

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

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

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

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

©2022 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2020.274258

Received: October 16, 2020.

Accepted: March 15, 2021.

Pre-published: April 1, 2021.

Correspondence: MICHAEL R. GREEN - mgreen5@mdanderson.org

---

## **SUPPLEMENTARY MATERIAL**

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

Man Chun John Ma<sup>\*</sup>, Saber Tadros<sup>\*</sup>, Alyssa Bouska, Tayla Heavican, Haopeng Yang, Qing Deng, Dalia Moore, Ariz Akhter, Keenan Hartert, Neeraj Jain, Jordan Showell, Sreejoyee Ghosh, Lesley Street, Marta Davidson, Christopher Carey, Joshua Tobin, Deepak Perumal, Julie M. Vose, Matthew A. Lunning, Aliyah R. Sohani, Benjamin J. Chen, Shannon Buckley, Loretta J. Nastoupil, R. Eric Davis, Jason R. Westin, Nathan H. Fowler, Samir Parekh, Maher Gandhi, Sattva Neelapu, Douglas Stewart, Kapil Bhalla, Javeed Iqbal, Timothy Greiner, Scott J. Rodig, Adnan Mansoor, Michael R. Green<sup>\*</sup>

\*Correspondence: [mgreen5@mdanderson.org](mailto:mgreen5@mdanderson.org)

## CONTENTS

### Supplementary Methods

Figure S1: Overview of approach

Figure S2: Validation of CopyWriteR DNA copy number calls.

Figure S3: Spectrum of mutational burden and copy number aberrant genome

Figure S4: GISTIC and integrative analysis

Figure S5: Comparison of DNA copy number alteration frequencies between NGS- and SNP microarray-based copy number analysis

Figure S6: Lollipop plots for select genes

Figure S7: Molecular classification of UNMC tumors

### Supplementary References

### Supplementary Tables

Table S1: Sample information

Table S2: LymphoSeq target genes

Table S3: Coding SNVs and InDels

Table S4: MutSig2CV results

Table S5: Gene-level frequency of mutations in AID motifs

Table S6: Frequency of AID mutations by histologic subtype

Table S7: Translocations of MYC, BCL2 and BCL6

Table S8: Genes in GISTIC2 copy gain peaks

Table S9: Genes in GISTIC2 copy loss peaks

Table S10: Integrative analysis of copy gains

Table S11: Integrative analysis of copy losses

Table S12: DAVID hypergeometric enrichment analysis

Table S13: Fisher tests of disease association

Table S14: Fisher tests of mutual exclusivity and co-occurrence

## SUPPLEMENTARY METHODS

### Target Selection for LymphoSeq

A catalogue of somatic mutations in 670 B-cell malignancies was compiled from previous high-throughput sequencing studies with sufficient quality and variant reporting, and with germline controls, that were available during the design phase in 2014. These included 53 Mantle cell lymphoma (MCL)(1-3), 82 Burkitt's lymphoma (BL)(4, 5), 155 Diffuse large B-cell lymphoma (DLBCL)(6-9), 65 Follicular lymphoma (FL)(6, 10-13), 214 Chronic lymphocytic leukemia (CLL)(14-17), 36 Marginal zone lymphoma(18, 19), 38 Multiple myeloma(20) (MM), 10 Waldenstrom's macroglobulinemia (WM)(21), 9 Primary mediastinal large B-cell lymphoma(22) (PMBCL), 8 Primary central nervous system lymphoma(23) (PCNSL). Genes that were mutated in >2% of any individual subtype (with WM, PMBCL and PCNSL being grouped together as rare lymphomas), and that were not previously identified as genes with recurrent false-positive mutation calls(10) (MUC4, RPTN, PRH1, DSPP, C9ORF96, NBPF1, MUC6, MAML2, MAN1B1, C9ORF84, PDE4DIP, PTPRN2, GPC1, TBP, PKP2, KIAA1683, TTC7B, OVOS2), were selected for targeting by hybrid capture. This resulted in the coding and untranslated regions of 380 genes (Table S2) totaling 3,232,925 bp.

### Detailed Next Generation Sequencing Methods

Library Preparation. Next generation sequencing libraries were prepared from 100 – 1000ng of genomic DNA (gDNA). High molecular weight gDNA was sheared by sonication in a Covaris S220 or M220 instrument (Covaris Inc.) to obtain an average molecular weight of 150-200bp. Sheared DNA was cleaned up with Ampure Beads (Beckman Coulter), and libraries prepared using Hyper Prep kits (KAPA Biosystems), according to the manufacturer's protocol except for an overnight ligation of adapters at 16°C. A maximum of 8 cycles of PCR was performed during library preparation (mean = 6 cycles). TruSeq adapters (Bioo Scientific) were utilized at the recommended ratio to input DNA. Library quality was assessed using TapeStation High-Sensitivity DNA 1000 (Agilent) and quantified by Qubit (Life Technologies).

