Am J Cancer Res 2024;14(4):1768-1783
www.ajcr.us /ISSN:2156-6976/ajcr0154768

Original Article

Roles of RNA m^{5}C modification patterns in prognosis and tumor microenvironment infiltration of diffuse large B-cell lymphoma

Wenting Cui[1,2*], Cancan Luo[1*], Lili Zhou[1*], Tiantian Yu[1], Yongsheng Meng[3], Qianqian Yu[3], Zhixiang Lei[1], Ya Wang[1], Lijuan Peng[1], Qingqing Luo[1], Duozhuang Tang[1], Ruifang Sun[3#], Li Yu[1#]

[1]Department of Hematology, The Second Affiliated Hospital, Jiangxi Medical College, Nanchang University, Key Laboratory of Hematology of Jiangxi Province, Nanchang, Jiangxi, China; [2]Department of Hematology, The First People's Hospital of Jiujiang, Jiujiang, Jiangxi, China; [3]Department of Tumor Biobank, Shanxi Province Cancer Hospital, Shanxi Hospital Affiliated to Cancer Hospital, Chinese Academy of Medical Sciences, Cancer Hospital Affiliated to Shanxi Medical University, Taiyuan, Shanxi, China. [*]Equal contributors and co-first authors. [#]Co-corresponding authors.

Received December 4, 2023; Accepted March 10, 2024; Epub April 15, 2024; Published April 30, 2024

Abstract: Genetic and epigenetic aberrations display an essential role in the initiation and progression of diffuse large B-cell lymphoma (DLBCL). 5-methylcytosine (m^{5}C), a common RNA modification, regulates various cellular processes and contributes to tumorigenesis and cancer progression. However, m^{5}C alterations in DLBCL remain unclear. Our research constructed an m^{5}C prognostic model utilizing GEO data sets, which can efficiently predict the prognosis of patients with DLBCL, and verified the m^{5}C prognostic model genes by immunohistochemistry analysis. This model was constructed using unsupervised consensus clustering analyses, Least Absolute Shrinkage and Selection Operator (LASSO), and multivariate Cox regression analyses. Based on the expression of m^{5}C genes in the model, patients with DLBCL could be effectively divided into groups with significant survival time differences. The m^{5}C risk-score signature demonstrated a highly significant independent prognostic value. Results from tumor microenvironment analyses revealed that m^{5}C genes altered the infiltration of eosinophils, Tregs, and M2 macrophages. Additionally, they regulated T cell activation by modulating the expression of CTLA4, PDL1, B2M, CD8A, ICOS, and other relevant immune checkpoint expressions. In conclusion, our study presents a robust m^{5}C prognostic model that effectively predicts prognosis in DLBCL. This model may offer a new approach for prognostic stratification and potential therapeutic interventions for patients with DLBCL.

Keywords: 5-methylcytosine, diffuse large B-cell lymphoma, prognostic model, tumor microenvironment, immune checkpoint

Introduction

Diffuse large B-cell lymphoma (DLBCL), the most prevalent lymphoid cancer exhibits heterogeneity in clinical presentation, morphology, and biology [1, 2]. Although standard-of-care chemoimmunotherapy achieves durable remissions in over 60% of patients, DLBCL continues to pose a difficult clinical challenge, with approximately one-third of patients experiencing poor outcomes when faced with relapsed or refractory disease [3, 4]. Recently, the utilization of epigenetic medications, such as compounds that inhibit histone deacetylase, enhancer of zeste homolog 2 (EZH2), and bromodomain, has played a vital role in enhancing treatment outcomes for individuals with DLBCL. However, the efficacy of these drugs and their specific groups remains limited [5]. Additional investigations into the biological roles of epigenetic changes in DLBCL could aid in the identification of possible targets for therapy and the implementation of personalized treatment for patients with DLBCL. Presently, contemporary genome-wide molecular examination of DLBCL has revealed a variety of modified cellular pathways that have crucial functions in the growth and sustenance of tumors, as well as in response to treatment.

RNA methylation, a crucial biological epigenetic mechanism, plays a significant role in governing

https://doi.org/10.62347/NXDR1826

transcriptional activation and inactivation [6, 7]. RNA stability and translation efficiency are influenced by a widespread epigenetic modification known as 5-methylcytosine (m^{5}C) [8]. Methyltransferases (known as 'Writers'), including the members of the NSUN and DNMT family, catalyze the formation of m^{5}C methylation. This process can be dynamically controlled by demethylases ('Erasers') and binding proteins ('Readers'), such as TET families or ALYREF [9]. According to previous studies, these regulators are responsible for m^{5}C alterations that function in cell differentiation and apoptosis [10]. Several cancers show alterations in the levels of m^{5}C methylation regulatory genes, which are linked to their pathogenesis and prognosis [11-14]. For example, regulation of target gene methylation by ALYREF promotes hepatocellular carcinoma progression [15]. Alterations in the m^{5}C genes can also adversely affect immune cells, resulting in tumor microenvironment (TME) transformation. For instance, the TET family could potentially affect the homeostasis of B cells and increase their susceptibility to B-cell cancers [16]. Additionally, TET proteins may affect the behavior of different types of immune cells like Tregs [17-19]. Nevertheless, the possible roles and mechanisms of m^{5}C methylation regulatory genes in cancer development, particularly in DLBCL, remain uncertain. Furthermore, the complete comprehension of the association between m^{5}C methylation, TME, and immunotherapy remains elusive. Hence, an in-depth examination of the TME characteristics facilitated by m^{5}C genes will enhance our understanding of the TME and offer valuable perspectives for immunotherapy in DLBCL.

To elucidate the role of m^{5}C genes in DLBCL, we conducted comprehensive evaluation of their expression profiles using the Gene Expression Omnibus database. Subsequently, we developed a predictive model and performed a functional analysis to explore the relationship between m^{5}C regulatory gene clusters and immunity. Our study established a robust m^{5}C prognostic model, serving as a reliable predictor of prognosis in DLBCL. This model may provide a new approach to prognostic stratification and potential therapeutics for DLBCL patients.

Materials and methods

Sources of data and the process of processing

Gene expression data and the clinical details of patients with DLBCL were obtained from the Gene Expression Omnibus (GEO) database. This study utilized two GEO datasets: the bulk sequence GSE10846 (with 414 tumor samples) [20] and GSE11318 (with 203 tumor samples) [21]. During model construction, GSE10846 was used as the training set, whereas GSE11318 served as the verification set. From the literature [22], a collection of 17 regulators of m^{5}C RNA methylation was acquired. The m^{5}C regulators' expression was extracted and organized using the 'limma' package in the training set. Patient clinical information is summarized in Supplementary Table 1.

Unsupervised clustering of m^{5}C regulator genes

Normalized RNA sequencing data was utilized to perform unsupervised consensus clustering analysis using the 'Consensusclusterplus' package. Principal Component Analysis (PCA) was used to confirm the classification results.

Construction of the prognostic risk model of m^{5}C regulator genes

Prognosis-related m^{5}C regulatory genes were screened using univariate Cox regression analysis based on the overall survival (OS) data obtained from the clinical information of the training set. To identify candidate genes for a predictive model related to m^{5}C, we performed LASSO regression analysis using the R package 'glmnet' (version 4.0.2). The risk scores for the genes involved in the m^{5}C predictive model were calculated as follows: Riskscore = $\sum i = 1n\text{Coef}(i) \times x(i)$.

Validation of the m^{5}C-regulated gene prognostic model

Survival analysis of the training set samples was performed using the 'survminer' package. The model efficient was evaluated by receiver-operating characteristic (ROC) curve. Subsequently, Cox analyses were conducted using univariate and multivariate approaches based on clinical characteristics. Similarly, the m^{5}C-

regulated gene prognostic model was used to calculate the risk scores for the verification sets. The validation set underwent survival analysis and ROC curve analysis to assess the model's reliability.

Constructing a nomogram and predicting clinical characteristics with the prognostic model of m^{5}C-regulated genes

The clinical significance of the risk model was evaluated through univariate and multivariate Cox regression analyses, examining the association between age, gender, pathological subtype, clinical stage, Eastern Cooperative Oncology Group (ECOG) status, Lactate dehydrogenase (LDH) ratio, extranodal sites, and risk score value, and OS in patients with DLBCL. A predictive nomogram was developed based on different clinical characteristics to determine the prognosis of high-risk and low-risk groups.

Relationship between the m^{5}C prognostic model and the TME

'ESTIMATE' [23] be used to compute the TME scores, which encompassed ESTIMATE score, immune score, stroma score, and tumor purity. Additionally, an analysis was conducted on the relationship between TME scores and survival time, as well as between the high- and low-risk subgroups.

Correlation between the m^{5}C prognostic model and the infiltration of immune cells

In the training set, 22 different types of immune cell infiltrations were scored using 'CIBERSORT' [24] and be used to calculate the proportion of immune cells in different group.

Identification of m^{5}C regulated gene expression in DLBCL tissues

Tissue microarray (TMA) was performed in 144 DLBCL tissues, which were provided by the Shanxi Cancer hospital between January 2015 and December 2023. Experiments were reviewed and approved by the Ethics Committee of the Second Affiliated Hospital of Nanchang University and the Shanxi Province Cancer Hospital, and were conducted in compliance with the Helsinki Declaration. TMA sections were immunestained with different primary antibodies: ALKBH1 (Rabbit monoclonal, ab126596, abcam, England); DNMT1 (Rabbit monoclonal, ab188453, abcam, England); DNMT3A (Rabbit polyclonal, ab188470, abcam, England); DNMT3B (Rabbit polyclonal, ab2851, abcam, England); NOP2 (Rabbit monoclonal, ab271075, abcam, England); NSUN3 (Rabbit monoclonal, ab272616, abcam, England); NSUN4 (Rabbit polyclonal, ab235430, abcam, England).

Statistical analysis

R package 'limma' (version 4.0.2) [25] was used to conduct all statistical analyses. Using the Kaplan-Meier technique, we evaluated the association between the high- and low-risk groups and OS. Univariate and multivariate Cox regression analysis was used to determine whether the risk score, age, stage, ECOG, and LDH could serve as autonomous prognostic factors. Statistical significance was set at $P < 0.05$.

