAD varied significantly. These findings underscore the precision molecular stratification, therapeutic implications, and prognostic value of R3HDM1 in diverse cancers. At last, the normalized RT-qPCR results show substantial variations in R3HDM1 expression levels between normal lung glandular epithelial cells and lung adenocarcinoma cells (Figure 7B). We assessed the efficiency of R3HDM1 knockdown in A549 and H1299 cells post transfection with small interfering RNA (siRNA) using RT-qPCR. The results indicated significant transfection efficiencies for all four siRNA sequences, with siRNA1 and siRNA3 showing the most effective knockdown efficiency ($P < 0.05$) (Figures 7C, D). Subsequently, siRNA1 and siRNA3 were selected for further cellular experiments. Following this, we conducted cell scratch and CCK8 cell viability assays to examine the impact of R3HDM1 knockdown on the migration and proliferation abilities of A549 and H1299 cells. The results demonstrated a significant decrease in migration ability in cell lines with R3HDM1 knockdown compared to the negative control group ($P < 0.05$) (Figure 7E). Furthermore, after 24 h of R3HDM1 knockdown expression, the proliferation capabilities in the treatment group exhibited a significant decrease compared to the negative control group ($P < 0.0001$). The experimental findings are in line with the aforementioned prognosis analysis (Figure 7A) and GSEA results (Figure 5A). In conclusion, our study highlights the significant role of R3HDM1 in the development and progression of lung adenocarcinoma.

4 Discussion

R3HDM1 is an uncharacterized RNA-binding protein containing an R3H domain. Previous studies have suggested R3HDM1 as a potential candidate oncogene (Tuupanen et al., 2014). Additionally, ARPP21 (also known as R3HDM3), belonging to the same protein family as R3HDM1, has been implicated as a potential key oncogene in tumors (Fukushi et al., 2021; He et al., 2021). Research has shown that ARPP21 can interact with mutant p53, leading to the inhibition of miR-128-2 targeting E2F5 protein, thereby enhancing the impact of mutant p53 on tumorigenesis, progression, and chemoresistance (He et al., 2021; Donzelli et al., 2012). However, the relationship between R3HDM1 of the same protein family and cancer remains unexplored.

This study, for the first time, comprehensively elucidated the dysregulated expression of R3HDM1 in pan-cancer using diverse tools, suggesting its potential as a diagnostic biomarker across multiple cancers. Additionally, it systematically analyzed the relationship between R3HDM1 expression and clinical outcomes in pan-cancer, identified potential drugs and small molecules targeting R3HDM1 from various dimensions, delineated cell populations expressing R3HDM1 at high resolution through extensive single-cell datasets, characterized R3HDM1-mediated pathways and metabolic disruptions, identified transcription factors regulating its expression through chip-seq analysis, associated R3HDM1 expression with various molecular subtypes, highlighting the potential for stratified precision therapy and exploring its underlying mechanisms. Our findings indicated upregulated R3HDM1 expression in most tumors, including

LUAD (Figure 2C). Moreover, significantly increased protein expression of R3HDM1 was noted in LUAD (Figure 2C). The study also revealed correlations between R3HDM1 expression and pathological staging in LUAD, ESCA, LIHC, PAAD, SKCM, and THCA (Supplementary Figure 5). Overall, these results strongly suggest that R3HDM1 may exhibit specific or contrasting functions across different tumor types, particularly in LUAD.

To further elucidate the functionality of R3HDM1 in various types of tumors, this study investigated the correlation between R3HDM1 expression and patient prognosis. It was observed that the expression levels of R3HDM1 were significantly correlated with the prognosis of 19 types of cancer when analyzing the prognostic predictive abilities using Overall Survival (OS), Disease-Specific Survival (DSS), Disease-Free Interval (DFI), and Progression-Free Interval (PFI) (Supplementary Figures 3A–D). High expression of R3HDM1 was identified as a risk factor for the occurrence and progression of tumors, particularly in six types of cancers (ACC, LUAD, MESO, UCEC, THCA, and KIRP). Additionally, the association between R3HDM1 expression levels and clinical characteristics in different types of cancers was evaluated. Consistent with the survival analysis results, R3HDM1 expression was correlated with staging in eight tumors, including LUAD, implying its involvement in the progression of these tumors (Supplementary Figure 5). The higher expression of R3HDM1 in the metastatic group in melanoma further indicates its role in tumor metastasis. Interestingly, differences in R3HDM1 expression were found across various molecular subtypes of 17 cancers, including LUAD, suggesting its involvement in tumor progression and its potential for precise molecular stratification therapy and predictive value (Supplementary Figure 6). These findings were further supported by the results of Gene Set Enrichment Analysis (GSEA) (Figure 5A). In summary, these results confirm that the expression levels of R3HDM1 can serve as a reliable biomarker for predicting the prognosis of patients with tumors, including LUAD.

