

RESEARCH

Open Access



Refining prognostic assessment of diffuse large B-cell lymphoma: insights from multi-omics and single-cell analysis unveil SRM as a key target for regulating immunotherapy

Xiaojie Liang^{1†}, Jia Guo^{1†}, Baiwei Luo^{2†}, Weixiang Lu¹, Qiumin Chen³, Yeling Deng³, Yunong Yang³ and Liang Wang^{1*}

[†]Xiaojie Liang, Jia Guo and Baiwei Luo contributed equally to this work.

*Correspondence: wangliangtrhos@126.com

¹ Department of Hematology, Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China

² Department of Hematology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China

³ The First School of Clinical Medicine, Guangdong Medical University, Zhanjiang 524000, China

Abstract

Purposes: Previous studies have demonstrated that proliferation, stroma or immunity strongly influence the prognosis and therapeutic resistance of diffuse large B-cell lymphoma (DLBCL). Herein, we aimed to integrate proliferation, stromal, and immune (PSI) features to systematically evaluate the risk stratification and explore novel therapeutic targets in DLBCL.

Methods: Using data from multiple researches, we comprehensively evaluated the characteristics and prognostic impact of PSI features in DLBCL, and developed a novel risk stratification model (PSI score) with a consistent cutoff value to stratify the risk of 3,229 DLBCL patients from different cohorts. Mechanisms underlying adverse prognosis in the high-risk DLBCLs were investigated through transcriptomic (n = 3,229), genomic (n = 576), and scRNA-seq (n = 20) analyses.

Results: We identified a high-risk DLBCL subgroup (HPSI, 36.1% of DLBCL). HPSI was characterized by upregulation of spermidine synthase (SRM) and cold tumor microenvironment (TME). Compared to low-risk group, HPSI exhibited poorer prognosis, with lower 3-year OS (51.7% vs. 78.1%, $P < 0.0001$) and PFS (48.9% vs. 72.6%, $P < 0.0001$) rates. HPSI shared malignant proliferative phenotype resembling Burkitt lymphoma. Genomic analysis revealed extensive copy-number loss in the chemokine and interleukin coding regions within HPSI. Bulk and scRNA-seq analyses indicated that upregulation of SRM might mediate cold TME in DLBCL, potentially through suppressing immune activation pathways, promoting dendritic cells (DCs) transformation into tolerogenic DCs, and facilitating M2 polarization of macrophages. Finally, for eventual clinical translation, we integrated the model with other clinical features to develop a comprehensive database for DLBCL.

Conclusion: Our study effectively simplifies risk stratification of DLBCL, revealing that immune microenvironment and SRM jointly shape a subgroup of DLBCL.

with extremely poor prognosis. Targeting SRM may become a potential strategy for modulating immunotherapy in DLBCL, providing new insight for immunotherapy.

Keywords: DLBCL, Proliferation, Stromal, Immune, Risk stratification, SRM, TME, Immunotherapy

Introduction

Diffuse large B-cell lymphoma (DLBCL) is a heterogeneous and aggressive form of lymphoma with a high mortality rate [1]. Despite advances in treatment modalities, approximately 40% of patients still demonstrate resistance to standard treatment and experience refractory or relapsed disease [2]. This necessitates clinical researchers to explore novel prognostic biomarkers and therapeutic targets.

Previous studies have emphasized that in addition to tumor cells, the immune and stromal component within the tumor microenvironment (TME) plays a vital role in the biological behavior and treatment resistance of DLBCL [3, 4]. The TME within DLBCL is a complex ecosystem consisting of various cell types, including tumor cells, immune cells, stromal cells, and extracellular matrix (ECM) components [4, 5]. Moreover, not only do the interactions between tumor and stromal cells play a pivotal role in DLBCL progression, but the abundance of stromal cells also leads to an increased production of various growth factors, activation of fibroblasts, and secretion of ECM proteins. These factors, in turn, promote treatment resistance, invasion, and proliferation in DLBCL [6, 7]. Additionally, elevated expression of c-Myc, along with several other genes involved in proliferation, has been linked to a poorer prognosis in DLBCL [8].

In support of these exciting scientific discoveries, comprehensive transcriptomic analyses conducted by multiple research groups have unveiled distinct molecular characteristics that shed light on the involvement of TME in DLBCL. For example, G. Lenz et al. revealed a close association between stromal signature and the prognosis of patients receiving rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP/CHOP) in DLBCL [9]; Nikita and colleagues identified four distinct lymphoma microenvironments (LMEs) that are associated with overall survival (OS) in DLBCL [10]; Chloe' B and colleagues revealed that context-specific cell states and ecosystems are associated with therapeutic responsiveness and resistance in DLBCL [4]. However, the application of these findings in clinical practice has encountered two primary obstacles. Firstly, despite the advancement in comprehending the unique TME within DLBCL through these endeavors, there is still a lack of consensus concerning the terminology and gene panel selection. Secondly, the utilization of gene expression-based signatures to identify distinct molecular characteristics within DLBCL necessitates the analysis of numerous genes using sophisticated bioinformatics methods, presenting a challenge for their routine integration into clinical settings. Nonetheless, considering the growing recognition of the tumor stroma and immune components in DLBCL, we hypothesize that the integration of the proliferation, stromal, and immune biomarkers would facilitate a more robust risk stratification, enabling clinicians to tailor personalized treatment regimens and improve therapeutic outcomes for individuals diagnosed with DLBCL.

In this study, we conducted a comprehensive investigation aimed at discovering and validating biomarkers for robust prognostication in DLBCL patients by considering

various components of the TME. Our research unveiled a novel 17-gene signature, referred to as the proliferation, stromal, and immune-related (PSI) signature, which demonstrated remarkable robustness in risk stratification for DLBCL. This unique gene signature enabled the identification of two distinct subgroups characterized by diverse clinical molecular features, transcriptional expression patterns, TME composition, and genomic alterations. The PSI gene signature provides an appealing platform for accurately stratifying the risk of DLBCL patients, thereby holding significant implications for the clinical management of this aggressive malignancy.

Methods

Data sources

The workflow of the study is illustrated in Figure S1. We systematically collected a total of 3229 DLBCL samples from multiple studies [9, 11–15]. Public multi-omics data can be accessed from the Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov>) database and the Gene Expression Omnibus (GEO: <https://www.ncbi.nlm.nih.gov/geo/>) database. Detailed data processing and analysis methods are described in supplementary methods.

Assessment of the enrichment levels of stromal activities, cellular proliferation, and immune components in DLBCL

We collected 468 human genes involved in cellular proliferation, stroma and immune components from previous researches (Table S1) [9, 16–18]. Gene set variation analysis (GSVA) was utilized to determine the enrichment abundance of cellular proliferation, stroma and immune components according to transcriptomics data [19]. Additionally, we used the GSEA function in the R package “clusterProfiler” to conduct functional enrichment analysis [20, 21].

Assessment of global differences of stromal activities, cellular proliferation, and immune components

We utilized two different methods to assess the global differences in stromal activities, cellular proliferation, and immune components between and within DLBCLs and normal samples [22]:

- (1) The Euclidean distance,

$$RMSD = \sqrt{\sum_{i=1}^n (\log_2 x_i - \log_2 y_i)^2 / n}$$

where x_i and y_i are the expression of gene i over two messenger RNA (mRNA) expression profiles, and n is the number of genes assigned to proliferation, stromal, and immune signatures.

(2) The correlation-based distance,

$$d_{cor} = 1 - cor(\log_2 x, \log_2 y)$$

where *cor* is Spearman's rank correlation coefficient between $\log_2$ expression values of the PSI genes in the two mRNA expression profiles.

Differential rank conservation (DIRAC) analysis

DIRAC algorithm evaluates the differential expression variability in distinct phenotypes using particular gene subsets from individual transcriptomes [23]. In this study, five PSI gene sets were employed to calculate the rank conservation indices (RCIs). According to the level (low/high) of conservation of transcript ordering (RCI), the gene set was defined as being loosely or tightly regulated.

Feature engineering for selecting PSI signature and model development and validation

To develop a risk stratification model that integrates tumor proliferation, stromal, and immune features, captures the heterogeneity of the TME, and reflects both intrinsic and extrinsic factors influencing DLBCL development, progression, and prognosis, we conducted a rigorous feature selection using machine learning algorithms. A total of 702 patients treated with R-CHOP from the GSE10846 ($n=232$) and GSE31312 ($n=470$) cohorts were combined to form the training set. The PCA plots and density plots show that, after correction of batch effects, the combined expression matrix achieves a greater degree of transcriptional homogeneity (Figures S2A and S2B). We firstly employed the least Absolute Shrinkage and Selection Operator (LASSO) to select PSI genes in the training cohort. For the LASSO selection of the 468 PSI genes, we subsampled 70 percent of the training cohort without replacement 1000 iterations and selected the markers with repeat occurrence frequency more than 500. The tuning parameters were determined using tenfold cross-validation, along with Akaike and Bayesian information criteria (AIC/BIC), and the lambda value within one standard error of the minimum. This process identified 22 genes from the initial 468 features. Subsequently, we incorporated these selected genes into a penalized stepwise Cox regression model (both forward and backward feature selection) in the training cohort, which further narrowed the selection to 17 genes that defined the final risk model (Figure S1).

Evaluate differences in transcription patterns among high- and low-PSI score DLBCL and BL

In order to investigate the potential mechanisms underlying high PSI score (HPSI) versus low PSI score (LPSI) DLBCL, we conducted GSVA on 23 gene sets associated with B-cell disorders from the normal and pathological lymphoid biology database [24] (Table S2). In addition, we obtained 29 functional gene expression signatures (Fges) that encompass key functional elements and immune, stromal, and other cell populations within tumors [25] (Table S3), as well as 28 gene sets of tumor-infiltrating immune cells representing adaptive and innate immunity [26] (Table S4) for additional investigation.

Characteristic genomic alterations of PSI-model-based subgroups in DLBCL

To assess the mutation frequency of genes associated with HPSI and LPSI DLBCL, we examined alterations in the cancer hallmark signaling pathways in each group. We utilized Fisher's exact test to compare the fraction of samples with at least one alteration in each pathway between the two groups. Copy number variation (CNV) data was annotated using Bedtools (v2.25.0) and GISTIC2.0 (v2.0.23), and we also performed a Fisher's exact test to compare somatic copy number alteration (CNAs) between the two groups.

Evaluation of immunotherapy response

We utilized two independent cohorts to evaluate the performance of the PSI model in immunotherapy. The first cohort (GSE78220) [27] consisted of patients with metastatic melanoma who were treated with pembrolizumab, an anti-programmed death-1 (PD-1) antibody. The second cohort (IMvigor210) [28] included patients with advanced urothelial cancer who received atezolizumab, an anti-PD-L1 antibody.

Statistical analysis

Spearman's correlation was used for correlation analysis. For comparison between two groups, either Student's t-test or Mann–Whitney test was applied. Differences among three or more groups were compared using either one-way analysis of variance or Kruskal–Wallis test [29]. The significance of differences in OS and progression-free survival (PFS) was determined by KM survival analysis and log-rank test. All reported P values were two-sided. Data processing was performed using R or Python software.

