HHS Public Access

Author manuscript

Clin Cancer Res. Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:

Clin Cancer Res. 2022 March 01; 28(5): 972–983. doi:10.1158/1078-0432.CCR-21-2949.

Genetic subtyping and phenotypic characterization of the immune microenvironment and MYC/BCL2 double expression reveal heterogeneity in diffuse large B-cell lymphoma

Zijun Y. Xu-Monette[1,#], **Li Wei**[1,2,#], **Xiaosheng Fang**[1,3,#], **Qingyan Au**[4,#], **Harry Nunns**[4,#], **Máté Nagy**[4], **Alexandar Tzankov**[5], **Feng Zhu**[1], **Carlo Visco**[6], **Govind Bhagat**[7], **Karen Dybkaer**[8], **April Chiu**[9], **Wayne Tam**[10], **Youli Zu**[11], **Eric D. Hsi**[12], **Fredrick B. Hagemeister**[13], **Xiaoping Sun**[14], **Xin Han**[14], **Heounjeong Go**[15], **Maurilio Ponzoni**[16], **Andrés J.M. Ferreri**[16], **Michael B. Møller**[17], **Benjamin M. Parsons**[18], **J. Han van Krieken**[19], **Miguel A. Piris**[20], **Jane N. Winter**[21], **Yong Li**[22], **Bing Xu**[23], **Maher Albitar**[24], **Hua You**[2,#], **Ken H. Young**[1,25,*]

[1]Hematopathology Division and Department of Pathology, Duke University Medical Center, Durham, NC, USA

[2]Children's Hospital of Chongqing Medical University, Chongqing, China

[3]Department of Hematology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, Shandong, China

[4]NeoGenomics Laboratories, Aliso Viejo, California, USA

[5]Institute of Pathology, University Hospital Basel, Switzerland

[6]University of Verona, Verona, Italy

[7]Columbia University Irving Medical Center and New York Presbyterian Hospital, New York, NY, USA

[8]Aalborg University Hospital, Aalborg, Denmark

[9]Mayo Clinic, Rochester, MN, USA

[10]Weill Medical College of Cornell University, New York, NY, USA

[11]The Methodist Hospital, Houston, TX, USA

[*]**Correspondence:** Ken H. Young, MD, PhD, Duke University Medical Center, Division of Hematopathology and Department of Pathology, Durham, NC 27710, USA. Phone: 1-919-668-7568; Fax: 1-919-684-1856; ken.young@duke.edu; Hua You, MD, PhD, Children's Hospital of Chongqing Medical University, Chongqing 400014, China. youhua307@163.com; Zijun Y. Xu-Monette, PhD, Duke University Medical Center, Division of Hematopathology and Department of Pathology, Durham, NC 27710, USA. zijun.xumonette@duke.edu.

[#]These authors contributed equally.

Authors' Contributions

Conception and design: ZYXM, HY, KHY.

Collection and assembly of data: ZYXM, LW, XF, QA, MN, AT, FZ, CV, GB, KD, AC, WT, YZ, EDH, FBH, XS, XH, HG, MP, AJMF, MBM, BMP, JHvK, MAP, JNW, YL, BX, MA, HY, KHY.

Data analysis and interpretation: ZYXM, QA, HN, MA, KHY

Manuscript writing: ZYXM, LW, XF, MA, HY, KHY.

Final approval of manuscript: All authors.

Conflict of interest statement: QA, HN and MN are employees of NeoGenomics Laboratories, Inc.. MA is an employee of Genomic Testing Corporative, LCA. Other authors declare no potential conflicts of interest.

[12]Cleveland Clinic, Cleveland, OH, USA

[13]Department of Lymphoma and Myeloma, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[14]Department of Laboratory Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

[15]Asan Medical Center, Ulsan University College of Medicine, Seoul, Korea

[16]San Raffaele H. Scientific Institute, Milan, Italy

[17]Odense University Hospital, Odense, Denmark

[18]Gundersen Lutheran Health System, La Crosse, WI, USA

[19]Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

[20]Hospital Universitario Marqués de Valdecilla, Santander, Spain

[21]Feinberg School of Medicine, Northwestern University, Chicago, IL, USA

[22]Department of Medicine, Baylor College of Medicine, Houston, TX, USA

[23]The First Affiliated Hospital of Xiamen University, Xiamen, Fujian, China

[24]Genomic Testing Cooperative, LCA, Irvine, CA, USA

[25]Duke University Cancer Institute, Durham, NC, USA

Abstract

Purpose: Diffuse large B-cell lymphoma (DLBCL) is molecularly and clinically heterogeneous, and can be subtyped according to genetic alterations, B cell-of-origin or microenvironmental signatures using high-throughput genomic data at the DNA or RNA level. Although high-throughput proteomic profiling has not been available for DLBCL subtyping, MYC/BCL2 protein double-expression (DE) is an established prognostic biomarker in DLBCL. The purpose of this study is to reveal the relative prognostic roles of DLBCL genetic, phenotypic, and microenvironmental biomarkers.

Experimental Design: We performed targeted next-generation sequencing, immunohistochemistry for MYC, BCL2, and FN1, and fluorescent multiplex immunohistochemistry for microenvironmental markers in a large cohort of DLBCL. We performed correlative and prognostic analyses within and across DLBCL genetic subtypes and MYC/BCL2 double-expressors.

Results: We found that MYC/BCL2 double-high-expression (DhE) had significant adverse prognostic impact within the EZB genetic subtype and LymphGen-unclassified DLBCL cases but not within MCD and ST2 genetic subtypes. Conversely, *KMT2D* mutations significantly stratified DhE but not non-DhE DLBCL. T-cell infiltration showed favorable prognostic effects within BN2, MCD, and DhE but unfavorable within ST2 and LymphGen-unclassified cases. FN1 and PD-1-high expression had significant adverse prognostic effects within multiple DLBCL genetic/phenotypic subgroups. The prognostic effects of DhE and immune biomarkers within DLBCL

genetic subtypes were independent although DhE and high Ki-67 were significantly associated with lower T-cell infiltration in LymphGen-unclassified cases.

Conclusions: Together, these results demonstrated independent and additive prognostic effects of phenotypic MYC/BCL2 and microenvironment biomarkers and genetic subtyping in DLBCL prognostication, important for improving DLBCL classification and identifying prognostic determinants and therapeutic targets.

Keywords

DLBCL; MYC; BCL2; double expressor; double hit; genetic subtype; tumor immune microenvironment; FN1; KMT2D; mIHC; GEP; gene signature; MCD; EZB

Introduction

Diffuse large B-cell lymphoma (DLBCL), the most common aggressive B-cell lymphoma, is biological and clinically heterogeneous. DLBCL can be classified into germinal center B-cell-like (GCB) and activated B-cell-like (ABC) subtypes according to unaltered B-cell cell-of-origin (COO) gene-expression signatures at the RNA level (1), or subtyped into various genetic subsets according to genetic alterations at the DNA level (2-5). Classification methods based on lymphoma microenvironmental (LME) gene-expression signatures with prognostic values are also emerging (6-8). For example, the ‘GC-like’ and ‘mesenchymal’ LME categories of DLBCL have significantly better clinical outcome than the ‘inflammatory’ and ‘depleted’ LME categories (6). Although genetic subtypes are pathogenically more homogenous than gene-expression-based COO subtypes, the EZB genetic subtype can be further divided into EZB-MYC^{+} and EZB-MYC^{-} subsets by DHIT gene-expression signatures (4), which are associated with ‘depleted’ and ‘mesenchymal’ immune-deserted microenvironment, respectively (6).

Note that in a LME gene-expression signature study, the ‘depleted’ microenvironment type has strong enrichment for the lymphoma cell proliferation gene signatures and increased tumor clonality (less intratumoral heterogeneity based on RNA sequences expressed from the immunoglobulin genes) (6). Lymphoma progression is accompanied by reduced microenvironment complexity and decreased tumor-infiltrating T cells in mouse models, and the ‘depleted’ type of microenvironment is recapitulated in a *Myc*/*Bcl2* mouse model (6). The inverse correlation between immune-cell signatures and tumor-cell proliferation signatures has also been shown in other cancer gene signature studies (9-11), which could suggest a driver role of tumor-intrinsic mechanisms in host immune response (or vice versa). However, it could also result from the mathematics of gene-expression analysis, and the cellularity analyses in the LME study (including those for mouse models) were all based on estimated cell types according to transcriptomic signatures rather than immunophenotying.

Although phenotypic proteomics-based DLBCL classification is not available currently, MYC/BCL2 protein double-expression (DE) is a WHO-recognized prognostic biomarker in DLBCL (12-16). MYC/BCL2-DE is predominately of ABC COO (proportion, 53-76% (12-15, 17, 18)), however, GCB and ABC subtypes of MYC/BCL2-DE have similarly poor prognosis (13, 17, 19). Compared with high-grade B-cell lymphoma (HGBCL, ref (16,

20, 21)) with *MYC*/*BCL2* genetic double-hit (*MYC*/*BCL2*-DH) (21), DE-DLBCL is more frequent (21%-34% compared with only 3-5% of DLBCL cases (13, 22-24)), clinically more heterogeneous, and with better overall survival (OS). The reported five-year OS has ranged from 30% (13) to 50% (25) for all DE-DLBCL cases, ~30% (13) to 36% (12) in the absence of *MYC*/*BCL2*-DH, compared to the 15% (13) to 27% (12, 26) rate for HGBCL-*MYC*/*BCL2*-DH and 70% (12) to 80% (14) for non-DE DLBCL in retrospective studies. In a prospective RICOVER-60 clinical trial, the five-year OS was 42% for DE-DLBCL versus ~82% for non-DE DLBCL (15, 27). However, MYC/BCL2-DE was not associated with significant inferior survival or prognostic impact in a REMoDL-B clinical trial (28), in young patients in a R-MegaCHOEP trial (29), and in a retrospective study in patients with stage I/II DLBCL (13, 30). Two recent studies indicated that a cut-off of ≥70% for MYC expression, which is the mean MYC expression level in *MYC*-rearranged DLBCL in our previous study (17), is more optimal than the ≥40% cutoff, with high concordance and reproducibility for diagnosis and prognosis in both retrospective and clinical trial study cohorts (31, 32).