Ten scoring methods correspond to reliable biomarkers for the prognosis of various cancers, serving as predictive factors for the effectiveness of many tumor immunotherapies, particularly TMB, MSI, and SNV neoantigens (Darragh et al., 2018; Chen et al., 1992; Capietto et al., 2022). Tumors with high TMB scores, high MSI scores, and high SNV neoantigens scores demonstrate better responses to immunotherapy (Reck et al., 2019; Repáraz et al., 2022; Chang et al., 2018; Li K. et al., 2020). Interestingly, the expression of R3HDM1 in KIRP, HNSC, and LUSC positively correlates with SNV neoantigens, TMB, and MSI (Figure 3A). Therefore, we speculate that high R3HDM1 expression in KIRP, HNSC, and LUSC may lead to greater survival benefits post-immunotherapy, suggesting R3HDM1 as a potential novel drug target for anticancer immunotherapy.

The tumor microenvironment (TME) comprises diverse cells and serves as the locale for tumor cell, invasive immune cell, and stromal cell accumulation (Wu and Dai, 2017). In recent years, advancements in immunotherapy have unveiled the significant association between immune cell infiltration and tumor initiation, progression, and treatment (Wu and Dai, 2017; Sadeghi Rad et al., 2021). Consequently, exploring the link between R3HDM1 and immunity is imperative. Our findings unveil the multifaceted involvement of R3HDM1 in immune microenvironment remodeling across distinct tumors. Remarkably, we observed a

significant positive correlation between R3HDM1 expression and B cells, CD4⁺T cells, CD8⁺T cells, and Treg cells in nearly all cancer types, hinting at R3HDM1's partial participation in immune evasion or immune tolerance, pivotal in immune-tumor interplay, notably immune escape. Analyses of immune regulatory function revealed associations between R3HDM1 expression and diverse tumor immune scores, indicating a potential role for R3HDM1 in modulating the tumor microenvironment. Subsequent scrutiny unveiled a positive correlation between R3HDM1 expression and Th2 and TFH cell infiltration, while displaying inverse relationships with central and effector memory CD4⁺T and NKT cells (Figure 4A). Th2 cells, recognized for their tumor-promoting and immunosuppressive functions, were notably highlighted (Liu et al., 2019; Frafjord et al., 2021). Conversely, CD4⁺T cells and NKT cells, pivotal in tumor eradication, showcased reduced infiltration in high R3HDM1 tumors, fostering an immunosuppressive tumor backdrop (Liu et al., 2021; Borst et al., 2018). Furthermore, we assessed the anti-cancer immune status across the seven stages of the cancer immune cycle (Figure 4D). Notably, in most cancer types, heightened R3HDM1 levels exhibited a negative correlation with immune cell infiltration within tumors and antigen-presenting molecule expression, contrasting with a positive correlation observed with various immune negative regulatory molecules. These observations align with the results from Gene Set Enrichment Analysis (GSEA), endorsing R3HDM1's role in mediating immune suppression and impacting the efficacy of immunotherapy within the tumor microenvironment. Analyzing an extensive corpus of single-cell data revealed that R3HDM1 exhibits high expression levels in malignant cells and T-related cells (Supplementary Figure 7). These findings underscore the close association between R3HDM1 expression and immune cell infiltration in tumors, influencing tumor onset, progression, and patient prognosis, thus offering novel targets for immunosuppressant development.

5 Conclusion

Multi omics studies have elucidated the dysregulation of R3HDM1 expression in pan cancer, which is a potential diagnostic marker for pan cancer and has been confirmed in lung adenocarcinoma cell lines. Elaborated on the pathways and metabolic disorders mediated by R3HDM1, and identified the transcription factors regulating its expression using chip-seq. By combining a large number of single-cell datasets, the cells expressing R3HDM1 were identified at high resolution. Clarified the possibility of targeting R3HDM1 from multiple dimensions, drugs and small molecules. The system describes the relationship between R3HDM1 and clinical outcomes in pan cancer. Linking R3HDM1 expression with multiple molecular subtypes demonstrates the potential for stratified precision therapy.

