

Original Article

Pan-cancer analysis reveals RPRD1B overexpression as a prognostic biomarker associated with altered immune cell infiltration in diffuse large B-cell lymphoma

Lu Xu^{1,2,3}, Zhihao Xie¹, Shi Tao¹, Yuwei Shen¹, Xinbao Hao^{1,2}

¹Department of Hematology, The First Affiliated Hospital of Hainan Medical University, Haikou 570102, Hainan, China; ²School of Medicine, Tsinghua University, Beijing 100084, China; ³Department of Hematology, Roswell Park Comprehensive Cancer Center, NY 14263, USA

Received April 21, 2024; Accepted June 10, 2026; Epub July 15, 2026; Published July 30, 2026

Abstract: Objective: RPRD1B plays an essential role in transcriptional and cell cycle regulation, and has been reported to influence the initiation and progression of certain tumors. This study was done to systematically evaluate RPRD1B function in human tumors, especially in diffuse large B-cell lymphoma (DLBCL), and correlation with prognosis. Methods: RPRD1B expression was analyzed using the Human Protein Atlas and TCGA datasets, with differential expression and survival analysis conducted in R software. Mutation and genomic heterogeneity were assessed with maftools. Immune infiltration correlations were evaluated using ESTIMATE and psych R packages. Cell colony formation and apoptosis assays were used to validate the analysis results. Statistical analysis was performed in R and GraphPad Prism. Results: RPRD1B mRNA was significantly upregulated in 16 cancer types and downregulated in 9, with markedly elevated expression in DLBCL compared to normal lymphoid tissues. Univariate Cox regression in TCGA cohorts identified RPRD1B as a prognostic factor in seven tumor types, including unfavorable associations in GBMLGG (HR = 1.87, $P = 1.8 \times 10^{-5}$) and LGG (HR = 2.57, $P = 2.5 \times 10^{-5}$), and a favorable association in KIRC (HR = 0.75, $P = 7.7 \times 10^{-3}$). This prognostic significance was further validated in an independent DLBCL cohort (GSE10846, n = 414; HR = 1.41, 95% CI 1.02-1.94, $P = 0.038$). RPRD1B expression was significantly correlated with copy number variation in 15 cancer types and associated with the ImmuneScore in all 30 cancer types analyzed, with predominantly negative correlations. In DLBCL, RPRD1B expression was negatively correlated with CD4⁺ T cells, $\gamma\delta$ T cells, and Th1 cells, and positively correlated with Th2 cells ($P < 0.05$ for all). *In vitro*, RPRD1B knock-down reduced colony formation in SU-DHL-2 cells ($P < 0.001$) and its overexpression suppressed early apoptosis in SU-DHL-4 cells ($P < 0.01$). Conclusion: RPRD1B is universally overexpressed in pan-cancer and is associated with the prognosis of multiple tumors, especially poor prognosis in DLBCL. RPRD1B may be used as an indicator for evaluating the efficacy of immunotherapy for DLBCL.

Keywords: RPRD1B, pan-cancer, diffuse large B-cell lymphoma, RNA polymerase II

Introduction

RPRD1B, which we previously identified and initially named CREPT (cell cycle-related and elevated-expression protein in tumors) [1], is a member of the RPRD protein family, which also includes RPRD1A and RPRD2 [2, 3]. RPRD1B is ubiquitous in the nuclei of human tissues. RPRD1B can bind to the phosphorylated C-terminal heptapeptide repeat domain (CTD) of POLR2A subunit of RNA polymerase II (RNAPII), regulate the transcriptional activity of RNAPII, and promote the binding of RNAPII to the promoter region and transcription termination

region of target genes to promote their transcription [4]. RPRD1B can stimulate the transcription of tumor genes, including the Wnt pathway and cell cycle-related genes, and is related to the malignant phenotype of tumor cells [5-11].

RPRD1B is highly expressed and associated with poor prognosis in some tumor types [7, 12-18]. RPRD1B can promote the proliferation, migration, and invasion of tumor cells *in vitro* and in animal experiments, but its specific role in tumorigenesis and development is still unclear [13, 19]. The expression of RPRD1B in

Table 1. Abbreviations of cancer types in TCGA database modified

Abbreviation	Cancer type
ACC	Adrenocortical carcinoma
BLCA	Bladder urothelial carcinoma
BRCA	Breast invasive carcinoma
CESC	Cervical and endocervical cancers
CHOL	Cholangiocarcinoma
COAD	Colon adenocarcinoma
COADREAD	Colorectal adenocarcinoma
DLBC	Lymphoid Neoplasm Diffuse Large B-cell Lymphoma
ESCA	Esophageal carcinoma
ESCC	Esophageal squamous cell carcinoma
GBM	Glioblastoma multiforme
GBMLGG	Glioma
HNSC	Head and Neck squamous cell carcinoma
KICH	Kidney Chromophobe
KIRC	Kidney renal clear cell carcinoma
KIRP	Kidney renal papillary cell carcinoma
KIPAN	Pan-kidney cohort (KICH+KIRC+KIRP)
LAML	Acute Myeloid Leukemia
LGG	Brain Lower Grade Glioma
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MESO	Mesothelioma
OV	Ovarian serous cystadenocarcinoma
PAAD	Pancreatic adenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate adenocarcinoma
READ	Rectum adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
STAD	Stomach adenocarcinoma
TGCT	Testicular Germ Cell Tumors
THCA	Thyroid carcinoma
THYM	Thymoma
UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma
UVM	Uveal Melanoma

hematologic tumors and its influence on prognosis and its relationship with tumor immune response have not been reported.

In this study, we aimed to analyze the expression of RPRD1B across various pan-cancer tumors, specifically focusing on diffuse large B-cell lymphoma (DLBCL), and to investigate its correlation with prognosis. Furthermore, we explored the mutations of the RPRD1B gene in

tumors and examined its relationship with genome heterogeneity. Using available sequencing data, we will study the expression of RPRD1B in immune cells and delve into possible mechanisms underlying its role in immune regulation. This study will provide essential insight to explore the potential of RPRD1B for prognosis of DLBCL.

Materials and methods

Cancer abbreviations and extraction of gene expression datasets

The abbreviations of cancers are as shown in **Table 1** (<https://stephenturner.github.io/tcga-codes/>).

We obtained a uniformly normalized pan-cancer dataset, The Cancer Genome Atlas Program (TCGA), Therapeutically Applicable Research to Generate Effective Treatments (TARGET), and Genotype-Tissue Expression (GTEx) project (PAN-CAN, N = 19131, G = 60499), from the UCSC database (<https://xenabrowser.net/>). Subsequently, we extracted the expression data of the RPRD1B gene from each sample in the dataset. We applied a $\log_2(x+0.001)$ transformation for each selected sample to the expression values. Additionally, we separately extracted the gene expression profiles for each tumor, which we

re used for the correlation analysis between RPRD1B expression and genomic heterogeneity and immune infiltration scores.

Distribution and differential expression of RPRD1B in normal and tumor tissues

We utilized the Human Protein Atlas (HPA, <https://www.proteinatlas.org/>) to analyze the protein and mRNA expression levels and tissue

distribution of RPRD1B in normal tissues, blood, and tumor samples. Furthermore, based on the RPRD1B gene expression dataset from TCGA Pan-Cancer patients that we obtained, we calculated the expression differences between normal samples and tumor samples for each type of tumor using R software (version 3.6.4). The statistical significance of the differences was analyzed using unpaired Wilcoxon Rank Sum and Signed Rank Tests.

Analysis of RPRD1B expression in pan-cancer and its association with patient prognosis

Based on the previously obtained expression data of RPRD1B across pan-cancers, we acquired high-quality prognostic data sets from TCGA [20]. We supplemented this with TARGET follow-up data from the UCSC Cancer Browser (<https://xenabrowser.net/datapages/>) and excluded samples with a follow-up duration of less than 30 days. Using the R survival package (version 3.2-7), we analyzed the association between RPRD1B expression and overall survival in each tumor type by univariate Cox proportional hazards regression, with results visualized as a forest plot [21]. The log-rank test was employed for statistical testing to determine prognostic significance. The Kaplan-Meier Plotter database (<https://kmplot.com/analysis/>) was further used to validate the prognostic significance of RPRD1B expression in pan-cancer, with patients in each tumor type stratified into high- and low-expression groups by the platform's auto-select best cutoff algorithm.

Analysis of RPRD1B mutation and its correlation with tumor genome heterogeneity

We acquired Simple Nucleotide Variation (SNV) datasets of all TCGA samples processed using the MuTect2 software [22] from the Genomic Data Commons (GDC, <https://portal.gdc.cancer.gov/>). Subsequently, we integrated the mutation data from these samples and obtained protein domain information using the R software package maftools (version 2.2.10). Additionally, we downloaded gene-level Copy Number Variation (CNV) datasets of all TCGA samples processed using the GISTIC software [23] from the GDC. We integrated the copy number data with gene expression data.

Furthermore, we obtained the Microsatellite Instability (MSI) scores for each tumor from a

previous study [24] and calculated the Tumor Mutation Burden (TMB) for each tumor using the R software package maftools (version 2.8.05). We then integrated the MSI and TMB data with gene expression data. Each expression value was normalized using the $\log_2(x+0.001)$ transformation and cancer types with fewer than three samples were excluded. Pearson correlations were computed within each tumor type.

Correlation analysis of RPRD1B expression with immune infiltration scores in tumors

We mapped each tumor's gene expression profiles to GeneSymbol. Furthermore, using the R package ESTIMATE (version 1.0.13) [25], we computed the STROMAL and IMMUNE scores for each tumor patient based on their gene expression profiles. Employing the R package psych (version 2.1.6), we calculated the Pearson's correlation coefficient between the RPRD1B gene and immune infiltration scores across the tumors to identify significantly correlated immune infiltration scores.

