Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Development and Validation of a Machine Learning–Based Cell Death–Immune Prognostic Model for Risk Stratification in Diffuse Large B-cell Lymphoma

**Zi-Yi Bai**
Jinan University
**Chao Yao**
Jinan University
**Yu-Qi Zhai**
Jinan University
**Xiao-Yuan Hou**
Jinan University
**Yu-Guang Li**
Jinan University
**Ze-Ping Weng**
Jinan University
**Jin-Qiong Xu**
Southern Medical University
**Xiao-Mei Xie**
Affiliated Hospital of Youjiang Medical University for Nationalities
**Ting-Zhuang Yi**
Affiliated Hospital of Youjiang Medical University for Nationalities
**Zhen-Yi Jin**
Jinan University
**Xiu-Li Wu**
siulier@163.com

Jinan University

**Research Article**

**Keywords:** Diffuse large B-cell lymphoma, programmed cell death, machine learning, immune microenvironment, drug sensitivity, prognostic model

**Posted Date:** October 1st, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-11002164/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

**Additional Declarations:** No competing interests reported.

## Abstract

# Background

Diffuse large B-cell lymphoma (DLBCL) exhibits marked prognostic heterogeneity, partly driven by interactions between regulated cell death and the tumour immune microenvironment. However, the interplay between immune status and cell death in DLBCL remains unclear.

# Methods

This study integrates bulk and single-cell transcriptomic data to characterise immune–cell death interactions in DLBCL. Immune–cell death–related genes are identified by intersecting differentially expressed genes with immune and cell death gene sets. A cell-death–immune score (CDIS) is constructed via 101 machine-learning algorithm combinations and validated in independent cohorts. We assess immune infiltration, cell death pathway activity, cell–cell communication, ligand–target interactions, pseudotime trajectories, drug connectivity, molecular docking, and virtual gene perturbation.

# Results

We identified 159 immune–cell death–related genes and derived a 16-gene CDIS from 25 candidates. The CoxBoost plus random survival forest (RSF) ensemble model achieved optimal performance. High CDIS was associated with suppressed immunogenic cell death programs and an immune-cold microenvironment, characterised by reduced effector immune infiltration and increased M2-like macrophages. Single-cell analysis revealed that model genes were expressed predominantly in macrophages and identified M2-like macrophage-centred pro-tumour survival signalling involving the BAFF/APRIL axis.

# Conclusions

CDIS offers a robust prognostic framework for DLBCL and uncovers immune escape mechanisms linked to macrophage-driven signalling and dysregulated cell death.

## Background

Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphoma and represents a biologically and clinically heterogeneous malignancy [1]. Although most patients receive rituximab-based chemoimmunotherapy as first-line treatment, a substantial proportion experience refractory disease or relapse, and durable benefit from subsequent therapeutic strategies remains limited [2]. Conventional prognostic systems, including the International Prognostic Index (IPI) and its refined versions (R-IPI), have improved clinical risk stratification but remain insufficient to fully capture the molecular and microenvironmental heterogeneity that drives treatment failure [3–6]. Therefore, identifying biologically informed biomarkers and actionable mechanisms is essential for improving prognostic assessment and developing precision therapeutic strategies in DLBCL.

The tumor immune microenvironment has emerged as a key determinant of lymphoma progression, therapeutic response, and clinical outcome [7, 8]. In DLBCL, immune infiltration [9], macrophage polarization, cytotoxic lymphocyte activity, stromal remodeling [10], and immune checkpoint-related dysfunction collectively shape tumor behavior [11]. However, the immune microenvironment does not act independently of tumor-intrinsic biological programs [12]. Among these programs, regulated cell death is particularly relevant because it influences antigen release, immune activation, inflammatory signaling, and therapeutic sensitivity [13]. Immunogenic cell death can convert dying tumor cells into an endogenous source of antitumor immune stimulation through the release of damage-associated molecular patterns, whereas defective immunogenic death may promote immune escape and resistance to immunotherapy [14].

Despite increasing interest in individual forms of regulated cell death, most studies in DLBCL have focused on a single death modality, such as ferroptosis [15], cuproptosis, or immunogenic cell death [16–18]. A systematic framework integrating multiple cell death programs with immune microenvironmental states is still lacking. This gap limits our ability to understand how death-related transcriptional states interact with immune suppression and how such interactions contribute to prognosis and therapeutic resistance.

In this study, we integrated bulk transcriptomic cohorts, single-cell RNA sequencing, machine learning, immune infiltration analysis, cell-cell communication inference, ligand-target prediction, drug connectivity analysis, molecular docking, and virtual gene perturbation to define an immune-cell death landscape in DLBCL. We developed a cell-death–immune score (CDIS) using 101 machine-learning algorithm combinations and validated its prognostic value across independent cohorts. We further found that high-CDIS tumors exhibited defective immunogenic cell death, together with an immune-cold microenvironment and an M1/M2 macrophage imbalance. At the single-cell level, we identified a macrophage-centered immunosuppressive communication network in which M2-like and IFN-related macrophages promote malignant B-cell survival through BAFF/APRIL signaling, while $CD4^+$ naive T cells may contribute to macrophage polarization through ANXA1–FPR signaling. These findings provide a multi-omics framework for DLBCL risk stratification and reveal a potentially targetable macrophage–BAFF/APRIL–malignant B-cell axis.

## Methods

To systematically investigate the immune-cell death landscape in DLBCL, we performed a multi-stage analysis encompassing prognostic model development, risk-group characterization, single-cell dissection of cell-cell communication, and virtual drug screening (Supplementary Figure S1).

## 2.1 Data acquisition and preprocessing

In this study, six DLBCL-related datasets were obtained from the NCBI GEO database, including five bulk transcriptomic datasets (GSE56315, GSE181063, GSE10846, GSE23501, and GSE32918) and one single-cell transcriptomic dataset (GSE182434). These bulk datasets were generated on microarray platforms: GSE56315, GSE10846, and GSE23501 were profiled using the Affymetrix Human Genome U133 Plus 2.0 Array (GPL570), while GSE181063 and GSE32918 were profiled using the Illumina HumanHT-12 WG-DASL platform. For each bulk dataset, expression values were log2-transformed where necessary and normalized across arrays; probe identifiers were collapsed to gene symbols by retaining the probe with the highest mean expression per gene, to ensure cross-platform comparability.

The single-cell transcriptomic dataset GSE182434, originally generated by Steen et al., was processed using the Seurat framework [19]. This dataset comprises sorted B cells and non-B cells from 7 DLBCL tumors, 3 follicular lymphoma tumors, and 1 reactive tonsillitis sample. In our analysis, we focused exclusively on DLBCL-derived cells to characterize the immune-cell death landscape in DLBCL. Quality control filtering was performed to exclude cells with fewer than 500 or more than 6,000 detected genes, as well as cells with a ribosomal RNA proportion greater than 10%. Doublets were further identified and removed using DoubletFinder, and only high-quality cells were retained for downstream analyses. A total of 3,000 highly variable genes were selected using “FindVariableFeatures”. After normalization with “NormalizeData”, scaled expression values were generated using “ScaleData”, followed by principal component analysis with “RunPCA”. Batch effects were corrected using the Harmony algorithm [20]. Louvain clustering was then conducted on Harmony-corrected principal components at a resolution of 0.5, and UMAP was applied for low-dimensional visualization. Cell-type annotation was initially performed using SingleR and subsequently refined by manual curation based on canonical marker genes. Copy number variation (CNV) inference in B-cell subsets was conducted using inferCNV [21], and k-means-based cross-validation was applied to improve classification accuracy, enabling robust discrimination of malignant B cells (Malig B). The resulting curated single-cell object provided the foundation for all subsequent analyses.

In addition, an independent anti–PD-L1 immunotherapy cohort of advanced urothelial carcinoma (IMvigor210) [22] was included for exploratory cross-cancer analysis; processed expression and clinical data were accessed through the IMvigor210CoreBiologies R package.

## 2.2 Screening of immune- and cell-death-related candidate genes

In the GSE56315 cohort (n = 88), differential expression analysis between patients with DLBCL and healthy controls was performed using the “limma” package [23]. Differentially expressed genes (DEGs) were identified using the thresholds of $P < 0.05$ and $|\log_2 FC| > 2$, and the results were visualized with a volcano plot. Genes associated with 31 cell death modalities were curated from the literature and the GeneCards database [24], yielding a total of 11,980 genes (Table S1). Immune-related genes were obtained from the ImmPort database [25], comprising 2,483 genes. By intersecting DEGs with the cell-death-related and immune-related gene sets, immune-cell death-related candidate genes were identified, and the overlap was illustrated using a Venn diagram.

## 2.3 Functional enrichment analysis

Functional enrichment analyses were conducted using the “clusterProfiler” package, including Gene Ontology (GO) categories—biological process (BP), cellular component (CC), and molecular function (MF)—as well as Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Terms with an adjusted P value < 0.05 were considered significantly enriched. The top 10 significantly enriched pathways were selected for presentation in each section.

## 2.4 Unsupervised clustering based on differentially expressed genes

In the tumor samples of the GSE56315 cohort, unsupervised clustering was performed based on immune-cell-death-related candidate genes. The optimal number of clusters was determined by integrating the cumulative distribution function (CDF) curve, relative change in the area under the CDF curve, and consensus heatmap patterns, resulting in two molecular subgroups (C1 and C2). Single-sample gene set enrichment analysis (ssGSEA) was subsequently conducted using the “GSVA” package to characterize differences in immune cell infiltration between the two subgroups. To further identify core genes, the overlap between immune-cell-death-related candidate genes and subgroup-specific differentially expressed genes was calculated, and the intersecting genes were defined as key immune-cell death genes, which were visualized using a Venn diagram. These intersecting genes were then submitted to the STRING database [26] to obtain a protein-protein interaction (PPI) network with interaction nodes and corresponding edge weights. The resulting network was further refined and visualized using the “igraph” and “ggraph” packages.

## 2.5 Feature prognostic gene selection and prognostic model construction using multiple machine learning algorithms

To systematically identify cell-death- and immune-related prognostic biomarkers in DLBCL and construct an efficient predictive model, we implemented nine classical machine learning algorithms: Lasso, StepCox, random survival forest (RSF), CoxBoost, GBM, SuperPC, Elastic Net (Enet), Ridge, and plsRcox. Through multi-algorithm integration and parameter optimization, a total of 101 predictive model combinations were developed [27]. Model training and key gene selection were performed in the GSE181063 training cohort, which included 1,303 patients after quality control filtering (the original dataset contained 1,311 samples; 8 were excluded due to incomplete survival data or expression outliers). Subsequently, model performance was rigorously evaluated in three independent validation cohorts—GSE10846 (n = 414), GSE23501 (n = 69), and GSE32918 (n = 244)—using the concordance index (C-index), and the optimal model was selected to calculate the cell-death–immune score (CDIS) for each sample. Patients were then stratified into high- and low-risk groups according to the median CDIS within each cohort. Kaplan-Meier survival analysis and time-dependent ROC analysis were conducted to comprehensively assess the prognostic value of CDIS at 1, 3, and 5 years. In addition, to improve model interpretability, we applied the SHAP framework to the random survival forest model [28]. SHAP values for all samples in the test sets were computed using KernelExplainer and visualized with a SHAP beeswarm plot. The final model genes and corresponding coefficients are provided in Supplementary Table S3.

