REGULAR ARTICLE

blood advances

Biological signatures of the International Prognostic Index in diffuse large B-cell lymphoma

Yue Wang,[1,*] Qing Shi,[1,*] Zi-Yang Shi,[1,*] Shuang Tian,[1,*] Mu-Chen Zhang,[1,*] Rong Shen,[1] Di Fu,[1] Lei Dong,[2] Hong-Mei Yi,[2] Bin-Shen Ouyang,[2] Rong-Ji Mu,[4] Shu Cheng,[1] Li Wang,[1,3] Peng-Peng Xu,[1] and Wei-Li Zhao[1,3]

[1]Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics; National Research Center for Translational Medicine at Shanghai, and [2]Department of Pathology, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China; [3]Pôle de Recherches Sino-Français en Science du Vivant et Génomique, Laboratory of Molecular Pathology, Shanghai, China; and [4]Clinical Research Institute, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Key Points

- IPI risk factors are associated with different genetic and tumor microenvironmental signatures in DLBCL.
- Combining the IPI with multi-omics information may provide more precise treatment rationale for patients with DLBCL.

Diffuse large B-cell lymphoma (DLBCL) is a highly aggressive subtype of lymphoma with clinical and biological heterogeneity. The International Prognostic Index (IPI) shows great prognostic capability in the era of rituximab, but the biological signatures of IPI remain to be discovered. In this study, we analyzed the clinical data in a large cohort of 2592 patients with newly diagnosed DLBCL. Among them, 1233 underwent DNA sequencing for oncogenic mutations, and 487 patients underwent RNA sequencing for lymphoma microenvironment (LME) alterations. Based on IPI scores, patients were categorized into 4 distinct groups, with 5-year overall survival of 41.6%, 55.3%, 71.7%, and 89.7%, respectively. MCD-like subtype was associated with age of >60 years, multiple extranodal involvement, elevated serum lactate dehydrogenase (LDH), and IPI scores ranging from 2 to 5, whereas ST2-like subtype showed an opposite trend. Patients with EZB-like MYC^{+} and $TP53^{Mut}$ subtypes exhibited poor clinical outcome independent of the IPI; integrating $TP53^{Mut}$ into IPI could better distinguish patients with dismal survival. The EZB-like MYC^{-}, BN2-like, N1-like, and MCD-like subtypes had inferior prognosis in patients with IPI scores of ≥2, indicating necessity for enhanced treatment. Regarding LME categories, the germinal center–like LME was more prevalent in patients with normal LDH and IPI scores of 0 to 1. The mesenchymal LME served as an independent protective factor, whereas the germinal center–like, inflammatory, and depleted LME categories correlated with inferior prognosis for IPI scores of 2 to 5. In summary, our work explored the biological signatures of IPI, thus providing useful rationale for future optimization of the IPI-based treatment strategies with multi-omics information in DLBCL.

Introduction

Diffuse large B-cell lymphoma (DLBCL) represents the most common subtype of aggressive lymphoma with clinical and biological heterogeneity.[1] The International Prognostic Index (IPI) is a powerful prognostic tool since establishment in 1993, based on 5 clinical parameters (age > 60 years, serum lactate

Submitted 10 August 2023; accepted 3 December 2023; prepublished online on *Blood Advances* First Edition 3 January 2024. https://doi.org/10.1182/bloodadvances.2023011425.

*Y.W., Q.S., Z.-Y.S., S.-T., and M.-C.Z. contributed equally to this study.

Genomic and gene expression data have been deposited on https://www.biosino.org/node in project OEP001143. Deidentified individual participant data that underlie the reported results will be made available 3 months after publication for a period of 5 years after the publication date at www.example.org. Proposals for access should be sent to the corresponding authors, Peng-Peng Xu (pengpeng_xu@126.com). The study protocol is included as a data supplement available with the online version of this article.

The full-text version of this article contains a data supplement.

dehydrogenase [LDH] above the upper limit of normal, Ann Arbor stage III-IV disease, Eastern Cooperative Oncology Group [ECOG] performance status score ≥2, and ≥2 sites with extranodal involvement) in patients with aggressive lymphoma treated with a chemotherapy regimen, primarily such as cyclophosphamide, doxorubicin, vincristine, and prednisone (CHOP).[2] During the past 30 years, clinical trials using the immunochemotherapy regimen rituximab plus CHOP (R-CHOP) have continued to underscore the prognostic significance of the IPI in the rituximab era.[3] However, the reason of the great prognostic capability and the underlying biological signatures of the IPI have not yet been elucidated, especially for risk stratification and treatment decision in the current context of precision medicine.

Notable achievements have been made to reveal the genetic heterogeneity of DLBCL, with genetic subtypes established using patterns of co-occurring gene alterations.[4-7] The LymphGen and Cluster (C0-C5) algorithm stratify comparable molecular features into genetic subtypes,[5,6] including the MCD/C5 cluster, BN2/C1 cluster, and EZB/C3 cluster, whereas subtypes such as the ST2/C4, N1, and *TP53*-associated/C2 clusters are slightly different.[5,6] Using information on mutations of 35 genes and rearrangements of 3 genes (*BCL2*, *BCL6*, and *MYC*), we developed the LymphPlex algorithm to identify 7 genetic subtypes, namely MCD-like, BN2-like, N1-like, EZB-like MYC^{+}, EZB-like MYC^{-}, ST2-like, and a specific subgroup *TP53*^{Mut} of patients with *TP53* mutations, displaying clinical feasibility in the selection of targeted agents.[5,6,8] In addition to genetic subtypes, 4 lymphoma microenvironment (LME) categories are identified as germinal center–like (GC), mesenchymal (MS), inflammatory (IN), and depleted (DP), also correlating with clinical behaviors and biological aberrations in DLBCL.[9]

In this study, we examined the clinical significance of the IPI in a large cohort of 2592 patients with newly diagnosed DLBCL, and performed genomic and transcriptomic analysis to illustrate the oncogenic mutations, genetic subtypes, tumor microenvironmental alterations, and LME categories associated with IPI risk factors, in order to explore the biological signatures of the IPI and optimize the IPI-based treatment strategies with multi-omics information in DLBCL.

Patients and methods

Patients

The selection process for patients in this study is outlined in Figure 1. From December 1997 to December 2020, 2592 patients with newly diagnosed DLBCL were screened, of whom 621 were from a nationwide multicenter clinical trial, NHL-001,[10] with the last follow-up through 30 October 2022. Two experienced pathologists (H.-M.Y. and B.-S.O.Y.) confirmed the histological diagnosis for each patient according to the World Health Organization classification.[11] Survival analysis was performed on 1932 patients receiving R-CHOP–based immunochemotherapy, excluding patients with central nervous system lymphoma (n = 42), or with immunochemotherapy other than R-CHOP-based regimens (n = 618). We performed DNA sequencing and RNA sequencing to detect oncogenic mutations (n = 1233), genetic subtypes (n = 963), and LME categories (n = 487). The study was approved by the Shanghai Ruijin Hospital review board with informed consent in accordance with the Declaration of Helsinki.

Immunohistochemistry and fluorescence in situ hybridization (FISH) analysis

Immunohistochemistry was performed on 5-μm paraffin sections by an indirect immunoperoxidase method using antibodies against CD10, BCL6, MUM1, BCL2, and MYC. GC B-cell (GCB) or non-GCB origin was determined using the Hans algorithm,[12] with 30% cutoff values of CD10, BCL6, and MUM1. As for BCL2/MYC double expressors, the cutoff values were 50% and 40%, respectively.[13] FISH was performed using fusion probe IGH/BCL2

