ONCOIMMUNOLOGY
2021, VOL. 10, NO. 1, e1928365 (12 pages)
https://doi.org/10.1080/2162402X.2021.1928365

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

OPEN ACCESS Check for updates

Genomic complexity is associated with epigenetic regulator mutations and poor prognosis in diffuse large B-cell lymphoma

Hua You[a,b], Zijun Y. Xu-Monette[b], Li Wei[a,b], Harry Nunns[c], Máté L. Nagy[c], Govind Bhagat[d], Xiaosheng Fang[b], Feng Zhu[b], Carlo Visco[e], Alexandar Tzankov[f], Karen Dybkaer[g], April Chiu[h], Wayne Tam[i], Youli Zu[j], Eric D. Hsi[k], Fredrick B. Hagemeister[l], Jooryung Huh[m], Maurilio Ponzoni[n], Andrés J.M. Ferreri[n], Michael B. Møller[o], Benjamin M. Parsons[p], J. Han Van Krieken[q], Miguel A. Piris[r], Jane N. Winter[s], Yong Li[t], Qingyan Au[c], Bing Xu[u], Maher Albitar[v], and Ken H. Young[b,w]

[a]Department of Hematology and Oncology, Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou, China; [b]Hematopathology Division and Department of Pathology, Duke University Medical Center, Durham, North Carolina, USA; [c]NeoGenomics Laboratories, Aliso Viejo, California, USA; [d]Department of Pathology and Cell Biology, Columbia University Irving Medical Center and New York Presbyterian Hospital, New York, New York, USA; [e]Department of Medicine, Section of Hematology, University of Verona, Verona, Italy; [f]Department of Pathology, Institute of Pathology,University Hospital Basel, Switzerland; [g]Clinical Department, Aalborg University Hospital, Aalborg, Denmark; [h]Hematopathology Department, Mayo Clinic, Rochester, Minnesota, USA; [i]Department of Pathology, Weill Medical College of Cornell University, New York, New York, USA; [j]Department of Pathology and Genomic Medicine, The Methodist Hospital, Houston, Texas, USA; [k]Department of Pathology, Cleveland Clinic, Cleveland, Ohio, USA; [l]Department of Lymphoma and Myeloma, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA; [m]Department of Pathology, Asan Medical Center, Ulsan University College of Medicine, Seoul, Korea; [n]Lymphoma Unit, Department of Onco-Hematology, IRCCS San Raffaele Scientific Institute, Milan, Italy; [o]Department of Pathology, Odense University Hospital, Odense, Denmark; [p]Hematology & Oncology, Gundersen Lutheran Health System, La Crosse, Wisconsin, USA; [q]Department of Pathology, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands; [r]Pathology Department, Hospital Universitario Marqués de Valdecilla, Santander, Spain; [s]Department of Medicine (Hematology and Oncology), Feinberg School of Medicine, Northwestern University, Chicago, Illinois,USA; [t]Department of Medicine, Baylor College of Medicine, Houston, Texas,USA; [u]Department of Hematology, The First Affiliated Hospital of Xiamen University, Xiamen, Fujian,China; [v]Genomic Testing Cooperative, LCA, Irvine, California, USA; [w]Duke Cancer Institute, Durham, North Caronlina, USA

ABSTRACT

Diffuse large B-cell lymphoma (DLBCL) is the most common type of lymphoma with high mutation burdens but a low response rate to immune checkpoint inhibitors. In this study, we performed targeted next-generation sequencing and fluorescent multiplex immunohistochemistry, and investigated the clinical significance and immunological effect of mutation numbers in 424 DLBCL patients treated with standard immunochemotherapy. We found that *KMT2D* and *TP53* nonsynonymous mutations (MUT) were significantly associated with increased nonsynonymous mutation numbers, and that high mutation numbers (MUT^{high}) were associated with significantly poorer clinical outcome in germinal center B-cell-like DLBCL with wild-type *TP53*. To understand the underlying mechanisms, we identified a gene-expression profiling signature and the association of MUT^{high} with decreased T cells in DLBCL patients with wild-type *TP53*. On the other hand, in overall cohort, MUT^{high} was associated with lower PD-1 expression in T cells and PD-L1 expression in macrophages, suggesting a positive role of MUT^{high} in immune responses. Analysis in a whole-exome sequencing dataset of 304 patients deposited by Chapuy et al. validated the correlation of MUT-*KMT2D* with genomic complexity and the significantly poorer survival associated with higher numbers of genomic single nucleotide variants in activated B-cell–like DLBCL with wild-type *TP53*. Together, these results suggest that KMT2D inactivation or epigenetic dysregulation has a role in driving DLBCL genomic instability, and that genomic complexity has adverse impact on clinical outcome in DLBCL patients with wild-type *TP53* treated with standard immunochemotherapy. The oncoimmune data in this study have important implications for biomarker and therapeutic studies in DLBCL.

ARTICLE HISTORY
Received 26 February 2021
Revised 6 May 2021
Accepted 6 May 2021

KEYWORDS
Tumor mutation burden; KMT2D; genomic instability; tumor microenvironment; PD-1; PD-L1; TP53; epigenetic; DLBCL; INDEL

Introduction

Nonsynonymous somatic mutations can not only contribute to cancer development but also produce tumor-specific neoantigens eliciting antitumor immune responses by the host. Tumor mutation burden (TMB) is a predictive biomarker for immune checkpoint blockade-based immunotherapy within and across multiple types of solid tumors.[1–3] Particularly, TMB of small insertions and deletions (INDELs) correlates with response to PD-1 blockade[4] and prognosis with other therapies in certain types of solid tumors,[5] with the hypothesis that INDELs may result in immunogenic neoantigens more often than single nucleotide variants (SNVs) do. TMB can be measured through whole-exome sequencing (WES) as the gold standard.

CONTACT Ken H. Young ken.young@duke.edu Duke University Medical Center, Division of Hematopathology and Department of Pathology, Duke Cancer Institute, Durham, NC 27710, USA; Hua You youhua307@163.com Affiliated Cancer Hospital & Institute of Guangzhou Medical University, Guangzhou, China; Maher Albitar malbitar@genomictestingcooperative.com Genomic Testing Cooperative LCA, 175 Technology Drive, Irvine, CA 92618, USA.
[†]These authors have contributed equally to this work and share first authorship.
This article has been republished with minor changes. These changes do not impact the academic content of the article.
Supplemental data for this article can be accessed on the publisher's website.

© 2021 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

However, WES is not practical in routine clinic, and studies have shown that TMB obtained from targeted next-generation sequencing (NGS) panels has high concordance with WES-derived TMB and significant predictive value for immunotherapy efficacy.[6–8]

Hematologic cancers generally have lower TMBs than solid tumors. The most common aggressive B-cell lymphoma diffuse large B-cell lymphoma (DLBCL) has a significantly higher TMB than chronic lymphocytic leukemia and acute myeloid leukemia as measured by WES (median, ~3 vs ~0.8 and 0.37 non-silent coding mutations per Mb, respectively).[9] However, WES of flow cytometry-sorted Hodgkin Reed–Sternberg cells found that Epstein-Barr virus-negative classical Hodgkin lymphoma (cHL) had a high median TMB (~9 mutations/Mb) which is comparable to that of lung squamous cell carcinoma.[10] Another study comprehensively profiled mutations in the exonic regions of 315 cancer-related genes, and showed that DLBCL is one of the TMB-high cancer types with a median of 10 synonymous/nonsynonymous mutations per Mb, and 18.4% of DLBCL patients showed a high TMB (>20 mutations/Mb of coding genome).[6] PD-1 blockade immunotherapy has very high efficacy in relapsed/refractory cHL in line with the high TMB,[11–14] but not in relapsed/refractory DLBCL despite the overall high TMB in DLBCL (response rate <10% in patients who were ineligible for or having failed autologous hematopoietic cell transplantation).[15] It is unknown whether relapsed/refractory DLBCL patients have lower TMBs than other DLBCL patients which may explain the low efficacy of PD-1 inhibitors in these patients.

To understand the clinical implication of TMB in DLBCL, in this study we performed NGS targeting 275 genes that are frequently mutated in hematologic neoplasms for 444 de novo DLBCL diagnostic samples, and analyzed the clinical impact and biological correlations of mutation numbers in our cohort and a publicly available WES dataset.[16] Patients were all treated with the standard first-line immunochemotherapy (rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone, R-CHOP). Rituximab (anti-CD20 antibody) addition to the CHOP regimen has significantly improved the clinical outcome of DLBCL, likely attributable to anti-CD20 antibody-dependent cellular cytotoxicity, complement-dependent cytotoxicity, and induction of apoptosis.[17,18]

Materials and methods

Patients

444 adult patients with *de novo* DLBCL were sequenced and 424 cases were included for final analysis in this study as part of the DLBCL Consortium Study Program.[19] Primary cutaneous DLBCL, primary mediastinal large B-cell lymphoma, and primary central nervous system lymphoma have been excluded. This study was conducted in accordance with the principles of the Declaration of Helsinki. Data collection protocols were approved as being of minimal to no risk or as exempt by the institutional review board of each participating institution.

Table 1. List of 275 genes in the NGS panel.