Results

Transcriptomic landscape uncovered dysregulation of stromal functionality, cellular proliferation, and immune components in DLBCL

In order to investigate the PSI characteristics of DLBCL, we obtained 468 PSI genes from previous studies (Table S1). Principal component analysis (PCA) showed clear differences in the PSI gene expression patterns between DLBCL and normal samples (Fig. 1A). The Euclidean distance (Fig. 1B) and correlation-based distance (Figure S3A) revealed that the expression distance of PSI genes between DLBCL and normal samples, as well as within DLBCL samples, was significantly larger than that within normal samples. Through differential expression analysis, we found heterogeneous expression in most of the proliferation, stromal, and immune genes between DLBCL and normal samples (Fig. 1C). In terms of gene set enrichment analysis (GSEA), we observed significant dysregulation [normalized enrichment score (NES) > 1.5 and adjusted p-value < 0.001] of PSI-related signalling pathways, including E2F targets, G2M checkpoint, MYC targets and Epithelial-mesenchymal transition in DLBCL samples relative to normal samples (Fig. 1D). Furthermore, we conducted differential rank conservation analysis to assess the conservation differences in PSI signatures (Fig. 1E). In comparison with normal samples, DLBCL showed significantly lower RCIs in PIS signatures, which implied that PSI signatures had higher heterogeneity and higher degree of dysregulation in the DLBCL.

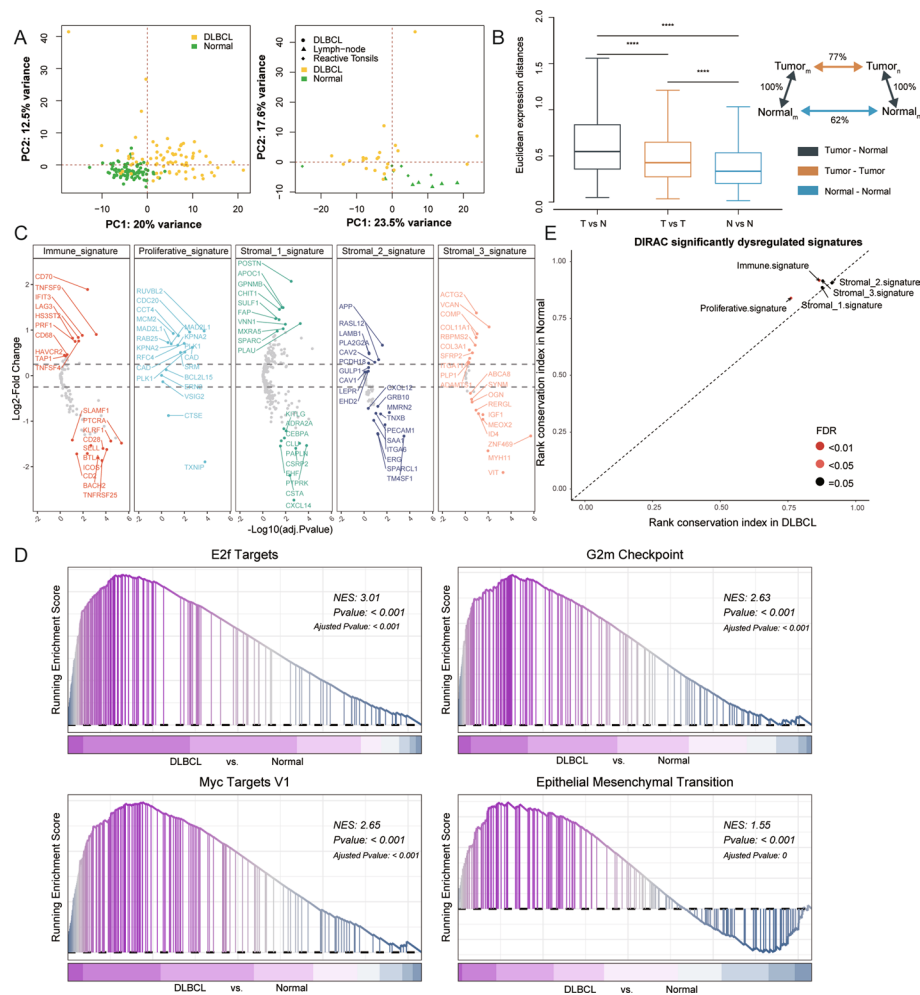


Fig. 1 Dysregulation of stromal functionality, cellular proliferation, and immune components in DLBCL. **A** Left: principal component analysis of normal and DLBCL peripheral blood samples in the GSE83632 cohort for the expression of PSI genes. Right: In the GSE32018 cohort, principal component analysis of normal tissues (lymph nodes and reactive tonsils) and DLBCL samples for the expression of PSI genes. **B** Global differences in PSI gene expression between DLBCL samples and normal tissues (lymph nodes and reactive tonsils). The Euclidean expression distances were estimated between DLBCL and normal samples (cyan), different DLBCL samples (orange), and different normal samples (blue). The legends summarize the average distances between different samples as a percentage of the average distance between DLBCL and normal samples. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$. **C** The volcano plot shows the results of differential expression analysis for PSI-associated genes between DLBCL samples and normal tissues (lymph nodes and reactive tonsils). The y-axis represents the log2 fold change, and the x-axis represents the $-\log_{10}$ adjusted P-value. **D** GSEA results showed significant upregulation of PSI-associated signaling pathways, such as E2f targets, G2m checkpoint, MYC targets, and Epithelial-mesenchymal transition in DLBCL compared to normal tissues (lymph nodes and reactive tonsils). **E** Comparison of the rank conservation index (RCI) of PSI signatures between DLBCL samples and normal tissues in the GSE32018 cohort. An RCI of 1.0 indicated that the ranks of signature genes were almost unchanged among samples, while an RCI of 0.5 indicated that the ranks of signature genes were greatly varied among samples of the same phenotype. FDR < 0.05 represents significantly dysregulated signatures between DLBCL and normal samples

Taken together, these data indicate that DLBCL exhibits significant PSI heterogeneity and dysregulation. And in view of such sharp heterogeneity and dysregulation, our hypothesis was that a gene expression signature associated with stromal functionality, cellular proliferation, and immune cells could serve as a compelling framework for identifying and selecting DLBCL patients with an unfavorable prognosis.

Identification of a 17-gene risk stratification signature for DLBCL

To further validate our hypothesis and develop a concise and robust gene panel for clinical applications, we performed the LASSO and stepwise Cox regression to analyse the 468 PSI genes in the training set. First, using the LASSO algorithm, we identified 22 PSI genes that appeared more than 500 times in 1000 repeated iterations. Next, we integrated these selected genes into a penalized stepwise Cox regression model to further narrow down to 17 and built the PSI score based on the following formula (*where $h_0(t)$ is the baseline risk function*):

$$\begin{aligned} \text{PSI score} = & h_0(t) \times \exp(0.70517 \times \text{SRM} + 0.16336 \times \text{FABP4} - 0.16527 \\ & \times \text{LYZ} - 0.80187 \times \text{VNN3} - 0.22452 \times \text{NAV2} - 0.14652 \\ & \times \text{FNDC1} + 0.50846 \times \text{DDX4} - 0.16316 \times \text{CD79A} \\ & - 0.0952 \times \text{FCRL3} + 0.36025 \times \text{CD68} + 0.53115 \times \text{PADI4} \\ & - 0.23496 \times \text{GSN} + 0.24376 \times \text{GABARAPL1} - 0.27031 \\ & \times \text{PTTG1IP} - 0.35819 \times \text{CTLA4} - 0.19987 \\ & \times \text{PDLIM3} + 0.20235 \times \text{SRPX}) \end{aligned}$$

In each cohort, DLBCL patients were categorized into low-PSI score group (LPSI) and high-PSI score group (HPSI) using the cutoff value 1.289567 determined by the *surv_cutpoint* function of the *Survminer* R package in the training cohort. Compared to LPSI-DLBCL, HPSI-DLBCL showed a worse prognosis with significantly lower 3-year OS rates (51.7% vs. 78.1%) and PFS rates (48.9% vs. 72.6%), as well as significantly reduced 5-year OS rates (42.7% vs. 72.5%) and PFS rates (45.7% vs. 69.1%) (Fig. 2A and B). Remarkably, the 17-gene model allowed for reliable stratification of DLBCL patient prognosis across multiple research centers utilizing a consistent cut-off point. This approach ensures the dependable assessment of OS and PFS in DLBCL (Fig. 2C–J; S3B–S3D). Additionally, univariate and multivariate Cox proportional hazard regression analyses demonstrated that our 17-gene model was an independent predictor of both favorable and unfavorable OS and PFS in DLBCL, even after adjusting for clinical covariates such as the International Prognostic Index (IPI) and cell-of-origin (COO) subtype (Figures S4 and S5).

PSI-based subgroups exhibit distinct clinicopathological features

We proceeded to examine the overlap between established clinicopathological features and PSI subgroups. HPSI-DLBCL had a higher prevalence of activated B-cell-like (ABC), unclassified (UNC), MCD, or molecular high-grade (MHG) subtypes, PFS events, and advanced stage, with a more frequent medium–high IPI. In comparison, LPSI-DLBCL had a higher proportion of germinal center B-cell-like (GCB), BN2 or EZB subtypes, low IPI score, early stage, and non-PFS events (Fig. 3A and B; S6A