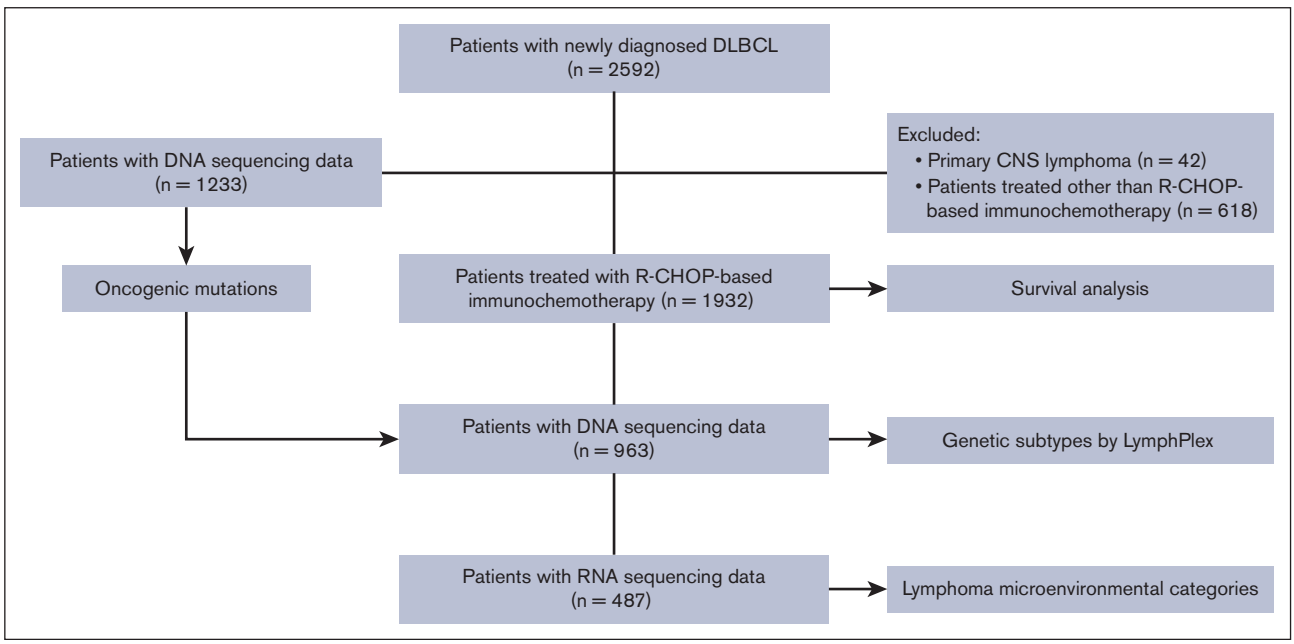

Figure 1. Patient flow diagram.

Table 1. Clinical and pathological characteristics of the patients with newly diagnosed DLBCL

Characteristic	IPI risk group				*P* value
	Low (n = 1229)	Low-intermediate (n = 521)	High-intermediate (n = 493)	High (n = 349)	
Sex					.116
Male	654 (53.2)	281 (53.9)	277 (56.2)	200 (57.3)	
Female	575 (46.8)	240 (46.1)	216 (43.8)	149 (42.7)	
Age, y					< .0001
≤60	913 (74.3)	262 (50.3)	229 (46.5)	49 (14.0)	
>60	316 (25.7)	259 (49.7)	264 (53.5)	300 (86.0)	
Ann Arbor stage					< .0001
I-II	1062 (86.4)	185 (35.5)	43 (8.7)	4 (1.1)	
III-IV	167 (13.6)	336 (64.5)	450 (91.3)	345 (98.9)	
Extranodal involvement					< .0001
0-1	1206 (98.1)	426 (81.8)	229 (46.5)	52 (14.9)	
≥2	23 (1.9)	95 (18.2)	264 (53.5)	297 (85.1)	
Serum LDH					< .0001
Normal	1049 (85.4)	212 (40.7)	85 (17.2)	16 (4.6)	
>1 × ULN	180 (14.6)	309 (59.3)	408 (82.8)	333 (95.4)	
ECOG performance status					< .0001
0-1	1212 (98.6)	478 (91.7)	400 (81.1)	158 (45.3)	
≥ 2	17 (1.4)	43 (8.3)	93 (18.9)	191 (54.7)	
Cell of origin (Han)					< .0001
GCB	431/1016 (42.4)	162/404 (40.1)	131/390 (33.6)	100/307 (32.6)	
non-GCB	585/1016 (57.6)	242/404 (59.9)	259/390 (66.4)	207/307 (67.4)	
BCL2/MYC double expressor					< .0001
Yes	161/782 (20.6)	80/281 (28.5)	76/275 (27.6)	79/216 (36.6)	
No	621/782 (79.4)	201/281 (71.5)	199/275 (72.4)	137/216 (63.4)	
***BCL2* and *MYC* translocation**					.057 (Fisher)
Yes	11/623 (1.8)	4/246 (1.6)	2/251 (0.8)	9/201 (4.5)	
No	612/623 (98.2)	242/246 (98.4)	249/251 (99.2)	192/201 (95.5)	

N = 2592; unless otherwise indicated, data are n (%).
ENI, extra-nodal involvements; ULN, upper limit of normal.

(14q32/18q21), break-apart probes BCL6 (3q27) and C-MYC (8q24; GP Medical Technologies, Ltd, Beijing, China) according to the guidelines of the manufacturer. *BCL2*, *BCL6*, and *MYC* translocations were performed on paraffin sections with cutoff values as 10%.[14,15]

DNA sequencing

Genomic DNA was extracted from frozen tumor tissue using a QIAamp DNA Mini Kit (Qiagen, Hilden, Germany), or from formalin-fixed paraffin-embedded tumor tissue by a GeneRead DNA formalin-fixed paraffin-embedded Tissue Kit (Qiagen), following the manufacturer's guidelines. Sequencing of DNA of 977 patients was reported in our previous study,[8] with 256 patients newly analyzed. Whole-genome sequencing, whole-exome sequencing, and targeted sequencing covering 55 lymphoma-associated genes were performed on 109, 227, and 897 patients, respectively, as previously described.[16-18] Using the GRCh37 human reference genome (version 2009-02), Samtools (version 0.1.18), Picard (version 1.93), and Genome Analysis Toolkit (version 4.1.4.0) were used for BAM file handling, local realignment, base recalibration, and calling variants, respectively. Mutations in the coding region were annotated using the Annovar software (version 2017-07-17). Variants were filtered according to the rules listed in our previous studies.[14,15,19] Details for DNA sequencing are provided in the supplementary material.

RNA sequencing

RNA was extracted from frozen tumor tissue with TRIzol and RNeasy Mini Kit (Qiagen, Hilden, Germany). RNA quantity was assessed on a Nanodrop instrument (NanoDrop Technologies, Wilmington, Delaware), and the integrity of RNA was qualified by a RNA 6000 Nano Kit on an Aligent 2100 Bioanalyzer (Agilent Technologies). RNA sequencing was performed on samples from 487 patients, including 400 patients from our previous studies,[8] and 87 patients newly analyzed. The paired-end reads were pseudoaligned to the GRCh38 transcriptome (Ensembl version 106) and quantified using Kallisto software (version 0.46.0).[20] R package "sva" was used to remove the batch effect. Raw counts

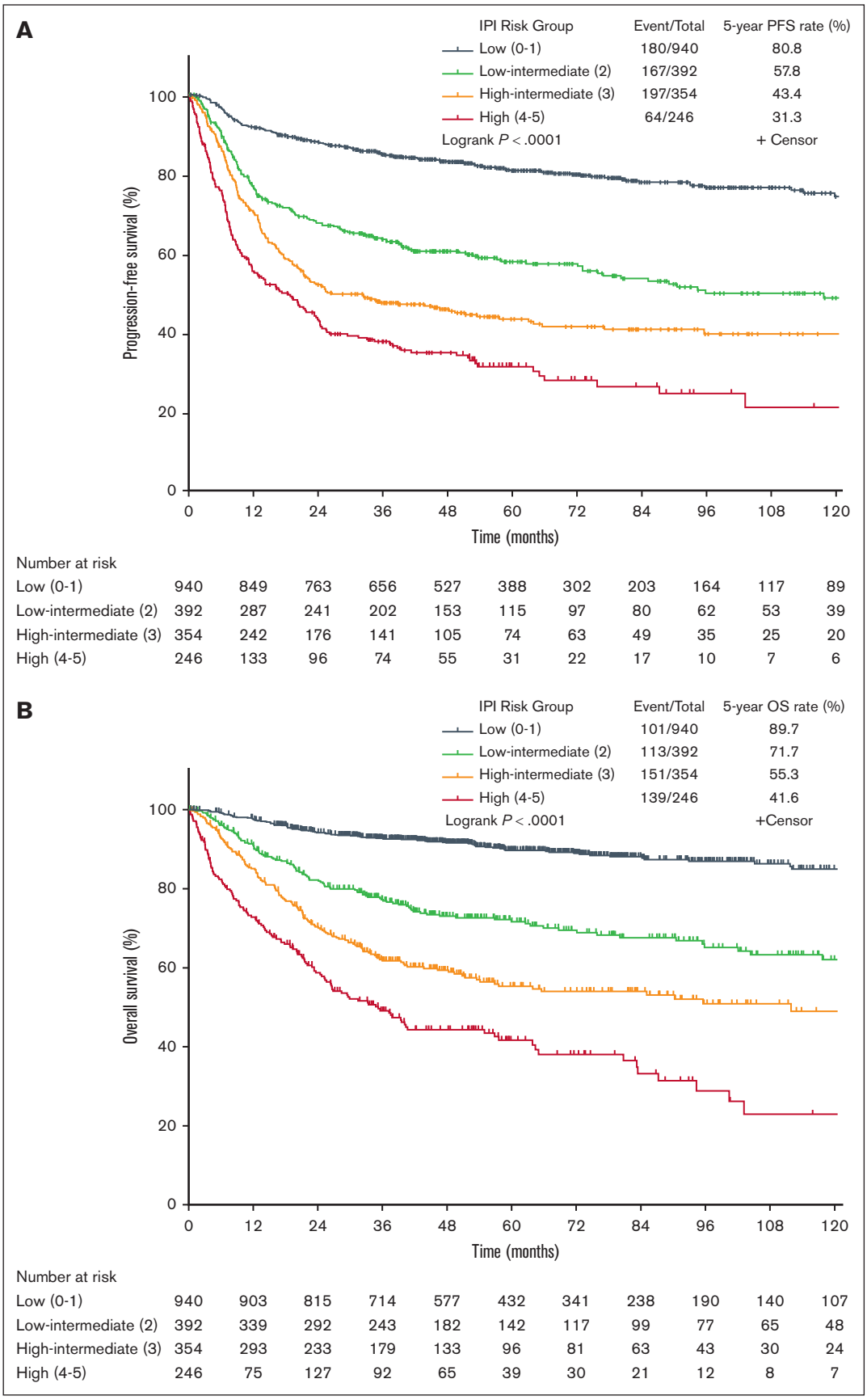

Figure 2. Survival curves of patients treated with R-CHOP–based immunochemotherapy. (A) PFS for IPI risk groups. (B) OS for IPI risk groups.