ABL1	*BIRC3*	*CREBBP*	*EZH2*	*GNAS*	*KMT2C*	*NF2*	*PPP2R1A*	*SMC3*	*XPO1*
ACVR1B	*BLM*	*CRLF2*	*FAM175A*	*GREM1*	*KMT2D*	*NFE2L2*	*PRDM1*	*SMO*	*XRCC2*
AKT1	*BRAF*	*CSF1R*	*FAM46C*	*GRIN2A*	*KRAS*	*NFKBIA*	*PRKAR1A*	*SOCS1*	*XRCC3*
AKT2	*BRCA1*	*CSF3R*	*FANCA*	*H3F3A*	*LRP1B*	*NKX2-1*	*PRKDC*	*SOX2*	*ZNF217*
AKT3	*BRCA2*	*CTCF*	*FANCC*	*HGF*	*MAP2K1*	*NOTCH1*	*PRSS1*	*SOX9*	*ZRSR2*
ALK	*BRIP1*	*CTNNA1*	*FANCD2*	*HIST1H3B*	*MAP2K2*	*NOTCH2*	*PTCH1*	*SPOP*	
AMER1	*BTK*	*CTNNB1*	*FANCE*	*HNF1A*	*MAP2K4*	*NOTCH3*	*PTEN*	*SRC*	
APC	*CALR*	*CUX1*	*FANCF*	*HOXB13*	*MAP3K1*	*NPM1*	*PTPN11*	*SRSF2*	
AR	*CARD11*	*CXCR4*	*FANCG*	*HRAS*	*MAP3K14*	*NRAS*	*RAC1*	*STAG2*	
ARAF	*CBL*	*CYLD*	*FAS*	*HSP90AA1*	*MAPK1*	*NSD1*	*RAD21*	*STAT3*	
ARID1A	*CBLB*	*DAXX*	*FBXW7*	*ID3*	*MCL1*	*NTRK1*	*RAD50*	*STK11*	
ARID1B	*CBLC*	*DDR2*	*FGF4*	*IDH1*	*MDM2*	*NTRK2*	*RAD51*	*SUFU*	
ARID2	*CCND1*	*DICER1*	*FGF6*	*IDH2*	*MDM4*	*NTRK3*	*RAF1*	*SUZ12*	
ASXL1	*CCND3*	*DNM2*	*FGFR1*	*IGF1R*	*MED12*	*PAK3*	*RB1*	*TAL1*	
ATM	*CCNE1*	*DNMT3A*	*FGFR2*	*IKZF1*	*MEF2B*	*PALB2*	*RET*	*TCF3*	
ATR	*CD274*	*DOT1L*	*FGFR3*	*IKZF3*	*MEN1*	*PAX5*	*RHEB*	*TERT*	
ATRX	*CD79A*	*EED*	*FGFR4*	*IL7R*	*MET*	*PBRM1*	*RHOA*	*TET2*	
AURKA	*CD79B*	*EGFR*	*FH*	*INHBA*	*MITF*	*PDGFRA*	*RIT1*	*TGFBR2*	
AURKB	*CDC73*	*EGLN1*	*FLCN*	*IRF4*	*MLH1*	*PDGFRB*	*RNF43*	*TNFAIP3*	
AURKC	*CDH1*	*EP300*	*FLT3*	*JAK1*	*MPL*	*PHF6*	*ROS1*	*TNFRSF14*	
AXIN1	*CDK12*	*EPAS1*	*FLT4*	*JAK2*	*MRE11A*	*PIK3CA*	*RUNX1*	*TP53*	
AXIN2	*CDK4*	*EPHA3*	*FOXL2*	*JAK3*	*MSH2*	*PIK3R1*	*SDHB*	*TRAF3*	
B2M	*CDK6*	*EPHA5*	*FUBP1*	*KAT6A*	*MSH6*	*PIK3R2*	*SETBP1*	*TSC1*	
BAP1	*CDKN2A*	*ERBB2*	*GALNT12*	*KDM5C*	*MTOR*	*PIM1*	*SETD2*	*TSC2*	
BCL2	*CDKN2B*	*ERBB3*	*GATA1*	*KDM6A*	*MUTYH*	*PLCG1*	*SF3B1*	*TSHR*	
BCL2L1	*CDKN2C*	*ERBB4*	*GATA2*	*KDR*	*MYC*	*PMS1*	*SMAD2*	*U2AF1*	
BCL6	*CEBPA*	*ERG*	*GATA3*	*KEAP1*	*MYCL*	*PMS2*	*SMAD4*	*U2AF2*	
BCOR	*CHEK1*	*ESR1*	*GEN1*	*KIT*	*MYCN*	*POLD1*	*SMARCA4*	*VHL*	
BCORL1	*CHEK2*	*ETV6*	*GNA11*	*KMT2A*	*MYD88*	*POLE*	*SMARCB1*	*WHSC1*	
BCR	*CIC*	*EXO1*	*GNAQ*	*KMT2B*	*NF1*	*PPM1D*	*SMC1A*	*WT1*	

Targeted NGS and mutation analysis

Genomic DNA was extracted from formalin-fixed, paraffin-embedded tissues and sequencing was performed on an Illumina NextSeq 550 System platform. Most cases were sequenced with a 275-gene panel in Table 1 whereas there were slight panel variations for 30 cases. The average sequencing depth was 700×, and a sequence coverage ≥100× (after removing duplicates) was required for mutation calling. The percent reads passing filter (Reads PF) was >80%. All coding exons of these genes were sequenced, along with 50 intronic nucleotides flanking each exon end.

Alignment of sequencing data and variant calling were performed with the DRAGEN Somatic Pipeline (Illumina) against the GRCh37 reference genome to identify SNVs and INDELs. Because we did not have matched normal samples, the DRAGEN tumor-only pipeline was used, and the output data were further refined using publicly available and in-house databases of germline single nucleotide polymorphisms and INDELs to remove germline variants with additional consideration of variant allele frequencies of the variants.

Immune profiling using fluorescent multiplex immunohistochemistry (mIHC)

Previously, we have performed fluorescent mIHC for 13 immune markers to quantitate the composition of the tumor immune microenvironment and immune checkpoint in a large DLBCL cohort[20] including 323 cases in the current study cohort. The details of mIHC staining and antibodies have been described previously. Here we examined the correlations between tumor genetic characteristics through NGS and the quantitated microenvironment immune traits through fluorescent mIHC, including the abundance of $CD3^+$ T cells (further subtyped into $CD8^+$ and $CD4^+$ T cells), $CD68^+$ macrophages, and $CD56^+$ natural killer cells in the 323 cases, and expression of the immune checkpoint molecules in tumor cells and immune cells (PD-L1/PD-L2/PD-1 expression in $CD20^+$ cells, CTLA-4/PD-1/PD-L1 in T cells, and PDL1/PD-L2 expression macrophages and natural killer cells; data for these immune checkpoint molecules were not available in 2–6 cases).

Gene-expression signature analysis

Gene expression profiling (GEP) data using the Affymetrix GeneChip Human Genome HG-U133 Plus 2.0 microarrays (GSE31312)[19] 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)[21] was performed. To visualize gene signatures with the set thresholds of false discover rate and fold change, CLUSTER software and JAVA TREEVIEW (https://www.java.com/en) were used.[22] The Expression Analysis Systematic Explorer[23] software was used to categorize over-represented biological pathways using Gene Ontology (GO) terms.

Statistical analysis

Fisher's exact test and unpaired (2-tailed) Student's *t*-test were used to compare clinical and molecular features between two groups. Overall survival and progression-free survival were compared with the Kaplan–Meier method and Log-rank (Mantel-Cox) test using GraphPad Prism software. Multivariate analysis was performed with Cox proportional hazards regression models using SPSS software. *P*-values ≤ 0.05 were considered to be statistically significant.

Results

High mutation numbers and INDELs in the absence of TP53 mutation correlate with poor prognosis

NGS targeting 275 lymphoma-related genes was successful in 424 patients, including 6 patients with high-grade B-cell lymphoma (HGBCL) with *MYC*/*BCL2*-double hit (DH) as determined by fluorescence *in situ* hybridization[24,25] and 418 patients with DLBCL, not otherwise specified (NOS). Among the 424 patients, 408 patients had 1–45 nonsynonymous mutations; 171 patients had 1–9 deletion mutations (as a frameshift, inframe deletion, larger deletion, or nonsense mutations) and only 64 patients had 1–2 insertion mutations (as a frameshift, inframe insertion, or nonsense mutations). The mean number for nonsynonymous mutations and non-silently mutated (MUT) genes in the study cohort was 4.2 and 3.75, respectively. HGBCL-*MYC*/*BCL2*-DH compared with DLBCL-NOS cases had a non-significantly higher mean mutation number (5.0 vs 3.8) but similar mean number of MUT genes.

Consistent with our earlier reports, in this sequencing cohort, activated B-cell-like (ABC) DLBCL had significantly worse survival than germinal center B-cell-like (GCB) DLBCL;[19] HGBCL-*MYC*/*BCL2*-DH had significantly worse survival than DLBCL-NOS despite the GCB cell-of-origin;[25] and *TP53* nonsynonymous mutations (MUT-*TP53*) correlated with significantly worse survival in both GCB and ABC subtypes and both HGBCL-*MYC*/*BCL2*-DH and DLBCL-NOS entities.[26,27] In correlating NGS mutation numbers to clinical outcome, we found that only in patients with wild-type (WT) *TP53*, and more particularly in the GCB subtype with WT-*TP53*, significantly poorer survival was associated with high numbers of nonsynonymous mutations (with all cutoffs ranged from >4 to >23) and MUT genes (with all cutoffs ranged from >4 to >18). The survival curves for ≥6 MUT genes (MUT^{high}) compared with 0–5 MUT genes (MUT^{low}) are shown in Figure 1(a) and Supplementary Figure 1a (excluding HGBCL-*MYC*/*BCL2*-DH cases). Similarly, presence of insertion/deletion mutations in sequenced genes also showed significant adverse impact in GCB-DLBCL with WT-*TP53* (Figure 1(b-c)), Supplementary Figure 1b-c).

