Supplementary Materials and Methods

Patient samples and B-cell lines

Pediatric BL specimens were obtained from the cooperative human tissue network pediatric NHL repository through the Children's Oncology Group (COG)[1] and adult BL specimens were collected from the Nebraska Lymphoma Study Group (NLSG) registry and Tissue Bank. Informed consent was obtained for all patients in the study. The clinical and pathological characteristics of the BL patients are tabulated in **Table-S1A&B**. We included 61 molecularly diagnosed BL (pediatric (<21 years), adult ≥21). The morphological diagnoses were performed by consensus pathology review by LLMPP members, and the molecular diagnoses were performed using the Dave et al. gene signature[2]. The study was approved by the Institutional Review Board of the University of the Nebraska Medical Center. Additional genomic data from previously published mBL series[3] (12 adults, and 18 pediatric), FL, tFL[4], and GCB-DLBCL[5,6] were included for comparative analysis.
BL cell lines, originally sourced from ATCC, were cultured in 10 mM HEPES-buffered RPMI 1640 (Gibco-Invitrogen, Carlsbad, CA) in (Ramos, Daudi, Raji, P3HR-1, and GA10) or IMDM (Namalwa) supplemented with 10% fetal bovine serum (FBS), penicillin G (100 U/ml) and streptomycin (100 µg/ml) and maintained at 37 ºC in 5% CO2. Cell lines were routinely tested for mycoplasma using the Universal Mycoplasm Detection Kit from ATCC. The characteristics of the BL cell lines including EBV status, MYC protein and mRNA, mutation status (*TCF3*, *TP53*, *ID3*, *CCND3*)[7,8], and *MIR17HG* copy number status are shown in **Table-S1B**. Mutations present in the cell lines were reported in previous studies [7-10].

MYC and BCL2 translocation status

MYC and BCL2 translocation status was known on a subset of cases based on clinical cytogenetics data or FISH. For the 52 cases that were sequenced the NGS data and Factera[11] was used to look for evidence of MYC, BCL2 or BCL6 translocations. **Table-S7** summarizes the cases identified as positive for MYC or the BCL2 translocation by NGS and denotes whether they were validated to be positive by FISH, cytogenetics data, or whether they were validated by Long Range PCR[12]. Of the 15 cases positive for MYC-IgH translocation by Factera,[11] 4 were confirmed positive based of previous cytogenetic or FISH information, and 4 were confirmed positive by long-range PCR, and 4 could not be confirmed by long-range PCR (Table-S7).

Structural and functional genomic analysis

Comparative genomic hybridization and copy number analysis:
Genomic DNA (gDNA) was isolated from fresh-frozen BL biopsies using the AllPrep DNA/RNA kit (Qiagen).The DNA (250ng) library was prepared using the NspI 250K SNP array protocol (Affymetrix, Inc), hybridized to GeneChips according to the manufacturer's protocol (Affymetrix, Inc), and scanned using a GeneChip Scanner 3000 (Affymetrix, Inc). CNA analysis was done as previously described[4]. Briefly, SNP genotypes and probe-intensity log2 ratios, generated using Genotyping Console 4.1 software (Affymetrix), were analyzed for CNAs using the DNA copy package[13] (Bioconductor); ≥4 copies were considered an amplification. Identified CNAs were aligned to determine the minimal common region (MCR), and genes within MCRs were identified as described. The percent aberrant genome per case was calculated by taking the total size of the aberrant regions divided by the total size of the genome using Nexus Copy Number software (Bio Discovery Inc). Of the 60 cases, 9 cases were excluded after initial analysis due to a suboptimal SNP array genomic data. We obtained Affymetrix 250K NspI SNP array data on 30 additional molecularly defined BL cases from a previously published series (GSE21597)[3], of which 22 used for analysis after the initial quality filter.

Target gene section and whole exome capture and data analysis:
A customized gene panel (n= 380) that targeted genes commonly mutated (≥5%) in NHLs, including BL, was designed using NimbleGen SeqCap Target Enrichment software (http://lifescience.roche.com/shop/en/us/products/seqcap-ez-human-exome-library-v30). Genomic DNA (2µg) was used for library preparations and

exomes were captured using the Kapa Biosystems exome enrichment kit, according to the manufacturer's protocol (Kapa Biosystems, Inc). Twelve samples were sequenced per lane on an Illumina HiSeq2500 sequencer for an average of 476-fold coverage of targeted regions after duplicate marking (range: 237x-805x). The data analysis of sequencing was performed as previously described[14]. Briefly, raw reads were mapped to the human genome (version HG19) using the Burrows-Wheeler Aligner-MEM (1.4.2). Genome Analysis Toolkit (1.0.2) was used for local realignment and base quality recalibration. Duplicate marking was done with Picard (v1.0.2). Variant calling and filtering were performed with VarScan 2 (v2.3.7) and Genome Analysis Toolkit UnifiedGenotyper. The variants were annotated using SeattleSeq (http://snp.gs.washington.edu/Seattle Seq Annotation138/index and Annovar[15]). Variants identified by both Varscan2 and GATK were selected for further analysis. Additional variant filters applied included requiring that variants should be supported by at least 5 reads with 100x coverage and a minimal variant allelic frequency of 10%. To obtain a confident variant list, a two-step filtering process was used. Variants were excluded if they 1) existed in the dbSNP database (v138), 2) did not change protein sequences or affect splice sites, 3) were also found in a set of unrelated normal samples, or 4) were shared by >30 cases unless they are from a gene known to have frequent mutations and thus present in the COSMIC database. These variants were designated as class 1 and represent high confidence somatic mutations. Class 2 mutations cannot be excluded as germline but are not annotated by annovar using the dbSNP138NotFlagged database (http://annovar.openbioinformatics.org/en/latest/user-guide/filter/). They include variants that 1) the positions were found in dbSNP but the changes were not 2) are flagged as clinically relevant in dbSNP or 3) have a <1% minor allele frequency in the general population. The class status of the mutations are noted in Table S3. Class 2 mutations include variants at known TP53 mutation hotspots R273 and R248. Given a sample size of n=52 (Adult+pediatric samples), we can detect mutations occurring at 10% frequency at 99.6% probability.

Gene and miRNA expression analysis:
Total RNA isolated from BL cases was used for GEP using Affymetrix U133Plus2.0 arrays and partly included in previous studies[16]. The raw data were normalized (MAS5.0) using BRB Array Tools (http://linus.nci.nih.gov /BRB-ArrayTools.html). GEP data from BL cases were used to classify samples as BL or DLBCL using a Bayesian classifier described previously[2]. The classification of the BL *versus* DLBCL cases was performed using the training cohort from a previous study[2]. The computational pathway and gene signature analysis was performed through GSEA using the lymphoid signature (http://lymphochip.nih.gov /signaturedb/) and Broad institute signature databases (http://software.broadinstitute.org/gsea/msigdb). Differential expression of genes was analyzed utilizing BRB arrayTools using T-test and other standard approaches described in earlier studies[16]. The microRNA (miRNA) expression data for a subset of the cases were generated using the Taqman® human microRNA array (V2.0, ABI, CA) and available from a previous publication[17].
We compared the average mRNA expression level of genes by copy number status (gain, loss, 2N) using one-sided Student's t-test to determine significance. The Ingenuity pathway analysis program (Qiagen, Inc) or DAVID (https://david.ncifcrf.gov/) was used for the molecular and functional annotation of the genes within aberrant regions.

Immunohistochemical staining of tissue sections
Tissue microarrays (TMAs) were prepared with adequate archival paraffin-embedded tissue. Five-micrometer sections were cut from each tissue microarray and stained with antibodies to MYC and BCL2, as described previously[18,19]. These TMAs were evaluated earlier by 2 pathologists for tumor content, CD20 stains and representative characteristics (KF, WCC). We used 50% or more staining as the positive threshold. A four-tiered scale (negative, 0; weak, 3; medium, 6; strong, 9) was used to grade the staining intensity of tumor cells. A 10-tiered scale (10% to 100%) was used to score the percentage of positive tumor cells. The product of the intensity and percentage of positive cells was used as the case score with a value ≥ 30 considered positive (e.g., ≥ 50% positive tumor cells with an intensity of 6, or 33% positive cells with an intensity of 9).

Survival outcome analysis

We have included the basic clinical characteristics of the BL in a previous study[17]. The overall survival (OS; death from any cause) was estimated using the Kaplan-Meier method, and differences were assessed using the log rank test. Statistical analyses were performed with the R-language survival package. Significant differences among groups were considered at P values below 0.05.

Graphing

Graphs were made using the R-language ggplot2 package, excel, or Circos plot (http://circos.ca).

Functional validation using *in vitro* assays in BL cell lines

Genomic status of loci encompassing MYC, and MIR17HG in BL cell lines:

The genomic DNA from BL cell lines was used to quantify the DNA copy number of *MYC* and the *MIR17HG* genomic locus in comparison to *GAPDH*, and tonsillar DNA was used a calibrator (normal). The qPCR experiment was performed using SYBR Green qPCR Kit (Bio Rad. Inc) using primers listed in **Table S7**. The standard $\Delta\Delta$ C_T ($C_{T\ (MYC\ or\ MIR17GH)}$–$C_{T\ (GAPDH)}$ method was used to estimate the gain/amplification status of genomic locus.

Quantitation of MYC and miRNA 17~92 expression levels:

MYC mRNA expression levels were estimated from total RNA by standard quantitative qRT-PCR procedure using the primers listed in **Table-S8,** the DyNAmo HS SYBR Green qPCR Kit (Thermo-Scientific Inc), and either the DNA Engine Opticon 2 or CFX Connect (Bio-Rad, Hercules, CA) Thermo-cycler. *GAPDH* or *RPL13A* were used to calibrate gene expression. The expression of three specific miRNAs (miR-17; miR-19; miR-92) was evaluated by TaqMan® small RNA assays using the TaqMan miRNA assays and 7900HT fast real-time PCR system (Applied Biosystems) and normalized to *RUN48* expression.

MYC and miRNA17~92 knockdown in BL cell lines:

Functional loss of miR17~92 in five BL cell lines (Raji, Namalwa, Daudi, P3HR1, and GA-10) was performed using a miRNA sponge, as detailed earlier[20]. Lentiviruses were generated by co-transfection of 5µg plasmid DNA (control, experimental) with 5µg of the packaging plasmid DNA into 293T cells. To estimate the sponge efficiency, protein expression of known miRNA target proteins (BIM and PTEN) was analyzed by western blotting.

CRISPR-mediated MYC or miRNA17~92 knockout:

For CRISPR-mediated knockout of MYC, a guide RNA targeting the MYC locus at the intron1-exon2 junction (**Table S8**) was cloned into a vector encoding espCas9 and GFP. Raji and Namalwa cells were transfected with the vectors by electroporation. GFP positive cells were selected by flow cytometry and expanded as single clones, MYC knockout was assessed by western blotting, and clones without MYC expression were selected for further study.

For CRISPR-mediated knockout of MIR17HG, two guide RNAs targeting upstream and downstream of the MIR17HG locus (**Table S8, Figure S10D**) were cloned into a vector encoding espCas9 and GFP. Daudi cells were transfected with the vector by electroporation. GFP positive cells were selected by flow cytometry and expanded as single clones. DNA from clones that grew out were isolated and the MIR17HG locus was examined by PCR. A clone with knockout of the MIR17HG locus was identified (**Figure S10E**) and selected for further study.

Evaluation of BCR signaling and in vitro Ibrutnib treatment

BL cell lines (with control and mirR17~92 sponge vectors) were treated with doxycycline (1µg/ml) for 48 hours and incubated with 10 µg/mL anti-IgM (Southern Biotech) to crosslink BCR for 5-10 minutes, followed by snap

freezing in liquid nitrogen. BLNK and SYK phosphorylation was quantified to assess BCR activation using western blots (for p-SYK, total SYK, p-BLNK, and total BLNK). An ABC-DLBCL cell line, TMD8, was used as a control.
BL cells lines (with control and miR17~92 sponge vectors) were treated with 2-12μM Ibrutnib (Cat#L124970, Toronto Research Chemical). The IC_{50} was estimated in parental cell lines (6μM for Raji and Daudi; 10μM for Namalwa and P3HR-1) with a cell viability curve (**Figure S11**) using Presto Blue™ cell viability reagent as per the manufacture's protocol instructions (Invitrogen Inc).

Western Blotting:
20-50μg of whole-cell extract was resolved by 10% Bis-Tris–polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane. The membrane was first blocked in Odyssey Blocking Buffer (LI-COR) at room temperature for 1 h and then incubated with specific primary antibody to (BIM (CST #2933), PTEN (CST #9559), Actin (Santa Cruz #1616), p-SYK (CST #2710), SYK (CST #2712), p-BTK (CST #5082), BTK (CST #8547), p-BLNK (CST #3601), BLNK (CST #3587), p-ERK (CST #4377), ERK (CST #9102), PTEN (CST #9188), MYC (CST #5605), BCL2 (CST #10571), BCLXL (CST #2764), and MCL1(CST #5453) at 4° C overnight, followed by treatment with specific secondary antibody. The immunoblots were visualized using the Odyssey CLX (LI-COR).

Quantitation of ITIM gene expression:
Sponge and control vector expressing cell lines were treated with Doxycycline (1μg/ml) for 48 hours and were treated with 10 μg/mL soluble anti-IgM (SouthernBiotech #202202) for 10 minutes followed by the snap freezing using liquid nitrogen. Total RNA was isolated using the RNeasy mini kit (Qiagen # 74104). The cDNA was generated with protoScript II, first strand synthesis super mix (New England BioLabs#0091702). q-PCR was performed using SsoAdvanced Universal SYBR Green supermix qPCR Kit (Bio Rad) using primers listed in **Table S8** and a CFX Connect (Bio-Rad, Hercules, CA) thermocycler.