In this study, first we aim to evaluate the associations and prognostic roles of MYC/BCL2 and microenvironment biomarkers at the protein level in genetically defined DLBCL subtypes. We evaluated not only the abundance of T cells, macrophages, and natural killer (NK) cells relative to B cells and expression of fibronectin (FN1, an extracellular matrix glycoprotein) in LME, but also immune checkpoint (PD-1, PD-L1/PD-L2, and CTLA-4) expression affecting the antitumor function of immune cells. The second objective of this study is to determine whether genetic subtyping and microenvironmental biomarkers can stratify patients with DE-DLBCL with either ≥40% or ≥70% cutoff for MYC overexpression. The results revealed the heterogeneity of DLBCL subgroups, gained insights into the relationship of genetic/phenotypic/microenvironment biomarkers, and demonstrated their independent prognostic impact within DLBCL subgroups.

Patients and Methods

Patients

We performed targeted next-generation sequencing (NGS) for 444 adult patients with *de novo* DLBCL as part of the DLBCL Consortium Study Program (33). 424 cases with quality sequencing data were included in this study. All patients were treated with the standard immunochemotherapy for DLBCL, rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP), or R-CHOP-like regimens. Primary cutaneous DLBCL, primary mediastinal large B-cell lymphoma, and primary central nervous system lymphoma have been excluded. Fluorescence *in situ* hybridization (FISH) (13, 34) identified HGBCL-*MYC*/*BCL2*-DH in 6 sequenced cases. The other 418 cases are referred as DLBCL, not otherwise specified (DLBCL-NOS). This retrospective study was conducted following data collection protocols involving no more than minimal risk to subjects with a waiver of written consent requirement approved by the institutional review boards of Duke University and each participating institution.

Targeted NGS and genetic alteration analysis

The Agencourt FormaPure Total 96-Prep Kit was used to extract DNA from formalin-fixed, paraffin-embedded tissue lysates using an automated KingFisher Flex and protocols as recommended by each manufacturer. Sequencing was performed on an Illumina NextSeq 550 System platform. Most cases were sequenced with a 275-gene NGS panel (35) which was expanded to 315 genes (Supplementary Table S1) for 30 cases. The DNA fragments were ligated at their 5' ends with a sequencing platform-specific adapter containing unique molecular markers and sample index. The average sequencing depth was 700×, and a ≥100× sequence coverage (after removing duplicates) was required for mutation calling. The percent reads passing filter (Reads PF) was >80%. All coding exons of targeted genes were sequenced, along with 50 intronic nucleotides flanking each exon end. Detected variants were annotated using COSMIC, dbSNP, ExAC, PolyPhen, and internal database accumulated from testing large number of various types of cancers. Variant allele frequency of each variant was considered along with chromosomal loss and male vs female if the variant was located on X chromosome. Somatic mutations were called when not listed in SNP databases, considered deleterious and previously reported in other cancers. Alignment of sequencing data and variant calling were performed with the DRAGEN Somatic Pipeline (Illumina) using tumor-only analysis against the GRCh37 reference genome to identify single nucleotide variants and indels. The VCF files were imported into VariantStudio (version 3.0) for evaluating variants in the bed file. Variants were annotated within Variant Studio software. Large gain and loss were evaluated using CNVkit software (36). Briefly, the software takes advantage of both on- and off-target sequencing reads, compares binned read depths in on- and off-target regions to pooled normal reference, and estimates the copy number at various resolutions. Fifteen normal samples were pooled and used for normalization, corrected for several systematic biases to result in a final table of $\log_2$ copy ratios. Copy number change in segments of various chromosomes were predicted and visualized by the software based on $\log_2$ ratio values.

Fluorescent multiplex immunohistochemistry (mIHC)

Immune and tumor cells in DLBCL samples were analyzed using the MultiOmyx platform (NeoGenomics Laboratories, Aliso Viejo, CA), as described previously (37). Briefly, fluorescent mIHC was performed on tissue microarrays using 13 immune marker antibodies conjugated to either cyanine 3 or cyanine 5 (cy3/5) and 4',6-diamidino-2-phenylindole (DAPI) over the course of eight staining rounds, and imaged on INCell analyzer 2200 microscopes (GE Healthcare Life Sciences). Single-cells were classified based on phenotypic marker expression, and phenotype-positive cell counts were quantified for each marker and coexpression of multiple markers. Moreover, immune cell infiltration was evaluated by the percentage of immune cells among the total number of $PAX5^+$ cells, $CD3^+$ cells, $CD68^+$ cells, and $CD56^+$ cells in each DLBCL sample, and immune checkpoint expression was evaluated by percentage of positive cells within a specific cell type (T, B, NK, macrophages). Uniform Manifold Approximation and Projection (UMAP) plots were generated from single-cell intensities of immune markers using a Python implementation of UMAP (38). Hierarchically clustered heatmaps (clustermaps) were generated from normalized phenotype counts across multiple genetic groups using the data visualization software package Seaborn (39).

Immunohistochemistry (IHC)

Conventional IHC for MYC, BCL2, Ki-67, and FN1 was performed on tissue microarray using methods described previously with MYC/BCL2/Ki-67 antibodies (13, 33) and a mouse monoclonal anti-Fibronectin antibody (clone# 568) (sc-52331; Santa Cruz) at a dilatation of 1:50. MYC, BCL2, and Ki-67 expression levels were evaluated by percentage of positive tumor cells. FN1 expression was evaluated by percentage of positive stromal cells.

Gene-expression signature analysis

Gene expression profiling (GEP) data from the Affymetrix GeneChip Human Genome HG-U133 Plus 2.0 microarrays (GSE31312) (33) were pre-processed and normalized by RMA (Robust Multi-chip Average) using the R package (version 1.65.1). To identify significantly differentially expressed genes between two groups, two-class unpaired Significance Analysis of Microarrays (SAMs) was performed. CLUSTER software and the average linkage metric were then used and the clustered gene signatures were displayed by JAVA TREEVIEW (https://www.java.com/en). The Expression Analysis Systematic Explorer (A desktop version of DAVID) (40) software was used to categorize over-represented biological pathways using Gene Ontology (GO) terms.

Statistical Analysis

OS duration was defined from time of diagnosis to death from any cause or censored at last follow-up. Progression-free survival (PFS) duration was defined from diagnosis to the time of progression or death from any cause or censored at last follow-up. OS and PFS were compared with the Kaplan-Meier method using GraphPad Prism 9 software. Fisher's exact test was used for comparing distribution of categorical variables, unpaired 2-tailed *t* test was used for comparing the means, and two-sided Mann-Whitney U test was used for differences comparing the distribution or the medians of two independent groups. *P*-values ≤ 0.05 were considered to be statistically significant.

Data Availability

GEP data analyzed in this study are publicly available in Gene Expression Omnibus (GEO) at GSE31312. Other data are available on request with approved IRB protocol.

Results

DLBCL genetic subtypes show distinct gene-expression signatures

The LymphGen algorithm web tool (https://llmpp.nih.gov/lymphgen/index.php) was implemented to predict genetic subtypes for 424 DLBCL cases using data obtained from NGS (including nonsynonymous gene mutations and copy number alterations) and FISH analysis (including *BCL2*/*BCL6* gene rearrangements and *TP53* deletion). With the cutoff of ≥2 features for allocating a case to a specific genetic subtype, totally 128 (30%) cases were classified as EZB (n = 73), MCD (n = 25), ST2 (n = 14), BN2 (n = 12), or A53 (n = 4), whereas the other 296 cases were classified as having an 'Other' genetic subtype by the LymphGen classifier (Supplementary Fig. S1A). Consistent with previous studies, the

MCD, BN2, and A53 genetic subtypes were predominantly of ABC COO, whereas EZB and ST2 were mostly GCB (Supplementary Table S2). Within the GCB subtype, patients with EZB showed inferior survival compared with genetically subtyped non-EZB cases (ST2, BN2, MCD, and A53) but the differences were not statistically significant (for PFS, P = 0.17, hazard ratio [HR] 2.67, 95% confidence interval [CI] 0.95-7.50, Supplementary Fig. S1B; for OS, P = 0.24, HR 2.36, 95% CI 0.79-7.02). Within the ABC subtype, combined MCD and EZB cases showed non-significant inferior survival compared with combined ST2, BN2, and A53 cases (for OS, P = 0.18, HR 1.92, 95% CI 0.80-4.58, Supplementary Fig. S1B; for PFS, P = 0.22, HR 1.81, 95% CI 0.76-4.37). In overall DLBCL, only MCD subtype showed non-significant unfavorable effect on OS (P = 0.13. In DLBCL-NOS, P = 0.11, HR 1.55, 95% CI 0.80-3.00. Supplementary Fig. S1C). The MCD cases included 12 of 13 DLBCL cases with concurrent *MYD88*/*CD79B* mutations, 12 of 69 (17.4%) of cases with only *MYD88* mutation, and one case with only *CD79B* mutation. Among all *MYD88*-mutated cases, MCD patients had non-significant inferior OS (P = 0.087, HR 1.80, 95% CI 0.85-3.80, Supplementary Fig. S1C).