The clinical features of WT-*TP53* GCB-DLBCL patients with MUT^{high} or MUT^{low} are shown in Table 2. Multivariate analysis adjusting clinical features confirmed the significant prognostic effects of MUT^{high} and INDELs (Supplementary Table 1). Although ≥5 nonsynonymous mutations and ≥5 MUT genes were associated with significantly better OS in

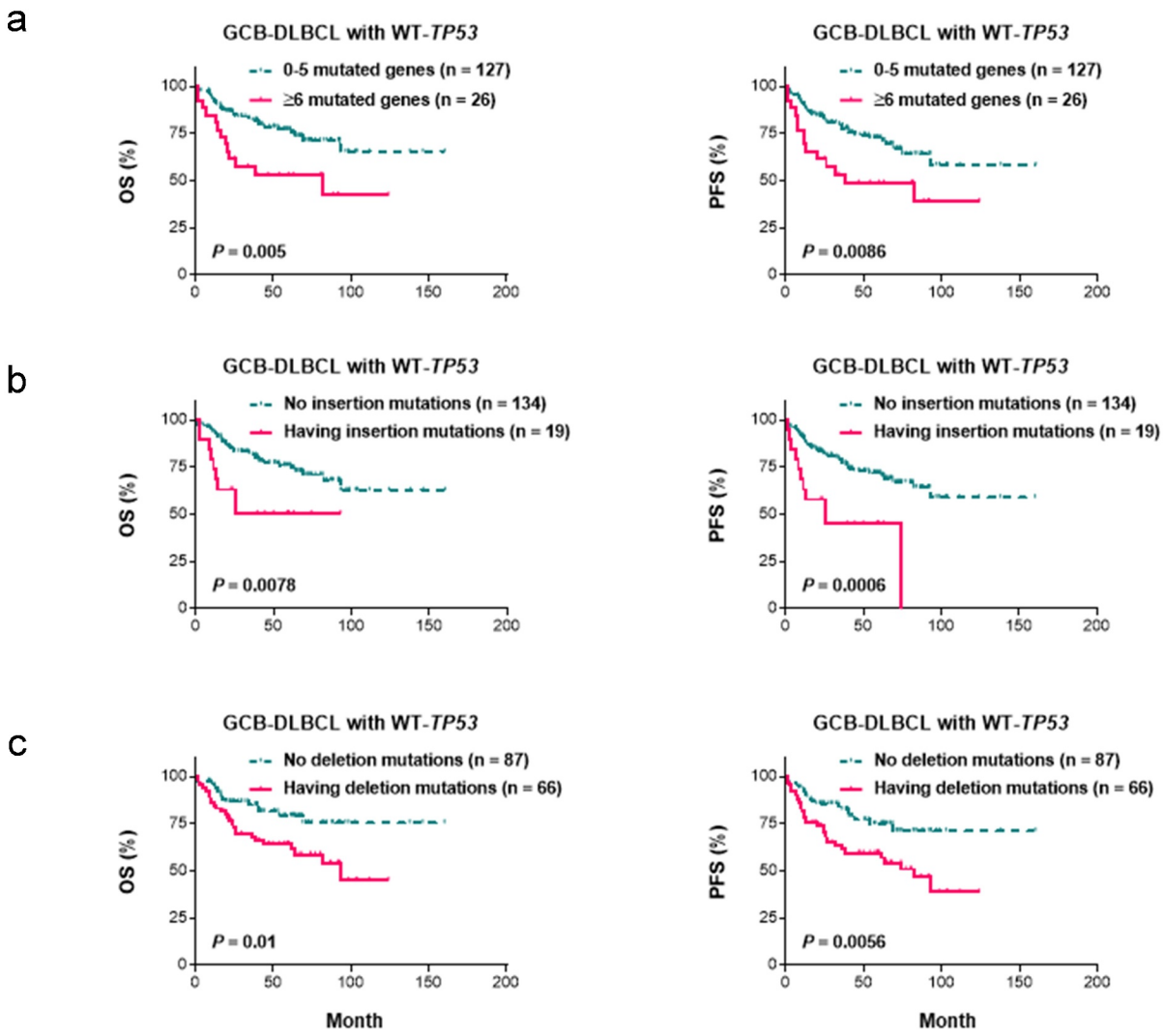

Figure 1. Prognostic analysis for mutation numbers in DLBCL. **(A)** In GCB-DLBCL with wild-type *TP53*, patients with nonsynonymous mutations in ≥6 genes had significantly poorer OS and PFS than patients with 0–5 non-silently mutated genes (totally 275 lymphoma-related genes were sequenced). **(B-C)** In GCB-DLBCL with wild-type *TP53*, patients with insertion or deletion mutations in sequenced genes had significantly poorer OS and PFS than patients without. GCB, germinal center B-cell–like; WT, wild-type; OS, overall survival; PFS, progression-free survival.

patients with mutant *TP53* (P = .039), the effect was not significant in the multivariate analysis.

Mutations in epigenetic regulators and TP53 correlate with high mutation numbers

To understand the prognostic effect of high mutation numbers, we first compared the genetic features of MUT^{high} patients with MUT^{low} patients with DLBCL-NOS. Distribution of frequent (occurred in ≥7 patients) gene mutations in MUT^{high} patients is displayed in Figure 2(a), and genes more frequently mutated in MUT^{high} versus MUT^{low} patients are shown in Figure 2(b) (in overall DLBCL-NOS) and Supplementary Table 2 (in GCB/ABC subtypes). Notably, by function many genes over-represented in MUT^{high} patients are involved in epigenetic regulation (such as *KMT2D, EZH2, CREBBP, TET2, SMARCA4, DNMT3A, EP300, KDM6A*, and *SMC3*). The most enriched gene was *KMT2D* (also known as *MLL2* or *MLL4*, encoding a histone methyltransferase for H3K4me; Figure 2(b-c)) in GCB (64.9% in MUT^{high} patients versus 26.7% in MUT^{low} patients) and *TP53* in ABC DLBCLs (46.4% in MUT^{high} patients versus 16.1% in MUT^{low} patients (Supplementary Table 2). The most common type of *KMT2D* mutations was nonsense mutations (48.7%), followed by missense (32.4%) and frameshift (20.9%) and inframe INDEL (2.0%) mutations (Figure 2(b)), in contrast with the predominant missense type of *TP53* mutations. *KMT2D* and *TP53* were also recurrently mutated in HGBCL-*MYC*/*BCL2*-DH patients despite the small number of cases in our cohort (Figure 2(b)).

Conversely, patients with *KMT2D* nonsynonymous mutations or MUT-*TP53* had significantly higher mean numbers of non-silent mutations/mutated genes in overall DLBCL-NOS and both the GCB/ABC subtypes (Figure 2(d), Supplementary Figure 2a). The increase remained to be highly significant after exclusion of patients with 0 MUT gene from the MUT^{low} group (P < .0001 for *KMT2D* mutations in all comparisons). In contrast, although GCB compared with ABC subtype was associated with increased mutation numbers in overall cohort, the association lost significance in both the WT-*KMT2D* and MUT-*KMT2D* subgroups (Supplementary Figure

Table 2. Clinical features of patients with DLBCL not otherwise specified with and without high mutation levels.

	GCB WT-*TP53*			ABC WT-*TP53*			
	0–5 MUT genes	≥6 MUT genes	*P*	0–5 MUT genes	≥6 MUT genes	*P*	*P**
Sex							
Male	65	16	0.38	84	14	0.48	1.0
Female	59	9		67	7		
Age, years							
≤60	61	13	0.83	51	9	0.47	0.57
>60	63	12		100	12		
Stage of disease							
I–II	68	12	0.65	57	11	0.23	1.0
III–IV	49	11		89	9		
Serum LDH level							
Normal	54	4	**0.011**	49	8	0.61	0.098
Elevated	62	19		92	11		
ECOG performance status							
0–1	92	18	1.0	111	18	0.54	1.0
≥2	15	3		27	2		
No. of extranodal sites involved							
0–1	95	18	1.0	105	17	0.60	1.0
≥2	20	4		39	4		
IPI risk group							
0–2	87	15	0.33	78	15	0.16	0.32
3–5	32	9		68	6		
B-symptoms							
Absence	90	13	0.08	91	15	0.48	0.36
Presence	30	10		54	6		
Tumor size							
<5 cm	58	13	0.59	76	6	0.25	0.15
≥5 cm	34	5		48	8		

GCB, germinal center B-cell–like; ABC, activated B-cell–like; WT, wild type; MUT, non-silently mutated; LDH, lactate dehydrogenase; ECOG, eastern cooperative oncology group; IPI: International Prognostic Index.
*P**: for GCB vs ABC subtype of WT-*TP53* patients with ≥6 non-silently mutated genes.
Significant *P* values are in bold.