## 2.6 Cell death score and immune infiltration assessment

Single-sample gene set enrichment analysis (ssGSEA) was performed using the "GSVA" R package to calculate enrichment scores for each sample, including 31 cell death patterns (Supplementary Table S1) and 29 immune-related gene sets. The score distributions were visualized using violin-box plots. In addition, the "estimate" package was applied to compare tumor purity, immune score, and stromal score between the high- and low-risk groups. The same gene set scoring approach was also used to evaluate M1/M2 polarization and functional characteristics of macrophages, as well as anti-apoptotic, BCR, MYC, NF-κB, and other functional modules in malignant B cells (Supplementary Table S2). In parallel, CIBERSORT with the LM22 signature matrix was applied to estimate the relative proportions of 22 immune cell subsets from the bulk expression profiles [29].

## 2.7 Construction and validation of a nomogram integrating CDIS and clinical features

After univariable and multivariable Cox regression analyses of CDIS and clinical characteristics (age, molecular subtype, sex, ECOG score, stage, LDH, IPI score, ALB, HB, number of lymph nodes, and WBC), all independent prognostic factors were integrated to construct a Nomogram model. Its predictive performance was evaluated using calibration plots, ROC curves, and decision curve analysis (DCA).

## 2.8 Cell–cell communication analysis

CellChat was used [30, 31] to evaluate the number and strength of intercellular communications among all cell subpopulations, which were visualized with circle plots. Chord diagrams were further generated to illustrate key signaling pathways, including BAFF, APRIL, and ANXA1–FPR. Based on the CellChat framework, NicheNet was subsequently applied [32] to systematically predict ligand–receptor regulatory interactions between monocytes/macrophages (Macro) and malignant B cells (Malig B). Ligands expressed in at least 10% of Macro cells (sender cells) and receptors expressed in at least 10% of Malig B cells (receiver cells) were used as the filtering criteria. Pearson correlation coefficients were then used as activity scores, and the top 20 ranked ligands were visualized in a heatmap. In addition, NicheNet target gene prediction was performed to identify downstream genes regulated by these ligands in Malig B cells, and the enrichment patterns of key target genes were displayed using heatmaps.

## 2.9 Pseudotime trajectory analysis

We used the "Monocle3" package [33] to analyze and visualize the differentiation trajectories of malignant/non-malignant B cells and monocyte/macrophage (Macro) subpopulations. For key genes in the risk model and signaling pathways, scatter plots with fitted smooth curves were generated to illustrate their dynamic expression patterns along the pseudotime axis.

## 2.10 Virtual drug screening

Based on "scFOCAL", single-cell transcriptomic profiles were mapped to L1000 drug perturbation signatures [34]. Mean connectivity scores were calculated within each cell subpopulation, followed by between-group differential analysis to identify candidate compounds with the most negative connectivity. Candidate drugs were visualized using pheatmap, and their distributional differences across subpopulations were further validated. In addition, drug sensitivity (half-maximal inhibitory concentration, IC50) for patients in the GSE181063 cohort was predicted using the oncoPredict R package with Genomics of Drug Sensitivity in Cancer (GDSC) training data [35], and predicted IC50 values were compared between the high- and low-risk groups.

## 2.11 Molecular docking analysis

The three-dimensional crystal structures of target proteins were obtained from the Protein Data Bank [36] in PDB format. Target compounds were retrieved from PubChem [37] in SDF format. For high-throughput evaluation of drug-protein binding potential, protein binding pockets were first identified using ProteinsPlus [38], followed by calculation of drug-protein binding scores in OmicsTools, with results visualized as heatmaps. For docking visualization, small-molecule compounds were energy-minimized in ChemBio3D Ultra and saved as MOL2 files. Protein structures were preprocessed in PyMOL by removing solvent molecules and then saved as PDB files. Molecular docking was subsequently performed using AutoDock Vina [39]. Favorable binding conformations were selected based on affinity scores and further visualized in PyMOL.

## 2.12 Virtual gene knockout analysis

To further investigate the functional roles of key genes in cells, we applied the scTenifoldKnk virtual gene knockout framework [40]. Specifically, we simulated knockout of TNFRSF13B, TNFRSF13C, and TNFRSF17 in malignant B cells; TNFSF13 and TNFSF13B, along with the top six model genes ranked by contribution, in monocytes/macrophages; and ANXA1 in $CD4^+$ naive T cells. Differentially regulated genes were ranked by log2FC and Z-score, and the top 20 genes were visualized using bar plots. Subsequently, enrichment analyses were performed on these top 20 regulated genes to characterize functional perturbations in each cell population.

## 2.13 Statistical analysis

Differences between two groups were assessed using the Wilcoxon rank-sum test or Student's t-test. For exploratory analyses involving multiple cell death and immune infiltration signatures, nominal P values are reported without adjustment for multiple comparisons, whereas FDR correction (Benjamini–Hochberg method) was applied where explicitly indicated (e.g., functional enrichment analysis).

# Results

## 3.1 Identification of immune–cell death–related genes and DLBCL subtypes

To identify immune–cell death–related genes associated with DLBCL, we first compared transcriptomic profiles between DLBCL samples and healthy controls in the GSE56315 cohort. Using $|\log_2 FC| > 2$ and $P < 0.05$ as thresholds, we identified 2,566 differentially expressed genes, including 1,761 downregulated and 805 upregulated genes (Fig. 1A). We then integrated these DEGs with gene sets related to 31 cell death modalities and immune-related genes from ImmPort, yielding 159 immune–cell death–related genes (Fig. 1B-C). Functional enrichment analysis showed that these genes were mainly

enriched in immune regulatory processes, including humoral immune response, leukocyte migration, chemokine activity, and cytokine receptor binding. KEGG analysis further indicated significant enrichment in cytokine-cytokine receptor interaction, chemokine signaling, T-cell receptor signaling, natural killer cell-mediated cytotoxicity, and NF-κB signaling pathways (Fig. 1D-E). Based on the expression profiles of these 159 genes, unsupervised clustering stratified DLBCL patients into two molecular subtypes, C1 and C2 (Supplementary Figure S2A-C). Principal component analysis showed a clear separation between the two subtypes (Fig. 1F). Differential expression analysis between C1 and C2 was then intersected with the 159 immune–cell death–related genes, yielding 25 candidate genes (Fig. 1G). Protein-protein interaction analysis revealed that CCL5, CCL2, CCR5, CCR1, CXCL10, CXCL9, and TYROBP occupied central positions within the network, suggesting their potential roles in immune-death regulation in DLBCL (Fig. 1H). Immune infiltration analysis demonstrated significant differences in multiple immune cell populations and immune functional signatures between C1 and C2, indicating that immune–cell death–related genes can distinguish DLBCL subtypes with distinct immune microenvironmental features (Fig. 1I).

## 3.2 Machine learning identifies CoxBoost plus RSF as the optimal model

Based on the 25 candidate immune–cell death–related genes, we constructed prognostic models in the training cohort GSE181063 and validated them in three independent cohorts, including GSE10846, GSE23501, and GSE32918. By integrating nine machine-learning algorithms and their combinations, a total of 101 predictive models were developed and compared using the average C-index (Fig. 2A). Among these models, the CoxBoost plus random survival forest (RSF) model achieved the best overall performance, with C-index values of 0.913, 0.620, 0.723, and 0.614 in the training and validation cohorts, respectively, and an average C-index of 0.718 (Fig. 2B). From the 25 candidate genes, the optimal model retained 16 as the final signature, combining CoxBoost's inherent variable selection (non-zero coefficients) with the RSF-derived importance ranking. The risk score derived from the optimal model was defined as the cell-death–immune score (CDIS). Patients were stratified into high- and low-risk groups according to the median CDIS. As CDIS increased, the number of death events rose and survival time shortened, indicating that CDIS effectively captured prognosis-related risk in DLBCL (Fig. 2C). SHAP-based model interpretation identified LYZ, CCL5, CCL18, GNLY, AQP9, CXCL11, and TLR8 as the most influential contributors to the model (Fig. 2D). Among them, LYZ, CCL5, and GNLY tended to act as protective factors, whereas CCL18, AQP9, CXCL11, and TLR8 were associated with increased risk. The expression heatmap further demonstrated distinct expression patterns of the 16 model genes between the high- and low-risk groups (Fig. 2E). In addition, all 16 genes were upregulated in DLBCL samples compared with healthy controls (Supplementary Figure S2F). Kaplan-Meier survival analysis demonstrated that patients in the high-CDIS group had significantly poorer overall survival than those in the low-CDIS group across the training cohort and all three validation cohorts (GSE181063, GSE10846, and GSE32918: $P < 0.001$; GSE23501: $P = 0.033$) (Fig. 2F, H, J, L). Time-dependent ROC analysis further confirmed the prognostic performance of CDIS. In the GSE181063 training cohort, the 1-, 3-, and 5-year AUCs were 0.973, 0.982, and 0.979, respectively. In the validation cohorts, the 1-, 3-, and 5-year AUCs ranged from 0.638 to 0.758, 0.637 to 0.748, and 0.593 to 0.661, respectively, supporting the robust prognostic discrimination of CDIS across independent cohorts (Fig. 2G, I, K, M).

## 3.3 Integration of CDIS with clinical variables improves prognosis prediction

To evaluate the independent clinical value of CDIS in prognostic stratification, we incorporated CDIS together with multiple clinical variables in the GSE181063 cohort, including age, molecular subtype, sex, ECOG performance status, B symptoms, Ann Arbor stage, LDH level, IPI score, albumin, hemoglobin, lymphocyte count, and white blood cell count. Univariable Cox regression analysis showed that CDIS and several clinical characteristics were significantly associated with overall survival (Fig. 3A). Further multivariable Cox regression analysis identified CDIS, age, ECOG performance status, and molecular subtype as independent prognostic factors in patients with DLBCL (Fig. 3B). Based on these independent prognostic factors, we constructed a Nomogram integrating CDIS and clinical variables to estimate 1-, 3-, and 5-year overall survival probabilities (Fig. 3C). Calibration curves demonstrated good agreement between the predicted and observed survival outcomes (Fig. 3D). Decision curve analysis further showed that the integrated model provided favorable clinical net benefit across different time points, supporting its potential clinical utility (Fig. 3E). In addition, Chi-square tests demonstrated significant associations between CDIS-based risk stratification (high/low) and clinical risk factors such as molecular subtype (GCB, ABC) and IPI score, reinforcing the clinical utility of CDIS (Fig. 3J-K). The nomogram achieved higher 1-, 3-, and 5-year areas under the receiver operating characteristic curve than CDIS or any individual clinical variable alone (Fig. 3F–I).

## 3.4 Defective immunogenic cell death and immunosuppressive microenvironment in high-CDIS tumors

To further characterize the biological features captured by CDIS, we compared 31 cell death modalities between the high- and low-risk groups. GSVA analysis showed that immunogenic cell-death-related pathways, including pyroptosis, necroptosis, and PANoptosis, were significantly suppressed in the high-risk group (Fig. 4A, D). These findings indicate that high-CDIS tumors are characterized by defective immunogenic cell death. We next investigated the immune infiltration landscape across CDIS-defined risk groups. Both single-sample gene set enrichment analysis (ssGSEA) and CIBERSORT [29] analyses revealed that effector immune components, including $CD8^+$ T cells, Th1/Th2 cells, and M1 macrophages, were markedly reduced in the high-risk group, whereas monocytes and M2 macrophages were increased (Fig. 4B, E; Supplementary Figure S2D-E). Notably, the high-risk group showed a pronounced inversion of the M1/M2 macrophage ratio, suggesting that macrophage polarization toward an M2-like protumor phenotype represents a key immunological feature of the high-CDIS state.