Results

Evaluation of m^{5}C-related modification patterns based on 17 regulators

Patients with DLBCL were classified using unsupervised clustering based on the expression of 17 m^{5}C genes. Finally, two patterns were identified, with clusters 1 and 2 having 195 and 217 patients, respectively (**Figure 1A**). **Figure 1B** shows the differentiation of the two m^{5}C modification patterns according to PCA analysis. This implies that patients with DLBCL can be accurately distinguished using m^{5}C-related genes. Survival analysis revealed that patients in cluster 2 exhibited a longer survival time than those in cluster 1 (**Figure 1C**). Examination of immune cells in clusters 1 and 2 confirmed that the proportion of highly responsive immune cells was notably greater in cluster 1. In contrast, the proportion of memory T cells and antigen-presenting cells (APC) in cluster 2 exceeded that in cluster 1 (**Figure 1D**). Subsequently, we attempted to elucidate the connections among 17 m^{5}C regulators. Spearman correlation analysis indicated that NSUN5 and NOP2, which are m^{5}C writers, could potentially serve as central genes among the m^{5}C-related genes and interact with the m^{5}C erasers TET2 and TET3 (**Figure 1E**). To examine the correlation between m^{5}C-asso-

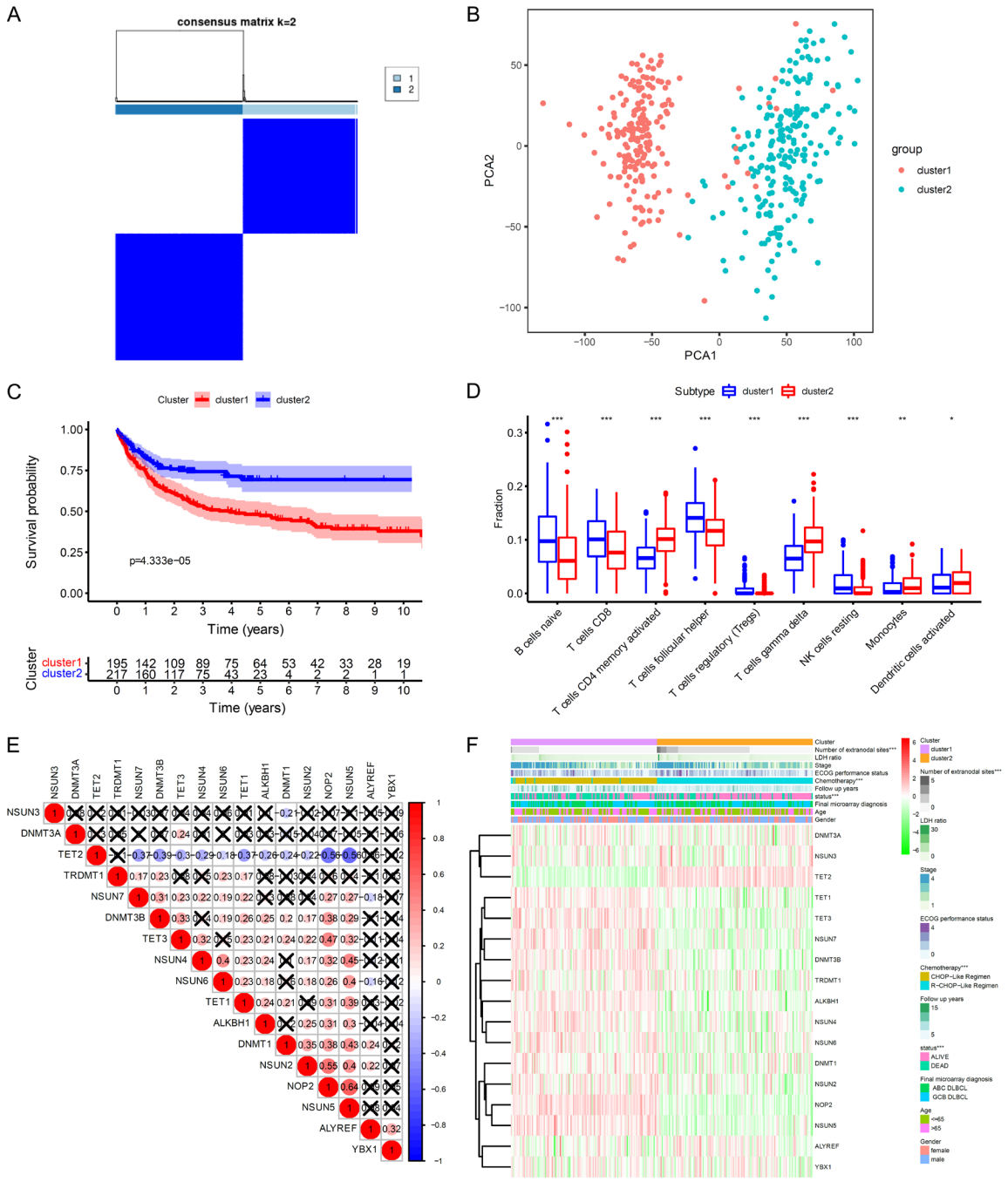


Figure 1. The clinical values in different m^{5}C modification patterns in DLBCL patients based on consensus clustering. A. The DLBCL patients were divided into 2 clusters based on the consensus clustering matrix (k = 2). B. PCA analysis on the two clusters. C. Kaplan-Meier survival analysis for the two clusters based on the training set. D. Differences in proportion of immune cells between the two clusters. E. Spearman correlation analysis of m^{5}C regulated gene. F. Heatmaps of different clinical pathological features and m^{5}C regulatory gene expression between two clusters. PCA, Principal Component Analysis; ECOG, Eastern Cooperative Oncology Group; ABC, activated B-cell-like; GCB, Germinal Center B-cell; LDH, Lactate Dehydrogenase; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone.

ciated genes and DLBCL, we assessed the expression of m^{5}C-related genes and their clinicopathological features in clusters 1 and 2 (**Figure 1F**). Data from Cluster 1 showed that most m^{5}C writers, including NSUN5, NOP2, and NSUN4, are abnormally expressed in DLBCL. In

cluster 2, the m^{5}C eraser TET2 was significantly overexpressed compared to cluster 1.

Screening and modelling of prognosis-related m^{5}C genes

Combining overall survival time in the training set (**Figure 2A**), we identified 11 m^{5}C regulators linked to DLBCL prognosis using univariate Cox regression analysis. These candidate genes were subjected to the LASSO Cox regression algorithm. Eventually, seven genes were used to construct the m^{5}C model, as shown in **Figure 2B**. The m^{5}C model genes included NOP2, DNMT1, NSUN4, ALKBH1, DNMT3B, NSUN3, and DNMT3A. Patients in the training set were categorized into high-risk and low-risk groups by using model. Analysis of survival characteristics indicated that individuals classified as high-risk experienced a considerably reduced survival duration and a greater mortality rate than those categorized as low-risk (**Figure 2C**, **2D**). The validation of the m^{5}C prognostic model's applicability in the test set is illustrated in **Figure 2E** and **2F**. The high-risk group exhibited a significantly lower OS than the low-risk group ($P < 0.05$). ROC analyses were conducted on both the training and testing sets to evaluate the specificity and sensitivity of the m^{5}C-score prognostic model. ROC curve analysis revealed that the m^{5}C modification signature had area under the curve (AUC) values of 0.678, 0.692, and 0.682 for 1-, 3-, and 5-year survival in the training set, respectively (**Figure 3A**, **3B**). In the validation set, the corresponding values were 0.629, 0.638, and 0.658, respectively (**Figure 3C**, **3D**). These results consistently verified that the m^{5}C-score prognostic model had good prediction accuracy. Since NOP2 mediated m^{5}C modification can alter c-Myc levels in an EIF3A dependent manner, we analyzed the impact of NOP2 expression levels on cell proliferation, and the results were consistent with our expectations. High expression of NOP2 was associated with strong cell proliferation (Supplementary Figure 3A).

Association between clinicopathological characteristics and m^{5}C risk score

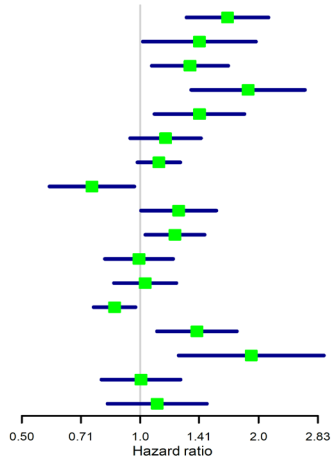
To validate the predictive significance of the m^{5}C prognostic model, both univariate and multivariate Cox regression analyses were conducted to investigate the association between clinicopathological features (sex, age, pathological subtype, ECOG status, Stage, LDH ratio, and extranodal sites) and the risk signature. Except for sex and extranodal sites, all clinical characteristics and risk scores showed a significant association with OS in both univariate (**Figure 3E**) and multivariate analyses (**Figure 3F**) ($P < 0.005$). Furthermore, we investigated the relationship between the clinical characteristics and prognosis of patients with DLBCL. The results indicated that being above 60 years old, receiving CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone)-like regimen treatment, having activated B-cell-like (ABC) type DLBCL, and having a higher ECOG level and stage were all associated with an unfavorable prognosis, whereas the prognosis of patients with DLBCL appeared to be unrelated to sex (Supplementary Figure 1A-F). Although these clinical characteristics can guide the prognosis of patients with DLBCL, they are relatively broad and limited in scope. In contrast, the m^{5}C model demonstrated higher accuracy in predicting the prognosis of patients with varying clinical characteristics, such as age (> 60 years, ≤ 60 years) (Supplementary Figure 2A, 2B), germinal center B-cell-like (GCB) type DLBCL (Supplementary Figure 2F), ECOG grade (1-2) (Supplementary Figure 2G), stage (1-2, 3-4) (Supplementary Figure 2I, 2J), and sex (male, female) (Supplementary Figure 2K, 2L). However, there were instances where this superiority was not consistently observed (Supplementary Figure 2C-E, 2H). According to the classic International Prognostic Index (IPI) scoring standard, we divided DLBCL patients into 4 groups: High, High-intermediate, Low-intermediate, and Low. Comparing the proportions of patients with different IPI scores in high-risk clusters and low-risk clusters, we found that the proportions of High, High-intermediate, Low-intermediate IPI scores in high-risk clusters were higher than those in low-risk clusters (Supplementary Figure 3B). These findings indicate that our constructed prognostic model possesses increased reliability and precision for predicting the prognosis of DLBCL patients with diverse clinical characteristics.

