Article

Discovery Medicine 2026; 38(210): 2110–2131
https://doi.org/10.24976/Discov.Med.202638210.165

RHBDF2 Is Correlated With Immune Infiltrates and Shows Prognostic Association in Acute Myeloid Leukemia

Qianni Jin[1,2], Zhifang Xu[1], Jianlan Li[2], Yan Huang[2], Huichao Wang[1], Bianbian Qiao[1], Shuo Li[1], Zhilei Jin[1], Yaofang Zhang[1], Fanggang Ren[1], Huijuan Hao[1], Hongwei Wang[1,3,*]

[1]Key Laboratory of Molecular Diagnosis and Treatment of Hematological Diseases of Shanxi Province, Institute of Hematology, The Second Hospital of Shanxi Medical University, 030001 Taiyuan, Shanxi, China
[2]Department of Internal Medicine Laboratory, The Second Hospital of Shanxi Medical University, 030001 Taiyuan, Shanxi, China
[3]College of Basic Medical Sciences, Shanxi University of Chinese Medicine, 030600 Jinzhong, Shanxi, China
*Correspondence: wanghw68@hotmail.com (Hongwei Wang)
Submitted: 1 March 2026 Revised: 14 May 2026 Accepted: 21 May 2026 Published: 23 July 2026

Background: Acute myeloid leukemia (AML) is a heterogeneous disease with a complex immunosuppression microenvironment, necessitating the identification of preferable biomarkers to enhance the efficacy of immunotherapy. Rhomboid 5 homolog 2 (*RHBDF2*), through its interaction with A Disintegrin And Metalloprotease 17 (ADAM17), has been implicated in immune responses and tumor progression; however, its role in AML remains insufficiently investigated.

Methods: This investigation delineated the expression profile of *RHBDF2* and its correlation with the clinical features of AML using The Cancer Genome Atlas (TCGA, n = 150 AML patients) and Genotype-Tissue Expression (GTEx, n = 337 non-malignant control individuals) databases. The RNA and protein expression levels of *RHBDF2* were validated in a single-center clinical cohort (105 bone marrow fluid samples and 125 bone marrow biopsy specimens) via quantitative reverse-transcription polymerase chain reaction (qRT-PCR) and immunohistochemistry (IHC). To elucidate the correlation between *RHBDF2* and the complex immune regulatory microenvironment of AML, we performed functional enrichment analyses (Gene Ontology [GO] and Kyoto Encyclopedia of Genes and Genomes [KEGG]), Gene Set Variation Analysis (GSVA), immune cell infiltration analysis, and single-cell sequencing t-distributed stochastic neighbor embedding (t-SNE) clustering. Furthermore, we assessed the prognostic value of *RHBDF2* using Kaplan-Meier survival analysis, Cox regression, and a prognostic nomogram model.

Results: Our findings indicated that *RHBDF2* was upregulated in AML compared to non-malignant control samples. Moreover, *RHBDF2* expression was associated with immune infiltration and exhibited a positive correlation with several immunological checkpoints including T-cell Immunoglobulin and Mucin-domain containing molecule 3 (TIM3), Programmed Cell Death 1 Ligand 2 (PD-L2), and Herpesvirus Entry Mediator (HVEM), with significant enrichment in the monocytes of AML patients. Multivariate Cox regression analysis revealed that elevated *RHBDF2* shows a marginally significant association with overall survival (OS) in AML patients (adjusted Hazard Ratio [HR] = 1.01, 95% Confidence Interval [CI] 1.00–1.02, $p = 0.048$). Kaplan-Meier analysis showed that high *RHBDF2* expression was associated with significantly shorter OS (HR = 1.87, 95% CI 1.22–2.86, $p = 0.004$). The constructed nomogram model exhibited preliminary prognostic stratification capacity for OS, with time-dependent receiver operating characteristic (ROC) curve analysis yielding area under the curve (AUC) values of 75.2, 80.1, and 95.3 for 1-year, 3-year, and 5-year OS, respectively, with a concordance index (C-index) of 0.76.

Conclusion: These correlative findings support that *RHBDF2* may be associated with prognosis in AML but requires further validation, offering insights for prognosis prediction and the development of innovative immunotherapy strategies for AML.

Keywords: acute myeloid leukemia; immune infiltration; prognosis; *RHBDF2*; biomarker

Introduction

Acute myeloid leukemia (AML) is characterized by complex genetic mutations and cytogenetic abnormalities, constituting an aggressive malignancy distinguished by significant heterogeneity, variable prognosis, and high fatality rates [1]. Current therapeutic interventions for AML are insufficient, with traditional chemotherapy regimens and hematopoietic stem cell transplantation having contributed modestly to enhanced patient survival over past decades [2]. Advances in targeted therapies, including fms-like tyrosine kinase 3 (FLT3) inhibitors and isocitrate dehydrogenase 1 and 2 (IDH1/IDH2) inhibitors, have introduced personalized treatment modalities for AML, contributing to prolonged survival to a certain extent [3]. However, existing therapeutic strategies still fall short of achieving satisfac-

tory outcomes. Due to primary resistance to initial treatment, therapy-related toxicity, and disease relapse, approximately 70% of adult AML patients in high-income countries do not survive [4,5]. Thus, there is an urgent need to develop breakthrough curative approaches to further improve treatment efficacy for AML patients.

The employment of immunotherapy has revolutionized the treatment of various cancers [6]. Immunotherapeutic strategies, particularly the employment of antibodies against inhibitory immune checkpoints and chimeric antigen receptor (CAR) T-cell therapy, have yielded substantial advancements in cancer care [7,8]. Regrettably, trials pertaining to T-cell-based immunotherapy in adult AML have yielded underwhelming outcomes [9,10]. The malignant cells of AML afflict T-cell and NK-cell functionality through the presentation of inhibitory ligands, they sidestep immune surveillance through down-modulating MHC-I/II molecules, they forestall phagocyte-mediated elimination by presenting CD47, and they drive myeloid-derived suppressor cell (MDSC) differentiation through the release of soluble factors like arginase, indoleamine 2,3-dioxygenase (IDO and NOS), thereby inducing immune escape [11]. In conclusion, the distinct immunosuppressive milieu and T-cell exhaustion markedly impair the efficacy of AML immunotherapy [12]. On the other hand, existing immune checkpoint inhibitors (ICIs) present significant clinical challenges in terms of toxicity and efficacy. Hence, a comprehensive characterization of the immunosuppressive mechanisms in AML, the identification of intervention targets to "warm" immunologically cold tumors, and the discovery of novel immune checkpoints are quintessential for the progression of AML diagnosis, prognostic stratification, monitoring of minimal residual disease (MRD), prediction of treatment response, and targeted drug development.

Rhomboid 5 Homolog 2 (*RHBDF2*), also known as iRhom2, is an inactive member of the rhomboid protein family [13]. It acts as an essential regulatory chaperone for A Disintegrin And Metalloprotease 17 (ADAM17), a key sheddase that mediates the cleavage and release of multiple inflammatory cytokines, chemokines, and immune checkpoint ligands, including TNF-α, EGFR, IL-6, and Notch signaling pathways [14–16]. Critically, *RHBDF2* is predominantly expressed in myeloid lineage cells, including monocytes, macrophages, and myeloid-derived suppressor cells (MDSCs), and has been identified as a master regulator of myeloid cell activation and inflammatory responses in multiple immune-related diseases [13,17–20]. In the tumor microenvironment, *RHBDF2* has been shown to modulate the immunosuppressive phenotype of tumor-associated macrophages (TAMs) and MDSCs in solid tumors, and is correlated with immune evasion and poor prognosis in hepatocellular carcinoma and lung cancer [21–24]. Given that myeloid cell-mediated immunosuppression is a core mechanism of immune evasion and immunotherapy resistance in AML, we hypothesize that *RHBDF2* may play a role in AML progression by modulating the myeloid cell compartment and potentially immunosuppressive bone marrow microenvironment. However, until now, the expression, prognostic value, and immune correlation of *RHBDF2* in AML have remained elusive.

This research investigates *RHBDF2* expression in AML patients and non-malignant control individuals by utilizing TCGA and GTEx databases, further examining its relationship with clinical features, immune cell infiltration, and patient prognosis. Quantitative PCR and immunohistochemistry were utilized to evaluate *RHBDF2* expression in the bone marrow of AML patients across various disease states. In addition, single-cell sequencing technology was employed to delineate *RHBDF2* expression distribution across AML cell subpopulations, providing preliminary correlative evidence and a testable hypothesis for future functional studies exploring the potential of *RHBDF2* as an immunotherapeutic target in AML.

Methods

Database

In this research, the raw RNA sequencing data and corresponding clinical features of AML patients were retrieved from the The Cancer Genome Atlas (TCGA)-LAML cohort via the UCSC Xena Browser (https://xenabrowser.net/), a public functional genomics explorer that allows standardized access and exploration of TCGA and GTEx genomic and phenotypic datasets. The gene expression matrix was in the format of Transcripts Per Million (TPM) normalized counts, generated from the STAR alignment pipeline. For non-malignant control samples, whole bone marrow RNA-seq data from non-malignant control individuals were retrieved from the GTEx database via the UCSC Xena Browser, with the same TPM normalized expression matrix format to ensure data consistency between the two cohorts. The preprocessing pipeline for the TCGA and GTEx datasets was as follows: (1) Uniform gene ID conversion: ENSEMBL gene IDs were converted to official HUGO gene symbols using the org.Hs.eg.db R package, with the probe with the highest average expression value retained for genes with multiple corresponding ENSEMBL IDs; (2) Log2 transformation: log2(TPM+1) transformation was applied to all expression values to reduce the impact of extreme expression outliers; (3) Low-expression gene filtering: genes detected in fewer than 10% of the total samples were filtered out to reduce statistical noise in downstream analysis; (4) Batch effect correction: the ComBat-seq function from the sva R package was used to correct for technical batch effects between the TCGA and GTEx datasets, to eliminate non-biological technical variation between the two cohorts. Notably, both the TCGA AML samples and GTEx non-malignant control samples were derived from whole bone marrow tissue, ensuring consistent tissue

origin between the two cohorts for differential expression analysis.

Moreover, the mRNA expression profile of *RHBDF2* in pan-cancer and AML FAB subtypes was visualized using the Gene Expression Profiling Interactive Analysis (GEPIA) database (http://gepia.cancer-pku.cn) and the University of Alabama at Birmingham Cancer data analysis Portal (UALCAN) database (http://ualcan.path.uab.edu/index.html), both of which are based on the same TCGA and GTEx datasets and were used solely for supplementary visualization, not independent validation. Additionally, two independent AML datasets (GSE12417 and GSE71014) were obtained from the Gene Expression Omnibus (GEO) database to corroborate the primary findings from the TCGA analysis (as illustrated in **Supplementary Figs. 1,2**). The raw gene expression matrix was downloaded, and probe IDs were converted to official HUGO gene symbols using the corresponding platform annotation file. For multiple probes mapping to the same gene symbol, the probe with the highest average expression value across all samples was selected as the representative for that gene. The expression matrix was then normalized using the quantile normalization method, and log2 transformation was performed for all expression values. Batch effects within each dataset were corrected using the ComBat function from the sva R package. Low-expression genes (detected in $<10\%$ of total samples) were filtered out prior to downstream analysis. These two datasets were external validation cohort to validate the prognostic and immune correlation findings from the TCGA discovery cohort (as illustrated in **Supplementary Fig. 1**). The GEO external validation cohorts utilized microarray platforms, which differ from the RNA-seq platform used for the TCGA discovery cohort. These platform differences may introduce technical variation that could affect the generalizability of our findings. The single-cell RNA sequencing data utilized in this study were derived from two independent public datasets: GSE116256 and GSE289435. GSE116256 dataset contains bone marrow scRNA-seq data from 15 patients with newly diagnosed de novo AML (non-APL) and 5 non-malignant control individuals. The data were generated using the 10x Genomics Chromium platform, with a total of 38,910 high-quality cells included after quality control filtering. This dataset was used as the primary discovery cohort for cell type clustering, annotation, and analysis of *RHBDF2* expression distribution across different cell populations in AML. GSE289435 dataset contains bone marrow scRNA-seq data from 7 patients with newly diagnosed de novo AML. The data were generated using the 10x Genomics Chromium platform, with a total of 22,648 high-quality cells included after quality control filtering. This dataset was used as an external validation cohort to confirm the cell type distribution pattern of *RHBDF2* expression observed in the GSE116256 dataset.

Differentially Expressed Gene (DEG) Analysis

This study extracted RNA-seq data of AML and normal volunteers from the TCGA and GTEx databases for statistical analysis of differentially expressed genes (DEGs). For gene annotation conversion, ENSEMBL gene IDs from the TCGA and GTEx datasets were mapped to official HUGO gene symbols using the org.Hs.eg.db R package. In cases where multiple ENSEMBL IDs corresponded to the same official gene symbol, we retained the ENSEMBL ID with the highest average expression value across all samples as the representative for that gene, and removed all other redundant ENSEMBL IDs for the same gene. This strategy ensures that each gene symbol is represented by a single expression value in the final matrix, avoiding duplicate gene entries in downstream differential expression analysis. Prior to differential expression analysis, we performed prefiltering of low-expression genes to reduce statistical noise. Genes that were detected (expression value >0) in fewer than 10% of the total samples in the analysis cohort were filtered out and excluded from downstream DEG analysis.

Two independent differential expression analyses were performed in this study, with consistent screening thresholds: Initial DEG screening between AML and non-malignant control samples: The screening criteria were set as raw Fold Change (FC) >2 or raw FC <0.5 (equivalent to $|\log2FC| >1$), and FDR <0.05 [25]. Here, raw FC refers to the original fold change calculated based on untransformed linear-scale TPM expression values, and log2FC is the value obtained by log2 transformation of the raw FC. The limma-voom pipeline was used for differential expression analysis based on the log2(TPM+1) transformed expression matrix, and the reported raw FC values were generated by inverse transformation of the model-estimated log2FC values to maintain the linear-scale biological interpretation of fold change. DEG screening between the high and low *RHBDF2* expression groups (per the median *RHBDF2* expression value) in the TCGA AML cohort: The screening criteria were set as raw FC >2 or raw FC <0.5, and FDR <0.05.