Correlation analysis further demonstrated that the seven most influential model genes were closely associated with multiple cell death programs and immune infiltration signatures, particularly M1- and M2-related macrophage features (Fig. 4C). Consistently, ESTIMATE analysis showed significantly lower ImmuneScore, StromalScore, and ESTIMATE score, together with increased tumor purity, in the high-risk group (Fig. 4F-I). Collectively, these results suggest that high-CDIS tumors are characterized by an immune-cold microenvironment marked by reduced immune infiltration, weakened stromal response, and increased tumor cellularity.

To examine whether the immune-cold phenotype captured by CDIS extends beyond DLBCL, we applied the model to an independent anti–PD-L1-treated urothelial carcinoma cohort (IMvigor210) [22] (Supplementary Figure S2G-H). In this exploratory cross-cancer analysis, patients in the high-CDIS group showed a lower objective response rate (15.2%) compared with the low-CDIS group (25.6%, P = 0.08), as assessed by RECIST criteria, and significantly worse overall survival. While this cross-cancer analysis does not directly address immunotherapy response in DLBCL, it suggests that the CDIS-defined immune-cold signature may have broader relevance across tumor types, warranting further investigation in DLBCL-specific immunotherapy cohorts.

## 3.5 Localization of CDIS model genes and identification of malignant B cells

To further determine the cellular origin of CDIS-related genes within the DLBCL tumor microenvironment, we performed an integrated analysis of the GSE182434 single-cell transcriptomic dataset. After quality control, dimensionality reduction, clustering, and cell annotation, major cell populations were identified, including B cells, $CD4^+$ naive T cells, regulatory T cells, $CD8^+$ naive T cells, $CD8^+$ effector T cells, monocytes/macrophages, natural killer cells, and plasmacytoid dendritic cells (Fig. 5A-D). The expression distribution of CDIS model genes showed that these genes were mainly enriched in monocytes/macrophages, natural killer cells, and $CD8^+$ naive T cells, rather than being predominantly expressed by B cells themselves (Fig. 5E-F). These findings suggest that CDIS primarily reflects the host antitumor immune state within the DLBCL microenvironment. To distinguish malignant from non-malignant B cells, we further extracted B cells and performed subclustering, identifying six B-cell subpopulations (Fig. 5G). We then inferred chromosomal copy number variation burden using inferCNV. Distinct CNV patterns were observed across B-cell subclusters, with some subpopulations showing high CNV burden consistent with malignant features, whereas clusters with lower CNV scores were more similar to non-malignant B cells (Fig. 5H-I). Cross-validation with unsupervised clustering further identified B-cell subsets with CNV profiles resembling normal cells, allowing B cells to be reclassified as malignant or non-malignant B cells (Fig. 5J-L).

Functional module scoring revealed that malignant B cells exhibited significantly higher activity of anti-apoptotic, BCR, MYC, NF-κB, and prosurvival transcription factor-related programs than non-malignant B cells, indicating enhanced survival and proliferative potential (Fig. 5M). When malignant B cells were further compared according to CDIS risk stratification, immunogenic death-related programs, including pyroptosis, paraptosis-like death and PANoptosis, were reduced in the high-risk group, consistent with the bulk RNA-seq findings (Fig. 5N). Together, these single-cell results validate the suppression of immunogenic cell death in high-CDIS tumors and provide a cellular basis for subsequent investigation of communication between malignant B cells and immune cells.

## 3.6 M2-like macrophages drive BAFF/APRIL-mediated prosurvival signaling to malignant B cells

To further investigate the interactions between malignant B cells and the tumor microenvironment, we projected the redefined malignant and non-malignant B cells back onto the integrated single-cell atlas and constructed a cell-cell communication network (Fig. 6A). Global communication analysis showed that $CD4^+$ naive T cells, malignant B cells, and monocytes/macrophages exhibited the highest number and strength of interactions, suggesting that these populations may cooperate to shape the immune microenvironment of DLBCL (Fig. 6B). At the pathway level, BAFF, APRIL, and ANXA1–FPR signaling displayed prominent communication activity (Fig. 6C). Among these, BAFF/APRIL signaling from monocytes/macrophages to malignant B cells was particularly strong. Ligand-receptor analysis showed that BAFF signaling was mainly mediated by TNFSF13B interacting with TNFRSF17, TNFRSF13C, and TNFRSF13B, whereas APRIL signaling was primarily mediated by TNFSF13 interacting with TNFRSF17 and TNFRSF13B (Fig. 6D). Given that the BAFF/APRIL axis can activate NF-κB signaling and induce anti-apoptotic BCL2 family members, these findings suggest that monocytes/macrophages may enhance malignant B-cell survival and suppress immunogenic cell death through BAFF/APRIL signaling. Functional role analysis further indicated that, in addition to $CD4^+$ naive T cells, monocytes/macrophages, and malignant B cells, NK cells and plasmacytoid dendritic cells also contributed to the regulation of BAFF-, APRIL-, and ANXA1–FPR-related communication networks (Fig. 6E, Supplementary Figure S3A).

To identify the macrophage subsets responsible for these prosurvival signals, we further subdivided monocytes/macrophages into five functional subsets: M0, M1, M2-LA, M2-RTM, and IFN macrophages [41, 42] (Fig. 6F, Supplementary Figure S3F). Communication analysis with malignant B cells revealed that BAFF and APRIL signaling mainly originated from M2-LA, M2-RTM, and IFN macrophage subsets, whereas classical M1 macrophages contributed minimally to this prosurvival communication axis (Fig. 6H). Functional scoring showed that M2-like subsets displayed stronger phagocytic, angiogenic, and immunosuppressive features. Although the IFN subset retained partial M1-like characteristics, it also exhibited high immunosuppressive and phagocytic scores, suggesting a protumor inflammatory macrophage state (Fig. 6I). In addition, strong interactions were also observed among multiple macrophage subgroups, suggesting the presence of a complex cooperative regulatory network within the macrophage compartment (Supplementary Figure S3G). We further stratified patients in the single-cell dataset into high- and low-risk groups based on the CDIS model. Cell-cell communication analysis showed that, in the high-risk group, both the number and strength of communications from monocytes/macrophages to the other subgroups were significantly increased (Supplementary Figure S3B). Moreover, the BAFF, APRIL, and ANXA1–FPR pathways, as well as the expression of their ligand–receptor pairs, were also significantly elevated in the high-risk group (Supplementary Figure S3D-E).

To delineate the dynamic process of cell-state transitions, we performed pseudotime analysis separately on macrophages and B cells. The pseudotime trajectory of B cells showed a progressive differentiation from non-malignant B cells to malignant B cells (Supplementary Figure S3C). Pseudotime analysis of macrophages indicated differentiation from M0 and M1 states toward the M2-RTM, M2-LA, and IFN subgroups (Fig. 6G). Along this trajectory, the BAFF/APRIL ligands TNFSF13B and TNFSF13 progressively increased, suggesting a trend that macrophages exhibit gradually enhanced prosurvival signaling toward malignant B cells along M2-like polarization (Fig. 6J). In parallel, the ANXA1–FPR receptors FPR1 and FPR2 also increased along the pseudotime axis, implicating ANXA1–FPR signaling in macrophage M2-like polarization. CDIS model genes showed risk-consistent dynamic patterns: protective genes such as LYZ gradually decreased, whereas risk-associated genes such as AQP9 increased along the trajectory (Fig. 6J). Together, these findings support a multilayered immunosuppressive model in which $CD4^+$ naive T cells may promote M2-like macrophage polarization through ANXA1–FPR signaling, while M2-

like and IFN-related macrophages further enhance malignant B-cell survival through the BAFF/APRIL axis, thereby establishing a cellular communication network that favors DLBCL progression and immune escape.

## 3.7 NicheNet analysis of M2-like macrophage–BAFF/APRIL regulation

To further dissect the downstream regulatory effects of BAFF/APRIL and ANXA1–FPR signaling, we performed NicheNet analysis to infer ligand-receptor-target gene networks. Using M2-LA, M2-RTM, and IFN macrophage subsets as sender cells and malignant B cells as receiver cells, we found that TNFSF13B and TNFSF13 exhibited high ligand activity in macrophage subsets and were linked to multiple prosurvival target genes in malignant B cells, including BCL2, BCL2L1, and NFKB1 (Fig. 7A, Supplementary Figure S3J-L). Consistently, BAFF/APRIL-related receptors, including TNFRSF13B, TNFRSF13C, and TNFRSF17, were actively expressed in malignant B cells, further supporting the hypothesis that M2-like macrophages promote malignant B-cell survival through the BAFF/APRIL axis (Fig. 7B). In the communication analysis between $CD4^+$ naive T cells and macrophages, ANXA1 showed high ligand activity and was predicted to regulate target genes associated with M2-like polarization and immune modulation in macrophages (Supplementary Figure S3H). Its receptors, FPR1 and FPR2, were highly expressed in monocytes/macrophages (Supplementary Figure S3I), suggesting that $CD4^+$ naive T cells may promote immunosuppressive macrophage polarization through ANXA1–FPR signaling.

## 3.8 Drug connectivity and molecular docking for BAFF/APRIL targeting

To explore potential therapeutic strategies targeting this cellular crosstalk network, we performed drug connectivity analysis using scFOCAL. Malignant B cells and M2-RTM/IFN macrophage subsets displayed similar drug connectivity patterns, suggesting that M2-like macrophage states may be functionally coupled with the drug sensitivity or resistance profile of malignant B cells (Fig. 7C). This observation further supports a functional association between M2-like macrophages and malignant B cells at the level of predicted drug response. Subsequently, we used oncoPredict [35] with the GDSC database to predict drug sensitivity (IC50) in patients from GSE181063 and compared differences in drug sensitivity between the high- and low-risk groups. Among the top 10 drugs ranked by scFOCAL connectivity (Capecitabine, Lenvatinib, Afatinib, Nintedanib, Vemurafenib, Letrozole, Vismodegib, Erlotinib, Gefitinib, and Tamoxifen), Afatinib, Erlotinib, Tamoxifen, and Gefitinib also showed significant differences between the high- and low-risk groups, with all four exhibiting greater sensitivity in the high-risk group (Fig. 7D-E; Supplementary Figure S2I-K). These findings, derived from large-scale pharmacogenomic data, constitute the principal computational evidence for CDIS-associated differential drug sensitivity.

We further performed molecular docking between the top 10 connectivity-ranked drugs and TNFRSF13B, TNFRSF13, and TNFRSF13C. The results showed that most drugs had predicted binding affinities of − 5 to − 7 kcal/mol with the three receptors (Fig. 7F). Notably, afatinib–TNFSF13B (− 6.86 kcal/mol), erlotinib–TNFRSF13B (− 6.00 kcal/mol), gefitinib–TNFSF13B (− 6.55 kcal/mol), and tamoxifen–TNFSF13B (− 5.36 kcal/mol) displayed the highest predicted binding affinities among the tested compounds, although these values fall within the moderate-to-weak range and are considered exploratory computational evidence rather than confirmation of direct targeting. Three-dimensional docking models showed that these drug molecules could be embedded into key binding pockets of the corresponding proteins to form stable complexes (Fig. 7G).

Notably, Tamoxifen has been previously reported to reduce cell viability in DLBCL cell lines through estrogen receptor beta (ERβ) targeting, and a large cohort analysis showed that tamoxifen-treated breast cancer patients had a significantly reduced risk of developing DLBCL. These prior findings support the biological plausibility of Tamoxifen's anti-lymphoma activity, although its interaction with BAFF/APRIL receptors identified in our docking analysis remains speculative and requires experimental validation. The primary evidence for Tamoxifen's potential relevance in high-risk DLBCL thus stems from the scFOCAL and oncoPredict sensitivity signals, while the docking results provide hypothesis-generating clues for further mechanistic investigation.

