



Subtype-specific and co-occurring genetic alterations in B-cell non-Hodgkin lymphoma

Man Chun John Ma,^{1*} Saber Tadros,^{1*} Alyssa Bouska,² Tayla Heavican,² Haopeng Yang,¹ Qing Deng,¹ Dalia Moore,³ Ariz Akhter,⁴ Keenan Hartert,³ Neeraj Jain,¹ Jordan Showell,¹ Sreejoyee Ghosh,¹ Lesley Street,⁵ Marta Davidson,⁵ Christopher Carey,⁶ Joshua Tobin,⁷ Deepak Perumal,⁸ Julie M. Vose,⁹ Matthew A. Lunning,⁹ Aliyah R. Sohani,¹⁰ Benjamin J. Chen,¹¹ Shannon Buckley,¹² Loretta J. Nastoupil,¹ R. Eric Davis,¹ Jason R. Westin,¹ Nathan H. Fowler,¹ Samir Parekh,⁸ Maher Gandhi,⁷ Sattva Neelapu,¹ Douglas Stewart,⁵ Kapil Bhalla,¹³ Javeed Iqbal,² Timothy Greiner,² Scott J. Rodig,¹⁴ Adnan Mansoor⁵ and Michael R. Green^{1,14,15}

Haematologica 2022
Volume 107(3):690-701

¹Department of Lymphoma and Myeloma, Division of Cancer Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; ²Department of Pathology and Microbiology, University of Nebraska Medical Center, Omaha, NE, USA; ³Eppley Institute for Research in Cancer and Allied Diseases, University of Nebraska Medical Center, Omaha, NE, USA; ⁴Department of Pathology and Laboratory Medicine, University of Calgary, Calgary, Alberta, Canada; ⁵Section of Hematology, Department of Medicine, University of Calgary, Calgary, Alberta, Canada; ⁶Northern Institute for Research, Newcastle University, Newcastle upon Tyne, UK; ⁷Diamantina Institute, University of Queensland, Queensland, Australia; ⁸Division of Hematology and Medical Oncology, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ⁹Department of Internal Medicine, Division of Hematology-Oncology, University of Nebraska Medical Center, Omaha, NE, USA; ¹⁰Department of Pathology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA; ¹¹Department of Pathology, University of Massachusetts Medical School, UMass Memorial Medical Center, Worcester, MA, USA; ¹²Department of Genetics, Cell Biology and Anatomy, University of Nebraska Medical Center, Omaha, NE, USA; ¹³Department of Pathology, Brigham and Womens Hospital, Boston, MA, USA; ¹⁴Department of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, TX, USA and ¹⁵Center for Cancer Epigenetics, University of Texas MD Anderson Cancer Center, Houston, TX, USA.

*MCJM and ST contributed equally as co-first authors.

ABSTRACT

B-cell non-Hodgkin lymphoma (B-NHL) encompasses multiple clinically and phenotypically distinct subtypes of malignancy with unique molecular etiologies. Common subtypes of B-NHL, such as diffuse large B-cell lymphoma, have been comprehensively interrogated at the genomic level, but rarer subtypes, such as mantle cell lymphoma, remain less extensively characterized. Furthermore, multiple B-NHL subtypes have thus far not been comprehensively compared using the same methodology to identify conserved or subtype-specific patterns of genomic alterations. Here, we employed a large targeted hybrid-capture sequencing approach encompassing 380 genes to interrogate the genomic landscapes of 685 B-NHL tumors at high depth, including diffuse large B-cell lymphoma, mantle cell lymphoma, follicular lymphoma, and Burkitt lymphoma. We identified conserved hallmarks of B-NHL that were deregulated in the majority of tumors from each subtype, including frequent genetic deregulation of the ubiquitin proteasome system. In addition, we identified subtype-specific patterns of genetic alterations, including clusters of co-occurring mutations and DNA copy number alterations. The cumulative burden of mutations within a single cluster were more discriminatory of B-NHL subtypes than individual mutations, implicating likely patterns of genetic cooperation that contribute to disease etiology. We therefore provide the first cross-sectional analysis of mutations and DNA copy number alterations across major B-NHL subtypes and a framework of co-occurring genetic alterations that deregulate genetic hallmarks and likely cooperate in lymphomagenesis.

Correspondence:

MICHAEL R. GREEN
mgreen5@mdanderson.org

Received: October 16, 2020.

Accepted: March 15, 2021.

Pre-published: April 1, 2021.

<https://doi.org/10.3324/haematol.2020.274258>

©2022 Ferrata Storti Foundation

Material published in *Haematologica* is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>. Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions: <https://creativecommons.org/licenses/by-nc/4.0/legalcode>, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.



Introduction

Non-Hodgkin lymphomas (NHL) are a heterogeneous group of lymphoid malignancies that predominantly arise from mature B cells (B-NHL). Although mature B-cell neoplasms encompass 38 unique diagnostic subtypes, over 85% of cases fall within only seven histologies.^{1,2} Recent next-generation sequencing studies have shed light onto the key driver mutations in many of these B-NHL subtypes; for example, large studies of diffuse large B-cell lymphoma (DLBCL) have led to proposed genomic subtypes that have unique etiologies.^{3,5} However, many less common NHL subtypes such as mantle cell lymphoma (MCL) have not been as extensively characterized.^{6,7} Furthermore, until recently^{3,4} genetic alterations have been considered in a binary fashion as either driver events, which directly promote disease genesis or progression, or passenger events, which have little or no impact on disease biology. In contrast to this principle, most B-NHL do not result from a single dominant driver but instead result from the serial acquisition of genetic alterations that cooperate in lymphomagenesis.⁸ The genetic context of each mutation likely determines its oncogenic potential, and groups of mutations should therefore be considered collectively rather than as singular events. For example, the 'C5' and 'MCD' clusters identified in DLBCL by Chapuy *et al.* and Schmitz *et al.*, respectively, are characterized by the co-occurrence of *CD79B* and *MYD88* mutations.^{3,4} In animal models, the *Myd88* L252P mutation (equivalent to human L265P) was found to promote downregulation of surface IgM and a phenotype resembling B-cell anergy.⁹ However, this effect could be rescued by the *Cd79b* mutation, showing that these co-occurring mutations cooperate.⁹ The characterization of other significantly co-occurring genetic alterations are therefore likely to reveal additional important cooperative relationships. We approached this challenge by performing genomic profiling of 685 B-NHL across different subtypes. Through this cross-sectional analysis we characterized genomic hallmarks of B-NHL and sets of significantly co-associated events that likely represent subtype-specific cooperating genetic alterations. This study therefore provides new insight into how co-occurring clusters of genetic alterations may contribute to molecularly and phenotypically distinct subtypes of B-NHL.

Methods

An overview of our approach is shown in *Online Supplementary Figure S1*. For detailed methods, please refer to the *Online Supplementary Information*.

Tumor DNA samples

We collected DNA from 685 B-NHL tumors, including 199 FL, 196 MCL, 148 DLBCL, 107 BL, 21 high-grade B-cell lymphoma not otherwise specified (HGBL-NOS), and 14 high-grade B-cell lymphoma with *MYC*, *BCL2* and/or *BCL6* rearrangement (DHL) (*Online Supplementary Table S1*). All samples were archival and de-identified. The study was approved by the institutional review board of the University of Nebraska Medical Center (203-15-EP) and performed in accordance with the Declaration of Helsinki. A total of 462 samples were obtained from the University of Nebraska Medical Center, and were prioritized for inclusion in this study if they had previously undergone pathology review and been interrogated by Affymetrix U133 Plus 2.0 gene expression

microarrays¹⁰⁻¹² (n=284). An additional series of 223 formalin-fixed paraffin-embedded tumors were provided by other centers. Samples were de-identified and accompanied by the patients' diagnosis from the medical records, plus overall survival time and status when available. Medical record diagnosis was used in all cases except for those with fluorescence *in situ* hybridization (FISH) showing translocations in *MYC* and *BCL2* and/or *BCL6*, which were amended to DHL. Sequencing results for a subset of 52 BL tumors were described previously.¹³ All MCL samples were either positive for *CCND1* translocation by FISH or positive for *CCND1* protein expression by immunohistochemistry, depending on the diagnostic practices of the contributing institution.