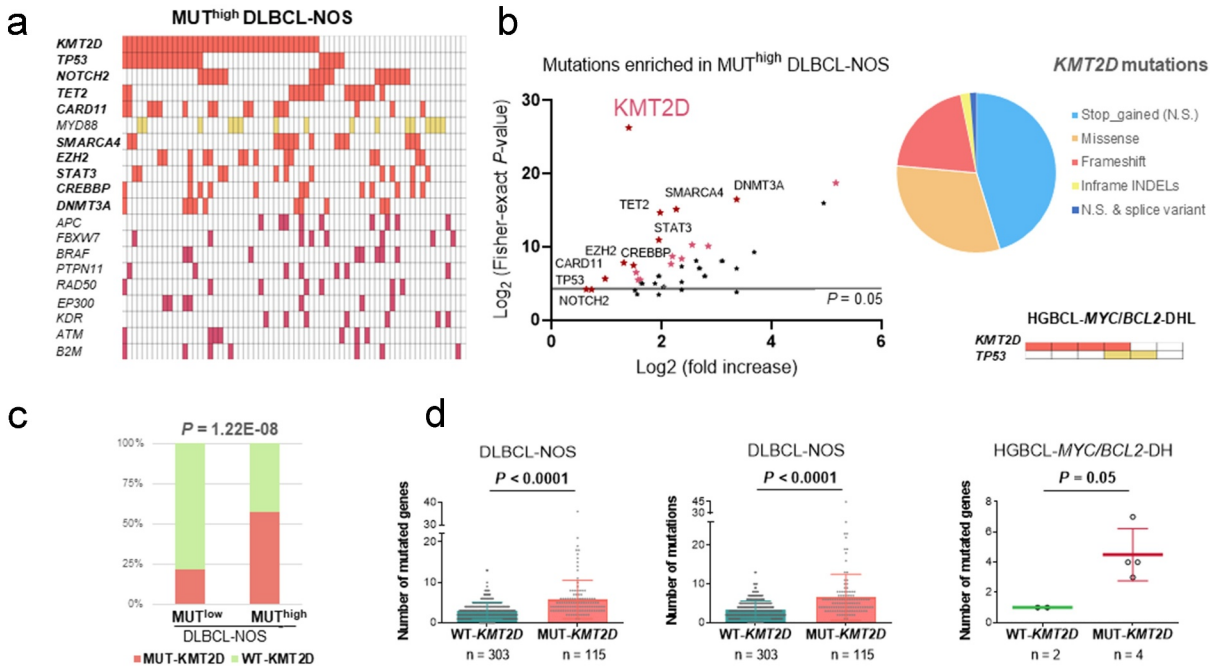

Figure 2. Mutations enriched in DLBCL with high number of mutations. **(A)** Genes non-silently mutated in ≥7 DLBCL-NOS patients who had ≥6 mutated genes (MUT^{high}, 68 patients) in our NGS analysis. Except for *MYD88*, the mutational frequencies of these genes were significantly higher in MUT^{high} vs MUT^{low} patients by Fisher's exact test. Each box represents one patient. **(B)** Left, mutations overrepresented in MUT^{high} vs MUT^{low} patients shown by the data dots. Red dots represents genes non-silently mutated in ≥7 MUT^{high} patients. Gene names are labeled for those mutated in ≥10 MUT^{high} patients i.e., bolded gene names in figure panel A. Right top, mutation types of *KMT2D* nonsynonymous mutations in DLBCL. Right bottom, recurrent gene mutations in HGBCL-*MYC*/*BCL2*-DH cases. Each box represents one patient. **(C)** Frequencies of *KMT2D* nonsynonymous mutations in MUT^{high} vs MUT^{low} patients. **(D)** *KMT2D* nonsynonymous mutations were significantly associated with higher numbers of non-silently mutated genes and nonsynonymous mutations in DLBCL-NOS, and associated with significantly higher numbers of non-silently mutated genes in HGBCL-*MYC*/*BCL2*-DH cases. DLBCL-NOS, diffuse large B-cell lymphoma, not otherwise specified; HGBCL-*MYC*/*BCL2*-DH, high-grade B-cell lymphoma with *MYC*/*BCL2* genetic double-hit; WT, wild-type; MUT, non-silently mutated; GCB, germinal center B-cell–like; ABC, activated B-cell–like.

2a). In the six HGBCL-*MYC*/*BCL2*-DH cases, only MUT-*KMT2D*(but not MUT-*TP53*) was significantly associated with increased numbers of MUT genes (but not total nonsynonymous mutation numbers Figure 2(d)).

When we performed the comparison in the absence and presence of MUT-*TP53*, respectively (Supplementary Tables 3–4), we found only mutations in a few epigenetic regulators (including *KMT2D, TET2, SMARCA4, DNMT3A*, and *SMC1A), MSH6* (a DNA mismatch repair gene), and *PTPN11* (a member of the protein tyrosine phosphatase family) were significantly associated with MUT^{high} independent of *TP53* mutation status. Among MUT^{high} patients overall and in the WT-*TP53* subset, ABC compared with GCB DLBCL patients had a significantly higher frequency of *MYD88* mutation whereas lower frequency of *EZH2* mutation (Supplementary Tables 2–3).

High mutation numbers, INDELs and KMT2D mutations are associated with lower T cell densities in patients with WT-TP53

Next, we examined the tumor immune microenvironment using fluorescent mIHC[20] and analyzed the relationship between genetics and immune characteristics in 323 cases (Supplementary Figures 2b and 3a). Case distribution of the absolute cell counts for 13 immune markers and comparisons between MUT^{high} and MUT^{low} patients are shown in Figure 3 (a-b). Figure 3(c) displays the single-cell intensities of CD20$^+$, CD3$^+$, CD68$^+$, CD56$^+$, PD-1$^+$, and PD-1$^+$ cells (each dot represents a cell) in a representative case with MUT^{high}, WT-*TP53* and MUT-*KMT2D*.

Regarding immune cell infiltration, we found that MUT^{high} was significantly associated with lower absolute T cell counts and cell densities in overall and the GCB subtype of DLBCL-NOS patients with WT-*TP53* (Figure 3(a-b, d)), whereas no significant difference in intratumoral macrophages and natural killer cells were observed. Similar results were found for presence of INDELs (P = .0096) and *KMT2D* nonsynonymous mutations (Supplementary Figure 3b-c). *TP53* mutations were associated with significantly decreased percentage of CD8$^+$ T cells in tumor/immune cells only when HGBCL cases were not excluded. Only in the GCB subtype of DLBCL-NOS patients with MUT-*TP53*, MUT^{high} was significantly associated with increased T cell densities (Figure 3(d)) and percentage in tumor/immune cells.

Regarding expression of immune checkpoint molecules in tumor and immune cells, we found that MUT^{high} in DLBCL-NOS and overall cases was significantly associated with lower PD-L1 expression in CD68$^+$ macrophages/CD20$^+$ B cells and lower PD-1 expression in CD4$^+$/CD8$^+$ T cells, evaluated by PD-L1$^+$/PD-1$^+$ percentage in CD68$^+$ cells, CD20$^+$ cells, CD3$^+$ cells, CD3$^+$CD4$^+$ cells, or CD3$^+$CD8$^+$ cells (Figure 3(d), Supplementary Figure 4a). More precisely, the association with decreased PD-L1 expression was significant in patients with a molecular background of ABC and WT-*TP53*, whereas the association with lower PD-1 expression was mainly in patients with MUT-*TP53* (Figure 3(e), Supplementary Figure 4a-b). Presence of INDELs was associated with a lower mean PD-L1$^+$ percentage in CD68$^+$ macrophages only in overall DLBCL-NOS (but not in DLBCL subsets) and a low mean PD-L1$^+$ percentage expression in B cells only in DLBCL patients with WT-*TP53* (Supplementary Figure 4c). In contrast, *KMT2D*/*TP53* mutations were not significantly associated with differential PD-1/PD-L1 expression evaluated by percentage in a specific cell type.

MUT^{high} shows prominent GEP signatures including p53-related genes

To gain further biological insight, we compared the gene expression profiles of MUT^{high} and MUT^{low} patients. Prominent GEP signatures were identified for MUT^{high} in overall cohort, the *WT-TP53* subset, and the *WT-TP53* GCB subset (Figure 4(a)). Notable signatures among the large number of upregulated genes in MUT^{high} WT-*TP53* GCB included *IGHM*, voltage-gated ion channel components/regulators (*CLCN1, CLCN2, KCNH4, KCNA4, CABP2*), p53 inhibitor *AGR2*, and paradoxically several tumor suppressors and positive regulators of the p53 pathway (*DHRS2* that attenuates MDM2-mediated p53 degradation, pro-apoptotic *BBC3, SIK1* with role in p53-dependent anoikis and metastasis suppression, *CADM4*, and *INSM2*). Downregulated genes included those functioning in tumor suppression (*CCDC6, RBL2, NEMF, BCLAF1, RASA1*), mRNA metabolism and/or translation regulation (*STAU1, SP3, DDX6, PABPC3*), cell cycle (*NSA2, ANAPC16, PCNP*), epigenetic regulation (*SMARCA5*), degradation (*UBE2D2, FEM1C*), and others (Table 3.). Among DNA repair genes, *HDAC1, ITM2A, PARP1, BCL11B, GATAD2B*, and *RAB27A* were downregulated whereas *PARP3, XRCC3, RAD54L*, and *ERCC2* were upregulated in MUT^{high} cases. The GO class and gene categories for differentially expressed genes are listed in Table 3. As the MUT^{high} gene signatures showed involvement of the p53 pathway, we compared the p53/MDM2 expression[26,28] in MUT^{high} and MUT^{low} patients with WT-*TP53*. Only in ABC-DLBCL with WT-*TP53*, MUT^{high} patients was significantly associated with higher mean levels of WT-p53 and MDM2 overexpression (Figure 4(b)).

In contrast, when we analyzed the GEP data for *KMT2D* mutations, only a few genes showed significant upregulation in MUT-*KMT2D* compared with WT-*KMT2D* patients (Supplementary Table 5), suggesting functional heterogeneity among MUT- or WT-*KMT2D* cases.

Validation in a WES cohort

We used the publicly available WES data and SNV/INDEL numbers in 304 DLBCL patients deposited by the Harvard study group[16] to validate our findings, including the full mutation annotation analyzed by Chapuy et al. for a subset of 134 non-microsatellite-instability (MSI) cases using matched tumor-normal samples. Totally 158 genetic drivers, including 85 driver gene mutations (including 29 genes in our 275-gene panel), 65 copy number alterations (CNAs) and 8 structural variants (SVs), have been identified by Chapuy et al in this WES cohort.