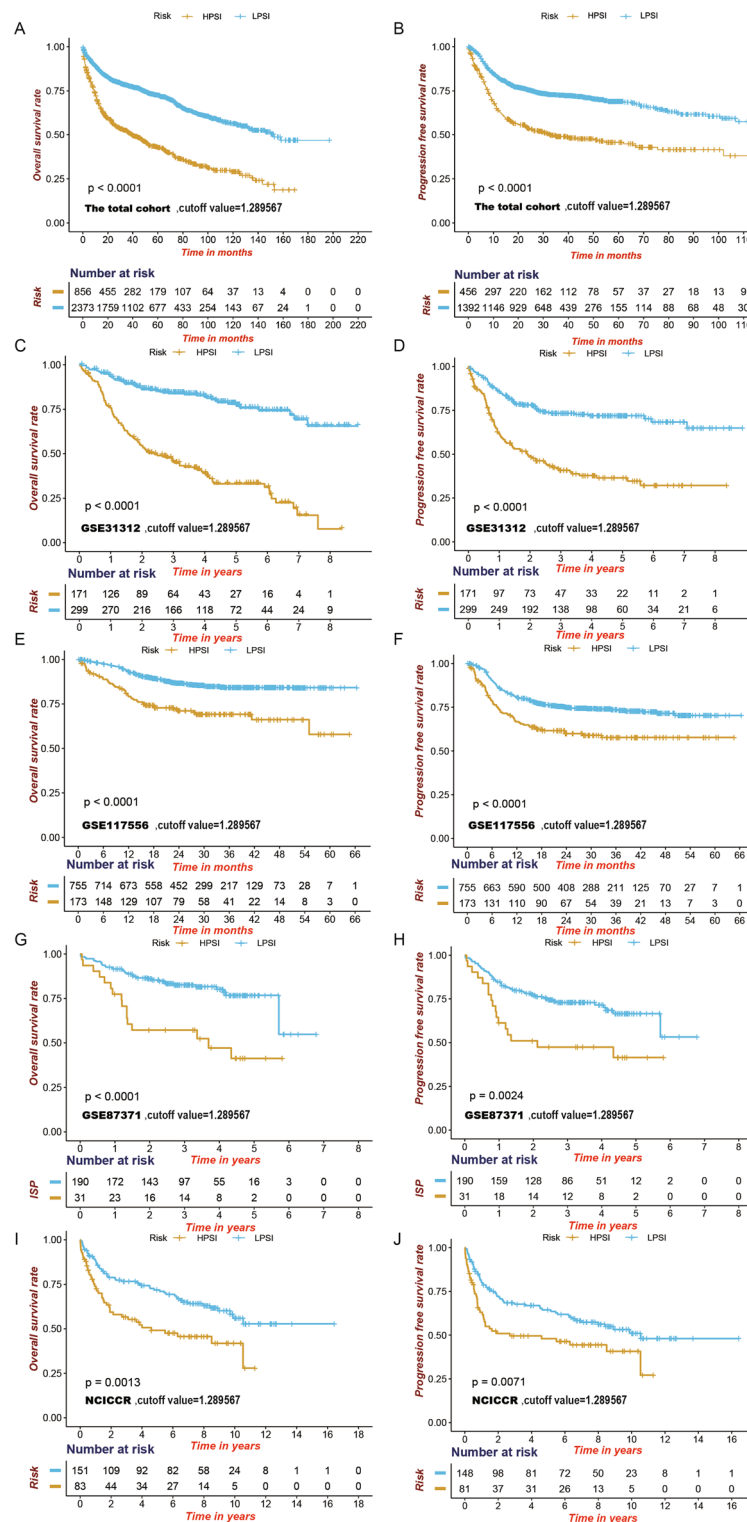


Fig. 2 17-gene-based risk stratification of DLBCL. **A–J** The risk stratification model exhibited strong and consistent ability to stratify the risk of multiple DLBCL cohorts, as evidenced by the KM curves for OS and PFS in the total cohort, internal validation cohort, and external validation cohorts. The significance of the results was assessed using the log-rank test. **A, B** the total cohort. **C, D** the internal cohort GSE31312. **E, F** the external cohort GSE117556. **G, H** the external cohort GSE87371. **I, J** the external cohort TCGA-NCICCR

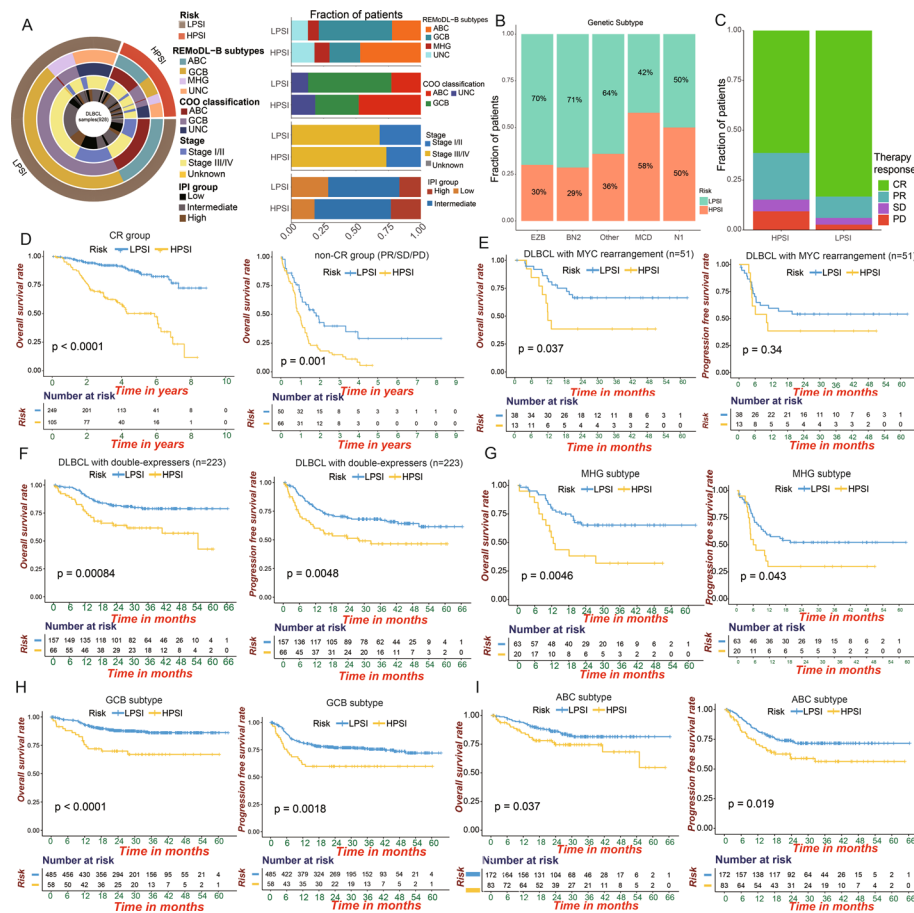


Fig. 3 Clinicopathological features in LPSI and HPSI DLBCLs. **A** Left: The association of 17-gene-based subgroups (outer rings) with established clinical and molecular subtypes (inner ring), such as COO classification, REMoDL-B trial subtypes, IPI groups, and tumor stage. Right: Bar charts illustrating the prevalence of COO subtypes, REMoDL-B trial subtypes, IPI groups, and tumor stage within 17-gene-based subgroups. REMoDL-B refers to the randomized evaluation of molecular-guided therapy for diffuse large B-cell lymphoma with bortezomib clinical trial. **B** The relative proportions of HPSI- and LPSI-DLBCL in various genetic subtypes, such as MCD, BN2, N1, and EZB. **C** Distribution of therapy response rates among patients categorized into HPSI and LPSI groups in the GSE31312 cohort (two-tailed Fisher exact test; $P = 1.8e-06$). CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease. **D** Survival plots for OS demonstrate the 17-gene model stratification for individual responses to R-CHOP among patients in the complete response (CR) group (left) and non-complete response (PR/SD/PD) group (right). **E** The OS and PFS KM curves were constructed for DLBCL patients with MYC rearrangement, stratified into HPSI and LPSI groups. MYC rearrangements were detected using fluorescent in situ hybridization (FISH). **F** The OS and PFS KM curves were generated for DLBCL patients with double-expressors, assessed by RNA-seq, characterized by high expression of both MYC and BCL2 at the mRNA level. Patients were stratified into HPSI and LPSI groups. **G-I** The OS and PFS KM curves were plotted for DLBCL patients with GCB, ABC or MHG subtypes in the GSE117556 cohort, categorized into HPSI and LPSI groups. For KM curves, the p-values were calculated using the log-rank test

and S6B). HPSI-DLBCL demonstrated a higher likelihood of experiencing progressive disease (PD) following treatment with R-CHOP or RB (Bortezomib)-CHOP. This finding indicates that our 17-gene model holds the potential to predict the therapeutic response in DLBCL (Fig. 3C and S6C). Notably, PSI subgroups were strongly associated with the outcome of DLBCL patients, irrespective of their therapy response (Fig. 3D and S6D-S6F). For instance, we noted a persistent pattern of mortality in

patients who achieved complete remission (CR), suggesting that CR may not always signify a complete cure (Fig. 3D and S6D–S6F). Our 17-gene model can identify individuals within the CR group who are at risk of early mortality. Remarkably, patients with high-risk factors such as MYC rearrangement, double-expressors and MHG subtype do not necessarily predict unfavorable outcomes in all cases. Of these patients, there were still some who have a favorable prognosis that can be identified by the 17-gene model (Fig. 3E–G). However, when we specifically analyzed survival outcomes in double-hit or triple-hit DLBCL (DHL/THL) patients, our 17-gene model did not achieve significant stratification (Figures S6G and S6H). While the OS curves for the HPSI and LPSI groups of DHL/THL patients showed a separation trend, this difference did not reach statistical significance. Notably, the 17-gene model effectively distinguishes the prognosis of DLBCL patients within ABC, GCB, or UNC subtypes, demonstrating its robustness and broad applicability in diverse biological contexts, reinforcing its potential clinical utility for DLBCL prognostic assessment (Fig. 3H, I, and S7A–S7F).

PSI-based subgroups show different transcriptional patterns

Given the critical role of transcriptional patterns in shaping the clinical and molecular characteristics of DLBCL [9, 10, 30, 31], we hypothesize that variations in transcriptional patterns contribute to the high-risk biological features and clinical prognosis of HPSI-DLBCL. To investigate our hypothesis, we systematically compared HPSI- and LPSI-DLBCLs with Burkitt lymphoma (BL) using the 23 gene signatures related to B-cell diseases. Transcriptome expression patterns revealed that HPSI-DLBCL and BL shared high expression of genes involved in cell cycle, proliferation, ribosomal proteins and MYC overexpression, indicating a common proliferative phenotype (Fig. 4A and S8A). Notable signatures that exhibited decreased expression in HPSI-DLBCL were associated with T/NK and dendritic cells, plasma cells, stromal, monocytes and inflammation, suggesting a potential cold TME phenotype in HPSI-DLBCL (Fig. 4A and S8A). Furthermore, we corroborated our hypothesis by analyzing the 29 functional gene expression signatures (Eges) from Bagaev's study [25], HPSI-DLBCL exhibited features consistent with an immune-depleted phenotype, characterized by the downregulation of fibrosis, the absence of most anti-tumor and pro-tumor immune infiltrating cells, and an upregulated malignant cell proliferation rate (Fig. 4B–E and S8B–S8E). Notably, there is an upregulation of tumor-associated macrophages, neutrophil signature, and Th2 signature

(See figure on next page.)