Cell proliferation, apoptosis and cell cycle assays:
Cell proliferation assays were performed using Presto Blue or CyQUANT® (ThermoFisher, Inc) assays, and apoptosis was quantified using a FACSCalibur flow cytometer (BD Biosciences) after staining the cells with Annexin-V (Apoptosis detection kit; BioLegend) according to the manufacturer's instructions. Cell cycle distribution was analyzed using Propidium Iodide (PI)(11348639001, Roche). The cells were washed once with PBS and fixed with 70% ethanol or 1:1 Methanol:Acetone at -20 C overnight. PI was added to the cells at a final concentration of 5μg/ml and cells then analyzed with a BD LSR II flow cytometer.

Data availability
Upon publication, all sequencing data will be deposited in the NCBI SRA database and all SNP array and GEP data will be deposited in GEO.

References

1. Deffenbacher KE, Iqbal J, Sanger W, et al. Molecular distinctions between pediatric and adult mature B-cell non-Hodgkin lymphomas identified through genomic profiling. *Blood.* Apr 19 2012;119(16):3757-3766.
2. Dave SS, Fu K, Wright GW, et al. Molecular diagnosis of Burkitt's lymphoma. *The New England journal of medicine.* Jun 8 2006;354(23):2431-2442.
3. Scholtysik R, Kreuz M, Klapper W, et al. Detection of genomic aberrations in molecularly defined Burkitt's lymphoma by array-based, high resolution, single nucleotide polymorphism analysis. *Haematologica.* Dec 2010;95(12):2047-2055.
4. Bouska A, McKeithan TW, Deffenbacher KE, et al. Genome-wide copy-number analyses reveal genomic abnormalities involved in transformation of follicular lymphoma. *Blood.* Mar 13 2014;123(11):1681-1690.
5. Lenz G, Wright GW, Emre NC, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proceedings of the National Academy of Sciences of the United States of America.* Sep 9 2008;105(36):13520-13525.
6. Scandurra M, Mian M, Greiner TC, et al. Genomic lesions associated with a different clinical outcome in diffuse large B-Cell lymphoma treated with R-CHOP-21. *Br J Haematol.* Nov 2010;151(3):221-231.
7. Love C, Sun Z, Jima D, et al. The genetic landscape of mutations in Burkitt lymphoma. *Nature genetics.* Dec 2012;44(12):1321-1325.
8. Schmitz R, Young RM, Ceribelli M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature.* Oct 4 2012;490(7418):116-120.
9. Greenough A, Dave SS. New clues to the molecular pathogenesis of Burkitt lymphoma revealed through next-generation sequencing. *Current opinion in hematology.* Jul 2014;21(4):326-332.
10. Richter J, Schlesner M, Hoffmann S, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nature genetics.* Dec 2012;44(12):1316-1320.
11. Newman AM, Bratman SV, Stehr H, et al. FACTERA: a practical method for the discovery of genomic rearrangements at breakpoint resolution. *Bioinformatics.* Dec 1 2014;30(23):3390-3393.
12. Basso K, Frascella E, Zanesco L, Rosolen A. Improved long-distance polymerase chain reaction for the detection of t(8;14)(q24;q32) in Burkitt's lymphomas. *The American journal of pathology.* Nov 1999;155(5):1479-1485.
13. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics.* Mar 15 2007;23(6):657-663.
14. Green MR, Kihira S, Liu CL, et al. Mutations in early follicular lymphoma progenitors are associated with suppressed antigen presentation. *Proceedings of the National Academy of Sciences of the United States of America.* Mar 10 2015;112(10):E1116-1125.
15. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research.* Sep 2010;38(16):e164.
16. Iqbal J, Shen Y, Liu Y, et al. Genome-wide miRNA profiling of mantle cell lymphoma reveals a distinct subgroup with poor prognosis. *Blood.* May 24 2012;119(21):4939-4948.
17. Iqbal J, Shen Y, Huang X, et al. Global microRNA expression profiling uncovers molecular markers for classification and prognosis in aggressive B-cell lymphoma. *Blood.* Feb 12 2015;125(7):1137-1145.
18. Iqbal J, Meyer PN, Smith LM, et al. BCL2 predicts survival in germinal center B-cell-like diffuse large B-cell lymphoma treated with CHOP-like therapy and rituximab. *Clinical cancer research : an official journal of the American Association for Cancer Research.* Dec 15 2011;17(24):7785-7795.
19. Johnson NA, Slack GW, Savage KJ, 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 : official journal of the American Society of Clinical Oncology.* Oct 1 2012;30(28):3452-3459.
20. Rao E, Jiang C, Ji M, et al. The miRNA-17 approximately 92 cluster mediates chemoresistance and enhances tumor growth in mantle cell lymphoma via PI3K/AKT pathway activation. *Leukemia.* May 2012;26(5):1064-1072.

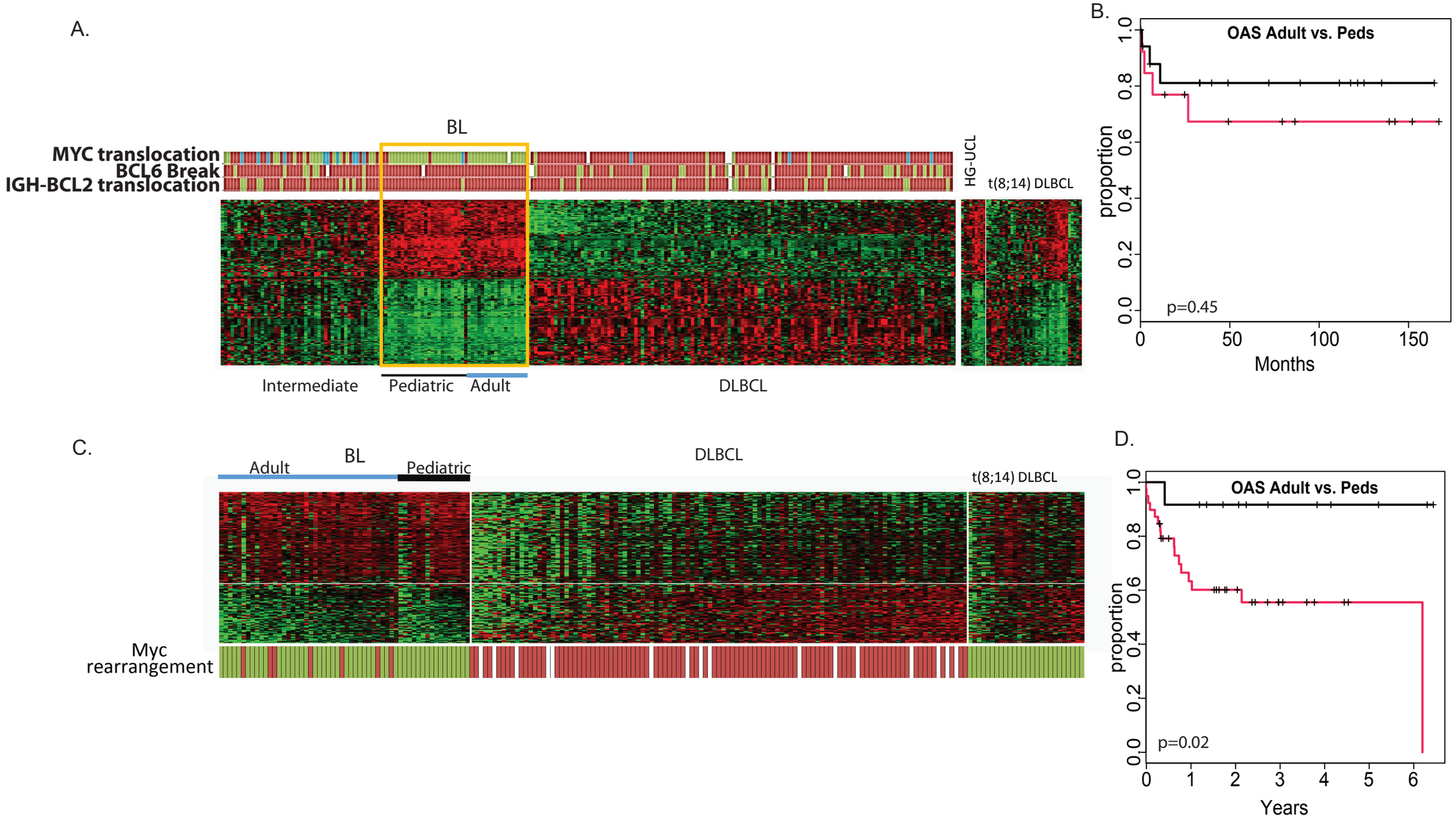


Figure S1. Validation of the Dave et al. Burkitt classifier (Dave, S.S, et al, NEJM 2006) in independently molecularly classified BL-DLBCL cohorts from GEO (http://www.ncbi.nlm.nih.gov/geo/). A. Heatmap of expression of genes in the Dave et al. Burkitt classifier (Dave, S.S, et al, NEJM 2006) in cases (GEO number GSE4475) run on the Affymetrix HG-U133A array. Cases are grouped by the molecular classification provided in GEO and are based on an independent molecular classifier (Hummel, et al NEJM 2006). MYC translocation, BCL6 translocation, and IGH-BCL2 translocation status is noted if available (green=positive, red=negative). B. Kaplan-Meier curves comparing overall survival for adult and pediatric Burkitt cases from the GSE4475 dataset. C. Heatmap of expression of genes in the Dave et al. Burkitt classifier (Dave, S.S, et al, NEJM 2006) in formalin fixed paraffin embedded cases (GEO number GSE69053) run using the Whole-Genome DASL HT platform. Cases are grouped by the molecular classification provided in GEO. MYC rearrangement status is noted if available (green=positive, red=negative).
D. Kaplan-Meier curves comparing overall survival for adult and pediatric Burkitt cases from the GSE69053 dataset.

