Supplementary Material

Mato *et al*

Supplementary Methods

Immunohistochemistry

The immunophenotype was studied using standard immunohistochemistry (IHC) protocols on an automated platform (Ventana BenchmarkUltra, Roche, Basel, Switzerland). Expression of MYC, BCL2 (1), BCL6 and MUM1 (2) was considered positive when >40% ≥70%, ≥30% or ≥60% of the tumor cells expressed the proteins, respectively. The cut offs were established according to previous reports of the literature.

Next generation sequencing (NGS) approach

DNA was extracted using Qiagen extraction kits (Qiagen, Hilden, Germany). For the study of structural variants (SV) and mutations, a custom capture panel (SureSelectXT, Agilent Technologies, Santa Clara, CA) was used for the analysis of 26 tumor samples (21 FFPE and 5 frozen tissues). This panel interrogates wide regions of the genome that cover all the chromosomal breakpoints involved in the 11 main B-cell lymphoma translocations and mutations on 168 genes covering the 114 genes necessary for the *LymphGen* prediction (3). The interrogation for the SV analysis was based on previously described design (4) with modifications covering all *MYC* breaks described in pediatric BL (5).

Pipeline of analysis and variant filtering

Quality assessment of the raw FASTQ-paired datasets (R1 and R2) was performed using the FastQC tool (https://www.bioinformatics.babraham.ac.uk/projects/fastqc). Sequence trimming was subsequently conducted using Trimmomatic (version 0.40), adhering to the default parameters (LEADING: 3, TRAILING: 3, SLIDINGWINDOW: 4:15) (6). Trimmed reads were then mapped to the human reference genome (GRCh37/hg19) using the Burrows-Wheeler Aligner-MEM (BWA-MEM) algorithm (7), considering soft clips (-Y) for Structural Variants (SVs) analysis. Consequent processing steps included in the Picard toolkit (version 2.24.0, https://broadinstitute.github.io/picard), specifically the MarkDuplicates, and CollectInsertSizeMetrics utilities to flag optical/PCR duplicates and compile insert size metrics, respectively. Low-quality, duplicated, and hard-clipped reads were removed using Samtools (version 1.9).

Tumor-only SV calling

SV were extracted using Delly (version 1.1.6) (8), SvABA (version 1.2.0) (9), and GRIDSS (version 2.13.2) (10). On-target variants, which attained a ‘PASS’ quality status from a minimum of two among the three algorithms within a 300 bp window and exhibited high quality in at least one (DELLY, MAPQ=60 or SRMAPQ=60; SVABA, MAPQ=60; GRIDSS, BQM=60) were selected for downstream analysis. In addition, IgCaller (version 1.2) (11) was employed for enhanced detection of variants in immunoglobulin loci. The analysis also involved a detailed visual inspection of SV breakpoints using Integrative Genomics Viewer (IGV, Broad Institute, version 2.15.1) (12). This contributed to the elimination of artifacts related to the source material (13), and to remove variants located in homologous or repetitive sequences.

Tumor-only variant calling

Preprocessing steps were conducted to hinder over-, or underestimation of quality scores originated by systematic bias (14). For this purpose, Base Quality Score Recalibration (BQSR, Genome Analysis Toolkit (GATK), version 4.0.3) (15), and ApplyBQSR (GATK) were applied. Afterwards, mutations were called using Mutect2 (GATK) (15), VarDictJava (version 1.8.3) (16), Pisces (version 5.3.0.0) (17), VarScan2 (version 2.3.9) (18), and Lofreq (version 2.1.5) (19). Normalization of the data was performed using bcftools (version 1.10.2). FilterMutectCalls (GATK) filtered low-quality variants identified by Mutect2. In the case of VarDictJava “-f 0.01” parameter was set when running both the VarDictJava program and the included var2vcf_valid.pl script. Preceding the implementation of Pisces, Gemini was executed for consensus sequence assembly (20). VarScan2 parameters included “-min-var-freq 0.01”, “p-value 0.05”, and “-strand-filter 1”. The thresholds set for LoFreq were “-q 15” for mapping quality, “–Q15” for base quality, and “-m 10” for read depth. Mutation annotation was conducted utilizing snpEff/snpSift (21). Variants were considered as significant if they achieved a ‘PASS’ status and were detected by at least three distinct algorithms.

In the absence of germline DNA, filtration steps excluded non-interrogated regions, focusing on non-synonymous variants not exceeding 0.1% global population frequency (as per dbSNP, 1000 Genome project, ExAC, or GnomAD). Other criteria included removing 3’UTR (except for *NOTCH1* (22)) and 5’UTR variants, those with mean depth below 20, and any classified as ‘Benign’ or ‘Likely Benign’ by COSMIC.

Functional prediction identified potential driver mutations as truncating variants, confirmed somatic variants associated with lymphoid neoplasms according to COSMIC, ‘Deleterious’ in two or more in silico predictors (SIFT, Polyphen2, MutationAssessor), or those with a CADD Phred quality score over 20. All variants were validated through visual inspection using IGV.

Accurate detection of low-frequency variants is challenged by inherent error rates of library preparation, sequencing chimeras, potential duplicated reads, and the lack of a standardized VAF cut-off (23–25). Observing this, both HG33 and HG51 exhibited skewed VAF distribution with peaks below 5%. Case HG51 presented a high-error rate (67%), and variants below the 5% threshold were excluded. Moreover, for case HG33, the availability of a paired-relapsed sample facilitated the confirmation of variants through comparative analysis (data not shown).

Copy number analysis

DNA from 35 samples were hybridized on an Oncoscan (n=30) or Cytoscan array (n=5) (ThermoFisher Scientific inc.). Gains and losses and copy number neutral-loss of heterozygosity (CNN-LOH) regions were evaluated and visually inspected using Nexus Copy Number version 9.0 (Bionano, San Diego, CA, USA). Human reference genome was GRCh37/hg19. The copy number alterations (CNA) with minimum size of 100 kb and CNN-LOH larger than 5 Mb, were considered informative. According to the literature (26), cases were considered to carry chromothripsis-like patterns when at least 7 switches between two or more CN states were observed on an individual chromosome. Published CN data from pediatric and young adult BL and DLBCL were used for comparison (27,28).

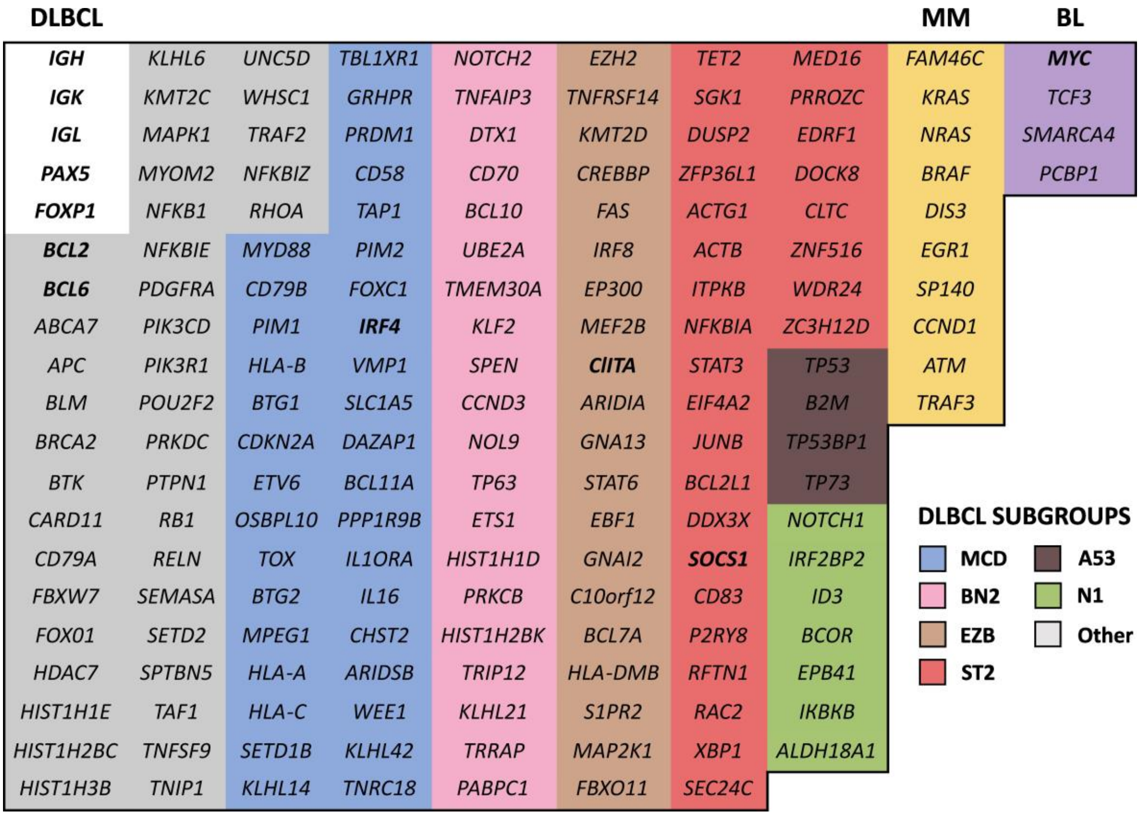
Digital gene expression profiling

RNA from FFPE and frozen tissues were extracted using Qiagen extraction kits (Qiagen, inc). Digital GEP was performed on 27 RNA samples (25 from FFPE and 2 frozen tissue) using the DLBCL90 assay (29) on the nCounter platform (NanoString Technologies, Seattle, WA) to assign COO and Dark zone signature (DZsig) status. In 5 additional cases, COO determination was performed using the Lymphoma Subtyping Test-Lymph2Cx (NanoString inc.) Data was normalized for loading and RNA integrity using standard normalization protocols in NSolver (NanoString inc.). The normalized data was then transformed log2 prior to analysis.