Table 2. Stratified models for PFS and OS of individual IPI risk factors

	PFS				OS			
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
Risk factor	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value	Hazard ratio (95% CI)	*P* value
Age, y								
≤60	Reference	< .0001	Reference	< .0001	Reference	< .0001	Reference	< .0001
>60	1.615 (1.393-1.872)		1.400 (1.206-1.626)		2.008 (1.682-2.398)		1.760 (1.472-2.104)	
Ann Arbor stage								
I-II	Reference	< .0001	Reference	< .0001	Reference	< .0001	Reference	< .0001
III-IV	3.157 (2.689-3.707)		2.083 (1.730-2.507)		3.703 (3.032-4.522)		2.228 (1.775-2.797)	
Extranodal involvement								
0-1	Reference	< .0001	Reference	.029	Reference	< .0001	Reference	.073
≥ 2	2.288 (1.965-2.664)		1.209 (1.020-1.433)		2.433 (2.036-2.908)		1.197 (0.984-1.456)	
Serum LDH								
Normal	Reference	< .0001	Reference	< .0001	Reference	< .0001	Reference	< .0001
>1 × ULN	2.928 (2.506-3.421)		1.960 (1.656-2.320)		3.765 (3.100-4.574)		2.416 (1.960-2.979)	
ECOG performance status								
0-1	Reference	< .0001	Reference	< .0001	Reference	< .0001	Reference	< .0001
≥2	2.721 (2.260-3.277)		1.712 (1.414-2.074)		3.159 (2.563-3.894)		1.888 (1.523-2.341)	

N = 1932
CI, confidence interval; ULN, upper limit of normal.

were normalized and applied to obtain differentially expressed genes using R package "limma" (version 3.38.3). Details for RNA sequencing are provided in the Supporting Information Methods.

Simplified LymphPlex algorithm for genetic subtyping

The LymphPlex algorithm was applied to paraffin samples with the information on mutation status of 35 genes and rearrangements revealed by targeted sequencing, and 3 genes (*BCL2*, *BCL6*, and *MYC*) by FISH. [8] A total of 963 patients receiving R-CHOP-based immunochemotherapy with DNA sequencing data were assigned to 1 of the following genetic subtypes: EZB-like (including EZB-like MYC^{-} and EZB-like MYC^{+}), ST2-like, BN2-like, *TP53*^{Mut}, N1-like, MCD-like, or not otherwise specified (NOS).

LME categories

LME subtypes of DLBCL were identified using the LME categories based on the tool as previously described.[9] Gene expression profiles in 487 patients were analyzed, calculated, and integrated, to define the diversity in the composition and functionality of the DLBCL microenvironment. Patients were assigned to 1 of the following LME categories: GC, MS, IN, and DP.

Statistical analysis

Categorical variables were summarized using frequencies and percentages, whereas continuous variables were summarized using medians and ranges. The baseline characteristics of patients were analyzed using 2-sided $\chi 2$ test. Progression-free survival (PFS) was calculated from the date of diagnosis to the date when disease progression or relapse were recognized or the date of the last follow-up. Overall survival (OS) was measured from the date of diagnosis to the date of death or the date of last follow-up. The Kaplan-Meier method was used to analyze survival functions, the log-rank test and weighted log-rank test were used to make the comparison. The Cox regression method was applied to analyze univariate and multivariate hazard ratios (HRs) and Wald z values. The oncogenic mutations, genetic subtypes, and LME categories according to 4 IPI risk groups were analyzed by Mantel-Haenszel $\chi 2$ test and Spearman correlation analysis. Differences between each IPI risk factor and oncogenic mutations were analyzed by Mann-Whitney *U* test. Pearson $\chi 2$ test was used to evaluate the relationship between each IPI factor and oncogenic mutations, genetic subtypes, and LME categories. Stratified models for each risk score were compared using Akaike information criterion (AIC) and the concordance index (C-index).[21] Overall statistical analyses were performed by Statistical Package for the Social Sciences (SPSS) 23.0 software and R version 4.2.2. Statistical significance was defined as $P < .05$ (2-sided).

Results

Clinical and prognostic significance of patients with DLBCL according to IPI risk groups

Among 2592 patients with newly diagnosed DLBCL, the IPI effectively categorized patients into 4 distinct groups: 1229 (47.4%) patients in the low-risk group (IPI = 0-1), 521 (20.1%) patients in the low-intermediate risk group (IPI = 2), 493 (19.0%) patients in the high-intermediate risk group (IPI = 3), and 349 (13.5%) patients in the high-risk group (IPI = 4-5). The clinical characteristics of the patients are summarized in Table 1. Compared with patients in the low and low-intermediate risk groups, patients in the high-intermediate and high-risk groups of the IPI had increased percentage of all the 5 IPI risk factors ($P < .0001$). Additionally, an increased prevalence of non-GCB

