Identification and characterization of pathogenetically relevant genes in mature aggressive B-cell lymphomas of childhood by a combination of genomic and functional analyses

Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Rabea Wagener

Kiel, 2015

Erster Gutachter: Prof. Dr. rer. nat. Manuela Dittmar

Zweiter Gutachter: Prof. Dr. med. Reiner Siebert

Tag der mündlichen Prüfung: 25.06.2015

Zum Druck genehmigt: 25.06.2015

Gez. der Dekan Prof. Dr. rer. nat. Wolfgang J. Duschl

Table of contents

Abbreviations and acronyms

A	adenine
ABC	activated B-cell like
AID	activation induced deaminase
AKT	v-akt murine thymoma viral oncogene homolog
ALL	acute lymphoblastic leukemia
aa	amino acids
Amp	ampere
ARID1A	AT rich interactive domain 1A (SWI-like)
ARID1B	AT rich interactive domain 1B (SWI1-like)
ATP5L	ATP synthase, H+ transporting, mitochondrial Fo complex, subunit G
AT/RT	atypical teratoid rhabdoid tumor
BAF	BRG-1 associated factor
BCAP	B-cell adaptor for PI3K
BCL2	B-cell CLL/lymphoma 2
BCL2L1	BCL2-like 1
BCR	B-cell receptor
bHLH	basic helix-loop-helix
BL	Burkitt lymphoma
BLAT	basic local alignment tool
bm	bone marrow
bp	base pairs
BRG1	brahma related gene 1
C	cytosine
CCG	Cologne center for genomics
CCLE	cancer cell line encyclopedia
CCND3	cyclin D3
CDK4	cyclin-dependent kinase 4
CDK6	cyclin-dependent kinase 6
CDKN2A	cyclin-dependent kinase inhibitor 2A
cDNA	complementary DNA
CDS	coding sequence
CEP164	centrosomal protein 164kDa
chr	chromosome
ChIP	chromatin immunoprecipitation

CNA	copy number alterations
CNV	copy number variations
COSMIC	catalogue of somatic mutations in cancer
CpG	cytosine-guanine-dinucleotide
Cq	quantification cycle
CREBBP	cAMP-response element binding protein (CREB) binding protein
C-terminal	carboxy-terminal
CTE	C-terminal extension
dbSNP	single nucleotide polymorphism database
ddH_2O	double-distilled water
DLBCL	diffuse-large B-cell lymphoma
DMSO	dimethyl sulfoxide
DNA	desoxyribonucleic acid
DNP	dinitrophenyl
dsDNA	double-stranded DNA
DZ	dark zone
eBL	endemic Burkitt lymphoma
EBV	Epstein-Barr virus
EDTA	ethylenediaminetetraacetate
e.g.	exempli gratia, for example
ELISA	enzyme-linked immunosorbent assay
EPHA7	EPH receptor 7
ESC	embryonic stem cell
ETS1	v-ets avian erythroblastosis virus E26 oncogene homolog 1
ex	exon
FBS	fetal bovine serum
FFPE	formalin-fixed, paraffin-embedded
FISH	fluorescence in situ hybridization
FLI1	friend leukemia virus integration 1
FL	follicular lymphoma
frs	frameshift
G	guanine
GC	germinal center
GCB	germinal center B-cell like
GCT	germ cell tumors

gDNA	genomic DNA
GFP	green fluorescent protein
GUSB	glucuronidase, beta
h	hour
H_2O	water
HAT	histone acetyltransferase
Helicase_C	helicase conserved C-terminal domain
HIV	human immunodeficiency virus
HLH	helix-loop-helix
HPRT1	hypoxanthine phosphoribosyltransferase 1
hPSC	human pluripotent stem cells
HRP	horseradish peroxidase
HSA	helicases and associated with SANT domains
ICGC	international cancer genome consortium
ID3	inhibitor of DNA binding 3, dominant negative helix-loop-helix protein
IG	integrative genomics
IL10	interleukin 10
IL10R	interleukin 10 receptor
IL10RA	interleukin 10 receptor subunit alpha
IGH	immunoglobulin heavy chain gene
IRF4	interferon regulating factor 4
IPTG	isopropyl-β-D-thiogalactoside
JAK1	janus kinase 1
kb	kilo bases
KH	K-homology domain
KMT2A	lysine (K)-specific metyhltransferase 2A
LB	lysogeny broth
lncRNA	long non coding RNA
LOH	loss-of-heterozygosity
LZ	light zone
MAF	mutated allele frequency
Max	MYC-associated factor X
Mb	mega bases
mBL	molecular Burkitt lymphoma
MDM2	MDM2 proto-oncogene, E3 ubiquitin protein ligase

min	minute
miRNA	microRNA
MIP	molecular inversion probe
MMML	molecular mechanisms in malignant lymphoma
mnBLL	*MYC*-negative Burkitt-like lymphoma
mTOR	mechanistic target of rapamycin (serine/threonine kinase)
MYC	v-Myc avian myelocytomatosis viral oncogene homolog
NaCl	sodium chloride
NGS	next-generation-sequencing
NHL	non-Hodgkin lymphoma
NLS	nuclear localization signal
NMD	nonsense mediated decay
N-terminal	amino-terminal
NTC	non-targeting control
o/n	over night
ORF	open reading frame
PAFAH1B2	platelet-activating factor acetylhydrolase 1b, catalytic subunit 2
pB	peripheral blood
PBS	phosphate buffered saline
PCBP1	poly(rC) binding protein 1
PCR	polymerase chain reaction
PCSK7	proprotein convertase subtilisin/kexin type 7
PDK1	phosphoinositide-dependent kinase 1
PFA	paraformaldehyde
PI3K	phosphatidylinositol-4,5-bisphosphate 3-kinase
PIP_2	phosphatidylinositol 4,5-diphosphate
PIP_3	phosphatidylinositol 3,4,5-triphosphate
PRDM1	PR domain containing 1, with ZNF domain
PRDM10	PR domain containing 10
PTEN	phosphatase and tensin homolog
qPCR	quantitative real-time PCR
RHOA	ras homology family member A
RNA	ribonucleic acid
RQI	RNA quality indicator
rs	reference SNP

s	second
sBL	sporadic Burkitt lymphoma
SCCOHT	small cell carcinoma of the ovary, hypercalcemic type
SDS	sodium dodecyl sulfate
shRNA	short hairpin RNA
SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4
SNF2_N	SNF2 family N-terminal domain
SNP	single nucleotide polymorphism
SNV	single-nucleotide variant
SNX19	sorting nexin 19
SOCS3	suppressor of cytokine signaling 3
ssDNA	single-stranded DNA
STAT3	signal transducer and activator of transcription 3
STR	short tandem repeat
SYK	spleen tyrosine kinase
T	thymine
TBE	TRIS-borat-EDTA buffer
TBS	TRIS-buffered saline
TCF3	transcription factor 3
TCF4	transcription factor 4
TEN	TERT essential N-terminal
TERC	telomerase RNA
TERT	telomerase reverse transcriptase
TP53	tumor protein p53
TRBD	TERT RNA binding domain
TRIS	tris(hydroxymethyl)aminomethane
TYK2	tyrosine kinase 2
UTR	untranslated region
UBE4A	ubiquitination factor E4A
V	volt
VCF	variant call files
w/o	without
wt	wildtype
ZNF259	zinc finger protein 259

1 Introduction

1.1 B-cell lymphomas

B-cell lymphomas are hematological malignancies developing from B-cells [1]. Traditionally, they are divided into Hodgkin and non-Hodgkin lymphoma (NHL) [1]. Another classification distinguishes low grade lymphomas, like follicular lymphoma (FL) from high grade lymphomas, like Burkitt lymphoma (BL) or diffuse large B-cell lymphoma (DLBCL) [1]. From a more clinical point of view, the high grade lymphomas are also referred to as aggressive lymphomas reflecting their noteworthy rapid growth but nevertheless good curability [1]. A further feature used for classification is the B-cell stage which they resemble, as for example B-lymphoblastic lymphoma is supposed to derive from immature B-cells whereas BL, FL and DLBCL resemble mature B-cells [1]. The mature B-cell malignancies represent 4 % of new cancers each year and 90 % of malignant lymphomas worldwide [2],[3]. The most common mature B-cell lymphoma subtypes are FL and DLBCL [4]. The relative frequency of the subtypes varies in different parts of the world as for example FL is more common in the US and Western Europe than in Africa or Asia [5]. The median age of diagnosis for all mature B-cell lymphomas is between 60 and 70 years [1]. But DLBCL and especially BL are described to occur also at considerable frequencies in children [6]. As the focus of this thesis is on BL which is a mature aggressive B-cell lymphoma predominantly of childhood, features of pediatric B-cell lymphoma will be presented in the following.

1.1.1 B-cell lymphomas in children

In general, the cancer spectrum in children differs from that in adults. For example, the most frequent non-leukemic malignancies in children are embryonal tumors like neuroblastoma, Wilms tumor, rhabdomyosarcoma and retinoblastoma which are rare or do not exist at all in adults [7],[8]. Furthermore, the biological characteristics of a tumor may differ between children and adults as has been shown for acute lymphoblastic leukemia (ALL). In ALL the most common genetic alteration in children is the *ETV6-RUNX1* fusion due to the translocation t(12;21) whereas in adult ALL, the most common translocation is the t(9;22) resulting in *BCR-ABL1* fusion [8]. The reasons for the different spectra of malignancies might be that malignancies of the childhood arise predominantly within developing tissues that still undergo formation, growth or maturation, reflected by the high incidence of embryonic tumors as described above [9]. In adults, on the other hand, the most prevalent tumors are carcinomas derived from epithelia of the gastrointestinal tract, lung, breast or prostate [8].