Hybrid capture and sequencing. Libraries were 12-plexed in equal quantities and a 1µg of pooled libraries were enriched by hybrid capture using a Nimblegen SeqCap EZ Custom reagent (Roche), according to the manufacturer's protocol. Each capture pool was sequenced on a single lane of a HiSeq 2500 in high output mode using 2 x 100bp at the Hudson Alpha Institute for Biotechnology. A total of 2.06 Tera-base-pairs of sequencing data was produced. The average

on-target rate and sequencing depth of the 685 samples presented in this study were 81.3% (range, 71.7% – 92.5%) and 624X (range, 101X – 1785X), respectively, determined as per below. Variant calling. Raw FASTQ files were assessed for quality using FASTQC. Samples with high quality metrics were run through our in-house pipeline. FASTQ files were (i) aligned to the human genome (hg19) using BWA-Mem(24), (ii) deduplicated using Picard MarkDuplicates, (iii) realigned around InDels using GATK(25), and (iv) recalibrated by base score using GATK(25). On-target rate and coverage over the targeted region were calculated by Picard CalculateHSMetrics. Only samples with >100X average coverage were utilized. If additional sequencing was required to obtain sufficient coverage, bam files from the same sample were merged using BamTools Mappings Merger following alignment. Variants were called by GATK Unified Genotyper and VarScan2. Only variants called by both tools were retained, which we have previously shown to provide a sensitivity of 96.7% and specificity of 92.9%(11). All variants were annotated using SeattleSeq(26).

AID mutations. Mutations were defined as products of AID activity if the wild-type allele was a cytosine within the context of a *WRCY* motif, as previously described(11). These data are reported in Table S5 as the total number of mutations in the coding sequence of each gene that fit this criteria.

Filtering of repetitive regions and potential germline variants. To avoid mapping artifacts in repetitive regions, all variants within RepeatMasker or tandemRepeat annotated regions were filtered from the dataset. All genes selected for capture had prior evidence of somatic mutation from whole exome or whole genome sequencing studies (see 'Target Selection for Hybrid Capture' section). In addition, all variants in dbSNP build 32 were removed, and all mutation hotspots (defined below) were manually searched in the ExAC browser(27) containing whole exome sequencing of 60,706 individuals. Genes with mutation hotspots present in  $\geq 1$  healthy individual were filtered from the dataset to remove the chance of these variants corresponding to a germline polymorphism. Mutation frequencies of significantly mutated genes (MutSigCV, below) were compared between fresh/frozen and FFPE tumors for BL and MCL using a Fisher exact test. No significant differences were detected, therefore filtering was not performed using this criteria. MutSig2CV. MAF files containing filtered variants were analyzed by MutSig2CV(28). This analysis was performed for the data set as a whole, and for the four major subtypes individually (DLBCL, BL, FL and MCL). Genes that reached a significance of  $Q < 0.25$  in any one of these analyses were included, as well as genes that did not reach this threshold but have a well-defined role in lymphoma and are targets of somatic hypermutation (eg. *BCL2* and *BCL6*; Table S5).

DNA copy number and GISTIC2 analysis. Genome-wide DNA copy number profiles were determined from mapped, realigned, sorted, deduplicated bam files using CopyWriteR(29). For the generation of CopywriteR library folder we utilized 100kb bins in hg19 assembly with contigs with a prefix of 'chr' [preCopywriteR (bin.size=100000, ref.genome='hg19', prefix='chr')]. For the identification of CNAs we used "input vs. none" [CopywriteR(sample.control, destination.folder, reference.folder, bp.param, capture.regions.file)], where sample.control is generated by data.frame(sample=[input\_bam], control=[input\_bam]), reference.folder is the folder generated by preCopywriteR above, bp.param involves parallelism, and capture.regions.file is the BED file containing the LymphoSeq panel's capture regions. Segmented data were analyzed using GISTIC2(30) in the GenePattern(31) environment, with copy number thresholds of 0.2 and marker number of 100. Peaks were considered to be significant if they had a residual Q-value < 0.1, to account for the significance in neighboring peaks. The gene-level data from GISTIC2 analysis have been uploaded to cBioPortal. The genes within each peak were used for integrative analysis with matched gene expression profiling data (below) to define the likely drivers of each alteration. For validation, DNA copy number alterations of genes shown in Figure 3 were evaluated in previously reported SNP microarray data from 85 BL, 694 DLBCL, 404 FL and 206 MCL. For BL(32), DLBCL(33, 34) and FL(35), the combined analysis of these datasets have been previously reported. For MCL, high-resolution (>200,000 markers) Affymetrix SNP microarray or Agilent CGH array DNA number data were downloaded for 206 tumors from the gene expression omnibus (GSE12906, GSE18820, GSE42854, GSE46969) and probe-level data segmented using the cbs tool in GenePattern(31). Gene level DNA copy number was calculated from segmented data using GISTIC2, and reported as copy loss or gain if absolute DNA copy number was <1.8 or >2.2, respectively, in line with the thresholds used for GISTIC analysis of the NGS cohort. Heatmaps were generated using the GENE-E tool.

Identification of B-NHL hallmarks. A gene list was created that included all genes that were significantly mutated, as determined by MutSig2CV, and/or targeted by DNA copy number gain or loss, as determined by GISTIC2 and integrative analysis. This gene list was interrogated using the 'Functional Annotation Clustering' function of DAVID(36). This identified 41 clusters of significantly enriched pathways, which were grouped together according to functional relatedness.