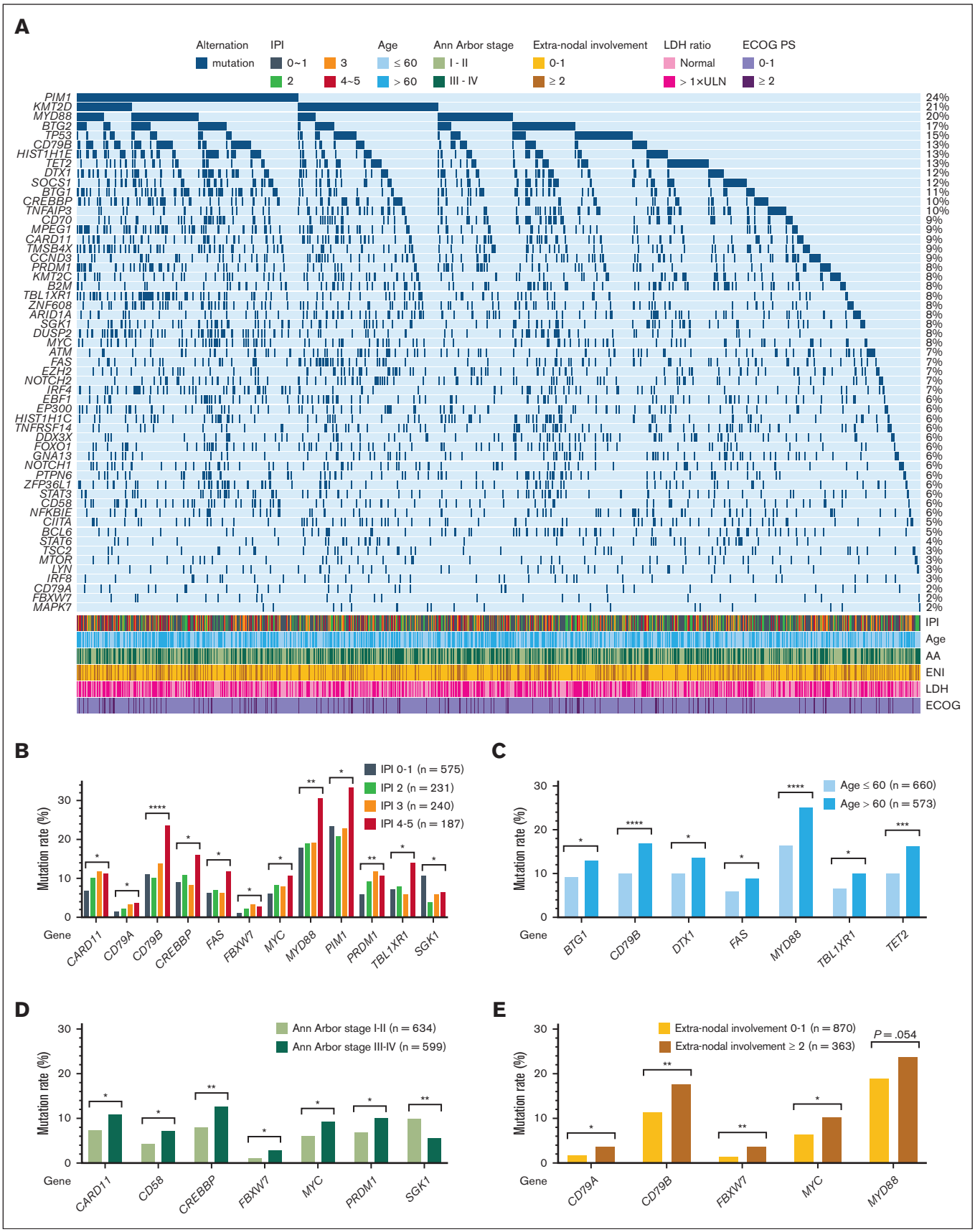
A
Alternation
mutation
IPI
0~1
2
3
4~5
Age
≤ 60
> 60
Ann Arbor stage
I - II
III - IV
Extra-nodal involvement
0-1
≥ 2
LDH ratio
Normal
> 1×ULN
ECOG PS
0-1
≥ 2
PIM1 24%
KMT2D 21%
MYD88 20%
BTG2 17%
TP53 15%
CD79B 13%
HIST1H1E 13%
TET2 13%
DTX1 12%
SOCS1 12%
BTG1 11%
CREBBP 10%
TNFAIP3 10%
CD70 9%
MPEG1 9%
CARD11 9%
TMSB4X 9%
CCND3 9%
PRDM1 8%
KMT2C 8%
B2M 8%
TBL1XR1 8%
ZNF609 8%
ARID1A 8%
SGK1 8%
DUSP2 8%
MYC 8%
ATM 7%
FAS 7%
EZH2 7%
NOTCH2 7%
IRF4 7%
EBF1 6%
EP300 6%
HIST1H1C 6%
TNFRSF14 6%
DDX3X 6%
FOXO1 6%
GNA13 6%
NOTCH1 6%
PTPN6 6%
ZFP36L1 6%
STAT3 6%
CD58 6%
NFKBIE 6%
CIITA 5%
BCL6 5%
STAT6 4%
TSC2 3%
MTOR 3%
LYN 3%
IRF8 3%
CD79A 2%
FBXW7 2%
MAPK7 2%
IPI
Age
AA
ENI
LDH
ECOG
B
IPI 0-1 (n = 575)
IPI 2 (n = 231)
IPI 3 (n = 240)
IPI 4-5 (n = 187)
Mutation rate (%)
Gene
CARD11
CD79A
CD79B
CREBBP
FAS
FBXW7
MYC
MYD88
PIM1
PRDM1
TBL1XR1
SGK1
C
Age ≤ 60 (n = 660)
Age > 60 (n = 573)
Mutation rate (%)
Gene
BTG1
CD79B
DTX1
FAS
MYD88
TBL1XR1
TET2
D
Ann Arbor stage I-II (n = 634)
Ann Arbor stage III-IV (n = 599)
Mutation rate (%)
Gene
CARD11
CD58
CREBBP
FBXW7
MYC
PRDM1
SGK1
E
Extra-nodal involvement 0-1 (n = 870)
Extra-nodal involvement ≥ 2 (n = 363)
P = .054
Mutation rate (%)
Gene
CD79A
CD79B
FBXW7
MYC
MYD88

Figure 3.

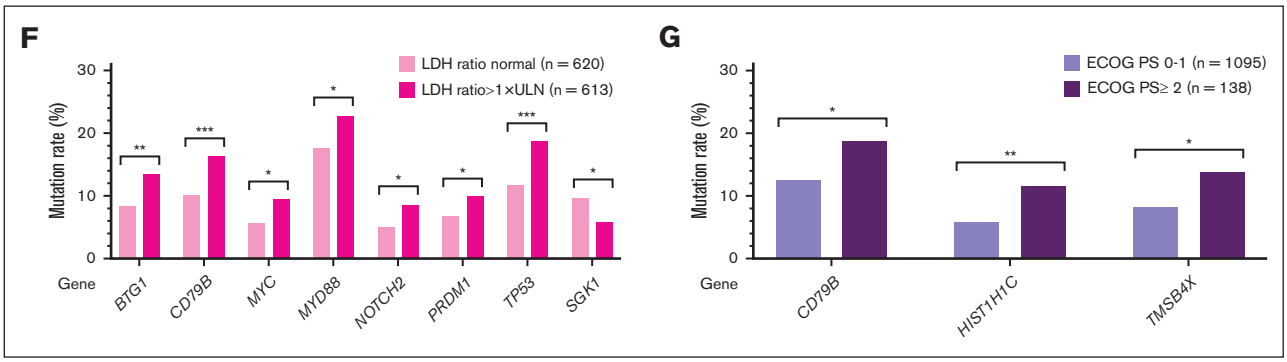

Figure 3 (continued) Mutation profile of patients with DLBCL. (A) Relationship between oncogenic mutations and IPI risk groups. (B) Mantel-Haenszel $\chi 2$ of oncogenic mutations according to the IPI. (C) Relationship between oncogenic mutations and age. (D) Relationship between oncogenic mutations and Ann Arbor stage. (E) Relationship between oncogenic mutations and extranodal involvement. (F) Relationship between oncogenic mutations and LDH ratio. (G) Relationship between oncogenic mutations and ECOG performance status (PS); $^{*}P < .05$, $^{**}P < .01$, $^{***}P < .001$, and $^{****}P < .0001$.

subtype and BCL-2/MYC double expression was observed in patients with high-intermediate and high-risk IPI scores ($P < .0001$).

As shown in Figure 2, 1932 patients treated with R-CHOP–based immunochemotherapy were categorized into 4 distinct groups by IPI. With a median follow-up of 59.6 months (range, 0.2-247.6 months), patients with low-risk and low-intermediate risk IPI scores had relatively favorable outcomes. Specifically, the 5-year PFS and OS rates were significantly higher in the low-risk group (80.8% and 89.7%) than in the low-intermediate (57.8% and 71.7%), high-intermediate (43.4% and 55.3%), and high-risk (31.3% and 41.6%) groups ($P < .0001$). The Cox proportional hazards model confirmed all the 5 IPI risk factors, including age, Ann Arbor stage, extranodal involvement, serum LDH, and ECOG performance status ($P < .05$; Table 2).

Oncogenic mutations related to different IPI risk groups

Oncogenic mutations of 55 genes involved in DLBCL tumorigenesis were analyzed in 1233 patients (Figure 3A; supplemental Table 1A-B). Significantly higher frequencies of *PIM1*, *MYD88*, *CD79B*, *CREBBP*, *TBL1XR1*, *FAS*, *MYC*, and *CD79A* mutations were observed in patients with an IPI score of 4 to 5. Significantly higher frequencies of *CARD11*, *PRDM1*, and *FBXW7* mutations were observed in patients with an IPI score of 3 to 5. Conversely, *SGK1* mutation was associated with an IPI score of 0 to 1 (Figure 3B). $MYD88^{L265P}$ mutation alone or in combination with *CD79B* mutation was also more frequently observed in patients with an IPI score of 4 to 5 (supplemental Table 1A).

Correlations between oncogenic mutations and each IPI risk factor were further analyzed. *BTG1*, *CD79B*, *DTX1*, *FAS*, *MYD88*, *TBL1XR1*, and *TET2* mutations were associated with age of >60 years (Figure 3C). Similar associations were observed for $MYD88^{L265P}$ mutation alone or in combination with *CD79B* mutation (supplemental Table 1B). Meanwhile, mutations in *CARD11*, *CD58*, *CREBBP*, *FBXW7*, *MYC*, and *PRDM1* were increased with Ann Arbor stage III to IV, whereas *SGK1* showed an opposite trend (Figure 3D). *CD79A*, *CD79B*, *FBXW7*, and *MYC* mutations, as well as $MYD88^{L265P}$ mutation alone or in combination with *CD79B* mutation, were associated with multiple extranodal involvement (Figure 3E; supplemental Table 1B). *BTG1*, *CD79B*, *MYC*, *MYD88*, *NOTCH2*, *PRDM1*, *TP53* mutations, and $MYD88^{L265P}$ with *CD79B* mutation were associated with elevated LDH (Figure 3E; supplemental Table 1B). Mutations in *CD79B*, *HIST1H1C*, and *TMSB4X* mutations were associated with ECOG performance status of ≥2 (Figure 3G).