Lymphoma is the third most frequent cancer in persons below the age of 18 years at diagnosis. It has accounted for 14 % of all childhood malignancies in the years 2009-2013 in Germany [10]. Refer to Table 1 for a comparison of the distribution of pediatric and adult lymphoma subtypes.

Table 1: Distribution of lymphoma subtypes in children and adults. Adapted from [11].

		Children	Adults
	Lymphoma subtypes	Distribution of lymphoma subtypes	
Hodgkin lymphoma	Hodgkin lymphoma	46.0 %	13.0 %
Non-Hodgkin lymphoma (NHL)	Burkitt lymphoma	25.0 %	0.4 %
	T-lymphoblastic lymphoma	8.0 %	1.3 %
	Anaplastic large cell lymphoma	6.5 %	1.7 %
	Diffuse-large B-cell lymphoma	4.0 %	26.0 %
	B-lymphoblastic lymphoma	3.0 %	0.5 %
	Primary mediastinal B-cell lymphoma	1.0 %	0.4 %
	Follicular lymphoma	0.5 %	26.0 %
	Other	6.0 %	30.7 %

About 5 % of the NHL in Western Europe are BL but the percentage differs between children and adults: 40 % of pediatric NHL are BL whereas only 1-2 % of adult NHL are BL [12].
The following chapter will focus on the epidemiology and clinical features of BL, as it is the major subtype of lymphoma studied within this thesis.

1.1.2 Burkitt lymphomas

1.1.2.1 Epidemiology of Burkitt lymphomas

Burkitt lymphomas belong to the mature aggressive B-cell lymphomas and as outlined above are the most common NHL of the childhood. In general, three clinical variants of BL are distinguished: endemic, sporadic and immunodeficiency-associated [1]. The epidemiologic features of each of them are shortly introduced in the following.
Endemic BL (eBL) is the most common childhood malignancy in equatorial Africa, Papua New Guinea and Brazil accounting in these countries for 30-50 % of all childhood cancers [13]. The occurrence of eBL correlates with the geographical distribution of endemic malaria [14] in this so called malaria belt [15]. The incidence peak is between the ages of 4-7 years at diagnosis [1]. A hallmark of the endemic variant is the presence of the Epstein-Barr virus (EBV) genome in the majority of neoplastic cells [16].
Immunodeficiency-associated BL is frequently seen in association with HIV infection and occurs as initial manifestation of AIDS [17]. But also iatrogenic causes for immunodeficiency as for example immunosuppression upon transplantations have been described [18],[19].
Sporadic BL (sBL), which is the main BL variant studied within this thesis, occurs throughout the world but predominantly in North America and Western Europe [1]. The age distribution of BL according to

the surveillance, epidemiology and end results registry (SEER) of the Unites States in children is trimodal showing peaks between the ages of 5-9 years at diagnosis as well as in middle aged adults between 40-44 years at diagnosis and in the elderly between 80-84 years at diagnosis [20]. According to the Netherland Cancer Registry, a bimodal age distribution is observed with one age peak between 6-10 years at diagnosis and the other above 60 years at diagnosis [21]. To sum up, based on the above described studies, the main incidence peak in children is between the ages of 5-10 years, whereas in adults BL is diagnosed between the ages of 40-80 years. In Germany 49 % of the children with non-Hodgkin lymphoma younger than 15 years at diagnosis have BL [22]. Sporadic BL in children predominantly manifests in males whereas in females an increase in BL presentation has been associated with advanced age [20]. Thus, the male to female ratio in children (<20 years at diagnosis) is 3.13 whereas it declines to 2.31 in middle-aged adults (20-59 years at diagnosis) and to 1.53 in elderly (>60 years at diagnosis) [23]. In contrast to the eBL, EBV is only detected in 15-30 % of sBL [16].

In high income countries pediatric sBL have a better outcome than their adult counterpart with an overall survival with current treatment of more than 90 % after 5 years [24] in contrast to 53-71 % in adults [25],[26].

1.1.2.2 Clinical features and pathology of Burkitt lymphomas

The focus of this thesis is on sporadic BL, which are referred to for the sake of convenience as BL. The following description of the features is confined to this variant.

In about 90 % of children BL manifests in the ileo-caecal region [27]. In pediatric BL extranodal sites are most often affected whereas in adults a nodal presentation is more common [1]. Manifestations within the ovaries, kidneys and breast are also observed. The frequent abdominal involvement is associated with a respective clinical presentation including abdominal pain, nausea and vomiting. The majority of BL show a bulky disease and a high tumor burden. BL cells show a short doubling time of 24-48 h [1], which in turn leads in pediatric BL to symptoms within a couple of weeks.

About 30-38 % of patients with BL have been reported to have a bone marrow involvement [28],[27]. Of those, only few cases present as Burkitt leukemia or acute lymphoblastic leukemia-L3 with more than 25 % of bone marrow cells being lymphoma cells at diagnosis [27].

The single BL cells present as medium-sized cells harboring round nuclei with finely clumped chromatin and multiple nucleoli [1]. Furthermore, BL show a diffuse pattern of growth with many mitoses and a so called “starry sky pattern” which refers to numerous benign macrophages that ingest apoptotic BL cells [1].

BL express moderate levels of IgM and the B-cell associated antigens as CD19 and CD20. Furthermore, BL show a germinal center-like immunophenotype being CD10 and BCL6 positive. The majority of pediatric BL are negative or only weakly positive for BCL2, whereas positivity is most often observed in adults. Key feature of BL is the high positivity for the proliferation marker Ki67 which reaches up to 100 % in the tumor cells [27].

A genetic hallmark of BL is the *IG-MYC* translocation which can be detected in almost 100 % of cases. This translocation juxtaposes the *MYC* gene to one of the *immunoglobulin* (IG) gene loci. The mechanism of the *MYC* translocation and its impact on Burkitt lymphomagenesis will be introduced in detail in 1.3.3.

1.2 B-cell development and lymphomagenesis

1.2.1 Stages of B-cell differentiation

As B-cell lymphoma develop from B-cells at different maturation stages, the knowledge of the features of normal B-cell development is important for the understanding of B-cell lymphomagenesis. Figure 1 depicts the different stages of B-cell development with focus on the germinal center.

B-cells derive from hematopoietic stem cells within the bone marrow which also give rise to T-cells and cells of the myeloid lineage. The B-cell lineage commitment is ensured upon expression of several transcription factors, like transcription factor 3 (TCF3), as well as by cytokine signaling via interleukin 7 [29],[30],[31]. Within the bone marrow the precursor cells differentiate via pro B-cell and pre B-cell stages to immature B-cells. During these steps, the V(D)J rearrangement of the immunoglobulin genes (1.2.2) takes place leading to the expression of a rearranged IgM on the surface of the immature B-cells. Subsequently, these B-cells are tested for successful rearrangement and autoreactivity. Self-reactive cells are negatively selected and undergo either (i) cell death by apoptosis or clonal deletion, (ii) produce a new receptor by receptor editing or (iii) enter a state of anergy or immunological tolerance [32]. The surviving, non-self-reactive B-cells leave the bone marrow and become mature, naïve B-cells. These cells circulate in the peripheral blood and are also found in primary lymphoid follicles [33]. Upon encounter of an antigen, the naïve B-cells might either mature directly into plasma cells which produce an early IgM antibody response or migrate to the center of the primary follicle and form the germinal center [33].

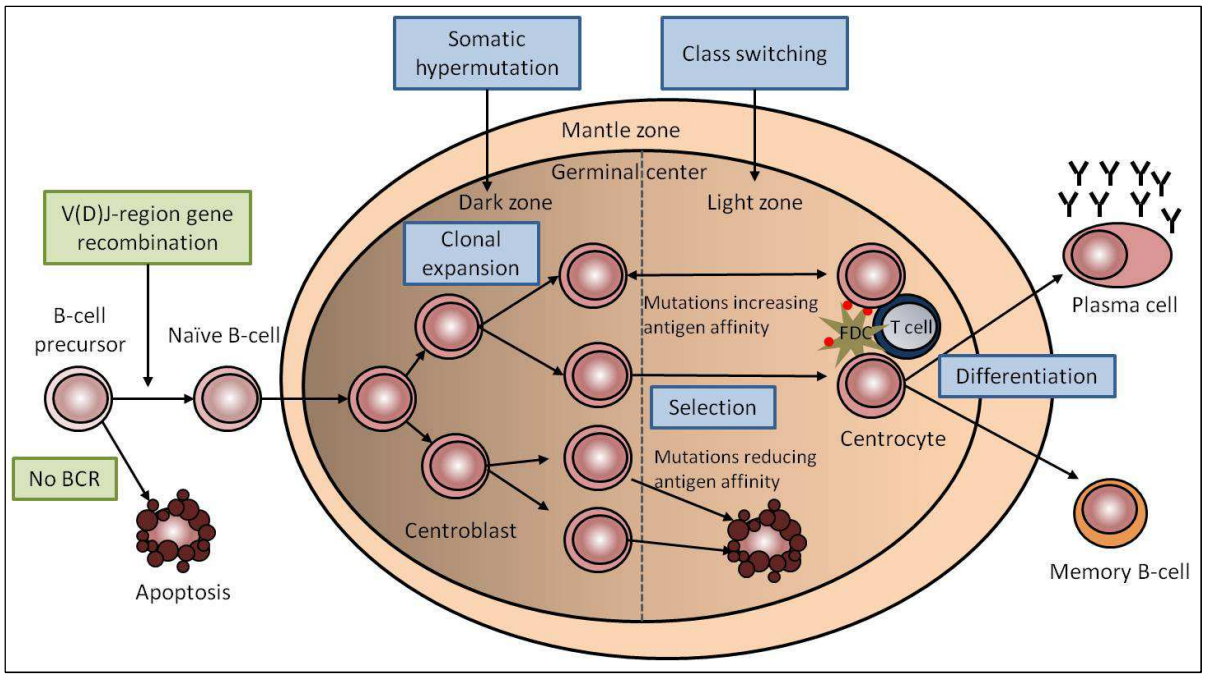


Figure 1: B-cell differentiation in the germinal center. Naïve B-cells derive from precursor B-cells within the bone marrow where the immunoglobulin gene has undergone V(D)J rearrangement and a functional IgM surface protein is expressed. Upon encounter with an antigen, the naïve B-cells enter the germinal center where they undergo clonal expansion. During this time the immunoglobulin V genes of the now called centroblasts accumulate mutations induced by the somatic hypermutation machinery which contribute to the increase of the antigen-specificity. The cells now enter the light zone, where they are called centrocytes. At this stage, they become selected for a high affinity to the antigen. Only those cells with a high affinity do either reenter the dark zone for some additional rounds of clonal expansion or differentiate into antibody secreting plasma cells or memory B-cells. The BCR of centrocytes which have a low antigen-affinity undergo apoptosis. A further mechanism which increases the affinity of the immunoglobulin is the class switching which takes place in the light zone. Modified from [4].