Next-generation sequencing

A total of 500-1000 ng of genomic DNA was sonicated using a Covaris S2 Ultrasonicator, and libraries prepared using KAPA Hyper Prep Kits (Roche) with TruSeq Adapters (Bio Scientific) and a maximum of eight cycles of polymerase chain reaction (average of 4 cycles). Libraries were qualified by TapeStation 4200, quantified by Qubit and 10- to 12-plexed for hybrid capture. Each multiplexed library was enriched using our custom LymphoSeq panel encompassing the full coding sequences of 380 genes that were determined to be somatically mutated in B-cell lymphoma (*Online Supplementary Table S2*, *Online Supplementary Methods*), as well as tiling recurrent translocation breakpoints. Enrichments were amplified with four to eight cycles of polymerase chain reaction and sequenced on a single lane of an Illumina HiSeq 4000 with 100PE reads in high-output mode at the Hudson Alpha Institute for Biotechnology or the MD Anderson Sequencing and Microarray Facility. Variants were called using our previously validated ensemble approach,^{13,14} germline polymorphisms were filtered using dbSNP annotation and the EXAC dataset containing 60,706 healthy individuals,¹⁵ and significantly mutated genes were defined by MutSig2CV.¹⁶ Copy number alterations (CNA) identified by CopyWriteR,¹⁷ which was validated using three FL tumors with matched Affymetrix 250K SNP array (*Online Supplementary Figure S2*), and significant DNA CNA were determined by GISTIC2.¹⁸ Translocations were called using FACTERA,¹⁹ which we previously validated against *MYC* translocation status determined by FISH.²⁰ Mutation and CNA data are publicly viewable through cBioPortal: https://www.cbioportal.org/study/summary?id=mbn_mdacc_2013. Matched gene expression microarray data are available through the Gene Expression Omnibus, accession GSE132929. For further details, refer to the *Online Supplementary Methods*.

Results

Identification of significantly mutated genes and structural alterations

We used a 380-gene custom targeted sequencing approach, LymphoSeq, to interrogate the genomes of 685 mature B-NHL, sequencing to an average depth of 578X (minimum, 101X; maximum, 1785X) (*Online Supplementary Table S1*) for a total yield of 1.81 Tbp. Somatic nucleotide variants and small insertions/deletions were identified using an ensemble approach that we have previously validated¹⁴ (*Online Supplementary Table S3*) and significantly mutated genes were identified using MutSig2CV (*Online Supplementary Table S4*). Matched germline DNA was available from purified T cells of 20 tumors (11 FL and 9 MCL) and sequenced to validate the filtering of germline variants; 0/632 variants called within these tumors were identified in the matched germline samples, which indicates that the fil-

tering of germline variants was effective. Genes that were significantly mutated in the full cohort or in any one of the four subtypes with more than 100 tumors (BL, DLBCL, FL, and MCL) were included, as well as frequently mutated genes that are targets of activation-induced cytidine deaminase (AID) (Figure 1, *Online Supplementary Table S5*). Predictably, the frequency of AID-associated mutations was higher among germinal center-derived lymphomas (BL, DLBCL, FL), but also accounted for 7.6% of all coding and non-coding mutations in MCL (*Online Supplementary Table S6*). The mutational burden calculated from our targeted region correlated significantly with that from the whole exome (*Online Supplementary Figure S3A*) and was significantly higher in DLBCL and other high-grade tumors than in FL and MCL (Figure 1, *Online Supplementary Figure S3B*).

The hybrid capture probes utilized in our design also targeted recurrent breakpoint regions in the immunoglobulin heavy- and light-chain loci, and recurrent breakpoints in or near the *BCL2*, *MYC* and *BCL6* genes, and translocations were called using a method that detects discordantly mapped reads¹⁹ (Figures 1 and 2A). Our prior validation of this approach in cases with matched FISH data for *MYC* showed that it is 100% specific, but only ~40% sensitive for translocation detection.¹⁵ This limit of sensitivity likely varies for different genes depending on how well the breakpoints are clustered into hotspots that are targeted by our capture probes. Nonetheless, we observed a significantly higher fraction of *BCL6* translocations (57% [27/47]) partnered to non-immunoglobulin loci (e.g., *CIITA*, *RHOH*, *EIF4A2*, and *ST6GAL1*) (*Online Supplementary Table S7*) compared to *BCL2* (1% [1/114]) and *MYC* (5% [2/38]) translocations (Figure 2A; Fisher *P*-value <0.001). These were more frequent in FL (88% [15/17] of *BCL6* translocations) than in DLBCL (39% [9/23] of *BCL6* translocations), presumably because the two immunoglobulin loci in FL are either translocated with the *BCL2* gene or functioning in immunoglobulin expression.²¹ We also employed off-target reads to detect DNA CNA in a manner akin to low-pass whole genome sequencing, identified significant peaks of copy gain and losses using GISTIC2¹⁸ (Figures 1 and 2A, *Online Supplementary Figure S4*, *Online Supplementary Tables S8* and *S9*), and defined the likely targets of these CNA by integrative analysis of matched gene expression profiling data from 290 tumors (Figure 2B, C, *Online Supplementary Figure S4*, *Online Supplementary Tables S10* and *S11*). This identified known CNA targets, including but not limited to deletion of *TNFAIP3* (6q24.2),²² *ATM* (11q22.3),²³ *B2M* (15q15.5),²⁴ and *PTEN* (10q23.21),²⁵ and copy gain of *REL* and *BCL11A* (2p15), and *TCF4* (18q23).²⁶ In addition, we identified novel targets such as deletion of *IBTK* (6q14.1), *UBE3A* (11q22.1) and *FBXO25* (8p23.3), and copy gain of *ATF7* (12q13.13), *UCLH5* (1q31.3), and *KMT2A* (11q23.3). Importantly, the frequency of DNA CNA in the target genes identified by next-generation sequencing-based analysis significantly correlated with those derived from single nucleotide polymorphism microarray-based measurements in independent cohorts of BL, DLBCL, FL and MCL tumors from previously published studies^{6,20,26-30} (*Online Supplementary Figure S5*), providing validation for the accuracy of this approach. The CNA peaks, defined as the smallest and most statistically significant region, included multiple genes that were significantly mutated (Figure 2D) as well as other genes for which we detected mutations at lower frequencies that were not significant by

MutSig2CV (*POU2AF1*, *TP53BP1*, *FAS*, *PTEN*). Deletions of *ATM*, *B2M*, *BIRC3* and *TNFRSF14* significantly co-associated with mutations of these genes, suggesting that these are complementary mechanisms contributing to biallelic inactivation.