E

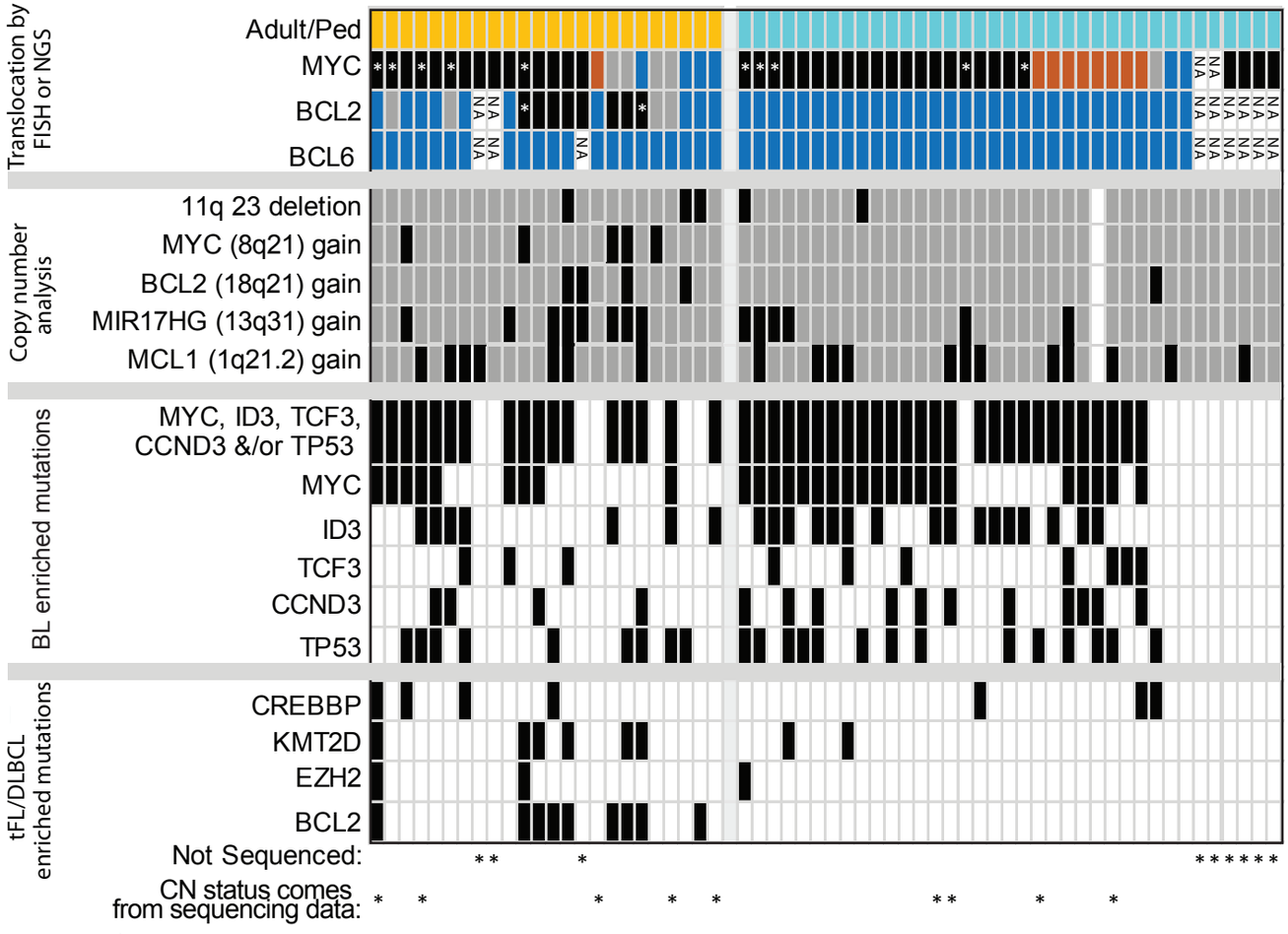
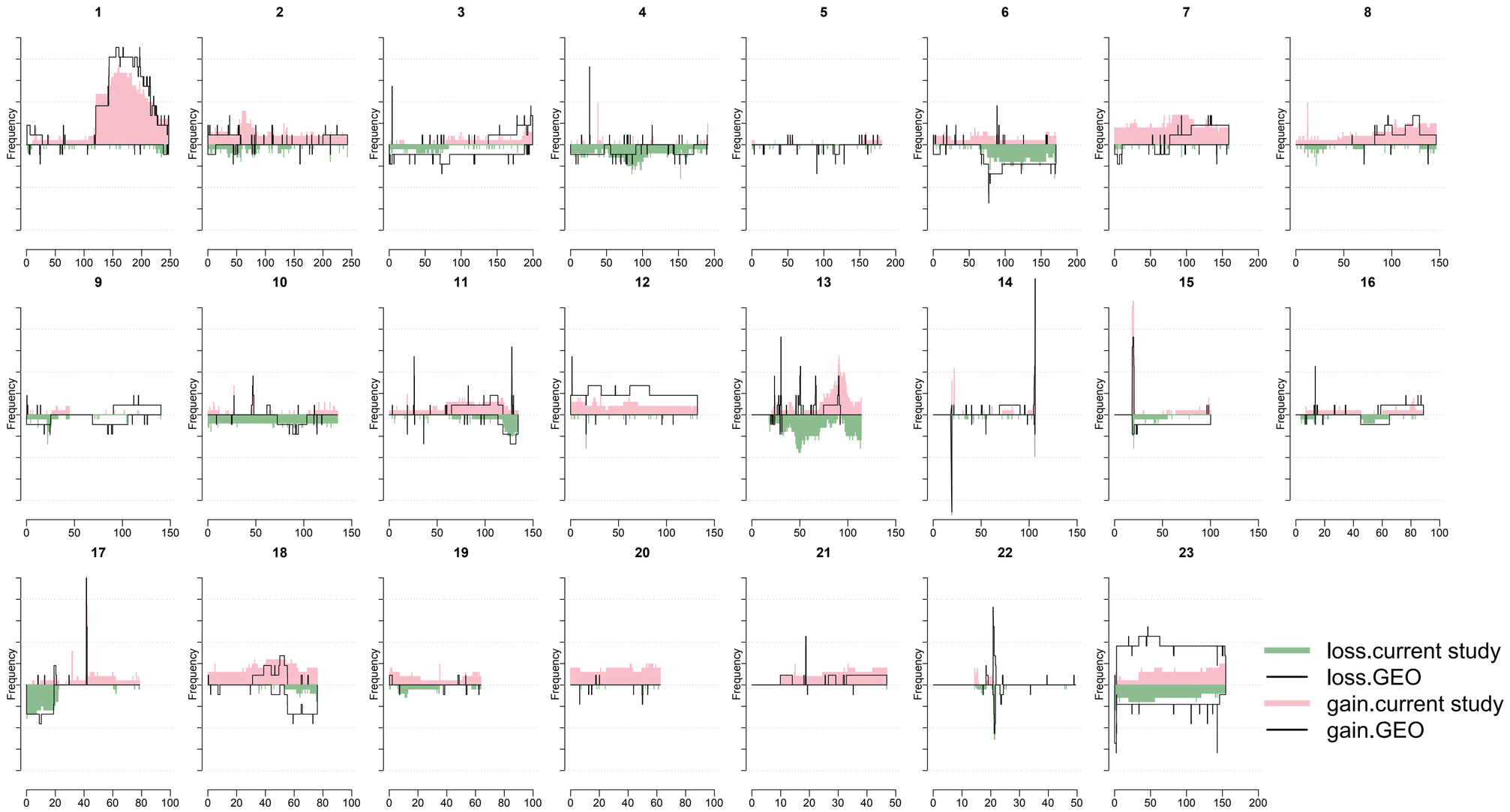


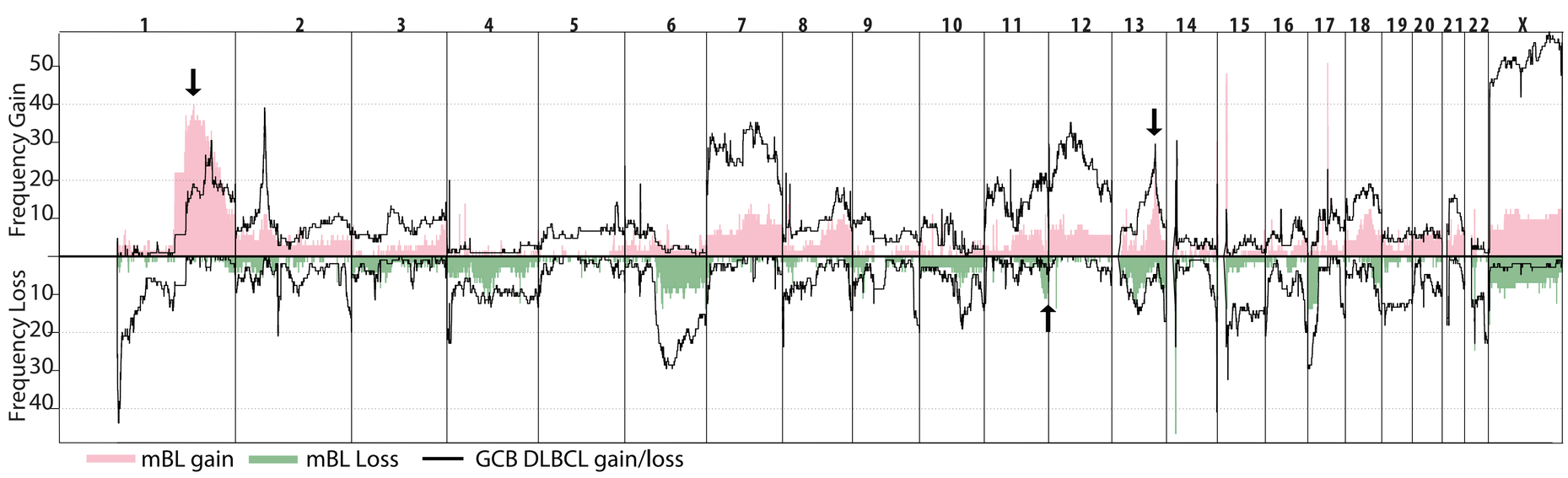
Figure S1E. Genetic characteristics of mBL cases included in the study. Myc hypermutation is defined by >10 mutations. On average in cases with known MYC translocation there were >10 coding+noncoding mutations in MYC, whereas the negative cases has ≤2 mutations. MYC, BCL2, and BCL6 translocation information was assessed by compiling available cytogenetics, FISH, and /or NGS data. NGS data was analyzed by FACTERA(1) for evidence of the listed translocations. Cases where evidence of the translocation comes only from NGS data and were not confirmed by either FISH, cytogenetics, or long range PCR (2) are noted with an asterisk. The CopywriteR analysis package (3) was used on sequencing data from cases without available 250K SNP array data to determine CN status of the listed genes/genomic regions.

1. Newman AM, Bratman SV, Stehr H, et al. FACTERA: a practical method for the discovery of genomic rearrangements at breakpoint resolution. Bioinformatics. Dec 1 2014;30(23):3390-3393.
2. Basso K, Frascella E, Zanesco L, and Rosolen, A. Improved long-distance polymerase chain reaction for the detection of t(8;14)(q24;q32) in Burkitt's lymphomas. Am J Pathol. 1999 Nov;155(5):1479-85.
3. Kuilman T, Velds A, Kemper K, Ranzani , et al. CopywriteR: DNA copy number detection from off-target sequence data. Genome Biol. 2015 Feb 27;16:49.

