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

Rabea Wagener,[1] Susanne Bens,[1] Umut H. Toprak,[2,3,4] Julian Seufert,[2,4] Cristina López,[1] Ingrid Scholz,[5] Heidi Herbrueggen,[6] Ilske Oschlies,[7] Stephan Stilgenbauer,[8] Matthias Schlesner,[2] Wolfram Klapper,[7] Birgit Burkhardt[6] and Reiner Siebert[1]

[1]Institute of Human Genetics, Ulm University and Ulm University Medical Center, Ulm; [2]German Cancer Research Center (DKFZ), Bioinformatics and Omics Data Analytics, Heidelberg; [3]German Caner Research Center (DKFZ), Division of Neuroblastoma Genomics Heidelberg; [4]Faculty of Biosciences, Heidelberg University, Heidelberg; [5]Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg; [6]Department of Pediatric Hematology and Oncology, NHL-BFM Study Center, University Children's Hospital, Münster; [7]Hematopathology Section, Christian-Albrechts University, Kiel and [8]Department of Internal Medicine III, University of Ulm, Ulm, Germany

Correspondence: REINER SIEBERT - reiner.siebert@uni-ulm.de
doi:10.3324/haematol.2018.208140

Supplementary Methods

Immunohistochemistry

Immunohistochemical analysis was performed on the cervical lymph node biopsy which was formalin fixed and embedded in paraffin (FFPE). The following antibodies have been applied for immunohistochemistry and in situ hybridization using a Leica/Bond automated stainer: BCL2 (Diagnostic Biosystems, clone 100/D5), CD10 (Novocastra, clone 56C6), CD20 (DAKO, clone L26), c-MYC (abcam, clone Y69), Ki67 (NeoMarkers, clone SP6), TDT (Novocastra, clone, SEN28), EBER-in situ hybridization (Leica).

Molecular cytogenetics

FISH analyses were performed as previously described [1,2] using FFPE tissue of the cervical lymph node biopsy. The probes Vysis LSI IGH/MYC/CEP8 Tri Color Dual Fusion Probe as well as the Vysis LSI Dual Color Break Apart Rearrangement Probes for MYC, BCL6 and BCL2 were applied which were obtained from Abbott Molecular Diagnostics. Furthermore, for the detection of IGK-*MYC* and IGL-*MYC* translocations as well as for the detection of an 11q aberration, typical for Burkitt-like lymphoma with 11q aberration, non-commercial probes were used. Evaluation of slides was performed using a Zeiss fluorescence microscope equipped with appropriate filter sets. For each probe at least 100 nuclei were examined whenever possible. Digital image acquisition, processing, and evaluation was performed using ISIS FISH Imaging System 5.8 (MetaSystems).

DNA extraction

The DNA extracted from the fresh frozen bone marrow biopsy using the FlexiGene DNA Kit (Qiagen) according to manufacturers´ instructions was used for whole genome sequencing, validation of the *MYC*-insertion as well as *MYC* and *ID3* mutations. In addition, we extracted DNA from the cervical lymph node fixed in paraffin using the QiAamp DNA FFPE Tissue Kit (Qiagen) according to manufacturers´ instructions. This DNA was used for IGHV gene sequence analysis, OncoScan analysis, screening for *ID3* as well as *MYC* mutations and validation of the *MYC* insertion into the IGH locus by Sanger sequencing. Using STR analysis applying the StemElite ID System (Promega), we confirmed that the DNA of both biopsies belong to the same patient.

IGHV gene sequence analysis

The IGHV gene sequence analysis was performed as described by Kröber *et* al [3].

Chromosomal imbalance profiling

Copy number analysis was performed using the OncoScan™ CNV FFPE Assay Kit according to standard protocols (Affymetrix) as recently published [4]. Briefly, chromosomal alterations were evaluated and visually inspected using OncoScan Console 1.3 (Affymetrix) and Nexus Express for OncoScan 3 (Biodiscovery). Human reference genome was GRCh37/hg19. Chromosomal imbalances encompassing at least 20 informative probes and having a minimum size of 100 kb as well as CNN-LOH larger than 5 Mb were considered informative.

Whole genome sequencing and data processing

Whole genome sequencing (WGS) of the bone marrow biospy was performed using protocols, pipelines and data analyses tools previously established and applied in framework of the ICGC MMML-Seq project [5,6]. The sequencing of the library was performed using an Illumina HiSeq X Ten sequencing instrument (Illumina). Alignment and single nucleotide variant/indel calling was performed as recently described [6].

The *MYC*-insertion into the IGH locus was detected using the SOPHIA algorithm (manuscript in preparation) by a workflow previously described in Sahm et al [7] using the genome sequencing data. For details refer to López *et* al. [8].

Mutation and breakpoint validation by Sanger sequencing

The breakpoints of the *MYC* insertion within the IGH gene locus identified by WGS were validated using Sanger sequencing with primers flanking the breakpoints (Supplementary Table S2). In addition, primers were designed (Supplementary Table S3) to cover the coding regions of the *ID3* gene and the second exon of the *MYC* gene which was screened for the occurrence of mutations. PCR products were subjected to Sanger sequencing using the Big Dye Terminator v1.1 Cycle Sequencing Kit (Life Technologies) and sequence analysis was performed using an ABI PRISM 3130 Genetic Analyzer.

Determining the allelic status of the *ID3* mutations

To determine the allelic status of the *ID3* mutations in the lymphoma sample of the lymph node, we cloned the PCR products using Strata Clone™ PCR Cloning kit (Agilent Technologies), and sequenced the colonies by Sanger sequencing.