As previously described (30), gene expression subgroups were assigned hierarchically, with COO taking precedence over DZsig status for ABC tumors. GCB and UNC tumors that were DZsig^{pos} were assigned to the DZsig^{pos} group, whereas DZsig^{ind} and DZsig^{neg} tumors were assigned to their respective COO subgroups.

Supplementary Figures

Supplementary Figure S1. Target NGS sequencing panel covering the mutational status of 168 genes and 11 structural variants (SV) associated to B-cell lymphomagenesis. Genes interrogated for SV determination are highlighted in bold. Genes associated with the genetic diffuse large B-cell lymphoma (DLBCL) subtypes according to the *LymphGen* (3) prediction and other lymphoma entities (MM, multiple myeloma; BL, Burkitt Lymphoma) are colored differently.

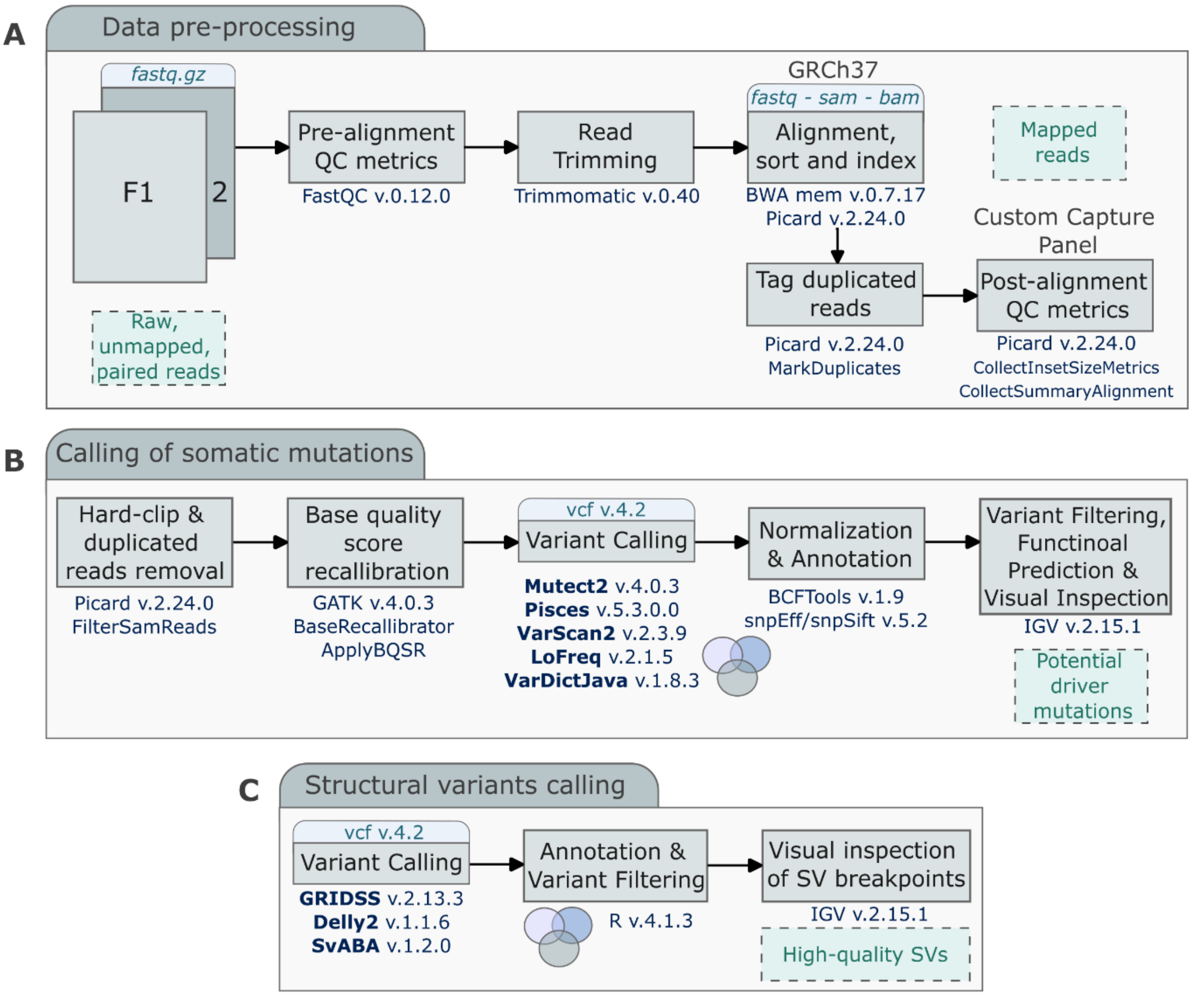


DLBCL								MM	BL
IGH	*KLHL6*	*UNC5D*	*TBL1XR1*	*NOTCH2*	*EZH2*	*TET2*	*MED16*	*FAM46C*	***MYC***
IGK	*KMT2C*	*WHSC1*	*GRHPR*	*TNFAIP3*	*TNFRSF14*	*SGK1*	*PRROZC*	*KRAS*	*TCF3*
IGL	*MAPK1*	*TRAF2*	*PRDM1*	*DTX1*	*KMT2D*	*DUSP2*	*EDRF1*	*NRAS*	*SMARCA4*
PAX5	*MYOM2*	*NFKBIZ*	*CD58*	*CD70*	*CREBBP*	*ZFP36L1*	*DOCK8*	*BRAF*	*PCBP1*
FOXP1	*NFKB1*	*RHOA*	*TAP1*	*BCL10*	*FAS*	*ACTG1*	*CLTC*	*DIS3*	
BCL2	*NFKBIE*	*MYD88*	*PIM2*	*UBE2A*	*IRF8*	*ACTB*	*ZNF516*	*EGR1*	
BCL6	*PDGFRA*	*CD79B*	*FOXC1*	*TMEM30A*	*EP300*	*ITPKB*	*WDR24*	*SP140*	
ABCA7	*PIK3CD*	*PIM1*	***IRF4***	*KLF2*	*MEF2B*	*NFKBIA*	*ZC3H12D*	*CCND1*	
APC	*PIK3R1*	*HLA-B*	*VMP1*	*SPEN*	***CIITA***	*STAT3*	*TP53*	*ATM*	
BLM	*POU2F2*	*BTG1*	*SLC1A5*	*CCND3*	*ARIDIA*	*EIF4A2*	*B2M*	*TRAF3*	
BRCA2	*PRKDC*	*CDKN2A*	*DAZAP1*	*NOL9*	*GNA13*	*JUNB*	*TP53BP1*		
BTK	*PTPN1*	*ETV6*	*BCL11A*	*TP63*	*STAT6*	*BCL2L1*	*TP73*		
CARD11	*RB1*	*OSBPL10*	*PPP1R9B*	*ETS1*	*EBF1*	*DDX3X*	*NOTCH1*		
CD79A	*RELN*	*TOX*	*IL1ORA*	*HIST1H1D*	*GNAI2*	***SOCS1***	*IRF2BP2*		
FBXW7	*SEMASA*	*BTG2*	*IL16*	*PRKCB*	*C10orf12*	*CD83*	*ID3*		
FOX01	*SETD2*	*MPEG1*	*CHST2*	*HIST1H2BK*	*BCL7A*	*P2RY8*	*BCOR*		
HDAC7	*SPTBN5*	*HLA-A*	*ARIDSB*	*TRIP12*	*HLA-DMB*	*RFTN1*	*EPB41*		
HIST1H1E	*TAF1*	*HLA-C*	*WEE1*	*KLHL21*	*S1PR2*	*RAC2*	*IKBKB*		
HIST1H2BC	*TNFSF9*	*SETD1B*	*KLHL42*	*TRRAP*	*MAP2K1*	*XBP1*	*ALDH18A1*		
HIST1H3B	*TNIP1*	*KLHL14*	*TNRC18*	*PABPC1*	*FBXO11*	*SEC24C*			