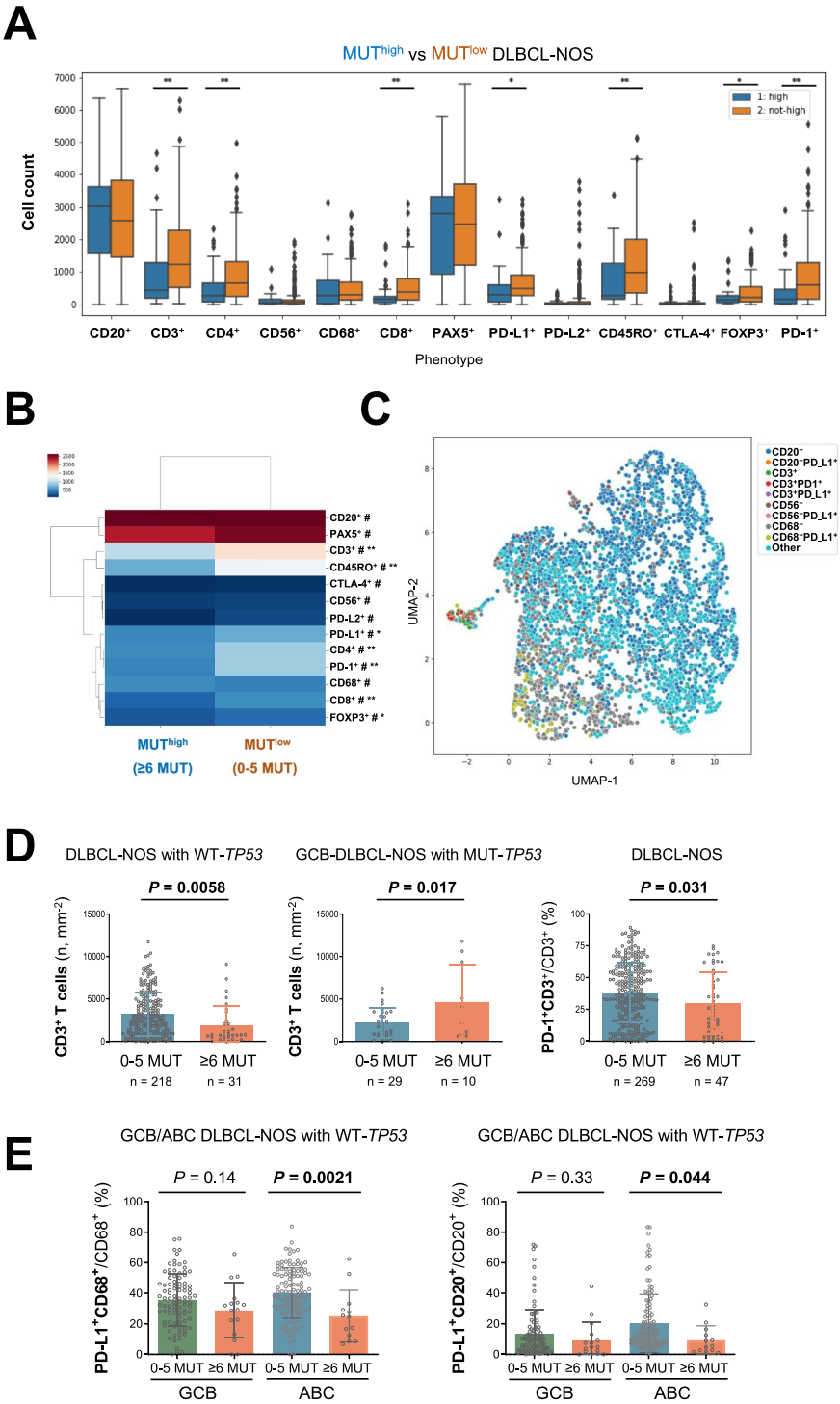

Figure 3. Immunological analysis for DLBCL-NOS patients with high mutation numbers. **(A)** Box plot showing the distribution of absolute cell counts for 13 immune markers in DLBCL-NOS patients with high (≥6) or low (<6) numbers of non-silently mutated genes by our NGS anaysis. Each dot represents one patient. Significant differences between two groups are marked by asterisks. *, $P \leq 0.05$; **, $P \leq 0.01$ by a two-sided Mann–Whitney U-test. **(B)** Clustermap visualization based on the mean absolute cell counts in two groups. **(C)** A representative Uniform Manifold Approximation and Projection (UMAP) plot generated from single-cell intensities for CD20, CD3, CD68, CD56, PD-1, and PD-L1 markers in a patient with mutated *KMT2D* and another five non-silently mutated genes and wild-type *TP53* through our 275-gene targeted sequencing. Each data point represents a cell, color labeled according to phenotype. In the legends, PD-L1 or PD-1-negative phenotype in the single-positive labels were omitted to avoid typing confusion. **(D)** High numbers (≥6) of non-silently mutated genes were significantly associated with lower cell densities of tumor-infiltrating T cells in DLBCL-NOS patients with wild-type *TP53*, higher T cell densities in GCB-DLBCL-NOS patients with mutant *TP53*, and lower PD-1 percentage expression in T cells in overall DLBCL-NOS patients. **(E)** High numbers of non-silently mutated genes were associated with significantly lower PD-L1 expression in $CD68^+$ macrophages and $CD20^+$ B cells in the ABC subtype of DLBCL-NOS with wild-type *TP53*. DLBCL-NOS, diffuse large B-cell lymphoma, not otherwise specified; WT, wild-type; MUT, non-silently mutated; GCB, germinal center B-cell–like, ABC, activated B-cell–like.

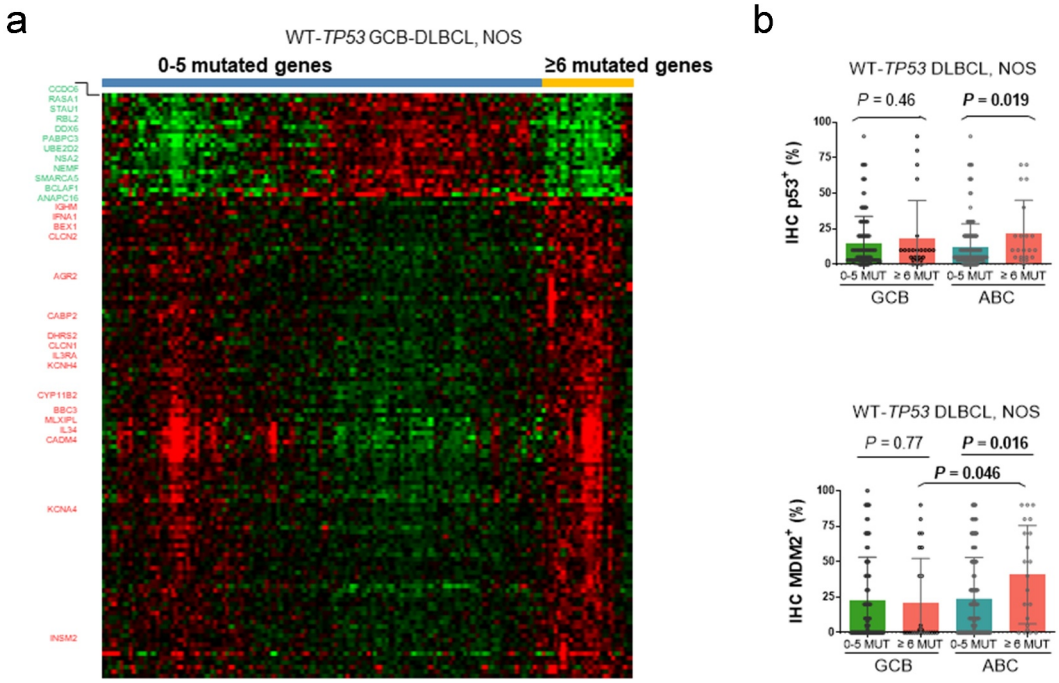

Figure 4. Gene expression analysis for high numbers of mutated genes. **(A)** Differentially expressed genes between patients with 0–5 and ≥6 non-silently mutated genes in GCB-DLBCL with WT-*TP53* (false discovery rate 0.0001, fold change ≥1.5, or false discovery rate 0.01, fold change ≥2). **(B)** Only in the ABC-DLBCL subset with WT-*TP53*, high numbers (≥6) of non-silently mutated genes were associated with increased p53 and MDM2 protein expression. WT, wild-type; DLBCL-NOS, diffuse large B-cell lymphoma, not otherwise specified; GCB, germinal center B-cell–like; ABC, activated B-cell–like; IHC, immunohistochemistry.

Supporting the results in our NGS cohort, MUT-*KMT2D* was associated with genomic complexity in the validation cohort, evidenced by significantly increased numbers of genomic SNVs (including synonymous variants, for the non-MSI cases only), insertion mutations, CNAs, and SVs (Figure 5(a)). High TMB (>75 SNVs) and INDEL numbers (≥5 insertion/deletion mutations) by WES analysis were associated with significantly poorer survival in ABC-DLBCL with WT-*TP53* (Supplementary Figure 5). Different from our cohort subtyped mainly by GEP, in the validation cohort GCB and ABC subtypes (mainly determined by NanoString) had similar prognosis. Analysis for 158 genetic drivers identified by Chapuy et al. also showed correlations between MUT-*KMT2D* and MUT^{high} and adverse prognostic effects of genomic complexity: MUT-*KMT2D* was significantly associated with increased numbers of MUT driver genes, driver CNAs, and driver SVs (Supplementary Figure 6a). Patients with ≥6 MUT driver genes had significantly poorer survival than those without in ABC-DLBCL and the C5 genetic subset treated with R-CHOP (Supplementary Figure 6b). In addition, ≥5 driver CNAs and ≥3 SVs were associated with significantly poorer survival in ABC-DLBCL (Supplementary Figure 6c).

In the 134 patients with tumor/normal paired samples available for genetic alteration analysis, *KMT2D* nonsynonymous mutations remained to be significantly associated with increased numbers of genomic SNVs and insertion mutations (Figure 5(a)). The significance of this association was enhanced ($P < .0001$, Supplementary Figure 7a) when MUT-*KMT2D* cases were combined with cases with nonsynonymous mutations in *EZH2, KMT2A, ARID1A, ARID1B, SMARCA4, KDM6A*, or *CHD2* in the validation set, which were gene mutations functioning in epigenetic regulation over-represented in our MUT^{high} versus MUT^{low} cases. These mutations were associated with both aging and non-canonical AID signatures (Supplementary Figure 7b); in contrast, *KMT2D* nonsynonymous mutations were only associated with the aging mutational signature. Furthermore, high numbers of SNVs, MUT genes, and INDELs and *KMT2D* mutations were all associated with significantly poorer survival in the WT-*TP53* ABC-DLBCL subset of these 134 patients (Figure 5 (b-e), Supplementary Figure 7c).