Supplementary Tables

Supplementary Table S1

Mutations identified in the *ID3* and *MYC* genes by whole genome sequencing as well as by verification and screening by Sanger sequencing (refer to Supplementary Figure S3) on the lymphoma biopsies from the bone marrow and the cervical lymph node. It has to be considered that mutations with a VAF below 10% might be missed by Sanger sequencing due to technical considerations.

Gene/ transcript	Genomic position (hg19)	Gene region	Exonic classification	Consequence on protein level	dbSNP (150) ID	Status by WGS	VAF based on WGS	Detected by Sanger sequencing	
								BM sample	LN sample
MYC ENST00000259523.6	chr8:128750456 A>C	intronic	-	-	-	mut	16.5%§	yes	yes
	chr8:128750475 T>C	intronic	-	-	-	mut	13.4%§	yes	yes
	chr8:128750519 A>G	exonic	missense	p.N4S	-	mut	18.4%§	yes	yes
	chr8:128750543 A>C	exonic	missense	p.Y12S	-	mut	20.2%§	yes	yes
	chr8:128750725 T>C	exonic	missense	p.S73P	-	mut	13.9%§	yes	yes
	chr8:128750747 T>C	exonic	missense	p.F80S	-	mut	7.5%§	yes	no
	chr8:128751015 C>A	exonic	missense	p.S169R	-	mut	12.5%§	yes	no
	chr8:128751183 C>T	exonic	synonymous	p.S225S	rs757865894	mut	12.4%§	yes	yes
	chr8:128751251 C>G	exonic	missense	p.T248S	-	mut	7.2%§	no	yes
	chr8:128751317 C>A	intronic	-	-	rs771286365	mut	6.3%§	n.a.	n.a.
ID3 ENST0000374561	chr1:23885774 G>C	exonic	stopgain	p.Y48X	-	wt	0%*	no	yes
	chr1:23885752 G>A	exonic	missense	p.P56S	-	wt	2.4%§	no	yes

BM: bone marrow; LN: lymph node; WGS: whole genome sequencing; mut: mutation found in more than 5% of reads; wt: wildtype nucleotide, or less than 5% of reads carrying alteration; n.a.: not analyzed by Sanger sequencing; VAF based on WGS: variant allele frequency based on whole-genome sequencing considering the amount of alleles carrying the alteration with respect to the overall read count at the nucleotide position. * The lack of this *ID3* mutation within the BM sample suggests this mutation to be of somatic origin. § The low VAF of these mutations indicates them to be of somatic origin.

Supplementary Table S2

PCR primers and PCR conditions for Sanger sequencing of breakpoints of the *MYC* gene insertion into the IGH locus.

BP	Gene	First base primer binds to (hg19)	Forward primer (5'-3') Reverse primer (5'-3')	Annealing temperature	Length PCR product
BP1	IGH	chr14:106,177,027	CAGGTCCAGCTTAGCCCAG	55°C	331 bp
	MYC	chr8:128,750,454	CGGGAGGCAGTCTTGAGTTA		
BP2	*MYC*	chr8: 128,758,759	ATCCATCTCCAGAAACGGCT	60°C	263 bp
	IGH	chr14: 106,177,246	CTGCCCTGGGTTGAGCTG		