Conserved functional hallmarks of B-cell non-Hodgkin lymphoma

To understand key hallmarks that are deregulated by genetic alterations, we performed hypergeometric enrichment analysis of genes targeted by recurrent mutations and DNA CNA using DAVID³¹ (*Online Supplementary Table S12*). This revealed a significant enrichment of multiple overlapping gene sets that could be summarized into hallmark processes associated with epigenetics and transcription (Figure 3A), apoptosis and proliferation (Figure 3B), signaling (Figure 3C), and ubiquitination (Figure 3D). One or more genes within these hallmarks was altered in the majority (>50%) of tumors from each of the four major histologies included in this study. Genes annotated in epigenetic-associated gene sets were altered in 72%, 70%, 93% and 50% of BL, DLBCL, FL, and MCL, respectively, whereas genes annotated in transcription-associated gene sets were altered in 94%, 91%, 95% and 88% of BL, DLBCL, FL, and MCL, respectively. However, there is an extremely high degree of functional overlap between epigenetics and transcriptional regulation, as well as overlapping gene set annotations for many genes, leading us to consider these categories collectively as a single hallmark. Collectively, genes involved in epigenetics and transcription were mutated in 94% of BL, 92% of DLBCL, 96% of FL and 89% of MCL, and included those that encode proteins that catalyze post-translational modifications of histones (*KMT2D*, *CREBBP*, *EZH2*, *EP300*, *WHSC1*, *ASH1L1L*, *KMT2A*), components of the BAF chromatin remodeling complex (*ARID1A*, *SMARCA4*, *BCL7A*, *BCL11A*), linker histones (*HIST1H1E*, *HIST1H1C*, *HIST1H1B*), and transcription factors (*BCL6*, *IRF4*, *IRF8*, *TCF3*, *TCF4*, *MYC*, *REL*, *PAX5*, *POU2AF2*, *FOXO1*, *CIITA*). Genes with a role in signaling included those involved in B-cell receptor signaling (*CD79B*, *ID3*, *TCF3*, *TCF4*, *RFTN1*), NFκB (*TNFAIP3*, *CARD11*, *NFKBIE*), NOTCH (*NOTCH1*, *NOTCH2*), JAK/STAT (*SOCS1*, *STAT6*), PI3K/mTOR (*FOXO1*, *ATP6V1B2*, *APT6AP1*) and G-protein signaling (*GNA13*, *GNAI2*). The *CD79A* and *BCL10* genes were also mutated at a lower frequency that was not significant by MutSig2CV (*Online Supplementary Figure S6A, B*). Among these, the *RFTN1* gene (*Online Supplementary Figure S6C*) is a novel recurrently mutated gene that was mutated in 7.4% of DLBCL and encodes a lipid raft protein that is critical for B-cell receptor signaling.³²

Deregulation of the ubiquitin proteasome system is important in many cancers,³³ but is not a well-defined hallmark of B-NHL. However, one or more genes with a role in regulating ubiquitination was genetically altered in 61% of BL, 79% of DLBCL, 61% of FL and 82% of MCL (Figure 3D). These included previously described genetic alterations such as amplification of *MDM2*,³⁴ deletions of *TNFAIP3*,³⁵ *CUL4A*,³⁶ and *RPL5*,³⁶ and mutations of *KLHL6*,³⁷ *DTX1*,³⁸ *UBR5*,³⁹ *SOCS1*,⁴⁰ and *BIRC3*.⁶ In addition, we identified novel targets such as recurrent deletions of *IBTK*, a negative regulator of Bruton tyrosine kinase,⁴¹ and somatic mutation of *CDC27* in 14% of MCL, which encodes an E3 ligase for CCND1.⁴² Therefore, common hallmark processes are targeted by genetic alterations in the majority of

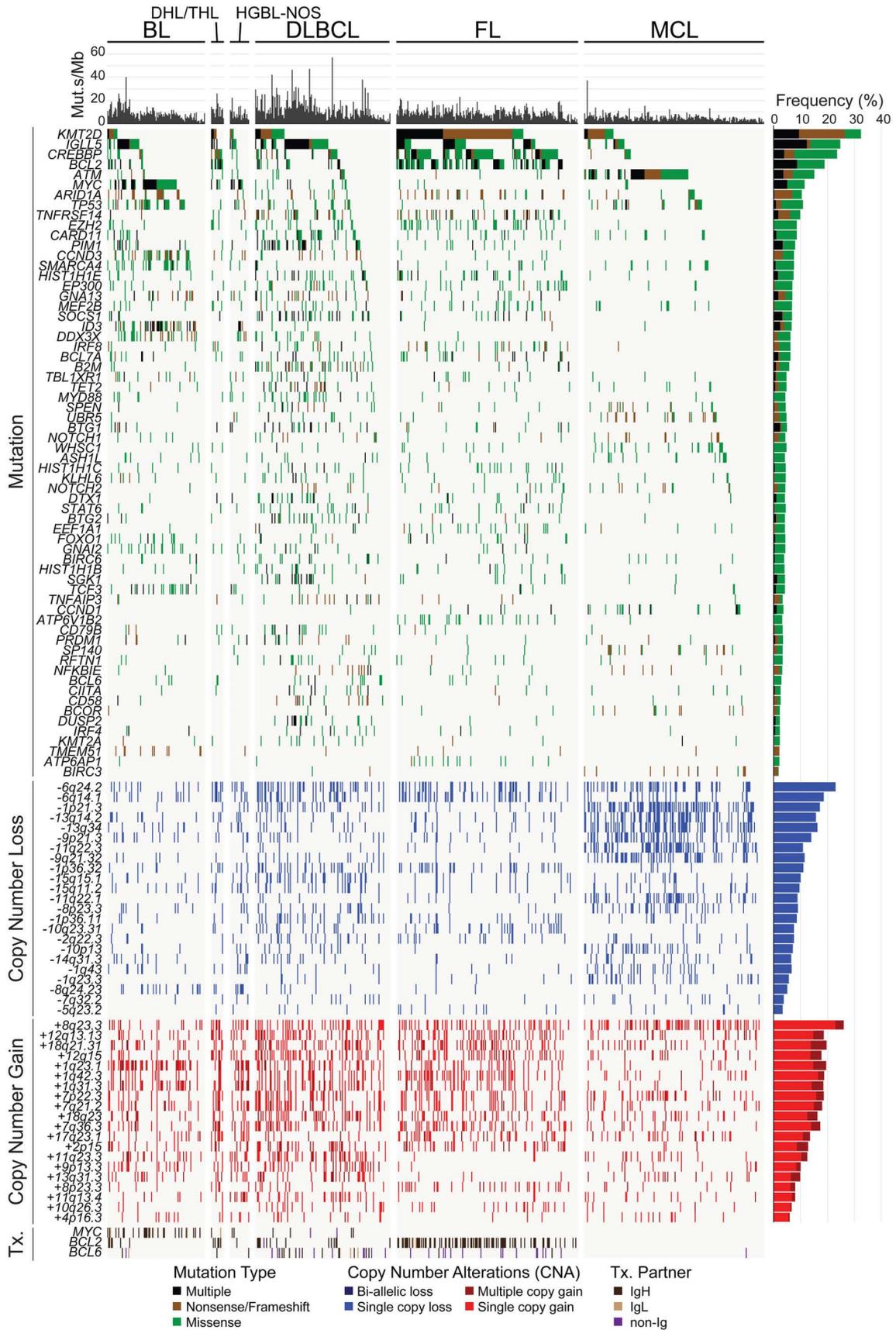


Figure 1. Recurrently mutated genes in subtypes of B-cell non-Hodgkin lymphoma. An oncoPrint shows significantly mutated genes, DNA copy number alterations (CNA) and translocations (Tx.) across our cohort of 685 B-cell non-Hodgkin lymphoma tumors. Mutation types and frequencies are summarized for each gene/CNA on the right, and the mutational burden for each case is shown at the top. DHL: double-hit lymphoma; THL: triple-hit lymphoma; HGBL-NOS: high-grade B-cell lymphoma not otherwise specified; BL: Burkitt lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; MCL: mantle cell lymphoma.