## 3.9 Virtual knockout of BAFF/APRIL axis in antigen presentation and immune recognition

To further evaluate the functional impact of key genes within the BAFF/APRIL axis, we performed virtual gene knockout analysis using scTenifoldKnk. Knockout of TNFSF13B in macrophages predominantly affected immune recognition-related processes, including MHC class II receptor activity, antigen processing and presentation, and immune receptor activity (Fig. 7H). Knockout of TNFSF13 similarly altered immune regulatory pathways, accompanied by perturbations in extracellular matrix- and fibroblast-associated functions (Fig. 7J). In malignant B cells, knockout of TNFRSF13B mainly affected antigen receptor binding, MHC class II complex-related processes, and protein translation programs (Fig. 7I). Similarly, TNFRSF17 and TNFRSF13C primarily perturb immune regulation and oxidative stress/detoxification processes (e.g., antigen/immunoglobulin complexes, MHC class II, and glutathione transferases) (Fig. 7K; Supplementary Figure S3M). In the ANXA1–FPR1/2 pathway, knockout of ANXA1 primarily perturbed lymphocyte migration, immune receptor signaling, and NK cell cytotoxicity in $CD4^+$ naive T cells, suggesting that $CD4^+$ T cells may influence macrophage function through the ANXA1–FPR pathway (Supplementary Figure S3N). In addition, knockout of LYZ, CCL18, GNLY, or AQP9 in macrophages produced a consistent perturbation profile, affecting genes involved in antigen processing/presentation and immune receptor signaling, with XCR1/CLEC9A serving as central nodes across multiple models (Supplementary Figure S3O-R).

Together, these findings indicate that the BAFF/APRIL axis may not only support malignant B-cell survival but also modulate antigen presentation, immune recognition, and adaptive immune responses within the DLBCL microenvironment. Collectively, Fig. 7 provides multi-layered computational evidence — spanning ligand–target prediction, drug connectivity, molecular docking, and virtual perturbation — that the M2-like macrophage–BAFF/APRIL–malignant B-cell axis is a central mechanism underlying immune escape and therapeutic tolerance in DLBCL.

## Discussion

In this study, we developed and validated CDIS, a prognostic model that integrates immune and regulated cell death features in DLBCL. The score stratifies patients into clinically distinct risk groups, with high-risk tumors characterized by defective immunogenic cell death, immune-cold microenvironmental remodeling, and M2-like macrophage-driven BAFF/APRIL prosurvival signaling. These findings provide a multi-omics framework that connects cell death dysregulation, immune suppression, and therapeutic resistance in DLBCL.

CDIS complements existing clinical risk stratification. Traditional prognostic systems such as IPI and NCCN-IPI are primarily built on clinical parameters and incompletely capture the biological heterogeneity of DLBCL. Although molecular classifications such as COO subtyping have improved our understanding of disease diversity, they do not systematically incorporate immune and cell death features [43]. CDIS addresses this gap by quantifying a transcriptional state shaped jointly by immune activity and cell death regulation. Notably, CDIS model genes are predominantly expressed in immune cells, particularly macrophages, rather than in malignant B cells themselves, suggesting that CDIS functions as a molecular readout of host antitumor immune competence rather than a tumor-intrinsic signature. The improved performance of the CDIS-integrated nomogram further supports its potential as a complementary tool for clinical decision-making.

CDIS also differs from previously reported DLBCL prognostic signatures in several respects. Earlier models were generally built on a single cell death modality, such as ferroptosis or cuproptosis [17, 18], whereas CDIS integrates 31 cell death programs together with immune features, capturing the qualitative balance between immunogenic and non-immunogenic death rather than the activity of any single pathway. Methodologically, model selection was benchmarked across 101 algorithm combinations and externally validated in three independent cohorts profiled on different platforms, reducing algorithm- and platform-specific bias. Finally, the signature was mechanistically anchored at single-cell resolution: rather than reporting an unannotated gene list, CDIS was mapped to a concrete, potentially druggable macrophage–BAFF/APRIL–malignant B-cell axis, providing a biological narrative that links risk stratification to therapeutic hypotheses.

A key biological feature of high-CDIS tumors is defective immunogenic cell death. We observed selective suppression of pyroptosis, necroptosis, and PANoptosis in high-risk patients—pathways known to release DAMPs and promote antitumor immunity through dendritic cell activation and T-cell priming [44, 45]. This defect likely contributes to the immune-cold microenvironment and primary resistance to immune checkpoint blockade, a hypothesis consistent with cross-cancer findings in an independent anti-PD-1/PD-L1-treated urothelial carcinoma cohort. While stress-associated death programs such as ferroptosis were concurrently upregulated, their activation appears insufficient to restore effective antitumor immunity. Thus, high-CDIS DLBCL is not characterized by a global reduction in cell death, but rather by a qualitative shift away from immunogenic death modalities. This distinction may have therapeutic implications: strategies aimed at restoring immunogenic cell death—rather than simply inducing tumor cell killing—may be more effective in high-risk patients.

High-CDIS tumors also exhibit a macrophage-skewed, immune-cold microenvironment. Effector T cells and M1 macrophages are reduced, whereas M2-like macrophages are enriched, resulting in a pronounced M1/M2 ratio inversion—one of the most prominent immune features associated with high CDIS. This polarization is biologically significant, as M2-like macrophages promote immune suppression, angiogenesis, and tumor progression, while M1 macrophages support antigen presentation and antitumor inflammation [46, 47]. The convergence of immunogenic death suppression and M2 macrophage enrichment suggests a self-reinforcing loop: defective immunogenic death limits T-cell recruitment and activation, while M2 macrophages further suppress the residual immune response.

Single-cell analysis reveals that M2-like and IFN-related macrophages drive malignant B-cell survival through the BAFF/APRIL axis. This pathway emerged as the strongest communication axis between macrophages and malignant B cells, with ligand expression increasing progressively along the M0→M2 differentiation trajectory. NicheNet analysis confirmed that BAFF/APRIL targets prosurvival genes including BCL2, BCL2L1, and NFKB1 in malignant B cells. Given that BAFF and APRIL are known to activate NF-κB and anti-apoptotic programs, this axis provides a mechanistic explanation for the coexistence of M2 macrophage enrichment and defective immunogenic death in high-CDIS DLBCL [48, 49]. The upstream ANXA1–FPR signaling from $CD4^+$ T cells may represent an additional layer of regulation promoting M2 polarization, suggesting a multi-cellular circuit that sustains immune suppression and tumor survival [50].

From a translational perspective, the BAFF/APRIL axis may be therapeutically targetable. Drug connectivity and molecular docking analyses identified tamoxifen as a candidate with predicted binding to BAFF/APRIL-related receptors. This computational finding is supported by prior evidence that tamoxifen reduces DLBCL cell viability through ERβ targeting and that tamoxifen-treated breast cancer patients have a lower risk of developing DLBCL [51]. While these observations remain hypothesis-generating, they provide a rationale for future investigation of tamoxifen or BAFF/APRIL-targeted agents in combination with immunotherapies or chemotherapy in high-risk DLBCL.

Several limitations should be acknowledged. First, although CDIS was validated across multiple retrospective cohorts, the higher C-index in the training cohort (0.913) relative to the validation cohorts (0.614–0.723) indicates partial overfitting, a common feature of high-dimensional transcriptomic signatures; the consistent survival separation observed across all independent cohorts nevertheless supports its generalizability, and prospective validation is warranted. Second, the median-based stratification cutoff, although applied uniformly within each cohort, is data-driven and should be refined or replaced by a prespecified threshold in future prospective studies. Third, the single-cell dataset included only a limited number of DLBCL samples, and the macrophage subsets identified here require spatial and functional validation. Finally, the computational predictions — including cell–cell communication, NicheNet, drug connectivity, and molecular docking — yield mechanistic hypotheses rather than causal evidence. Functional experiments, including ligand-receptor blockade and co-culture systems, are essential to establish causality.

## Conclusions

In conclusion, CDIS provides an immune–cell-death-based framework for prognostic stratification in DLBCL. High-risk tumors are defined by defective immunogenic cell death, immune-cold microenvironmental remodeling, and M2-like macrophage-driven BAFF/APRIL prosurvival signaling. These findings offer a multi-omics perspective on DLBCL progression and suggest that targeting macrophage polarization and BAFF/APRIL-mediated malignant B-cell survival may represent a promising strategy for improving outcomes in high-risk patients.

## Abbreviations

ABC
activated B-cell
AUC
area under the curve
BAFF
B-cell activating factor
APRIL
a proliferation-inducing ligand
BP
biological process
CC
cellular component
CDF
cumulative distribution function
CDIS
cell-death–immune score
C-index
concordance index
CNV
copy number variation
COO
cell of origin
DAMP
damage-associated molecular pattern
DCA
decision curve analysis
DEG
differentially expressed gene
DLBCL
diffuse large B-cell lymphoma
ECOG
Eastern Cooperative Oncology Group
FDR
false discovery rate
GCB
germinal center B-cell
GDSC
Genomics of Drug Sensitivity in Cancer
GO
Gene Ontology
ICD
immunogenic cell death
IPI
International Prognostic Index
KEGG
Kyoto Encyclopedia of Genes and Genomes
LDH
lactate dehydrogenase
MF
molecular function
NK
natural killer
PCA

principal component analysis
pDC
plasmacytoid dendritic cell
PPI
protein-protein interaction
ROC
receiver operating characteristic
RSF
random survival forest
SHAP
SHapley Additive exPlanations
ssGSEA
single-sample gene set enrichment analysis
Treg
regulatory T cell
UMAP
uniform manifold approximation and projection

## Declarations

# Ethics approval and consent to participate

All datasets analyzed in this study were obtained from publicly available sources (GEO). The original studies had obtained institutional ethics approval and informed consent. The secondary analysis was approved by the Research Ethics Committee of the Medical School of Jinan University (approval number: JNUKY-2023-0104), which waived the need for additional consent for the use of de-identified public data.

**Consent for publication**

All authors have read the manuscript and agree for publication.

# Competing interests

The authors declare no competing interests.

# Funding

This study was funded by the Natural Science Foundation of Guangdong Province (Nos. 2025A1515010192 and 2023A1515030271, ZYJ), the National Innovation and Entrepreneurship Training Program For Undergraduate (Nos. 202510559116 and 202610559049, YGL), Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation (“Climbing Program” Special Funds) (No. pdjh2026bk028, YGL), Jinan University “Challenge Cup” and other student extracurricular academic, scientific, technological innovation and entrepreneurship competition funded projects (No. 20252003, YGL), the Fund of Jinan-Puhua Joint Laboratory (JNUC-2025-0456, XLW), and the 2025–2026 Youth Science and Technology Talent Cultivation Program of Guangdong Provincial Association for Science and Technology (No. SKXRC2025351), Guangxi Natural Science Foundation under Grant (No. 2025GXNSFHA069247, YTZ).

# Author Contribution

ZYB and CY contributed to concept development and study design. YQZ and XYH acquired and curated the public datasets. ZYB and CY analyzed the data and wrote the manuscript. YQZ, XYH, YGL, ZPW, JQX, and XMX assisted in data analysis. YTZ, ZYJ, and XLW supervised the study, acquired funding, interpreted the data, and revised the manuscript. All authors read and approved the final manuscript.

# Acknowledgements

The authors would like to thank the patients and their families for consenting to the use of the sequencing data.

# Data Availability

The bulk transcriptomic and single-cell RNA sequencing datasets analyzed in this study are publicly available from the NCBI GEO database under accession numbers GSE56315, GSE181063, GSE10846, GSE23501, GSE32918, and GSE182434.