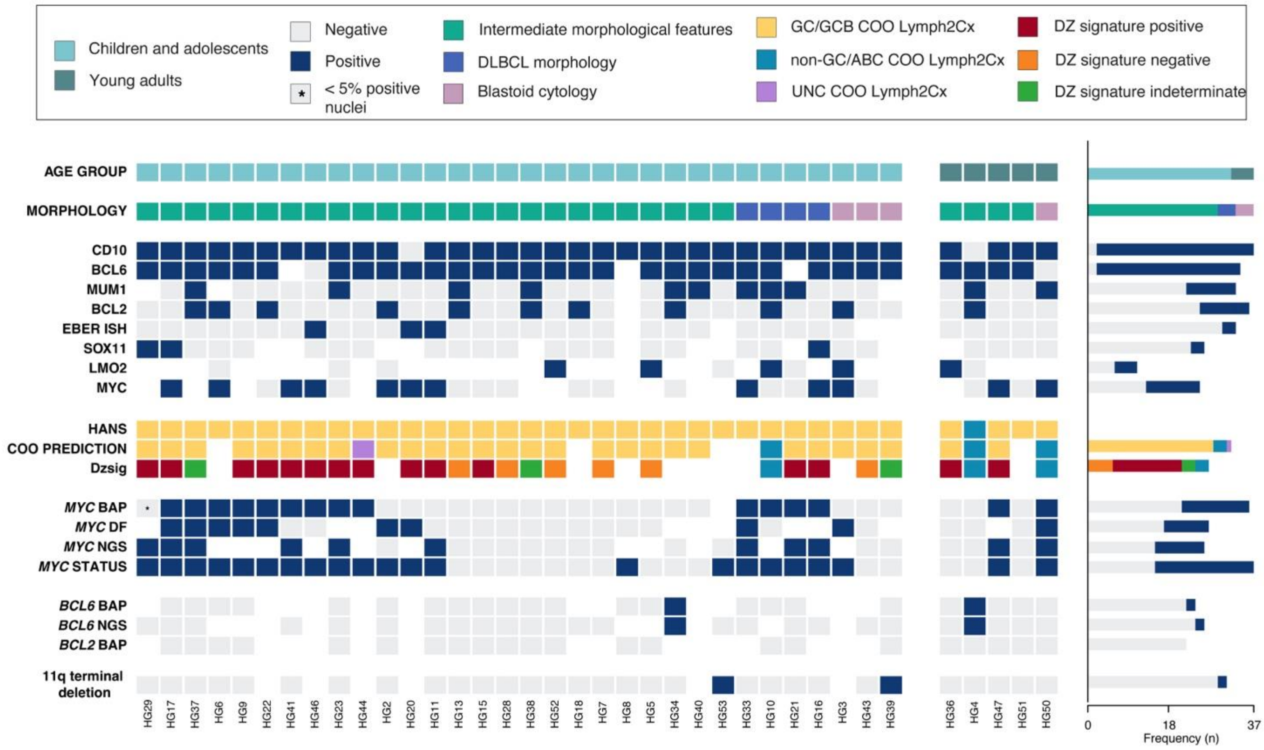
DLBCL SUBGROUPS

🟦	**MCD**	■	**A53**
■	**BN2**	🟩	**N1**
■	**EZB**	☐	**Other**
🟥	**ST2**		

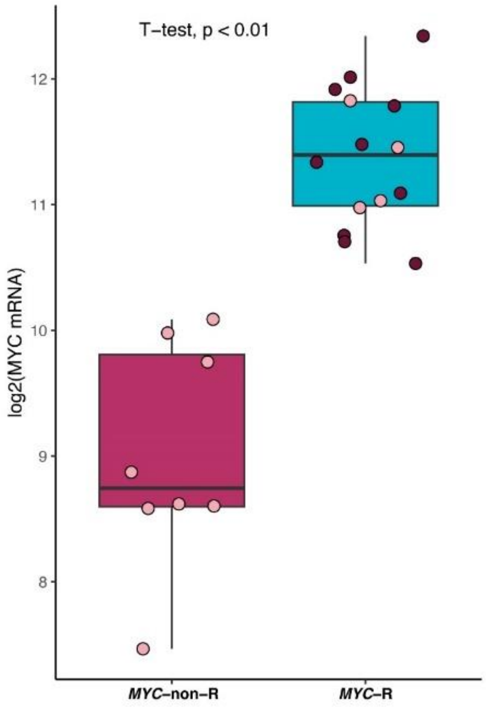
Supplementary Figure S2. Pipeline for NGS analysis for each individual sample. **(A)** Alignment and pre-processing steps of raw reads. **(B)** Mutation detection from deduplicated reads to the prediction of potential driver mutations. **(C)** Structural variant identification from initial discovery to the manual curation of the breakpoints.



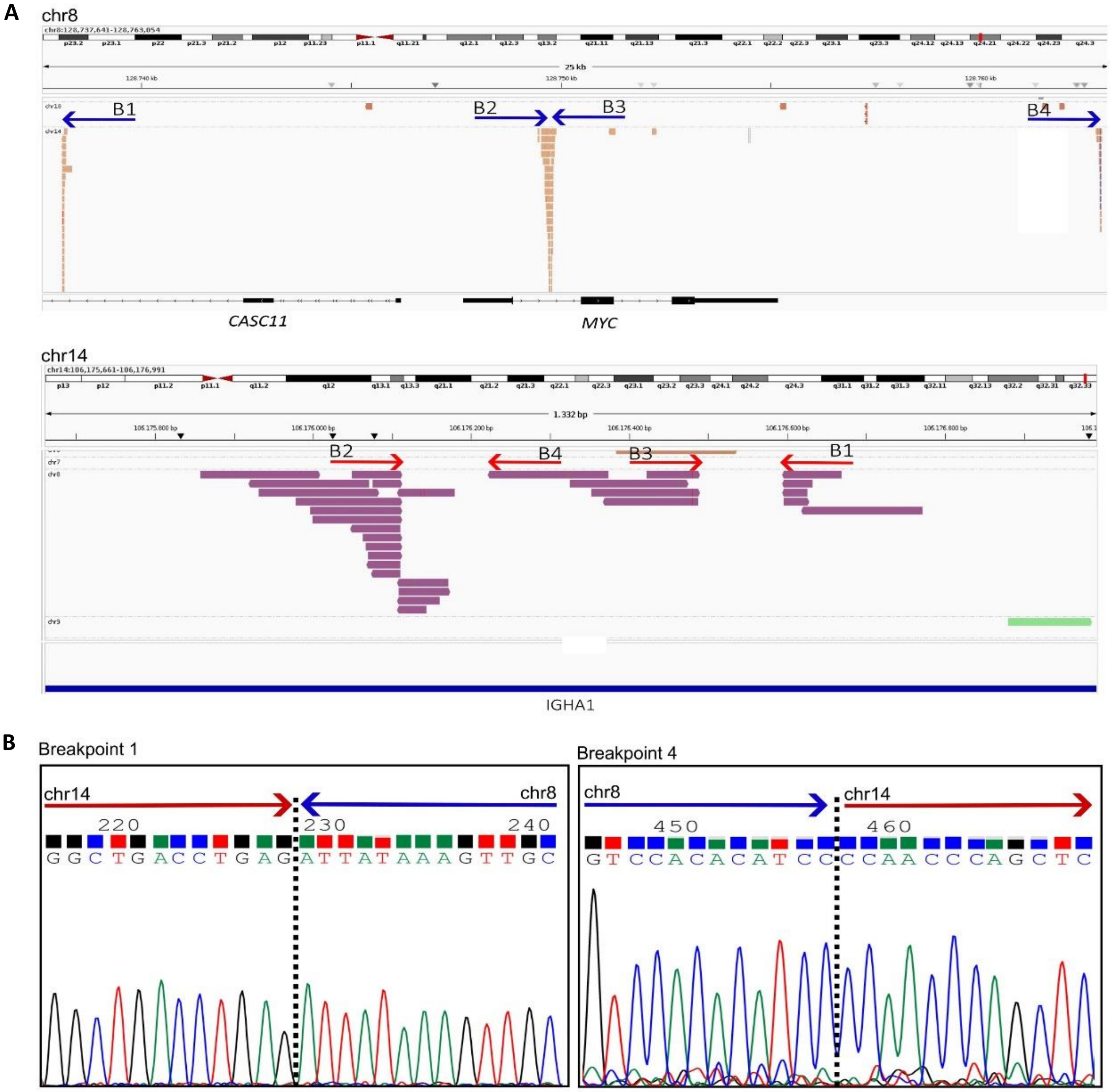
Supplementary Figure S3. Overview of histological, immunophenotypic and molecular findings in 37 B-NHL in CAYA with overlapping features between DLBCL and BL. Each column of the oncoprint represents one case and each line a specific analysis. On the right side of the figure, the frequency of each analysis is shown.