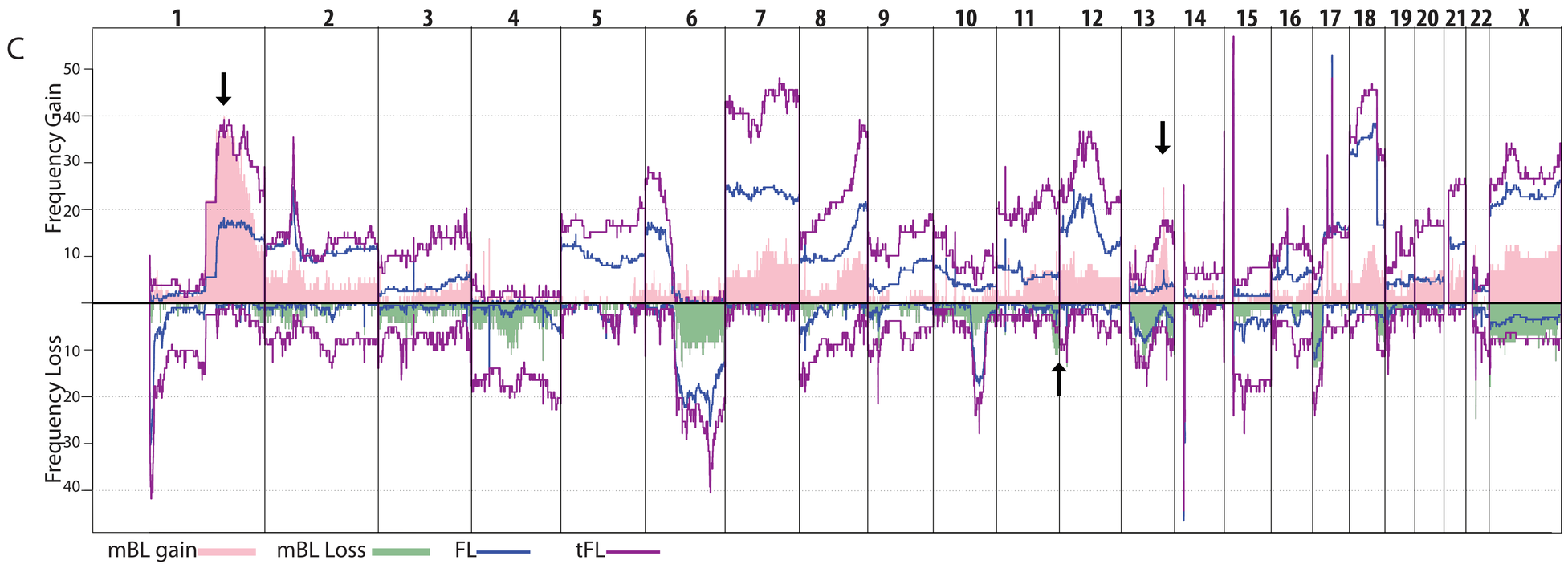
A



B



C



D

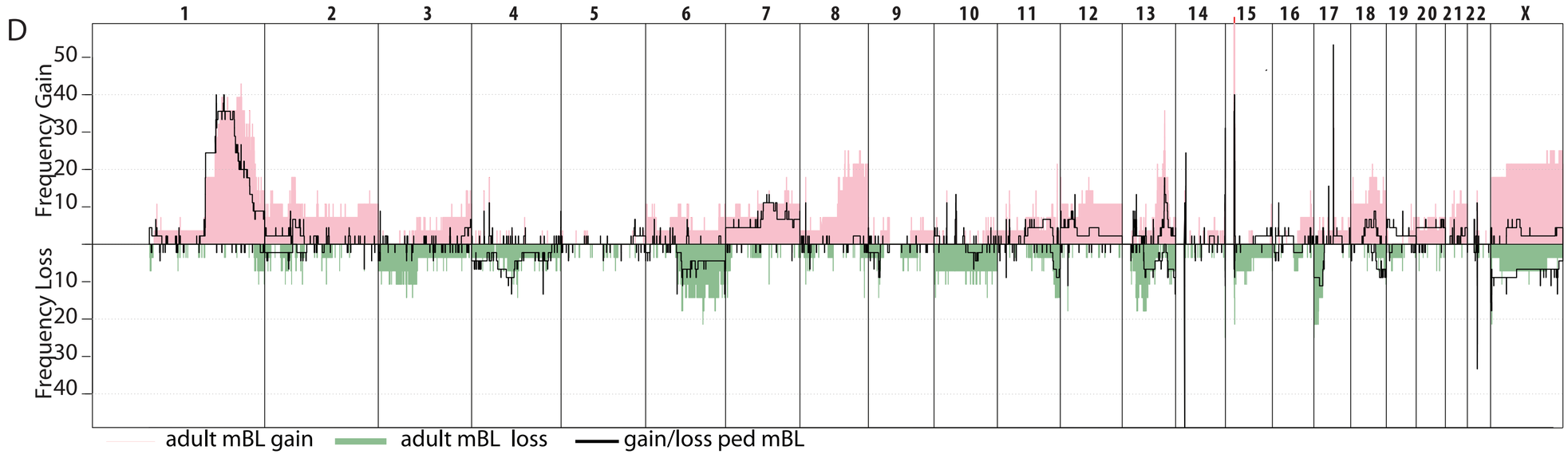
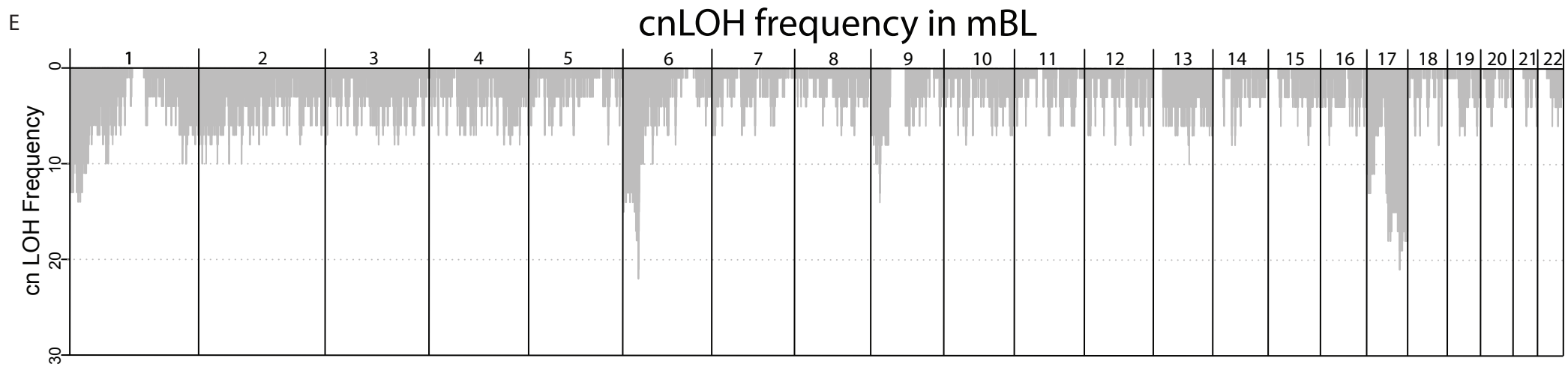
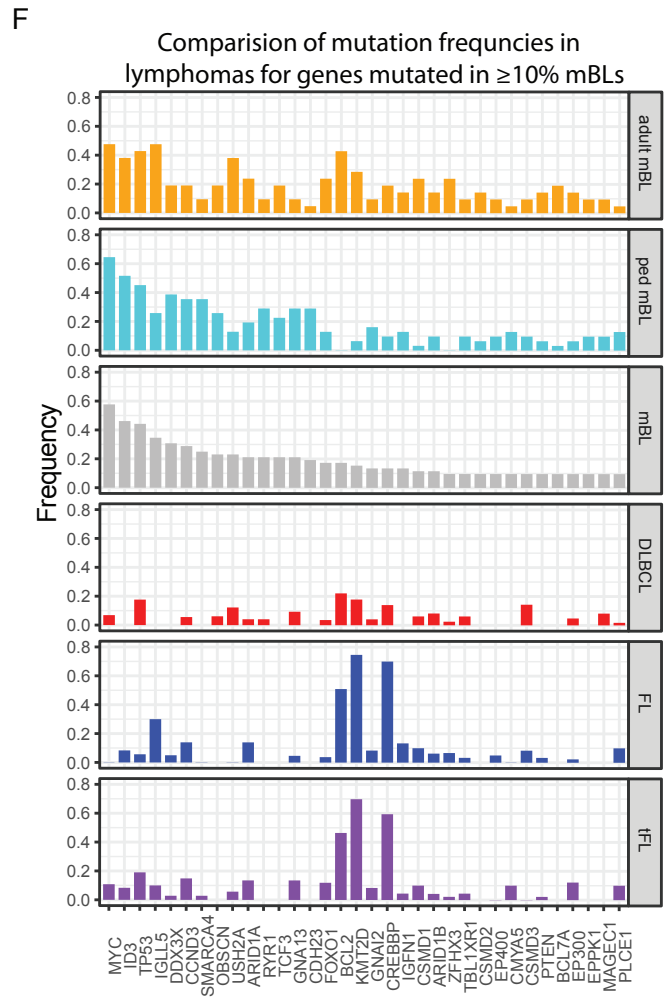


Figure S2. A. Comparison of the frequency of gains, and losses found in mBLs from this study and a GEO dataset (GSE21597). B&C. Comparison of the frequency of gains, and losses found in mBLs and GCB-DLBCL(B) or FL and tFL (C). D. Comparison of the frequency of gains, and losses found in Adult mBLs and pediatric mBLs.





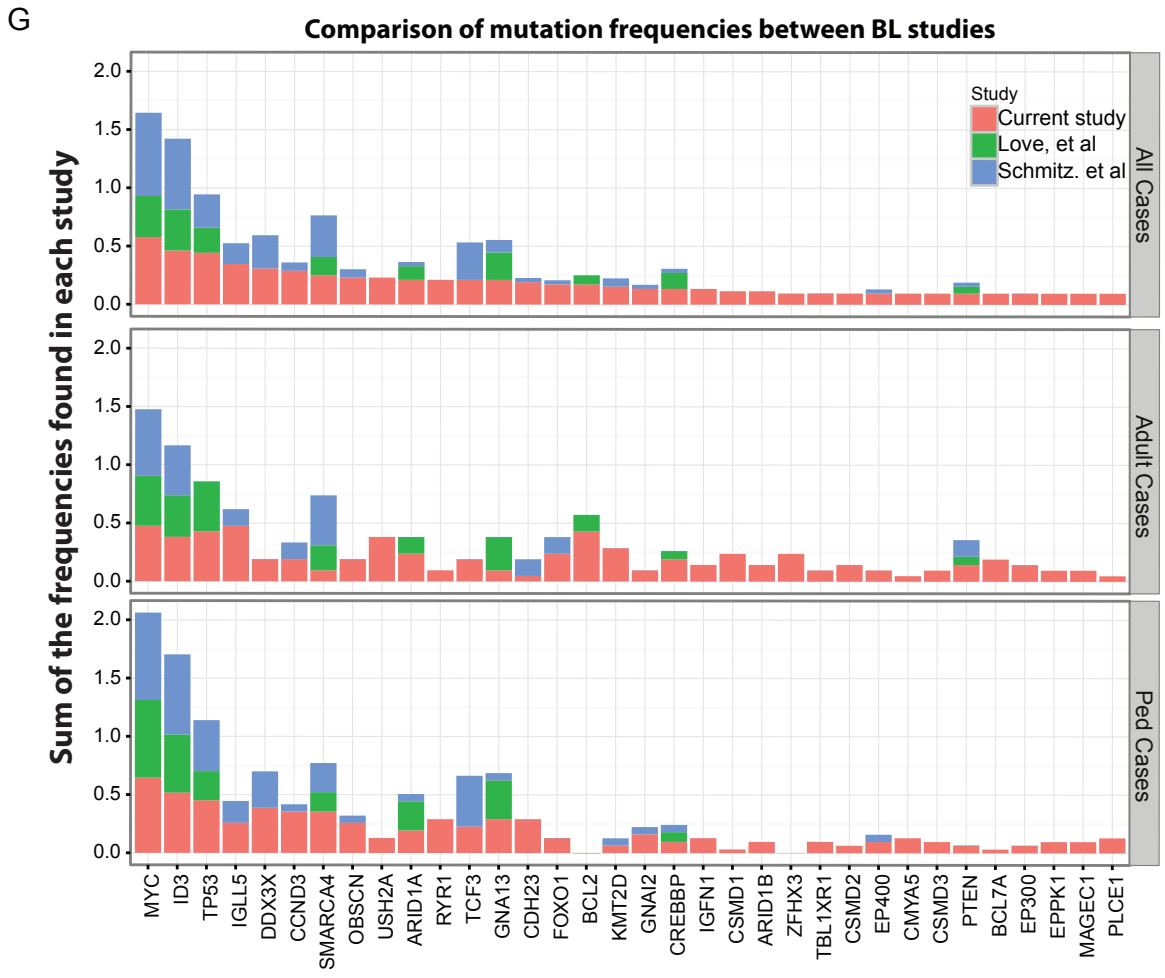


Figure S2. E. Frequency of copy neutral loss of heterozygosity in mBLs. F. Comparision of the frequency of mutation in the noted lymphomas for mutation identified in at least 10% of mBL cases in the current study. G. Comparison of the frequency of mutations found in 10% of cases in the current study with 2 previously published BL sequencing studies (PMID 23143597 & 22885699). The Y axis is the sum of the frequencies of the mutations found in each of the 3 studies.

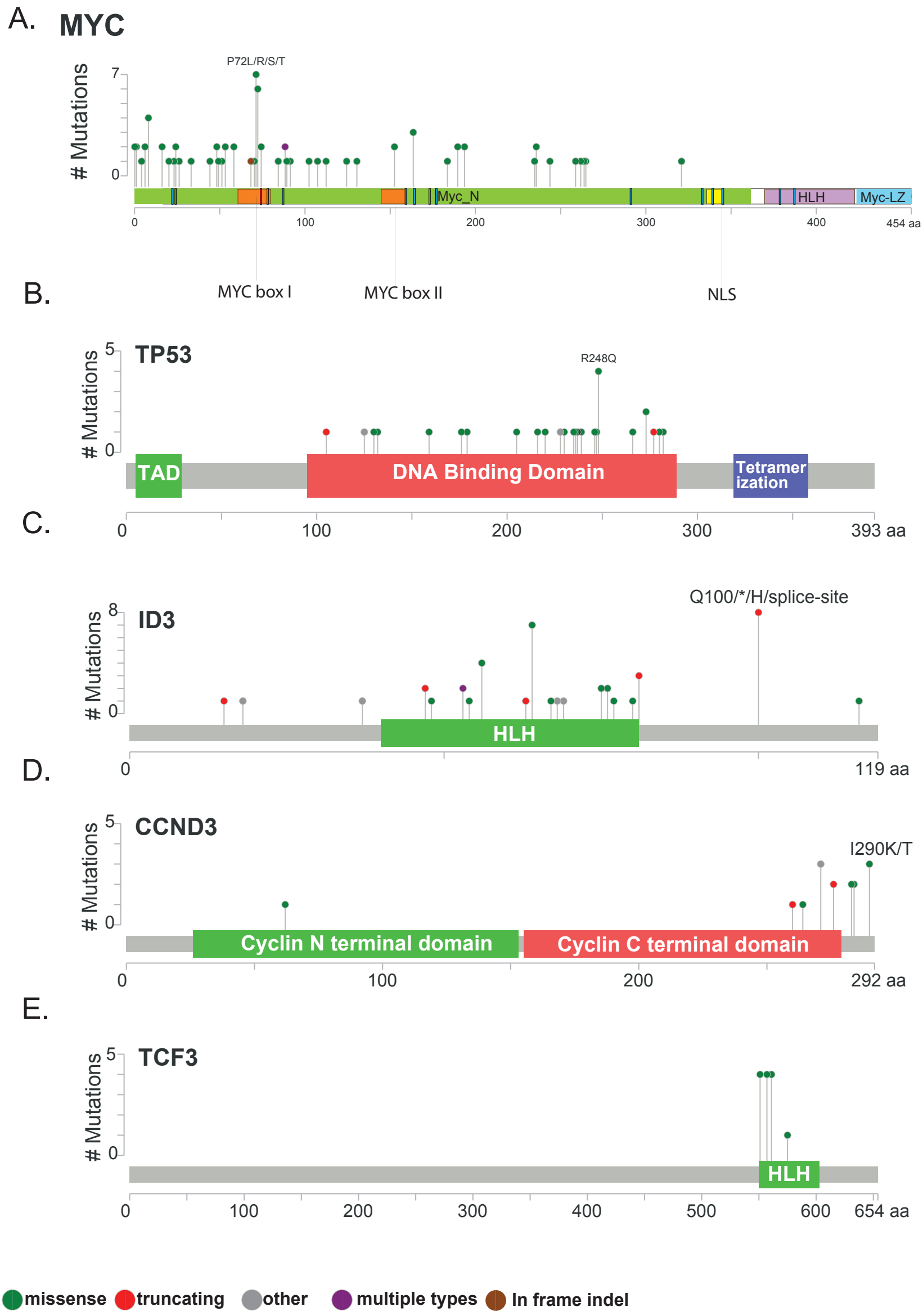
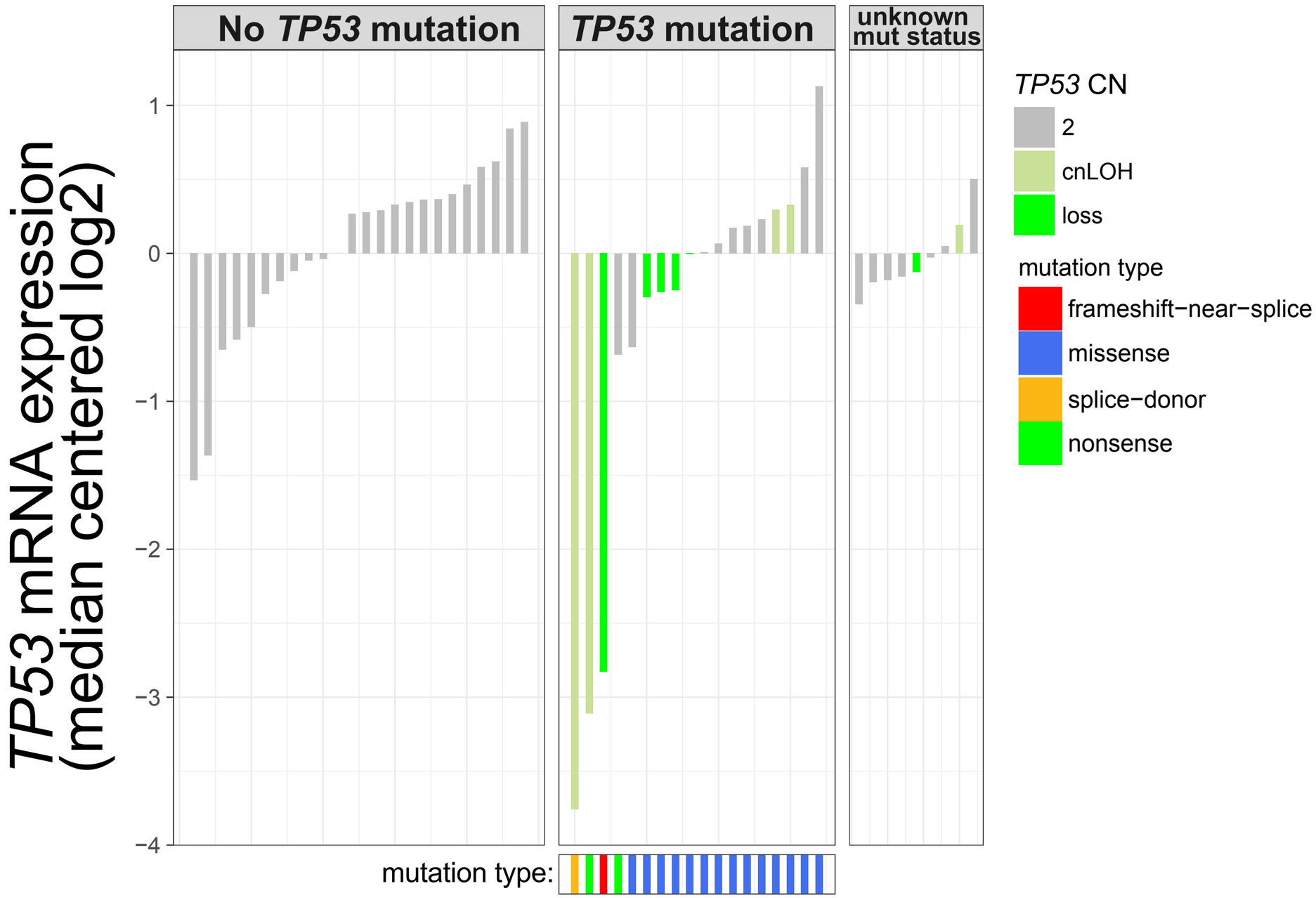