TME characteristics of m^{5}C prognostic model

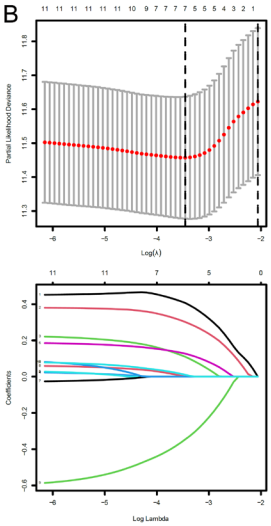
The ESTIMATE algorithm was used to investigate the heterogeneity of the TME in different m^{5}C risk groups, considering the significance of the TME in the development and progression

A

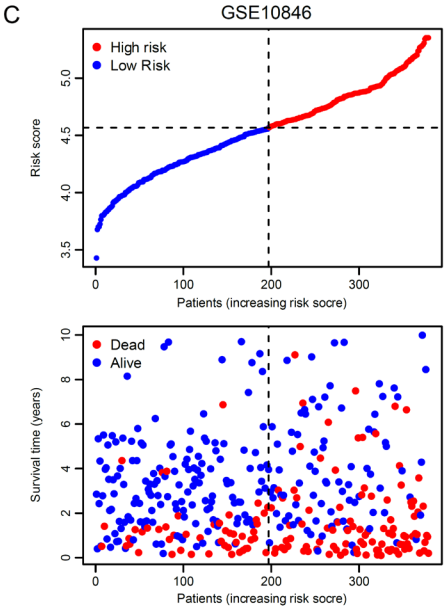
	pvalue	Hazard ratio
NOP2	<0.001	1.667(1.311−2.119)
NSUN2	0.040	1.414(1.016−1.969)
NSUN3	0.011	1.339(1.070−1.674)
NSUN4	<0.001	1.881(1.348−2.624)
NSUN5	0.010	1.415(1.087−1.841)
NSUN6	0.160	1.160(0.943−1.428)
NSUN7	0.088	1.116(0.984−1.265)
DNMT1	0.026	0.754(0.588−0.966)
DNMT3A	0.046	1.252(1.004−1.562)
DNMT3B	0.021	1.226(1.031−1.458)
TRDMT1	0.946	0.993(0.813−1.213)
TET1	0.754	1.030(0.858−1.236)
TET2	0.016	0.861(0.761−0.973)
TET3	0.005	1.394(1.103−1.761)
ALKBH1	0.003	1.916(1.252−2.930)
ALYREF	0.969	1.005(0.797−1.267)
YBX1	0.502	1.105(0.826−1.478)



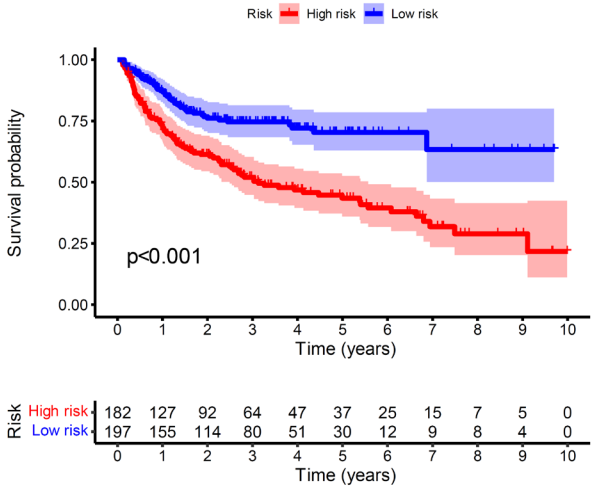
B



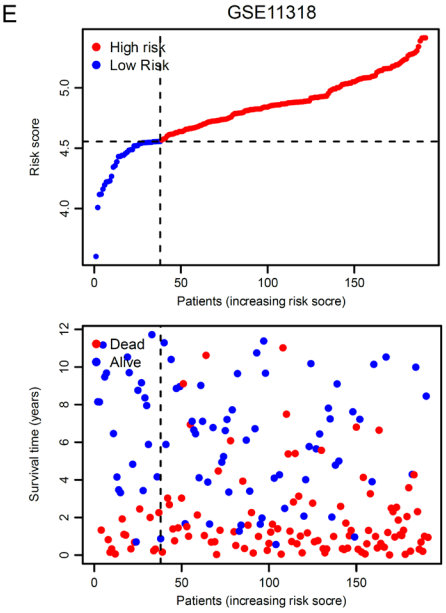
C



D



E



F

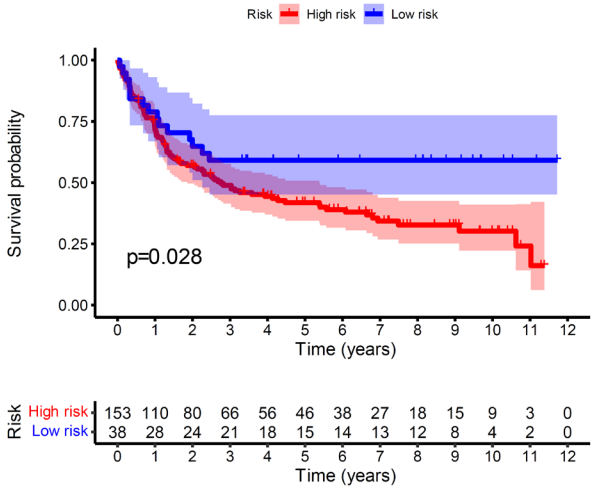


Figure 2. The construction of the m^{5}C prognostic model in DLBCL. A. The prognostic value of m^{5}C regulated genes were identified by univariable Cox regression in a forest map. B. The cross-validation and LASSO coefficient profiles were used for the parameter selection in the LASSO analysis. C, D. The association of m^{5}C risk score and prognosis of DLBCL patients in training set. E, F. The value of the m^{5}C prognostic model was verified in the validation set.

of tumors and its prognostic influence. In line with our original hypothesis, the high-risk group exhibited notably greater tumor purity but lower stroma and ESTIMATE scores than the low-risk group (**Figure 4A**). Survival analysis also indicated that elevated tumor purity scores indicated an unfavorable prognosis (**Figure 4B**), whereas higher stromal scores or ESTIMATE scores were associated with a more favorable prognosis (**Figure 4C-E**). Next, the CIBERSORT analysis was utilized to approximate the ratios of 22 unique immune cell characteristics in the low- and high-risk subtypes (**Figure 5A**). We observed that the high-risk subgroup was remarkably enriched in naive B cells, resting NK cells, CD8 T cells, follicular helper T cells, Tregs, M2 macrophages, and eosinophils, while the low-risk subgroup is notably enriched in activated memory CD4 T cells, γδT cells, M0 macrophages, monocytes, and resting mast cells. Poor prognosis was correlated with high ratios in eosinophils, Tregs, and M2 macrophages (**Figure 5B**, **5D**, **5F**), while good prognosis was associated with high ratios of M0 macrophages and γδT cells (**Figure 5C**, **5E**). This discovery suggests a potential connection between m^{5}C genes that affect the function and infiltration of immune cells, thereby ultimately influencing the prognosis of patients with DLBCL.

Correlation between immune checkpoints and m^{5}C prognostic model

Immune checkpoints are a group of proteins and signaling pathways that play a critical role in regulating the activity and balance of the immune system. Blocking immune checkpoint signaling pathways can stimulate T-cells and boost their capacity to attack tumor cells. Hence, we investigated the manifestations of primary immune checkpoints in high-risk and low-risk groups. As anticipated, immune checkpoints such as CTLA4, CD274 (PDL1), PDCD1, and LGALS9, which hinder T-cell activity, exhibited a notable increase in expression within the high-risk category. In contrast, the low-risk group exhibited the upregulation of genes related to antigen presentation and T-cell activation, including B2M, CD8A, ICOS, and IL12B (**Figure 6A**).

Additional investigations revealed that high expression of B2M, CD8A, CD40LG, ICOS, IL23A, LDHA, PTPRC, SIGLEC15, and TNFRSF9 was associated with a favorable prognosis, and these specific genes were upregulated in the low-risk category (**Figure 6B-I**). Conversely, genes such as FGL1, IL12A, LGALS9, PVR, TNFSF9, and YTHDF1 correlated with an unfavorable prognosis and demonstrated increased expression in the high-risk group (**Figure 6J-O**). These findings indicate that m^{5}C may regulate T cell activation by altering the expression of relevant immune checkpoint molecules rather than having a direct impact on the number of immune cells within the TME.

Validation of the epigenetic changes of m^{5}C genes by immunohistochemistry in patients with DLBCL

To delve deeper into the role of m^{5}C-related genes in DLBCL, the TMA from 144 DLBCL tumors and 48 normal lymph nodes was used for immunohistochemistry (IHC) analysis. The clinical features of 144 DLBCL tumors are displayed in Supplementary Table 2. We detected the protein expression levels of NOP2, DNMT1, NSUN4, ALKBH1, DNMT3B, NSUN3, and DNMT3A. The results of IHC analysis demonstrated that ALKBH1, NOP2, DNMT1 and NSUN4 were significantly increased in DLBCL, compared with normal lymph node tissues ($P < 0.001$), while the expression levels of DNMT3A and DNMT3B showed no obvious difference, and NSUN3 was not evaluated due to non-specific binding (Supplementary Figure 4A-C). The m^{5}C gene is mainly responsible for post-transcriptional modifications, DNMT3A and DNMT3B are the genes responsible for Writers. These genes are often regulated by Erasers genes. Therefore, we speculate that the inconsistency between DNMT3A and DNMT3B protein expression and mRNA expression may be caused by post-transcriptional regulation. Since these precious tissue sections had been paraffin-embedded, we unfortunately did not perform RNA-level validation.