The Benjamini-Hochberg (BH) method was used for FDR correction (multiple testing correction) for all differential expression analyses, to strictly control the false discovery rate of DEG identification. The pheatmap R package was used to generate heatmaps for visualization of the identified DEGs. Given the sample size imbalance between the TCGA AML cohort ($n = 150$) and GTEx control cohort ($n = 337$), we performed a sensitivity analysis by randomly downsampling the GTEx cohort to match the AML cohort size for 100 iterations. However, we emphasize that the primary limitation of this comparison is the fundamental difference in cellular composition between leukemic and non-malignant bone marrow, which cannot be fully addressed by downsampling.

Functional Enrichment Analysis

This study utilized the ClusteProfiler package [26] in R to conduct a systematic analysis of the obtained differentially expressed genes (DEGs) from two dimensions: Gene Ontology (GO) function and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. The GO analysis covered three major functional annotations: Cellular Component (CC), Molecular Function (MF), and Biological Process (BP), while the KEGG analysis focused on revealing the enrichment of related pathways. The significance threshold for enriched terms was set as adjusted p-value < 0.05, with the Benjamini-Hochberg (BH) method used for multiple testing correction (FDR correction) to control for false positive results from multiple comparisons.

Clinical Sample Collection

This study was approved by the Ethics Committee of the Second Hospital of Shanxi Medical University (Ethics Approval No.2023-YX-255), and all subjects signed written informed consent before sample collection, in accordance with the Declaration of Helsinki.

A total of 306 clinical bone marrow fluid samples and bone marrow biopsy samples were collected from the center, including AML patients at different disease stages and non-malignant controls.

Inclusion criteria for AML patients: (1) De novo AML diagnosed per the 2016 NCCN AML Guidelines (2nd Edition); (2) Aged 18–75 years; (3) Complete clinical and follow-up data available.

Exclusion criteria for AML patients: (1) Secondary/therapy-related AML, or AML transformed from MDS/MPN; (2) Concurrent malignancies; (3) Severe autoimmune disease or active infection; (4) Prior anti-tumor treatment before sampling.

A total of 139 fresh bone marrow fluid samples (for qRT-PCR: 63 pAML, 26 CR-AML, 16 rAML, 34 non-malignant controls) and 167 FFPE bone marrow biopsy specimens (for IHC: 91 pAML, 25 CR-AML, 9 rAML, 42 non-malignant controls) were collected. All pAML samples were obtained at initial diagnosis pretreatment; CR-AML samples at morphological remission post-consolidation chemotherapy; rAML samples at hematological relapse pre-salvage treatment. There was partial samples overlap between the two cohorts: 42 of the 63 pAML patients in the qRT-PCR cohort had matched FFPE specimens in the IHC cohort, and all other samples were from independent patients.

Non-malignant controls included patients with iron-deficiency anemia (IDA) and megaloblastic anemia (MA), with no hematopoietic malignant clones, concurrent malignancies, autoimmune diseases or active infection. For the qRT-PCR cohort, controls consisted of 21 IDA and 13 MA patients; for the IHC cohort, controls consisted of 25 IDA and 17 MA patients The control cohort was age- and gender-frequency matched with the AML cohort, and all samples were processed with the same standardized procedure as AML samples to minimize technical bias. Our non-malignant control group consisted of patients with IDA and MA, whose bone marrow microenvironment differs from that of healthy individuals. This may introduce biological confounding factors that could affect gene expression and immune-related analyses.

Fresh bone marrow samples were stored in RNAlater at –80 °C for RNA extraction, and FFPE specimens were fixed in 10% neutral buffered formalin for subsequent IHC staining.

Quantitative Reverse-Transcription Polymerase Chain Reaction (qRT-PCR)

Total RNA was extracted from bone marrow mononuclear cells using the Trizol reagent (Invitrogen, Carlsbad, CA, USA) following the manufacturer's standard protocol. RNA concentration and purity were assessed using the NanoDrop One spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA), with only RNA samples with an OD 260/280 ratio of 1.8–2.1 used for subsequent experiments. For cDNA synthesis, 1 µg of total RNA was reverse transcribed into cDNA using the HiScript III 1st Strand cDNA Synthesis Kit (Vazyme Biotech, Nanjing, China) following the manufacturer's instructions.

qRT-PCR was performed using the SYBR Green I fluorescent dye method on the ABI 7500 Real-time PCR System (Thermo Fisher Scientific, Waltham, MA, USA). The 20 µL reaction system contained: 2 µL cDNA template (100 ng/µL), 10 µL 2× ChamQ Universal SYBR qPCR Master Mix (Vazyme Biotech, Nanjing, China), 0.4 µL forward primer (10 µM), 0.4 µL reverse primer (10 µM), and 7.2 µL RNase-free ddH2O. The thermal cycling parameters were as follows: 95 °C for 5 minutes of pre-denaturation; followed by 40 cycles of 95 °C for 10 seconds and 60 °C for 30 seconds. After amplification, melting curve analysis was performed to verify the specificity of the amplification product, with the following program: 95 °C for 15 seconds, 60 °C for 1 minute, and 95 °C for 15 seconds. Each patient sample was analyzed with three technical replicates; independent patient samples served as biological replicates. The relative mRNA expression levels of *RHBDF2* were calculated by the $2^{-\Delta\Delta Ct}$ method, with the GAPDH gene used as the endogenous reference for normalization. The primer sequences were as follows: *RHBDF2* Forward Primer: 5′-GCCCGATGACATCACTAA-3′; *RHBDF2* Reverse Primer: 5′-TGAAAGACCACAGACACGA-3′; GAPDH Forward Primer: 5′-GAAGGTGAAGGTCGGAGTC-3′; GAPDH Reverse Primer: 5′-GAAGATGGTGATGGGATTTC-3′. Statistical analyses and visualization of qRT-PCR results were performed in GraphPad Prism 8.3.0 (San Diego, California, USA). Normality of continuous variables was tested using the Shapiro-Wilk test, and homogeneity of variance was tested using Levene's test. The Student's

t-test was used only when both assumptions were satisfied; otherwise, the non-parametric Mann-Whitney U test was employed to test the significance of the difference between two independent groups. One-way ANOVA followed by Tukey's HSD post-hoc test was used to test the significance of the difference between more than two groups.

Immunohistochemistry (IHC) Staining and Hematoxylin and Eosin (HE) Staining

This study utilized immunohistochemical staining for histological analysis. All FFPE bone marrow biopsy specimens underwent decalcification prior to sectioning, using a 10% ethylenediaminetetraacetic acid (EDTA) decalcification buffer (pH 8.0) at room temperature for 24 hours, with gentle agitation. After decalcification, the specimens were rinsed with running tap water for 2 hours, followed by standard dehydration, paraffin embedding, and sectioning (4 µm thickness). For antigen retrieval, dewaxed and rehydrated tissue sections were immersed in 10 mM sodium citrate buffer (pH 6.0), and subjected to high-temperature high-pressure antigen retrieval for 15 minutes (121 °C, 100 kPa). After heating, the sections were allowed to cool naturally to room temperature for 30 minutes, followed by rinsing with phosphate-buffered saline (PBS). Subsequently, the sections were treated with 3% hydrogen peroxide, blocked, and incubated overnight at 4 °C with a rabbit anti-human *RHBDF2* polyclonal primary antibody (Proteintech Group, Inc., Wuhan, China, Catalog No.: 23181-1-AP) diluted at a ratio of 1:500. After washing, the sections were incubated with a horseradish peroxidase (HRP)-conjugated goat anti-rabbit IgG (H+L) secondary antibody (Proteintech Group, Inc., Wuhan, China, Catalog No.: RGAR001) at 37 °C for 30 minutes, followed by washing to remove non-specific binders. Staining was visualized using 3,3'-diaminobenzidine (DAB) and counterstained with hematoxylin. All IHC staining results were independently evaluated by two qualified clinical pathologists who were blinded to the patient clinical data. The Immunoreactive Score (IRS) system was used for semi-quantitative evaluation of *RHBDF2* staining, which combines the staining intensity and the proportion of positive cells. The staining intensity was scored as follows: 0 = no staining, 1 = weak staining (light yellow), 2 = moderate staining (brown), 3 = strong staining (dark brown). The proportion of positive cells was scored as follows: 0 = 0% positive cells, 1 = 1–25% positive cells, 2 = 26–50% positive cells, 3 = 51–75% positive cells, 4 = 76–100% positive cells. The final IRS score was calculated by multiplying the staining intensity score by the positive cell proportion score, with a possible range of 0 to 12. Positive *RHBDF2* staining was defined as brown-yellow signal localized to the cytoplasm and cell membrane of bone marrow cells. The cutoff value for positive expression was set as an IRS score $\geq$3. For negative controls, the primary *RHBDF2* antibody was replaced with PBS, with all other staining procedures identical to the experimental slides. In cases of scoring discrepancies between the two pathologists, a third senior hematopathologist was consulted to reach a consensus score [27]. Additionally, dewaxed tissue sections were subjected to hematoxylin-eosin staining, followed by standard dehydration and mounting procedures.

Gene Set Variation Analysis (GSVA)

The immune-related gene sets used for GSVA analysis were retrieved from the AmiGO2 portal (http://amigo.geneontology.org/amigo), specifically the Gene Ontology (GO) biological process term "immune system process" (GO:0002376). The selection criteria for the gene set were: (1) human protein-coding genes with well-annotated immune-related functions; (2) genes with detectable expression in the TCGA AML cohort. A total of 2483 immune-related genes were included in the final gene set. GSVA enrichment scores were calculated using the GSVA R package with default parameters, using the Gaussian kernel model for continuous expression data. The input matrix was the log2(TPM + 1) transformed, batch-corrected gene expression matrix from the TCGA AML cohort, with low-expression genes (detected in $<$10% of samples) filtered out prior to analysis, as described in the preprocessing pipeline. Pearson correlation analysis was used to assess the correlation between *RHBDF2* expression and GSVA enrichment scores, as both variables were continuous and conformed to a normal distribution (verified via the Shapiro-Wilk normality test). The Benjamini-Hochberg (BH) method was applied for multiple testing correction of all correlation analyses, with adjusted p-value < 0.05 set as the threshold for statistical significance.

Immune Cell Infiltration Analysis

Immune cell infiltration analysis was performed using the CIBERSORT algorithm, which estimates the relative abundance of 22 human immune cell types from gene expression profiles [28]. The input data for CIBERSORT was the TPM-normalized gene expression matrix from the TCGA AML cohort, with log2 transformation and low-expression gene filtering performed as described in the preprocessing pipeline. The LM22 signature matrix was used as the reference for immune cell type deconvolution. Key parameter settings for CIBERSORT were: 1000 permutations for statistical significance testing, and quantile normalization disabled for RNA-seq data. A total of 136 samples with CIBERSORT output $p < 0.05$ were included in downstream immune infiltration analysis. The Tumor Microenvironment (TME) scores, including immune score, stromal score, ESTIMATE total score, and tumor purity, were calculated for each sample using the ESTIMATE R package, based on the TPM-normalized gene expression matrix. The applicability of the ESTIMATE algorithm in hematological malignancies (including AML) has been validated in multiple published studies, as it can reliably es-

timate the abundance of immune and stromal cells in the bone marrow microenvironment. However, we acknowledge the limitations of this method: it is a computational inference method based on bulk RNA-seq data, and cannot directly quantify the absolute number of immune cells, with potential discrepancies compared to flow cytometry or immunohistochemical validation. For comparisons of immune cell abundance between the high and low *RHBDF2* expression groups (per the median *RHBDF2* expression value), the Wilcoxon rank-sum test was used for pairwise comparisons. The Benjamini-Hochberg (BH) method was applied for multiple testing correction to account for multiple immune cell types being compared, with adjusted *p*-value < 0.05 set as the threshold for statistical significance.