To analyze RPRD1B expression in immune cells and its correlation with immune cell infiltration in DLBCL, we mapped the gene expression profiles of DLBCL patients to GeneSymbol. Subsequently, using the deconvo_cibersort from the R package IOBR (Immuno-Oncology Biological Research) (version 0.99.9) [26, 27], we evaluated the infiltration scores of 22 types of immune cells in DLBCL patients. Employing the R package psych (version 2.1.6), we calculated the Pearson's correlation coefficient between RPRD1B gene expression and immune cell infiltration scores across the tumors and assessed their statistical significance.

Analysis of RPRD1B expression and its prognostic implication in lymphoma

The results of RPRD1B immunohistochemical staining in normal lymph node tissue and lymphoma were obtained from the HPA database. RPRD1B mRNA analysis of lymphoma and normal lymphoid tissues, expression data obtained from UCSC XENA (<https://xenabrowser.net/datapages/>) by the Toil process unified processing of TCGA, and GTEx transcripts per million (TPM)-formatted RNA-seq data [28]. Data corresponding to DLBCL in TCGA and normal tissue in GTEx were extracted, and the ggplot2

package was used to visualize the data. The DLBCL expression data of GSE 10846 (n = 414) were downloaded from the Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo) database. Patients were stratified into high- and low-expression groups (n = 207 each) by the median value of RPRD1B mRNA expression. The prognostic significance of RPRD1B was assessed using the R survival package (version 3.3.1) with a univariate Cox proportional hazards regression model, and Kaplan-Meier survival curves were compared by the log-rank test.

Experimental validation of RPRD1B for promoting cell colony formation and inhibiting apoptosis

Plasmids and cell lines: Expression plasmids for human CREPT are pCMV/HACREPT constructed in our lab. The plasmid pBS/U6/CREPT-si was developed as previously described [1]. The target sequence for siRNA (CREPT-si), GGACCTGAATTCAGTACAGAGA, was originally designed and validated by Lu et al. [1] and is conserved between human and mouse. B cell lymphoma cell lines SU-DHL-2 and SU-DHL-4 were purchased from the Shanghai Cell Bank. Cells used for transfection and apoptosis assays were cultured based on prior research [1]. Specific experiments were replicated three times.

Cell colony formation assay: For anchorage-independent growth in soft agar, 7×10^3 cells were cultured in a medium containing 0.35% agar and placed in a six-well plate with 0.7% agar medium. Cells were cultured for two weeks or four weeks. Cell colonies were enumerated and data are presented as mean $\pm$ SD from three independent biological replicates. Differences between the control group and each shRNA group at each time point (14 and 28 days) were assessed by independent-samples t-test using GraphPad Prism (version 8.0).

Cell apoptosis assay using flow cytometry: Cells were stained using the APC Annexin V Apoptosis Detection Kit (BioLegend) per the manufacturer's instructions and then analyzed using a FACSCalibur flow cytometer (BD Biosciences, San Jose, CA, USA).

Statistical analysis

All statistical analyses were performed in R (v3.6.3 and v4.2.1) and GraphPad Prism (ver-

sion 8.0). Statistical significance was assessed using an independent samples t-test or the Wilcoxon rank-sum test for expression data. For correlation analysis, Pearson or Spearman correlation coefficient were used according to data distribution. Survival analyses were performed using the R survival package; Kaplan-Meier curves were compared by the log-rank test, and univariate Cox proportional hazards regression was used to estimate hazard ratios (HRs) and 95% confidence intervals (CIs). A two-sided P value < 0.05 was considered significant.

Results

RPRD1B distribution in normal tissues and differential expression in tumors

According to the HPA database, the protein expression of RPRD1B was detected in 44 tissues, except for adipose tissue. The expression score was the lowest in the liver, cerebellum, and hippocampus and moderate or high in the other 41 tissues (**Figure 1A**). Compared to the distribution of the protein in tissues, we found that the mRNA abundance of RPRD1B in 54 tissues showed roughly the same trend. Still, the content was higher in the myocardium, skeletal muscle, and lymphatic blood system. In contrast to the lowest content of RPRD1B protein in the liver, its mRNA level was significantly higher than all other tissues (**Figure 1B**). These results suggest that RPRD1B has tissue differences in protein translation and post-translational regulation in liver tissues.

Immunofluorescence staining of colon cancer and rhabdomyosarcoma cell lines showed that RPRD1B protein expression was localized to the nucleus (**Figure 1C**). Immunohistochemical staining results of 20 tumor tissues in the HPA database showed that all tumor tissues had moderate or higher intensity nuclear staining, and the positive percentage was 72-100% in patients with various tumors (**Figure 1D**).

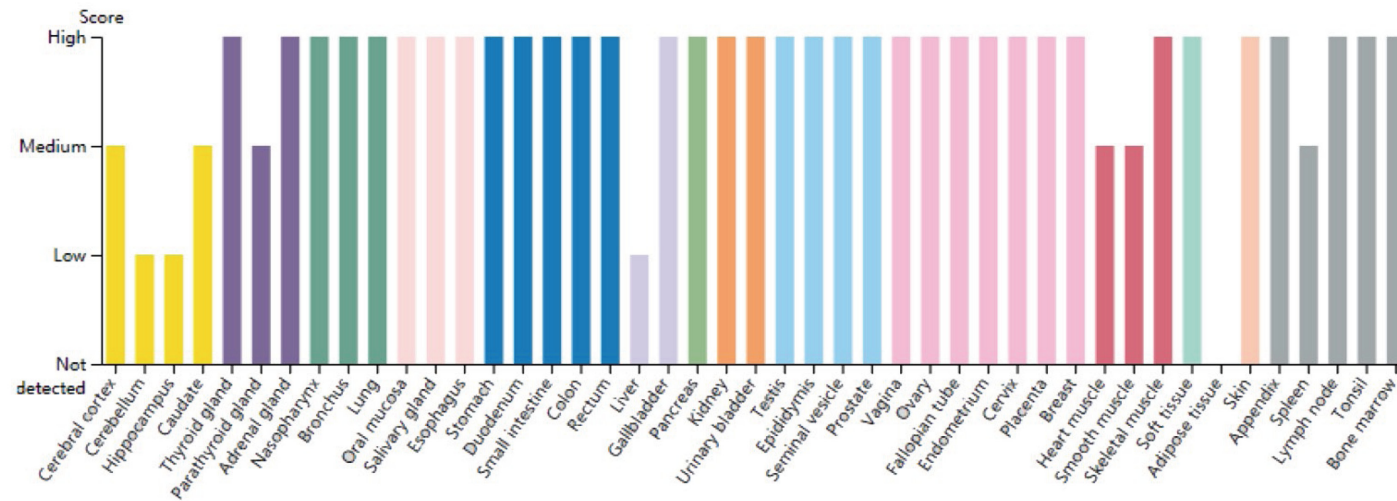
Using data from the TCGA and GTEx databases, the differential expression between normal and tumor samples for each tumor type was calculated. The results showed that RPRD1B was significantly upregulated in 16 tumor types and downregulated in 9, as shown in **Figure 1E**.

RPRD1B expression is correlated to the prognosis in pan-cancers

The correlation analysis between gene expression and prognosis revealed that in four tumor