Figure S3 A-E Schematics of the location of coding mutations identified MYC, TP53, ID3, CCND3, and TCF3. Figures were generated using MutationMapper (http://www.cbioportal.org/mutation_mapper.jsp).

A



B

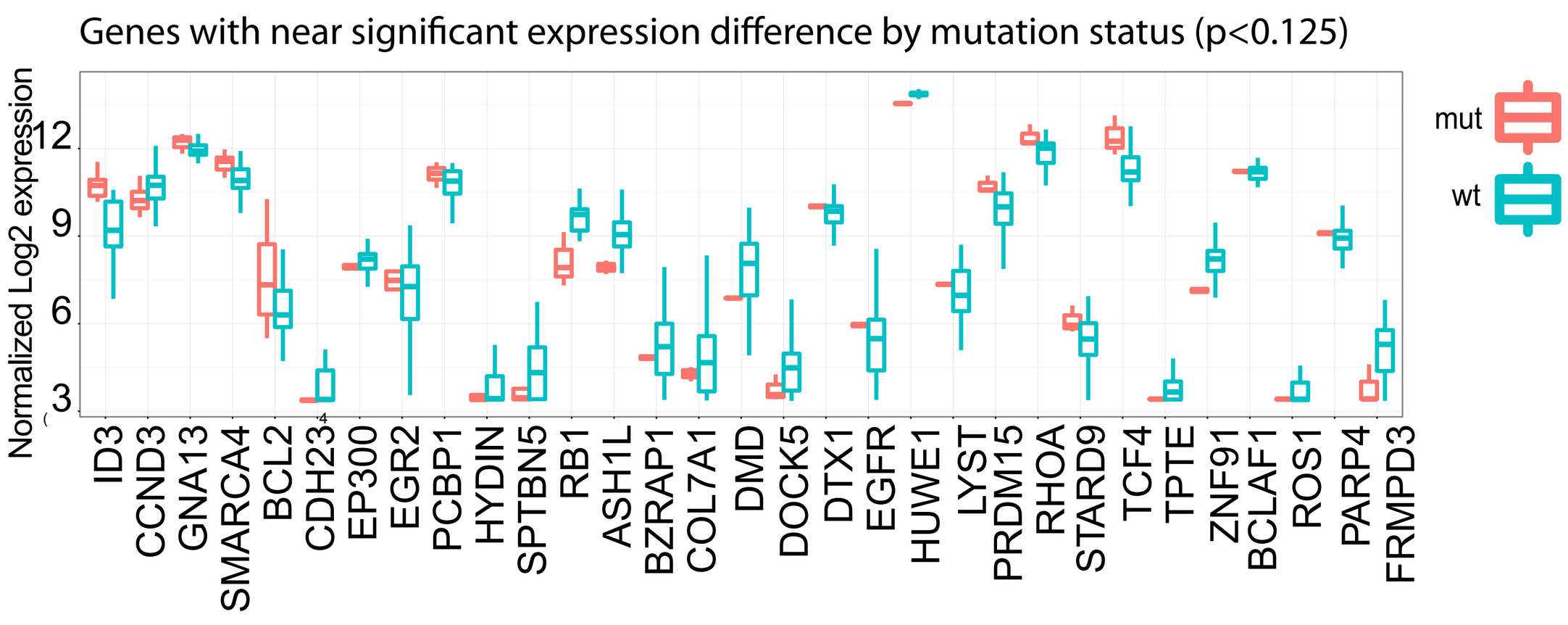
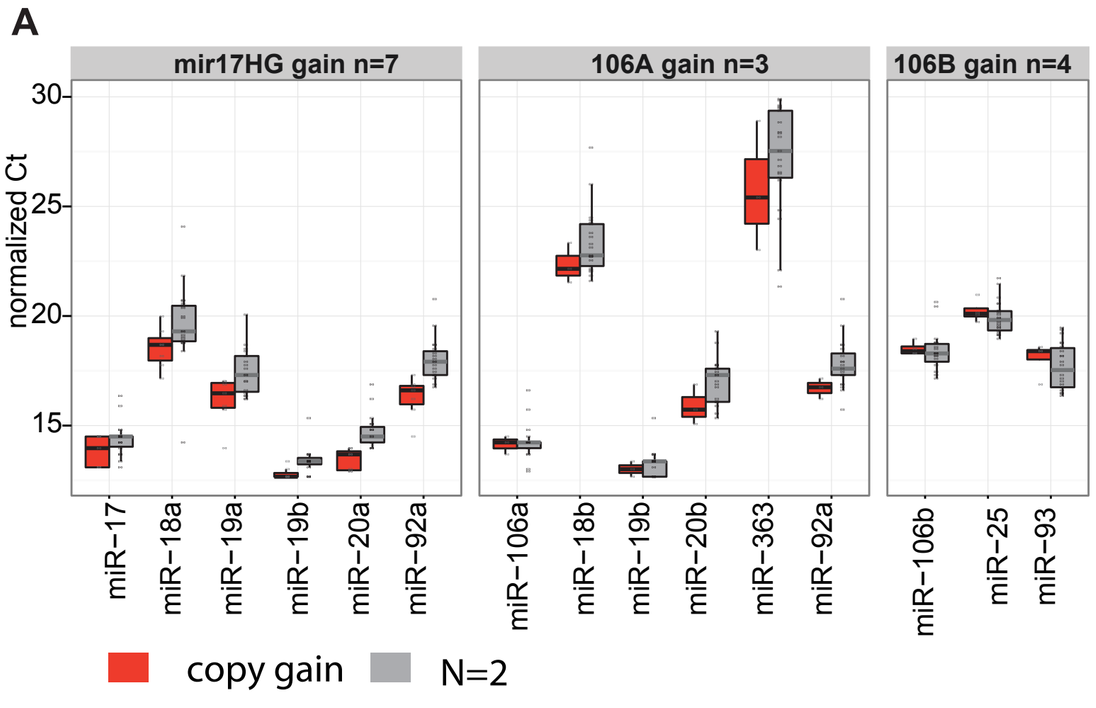
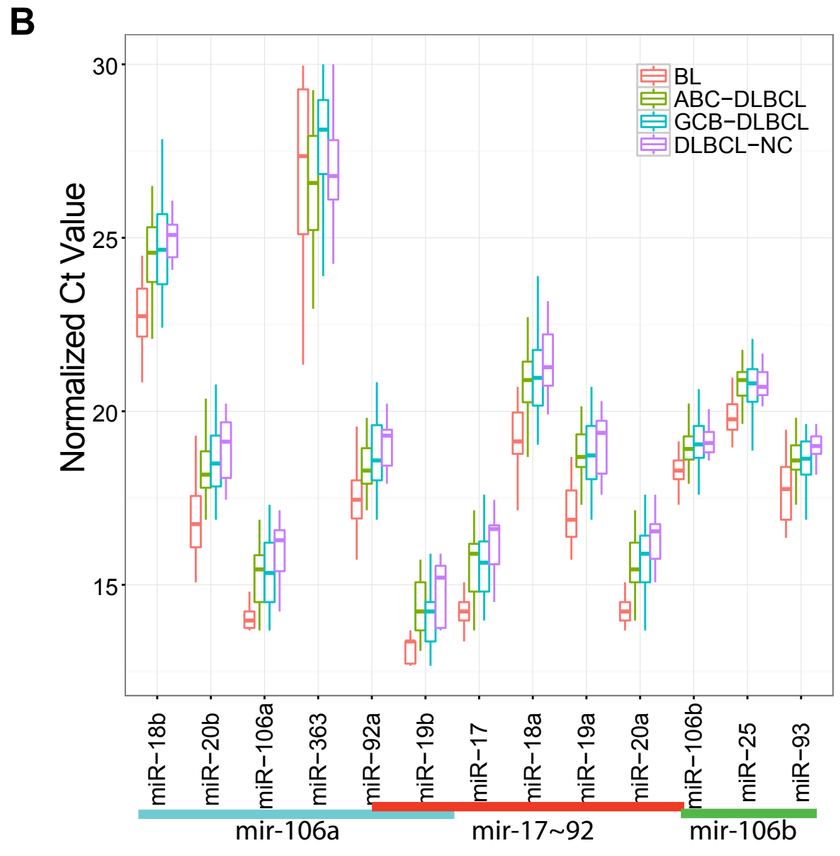


Figure S4: mRNA expression level measured by Affymetrix U133+2 array. A. *TP53* mRNA expression level by mutation and copy number abnormality status. Cases with copy loss or cnLOH has significantly lower expression of TP53 than those without mutation or CNAs (mean log2 WT=8.6, mean Loss= 7.67, p=0.04) B. Comparison of mRNA expression levels in cases with or without mutations in the noted genes.