Although genetic subtypes did not show strong impact on prognosis, they showed distinct gene-expression signatures (Supplementary Fig. S2). Supplementary Table S3 lists enriched GO class and gene categories of differentially expressed genes between MCD, EZB and 'Other' subtypes. For MCD gene-expression signatures, enriched GO biological process included upregulation of metabolism and mitotic cell cycle and downregulation of cellular process, cell communication and cell adhesion. For EZB gene-expression signatures, enriched GO biological process included upregulation of antigen presentation/processing and protein kinase cascade and downregulation of cell communication. Phosphorylation process showed both upregulation and downregulation in EZB.

MYC/BCL2 IHC can stratify EZB, BN2 and genetically unclassified DLBCL

141 DLBCL-NOS cases were categorized as $MYC^{+}BCL2^{+}$ DE and 262 DLBCL-NOS cases were categorized as nonDE using the ≥40% cutoff for MYC expression. With a ≥70% cutoff for MYC-high expression (MYC^{hi}, ref. (17)), DE cases were further divided into two groups: 73 cases with MYC/BCL2 double-high expression ($MYC^{hi}BCL2^{hi}$, DhE) and 68 DE/non-DhE cases with MYC^{inter} (40-60% MYC^{+} expression)/$BCL2^{hi}$. Clinical features of these three DLBCL-NOS subgroups and HGBCL-*MYC*/*BCL2*-DH cases are shown in Supplementary Table S4.

MYC/BCL2 expression levels in DLBCL genetic subtypes are shown in Fig. 1A. MCD, BN2, and A53 genetic subtypes had high frequencies of DE (50%-58.3%), whereas only MCD had a high frequency of DhE (45.8%; significantly higher than frequencies in other subtypes. Supplementary Table S2). In the prognostic analysis within distinct EZB, MCD, ST2, BN2, and A53 genetic subtypes, *MYC*/*BCL2* double-hit had significant adverse impact on both OS and PFS within the EZB subtype; MYC^{hi} expression and $MYC^{hi}BCL2^{hi}$ DhE (but not $MYC^{+}BCL2^{+}$ DE) had significant adverse impact on OS only in EZB patients and on PFS within EZB and BN2 subtypes (Fig. 1B-C, Supplementary Fig. S3A). In contrast, MYC/BCL2 expression did not show prognostic effects within ST2 and MCD genetic subtypes (Supplementary Fig. S3B; A53 subtype had four non-DhE cases and no DhE cases

for comparison). In LymphGen-unclassified cases/'Other' genetic subtype, both DhE and DE (but not single MYC^{hi} expression) showed significant adverse effects on OS and PFS (Supplementary Fig. S3C). These results indicated that EZB and LymphGen-unclassified genetic subsets can be further stratified by routine IHC for MYC/BCL2, whereas the unfavorable prognosis of MCD was independent of MYC/BCL2 expression.

To understand the molecular mechanisms underlying the prognostic effects, we performed GEP and enrichment analysis for DhE. GEP signatures identified for DhE in EZB patients were featured by significant downregulation of antigen presentation pathway in immune responses (Fig. 1D, Table 1). In contrast, the DhE signatures identified in LymphGen-unclassified cases were enriched with metabolism, ribosome and growth-related GO categories (Table 1). To gain further insights into the immune-related signature in EZB, we compared immunophenotypes in EZB with and without DhE. Fig. 1E shows three representative UMAP plots for EZB cases with nonDE, DhE, and *MYC*/*BCL2*-DH, respectively. DhE in EZB-DLBCL-NOS cases was significantly associated with lower PD-1 expression in T cells ($P = 0.02/0.018$ by 2-tailed unpaired *t*/Mann-Whitney test, Supplementary Fig. S3D). Similar association was observed within ST2 (however statistically not significant, $P = 0.088$, only two DhE cases) but not within BN2/MCD subtypes.

Immune dysregulation and prognostic significance of immune microenvironment biomarkers vary in different DLBCL genetic subtypes

We quantitated the absolute cell counts immunophenotypically identified by fluorescent mIHC; the mean cell counts in genetic subtypes are shown by a clustermap in Supplementary Fig. S4A. To better visualize the differences between DLBCL genetic subtypes, Z-score clustermaps were generated using normalized median or mean counts of genetic subtypes for each phenotype (Fig. 2A), in which the immune microenvironment in A53, MCD, EZB and ST2 subtypes appeared "colder" compared with that in BN2 and 'Other' subtypes, and the A53 subtype more resembled the 'depleted' immune category (6). EZB compared with MCD had significantly lower $CD8^+$ and $PD\text{-}L1^+$ counts and higher $FOXP3^+$ counts (Mann-Whitney U test, Supplementary Fig. S2). We further calculated T cell proportion and percentage of $PD\text{-}1/L1^+$ cells in specific cell types. A53 and MCD had lower $CD3^+$ cells than other genetic subtypes (*t* test, Fig. 2A), A53, MCD, and ST2 had lower $CD4^+$ cells than other genetic subtypes, and A53 and EZB had lower $CD8^+$ T cells than BN2 and LymphGen-unclassified cases (*t* and Mann-Whitney tests, Supplementary Fig. S2). PD-1 expression in T cells was significantly higher in BN2; PD-L1 expression in macrophages was significantly higher in ABC-associated (MCD, BN2 and A53) than GCB-associated (EZB and ST2) genetic subtypes; and PD-L1 expression in B cells was significantly higher in MCD (both *t* and Mann-Whitney tests, Fig. 2B). In addition to immune characteristics of the LME, FN1 expression was evaluated, and A53 and ST2 showed comparably higher FN1 levels than other genetic subtypes (Fig. 2A). However, heterogeneity existed within each subtype.

Prognostic analysis in EZB subset found DhE-independent significant prognostic effects by $CD56^+$ NK cell infiltration (favorable), $FN1^+$ expression (unfavorable) and PD-1-high

expression (>55%) in T cells (unfavorable; Fig. 2C, Supplementary Fig. S4B). There was a negative association between $FN1^+$ expression and NK cell proportion in EZB (Supplementary Fig. S4B). In BN2 subset, B-cell PD-L1 expression and $CD4^+$ T cell infiltration were associated with significantly better OS/PFS (Fig. 2B and 2D). In MCD subset, $CD8^+$ T cell infiltration was associated with better PFS (Fig. 2D) and OS (P = 0.029, not significant with the Bonferroni correction for multiple testing), which effect was restricted in DhE cases however (Supplementary Fig. S4C). In contrast with these favorable prognostic effects within EZB, BN2, and MCD genetic subtypes, within the ST2 subset significantly poorer OS and PFS were associated with NK cell infiltration, high T cell/macrophage infiltration, and PD-L1 expression in B cells (Supplementary Fig. S4D). Although FN1-positivity did not show significant prognostic effect within MCD, BN2, and ST2 (all A53 cases were $FN1^+$), FN1-high expression was associated with significantly poorer survival in MCD and BN2 (P = 0.038 and 0.019 for OS and P = 0.038 and 0.012 for PFS with an $FN1^{high}$ cutoff of >30% and >50%, respectively, Supplementary Fig. S4C. Non-significant P = 0.11 in ST2).

In LymphGen-unclassified cases (majority of this study cohort), we found that the prognostic effects of immune biomarkers were similar to those in overall cohort and all immune-profiled DLBCL cases as reported previously (37), including the adverse prognostic effects associated with high levels of $CD8^+$ and memory T cell infiltration, low NK cell infiltration, low $CD3^+CD4^+FOXP3^+$ cell infiltration, PD-1 expression in $CD4^+$ T cells, PD-1-high expression in $CD8^+$ T cells, PD-L1-high expression in T cells and $CD68^+$ macrophages, and low PD-L2 expression in $CD20^+$ B cells (Supplementary Fig. S5; Supplementary Table S5, multiple testing correction with the Benjamini-Hochberg procedure). However, differently from the effects in overall cohort and all immune-profiled cases (37), in these genetically unclassified cases, high macrophage levels were associated with poorer PFS (P = 0.0032) and OS (P = 0.034), and PD-L1 expression in $CD20^+$ B cells was associated with significantly poorer OS (Fig. 2B; specifically in ABC, Supplementary Fig. S5). Together, these results suggest that the prognostic significance of immune biomarkers was affected by genetic contexts, and that immune targets existed but varied in different genetic subtypes.

For FN1 IHC performed in this study, FN1-positivity was associated with significantly poorer OS (P = 0.0026, Supplementary Fig. S5) in LymphGen-unclassified cases as in overall DLBCL (P < 0.0001). Surprisingly, high *FN1* mRNA expression was significantly associated with favorable OS/PFS in overall DLBCL, the EZB subset (Supplementary Fig. S3D), and LymphGen-unclassified cases with cutoffs in a large range, even though overall *FN1* mRNA and FN1 protein expression were positively correlated in overall DLBCL (Pearson r 0.4031, two-tailed P < 0.0001), EZB (Pearson r 0.3833, P = 0.003), and LymphGen-unclassified cases (Pearson r 0.3976, P < 0.0001), and GEP comparison between $FN1^+$ and $FN1^-$ DLBCL confirmed upregulation of *FN1* in $FN1^+$ cases (3.9 fold, false discovery rate [FDR] 0.01). Only in MCD, high *FN1* mRNA expression was associated with unfavorable survival consistent with the effect of FN1 protein expression, which was not significant however (P = 0.057 for OS, P = 0.067 for PFS).