Genetic subtypes related to different IPI risk groups

Survival and regression analyses in genetic subtypes and the IPI. Among 963 patients with DNA sequencing data, 518 patients were classified into genetic subtypes by LymphPlex, with the remaining patients as NOS subtype. We analyzed survival outcomes of patients with other genetic subtypes stratified by different IPI risk groups. The outcomes of patients with genetic subtypes (n = 963; supplemental Figure 1A-B) were comparable with the overall study population (n = 1932). Compared with EZB-like MYC^{-}, ST2-like, BN2-like, and NOS subtypes, patients with EZB-like MYC^{+}, $TP53^{Mut}$, N1-like, and MCD-like subtypes had inferior prognosis **(**supplemental Figure 1C-D**).**

The relationship between genetic subtypes and IPI is illustrated in Figure 4A. The MCD-like subtype was positively correlated with IPI risk scores ($P = .002$; Spearman R = 0.122), whereas the ST2-like subtype showed the opposite trend ($P = .027$; Spearman R = −0.113). Univariate and multivariate models were constructed between IPI and genetic subtypes with IPI low-risk group and NOS subtype set as references. EZB-like MYC^{+} and $TP53^{Mut}$ subtypes presented inferior prognosis, independent of the IPI (Figure 4B-C), similar to the results in an external cohort from BC Cancer (BCC; 296 newly diagnosed patients with DLBCL received R-CHOP and had available data of the IPI; supplemental Table 2).[22] We also examined all the genetic subtypes and 5 risk factors within the IPI using a Cox model, EZB-like MYC^{+} and $TP53^{Mut}$ subtypes also add novel prognostic information beyond the IPI risk factors (supplemental Table 3).

Survival analysis was conducted on the patients with remaining genetic subtypes. The results demonstrated that, with the exception of the ST2-like subtype, the remaining genetic subtypes could further be classified into 4 groups with different survival according

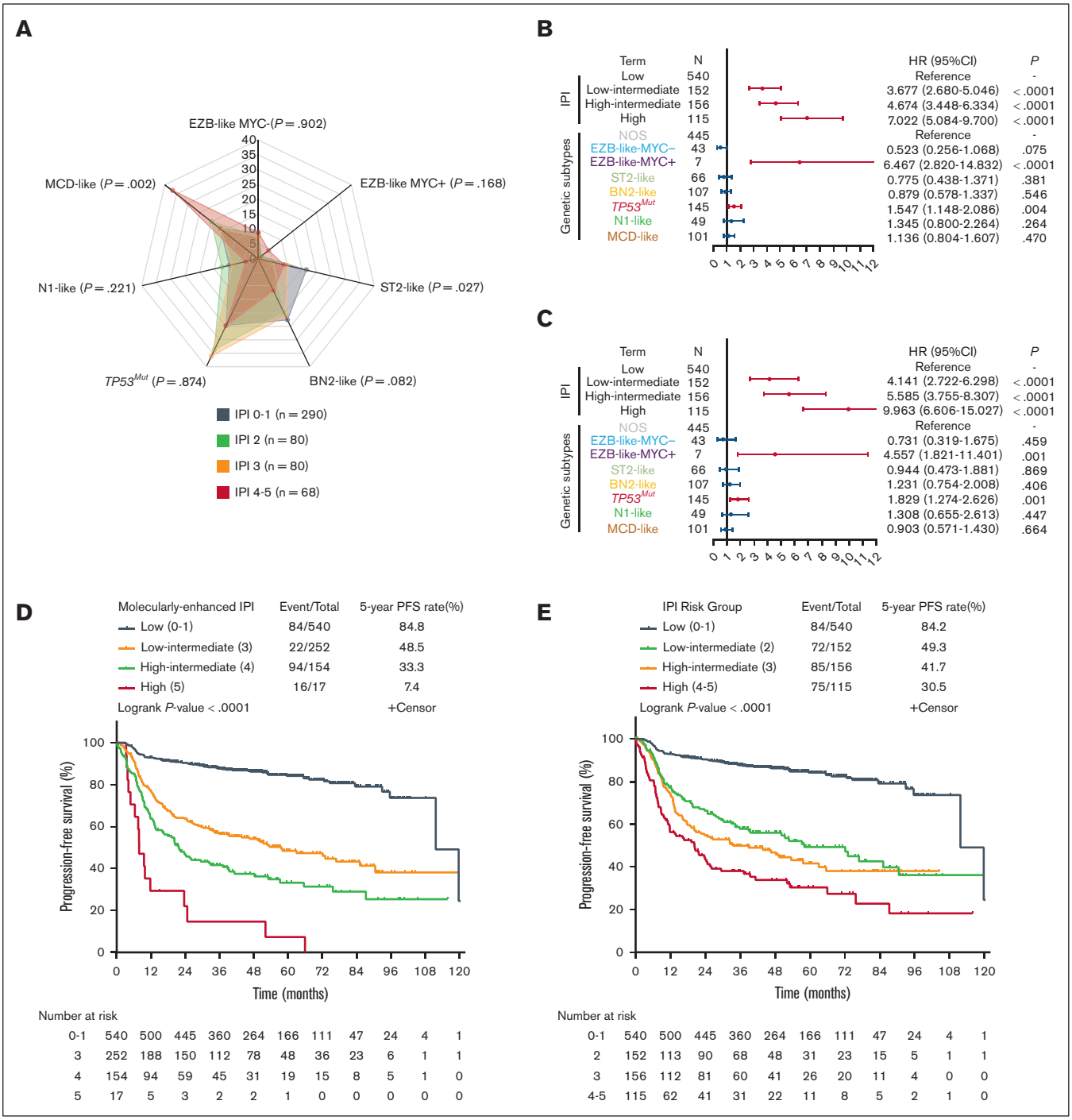

Figure 4. Genetic subtypes of patients with DLBCL. (A) Mantel-Haenszel $\chi 2$ of genetic subtypes related to IPI risk group. (B-C) Forest plots visualize hazard ratios (HRs) and *P* values obtained from the multivariate analysis of genetic subtypes and IPI for PFS and OS. (D) PFS into 4 main risk groups stratified by molecularly-enhanced IPI scores in training cohort. (E) PFS for IPI risk groups in training cohort. (F) Distribution of IPI risk factors across different genetic subtypes (*$P < .05$ and **$P < .01$).

to IPI score, and correlated with inferior prognosis for patients with IPI scores of ≥2 (supplemental Figure 2).

Molecularly-enhanced IPI. Due to the significant prognostic role of EZB-like-MYC+ and *TP53*^{Mut} subtypes, incorporating them into the IPI was essential. EZB-like MYC+ (n = 7, 0.7%) was excluded in the subsequent analyses because of the limited representation, while IPI risk scores and *TP53*^{Mut} were considered prognostic factors in multivariable Cox proportional hazard (PH) regression model. The prognostic score was obtained giving a weight to each variable according to its relative importance, derived from the z-Wald values found in the Cox PH model (supplemental

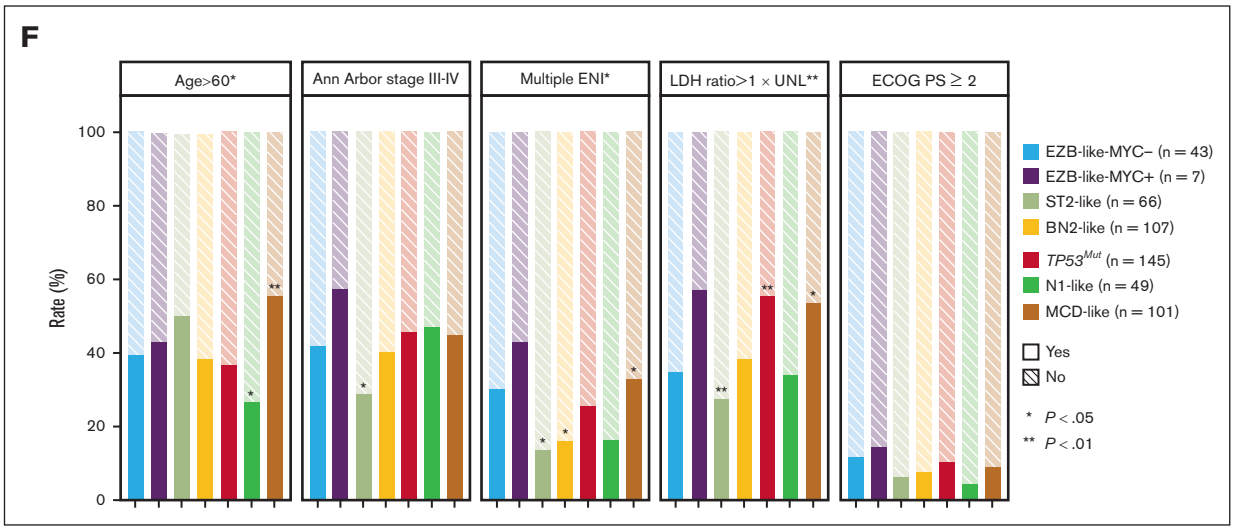

Figure 4 (continued)