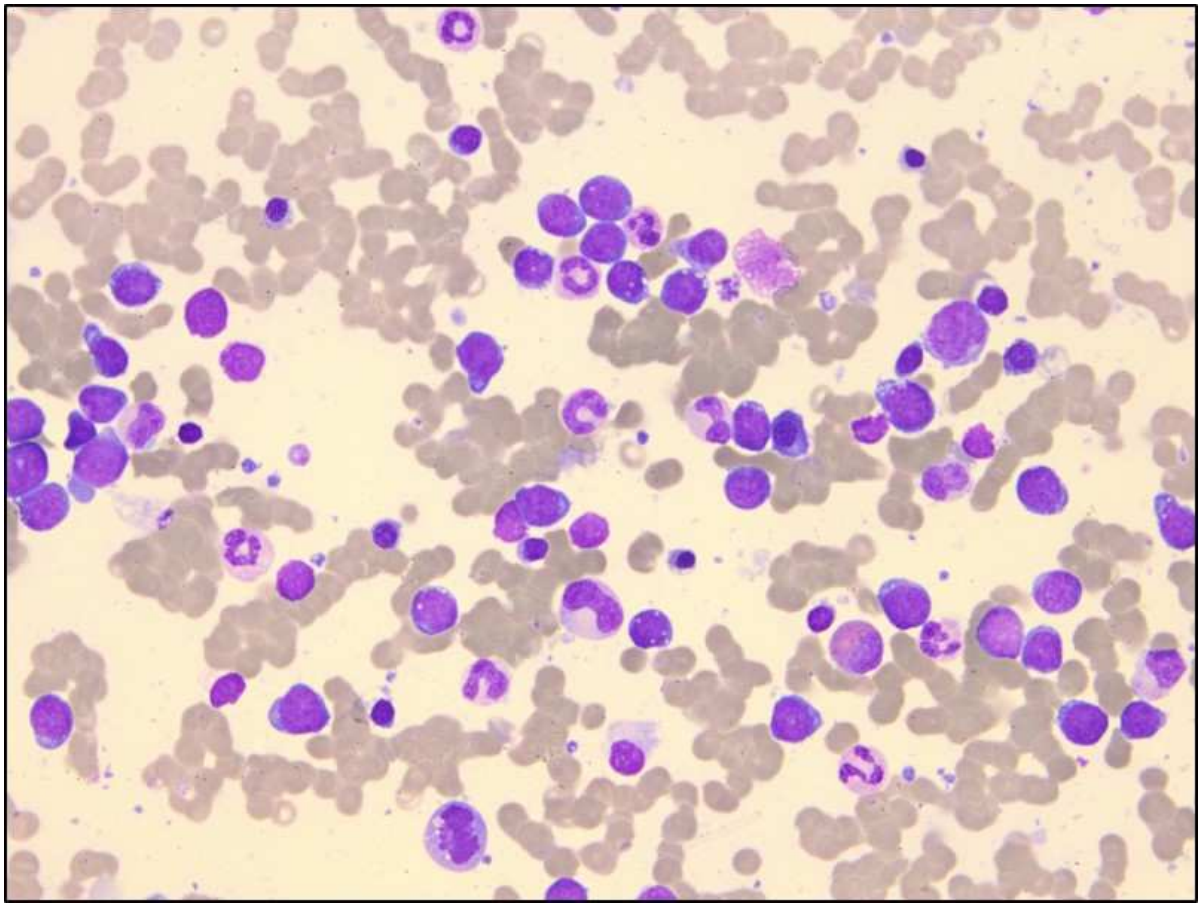
BP: breakpoint

Supplementary Table S3

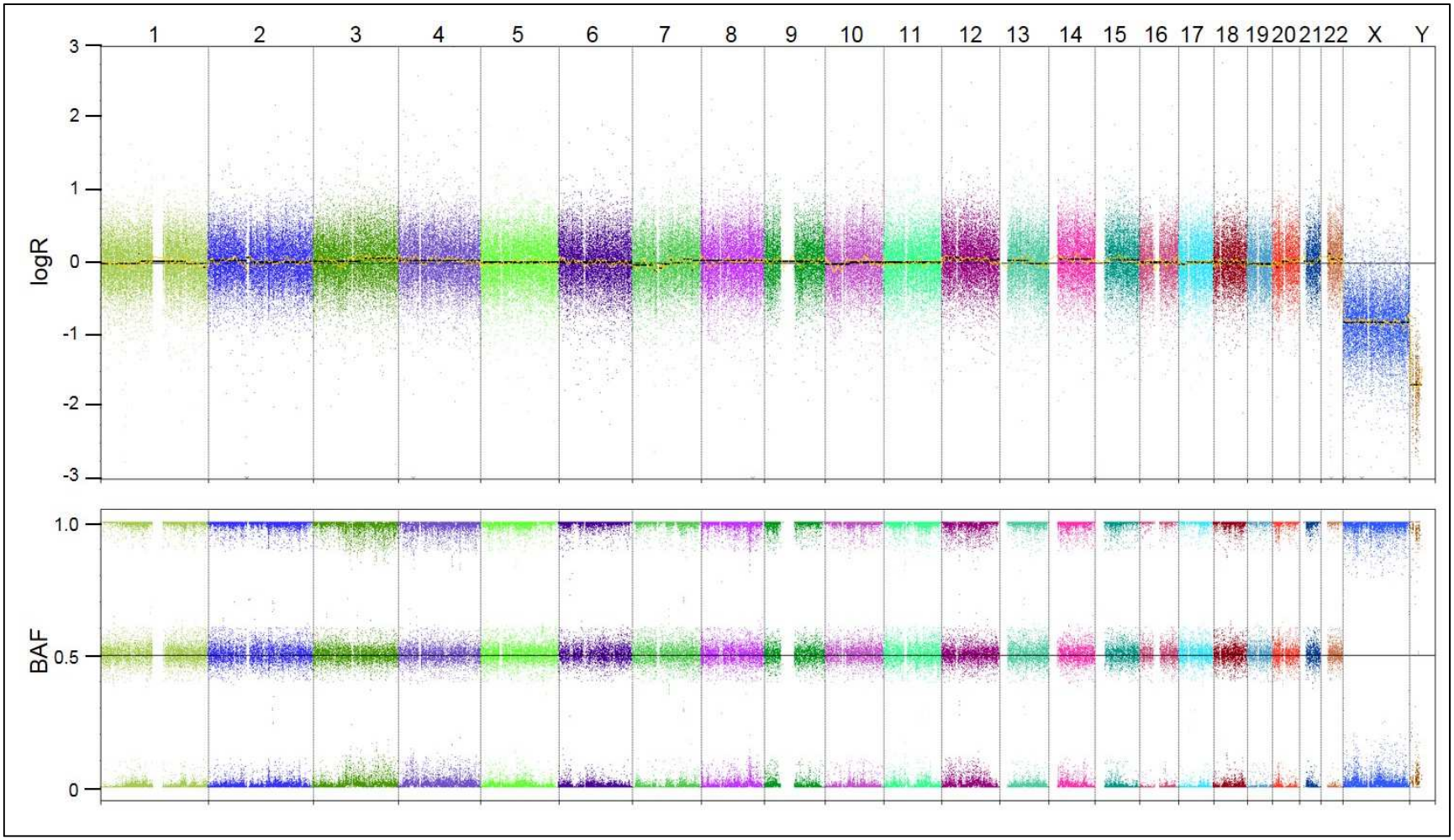
PCR primers and PCR conditions for verification of mutations within the *ID3* (NM_002167.4) and *MYC* (NM_002467) genes by Sanger sequencing.