MYC/BCL2 DhE-DLBCL can be stratified by *KMT2D* mutation status

Next, we analyzed the genetic heterogeneity of MYC/BCL2 DE-DLBCL-NOS cases, which were further grouped into MYC^{hi}-DhE (Fig. 3A) and MYC^{inter}-DE (Fig. 3B) subgroups. The DhE DLBCL-NOS subgroup had significantly higher frequencies of *MYD88*, *CD79B*, *PIM1,* and *PRDM1* mutations compared with nonDE subgroup, significantly higher frequencies of *CREBBP*, *CD79B*, and *PRDM1* mutations compared with MYC^{inter}-DE subgroup, and a significantly higher frequency of MCD subtype than both MYC^{inter}-DE and nonDE DLBCL-NOS subgroups (Supplementary Fig. S6). Compared with overall DLBCL-NOS cases, HGBCL-*MYC*/*BCL2*-DH cases (Fig. 3A) had a higher frequency of *KMT2D* mutations (67% *vs*. 27.5%, $P = 0.055$) and EZB subtype (67% *vs*. 16.5%, $P = 0.0009$).

Prognostic analysis for genetic subtypes found that within the MYC^{hi}-DhE subgroup, EZB patients had statistically non-significant inferior OS compared with non-EZB patients when HGBCL-*MYC*/*BCL2*-DH cases were also included ($P = 0.055$, HR 1.89, 95% CI 0.83-4.32, Fig. 3C), whereas within the nonDE subgroup, MCD and A53 cases had non-significant poorer OS than non-MCD/A53 cases ($P = 0.19$, HR 1.73, 95% CI 0.60-4.99; Supplementary Fig. S7). To gain molecular insights, GEP analysis was performed for EZB status in DhE patients (Fig. 1D). Intracellular signaling cascade (upregulated) by GO biological process category was significantly enriched in EZB-DhE patients (Table 2).

Prognostic analysis for individual genetic alterations found that *KMT2D* (Fig. 3C), *ARID1A*, *SMARCA4*, and *TET2* mutations were associated with significant poorer OS or PFS in patients with DhE after multiple testing corrections by the Benjamini-Hochberg method (Supplementary Table S6, Supplementary Fig. S8A). KMT2D (a histone H3 lysine 4 methyltransferase), ARID1A (functioning in chromatin remodeling), TET2 (a methylcytosine dioxygenase with key role in DNA demethylation) and SMARCA4 (component of SWI/SNF chromatin remodeling complexes) were all involved in epigenetic regulation. Except for *TET2*, these epigenetic gene mutations did not have significant prognostic impact in the non-DhE subgroup (Supplementary Fig. S8B). In contrast, the adverse prognostic impact of DhE was independent of these mutations (significant in both wild-type and mutant cases, Supplementary Fig. S8B).

The most frequent *KMT2D* mutations were associated with EZB subtype in both DhE and non-DhE subgroups ($P < 0.0001$, Supplementary Fig. S8C). We analyzed the contribution of *KMT2D*-mutated DhE-DLBCL cases to the prognostic effect of EZB genetic subtype within DhE-DLBCL and to that of DhE within EZB and ‘Other’ genetic subsets. Only the ten *KMT2D*-mutated but not the four *KMT2D*-wild-type EZB cases of DhE cases had significantly poorer survival than DhE-DLBCL cases with non-EZB genetic subtypes ($P = 0.007$, Fig. 3C) and EZB cases without DhE ($P < 0.0001$, Supplementary Fig. S8D). In contrast, in ‘Other’/genetically unsubtyped DLBCL cases, *KMT2D* mutations were infrequent in DhE cases (significantly less than in non-DhE ‘Other’ cases, $P = 0.03$), and the adverse effect of DhE in this ‘Other’ genetic subset was independent of *KMT2D* mutations (Supplementary Fig. S8D).

Prognostic effects of MYC$^+$BCL2$^+$ in DLBCL genetic subtypes and prognostic effects of *KMT2D* mutations in MYC$^+$BCL2$^+$ DLBCL in validation cohorts

We repeated our prognostic analysis in two publicly available cohorts with both mutation data and MYC/BCL2 protein data available: a cohort from British Columbia Cancer Agency (BCCA) (4, 41), and a multicenter cohort analyzed by Reddy et al (42). In both two cohorts, MYC/BCL2 double-positive expression (DpE/DE) was defined by a ≥40% cutoff for MYC$^+$ and a 50% cutoff for BCL2$^+$. Only the BCCA cohort has been genetically subtyped by the LymphGen algorithm (4)

In the BCCA cohort, consistent with our findings, only EZB and LymphGen-unclassified cases can be stratified by MYC$^+$BCL2$^+$ protein DpE ($P = 0.037$ and $P < 0.0001$, respectively. Supplementary Fig. S9A). DpE had no significant effects in the MCD/A53/BN2/ST2/N1 genetic subtypes (non-significantly unfavorable in A53, $P = 0.11$, and BN2 with only 4 DpE cases, $P = 0.17$). Conversely, MYC$^+$BCL2$^+$ DpE cases can be stratified genetically. Although MYC$^+$BCL2$^+$ was associated with MCD, LymphGen-unclassified cases but not MCD cases had significantly poorer survival among MYC$^+$BCL2$^+$ cases ($P = 0.010$, Supplementary Fig. S9B). *KMT2D* mutations showed significant adverse prognostic impact in a prognostic unfavorable genetic subset of DpE cases which combined A53, composite EZB/A53, and LymphGen-unclassified cases ($P = 0.045$, Supplementary Fig. S9B). In the other validation cohort (42), *KMT2D* mutations showed significant adverse prognostic effect in MYC$^+$BCL2$^+$ cases ($P = 0.015$) but not nonDpE cases (Supplementary Fig. S9C), similar to the prognostic effects of *KMT2D* mutations in our DhE and non-DhE cases.

DE-DLBCL can be stratified by FN1 expression in the tumor microenvironment

Finally, we analyzed the role of LME in DhE-DLBCL. The mean absolute immunophenotypic cell counts in DhE, non-DhE, and *MYC*/*BCL2*-DH groups are shown by a clustermap (Supplementary Fig. S10A), and Z-score clustermaps based on normalized median and mean immunophenotypic cell counts are shown in Fig. 4A. Consistent with the LME gene-expression signature study (6), *MYC*/*BCL2*-DH exhibited an immune-'depleted' phenotype (significantly lower CD3$^+$, CD4$^+$, CD8$^+$, CD68$^+$, PD-L1$^+$, FOXP3$^+$, and CD45RO$^+$ cells). Similarly, DhE compared with non-DhE cases had significantly lower counts of CD3$^+$, CD4$^+$, PD-1$^+$, FOXP3$^+$, and CD45RO$^+$ cells whereas higher CD20$^+$/PAX5$^+$ cell counts (Fig. 4B); significantly lower T cell proportion in overall cohort (Fig. 4B; significant for both CD4 and CD8 T) and within 'Other'/LymphGen-unclassified subset (significant for CD4 T only, Supplementary Fig. S3D. Representative UMAP plots in Fig. 4A); and significantly lower PD-1$^+$ percentage expression in T cells in overall DLBCL-NOS (significant for CD4 T cells only) and GCB-DLBCL-NOS (Fig. 4B; significant for both CD4 and CD8 T cells). Given the role of MYC in proliferation and recent demonstration of negative associations between immune microenvironment and cell proliferation GEP signatures (6, 9, 11), we analyzed the Ki-67 proliferation index and found that high Ki-67 scores were also significantly associated with lower T cells (but not PD-1 expression) in overall DLBCL (Fig. 4B) and among LymphGen-unclassified cases.

Prognostic analysis in DhE cases found that $FN1^+$ expression, T cell deficiency, NK cell deficiency, PD-1-high expression in T cells, $PD\text{-}L1^+$ expression in T cells, and CTLA-4-negative expression in $CD4^+$ T cells were associated with significantly poorer OS (Fig. 4C, Supplementary Fig. S10B); these effects were independent of MCD/EZB subtypes in which T/NK/PD-1 showed similar effects. After multiple testing correction with a Benjamini-Hochberg FDR threshold of 0.05, only the adverse effect of $FN1^+$ expression on OS was significant (Supplementary Table S7. The Benjamini-Hochberg *P* values for other biomarkers were 0.06).

Prognostic analysis in MYC^{inter}-DE cases found that only FN1-high expression (≥50%) and PD-L1 expression in immune cells (T, NK, and macrophages) had significant adverse prognostic effects (Supplementary Fig. S10C), and only the adverse effect of FN1-high expression on PFS was significant after multiple testing correction (Supplementary Table S7). Also notably, FN1-high expression was associated with significantly poorer PFS in HGBCL-*MYC*/*BCL2*-DH cases (Fig. 4C, small case numbers).

Prognostic analysis in nonDE cases found that FN1-high expression, PD-1-high expression in T cells, and $PD\text{-}L1^+$ expression in T cells and macrophages were associated with significantly poorer OS, whereas high NK cell infiltration and PD-L2 expression in B cells/macrophages were associated with significantly better OS (Supplementary Fig. S10D). The effects of FN1 and PD-L2 expression were significant after multiple testing corrections with the Benjamini-Hochberg procedure (FDR threshold 0.05).

Discussion

DLBCL is a heterogeneous disease. COO and microenvironment transcriptional signatures, genetic driver signatures, and MYC/BCL2 protein double-expression have been used for DLBCL classification or prognostic prediction (3, 43). A negative correlation between proliferation and immune signatures was found (6, 9, 11), making one wonder whether cell proliferation or depleted microenvironment is an ultimate prognostic determinant and effective therapeutic target. However, high-throughput gene signatures often do not precisely inform critical determinants and therapeutic targets. For example, the ‘inflammatory’ LME category associated with unfavorable prognosis does not isolate therapeutic targets from the enriched signatures of tumor-infiltrating lymphocytes (6); the “stromal-1” signature is associated with favorable prognosis but includes genes characteristically expressed in monocytes/histiocytes (44). A study found that mIHC-detected immune checkpoint expression in T cells but not the immune cell transcriptional signature correlated with DLBCL clinical outcome (11). Moreover, gene-expression cutoffs and identification of signatures are often cohort-dependent, not applicable for samples measured individually in the clinic. In this study, we dissected the prognostic role of individual phenotypic LME biomarkers and MYC/BCL2 protein expression within DLBCL genetic subtypes, and further stratified MYC/BCL2-DE cases by genetic subtypes and LME biomarkers. The results overall showed the stratificability of DLBCL subtypes/subsets and context-dependent but independent prognostic effects of phenotypic/genetic biomarkers.