RPRD1B in pan-cancer prognosis and DLBCL immunity

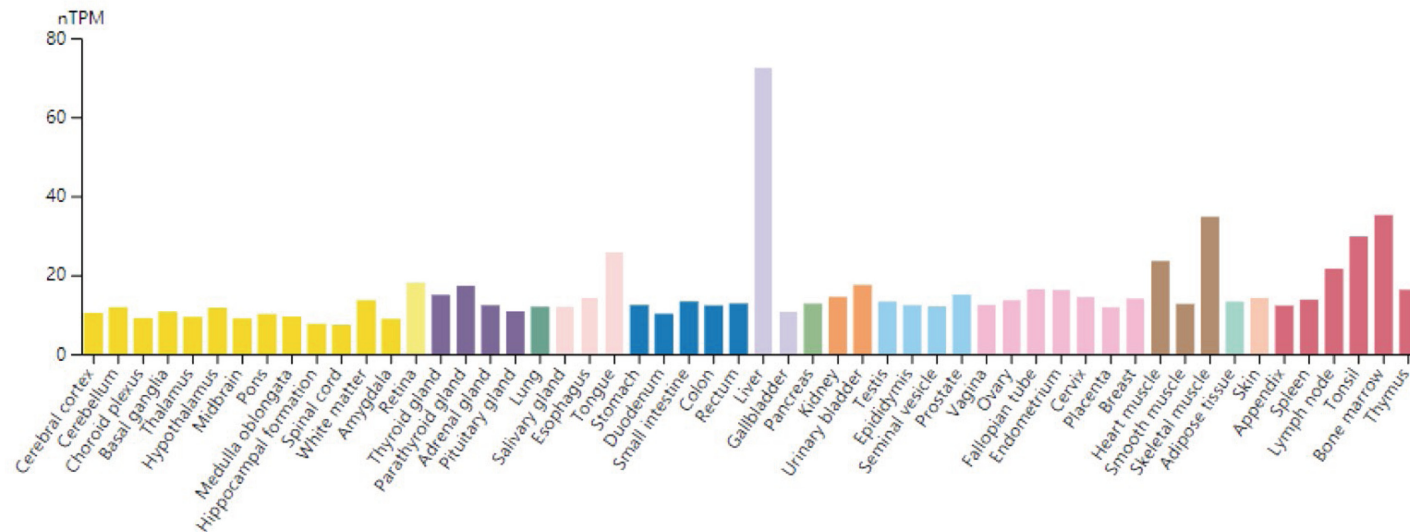
A



B

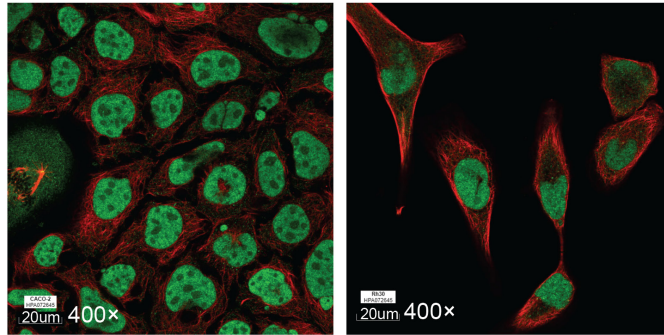
Consensus dataset¹

RNA Tissue Specificity

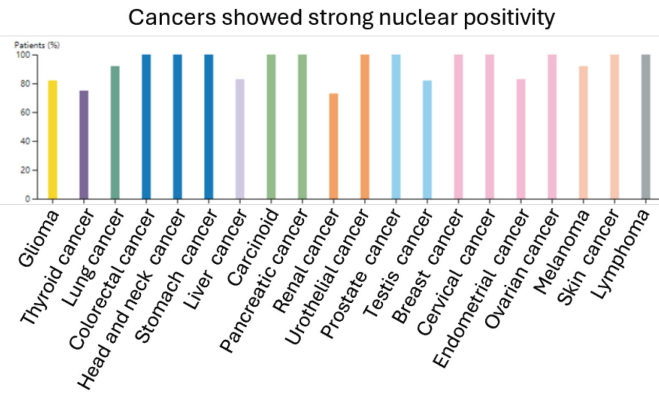


RPRD1B in pan-cancer prognosis and DLBCL immunity

C



D



E

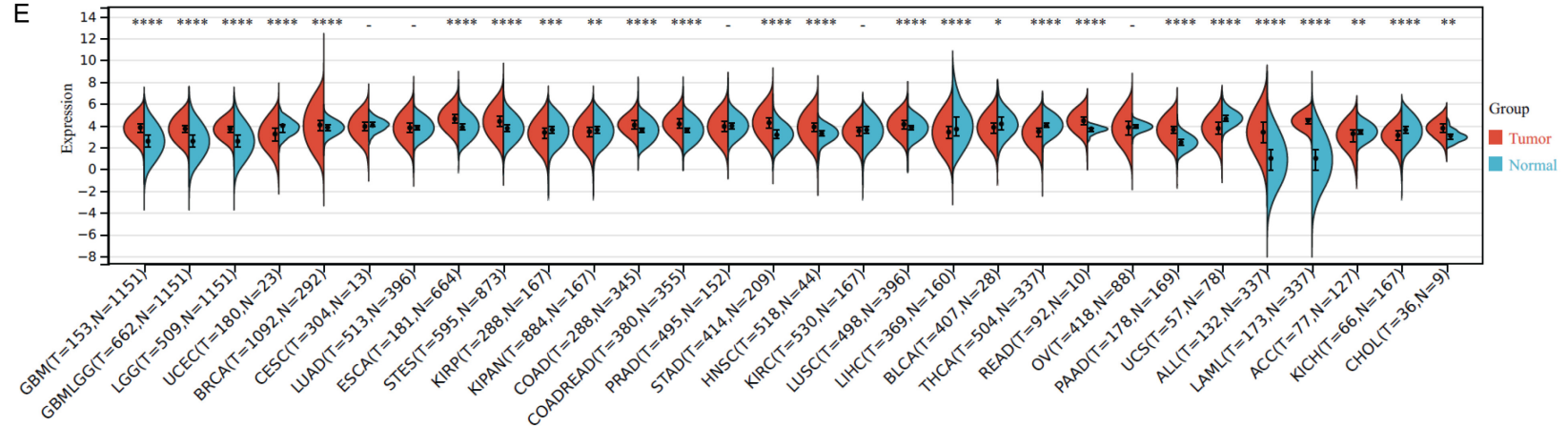


Figure 1. RPRD1B expression in normal and cancer tissues. A. Distribution level of RPRD1B protein in various organs. B. RPRD1B mRNA in normal tissues (55 tissue types). C. RPRD1B protein is localized in the nucleus of the CACO-2 cell (left) and RH-30 cell (right). D. Positive expression percentage of RPRD1B protein in tumor tissues of patients (20 tumor types). E. Differential expression of RPRD1B in pan-cancers (30 tumor types).

types (GBMLGG, LGG, TARGET-LAML, LAML), high RPRD1B expression was associated with poor prognosis. Conversely, in two tumor types (KIRC, READ), low RPRD1B expression correlated with unfavorable outcomes, as shown in **Figure 2A**.

We used the Kaplan-Meier Plotter database to produce survival curves for pan-cancer patients, stratifying them based on their RPRD1B expression levels. The analysis demonstrated that for tumors including BRCA, ESCA, HNSC, LIHC, KIRP, THCA, and UCEC, increased RPRD1B expression was linked to less favorable overall survival (OS) outcomes ($P < 0.05$). Conversely, for ESCC, STAD, and KIRC, higher RPRD1B expression was connected to improved OS ($P < 0.05$) (**Figure 2B**).

Across the nine tumor types analyzed by both panels, the directions of association between RPRD1B expression and overall survival were consistent, although the levels of statistical significance differed. These differences likely reflect the use of distinct datasets (TCGA/TARGET versus Kaplan-Meier Plotter cohorts) and stratification strategies (median dichotomization with univariate Cox regression versus the data-driven best-cutoff algorithm). The TCGA-based results (**Figure 2A**) should therefore be considered the primary evidence, with **Figure 2B** serving as supportive validation; KIRC, where both analyses concurred, represents the most robust association.

RPRD1B is mutated in tumors, and its expression favors tumor genomic heterogeneity

By analyzing gene sequencing data from relevant tumors in the GDC database, we identified the types of mutations of RPRD1B in 20 cancers and calculated their mutation rates. By aligning the nucleotide mutation locations with the protein sequence, it can be observed that all types of mutations can occur in the CID, Linker, and CCT regions of RPRD1B. The lowest mutation rate was 0.1% in KIPAN, while the highest was 3.4% in UCEC (**Figure 3A**). Regarding the relationship between RPRD1B gene expression and CNV, we ultimately obtained expression data for 20 types of cancer using the TCGA database. We computed the expression differences of the RPRD1B gene in various CNV samples within each tumor. Significant differences were observed in 15 types of cancers,

with the most marked differences seen in LUAD, BRCA, and LUSC (**Figure 3B**). The expression data of 37 cancer types were obtained to analyze the correlation between RPRD1B expression and MSI and TMB in pan-cancer.

Correlation analysis between RPRD1B expression and tumor MSI showed a significant correlation in 13 tumors, including five tumors with a positive correlation (LUAD, KIRC, LUSC, TGCT, UVM) and eight tumors with a negative correlation (COAD, COADREAD, STES, STAD, PRAD, HNSC, THCA, DLBC) (**Figure 3C**). Pearson correlation of TMB and RPRD1B was calculated in each tumor, and significant correlation was observed in five tumors, among which a significant positive correlation was observed in LUAD tumors, and negative correlation was observed in the remaining four tumors (COAD, COADREAD, STES, THCA) (**Figure 3D**).

RPRD1B expression suppresses tumor immune cell infiltration

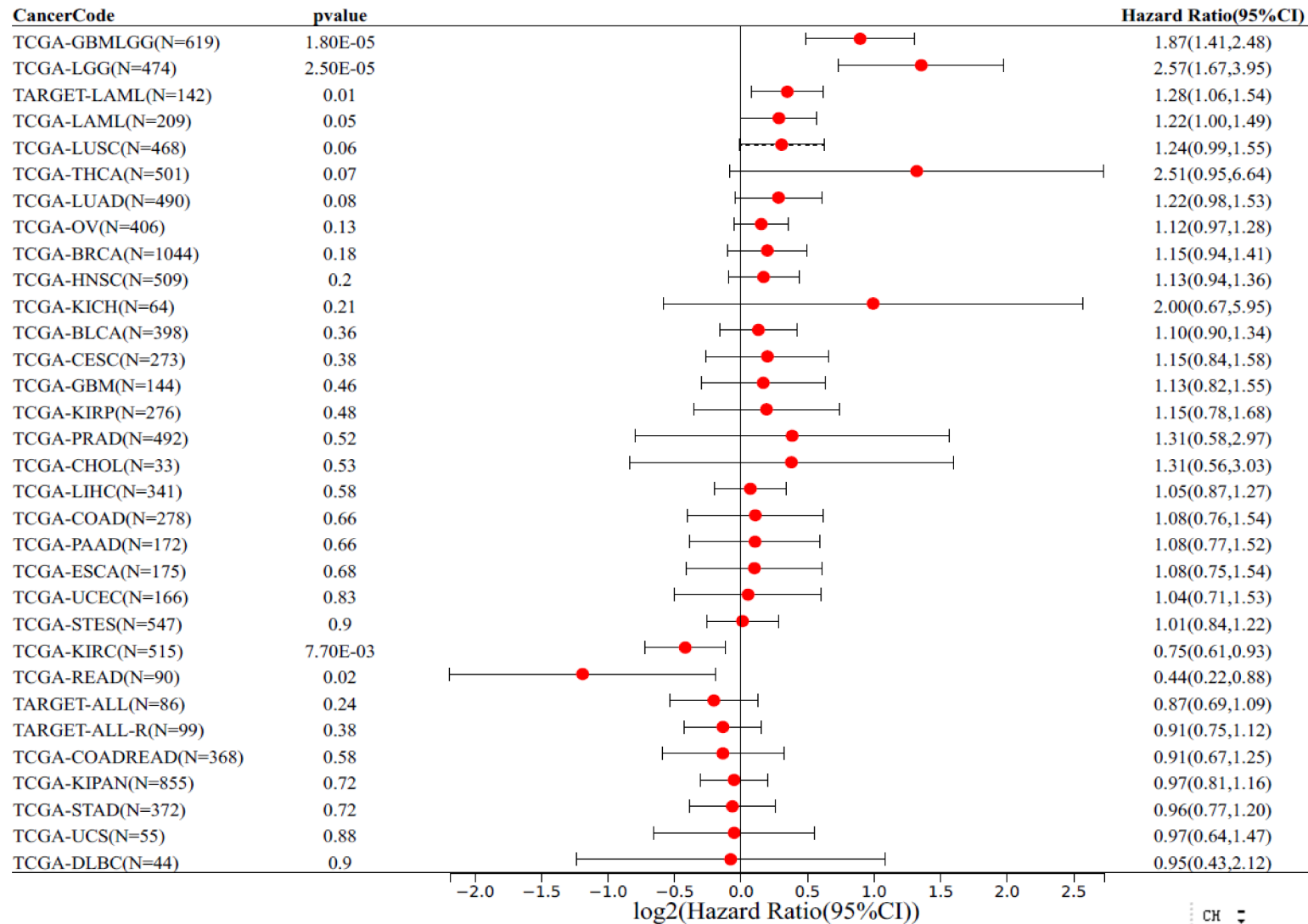
Using the ESTIMATE immune scoring system, we conducted a statistical analysis on the correlation between RPRD1B gene expression and tumor microenvironment scores. Analysis of the association between RPRD1B expression and the ImmuneScore, which represents immune cell infiltration, indicated that RPRD1B expression was significantly correlated with the ImmuneScore across all 30 cancer types analyzed, predominantly in a negative direction (**Figure 4A**). We further analyzed the correlation between RPRD1B expression and the StromalScore, which represents the infiltration of stromal cells. We observed a significant correlation between RPRD1B expression and stromal cell infiltration in 20 types of cancer, with six tumors showing a significant positive correlation and 14 tumors showing a significant negative correlation (**Figure 4B**). These findings suggest that the RPRD1B protein exhibits a negative regulatory function in antitumor immunity.