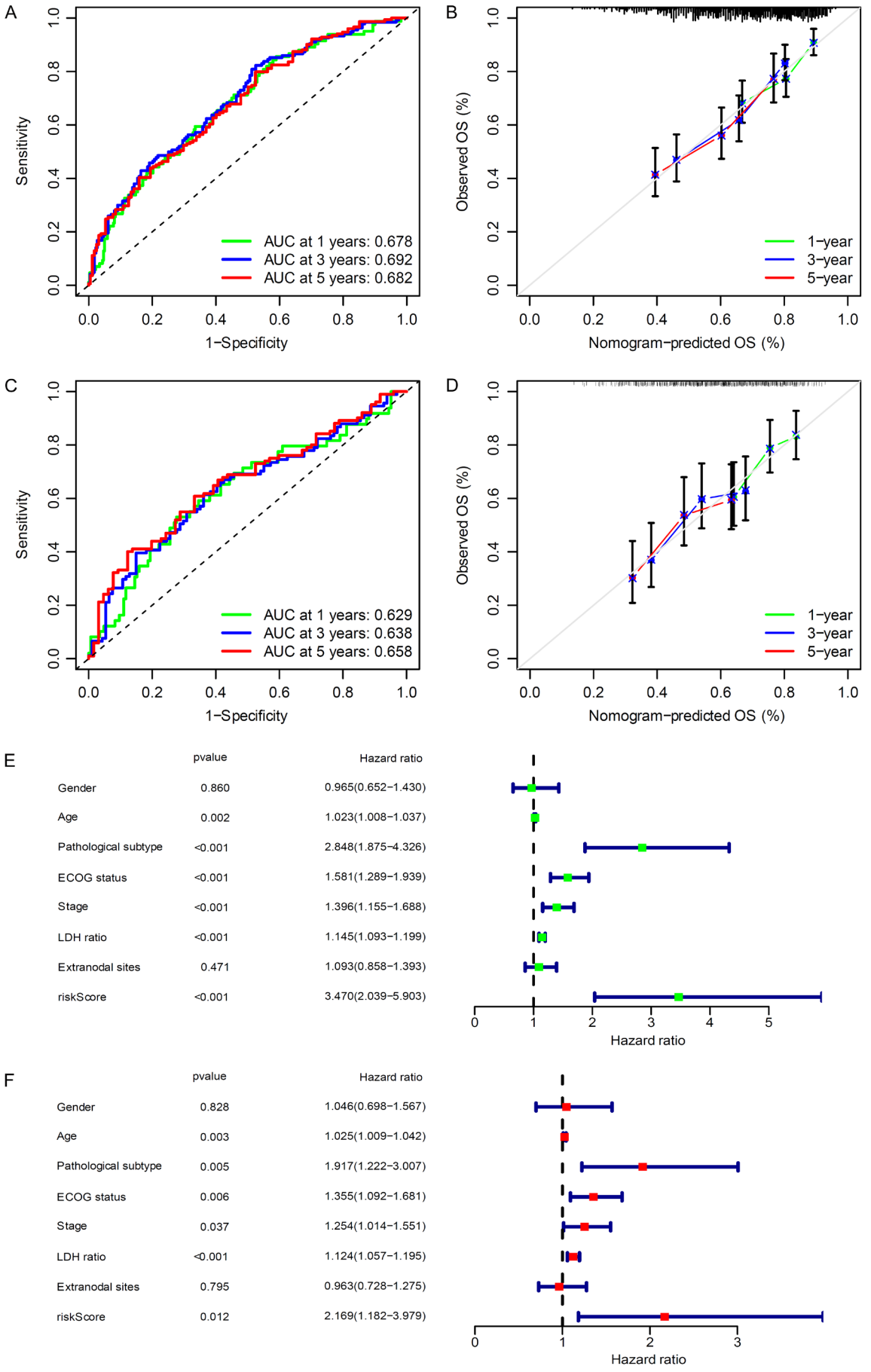

A
Sensitivity
1−Specificity
AUC at 1 years: 0.678
AUC at 3 years: 0.692
AUC at 5 years: 0.682
B
Observed OS (%)
Nomogram−predicted OS (%)
1−year
3−year
5−year
C
Sensitivity
1−Specificity
AUC at 1 years: 0.629
AUC at 3 years: 0.638
AUC at 5 years: 0.658
D
Observed OS (%)
Nomogram−predicted OS (%)
1−year
3−year
5−year
E
pvalue
Hazard ratio
Gender 0.860 0.965(0.652−1.430)
Age 0.002 1.023(1.008−1.037)
Pathological subtype <0.001 2.848(1.875−4.326)
ECOG status <0.001 1.581(1.289−1.939)
Stage <0.001 1.396(1.155−1.688)
LDH ratio <0.001 1.145(1.093−1.199)
Extranodal sites 0.471 1.093(0.858−1.393)
riskScore <0.001 3.470(2.039−5.903)
Hazard ratio
F
pvalue
Hazard ratio
Gender 0.828 1.046(0.698−1.567)
Age 0.003 1.025(1.009−1.042)
Pathological subtype 0.005 1.917(1.222−3.007)
ECOG status 0.006 1.355(1.092−1.681)
Stage 0.037 1.254(1.014−1.551)
LDH ratio <0.001 1.124(1.057−1.195)
Extranodal sites 0.795 0.963(0.728−1.275)
riskScore 0.012 2.169(1.182−3.979)
Hazard ratio

Figure 3. Accuracy and prognostic value of the model in DLBCL. (A) ROC analysis in the training group. (B) Nomogram in the training group. (C) ROC analysis in the validation group. (D) Nomogram in the validation group. Univariate regression (E) and multivariate regression (F) analysis of clinical features and the predictive model. ROC, Receiver-operating Characteristic; AUC, Area under curve; ECOG, Eastern Cooperative Oncology Group; LDH, Lactate Dehydrogenase.

Discussion

The findings of previous studies indicate that epigenetic modifications can drive tumorigenesis and pathogenesis in DLBCL [26]. However, the role of m^{5}C modifications in DLBCL has not been fully studied. Recently, the biological functions of m^{5}C modifications in various tumor types have been revealed. In this study, we aimed to determine whether m^{5}C modifications serve as independent biomarkers and prognostic factors for DLBCL.

Our study reveals the epigenetics of m^{5}C-related genes in patients with DLBCL. Based on the expression of m^{5}C, patients with DLBCL can be effectively distinguished, and these distinguished patients show significant differences in survival time. Furthermore, we constructed a prognostic model based on m^{5}C genes associated with the prognosis of patients with DLBCL, including DNMT1, DNMT3A, DNMT3B, NOP2, NSUN3, NSUN4, and ALKBH1. Simultaneously, we validated these m^{5}C prognostic model genes using 144 cases of DLBCL histopathological sections. The results revealed a significant increase in DNMT1, NOP2, NSUN4, and ALKBH1 in DLBCL tissues. DNMT1, a DNA methyltransferase, plays a crucial role in maintaining the stability of DNA methylation status, ensuring the inheritance of the methylation pattern to newly synthesized DNA strands during cell division and replication. Alterations in DNMT1 expression are frequently observed in hematological malignancies. Studies have shown that DNMT1 is associated with advanced clinical stages and drug resistance in DLBCL [27]. DNMT1 is more abundant in the GCB DLBCL subtype than in non-GCB DLBCL [28]. Most patients with GCB have a better prognosis than patients with ABC [29], and our study also suggests a correlation between DNMT1 and good prognosis. These findings imply that DNMT1 may be a potential target for improving the prognosis of patients with DLBCL.

Similarly, as m^{5}C "writers", DNMT3A and DNMT3B play a crucial role in DLBCL. An analysis of cell-free DNA identified seven mutations in DNMT3A in the serum during complete metabolic responses, which may originate from clonal haematopoiesis of indeterminate potential (CHIP) rather than residual disease. This suggests that DNMT3A mutations could be evidence of CHIP-related mutations in B-cell lymphomas [30]. In another clinical study, the overexpression of DNMT3B was significantly associated with advanced clinical staging and treatment resistance [27]. Although the protein expression levels of DNMT3A and DNMT3B in the 144 DLBCL tissues we examined showed no significant difference compared to normal lymph nodes, we speculate that this might be regulated by m^{5}C erasers genes, particularly DNMT3A and DNMT3B, which, as methyltransferases, may exert their functions without requiring significant changes in expression levels. Alternatively, DNMT3A and DNMT3B may interact or recruit with other epigenetic enzymes and transcription factors to promote different chromatin states and tumor development. Although there is a lack of research on NOP2 expression in DLBCL, its presence has been identified in hepatocellular carcinoma. In this context, NOP2-mediated m^{5}C modification can alter c-Myc levels in an EIF3A-dependent manner, thereby influencing glucose metabolism. Notably, the overexpression or mutation of c-Myc is considered a significant driver of DLBCL [31]. Furthermore, Wang et al. reported a strong association between the NSUN4 rs10252 variant and childhood leukemia risk [32]. Previous studies have confirmed that m^{5}C modification in B cells can affect antibody production, influencing the immune response to antigens [33-35].

In macrophages and dendritic cells, m^{5}C modification can impact antigen presentation and cytokine production [36, 37]. Hence, we investigated the correlation between immune cells and m^{5}C expression. M2 macrophages and Tregs were proportionally increased in the high-risk m^{5}C group and were associated with a poor prognosis. Macrophage M2 is generally considered an anti-inflammatory macrophage that plays an anti-inflammatory role in the tumor immune microenvironment and aids

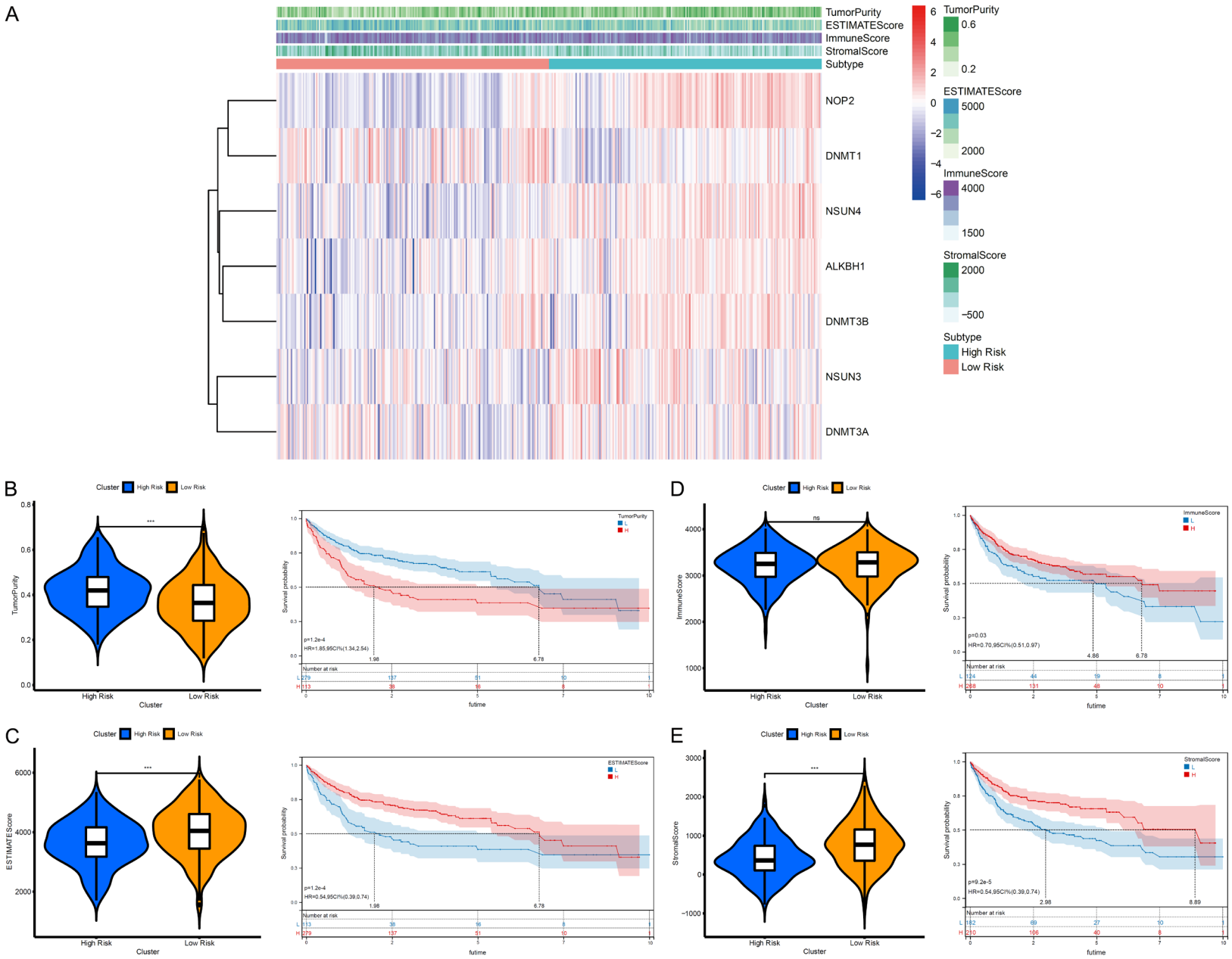

A
TumorPurity
ESTIMATEScore
ImmuneScore
StromalScore
Subtype
NOP2
DNMT1
NSUN4
ALKBH1
DNMT3B
NSUN3
DNMT3A
B
C
D
E
Cluster
High Risk
Low Risk
Survival probability
futime
Number at risk

Figure 4. Analysis of the tumor immune microenvironment using the m^{5}C prediction model. A. Heatmap of genes related to the tumor immune microenvironment predicted by the m^{5}C prediction model. B. Assessment of tumor purity and survival analysis in high and low-risk groups. C. Evaluation of ESTIMATE scores and survival analysis in high and low-risk groups. D. Immune scores and survival analysis in high and low-risk groups. E. Stromal scores and survival analysis in high and low-risk groups.