Discussion

Previous studies have shown that approximately one-fifth of DLBCL has a high TMB, however, less than 10% relapsed/refractory DLBCL patients responded to immune checkpoint inhibitors. In this study, we investigated the clinical significance of tumor mutation numbers in DLBCL treated with standard R-CHOP immunochemotherapy, and correlated with our immune profiling data[20] to understand the interaction between tumor genomics and the host immune responses. We found that high mutation numbers of DLBCL, either measured by numbers of lymphoma-driver genes or by genomic SNVs were associated with poorer survival in patients with WT-*TP53*, significantly in the GCB molecular subset in our cohort and the ABC subset of the Harvard WES cohort, not supporting the hypothesis that DLBCL patients with low TMB are enriched in relapsed/refractory patients to explain the low efficacy of PD-1 inhibitors in DLBCL clinical trials. High MUT gene numbers and INDELs were associated with decreased T cells and seemly lower T cell responses (Figure 3 (b)) in patients with WT-*TP53*, which may suggest that aggressive lymphoma tumors often harbor immune-escaping

Table 3. Gene expression profiling analysis for high numbers (≥6 of sequenced genes) of non-silently mutated genes in GCB-DLBCL-NOS with wild-type *TP53*.

	Downregulated	Upregulated
List of genes:	(FDR 0.01, fold change ≥2) *DDX6, NSA2, SNX3, PABPC3, FEM1C, GDI2, MGEA5, SMARCA5, YWHAQ, RASA1, UBE2D2, CCDC6, NEMF, STAU1, RBL2, PCNP, SP3, ANAPC16, ESYT2, BCLAF1, STAM2, GNAS*	(FDR 0.0001, fold change ≥1.5) *PARD6A, CGB, MUC3B, DOLPP1, PIP5KL1, ARMC5, SHANK1, DEFA5, FOLR3, FAM196A, PDE1B, TBL1Y, CRYGB, CABP2, IL3RA, CPNE9, MYOD1, SCGN, HPGD, NGFR, B3GAT1, DOHH, TPH2, FNDC8, PTPRU, ODF3L2, LRRC36, FAM153A, DHRS2, CRYBB1, STC2, CNTD1, REG3A, FMN2, INGX, TMCO5B, KCNH4, KLF16, NR2E1, IFNA1, XAGE2, CLCN1, MGC13053, FEV, LLGL1, INSM2, GSTA3, AGR2, ANXA8, LGALS8-AS1, LINC00520, LINC00652, DGCR5, ANKS4B, CLCN2, ZNF503-AS1, KCNA4, C1orf158, FAM151A, EPX, C11orf42, ADORA1, LINC00658, ILVBL, ST8SIA2, TDRG1, LINC00545, ELANE, CPLX1, CADM4, FLJ38576, FABP1, KIAA1656, SNTN, MYCL, TSLP, MAGEB2, BBC3, SPRED3, MLXIPL, SLC7A11-AS1, RBP2, TM4SF20, B4GALT2, SRCRB4D, CYP11B2, C1orf170, SNAPC2, PLA2G4F, SIK1, BEX1, IGHM, IL34*
Gene category of GO terms:	Protein transport; Cell growth and/or maintenance; Intracellular transport; Vesicle-mediated transport; Intracellular protein transport; Intracellular signaling cascade; Transport; Intracellular	Membrane; Voltage-gated potassium channel complex; Integral to membrane; Glycosaminoglycan binding; Plasma membrane; Hyaluronic acid binding; Serine-type endopeptidase activity; Integral to plasma membrane; Receptor activity; Voltage-gated ion channel activity; Serine-type peptidase activity; Transmembrane receptor activity; Anion transport; Calmodulin-dependent cyclic nucleotide phosphodiesterase activity; Inorganic anion transport; Lipid metabolism; Monooxygenase activity; Oxidoreductase activity/ activing on paired donors/with incorporation or reduction of molecular oxygen; Chymotrypsin activity

GCB, germinal center B-cell–like; DLBCL, NOS, diffuse large B-cell lymphoma, not otherwise specified; FDR, false discovery rate; GO, Gene Oncology.

mutations instead of immunogenic mutations. Indeed, previous studies have shown that multiple genetic lesions implicated in immune escape are frequent in DLBCL, such as genetic deletion/mutations that inactivate B2M (component of the major histocompatibility complex class I), CD58 (important for adhesion and activation of T cells and natural killer cells), and CREBBP/EP300 (histone acetyltransferases that also acetylate BCL6 and p53) in 29%, 21%, and 39% of DLBCL, respectively,[29,30] and that *HLA-A* mutation burdens and loss-of-heterozygosity were increased in the diagnostic samples of patients who later experienced relapse after R-CHOP treatment than patients who had durable therapeutic responses.[31] In our cohort, *B2M* mutations were significantly enriched in MUT^{high} cases with GCB and WT-*TP53* molecular background (the subset in which MUT^{high} showed significant adverse prognostic effect).

In contrast, in the GCB subtype of DLBCL-NOS patients with MUT-*TP53*, MUT^{high} was associated with increased tumor-infiltrating T cells. Moreover, MUT^{high} was associated with lower PD-1 expression in T cells in overall cohort (high PD-1 expression had adverse prognostic impact[20] in this study cohort) and lower PD-L1 expression in macrophages and B cells in ABC-DLBCL with WT-*TP53* (high PD-L1 expression in macrophages was associated with poorer survival[20] in overall cohort and the ABC-DLBCL subset with WT-*TP53*, $P = .0069$), which appeared to suggest a favorable role of MUT^{high} in T cell responses. Notably, only in the ABC subtype of DLBCL, genomic MUT^{high} and *KMT2D* mutations showed significant association with high degree of somatic hypermutations (SHM) in the immunoglobulin heavy chain variable region (IGHV, Supplementary Figure 2b).[32] In our previous study, high degree of IGHV SHM was associated with significantly better survival and lower PD-1 expression in CD4$^+$/CD8$^+$ T cells in ABC-DLBCL.[32] It will be interesting to address the different roles of heterogeneous *TP53* mutations, IGHV SHM, and other non-IG mutations as neoantigens versus oncogenic drivers in future DLBCL studies. These oncoimmune data have important implications for future therapeutic strategies and biomarker studies for PD-1/PD-L1 inhibitors since the immune checkpoint blockade clinical trial data are limited in DLBCL.

Moreover, in this study, we showed that nonsynonymous mutations in the histone methyltransferase *KMT2D* gene were significantly associated with increased numbers of mutated genes in our cohort and higher numbers of genetic drivers (mutations, CNAs and SVs) and genomic SNVs/MUT genes by WES in DLBCL excluding MSI cases in the Harvard cohort.[16] Distinct GEP signatures were identified for MUT^{high} (rather than for *KMT2D* mutations) including genes involved in the p53 and apoptotic pathways, reminiscent of the transcriptional signatures in DLBCLs with complex CNAs of p53 and cell cycle genes,[33] which may suggest convergence of oncogenic pathways in MUT^{high} DLBCLs despite the diverse mutational profiles, as well as the importance of the p53 pathway for clinical outcome.[34] Epigenetic regulation has important role in genomic stability by regulating the chromatin accessibility and DNA repair machinery[35] and by mitigating transcription-replication conflicts in the presence of H3K4 methylation.[36] Therefore, the correlations shown in two study cohorts may help understand the origin of DLBL genomic complexity and instability in patients. Because *KMT2D* is a large gene encoding a protein of 5537 amino acids (aa), one may question whether the correlation between high mutation burdens and *KMT2D* mutations was merely due to its large size or its high mutational frequency. However, KMT2C (MLL3) is also large (4911 aa), but *KMT2C* nonsynonymous mutations were not detected in our NGS cohort and were not associated with significantly increased SNV numbers in the Harvard WES cohort, whereas mutations in KMT2A (MLL1; 1162 aa) and EZH2 (for H3K9me and K3K27me; 746 aa) of smaller sizes were significantly associated with higher SNV/MUT numbers in both our

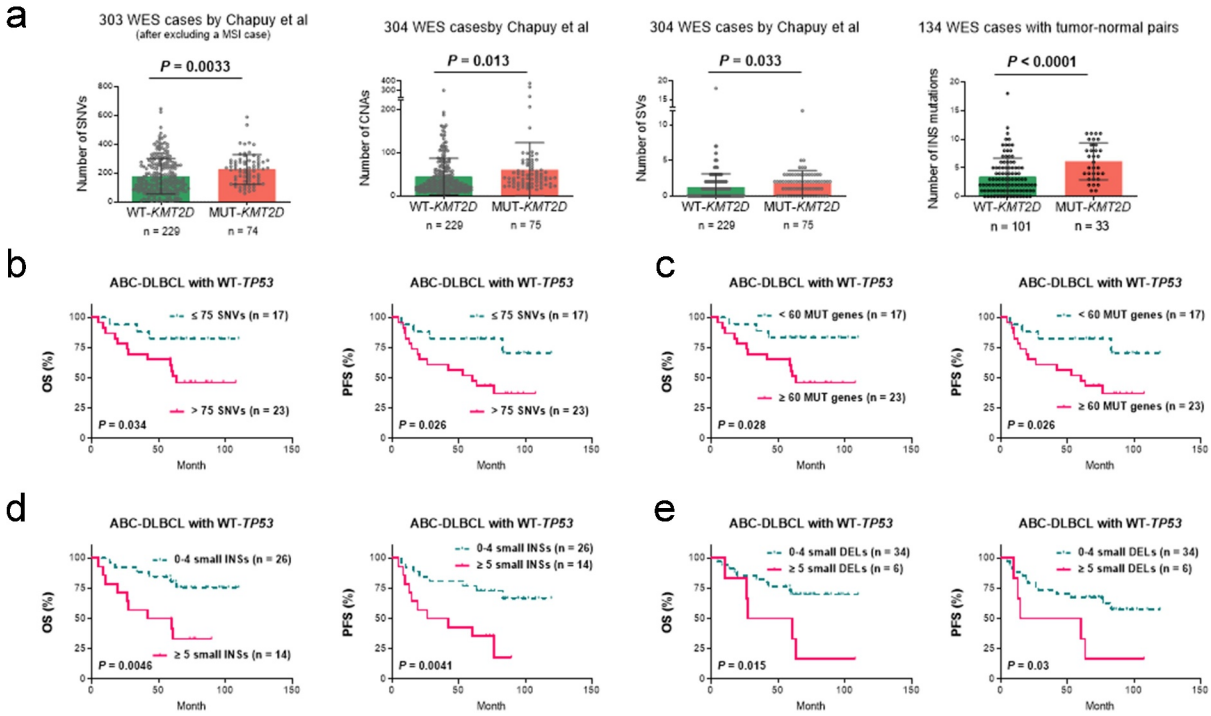

Figure 5. Correlative and prognostic analysis for genomic complexity in a validation cohort from the Harvard study group. **(A)** *KMT2D* nonsynonymous mutations were significantly associated with increased numbers of SNVs and insertion mutations by WES and copy number alterations (CNAs) and structural variants (SVs) by targeted sequencing of DLBCL tumor samples with or without paired normal samples. **(B-E)** Prognostic effects of genomic mutation numbers by WES analysis of paired tumor-normal samples in ABC-DLBCL patients with wild-type *TP53* treated with R-CHOP. WES, whole-exome sequencing; SNVs, single nucleotide variants; MSI, microsatellite instability; WT, wild-type; MUT, non-silently mutated; INS, insertion mutation; ABC, activated B-cell–like; OS, overall survival; PFS, progression-free survival; DEL, deletion mutation.