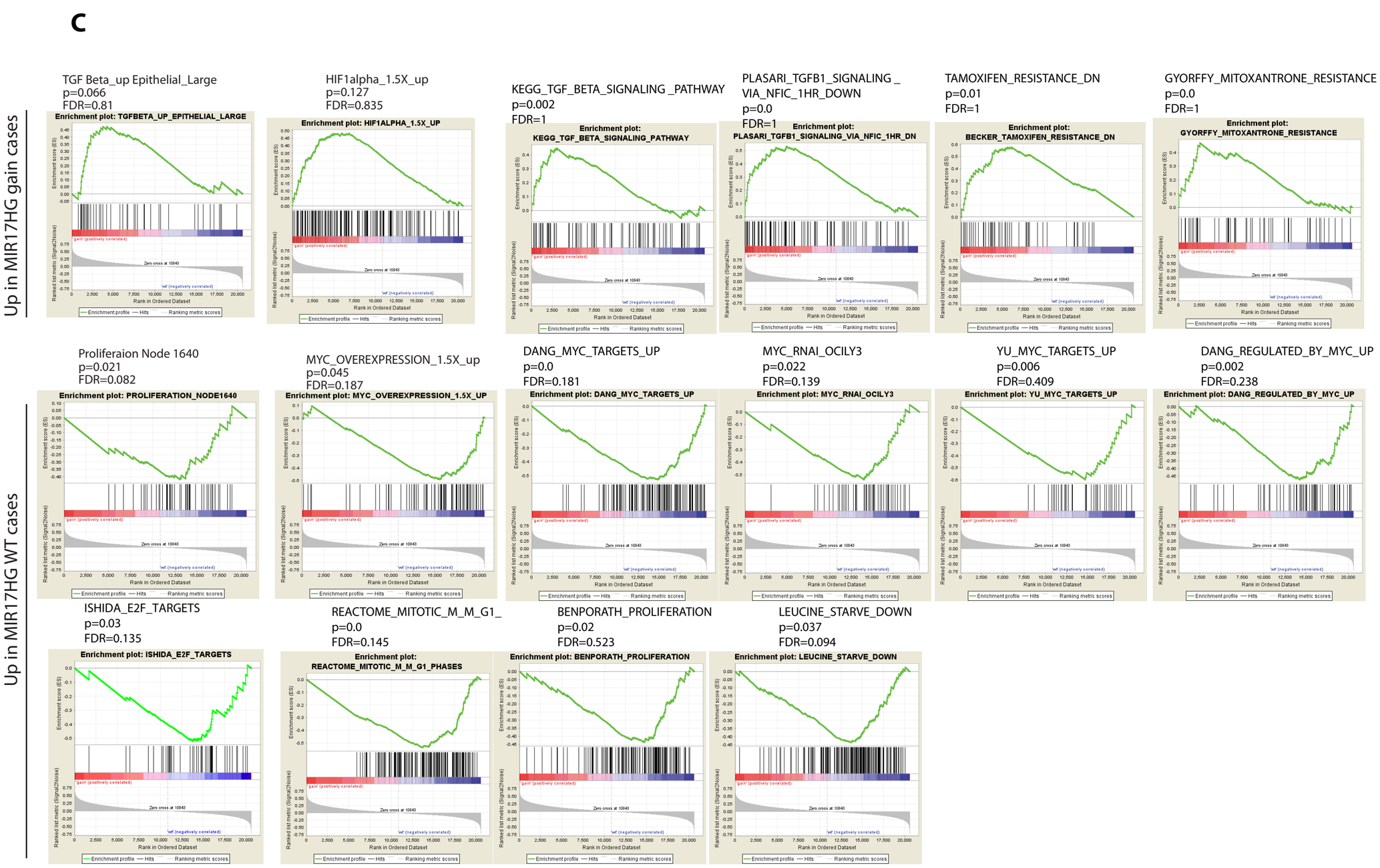
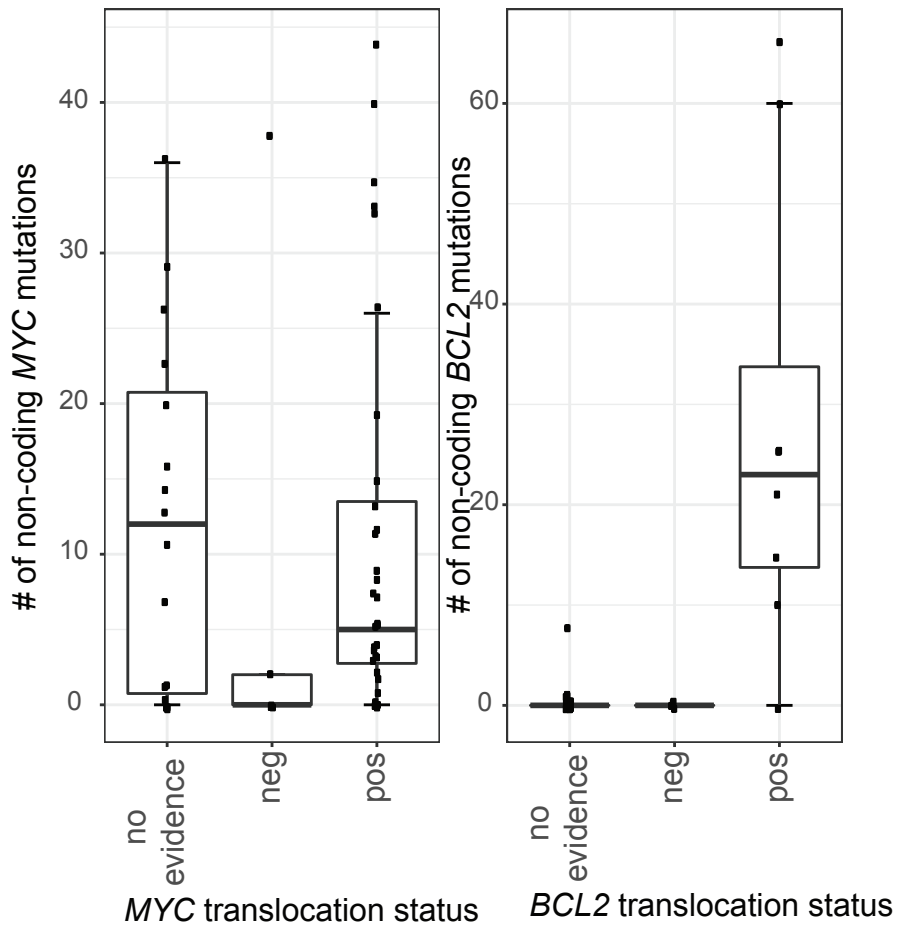


Figure S5. A. Boxplot of normalized Ct values for microRNAs in the MIR17~92, MIR106A, and MIR106B clusters by copy number status of the cluster. Low Ct values indicate high expression. B. Boxplot of normalized Ct values for microRNAs in the MIR17~92, MIR106A, and MIR106 B clusters in BL and DLBCL cases. Low Ct values indicate high expression. C. Gene set enrichment analysis results for MIR17HG gain vs MIR17HG wt cases.

Figure S6. Bar graphs of the fold increase in expression of genes along chromosome that are significantly up-regulated in cases with DNA copy gains (A, C, F) or fold decrease in expression of genes along the chromosome that are significantly downregulated in cases with DNA copy losses (B, D-E). Frequency of gains/losses are denoted by the orange line.

A



B

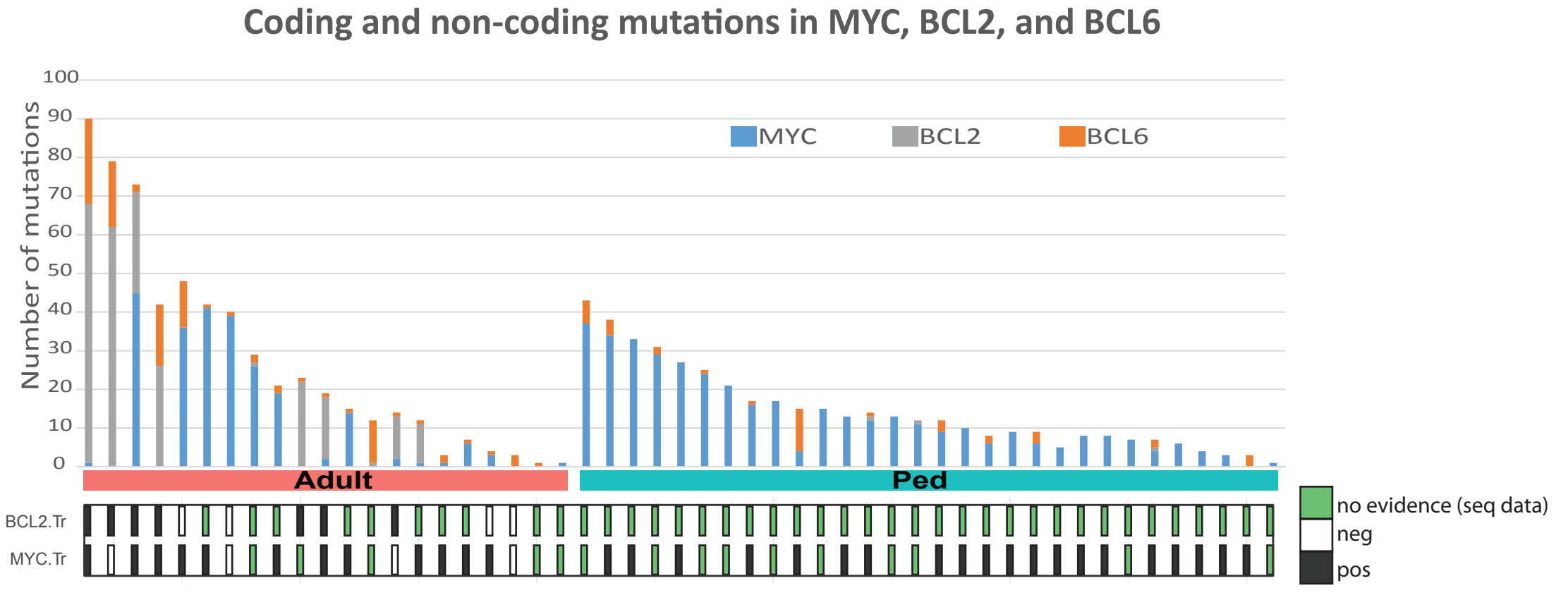
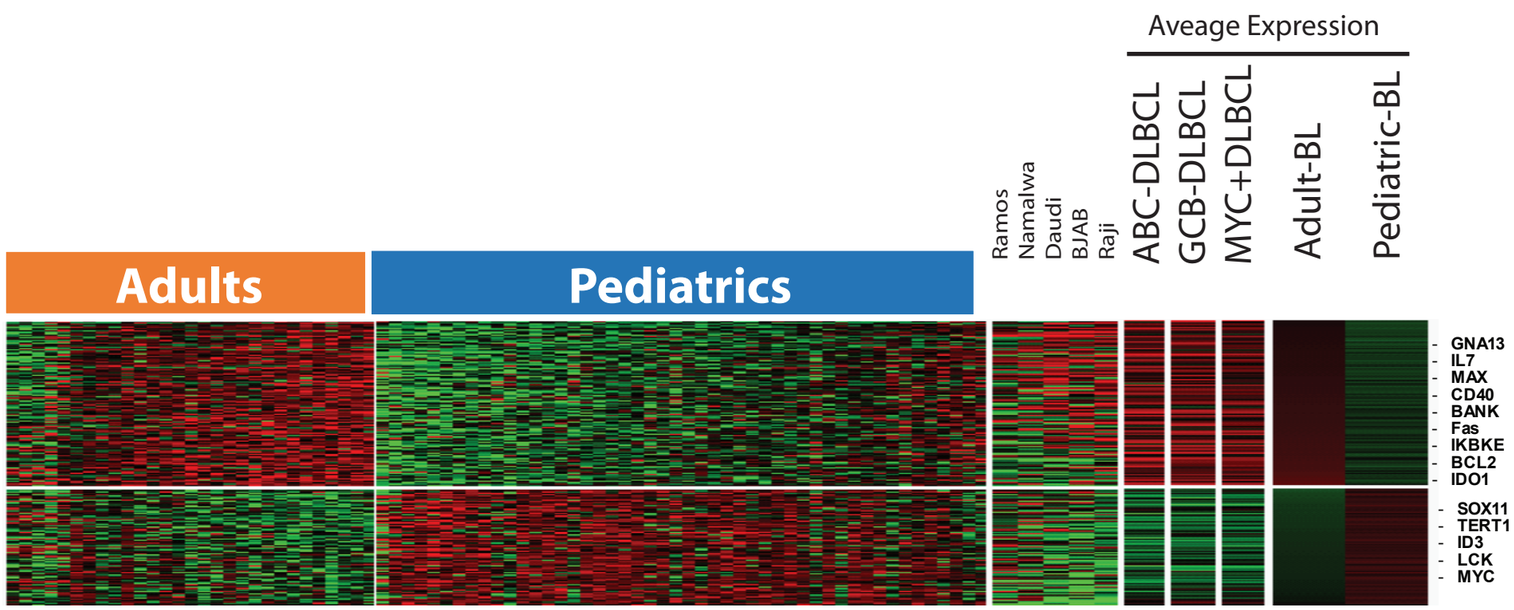


Figure S7. A. Boxplot of non-coding mutations by *MYC* or *BCL2* translocation status. Cases that no translocation was detected in the sequencing data are noted as no evidence. B.Number of coding and non-coding mutations in *MYC, BCL2, and BCL6.*

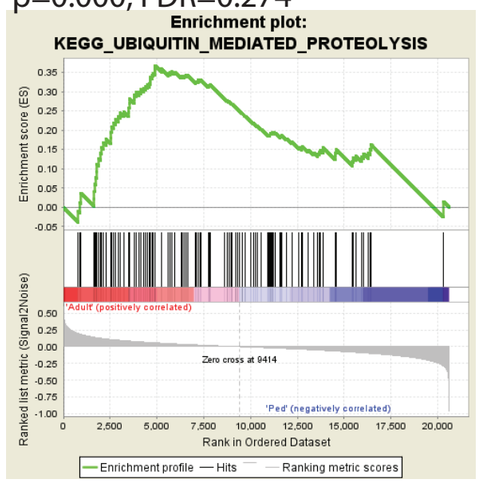
A



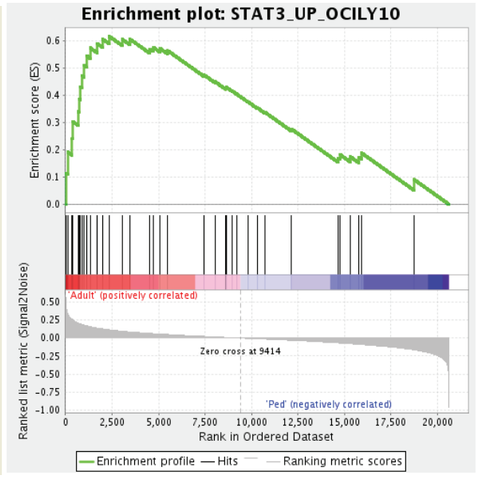
B

Enriched in Adults

KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS
p=0.000, FDR=0.274

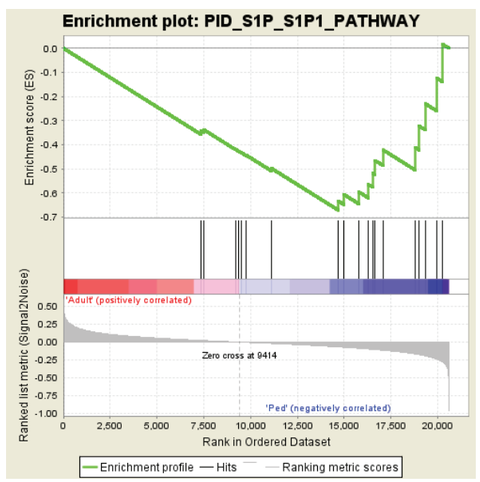


STAT3_UP_OCILY10
p=0.022
FDR= 0.229



Enriched in Peds

PID_S1P_S1P1_PATHWAY
p=0.000
FDR= 1



MYC_OVEREXPRESSION_2X_UP
p=0.04
FDR=0.88

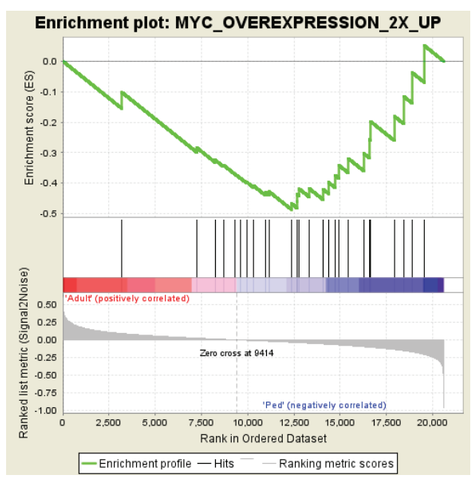


Figure S8: A : Heatmap of genes differentially expressed in adults vs peds (p=0.01, 2x difference). B. Signatures enriched in adult or pediatric mBLs.