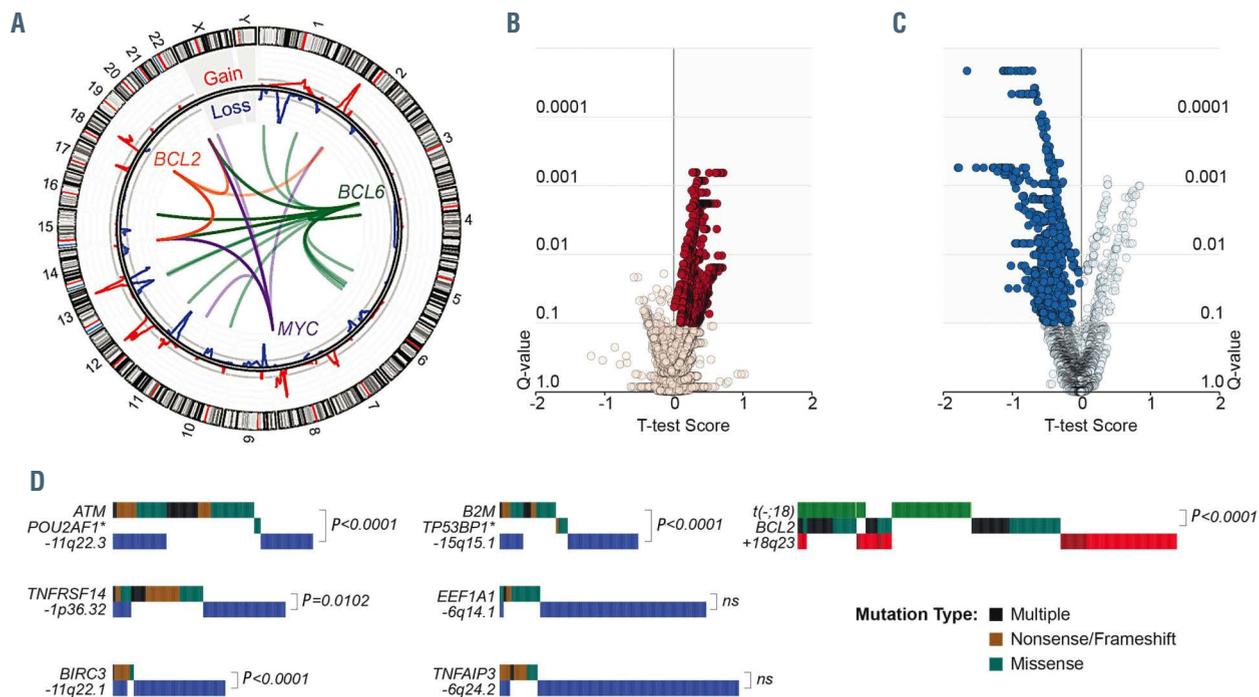


Figure 2. Structural alterations in subtypes of B-cell non-Hodgkin lymphoma. (A) A circos plot shows translocations of *MYC* (purple), *BCL2* (orange) and *BCL6* (green) genes, and GISTIC tracks of DNA copy number gains (red) and losses (blue). (B, C) Volcano plots of integrative analysis results showing the changes in gene expression of genes within peaks of DNA copy number gain (B) or loss (C). A positive T-test score indicates increased expression in tumors with a given copy number alteration, and vice versa. Significantly expressed genes with the correct directionality are highlighted in the shaded areas. (D) Oncoplots show the overlap of structural alterations and mutations that target the same genes. P-values are derived from a Fisher exact test (ns: not significant).

major B-NHL subtypes, including genes with a role in regulating protein ubiquitination.

Subtype-specific patterns of genetic alterations

We formally tested the over- or under-representation of recurrent genetic alterations in each of the four subtypes with more than 100 samples (BL, DLBCL, FL, MCL), compared to all other tumors in the study (Figure 4, *Online Supplementary Table S13*). We observed some interesting patterns within hallmark characteristics that differ between subtypes. An illustrative example of this is the alternative avenues for BAF complex perturbation between different histologies (Figure 5). Specifically, mutations of the *SMARCA4* (aka. BRG1) component of the ATPase module were significantly enriched in BL (24%) compared to other subtypes (4%, Q-value <0.001), while mutations of the *BCL7A* component of the ATPase module were significantly enriched in FL (11%) compared to other subtypes (4%, Q-value=0.007). In contrast, mutations of *ARID1A* were frequent in both BL (19%) and FL (15%), and DNA copy number gains of *BCL11A* were frequent in both DLBCL (28%) and FL (22%). The BAF complex is therefore a target of recurrent genetic alterations, as previously suggested,⁴³ but the manner in which this complex is perturbed varies between B-NHL subtypes (Figure 5). Similar disease-specific patterns were also observed for signaling genes; for example, *TCF3* and *ID3* have important functions in normal germinal center B cells (GCB),⁴⁴ but mutations of these genes are specifically enriched within BL and are rarely found in the other GCB-derived malignancies, DLBCL and FL. Similarly, the *ATP6AP1* and *ATP6V1B2* genes that function in mTOR signaling^{45,46} are specifically mutated in FL, and the *DUSP2* gene which inactivates ERK1/2⁴⁷ and

STAT3⁴⁸ is specifically mutated in DLBCL. The disease-specific patterns of genetic alterations therefore reveal subtle but important differences in how each subtype of B-NHL perturbs hallmark features.

Clusters of co-associated genomic alterations in subtypes of B-cell non-Hodgkin lymphoma

We next defined how each genetic alteration co-associated with or mutually excluded other genetic alterations by pairwise assessments using false-discovery rate (FDR)-corrected Fisher tests (*Online Supplementary Table S14*). A matrix of the transformed FDR Q-values (-logQ) was used for unsupervised hierarchical clustering to identify clusters of co-associated genetic alterations. Together with patterns of disease-specificity, unsupervised clustering revealed clear groupings of co-associated events for BL, DLBCL, FL and MCL (Figure 4). We identified a single cluster of significantly co-associated genetic alterations that was specifically enriched in BL (Cluster 1), including mutations and translocations of *MYC*, and mutations of *CCND3*, *SMARCA4*, *TCF3* and *ID3* which have been previously reported in BL.⁴ A single cluster was significantly enriched in MCL (Cluster 7), with a high frequency of *ATM* mutations and deletions, as well as other DNA CNA. Other mutations that were not significantly co-associated were also enriched in MCL (Cluster 6), such as those in *WHSC1*, *NOTCH1*, *NOTCH2*, *BCOR* and *UBR5*, although statistical assessment of co-association may be hampered in this context by the low frequencies of mutations within these genes. A single cluster was also enriched in FL (Cluster 4), with a high prevalence of *KMT2D*, *BCL2*, *CREBBP*, *EZH2* and *TNFRSF14* mutations and *BCL2* translocations. The genes within Cluster 4 also significantly overlapped with the previously reported

C3, EZB and BCL2 clusters from prior whole exome sequencing studies of DLBCL^{3,49,50} (Fisher test *P*-values: *P*=0.0006, *P*=0.0148 and *P*=0.0173, respectively). Two clusters (Clusters 2 and 3) were enriched in DLBCL, with lower frequencies of mutations in a larger number of genes, in line with the genetic heterogeneity of this disease.^{3,4} Cluster 2 includes co-associated genetic alterations that overlapped with the previously described C5, MCD, and MYD88 clusters^{3,4} (Fisher *P*-values: *P*=0.0004, *P*=0.0002 and *P*=0.0007, respectively) including *CD79B*, *MYD88* and *TBL1XR1* mutations. Genes within Cluster 3 overlapped in a statistically significant manner with those in the previously

described C4 and SOCS1/SGK1 clusters (Fisher *P*-values: *P*=0.0002 and *P*=0.0074, respectively), including *SGK1*, *TET2*, *SOCS1* and histone H1 genes. We also identified a cluster consisting of *TP53* mutations and multiple CNA (Cluster 5) similar to the genetically complex C2/A53 subtype reported in DLBCL,^{3,49} however, the overlap of features within these clusters could not be formally assessed due to differing annotations. The CNA captured in this cluster were variably represented across B-NHL subtypes, but were most frequent in DLBCL. B-NHL subtypes therefore harbor characteristic clusters of co-associated genetic alterations that likely cooperate in disease etiology.