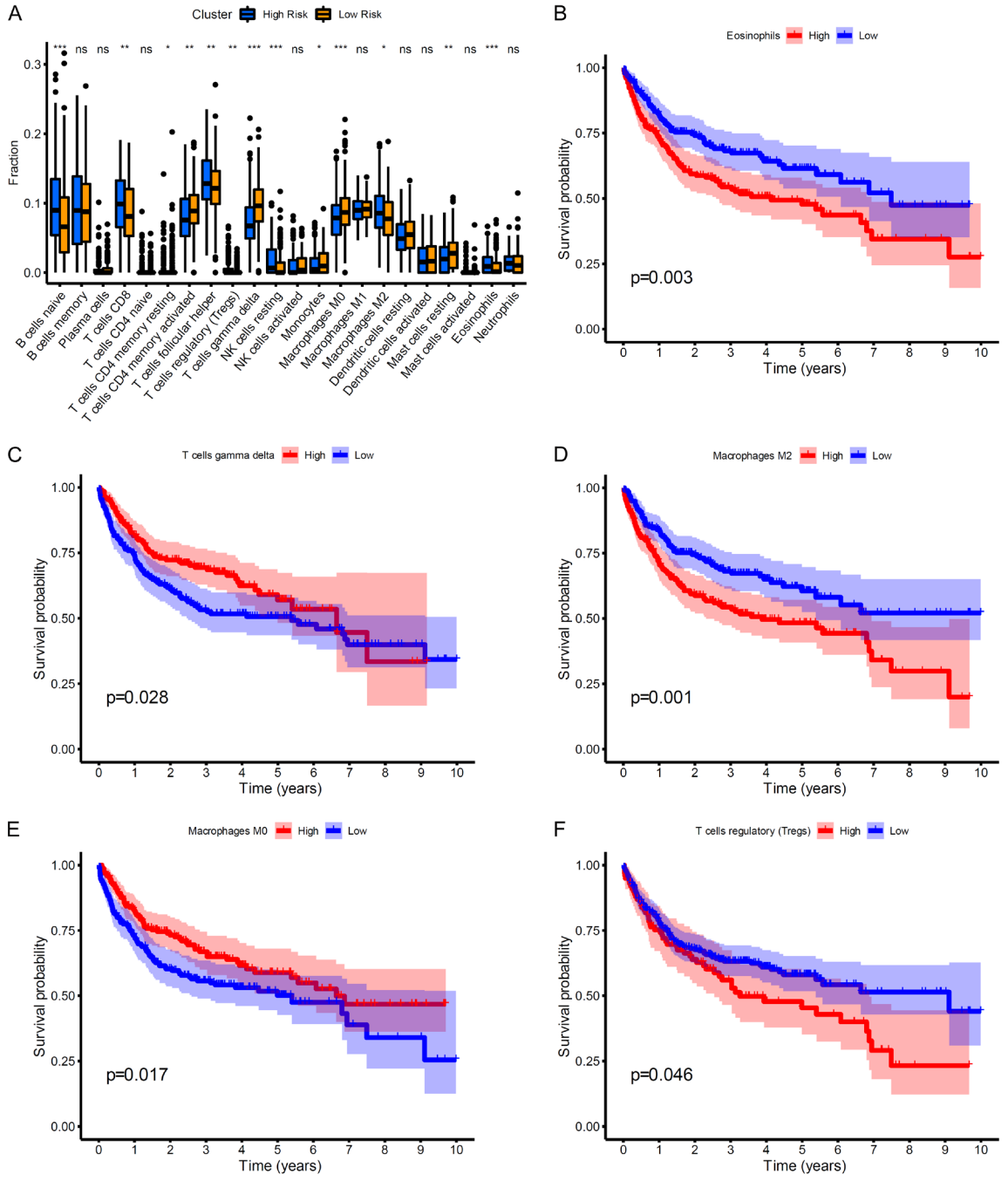


Figure 5. Association analysis between the m^{5}C prediction model and immune cell infiltration. A. Proportions of 22 immune cells infiltration in high and low-risk groups predicted by the m^{5}C prediction model. B-F. Survival analysis based on immune cells expression in DLBCL patients: eosinophils, macrophages M0, macrophages M2, gamma delta T cells and Tregs.

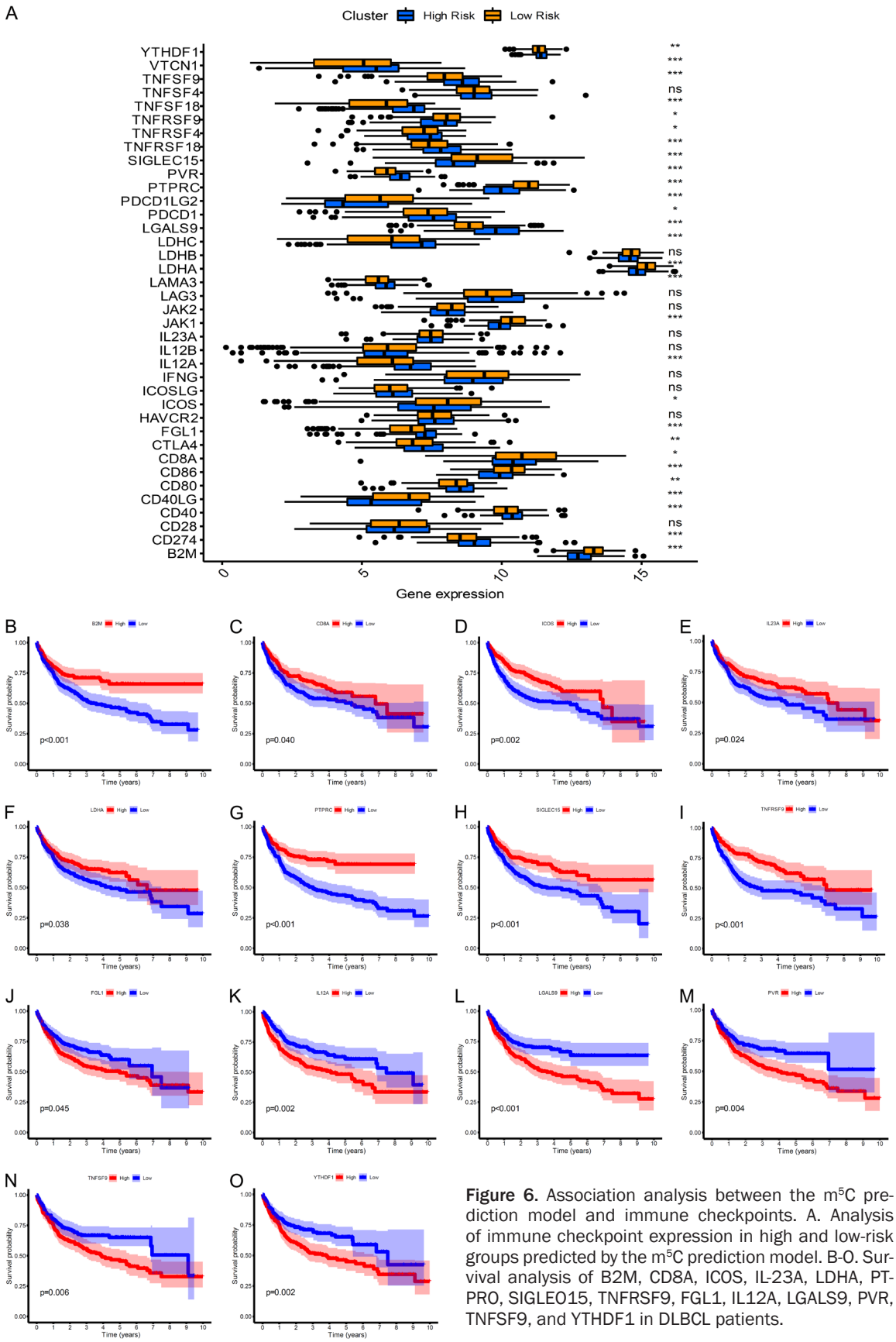


Figure 6. Association analysis between the m^{5}C prediction model and immune checkpoints. A. Analysis of immune checkpoint expression in high and low-risk groups predicted by the m^{5}C prediction model. B-O. Survival analysis of B2M, CD8A, ICOS, IL-23A, LDHA, PTPRO, SIGLEO15, TNFRSF9, FGL1, IL12A, LGALS9, PVR, TNFSF9, and YTHDF1 in DLBCL patients.

tumor cells in immune evasion [38, 39]. Stirm et al. found that Tregs, in cooperation with CD40 activation, sustained responses in lymphoma models. In addition, large-scale clinical studies have shown that Treg enrichment is associated with adverse outcomes, such as multiple extranidal involvement [40, 41]. These findings indicate that m^{5}C is closely associated with the activation and infiltration of immune cells, which can influence the development of DLBCL by modulating immune cell regulation.

The treatment of lymphoma has undergone notable advancements with the introduction of immune checkpoint therapies. Hence, the identification of dependable prognostic biomarkers and possible targets aids in minimizing adverse reactions to immunosuppressive treatment and broadening its suitability for patients with DLBCL. Upon examining the correlation between the m^{5}C pattern and immune checkpoints, it was discovered that the high-risk group exhibited a notable increase in the expression of immunosuppressive genes such as CTLA4 and CD274 (PDL1). In contrast, genes associated with antigen presentation or T-cell activation, such as B2M, CD8A, and ICOS, were upregulated in the low-risk group, and the expression of these immune checkpoints was closely related to patient prognosis. Multiple studies have shown that mutations in B2M are frequent in DLBCL and correlate with rituximab plus CHOP (R-CHOP) treatment outcomes [42]. This suggests that m^{5}C may influence DLBCL progression by affecting RNA's translation, stability, and degradation during immune checkpoints.

Although we investigated the correlation between m^{5}C-associated genes and DLBCL, our study has certain limitations. Considering this was a retrospective study, conducting a prospective study would be more persuasive. Furthermore, experimental studies on the functions of m^{5}C genes on immune cells and hematopoietic cell will be a focus of future research.

Conclusion

Our study successfully developed an m^{5}C model for predicting the prognosis of patients with DLBCL along with their various clinical characteristics. We present a genetic framework of m^{5}C genes for DLBCL, revealing the genetic and clinical heterogeneity of DLBCL and the treatment responses of its tumor subgroups defined by common pathogenic mechanisms. This categorization separated DLBCL into two genetic subcategories, showing variations in m^{5}C gene expression, tumor surroundings, and survival rates. Furthermore, we investigated the correlation between m^{5}C and immune checkpoints in DLBCL to identify potential therapeutic immune checkpoints.

Acknowledgements

We thank the Shanxi Province Cancer Hospital for providing the pathological sections for this research. We thank the Editage (www.editage.com) for its linguistic assistance during the preparation of this manuscript. This research was funded by the National Natural Science Foundation of China (82260043), the Jiangxi Provincial Nature Science Foundation (2023-2ACB206016, 20204BCJ22030) and the Main academic disciplines and technical leaders training program leading talents project of Jiangxi Province (20225BCJ22001).

Written informed consent was provided from the patients.

Disclosure of conflict of interest

None.