Table 4A). The z-score for any factor was divided by the minimum z-score observed (*TP53*^{Mut}, considered as reference) to obtain the ratio. Finally, the weights were obtained rounding the ratio, and the score was the sum of weights.[23]

We thus obtained a new model called molecularly-enhanced IPI, with scores ranging from 0 to 5, demonstrating a good correlation with PFS and OS (supplemental Figure 3A-B). Patients were classified into 4 main risk groups: low (score 0-1, 56.1% of patients), low-intermediate (score 3, 26.2% of patients), high-intermediate (score 4, 16.0% of patients), and high (score 5, 1.8% of patients). Notably, the 5-year PFS and OS rates were significantly higher in the low-risk group (84.2% and 92.0%) than the low-intermediate (48.5% and 66.2%), high-intermediate (33.3% and 50.6%), and high-risk (7.4% and 24.8%) groups ($P < .0001$ for all) (Figure 4D and supplemental Figure 3C). HRs among the 4 risk groups were shown in supplemental Table 4B. Furthermore, the molecularly-enhanced IPI also discriminated the patients with more dismal PFS and OS, and displayed a more distinct survival difference between the intermediate-low-risk group and the intermediate-high-risk group compared to IPI (P value in molecularly-enhanced IPI for PFS and OS: 0.001, compared to P value in IPI for PFS and OS: 0.186 and 0.193, respectively) (Figure 4D-E, supplemental Figure 3C-D). Although the differences in the C-index were marginal, molecularly-enhanced IPI provided the better fit for the data than IPI (AIC value for PFS: 3898.918 vs 3914.485, AIC value for OS: 2462.153 vs 2472.199, respectively; supplemental Table 5).

To evaluate the model of molecularly-enhanced IPI, we collected a total of 795 cases for external validation from BCC (296 newly diagnosed patients with DLBCL received R-CHOP and had available data of IPI),[22] and the UK population-based Hematological Malignancy Research Network which consisted of 499 patients with DLBCL had available information of IPI and treated with R-CHOP.[24] The LymphPlex was calculated in our previous research.[8] Applying the molecularly-enhanced IPI to the validation cohorts, patients survival were also classified into 4 risk groups (supplemental Figure 3E-F), and the HRs between the 4 risk groups were displayed on supplemental Table 6. Compared with the IPI, the molecularly-enhanced IPI discriminated the patients with more dismal PFS and OS (supplemental Figure 3G-H), and the PFS differentiation between the intermediate-low-risk group and the intermediate-high-risk group was more discernible (P value in molecularly-enhanced IPI was <.0001, as P value in the IPI was .156, respectively). Although the differences in the C-index were marginal, the molecularly-enhanced IPI provided better fit for the data than IPI (AIC value for PFS: 4955.806 vs 4962.667, AIC value for OS: 4942.604 vs 4443.204, respectively; supplemental Table 6).

Distribution of genetic subtypes from 5 risk factors within IPI. The distribution of genetic subtypes from 5 IPI risk factors was also analyzed. The MCD-like subtype increased with older age, whereas the N1-like subtype had decreased prevalence with age ($P = .011$). Multiple extranodal involvement tended to be clustered in the MCD-like subtype, whereas the ST2-like and BN2-like subtypes were often associated with fewer extranodal involvement ($P = .013$). *TP53*^{Mut} and MCD-like subtypes had elevated LDH, whereas ST2-like had an opposite trend ($P = .001$; Figure 4F).

Tumor microenvironment alterations related to different IPI risk groups

Among 487 patients with RNA sequencing data, 4 LME categories were assigned: 69 patients in GC-LME, 166 patients in MS-LME, 122 patients in IN-LME, and 130 patients in DP-LME. We analyzed the survival time of different IPI risk groups within LME categories (n = 487), who have a comparable OS rate to that of the overall study population (n = 1932; supplemental Figure 4). Patients with GC-LME and MS-LME had favorable PFS and OS, whereas IN-LME and DP-LME presented inferior prognoses (Figures 5A-B). GC-LME was negatively correlated with IPI risk scores ($P = .006$; Spearman R = −0.128; Figure 5C). Besides, multivariate regression analysis showed that MS-LME was an independent protective factor of the IPI and factors within the IPI (Figure 5D-E; supplemental Table 7). Survival analysis of different IPI risk groups within the GC-LME, IN-LME, and DP-LME was conducted (n = 321). The results showed that these patients

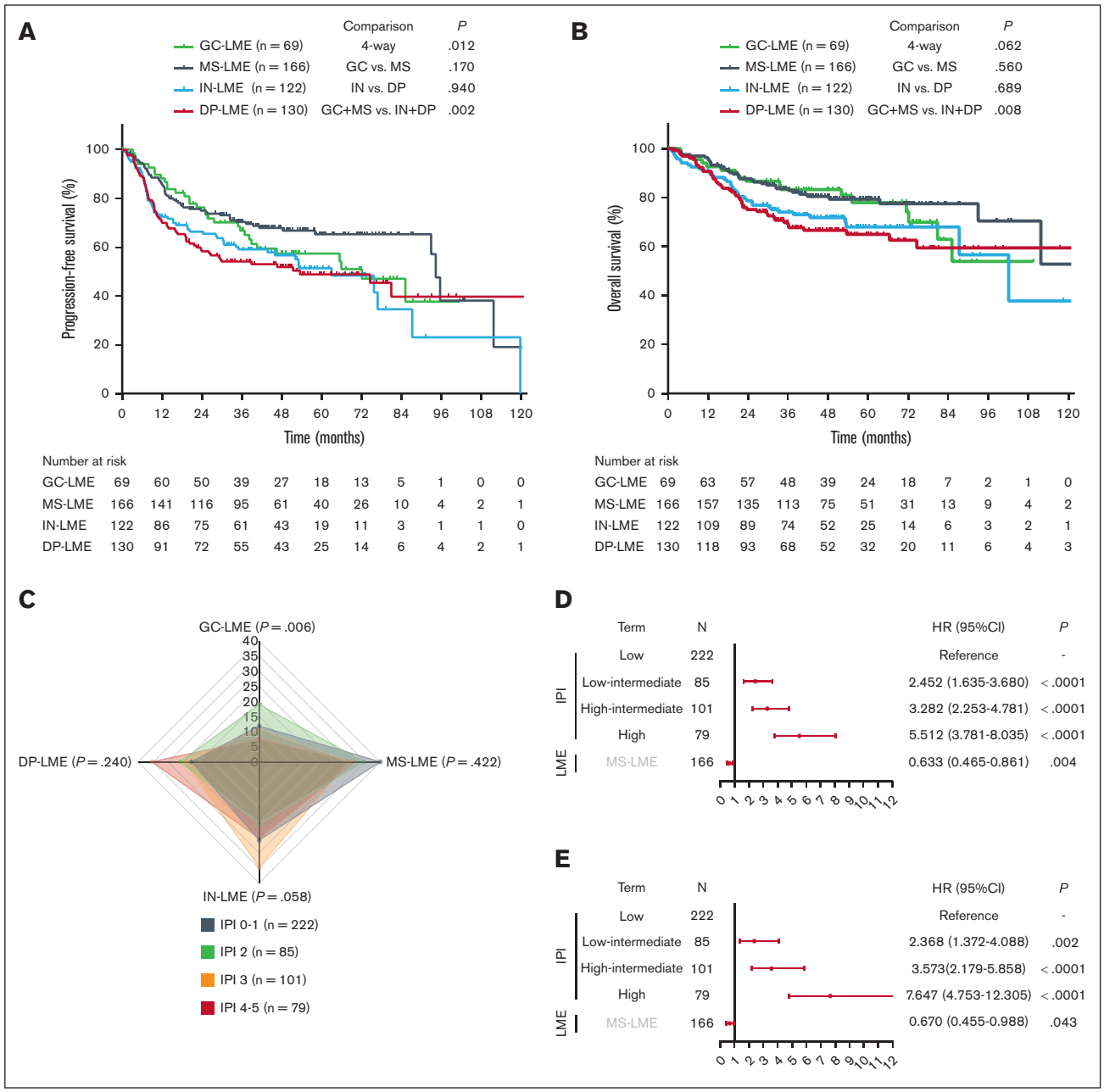

Figure 5. LME categories of patients with DLBCL. (A) Kaplan-Meier models of PFS according to LME categories. (B) Kaplan-Meier models of OS according to LME categories. (C) Mantel-Haenszel $\chi 2$ of LME categories related to IPI risk group. (D-E) Forest plots visualize hazard ratios (HRs) and *P* values obtained from the multivariate analysis of LME categories and IPI for PFS and OS. (F) Distribution of IPI risk factors across different LME categories; $^{*}P < .05$, $^{**}P < .01$, and $^{***}P < .001$.

correlated with inferior prognosis for IPI scores of ≥2 (supplemental Figure 5). Moreover, the distribution of IPI risk factors across various LME categories showed that elevated LDH was more frequently observed in IN-LME, whereas GC-LME had normal LDH ($P = .003$; Figure 5F).