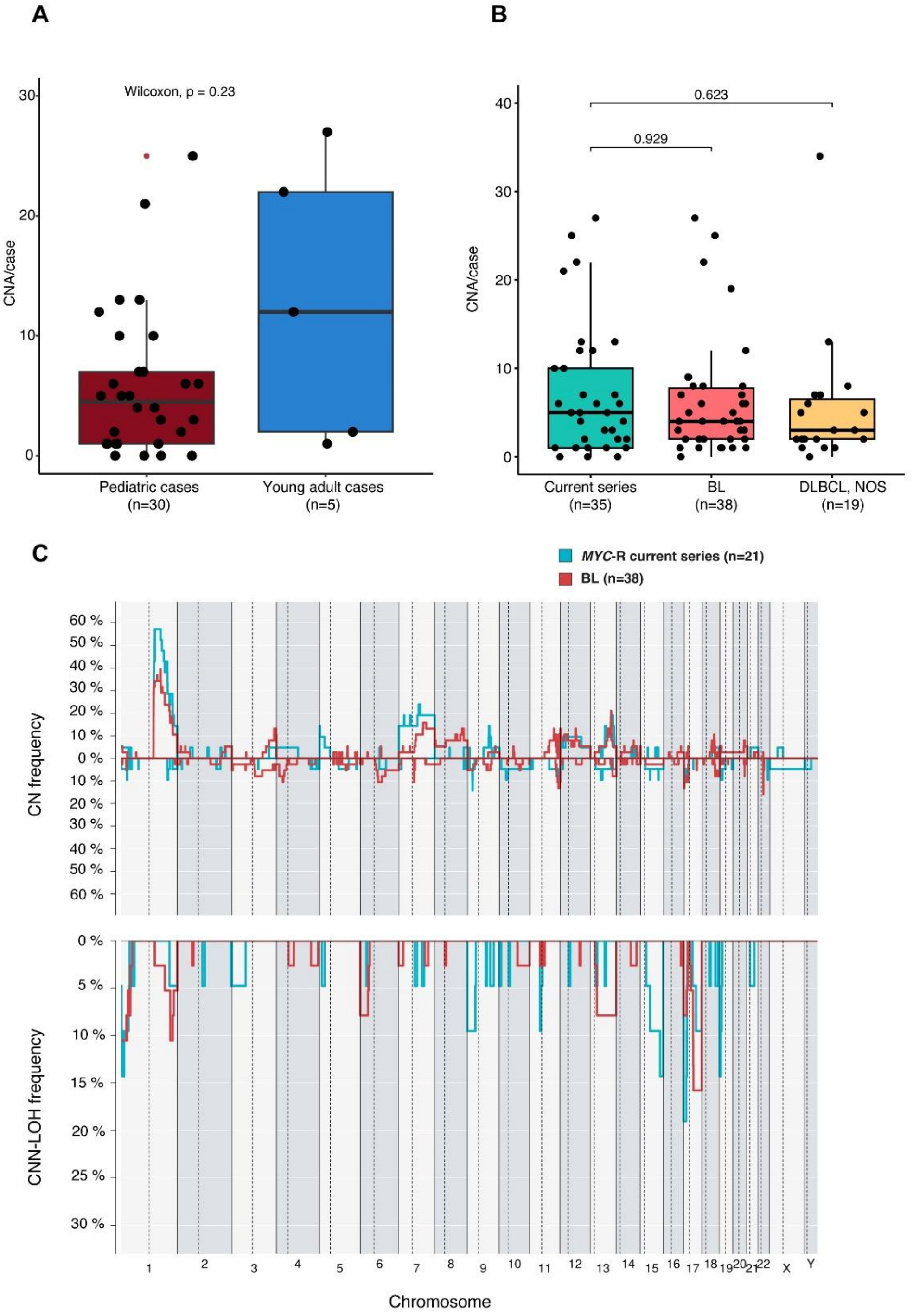
Supplementary Figure S4. Box plots representing the gene expression (log2 of normalized number of counts) of *MYC* gene (NM_002467.3 from DLBCL90 assay). Dots are labelled according to IHC protein expression (Clone Y69, Ventana, Roche). Cases with MYC positivity by IHC are indicated in dark red dots, whereas cases negative by IHC expression are labelled in light pink. Differences in mRNA MYC expression were observed between *MYC*-R and *MYC*-non-R cases (*P*<0.05).



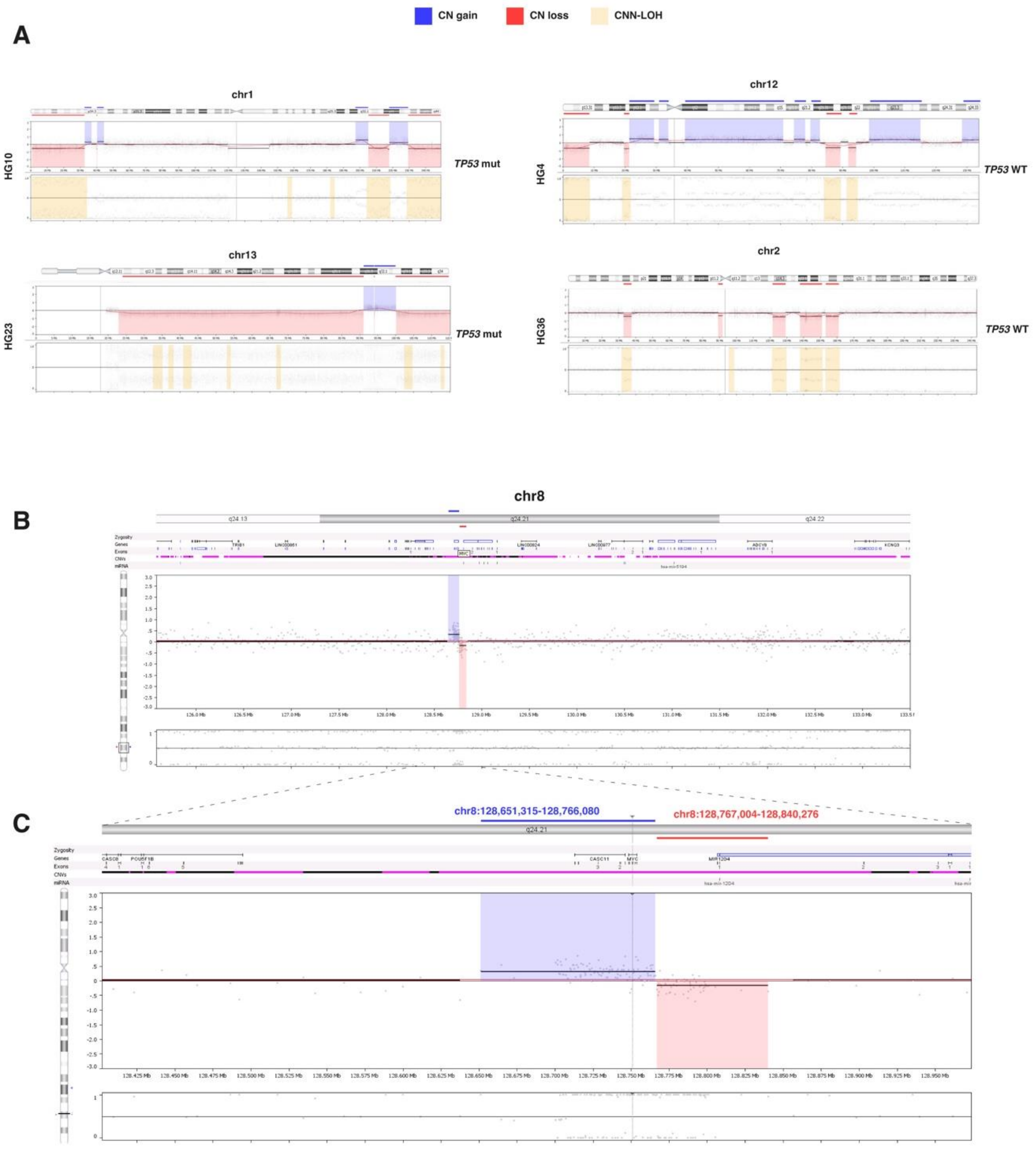
Supplementary Figure S5. Schematic outline of the cryptic four-breakpoint complex *IGH::MYC* rearrangement that juxtaposed the *MYC* coding exons 2 and 3 and the class-switch region (CSR) of IGHA1 (IGHA1 switch α1, chr14:106,175,034-106,178,629; hg19). **(A)** Schematic overview of the complex translocation using IGV. B, breakpoint. **(B)** Verification of the breakpoint junctions by Sanger sequencing (primers used not shown).