First, we found that in both our cohort and the BCCA validation cohort, phenotypic MYC/BCL2-DE remained to show significant prognostic impact in two LymphGen genetic subsets, EZB and 'Other', but not MCD/ST2 (larger studies are needed for BN2/A53 subtypes), using either the ≥70% cutoff for MYC^{hi} in our cohort (designated as DhE) or ≥40% cutoff for MYC^{+} in the BCCA cohort (designated as DpE). The prognostic effects of DhE could result from both tumor growth/metabolism and antiapoptotic mechanisms as well as COO and microenvironment-related mechanisms in the EZB subtype (8). These results may suggest that MYC/BCL2 is an important prognostic factor and IHC still has clinical utility in the genomic era. On the other hand, genetic subtyping can identify DLBCL patients with poor prognosis independent of MYC/BCL2, such as MCD and A53 subtypes.

Second, we found that LME composition (immune cells and FN1) and immune checkpoint expression had variable prognostic effects within DLBCL genetic subtypes and MYC/BCL2 subgroups, which may suggest rationale therapeutic strategies for clinically unfavorable DLBCL (such as DhE, EZB and MCD). NK cell infiltration showed DhE-independent favorable prognostic effects in EZB, LymphGen-unclassified cases, nonDE, and DhE patients. T cell infiltration showed favorable prognostic effects within BN2, MCD (DhE subset), and DhE cases. In contrast, unfavorable prognostic effects were associated with high levels of T/macrophage/NK cell infiltration within ST2, CD8 T cell/macrophage infiltration among LymphGen-unclassified cases, and T cell densities among nonDE cases (data not shown). Opposite prognostic effects of PD-L1 expression were shown in BN2 (favorable) and ST2/LymphGen-unclassified cases (unfavorable). These results raised a question whether the efficacy of T cell therapy and PD-1/PD-L1 inhibitors varies in different genetic subtypes. Although tumor MYC/BCL2-DhE and Ki-67-high scores were associated with lower T cell proportion and cell densities (data not shown) in overall DLBCL and LymphGen-unclassified (but not EZB) cases, the prognostic effects of DhE were different from those of T cell infiltration in DLBCL genetic subtypes and T cell deficiency could still stratify DhE cases, suggesting that their independent prognostic roles and that the unfavorable prognostic effect of T cell deficiency may depend on DhE. In contrast, PD-1 expression in T cells showed DhE-independent significant adverse prognostic effects in DhE, nonDE, EZB and LymphGen-unclassified cases, while DhE was associated with lower PD-1 expression in DLBCL and EZB/GCB subtypes. FN1 protein expression showed adverse prognostic effects in our cohort largely regardless of MYC/BCL2-DE/DH status and genetic subtypes (however requiring different cutoffs), which resembled the effects associated with fibrotic gene-expression signatures (45) and FN1 protein in solid tumors, suggesting a role of fibronectin or tumor-protective stromal cells in mediating chemo or rituximab-resistance (46). Oppositely, favorable prognostic effects were found to be associated with high *FN1*/FN1 expression based on *FN1* mRNA levels in our and other DLBCL cohorts (44, 47) or FN1 staining intensity of fibrous strands in the extracellular matrix in a previous study (48).

Third, we found that MYC/BCL2-DhE DLBCL could be stratified by *KMT2D* mutations in our cohort and a validation cohort. The poorer survival of EZB (but not MCD) genetic subtype among DhE patients in our cohort and A53/unsubtyped cases among DpE patients in the BCCA cohort was mostly attributable to cases with *KMT2D* mutations. *KMT2D* mutations (and EZB, $P = 0.0027$) were associated with decreased T cells in overall DLBCL

with wild-type *TP53* (35) but not in DhE cases; *KMT2D* mutations (and ST2/EZB/MCD subtypes) were significantly associated with higher numbers of mutated genes in DhE ($P < 0.0001$, data not shown) as in overall DLBCL (35). In solid tumor models, KMT2D and KMT2C are components of a p53 coactivator complex which is required for H3K4-trimethylation and expression of p53-target genes in response to doxorubicin treatment (49). These results, together with the distinct GEP signatures and impact of genetic subtypes on prognostic significance of immune biomarkers, may suggest a role of genetic background in defining prognosis and potential of novel therapies. However, *KMT2D* mutations had no significant prognostic effects in non-DhE patients; DhE patients had poorer survival regardless of *KMT2D* mutation status and genetic subtypes (except for ST2-DhE) compared with overall non-DhE patients; and genetic subtypes showed weaker prognostic effects than the DhE biomarker in our DLBCL cohort. However, we acknowledge that only 30% of our cohort were able to be subtyped by the LymphGen classifier, which is much lower than the previously reported 55-72% prediction rates with proper adaption to the genetic alteration results generated by different platforms (4). Many non-significant trends of prognostic effects of genetic subtypes in our cohort were significant in the BCCA cohort with better subtyping (data not shown). We have tried to add more FISH data including those for *MYC*, *BCL2*, *BCL6*, *PD-L1*/*L2*, *PRDM1*, *PTEN*, *CDKN2A*, *MDM2*, *MDM4*, *CCND1*, and *REL* to LymphGen input files, which however only gave additional composite genetic features (mostly EZB) to the already identified cases presented in this study. Optimizing the LymphGen algorithm for different NGS panels in the future may identify more genetic subsets with homogenous therapeutic vulnerabilities.

In conclusion, our comprehensive analyses indicated that MYC/BCL2-DhE, genetic subtypes, and the tumor microenvironment have independent and interacting roles in defining DLBCL prognosis, and complete molecular assessments will improve DLBCL stratification and precision medicine in clinic. How to classify DLBCL with composite methods that simultaneously evaluate characteristics at the DNA, RNA, and protein levels (50) may be addressed in future studies. This study also gained insights into context-dependent therapeutic targets in DLBCL subsets, including those with unfavorable prognosis with the standard treatment.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by The Duke University Institutional Research Grant Award and the Hagemeister Lymphoma Foundation.

References

1. 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;403:503–511. [PubMed: 10676951]

2. 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;378:1396–1407. [PubMed: 29641966]

3. 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;24:679–690. [PubMed: 29713087]

4. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. Cancer Cell 2020;37:551–568 e514. [PubMed: 32289277]

5. 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;135:1759–1771. [PubMed: 32187361]

6. 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;11:1468–1489. [PubMed: 33541860]

7. Staiger AM, Altenbuchinger M, Ziepert M, Kohler C, Horn H, Huttner M, et al. A novel lymphoma-associated macrophage interaction signature (LAMIS) provides robust risk prognostication in diffuse large B-cell lymphoma clinical trial cohorts of the DSHNHL. Leukemia 2020;34:543–552. [PubMed: 31530861]

8. Tripodo C, Zanardi F, Iannelli F, Mazzara S, Vegliante M, Morello G, et al. A spatially resolved dark- versus light-zone microenvironment signature subdivides germinal center-related aggressive B cell lymphomas. iScience 2020;23:101562. [PubMed: 33083730]

9. Sarver AL, Xie C, Riddle MJ, Forster CL, Wang X, Lu H, et al. Retinoblastoma tumor cell proliferation is negatively associated with an immune gene expression signature and increased immune cells. Lab Invest 2021;101:701–718. [PubMed: 33658609]

10. Cuadros M, Dave SS, Jaffe ES, Honrado E, Milne R, Alves J, et al. Identification of a proliferation signature related to survival in nodal peripheral T-cell lymphomas. J Clin Oncol 2007;25:3321–3329. [PubMed: 17577022]

11. Autio M, Leivonen SK, Brück O, Mustjoki S, Mészáros Jørgensen J, Karjalainen-Lindsberg ML, et al. Immune cell constitution in the tumor microenvironment predicts the outcome in diffuse large B-cell lymphoma. Haematologica 2021;106:718–729. [PubMed: 32079690]

12. Johnson NA, Slack GW, Savage KJ, Connors JM, Ben-Neriah S, Rogic S, et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol 2012;30:3452–3459. [PubMed: 22851565]

13. Hu S, Xu-Monette ZY, Tzankov A, Green T, Wu L, Balasubramanyam 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;121:4021–4031; quiz 4250. [PubMed: 23449635]

14. Green TM, Young KH, Visco C, Xu-Monette ZY, Orazi A, Go RS, et al. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol 2012;30:3460–3467. [PubMed: 22665537]

15. Staiger AM, Ziepert M, Horn H, Scott DW, Barth TFE, Bernd HW, et al. Clinical impact of the cell-of-origin classification and the MYC/ BCL2 dual expresser status in diffuse large B-cell lymphoma treated within prospective clinical trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group. J Clin Oncol 2017;35:2515–2526. [PubMed: 28525305]

16. Swerdlow SH, Campo E, Pileri SA, Harris NL, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. Blood 2016;127:2375–2390. [PubMed: 26980727]

17. Xu-Monette ZY, Dabaja BS, Wang X, Tu M, Manyam GC, Tzankov A, et al. Clinical features, tumor biology, and prognosis associated with MYC rearrangement and Myc overexpression in diffuse large B-cell lymphoma patients treated with rituximab-CHOP. Mod Pathol 2015;28:1555–1573. [PubMed: 26541272]