Within the germinal center the immunoglobulins undergo affinity maturation. Initially, the B-cells enter the dark zone (DZ) where they proliferate extensively and are called centroblasts. During this stage, the immunoglobulin variant region genes accumulate mutations due to the activity of the somatic hypermutation machinery (1.2.2). After several rounds of cell division, the B-cells, now called centrocytes, enter the light zone (LZ). At this stage the class switch recombination (1.2.2) of the immunoglobulin gene takes place which is a further mechanism for affinity maturation [34]. In the LZ, the centrocytes are surrounded by follicular dendritic cells which present different antigens on their surface. The centrocytes compete with each other for the specific recognition of the antigen [33]. If a centrocyte recognizes the antigen it receives a signal to reenter the dark zone for some additional rounds of clonal expansion and further affinity maturation. A role for MYC in this so called cyclic reentry model for germinal center B-cells and, thus, for affinity maturation has been recently proposed by two studies [35],[36]. MYC has been shown to be expressed in mature B-cells and, thus, to be important for the initiation of the GC formation. But the expression of MYC is inhibited in the centroblasts of the DZ likely due to the expression of BCL6, which can bind to the *MYC* promoter. In the LZ the expression of MYC is induced in a small subset of B-cells likely by a T-cell- and antigen-dependent interaction of the respective centrocytes [35]. This small subset of LZ B-cells is likely

positively selected for and, thus, expresses a high-affinity B-cell receptor (BCR). Subsequently, these B-cells can reenter in a MYC-dependent manner the DZ for an additional round of proliferation to acquire more mutations increasing the BCR affinity for its antigen. The other centrocytes, not re-expressing MYC, exit the germinal center and become either plasma cells or memory B-cells. Different factors direct the cell fate of the plasma cells. Plasma cells express low levels of PAX5 and BCL6 but high levels of IRF4 [37]. IRF4 induces the expression of PRDM1, which is also called BLIMP1. PRDM1 is a transcriptional repressor that switches off genes required for B-cell proliferation and affinity maturation [38]. Some of the plasma cells migrate into the bone marrow where a subset lives for a long period due to pro-survival signals by the surrounding stroma cells. Others migrate into the medullary cords in the lymph nodes or in the spleen. Irrespective of their location, they produce high levels of antibody. Memory B-cells are long-lived and divide - if at all - very slowly [39]. They recircle within the peripheral blood and migrate into the bone marrow [40] or splenic marginal zone [41]. In the initial phase of the secondary immune response memory B-cells provide the antigens to pathogens and, thus, are an important part of the adaptive immune response [42].

1.2.2 Molecular processes remodeling the immunoglobulin genes

The BCR undergoes complex rearrangements during B-cell development. In the following, the structure of the BCR and the molecular processes of the remodeling will by briefly presented.

Antibodies or immunoglobulins are the secreted form of the BCR. In Figure 2 the structure of an immunoglobulin is exemplary depicted. IGs are Y-shaped and consist of two heavy and two light chains which are made of four and two immunoglobulin domains, respectively [43]. The variable domains of the light chain (V_L) and of the heavy chain (V_H) make up the variable region which confers the ability to bind specific antigens [43]. The constant domains of the heavy chain (C_H1-3) and light chain (C_L) make up the constant region which is important for the interaction with the effector cells and molecules [43]. The BCR harbors in addition a C-terminal polypeptide which anchors the receptor in the membrane of the cells [44],[45].

The heavy and light chains are encoded by separate multigene families localized on different chromosomes. Two different types of light chains exist: lambda (λ) light chain encoded on chromosome 22q11 and kappa (κ) light chain encoded on chromosome 2p11. The heavy chain is encoded by a gene cluster on chromosome 14q32. Each of these gene clusters contains several coding sequences, so called gene segments. The light chains contain variable (V), joining (J) and constant (C) gene segments, whereas the heavy chains contain additionally a diversity (D) gene segment. The C gene segment in the heavy chain does not only contain a single gene segment but a series of regions as μ, δ, γ, α and ε which correspond to the different isotypes IgM, IgD, IgG, IgA and IgE of the immunoglobulins [39].

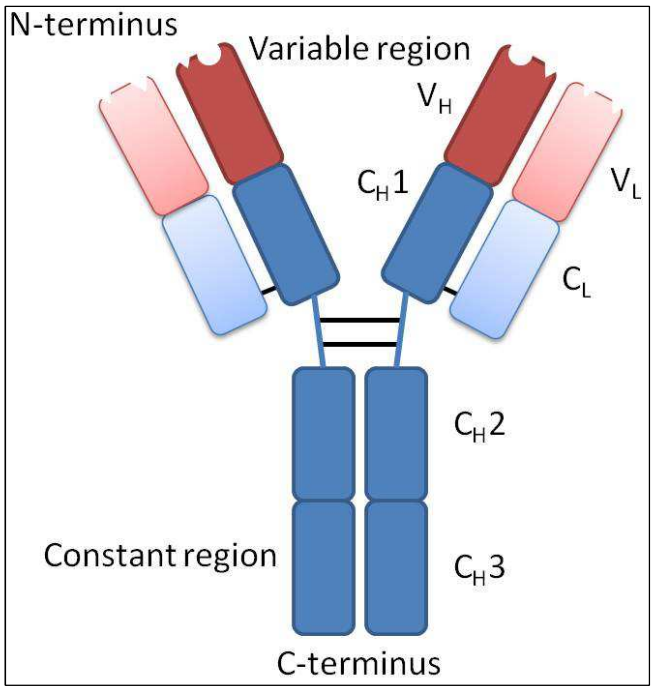


Figure 2: Exemplary structure of an immunoglobulin molecule (IgG). The Y-shaped immunoglobulin consists of two heavy chains and two light chains, the latter are indicated by the lighter color. The heavy and the light chains are made of four and two immunoglobulin domains, respectively. The variable domains of the light chain (V_L) and of the heavy chain (V_H) make up the variable region, whereas the constant domains of the heavy chain (C_H1-3) and light chain (C_L) make up the constant region. Modified from [39].

The V-J in the light chain and the V-D-J gene segments in the heavy chain build the variable region and the C gene segment the constant region of the immunoglobulin gene. As the gene segments are separated by non-coding and other coding gene segments, a rearrangement of the DNA sequence is necessary to generate a functional coding exon of the immunoglobulin. This process is called V(D)J rearrangement and takes place during the B-cell maturation in the bone marrow [46]. This random recombination is carried out by a complex of enzymes called V(D)J recombinase. This protein complex binds to the conserved recombination signal sequences which lie adjacent to each gene segment and joins initially the D to the J gene segment in the heavy chain. In a second step the V gene segment is joined to the DJ gene segment (Figure 3A). In the light chain, the V gene segment is directly joined to the J segment. Due to the high number of functional gene segments, many combinations for the V(D)J recombination are possible leading to a high combinatorial diversity of immunoglobulins. This diversity is further increased by the so called junctional diversity, as during the recombination processes nucleotides at the joint region between two gene segments might be added or deleted [47].

Additional mechanisms taking place in the germinal center reaction lead to secondary diversification of the immunoglobulins. Among these mechanisms is the somatic hypermutation of the immunoglobulin genes (Figure 3B). During the germinal center reaction point mutations are introduced in centroblasts at a high rate into the gene segment encoding the V region [48]. The process of somatic hypermutation is mainly conducted by the activation induced deaminase (AID)

which is expressed in activated germinal center B-cells [49],[50]. Upon binding of this enzyme to single stranded DNA, it deaminates cytosine to uracil [49]. The deamination induces mismatch repair or base-excision repair mechanisms which further alter the DNA sequence [51]. Alterations of the DNA sequence do not only diversify the immunoglobulins but also alter the affinity of the BCR for the antigen. The survival of those B-cells is favored which harbor *IG* mutations increasing the affinity of the immunoglobulin.