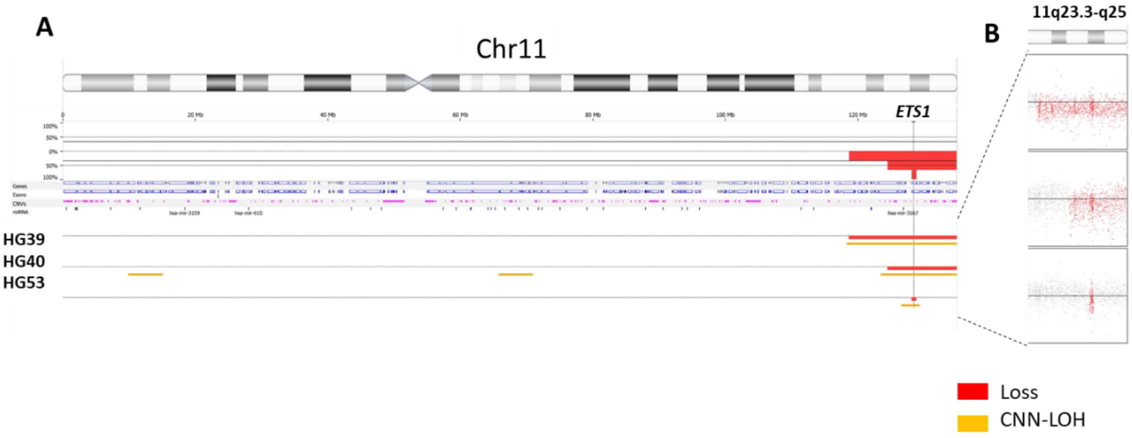
Supplementary Figure S6. Comparison of number of copy number alterations (CNA). **(A)** according to age groups and **(B)** with previously published data on BL (28) and pediatric and young adult DLBCL (27). The vertical axis represents number of alterations, and the different groups are separated in the X-axis. No significant differences according to Wilcoxon rank-sum test were observed **(C)** Comparative plot of CN and CNN-LOH. Light blue identifies *MYC*-rearranged (*MYC*-R) cases of the current series and red BL (28). No significant differences were observed according to Fisher's exact test.



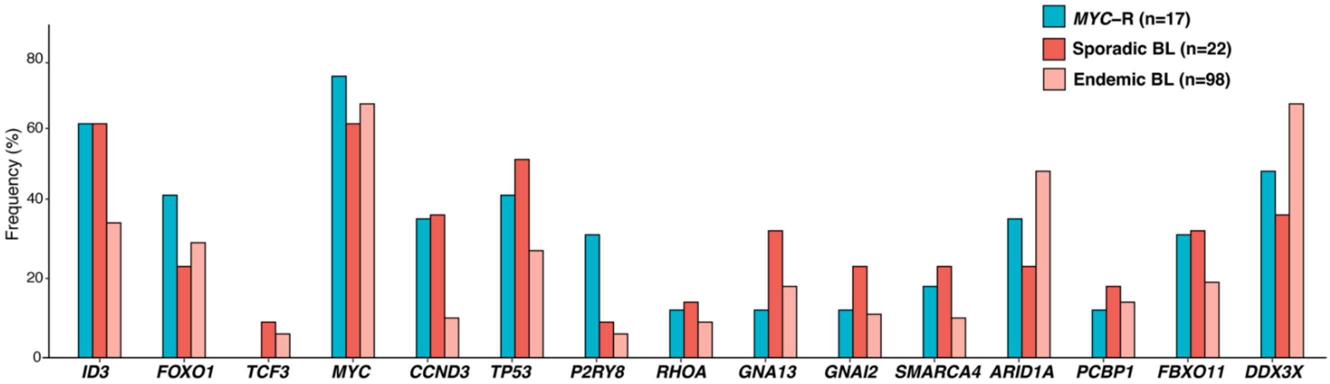
Supplementary Figure S7. (A) Chromotripsis-like patterns affecting chromosomes 1, 13, 12, and 2 identified by OncoScan arrays in cases HG10, HG23, HG4, and HG36, respectively. **(B)** Copy number pattern by OncoScan array indicative of a cryptic *MYC* rearrangement with a gain and loss pattern at 8q24.21 in case HG8. **(C)** In detail, the pattern suggests the existence of a chromosomal break 3’ of the *MYC* gene (chr8:128840276-128767004; GRCh37/hg19). CN gains are displayed in blue, CN losses in red and CNN-LOH in yellow.



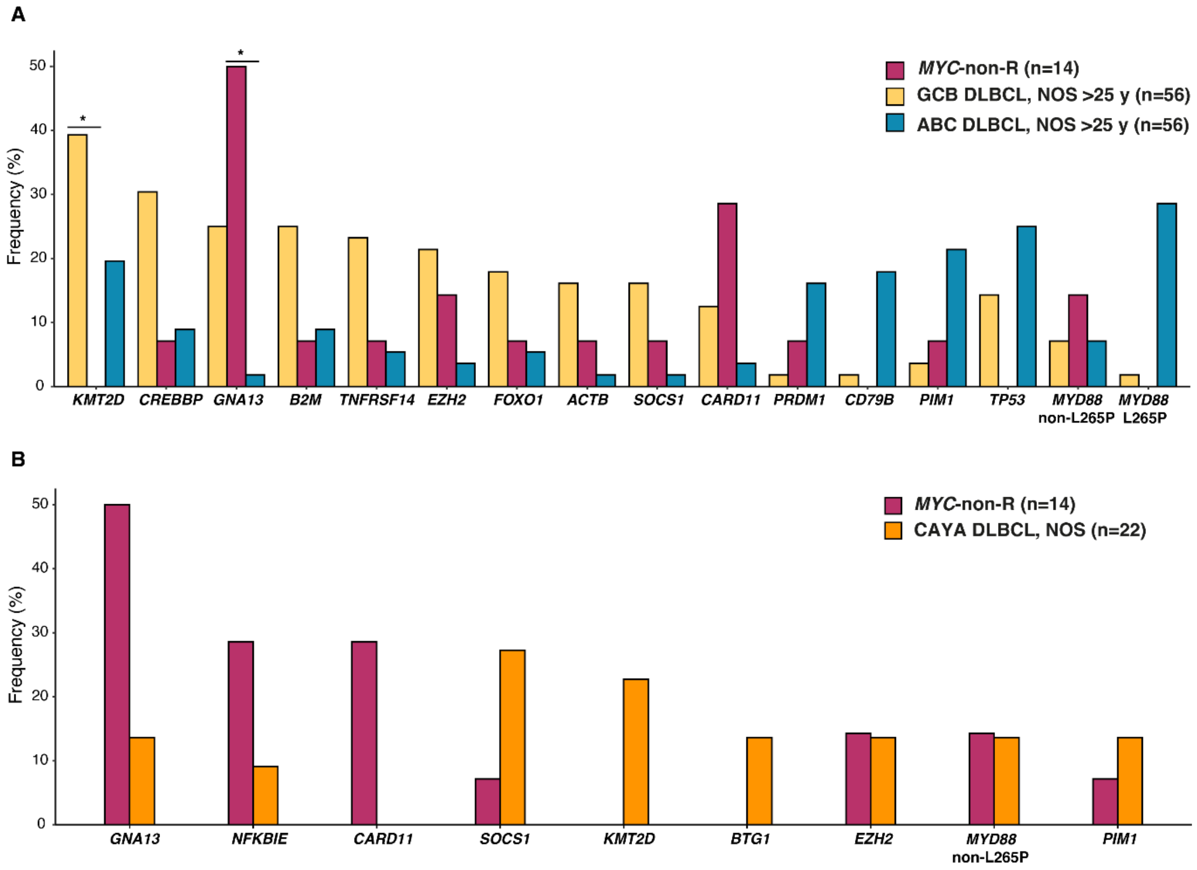
Supplementary Figure S8. (A) Copy number profile of chr11 in the three cases with terminal 11q deletion by OncoScan array (GRCh37/hg19) including **(B)** zoom in in the deleted region visualizing probes. Copy number losses are depicted in red and stretches of loss of heterozygosity associated to the CN losses are indicated in yellow.



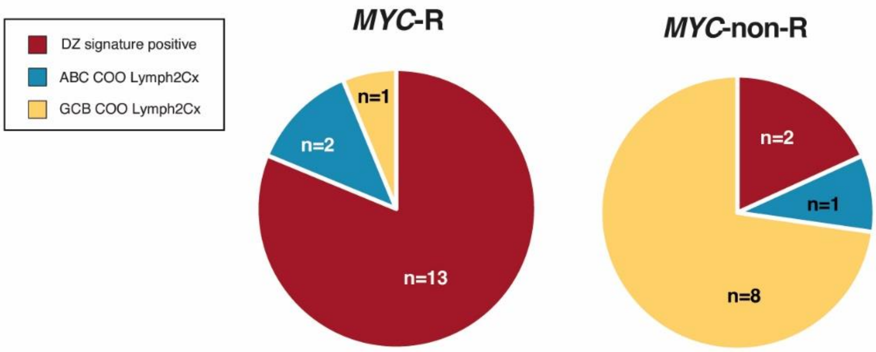
Supplementary Figure S9. Mutation frequency (%) of genes recurrently mutated in our *MYC*-rearranged (*MYC*-R) cases in comparison to sporadic BL (sBL) and endemic BL (eBL) (31). To perform statistical analysis using comparable data, only exonic *MYC* mutations were considered. No significant differences were observed according to adjusted (FDR) Fisher's exact test.