Assessment of co-association and clustering. Genetic alterations that were present in  $\geq 5\%$  of any major histologic subtype (BL, DLBCL, FL, MCL) were assessed for mutual co-association or exclusion using a Fisher exact test corrected for multiple hypothesis testing by the Benjamini-

Hochberg method (Table S14). Co-association q-values were transformed by log<sub>10</sub> and directionality (co-association assigned as positive values, exclusion as negative values) and then a matrix of transformed values was clustered in GENE-E using Pearson's correlation coefficient with complete linkage. Genes within the clusters reported in this study were compared to those from other studies (Lacy, Chapuy, Wright) using a Fisher exact test. DNA copy number alterations were omitted for this comparison due to absence (Lacy) or differences in the nomenclature used (Wright).

### **Gene Expression Analysis**

Affymetrix U133 plus 2.0 gene expression microarray data were available for 284 tumors from the University of Nebraska Medical Center (Table S1)(37-39). Raw cel files were RMA normalized using the ExpressionFileCreator module of GenePattern(31), and quality checked for batch effects by unsupervised clustering of the 3,000 most variably expressed genes across the dataset. Integrative analysis of DNA copy number alterations. Integrative analysis was performed using tumor subsets with  $\geq 3$  tumors with a given copy number alterations. If tumor subtypes did not harbor the copy number alteration of interest, they were excluded so as not to confound the analysis with subtype-specific gene expression differences. For each DNA copy number alteration, the expression of genes within the peak was compared between diploid and gain/loss tumors using a Student's T-test corrected for multiple hypothesis testing with a Benjamini-Hochberg correction. Genes with a directionality of change matching the DNA copy number alteration type (i.e. increased expression with DNA copy gain, decreased expression with DNA copy loss) and a Q-value  $< 0.25$  were considered significant.

Molecular Burkitt lymphoma (mBL) classification. For mBL classification, we built a Bayesian classifier using previously published dataset(40). The model was created on the discovery set from the original paper and evaluated on the test set. Classifications based upon our model had minor differences to the originally reported classification, primarily in reducing the number of unclassified tumors, and improved upon the significance in overall survival difference between mBL and non-mBL subgroups (Figure S6). We therefore applied this model to the UNMC dataset, including DLBCL, BL and HGBL-NOS tumors (n=159).

Cell of origin subtyping. For COO subtyping, we utilized a previously described 140 gene Bayesian classifier(41). This classification was applied to non-mBL tumors from the molecular Burkitt classifier, and overall survival assessed in the subgroups as a sanity check (Figure S6).

**Data Sharing.** All gene expression data has been consolidated under a single Gene Expression Omnibus Accession (GSE132929). Mutation, DNA copy number alteration, translocation and gene expression data have been uploaded to cBioPortal for easy access ([https://www.cbioportal.org/study/summary?id=mbn\\_mdacc\\_2013](https://www.cbioportal.org/study/summary?id=mbn_mdacc_2013)). Raw next generation sequencing data will be made available to investigators at academic or not-for-profit research institutions upon reasonable request.