Discussion

In addition to the prognostic capability of IPI upon treatment with R-CHOP–based immunochemotherapy,[25-27] we found a correlation of IPI with genetic subtypes and LME categories based on a large cohort of patients with newly diagnosed DLBCL, demonstrating different genetic and tumor microenvironmental signatures of the IPI in DLBCL.

Specific oncogenic mutations are linked to DLBCL progression.[1,14,16] As revealed by DNA sequencing, mutations involving the BCR/NF-κB signaling pathway, B-cell differentiation, and histone acetylation (such as *BTG1*, *CD79B*, *FAS*, *MYD88*, *MYD88*^{L265P}, *TBL1XR1*, and *TET2*) showed significant association with age, consistent with our previous study.[14] Multiple extranodal involvement was often related to the MCD subtype, with increased prevalence of

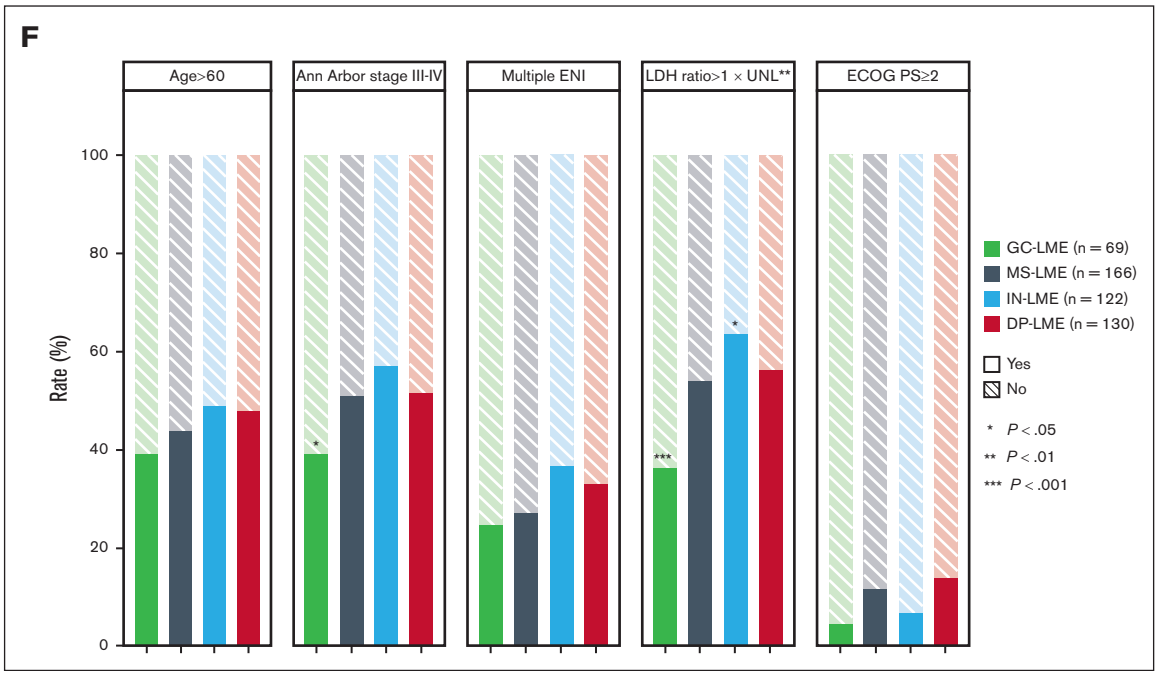

Figure 5 (continued)

MYD88^{L265P} and *CD79B* mutations.[16] Elevated LDH represents tumor growth and invasive potential,[28] co-occurring with unfavorable markers, such as *TP53* mutation,[29] and BCR/NF-κB–associated mutations (such as *BTG1*, *CD79B*, *MYD88*, and *PRDM1*). Accordingly, higher frequencies of *PIM1*, *MYD88*, and *CD79B* mutations (clustered in the MCD-like subtype) were observed in patients with IPI score of 4 to 5, whereas lower frequency of *SGK1* mutation (clustered in the ST2-like subtype) was shown in patients with IPI scores of 0/1. Together, oncogenic mutations and corresponding genetic subtypes contributed to distinct biological signatures underlying IPI risk factors.

IPI-based risk stratification is widely applied for treatment decision in DLBCL.[30,31] Of note, *TP53*^{Mut} was an independent prognostic factor of PFS and OS, representing a distinct subtype with inferior prognosis irrespective of IPI score.[5,32] Less likely to benefit from R-CHOP–based immunochemotherapy,[33-35] novel targeted therapies, such as decitabine,[19] are required for patients with *TP53*^{Mut}, even stratified as low-risk IPI. Furthermore, our study demonstrated that integrating *TP53*^{Mut} into the IPI enhanced its capacity to differentiate patient survival outcomes, which has also been validated by external cohorts. Similarly, patients with EZB-like MYC^{+} subtype, characterized by *BCL2* translocation, *EZH2* mutation, and *MYC* rearrangement, exhibited inferior prognosis and CD8^{+} T-cell deficiency within the tumor microenvironment.[8] Prospective trials should explore the potential impact of BCL2 inhibitors, EZH2 inhibitors, or immunomodulatory agents in patients with EZB-like MYC^{+} subtype.[8,36] Moreover, the ST2-like subtype, which was associated with a favorable prognosis, warranted further clinical investigation focusing on therapy deescalation.[24] The EZB-like MYC^{-}, BN2-like, N1-like, and MCD-like subtypes were associated with inferior prognosis only for IPI scores of ≥2, suggesting the addition of targeted agents in an IPI-dependent manner, such as EZH2 inhibitors in EZB-like MYC^{-} subtype,[37] and Bruton tyrosine kinase inhibitors in BN2-like, N1-like, and MCD-like subtypes.[6,38-40]

More recently, LME categories have been implicated in immunotherapy of DLBCL.[3] MS-LME was an independent protective factor in the IPI. MS-LME exhibited enriched signatures from vascular endothelial cells, cancer-associated fibroblasts, fibroblastic reticular cells, and extracellular matrix,[9] all of which are linked to favorable outcome in DLBCL.[9] For IPI scores of ≥2, GC-LME, IN-LME, and DP-LME categories correlated with inferior prognosis, indicating a necessity for more precise treatment within these patients: GC-LME represented transformed variants of indolent, GC-derived follicular lymphoma, and may be sensitive to lenalidomide plus R-CHOP[9,41]; IN-LME exhibited an enrichment of neutrophils, macrophages, CD8^{+} T cells, CD8^{+} T cells with high programmed cell death protein 1 expression, as well as immune-suppressive cytokines. Using Bruton tyrosine kinase inhibitors or other potential agents to modulate the immunosuppressive microenvironment could be potential therapeutic approaches.[40] DP-LME was characterized by minimal presence of microenvironmental cells, DNA hypermethylation, and specific SMAD1 promoter methylation,[42] which could be targeted by epigenetic repression of SMAD1,[43,44] and PI3K inhibitors.[9] Therefore, LME categories should also be taken account into IPI-based risk stratification and treatment selection.

In conclusion, by exploring the biological signatures of the IPI and revealing the correlation of the IPI with different genetic and environmental background, the IPI score in combination with multi-omics information may represent a further precision treatment strategy in DLBCL.

Acknowledgments

The authors appreciate the effort of the physicians in enrolling patients; and thank all the patients involved for consent to analysis of their clinical data.

This study was supported in part by research funding from the National Key Research and Development Program of China (2022YFC2502600), the National Natural Science Foundation of China (82370187, 81830007, 82130004, 82000192, and 82070204), the Chang Jiang Scholars Program, the Shanghai Municipal Education Commission Gaofeng Clinical Medicine Grant Support, the Clinical Research Plan of Shanghai Hospital Development Center (SHDC2020CR1032B and SHDC2022CRD033), and the Multicenter Clinical Research Project by Shanghai Jiao Tong University School of Medicine (DLY201601).

Authorship

Contribution: W.-L.Z. and P.-P.X. designed and supervised the study; Y.W., Q.S., Z.-Y.S., S.-T., and M.-C.Z. collected the data and performed the analyses; L.D., H.-M.Y., and B.-S.O. reviewed the histopathologic diagnoses and gave technical support; R.S. and D.F. prepared biological samples; R.-J.M., S.C., and L.W. provided executive support and data surveillance; Y.W., P.-P.X., W.-L.Z. interpreted the results and wrote the manuscript; and P.-P.X. and W.-L.Z. verified the underlying data.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: S.T., 0000-0002-3845-9891; W.-L.Z., 0000-0002-6834-1616.