Data availability statement

Data for the TCGA pan-cancer cohort were obtained from the Firehose database (<http://gdac.broadinstitute.org>), comprising both

tumor and normal samples. MiRNA, TCGA, mutation data, molecular subtypes, and clinical information were downloaded from the UCSC Xena database (<https://xenabrowser.net/datapages/>). External validation data for mRNA levels were sourced from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>), while protein-level mass spectrometry data were retrieved from the CPTAC database (<https://proteomics.cancer.gov/programs/cptac>). Pan-cancer immune infiltration results, assessed using various algorithms, were obtained from the TIMER2.0 database (<http://timer.cistrome.org/>). It is important to highlight that these public databases are accessible to all users at no cost. This study strictly adheres to the data extraction policies of each database and therefore does not necessitate ethical review and approval from an ethics committee.

Ethics statement

Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

Author contributions

JL: Writing—original draft. ZB: Writing—review and editing. JW: Writing—review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by Gansu Laboratory of Isotope Phase I Project (0300000006).

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/fgene.2024.1404348/full#supplementary-material>

References

- Borst, J., Ahrends, T., Bąbala, N., Melief, C. J. M., and Kastenmüller, W. (2018). CD4⁺ T cell help in cancer immunology and immunotherapy. *Nat. Rev. Immunol.* 18, 635–647. doi:10.1038/s41577-018-0044-0
- Capietto, A.-H., Hoshyar, R., and Delamarre, L. (2022). Sources of cancer neoantigens beyond single-nucleotide variants. *Int. J. Mol. Sci.* 23, 10131. doi:10.3390/ijms231710131
- Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2, 401–404. doi:10.1158/2159-8290.CD-12-0095
- Chang, L., Chang, M., Chang, H. M., and Chang, F. (2018). Microsatellite instability: a predictive biomarker for cancer immunotherapy. *Appl. Immunohistochem. Mol. Morphol.* 26, e15–e21. doi:10.1097/PAL.0000000000000575
- Chen, L., Ashe, S., Brady, W. A., Hellström, I., Hellström, K. E., Ledbetter, J. A., et al. (1992). Costimulation of antitumor immunity by the B7 counterreceptor for the T lymphocyte molecules CD28 and CTLA-4. *Cell* 71, 1093–1102. doi:10.1016/s0092-8674(05)80059-5
- Darragh, L. B., Oweida, A. J., and Karam, S. D. (2018). Overcoming resistance to combination radiation-immunotherapy: a focus on contributing pathways within the tumor microenvironment. *Front. Immunol.* 9, 3154. doi:10.3389/fimmu.2018.03154
- Donzelli, S., Fontemaggi, G., Fazi, F., Di Agostino, S., Padula, F., Biagioni, F., et al. (2012). MicroRNA-128-2 targets the transcriptional repressor E2F5 enhancing mutant p53 gain of function. *Cell. Death Differ.* 19, 1038–1048. doi:10.1038/cdd.2011.190
- Frafjord, A., Buer, L., Hammarström, C., Aamodt, H., Woldbæk, P. R., Brustugun, O. T., et al. (2021). The immune landscape of human primary lung tumors is Th2 skewed. *Front. Immunol.* 12, 764596. doi:10.3389/fimmu.2021.764596
- Fukushi, D., Inaba, M., Katoh, K., Suzuki, Y., Enokido, Y., Nomura, N., et al. (2021). R3HDM1 haploinsufficiency is associated with mild intellectual disability. *Am. J. Med. Genet. Part A* 185, 1776–1786. doi:10.1002/ajmg.a.62173
- Han, Y., Wang, Y., Dong, X., Sun, D., Liu, Z., Yue, J., et al. (2023). TISCH2: expanded datasets and new tools for single-cell transcriptome analyses of the tumor microenvironment. *Nucleic Acids Res.* 51, D1425–D1431. doi:10.1093/nar/gkac959
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674. doi:10.1016/j.cell.2011.02.013
- He, S., Yang, L., Xiao, Z., Tang, K., and Xu, D. (2021). Identification of key carcinogenic genes in Wilms' tumor. *Genes. Genet. Syst.* 96, 141–149. doi:10.1266/ggs.21-00015
- Jeschke, J., Bizet, M., Desmedt, C., Calonne, E., Dedeurwaerder, S., Garaud, S., et al. (2017). DNA methylation-based immune response signature improves patient diagnosis in multiple cancers. *J. Clin. Investig.* 127, 3090–3102. doi:10.1172/JCI91095
- Lapiente-Santana, Ó., van Genderen, M., Hilbers, P. A. J., Finotello, F., and Eduati, F. (2021). Interpretable systems biomarkers predict response to immune-checkpoint inhibitors. *Patterns (N Y)* 2, 100293. doi:10.1016/j.patter.2021.100293
- Lee, E., Chuang, H.-Y., Kim, J.-W., Ideker, T., and Lee, D. (2008). Inferring pathway activity toward precise disease classification. *PLoS Comput. Biol.* 4, e1000217. doi:10.1371/journal.pcbi.1000217
- Li, J., Akbani, R., Zhao, W., Lu, Y., Weinstein, J. N., Mills, G. B., et al. (2017). Explore, visualize, and analyze functional cancer proteomic data using the cancer proteome atlas. *Cancer Res.* 77, e51–e54. doi:10.1158/0008-5472.CAN-17-0369
- Li, K., Luo, H., Huang, L., Luo, H., and Zhu, X. (2020). Microsatellite instability: a review of what the oncologist should know. *Cancer Cell. Int.* 20, 16. doi:10.1186/s12935-019-1091-8
- Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., et al. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res.* 48, W509–W514. doi:10.1093/nar/gkaa407
- Liu, C.-J., Hu, F.-F., Xie, G.-Y., Miao, Y.-R., Li, X.-W., Zeng, Y., et al. (2023). GSCA: an integrated platform for gene set cancer analysis at genomic, pharmacogenomic and immunogenomic levels. *Brief. Bioinform.* 24, bbac558. doi:10.1093/bib/bbac558
- Liu, X., Li, L., Si, F., Huang, L., Zhao, Y., Zhang, C., et al. (2021). NK and NKT cells have distinct properties and functions in cancer. *Oncogene* 40, 4521–4537. doi:10.1038/s41388-021-01880-9
- Liu, X., Xu, J., Zhang, B., Liu, J., Liang, C., Meng, Q., et al. (2019). The reciprocal regulation between host tissue and immune cells in pancreatic ductal adenocarcinoma: new insights and therapeutic implications. *Mol. Cancer* 18, 184. doi:10.1186/s12943-019-1117-9
- Malta, T. M., Sokolov, A., Gentles, A. J., Burzykowski, T., Poisson, L., Weinstein, J. N., et al. (2018). Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* 173, 338–354. doi:10.1016/j.cell.2018.03.034
- Mei, S., Qin, Q., Wu, Q., Sun, H., Zheng, R., Zang, C., et al. (2017). Cistrome Data Browser: a data portal for ChIP-Seq and chromatin accessibility data in human and mouse. *Nucleic Acids Res.* 45, D658–D662. doi:10.1093/nar/gkw983
- Nassar, D., and Blanpain, C. (2016). Cancer stem cells: basic concepts and therapeutic implications. *Annu. Rev. Pathol.* 11, 47–76. doi:10.1146/annurev-pathol-012615-044438
- Phillips, K.-A., Jenkins, M., Lindeman, G., McLachlan, S., McKinley, J., Weideman, P., et al. (2006). Risk-reducing surgery, screening and chemoprevention practices of BRCA1 and BRCA2 mutation carriers: a prospective cohort study. *Clin. Genet.* 70, 198–206. doi:10.1111/j.1399-0004.2006.00665.x
- Reck, M., Schenker, M., Lee, K. H., Provencio, M., Nishio, M., Lesniewski-Kmak, K., et al. (2019). Nivolumab plus ipilimumab versus chemotherapy as first-line treatment in advanced non-small-cell lung cancer with high tumour mutational burden: patient-reported outcomes results from the randomised, open-label, phase III CheckMate 227 trial. *Eur. J. Cancer* 116, 137–147. doi:10.1016/j.ejca.2019.05.008
- Reinhold, W. C., Wilson, K., Elloumi, F., Bradwell, K. R., Ceribelli, M., Varma, S., et al. (2023). CellMinerCDB: NCATS is a web-based portal integrating public cancer cell line databases for pharmacogenomic explorations. *Cancer Res.* 83, 1941–1952. doi:10.1158/0008-5472.CAN-22-2996
- Repáraz, D., Ruiz, M., Llopiz, D., Silva, L., Vercher, E., Aparicio, B., et al. (2022). Neoantigens as potential vaccines in hepatocellular carcinoma. *J. Immunother. Cancer* 10, e003978. doi:10.1136/jitc-2021-003978
- Sadeghi Rad, H., Monkman, J., Warkiani, M. E., Ladwa, R., O'Byrne, K., Rezaei, N., et al. (2021). Understanding the tumor microenvironment for effective immunotherapy. *Med. Res. Rev.* 41, 1474–1498. doi:10.1002/med.21765
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *P Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi:10.1073/pnas.0506580102
- Tuupanen, S., Hänninen, U. A., Kondelin, J., Von Nandelstadh, P., Cajuso, T., Gylfe, A. E., et al. (2014). Identification of 33 candidate oncogenes by screening for base-specific mutations. *Br. J. Cancer* 111, 1657–1662. doi:10.1038/bjc.2014.429
- Veres, D. V., Gyurkó, D. M., Thaler, B., Szalay, K. Z., Fazekas, D., Korcsmáros, T., et al. (2015). ComPPI: a cellular compartment-specific database for protein-protein interaction network analysis. *Nucleic Acids Res.* 43, D485–D493. doi:10.1093/nar/gku1007
- Wang, S., Sun, H., Ma, J., Zang, C., Wang, C., Wang, J., et al. (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat. Protoc.* 8, 2502–2515. doi:10.1038/nprot.2013.150
- Wu, T., and Dai, Y. (2017). Tumor microenvironment and therapeutic response. *Cancer Lett.* 387, 61–68. doi:10.1016/j.canlet.2016.01.043
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innov. (Camb)* 2, 100141. doi:10.1016/j.xinn.2021.100141
- Xu, L., Deng, C., Pang, B., Zhang, X., Liu, W., Liao, G., et al. (2018). TIP: a web server for resolving tumor immunophenotype profiling. *Cancer Res.* 78, 6575–6580. doi:10.1158/0008-5472.CAN-18-0689
- Yang, C., Zhang, H., Chen, M., Wang, S., Qian, R., Zhang, L., et al. (2022). A survey of optimal strategy for signature-based drug repositioning and an application to liver cancer. *Elife* 11, e71880. doi:10.7554/eLife.71880
- Yuan, H., Yan, M., Zhang, G., Liu, W., Deng, C., Liao, G., et al. (2019). Cancer SEA: a cancer single-cell state atlas. *Nucleic Acids Res.* 47, D900–D908. doi:10.1093/nar/gky939
- Zheng, R., Wan, C., Mei, S., Qin, Q., Wu, Q., Sun, H., et al. (2019). Cistrome Data Browser: expanded datasets and new tools for gene regulatory analysis. *Nucleic Acids Res.* 47, D729–D735. doi:10.1093/nar/gky1094