18. Scott DW, Mottok A, Ennishi D, Wright GW, Farinha P, Ben-Neriah S, et al. Prognostic significance of diffuse large B-cell lymphoma cell of origin determined by digital gene expression in formalin-fixed paraffin-embedded tissue biopsies. J Clin Oncol 2015;33:2848–2856. [PubMed: 26240231]
19. Meriranta L, Pasanen A, Alkodsi A, Haukka J, Karjalainen-Lindsberg ML, Leppä S. Molecular background delineates outcome of double protein expressor diffuse large B-cell lymphoma. Blood Adv 2020;4:3742–3753. [PubMed: 32780847]
20. Aukema SM, Siebert R, Schuuring E, van Imhoff GW, Kluin-Nelemans HC, Boerma EJ, et al. Double-hit B-cell lymphomas. Blood 2011;117:2319–2331. [PubMed: 21119107]
21. Jaffe ES, Barr PM, Smith SM. Understanding the new WHO classification of lymphoid malignancies: Why it's important and how it will affect practice. Am Soc Clin Oncol Educ Book 2017;37:535–546. [PubMed: 28561690]
22. 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;131:2060–2064. [PubMed: 29475959]
23. Oki Y, Noorani M, Lin P, Davis RE, Neelapu SS, Ma L, et al. Double hit lymphoma: the MD Anderson Cancer Center clinical experience. Br J Haematol 2014;166:891–901. [PubMed: 24943107]
24. Friedberg JW. How I treat double-hit lymphoma. Blood 2017;130:590–596. [PubMed: 28600333]
25. Savage KJ, Slack GW, Mottok A, Sehn LH, Villa D, Kansara R, et al. Impact of dual expression of MYC and BCL2 by immunohistochemistry on the risk of CNS relapse in DLBCL. Blood 2016;127:2182–2188. [PubMed: 26834242]
26. Johnson NA, Savage KJ, Ludkovski O, Ben-Neriah S, Woods R, Steidl C, et al. Lymphomas with concurrent BCL2 and MYC translocations: the critical factors associated with survival. Blood 2009;114:2273–2279. [PubMed: 19597184]
27. Horn H, Ziepert M, Becher C, Barth TF, Bernd HW, Feller AC, et al. MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. Blood 2013;121:2253–2263. [PubMed: 23335369]
28. Davies A, Cummin TE, Barrans S, Maishman T, Mamot C, Novak U, et al. Gene-expression profiling of bortezomib added to standard chemoimmunotherapy for diffuse large B-cell lymphoma (REMoDL-B): an open-label, randomised, phase 3 trial. Lancet Oncol 2019;20:649–662. [PubMed: 30948276]
29. Horn H, Ziepert M, Wartenberg M, Staiger AM, Barth TF, Bernd HW, et al. Different biological risk factors in young poor-prognosis and elderly patients with diffuse large B-cell lymphoma. Leukemia 2015;29:1564–1570. [PubMed: 25687653]
30. Barraclough A, Alzahrani M, Ettrup MS, Bishton M, van Vliet C, Farinha P, et al. COO and MYC/BCL2 status do not predict outcome among patients with stage I/II DLBCL: a retrospective multicenter study. Blood Adv 2019;3:2013–2021. [PubMed: 31285189]
31. Ambrosio MR, Lazzi S, Bello GL, Santi R, Porro LD, de Santi MM, et al. MYC protein expression scoring and its impact on the prognosis of aggressive B-cell lymphoma patients. Haematologica 2019;104:e25–e28. [PubMed: 29954940]
32. Ziepert M, Lazzi S, Santi R, Vergoni F, Granai M, Mancini V, et al. A 70% cut-off for MYC protein expression in diffuse large B cell lymphoma identifies a high-risk group of patients. Haematologica 2020;105:2667–2670. [PubMed: 33131258]
33. Visco C, Li Y, Xu-Monette ZY, Miranda RN, Green TM, Li Y, et al. Comprehensive gene expression profiling and immunohistochemical studies 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;26:2103–2113. [PubMed: 22437443]
34. Tzankov A, Xu-Monette ZY, Gerhard M, Visco C, Dirnhofer S, Gisin N, et al. Rearrangements of MYC gene facilitate risk stratification in diffuse large B-cell lymphoma patients treated with rituximab-CHOP. Mod Pathol 2014;27:958–971. [PubMed: 24336156]

35. You H, Xu-Monette ZY, Wei L, Nunns H, Nagy ML, Bhagat G, et al. Genomic complexity is associated with epigenetic regulator mutations and poor prognosis in diffuse large B-cell lymphoma. Oncoimmunology 2021;10:1928365. [PubMed: 34350060]
36. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-wide copy number detection and visualization from targeted DNA sequencing. PLoS Comput Biol 2016;12:e1004873. [PubMed: 27100738]
37. Xu-Monette ZY, Xiao M, Au Q, Padmanabhan R, Xu B, Hoe N, et al. Immune profiling and quantitative analysis decipher the clinical role of immune-checkpoint expression in the tumor immune microenvironment of DLBCL. Cancer Immunol Res 2019;7:644–657. [PubMed: 30745366]
38. Leland M, John H, Nathaniel S, Lukas G. UMAP: Uniform Manifold Approximation and Projection. J Open Source Softw 2018;3(29):861.
39. Waskom ML. seaborb: statistical data visualization. J Open Source Softw 2021;6(60):3021.
40. Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. Genome Biol 2003;4:R70. [PubMed: 14519205]
41. Ennishi D, Jiang A, Boyle M, Collinge B, Grande BM, Ben-Neriah S, 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;37:190–201. [PubMed: 30523716]
42. Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, et al. Genetic and functional drivers of diffuse large B cell lymphoma. Cell 2017;171:481–494 e415. [PubMed: 28985567]
43. Wright GW, Wilson WH, Staudt LM. Genetics of diffuse large B-cell lymphoma. N Engl J Med 2018;379:493–494.
44. 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;359:2313–2323. [PubMed: 19038878]
45. 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;39:845–865 e847. [PubMed: 34019806]
46. Mraz M, Zent CS, Church AK, Jelinek DF, Wu X, Pospisilova S, et al. Bone marrow stromal cells protect lymphoma B-cells from rituximab-induced apoptosis and targeting integrin α−4-β−1 (VLA-4) with natalizumab can overcome this resistance. Br J Haematol 2011;155:53–64. [PubMed: 21749361]
47. Lossos IS, Czerwinski DK, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D, et al. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med 2004;350:1828–1837. [PubMed: 15115829]
48. Brandt S, Montagna C, Georgis A, Schüffler PJ, Bühler MM, Seifert B, et al. The combined expression of the stromal markers fibronectin and SPARC improves the prediction of survival in diffuse large B-cell lymphoma. Exp Hematol Oncol 2013;2:27. [PubMed: 24499539]
49. Lee J, Kim DH, Lee S, Yang QH, Lee DK, Lee SK, et al. A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. Proc Natl Acad Sci U S A 2009;106:8513–8518. [PubMed: 19433796]
50. Xu-Monette ZY, Zhang H, Zhu F, Tzankov A, Bhagat G, Visco C, et al. A refined cell-of-origin classifier with targeted NGS and artificial intelligence shows robust predictive value in DLBCL. Blood Adv 2020;4:3391–3404. [PubMed: 32722783]

Translational relevance

DLBCL is a pathogenically heterogeneous disease and can be stratified by either high-throughput genomic signatures or cost-effective immunohistochemistry-based biomarkers such as MYC/BCL2 double protein expression. How these stratified subgroups overlap and whether one stratification method is superior to the other for prognostic and therapeutic prediction are elusive. This study analyzed the associations and relative prognostic values of LymphGen genetic subtypes, MYC/BCL2, and tumor-microenvironmental biomarkers in a large cohort of DLBCL. Results showed heterogeneity and shared characteristics at the DNA, RNA, and protein levels within DLBCL LymphGen genetic subtypes, cell-of-origin subtypes, and MYC/BCL2 subgroups, and identified significant independent prognostic biomarkers within DLBCL subsets stratified by different methods. These results are important for understanding the prognostic determinants in DLBCL, improving DLBCL classification, and developing immunotherapies and targeted therapies.

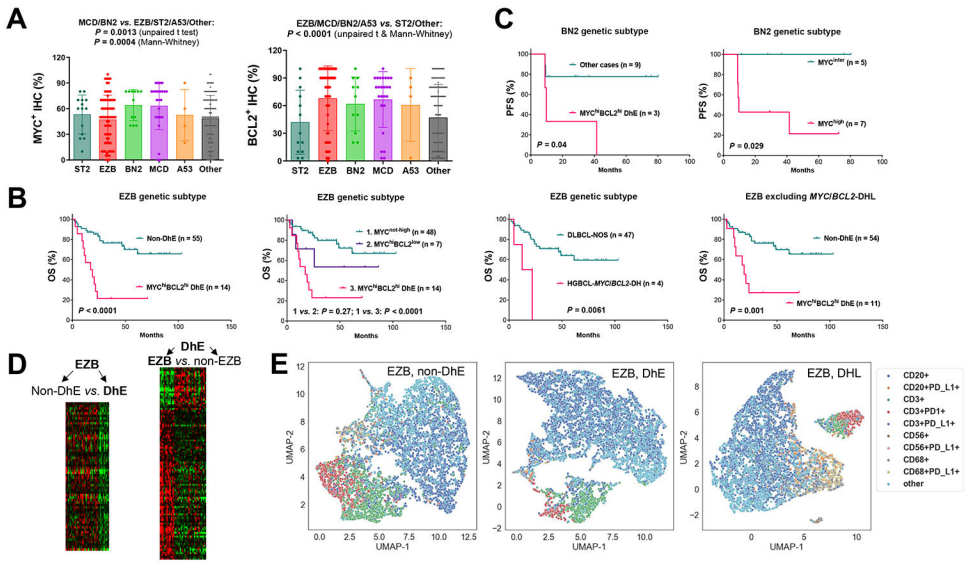

Fig. 1. Heterogeneity and prognostic effects of MYC/BCL2 expression in DLBCL genetic subsets.