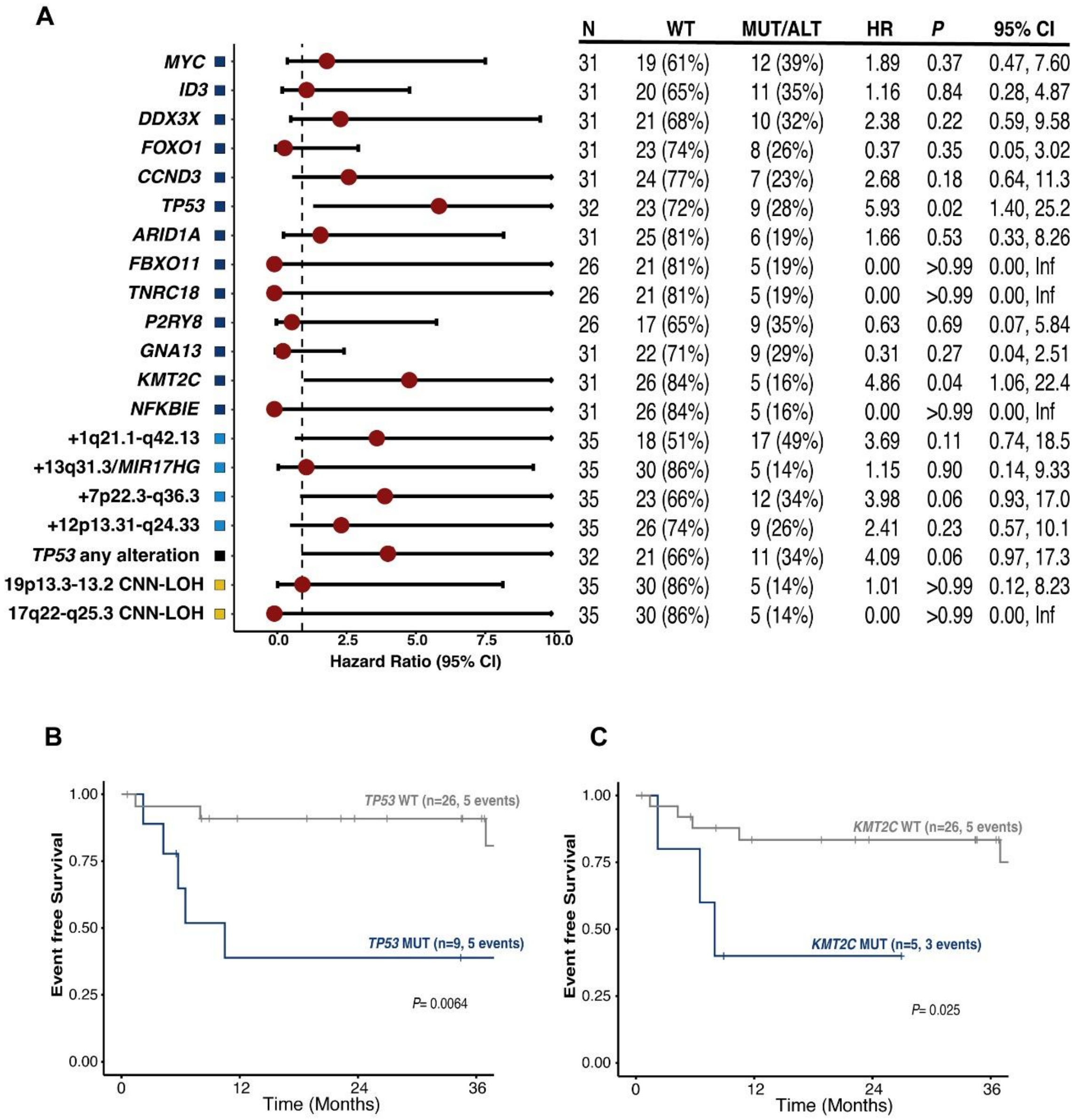
Supplementary Figure S10. Mutation frequency (%) of genes recurrently mutated in our *MYC*-non-R cases in comparison to **(A)** adult GCB and ABC DLBCL, NOS (32) and **(B)** CAYA DLBCL, NOS (27). Asterisks identify significant differences according to adjusted (FDR) Fisher's exact test. (*P-adjusted* < 0.05).



Supplementary Figure S11. GEP profiling according to *MYC*-R. Dark red represents expression of DZsig. Blue indicates an ABC COO and yellow indicates GCB COO.

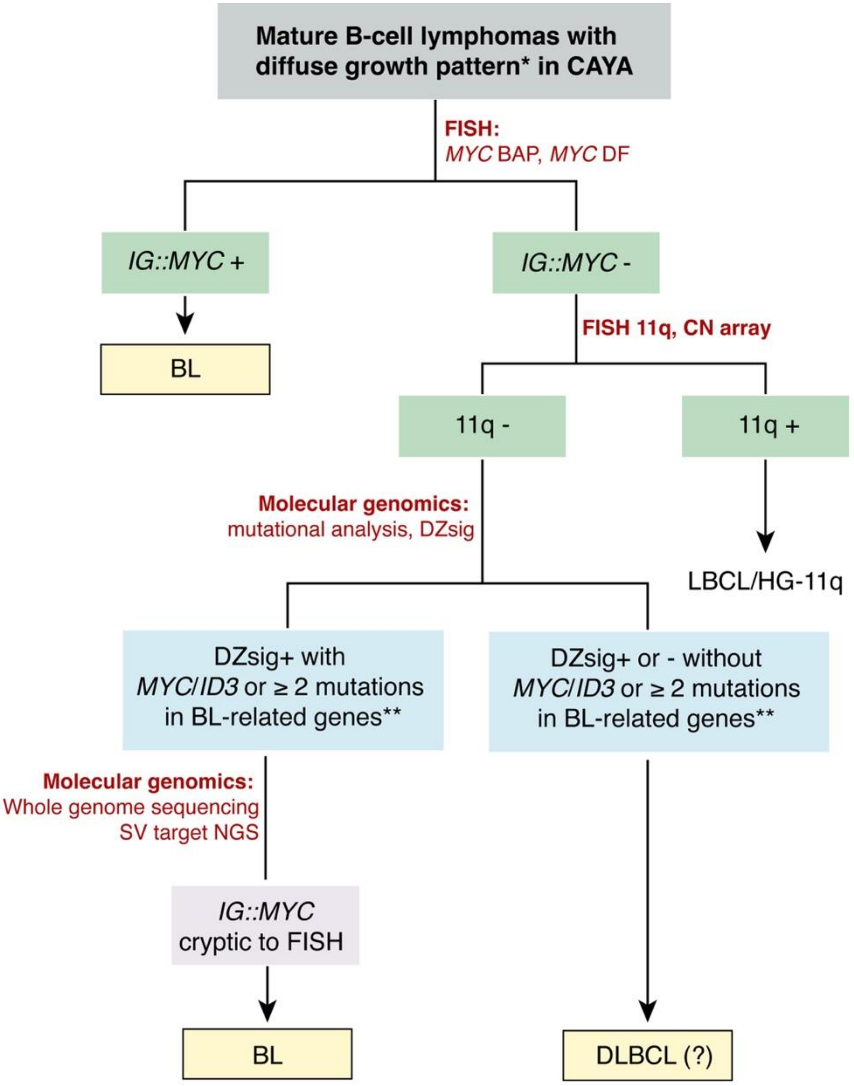


Supplementary Figure S12. (A) Impact of genetic alterations on EFS. The impact is quantified with the hazard ratio and its 95% confidence interval. The dark blue, blue, and yellow boxes indicate the type of genetic alteration (SNVs/indel, gain, and CNN-LOH, respectively). The right columns show: the number of cases (N), the number of non-altered cases (WT) and the number of mutated/altered cases (MUT/ALT), and the *P*-value (*P*) of the log-rank test. Only alterations with at least 5 altered cases are shown **(B-C)** Kaplan-Meier curves of EFS based on (**B**) the presence of *TP53* or (**C**) *KMT2C* mutations.