Fig. 4 Distinct transcriptional patterns are observed in 17-gene-based subgroups. **A** The heatmap illustrates the enrichment scores (blue = low, red = high) of selected signatures in LPSI and HPSI DLBCL subgroups of the GSE117556 cohort, while also comparing gene expression patterns in 70 patients with BL. Fb, fibroblast; NF- κ B, nuclear factor kappa-lightchain-enhancer of activated B cell; NK, natural killer cell. **B–E** Differences of angiogenesis and fibrosis (**B**), anti-tumor immune infiltrate (**C**), tumor proliferation rate (**D**) and pro-tumor immune infiltrate (**E**) among HPSI and LPSI DLBCL samples of the GSE117556 cohort. **F** Abundance of TME-infiltrating innate and adaptive immune cells among HPSI and LPSI DLBCL samples of the GSE117556 cohort. **G** Differences of immune score and stromal score between HPSI and LPSI DLBCL of the training cohort. **H** Correlation between immune score and PSI score in the training cohort. For boxplots, the lines in the boxes represent median values and the lower and upper ends of the box represent the interquartile range of values. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$

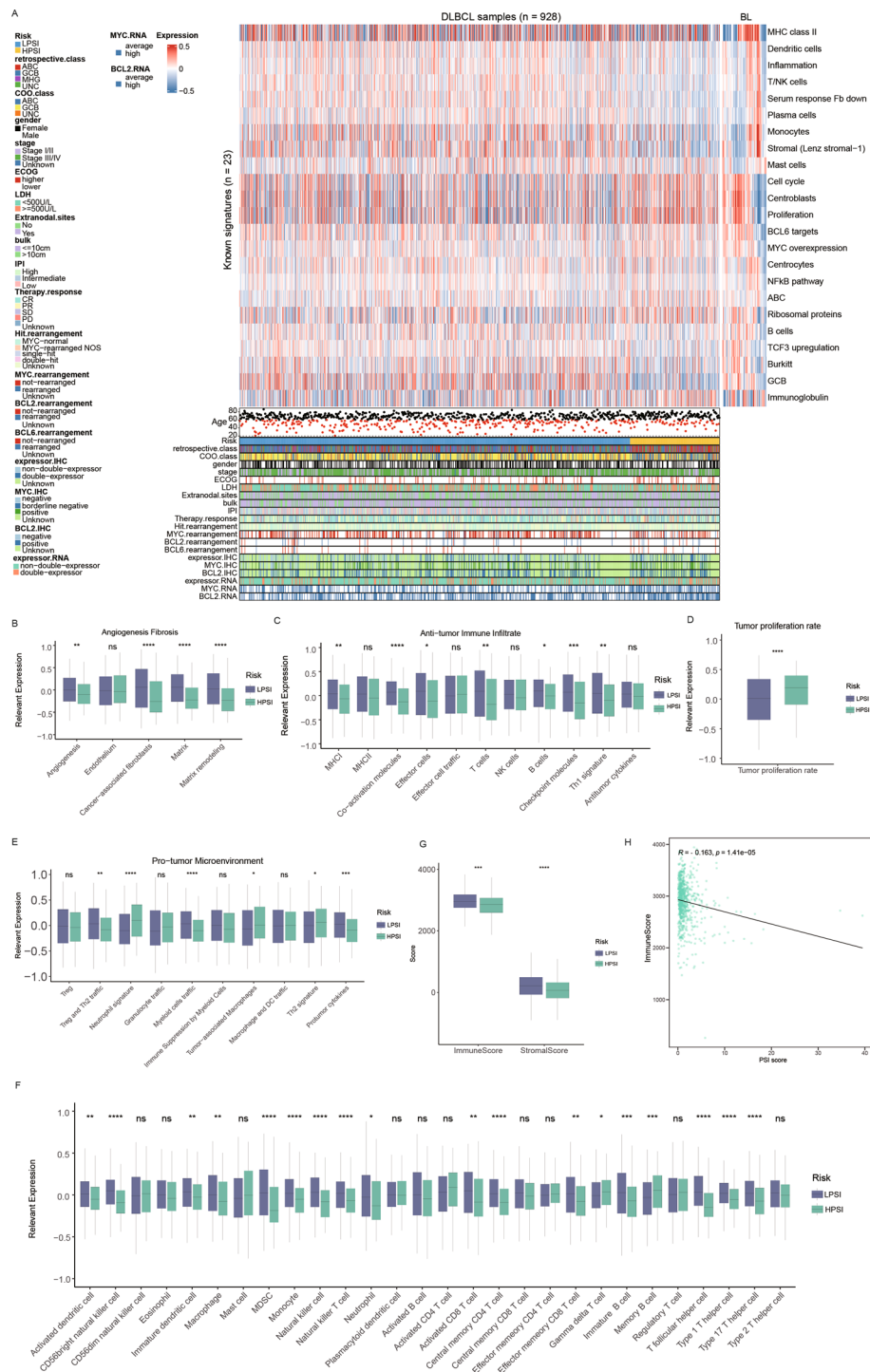


Fig. 4 (See legend on previous page.)

in HPSI-DLBCL, which may partly explain the formation of its cold immune microenvironment (Fig. 4E and S8E). And there is a decrease in the infiltration of most innate and adaptive immune cells, as well as immune and stromal scores in HPSI-DLBCL (Fig. 4F,

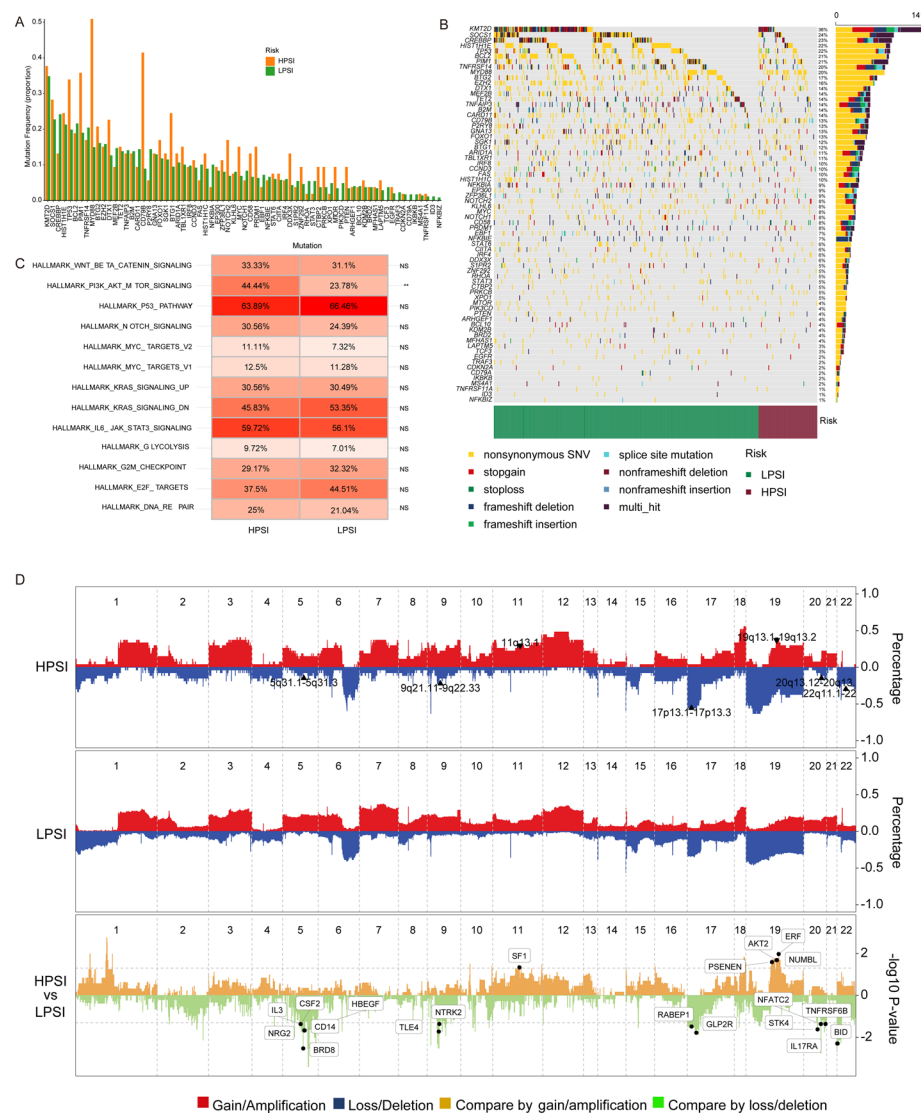


Fig. 5 Genomic changes of 17-gene-based subgroups in DLBCL. **A** Mutation frequencies for the 70-gene panel of 400 patients belonging to the LPSI and HPSI subgroups in the REMoDL-B cohort. **B** The OncoPrint displays the mutation types and their distribution for the 70 genes in LPSI and HPSI subgroups. The right-side bar chart presents the proportion of each mutation type for every gene. **C** The frequency of genomic alterations in 13 cancer hallmark pathways was examined among LPSI and HPSI DLBCL samples in the REMoDL-B cohort, with color saturation representing the frequency. **D** Somatic copy number alterations (CNAs) were compared between LPSI and HPSI subgroups. The upper and middle diagrams depict the frequency of gene loss (blue) and gain (red) in HPSI and LPSI. The lower diagram displays the $\log_{10}$ (p-value) for each gene when comparing HPSI and LPSI, using gain-centric (yellow) or loss-centric (green) calculations

G and S8F). Negative correlations were observed between PSI score and immune score (Fig. 4H).

PSI-based subgroups demonstrate different featured genomic alterations

Genomic alterations, including MYC amplification and mutations in WNT or KRAS signaling pathways, could drive immune microenvironment reprogramming in tumors [32–35]. We hypothesized that the genotypes could shape the immunophenotype of

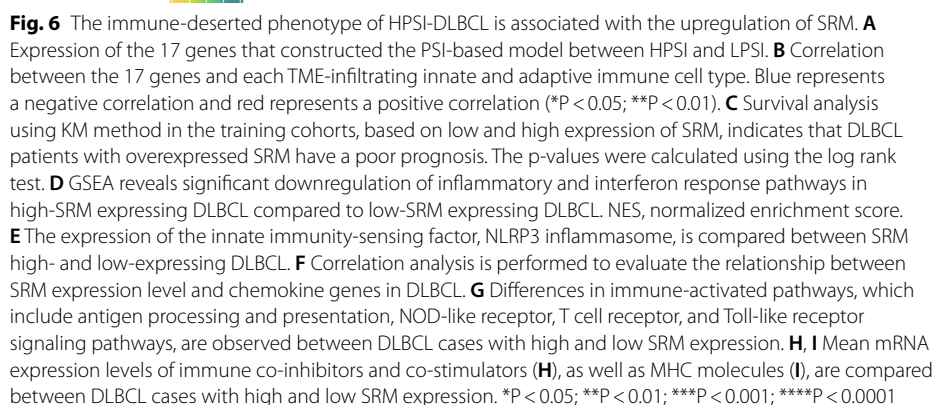
DLBCL. To test this, we further investigated the somatic mutations and CNAs between HPSI- and LPSI-DLBCLs (Fig. 5A and B). HPSI-DLBCL showed increased mutations in TP53, MYD88, PIM1, DTX1, CD79B, BTG1, MYC, PRDM1 and DDX3X. LPSI-DLBCL had mutations in CREBBP, P2RY8, CARD11, NFKBIA, STAT3, and PIK3CD, which were infrequent in HPSI-DLBCL. HPSI-DLBCL shared genomic mutations with BL [36, 37], including KMT2D, MYC, DDX3X, and CCND3, which might account for their shared proliferative phenotype. We further investigated the genomic alterations in cancer hallmark signaling pathways between the two groups. PI3K-AKT-MTOR pathway predominantly mutated in HPSI-DLBCL ($P < 0.01$) and the abnormal activation of this pathway can lead to enhanced cell proliferation and anti-apoptosis [38]. And HPSI-DLBCL exhibited higher frequency of mutations in MYC, WNT-BETA, and DNA repair pathways, although statistical significance was not observed (Fig. 5C). Consistently, we observed a similar pattern of genomic alterations in the cohort from Chapuy et al. [39], further corroborating our findings (Figures S9A and S9B).