(**A**) Combined MCD and BN2 cases compared with combined EZB, ST2, A53, and 'Other' subtype cases had significantly higher MYC expression evaluated by immunohistochemistry (IHC). Combined EZB, MCD, BN2, and A53 cases compared with combined ST2 and 'Other' subtype cases had significantly higher BCL2 expression. Each dot represents one patient. *P* values are 2-tailed (unpaired *t* and Mann-Whitney) and exact (Mann-Whitney). (**B**) MYC/BCL2 double-high-expression (DhE) and *MYC*/*BCL2* genetic double-hit (DH) had significant adverse impact on overall survival (OS) in the EZB genetic subtype. Single MYC^{hi} (without concurrent $BCL2^{hi}$) expression did not have significant prognostic effect in EZB. (**C**) DhE and high MYC expression were associated with significantly poorer progression-free survival (PFS) in the BN2 genetic subset. (**D**) Heatmaps for significantly differentially expressed genes between DhE and non-DhE cases in the EZB subset and between EZB and non-EZB cases in the DhE subgroup. (**E**) Representative Uniform Manifold Approximation and Projection (UMAP) plots generated from single-cell intensities for CD20, CD3, CD68, CD56, PD-1, and PD-L1 markers in three EZB cases with non-DhE, DhE, and *MYC*/*BCL2*-DH, respectively. Each datapoint represents a cell, labeled according to phenotype. In the legends, PD-L1/PD-1-negative phenotypes in $CD20^+$, $CD3^+$, $CD56^+$, and $CD68^+$ were omitted.

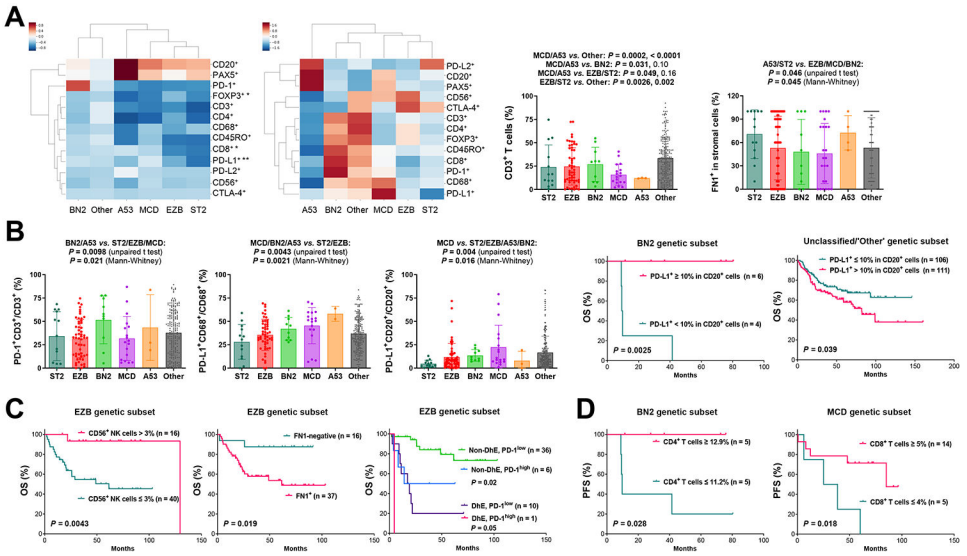

Fig. 2. Heterogeneity and prognostic effects of immune/microenvironment biomarkers in DLBCL genetic subsets.

(**A**) Left, a Z-score clustermap based on median immunofluorescence counts; the Z-score was computed for each phenotype count across all samples, and then the median Z-score was computed within each group. Significant differences between MCD and EZB subtypes by two-sided Mann-Whitney U test are marked by asterisks. *: $P \leq 0.05$, **: $P \leq 0.01$. Middle, a Z-score clustermap to show clusters based on mean cell counts for each group; the mean phenotype count was computed within each group, and then the Z-score was computed across all groups. Right, two scatter plots for T cell percentage and FN1 expression levels, respectively. T cell percentage was calculated by T cell counts divided by total counts of B cells, T cells, macrophages, and NK cells; two *P* values for each comparison are by unpaired *t* test and Mann-Whitney test, respectively, and statistically significant *P* values are in bold. (**B**) Left, three scatter plots for PD-1 expression in T cells and PD-L1 expression in macrophages and B cells in DLBCL genetic subtypes. In all scatter plots, each dot represents one patient; *P* values are 2-tailed (unpaired *t* and Mann-Whitney) and exact (Mann-Whitney). Right, PD-L1 expression in B-cells was significantly associated with better overall survival (OS) in the BN2 genetic subset and worse OS in genetically unsubtyped cases. (**C**) In the EZB genetic subset, NK-cell infiltration was associated with significantly better OS, whereas FN1 expression in the tumor microenvironment and PD-1-high expression in T cells (cutoff, 55%) were associated with significantly poorer OS independent of MYC/BCL2 DhE status. (**D**) Higher CD4 T cell infiltration and CD8 T cell infiltration levels were associated with significantly better progression-free survival (PFS) within the BN2 and MCD genetic subset, respectively.

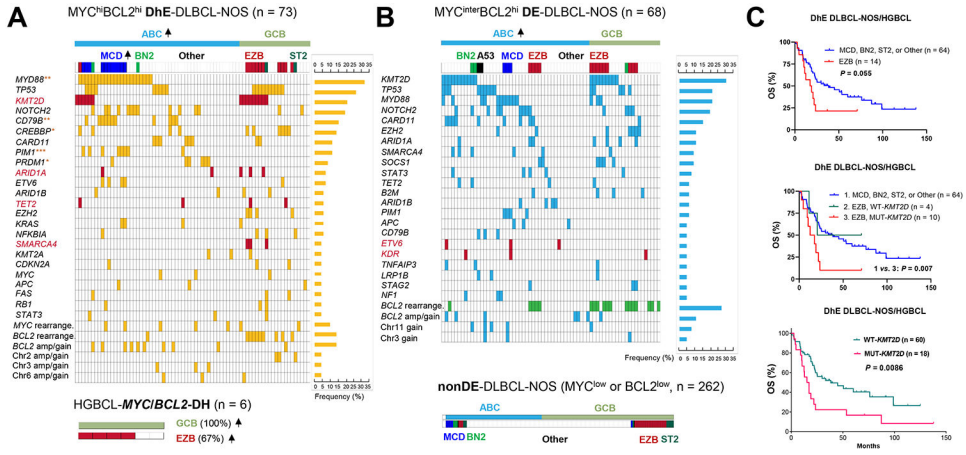

Fig. 3. Molecular and genetic analysis in DLBCL, not otherwise specified (NOS) stratified by MYC/BCL2 expression status and high grade B-cell lymphoma (HGBCL) with *MYC*/*BCL2* double-hit (DH).

(**A**) Distribution plot for DLBCL-NOS with MYC/BCL2 double-high expression (DhE) and HGBCL-*MYC*/*BCL2*-DH cases. (**B**) Distribution plots in DLBCL-NOS cases with MYC-intermediate expression (MYC^{inter}) and BCL2-high expression (MYC/BCL2 double-expression, DE) and nonDE-DLBCL-NOS cases. Each column in the oncoplots represents one patient. Each row shows distribution of a genetic alteration with positive cases highlighted and prevalence shown on the right. Genetic alterations shown in the plots include non-synonymous mutations and copy number alterations detected by targeted next-generation sequencing and *MYC* and *BCL2* alterations detected by fluorescence *in situ* hybridization occurring in ≥3 patients. Enriched genes in DhE with significantly higher mutation frequencies in DhE than in MYC^{inter}-DE cases or nonDE cases by Fisher's exact test are marked by red asterisks. *: $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$. Up arrows next to the ABC/GCB molecular subtypes and MCD/EZB genetic subtypes indicate their increased frequencies in the subgroup (significance was determined by Fisher's exact test). *KMT2D*, *ARID1A*, *TET2*, *SMARCA4*, *ETV6*, and *KDR* are highlighted in red to indicate that their mutations had significant adverse prognostic effects in DE-DLBCL cases. The distribution of *BCL2* gene rearrangement in MYC^{inter}-DE-DLBCL-NOS cases is highlighted in green, to indicate its non-significant association with a better overall survival (OS, $P = 0.069$). Abbreviations: rearrange., rearrangement; amp, amplification. (**C**) In DhE patients, EZB subtype and *KMT2D* mutation were associated with poorer OS with a statistically non-significant and significant *P* value, respectively. DhE patients with both EZB subtype and *KMT2D* mutations had significantly poorer OS compared with DhE patients with non-EZB subtypes.

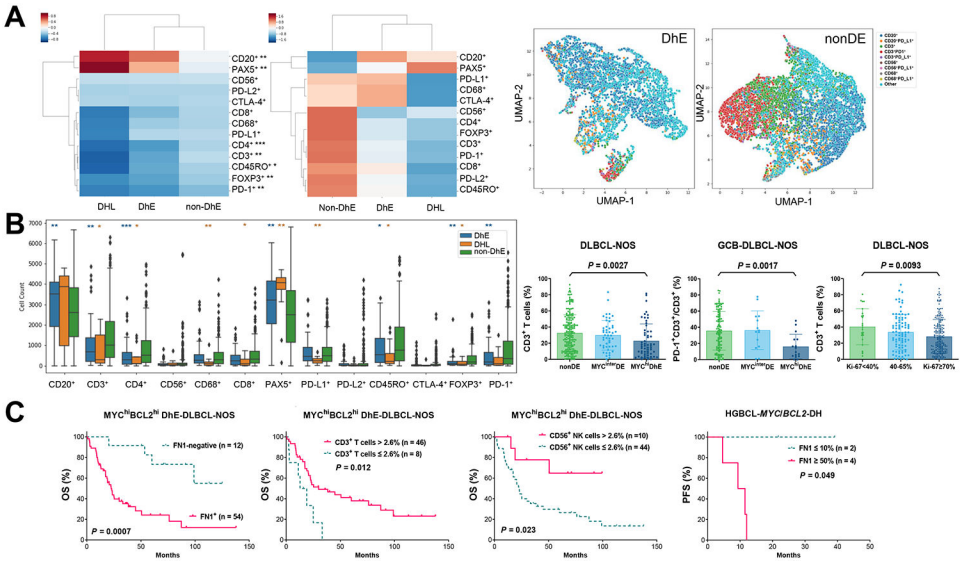

Fig. 4. immune microenvironment analysis for MYC/BCL2 double-high expression (DhE) and *MYC*/*BCL2* double-hit (DH) in DLBCL and HGBCL.