RPRD1B is markedly expressed in lymphoma tissues and is associated with poor prognosis

In the HPA database, all lymphoma tissues showed moderate expression or above of RPRD1B compared to normal lymph nodes (**Figure 5A**). Sequencing data from TCGA and GTEx databases showed that the mRNA expression level of RPRD1B in DLBCL patients was signifi-

RPRD1B in pan-cancer prognosis and DLBCL immunity

A



RPRD1B in pan-cancer prognosis and DLBCL immunity

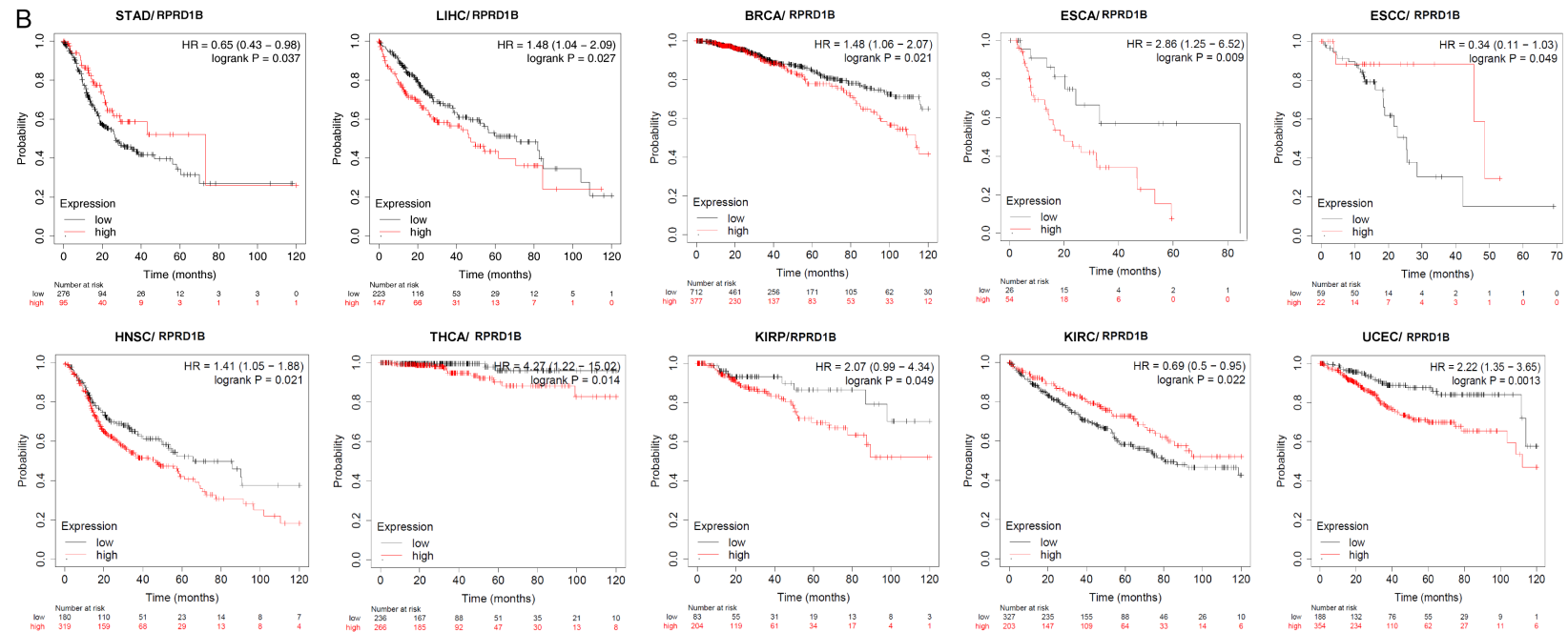
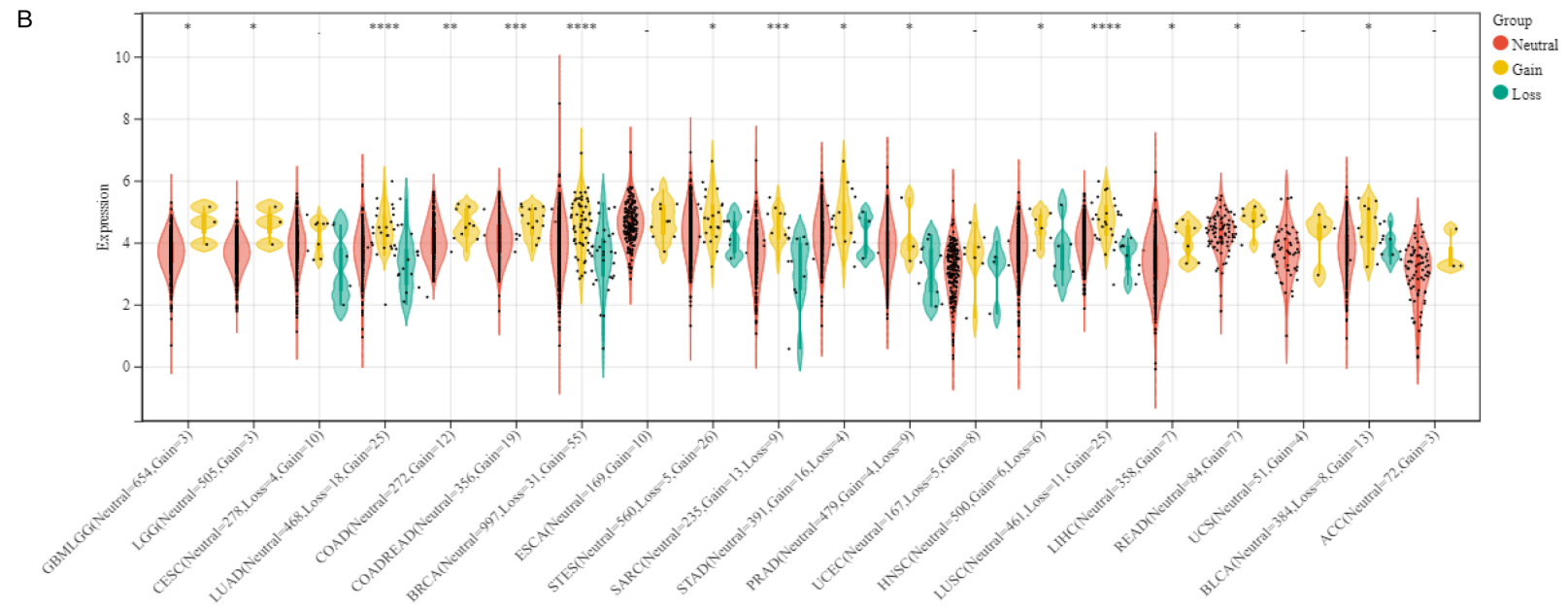
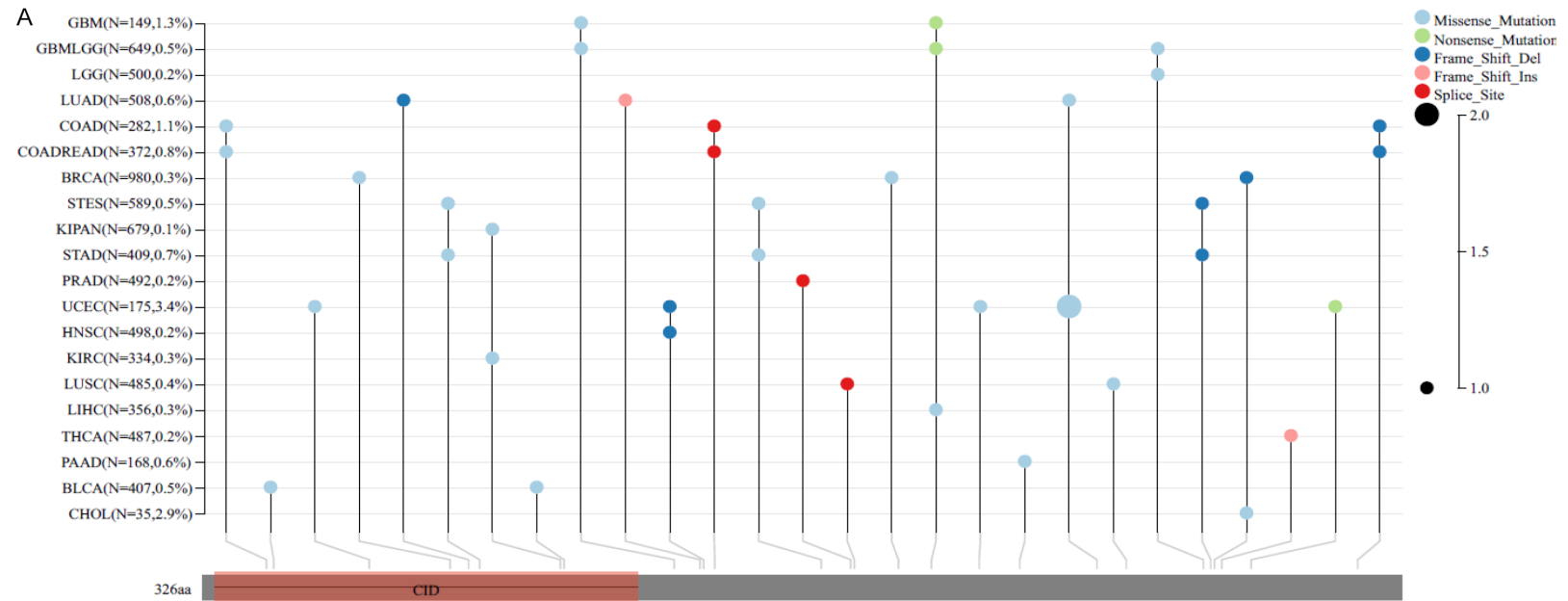