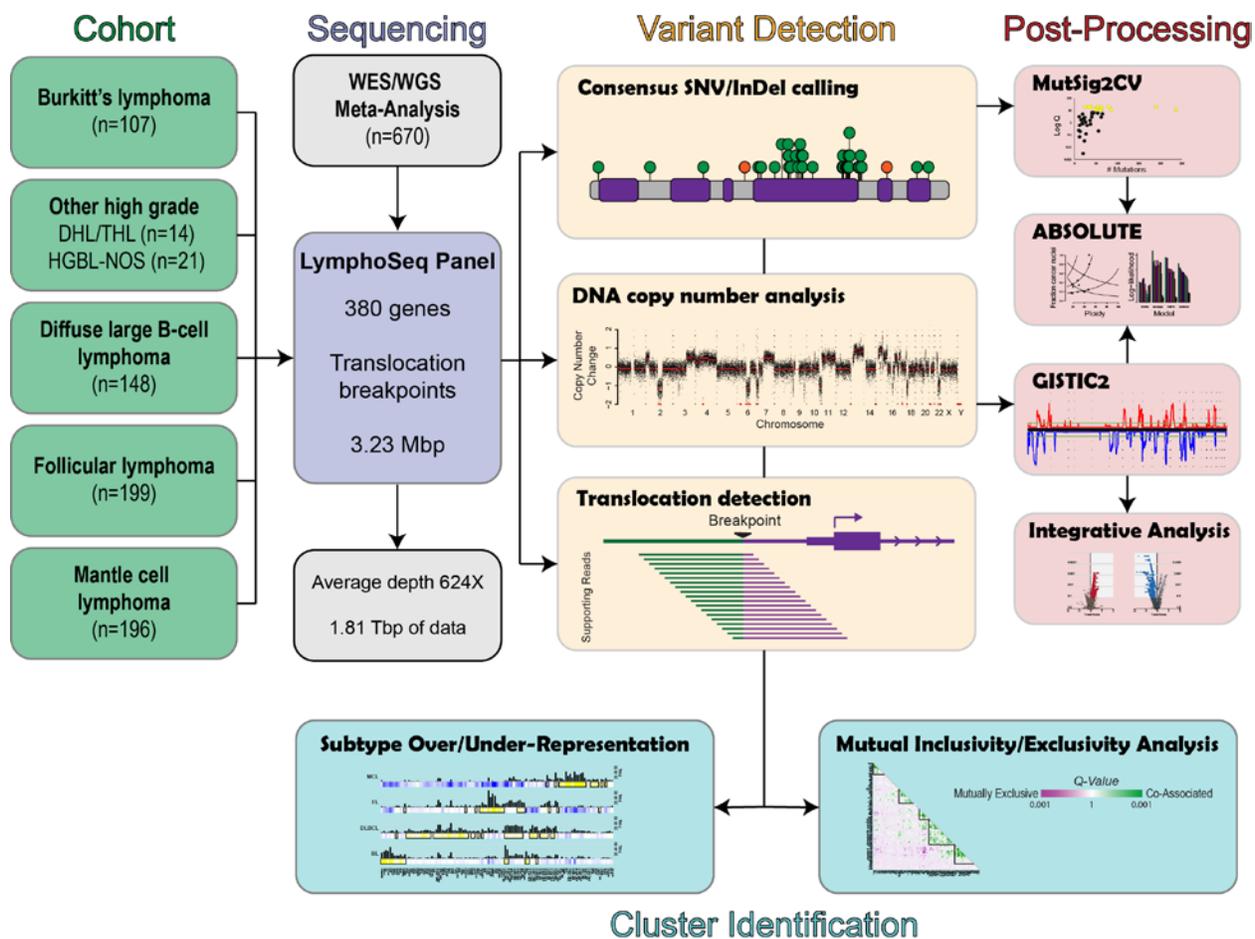
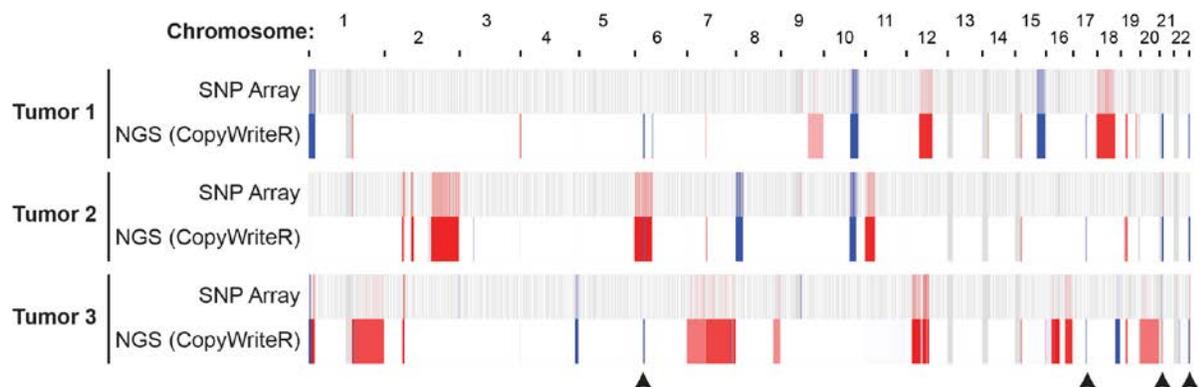
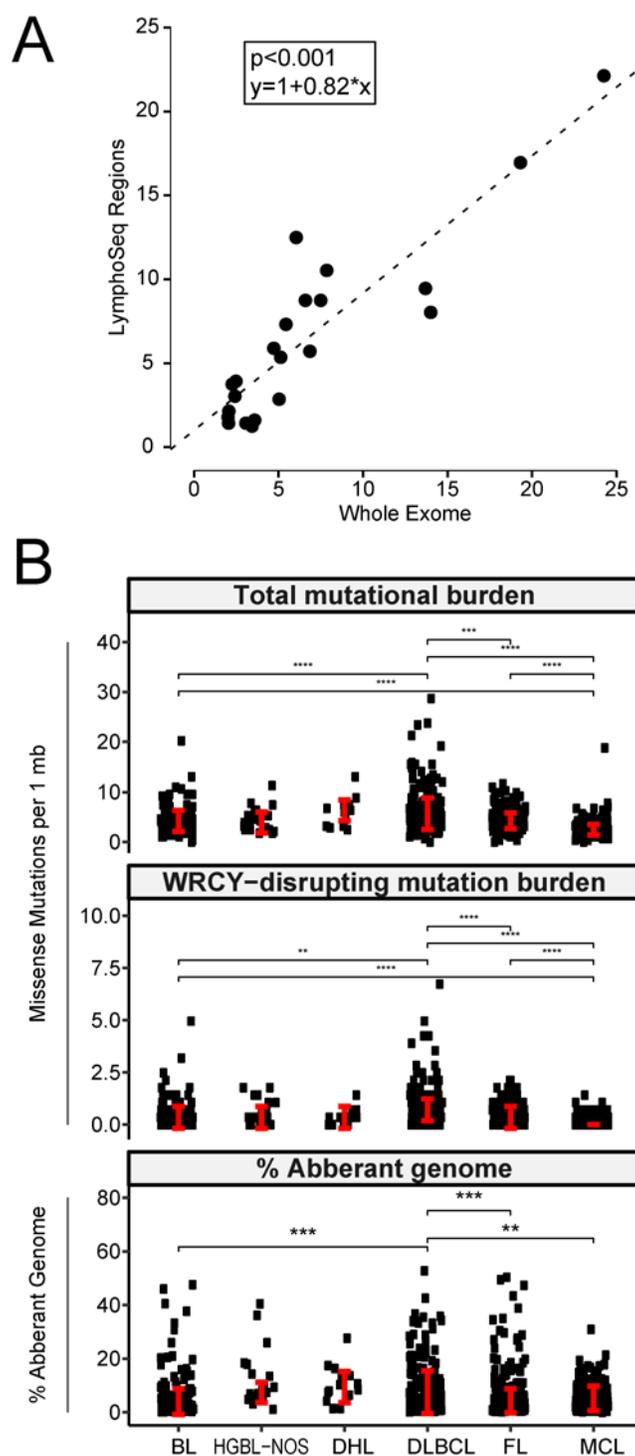