Correspondence: Wei-Li Zhao, Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics; National Research Center for Translational Medicine at Shanghai, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, 197 Rui Jin Er Rd, Shanghai 200025, China; email: zhao.weili@yahoo.com; and Peng-Peng Xu, Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics; National Research Center for Translational Medicine at Shanghai, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, 197 Rui Jin Er Rd, Shanghai 200025, China; email: pengpeng_xu@126.com.

References

1. Sehn LH, Salles G. Diffuse large B-cell lymphoma. *N Engl J Med*. 2021;384(9):842-858.
2. International Non-Hodgkin's Lymphoma Prognostic Factors Project. A predictive model for aggressive non-Hodgkin's lymphoma. *N Engl J Med*. 1993;329(14):987-994.
3. Xu PP, Huo YJ, Zhao WL. All roads lead to targeted diffuse large B-cell lymphoma approaches. *Cancer Cell*. 2022;40(2):131-133.
4. Miao Y, Medeiros LJ, Li Y, Li J, Young KH. Genetic alterations and their clinical implications in DLBCL. *Nat Rev Clin Oncol*. 2019;16(10):634-652.
5. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med*. 2018;24(5):679-690.
6. Wright GW, Huang DW, Phelan JD, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell*. 2020;37(4):551-568.e14.
7. Schmitz R, Wright GW, Huang DW, et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. *N Engl J Med*. 2018;378(15):1396-1407.
8. Shen R, Fu D, Dong L, et al. Simplified algorithm for genetic subtyping in diffuse large B-cell lymphoma. *Signal Transduct Target Ther*. 2023;8(1):145.
9. Kotlov N, Bagaev A, Revuelta MV, et al. Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures. *Cancer Discov*. 2021;11(6):1468-1489.
10. Xu PP, Fu D, Li JY, et al. Anthracycline dose optimisation in patients with diffuse large B-cell lymphoma: a multicentre, phase 3, randomised, controlled trial. *Lancet Haematol*. 2019;6(6):e328-e337.
11. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
12. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103(1):275-282.
13. Hu S, Xu-Monette ZY, Tzankov A, et al. MYC/BCL2 protein coexpression contributes to the inferior survival of activated B-cell subtype of diffuse large B-cell lymphoma and demonstrates high-risk gene expression signatures: a report from The International DLBCL Rituximab-CHOP Consortium Program. *Blood*. 2013;121(20):4021-4031. quiz 4250.
14. Zhu Y, Fu D, Shi Q, et al. Oncogenic mutations and tumor microenvironment alterations of older patients with diffuse large B-cell lymphoma. *Front Immunol*. 2022;13:842439.
15. Huo YJ, Xu PP, Fu D, et al. Molecular heterogeneity of CD30+ diffuse large B-cell lymphoma with prognostic significance and therapeutic implication. *Blood Cancer J*. 2022;12(3):48.
16. Shen R, Xu PP, Wang N, et al. Influence of oncogenic mutations and tumor microenvironment alterations on extranodal invasion in diffuse large B-cell lymphoma. *Clin Transl Med*. 2020;10(7):e221.
17. Huang YH, Cai K, Xu PP, et al. 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(1):10.

18. Qin W, Fu D, Shi Q, et al. Molecular heterogeneity in localized diffuse large B-cell lymphoma. *Front Oncol*. 2021;11:638757.
19. Zhang MC, Fang Y, Xu PP, et al. Clinical efficacy and tumour microenvironment influence of decitabine plus R-CHOP in patients with newly diagnosed diffuse large B-cell lymphoma: phase 1/2 and biomarker study. *Clin Transl Med*. 2021;11(12):e584.
20. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34(5):525-527.
21. Ruppert AS, Dixon JG, Salles G, et al. International prognostic indices in diffuse large B-cell lymphoma: a comparison of IPI, R-IPI, and NCCN-IPI. *Blood*. 2020;135(23):2041-2048.
22. Ennishi D, Jiang A, Boyle M, et al. Double-hit gene expression signature defines a distinct subgroup of germinal center B-cell-like diffuse large B-cell lymphoma. *J Clin Oncol*. 2019;37(3):190-201.
23. Merli F, Luminari S, Tucci A, et al. Simplified geriatric assessment in older patients with diffuse large B-cell lymphoma: the Prospective Elderly Project of the Fondazione Italiana Linfomi. *J Clin Oncol*. 2021;39(11):1214-1222.
24. Lacy SE, Barrans SL, Beer PA, et al. Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. *Blood*. 2020;135(20):1759-1771.
25. Zhou Z, Sehn LH, Rademaker AW, et al. An enhanced International Prognostic Index (NCCN-IPI) for patients with diffuse large B-cell lymphoma treated in the rituximab era. *Blood*. 2014;123(6):837-842.
26. Cai J, Tian X, Ma S, et al. A nomogram prognostic index for risk-stratification in diffuse large B-cell lymphoma in the rituximab era: a multi-institutional cohort study. *Br J Cancer*. 2021;125(3):402-412.
27. Marcheselli R, Bari A, Tadmor T, et al. Improving the international prognostic index score using peripheral blood counts: results of a large multicenter study involving 520 patients with diffuse large B cell lymphoma. *Hematol Oncol*. 2020;38(4):439-445.
28. Park JH, Yoon DH, Kim DY, et al. The highest prognostic impact of LDH among International Prognostic Indices (IPIs): an explorative study of five IPI factors among patients with DLBCL in the era of rituximab. *Ann Hematol*. 2014;93(10):1755-1764.
29. Zenz T, Kreuz M, Fuge M, et al. TP53 mutation and survival in aggressive B cell lymphoma. *Int J Cancer*. 2017;141(7):1381-1388.
30. Nowakowski GS, Hong F, Scott DW, et al. Addition of lenalidomide to R-CHOP improves outcomes in newly diagnosed diffuse large B-cell lymphoma in a randomized phase II US intergroup study ECOG-ACRIN E1412. *J Clin Oncol*. 2021;39(12):1329-1338.
31. Murawski N, Pfreundschuh M, Zeynalova S, et al. Optimization of rituximab for the treatment of DLBCL (I): dose-dense rituximab in the DENSE-R-CHOP-14 trial of the DSHNHL. *Ann Oncol*. 2014;25(9):1800-1806.
32. Liu C, Shi P, Li Z, Li B, Li Z. A nomogram for predicting the rapid progression of diffuse large B-cell lymphoma established by combining baseline PET/CT total metabolic tumor volume, lesion diffusion, and TP53 mutations. *Cancer Med*. 2023;12(16):16734-16743.
33. Ichikawa A, Kinoshita T, Watanabe T, et al. Mutations of the p53 gene as a prognostic factor in aggressive B-cell lymphoma. *N Engl J Med*. 1997;337(8):529-534.
34. Xu-Monette ZY, Wu L, Visco C, et al. Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. *Blood*. 2012;120(19):3986-3996.
35. Dodero A, Guidetti A, Marino F, et al. Dose-adjusted EPOCH and rituximab for the treatment of double expressor and double-hit diffuse large B-cell lymphoma: impact of TP53 mutations on clinical outcome. *Haematologica*. 2022;107(5):1153-1162.
36. Alduaij W, Collinge B, Ben-Neriah S, et al. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. *Blood*. 2023;141(20):2493-2507.
37. McCabe MT, Ott HM, Ganji G, et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature*. 2012;492(7427):108-112.
38. Wilson WH, Young RM, Schmitz R, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. *Nat Med*. 2015;21(8):922-926.
39. Davis RE, Ngo VN, Lenz G, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature*. 2010;463(7277):88-92.
40. Wilson WH, Wright GW, Huang DW, et al. Effect of ibrutinib with R-CHOP chemotherapy in genetic subtypes of DLBCL. *Cancer Cell*. 2021;39(12):1643-1653.e3.
41. Desai SH, LaPlant B, Macon WR, et al. Lenalidomide in combination with R-CHOP produces high response rates and progression-free survival in new, untreated diffuse large B-cell lymphoma transformed from follicular lymphoma: results from the Phase 2 MC078E study. *Blood Cancer J*. 2021;11(9):160.
42. Calon A, Tauriello DV, Batlle E. TGF-beta in CAF-mediated tumor growth and metastasis. *Semin Cancer Biol*. 2014;25:15-22.
43. Clozel T, Yang S, Elstrom RL, et al. Mechanism-based epigenetic chemosensitization therapy of diffuse large B-cell lymphoma. *Cancer Discov*. 2013;3(9):1002-1019.
44. Stelling A, Wu CT, Bertram K, et al. Pharmacological DNA demethylation restores SMAD1 expression and tumor suppressive signaling in diffuse large B-cell lymphoma. *Blood Adv*. 2019;3(20):3020-3032.