Figure 2. Correlation between RPRD1B and pan-cancer prognosis. A. Elevated or decreased RPRD1B in a variety of tumor tissues is significantly associated with poor prognosis in cancer patients. B. The correlation between RPRD1B and OS of cancers was analyzed using a Kaplan plotter (10 tumors with significant differences presented: 8 tumors $P < 0.05$, 2 tumors $P < 0.01$).

RPRD1B in pan-cancer prognosis and DLBCL immunity



RPRD1B in pan-cancer prognosis and DLBCL immunity

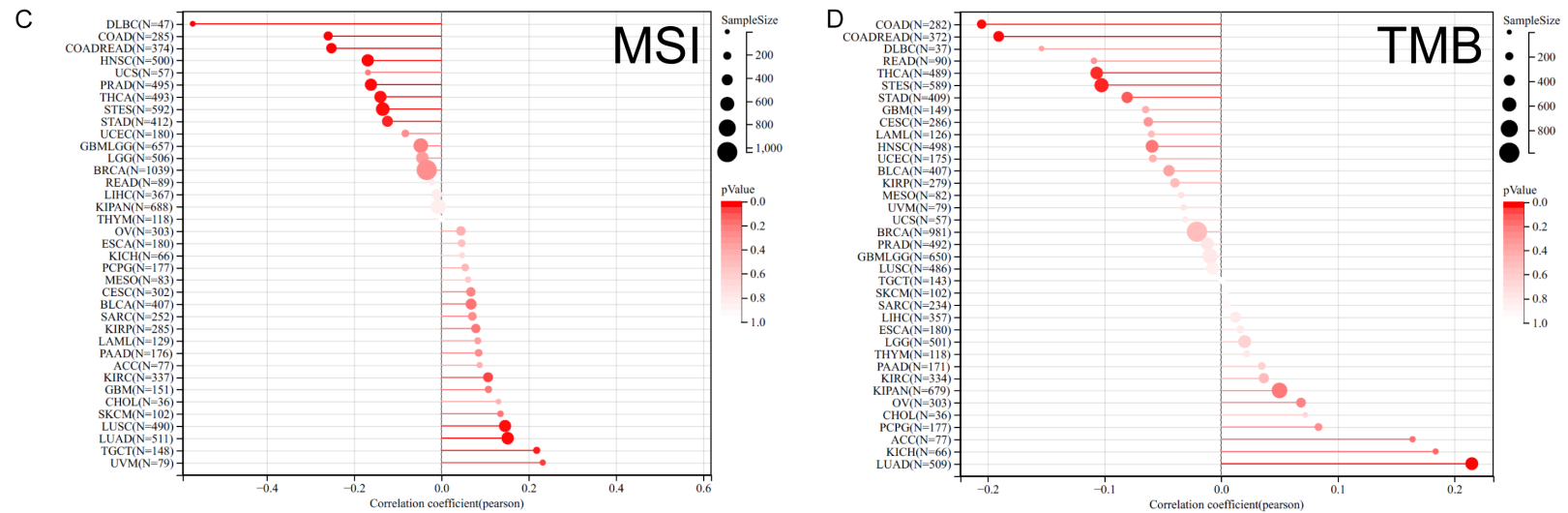
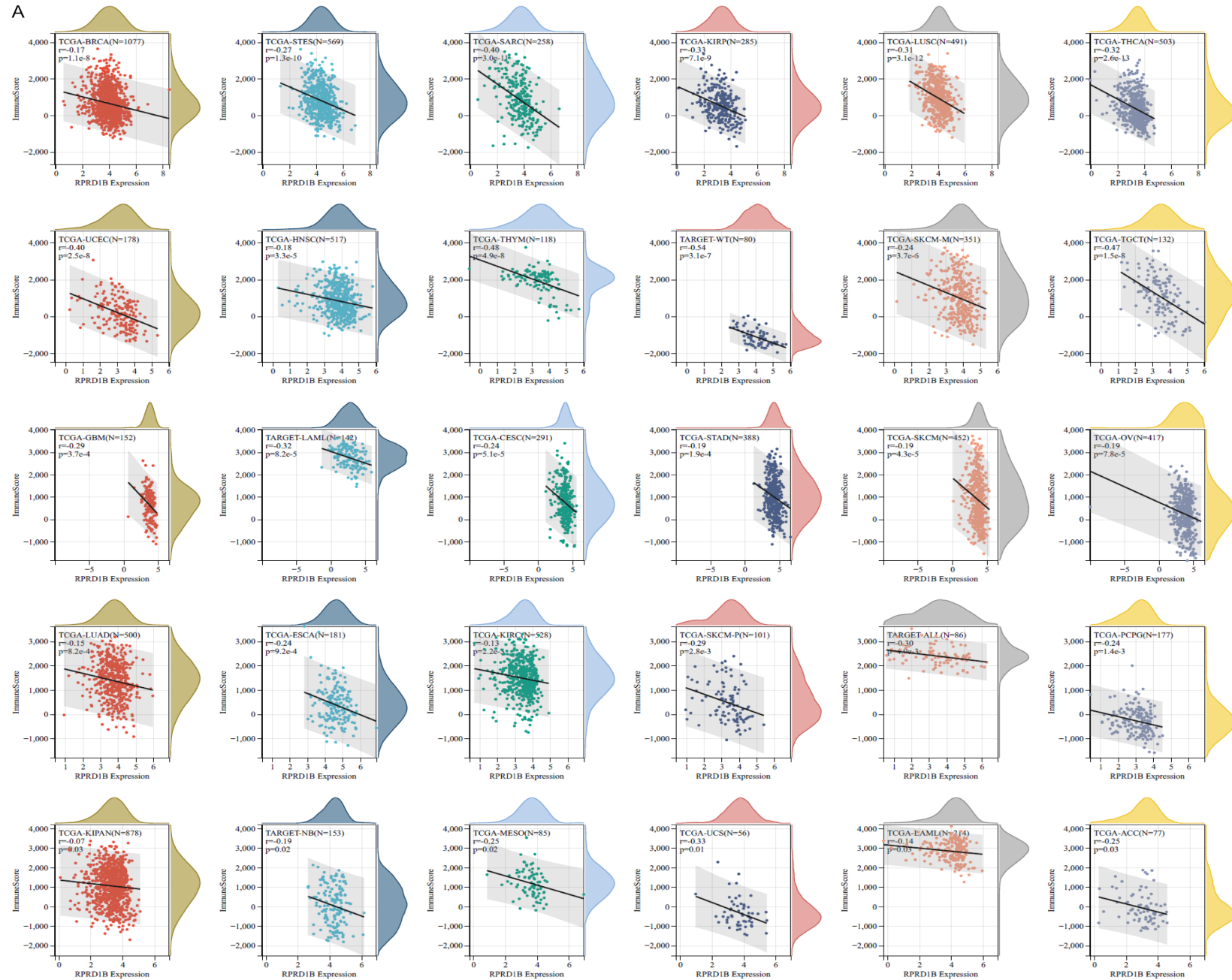


Figure 3. Mutation analysis of RPRD1B across various cancers and correlation of its expression with CNV and genomic heterogeneity. A. RPRD1B mutation types and total mutation rates by tumor types. B. Correlation between RPRD1B expression and CNV in pan-cancers. C. Correlation analysis of RPRD1B expression and MSI in pan-cancers. D. Correlation analysis of RPRD1B expression and TMB in pan-cancers.