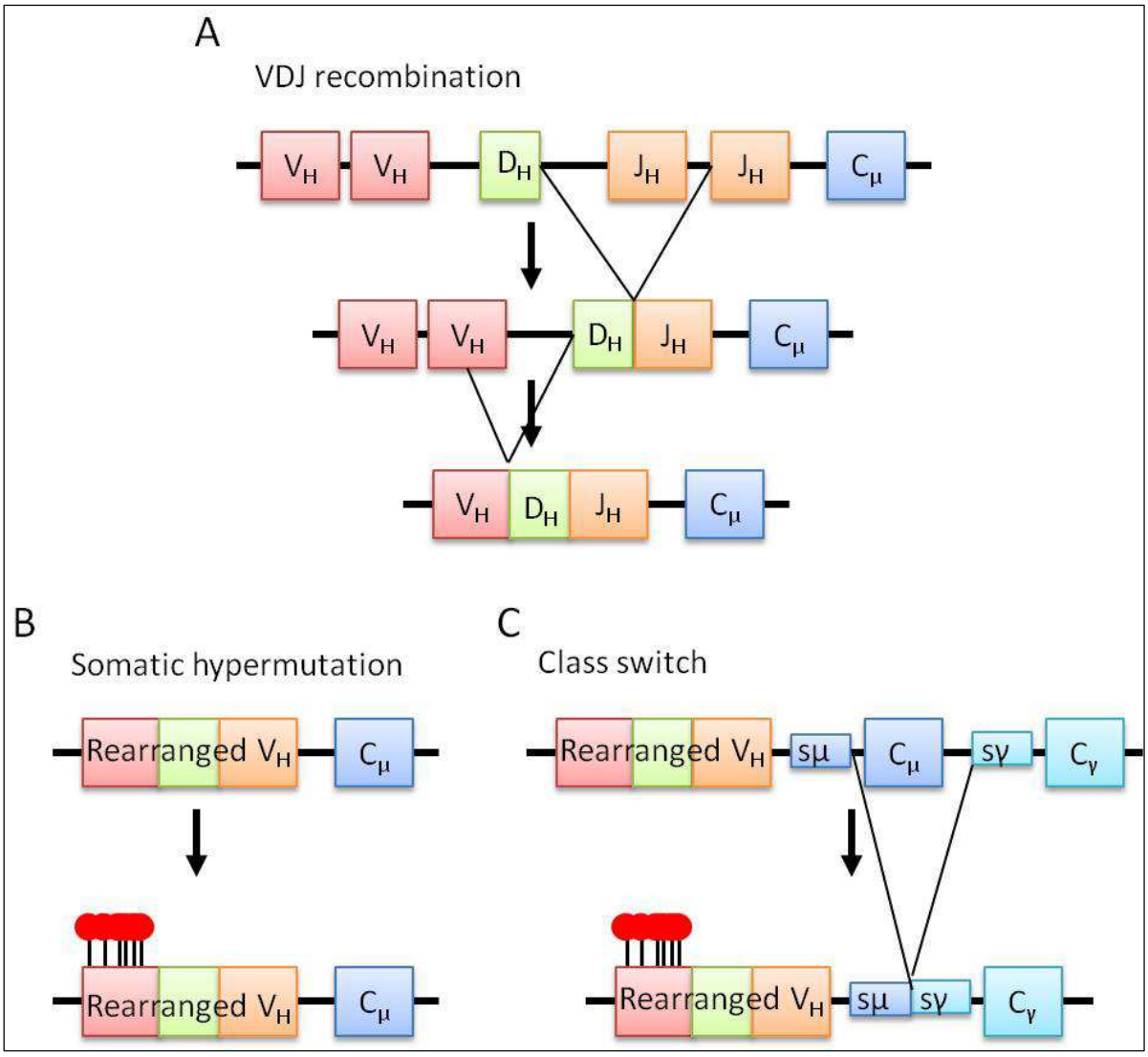


Figure 3: Overview on the molecular processes remodeling the immunoglobulin genes using the example of the immunoglobulin heavy chain (*IGH*) locus. A: Depicted is the schematic overview on the VDJ recombination. Initially the D gene segment is joined randomly to one of the J gene segments. In a next step one of the V gene segments is joined to the DJ segment. B: After VDJ recombination and after encounter of an antigen, the somatic hypermutation machinery is activated in the germinal center. This leads to the introduction of mutations at a high rate in the V region of the heavy and the light chain (indicated by the lollipops). C: The final step in the affinity maturation of the immunoglobulin genes is the class switch which only occurs on the heavy chain locus. This mechanism is based on a type of non-homologous end joining between the so called switch regions (here shown for sμ and sγ). These lie adjacent to each C gene segment and direct the class switch. Modified from [4].

Another mechanism which leads to the diversification of the immunoglobulins is the so called class switch recombination or class switching (Figure 3C). During this process - occurring following antigen

encounter of the B-cell - the constant gene segment of the heavy chain is switched. Initially, the naïve B-cell express IgM or IgD immunoglobulins. During affinity maturation, the IG class switches in the majority of cases to IgG resulting in an antibody with different effector functions [52]. Class switching involves irreversible DNA recombination, which is a type of non-homologous DNA recombination [53]. Switching is guided by stretches of repetitive DNA, so called switch regions, which lie in the intron between the J_H and the Cμ gene segment as well as upstream of the other C gene segments [54]. AID induces single-strand nicks in these switch regions leading to double-strand breaks [55]. By non-homologous recombination between the switch regions the breaks are repaired leading to a joining of the VDJ segment with the switched C segment [53].

1.2.3 Cellular origin of germinal center B-cell lymphomas

The cell-of-origin is defined as the cellular stage (normal stage of B-cell differentiation) resembling best the phenotype of a tumor cell. The analysis of the cell-of-origin in malignant B-cell lymphoma is based on the finding that the tumor cells seem to be frozen at a particular differentiation stage [56]. This does not necessarily reflect the real cell of origin as has been shown for FL. Although in this lymphoma the initiating translocation of the *BCL2* gene is likely created during the VDJ recombination in the bone marrow [57],[58], FL cells display features of mature germinal center B-cells [1]. Despite this, for the sake of convenience, the term cell-of-origin is used in the following.

The analysis of the cell-of-origin is important as it will likely influence the physiology of the tumor cell. Therefore, its definition is important to understand the biology of the tumor. Nevertheless, the approach to define the cell-of-origin has changed in the last decades.

Initially, the definition of the cell-of-origin was based on the fact that the BCR of mature B-cells undergoes affinity selection within the germinal center of secondary lymphoid organs, which is reflected by an increase of mutations within the *IG* genes [59]. Based on the *IG* mutation status the origin of the lymphoma used to be defined. In general, lymphomas harboring no *IG* gene mutations were defined to derive from pre-GC B-cells. Lymphomas harboring mutated *IG* genes on the other hand were defined to derive from B-cells of the GC or post-GC [56]. But this is not a reliable definition as for example a group of memory B-cells has been described to have acquired *IG* mutations by somatic hypermutation outside the GC [60]. Thus, the presence of *IG* mutations in a lymphoma does solely indicate that it derives from cells which had passed through a differentiation stage during which somatic hypermutation occurred [61]. Furthermore, some lymphoma show ongoing somatic hypermutation of the *IG* genes due to the high AID activity, as has been described for FL [62] and germinal center like DLBCL [63].

Later on, gene expression analyses of lymphomas in comparison to normal B-cells of different maturation stages have been used to define the cell of origin. This was based on the finding that at

distinct stages of differentiation a specific set of marker genes is expressed [64],[65]. Using this approach, the majority of mature B-cell lymphomas including BL have been described to harbor a gene expression signature similar to normal germinal center B-cells [66],[67].

Victora *et al.* [68] have shown that based on differential expression of a few genes, dark zone cells can be differentiated from light zone cells. Included in this gene expression signature were for example *CXCR4*, *CCND3*, *AURKA* which were up-regulated in DZ cells as well as *CD40*, *MYC* or *EGR3* which were up-regulated in LZ cells [68]. Using these expression signatures, the majority of analyzed germinal center derived B-cell lymphomas like FL and DLBCL were shown to harbor a gene expression signature similar to light zone B-cells. The majority of analyzed BL (75 %) on the other hand have been shown to be more similar with regard to their gene expression signature to dark zone B-cells. Remarkably, the remaining 25 % of BL had a higher likelihood to derive from light zone B-cells [68]. Refer to Figure 4 for an overview on the cell of origin of germinal center B-cell lymphoma.