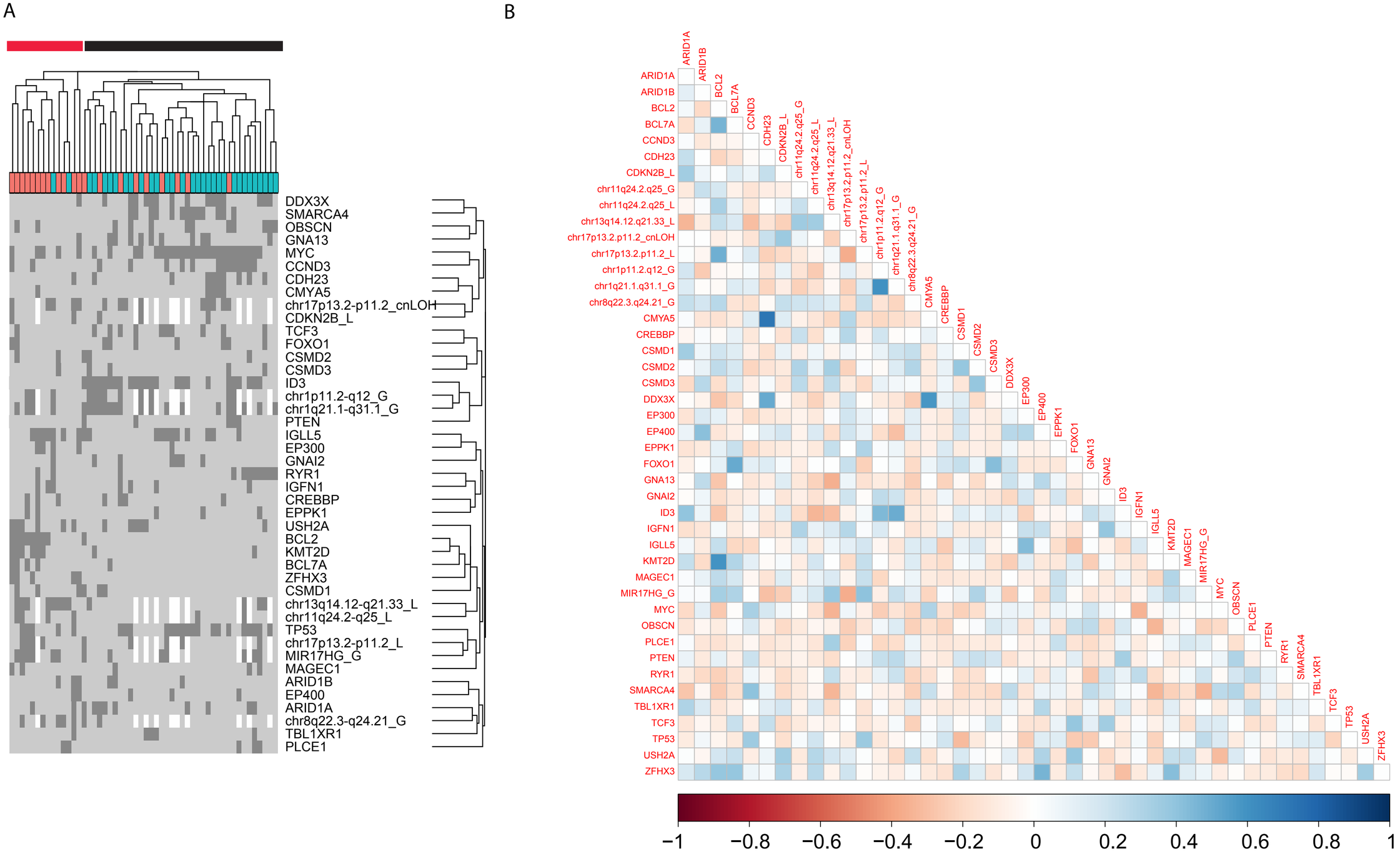
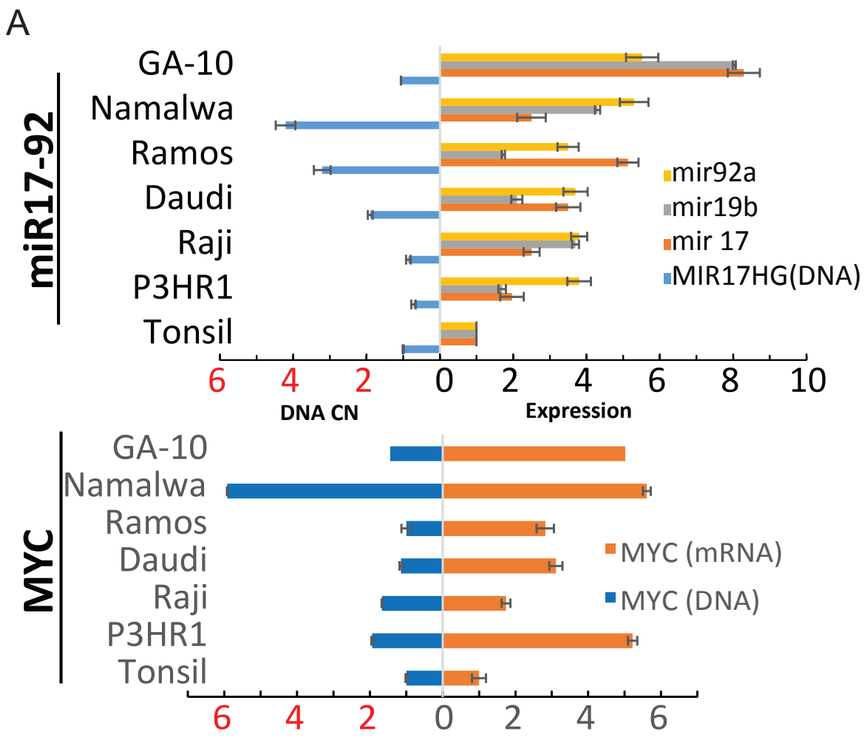
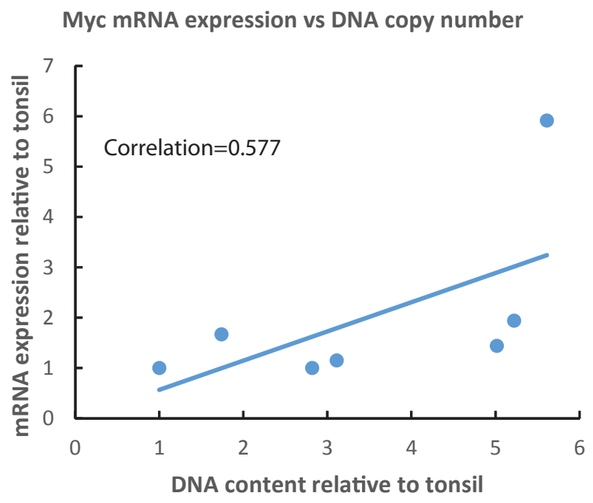
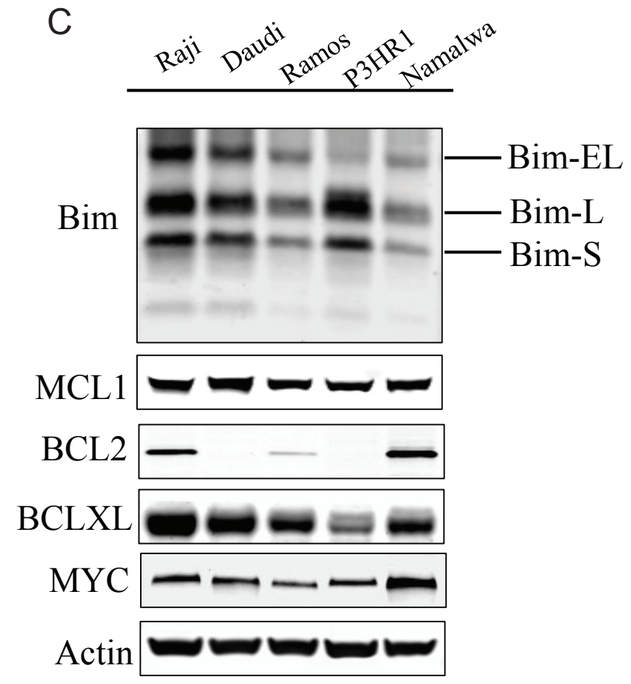
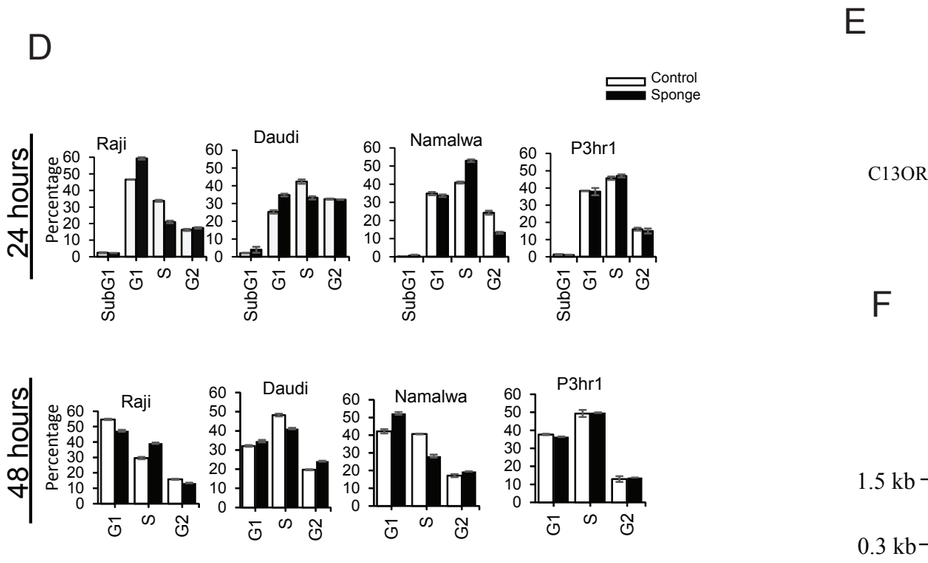


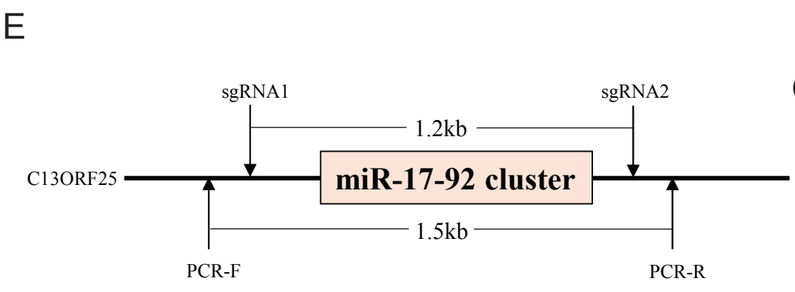
Figure S9. A. Clustering of mutations and CNAs found in >/=5 cases is dominaged by age group (centered correlation, average linkage). B. Heatmap of pearson correlation coefficients.