RPRD1B in pan-cancer prognosis and DLBCL immunity

A



RPRD1B in pan-cancer prognosis and DLBCL immunity

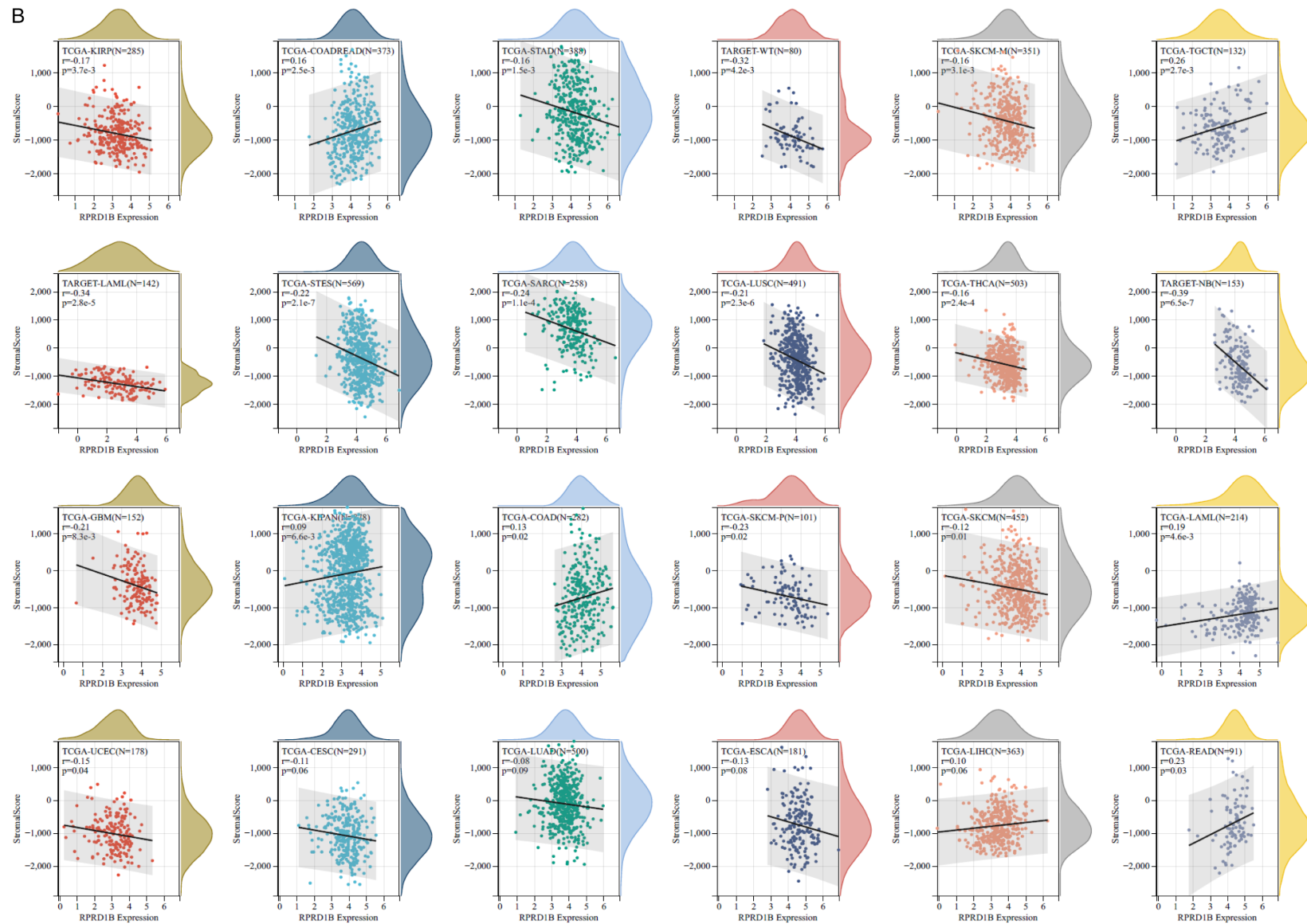


Figure 4. Immune infiltration scores in various cancers correlated with RPRD1B expression. A. Types of cancer where immune cell infiltration scores significantly correlate with RPRD1B expression. B. Types of cancer where stromal cell infiltration scores significantly correlate with RPRD1B expression.

cantly higher than in the normal control group ($P < 0.01$, **Figure 5B**). To validate the analysis result, we incorporated a cohort, GSE10846, from the Gene Expression Omnibus (GEO) database. Following stringent data preprocessing, we extracted overall survival (OS) data for the subjects within the cohort. Stratified analysis based on RPRD1B expression consistently demonstrated that patients with elevated RPRD1B expression had less favorable prognostic outcomes. Notably, there was a significant disparity in the median survival times between the respective subgroups, as illustrated in **Figure 5C**.

RPRD1B inhibits the tumor immune cells in DLBCL and promotes lymphoma cell growth

Utilizing the CIBERSORT platform, we analyzed the correlation between the expression of the RPRD1B gene and various immune cells in DLBCL. The results revealed that the expression of RPRD1B in DLBCL was significantly negatively correlated with CD4⁺ T cells, Th1 cells, and major histocompatibility complex (MHC)-restricted antigen-presenting cells. At the same time, it demonstrated a significant positive correlation with Th2 cells (**Figure 5D**).

We knocked down the expression of endogenous RPRD1B in SU-DHL-2 and overexpressed the RPRD1B protein in SU-DHL-4 using a lentivirus vector to observe the effect of RPRD1B on cell growth. The results indicated that the clonogenic capacity of SU-DHL-2 cells with RPRD1B knockdown was also markedly reduced (**Figure 5E**). To explore the mechanism by which elevated RPRD1B expression enhances the growth ability of lymphoma cells, we examined its effect on cell apoptosis. Compared to wild-type cells, apoptosis decreased in SU-DHL-4 cells overexpressing RPRD1B (**Figure 5F**).

Discussion

Although studies on RPRD1B have shown that it is closely related to tumor progression [1, 29, 30], further studies across pan-cancers have yet to be conducted. This study utilized public databases to systematically analyze the distribution of RPRD1B as a nuclear protein in normal tissues and organs. The results showed that it commonly exists in normal tissues and organs, and is especially highly expressed in the muscular system and blood lymphatic tissue, suggesting that it may have essential

physiologic functions [29]. In addition, RPRD1B also has an elevated expression in most tumor tissues. From the 30 types of tumors analyzed in the TCGA pan-cancer database, RPRD1B displayed significant expression deviations in 25 tumors compared to normal tissues: 16 were upregulated and nine downregulated. Correlation analysis between RPRD1B expression levels and tumor prognoses revealed some significant correlations in the TCGA database. The validated unfavorable prognostic associations of RPRD1B expression in LIHC, ESCA, BRCA, HNSC, THCA, KIRP, and UCEC, and favorable associations in STAD, ESCC, and KIRC, were achieved using Kaplan-Meier Plotter. The results suggest that RPRD1B plays diverse roles in the pathogenesis and development of different tumors.

Mutations in cancer-related genes can also lead to abnormalities in cell function. The result of mutation analysis showed that mutations in the CID, Linker, and CCT structural regions of RPRD1B protein can occur in 20 types of tumors, with the lowest mutation rate being in pan-renal cancer KIPAN at 0.1% and the highest being in UCEC at 3.4%, implying that the mutation rate of the RPRD1B coding region varies by more than 30 times in different tumors. Still, its clinical significance for prognosis remains unclear. CNV is widely present in the human genome, which can lead to the activation of oncogenes or the inactivation of tumor suppressor genes, resulting in the occurrence and development of tumors. In different CNV samples in each tumor, significant differences expression of RPRD1B were observed in 15 tumors over 20 cancer types, suggesting that RPRD1B may affect genome CNV formation [31-33].

MSI and TMB, associated with DNA damage and mismatch repair defects, are important tumor markers in clinical practice, reflecting the ability of tumors to produce new antigens and can serve as predictive indicators for immunotherapy efficacy [34-36]. Among 37 tumor datasets, RPRD1B expression significantly correlated with MSI in 13 tumors and TMB in five tumors, respectively. The previous research from our and other labs has confirmed that RPRD1B participates in DNA break and mismatch repair through non-homologous end joining (NHEJ), homologous recombination (HR), and mismatch repair (MMR), which may be the intrinsic mechanism related to chromo-