NGS cohort and the Harvard WES cohort. Mutations in many other epigenetic regulators showed significant or a consistent trend of enrichment in MUT^{high} cases of our and the Harvard cohorts (such as KMT2B, MLL2/MLL4, 2715 aa; KDM6A, 1401 aa; ARID1A, 2285 aa; TET2, 2002 aa; SMARCA4, 1647 aa; CHD2, 1828 aa and DNMT3A, 912 aa), although these genes had much lower mutational frequencies in DLBCL than *KMT2D*. In addition, gene sizes and domain function have been taken into account while identifying candidate cancer driver genes by Chapuy et al. using MutSig2CV.[9,16]

In this study, MUT-*KMT2D* was associated with significantly decreased tumor-infiltrating T cells despite increased TMB, independent of GCB/ABC subtypes in patients with WT-*TP53*. Previous functional studies have shown that KMT2D is a tumor suppressor repressing germinal center B-cell lymphoma development, and KMT2D inactivation affects growth and survival pathways including BCR, CD40, and JAK-STAT signaling in lymphoma cells;[37,38] our findings add to a possible role of KMT2D in the T cell response. Intriguingly, a recent *in vivo* screening with CRISPR identified KMT2D loss-of-function mutations as a major biomarker for PD-1 blockade therapy across multiple solid tumor types, and *Kmt2d* loss led to increased DNA damage, elevated mutation burden, activation of transposable elements, and increased immune cell infiltration, underlying the sensitivity to anti-PD-1 treatment in that study.[39] This *in vivo* study supports the association of KMT2D mutations with high mutation burdens but not the T cell infiltration results in our lymphoma patients. The discrepancy on immune cell infiltration could be related to differences in cancer types, the p53 status, and/or Myc overexpression (for example, in that *in vivo* study T cell infiltrate was not increased in a p53-competent LLC model, and the mouse model used to demonstrate the anti-PD-1 efficacy had Myc overexpression and *Trp53* knockout). Restricting our analysis to DLBCL patients with MYC overexpression and *TP53* mutation, *KMT2D* mutation continued to correlate with increased numbers of mutated genes (P = .0026) but no longer with lower T cell densities (P = .68). Moreover, T cell density had no significant prognostic effects in DLBCL-NOS patients treated with R-CHOP if the mean T cell densities in WT/MUT-*KMT2D* or MUT$^{high/low}$ patients are used as cutoffs (data not shown). Different from MUT^{high}, *KMT2D* mutations were not significantly associated with decreased PD-1/PD-L1 expression. Together, the role of *KMT2D* mutations in T cell responses and immunotherapy may need further elucidation in DLBCL,[40] and the biomarker values of T cells, PD-L1/PD-1 expression, and TMB in DLBCL need to be studied in patients treated with PD-1 inhibitors. Also noteworthy, in non-small-cell lung cancer, KMT family member mutations were associated with higher TMB and PD-L1 expression,[41] whereas *KMT2D* mutation was an unfavorable prognostic factor.[42] Mutations in several epigenetic genes have been reported to be associated with high TMB and/or efficacy of immune checkpoint blockade immunotherapy in solid tumors, including *ARID1A*,[43–45] *TET1*,[46] *KMT2A*/*C*,[47] and *EP300*.[48] In contrast, another recent study found that

DLBCL patients with *CREBBP/EP300* mutations had significantly poorer survival and decreased peripheral blood lymphocyte to monocyte ratios, and that in DLBCL xenograft murine models, mutations in CREBBP and EP300 activated the NOTCH signaling pathway and promoted macrophage polarization to M2 phenotype.[49]

In summary, *KMT2D* nonsynonymous mutations are associated with DLBCL genomic instability, and genomic complexity is associated with poor prognosis and decreased T cells and PD-L1 expression in macrophages and B cells in DLBCL with wild-type *TP53*. Further studies elucidating the oncogenic and neoantigen roles of DLBCL mutations in DLBCL patients are needed, as well as the therapeutic implications of genetic and immune biomarkers.[50]

Notes on contributors

Conception and design: ZYXM, KHY.

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

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

Manuscript writing: HY, ZYXM, MA, KHY.

Final approval of manuscript: All authors.

Acknowledgments

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

Disclosure statement

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

ORCID

Zijun Y. Xu-Monette http://orcid.org/0000-0002-7615-3949
Xiaosheng Fang http://orcid.org/0000-0002-4515-9731
Benjamin M. Parsons http://orcid.org/0000-0002-5634-2937
Ken H. Young http://orcid.org/0000-0002-5755-8932