(**A**) Left, Z-score clustermaps based on median and mean cell counts in DhE, non-DhE, and *MYC*/*BCL2* double-hit lymphoma (DHL) cases, respectively; right, representative Uniform Manifold Approximation and Projection (UMAP) plots generated from single-cell intensities for six markers (CD20, CD3, CD68, CD56, PD-1, and PD-L1) in two LymphGen-unclassified cases. UMAP legends of $CD20^+$, $CD3^+$, $CD56^+$, and $CD68^+$ meant single positive. (**B**) Left, a box plot to show the distribution and significant differences in absolute cell counts of immunophenotypes between DhE/DH and non-DhE cases by two-sided Mann-Whitney U test. Asterisks mark significant differences. *: $P \leq 0.05$, **: $P \leq 0.01$, ***: $P \leq 0.001$. Significant differences between DhE and non-DhE cases are also marked by asterisks in the median Z-score clustermap above. Right, three scatter plots showing that DhE and high Ki-67 scores were significantly associated with lower T cell percentages in overall DLBCL-NOS and that DhE was associated with lower PD-1 expression in T cells in GCB-DLBCL-NOS cases. *P* values are by 2-tailed unpaired *t* test (also significant by Mann-Whitney U test). Each dot represents one patient. (**C**) In DhE-DLBCL-NOS and *MYC*/*BCL2*-DHL cases, FN1-positive and FN1-high expression in the tumor microenvironment were associated with significantly poorer overall survival (OS) and progression-free survival (PFS), respectively. In DhE-DLBCL-NOS, deficiency in T cell and NK cell infiltration was associated with significantly poorer OS.

Table 1.

Gene-expression analysis for MYC/BCL2 double-high expression (DhE) in DLBCL within the EZB genetic subset and LymphGen-unclassified cases (false discovery rate 0.10)

Gene-expression in DhE compared with non-DhE	
In the EZB subset	
Gene name	Downregulated: *CLECL1*, *GZMK*, *TRBC1*, *SH3BGRL3*, *IL10RA*, *PFN1*, *DENND4B*, *RGS1*, *HLA-B*, *RASSF2*, *UHMK1*, *B2M*, *SCAMP2*, *COMMD3*, *RNF19A*, *NRBP1*, *BCL11B*, *BTN3A1*, *HLA-F*, *HLA-E*, *RASAL3*, *ANKRD12*, *USP32*, *CEBPG*, *MKRN1*, *UBXN4*, *MPPE1*, *LCP2*, *BTN2A2*, *CALM2*, *CLIP4*, *ATXN7*, *CHIC1*, *CPEB2*, *STAMBPL1*, *RFFL*, *SLC4A1AP*, *CSNK1D*, *NBR1*, *IDS*, *CD84*, *TFPT*, *ZCCHC6*, *GTDC1*, *AKAP13*, *CYLD*, *ASXL2*, *CHTOP*, *TDRD3*, *TGOLN2*, *AAK1*, *LOC401320*
Gene category of GO terms	GO Biological Process: "antigen presentation\, endogenous antigen″, "antigen processing\, endogenous antigen via MHC class I″, "antigen presentation″, "antigen processing", "defense response″, "response to biotic stimulus″, "immune response″, "response to external stimulus″ GO Cellular Component: "membrane″, "cell″ GO Molecular Function: "MHC class I receptor activity″, "receptor activity″
In LymphGen-unclassified cases	
Gene name	Upregulated, fold change ≥1.5: *MYC*, *KIAA0226L*, *PMAIP1*, *CLECL1*, *FUT8*, *PUS7*, *PIGW*, *TMEM97*, *BCL2*, *HIST1H2AC*, *ADTRP*, *ZNF385C*, *MRPL3*, *FABP5*, *IGF2BP3*, *SNHG1*, *CCDC86*, *NOC3L*, *P2RX5*, *CYB5R2*, *RPL17*, *EBNA1BP2*, *FOXC1*, *CCDC113*, *ZNF260*, *ZNF320*, *ARHGAP24*, *SPIB*, *PSORS1C2*, *NOL11*, *SLC25A32*, *DBN1*, *NOP2*, *DCTPP1*, *GCSH*, *GAS5*, *ZNF587B*, *PDCD2L*, *TMCC3*, *GAR1*, *TRIT1*, *BTBD19*, *LOC101928702*, *FSTL5*, *VSNL1*, *POLR3G*, *NME2*, *CISD1*, *DDX21*, *C10orf2*, *IPO4*, *C12orf45*, *POLR1D*, *APEX1*, *DIRF-AS1*, *MRPS33*, *GPX7*, *RUVBL2*, *PRMT5*, *POLR2H*, *MAT2A*, *ZNF639*, *NPM3*, *TEX10*, *IER3IP1*, *CDC123*, *PLEKHJ1*, *LYAR*, *BMF*, *ATP5G1*, *EME1*, *TOMM40*, *PSME3*, *ZNF784*, *NAT9*, *C7orf41*, *SNHG8*, *TCTN3*, *ZNF480*, *CLPX*, *LOC100507388*, *MSRB1*, *PNP*, *PLIN1*, *TIMM10B*
Gene category of GO terms	GO Biological Process: "rRNA metabolism″, "RNA metabolism″, "ribosome biogenesis″, "ribosome biogenesis and assembly″, "RNA processing″, "rRNA processing″, "cell growth and/or maintenance″ GO Cellular Component: "nucleolus″, "mitochondrion", "intracellular", "mitochondrial membrane", "RNA polymerase complex", "mitochondrial inner membrane", "inner membrane", "cytoplasm", "outer membrane", "organellar ribosome", "mitochondrial ribosome", "nucleus", "DNA-directed RNA polymerase I complex", "proteasome complex (sensu Eukarya)" GO Molecular Function: "DNA-directed RNA polymerase activity", "RNA binding", "nucleotidyltransferase activity", "transferase activity", "lipid binding", "nucleic acid binding"

Abbreviation: GO, Gene Oncology.

Enriched Go categories were identified by Expression Analysis Systematic Explorer (EASE, A desktop version of DAVID) with EASE scores <0.05.

Table 2.

Significantly differentially expressed genes in EZB compared with non-EZB genetic subtypes in DLBCL with MYC/BCL2 double-high expression

	Downregulated	Upregulated
Gene name	FDR 0.10: *TNFRSF13B*, *CLECL1*, *PHF16*, *P2RX5*, *ERP29*, *SLA*; FDR 0.15: *PIM2*, *POP4*, *HCK*, *SLC5A6*, *PLEKHO1*, *STAMBPL1*, *ADAM8*, *DLGAP1-AS1*, *LRRC33*, *IL10RA*, *LIMD1*, *POU2F2*, *FOXP1*, *PARP15*, *DOCK10*, *TGIF1*, *IRF4*	FDR 0.10: *STAP1*, *MME*, *SPON1*, *TNKS*, *PTK2*, *CCNG2*, *S1PR2*, *ICOSLG*, *RRAS2*, *FAM208B*, *SWAP70*, *MAML3*, *ZFAND4*, *PLEKHF2*, *FAM134B*, *CDK14*, *PACS1*, *NDUFAF6*, *BRWD1*, *RASL11A*, *LOC101929456*, *RAPGEF5*, *LRMP*, *ATPIF1*, *SLC25A27*, *CPNE3*, *HGSNAT*, *DEF8*, *MYBL1*, *LOC286149*, *GFOD1*, *SSBP3*, *SPINK2*, *LHPP*, *MARCKSL1*, *PCDHGC3*, *WEE1*, *LONRF1*, *HIP1R*, *BRIP1*, *REL*, *MFHAS1*, *ASB13*, *HERC2P3*, *SLC30A4*, *STK17A*, *PALD1*, *SERPINA9*, *SPIRE2*, *PDK3*, *SSBP2*, *ENPP3*, *C8orf37*, *ITPKB*, *FLJ31485*, *SLC24A3*, *SEL1L3*, *FADS3*, *KIF3A*, *TP53INP1*, *CLIC4*, *COL14A1*, *ZDHHC2*, *ANKH*, *RNGTT*, *PQLC2*, *TSPAN15*, *CCDC85A*, *NAP1L3*, *GNA13*, *FNDC1*, *GBA3*, *USP34*, *CENPM*, *ITSN2*, *ARL14EP*, *VCL*, *LOC100131581*, *ARL2BP*, *IZUMO4*, *RNF8*
Gene category of GO terms	GO Biological Process: "regulation of transcription\, DNA-dependent″, "regulation of transcription″, "transcription\, DNA-dependent″	GO Biological Process: "intracellular signaling cascade″ GO Molecular Function: ″kinase activity″, "transferase activity\, transferring phosphorus-containing groups″

Abbreviations: FDR, false discovery rate; GO, Gene Oncology.

Enriched Go categories were identified by Expression Analysis Systematic Explorer (EASE, A desktop version of DAVID) with EASE scores <0.05