Region	Genomic region (hg19) in bp	Forward primer (5'-3') Reverse primer (5'-3')	Annealing temperature	Length PCR product
ID3 Exon 1	chr1:23,885,690-23,886,034	TCCAGGCAGGCTCTATAAGTG GTAGTCGATGACGCGCTGTA	65°C	345 bp
ID3 Exon 1-2	chr1:23,8885,341–23,885,797	CTGGACGACATGAACCACTG CCGAGTGAGTGGCAATTTTTC	55°C	457 bp
MYC Exon 2.1	chr8:128,750,430-128,750,810	CAGGTTTCCGCACCAAGAC ATCTCCAGCTGGTCGGC	60°C	381 bp
MYC Exon 2.2	chr8:128,750,946-128,751,321	AGAGAAGCTGGCCTCCTACC AATGGGAAAGGTATCCAGCC	60°C	376 bp

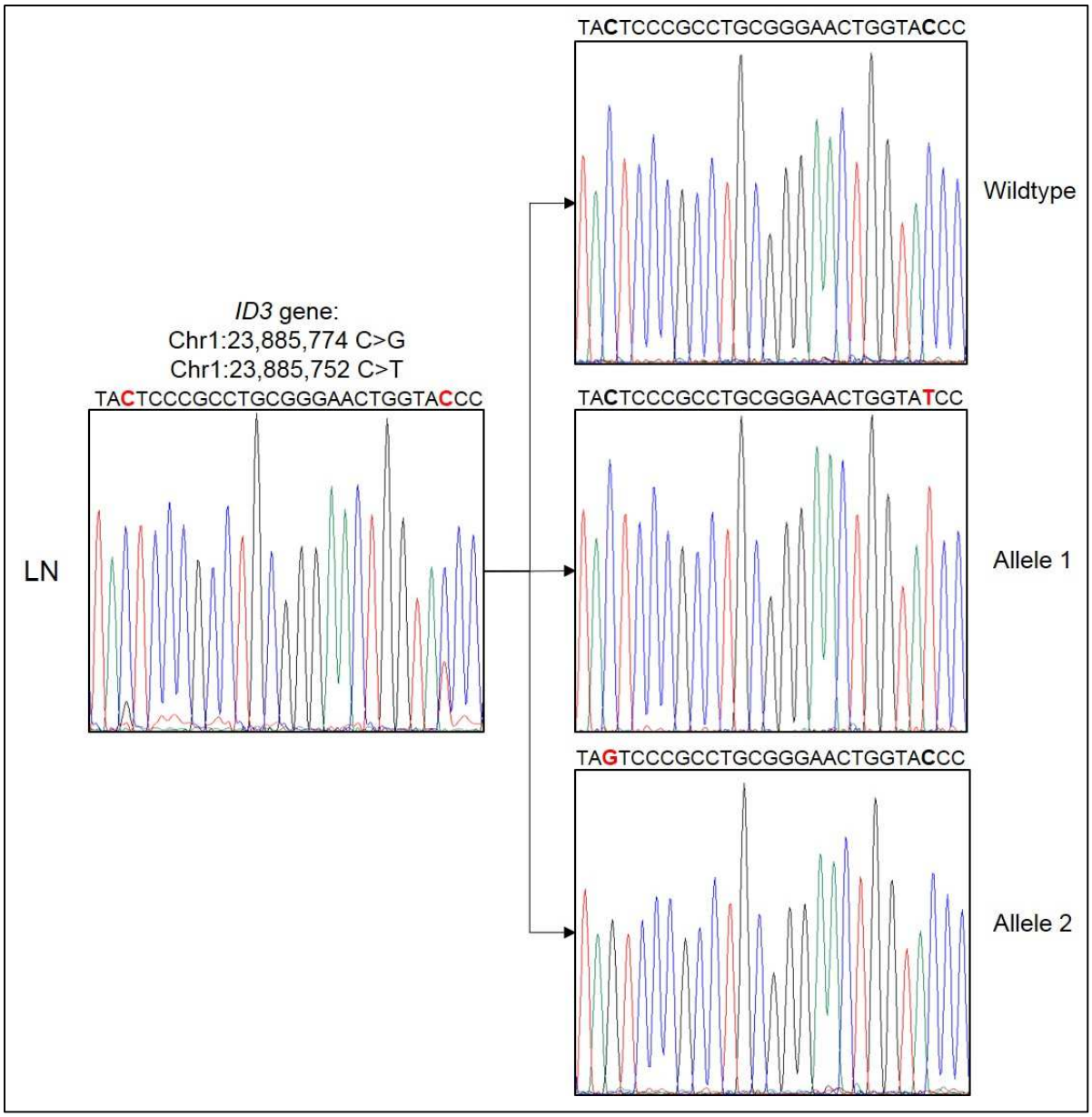
Supplementary Figures



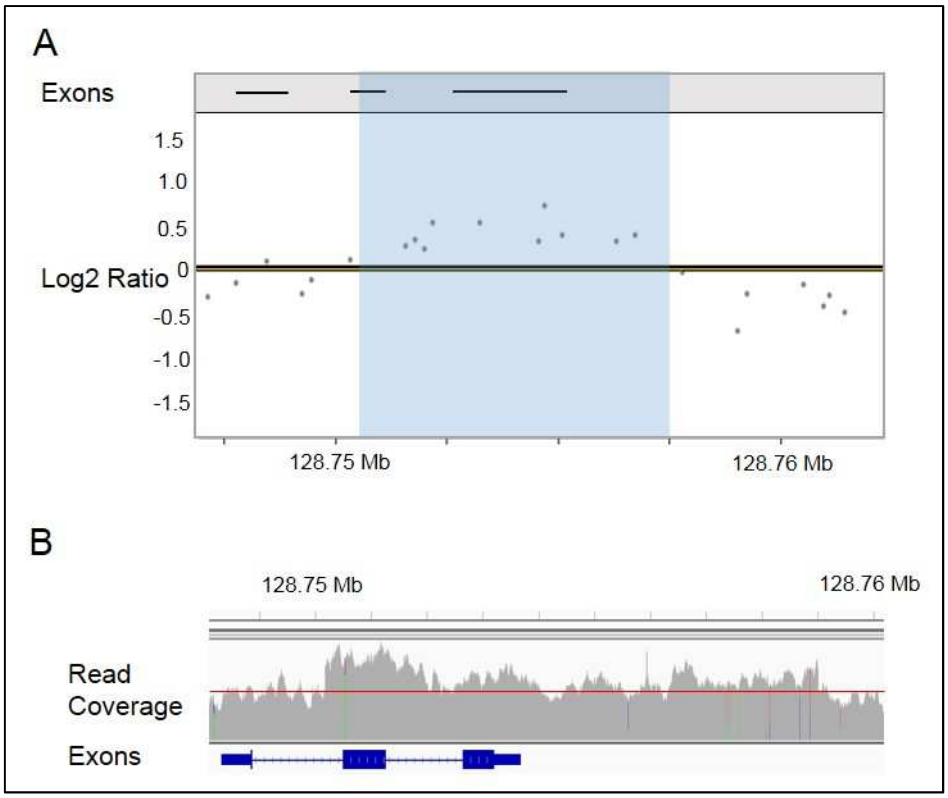