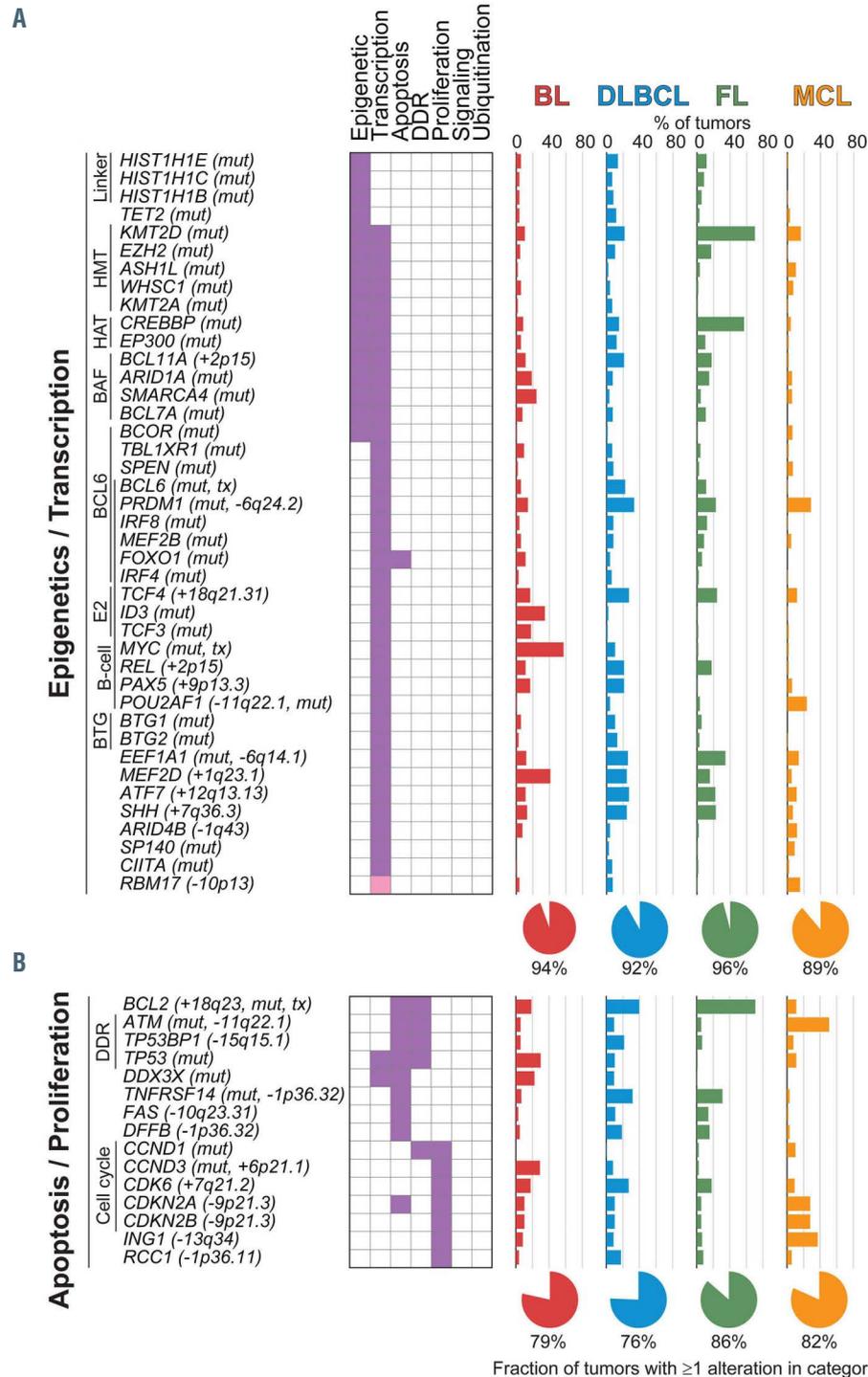


Figure 3. Figure continued on the following page.

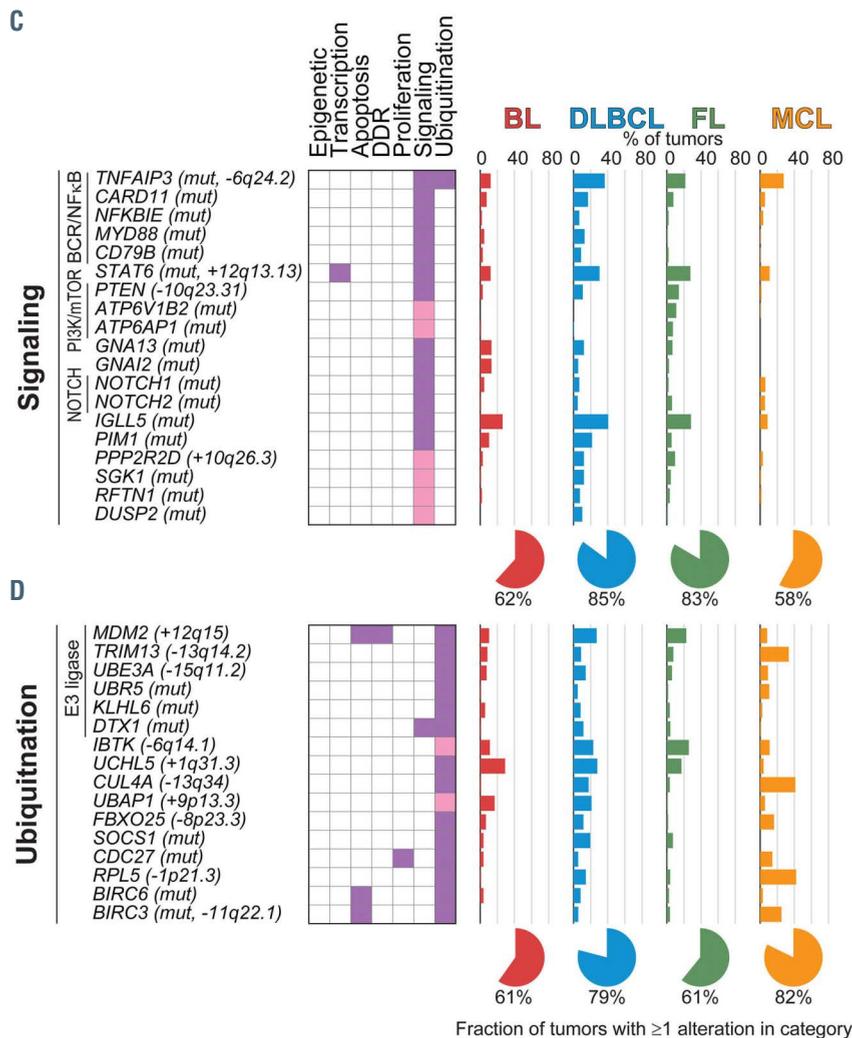


Figure 3. Functional enrichment of targets of somatic mutations and DNA copy number alterations. Genes targeted by somatic mutation and/or DNA copy number alteration were evaluated for enrichment in curated gene sets, and significant gene sets subsequently grouped according to overlapping gene set membership and functional similarity. In addition to genes assigned by DAVID (purple), some genes were manually curated into hallmark processes by literature review of their function (pink). Enriched gene sets could be summarized into four major hallmark processes, including (A) epigenetic and transcriptional control of gene expression, (B) regulation of apoptosis and proliferation, (C) regulation of signaling pathway activity, and (D) regulation of protein ubiquitination. The frequency of each genetic alteration is shown for each of the four major histologies included in this study, and the fraction of tumors in each histology bearing genetic alterations of one or more of the genes is summarized by a pie graph at the bottom for each hallmark. BL: Burkitt lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; MCL: mantle cell lymphoma; HMT, histone methyltransferase. HAT, histone acetyltransferase. DDR, DNA damage response. BCR, B-cell receptor.