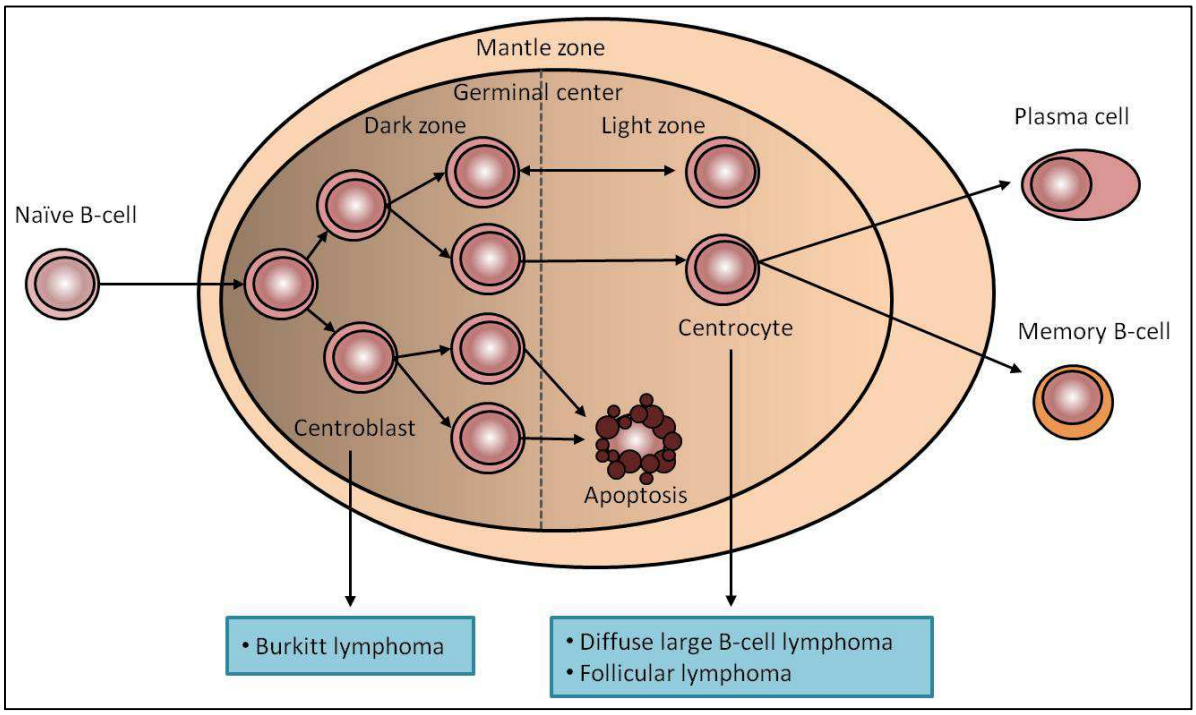


Figure 4: Cellular origin of germinal center B-cell lymphoma. The majority of Burkitt lymphomas have been suggested to derive from dark zone B-cells, whereas the diffuse large B-cell lymphomas and follicular lymphomas are supposed to derive from light zone B-cells [68]. Modified from [4].

As described in 1.2.1, MYC is an important regulator of the cyclic reentry of centrocytes into the dark zone. Strikingly, it is lowly expressed within cells of the dark zone. Thus, the finding that most of the BL derive from DZ B-cells seems paradox. The *MYC* translocation might take place within the LZ B-cells probably as a result of abnormal AID activity [69],[70]. This is in line with the finding that some of the BL harbor a LZ phenotype [68]. Those cells which harbor a high MYC expression due to the translocation might (re)enter the DZ. In the DZ the downregulation of MYC is either prevented due to

the loss of the BCL6 binding site during the translocation event. Alternatively or cooperating, the induction of MYC expression by the *IG* enhancers is stronger than the repression by BCL6 [71]. As a consequence, MYC expression is uncoupled from the positive selection within the light zone. Due to the reentry within the DZ, a DZ-specific gene expression pattern is activated which is reflected by the DZ phenotype in BL [68]. Among the genes of the DZ signature is the bHLH transcription factor TCF3 which induces several downstream pro-survival signaling pathways. These contribute to the tolerance against MYC induced apoptosis and, thus, the survival of the transformed GC B-cell [72]. Moreover, Schmitz *et al.* [72] identified activating *TCF3* mutations in BL, supporting the role of TCF3-activated pathways within the development of BL.

1.3 Molecular pathomechanisms in lymphomagenesis

As described above, MYC has been shown to be important for the germinal center reaction by regulating the cyclic reentry of activated B-cells. In BL, the *IG-MYC* translocation and subsequent MYC overexpression has been described as its hallmark alteration. In the following, the mechanisms and the role of MYC deregulation in BL will be introduced as well as tumorigenic pathomechanisms which could be already linked to Burkitt lymphomagenesis. Initially, the general concepts of oncogene activation and tumor suppressor gene inactivation will be presented.

1.3.1 Oncogene activation

Genes, which promote the tumorigenesis by mediating a proliferative advantage upon overexpression or –activity are called oncogenes [73]. Thus, the general function of oncogenes is to promote the proliferation and the consequent uncontrolled cell division further supports the accumulation of genomic lesions. Oncogenes belong to different protein families like growth factors and their receptors, proteins with GTPase-activity, tyrosine kinase receptors, DNA-binding transcription factors and proteins of the cell cycle control like cyclin dependent kinases or kinase inhibitors. Furthermore, oncogenes can be classified depending on their physiological functions in proliferation or differentiation [74].

An oncogene becomes activated by a dominant gain-of-function [75], meaning that the alteration of one allele is sufficient to exert the oncogenic effect. Different gain-of-function mechanisms are known: (i) Amplification of a gene, not changing the gene structure, might contribute to its overexpression [76]. (ii) A translocation might place the proto-oncogene under the influence of the promoter or enhancer of another gene. This could lead either to a constitutive overexpression, like in the case of translocations involving the *IG* locus, or to an expression in tissues or in differentiation stages in which the gene is normally silenced. Furthermore, the fusion of two genes can lead to the

generation of an abnormal protein which is more active than the non-fused genes as in the case of the BCR-ABL1 fusion due to the so called Philadelphia translocation [77]. (iii) Various types of mutations can be activating by increasing the stability or functionality of the protein. Alternatively, like in the case of a nonsense mutation, mutations might lead to the ablation of interaction domains with binding partners or regulatory sites [75]. (iv) Finally, several genes including *MYC* [78] have been described to become activated upon insertion of tumor viruses.

Classical proto-oncogenes in B-cell lymphoma are *MYC* [79],[80], *BCL2* [81],[82], *CCND1* [83],[84] and *REL* [85].

1.3.2 Tumor suppressor gene inactivation

Tumor suppressor genes are important negative regulators of the cell cycle as well as cell division and, thus, are important counter-actors of oncogenes [86]. Tumor suppressors control genes which are essential for regulating progression through the cell cycle, connect the cell cycle to the DNA damage response and promote apoptosis in case of not repairable damage [75]. Additionally, they are caretakers of the genomic integrity and, hence, reduce the mutation frequency which influences the cancer onset. Thus, important tumor suppressors belong to genes encoding proteins important for genome stability (*ATM*) [87], to negative regulators of cell cycle (*CDKN2A*, *TP53*) [88],[89] and to chromatin remodelers (*ARID1A*) [90],[91].

According to Knudson´s two hit theory [92], both alleles of a tumor suppressor need to be inactivated before its function is lost and the cells become constitutively mitotic active. This is nowadays partially obsolete. Recent findings have shown that the loss of one allele can give the cells for example a selective growth advantage [93],[94],[95] if the expression of the remaining allele cannot reconstitute the wildtype condition [96]. This mechanism is called haploinsufficiency.

Several inactivating mechanisms can lead to a so called loss-of-function: (i) Missense mutations might alter functional important domains of the protein; (ii) nonsense mutations truncating the protein might lead to a loss of important domains or to nonsense mediated decay (NMD); (iii) frameshift deletions or insertions might disrupt the gene structure and (iv) epigenetic mechanisms might silence the transcription [75].

Initial descriptions of tumor suppressors were based on the finding that they were localized within focal homozygous deletions [97],[98]. In line with this most of the tumor suppressors are inactivated due to a deletion of one allele or loss of heterozygosity whereas the remaining allele harbors inactivating mutations or was epigenetically silenced. Some tumor suppressors are already mutated in the germline, leading to a higher cancer predisposition, as in the case of *ATM* germline mutations which are associated with the development of lymphomas [99],[100]. Typical tumor suppressor genes in B-cell lymphoma are *TP53* [101],[102] or *ATM* [103].

1.3.3 MYC deregulation in Burkitt lymphoma

1.3.3.1 Mechanisms of MYC deregulation in Burkitt lymphoma

The genetic hallmark of BL is the translocation of the *MYC* gene which is identified in 90-100 % of cases [104],[105]. The translocation leads to the juxtaposition of the *MYC* gene to sequences from immunoglobulin (IG) genes. In 80 % of the BL the translocation partner is the immunoglobulin heavy chain (*IGH*) gene whereas in 20 % the translocation partner is either the gene encoding the immunoglobulin *κ* or *λ* light chain locus. Three classes of breakpoints in *MYC* exist and the position of the breakpoints correlates with the BL variants: Class I breakpoints map within the first exon or intron of the *MYC* gene. This class predominantly occurs in sporadic as well as in AIDS-associated BL. Class II breakpoints map immediately within a few kilobases (kb) 5' to the *MYC* gene. These are predominantly detected in endemic BL. Class III breakpoints are located in some distance upstream to the *MYC* gene [105].

The breakpoints in the light chain gene loci map 5' of the constant region gene segments [104]. The breakpoints in the *IGH* locus differ between the BL variants. In eBL the breakpoints occur mainly in the J_H gene segments, indicating that the *IG-MYC* translocation was created during a VDJ rearrangement as suggested by Haluska *et al.* [106],[107]. This further suggests that this initiating event of the eBL does not occur in the germinal center itself which is the supposed cell of origin (1.2.3) but rather in bone marrow precursor B-cells. In contrast, the breakpoint region in sBL and AIDS-associated BL was determined to fall within the *IGH* sμ switch region [104]. This indicates that these translocations are created during class switch recombination in the light zone of the germinal center. In line with this, several studies have shown that the induction of *IG-MYC* translocations *in vitro* and *in vivo* are dependent on AID which induces double-strand breaks during class switching and somatic hypermutation [69],[70]. Hence, although several studies described BL as a homogenous lymphoma entity [67],[108], it seems that the B-cell stage, at which the malignant transformation is induced, differs between eBL and sBL or AIDS-associated BL.

Subsequent to the *IG-MYC* translocation, the expression of MYC is regulated by the enhancers of the *IG* genes, which are especially active in mature B cells [109],[110]. Of note is that the "normal" *MYC* allele is usually transcriptionally silenced in BL and the transcribed MYC derives solely from the translocated allele [109],[110]. *MYC* itself becomes target of the somatic hypermutation machinery by AID due to the translocation into the *IG* locus and the subsequent increased transcription [111]. The mutations within *MYC* contribute to its deregulation, as mutations within the promoter region, first exon or intron prevent the inhibition of *MYC* transcription by other transcription factors as for example BCL6 (1.2.3) [112],[113]. Furthermore, mutations within the *MYC* coding sequence, like e.g.

mutations affecting threonine 58 prevent its degradation induced by GSK3β and, thus, increase the stability of the MYC protein [114].