	N	WT	MUT/ALT	HR	*P*	95% CI
MYC	31	19 (61%)	12 (39%)	1.89	0.37	0.47, 7.60
ID3	31	20 (65%)	11 (35%)	1.16	0.84	0.28, 4.87
DDX3X	31	21 (68%)	10 (32%)	2.38	0.22	0.59, 9.58
FOXO1	31	23 (74%)	8 (26%)	0.37	0.35	0.05, 3.02
CCND3	31	24 (77%)	7 (23%)	2.68	0.18	0.64, 11.3
TP53	32	23 (72%)	9 (28%)	5.93	0.02	1.40, 25.2
ARID1A	31	25 (81%)	6 (19%)	1.66	0.53	0.33, 8.26
FBXO11	26	21 (81%)	5 (19%)	0.00	>0.99	0.00, Inf
TNRC18	26	21 (81%)	5 (19%)	0.00	>0.99	0.00, Inf
P2RY8	26	17 (65%)	9 (35%)	0.63	0.69	0.07, 5.84
GNA13	31	22 (71%)	9 (29%)	0.31	0.27	0.04, 2.51
KMT2C	31	26 (84%)	5 (16%)	4.86	0.04	1.06, 22.4
NFKBIE	31	26 (84%)	5 (16%)	0.00	>0.99	0.00, Inf
+1q21.1-q42.13	35	18 (51%)	17 (49%)	3.69	0.11	0.74, 18.5
+13q31.3/*MIR17HG*	35	30 (86%)	5 (14%)	1.15	0.90	0.14, 9.33
+7p22.3-q36.3	35	23 (66%)	12 (34%)	3.98	0.06	0.93, 17.0
+12p13.31-q24.33	35	26 (74%)	9 (26%)	2.41	0.23	0.57, 10.1
***TP53* any alteration**	32	21 (66%)	11 (34%)	4.09	0.06	0.97, 17.3
19p13.3-13.2 CNN-LOH	35	30 (86%)	5 (14%)	1.01	>0.99	0.12, 8.23
17q22-q25.3 CNN-LOH	35	30 (86%)	5 (14%)	0.00	>0.99	0.00, Inf

Supplementary Figure S13. Suggested molecular, cytogenetic, and genetic algorithm for the diagnosis of mature B-cell lymphomas with diffuse growth pattern in children, adolescents, and young adults (CAYA). *IG::MYC* detected by *MYC* break-apart and/or t(8;14) dual-fusion FISH probes. The 11q aberration is defined as the presence of 11q24.3-q23.3 gains followed by telomeric loss in 11q24-qter detected by FISH and/or copy number arrays. As described, in CAYA tumors negative for *MYC*-R and 11q aberrations by FISH, the presence of BL related mutations (*MYC*/*ID3* or 2 other BL-related mutated genes [*DDX3X*, *SMARCA4*, *CCND3*, and *TCF3*]) (33,34) and DZ signature expression suggest a BL diagnosis with a cryptic *IG::MYC* translocation. An NGS-based SV analysis, if available, could be helpful to confirm *MYC* rearrangement and refine BL diagnosis.



*Excluding LBCL-*IRF4* by IHC and/or FISH

** BL-related genes: *DDX3X, CCND3, SMARCA4, TCF3*

Supplementary Tables

Supplementary Table S1. Clinicopathological features of 37 aggressive B-cell lymphomas in CAYA with overlapping features between DLBCL and BL.

Provided in excel format

Supplementary Table S2. Morphological features of 37 B-NHL in CAYA with overlapping features between DLBCL and BL

Case	Starry sky	Nucleus size	Pleomorphism	Irregular nuclei	Burkitt-like chromatin	Necrosis	Morphology
HG2	no	medium/large	yes	yes	no	apoptosis	intermediate
HG3	yes	small	yes	yes	no	no	blastoid
HG4	yes	medium	yes	yes	no	no	intermediate
HG5	no	medium	yes	yes	no	apoptosis	intermediate
HG6	yes	medium	yes	yes	no	no	intermediate
HG7	NA	medium/large	yes	yes	no	apoptosis	intermediate
HG8	yes	medium	yes	yes	no	no	intermediate
HG9	yes	medium	yes	yes	yes	no	intermediate
HG10	no	large	yes	yes	no	yes	DLBCL
HG11	yes	medium	yes	yes	no	yes	intermediate
HG13	no	medium/large	yes	yes	no	no	intermediate
HG15	yes	medium	yes	yes	no	no	intermediate
HG16	yes	large	yes	yes	no	no	DLBCL
HG17	yes	medium	low	no	no	no	intermediate
HG18	no	medium	yes	yes	no	no	intermediate
HG20	yes	medium/large	yes	yes	no	no	intermediate
HG21*	NA	large	NA	NA	NA	NA	DLBCL
HG22	no	large	yes	yes	no	no	intermediate
HG23	yes	medium	yes	yes	no	no	intermediate
HG28	no	medium	yes	yes	no	no	intermediate
HG29	yes	medium	low	yes	partial	apoptosis	intermediate
HG33	no	large	yes	yes	no	no	DLBCL
HG34	no	medium	yes	yes	no	apoptosis	intermediate
HG36	yes	medium	yes	yes	no	no	intermediate
HG37	yes	medium	yes	yes	no	yes	intermediate
HG38	no	medium	yes	yes	NE	yes	intermediate
HG39	yes	medium	no	yes	no	no	blastoid
HG40	focal	medium	yes	yes	no	apoptosis	intermediate
HG41	yes	medium	no	yes	no	no	intermediate
HG43	no	large	yes	yes	no	no	blastoid
HG44	no	medium	yes	yes	yes	no	intermediate
HG46	no	medium	low	yes	no	apoptosis	intermediate
HG47	no	medium/large	yes	yes	NE	no	intermediate
HG50	no	medium/large	yes	no	no	no	blastoid
HG51	no	medium	no	yes	no	apoptosis	intermediate
HG52	yes	medium/large	yes	yes	no	no	intermediate
HG53	yes	medium/large	yes	yes	no	no	intermediate

*HG21 evaluation was based on the local pathology report; NA: not assessable; NE: not evaluable due to quality.

Supplementary Table S3. Details of all antibodies source and conditions of use.

Antibody	Clone	Source	Antigen retrieval/visualization	Dilution
CD20	L26	DAKO, (Copenhagen, Denmark)	EDTA 1 mM pH 9/ ENVISION FLEX (DAKO)	RTU
CD10	56C6	DAKO	EDTA 1 mM pH 9/ ENVISION FLEX (DAKO)	RTU
BCL6	PG-B6p	DAKO	EDTA 1 mM pH 9/ ENVISION FLEX (DAKO)	RTU
BCL2	124	DAKO	EDTA 1 mM pH 9/ ENVISION FLEX (DAKO)	RTU
Ki67	Mib-1	DAKO	Citrate 10 mM pH 6/ ENVISION FLEX (DAKO)	RTU
MUM1	MRQ-43	Ventana, Roche (Oro Walley, AR, USA)	CC1 solution / ultraView Universal DAB Detection Kit. Automated immunostainer (Benchmark XT; Ventana)	RTU
MYC	Y69	Ventana, Roche	CC1 solution / ultraView Universal DAB Detection Kit. Automated immunostainer (Benchmark XT; Ventana)	RTU
LMO2	SP51	Cell Marque	CC1 solution / ultraView Universal DAB Detection Kit. Automated immunostainer (Benchmark XT; Ventana)	RTU
SOX11	MRQ-58	Cell Marque	CC1 solution / ultraView Universal DAB Detection Kit. Automated immunostainer (Benchmark XT; Ventana)	RTU

RTU, ready to use.

According to previous reports MYC, BCL2 (1), BCL6 and MUM1 (2) were considered positive when >40% ≥70%, ≥30% or ≥60% of the cells were positive, respectively.

Supplementary Table S4. Genes included in the SureSelectXT custom panel (GRCh37/hg19) used for the NGS analysis.

Provided in excel format

Supplementary Table S5. List of structural variants identified in 26 aggressive B-cell lymphomas with overlapping features between BL and DLBCL.

Provided in excel format

Supplementary Table S6. Global table of copy number and copy number neutral of heterozygosity (CNN-LOH) alterations identified in 35 aggressive B-cell lymphomas with overlapping features between BL and DLBCL.

Provided in excel format

Supplementary Table S7. List of somatic mutations identified in 31 aggressive B-cell lymphomas with overlapping features between BL and DLBCL in children and young adults including prediction of amino acid changes that affect protein function (MutationAssessor, SIFT, Polyphen2).

Provided in excel format

Supplementary Table S8. KEGG pathway enrichment analysis on *MYC*-non-R cases (n=14). Only the top 10 significant results (*P-adjusted*<0.05) are displayed, using the Benjamini-Hochberg (BH) method.

KEGG Pathway	**Term**	**Gene Symbol**	***P*-value**	**Adjusted *P*-value**	**q-value**
hsa05166	**Human T-cell leukemia virus 1 infection**	*CREBBP, EGR1, TRRAP, ETS1, CCND3, HLA-A, IKBKB, MYC, ATM, B2M, TCF3, NFKBIA, MAP2K1*	1.642E-11	2.923E-09	1.763E-09
hsa04218	**Cellular senescence**	*ETS1, CCND3, HLA-A, ZFP36L1, FOXO1, MYC, ATM, MAP2K1*	5.907E-07	5.257E-05	3.171E-05
hsa04662	**B-cell receptor signaling pathway**	*NFKBIE, CARD11, IKBKB, BCL10, NFKBIA, MAP2K1*	3.844E-06	1.846E-04	1.113E-04
hsa05169	**Epstein-Barr virus infection**	*NFKBIE, CCND3, HLA-A, IKBKB, MYC, B2M, NFKBIA, MYD88*	4.147E-06	1.846E-04	1.113E-04
hsa04064	**NF-kappa B signaling pathway**	*CARD11, IKBKB, BCL10, ATM, NFKBIA, MYD88*	9.422E-06	2.991E-04	1.804E-04
hsa05161	**Hepatitis B**	*CREBBP, DDX3X, IKBKB, MYC, NFKBIA, MYD88, MAP2K1*	1.008E-05	2.991E-04	1.804E-04
hsa05164	**Influenza A**	*CREBBP, CCND3, IKBKB, ACTB, NFKBIA, MYD88, MAP2K1*	1.436E-05	3.651E-04	2.203E-04
hsa04660	**T-cell receptor signaling pathway**	*NFKBIE, CARD11, IKBKB, BCL10, NFKBIA, MAP2K1*	2.126E-05	4.731E-04	2.854E-04
hsa04068	**FoxO signaling pathway**	*CREBBP, BCL6, IKBKB, FOXO1, ATM, MAP2K1*	3.338E-05	6.601E-04	3.982E-04

Supplementary Table S9. Event-free Survival and Hazard Ratio of 37 aggressive B-cell lymphomas with overlapping features between BL and DLBCL in children and young adults.