Combinations of genetic alterations define molecular subtypes of B-cell non-Hodgkin lymphoma

Our data have revealed statistical enrichment of individual genetic alterations in subtypes of B-NHL, and pairwise relationships between different genetic alterations that define clusters of subtype-specific events. To validate and expand upon these observations we leveraged gene expression microarray data from 284 tumors that underwent pathology review and were profiled as part of prior studies.¹⁰⁻¹² We utilized BL, DHL, HGBL-NOS and DLBCL tumors to perform classification into molecularly-defined BL (mBL) and non-mBL using a Bayesian classifier with previously described marker genes,⁵¹ and subclassified non-mBL into activated B-cell (ABC)-like and GCB-like subtypes as we have described previously²⁶ (Online Supplementary Figure S7). We evaluated the frequency of cumulative (≥ 2) genetic alterations within each cluster among mBL, ABC-like DLBCL, GCB-like DLBCL, FL and MCL (Figure 6). This showed that Cluster 1 genetic alterations that were individually enriched in BL are cumulatively acquired in mBL, with 87% of tumors having ≥ 2 of these alterations compared to only 22% of GCB-like DLBCL. Similarly, Cluster 4 and Cluster 7 alterations were cumulatively acquired in 77% and 72% of molecularly-annotated FL and MCL, respectively. Cluster 4 mutations were also cumulatively acquired in 51% of GCB-like DLBCL, likely capturing the C3/EZB/BCL2 subtype that has genetic similarities to FL.^{3,4,50}

Furthermore, Cluster 2 and Cluster 4 alterations were cumulatively acquired in 58% of ABC-like DLBCL and 60% of GCB-like DLBCL, respectively, further supporting their respective overlap with the C5/MCD/MYD88 and C4/ST2 subtypes of DLBCL. CNA within Cluster 5 were cumulatively acquired at high but variable frequencies in all of the subtypes, but showed subtype-specific patterns within this cluster such as higher frequencies of 18q21 and 18q23 gains in ABC-like DLBCL, and higher frequencies of chromosome 7 gains in GCB-like DLBCL and FL. B-NHL tumors therefore cumulatively acquire co-associated sets of genetic alterations in a manner that is characteristically associated with histologically- and molecularly-defined subsets of disease.

Discussion

By performing cross-sectional genomic profiling of a large cohort of tumors, we have developed a resource of genes and functional hallmarks that are recurrently targeted by genetic alterations in B-NHL, and have shown that the cumulative acquisition of combinations of genetic alterations are characteristic of histological and molecular subtypes of disease. Some of the functional hallmarks that we identified have been previously appreciated, with a few exceptions. For example, the mutation of genes with roles

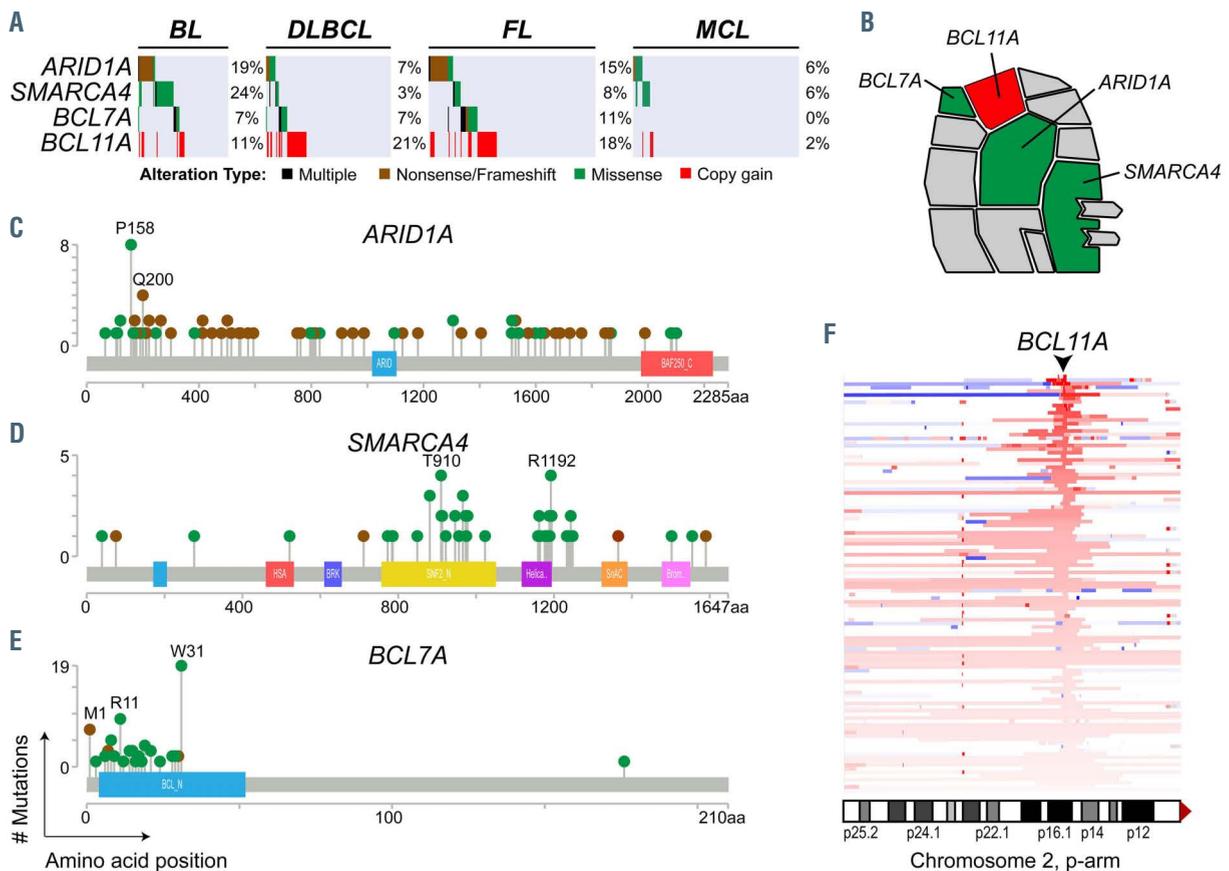


Figure 5. Subtype-specific patterns of BAF complex mutations. (A) An oncoplot shows the frequency of genetic alterations in genes that encode components of the BAF complex. (B) A schematic of the BAF complex shows recurrently mutated genes, *ARID1A*, *SMARCA4* and *BCL7A*, and the *BCL11A* gene that is targeted by 2p15 DNA copy number gains. (C-E) Lollipop plots show the distribution of mutations in the BAF components *ARID1A* (C), *SMARCA4* (D), and *BCL7A* (E). (F) A heatmap shows the location of chromosome 2p DNA copy number gains (red) ordered from highest DNA copy number (top) to lowest (bottom, copy number = 2.2). The *BCL11A* gene is in the peak focal copy gain. BL: Burkitt lymphoma; DLBCL: diffuse large B-cell lymphoma; FL: follicular lymphoma; MCL: mantle cell lymphoma.

in epigenetic and transcriptional control of gene expression are known to be a hallmark of FL⁵² and we observed that 96% of FL tumors possessed mutations in one or more of the genes in this category. However, mutations within these genes were also observed in the majority of BL, DLBCL and MCL tumors, highlighting the conservation of this functional hallmark across B-NHL subtypes. There are subtype-specific patterns of chromatin-modifying gene alterations, such as those that we highlighted for BAF complex mutations, but we suggest that the genetic deregulation of epigenetic and transcriptional control of gene expression should be considered a general hallmark of B-NHL. In addition, we suggest that the deregulation of the ubiquitin proteasome system is a hallmark of B-NHL that requires further investigation. Mutations in genes such as *KLHL6*⁵⁷ and *UBR5*⁵⁹ have recently been shown to play an important role in B-cell lymphoma, while the roles of other frequently mutated genes such as *DTX1* and *SOCS1* have not yet been functionally dissected. Furthermore, while the nature of AID-driven mutations in genes such as *DTX1* and *SOCS1* remain to be defined, other genes that are recurrently mutated by AID such as *BCL7A*⁵³ and linker histone genes⁵⁴ have been shown to play driving roles in lymphomagenesis. Genetic deregulation of the ubiquitin proteasome system has the potential to influence the activity or abundance of a range of substrate proteins, and represents a current gap in our knowledge of B-NHL etiology.