Nomenclature

ACC	Adrenocortical carcinoma	SARC	Sarcoma
AUC	Area under the curve	SKCM	Skin cutaneous melanoma
BLCA	Bladder urothelial carcinoma	STAD	Stomach adenocarcinoma
BRCA	Breast invasive carcinoma	STES	Stomach and esophageal carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma	TCGA	The Cancer Genome Atlas
CHOL	Cholangiocarcinoma	TIICs	Tumor-infiltrating immune cells
CCLE	Cancer cell line encyclopedia	TME	Tumor microenvironment
COAD	Colon adenocarcinoma	TGCT	Testicular Germ Cell Tumors
COAD/ READ	Colon adenocarcinoma/Rectum adenocarcinoma esophageal carcinoma	THCA	Thyroid carcinoma
DSS	Disease-specific survival	THYM	Thymoma
DFI	Disease-free interval	UCEC	Uterine corpus endometrial carcinoma
DLBC	Lymphoid neoplasm diffuse large B-cell lymphoma	UCS	Uterine carcinosarcoma
ESCA	Esophageal carcinoma	UVM	Uveal melanoma
FAP	Fibroblast activation protein- α		
GTE_x	Genotype Tissue-Expression		
GSEA	Gene set enrichment analysis		
GSVA	Gene set variation analysis		
GBM	Glioblastoma multiforme		
GBMLGG	Glioma		
HNSC	Head and Neck squamous cell carcinoma		
ICIs	Immune checkpoint inhibitors		
KICH	Kidney chromophobe		
KM	Kaplan-Meier		
KIRC	Kidney renal clear cell carcinoma		
KIRP	Kidney renal papillary cell carcinoma		
LAML	Acute myeloid leukemia		
LGG	Brain lower grade glioma		
LIHC	Liver hepatocellular carcinoma		
LUAD	Lung adenocarcinoma		
LUSC	Lung squamous cell carcinoma		
MESO	Mesothelioma		
NCI	National Cancer Institute		
OS	Overall survival		
OV	Ovarian serous cystadenocarcinoma		
PFI	Progress-free interval		
PPI	Protein-protein interaction		
PAAD	Pancreatic adenocarcinoma		
PCPG	Pheochromocytoma and paraganglioma		
PRAD	Prostate adenocarcinoma		
ROC	Receiver operating characteristic curve		