Supplementary Figure S1: Photomicrographs of the diagnostic bone marrow aspirate from the right posterior iliac crest (air-dried smears, Wright Giemsa stain, magnification 400x).



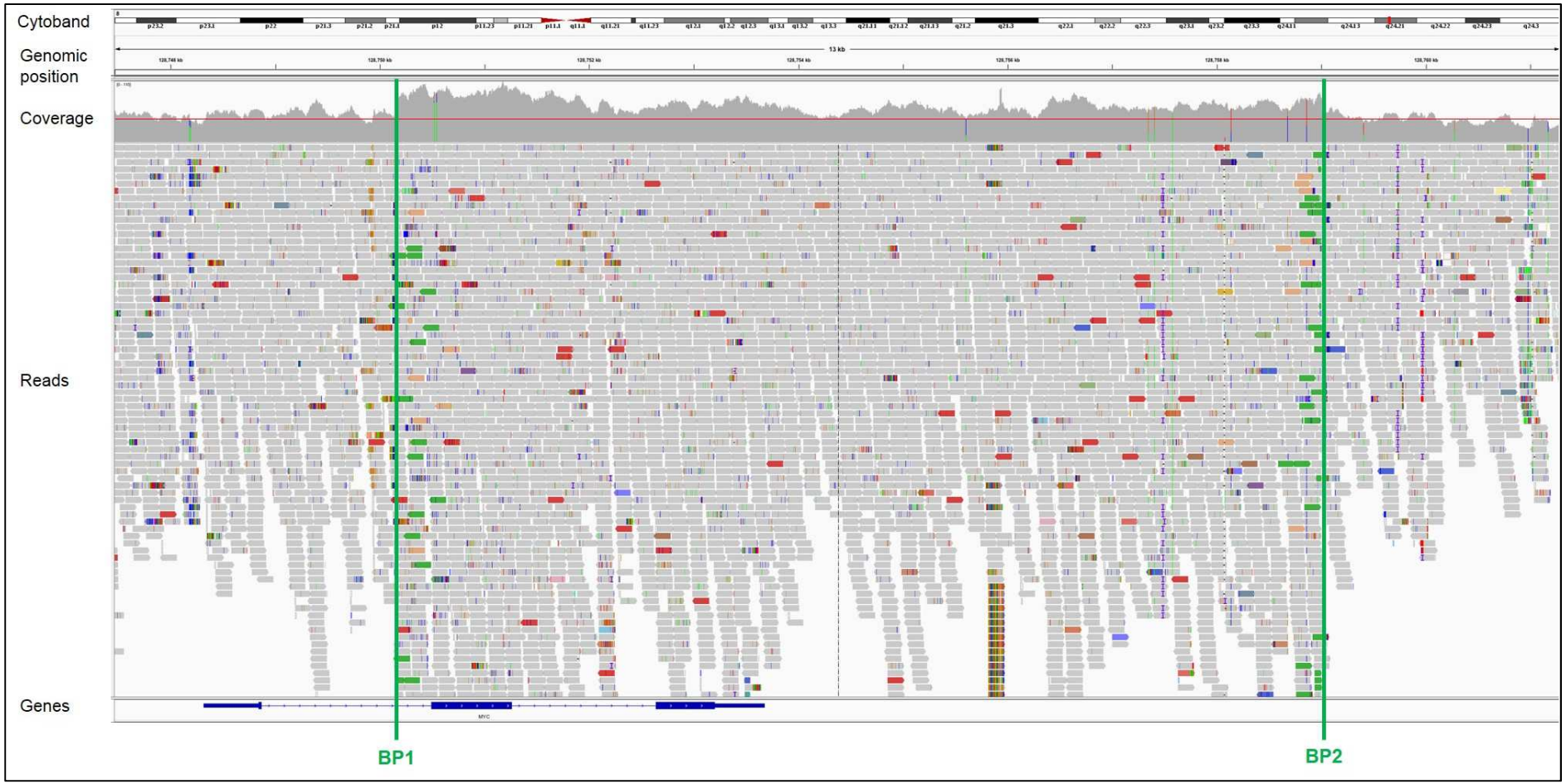
Supplementary Figure S2: Genome wide copy number profile for chromosomes 1-22 as well as X and Y of cervical LN lymphoma biopsy. Displayed are in the upper panel the logR values of the single probes and in the lower panel the B-allele frequency (BAF). A gain would be represented as an increase in the logR values above 0.3, whereas a loss would be indicated by logR value smaller -0.3. Besides a subclonal loss of chromosome Y, no alterations could be detected within this case.