## References


1. ChapuyB,StewartC,DunfordAJ,etal.Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes.Nat Med.2018;24:679 – 90.https://doi.org/10.1038/s41591-018-0016-8.
2. BarracloughA,HawkesE,SehnLH,SmithSM.Diffuse large B-cell lymphoma.Hematol Oncol.2023;42:e3202.https://doi.org/10.1002/hon.3202.
3. ZhouZ,SehnLH,RademakerAW,etal.An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era.Blood.2014;123:837 – 42.https://doi.org/10.1182/blood-2013-09-524108.
4. RuppertAS,DixonJG,SallesG,etal.International prognostic indices in diffuse large B-cell lymphoma: a comparison of IPI, R-IPI, and NCCN-IPI.Blood.2020;135:2041-8.https://doi.org/10.1182/blood.2019002729.
5. JelicicJ,Juul-JensenK,BukumiricZ,etal.Prognostic indices in diffuse large B-cell lymphoma: a population-based comparison and validation study of multiple models.Blood Cancer J.2023;13:157.https://doi.org/10.1038/s41408-023-00930-7.
6. Huo YJ, Xu PP, Wang L, Zhong HJ, Fu D, Shi Q, et al. Dynamic change of soluble interleukin-2 receptor distinguished molecular heterogeneity and microenvironment alterations in diffuse large B-cell lymphoma. Biomark Res. 2022. https://doi.org/10.1186/s40364-022-00401-4.
7. WangPL,LaiW,ZhengJM,etal.Heterogeneous characteristics of γδ T cells in peripheral blood of diffuse large B-cell lymphoma.Biomark Res.2025;13:82.https://doi.org/10.1186/s40364-025-00795-x.
8. SchollerN,PerbostR,LockeFL,etal.Tumor immune contexture is a determinant of anti-CD19 CAR T cell efficacy in large B cell lymphoma.Nat Med.2022;28:1872-82.https://doi.org/10.1038/s41591-022-01916-x.
9. DaiY,KizhakeyilA,ChiharaD,etal.Multi-modal spatial characterization of tumor immune microenvironments identifies targetable inflammatory niches in diffuse large B cell lymphoma.Nat Genet.2025;57:2715-27.https://doi.org/10.1038/s41588-025-02353-5.
10. CzernilofskyF,MathioudakiA,Jopp-SaileL,etal.Reprogramming of stroma-derived chemokine networks drives the loss of tissue organization in nodal B cell lymphoma.Nat Cancer.2026;7:538 – 52.https://doi.org/10.1038/s43018-026-01136-z.
11. WenZF,HuangQT,WangYY,etal.MCP-1-CCR2-M2 macrophages axis contributes to diffuse large B-cell lymphoma progression and inhibits antitumor immune response.Sci Rep.2025;15:29044.https://doi.org/10.1038/s41598-025-14527-6.
12. ScottDW,GascoyneRD.The tumour microenvironment in B cell lymphomas.Nat Rev Cancer.2014;14:517–534.https://doi.org/10.1038/nrc3774.
13. GaoW,WangX,ZhouY,etal.Autophagy, ferroptosis, pyroptosis, and necroptosis in tumor immunotherapy.Signal Transduct Target Ther.2022;7:196.https://doi.org/10.1038/s41392-022-01046-3.
14. GalluzziL,HumeauJ,BuquéA,etal.Immunostimulation with chemotherapy in the era of immune checkpoint inhibitors.Nat Rev Clin Oncol.2020;17:725 – 41.https://doi.org/10.1038/s41571-020-0413-z.
15. DaiZ,LiuJ,ZengL,etal.Targeting ferroptosis in cancer therapy: mechanisms, strategies, and clinical applications.Cell Investigation.2025;1:100049.https://doi.org/10.1016/j.clnves.2025.100049.
16. WuJ,YeY,RanD,etal.Chidamide synergizes with cisplatin-etoposide to trigger pyroptosis and anti-tumor immunity in diffuse large B-cell lymphoma.Commun Med (Lond).2026.https://doi.org/10.1038/s43856-026-01598-3.
17. Jia Z, Zhang J, Li Z, Ai L. Identification of ferroptosis-related genes associated with diffuse large B-cell lymphoma via bioinformatics and machine learning approaches. Int J Biol Macromol. 2024;282:137117. https://doi.org/10.1016/j.ijbiomac.2024.137117.
18. WangN,ShiS,LiM,etal.Development and validation of a combined cuproptosis and immunogenic cell death prognostic model for diffuse large B-cell lymphoma.Aging (Albany NY).2024;16:1218-36.https://doi.org/10.18632/aging.205399.
19. SteenCB,LucaBA,EsfahaniMS,etal.The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma.Cancer Cell.2021;39:1422-37.e10.https://doi.org/10.1016/j.ccell.2021.08.011.
20. KorsunskyI,MillardN,FanJ,etal.Fast, sensitive and accurate integration of single-cell data with Harmony.Nat Methods.2019;16:1289-96.https://doi.org/10.1038/s41592-019-0619-0.
21. PatelAP,TiroshI,TrombettaJJ,etal.Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma.Science.2014;344:1396 – 401.https://doi.org/10.1126/science.1254257.
22. MariathasanS,TurleySJ,NicklesD,etal.TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells.Nature.2018;554:544-8.https://doi.org/10.1038/nature25501.
23. RitchieME,PhipsonB,WuD,etal.limma powers differential expression analyses for RNA-sequencing and microarray studies.Nucleic Acids Res.2015;43:e47.https://doi.org/10.1093/nar/gkv007.
24. RebhanM,Chalifa-CaspiV,PriluskyJ,LancetD.GeneCards: integrating information about genes, proteins and diseases.Trends Genet.1997;13:163.https://doi.org/10.1016/s0168-9525(97)01103-7.
25. BhattacharyaS,DunnP,ThomasCG,etal.ImmPort, toward repurposing of open access immunological assay data for translational and clinical research.Sci Data.2018;5:180015.https://doi.org/10.1038/sdata.2018.15.
26. SzklarczykD,KirschR,KoutrouliM,etal.The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest.Nucleic Acids Res.2023;51:D638-46.https://doi.org/10.1093/nar/gkac1000.
27. YinW,ChenG,LiY,etal.Identification of a 9-gene signature to enhance biochemical recurrence prediction in primary prostate cancer: a benchmarking study using ten machine learning methods and twelve patient cohorts.Cancer Lett.2024;588:216739.https://doi.org/10.1016/j.canlet.2024.216739.
28. LundbergSM,LeeSI.AunifiedapproachtointerpretingmodelpredictionsIn:Proceedingsofthe31stInternationalConferenceonNeuralInformationProcessingSys 77.

29. NewmanAM,LiuCL,GreenMR,etal.Robust enumeration of cell subsets from tissue expression profiles.Nat Methods.2015;12:453-7.https://doi.org/10.1038/nmeth.3337.
30. JinS,Guerrero-JuarezCF,ZhangL,etal.Inference and analysis of cell-cell communication using CellChat.Nat Commun.2021;12:1088.https://doi.org/10.1038/s41467-021-21246-9.
31. JinS,PlikusMV,NieQ.CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics.Nat Protoc.2025;20:180–219.https://doi.org/10.1038/s41596-024-01045-4.
32. BrowaeysR,SaelensW,SaeysY.NicheNet: modeling intercellular communication by linking ligands to target genes.Nat Methods.2020;17:159 – 62.https://doi.org/10.1038/s41592-019-0667-5.
33. CaoJ,SpielmannM,QiuX,etal.The single-cell transcriptional landscape of mammalian organogenesis.Nature.2019;566:496–502.https://doi.org/10.1038/s41586-019-0969-x.
34. SuterRK,JermakowiczAM,VeeramachaneniR,etal.Drug and single-cell gene expression integration identifies sensitive and resistant glioblastoma cell populations.Nat Commun.2026;17:99.https://doi.org/10.1038/s41467-025-67783-5.
35. MaeserD,GruenerRF,HuangRS.oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data.Brief Bioinform.2021;22:bbab260.https://doi.org/10.1093/bib/bbab260.
36. BermanHM,WestbrookJ,FengZ,etal.The Protein Data Bank.Nucleic Acids Res.2000;28:235 – 42.https://doi.org/10.1093/nar/28.1.235.
37. KimS,ChenJ,ChengT,etal.PubChem 2023 update.Nucleic Acids Res.2023;51:D1373-80.https://doi.org/10.1093/nar/gkac956.
38. Schöning-StierandK,DiedrichK,EhrtC,etal.ProteinsPlus: a comprehensive collection of web-based molecular modeling tools.Nucleic Acids Res.2022;50:W611-5.https://doi.org/10.1093/nar/gkac398.
39. EberhardtJ,Santos-MartinsD,TillackAF,ForliS.AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings.J Chem Inf Model.2021;61:3891-8.https://doi.org/10.1021/acs.jcim.1c00203.
40. OsorioD,ZhongY,LiG,etal.scTenifoldKnk: an efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. Patterns (N Y).2022;3:100434.https://doi.org/10.1016/j.patter.2022.100434.
41. MaRY,BlackA,QianBZ.Macrophage diversity in cancer revisited in the era of single-cell omics.Trends Immunol.2022;43:546 – 63.https://doi.org/10.1016/j.it.2022.04.008.
42. LiYS,LaiWP,YinK,etal.Lipid-associated macrophages for osimertinib resistance and leptomeningeal metastases in NSCLC.Cell Rep.2024;43:114613.https://doi.org/10.1016/j.celrep.2024.114613.
43. KotlovN,BagaevA,RevueltaMV,etal.Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures.Cancer Discov.2021;11:1468-89.https://doi.org/10.1158/2159-8290.CD-20-0839.
44. GalluzziL,VitaleI,AaronsonSA,etal.Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018.Cell Death Differ.2018;25:486–541.https://doi.org/10.1038/s41418-017-0012-4.
45. KroemerG,GalluzziL,KeppO,ZitvogelL.Immunogenic cell death in cancer therapy.Annu Rev Immunol.2013;31:51–72.https://doi.org/10.1146/annurev-immunol-032712-100008.
46. LinM,MaS,SunL,QinZ.The prognostic value of tumor-associated macrophages detected by immunostaining in diffuse large B cell lymphoma: a meta-analysis.Front Oncol.2022;12:1094400.https://doi.org/10.3389/fonc.2022.1094400.
47. XieM,WeiY,YangL,etal.The prognostic significance of tumor-associated macrophages and NK cells in patients with diffuse large B-cell lymphoma treated with RCHOP.Blood.2014;124:5390.https://doi.org/10.1182/blood.V124.21.5390.5390.
48. UllahMA,MackayF.The BAFF-APRIL system in cancer.Cancers (Basel).2023;15:1791.https://doi.org/10.3390/cancers15061791.
49. VareliA,Vaidehi NarayananH,ClarkH,etal.Systems biology-enabled targeting of NF-κB and BCL2 overcomes microenvironment-mediated BH3-mimetic resistance in DLBCL.Cell Death Dis.2025;16:620.https://doi.org/10.1038/s41419-025-07942-0.
50. McArthurS,JubanG,GobbettiT,etal.Annexin A1 drives macrophage skewing to accelerate muscle regeneration through AMPK activation.J Clin Invest.2020;130:1156-67.https://doi.org/10.1172/JCI124635.
51. LangendonkM,deJongMRW,SmitN,etal.Identification of the estrogen receptor beta as a possible new tamoxifen-sensitive target in diffuse large B-cell lymphoma.Blood Cancer J.2022;12:36.https://doi.org/10.1186/s41408-022-00631-7.