Abbreviations

m^{5}C, 5-methylcytosine; DLBCL, Diffuse large B-cell lymphoma; LASSO, Least Absolute Shrinkage and Selection Operator; TME, Tumor microenvironment; R-CHOP, Rituximab, cyclophosphamide, adriamycin, vincristine, prednisone; LDH, Lactate dehydrogenase; ECOG, Eastern Cooperative Oncology Group; PCA, Principal component analysis; ROC, Receiver operating characteristic; AUC, Area under curve; ABC, Activated B-cell; GCB, Germinal center B-cell-like; OS, Overall survival; IPI, International Prognostic Index.

Address correspondence to: Li Yu, Department of Hematology, The Second Affiliated Hospital of Nanchang University, Hematology Institute, Key Laboratory of Hematology of Jiangxi Province, Nanchang 330031, Jiangxi, China. Tel: +86-791-86292870; +86-18679124515; E-mail: ndefy020-21@ncu.edu.cn; zengyulii@126.com; Ruifang Sun, Department of Tumor Biobank, Shanxi Province Cancer Hospital, Shanxi Hospital Affiliated to Can-

cer Hospital, Chinese Academy of Medical Sciences, Cancer Hospital Affiliated to Shanxi Medical University, Taiyuan 030013, Shanxi, China. Tel: +86-13513611491; E-mail: sunruifang@sxmu.edu.cn

References

[1] Sehn LH and Salles G. Diffuse large B-cell lymphoma. N Engl J Med 2021; 384: 842-858.

[2] Alduaij W, Collinge B, Ben-Neriah S, Jiang A, Hilton LK, Boyle M, Meissner B, Chong L, Miyata-Takata T, Slack GW, Farinha P, Craig JW, Lytle A, Savage KJ, Villa D, Gerrie AS, Freeman CL, Gascoyne RD, Connors JM, Morin RD, Sehn LH, Mungall AJ, Steidl C and Scott DW. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. Blood 2023; 141: 2493-2507.

[3] Susanibar-Adaniya S and Barta SK. 2021 update on diffuse large B cell lymphoma: a review of current data and potential applications on risk stratification and management. Am J Hematol 2021; 96: 617-629.

[4] Poletto S, Novo M, Paruzzo L, Frascione PMM and Vitolo U. Treatment strategies for patients with diffuse large B-cell lymphoma. Cancer Treat Rev 2022; 110: 102443.

[5] Wang L, Li LR and Young KH. New agents and regimens for diffuse large B cell lymphoma. J Hematol Oncol 2020; 13: 175.

[6] Cui L, Ma R, Cai J, Guo C, Chen Z, Yao L, Wang Y, Fan R, Wang X and Shi Y. RNA modifications: importance in immune cell biology and related diseases. Signal Transduct Target Ther 2022; 7: 334.

[7] Shi H, Chai P, Jia R and Fan X. Novel insight into the regulatory roles of diverse RNA modifications: re-defining the bridge between transcription and translation. Mol Cancer 2020; 19: 78.

[8] Xue C, Chu Q, Zheng Q, Jiang S, Bao Z, Su Y, Lu J and Li L. Role of main RNA modifications in cancer: N(6)-methyladenosine, 5-methylcytosine, and pseudouridine. Signal Transduct Target Ther 2022; 7: 142.

[9] Chen YS, Yang WL, Zhao YL and Yang YG. Dynamic transcriptomic m(5) C and its regulatory role in RNA processing. Wiley Interdiscip Rev RNA 2021; 12: e1639.

[10] Haran V and Lenka N. Deciphering the epitranscriptomic signatures in cell fate determination and development. Stem Cell Rev Rep 2019; 15: 474-496.

[11] Song H, Zhang J, Liu B, Xu J, Cai B, Yang H, Straube J, Yu X and Ma T. Biological roles of RNA m(5)C modification and its implications in cancer immunotherapy. Biomark Res 2022; 10: 15.

[12] Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, Chen RX, Wei WS, Liu Y, Gao CC, Chen YS, Zhang M, Ma XD, Liu ZW, Luo JH, Lyu C, Wang HL, Ma J, Zhao YL, Zhou FJ, Huang Y, Xie D and Yang YG. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol 2019; 21: 978-990.

[13] Sjöström M, Zhao SG, Levy S, Zhang M, Ning Y, Shrestha R, Lundberg A, Herberts C, Foye A, Aggarwal R, Hua JT, Li H, Bergamaschi A, Maurice-Dror C, Maheshwari A, Chen S, Ng SWS, Ye W, Petricca J, Fraser M, Chesner L, Perry MD, Moreno-Rodriguez T, Chen WS, Alumkal JJ, Chou J, Morgans AK, Beer TM, Thomas GV, Gleave M, Lloyd P, Phillips T, McCarthy E, Haffner MC, Zoubeidi A, Annala M, Reiter RE, Rettig MB, Witte ON, Fong L, Bose R, Huang FW, Luo J, Bjartell A, Lang JM, Mahajan NP, Lara PN, Evans CP, Tran PT, Posadas EM, He C, Cui XL, Huang J, Zwart W, Gilbert LA, Maher CA, Boutros PC, Chi KN, Ashworth A, Small EJ, He HH, Wyatt AW, Quigley DA and Feng FY. The 5-hydroxymethylcytosine landscape of prostate cancer. Cancer Res 2022; 82: 3888-3902.

[14] Lopez-Bertoni H, Johnson A, Rui Y, Lal B, Sall S, Malloy M, Coulter JB, Lugo-Fagundo M, Shudir S, Khela H, Caputo C, Green JJ and Laterra J. Sox2 induces glioblastoma cell stemness and tumor propagation by repressing TET2 and deregulating 5hmC and 5mC DNA modifications. Signal Transduct Target Ther 2022; 7: 37.

[15] Xue C, Gu X, Zheng Q, Shi Q, Yuan X, Su Y, Jia J, Jiang J, Lu J and Li L. ALYREF mediates RNA m(5)C modification to promote hepatocellular carcinoma progression. Signal Transduct Target Ther 2023; 8: 130.

[16] Shukla V, Samaniego-Castruita D, Dong Z, González-Avalos E, Yan Q, Sarma K and Rao A. TET deficiency perturbs mature B cell homeostasis and promotes oncogenesis associated with accumulation of G-quadruplex and R-loop structures. Nat Immunol 2022; 23: 99-108.

[17] Fraietta JA, Nobles CL, Sammons MA, Lundh S, Carty SA, Reich TJ, Cogdill AP, Morrissette JJD, DeNizio JE, Reddy S, Hwang Y, Gohil M, Kulikovskaya I, Nazimuddin F, Gupta M, Chen F, Everett JK, Alexander KA, Lin-Shiao E, Gee MH, Liu X, Young RM, Ambrose D, Wang Y, Xu J, Jordan MS, Marcucci KT, Levine BL, Garcia KC, Zhao Y, Kalos M, Porter DL, Kohli RM, Lacey SF, Berger SL, Bushman FD, June CH and Melenhorst JJ. Disruption of TET2 promotes the therapeutic efficacy of CD19-targeted T cells. Nature 2018; 558: 307-312.

[18] Nakatsukasa H, Oda M, Yin J, Chikuma S, Ito M, Koga-Iizuka M, Someya K, Kitagawa Y, Ohkura N, Sakaguchi S, Koya I, Sanosaka T, Kohyama J, Tsukada YI, Yamanaka S, Takamura-Enya T, Lu Q and Yoshimura A. Loss of TET proteins in regulatory T cells promotes abnor-

mal proliferation, Foxp3 destabilization and IL-17 expression. Int Immunol 2019; 31: 335-347.
[19] Tsiouplis NJ, Bailey DW, Chiou LF, Wissink FJ and Tsagaratou A. TET-mediated epigenetic regulation in immune cell development and disease. Front Cell Dev Biol 2020; 8: 623948.
[20] Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, Xu W, Tan B, Goldschmidt N, Iqbal J, Vose J, Bast M, Fu K, Weisenburger DD, Greiner TC, Armitage JO, Kyle A, May L, Gascoyne RD, Connors JM, Troen G, Holte H, Kvaloy S, Dierickx D, Verhoef G, Delabie J, Smeland EB, Jares P, Martinez A, Lopez-Guillermo A, Montserrat E, Campo E, Braziel RM, Miller TP, Rimsza LM, Cook JR, Pohlman B, Sweetenham J, Tubbs RR, Fisher RI, Hartmann E, Rosenwald A, Ott G, Muller-Hermelink HK, Wrench D, Lister TA, Jaffe ES, Wilson WH, Chan WC and Staudt LM; Lymphoma/Leukemia Molecular Profiling Project. Stromal gene signatures in large-B-cell lymphomas. N Engl J Med 2008; 359: 2313-2323.
[21] Lenz G, Wright GW, Emre NC, Kohlhammer H, Dave SS, Davis RE, Carty S, Lam LT, Shaffer AL, Xiao W, Powell J, Rosenwald A, Ott G, Muller-Hermelink HK, Gascoyne RD, Connors JM, Campo E, Jaffe ES, Delabie J, Smeland EB, Rimsza LM, Fisher RI, Weisenburger DD, Chan WC and Staudt LM. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proc Natl Acad Sci U S A 2008; 105: 13520-13525.
[22] Nombela P, Miguel-López B and Blanco S. The role of m(6)A, m(5)C and Ψ RNA modifications in cancer: novel therapeutic opportunities. Mol Cancer 2021; 20: 18.
[23] Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB and Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013; 4: 2612.
[24] Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M and Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015; 12: 453-457.
[25] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015; 43: e47.
[26] Bakhshi TJ and Georgel PT. Genetic and epigenetic determinants of diffuse large B-cell lymphoma. Blood Cancer J 2020; 10: 123.
[27] Amara K, Ziadi S, Hachana M, Soltani N, Korbi S and Trimeche M. DNA methyltransferase DNMT3b protein overexpression as a prognostic factor in patients with diffuse large B-cell lymphomas. Cancer Sci 2010; 101: 1722-1730.
[28] Loo SK, Ch'ng ES, Lawrie CH, Muruzabal MA, Gaafar A, Pomposo MP, Husin A, Md Salleh MS, Banham AH, Pedersen LM, Møller MB, Green TM and Wong KK. DNMT1 is predictive of survival and associated with Ki-67 expression in R-CHOP-treated diffuse large B-cell lymphomas. Pathology 2017; 49: 731-739.
[29] Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, Lawrence MS, Roemer MGM, Li AJ, Ziepert M, Staiger AM, Wala JA, Ducar MD, Leshchiner I, Rheinbay E, Taylor-Weiner A, Coughlin CA, Hess JM, Pedamallu CS, Livitz D, Rosebrock D, Rosenberg M, Tracy AA, Horn H, van Hummelen P, Feldman AL, Link BK, Novak AJ, Cerhan JR, Habermann TM, Siebert R, Rosenwald A, Thorner AR, Meyerson ML, Golub TR, Beroukhim R, Wulf GG, Ott G, Rodig SJ, Monti S, Neuberg DS, Loeffler M, Pfreundschuh M, Trümper L, Getz G and Shipp MA. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat Med 2018; 24: 679-690.
[30] Suehara Y, Sakata-Yanagimoto M, Hattori K, Kusakabe M, Nanmoku T, Sato T, Noguchi M and Chiba S. Mutations found in cell-free DNAs of patients with malignant lymphoma at remission can derive from clonal hematopoiesis. Cancer Sci 2019; 110: 3375-3381.
[31] Zhang H, Zhai X, Liu Y, Xia Z, Xia T, Du G, Zhou H, Franziska Strohmer D, Bazhin AV, Li Z, Wang X, Jin B and Guo D. NOP2-mediated m5C modification of c-Myc in an EIF3A-dependent manner to reprogram glucose metabolism and promote hepatocellular carcinoma progression. Research (Wash D C) 2023; 6: 0184.
[32] Wang X, Deng D, Yan Y, Cai M, Liu X, Luo A, Liu S, Zhang X, Jiang H and Liu X. Genetic variants in m5C modification core genes are associated with the risk of Chinese pediatric acute lymphoblastic leukemia: a five-center case-control study. Front Oncol 2022; 12: 1082525.
[33] Lu Z, Liu R, Wang Y, Jiao M, Li Z, Wang Z, Huang C, Shi G, Ke A, Wang L, Fu Y, Xia J, Wen H, Zhou J, Wang X, Ye D, Fan J, Chu Y and Cai J. Ten-eleven translocation-2 inactivation restrains IL-10-producing regulatory B cells to enable antitumor immunity in hepatocellular carcinoma. Hepatology 2023; 77: 745-759.
[34] Tanager KS, Yu J, Chiu BC, Carll TC, Tatarian AH, Riedell P, Smith S, Kline J and Venkataraman G. Loss of 5-hydroxymethylcytosine expression is near-universal in B-cell lymphomas with variable mutations in epigenetic regulators. Haematologica 2022; 107: 966-969.