1.3.3.2 Impact of MYC deregulation in Burkitt lymphoma

All of the above described genetic alterations lead to the overexpression of MYC which belongs to the basic helix-loop-helix (bHLH) family of transcription factors [115]. The *MYC* gene, also named *C-MYC* to differentiate it from the other *MYC* family members *N-MYC* and *L-MYC* [116], consists of three exons [117]. Only the second and third exons are coding. The long first exon contains several negative regulatory sequences. The first intron harbors binding sites for nuclear proteins [104]. MYC has been proposed to be a general amplifier of the active transcriptional program of the cell [118],[119]. This might explain the observation that the MYC-regulated transcriptional gene profile differs in the various studied cell types [120]. In contrast, around one third of the MYC target genes are transcriptionally repressed [118]. But this might be an indirect effect, as for example EZH2 is upregulated by MYC but represses subsequently the expression of other genes [121]. Nevertheless, the main functions of MYC are implicated within the cellular homeostasis. Thus, it regulates as a heterodimer complex with the MYC-associated factor X (Max) the transcription of genes involved in protein synthesis, cell cycle, translation or proliferation [122],[123]. Paradoxically, among the MYC target genes are the pro-apoptotic genes *TP53* and *CDKN2A* ($p14^{ARF}$) which are directly or indirectly upregulated [124],[125]. This is thought to be a protective regulatory feedback mechanism by the cell to counteract the oncogenic activation and prevent the transformation of the cell. Thus, the MYC overexpression in primary cells including lymphocytes results in apoptosis or cell cycle arrest [125]. To overcome this pro-apoptotic program, the cells need to acquire secondary alterations to survive and transform. In line with this, inactivating mutations within the TP53-ARF-MDM2 tumor suppressor pathway have been described for BL [126]. The necessity for secondary alterations to induce lymphomagenesis in addition to the MYC deregulation is further supported by the study of transgenic Eμ-MYC mouse models. In these mouse models the human *IG-MYC* translocation is mimicked leading to the development of pre B-cell and B-cell lymphoma [127],[128]. Furthermore, Park *et al.* [129] have shown in their transgenic Eμ-MYC mouse model that the development of B-cell and plasma neoplasms was accompanied by inactivating mutations of the already above described TP53-ARF-MDM2 pathway. In line with this, further studies in Eμ-MYC mice showed that mutations in for example *TP53* [130] accelerated the development of lymphoma. Thus, the *IG-MYC* translocation alone is not sufficient to induce full-blown lymphomagenesis. Instead, additional secondary alterations are necessary to induce Burkitt lymphomagenesis. In line with this, a few recurrent secondary alterations in BL have been described. On chromosomal level, recurrent copy number gains in chromosome 1q, 7 and 12 as well as losses in 6q, 13q32 and 17p were identified in BL

[131],[132]. On mutational level, BL have been shown to carry recurrent mutations in the TP53-ARF-MDM2 pathway as already described above. Recurrent activating mutations in *CCND3* have been identified in BL [133]. Moreover, epigenetic inactivation of the cell cycle regulator *CDKN2A* has been described in BL [134]. A major drawback of those analyses was that only single genes were analyzed for recurrent mutations. Thus, usage of high throughput sequencing techniques will help to decipher the global mutational landscape and hence to identify genetic mutations contributing in addition to the *IG-MYC* translocation to Burkitt lymphomagenesis.

As mentioned in 1.2.3, the majority of BL have been suggested to derive from the DZ. This is in contrast to the finding that *MYC*-translocations of sBL seem to occur during class switch recombination in the LZ of the germinal center (1.3.3.1). But as MYC plays a role in the cyclic reentry of the cells (1.2.1), it is conceivable that *IG-MYC* translocation induced MYC overexpression leads to a reentry of the cells in the DZ. In the DZ, MYC augments the transcription of those genes within the B-cells which are already actively transcribed at this stage. Thus, the amplification of the transcriptional program of the centroblasts might hence lead to the DZ cell phenotype of BL [135].

Taken together, the MYC overexpression in BL is a key oncogenic mechanism augmenting the proliferative and survival pathways whereas further alterations are needed for full transformation of the B-cells.

1.3.4 Other molecular pathomechanisms in Burkitt lymphoma

In addition to the above described chromosomal and genetic alterations, several additional molecular pathomechanisms, most of them induced by MYC, have been described to contribute to the Burkitt lymphomagenesis. Refer to Figure 5 for a schematic overview on the already described MYC-related functions (1.3.3) and the associated pathomechanisms in BL which will be presented in the following.

The attrition of the telomeres is an intrinsic life span clock. Thus, when the telomeres reach a critical length after several rounds of replication, the cells normally progress into senescence [136]. The telomerase activity is normally restricted to embryonic cells as well as to some fast dividing cells of the immune system as for example the extensive proliferating cells of the germinal center [137],[138]. Due to the high telomerase activity those cells are protected from telomere shortening. Lymphoma cells overcome this restricted telomerase activity as has been described for BL by Klapper *et al.* [139]. Hence, BL have been shown in comparison to for example FL or DLBCL to harbor a significant higher activity of the telomerase [139]. Furthermore, the high activity correlated with a high expression of the telomerase reverse transcriptase (TERT) in BL which is the catalytic subunit of the telomerase [140],[141],[142]. The high expression of TERT might be a result of the overexpression of MYC in BL, as a MYC-dependent activation of TERT expression in lymphocytes has

been described [143],[144]. Hence, this finding indicates a link between MYC deregulation and increased telomerase activity, contributing to the immortalization of BL cells.

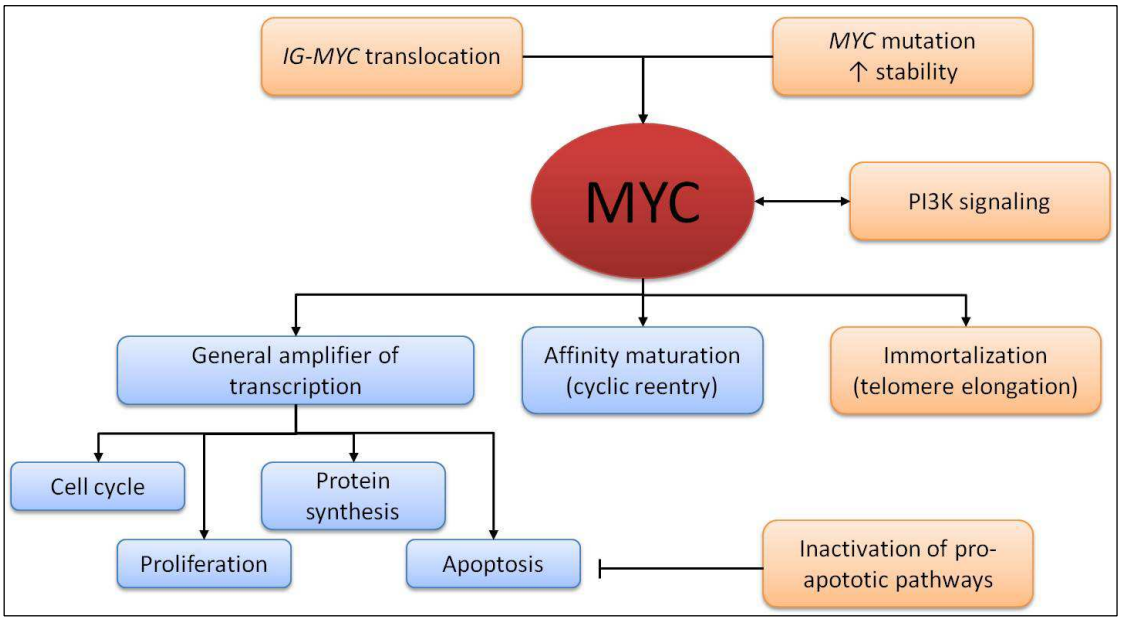


Figure 5: Schematic overview on MYC-related functions with regard to the pathomechanisms in BL. Blue boxes depict the physiological functions of MYC as general amplifier of transcription inducing cell cycle, proliferation, protein synthesis and apoptosis, as well as its role in the cyclic reentry in the germinal centre during affinity maturation. Alterations and pathomechanisms related to BL with regard to MYC are depicted by orange boxes.

The infection with EBV is a risk factor for the development of BL. Almost in all cases of eBL, as well as in 15-30 % of sBL, the genome of this virus is detectable [16]. The role of EBV in BL is still unclear. Although EBV has been shown to be able to induce the proliferation of resting human B cells [145], it is not sufficient to induce their malignant transformation as has been shown in a mouse model [146]. It is thought that EBV inhibits apoptotic signal pathways in BL and might hence contribute to the lymphomagenesis [147],[148],[149].