References

1. Valero C, Lee M, Hoen D, Weiss K, Kelly DW, Adusumilli PS, Paik PK, Plitas G, Ladanyi M, Postow MA, *et al.* Pretreatment neutrophil-to-lymphocyte ratio and mutational burden as biomarkers of tumor response to immune checkpoint inhibitors. Nat Commun. 2021;12(1):729. 10.1038/s41467-021-20935-9.
2. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, Sher X, Liu XQ, Lu H, Nebozhyn M, *et al.* Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. Science. 2018;362(6411):362. 10.1126/science.aar3593.
3. Wang Z, Duan J, Cai S, Han M, Dong H, Zhao J, Zhu B, Wang S, Zhuo M, Sun J, *et al.* Assessment of blood tumor mutational burden as a potential biomarker for immunotherapy in patients with non-small cell lung cancer with use of a next-generation sequencing cancer gene panel. JAMA Oncol. 2019;5(5):696–702. 10.1001/jamaoncol.2018.7098.
4. Turajlic S, Litchfield K, Xu H, Rosenthal R, McGranahan N, Reading JL, Wong YNS, Rowan A, Kanu N, Al Bakir M, *et al.* Insertion-and-deletion-derived tumour-specific neoantigens and the immunogenic phenotype: a pan-cancer analysis. Lancet Oncol. 2017;18(8):1009–1021. 10.1016/S1470-2045(17)30516-8.
5. Wu HX, Wang ZX, Zhao Q, Chen DL, He MM, Yang LP, Wang YN, Jin Y, Ren C, Luo HY, *et al.* Tumor mutational and indel burden: a systematic pan-cancer evaluation as prognostic biomarkers. Ann Transl Med. 2019;7(22):640. 10.21037/atm.2019.10.116.
6. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, Schrock A, Campbell B, Shlien A, Chmielecki J, *et al.* Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med. 2017;9(1):34. 10.1186/s13073-017-0424-2.
7. Bevins N, Sun S, Gaieb Z, Thorson JA, Murray SS. Comparison of commonly used solid tumor targeted gene sequencing panels for estimating tumor mutation burden shows analytical and prognostic concordance within the cancer genome atlas cohort. J Immunother Cancer. 2020;8(1):8. 10.1136/jitc-2020-000613.
8. Fancello L, Gandini S, Pelicci PG, Mazzarella L. Tumor mutational burden quantification from targeted gene panels: major advancements and challenges. J Immunother Cancer. 2019;7(1):183. 10.1186/s40425-019-0647-4.
9. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL, Stewart C, Mermel CH, Roberts SA, *et al.* Mutational heterogeneity in cancer and the search for new cancer genes. Nature. 2013;499(7457):214–218. 10.1038/nature12213.
10. Wienand K, Chapuy B, Stewart C, Dunford AJ, Wu D, Kim J, Kamburov A, Wood TR, Cader FZ, Ducar MD, *et al.* Genomic analyses of flow-sorted Hodgkin Reed-Sternberg cells reveal complementary mechanisms of immune evasion. Blood Adv. 2019;3(23):4065–4080. 10.1182/bloodadvances.2019001012.
11. Xu-Monette ZY, Zhou J, Young KH. PD-1 expression and clinical PD-1 blockade in B-cell lymphomas. Blood. 2018;131(1):68–83. 10.1182/blood-2017-07-740993.
12. Armand P, Chen YB, Redd RA, Joyce RM, Bsat J, Jeter E, Merryman RW, Coleman KC, Dahi PB, Nieto Y, *et al.* PD-1 blockade with pembrolizumab for classical Hodgkin lymphoma after autologous stem cell transplantation. Blood. 2019;134(1):22–29. 10.1182/blood.2019000215.
13. Chen R, Zinzani PL, Lee HJ, Armand P, Johnson NA, Brice P, Radford J, Ribrag V, Molin D, Vassilakopoulos TP, *et al.* Pembrolizumab in relapsed or refractory Hodgkin lymphoma: 2-year follow-up of KEYNOTE-087. Blood. 2019;134(14):1144–1153. 10.1182/blood.2019000324.
14. Xu-Monette ZY, Zhang M, Li J, Young KH. PD-1/PD-L1 blockade: have we found the key to unleash the antitumor immune response? Front Immunol. 2017;8:1597. 10.3389/fimmu.2017.01597.
15. Ansell SM, Minnema MC, Johnson P, Timmerman JM, Armand P, Shipp MA, Rodig SJ, Ligon AH, Roemer MGM, Reddy N, *et al.* Nivolumab for relapsed/refractory diffuse large B-cell lymphoma in patients ineligible for or having failed autologous transplantation: a Single-Arm, phase II study. J Clin Oncol. 2019;37(6):481–489. 10.1200/JCO.18.00766.
16. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, Lawrence MS, Roemer MGM, Li AJ, Ziepert M, *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. 10.1038/s41591-018-0016-8.
17. Weiner GJ. Rituximab: mechanism of action. Semin Hematol. 2010;47(2):115–123. 10.1053/j.seminhematol.2010.01.011.
18. Pérez-Callejo D, González-Rincón J, Sánchez A, Provencio M, Sánchez-Beato M. Action and resistance of monoclonal CD20 antibodies therapy in B-cell non-Hodgkin lymphomas. Cancer Treat Rev. 2015;41(8):680–689. 10.1016/j.ctrv.2015.05.007.
19. Visco C, Li Y, Xu-Monette ZY, Miranda RN, Green TM, Li Y, Tzankov A, Wen W, Liu WM, Kahl BS, *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(9):2103–2113. 10.1038/leu.2012.83.
20. Xu-Monette ZY, Xiao M, Au Q, Padmanabhan R, Xu B, Hoe N, Rodriguez-Perales S, Torres-Ruiz R, Manyam GC, Visco C, *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(4):644–657. 10.1158/2326-6066.CIR-18-0439.
21. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001;98(9):5116–5121. 10.1073/pnas.091062498.
22. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998;95(25):14863–14868. 10.1073/pnas.95.25.14863.
23. Hosack DA, Dennis G Jr., Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. Genome Biol. 2003;4(10):R70. 10.1186/gb-2003-4-10-r70.
24. Hu S, Xu-Monette ZY, Tzankov A, Green T, Wu L, Balasubramanyam A, Liu WM, Visco C, Li Y, Miranda RN, *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. 10.1182/blood-2012-10-460063.
25. Tzankov A, Xu-Monette ZY, Gerhard M, Visco C, Dirnhofer S, Gisin N, Dybkaer K, Orazi A, Bhagat G, Richards KL, *et al.* Rearrangements of MYC gene facilitate risk stratification in diffuse large B-cell lymphoma patients treated with rituximab-CHOP. Mod Pathol. 2014;27(7):958–971. 10.1038/modpathol.2013.214.
26. Xu-Monette ZY, Wu L, Visco C, Tai YC, Tzankov A, Liu WM, Montes-Moreno S, Dybkaer K, Chiu A, Orazi A, *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. 10.1182/blood-2012-05-433334.
27. Deng M, Xu-Monette ZY, Pham LV, Wang X, Tzankov A, Fang X, Zhu F, Visco C, Bhagat G, Dybkaer K, *et al.* Aggressive B-cell lymphoma with MYC/TP53 dual alterations displays distinct clinicopathobiological features and response to novel targeted agents. Mol Cancer Res. 2021;19(2):249–260. 10.1158/1541-7786.MCR-20-0466.
28. Xu-Monette ZY, Moller MB, Tzankov A, Montes-Moreno S, Hu W, Manyam GC, Kristensen L, Fan L, Visco C, Dybkaer K, et al. MDM2 phenotypic and genotypic profiling, respective to TP53 genetic status, in diffuse large B-cell lymphoma patients treated with rituximab-CHOP immunochemotherapy: a report from the International DLBCL Rituximab-CHOP Consortium Program. Blood. 2013;122(15): 2630–2640. 10.1182/blood-2012-12-473702.
29. Challa-Malladi M, Lieu YK, Califano O, et al. Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. Cancer Cell. 2011;20: 728-740. 10.1016/j.ccr.2011.11.006.
30. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, Kasper LH, Lerach S, Tang H, Ma J, *et al.* Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature. 2011;471(7337):189–195. 10.1038/nature09730.
31. Wise JF, Nakken S, Steen CB, Vodák D, Trøen G, Johannessen B, Lingjærde OC, Hilden V, Blaker YN, Bai B, *et al.* Mutational dynamics and immune evasion in diffuse large B-cell lymphoma explored in a relapse-enriched patient series. Blood Adv. 2020;4(9):1859–1866. 10.1182/bloodadvances.2019001325.
32. Xu-Monette ZY, Li J, Xia Y, Crossley B, Bremel RD, Miao Y, Xiao M, Snyder T, Manyam GC, Tan X, *et al.* Immunoglobulin somatic hypermutation has clinical impact in DLBCL and potential implications for immune checkpoint blockade and neoantigen-based immunotherapies. J Immunother Cancer. 2019;7(1):272. 10.1186/s40425-019-0730-x.
33. Monti S, Chapuy B, Takeyama K, Rodig SJ, Hao Y, Yeda KT, Inguilizian H, Mermel C, Currie T, Dogan A, *et al.* Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. Cancer Cell. 2012;22(3):359–372. 10.1016/j.ccr.2012.07.014.
34. Miao Y, Xu-Monette ZY, Li J, Young KH. Dysregulation of cell survival in diffuse large B cell lymphoma: mechanisms and therapeutic targets. Front Oncol. 2019;9:107. 10.3389/fonc.2019.00107.
35. Putiri EL, Robertson KD. Epigenetic mechanisms and genome stability. Clin Epigenetics. 2011;2(2):299–314. 10.1007/s13148-010-0017-z.
36. Chong SY, Cutler S, Lin JJ, Tsai CH, Tsai HK, Biggins S, Tsukiyama T, Lo YC, Kao CF. H3K4 methylation at active genes mitigates transcription-replication conflicts during replication stress. Nat Commun. 2020;11(1):809. 10.1038/s41467-020-14595-4.
37. Ortega-Molina A, Boss IW, Canela A, Pan H, Jiang Y, Zhao C, Jiang M, Hu D, Agirre X, Niesvizky I, *et al.* The histone lysine methyltransferase KMT2D sustains a gene expression program that represses B cell lymphoma development. Nat Med. 2015;21(10):1199–1208. 10.1038/nm.3943.
38. Zhang J, Dominguez-Sola D, Hussein S, Lee JE, Holmes AB, Bansal M, Vlasevska S, Mo T, Tang H, Basso K, *et al.* Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. Nat Med. 2015;21(10):1190–1198. 10.1038/nm.3940.
39. Wang G, Chow RD, Zhu L, Bai Z, Ye L, Zhang F, Renauer PA, Dong MB, Dai X, Zhang X, *et al.* CRISPR-GEMM Pooled mutagenic screening identifies KMT2D as a major modulator of immune checkpoint blockade. Cancer Discov. 2020;10(12):1912–1933. 10.1158/2159-8290.CD-19-1448.
40. Miao Y, Li Y, Li J, Young KH. Genetic alterations and their clinical implications in DLBCL. Nat Rev Clin Oncol. 2019;16(10):634–652.
41 Shi Y, Lei Y, Liu L, Zhang S, Wang W, Zhao J, Zhao S, Dong X, Yao M, Wang K, et al. Integration of comprehensive genomicprofiling, tumor mutational burden, and PD-L1 expression toidentify novel biomarkers of immunotherapy in non-small celllung cancer. Cancer Med. 2021;10(7):2216–2231. 10.1002/cam4.3649
42. Ardeshir-Larijani F, Bhateja P, Lipka MB, Sharma N, Fu P, Dowlati A. KMT2D mutation is associated with poor prognosis in non-small-cell lung cancer. Clin Lung Cancer. 2018;19(4):e489–e501. 10.1016/j.cllc.2018.03.005
43. Shen J, Ju Z, Zhao W, Wang L, Peng Y, Ge Z, Nagel ZD, Zou J, Wang C, Kapoor P, *et al.* ARID1A deficiency promotes mutability and potentiates therapeutic antitumor immunity unleashed by immune checkpoint blockade. Nat Med. 2018;24(5):556–562. 10.1038/s41591-018-0012-z.
44. Okamura R, Kato S, Lee S, Jimenez RE, Sicklick JK, Kurzrock R. ARID1A alterations function as a biomarker for longer progression-free survival after anti-PD-1/PD-L1 immunotherapy. J Immunother Cancer. 2020;8(1):e000438. 10.1136/jitc-2019-000438.
45. Wang L, Qu J, Zhou N, Hou H, Jiang M, Zhang X. Effect and biomarker of immune checkpoint blockade therapy for ARID1A deficiency cancers. Biomed Pharmacother. 2020;130:110626. 10.1016/j.biopha.2020.110626.
46. Wu H-X, Chen Y-X, Wang Z-X, Zhao Q, He -M-M, Wang Y-N, Wang F, Xu R-H. Alteration in TET1 as potential biomarker for immune checkpoint blockade in multiple cancers. J Immunother Cancer. 2019;7(1):264. 10.1186/s40425-019-0737-3.
47. Zhang R, Wu HX, Xu M, Xie X. KMT2A/C mutations function as a potential predictive biomarker for immunotherapy in solid tumors. Biomark Res. 2020;8(1):71. 10.1186/s40364-020-00241-0.
48. Zhu G, Pei L, Li Y, Gou X. EP300 mutation is associated with tumor mutation burden and promotes antitumor immunity in bladder cancer patients. Aging (Albany NY). 2020;12(3):2132–2141. 10.18632/aging.102728.
49. Huang YH, Cai K, Xu PP, Wang L, Huang CX, Fang Y, Cheng S, Sun XJ, Liu F, Huang JY, *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. 10.1038/s41392-020-00437-8.
50. Xu-Monette ZY, Young KH. Therapeutic vaccines for aggressive B-cell lymphoma. Leuk Lymphoma. 2020;61(13):3038–3051. 10.1080/10428194.2020.1805113.