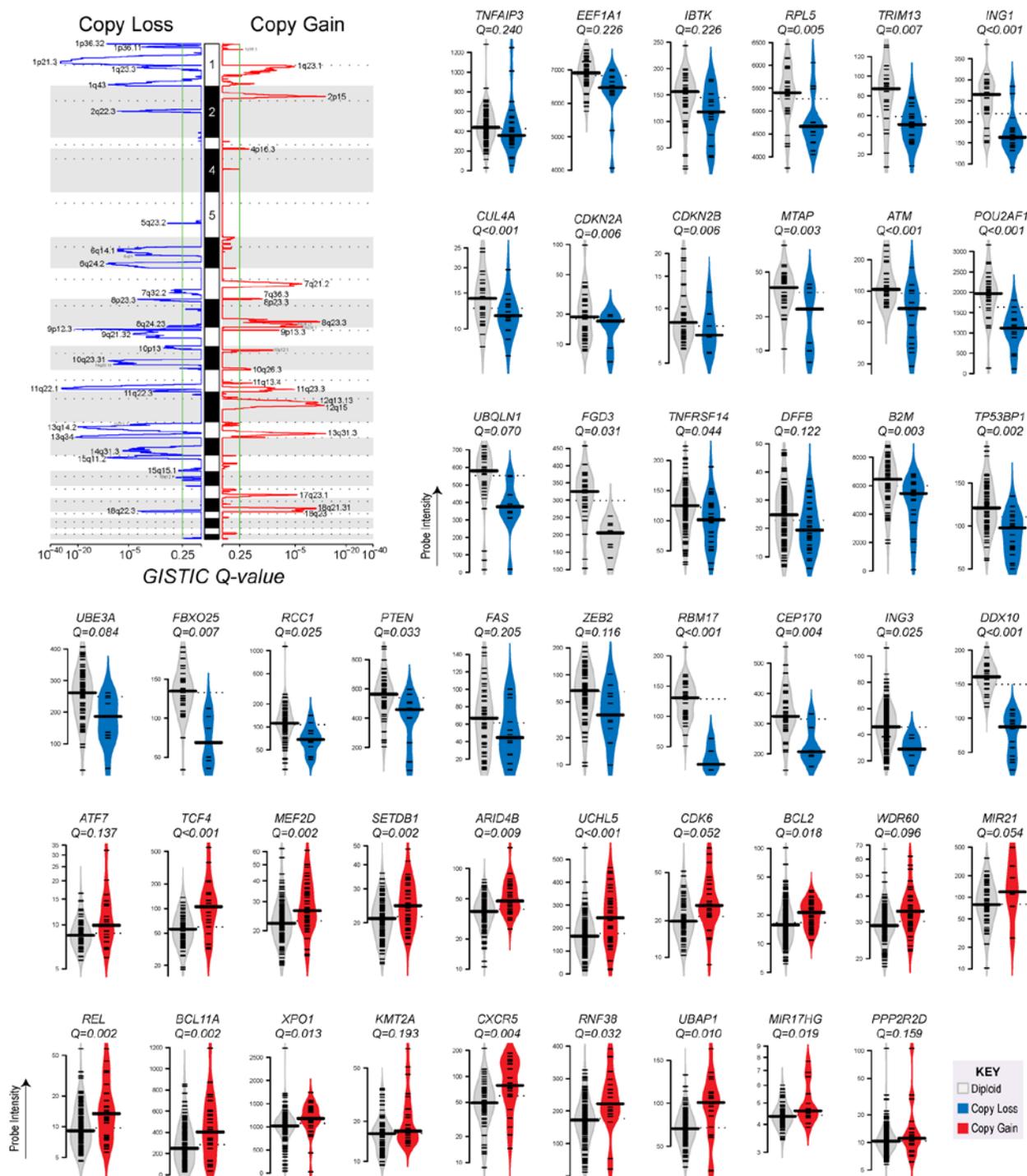
Figure S1: A schematic overview of the study design.