Single Cell Sequence Analysis and Cell Clustering

Single-cell RNA sequencing (scRNA-seq) analysis was performed to characterize the expression distribution of *RHBDF2* across distinct cell types in the acute myeloid leukemia (AML) bone marrow microenvironment, with all analytical procedures conducted using the Seurat R package (version 5.2.1) [29]; two public AML scRNA-seq datasets generated via the 10x Genomics platform (GSE116256 and GSE289435) were downloaded from the Gene Expression Omnibus (GEO) database, and the two datasets were analyzed independently without cross-dataset integration to eliminate inter-study batch effects, with a standardized analytical pipeline applied to each dataset separately. Strict quality control filtering was first performed on the raw gene-cell expression matrix of each dataset, where genes detected in 3 or fewer cells were excluded from downstream analysis, and cells were retained only if they met all predefined high-quality criteria: a minimum of 200 detected genes, total unique molecular identifier (UMI) counts ranging from 200 to 7000, and a mitochondrial UMI proportion not exceeding 20%. For each qualified dataset, independent batch correction and normalization were performed using the SCTransform function in Seurat to mitigate technical variation between individual samples within the dataset; following preprocessing and normalization, principal component analysis (PCA) was conducted for linear dimensionality reduction, the number of significant principal components (PCs) was determined via JackStraw test and ElbowPlot method with a threshold of JackStraw adjusted *p*-value < 0.05, and the top 11 significant PCs were used for downstream cell clustering and non-linear dimensionality reduction for both datasets, with t-distributed stochastic neighbor embedding (t-SNE) applied for 2-dimensional visualization of single-cell transcriptome profiles. Cell clustering was performed using the FindClusters function in Seurat with the resolution parameter set to 0.8, and differentially expressed cluster-specific marker genes were identified using the FindAllMarkers function with the following predefined parameters: test.use = "wilcox" (Wilcoxon rank-sum test), only.pos = TRUE (only positive cluster-specific markers were considered), min.pct = 0.25 (the gene must be detected in at least 25% of cells in the target cluster), and logfc.threshold = 0.25 (minimum log2 fold change of 0.25 between the target cluster and all other cells), with raw *p*-values corrected for multiple testing using the Benjamini-Hochberg (BH) method and an adjusted *p*-value < 0.05 set as the threshold for statistically significant marker genes. Cell type annotation was performed via a two-step approach to ensure high classification accuracy: initial automated cell type assignment was conducted using the SingleR R package by matching the gene expression profiles of query cells to a reference hematopoietic cell dataset, followed by manual refinement, validation, and supplementation of the initial annotations through cross-referencing with canonical cell type marker genes from published AML scRNA-seq studies, including the following validated markers: CD14, CD68, CSF1R, CD163 for monocytes; MPO, ELANE, PRTN3, CTSG for granulocyte-macrophage progenitors (GMP); MPO, AZU1, PRTN3 for pro-myelocytes; CD34, GATA2, PROM1 for common myeloid progenitors (CMP); HBB, HBA1, GYPA for erythroblasts; CD3D, CD3E, CD8A, CD4 for T cells; CD19, CD79A, MS4A1 for B cells; CD34, CD38, PROM1, HOXA9 for hematopoietic stem and progenitor cells (HSPC); and GATA1, KLF1, TAL1 for megakaryocyte-erythrocyte progenitors (MEP). After completion of cell type annotation, the expression distribution of *RHBDF2* across all annotated cell types was visualized and systematically analyzed.

Prognostic Model Generation and Prediction

This study established a predictive model for individualized assessment of overall survival (OS) in AML patients. OS was defined as the time from the initial diagnosis of AML to the date of all-cause death. The start time point for OS calculation was the date of AML diagnosis, and the cutoff date for follow-up was the latest available follow-up time point provided in the public TCGA and GEO datasets. Patients who were alive at the follow-up cutoff date, or who were lost to follow-up before the cutoff date, were censored in the survival analysis. Variables with $p < 0.10$ in univariate Cox regression analysis were entered into the multivariate Cox regression model. Clinically relevant factors (age, gender, ELN 2022 risk category) were also considered, provided no collinearity (variance inflation factor, VIF <2.5). Collinearity was assessed using the variance inflation factor (VIF) calculated with the car R package. The final multivariate model included *RHBDF2* (continuous), age (continuous), gender, ELN 2022 risk category, and white blood cell (WBC) count (continuous). The proportional hazards assumption was verified using the Schoenfeld residual test. Utilizing the Free Statistics analysis platform, the model was constructed by integrating all independent prognostic factors identified in the multivariate Cox regression analysis: *RHBDF2* expression, age, ELN 2022 risk category, and WBC count at diagnosis. To evaluate predictive per-

formance, calibration curves were plotted to compare predicted probabilities with observed outcomes, where the 45° diagonal line represented ideal calibration. Additionally, the model's discriminative ability was assessed by calculating the concordance index (C-index). Time-dependent Receiver Operating Characteristic (ROC) curve analysis was performed using the timeROC R package, to evaluate the predictive performance of the nomogram model for OS at 1 year, 3 years, and 5 years after diagnosis. The reported Area Under the Curve (AUC) values are the bias-corrected AUC values, calculated via 1000 iterations of bootstrap resampling, to adjust for optimism bias in model performance estimation. AML patients were uniformly stratified into high and low *RHBDF2* expression groups using the median *RHBDF2* expression value of the corresponding cohort: samples above the median were classified as the high expression group, and those below or equal to the median as the low expression group. Kaplan-Meier survival curves were generated to illustrate OS differences between the two groups, while the log-rank test assessed the correlation between expression levels and survival time alongside its statistical significance.

Statistical Analysis

All statistical analyses were performed using R Statistical Software (Version 4.2.2, http://www.R-project.org, The R Foundation) and Free Statistics analysis platform (Version 2.0, Beijing, China, http://www.clinicalscientists.cn/freestatistics). FreeStatistics is a software package with intuitive graphical user interfaces (GUI written in Python) for common analyses and data visualization, using R as its underlying statistical engine. A two-sided p-value < 0.05 was considered statistically significant for all analyses. For subgroup analyses comparing high and low *RHBDF2* expression groups, samples were uniformly stratified using the median *RHBDF2* expression value of the corresponding cohort: samples above the median were classified as the high expression group, and those below or equal to the median as the low expression group. To ensure result comparability across different technical platforms (TCGA RNA-seq, GEO microarray, single-cell RNA-seq), consistent preprocessing steps were performed for all datasets independently, including uniform gene ID conversion, log2 transformation of expression values, low-expression gene filtering, and batch effect correction using the ComBat-seq/ComBat method; all downstream analyses (differential expression analysis, survival analysis, correlation analysis) were conducted using consistent statistical methods and thresholds across datasets to avoid cross-platform batch effects. Specific statistical methods for each analysis type were as follows: differential expression analysis between two groups was performed via the limma-voom pipeline, with the Benjamini-Hochberg (BH) method for multiple testing correction and significance thresholds set as FDR <0.05 and $|\log2FC| > 1$; GO and KEGG functional enrichment analyses were conducted using the hypergeometric test, with BH correction and adjusted p-value < 0.05 as the threshold for significant enrichment; Pearson correlation analysis was used for normally distributed continuous variables, while Spearman rank correlation analysis was applied for non-normally distributed variables, with BH correction for multiple correlation analyses and adjusted p-value < 0.05 as the significance threshold; survival analysis included Kaplan-Meier survival curves (with log-rank test for significance) to compare overall survival (OS) between high and low *RHBDF2* expression groups, as well as univariate and multivariate Cox proportional hazards regression analyses to calculate hazard ratios (HR) and 95% confidence intervals (CI), with the Schoenfeld residual test verifying the proportional hazards assumption ($p > 0.05$ indicating satisfaction); multi-group comparisons used one-way analysis of variance (ANOVA) followed by Tukey's HSD post-hoc test for normally distributed continuous variables with homogeneous variance, and Kruskal-Wallis H test followed by Dunn's post-hoc test with BH correction for non-normally distributed or heteroscedastic variables; two-group comparisons employed independent samples Student's t-test for normally distributed continuous variables, Wilcoxon rank-sum test for non-normally distributed continuous variables, and chi-square test or Fisher's exact test for categorical variables as appropriate; normality of continuous variables was tested via the Shapiro-Wilk test, and homogeneity of variance between groups via Levene's test; data description followed standard conventions: normally distributed continuous variables as mean $\pm$ standard deviation (SD), non-normally distributed continuous variables as median (interquartile range, IQR), and categorical variables as frequency (percentage, %); the BH method was uniformly applied for multiple testing correction in all analyses involving multiple comparisons to control the false discovery rate.

Results

Expression of RHBDF2 Is Elevated in AML Patients and Associated With Immune-Related Pathways and Overall Survival (The Screening Process Of RHBDF2)

To ascertain the transcriptomic variations between AML patients and non-malignant control individuals and to pinpoint key regulatory elements involved in AML progression, we initiated our investigation by retrieving transcriptomic sequencing data and clinical characteristics from the TCGA database for 150 AML patients' bone marrow samples. Furthermore, transcriptomic sequencing data from 337 non-malignant control individuals were sourced from the GTEx database for comparative analysis. We amalgamated and batch-normalized the two datasets, resulting in the identification of 2680 differentially expressed genes (DEGs) based on the criteria of |logFC| (log2 fold-change)

>1 and FDR <0.05, and depicted these genes in a heatmap (Fig. 1B). These DEGs underwent further enrichment analyses through GO (Fig. 1C,D) and KEGG (Fig. 1E,F), represented using bubble plots (Fig. 1C,E), chord plots, and cluster plots (Fig. 1D,F). The prominent biological processes that were enriched included those involving immune response-regulating signaling and leukocyte-mediated immunity. The foremost cellular components were identified as secretory granule lumen and cytoplasmic vesicle lumen, with the main molecular functions being immune receptor activity and antioxidant activity. The principal pathways were those involving hematopoietic cell lineage and transcriptional misregulation in cancer. These observations infer that the DEGs between AML and non-malignant control individuals could potentially play pivotal roles in immune response and regulation. We then screened the top 10 immune-related biological process (BP) gene sets from the GO enrichment analysis of AML vs. control DEGs, which were all associated with immune response regulation and leukocyte-mediated immunity. All genes within these 10 BP gene sets were extracted, and after removing duplicate genes, a total of 248 immune-related genes were obtained. Concurrently, we excluded 56 genes with low or uniform expression across all samples from the initial 2680 DEGs, isolating 2624 genes. These genes were segregated into high- and low-expression groups based on median expression and underwent high-throughput Kaplan-Meier survival analysis for OS. The Benjamini-Hochberg (BH) method was applied for multiple testing correction during this screening step, with FDR <0.05 set as the threshold for significant prognostic association. This analysis led to the identification of 581 genes with a significant association with OS (adjusted $p < 0.05$). The intersection of these 581 survival-related genes and the 248 immune-related genes resulted in 54 candidate genes. A further filtering for univariate Cox regression p-value < 0.01 reduced this number to 20 candidate genes. From these 20 candidates, we performed a comprehensive literature review and applied three additional selection criteria: (1) the gene has been reported in multiple solid tumors; (2) no published studies exist on its role in any hematological malignancies; and (3) it is associated with adverse prognosis in cancer. Through this rigorous screening, we selected *RHBDF2* for further investigation. *RHBDF2*, which exhibited significantly higher expression in TCGA AML bone marrow samples compared to GTEx non-malignant bone marrow controls (consistent tissue origin), was correlated with both immune-related pathways and patient survival, and had previously been unreported in acute leukemia. This observational finding forms the basis of the current correlative study, with no functional validation of mechanisms. The screening process is outlined in a flowchart (Fig. 1A).

Correlation Between RHBDF2 Expression and the Clinical-Pathological Attributes of AML

Then, Initial studies on the mRNA profile of *RHBDF2* in human malignancies were conducted using the GEPIA database, revealing that *RHBDF2* exhibited increased expression in a majority of tumor types, including AML (Fig. 2A,B $p < 0.05$). The present study primarily focus on elucidating the role of *RHBDF2* in AML. Utilizing the TCGA dataset, we visualized the distribution of patient survival status, FAB classification, leukocyte count, age, and ELN 2022 risk stratification across *RHBDF2* expression levels (Fig. 2D). Formal statistical testing confirmed that high *RHBDF2* expression was significantly associated with adverse clinicopathological features, as detailed below and in Table 1. Subgroup analysis of these samples was conducted. Notably, *RHBDF2* overexpression was predominantly observed in acute monocytic leukemia (M5) as compared to M0, M2, and M3 subtypes ($p < 0.01$ for each comparison). No statistically significant difference was detected between M1 and M5 ($p = 0.12$) or between M4 and M5 ($p = 0.07$) (Fig. 2C). *RHBDF2* was also significantly more abundant in patients with a leukocyte count of $\geq 9.5 \times 10^9$/L (Fig. 2F, $p = 0.023$) and in those above the age of 65 years (Fig. 2G, $p = 0.014$). Moreover, *RHBDF2* expression was found to be higher in patients with intermediate-risk compared to those with low-risk, although no significant difference was detected between high- and intermediate-risk patients (Fig. 2H, $p = 0.028$). Additionally, *RHBDF2* expression was notably higher in deceased patients in comparison to their surviving counterparts (Fig. 2I, $p = 0.014$). Consequently, we performed a restricted cubic spline (RCS) Cox regression model to assess the non-linear association between *RHBDF2* expression and the hazard of death in AML patients. Restricted cubic spline (RCS) Cox regression analysis showed a non-significant nonlinear trend between *RHBDF2* expression and the hazard of death (p for non-linearity = 0.091) (Fig. 2E). The proportional hazards assumption of the Cox model was verified using the Schoenfeld residual test, with a global p-value of 0.32, indicating that the proportional hazards assumption was satisfied. Compared to the low *RHBDF2* group, the high *RHBDF2* expression group exhibited a greater abundance of poor-prognosis genes, including IDH1 ($p < 0.001$), TET2 ($p < 0.01$), and ZRSR2 ($p < 0.001$) (Fig. 2J,L,M), while the favorable-prognosis gene NPM1 was predominantly observed in the low *RHBDF2* group (Fig. 2K, $p < 0.001$).The baseline data analysis was elucidated in Table 1. In summary, these findings suggest that *RHBDF2* expression is notably more abundant in AML samples with increased malignancy and is a significant predictor of poorer prognosis.