The role of cooperative interactions between co-occurring genetic alterations is also an emerging field that requires further investigation. These interactions are not uncommon in cancer,⁵⁵ and have been recently highlighted in DLBCL,^{3,4} but our data show that they are pervasive and characteristic features of the B-NHL genetic landscape. Cooperation between co-associated genetic alterations identified in this study requires formal validation in cell lines and/or animal models. However, there are many instances in which co-occurring genetic alterations that we observed have already been shown to cooperate in lymphomagenesis. In addition to the aforementioned example of *MYD88* and *CD79B* mutations, transgenic mouse models of *Ezh2* activating mutations or conditional deletion of *Crebbp* or *Kmt2d* have shown that these events are not alone sufficient for lymphomagenesis.⁵⁶⁻⁶¹ We and others have observed a co-association between mutation of these genes and *BCL2* translocations,^{14,62} and the addition of a *Bcl2* transgene to these murine models indeed promoted lymphoma at a significantly higher rate than that observed with the *Bcl2* transgene alone.⁵⁶⁻⁶¹ These genetic alterations are therefore significantly more lymphomagenic in combination than they are alone, which provides proof of principle that a cooperative relationship exists between these co-occurring genetic alterations. Future studies focusing on other co-occurring mutations, such as *MYC* translocation and *SMARCA4* mutation in BL, *CREBBP* and *KMT2D*

mutation in *FL*, *TCF4* copy gain and *MYD88* mutation in DLBCL, and *ATM* mutation and *RPL5* deletion in MCL, should therefore be performed to further explore these concepts and define their underlying functional relationship. We suggest that combinations of genetic alterations are likely to recapitulate the biology of B-NHL more accurately than are single gene models, and may reveal contextually different functional roles of genetic alterations depending on the co-occurring events.

The caveats regarding this study include the targeted nature of the LymphoSeq platform which may preclude consideration of a subset of important genes, the lack of germline DNA for the majority of samples that may lead to a small number of germ-line variants being falsely assigned as somatic, and the sample size for any given histological subtype being below that required to identify genes that are mutated at low frequency. Nonetheless, these data represent the first broad cross-sectional analysis of multiple histological and molecular subtypes of B-NHL using the same methodology and provide a framework of functional hallmarks and co-occurring genetic alterations that are enriched within these subtypes of B-NHL. These functional hallmarks are genetically perturbed in the majority of B-NHL, but our cross-sectional approach enabled us to elucidate subtype-specific preferences for genetic alterations within each functional hallmark. Furthermore, the subtype-specific clusters of co-occurring genetic alterations likely represent cooperative interactions that underpin the biology of different subtypes of B-NHL. These combinations identify opportunities for moving from single-allele to multi-allele designs in cell line or animal models to better understand the molecular etiology of B-NHL subtypes. Together,

these hallmarks and clusters of co-associated genetic alterations represent processes that are potentially druggable with targeted therapies,^{65,66} but that are likely influenced in a non-binary fashion by different combinations of genetic alterations. Deciphering the relationships between complex sets of genetic alterations and targetable dependencies will be a next step towards developing new rationally targeted therapeutic strategies in B-NHL.

Disclosures

No conflicts of interest to disclose.

Contributions

MCJM and ST performed experiments, analyzed data and wrote the manuscript. AB analyzed data. TH, HY, QD, DM, KH, NJ, JS, and SG performed experiments. AA, LS, MD, CC, JT, DP, KMV, MAL, ARS, BJC, RB, SN, LN, RED, JW, SP, MG, DS, KB, JI, SR, and AM provided samples and/or clinical data. MRG conceived and supervised the study, performed experiments, analyzed the data and wrote the manuscript. All authors reviewed and approved the manuscript.

Funding

This research was supported by NCI R01CA201380 (MRG), the Nebraska Department of Health and Human Services (LB506 2016-17; MRG), and NCI cancer center support grants to the University of Texas MD Anderson Cancer Center (P30 CA016672) and the Fred & Pamela Buffet Cancer Center (P30 CA036727). HY is supported by a Fellow award from the Leukemia and Lymphoma Society. MRG is supported by a Scholar award from the Leukemia and Lymphoma Society and an Andrew Sabin Family Foundation Fellow award.

References

- Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
- Armitage JO, Gascoyne RD, Lunning MA, Cavalli F. Non-Hodgkin lymphoma. *Lancet*. 2017;390(10091):298-310.
- Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med*. 2018;24(5):679-690.
- Schmitz R, Wright GW, Huang DW, et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. *N Engl J Med*. 2018;378(15):1396-1407.
- Reddy A, Zhang J, Davis NS, et al. Genetic and functional drivers of diffuse large B cell lymphoma. *Cell*. 2017;171(2):481-494.
- Bea S, Valdes-Mas R, Navarro A, et al. Landscape of somatic mutations and clonal evolution in mantle cell lymphoma. *Proc Natl Acad Sci U S A*. 2013;110(45):18250-18255.
- Zhang J, Jima D, Moffitt AB, et al. The genomic landscape of mantle cell lymphoma is related to the epigenetically determined chromatin state of normal B cells. *Blood*. 2014;123(19):2988-2996.
- Green MR, Alizadeh AA. Common progenitor cells in mature B-cell malignancies: implications for therapy. *Curr Opin Hematol*. 2014;21(4):333-340.
- Wang JQ, Jeelall YS, Humburg P, et al. Synergistic cooperation and crosstalk between MYD88L265P and mutations that dysregulate CD79B and surface IgM. *J Exp Med*. 2017;214(9):2759-2776.
- Lenz G, Wright G, Dave SS, et al. Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med*. 2008;359(22):2313-2323.
- Dave SS, Fu K, Wright GW, et al. Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med*. 2006;354(23):2431-2442.
- Iqbal J, Shen Y, Liu Y, et al. Genome-wide miRNA profiling of mantle cell lymphoma reveals a distinct subgroup with poor prognosis. *Blood*. 2012;119(21):4939-4948.
- Bouska A, Bi C, Lone W, et al. Adult high grade B-cell lymphoma with Burkitt lymphoma signature: genomic features and potential therapeutic targets. *Blood*. 2017;130(16):1819-1831.
- Green MR, Kihira S, Liu CL, et al. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proc Natl Acad Sci U S A*. 2015;112(10):E1116-1125.
- Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291.
- Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214-218.
- Kuilman T, Velds A, Kemper K, et al. CopywriteR: DNA copy number detection from off-target sequence data. *Genome Biol*. 2015;16(1):49.
- Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhi R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011;12(4):R41.
- Newman AM, Bratman SV, Stehr H, et al. FACTER: a practical method for the discovery of genomic rearrangements at breakpoint resolution. *Bioinformatics*. 2014;30(23):3390-3393.
- Bouska A, Bi C, Lone W, et al. Adult high-grade B-cell lymphoma with Burkitt lymphoma signature: genomic features and potential therapeutic targets. *Blood*. 2017;130(16):1819-1831.
- Akasaka T, Lossos IS, Levy R. BCL6 gene translocation in follicular lymphoma: a harbinger of eventual transformation to diffuse aggressive lymphoma. *Blood*. 2003;102(4):1443-1448.
- Kato M, Sanada M, Kato I, et al. Frequent inactivation of A20 in B-cell lymphomas. *Nature*. 2009;459(7247):712-716.
- Greiner TC, Dasgupta C, Ho VV, et al. Mutation and genomic deletion status of ataxia telangiectasia mutated (ATM) and p53 confer specific gene expression profiles in mantle cell lymphoma. *Proc Natl Acad Sci U S A*. 2006;103(7):2352-2357.
- 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(6):728-740.
- Pfeifer M, Grau M, Lenze D, et al. PTEN loss defines a PI3K/AKT pathway-dependent germinal center subtype of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A*. 2013;110(30):12420-12425.
- Jain N, Hartert K, Tadros S, et al. Targetable