As for somatic CNA differences, HPSI-DLBCL showed more frequent of amplification or gain of specific chromosomal regions, including 11q13.1 and 19q13.1-19q13.2 (Fig. 5D). On these chromosome regions, there were several cancer-related genes, including SF1, ERF, PSENEN, AKT2, and NUMBL. The deletion or loss of chromosomal regions 5q31.1-5q31.3, 9q21.11-9q22.33, 17p13.1-17p13.3, 20q13.12-20q13.33, and 22q11.1-22q13.32 was also more frequent in HPSI-DLBCL (Fig. 5D). On these chromosome regions, there were a significant number of genes related to antimicrobials, antigen processing and presentation, interleukins and interleukin receptors, cytokines and cytokine receptors, such as IL3, CSF2, BRD8, NRG2, HBEGF, CD14, GLP2R, RABEP1, TNFRSF6B, IL17RA, NFATC2 and BID, as well as cancer-related genes including NTRK2, STK4 and TLE4. Recent researches have confirmed that the loss of interleukin and cytokine pathway genes serves as a mechanism to suppress antitumor immunity and confer resistance to immunotherapy [40].

Conclusively, our analysis revealed that alterations in specific chromosomal regions might influence the TME reprogramming and heterogeneity of DLBCL, and HPSI-DLBCL could exhibit resistance to immune checkpoint therapy.

Upregulation of spermidine synthase (SRM) mediates the cold TME of HPSI-DLBCL

As the PSI subgroups were established based on the 17 genes, to further elucidate the mechanisms underlying HPSI-DLBCL and identify potential therapeutic targets, we compared the expression of these genes between the subgroups (Fig. 6A). SRM was overexpressed in HPSI-DLBCL and negatively correlated with the majority of immune cells (Fig. 6B). The upregulation of SRM in DLBCL is accompanied with a decrease in most innate and adaptive immune cells as well as immune scores (Figures S10A and S10B). Notably, DLBCL patients with SRM overexpression had reduced OS rates (Fig. 6C). Therefore, we focused on SRM. SRM overexpression in DLBCL can result in the suppression of interferon and inflammatory response pathways (Fig. 6D). In contrast, MYC targets, oxidative phosphorylation, glycolysis, and cell cycle-related pathways were upregulated in DLBCL patients with SRM overexpression (Figure S10C), and a positive correlation between the expression of SRM and these pathways was observed (Figure S10D-S10J). Meanwhile, SRM overexpression in DLBCL was linked to weaker



immunity activation based on comparison of innate immunity-sensing factor NLRP3 inflammasome (Fig. 6E). In addition, SRM expression was strongly anti-correlated with most chemokine genes in DLBCL (Fig. 6F), which is consistent with the downregulation of immunoactivation pathways in SRM-high expression DLBCL patients (Fig. 6G). Notably, in DLBCL cases with low SRM expression, there is an enhancement of immune activation pathways, accompanied by an increase in the expression of immune checkpoints, co-stimulatory molecules and MHC molecules (Fig. 6G–I).

Overall, our analysis suggests that SRM overexpression mediates the cold TME phenotype in DLBCL by promoting immune evasion and attenuating innate immune sensing.

Single-cell analysis revealed that SRM facilitates proliferation and immunosuppression of DLBCL tumor cells

Mechanisms of immune escape involve tumor-intrinsic and -extrinsic factors. Does SRM impact tumor cells or immune cells to contribute to immune evasion in DLBCL? We analyzed 20 single-cell RNA-sequencing (scRNA-seq) samples to decipher this question [3, 4]. Detailed data processing and analysis methods of scRNA-seq samples are described in supplementary methods. SRM expression was highest in B cells (Fig. 7A). SRM expression in DLBCL tumor cells was higher than in normal cells, while the opposite trend was observed in follicular lymphoma (FL) and transformed FL (tFL) cells (Fig. 7A). We further observed a gradual increase in SRM expression in FL, tFL, and DLBCL tumor cells (Fig. 7B). Subsequently, DLBCL tumor cells were stratified into SRM-high and SRM-low subclusters based on their SRM expression. In consistent with bulk cohort, most cancer hallmark signaling pathways were upregulated in SRM-high tumor cells, whereas interferon and inflammatory response pathways were upregulated in SRM-low tumor cells (Fig. 7C). Notably, SRM-high tumor cells exhibited greater total read counts in comparison to SRM-low tumor cells, indicating enhanced transcriptional activity (Fig. 7D). The pathways associated with glycolysis and cell cycle demonstrated enrichment, providing additional evidence for the heightened demand for transcription (Fig. 7C and E). These findings indicate that SRM overexpression might promote

(See figure on next page.)

Fig. 7 Influences of SRM upregulation on DLBCL tumor cells. **A** Comparison of SRM expression in different cell types across 20 scRNA-seq samples and between tumor (malignant) and normal (non-malignant) B cells in various B cell lymphomas. Mac, macrophage cell; Mono, monocyte cell; DC, dendritic cell. **B** Comparison of SRM expression in tumor cells (malignant B cells) among DLBCL, FL and tFL. **C** Heatmap show GSEA normalized enrichment scores (NES) (blue = low to red = high) of cancer hallmark signatures in SRM high- versus low-expressing tumor cells of DLBCL. **D** Boxplots showing the increased transcript count in SRM high-expressing tumor cells of DLBCL. **E** Heatmap shows GSEA NES of the cell cycle, chemokines, and stem-cell associated pathways in SRM high- and low-expressing tumor cells of DLBCL. **F** Heatmap displays the average expression of genes involved in immune activation and immune cell chemotaxis pathways between SRM high- and low-expressing tumor cells. **G** Heatmap of the area under the curve (AUC) scores (green = low to red = high) of expression regulation by transcription factors, as estimated using SCENIC. Shown are the transcription factors having the higher difference in expression regulation estimates between SRM high—and low-expressing tumor cells of DLBCL. **H** tSNE analysis for SRM high- and low-expressing tumor cells, for the expression of MYC, SMARCA4, NFATC1, TCF3, FOS, JUN, JUNB and HDAC2, and for the AUC of the estimated regulon activity of these transcription factors, corresponding to the degree of expression regulation of their target genes. **I** Violin plots showing the GSVA enrichment scores of stemness signatures in SRM high- versus low-expressing tumor cells of DLBCL. **J** Survival analysis demonstrated that SRM high-expressing tumor cells were associated with poor prognosis of DLBCL

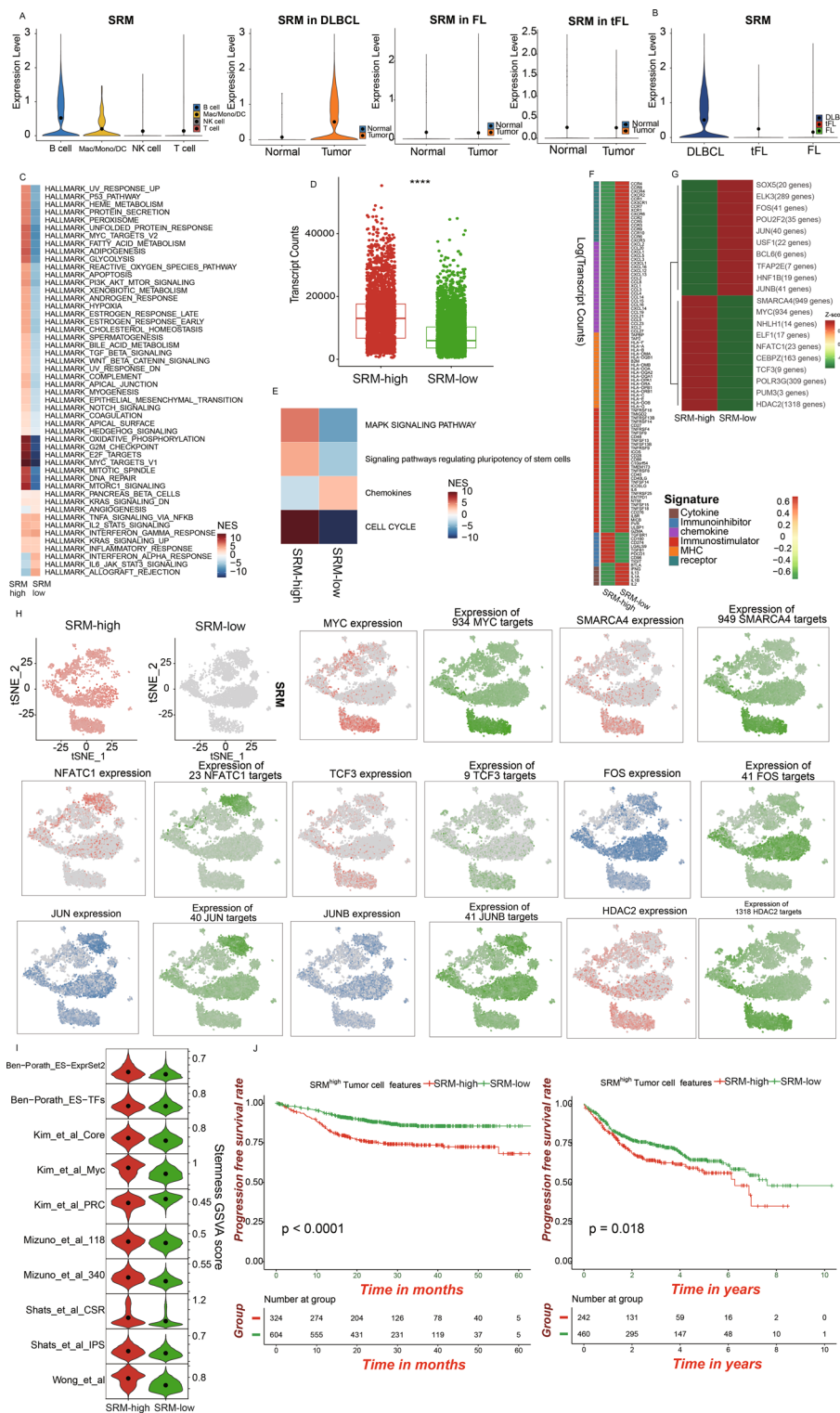


Fig. 7 (See legend on previous page.)

the proliferation of tumor cells, underlining a potential vulnerability of tumor cells to SRM inhibition in DLBCL. In contrast, chemokine pathway was downregulated in SRM-high tumor cells (Fig. 7E). A more comprehensive analysis uncovered downregulation

of genes associated with immune cell chemotaxis and immune activation in SRM-high tumor cells (Fig. 7F). These data suggest that tumor cells with SRM upregulation are remodeled, thereby downregulating their antigen presentation and immune cell chemotaxis activities to facilitate immune tolerance.

Additionally, to explore the transcriptional regulatory mechanisms underlying the differences between high- and low-SRM tumor cells, we performed single-cell regulatory network inference and clustering (SCENIC) analysis. The downregulation of FOS/JUN/JUNB and upregulation of TCF3, SMARCA4, NFATC1, HDAC2, NHLH1 and MYC seemed responsible for tumor-specific SRM-high cell phenotypes (Fig. 7G and H). These upregulated regulators have been reported to be associated with cell proliferation and differentiation [41–45].

Of interest, genes implicated in immune activation showed FOS/JUN/JUNB binding sites, implying that FOS/JUN/JUNB deletion underlies the reduced immune stimulatory phenotype of SRM-high tumor cells. We detected elevated levels of stemness signatures in SRM-high tumor cells (Fig. 7I) and found upregulated stem-cell-related pathways, such as PI3K-AKT, TGF-beta, and signaling pathways regulating pluripotency of stem cells (Fig. 7C and E), indicating a stem-cell-like characteristic was exhibited in SRM-high tumor. Moreover, SRM-high tumor cells were linked to unfavorable outcomes in DLBCL (Fig. 7J).