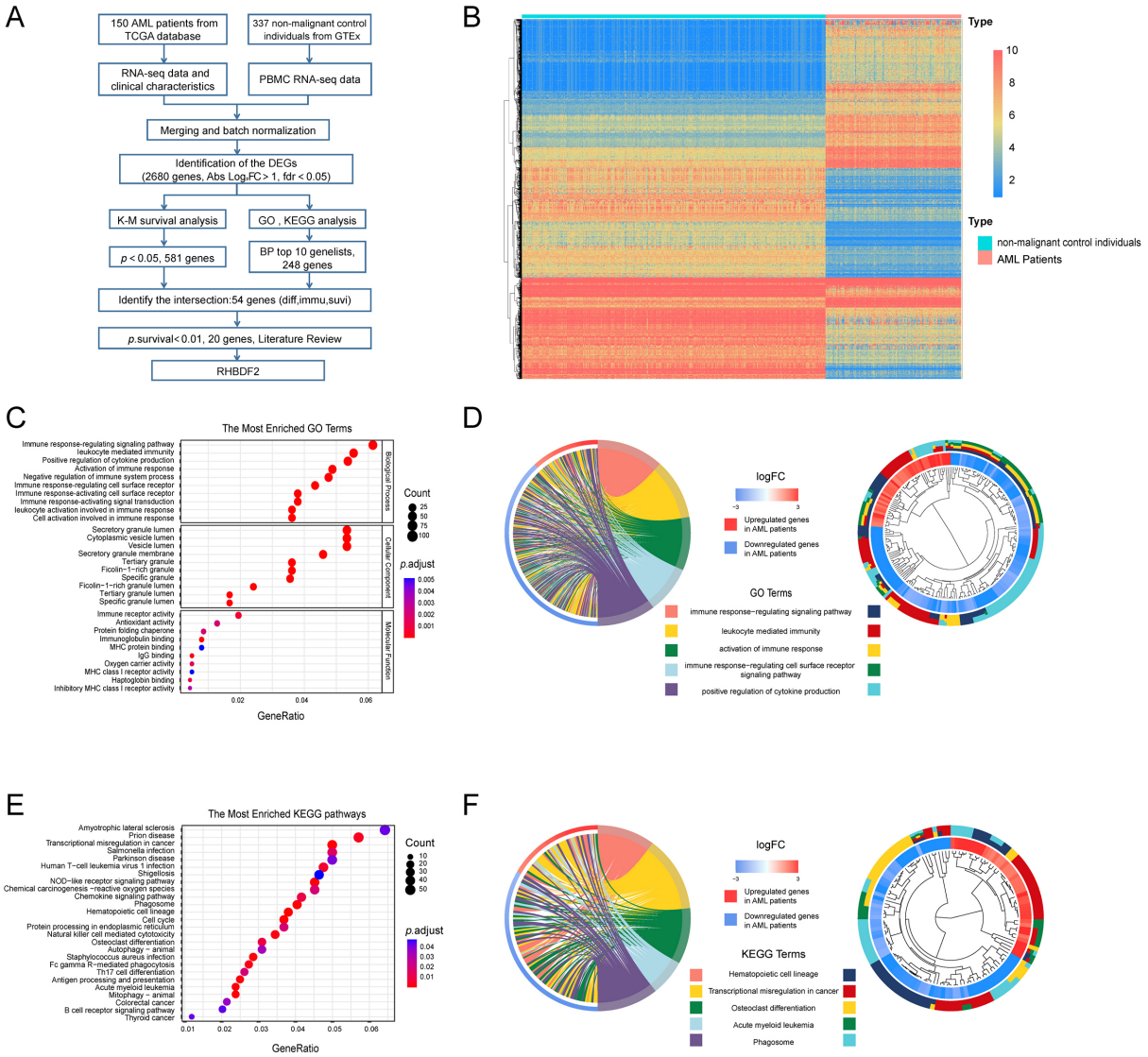

Fig. 1. Compared with non-malignant control individuals, *RHBDF2* was highly expressed in AML patients and was closely related to the immune response process. (A) The flowchart illustrated the screening process of *RHBDF2*. The AML samples were derived from the bone marrow of newly diagnosed de novo AML patients in the TCGA-LAML cohort, and the non-malignant control samples were derived from the whole bone marrow of non-malignant control individuals in the GTEx cohort. (B) Heatmap of the 2680 differentially expressed genes (DEGs) between AML patients in the TCGA cohort and non-malignant control individuals in the GTEx cohort. Row-wise Z-score transformation was applied to the expression values, with blue representing low expression and red representing high expression. Hierarchical clustering was performed using the ward. D2 method for both rows and columns. (C,D) Gene ontology (GO) functional enrichment analysis of DEGs between AML patients and non-malignant control individuals displayed with bubble plot (C), cluster plot and chord plot (D). (E,F) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of DEGs between AML patients and non-malignant control individuals displayed with bubble plot (E), chord plot and cluster plot (F). Chord diagram clustered by colors. The left semicircle in blue represented downregulated genes in AML (LogFC <0), while red represented upregulated genes (LogFC >0). Darker red indicated higher LogFC values, and darker blue indicated lower LogFC values. The right semicircle features bands of different colors represented various GO/KEGG terms, with connecting lines linking each gene to the GO/KEGG terms to which it is related. In the cluster plot, the inner circles depict genes, with colors corresponding to the abundance fold-change observed in the study, as per the heat map scale. The outer circles, in turn, denote different GO or KEGG enrichment categories. *RHBDF2*, Rhomboid 5 Homolog 2; AML, Acute myeloid leukemia; TCGA, The Cancer Genome Atlas; LAML, Acute Myeloid Leukemia-like; GTEx, Genotype-Tissue Expression.

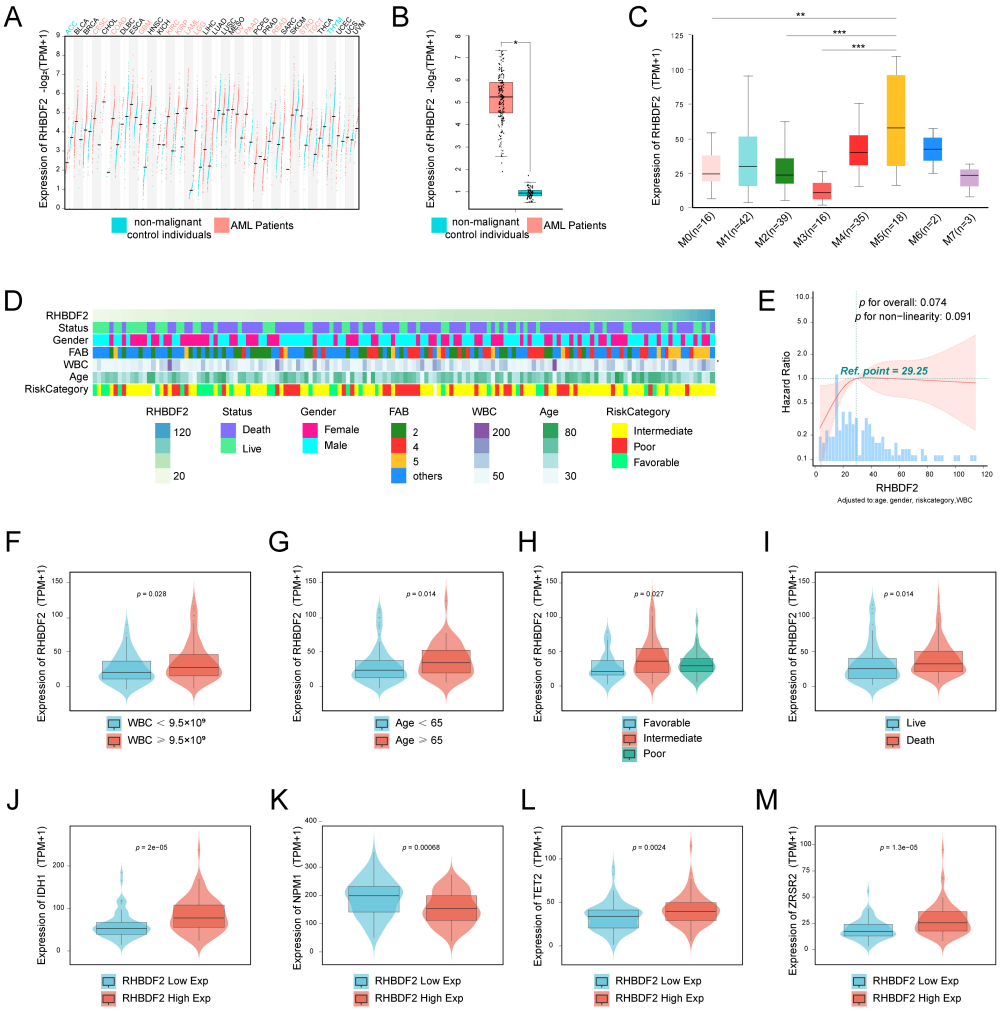

Fig. 2. The expression of *RHBDF2* in AML and its correlation with clinicopathological characteristics. (A) The expression of *RHBDF2* in paired normal and pan-cancer samples using the GEPIA database. (B) The expression of *RHBDF2* paired normal and AML samples from GEPIA database. (C) The expression of *RHBDF2* in AML based on French American British classifiation from UALCAN database. M0 n = 16, M1 n = 42, M2 n = 39, M3 n = 16, M4 n = 35, M5 n = 18, M6 n = 2, M7 n = 3, The values displayed in the box plots represent the median expression level of each subgroup, with *p*-values calculated via the Kruskal-Wallis H test. Pairwise comparisons were performed using Dunn's post-hoc test with Benjamini-Hochberg correction for multiple testing. (D) The Landscape of *RHBDF2* related clinicopathological features of AML patients from TCGA database. (E) Association of *RHBDF2* mRNA expression with Hazard Ratio (HR) for OS in AML patients from the TCGA database, analyzed using a restricted cubic spline (RCS) Cox regression model with 3 knots. The shaded area represents the 95% confidence interval of the HR. The RCS analysis showed a non-significant nonlinear trend between *RHBDF2* expression and the hazard of death (p for non-linearity = 0.091). (F–I) *RHBDF2* expression levels in subgroups stratified by white blood cell count, age, ELN 2022 risk category, and survival status. Normality of gene expression data was tested using the Shapiro-Wilk test, and the results showed that *RHBDF2* expression data in all subgroups did not conform to normal distribution (all $p < 0.05$). Therefore, the non-parametric Mann-Whitney U test was used for between-group comparisons. $p < 0.05$ was considered statistically significant. For the multiple group expression comparisons in subfigures (H) p values were adjusted for multiple testing using the Benjamini-Hochberg (BH) method, with adjusted $p < 0.05$ set as the threshold for statistical significance. (J–M) AML patients were divided into high and low expression groups based on the median value of *RHBDF2* expression in TCGA database. IDH1, TET2, and ZRSR2 were significantly increased in the high-expression group, while NPM1 was significantly enriched in the low-expression group. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$. GEPIA, Gene Expression Profiling Interactive Analysis; UALCAN, University of Alabama at Birmingham Cancer data analysis Portal; ELN, European LeukemiaNet; IDH1, isocitrate dehydrogenase 1; TET2, Ten-Eleven Translocation 2; ZRSR2, Zinc Finger CCCH-Type, RNA Binding Motif and Serine/Arginine Rich 2; NPM1, Nucleophosmin 1.

Table 1. Baseline patient characteristics according to *RHBDF2* expression level in AML patients from the TCGA database.

Variables	Total (n = 150)	*RHBDF2* Low Exp (n = 75)	*RHBDF2* High Exp (n = 75)	Statistical Test	Test Statistic	*p* value
Age, years, n (%)				Chi-square test	$\chi^2 = 4.63$	0.031
<65	106 (70.7)	59 (78.7)	47 (62.7)			
≥65	44 (29.3)	16 (21.3)	28 (37.3)			
Gender, n (%)				Chi-square test	$\chi^2 = 0.03$	0.870
Male	83 (55.3)	41 (54.7)	42 (56.0)			
Female	67 (44.7)	34 (45.3)	33 (44.0)			
Risk Category, n (%)				Chi-square test	$\chi^2 = 7.23$	0.014
Favorable	31 (20.7)	22 (29.3)	9 (12.0)			
Intermediate	82 (54.7)	35 (46.7)	47 (62.7)			
Poor	37 (24.7)	18 (24.0)	19 (25.3)			
WBC, n (%)				Chi-square test	$\chi^2 = 4.85$	0.028
$<9.5 \times 10^9$/L	55 (36.7)	34 (45.3)	21 (28.0)			
$\geq 9.5 \times 10^9$/L	95 (63.3)	41 (54.7)	54 (72.0)			
Hb, g/L, Mean ± SD	9.5 ± 1.4	9.7 ± 1.3	9.3 ± 1.5	Independent samples *t*-test	$t = 1.81$	0.072
PLT, $\times 10^9$/L, Median (IQR)	45.5 (28.5, 86.0)	42.0 (24.0, 76.0)	57.0 (32.0, 89.0)	Wilcoxon rank-sum test	Z = –1.48	0.139
BM Blast, %, Median (IQR)	39.0 (8.2, 66.2)	41.0 (6.5, 73.0)	37.0 (10.0, 58.5)	Wilcoxon rank-sum test	Z = 0.59	0.550

Note: *p*-values for categorical variables were calculated using the chi-square test; for continuous variables, the Mann-Whitney U test was used. *RHBDF2*, Rhomboid 5 Homolog 2; AML, Acute myeloid leukemia; TCGA, The Cancer Genome Atlas; WBC, white blood cell; SD, standard deviation; PLT, Platelet Count; IQR, interquartile range; BM, Bone Marrow.

Confirmation of Altered RHBDF2 Gene Expression in AML

To further analyze the mRNA and protein expression levels of *RHBDF2* in AML patients, we gathered a dataset comprising 306 bone marrow samples from the Second Hospital of Shanxi Medical University. Among these, 139 bone marrow fluid samples were used for mRNA expression analysis, including 63 newly diagnosed AML (pAML) cases, 26 patients in complete remission (CR-AML) cases, 16 relapsed (rAML) cases and 34 non-malignant control individuals. Additionally, 167 bone marrow biopsy specimens were subjected to protein expression analysis via immunohistochemistry, consisting of 91 pAML, 25 CR-AML, 9 rAML and 42 non-malignant control individuals.