[35] Chen HY, Zhang WL, Zhang L, Yang P, Li F, Yang ZR, Wang J, Pang M, Hong Y, Yan C, Li W, Liu J, Xu N, Chen L, Xiao XB, Qin Y, He XH, Liu H, Zhu HC, He C, Lin J and Jing HM. 5-hydroxymethylcytosine profiles of cfDNA are highly predictive of R-CHOP treatment response in diffuse large B cell lymphoma patients. Clin Epigenetics 2021; 13: 33.

[36] Wang D, Wu W, Callen E, Pavani R, Zolnerowich N, Kodali S, Zong D, Wong N, Noriega S, Nathan WJ, Matos-Rodrigues G, Chari R, Kruhlak MJ, Livak F, Ward M, Caldecott K, Di Stefano B and Nussenzweig A. Active DNA demethylation promotes cell fate specification and the DNA damage response. Science 2022; 378: 983-989.

[37] Yan D, Xie Y, Huang L, Zhang Y, Gu R, Xie H, Huang X and Luo H. RNA m5C methylation orchestrates BLCA progression via macrophage reprogramming. J Cell Mol Med 2023; 27: 2398-2411.

[38] Autio M, Leivonen SK, Brück O, Karjalainen-Lindsberg ML, Pellinen T and Leppä S. Clinical impact of immune cells and their spatial interactions in diffuse large B-cell lymphoma microenvironment. Clin Cancer Res 2022; 28: 781-792.

[39] Huang YH, Cai K, Xu PP, Wang L, Huang CX, Fang Y, Cheng S, Sun XJ, Liu F, Huang JY, Ji MM and Zhao WL. CREBBP/EP300 mutations promoted tumor progression in diffuse large B-cell lymphoma through altering tumor-associated macrophage polarization via FBXW7-NOTCH-CCL2/CSF1 axis. Signal Transduct Target Ther 2021; 6: 10.

[40] Stirm K, Leary P, Wüst D, Stark D, Joller N, Karakus U, Boyman O, Tzankov A and Müller A. Treg-selective IL-2 starvation synergizes with CD40 activation to sustain durable responses in lymphoma models. J Immunother Cancer 2023; 11: e006263.

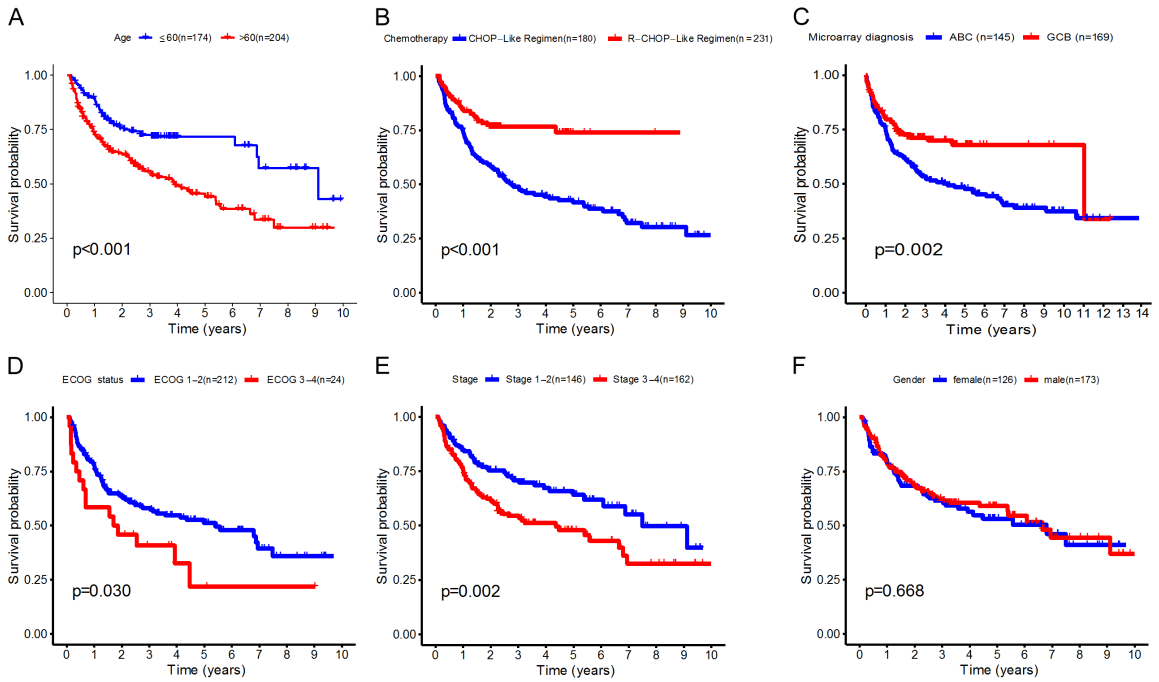
[41] Shen R, Xu PP, Wang N, Yi HM, Dong L, Fu D, Huang JY, Huang HY, Janin A, Cheng S, Wang L and Zhao WL. Influence of oncogenic mutations and tumor microenvironment alterations on extranodal invasion in diffuse large B-cell lymphoma. Clin Transl Med 2020; 10: e221.

[42] Higuchi T, Matsuo K, Hashida Y, Kitahata K, Ujihara T, Taniguchi A, Yoshie O, Nakayama T and Daibata M. Epstein-Barr virus-positive pyothorax-associated lymphoma expresses CCL17 and CCL22 chemokines that attract CCR4-expressing regulatory T cells. Cancer Lett 2019; 453: 184-192.

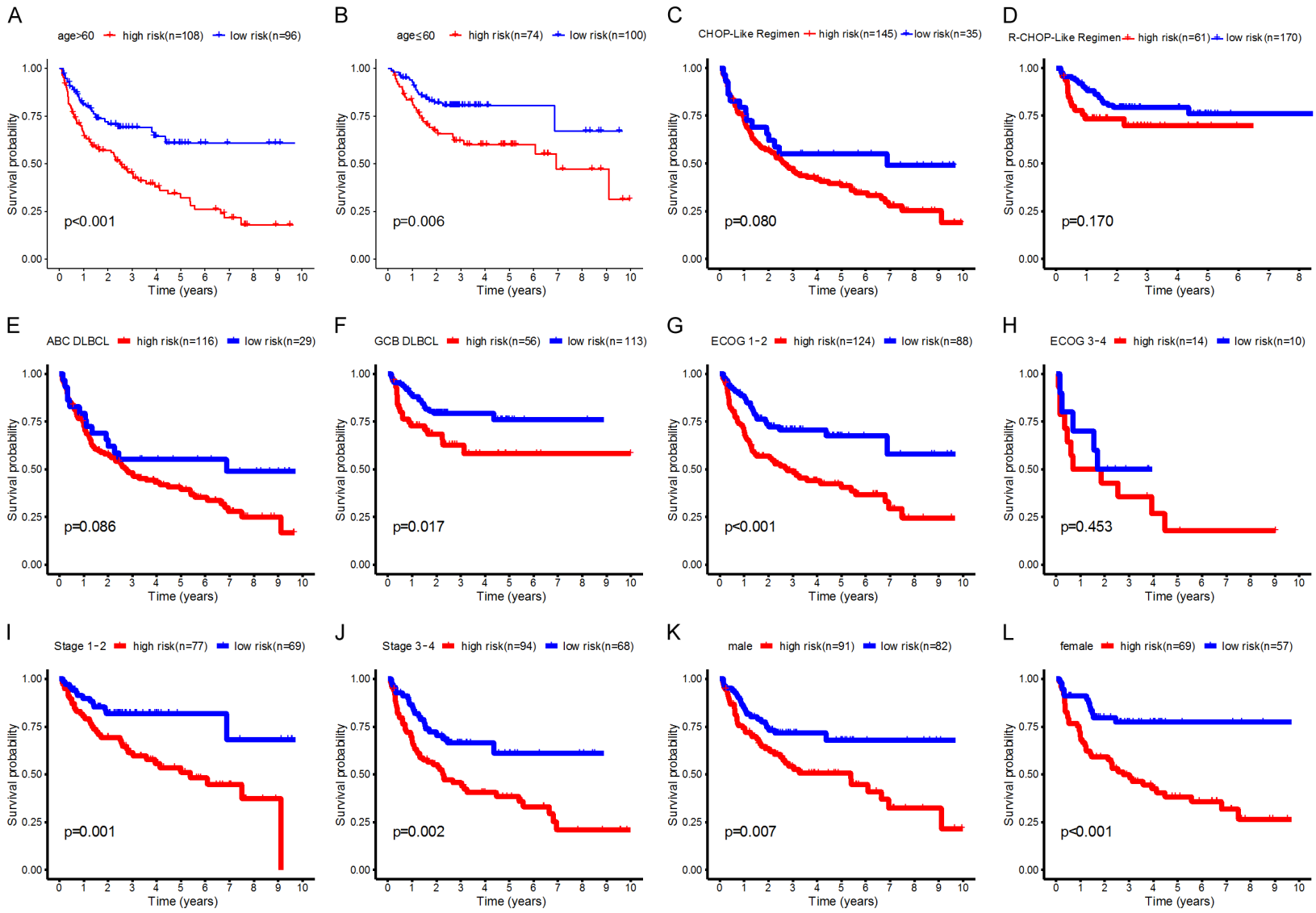
Supplementary Table 1. Clinical information of patients in the training set (GSE10846, N = 414) and validation set (GSE11318, N = 203)