## Figures

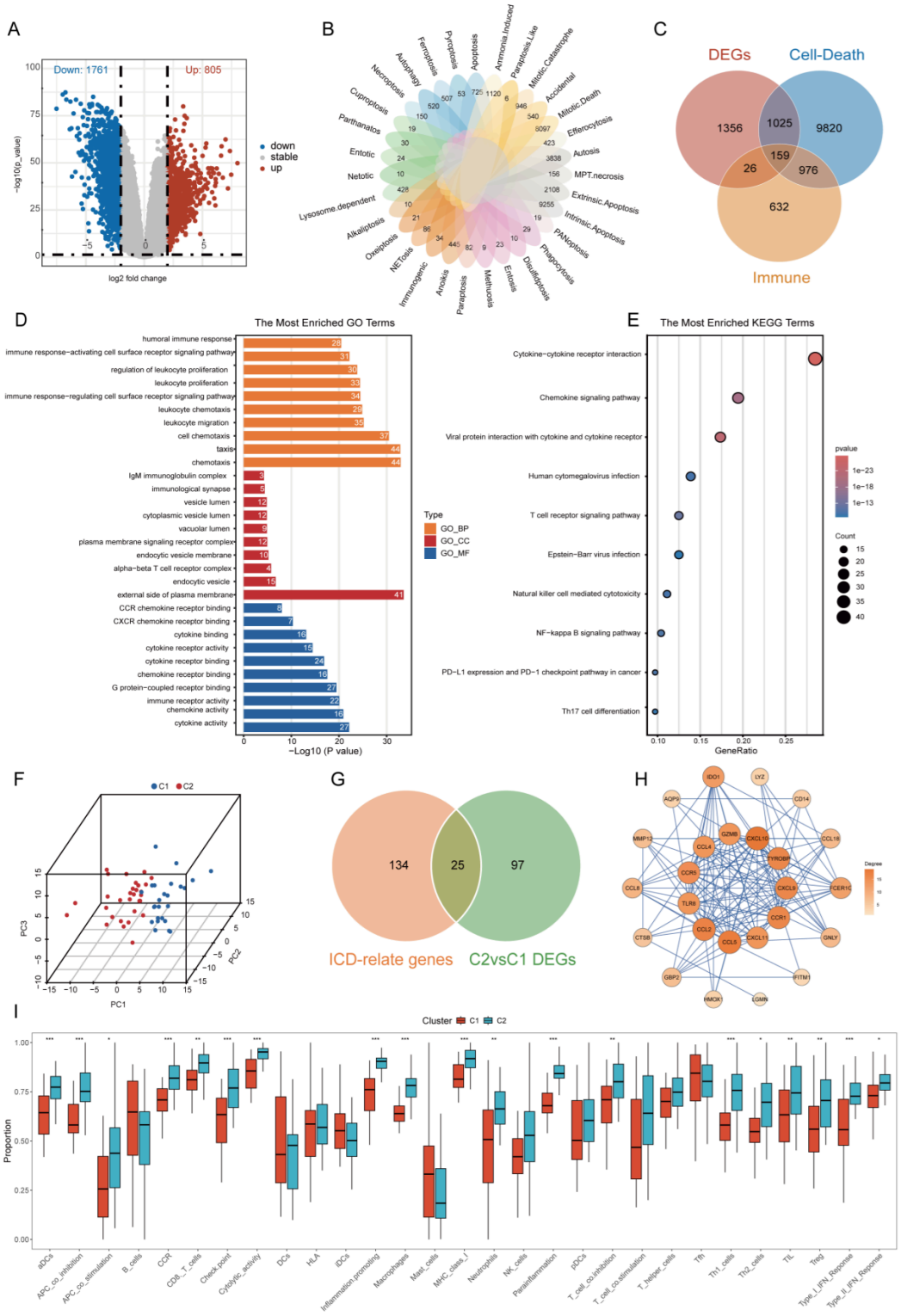


**Figure 1**

**Screening and functional analysis of key immune– and cell death–related genes.**

(A) The volcano plot of DEGs in the GSE56315 cohort; (B) petal plot of the number of related genes corresponding to 31 cell death modes; (C) Venn diagram of the intersection of cell-death-related genes, immune-related genes and DEGs; (D) bar plot of GO pathway enrichment analysis for cell-death- and immune-related DEGs; (E) bubble plot of KEGG pathway enrichment analysis for cell-death- and immune-related DEGs; (F) Principal Component Analysis (PCA) cluster plots of C1 and C2 subtypes; (G) Venn diagram of the overlap between Immune-related cell death genes and C2_vs_C1 differentially expressed genes; (H) PPI network analysis results of 25 key genes; (I) boxplot comparison of immune cell infiltration abundance between C1 and C2 subtypes. ∗ indicates P < 0.05, ∗∗ indicates P < 0.01, ∗∗∗ indicates P < 0.001

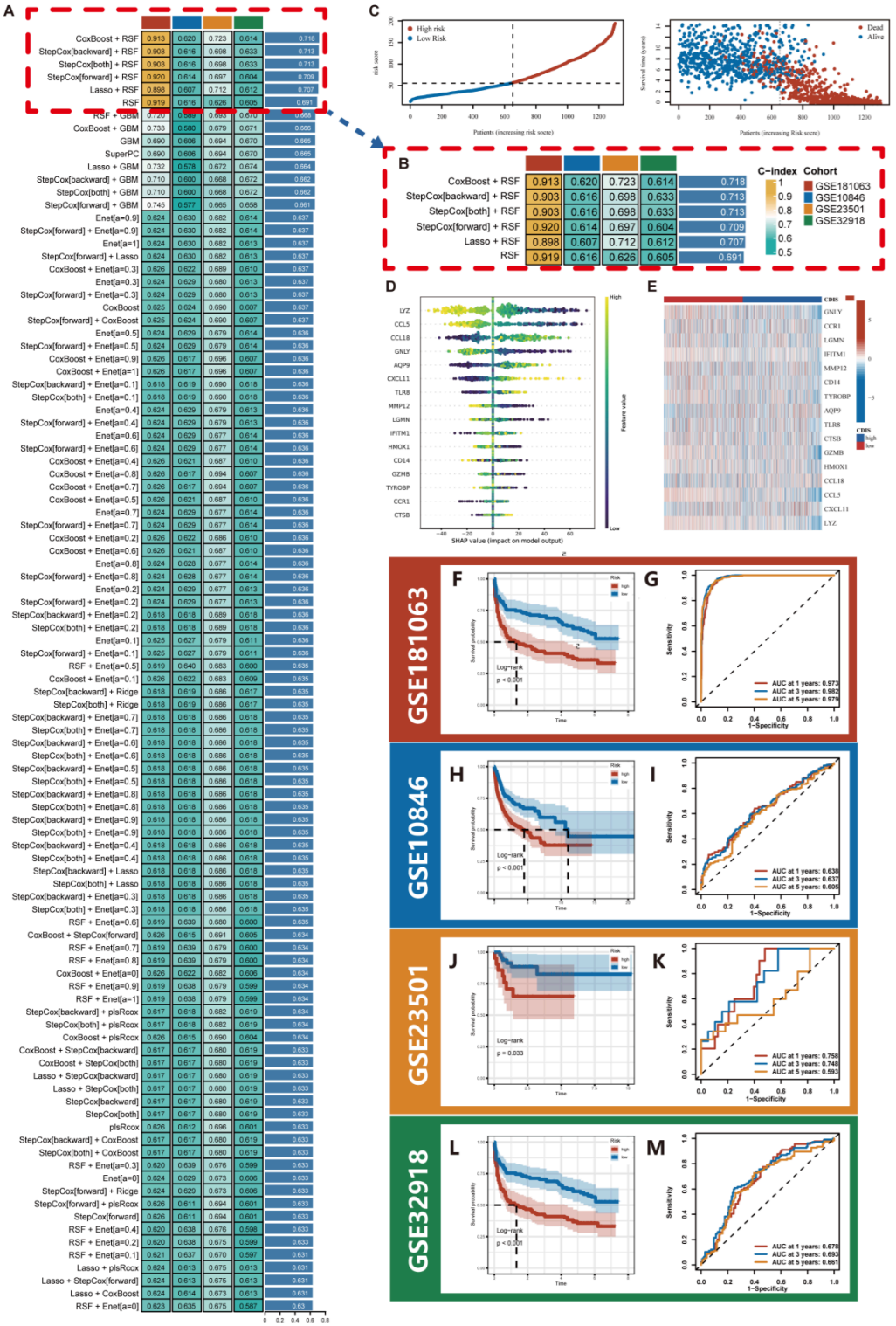


**Figure 2**

**Construction and validation of the cell-death–immune score (CDIS) model.**

(A) Heatmap of the average C-index of 101 machine-learning algorithm combinations based on 25 key genes in the training and validation cohorts; (B) heatmap of the top 6 machine-learning models ranked by average C-index; (C) scatter plot of CDIS distribution and sample distribution in high- and low-score groups in the GSE181063 cohort; (D) SHAP summary plot showing the distribution of SHAP values for the 16 genes ranked by mean absolute importance; (E) heatmap of differential expression of top 16 model genes in CDIS between high and low risk groups in GSE181063; (F) Kaplan-Meier curves of patients in CDIS high- and low-risk groups in the GSE181063 cohort; (G) Receiver operating characteristic (ROC) curves at 1, 3, and 5 years in the GSE181063 cohort; (H) Kaplan-Meier curves of patients in CDIS high- and low-risk groups in the GSE10846 cohort; (I) ROC curves at 1, 3, and 5 years in the GSE10846 cohort; (J) Kaplan-Meier curves of patients in CDIS high- and low-risk groups in the GSE23501 cohort; (K) ROC curves at 1, 3, and 5 years in the GSE23501 cohort; (L)

Kaplan-Meier curves of patients in CDIS high- and low-risk groups in the GSE32918 cohort; (M) Receiver operating characteristic (ROC) curves at 1, 3, and 5 years in the GSE32918 cohort.