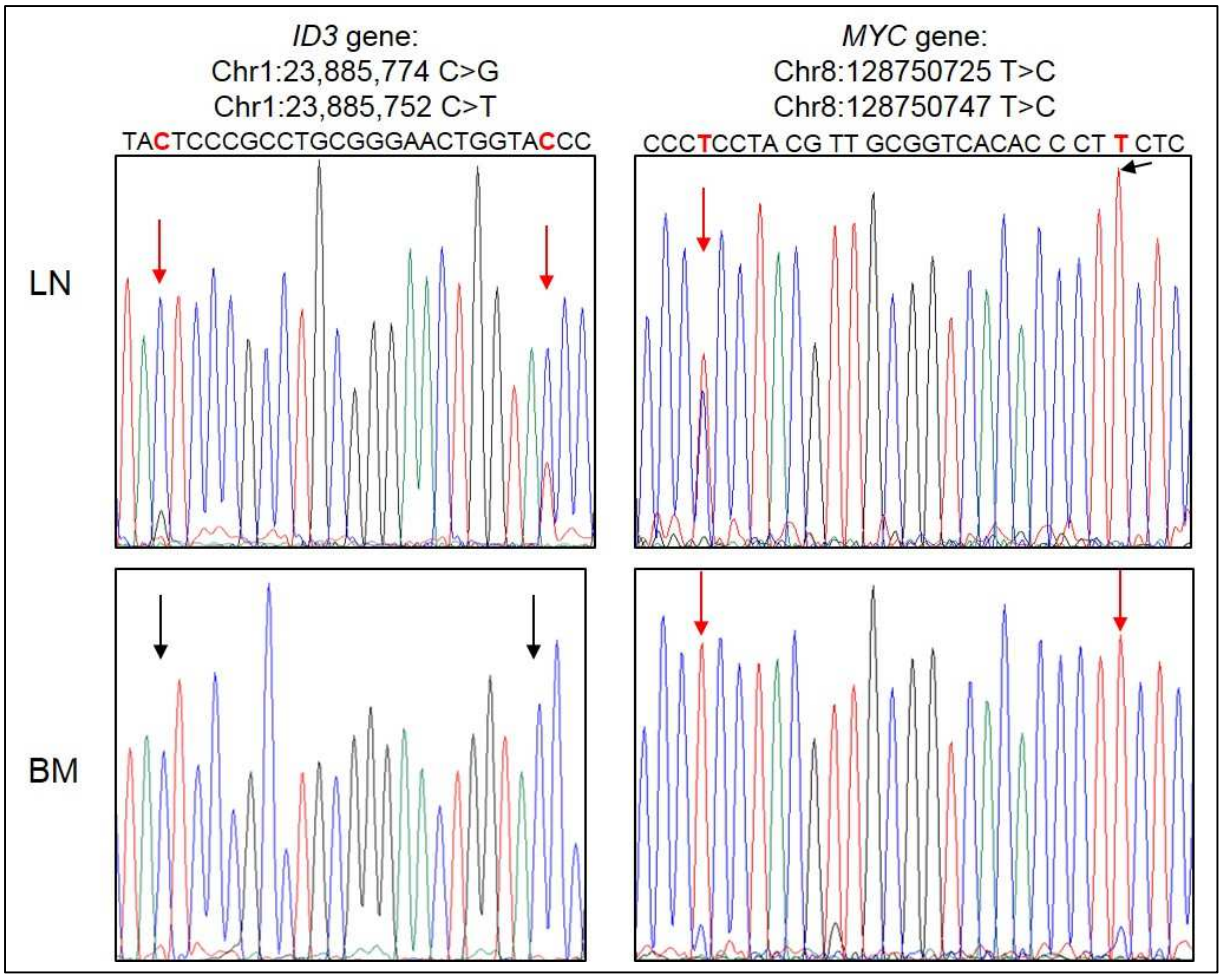
Supplementary Figure S3: Analysis of the allelic status of the *ID3* mutations in the lymphoma biopsy from the lymph node (LN). The left panel shows the initial detected *ID3* mutations (highlighted in red) as visualized in Supplementary Figure S6. Cloning of the PCR product covering both *ID3* mutations revealed that the mutations occur on different alleles with Allele 1 harboring the mutation at position chr1:23,885,752C>T and Allele 2 carrying the mutation at position chr1:23,885,774C>G. Moreover, we could detect one wildtype allele, which does not carry an *ID3* mutation.