Characteristics	DLBCL (N = 203)	DLBCL (N = 414)
Age at diagnosis		
≤ 60	69	188
> 60	94	226
Unknown	40	0
Gender		
Male	113	224
Female	90	171
Unknown	0	19
Microarray diagnosis		
GCB	71	183
Non-GCB	73	166
Unknown	59	65
ECOG performance status		
0-1	122	296
2-4	39	93
Unknown	42	25
Stage		
I-II	75	188
III-IV	87	218
Unknown	41	8
LDH ratio		
Normal	68	173
Upper limit of normal	76	178
Unknown	59	63
Number of extranodal sites		
< 2	161	353
≥ 2		30
Unknown	42	31
Chemotherapy		
CHOP-Like Regimen		181
R-CHOP-Like Regimen		233
Unknown	203	

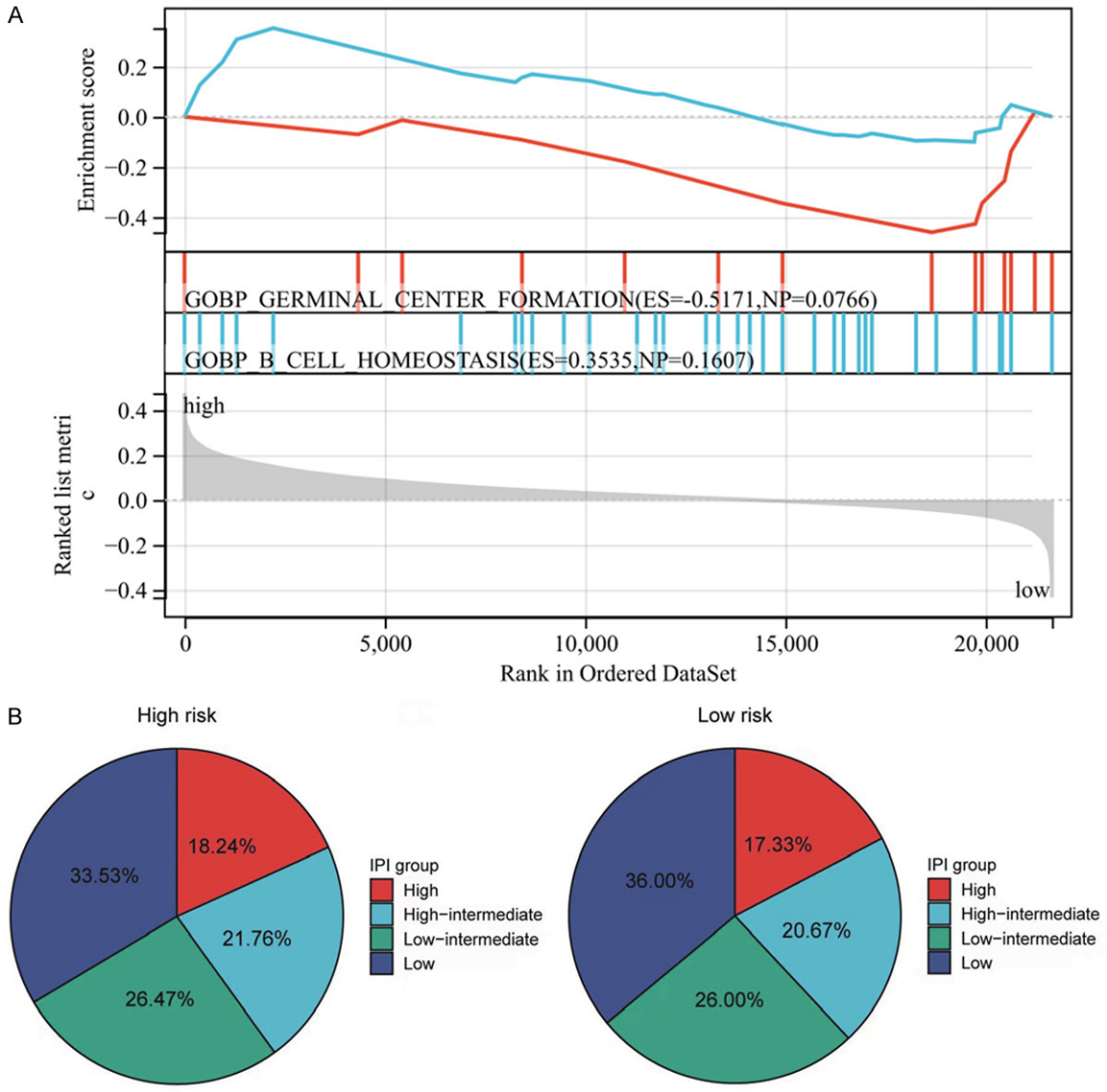
DLBCL, Diffuse Large B Cell Lymphoma; ECOG, Eastern Cooperative Oncology Group; GCB, Germinal Center B-cell; LDH, Lactate Dehydrogenase; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone.



Supplementary Figure 1. The association between clinical features and the prognosis of DLBCL patients. A-F. Age (> 60 years, ≤ 60 years); treatment (CHOP, R-CHOP); subtype (ABC, GCB); ECOG grade (1-2, 3-4); Stage (1-2, 3-4); gender (female, male). R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; ABC, activated B-cell-like; GCB, germinal-center B-cell-like; ECOG, Eastern Cooperative Oncology Group.



Supplementary Figure 2. The m^{5}C prognostic model's prognosis prediction for DLBCL patients across various clinical characteristics. A-L. Kaplan-Meier curves showing the relationship between age (> 60 years, ≤ 60 years), treatment (CHOP, R-CHOP), subtype (ABC, GCB), ECOG grade (G1-2, G3-4), stage (1-2, 3-4), gender (female, male) and survival time of patients with DLBCL in the high- and low-risk groups. R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone; ABC, activated B-cell-like; GCB, germinal-center B-cell-like; ECOG, Eastern Cooperative Oncology Group.

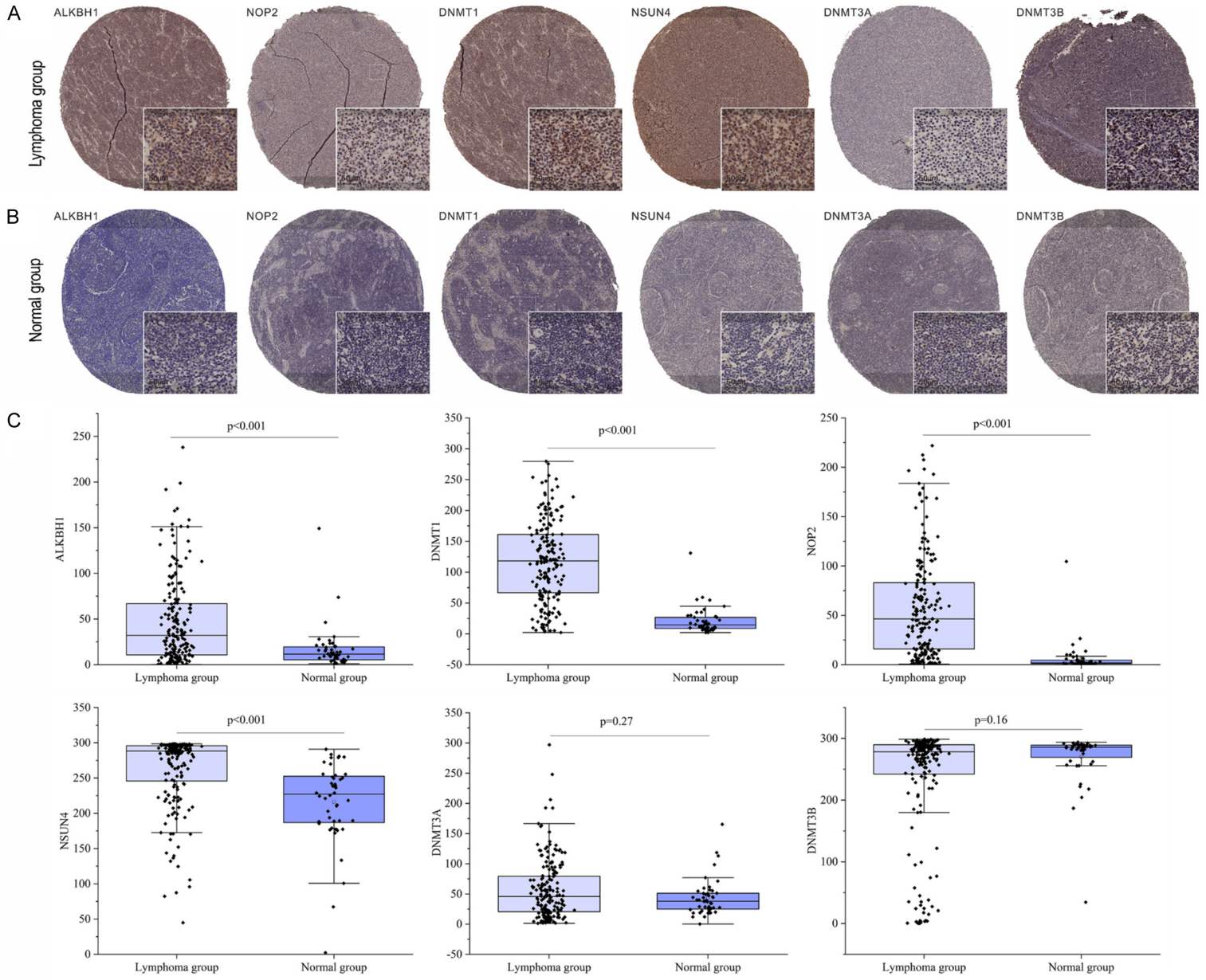


Supplementary Figure 3. A. The impact of NOP2 expression levels on cell proliferation. B. The IPI scores between high and low group. IPI, International Prognostic Index.

Supplementary Table 2. The clinical features of 144 DLBCL tissues for immunohistochemistry

Characteristics	DLBCL (N = 144)
Age at diagnosis	
≤ 60	65
> 60	79
Gender	
Male	63
Female	81
Microarray diagnosis	
GCB	27
Non-GCB	39
Unknown	78
ECOG performance status	
0-1	17
2-4	92
Unknown	35
Stage	
I-II	69
III-IV	75
LDH ratio	
Normal	58
Upper limit of normal	86
Number of extranodal sites	
< 2	90
≥ 2	54
Chemotherapy	
CHOP-Like Regimen	32
R-CHOP-Like Regimen	79
Unknown	33

DLBCL, Diffuse Large B Cell Lymphoma; ECOG, Eastern Cooperative Oncology Group; GCB, Germinal Center B-cell; LDH, Lactate Dehydrogenase; R-CHOP, rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone.



Supplementary Figure 4. The immunohistochemical analysis of m^{5}C genes in DLBCL samples. A, B. Immunohistochemical staining of m^{5}C genes in normal lymph nodes and DLBCL tissue. C. The differential expression of NOP2, DNMT1, NSUN4, ALKBH1, DNMT3B and DNMT3A between normal lymph nodes and DLBCL tissue.