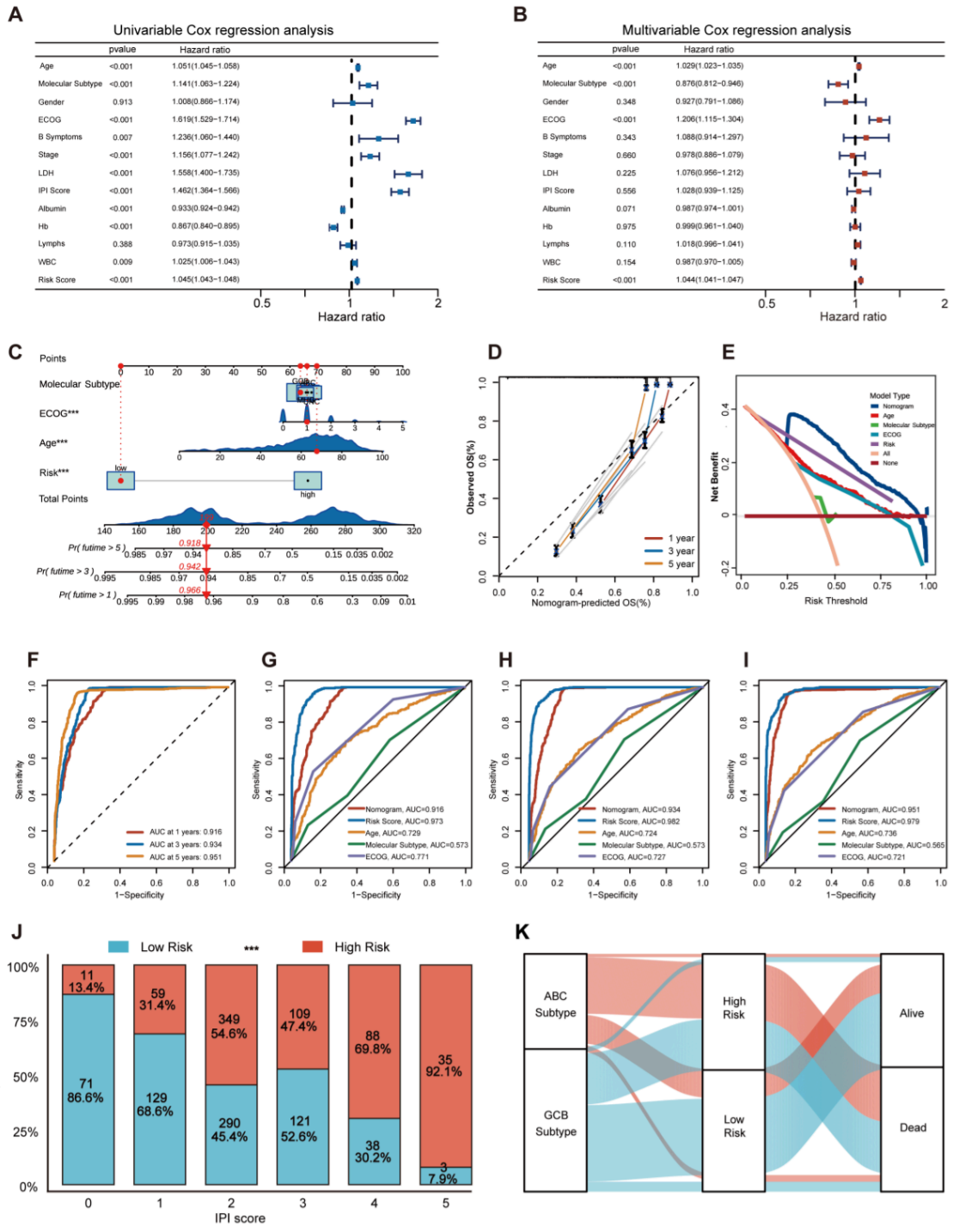


**Figure 3**

**Construction and validation of a nomogram integrating CDIS and clinical features.**

(A) Forest plot of univariable Cox regression analyses of CDIS score and clinical factors in the prognosis of patients; (B) forest plot of multivariable Cox regression analyses of CDIS score and clinical factors in the prognosis of patients; (C) Nomogram developed to predict overall survival probability based on clinical characteristics and CDIS; (D) calibration curves of the Nomogram for predicting 1-, 3-, and 5-year overall survival (OS); (E) decision curve analysis (DCA) of evaluating and comparing the clinical utility of the Nomogram in patients across different time points; (F) receiver operating characteristic curve

(ROC) curves of the Nomogram model for 1-, 3-, and 5-year survival prediction; (G) ROC curves of the Nomogram model and its individual factors for 1-year survival prediction; (H) ROC curves of the Nomogram model and its individual factors for 3-year survival prediction; (I) ROC curves of the Nomogram model and its individual factors for 5-year survival prediction; (J) stacked bar plot of the proportion of low- and high-risk patients across different IPI scores; (K) alluvial plot of the relationship among molecular subtype, risk group, and survival status.

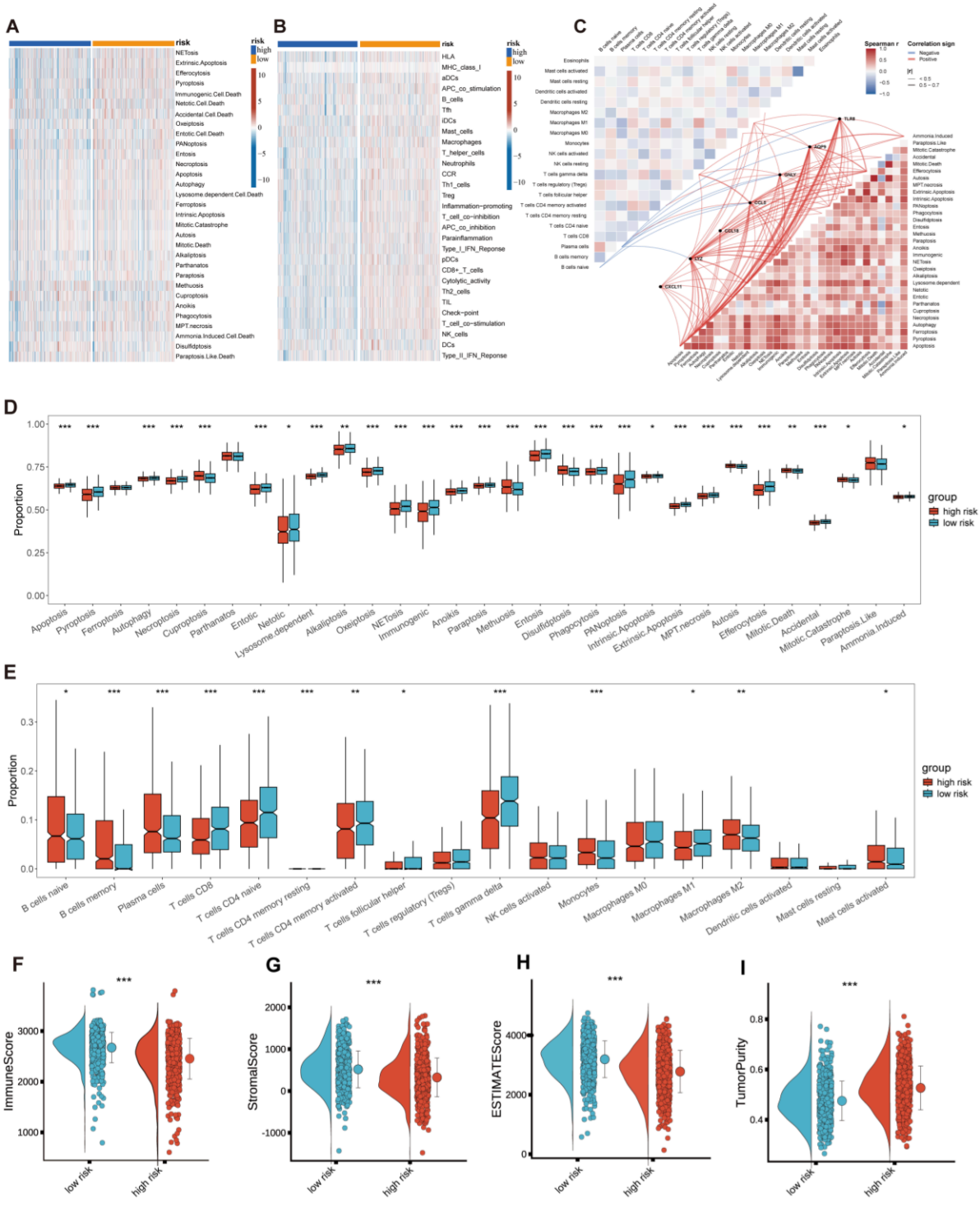


**Figure 4**

**The CDIS high- and low-risk groups exhibited significant differences in cell death patterns and immune microenvironment characteristics.**

(A) Heatmap of 31 cell death score differences between CDIS high- and low-risk groups; (B) heatmap of immune cell infiltration scores in the CDIS high- and low-risk groups; (C) correlation analysis of the top 7 model genes, immune cell infiltration, and cell death patterns; (D) box plot of 31 cell death scores

between CDIS high- and low-risk groups; (E) box plot of CIBERSORT-based analysis of immune infiltration between the CDIS high- and low-risk groups; (F-I) violin plot of immune score, stromal score, ESTIMATE score, and tumor purity differences between CDIS high- and low-risk groups. ∗ indicates P < 0.05, ∗∗ indicates P < 0.01, ∗∗∗ indicates P < 0.001

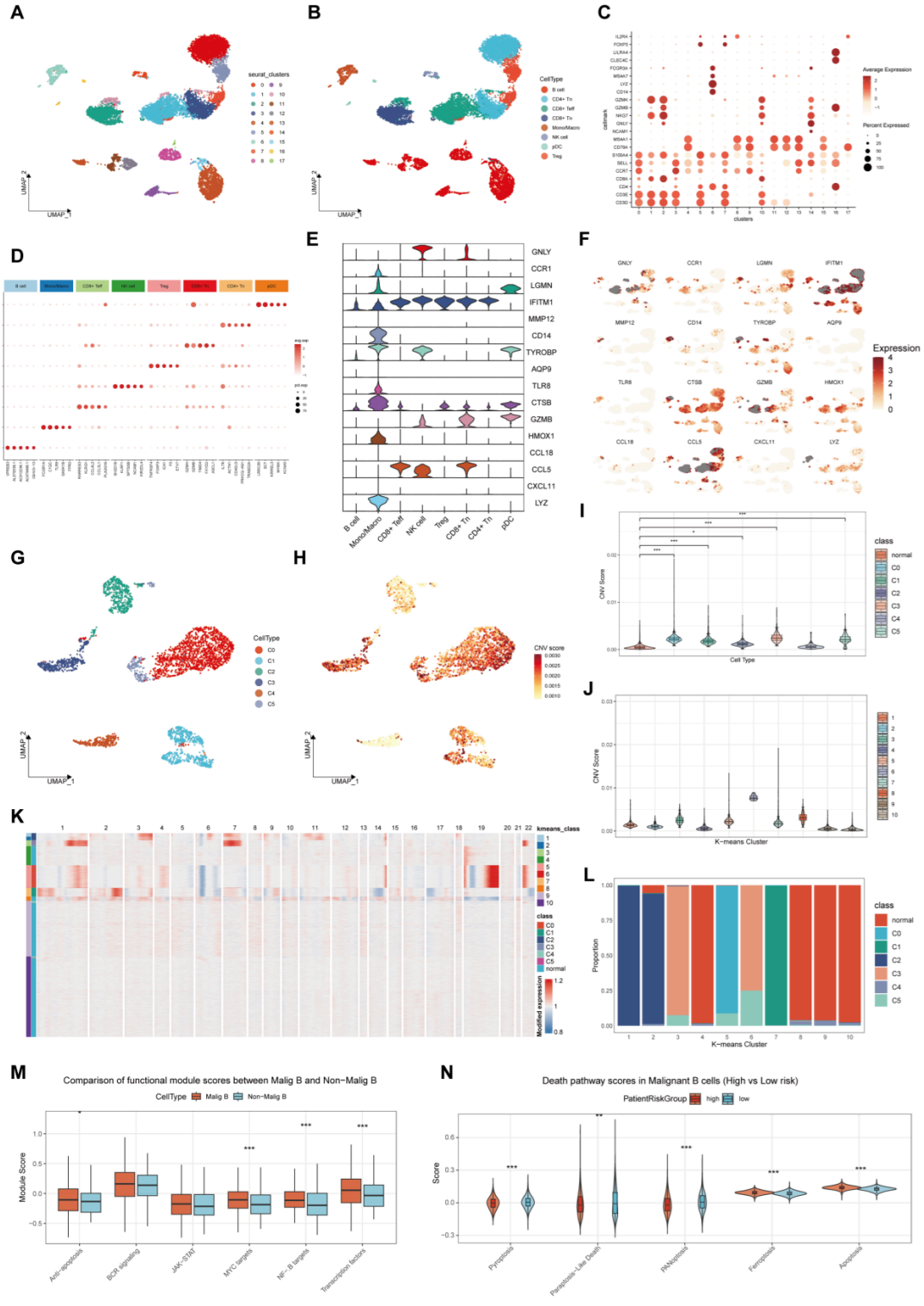


**Figure 5**

**Cell clustering annotation of single-cell data and identification of tumor versus normal cells.**

(A-B) UMAP visualization of 18 identified cell clusters; (C) bubble plot of expression levels of marker genes for each cell cluster; (D) bubble plot of the top 5 differentially expressed genes for each cell cluster compared with other clusters; (E) violin plot of the expression of the 16 key genes in the CDIS model across cell subclusters; (F) UMAP visualization of the expression patterns of the 16 key genes in the CDIS model across different cell subclusters; (G-H)