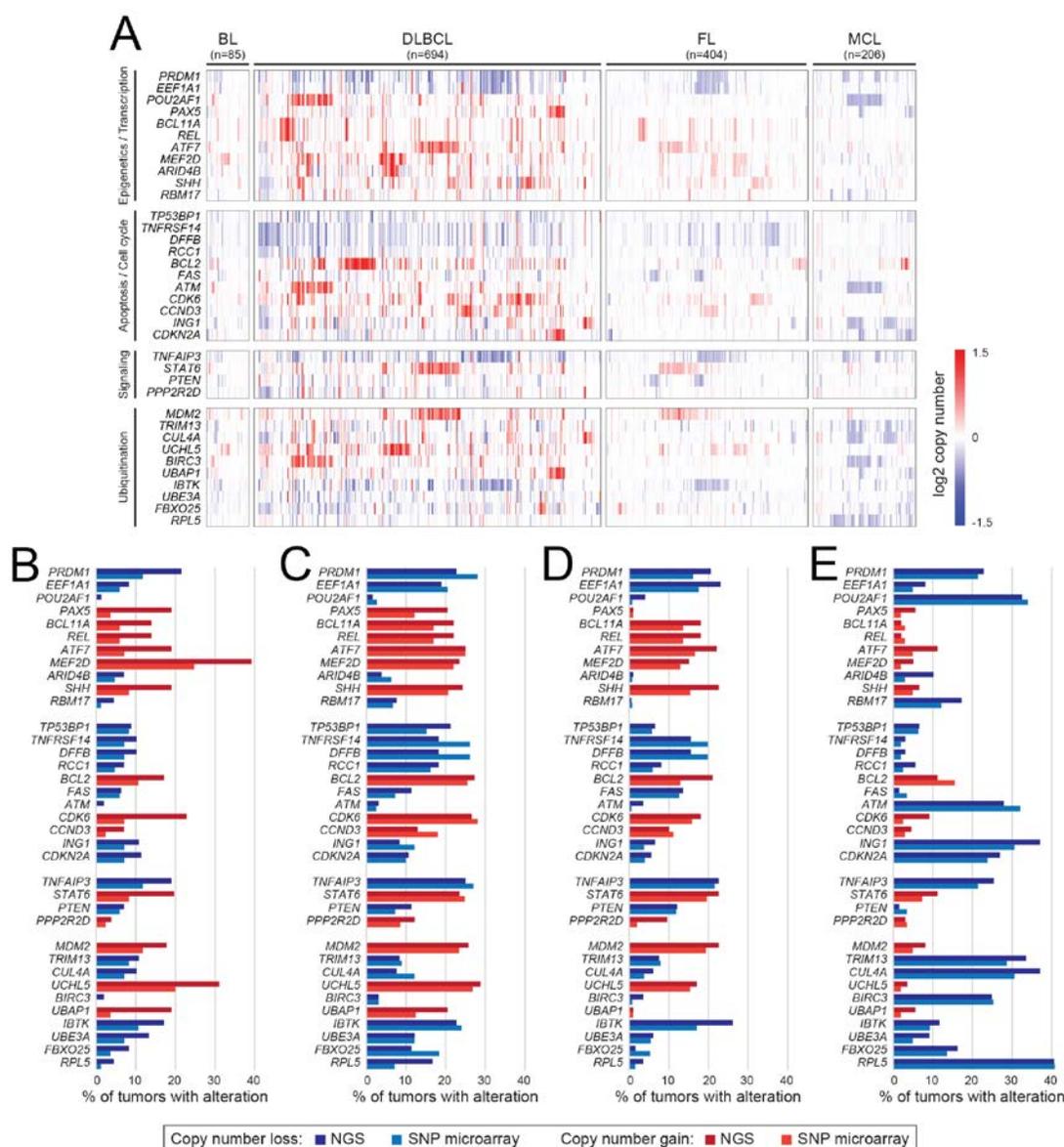
**Figure S2: Validation of CopyWriteR DNA copy number calls.** DNA copy number calculated from off-target NGS reads by CopyWriteR was compared to that calculated using Affymetrix 250K Nsp SNP microarrays for 3 FL tumors with available data. The copy number calls were highly concordant with the exception of 4 repetitive regions (arrows) that were identified as copy number losses by CopyWriteR, but not in the SNP microarray data. These regions were masked for the GISTIC2 analysis.



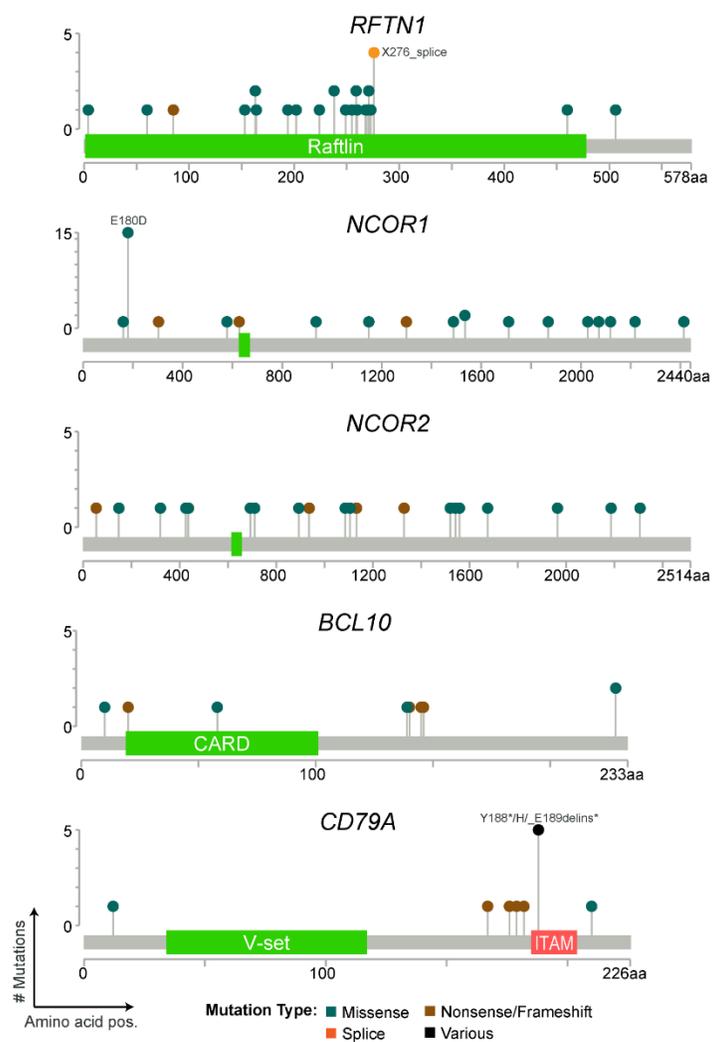
**Figure S3: Spectrum of mutational burden and copy number aberrant genome. A)** Mutational burden was calculated for the whole exome and from LymphoSeq targeted regions using previously published whole-exome data(11), which showed a significant correlation. **B)** The mutational burden, AID-driven mutational burden, and percentage of copy number aberrant genome are shown for each tumor grouped by diagnosis. (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ )