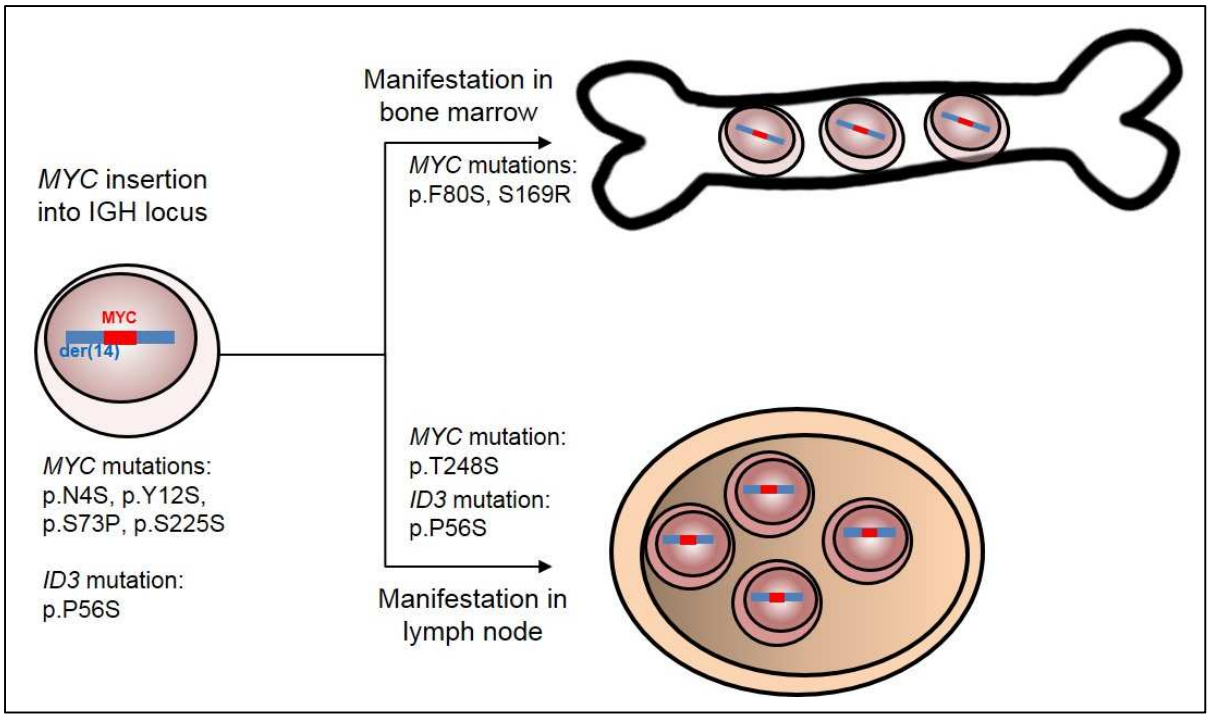
Supplementary Figure S4: Copy number gain of *MYC* gene locus as detected by OncoScan SNP-array (A) and whole genome sequencing (B). (A) Displays the log2 ratio (lower track) of the OncoScan probes for the *MYC* gene locus analyzed within the lymph node sample. The upper track displays the location of the exons. Highlighted in blue are 14 probes showing an increased log2 ratio indicating a copy number gain in this region. Since this region is too small to be picked up with the applied analysis filter, this alteration is not reported in the copy number analysis. Of note is that due to the accuracy of the technique the breakpoint borders might vary by a few tags. (B) The track displays the read coverage of the *MYC* gene locus as analyzed by whole genome sequencing of the bone marrow sample. The red line depicts the average coverage of the sample. The gained region is about 8.8 kb in size and was not picked up by the applied copy number alteration algorithms. Accordingly, it needs to be taken into consideration that even by whole genome sequencing or copy number profiling small copy number alterations might be overlooked depending on the algorithms and thresholds applied.