UMAP visualization of cell subclusters and CNV scores in the inferCNV analysis; (I) violin plot of CNV scores across different cell subclusters in the inferCNV analysis; (J) violin plot of inferCNV analysis scores across unsupervised clusters; (K) cell proportion plot from inferCNV analysis; (L) heatmap of chromosomal copy number variations identified by inferCNV analysis; (M) box plot of functional module scores between malignant and non-malignant B cells; (N) violin plot of cell death pathway scores between high-risk and low-risk malignant B cells. ∗ indicates P < 0.05, ∗∗indicates P < 0.01, ∗∗∗ indicates P < 0.001

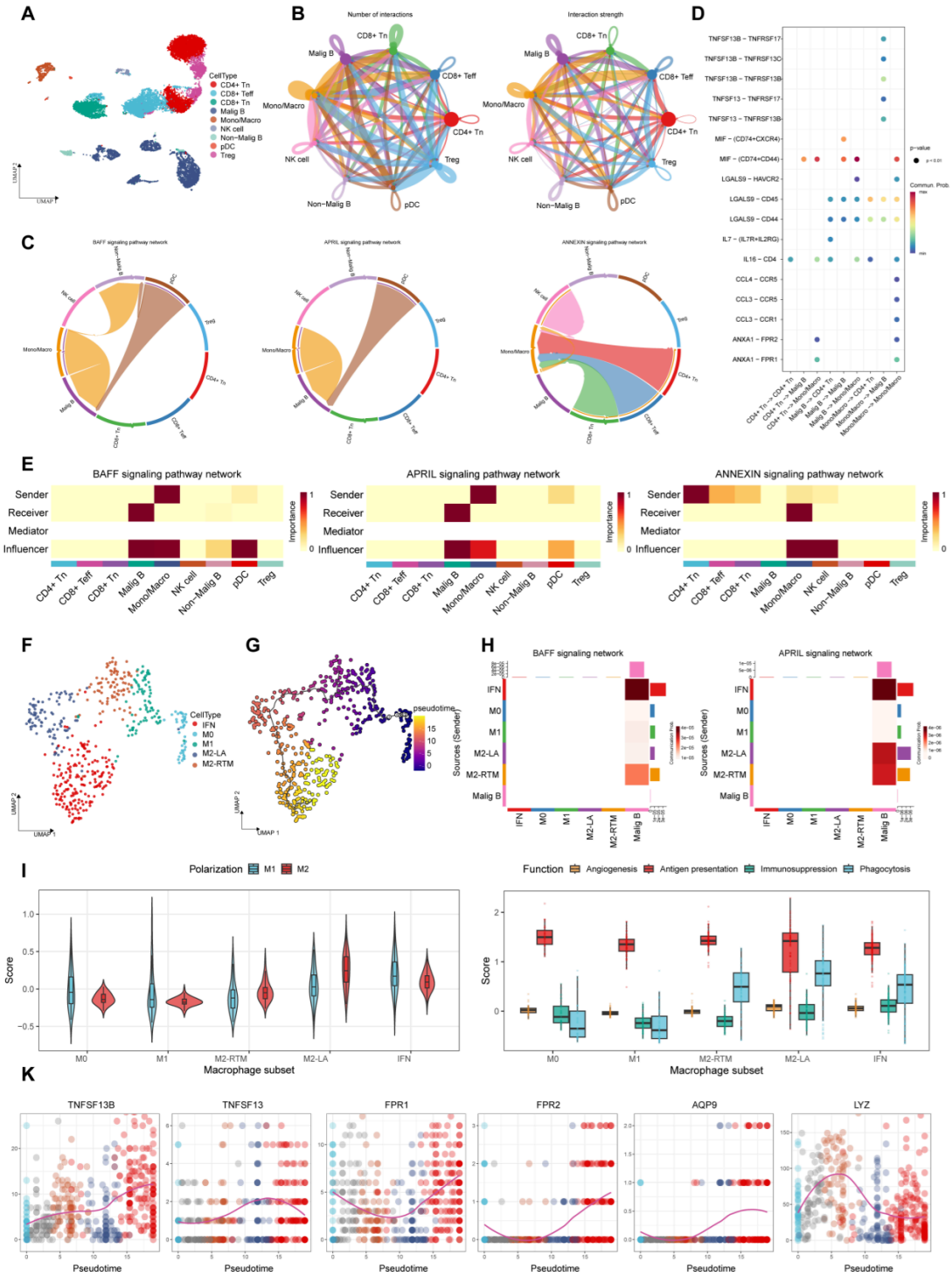


**Figure 6**

**Single-cell analysis identifies M2-like macrophages and BAFF/APRIL signaling in DLBCL.**

(A) UMAP plot of the distribution of reannotated cell types; (B) number and strength of interactions between 9 cell subtypes; (C) chord diagrams of BAFF, APRIL, and ANXA1–FPR signaling network communication; (D) ligand-receptor interaction pairs for communication among CD4 Tn cells, malignant B cells, and macrophages; (E) heatmap of functional role analysis of cell types in BAFF, APRIL, and ANXA1–FPR signaling pathway networks; (F-G) Annotation of macrophage subclusters and pseudotime trajectory plots depicting their developmental progression along the pseudotime axis; (H) BAFF and APRIL signaling networks between macrophage subsets and malignant B cells; (I) polarization status and functional characterization of macrophage subsets; (J) Dynamic expression patterns of key ligand-receptor pairs and model genes along the macrophage differentiation trajectory.

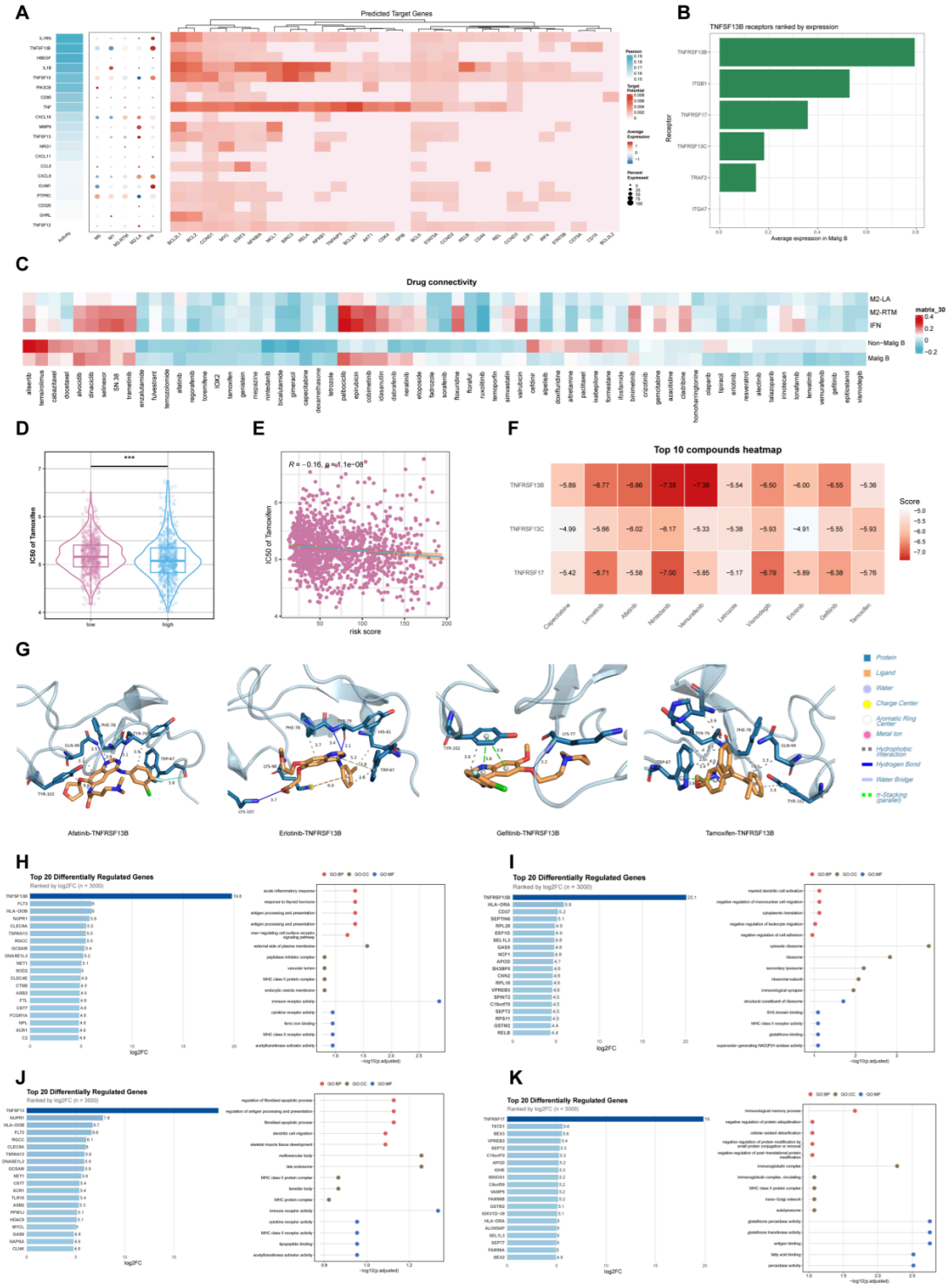


**Figure 7**

**Virtual screening and molecular docking identify potential therapeutic agents.**

(A) Left bar plot: Ligands are ranked from high to low by activity score; a higher score indicates stronger potential of M2/IFN macrophage subsets to send signals to malignant B cells. Middle dot plot: For the same ligands, dot size represents the proportion of expressing cells within each subset, and color intensity indicates the average expression level. Right matrix: Shows the regulatory strength between these ligands and downstream target genes; darker color represents stronger evidence of regulation. Together, the three panels reveal how macrophage subsets may regulate malignant B cells through specific ligand-target gene axes; (B) ranked expression of TNFSF13B receptors in malignant B cells; (C) Heatmap showing single-cell drug reverse connectivity (linear model, Benjamini-Hochberg corrected *p* < 0.05) of FDA-approved oncology compounds across macrophage and B-cell states. Color indicates predicted cell-state sensitivity: red for relative resistance, blue for relative sensitivity; (D) violin plots of Tamoxifen IC50 across high- and low-risk subgroups; (E) correlation analysis between Tamoxifen IC50 and CDIS scores; (F) heatmap of the binding affinities of TNFRSF13B, TNFRSF13C, and TNFRSF17 with the top 10 sensitive drugs; (G) molecular docking diagrams of the interactions between TNFRSF13B and afatinib, erlotinib, gefitinib, and tamoxifen; (H) bar plot of the top 20 differentially regulated genes following virtual knockout of TNFSF13B and their enrichment results; (I) bar plot of the top 20 differentially regulated genes following virtual knockout of TNFRSF13B and their enrichment results; (J) bar plot of the top 20 differentially regulated genes following virtual knockout of TNFSF13 and their enrichment results; (K) bar plot of the top 20 differentially regulated genes following virtual knockout of TNFRSF17 and their enrichment results.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigureS1.tif
- SupplementaryFigureS2.tif
- SupplementaryFigureS3.tif
- SupplementaryTableS131celldeathgenesets.xlsx
- SupplementaryTableS2functionalscoringgenesets.xlsx
- SupplementaryTableS3CDISmodelgenescoefficients.xlsx