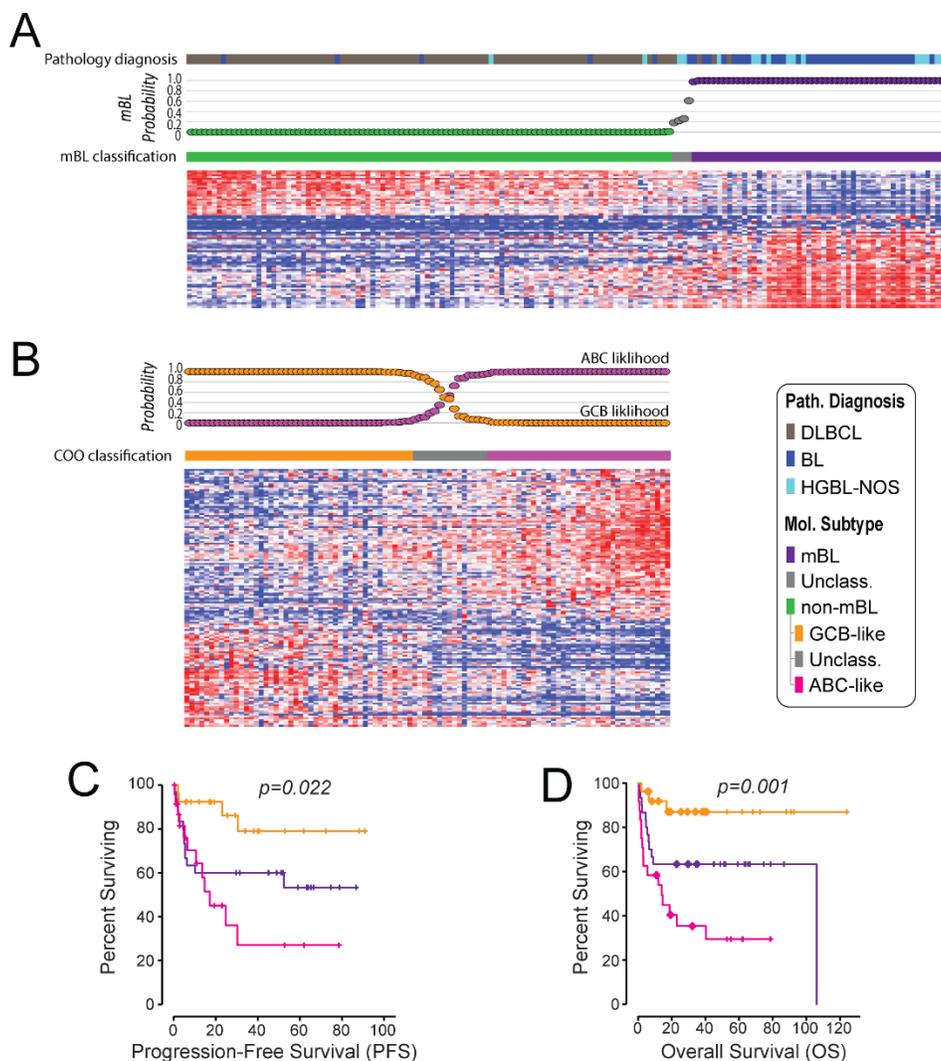
**Figure S4: GISTIC and integrative analysis.** A standard GISTIC plot for figure 2A is shown on the top left. Individual violin plots for the integrative analysis of genes appearing in figure 3 are also provided, colored by whether the gene is targeted by copy loss (blue) or copy gain (red).



**Figure S5: Comparison of DNA copy number alteration frequencies between NGS- and SNP microarray-based copy number analysis. A)** A heatmap shows the absolute DNA copy number of genes shown in Figure 3, as determined by SNP microarray analysis of 85 BL, 694 DLBCL, 404 FL and 206 MCL tumors. These same genes show frequent DNA copy number alterations in this independent cohort, as measured by SNP microarray. **B-E)** The frequencies of DNA copy number alterations for genes shown in Figure 3 are compared between NGS-based and SNP microarray-based DNA copy number analysis of independent cohorts. The frequencies significantly correlate in BL (B; Pearson's correlation p-value < 0.0001,  $r = 0.8635$ ), DLBCL (C; Pearson's correlation p-value < 0.0001,  $r = 0.8519$ ), FL (D; Pearson's correlation p-value < 0.0001,  $r = 0.9233$ ), and MCL (E; Pearson's correlation p-value < 0.0001,  $r = 0.9683$ ), providing validation of the NGS-based approach for DNA copy number analysis.