Furthermore, an oncogenic role for the PI3 kinase (PI3K) pathway has been implicated in the pathogenesis of BL. A link between BL and PI3K activity has been long proposed [150],[151],[152] as well as the dependence of BL on the signalling for proliferation and survival [153]. But a study by Sander *et al.* was the first to show in a mouse model that constitutively expressed MYC together with PI3K activity in B-cells undergoing GC reaction leads to the formation of BL-like tumors [133]. Further evidence for a role of PI3K in BL comes from the analysis of gene expression signatures of the PI3K activity which is higher in human BL biopsies than in other lymphoma subtypes [72],[133]. Figure 6 gives a schematic overview on the PI3K signaling in B-cells and BL which is shortly outlined in the following. Interaction of an antigen with the BCR or in the case of BL tonic BCR signaling [72] activates the tyrosine kinase SYK. This kinase activates the co-receptor CD19 which is one of the main regulators of PI3K activity in B cells [154] and the B-cell adaptor for PI3K (BCAP). Both activate

subsequently PI3K. Upon activation, PI3K produces phosphatidylinositol 3,4,5-triphosphate (PIP_3) from phosphatidylinositol 4,5-diphosphate (PIP_2) which is a lipid second-messenger. PIP_3 mediates the recruitment of effector proteins including the serine-threonine protein kinase Akt, also known as protein kinase B. In this manner Akt becomes activated and promotes via various downstream effectors cell growth, survival and differentiation [155].

A counter-actor of the PI3K signaling is the phosphatase and tensin homolog (PTEN). PTEN dephosphorylates PIP_3 and thereby inhibits the PI3K-induced signaling pathway. Thus, PTEN is an important tumor suppressor gene frequently inactivated in many solid tumors as well as in lymphoma [156],[157]. Recently, Schmitz *et al.* [72] have described sporadic mutations of *PTEN* in BL, likely augmenting the PI3K activity in these lymphomas.

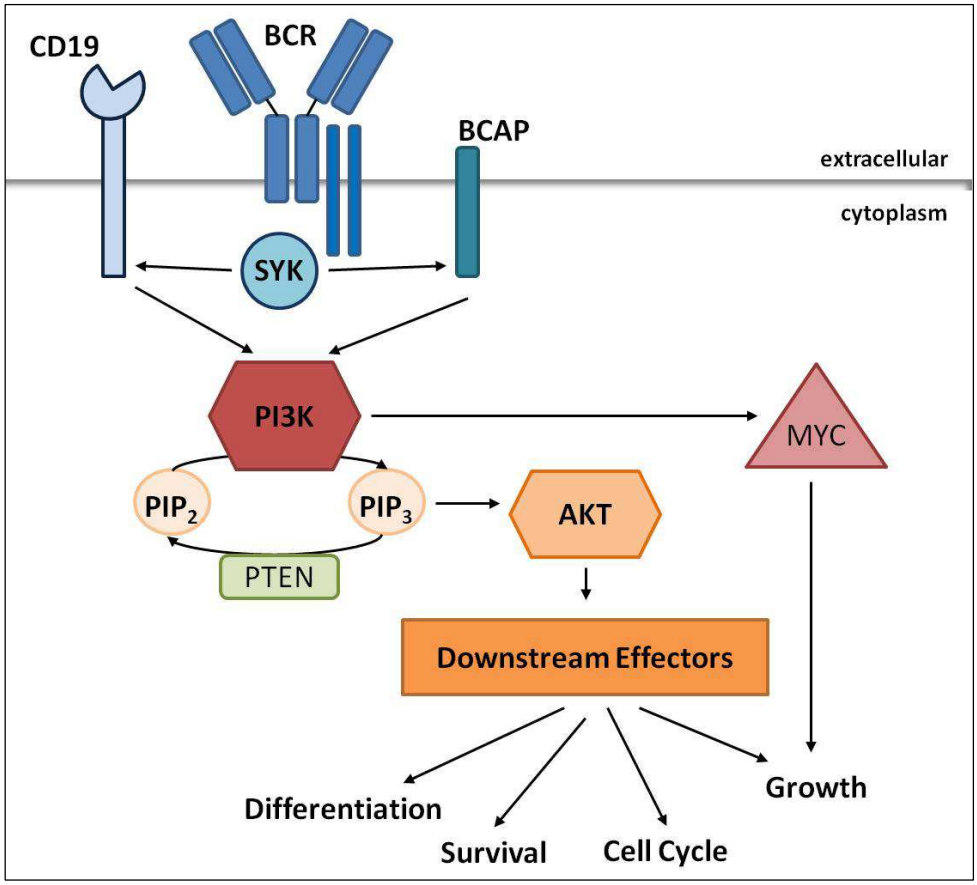


Figure 6: Schematic overview on PI3K activation and signaling in B-cells and BL. Stimulation of the BCR activates PI3K signaling via CD19 and BCAP. Active PI3K synthesizes PIP_3 which supports the recruitment of effector proteins to PI3K including Akt. In this manner, Akt becomes activated and induces several downstream effectors mediating differentiation, cell cycle and growth as well as the inactivation of anti-proliferative proteins and, thus, supports the survival. A counter-actor of PI3K is PTEN, which dephosphorylates PIP_3 to PIP_2. Furthermore, PI3K signaling has been shown to stabilize the MYC protein, thus, promoting its tumorigenic effect. Modified from [158].

In addition to mediating the downstream signaling pathway, PI3K has been shown to promote the stabilization of MYC via inhibition of GSK3β which mediates its degradation [159]. This finding further supports an oncogenic role for the cooperation of PI3K signaling with MYC deregulation.

Although all above described mechanisms have been linked to Burkitt lymphomagenesis, the underlying mechanisms leading for example to the tonic BCR signaling and, thus, to the constant activation of PI3K signaling remain unclear. It might be possible that secondary genetic alterations like inactivating or activating mutations of the PI3K pathway members might contribute to the PI3K activity. Thus, the study of the complete mutational landscape in BL might be a conclusive approach to identify such secondary alterations contributing to Burkitt lymphomagenesis. In line with this, the systematic studies of more than 25,000 cancer genomes of in total 50 different cancer types including BL is the aim of the International Cancer Genome Consortium [160].

1.3.5 MYC properties in stemness

MYC is one of the four reprogramming genes sufficient for transformation of somatic cells to inducible pluripotent stem cells [161]. The induction of pluripotency in somatic cells shares several molecular mechanisms with malignant transformation. Remarkably, it has been shown in studies of chicken´s bursa of Fabricia that myc overexpression induces preneoplastic lesions which subsequently give rise to lymphoma in a cell population which is only present during B-cell differentiation at the bursal stem cell age [162],[163]. Furthermore, the incidence peak of BL between the ages of 5 and 10 years at diagnosis might indicate an early initiating transforming event like for example in neuroblastoma which have a median incidence peak at 2 years of diagnosis. Based on this, it is conceivable that pediatric BL might derive from or resemble B-lymphoid cells with –potential- stem cell features.

1.4 *MYC*-negative Burkitt-like lymphoma

In addition to *MYC*-positive Burkitt lymphoma, the newly identified subtype of the *MYC*-negative Burkitt-like lymphoma was extensively studied within this thesis. The identification and characterization of this new lymphoma subtype is presented in the following.

1.4.1 Identification of *MYC*-negative Burkitt-like lymphoma

As described above the hallmark of BL is the *IG-MYC* translocation. There is an ongoing discussion if a subset of true BL might exist lacking an *IG-MYC* or non-*IG-MYC* translocation and, thus, are *MYC*-negative. In the literature 2-22 % of BL have been described to be MYC-negative [67],[108], [164],[165]. The discrepant frequencies might be due to a high rate of false-negative classification, as the breakpoints in the *MYC* and *IG* loci scatter widely. This renders the identification of a *MYC* break difficult and sometimes impossible unless an extensive set of FISH probes is used. Thus, the discrimination between true *MYC*-positive and *MYC*-negative BL is difficult. Recently, our group at

the Institute of Human Genetics, Kiel has shown the existence of a group of *MYC*-negative Burkitt lymphomas which displays a recurrent pattern of alteration on chromosome 11q [166]. In the following the approach leading to the identification of this lymphoma subtype is presented.

Initially, two gene expression classifiers: the mBL index [108] and the BL-PAP index [167], were applied to identify molecular BL (mBL). A total of 58 genes make up the mBL gene expression signature which has been defined by Hummel *et al.* [108]. This mBL index is used to classify mature aggressive B-cell lymphomas in mBL (mBL index >0.95), non-BL (mBL index<0.05) and intermediate B-cell lymphoma (0.05≤mBL index≤0.95) [108]. The BL-PAP index developed by Bentink *et al.* [167] identifies BL (BL-PAP) as well as four distinct groups of non-BL based on a signature of eight oncogene-inducible pathways. Applying these classifiers on the gene expression of the 754 aggressive B-cell lymphomas of the Molecular Mechanisms in Malignant Lymphoma (MMML) network in combination with data of their *MYC* translocation status, revealed that 90 % of the mBL and 92 % of the BL-PAP carried an *IG-MYC* translocation or non-*IG-MYC* translocation. Hence, a total of 6 mBL and 4 BL-PAP cases lacked a *MYC*-translocation. After excluding those cases which were only defined as molecular BL by one classifier, two cases lacking a *MYC*-translocation remained. Interestingly, both cases harbored a peculiar pattern of 11q22-q24 copy gain and 11q24-qter copy loss. In line with this, cytogenetic studies have described *MYC*-negative lymphoma classified as either intermediate lymphoma or BL which also harbored recurrent abnormalities within chromosome 11 such as dup(11)(q23q13) or der(11)(q) [168],[169]. Including five *MYC*-negative BL of the Pienkowska-Grela *et al.* study [168] and seven additional *MYC*-negative lymphomas identified at our Institute, a cohort of 14 B-cell lymphomas was used for characterization of the alteration on chromosome 11q. All but one case harbored an interstitial gain in 11q23.2-q23.3 within a minimal region of gain chr11:114,530,818-117,939,359 bp (hg18). Four cases harbored a focal high-level amplification in this minimal region in 11q23.3 leading to the detection of a minimal region of amplification in chr11:117,107,631-117,939,359 bp (hg18). Furthermore, all cases harbored a telomeric loss in 11q24.1-qter, defined as minimal region of loss chr11:126,977,015-134,445,937 bp (hg18). Strikingly, one case harbored a focal region of homozygous loss located at chr11:127,322,011-128,846,569 bp (hg18). After defining the aberration on 11q, the array-comparative genomic hybridization (CGH) data of cases from the MMML cohort have been mined for these alterations leading to the identification of three additional cases lacking a *MYC*-translocation. Taken together, in total 17 *MYC*-negative B-cell lymphomas have been identified which harbored a recurrent gain-and-loss pattern on chromosome 11q.