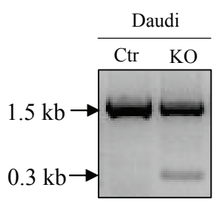


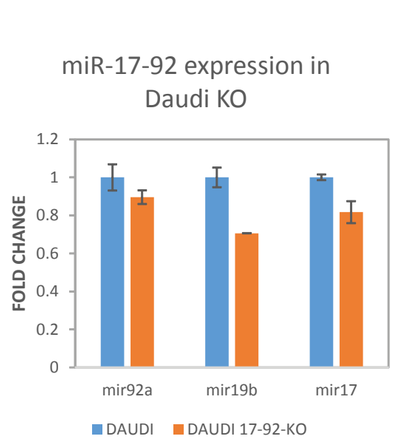


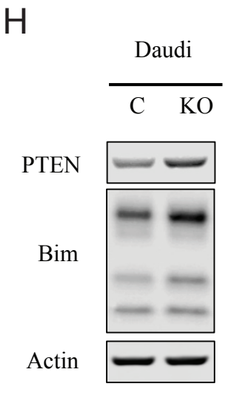


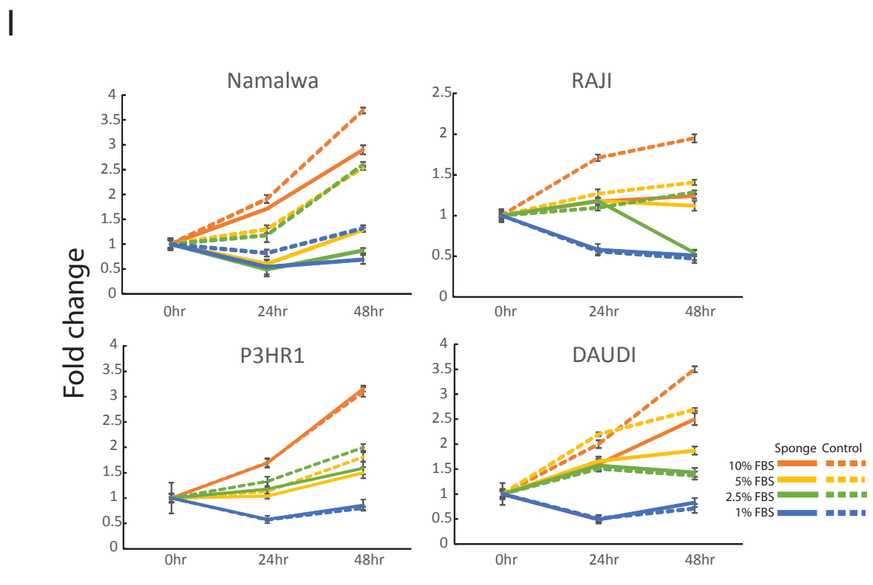


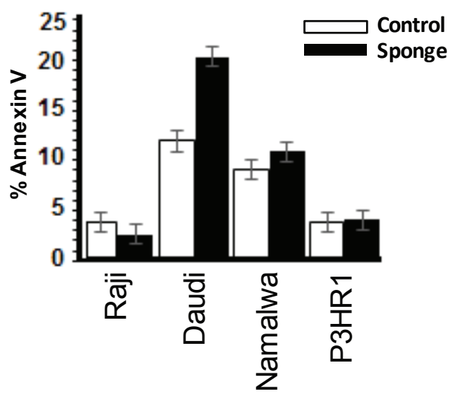


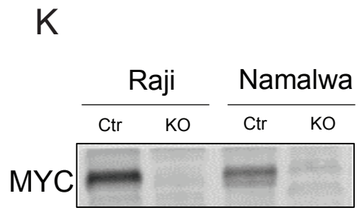


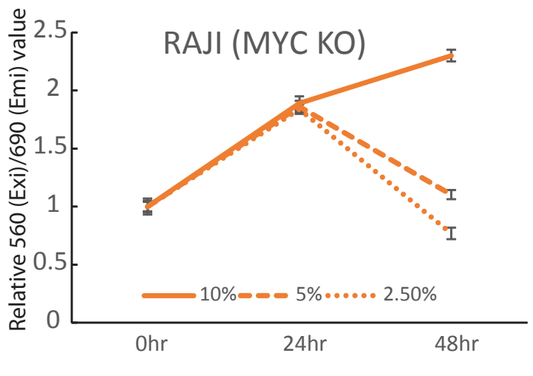












M

698-704 of CD22 3' UTR

5’ GUCACUUAUUCAGCAUUGCACAA 3’

3’ AGUCAAAACGUAUCUAAACGUGU 5’

miR-19a

198-204 of FCGR2B 3' UTR

5’ CCCAAAUAACCGACUGCACCUUC 3’

3’ GAUAGACGUGAUCUACGUGGAAU 5’

miR-18a

974-980 of FCRL4 3' UTR

5’ UAUGAGAAUAAUUCUUGCAAUAA 3’

3’ UGUCCGGCCCUGUUCACGUUAU 5’

miR-92a

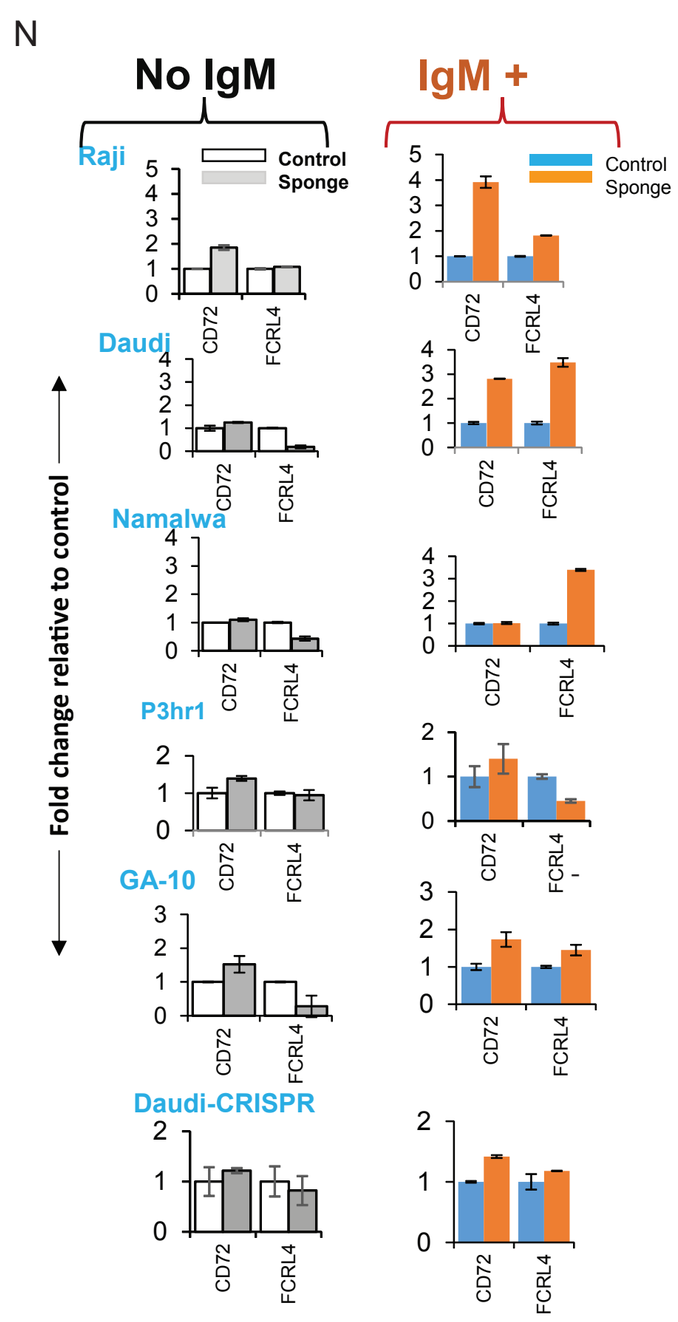
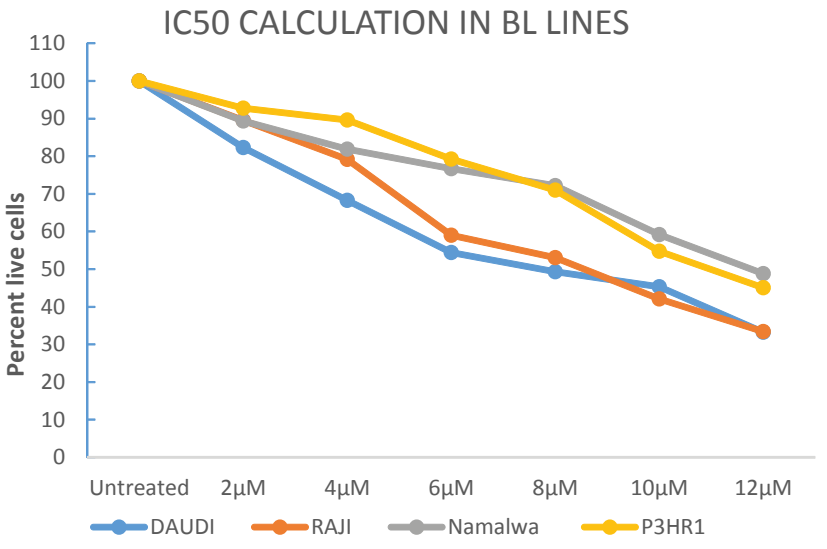


Figure S10 A. Quantification of DNA copy number and expression of MIR17HG/mir-17~92 cluster members and MYC in BL cell lines. MYC DNA and MYC mRNA expression was positively in BL lines (r=0.699). MYC mRNA expression and the expression of miRi19b (r=0.1) and miR-92a (r=0.664) were positively correlated, while as negative correlation was found between MYC mRNA and miR-17 (r=-0.448). Bars represent the mean of 3 replicates from one representative experiment of at least 3 separate experiments and error bars represent the standard error of the mean. B. Comparison of MYC mRNA expression to MYC DNA copy number in Raji, Daudi, Namalwa, Ramos, P3HR1, and GA-10 cells. C. Whole cell lysate was western blotted for BIM, MCL1, BCL2, BCLXL,MYC, and Actin expression in the noted BL cell lines D) Cell cycle analysis of BL cell lines 24 and 48 hours after induction of sponge expression. Bars represent the mean of 3 biological replicates from one representative experiment of 3 separate experiments and error bars represent the standard error of the mean. E) Schematic of CRISPR/Cas9 targeting strategy for knock out of the mir17-92 cluster. F) Copy loss of the MIR17HG locus was confirmed by PCR. G) qRT-PCR quantification of miR92a, miR19b, and miR17 expression in Daudi and Daudi cells with CRISPR-mediated knockout of MIR17HG. H) Whole cell lysate from Daudi parental. (C) or MIR17HG KO (KO) cells was western blotted for PTEN, BIM, and Actin expression. I) Effect of Serum concentration on growth of miR17-92 sponge-expressing cells. J) Apoptosis measured by Annexin-V staining in miR17~92 sponge transduced cell lines compared to vector control. Bars represent the mean of 3 biological replicates from one representative experiment of 3 separate experiments and error bars represent the standard error of the mean. K) Whole cell lysate was western blotted for MYC and Actin expression control BL cell lines or cell lines with Crisper mediated KO of MYC. L) Effect of Serum concentration on growth of Raji-MYC-KO cells. M) Schematic depicting miR17~92 member binding location in the 3'-UTRs of CD22, FCGR2B, and FCRL4. N) Comparison of mRNA levels of CD72 and FCRL4 in control (vector or parental) or miR17~92 sponge or knockout (Daudi-CRISPR) cells with or without IgM stimulation. Fold change was calculated relative to the vector or parental controls. Graph depicts one representative experiment of 3 separate experiments. Error bars depict one standard deviation of 3 technical replicates.

A



B

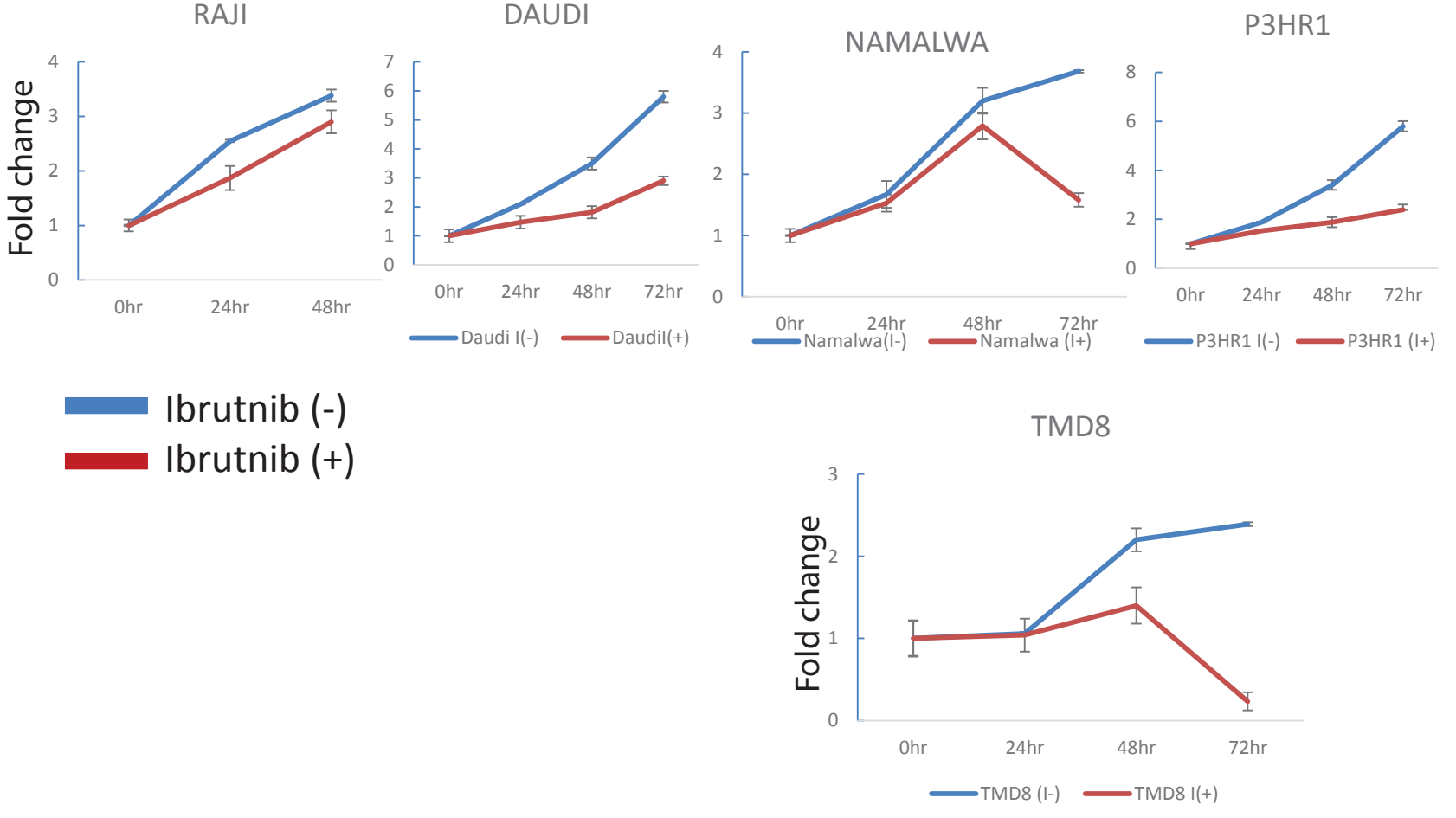


Figure S11. A. IC50 of Ibrutinib. B. Growth curve of cell lines upon treatment with Ibrutinib at theIC50 concentration.