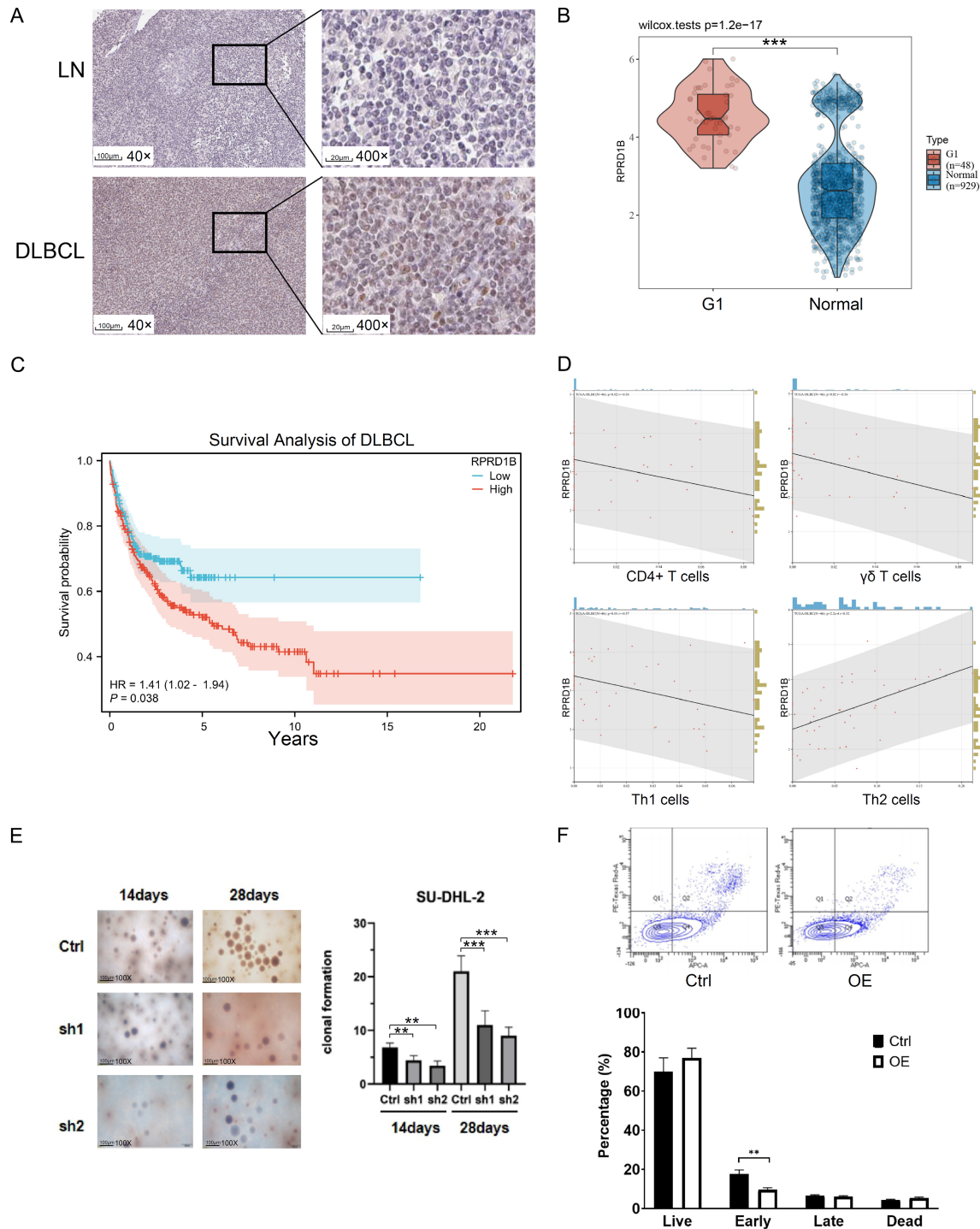


Figure 5. Comparative and correlational analysis of RPRD1B expression and effect on survival and cellular behavior in DLBCL. (A) The expression of RPRD1B is higher in the lymphoma (lower) than in the lymph nodes (upper). Data sourced from the HPA database. (B) RPRD1B expression is significantly elevated in DLBCL compared to normal lymphoid tissues, based on data from the TCGA and GTEx databases. (C) Correlation between overall survival (OS) and RPRD1B expression in DLBCL (GSE10846). (D) RPRD1B was significantly negatively associated with CD4⁺ T, $\gamma\delta$ T, and Th1 cells and positively associated with Th2 cells. (E) RPRD1B promotes colony formation of lymphoma cells in soft agar. (F) Overexpression of RPRD1B in SU-DHL-4 inhibits cell apoptosis. Statistical significance is indicated as ** $P < 0.01$ and *** $P < 0.001$; statistical analyses were performed using the Wilcoxon rank-sum test (B) and independent-samples t-test (E, F). The exact P value for (B) is shown in the figure.

somal instability and heterogeneity, and its expression level may affect the efficacy of tumor immune checkpoint therapy [29].

The immune cells play a significant role in the outcome of tumors, and RPRD1B may cause poor prognosis in patients by promoting the growth of tumor cells and inhibiting anti-tumor immunity. Using the ESTIMATE immune scoring system, RPRD1B gene expression and immune cell infiltration (ImmuneScore) correlation analysis across various tumor types displayed a significant correlation in 30 cancers, predominantly in a negative direction, suggesting that RPRD1B inhibits immune cell infiltration and tumor immunity function. About the stromal cell infiltration score (StromalScore), out of 20 cancers, six showed significant positive correlations with RPRD1B expression, and 14 displayed negative correlations, which may be a bidirectional regulation in different tumors.

Current research into RPRD1B's role in hematologic cancers is limited. Our analysis revealed significantly higher RPRD1B expression in DLBCL tissues than normal lymph nodes. Patients with high RPRD1B expression had notably poorer prognoses than those with lower levels (see **Figure 5C**). We investigated how RPRD1B expression affects different immune cells in DLBCL. Our findings indicate that high RPRD1B levels are associated with decreased in CD4⁺ T cells, Th1 cells, and $\gamma\delta$ T cells while Th2 cell numbers increase. This shift may lead to a Th2-dominant environment that suppresses cytotoxic T and NK cells, enhances M2 macrophage polarization, and ultimately undermines cell-mediated immune responses, promoting tumor escape. This mechanism underscores the potential of RPRD1B as a target for immune therapy.

To assess and verify its potential to promote the growth of lymphoma cells, we overexpressed the RPRD1B protein in the SU-DHL-4 cell line and knocked down the expression of RPRD1B in the SU-DHL-2 cell line. After the RPRD1B expression level in SU-DHL-2 decreased, the cellular colony formation ability decreased significantly, indicating that RPRD1B indeed promotes the establishment of the biological phenotype of lymphoma cells.

To explore the mechanism by which RPRD1B expression promotes the growth of lymphoma cells, cell apoptosis analysis was conducted.

The results showed that apoptosis in SU-DHL-4 cells overexpressing the RPRD1B protein significantly decreased compared to wild type cells, indicating that RPRD1B might promote tumor growth by inhibiting the cell apoptosis pathway.

Several limitations should be acknowledged. First, the bioinformatics analyses of public datasets establish statistical correlations between RPRD1B and tumor prognosis, genomic heterogeneity, and immune infiltration, but do not directly prove causation. Second, *in vitro* validation was performed in only two DLBCL cell lines (SU-DHL-2 and SU-DHL-4), and confirmation in additional cell lines and *in vivo* models is warranted. Third, the best-cutoff stratification used in **Figure 2B** may overestimate statistical significance relative to median-based analysis (**Figure 2A**); concordant findings between the two strategies should therefore be considered the most reliable. Although our study links RPRD1B with tumor immune resistance, further research is needed to elucidate the exact mechanisms in different tumor microenvironments and to develop effective immune-targeted therapies against RPRD1B. Additionally, our conclusions are based on retrospective analyses and require validation through prospective clinical trials with broader patient samples.

Conclusion

Our study demonstrated that RPRD1B is broadly overexpressed across multiple cancer types and was significantly associated with poor prognosis in several malignancies, particularly in DLBCL. RPRD1B not only correlates with genomic heterogeneity and reduced immune cell infiltration but also directly promotes lymphoma cell proliferation and suppresses apoptosis. Mechanistically, RPRD1B appears to shift the tumor immune microenvironment toward an immunosuppressive Th2-dominant profile, facilitating tumor immune evasion. These findings suggest that RPRD1B is not only a promising prognostic biomarker but also a potential therapeutic target for pan-cancers, especially for DLBCL. Further prospective clinical and mechanistic studies are warranted to validate its role and application in precision oncology.

Acknowledgements

The authors thank Weifeng Lu and Song Chen for assisting with project management and

financial reimbursement. The authors also sincerely acknowledge the TCGA and GEO databases for providing their platforms and the contributors for uploading meaningful datasets. This research was funded by Hainan Provincial Natural Science Foundation of China, 820QN-401, 822QN468; Science and Technology special fund of Hainan Province, China, ZDYF2024-SHFZ114; Health Science and Technology Innovation Joint Project of Hainan Province, China, WSJK2024MS231; and supported by Hainan Province Clinical Medical Center Construction Project, [2022] 276.

Disclosure of conflict of interest

None.

Address correspondence to: Drs. Xinbao Hao and Lu Xu, Department of Hematology, The First Affiliated Hospital of Hainan Medical University, 31 Longhua Road, Haikou 570102, Hainan, China. Tel: +86-898-66528100; E-mail: haoxinbao@tsinghua.org.cn (XBH); xulu2324@163.com (LX)