**Figure S6: Lollipop plots for select genes.**



**Figure S7: Molecular classification of UNMC tumors. A)** Application of the molecular Burkitt lymphoma (mBL) Bayesian classifier to the tumors from the UNMC cohort. **B)** Application of the previously described cell of origin Bayesian classifier(41) to the UNMC cohort. **C)** Progression-free survival of CHOP-treated patients according to their tumor's molecular classification. **D)** Overall survival of CHOP-treated patients according to their tumor's molecular classification.

## Supplementary References

1. Khodadoust MS, Olsson N, Wagar LE, *et al.* Antigen presentation profiling reveals recognition of lymphoma immunoglobulin neoantigens. *Nature*. 2017;543(7647):723-7.
2. 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):18250-5.
3. 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):2988-96.
4. Schmitz R, Young RM, Ceribelli M, *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature*. 2012;490(7418):116-20.
5. Love C, Sun Z, Jima D, *et al.* The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet*. 2012;44(12):1321-5.
6. 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.
7. Morin RD, Mungall K, Pleasance E, *et al.* Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood*. 2013;122(7):1256-65.
8. Pasqualucci L, Trifonov V, Fabbri G, *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet*. 2011;43(9):830-7.
9. Lohr JG, Stojanov P, Lawrence MS, *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A*. 2012;109(10):3879-84.
10. Green MR, Gentles AJ, Nair RV, *et al.* Hierarchy in somatic mutations arising during genomic evolution and progression of follicular lymphoma. *Blood*. 2013;121(9):1604-11.
11. 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):E1116-25.
12. Okosun J, Bodor C, Wang J, *et al.* Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. *Nat Genet*. 2014;46(2):176-81.
13. Pasqualucci L, Khiabanian H, Fangazio M, *et al.* Genetics of follicular lymphoma transformation. *Cell Rep*. 2014;6(1):130-40.
14. Landau DA, Carter SL, Stojanov P, *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*. 2013;152(4):714-26.
15. Puente XS, Pinyol M, Quesada V, *et al.* Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2011;475(7354):101-5.
16. Quesada V, Conde L, Villamor N, *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nat Genet*. 2012;44(1):47-52.
17. Wang L, Lawrence MS, Wan Y, *et al.* SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N Engl J Med*. 2011;365(26):2497-506.
18. Kiel MJ, Velusamy T, Betz BL, *et al.* Whole-genome sequencing identifies recurrent somatic NOTCH2 mutations in splenic marginal zone lymphoma. *J Exp Med*. 2012;209(9):1553-65.
19. Rossi D, Trifonov V, Fangazio M, *et al.* The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. *J Exp Med*. 2012;209(9):1537-51.
20. Chapman MA, Lawrence MS, Keats JJ, *et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011;471(7339):467-72.
21. Treon SP, Xu L, Yang G, *et al.* MYD88 L265P somatic mutation in Waldenstrom's macroglobulinemia. *N Engl J Med*. 2012;367(9):826-33.

22. Gunawardana J, Chan FC, Telenius A, *et al.* Recurrent somatic mutations of PTPN1 in primary mediastinal B cell lymphoma and Hodgkin lymphoma. *Nat Genet.* 2014;46(4):329-35.
23. Vater I, Montesinos-Rongen M, Schlesner M, *et al.* The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing. *Leukemia.* 2015;29(3):677-85.
24. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv.* 2013;1303.3997.
25. McKenna A, Hanna M, Banks E, *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297-303.
26. Ng SB, Turner EH, Robertson PD, *et al.* Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009;461(7261):272-6.
27. Lek M, Karczewski KJ, Minikel EV, *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-91.
28. Lawrence MS, Stojanov P, Polak P, *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature.* 2013;499(7457):214-8.
29. Kuilman T, Velds A, Kemper K, *et al.* CopywriteR: DNA copy number detection from off-target sequence data. *Genome Biol.* 2015;16:49.
30. 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.
31. Reich M, Liefeld T, Gould J, Lerner J, Tamayo P, Mesirov JP. GenePattern 2.0. *Nat Genet.* 2006;38(5):500-1.
32. 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):1819-31.
33. 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.
34. Li M, Chiang YL, Lyssiotis CA, *et al.* Non-oncogene Addiction to SIRT3 Plays a Critical Role in Lymphomagenesis. *Cancer Cell.* 2019;35(6):916-31 e9.
35. Parsa S, Ortega-Molina A, Ying H-Y, *et al.* The serine hydroxymethyltransferase-2 (SHMT2) initiates lymphoma development through epigenetic tumor suppressor silencing. *Nature Cancer.* 2020;1(6):653-64.
36. Huang DW, Sherman BT, Tan Q, *et al.* DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res.* 2007;35(Web Server issue):W169-75.
37. Lenz G, Wright G, Dave SS, *et al.* Stromal gene signatures in large-B-cell lymphomas. *N Engl J Med.* 2008;359(22):2313-23.
38. Dave SS, Fu K, Wright GW, *et al.* Molecular diagnosis of Burkitt's lymphoma. *N Engl J Med.* 2006;354(23):2431-42.
39. 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):4939-48.
40. 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):2419-30.
41. Wright G, Tan B, Rosenwald A, Hurt EH, Wiestner A, Staudt LM. A gene expression-based method to diagnose clinically distinct subgroups of diffuse large B cell lymphoma. *Proc Natl Acad Sci U S A.* 2003;100(17):9991-6.