Our immunohistochemistry (IHC) analysis showed that RHBDF2 protein expression was significantly higher in pAML compared to non-malignant controls ($p < 0.0001$) and in pAML compared to CR-AML ($p < 0.0001$). rAML samples showed higher expression than non-malignant controls ($p < 0.0001$) and than CR-AML ($p < 0.05$). No significant difference was observed between pAML and rAML ($p = 0.67$). The qRT-PCR results consistently showed that RHBDF2 mRNA expression was significantly higher in pAML compared to non-malignant controls ($p < 0.001$), significantly lower in CR-AML compared to pAML ($p < 0.05$), and significantly higher in rAML compared to both non-malignant controls ($p < 0.0001$) and CR-AML ($p < 0.01$) (Fig. 3A–C). These unpaired cross-sectional results indicate that *RHBDF2* expression is associated with AML disease status. Hematoxylin and eosin (H&E) staining was performed to assess the morphological structure of the bone marrow biopsy specimens, confirm the presence of leukemic blasts in AML samples, and verify the absence of malignant infiltration in non-malignant control samples. Representative H&E staining images are shown in Fig. 3A, confirming the morphological diagnosis of AML and the normal bone marrow structure in control samples.

Elevated RHBDF2 Expression Was Positively Correlated With the Expression of Suppressive Immune Checkpoints, and Showed a Significant Asssociation With Inflammatory Dysregulation in AML

Subsequently, we performed differential expression analysis between the high and low *RHBDF2* expression groups (total n = 150, high n = 75, low n = 75, stratified by median *RHBDF2* expression) using TCGA AML RNA-seq data. A total of 1709 statistically significant differentially expressed genes (DEGs) were identified (|log2FC| >1, FDR <0.05), including 1055 upregulated genes and 654 downregulated genes. A heatmap of the top 50 most significantly upregulated and downregulated DEGs is presented in Fig. 4A. GO functional enrichment and KEGG pathway analyses were performed on these 1709 DEGs (Fig. 4C,D). The enriched biological processes were primarily related to positive regulation of cytokine production and negative regulation of immune system processes; enriched cellular components included the extracellular side of the plasma membrane and secretory granule membrane; enriched molecular functions included immune receptor activity and inhibitory MHC class I receptor activity; significantly enriched KEGG pathways included osteoclast dif-

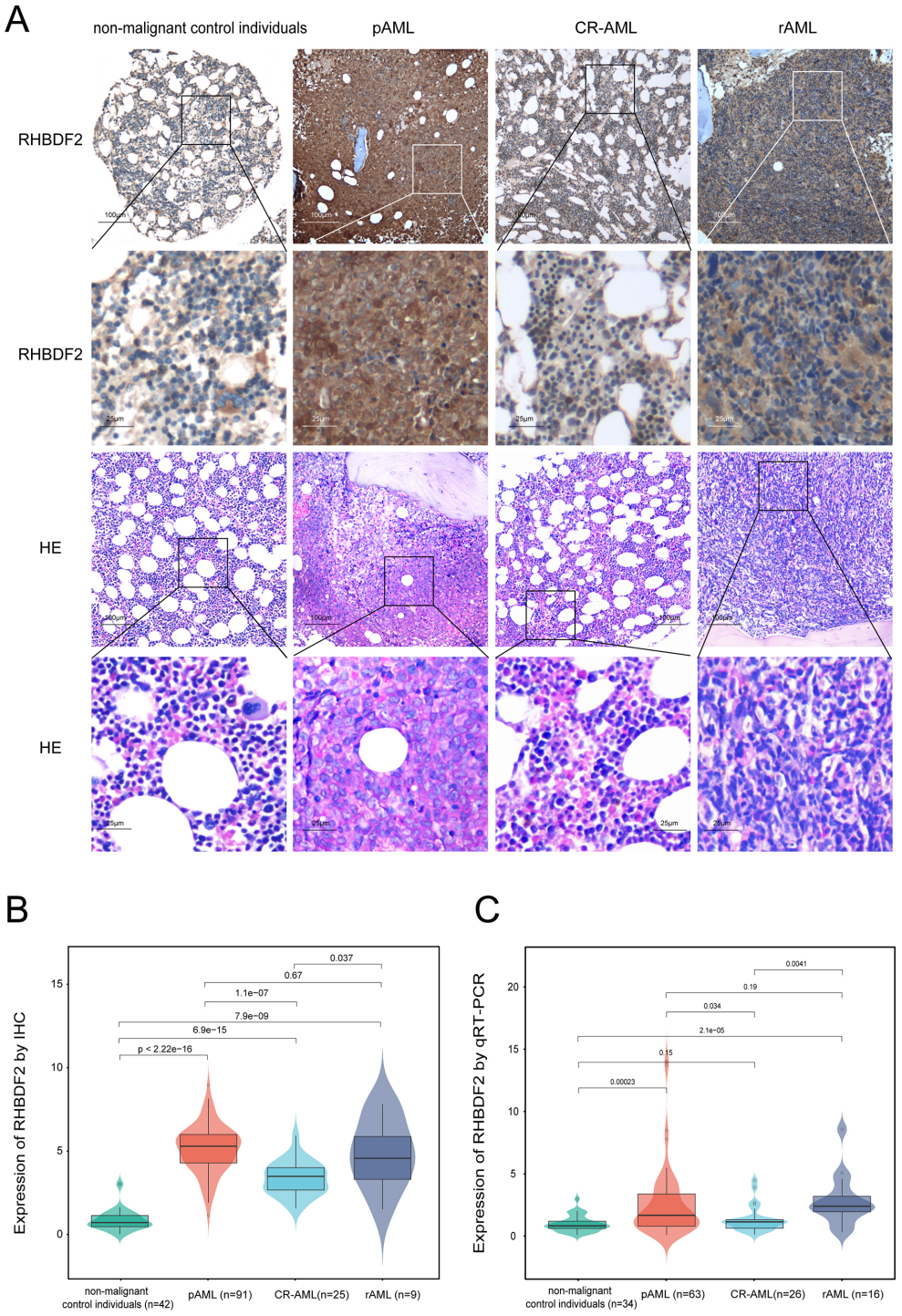

Fig. 3. Validation of *RHBDF2* expression in AML. (A,B) Immunohistochemical staining and hematoxylin and eosin staining of *RHBDF2* were performed in bone marrow biopsy of AML patients at initial diagnosis (pAML, n = 91), complete remission (CR-AML, n = 25), relapse (rAML, n = 9) and non-malignant control individuals (n = 42). Representative images were shown. Scale bars: 100 μm (main image); 25 μm (magnified inset). The statistical results are presented in the Fig. 3B. (C) qRT-PCR were used to analyzed the mRNA expression of *RHBDF2* in different stage of AML patients (pAML, n = 63; CR-AML, n = 26; rAML, n = 16) and non-malignant control individuals (n = 34). Data are presented as median (interquartile range, IQR). Statistical significance was determined using the Kruskal-Wallis test followed by Dunn's post-hoc test with Benjamini-Hochberg correction for multiple comparisons.

ferentiation and cytokine-cytokine receptor interaction signaling pathways. These correlative findings indicate that *RHBDF2* expression is associated with immune-related biological processes and pathways in AML. To validate these conclusions, we assessed the impact of *RHBDF2* activation on immune pathways and cytokine signatures. Gene Set Variation Analysis (GSVA) was utilized to calculate enrichment scores for immune processes. Correlation analysis between *RHBDF2* expression and GSVA enrichment scores revealed that *RHBDF2* was significantly associated with multiple immune-related processes (Fig. 4B). Full correlation coefficients and adjusted *p*-values are provided in **Supplementary Table 1**. Moreover, we investigated the relationship between *RHBDF2* and known suppressive immune checkpoints (including CD200R1, CD40, CD48, CD86, TIM-3, HHLA2, LAIR1, LGALS9, PD-L2, HVEM, CD30, and TL1A) (Fig. 4E). The similar results were found in GSE12417 and GSE71014 database (**Supplementary Fig. 1C,D**). *RHBDF2* showed significant positive correlations with these immune checkpoints, and is associated with the suppressed immune response against AML. Additionally, we selected seven immune system-related metagene clusters as indicators of immune and inflammatory status in AML, including hematopoietic cell kinase (HCK), lymphocyte-specific kinase (LCK), major histocompatibility complex (MHC), signal transducer and activator of transcription 1/2 (STAT1/2), interferon, and IgG3/4. These metagene clusters are well-validated immune signatures in published AML studies, with defined gene sets that reflect the activity of specific immune and inflammatory pathways in the tumor microenvironment. Correlation analysis revealed that *RHBDF2* was significantly positively correlated with all seven metagene clusters (Fig. 4F), indicating a robust positive association between *RHBDF2* expression and most inflammatory responses in AML. These correlative results show that *RHBDF2* expression is positively correlated with the expression of multiple suppressive immune checkpoints in AML, and is associated with a potentially immunosuppressive bone marrow microenvironment in AML patients. As this is an observational study, no causal relationships or functional mechanisms are inferred.

Elevated RHBDF2 Expression in AML Is in Connection With a Heightened Infiltration of Monocyte-Based Immune Cells

To further probe into the role of *RHBDF2* within the tumor's immune microenvironment, we undertook an assessment of the infiltrating immune cells in the AML TME. Our initial examination of the TCGA database via the CIBERSORT algorithm unveiled that monocytes exhibited a negative correlation with11 different types of immune cells, encompassing naive B cells, CD4+ memory resting T cells, and CD8+ T cells (Fig. 5A). We conducted an analysis of the tumor microenvironment's disparities between patients with high and low *RHBDF2* expression utilizing the ESTIMATE package to compute stromal scores, immune scores, and ESTIMATE scores for every sample, in addition to determining tumor purity. As shown in Fig. 5B, the high RHBDF2 expression group had significantly higher stromal, immune, and ESTIMATE scores (all $p < 0.001$), indicating a more abundant immune cell infiltration in the bone marrow microenvironment of AML patients with high *RHBDF2* expression. This hypothesis was corroborated by our analysis of the GSE12417 and GSE71014 datasets, as evidenced in **Supplementary Fig. 2A,B**.

Analysis of immune cell composition (Fig. 5C) revealed that monocytes were significantly more abundant in the high *RHBDF2* expression group ($p < 0.001$). This observation was corroborated by our analysis of the GSE12417 and GSE71014 datasets (**Supplementary Fig. 2C,D**). We conducted correlation analysis between the expression levels of target genes and immune cells, presented in a lollipop plot (Fig. 5D). The findings revealed a positive correlation between *RHBDF2* expression and the presence of monocytes, M1 macrophages, and Tregs, whereas CD4 memory resting T cells, follicular helper T cells, resting mast cells, activated dendritic cells, plasma cells, activated mast cells, and naive B cells exhibited an inverse correlation with *RHBDF2* expression. The corresponding data from the GSE12417 and GSE71014 datasets can be found in the **Supplementary Fig. 2E–H**. These findings are based on computational deconvolution of bulk RNA-seq data using the CIBERSORT algorithm, which provides an inference of the relative abundance of immune cell subsets in the AML bone marrow microenvironment. We acknowledge that this *in silico* inference may have potential discrepancies compared to the actual immune cell composition measured via direct experimental methods such as flow cytometry or immunohistochemistry, and the results should be interpreted with appropriate caution.

Analyses of Single-Cell Sequencing Data Suggested That RHBDF2 Is Predominantly Expressed in the Monocytic Lineage in AML Patients

By utilizing single-cell sequencing data obtained from AML patients in the GSE116256 and GSE289435 datasets, we applied the t-SNE algorithm for the clustering of cells and to visually represent their similarities. Cells with similar attributes were located closer in the t-SNE plot, while cells with dissimilar attributes were positioned further apart. Consequently, 12 (Fig. 6A) and18 (Fig. 6E) distinct clusters were discerned through t-SNE, respectively. Each cluster was annotated using the R package Single R. The12 clusters from the GSE116256 dataset were categorized into 7 cell types: Monocyte, pro-Myelocyte, Granulocyte-Macrophage Progenitor (GMP), T cells, B cells, Common Myeloid Progenitor (CMP), and Erythroblast (Fig. 6B). In contrast, the18 clusters from the GSE289435 dataset were classified into 7 cell types: GMP, Monocyte, Bone Marrow Progenitors (BM&prog), also known as Hematopoi-