- genetic alterations of TCF4 (E2-2) drive immunoglobulin expression in the activated B-cell subtype of diffuse large B-cell lymphoma. *Sci Transl Med.* 2019;11(497):eaav5599.
27. Kim D, Fiske BP, Birsoy K, et al. SHMT2 drives glioma cell survival in ischaemia but imposes a dependence on glycine clearance. *Nature.* 2015;520(7547):363-367.
 28. Rosenwald A, Wright G, Wiestner A, et al. The proliferation gene expression signature is a quantitative integrator of oncogenic events that predicts survival in mantle cell lymphoma. *Cancer Cell.* 2003;3(2):185-197.
 29. Salaverria I, Royo C, Carvajal-Cuenca A, et al. CCND2 rearrangements are the most frequent genetic events in cyclin D1(-) mantle cell lymphoma. *Blood.* 2013;121(8):1394-1402.
 30. Green MR, Vicente-Duenas C, Romero-Camarero I, et al. Transient expression of Bcl6 is sufficient for oncogenic function and induction of mature B-cell lymphoma. *Nat Commun.* 2014;5:3904.
 31. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4(1):44-57.
 32. Saeki K, Miura Y, Aki D, Kurosaki T, Yoshimura A. The B cell-specific major raft protein, Raftlin, is necessary for the integrity of lipid raft and BCR signal transduction. *EMBO J.* 2003;22(12):3015-3026.
 33. Senft D, Qi J, Ronai ZA. Ubiquitin ligases in oncogenic transformation and cancer therapy. *Nat Rev Cancer.* 2018;18(2):69-88.
 34. Monti S, Chapuy B, Takeyama K, 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.
 35. Honma K, Tsuzuki S, Nakagawa M, et al. TNFAIP3/A20 functions as a novel tumor suppressor gene in several subtypes of non-Hodgkin lymphomas. *Blood.* 2009;114(12):2467-2475.
 36. Hartmann EM, Campo E, Wright G, et al. Pathway discovery in mantle cell lymphoma by integrated analysis of high-resolution gene expression and copy number profiling. *Blood.* 2010;116(6):953-961.
 37. Choi J, Lee K, Ingvarsdottir K, et al. Loss of KLHL6 promotes diffuse large B-cell lymphoma growth and survival by stabilizing the mRNA decay factor roquin2. *Nat Cell Biol.* 2018;20(5):586-596.
 38. Meriranta L, Pasanen A, Louhimo R, et al. Deltex-1 mutations predict poor survival in diffuse large B-cell lymphoma. *Haematologica.* 2017;102(5):e195-e198.
 39. Swenson SA, Gilbreath TJ, Vahle H, et al. UBR5 HECT domain mutations identified in mantle cell lymphoma control maturation of B cells. *Blood.* 2020;136(3):299-312.
 40. Mottok A, Renne C, Seifert M, et al. Inactivating SOCS1 mutations are caused by aberrant somatic hypermutation and restricted to a subset of B-cell lymphoma entities. *Blood.* 2009;114(20):4503-4506.
 41. Liu W, Quinto I, Chen X, et al. Direct inhibition of Bruton's tyrosine kinase by IBtk, a Btk-binding protein. *Nat Immunol.* 2001;2(10):939-946.
 42. Pawar SA, Sarkar TR, Balamurugan K, et al. C/EBP δ targets cyclin D1 for proteasome-mediated degradation via induction of CDC27/APC3 expression. *Proc Natl Acad Sci U S A.* 2010;107(20):9210-9215.
 43. Krysiak K, Gomez F, White BS, et al. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. *Blood.* 2017;129(4):473-483.
 44. Gloury R, Zotos D, Zuidschewoude M, et al. Dynamic changes in Id3 and E-protein activity orchestrate germinal center and plasma cell development. *J Exp Med.* 2016;213(6):1095-1111.
 45. Okosun J, Wolfson RL, Wang J, et al. Recurrent mTORC1-activating RRAGC mutations in follicular lymphoma. *Nat Genet.* 2016;48(2):183-188.
 46. Wang F, Gatica D, Ying ZX, et al. Follicular lymphoma-associated mutations in vacuolar ATPase ATP6V1B2 activate autophagic flux and mTOR. *J Clin Invest.* 2019;130:1626-1640.
 47. Hu J, Li L, Chen H, et al. MiR-361-3p regulates ERK1/2-induced EMT via DUSP2 mRNA degradation in pancreatic ductal adenocarcinoma. *Cell Death Dis.* 2018;9(8):807.
 48. Lu D, Liu L, Ji X, et al. The phosphatase DUSP2 controls the activity of the transcription activator STAT3 and regulates TH17 differentiation. *Nat Immunol.* 2015;16(12):1263-1273.
 49. Wright GW, Huang DW, Phelan JD, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell.* 2020;37(4):551-568.
 50. Lacy SE, Barrans SL, Beer PA, et al. Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. *Blood.* 2020;135(20):1759-1771.
 51. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med.* 2006;354(23):2419-2430.
 52. Green MR. Chromatin modifying gene mutations in follicular lymphoma. *Blood.* 2018;131(6):595-604.
 53. Balinas-Gavira C, Rodriguez MI, Andrades A, et al. Frequent mutations in the amino-terminal domain of BCL7A impair its tumor suppressor role in DLBCL. *Leukemia.* 2020;34(10):2722-2735.
 54. Yusufova N, Kloetgen A, Teater M, et al. Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. *Nature.* 2021;589(7841):299-305.
 55. Ashworth A, Lord CJ, Reis-Filho JS. Genetic interactions in cancer progression and treatment. *Cell.* 2011;145(1):30-38.
 56. Beguelin W, Popovic R, Teater M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell.* 2013;23(5):677-692.
 57. Garcia-Ramirez I, Tadros S, Gonzalez-Herrero I, et al. Crebbp loss cooperates with Bcl2 over-expression to promote lymphoma in mice. *Blood.* 2017;129(19):2645-2656.
 58. Zhang J, Vlasevska S, Wells VA, et al. The Crebbp acetyltransferase is a haploinsufficient tumor suppressor in B cell lymphoma. *Cancer Discov.* 2017;7(3):322-337.
 59. Jiang Y, Ortega-Molina A, Geng H, et al. CREBBP inactivation promotes the development of HDAC3-dependent lymphomas. *Cancer Discov.* 2017;7(1):38-53.
 60. Zhang J, Dominguez-Sola D, Hussein S, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med.* 2015;21(10):1190-1198.
 61. Ortega-Molina A, Boss IW, Canela A, 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.
 62. Morin RD, Mendez-Lago M, Mungall AJ, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature.* 2011;476(7360):298-303.
 63. Sermer D, Pasqualucci L, Wendel HG, Melnick A, Younes A. Emerging epigenetic-modulating therapies in lymphoma. *Nat Rev Clin Oncol.* 2019;16(8):494-507.
 64. Shen M, Schmitt S, Buac D, Dou QP. Targeting the ubiquitin-proteasome system for cancer therapy. *Expert Opin Ther Targets.* 2013;17(9):1091-1108.
 65. Merino D, Kelly GL, Lessene G, Wei AH, Roberts AW, Strasser A. BH3-mimetic drugs: blazing the trail for new cancer medicines. *Cancer Cell.* 2018;34(6):879-891.
 66. Roschewski M, Staudt LM, Wilson WH. Diffuse large B-cell lymphoma-treatment approaches in the molecular era. *Nat Rev Clin Oncol.* 2014;11(1):12-23.