Effects of SRM upregulation on immune cells within TME

Dendritic cells (DCs) play a crucial role in the activation and recruitment of T cells [46, 47]. Using the same method, we classified DCs into SRM-high and SRM-low expressing groups. The total read counts in SRM-high DCs were increased (Figure S11A). We observed a significant reduction in the expression of co-stimulatory and MHC molecules, as well as chemokines, in SRM-high DCs (Figures S11B and S11C). These results suggest a diminished ability of SRM-high DCs to activate T cells and present antigens. These findings align with the observations in Bulk cohort (Fig. 6H, I and S10A). Functional enrichment analysis revealed that SRM-high DCs displayed upregulation of cancer hallmark pathways (Figure S11D). SRM-high DCs exhibited increased expression of multiple metabolic pathways, including fatty acid metabolism, nucleotide metabolism, the TCA cycle, glycolysis and oxidative phosphorylation (Figure S11E). Previous research has demonstrated that the shift from glucose dependence to Fatty Acyl-CoA Oxidase (FAAO) and the upregulation of various metabolic pathways contribute to the immunosuppressive phenotype in DCs [48]. Moreover, heightened glycolysis activity promotes the maturation resistance and de-differentiation of immune-suppressive tolerogenic DCs (tol-DCs) [48]. In SRM-high DCs, the expression of multiple markers of tol-DCs, including CCR7, CD14, CD141 (THBD), and CD274 (PD-L1), were upregulated (Figure S11B). Together, these findings indicate that SRM-high DCs demonstrate an anti-inflammatory tol-DC phenotype.

We conducted re-clustering on macrophages and identified two distinct clusters: C0 and C1 (Fig. 8A). C1 exhibited significantly higher expression of SRM compared to C0 (Fig. 8B). Functional enrichment analysis revealed that C0 displayed stronger activity of MHC protein binding and Toll-like receptor binding, while C1 showed enhanced ribosome metabolism and rRNA binding activity (Fig. 8C). Moreover, C1 exhibited a

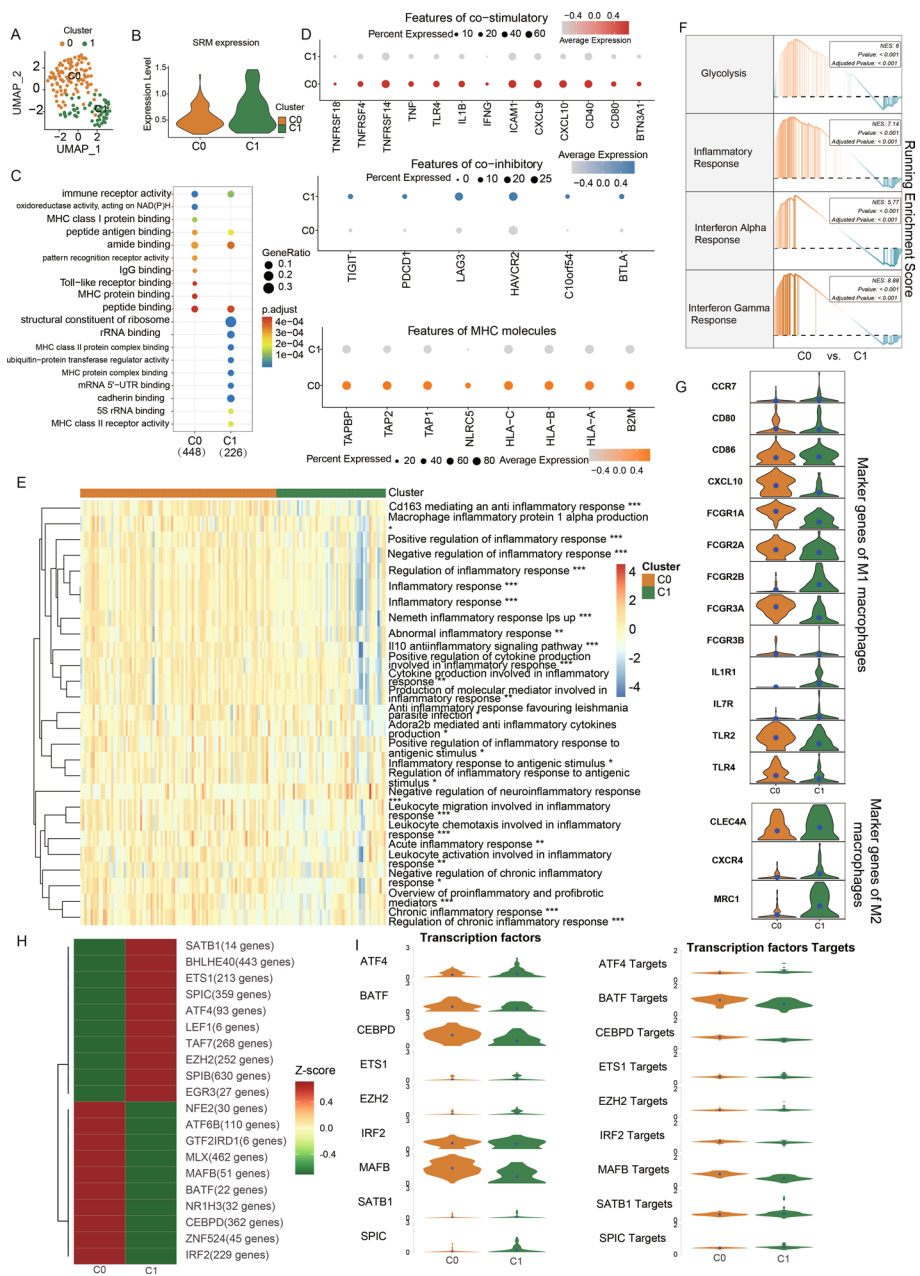


Fig. 8 Overexpression of SRM induces an M2-like anti-inflammatory phenotype in macrophages. **A** Reclustering of macrophages in single-cell samples of DLBCL. **B** Violin plot show the increased mRNA level of SRM in C1 macrophages compared to C0 macrophages. **C** GO enrichment analysis shows that C0 exhibited stronger immune activity, while C1 showed stronger ribosome metabolism and rRNA binding activity. **D** Dot plot shows the expression of marker genes of co-stimulatory and MHC molecules, as well as co-inhibitory in C0 and C1 macrophages. **E** Heatmap shows GSVA enrichment scores of immune response pathways in C0 and C1 macrophages. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$. **F** GSEA of glycolysis, inflammatory and interferon response pathways in C0 and C1 macrophages. NES, normalized enrichment score. Adjusted P value represents false discovery rate (FDR). **G** Violin plots show the expression of marker genes of M1 and M2 macrophages in C0 and C1 macrophages. **H** Heatmap of the area under the curve (AUC) scores (green = low to red = high) of expression regulation by transcription factors, as estimated using SCENIC. Shown are the transcription factors having the higher difference in expression regulation estimates between C0 and C1 macrophages of DLBCL. **I** Violin plots of C0 and C1 macrophages, for the expression of SPIC, SATB1, MAFB, IRF2, EZH2, ETS1, CEBPD, BATF and ATF4 (left), and for the AUC of the estimated regulon activity of these transcription factors (right), corresponding to the degree of expression regulation of their target genes

decrease in the expression of co-stimulatory molecules, MHC molecules, and immune response pathways, along with an increase in co-inhibitory molecule expression (Fig. 8D and E), indicating an anti-inflammatory phenotype. Conversely, C0 exhibited a pro-inflammatory phenotype and were significantly enriched in interferon and inflammatory response pathways, as well as glycolysis (Fig. 8F). Previous studies have demonstrated that macrophages with upregulated glycolysis exhibit enhanced phagocytic capacity and an M1-like proinflammatory phenotype [49, 50]. Additionally, C1 displayed an M2-like immunosuppressive phenotype, characterized by the significant downregulation of M1 macrophage markers (CXCL10, TLR2, FCGR1A, FCGR2A, FCGR3A, and TLR4) and upregulation of M2 macrophage markers (CLEC4A, CXCR4 and MRC1) (Fig. 8G).

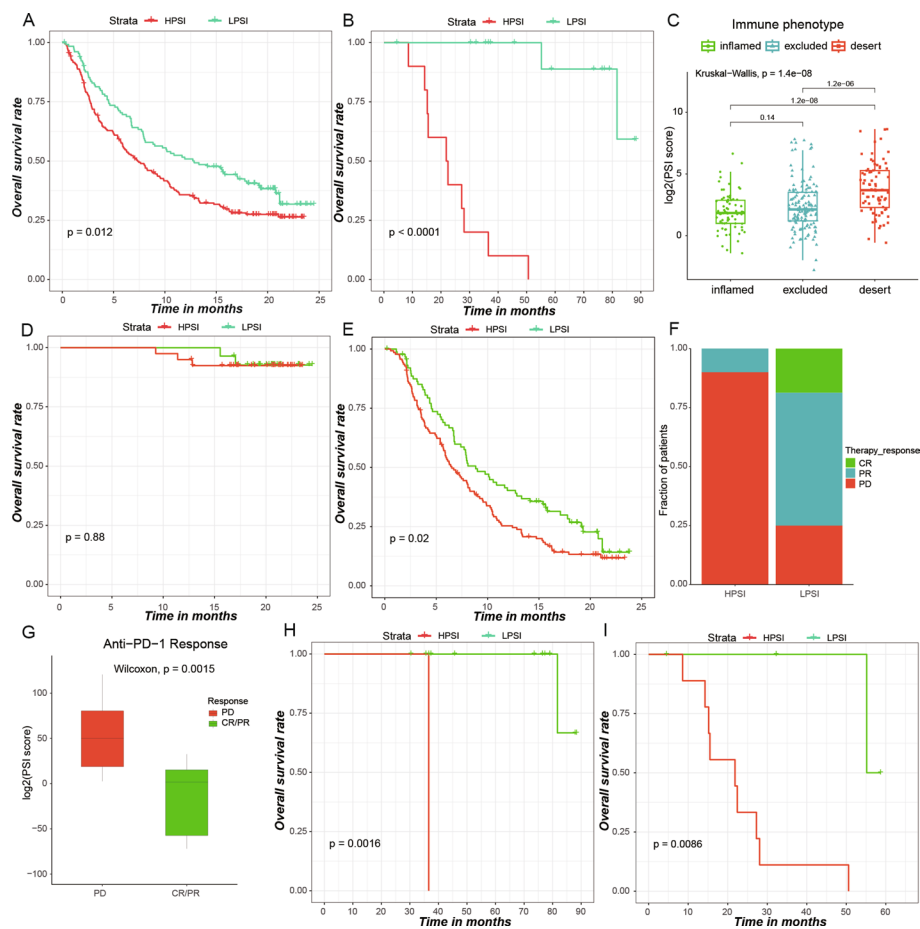


Fig. 9 Role of the 17-gene-based model in anti-PD-1/L1 immunotherapy. **A** Survival analysis for low and high PSI scores patients in the anti-PD-L1 immunotherapy cohort using KM curves (IMvigor210 cohort; log-rank test, $P = 0.012$). **B** KM survival analysis for low and high PSI scores groups in the anti-PD1 immunotherapy cohort (GSE78220 cohort; log-rank test, $P < 0.0001$). **C** Differences in PSI scores among patients with different immune phenotypes in the IMvigor210 cohort. **D, E** In the IMvigor210 cohort, KM survival curves for low and high PSI scores patients within (D) the complete response (CR)/partial response (PR) group (log-rank test, $P = 0.88$), and (E) patients within the stable disease (SD)/progressive disease (PD) group (log-rank test, $P = 0.02$). **F** The percentage of patients exhibiting different treatment responses in the high and low PSI scores groups of the GSE78220 cohort showed a significant difference ($P = 0.0053$, two-tailed Fisher's exact test). **G** Difference of PSI scores between different clinical response groups in the GSE78220 cohort ($P = 0.0015$, Wilcoxon test). **H, I** In the GSE78220 cohort, KM survival curves for low and high PSI scores patients within (H) the CR/PR group (log-rank test, $P = 0.0016$) and (I) patients within the PD group (log-rank test, $P = 0.0086$).