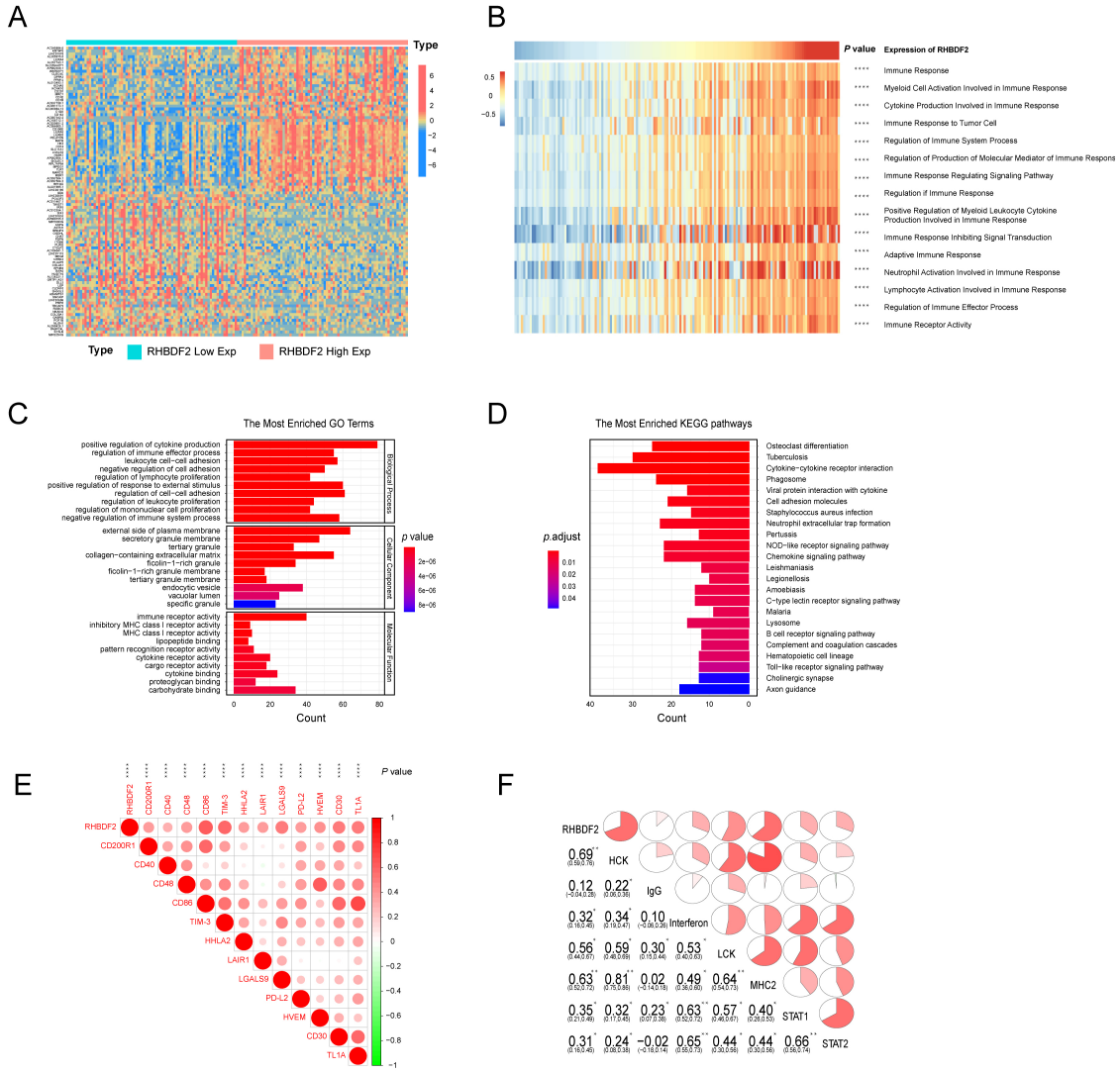

Fig. 4. The correlation between *RHBDF2* expression and immune infiltration and inflammatory activities in TCGA cohort. (A) Heatmap showed the DEGs in AML patients between high and low *RHBDF2* expression groups. (B) Heatmap showing the correlation between *RHBDF2* expression and GSVA enrichment scores for immune-related processes. Only correlations with adjusted $p < 0.05$ (Benjamini-Hochberg corrected) are displayed. Complete statistical data including Pearson correlation coefficients and exact adjusted *p*-values are available in **Supplementary Table 1**. (C,D) Gene Ontology (GO) functional enrichment analysis (C) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis (D) of DEGs between high and low *RHBDF2* expression groups. Only statistically significantly enriched pathways (adjusted $p < 0.05$, Benjamini-Hochberg corrected for multiple testing) are shown. The color gradient represents the *p*-value of each pathway. All modifications ensure that the statistical significance of each presented pathway is clearly and unambiguously demonstrated, and the rationale for our significance threshold and figure presentation is fully justified. (E) Pearson correlation between *RHBDF2* and inhibitory immune checkpoints. The Green represented a negative correlation while red represents a positive correlation. (F) Correlation matrix of *RHBDF2* and inflammatory-related metagenes. The bottom left showed the correlation coefficient. The correlation coefficients were demonstrated as the proportion of the pie charts. The red parts represented a positive correlation, while the green parts represented a negative correlation. All correlation analyses were performed using Pearson correlation analysis, with *p*-values corrected for multiple testing using the Benjamini-Hochberg (BH) method. Adjusted *p*-value < 0.05 was considered statistically significant. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{****}p < 0.0001$.

etic Stem/Progenitor Cells (HSPC), Erythroblast, T cells, B cells, and Megakaryocyte-Erythrocyte Progenitor (MEP) (Fig. 6F).

We employed a dot plot to depict the expression of *RHBDF2* in each cell type. The findings from both datasets indicated that *RHBDF2* exhibited the highest expression

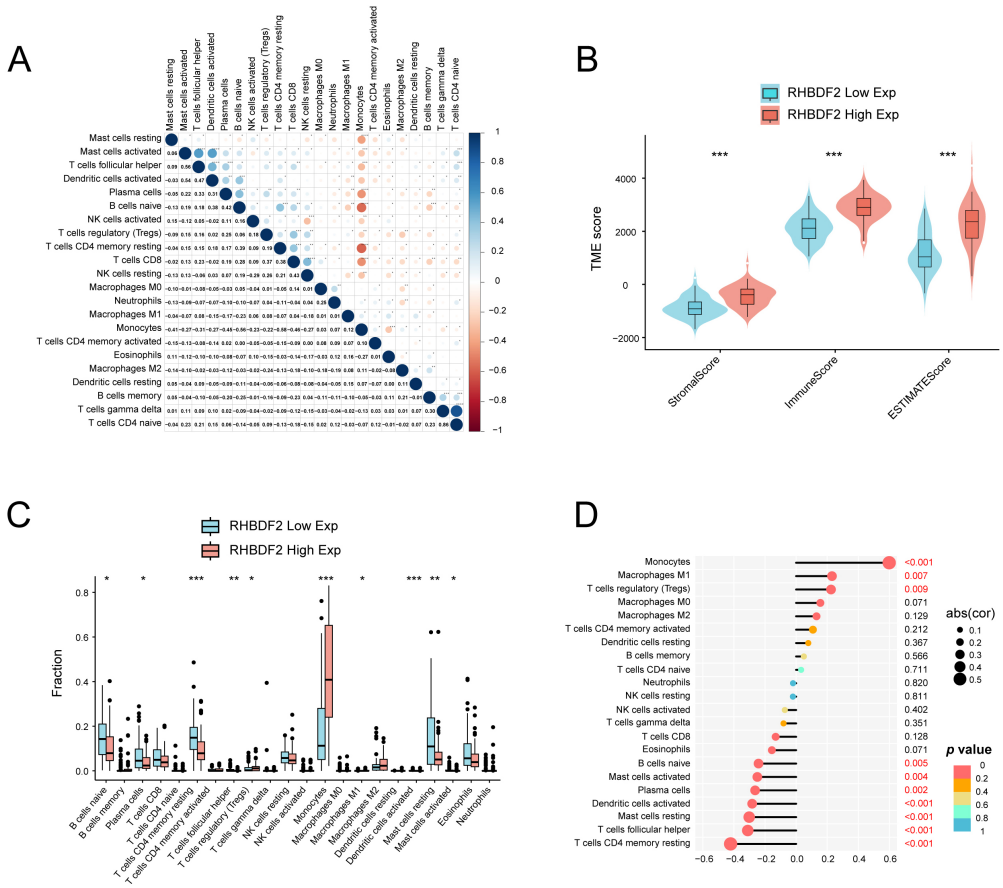

Fig. 5. The correlation of immune cells infiltrated in TME and different immune profiles between *RHBDF2* high and low expression groups in the TCGA dataset. (A) Diagram of correlations between the levels of various tumor infiltrated immune cells. (B) Comparison of stromal score, immune score, and ESTIMATE score between *RHBDF2* high and low expression groups. (C,D) Comparison of immune cell infiltration between high and low expression groups (C, box plot; D, lollipop diagram). *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

ratio and average level in monocytes (Fig. 6C,G). Scatter plot analyses uncovered that *RHBDF2* was predominantly present in Monocyte, GMP, and pro-myelocyte cells (Fig. 6D,H).

Construction and Validation of a Nomogram Integrating RHBDF2 Expression and Clinicopathological Characteristics for Assessing Predictive Modeling and Diagnostic Performance

Consequently, we remain curious as to whether *RHBDF2* might emerge as a potential prognostic indicator for AML patients. A nomogram, developed by integrating outcomes from both univariate and multivariate Cox regression analyses (as outlined in Table 2), visually mapped the effects of *RHBDF2* expression, gender, age, risk category (extrapolated from ELN2022), and WBC counts on the prognosis of overall survival (OS) in individuals with AML. The binary WBC variable was not significant in univariate analysis but became significant in multivariate analysis. This may be due to the adjustment for confounding factors such as age and ELN risk category, which revealed the independent prognostic effect of WBC count after controlling for these variables. This was illustrated in Fig. 7A. Moreover, Kaplan-Meier survival analysis was performed on 140 AML patients with complete overall survival data from the TCGA cohort (10 patients were excluded due to missing survival information) showed that high *RHBDF2* expression was associated with significantly shorter overall survival in AML patients as shown in Fig. 7B. We further validated the prognostic value of *RHBDF2* itself in two independent GEO cohorts, with patients stratified into high and low *RHBDF2* expression groups by cohort-specific median expression: ① GSE12417 cohort (total n = 79, high expression group n = 40, low expression group n = 39); ② GSE71014 cohort (total n = 104, high expression group n = 52, low expression group n = 52). The adverse prognostic effect of high *RHBDF2* expression was substantiated in both GEO cohorts. It should be noted that the constructed prognostic nomogram model was only validated via internal bootstrap resampling in the TCGA cohort, and no ex-

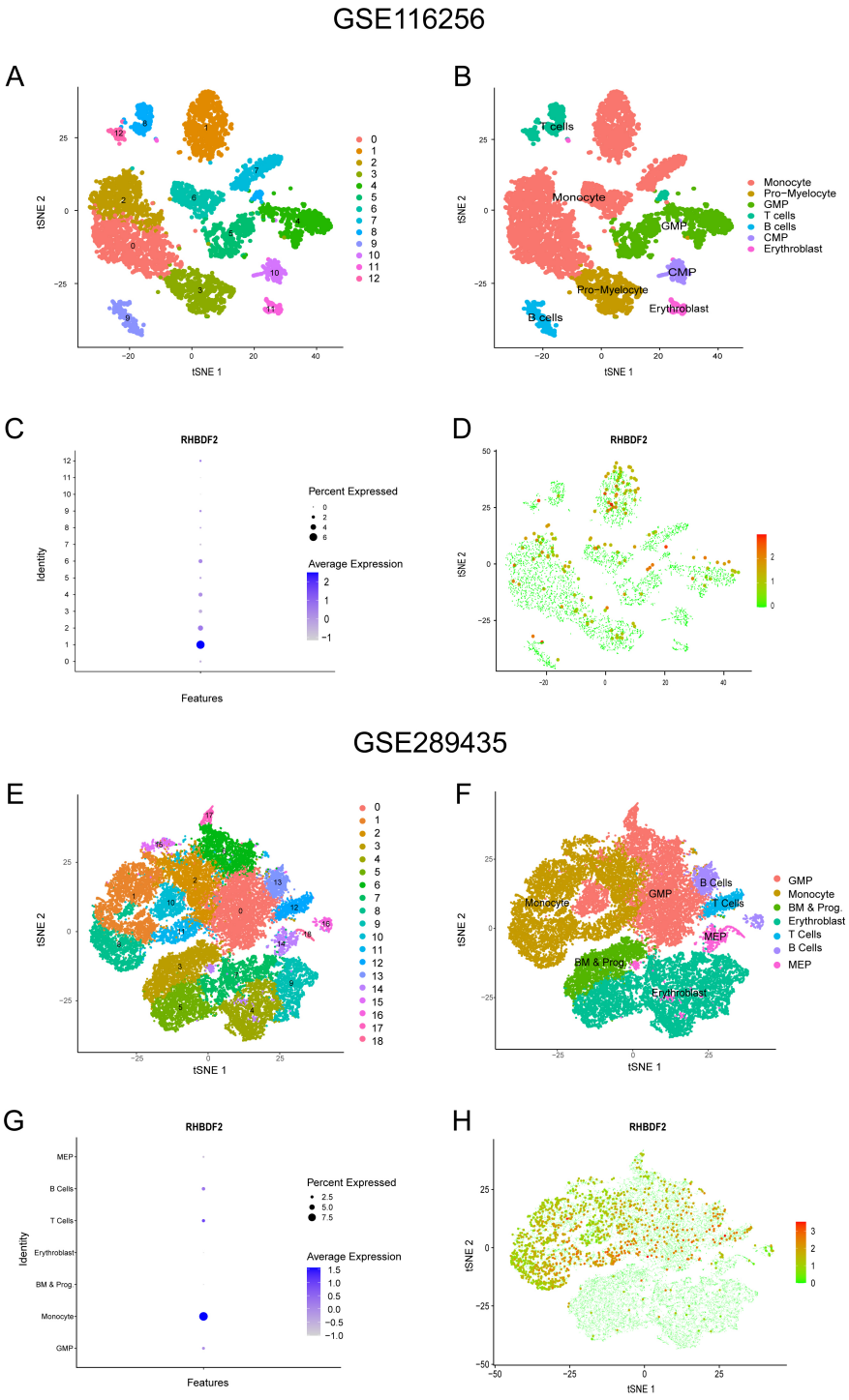

Fig. 6. Single-cell RNA sequence data analysis revealed that *RHBDF2* is highly expressed in monocytes. (A,E) t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm was performed and cell clusters were classified based on the available significant components from principal component analysis (PCA) in a public database GSE116256 (A) and GSE289435 (E). (B,F) different clusters were annotated based on specific gene expression profiles and were assigned to seven categories. (C,G) The dot plot displayed the expression levels of *RHBDF2* in each cluster. The size of the dots represented the percentage of cells expressing *RHBDF2* within a cluster. The color of the dot encodes the average expression level across all cells within a cluster (blue denotes high expression). (D,H) The scatter plot showed the expression of *RHBDF2* in each cell cluster, where low expression was represented in green and high expression in red.

ternal validation of the entire nomogram model was performed in the GEO cohorts (**Supplementary Fig. 1A,B**). Notably, the nomogram showed time-dependent AUC values of 75.2, 80.1, and 95.3 for 1-year, 3-year, and 5-year OS, respectively (as depicted in Fig. 7C,D), indicating the potential predictive value of the model in the TCGA discovery cohort. Notably, the high 5-year AUC value should be interpreted with extreme caution due to the limited number of patients remaining at risk at 5 years in the TCGA cohort, which may indicate model instability and potential overfitting. The generalizability of the model requires further validation in independent external AML cohorts. The patients were subsequently divided into high- and low-risk categories based on the median value of the risk score from our nomogram. As the risk scores of the patients increased in both categories, the number of fatalities also increased (as depicted in Fig. 7E). Furthermore, the calibration curve, illustrated in Fig. 7F–H, demonstrated a satisfactory concordance between the predicted and observed probabilities of 1-year, 3-year, and 5-year OS, which was affirmed through 1000 cycles of bootstrap resampling. The nomogram underwent further evaluation through DCA, which revealed that it provided an enhanced net benefit across a wider spectrum of threshold probabilities for predicting the risk of OS. As depicted in Fig. 7I–K, this analysis underscored the clinical utility of the nomogram in predicting the OS of AML patients.