References

- [1] Lu D, Wu Y, Wang Y, Ren F, Wang D, Su F, Zhang Y, Yang X, Jin G, Hao X, He D, Zhai Y, Irwin DM, Hu J, Sung JJ, Yu J, Jia B and Chang Z. CREPT accelerates tumorigenesis by regulating the transcription of cell-cycle-related genes. *Cancer Cell* 2012; 21: 92-104.
- [2] Liu J, Liu H, Zhang X, Gao P, Wang J and Hu Z. Identification and characterization of P15RS, a novel P15(INK4b) related gene on G1/S progression. *Biochem Biophys Res Commun* 2002; 299: 880-885.
- [3] Ni Z, Olsen JB, Guo X, Zhong G, Ruan ED, Marcon E, Young P, Guo H, Li J, Moffat J, Emili A and Greenblatt JF. Control of the RNA polymerase II phosphorylation state in promoter regions by CTD interaction domain-containing proteins RPRD1A and RPRD1B. *Transcription* 2011; 2: 237-242.
- [4] Ni Z, Xu C, Guo X, Hunter GO, Kuznetsova OV, Tempel W, Marcon E, Zhong G, Guo H, Kuo WW, Li J, Young P, Olsen JB, Wan C, Loppnau P, El Bakkouri M, Senisterra GA, He H, Huang H, Sidhu SS, Emili A, Murphy S, Mosley AL, Arrow-smith CH, Min J and Greenblatt JF. RPRD1A and RPRD1B are human RNA polymerase II C-terminal domain scaffolds for Ser5 dephosphorylation. *Nat Struct Mol Biol* 2014; 21: 686-695.
- [5] Jiang J, Yang X, He X, Ma W, Wang J, Zhou Q, Li M and Yu S. MicroRNA-449b-5p suppresses the growth and invasion of breast cancer cells via inhibiting CREPT-mediated Wnt/beta-catenin signaling. *Chem Biol Interact* 2019; 302: 74-82.
- [6] Jin K, Chen H, Zuo Q, Huang C, Zhao R, Yu X, Wang Y, Zhang Y, Chang Z and Li B. CREPT and p15RS regulate cell proliferation and cycling in chicken DF-1 cells through the Wnt/beta-catenin pathway. *J Cell Biochem* 2018; 119: 1083-1092.
- [7] Wei M, Cao Y, Jia D, Zhao H and Zhang L. CREPT promotes glioma cell proliferation and invasion by activating Wnt/beta-catenin pathway and is a novel target of microRNA-596. *Biochimie* 2019; 162: 116-124.
- [8] Zhai W, Ye X, Wang Y, Feng Y, Wang Y, Lin Y, Ding L, Yang L, Wang X, Kuang Y, Fu X, Eugene Chin Y, Jia B, Zhu B, Ren F and Chang Z. CREPT/RPRD1B promotes tumorigenesis through STAT3-driven gene transcription in a p300-dependent manner. *Br J Cancer* 2021; 124: 1437-1448.
- [9] Zhang Y, Liu C, Duan X, Ren F, Li S, Jin Z, Wang Y, Feng Y, Liu Z and Chang Z. CREPT/RPRD1B, a recently identified novel protein highly expressed in tumors, enhances the beta-catenin. TCF4 transcriptional activity in response to Wnt signaling. *J Biol Chem* 2014; 289: 22589-22599.
- [10] Zhang Y, Wang S, Chang Z, Sung J and Yu J. 674 CREPT plays an oncogenic role in colorectal cancer through promoting Wnt/ β -catenin pathway via enhancing acetylation of β -catenin. *Gastroenterology* 2016; 150: S138.
- [11] Zhang Y, Wang S, Kang W, Liu C, Dong Y, Ren F, Wang Y, Zhang J, Wang G, To KF, Zhang X, Sung JJ, Chang Z and Yu J. CREPT facilitates colorectal cancer growth through inducing Wnt/beta-catenin pathway by enhancing p300-mediated beta-catenin acetylation. *Oncogene* 2018; 37: 3485-3500.
- [12] Li W, Zheng G, Xia J, Yang G, Sun J, Wang X, Wen M, Sun Y, Zhang Z and Jin F. Cell cycle-related and expression-elevated protein in tumor overexpression is associated with proliferation behaviors and poor prognosis in non-small-cell lung cancer. *Cancer Sci* 2018; 109: 1012-1023.
- [13] Liu H, Seynhaeve ALB, Brouwer RWW, van IJcken WFJ, Yang L, Wang Y, Chang Z and Ten Hagen TLM. CREPT promotes melanoma progression through accelerated proliferation and enhanced migration by rhoa-mediated actin filaments and focal adhesion formation. *Cancers (Basel)* 2019; 12: 33.
- [14] Siril YJ, Kouketsu A, Saito H, Takahashi T and Kumamoto H. Immunohistochemical expression levels of cyclin D1 and CREPT reflect the course and prognosis in oral precancerous lesions and squamous cell carcinoma. *Int J Oral Maxillofac Surg* 2022; 51: 27-32.

- [15] Wen N, Bian L, Gong J and Meng Y. Overexpression of cell-cycle related and expression-elevated protein in tumor (CREPT) in malignant cervical cancer. *J Int Med Res* 2020; 48: 300060519895089.
- [16] Yang G, Wang Y, Xiao J, Zhao F, Qiu J, Liu Y, Chen G, Cao Z, You L, Zheng L, Zhang T and Zhao Y. CREPT serves as a biomarker of poor survival in pancreatic ductal adenocarcinoma. *Cell Oncol (Dordr)* 2021; 44: 345-355.
- [17] Yu S, Huang H, Wang S, Xu H, Xue Y, Huang Y, He J, Xu X, Wu Z, Wu J, Zhang Y, Huang Q, Chang Z, Li E and Xu L. CREPT is a novel predictor of the response to adjuvant therapy or concurrent chemoradiotherapy in esophageal squamous cell carcinoma. *Int J Clin Exp Pathol* 2019; 12: 3301-3310.
- [18] Zhang Z, Shao L, Wang Y and Luo X. MicroRNA-501-3p restricts prostate cancer growth through regulating cell cycle-related and expression-elevated protein in tumor/cyclin D1 signaling. *Biochem Biophys Res Commun* 2019; 509: 746-752.
- [19] Sun M, Si G, Sun HS and Si FC. Inhibition of CREPT restrains gastric cancer growth by regulation of cycle arrest, migration and apoptosis via ROS-regulated p53 pathway. *Biochem Biophys Res Commun* 2018; 496: 1183-1190.
- [20] Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, Kovatich AJ, Benz CC, Levine DA, Lee AV, Omberg L, Wolf DM, Shriver CD, Thorsson V; Cancer Genome Atlas Research Network; Hu H. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 2018; 173: 400-416 e411.
- [21] Andersen PK and Gill RD. Cox's regression model for counting processes: a large sample study. *Ann Statist* 1982; 10: 1100-1120.
- [22] Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES and Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013; 31: 213-219.
- [23] Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhi R and Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol* 2011; 12: R41.
- [24] Bonneville R, Krook MA, Kautto EA, Miya J, Wing MR, Chen HZ, Reeser JW, Yu L and Roychowdhury S. Landscape of microsatellite instability across 39 cancer types. *JCO Precis Oncol* 2017; 2017: PO.17.00073.
- [25] Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB and Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* 2013; 4: 2612.
- [26] Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M and Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015; 12: 453-457.
- [27] Zeng D, Ye Z, Shen R, Yu G, Wu J, Xiong Y, Zhou R, Qiu W, Huang N, Sun L, Li X, Bin J, Liao Y, Shi M and Liao W. IOBR: multi-omics immuno-oncology biological research to decode tumor microenvironment and signatures. *Front Immunol* 2021; 12: 687975.
- [28] Vivian J, Rao AA, Nothhaft FA, Ketchum C, Armstrong J, Novak A, Pfeil J, Narkizian J, Deran AD, Musselman-Brown A, Schmidt H, Amstutz P, Craft B, Goldman M, Rosenbloom K, Cline M, O'Connor B, Hanna M, Birger C, Kent WJ, Patterson DA, Joseph AD, Zhu J, Zaranek S, Getz G, Haussler D and Paten B. Toil enables reproducible, open source, big biomedical data analyses. *Nat Biotechnol* 2017; 35: 314-316.
- [29] Li M, Ma D and Chang Z. Current understanding of CREPT and p15RS, carboxy-terminal domain (CTD)-interacting proteins, in human cancers. *Oncogene* 2021; 40: 705-716.
- [30] Yeh IJ, Motea EA and Boothman DA. Kub5-Her-RPRD1B regulates CDK1 expression through the DREAM complex. *FASEB J* 2019; 33: 458.419.
- [31] Calabresi L, Carretta C, Romagnoli S, Rotunno G, Parenti S, Bertesi M, Bartalucci N, Rontauroli S, Chiereghin C, Castellano S, Gentili G, Maccari C, Vanderwert F, Mannelli F, Della Porta M, Manfredini R, Vannucchi AM and Guglielmelli P. Clonal dynamics and copy number variants by single-cell analysis in leukemic evolution of myeloproliferative neoplasms. *Am J Hematol* 2023; 98: 1520-1531.
- [32] Guglielmelli P, Calabresi L, Carretta C, Rotunno G, Parenti S, Romagnoli S, Mallia S, Bartalucci N, Rontauroli S, Chiereghin C, Castellano S, Mannelli F, Della Porta MG, Manfredini R and Vannucchi A. Single cell mutation analysis delineates clonal architecture in leukemic transformation of myeloproliferative neoplasms. *Blood* 2021; 138: 56-56.
- [33] Li Y, Porta-Pardo E, Tokheim C, Bailey MH, Yaron TM, Stathias V, Geffen Y, Imbach KJ, Cao S, Anand S, Akiyama Y, Liu W, Wyczalkowski MA, Song Y, Storrs EP, Wendl MC, Zhang W, Sibai M, Ruiz-Serra V, Liang WW, Terekhanova NV, Rodrigues FM, Clauser KR, Heiman DI, Zhang Q, Aguet F, Calinawan AP, Dhanasekaran SM, Birger C, Satpathy S, Zhou DC, Wang LB, Baral J, Johnson JL, Huntsman EM, Pugliese P, Colaprico A, Iavarone A, Chheda MG,

- Ricketts CJ, Fenyő D, Payne SH, Rodriguez H, Robles AI, Gillette MA, Kumar-Sinha C, Lazar AJ, Cantley LC, Getz G and Ding L; Clinical Proteomic Tumor Analysis Consortium. Pan-cancer proteogenomics connects oncogenic drivers to functional states. *Cell* 2023; 186: 3921-3944 e3925.
- [34] Bakhoum SF and Cantley LC. The multifaceted role of chromosomal instability in cancer and its microenvironment. *Cell* 2018; 174: 1347-1360.
- [35] Jiang M, Jia K, Wang L, Li W, Chen B, Liu Y, Wang H, Zhao S, He Y and Zhou C. Alterations of DNA damage response pathway: Biomarker and therapeutic strategy for cancer immunotherapy. *Acta Pharm Sin B* 2021; 11: 2983-2994.
- [36] Palmeri M, Mehnert J, Silk AW, Jabbour SK, Ganesan S, Popli P, Riedlinger G, Stephenson R, de Meritens AB, Leiser A, Mayer T, Chan N, Spencer K, Girda E, Malhotra J, Chan T, Subbiah V and Groisberg R. Real-world application of tumor mutational burden-high (TMB-high) and microsatellite instability (MSI) confirms their utility as immunotherapy biomarkers. *ESMO Open* 2022; 7: 100336.