SCENIC analysis also showed that regulators such as SATB1, ETS1, ATF4 and EZH2 associated with cell differentiation and promoting cancer development and progression in C1 were enriched compared to C0, while regulators that suppress tumors such as IRF2, BATF and CEBPD were significantly downregulated (Fig. 8H and I) [51–53].

Evaluation of immunotherapy response as well as nomogram development and validation based on the 17-gene model

Based on the above findings, we speculate that the 17-gene model can assess the efficacy of immunotherapy in tumor patients. Although we were unable to access publicly available data of immunotherapy specifically for DLBCL, we validated the utility of the 17-gene model for assessing immunotherapy response and prognosis in cohorts treated with different immune checkpoint inhibitors for two solid tumors. Low PSI score was associated with greater clinical benefit and prolonged survival in both anti-PD-L1 and anti-PD-1 cohorts (Fig. 9A and B). The immune phenotypes of patients in the IMvigor210 cohort has been detected, so we explored the difference of PSI score among different phenotypes. We found that higher PSI score was remarkably associated with desert and exclusion immune phenotypes, and checkpoint inhibitors were difficult to exert antitumor effect in these phenotype (Fig. 9C). And similarly to the results observed in DLBCL cohorts, 17-gene model exhibited significant prognostic value that remained independent of the clinical response stratification in immunotherapy cohorts (Fig. 9D–I).

Finally, we developed an integrated nomogram comprising PSI score and clinico-pathological features for DLBCL (Figure S12A). Receiver operating characteristic (ROC) curve analysis confirmed that the combined nomogram offers a robust prognostic assessment for DLBCL patients, outperforming age, stage, subtypes, and Eastern Cooperative Oncology Group (ECOG) performance status alone in both the training and validation cohorts (internal and external). Additionally, the area under the curve (AUC) values demonstrate that the 17-gene model surpasses the predictive ability of the IPI (Figures S12B–S12E).

Discussion

Previous studies have shown that TME plays an important role in the pathogenesis of DLBCL [4], so we focused on the gene expression profile that is closely related to the three main cell types (cell proliferation, stroma, and immunity) [54] in TME, which can provide a more personalized risk assessment for DLBCL.

In this study, we developed a unique risk stratification system for DLBCL and assessed its clinical relevance in disease prognosis and biology. Our 17-gene model exhibited a robust ability to stratify risk, ensuring stable prognostic stratification of diverse DLBCL cohorts across international multiple studies using a consistent cutoff value. We defined HPSI-DLBCL as a spermidine metabolism-cold immune tumor with upregulated SRM and desert-like immune infiltration, which showed poor prognosis with lower 3- and 5-year OS/PFS rates than LPSI-DLBCL. Furthermore, this new model demonstrates even better performance in prognostic stratification than the traditional classification based on the COO [31, 55]. Both LPSI- and HPSI-DLBCL included MHG, GCB, ABC, and UNC subtypes, but they exhibited varied prognosis within the same subtype,

indicating that the 17-gene model identify previously undetected biological features. Moreover, the 17-gene model effectively stratifies patients within high-risk categories, such as those with MYC rearrangements, double-expressors, and MHG group. Specifically, our model highlights the necessity for a more nuanced approach to risk assessment and management, even among patients who are currently categorized as high-risk based on existing criteria. This has important implications for optimizing therapeutic strategies, as it suggests that not all patients with high-risk features (like MYC rearrangements or double-expression status) uniformly experience poor outcomes. We recognize the heterogeneity in high-grade lymphomas, such as DHL/THL. However, when we analyzed survival outcomes in these patients, our 17-gene model did not demonstrate statistically significant stratification. This is likely attributable to the small sample size of DHL/THL cases, with only 36 patients (9 in the HPSI group and 27 in the LPSI group). This limitation highlights an important area for future research, and further studies should focus on validating the model's risk stratification capacity in these specific subgroups in well-designed clinical trials.

Previous studies have predominantly focused on immune cells, stromal components, or tumor cells individually, without integrating these three crucial factors to comprehensively evaluate DLBCL treatment response and prognosis, or to develop a robust prognostic scoring system. The 17-gene based PSI score, which incorporates tumor proliferation, stromal, and immune characteristics, effectively captures the heterogeneity of the TME and reflects both intrinsic and extrinsic factors influencing DLBCL development, progression, and prognosis. This is also why we did not choose a simpler scoring system; simpler models typically fail to consider all three factors simultaneously, thereby inadequately capturing the comprehensive influence of both intrinsic and extrinsic tumor characteristics on disease outcomes. As a result, simpler models have often failed to provide consistent prognostic risk stratification across multiple independent cohorts. In contrast, the 17-gene model offers superior stability and robustness, enabling reliable prognostic risk stratification across cohorts from different research centers using the same cutoff value. A reliable cutoff value for risk stratification is essential for the clinical application of a scoring system. The robust cutoff value provides a standard for further clinical validation and implementation of the 17-gene scoring system.

Transcriptional profile largely determines the biological behaviors and molecular subtypes of DLBCL and can reflect the immune infiltrate pattern of DLBCL. By performing transcriptional profiling and genomic alteration analyses, we have observed that HPSI-DLBCL displays characteristics of malignant proliferative phenotype similar to BL and desert-like immune phenotype. Furthermore, HPSI-DLBCL exhibited a higher frequency of chromosomal region alterations. The loss and deletion of immunomodulatory genes on specific chromosome regions might contribute to the cold TME. In contrast, gains or amplifications of chromosomal regions 11q13.1 and 19q13.1-19q13.2 could potentially promote tumor progression by influencing specific cancer-related genes. These data suggest the potential role of the 17-gene risk model in immunotherapy. Aligned with our hypothesis, we found a consistent association in the anti-PD-L1 and anti-PD-1 cohorts, where tumor patients with elevated PSI scores exhibited a worse prognosis. These

findings underscore the robust applicability of the 17-gene model, effectively evaluating the response and prognosis of immune therapy among cancer patients.

As a proto-oncogene, MYC can contribute to multiple cancer hallmarks, such as self-renewal, proliferation, genomic instability, metabolism as well as immune evasion. However, despite decades of effort and numerous experiments, MYC has not yet been successfully targeted therapeutically due to the lack of a binding pocket or specific enzymatic activity and the highly disordered protein structure. An important therapeutic approach for treating tumors with MYC overexpression is to explore the regulation of crucial molecules upstream or downstream of MYC oncogenic pathway. Insufficient research has been conducted on the role of SRM in DLBCL, despite its significance as a vital enzyme in spermidine metabolism and a downstream target of MYC [56, 57]. Our investigation, leveraging multi-omics data, revealed the potential role of SRM to drive aggressive proliferation and induce immunosuppressive TME in DLBCL. Importantly, by identifying HPSI-DLBCL, targeting SRM may improve the immune therapy response and help resolve the issue of low overall efficacy rate in immune therapy for DLBCL patients. This offers new hope for patients with refractory or relapsed DLBCL.

Although our study did not directly assess the pathological structure of the TME in DLBCL, we indirectly reflected the TME composition of DLBCL using 23 gene sets related to B-cell disorders from the normal and pathological lymphoid biology database, 28 gene sets representing innate and adaptive immune cells, and typical gene enrichment analysis algorithms. Our findings showed that LPSI-DLBCL was enriched with normal germinal center (GC) microenvironmental cells—such as macrophages/dendritic cells, follicular T cells, and stromal cells—while these cells were lacking in HPSI-DLBCL. Consistent with our findings, Miyawaki et al. employed gene expression profiling and multispectral fluorescence imaging to demonstrate that the presence of normal GC microenvironmental cells—such as macrophages/dendritic cells, follicular T cells, and stromal cells—within lymphoma tissues is indicative of a favorable therapeutic response. Conversely, DLBCL tissues lacking GC-microenvironmental cells exhibited a higher frequency of genomic alterations and dysregulated gene expression patterns associated with poor prognosis [58]. These observations provide additional pathological evidence supporting our findings.

Finally, we developed a user-friendly tool, which integrates multi-omics datasets, the 17-gene model, and clinicopathological features to generate a comprehensive portrait of DLBCL for clinical decision support. The tool is accessible in real-time and aims to provide an intuitive interface to facilitate the analysis and interpretation of complex data. The tool can be accessed at <http://yxszxssx.com:3838/psidb/> or <http://shiny.yxszxssx.net:3838/psidb/>.

Conclusions

In summary, our study employed a thorough and systematic approach to biomarker discovery and validation. Our results highlight a novel 17-gene PSI signature, which effectively stratifies the risk of DLBCL and reveals that poor prognosis in a subset of patients is associated with dysregulated spermidine metabolism and a cold TME. Targeting SRM

could potentially enhance the efficacy of immune therapy and improve the prognosis for DLBCL patients, thus providing novel insights for immunotherapy strategies. Further research will focus on the functional interpretation of our results and validation of our findings in well-designed clinical trials.

Abbreviations

DLBCL	Diffuse large B-cell lymphoma
PSI	Proliferation, stromal, and immune
TME	Tumor microenvironment
SRM	Spermidine synthase
ECM	Extracellular matrix
OS	Overall survival
PFS	Progression-free survival
GSVA	Gene set variation analysis
GSEA	Gene set enrichment analysis
NES	Normalized enrichment score
DIRAC	Differential rank conservation
RCIs	Rank conservation indices
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone
LASSO	Least absolute shrinkage and selection operator
AIC/BIC	Akaike and Bayesian information criteria
CNV	Copy number variation
CNAs	Copy number alteration
IPI	International Prognostic Index
COO	Cell-of-origin
ABC	Activated B-cell-like
UNC	Unclassified
MHG	Molecular high-grade
GCB	Germinal Center B-cell-like
EZB subtype	Based on EZH2 mutations and BCL2 translocations
BN2 subtype	Based on BCL6 fusions and NOTCH2 mutations
MCD subtype	Based on the co-occurrence of MYD88L265P and CD79B mutations
N1 subtype	Based on NOTCH1 mutations
CR	Complete response
PR	Partial response
SD	Stable disease
PD	Progressive disease
DHL/THL	Double-hit or triple-hit lymphoma
SCENIC	Single-cell regulatory network inference and clustering
ECOG	Eastern Cooperative Oncology Group

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40537-025-01067-z>.