Provided in excel format

Supplementary Discussion

As prior described in the literature (35,36), HG11 displayed a cryptic four-breakpoint complex *IGH*::*MYC* rearrangement that juxtaposed the *MYC* coding exons 2 and 3 and the class-switch region (CSR) of IGHA1. The size of the involved chromosomal fragments and complexity of this structural abnormality explains why it was not detected by FISH (**Supplementary Figure S5**).

Otherwise, HG29 *MYC*-negative by FISH, showed a classical t(8;14) translocation with a class I breakpoint within the region of detection of the *MYC* BAP probe used. Therefore, this discrepancy could be attributed to technical issues in a tissue with a 70% tumor cell content (TCC). In fact, although the case was considered *MYC*-negative by FISH, isolated cells with *MYC* breaks were detected (not reaching the 5% threshold to consider it positive) (**Supplementary Figure S3**). Of note, both cases carried *MYC*, *DDX3X* mutations and were positive for the DZ signature. Moreover, although the capture SV-NGS strategy allows a resolution not reachable by FISH, it still has some limitations. In our series, one case (HG2) with an *IGH*::*MYC* translocation detected by FISH was not identified by the SV-NGS strategy. This discrepancy could be explained, as previously reported (4), by a potential FFPE DNA degradation on the regions of breakpoint.

Supplementary References

1. Johnson NA, Slack GW, Savage KJ, Connors JM, Ben-Neriah S, Rogic S, et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. Journal of Clinical Oncology. 2012;30:3452–9.
2. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood. 2004;103:275–82.
3. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, et al. A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. Cancer Cell. 2020;37:551–68.
4. Chong LC, Ben-Neriah S, Slack GW, Freeman C, Ennishi D, Mottok A, et al. High-resolution architecture and partner genes of MYC rearrangements in lymphoma with DLBCL morphology. Blood Adv. 2018;2:2755–65.
5. López C, Burkhardt B, Chan JKC, Leoncini L, Mbulaiteye SM, Ogwang MD, et al. Burkitt lymphoma. Nat Rev Dis Primers. 2022;8:78.
6. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
7. Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics. 2009;25:1754–60.
8. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics. 2012;28:i333–9.
9. Wala JA, Bandopadhayay P, Greenwald NF, O’Rourke R, Sharpe T, Stewart C, et al. SvABA: genome-wide detection of structural variants and indels by local assembly. Genome Res. 2018;28:581–91.
10. Cameron DL, Schröder J, Penington JS, Do H, Molania R, Dobrovic A, et al. GRIDSS: Sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly. Genome Res. 2017;27:2050–60.
11. Nadeu F, Mas-de-les-Valls R, Navarro A, Royo R, Martín S, Villamor N, et al. IgCaller for reconstructing immunoglobulin gene rearrangements and oncogenic translocations from whole-genome sequencing in lymphoid neoplasms. Nat Commun. 2020;11:3390.
12. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. Nat Biotechnol. 2011;29:24–6.
13. Haile S, Corbett RD, Bilobram S, Bye MH, Kirk H, Pandoh P, et al. Sources of erroneous sequences and artifact chimeric reads in next generation sequencing of genomic DNA from formalin-fixed paraffin-embedded samples. Nucleic Acids Res. 2019;47:e12.
14. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, et al. From FastQ data to high-confidence variant calls: the Genome Analysis Toolkit best practices Pipeline. Curr Protoc Bioinformatics. 2013;43:11.10.1-11.10.33.

15. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297.

16. Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, Mcewen R, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res. 2016;44:e108.

17. Dunn T, Berry G, Emig-Agius D, Jiang Y, Lei S, Iyer A, et al. Pisces: an accurate and versatile variant caller for somatic and germline next-generation sequencing data. Bioinformatics. 2019;35:1579.

18. Koboldt DC, Larson DE, Wilson RK. Using VarScan 2 for Germline Variant Calling and Somatic Mutation Detection. Curr Protoc Bioinformatics. 2013;44:15.4.1-15.4.17.

19. Wilm A, Aw PPK, Bertrand D, Yeo GHT, Ong SH, Wong CH, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res. 2012;40:11189–201.

20. Paila U, Chapman BA, Kirchner R, Quinlan AR. GEMINI: integrative exploration of genetic variation and genome annotations. PLoS Comput Biol. 2013;9:e1003153.

21. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012;6:80–92.

22. Puente XS, Beà S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015;526:519–24.

23. Bastos-Oreiro M, Suárez-González J, Andrés-Zayas C, Carrión NC, Moreno S, Carbonell D, et al. Incorporation of next-generation sequencing in clinical practice using solid and liquid biopsy for patients with non-Hodgkin's lymphoma. Sci Rep. 2021;11:22815.

24. Petrackova A, Vasinek M, Sedlarikova L, Dyskova T, Schneiderova P, Novosad T, et al. Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics. Front Oncol. 2019;9::851.

25. Jennings LJ, Arcila ME, Corless C, Kamel-Reid S, Lubin IM, Pfeifer J, et al. Guidelines for Validation of Next-Generation Sequencing–Based Oncology Panels: A Joint Consensus Recommendation of the Association for Molecular Pathology and College of American Pathologists. The Journal of molecular diagnosis. 2017;19:341–65.

26. Edelmann J, Holzmann K, Miller F, Winkler D, Bühler A, Zenz T, et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. Blood. 2012;120:4783–94.

27. Ramis-Zaldivar JE, Gonzalez-Farré B, Balagué O, Celis V, Nadeu F, Salmerón-Villalobos J, et al. Distinct molecular profile of IRF4-rearranged large B-cell lymphoma. Blood. 2020;135:274–86.

28. Scholtysik R, Kreuz M, Klapper W, Burkhardt B, Feller AC, Hummel M, et al. Detection of genomic aberrations in molecularly defined Burkitt's lymphoma by array-based, high resolution, single nucleotide polymorphism analysis. Haematologica. 2010;95:2047–55.

29. Ennishi D, Jiang A, Boyle M, Collinge B, Grande BM, Ben-Neriah S, et al. Double-hit gene expression signature defines a distinct subgroup of germinal center B-cell-like diffuse large B-cell lymphoma. Journal of Clinical Oncology. 2019;37:190–201.

30. Alduaij W, Collinge B, Ben-Neriah S, Jiang A, Hilton LK, Boyle M, et al. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. Blood. 2023;141:2493–507.

31. Grande BM, Gerhard DS, Jiang A, Griner NB, Abramson JS, Alexander TB, et al. Genome-wide discovery of somatic coding and noncoding mutations in pediatric endemic and sporadic Burkitt lymphoma. Blood. 2019;133:1313–24.

32. Karube K, Enjuanes A, Dlouhy I, Jares P, Martin-Garcia D, Nadeu F, et al. Integrating genomic alterations in diffuse large B-cell lymphoma identifies new relevant pathways and potential therapeutic targets. Leukemia. 2018;32:675–84.

33. Thomas N, Dreval K, Gerhard DS, Hilton LK, Abramson JS, Ambinder RF, et al. Genetic subgroups inform on pathobiology in adult and pediatric Burkitt lymphoma. Blood. 2023;141:904–16.

34. Burkhardt B, Michgehl U, Rohde J, Erdmann T, Berning P, Reutter K, et al. Clinical relevance of molecular characteristics in Burkitt lymphoma differs according to age. Nat Commun. 2022;13:3881.

35. King RL, McPhail ED, Meyer RG, Vasmatzis G, Pearce K, Smadbeck JB, et al. False-negative rates for MYC fluorescence in situ hybridization probes in B-cell neoplasms. Haematologica. 2019;104:e248–51.

36. Wagener R, Bens S, Toprak UH, Seufert J, López C, Scholz I, et al. Cryptic insertion of MYC exons 2 and 3 into the immunoglobulin heavy chain locus detected by whole genome sequencing in a case of “MYC-negative” Burkitt lymphoma. Haematologica. 2020;105:e202–5.