Discussion

AML, a notably heterogeneous hematological malignancy, has been witnessing an escalation in both incidence and mortality rates globally [30]. Over the last decade, the five-year survival rate for individuals diagnosed with AML and treated with traditional chemotherapy and stem cell transplantation has only been reported to be 29% [3]. Although novel targeted therapies and immunotherapeutic strategies have provided patients with additional treatment options and optimism, they also come with their own unique degrees of side effects and drug resistance. Presently, the therapeutic approaches for AML have not yet fulfilled their potential in achieving the desired outcomes. And there's an urgent need to explore new prognostic biomarkers for this condition.

The iRhom2 protein, encoded by *RHBDF2* and being an inactive member of the rhomboid family, plays an important role in regulating an array of vital biological processes, including those involving ADAM17 [15], TNF signaling [16], and EGFR pathway activation [31]. In order to gain a deeper understanding of *RHBDF2*'s role in AML, including its expression level, prognostic relevance, underlying biological mechanisms, and its connection with tumor immune cell infiltration, we undertook relevant analyses. The screening rationale for *RHBDF2* is as follows: we first identified differentially expressed genes between AML patient samples and non-malignant control samples from the TCGA and GTEx databases. GO enrichment analysis showed that these DEGs were predominantly enriched in immune-related signaling pathways. Meanwhile, high-throughput Kaplan-Meier survival screening identified 581 genes significantly associated with AML patient overall survival. The intersection of immune-related DEGs and survival-related genes yielded 54 candidate genes, among which *RHBDF2* was selected for further analysis due to its significant upregulation in AML, robust correlation with immune signatures and patient survival, and the absence of published studies investigating its role in AML.

We further verified the expression pattern of *RHBDF2* using the GEPIA database and our single-center clinical cohort. IHC staining and qRT-PCR results showed that *RHBDF2* expression was significantly lower in AML patients in complete remission compared to newly diagnosed AML patients, and significantly higher in relapsed AML patients compared to remission patients and non-malignant controls. These findings are consistent with previous studies reporting that *RHBDF2* is upregulated and associated with poor prognosis in multiple solid tumors, including hepatocellular carcinoma, lung cancer, and cervical cancer.

Kaplan-Meier survival analysis showed that high *RHBDF2* expression was correlated with shorter overall survival in AML patients, which was further validated in two independent GEO cohorts (GSE12417 and GSE71014). Multivariate Cox regression analysis, adjusted for age, ELN 2022 risk category, and white blood cell (WBC) count, revealed that continuous *RHBDF2* expression was a marginally significant independent prognostic factor for overall survival (OS) in AML patients. The binary *RHBDF2* expression group (high vs. low, stratified by median expression) showed a trend towards adverse prognosis in the multivariate model, but did not reach statistical significance. We further constructed a prognostic nomogram integrating *RHBDF2* and clinical prognostic factors, with time-dependent AUC values of 75.2, 80.1, and 95.3 for 1-year, 3-year, and 5-year overall survival, respectively. However, the high 5-year AUC should be interpreted with caution due to the limited number of patients at risk at 5 years ($n < 20$), and potential overfitting cannot be excluded. External validation in larger multi-center cohorts is required before clinical application. These results suggest that the nomogram has potential predictive value for AML patient survival, and *RHBDF2* may be associated with prognosis in AML, but requires further validation.

AML is an inflammation-associated hematologic malignancy, whose onset and progression are closely correlated with immune cell infiltration in the bone marrow microenvironment [2]. Immunotherapeutic strategies, including allogeneic hematopoietic stem cell transplantation, chimeric antigen receptor T cell (CAR-T) therapy, and immune checkpoint inhibitors, have greatly advanced the treatment of hematologic malignancies [12]. However, the

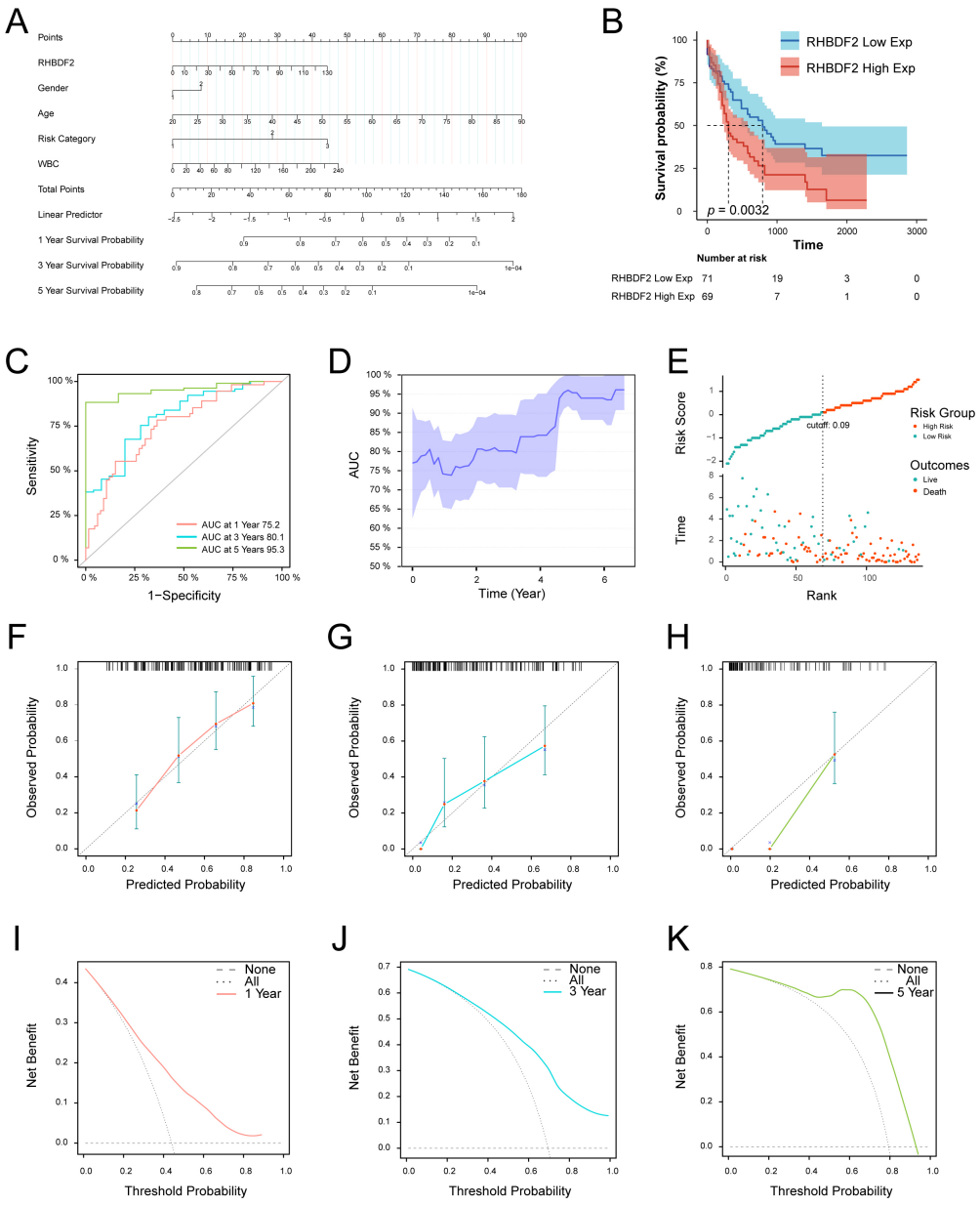

Fig. 7. A prognostic predictive model of *RHBDF2* in AML. (A) Nomogram for predicting the probability of 1-, 3-, 5-year OS for AML. Each relevant patient value was aligned and plotted on its corresponding variable axis, alongside a point axis that quantified the score for each variable. The cumulative score was pinpointed on the total axis, from which a line was drawn to determine the patient's 1-, 3-, and 5-year survival rates. (B) Kaplan-Meier OS analysis of AML patients between high and low *RHBDF2* expression groups. *p* value was calculated using the log-rank test. (C) Time-dependent ROC curves for OS of the AML patients. The AUC was assessed at 1, 3 and 5 years. (D) Changes in AUC over time. (E) Relationship between the survival status and risk score rank. (F–H) Calibration curve demonstrating the agreement between the predicted and observed probabilities of OS at 1 (F), 3 (G), and 5 (H) years. (I–K) Decision curve analysis (DCA) for the predictive model and its validation at 1 (I), 3 (J), and 5 (K) years.

complex immune evasion mechanisms and highly heterogeneous immune microenvironment of AML leads to unsatisfactory clinical outcomes of immunotherapy in AML patients.

In this study, GSVA enrichment analysis showed that *RHBDF2* expression was positively correlated with multiple immune-related biological processes in AML. GO and KEGG enrichment analyses showed that DEGs between high and low *RHBDF2* expression groups were enriched in immune receptor activity, cytokine signaling, and immune system regulation-related pathways. We also found that *RHBDF2* expression was positively correlated with the

Table 2. Univariable and Multivariable Cox analysis of putative prognostic parameters for overall survival (OS) in the AML patients from the TCGA database (n = 150).

Variable	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*p* value	HR (95% CI)	*p* value
RHBDF2	1.01 (1.00, 1.02)	0.010	1.01 (1.00, 1.02)	0.048
Binary variable				
RHBDF2 Low Exp	1 (Ref)		1 (Ref)	
RHBDF2 High Exp	1.87 (1.22, 2.86)	0.004	1.52 (0.98, 2.35)	0.063
Age (years)	1.04 (1.02, 1.05)	<0.001	1.04 (1.02, 1.05)	<0.001
Binary variable				
Age <65 years	1 (Ref)		1 (Ref)	
Age ≥65 years	3.29 (2.14, 5.06)	<0.001	3.40 (2.12, 5.45)	<0.001
Gender				
Male	1 (Ref)		1 (Ref)	
Female	0.99 (0.85, 1.20)	0.091	1.23 (0.80, 1.90)	0.350
Risk Category				
Favorable	1 (Ref)		1 (Ref)	
Intermediate	2.95 (1.50, 5.82)	0.002	2.24 (1.11, 4.52)	0.024
Poor	4.28 (2.01, 9.11)	<0.001	2.97 (1.37, 6.43)	0.006
Trend.test	1.91 (1.38, 2.65)	<0.001	1.62 (1.15, 2.28)	0.006
WBC count	1.01 (1.00, 1.01)	0.035	1.01 (1.00, 1.01)	0.019
Hb (g/L)	1.10 (0.95, 1.28)	0.201		
PLT ($\times 10^9$/L)	1.00 (1.00, 1.00)	0.920		
BM Blast	1.00 (1.00, 1.01)	0.464		

Multivariate analysis was adjusted for age, gender, ELN 2022 risk category, and white blood cell (WBC) count. Patients were stratified into high and low *RHBDF2* expression groups by the median *RHBDF2* expression value (high n = 75, low n = 75). Hb, PLT, and BM Blast were not included in the final multivariate model due to lack of univariate significance ($p > 0.1$).

expression of multiple immune-regulatory molecules including HVEM, TIM3, and CD86, which have been reported to be associated with tumor immune evasion in AML [32]. CIBERSORT deconvolution analysis showed that *RHBDF2* expression was significantly associated with the abundance of monocytes, M1 macrophages, and regulatory T cells in AML bone marrow. Single-cell sequencing analysis further confirmed that *RHBDF2* was predominantly enriched in monocytes from AML patients. These observations indicate that *RHBDF2* upregulation is significantly correlated with immune cell infiltration in the AML bone marrow microenvironment, which is consistent with previous studies reporting the correlation between *RHBDF2* expression and immune cell infiltration in hepatocellular carcinoma [21,33].

Myeloid-derived suppressor cells (MDSCs) contribute to T cell tolerance via multiple mechanisms, including the expression of VISTA, PD-L1, IDO1, and arginase, along with the generation of reactive oxygen species (ROS), peroxynitrite, and various cytokines (TGF-β and IL-10), thereby facilitating AML immune evasion [34]. AML tumor cells may also employ extracellular vesicles or other mechanisms to endow monocytes with a suppressive phenotype characterized by CD14+HLA-DR- expression and upregulate MDSC signature genes [35]. Our research found that *RHBDF2* is highly enriched in monocytes from AML patients, and its expression is positively correlated with immunosuppressive checkpoints and inflammatory signatures. Based on the known function of *RHBDF2* as a regulator of ADAM17 in myeloid cells, we propose an unvalidated hypothesis that *RHBDF2* may modulate the immune regulatory and potentially immunosuppressive states of monocytes in AML via the ADAM17 pathway. This hypothesis will be systematically tested in our ongoing *in vitro* and *in vivo* functional studies, which are not presented in the current manuscript.

Indeed, this observational, correlative study presents certain inherent limitations. First, this is a hypothesis-generating study based on bioinformatic analysis and clinical sample expression validation, with no *in vitro* or *in vivo* functional assays performed to validate the causal relationship between *RHBDF2* and AML immune dysregulation. Mechanistic exploration of the association between *RHBDF2* and AML prognosis is still at an early stage.