Additional file 1.
Additional file 2.
Additional file 3.

Acknowledgements

Thank all the datasets' owners who made the valuable data public.

Author contributions

L.W designed the study, interpreted data, and revised the manuscript. X.J.L, J.G, and B.W.L performed statistical analysis, interpreted data, and wrote the manuscript. W.X.L, Q.M.C, Y.L.D, and Y.N.Y provided bioinformatic data analysis and contributed to manuscript revision. All authors have read and approved the final version of the manuscript.

Funding

This work was supported by grants from the National Natural Science Foundation of China (Nos. 82170181), Beijing Natural Science Foundation (No. 7222027), Beijing Physician Scientist Training Project (BJPSTP-2024-01), and the National Key R&D Program of China (Grant No. 2022YFF1502000) to Liang Wang.

Availability of data and materials

All the datasets could be downloaded directly from the indicated websites. Datasets and custom scripts are available upon request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 5 November 2023 Accepted: 9 January 2025

Published online: 23 January 2025

References

- Baecklund E, Backlin C, Iliadou A, Granath F, Ekblom A, Amini RM, et al. Characteristics of diffuse large B cell lymphomas in rheumatoid arthritis. *Arthritis Rheum*. 2006;54(12):3774–81.
- Wang L, Li LR, Young KH. New agents and regimens for diffuse large B cell lymphoma. *J Hematol Oncol*. 2020;13(1):175.
- Roider T, Seufert J, Uvarovskii A, Frauhammer F, Bordas M, Abedpour N, et al. Dissecting intratumour heterogeneity of nodal B-cell lymphomas at the transcriptional, genetic and drug-response levels. *Nat Cell Biol*. 2020;22(7):896–906.
- Steen CB, Luca BA, Esfahani MS, Azizi A, Sworder BJ, Nabet BY, et al. The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. *Cancer Cell*. 2021;39(10):1422–37.e10.
- Xu B, Wang T. Intimate cross-talk between cancer cells and the tumor microenvironment of B-cell lymphomas: the key role of exosomes. *Tumour Biol*. 2017;39(6):1010428317706227.
- Höpken UE, Rehm A. Targeting the tumor microenvironment of leukemia and lymphoma. *Trends Cancer*. 2019;5(6):351–64.
- Kumar D, Xu ML. Microenvironment cell contribution to lymphoma immunity. *Front Oncol*. 2018;8:288.
- Scott DW, King RL, Staiger AM, Ben-Neriah S, Jiang A, Horn H, et al. High-grade B-cell lymphoma with MYC and BCL2 and/or BCL6 rearrangements with diffuse large B-cell lymphoma morphology. *Blood*. 2018;131(18):2060–4.
- Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, et al. Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med*. 2008;359(22):2313–23.
- Kotlov N, Bagaev A, Revuelta MV, Phillip JM, Cacciapuoti MT, Antysheva Z, et al. Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures. *Cancer Discov*. 2021;11(6):1468–89.
- Visco C, Li Y, Xu-Monette ZY, Miranda RN, Green TM, Li Y, et al. Comprehensive gene expression profiling and immunohistochemical findings support application of immunophenotypic algorithm for molecular subtype classification in diffuse large B-cell lymphoma: a report from the International DLBCL Rituximab-CHOP Consortium Program Study. *Leukemia*. 2012;26(9):2103–13.
- Dubois S, Viailly PJ, Bohers E, Bertrand P, Ruminy P, Marchand V, et al. Biological and clinical relevance of associated genomic alterations in MYD88 L265P and non-L265P-mutated diffuse large B-cell lymphoma: analysis of 361 cases. *Clin Cancer Res*. 2017;23(9):2232–44.
- Sha C, Barrans S, Cucco F, Bentley MA, Care MA, Cummin T, et al. Molecular high-grade B-cell lymphoma: defining a poor-risk group that requires different approaches to therapy. *J Clin Oncol*. 2019;37(3):202–12.
- Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. *N Engl J Med*. 2018;378(15):1396–407.
- Lacy SE, Barrans SL, Beer PA, Painter D, Smith AG, Roman E, et al. Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. *Blood*. 2020;135(20):1759–71.
- Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160(1–2):48–61.
- Bian B, Bigonnet M, Gayet O, Loncle C, Maignan A, Gilabert M, et al. Gene expression profiling of patient-derived pancreatic cancer xenografts predicts sensitivity to the BET bromodomain inhibitor JQ1: implications for individualized medicine efforts. *EMBO Mol Med*. 2017;9(4):482–97.
- Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SGH, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet*. 2015;47(10):1168–78.
- Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*. 2013;14:7.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–50.
- Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–7.
- Hu J, Locasale JW, Bielas JH, O'Sullivan J, Sheahan K, Cantley LC, et al. Heterogeneity of tumor-induced gene expression changes in the human metabolic network. *Nat Biotechnol*. 2013;31(6):522–9.
- Eddy JA, Hood L, Price ND, Geman D. Identifying tightly regulated and variably expressed networks by Differential Rank Conservation (DIRAC). *PLoS Comput Biol*. 2010;6(5): e1000792.
- Shaffer AL, Wright G, Yang L, Powell J, Ngo V, Lamy L, et al. A library of gene expression signatures to illuminate normal and pathological lymphoid biology. *Immunol Rev*. 2006;210:67–85.

25. Bagaev A, Kotlov N, Nomie K, Svekolkin V, Gafurov A, Isaeva O, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell*. 2021;39(6):845–65.e7.
26. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep*. 2017;18(1):248–62.
27. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and transcriptomic features of response to Anti-PD-1 therapy in metastatic melanoma. *Cell*. 2016;165(1):35–44.
28. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGF β attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature*. 2018;554(7693):544–8.
29. Hazra A, Gogtay N. Biostatistics series module 3: comparing groups: numerical variables. *Indian J Dermatol*. 2016;61(3):251–60.
30. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature*. 2000;403(6769):503–11.
31. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med*. 2002;346(25):1937–47.
32. Jiménez-Sánchez A, Cybulska P, Mager KL, Koplev S, Cast O, Couturier DL, et al. Unraveling tumor-immune heterogeneity in advanced ovarian cancer uncovers immunogenic effect of chemotherapy. *Nat Genet*. 2020;52(6):582–93.
33. Kortlever RM, Sodir NM, Wilson CH, Burkhart DL, Pellegrinet L, Brown Swigart L, et al. Myc cooperates with ras by programming inflammation and immune suppression. *Cell*. 2017;171(6):1301–15.e14.
34. Casey SC, Tong L, Li Y, Do R, Walz S, Fitzgerald KN, et al. MYC regulates the antitumor immune response through CD47 and PD-L1. *Science*. 2016;352(6282):227–31.
35. Meyer N, Penn LZ. Reflecting on 25 years with MYC. *Nat Rev Cancer*. 2008;8(12):976–90.
36. Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature*. 2012;490(7418):116–20.
37. Love C, Sun Z, Jima D, Li G, Zhang J, Miles R, et al. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet*. 2012;44(12):1321–5.
38. Engelman JA. Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat Rev Cancer*. 2009;9(8):550–62.
39. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med*. 2018;24(5):679–90.
40. Eissmann MF, Dijkstra C, Wouters MA, Baloyan D, Mouradov D, Nguyen PM, et al. Interleukin 33 signaling restrains sporadic colon cancer in an interferon- γ -dependent manner. *Cancer Immunol Res*. 2018;6(4):409–21.
41. Somasundaram R, Prasad MA, Ungerback J, Sigvardsson M. Transcription factor networks in B-cell differentiation link development to acute lymphoid leukemia. *Blood*. 2015;126(2):144–52.
42. Li L, Bhatia R. Stem cell quiescence. *Clin Cancer Res*. 2011;17(15):4936–41.
43. Tian Y, Xu L, Li X, Li H, Zhao M. SMARCA4: current status and future perspectives in non-small-cell lung cancer. *Cancer Lett*. 2023;554: 216022.
44. West AC, Johnstone RW. New and emerging HDAC inhibitors for cancer treatment. *J Clin Invest*. 2014;124(1):30–9.
45. Wadman I, Li J, Bash RO, Forster A, Osada H, Rabbitts TH, et al. Specific in vivo association between the bHLH and LIM proteins implicated in human T cell leukemia. *Embo j*. 1994;13(20):4831–9.
46. Banchereau J, Steinman RM. Dendritic cells and the control of immunity. *Nature*. 1998;392(6673):245–52.
47. Qian C, Cao X. Dendritic cells in the regulation of immunity and inflammation. *Semin Immunol*. 2018;35:3–11.
48. Adamik J, Munson PV, Hartmann FJ, Combes AJ, Pierre P, Krummel MF, et al. Distinct metabolic states guide maturation of inflammatory and tolerogenic dendritic cells. *Nat Commun*. 2022;13(1):5184.
49. Li S, Yu J, Huber A, Kryczek I, Wang Z, Jiang L, et al. Metabolism drives macrophage heterogeneity in the tumor microenvironment. *Cell Rep*. 2022;39(1): 110609.
50. Wculek SK, Heras-Murillo I, Mastrangelo A, Mañanes D, Galán M, Miguel V, et al. Oxidative phosphorylation selectively orchestrates tissue macrophage homeostasis. *Immunity*. 2023;56(3):516–30.e9.
51. Pan YC, Li CF, Ko CY, Pan MH, Chen PJ, Tseng JT, et al. CEBPD reverses RB/E2F1-mediated gene repression and participates in HMDB-induced apoptosis of cancer cells. *Clin Cancer Res*. 2010;16(23):5770–80.
52. Kayagaki N, Lee BL, Stowe IB, Kornfeld OS, O'Rourke K, Mirrashidi KM, et al. IRF2 transcriptionally induces GSDMD expression for pyroptosis. *Sci Signal*. 2019;12(582).
53. Delacher M, Simon M, Sanderink L, Hotz-Wagenblatt A, Wuttke M, Schambeck K, et al. Single-cell chromatin accessibility landscape identifies tissue repair program in human regulatory T cells. *Immunity*. 2021;54(4):702–20.e17.
54. Kandimalla R, Tomihara H, Banwait JK, Yamamura K, Singh G, Baba H, et al. A 15-gene immune, stromal, and proliferation gene signature that significantly associates with poor survival in patients with pancreatic ductal adenocarcinoma. *Clin Cancer Res*. 2020;26(14):3641–8.
55. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103(1):275–82.
56. Fernandez PC, Frank SR, Wang L, Schroeder M, Liu S, Greene J, et al. Genomic targets of the human c-Myc protein. *Genes Dev*. 2003;17(9):1115–29.
57. Forshell TP, Rimpf S, Nilsson JA. Chemoprevention of B-cell lymphomas by inhibition of the myc target spermidine synthase. *Cancer Prev Res*. 2010;3(2):140–7.
58. Miyawaki K, Kato K, Sugio T, Sasaki K, Miyoshi H, Semba Y, et al. A germinal center-associated microenvironmental signature reflects malignant phenotype and outcome of DLBCL. *Blood Adv*. 2022;6(7):2388–402.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.