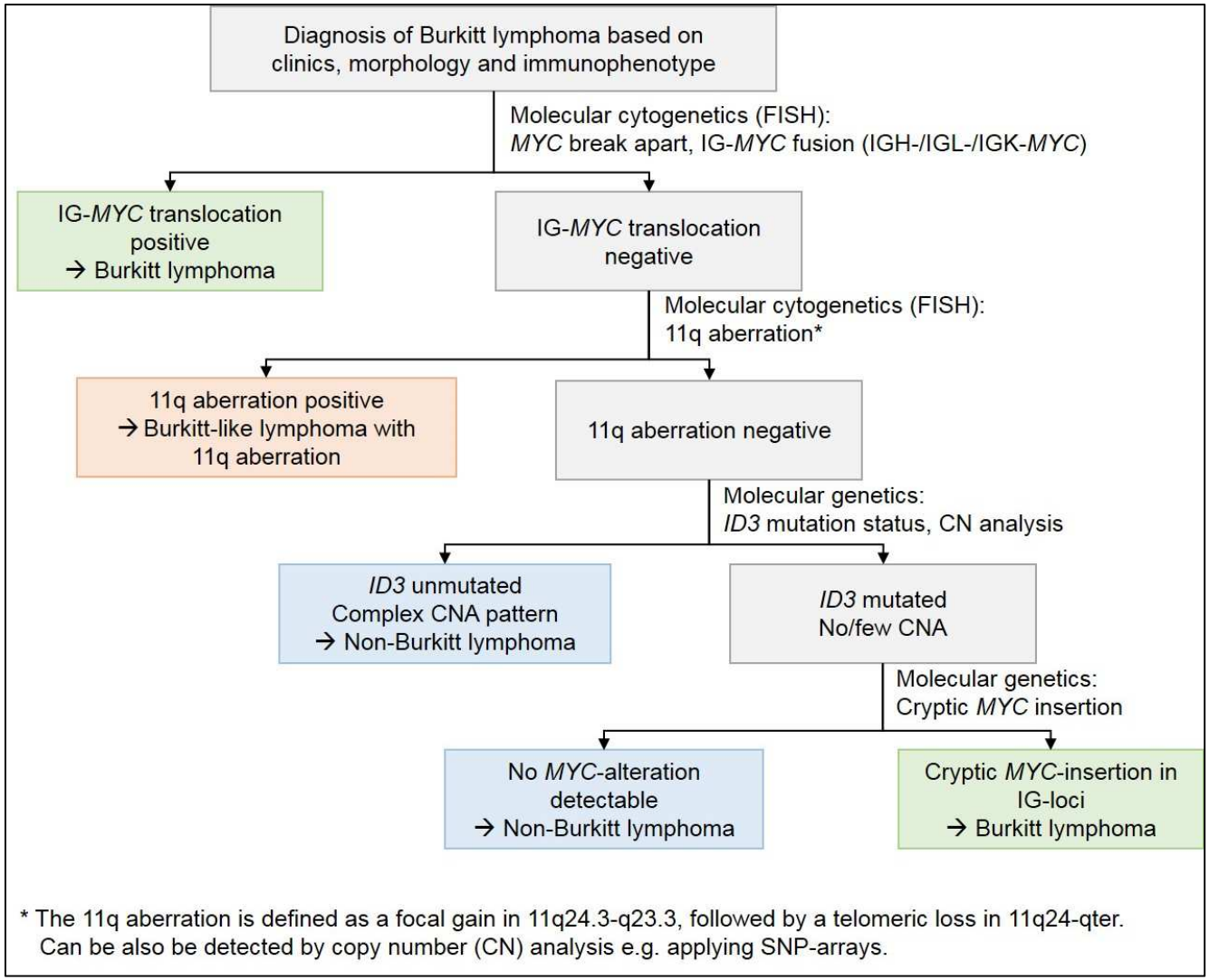
Supplementary Figure S5: Integrative Genomics Viewer screenshot of the *MYC* gene locus, which is gained and inserted within the IGH locus. The coverage tracks depicts the read coverage of the nucleotides, with the red line showing the average coverage of the sample and hence, indicating the copy number gain of this region as depicted in Supplementary Figure S4. The "Reads" track below shows the single reads for this genomic region with the colored lines within the reads indicating mismatched nucleotides. The green color code of the reads represent the discordant mate reads for which the mate is located on chromosome 14 and thus, span the breakpoint. The green line within the figure represents the breakpoint regions (BP= breakpoint) with BP1 being covered by 13 split reads and 53 normal reads and BP2 by 15 split reads and 53 normal reads. Hence, this view demonstrates that there is at least one intact *MYC* allele in the analyzed sample.



Supplementary Figure S6: Verification of *ID3* (left panel) and *MYC* (right panel) mutations by Sanger sequencing. Verification was performed on the lymphoma biopsy from the lymph node (LN) and the bone marrow (BM). The *ID3* mutations were detected by Sanger sequencing within the LN biopsy (indicated by red arrows) but not within the BM biopsy (indicated by black arrows). The *MYC* mutations were detected by Sanger sequencing in the bone marrow samples (indicated by red arrows), whereby the height of the peaks suggests a subclonal mutation pattern. The mutation at position chr8:128750725 bp (hg19) detected within the BM biopsy was detected in the LN sample, whereas the mutation at position chr8:128750747 was not detectable.



Supplementary Figure S7: Clonal evolution of the tumor samples in the patient. As verified in both tumor samples from the bone marrow and the lymph node, they derive from the same ancestor carrying a *MYC* insertion within the IGH locus and share four *MYC* mutations. During tumor evolution, the tumor sample of the bone marrow acquired two more *MYC* mutations not detectable in the lymph node sample. In contrast, the tumor sample of the lymph node acquired another *MYC* mutation as well as *ID3* mutations, not detectable in the bone marrow sample.



Supplementary Figure S8: Suggested gradually molecular cytogenetic and genetic diagnostics to differentiate *MYC*-positive Burkitt lymphoma from *MYC*-negative Burkitt-like lymphoma with 11q aberration and other non-Burkitt lymphoma. Abbreviations: CN: copy number, CNA: copy number alterations.

References

1. Schlegelberger B, Zwingers T, Harder L, et al. Clinicopathogenetic significance of chromosomal abnormalities in patients with blastic peripheral B-cell lymphoma. Kiel-Wien-Lymphoma Study Group. Blood 1999;94(9):3114–3120.

2. Ventura RA, Martin-Subero JI, Jones M, et al. FISH analysis for the detection of lymphoma-associated chromosomal abnormalities in routine paraffin-embedded tissue. J Mol Diagn 2006;8(2):141–151.

3. Kröber A, Seiler T, Benner A, et al. V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. Blood 2002;100(4):1410–1416.

4. Wagener R, López C, Kleinheinz K, et al. IG-MYC+ neoplasms with precursor B-cell phenotype are molecularly distinct from Burkitt lymphomas. Blood 2018;132(21):2280–2285.

5. 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. Nat Genet 2012;44(12):1316–1320.

6. Kretzmer H, Bernhart SH, Wang W, et al. DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. Nat Genet 2015;47(11):1316–1325.

7. Sahm F, Toprak UH, Hübschmann D, et al. Meningiomas induced by low-dose radiation carry structural variants of NF2 and a distinct mutational signature. Acta Neuropathol 2017;134(1):155–158.

8. López C, Kleinheinz K, Aukema SM, et al. Genomic and transcriptomic changes complement each other in the pathogenesis of sporadic Burkitt lymphoma. Nat Commun 2019;10(1):1459.