Second, the prognostic nomogram model constructed in this study was only validated via internal bootstrap resampling (1000 iterations) in the TCGA AML cohort, with no independent external validation performed using a multi-

center clinical cohort. Although internal bootstrap validation confirmed the stability of the model, the generalizability of the nomogram across different clinical demographics, ethnic groups, and treatment centers remains unproven. Third, the exceptionally high 5-year AUC value of the model may also be partially attributed to the limited number of patients with complete 5-year follow-up data in the TCGA cohort, which may introduce a risk of overfitting. Future large-scale, multi-center, prospective clinical studies are required to validate the clinical utility and generalizability of this prognostic model in real-world AML patient populations. Fourth, the subgroup analysis of FAB classification subtypes included extremely small sample sizes for certain rare subtypes (e.g., M6, M7), which limits the statistical power of these subgroup analyses. The association between *RHBDF2* expression and FAB classification should be interpreted with caution, and requires validation in a larger cohort with sufficient sample sizes for all FAB subtypes. Fifth, the most significant limitation of our differential expression analysis is the inherent difference in cellular composition between AML bone marrow (dominated by blasts) and non-malignant bone marrow (composed of normal hematopoietic cells). This composition bias may affect the interpretation of gene expression differences. Sixth, we used the median *RHBDF2* expression value as the cutoff to stratify patients into high and low expression groups for survival and correlation analyses. This median-based dichotomy method may lead to the loss of information from the continuous variable, and may affect the stability of the grouping results in different cohorts. Future studies should explore the optimal cutoff value of *RHBDF2* expression through large-scale multi-center AML cohorts to improve the robustness of the results.

Conclusion

This observational, correlative study demonstrates that *RHBDF2* is upregulated in newly diagnosed AML patients, with dynamic expression changes across different disease states. High *RHBDF2* expression is correlated with adverse clinicopathological features and shorter overall survival in AML patients. Furthermore, *RHBDF2* expression is enriched in monocytes in the AML bone marrow microenvironment, and is positively correlated with immune cell infiltration and the expression of multiple immunosuppressive checkpoints. In conclusion, *RHBDF2* may be associated with prognosis in AML, but its clinical value as a prognostic biomarker requires further validation. Our correlative findings generate a testable hypothesis for future functional studies to validate the role of *RHBDF2* in AML immune dysregulation.

Abbreviations

AML, Acute myeloid leukemia; *RHBDF2*, Rhomboid 5 homolog 2; TCGA, The Cancer Genome Atlas; GTEx, The Genotype-Tissue Expression; UCSC, University of California Santa Cruz; GEPIA, The Gene Expression Profiling Interactive Analysis; UALCAN, University of Alabama at Birmingham Cancer data analysis Portal; DEG, Differentially expressed gene; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; BP, Biological Process; CC, Cellular Component; MF, Molecular Function; GSVA, Gene Set Variation Analysis.

Availability of Data and Materials

The datasets used or analyzed during the current study are available from the corresponding author via email request.

Author Contributions

QNJ, ZFX and HWW conceptualized and designed the whole research scheme. QNJ, JLL, YH, HCW, FGR, HJH and ZLJ conducted all experimental operations, processed raw data and interpreted relevant research results. QNJ, HCW and HJH drafted the original manuscript. SL, BBQ and YFZ analysed experimental data, supplemented and revised the content of experimental procedures and results sections of the manuscript. HWW and FGR acquired research funding, supervised the whole study, and polished the full text. All authors critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

This study was approved by the Institutional Review Board (IRB) Institutional of the Second Hospital of Shanxi Medical University (No.2023-YX-255) and all subjects signed written informed consent before sample collection. The clinical procedures conformed to the ethical guidelines set forth in the Declaration of Helsinki.

Acknowledgment

We gratefully thank Dr. Jie Liu of Department of Vascular and Endovascular Surgery, Chinese PLA General Hospital for his contribution to the statistical support and study design consultations.

Funding

This paper was funded by the National Natural Science Foundation of China (81670126) and The Shanxi Natural Science Foundation of China (202403021221321).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.24976/Discov.Med.202638210.165.

References

[1] Döhner H, Wei AH, Appelbaum FR, Craddock C, DiNardo CD, Dombret H, *et al*. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood. 2022; 140: 1345–1377. https://doi.org/10.1182/blood.2022016867.

[2] Liu H. Emerging agents and regimens for AML. Journal of Hematology & Oncology. 2021; 14: 49. https://doi.org/10.1186/s13045-021-01062-w.

[3] Forsberg M, Konopleva M. AML treatment: conventional chemotherapy and emerging novel agents. Trends in Pharmacological Sciences. 2024; 45: 430–448. https://doi.org/10.1016/j.tips.2024.03.005.

[4] Perzolli A, Koedijk JB, Zwaan CM, Heidenreich O. Targeting the innate immune system in pediatric and adult AML. Leukemia. 2024; 38: 1191–1201. https://doi.org/10.1038/s41375-024-02217-7.

[5] Kantarjian H, Kadia T, DiNardo C, Daver N, Borthakur G, Jabbour E, *et al*. Acute myeloid leukemia: current progress and future directions. Blood Cancer Journal. 2021; 11: 41. https://doi.org/10.1038/s41408-021-00425-3.

[6] De Martino M, Rathmell JC, Galluzzi L, Vanpouille-Box C. Cancer cell metabolism and antitumour immunity. Nature Reviews. Immunology. 2024; 24: 654–669. https://doi.org/10.1038/s41577-024-01026-4.

[7] Liang C, Wang X, Zhang Z, Xiao F, Feng H, Ma Q, *et al*. ACOT11 promotes cell proliferation, migration and invasion in lung adenocarcinoma. Translational Lung Cancer Research. 2020; 9: 1885–1903. https://doi.org/10.21037/tlcr-19-509.

[8] Di W, Fan W, Wu F, Shi Z, Wang Z, Yu M, *et al*. Clinical characterization and immunosuppressive regulation of CD161 (KLRB1) in glioma through 916 samples. Cancer Science. 2022; 113: 756–769. https://doi.org/10.1111/cas.15236.

[9] Isidori A, Cerchione C, Daver N, DiNardo C, Garcia-Manero G, Konopleva M, *et al*. Immunotherapy in Acute Myeloid Leukemia: Where We Stand. Frontiers in Oncology. 2021; 11: 656218. https://doi.org/10.3389/fonc.2021.656218.

[10] Subklewe M, Bücklein V, Sallman D, Daver N. Novel immunotherapies in the treatment of AML: is there hope? Hematology. American Society of Hematology. Education Program. 2023; 2023: 691–701. https://doi.org/10.1182/hematology.2023000455.

[11] Menter T, Tzankov A. Tumor Microenvironment in Acute Myeloid Leukemia: Adjusting Niches. Frontiers in Immunology. 2022; 13: 811144. https://doi.org/10.3389/fimmu.2022.811144.

[12] Restelli C, Ruella M, Paruzzo L, Tarella C, Pelicci PG, Colombo E. Recent Advances in Immune-Based Therapies for Acute Myeloid Leukemia. Blood Cancer Discovery. 2024; 5: 234–248. https://doi.org/10.1158/2643-3230.BCD-23-0202.

[13] Al-Salihi MA, Lang PA. iRhom2: An Emerging Adaptor Regulating Immunity and Disease. International Journal of Molecular Sciences. 2020; 21: 6570. https://doi.org/10.3390/ijms21186570.

[14] Dulloo I, Muliyil S, Freeman M. The molecular, cellular and pathophysiological roles of iRhom pseudoproteases. Open Biology. 2019; 9: 190003. https://doi.org/10.1098/rsob.190003.

[15] Kubo S, Fritz JM, Raquer-McKay HM, Kataria R, Vujkovic-Cvijin I, Al-Shaibi A, *et al*. Congenital iRHOM2 deficiency causes ADAM17 dysfunction and environmentally directed immunodysregulatory disease. Nature Immunology. 2022; 23: 75–85. https://doi.org/10.1038/s41590-021-01093-y.

[16] Lu F, Zhao H, Dai Y, Wang Y, Lee CH, Freeman M. Cryo-EM reveals that iRhom2 restrains ADAM17 protease activity to control the release of growth factor and inflammatory signals. Molecular Cell. 2024; 84: 2152–2165.e5. https://doi.org/10.1016/j.molcel.2024.04.025.

[17] Rosetti F, de la Cruz A, Crispín JC. Gene-function studies in systemic lupus erythematosus. Current Opinion in Rheumatology. 2019; 31: 185–192. https://doi.org/10.1097/BOR.0000000000000572.

[18] Ni R, Song G, Fu X, Song R, Li L, Pu W, *et al*. Reactive oxygen species-responsive dexamethasone-loaded nanoparticles for targeted treatment of rheumatoid arthritis via suppressing the iRhom2/TNF-α/BAFF signaling pathway. Biomaterials. 2020; 232: 119730. https://doi.org/10.1016/j.biomaterials.2019.119730.

[19] Geesala R, Schanz W, Biggs M, Dixit G, Skurski J, Gurung P, *et al*. Loss of RHBDF2 results in an early-onset spontaneous murine colitis. Journal of Leukocyte Biology. 2019; 105: 767–781. https://doi.org/10.1002/JLB.4A0718-283RR.

[20] Xu M, Tan J, Zhu L, Ge C, Zhang Y, Gao F, *et al*. Palmitoyltransferase ZDHHC3 Aggravates Nonalcoholic Steatohepatitis by Targeting S-Palmitoylated IRHOM2. Advanced Science. 2023; 10: e2302130. https://doi.org/10.1002/advs.202302130.

[21] Gong H, Xie H, Huangfu Z, Zhang S, Tang Y, Xiao M, *et al*. RHBDF2 is correlated with immune infiltrates in hepatocellular carcinoma and may have potential as a biomarker. FEBS Open Bio. 2023; 13: 881–897. https://doi.org/10.1002/2211-5463.13598.

[22] Sieber B, Lu F, Stribbling SM, Grieve AG, Ryan AJ, Freeman M. iRhom2 regulates ERBB signalling to promote KRAS-driven tumour growth of lung cancer cells. Journal of Cell Science. 2022; 135: jcs259949. https://doi.org/10.1242/jcs.259949.

[23] Xu Q, Chen C, Liu B, Lin Y, Zheng P, Zhou D, *et al*. Association of iRhom1 and iRhom2 expression with prognosis in patients with cervical cancer and possible signaling pathways. Oncology Reports. 2020; 43: 41–54. https://doi.org/10.3892/or.2019.7389.

[24] Gong H, Zhang Y, Chen X, Cao X, Tang L, Wang Y. The prognostic value of RHBDF2 in Pan-Cancer, and its correlation with cell Adhesion of Hepatocellular Carcinoma. Biotechnology & Genetic Engineering Reviews. 2024; 40: 1024–1048. https://doi.org/10.1080/02648725.2023.2191092.

[25] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014; 15: 550. https://doi.org/10.1186/s13059-014-0550-8.

[26] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics. 2012; 16: 284–287. https://doi.org/10.1089/omi.2011.0118.

[27] Song Y, Wang J, Sun J, Chen X, Shi J, Wu Z, *et al*. Screening of Potential Biomarkers for Gastric Cancer with Diagnostic Value Using Label-free Global Proteome Analysis. Genomics, Proteomics & Bioinformatics. 2020; 18: 679–695. https://doi.org/10.1016/j.gpb.2020.06.012.

[28] Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling Tumor Infiltrating Immune Cells with CIBERSORT. Methods in Molecular Biology. 2018; 1711: 243–259. https://doi.org/10.1007/978-1-4939-7493-1_12.

[29] Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, *et al*. Integrated analysis of multimodal single-cell data. Cell. 2021; 184: 3573–3587.e29. https://doi.org/10.1016/j.cell.2021.04.048.

[30] Zhang N, Wu J, Wang Q, Liang Y, Li X, Chen G, *et al*. Global burden of hematologic malignancies and evolution patterns over the past 30 years. Blood Cancer Journal. 2023; 13: 82. https://doi.org/10.1038/s41408-023-00853-3.

[31] Hosur V, Johnson KR, Burzenski LM, Stearns TM, Maser RS, Shultz LD. Rhbdf2 mutations increase its protein stability and drive EGFR hyperactivation through enhanced secretion of amphiregulin. Proceedings of the National Academy of Sciences of the United States of America. 2014; 111: E2200–E2209. https://doi.org/10.1073/pnas.1323908111.

[32] Abril-Rodriguez G, Ribas A. SnapShot: Immune Checkpoint Inhibitors. Cancer Cell. 2017; 31: 848–848.e1. https://doi.org/10.1016/j.ccell.2017.05.010.

[33] Alvarez Calderon F, Kang BH, Kyrysyuk O, Zheng S, Wang H, Mathewson ND, *et al*. Targeting of the CD161 inhibitory receptor enhances T-cell-mediated immunity against hematological malignancies. Blood. 2024; 143: 1124–1138. https://doi.org/10.1182/blood.2023022882.

[34] Yang Y, Li C, Liu T, Dai X, Bazhin AV. Myeloid-Derived Suppressor Cells in Tumors: From Mechanisms to Antigen Specificity and Microenvironmental Regulation. Frontiers in Immunology. 2020; 11: 1371. https://doi.org/10.3389/fimmu.2020.01371.

[35] Tohumeken S, Baur R, Böttcher M, Stoll A, Loschinski R, Panagiotidis K, *et al*. Palmitoylated Proteins on AML-Derived Extracellular Vesicles Promote Myeloid-Derived Suppressor Cell Differentiation via TLR2/Akt/mTOR Signaling. Cancer Research. 2020; 80: 3663–3676. https://doi.org/10.1158/0008-5472.CAN-20-0024.