1.4.2 Molecular and pathological characteristics of *MYC*-negative Burkitt-like lymphoma

To further characterize these 17 B-cell lymphomas with 11q alteration, diverse molecular and pathological data were collected. The lymphomas were identified in children and young adults (median age at diagnosis: 16 years ranging from 6 to 76 years). The majority of patients were male (male:female ratio 4.7:1). The survival was excellent. In comparison to an age-matched group of *MYC*-positive BL (<40 years at diagnosis) no significant differences regarding the prognosis were observed.

The presentation of *MYC*-negative lymphoma was more often nodal (82 %, 14/17) than in patients with BL younger than 40 years at diagnosis (55 %) [166]. Other than that, the cases resembled BL with regard to (i) morphological features including "starry sky" pattern, many mitoses and few infiltrating reactive lymphocytes, (ii) cytogenetic features including lack of *IGH-BCL2* translocations and *BCL2* breaks and (iii) immunophenotype being $CD20^+$, $CD10^+$, $BCL2^-$ and showing Ki67 positivity in >90 % of cells. Based on the available gene expression data from six cases of the MMML cohort the cell of origin could be defined as germinal center-like as it has been for BL (1.2.3). Furthermore, those cases showed a low MYC expression, confirming that *MYC* is not genomically activated in those cases. The array-CGH data showed a significantly higher genomic complexity in the *MYC*-negative lymphomas than in mBL and DLBCL. Concomitant to the 11q alteration, recurrent gains were observed in 7q34-qter, 12pter-p12.2, 18q21.2, 19pter-p13.2 and a recurrent loss in 6q14.3-q22.2.

All these data show that the *MYC*-negative lymphomas resemble more BL than DLBCL, suggesting that this group might represent a genetic variant of the BL. Hence, they were termed "*MYC*-negative high grade B-cell lymphoma resembling Burkitt lymphoma" or shortened as *MYC*-negative Burkitt-like lymphoma (mnBLL).

The alteration on chromosome 11q consists of a copy number gain and loss. Thus, it is conceivable that due to the amplification an oncogene is activated whereas due to the loss a tumor suppressor becomes inactivated which might provide to the lymphomagenesis of mnBLL. Combining gene expression data of the cases from the MMML cohort with the copy number data led to the identification of a few potent candidate genes like *PAFAH1B2* or *ETS1*. Nevertheless, the underlying pathomechanisms of these potential candidate genes remained unclear. Thus, the aims of the studies conducted within this thesis are (i) to screen the region for additional potential candidate genes and (ii) to analyze the role of these potential candidate genes for the development or progression of mnBLL.

1.5 Aims of this thesis

The aim of this thesis is the identification and characterization of new functionally relevant candidate genes in two mature, aggressive B-cell lymphomas of the childhood: *MYC*-positive Burkitt lymphoma and *MYC*-negative Burkitt-like lymphoma.

The hallmark aberration of Burkitt lymphoma is the *IG-MYC* translocation. Nevertheless, it is known that secondary alterations are necessary to induce full-blown lymphomagenesis. Furthermore, as *MYC* is one of the four reprogramming factors sufficient to transform somatic cells to pluripotent stem cells, it was questioned if an association between pluripotency and BL does exist.

Thus, the aims of the analyses of *MYC*-positive Burkitt lymphoma were:

- Determination if *MYC*-positive Burkitt lymphoma possess potential traces of pluripotency
- Analysis of the expression of the alternative transcripts of the *TERT* gene and their implications for BL
- Identification of recurrently mutated genes and analysis of their tumorigenic role for BL

The existence of *MYC*-negative Burkitt-like lymphoma, abbreviated as mnBLL, has been recently described by our group. The main focus has been on the characterization of this new B-cell lymphoma subtype. Thus, the aims of the analyses of *MYC*-negative Burkitt-like lymphoma were:

- Analysis of the chromosomal landscape
  - Characterization of the alteration on chromosome 11q of newly recruited mnBLL
  - Definition of the minimal gain and loss region in 11q based on all mnBLL cases
  - Identification of further recurrent, secondary genomic alterations and involved genes
- Analysis of the mutational landscape
  - Identification of recurrently mutated genes within the altered 11q region indicating roles as oncogenes or tumor suppressors
  - Identification of secondary alterations potentially contributing to the lymphomagenesis of mnBLL
- Analysis of the transcriptional landscape
  - Identification of differentially expressed genes of the altered 11q region potentially contributing to the pathogenesis of mnBLL
- Characterization and functional analyses of possible candidate genes and their potential role in the lymphomagenesis of mnBLL

2 Materials and Methods

This chapter presents the materials used and methods applied in this thesis. Standard laboratory equipment, reagents and instruments are not mentioned. It is important to note, that the molecular genetic analyses at the Institute of Human Genetics of the Christian-Albrechts-University of Kiel were performed according to the determinations of the DIN EN ISO 15189:2007 and, thus, the respective protocols of the standard operating procedures have been used within this thesis.

2.1 Materials

2.1.1 Patient materials

Most of the material from patients analyzed within this thesis was provided in the framework of the Molecular Mechanism in Malignant Lymphoma (MMML) project (n=31 B-cell lymphomas) and the International Cancer Genome Consortium (ICGC) MMML-Seq project (n=27 B-cell lymphomas). Some of the Burkitt lymphomas (BL) as well as *MYC*-negative Burkitt-like lymphomas (mnBLL) which were analyzed in the framework of this thesis have been extensively studied in already published manuscripts [108],[166],[170]. Thus, the reader is referred to the respective publications for an overview on the molecular and pathological characteristics of those cases.

The ICGC MMML-Seq study was approved by the Institutional Review Boards of the Medical Faculty of the University of Kiel (A150/10) and of the recruiting centers. The protocols of the MMML network have been approved by central (University of Göttingen) and local review boards (Institutional Review Boards) of the Medical Faculty of the University of Kiel (D403/05). The lymphomas included in both projects were partly derived from the Berlin-Frankfurt-Münster (BFM) clinical trials which have also been approved by central and local institutional review boards.

2.1.1.1 Characteristics of *MYC*-positive Burkitt lymphomas

The material from 21 patients with Burkitt lymphoma (male to female ratio 6:1, median age at diagnosis 10 years, range 4-18 years) was provided for this study in the framework of the ICGC MMML-Seq (case identifier with seven-digit ID). Furthermore, material from 31 patients with Burkitt lymphoma (male to female ratio 2.75:1, median age 12.5 years at diagnosis, range 2-76 years at diagnosis) from the MMML project (case identifier starting with MPI) was studied. The tumor material for the analyses was obtained from the affected tissue of the patients´ lymph nodes, other lymphatic tissues or tumor manifestations. For an overview on the cases analyzed in this thesis refer to the appendices 8.1 and 8.2.

Moreover, whole-genome and transcriptome data, as well as the clinico-pathological data of the analyzed BL samples of the ICGC MMML-Seq project were available for further analyses in the framework of this thesis. Clinico-pathologic data of the analyzed cases of the MMML cohort were kindly provided by Dr. Markus Kreuz, Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig.

In this thesis the number of analyzed BL cases in the different projects differed as the recruitment of new cases within the ICGC MMML-Seq project has been still ongoing.

2.1.1.2 Characteristics of *MYC*-negative Burkitt-like lymphomas

Table 2 gives an overview on the characteristics of the mnBLL and the techniques applied for their characterization performed in this study. Some of those cases have been already described by our Institute [166].

Table 2: Overview on characteristics of mnBLL.

Case	Initial diagnosis	Age (yrs)	Sex	Tissue/ material	450k	Onco Scan	NGS
1[1,2] (MPI-626)	Aggressive B-NHL, BL	7	m	LN/ Cryo	x		x[3]
2[1]	BL	7	m	LN/ Cryo	x		x[4]
3[1]	High-grade B-NHL, atypical BL	16	f	LN/ Cryo	x		x[4]
8[1]	Mature B-cell lymphoma/ lymphoblastic B-lymphoma	11	m	LN/ Cryo	x		x[4]
MPI-078[1,2]	Centroblastic, DLBCL	8	m	LN/ Cryo	x		
MPI-086[1,2]	High-grade B-cell lymphoma	14	m	LN/ Cryo	x		
MPI-148[1,2]	Atypical BL	26	m	extranodal/ Cryo	x		
MPI-315[1,2]	Centroblastic, DLBCL	49	f	intestine/ Cryo	x		
18	Atypical BL	14	m	nasopharynx/ FFPE		x	
19	BL	4	m	LN/ FFPE		x	
20	Aggressive B-cell lymphoma	23	f	BM/ FFPE		x	
21	BL	8	m	LN/ FFPE		x	
22	BL	27	m	LN/ FFPE		x	
23	FL grade IIIb, DLBCL	17	m	LN/ FFPE		x	
24	ALL	10	m	BM/ FFPE		x	
25	B-cell lymphoma, unclassifiable	13	m	LN/FFPE		x	

Cases denoted with [1] were published as part of the initial cohort [166]. Cases denoted with [2] were analyzed within the framework of the MMML project. Age, patients´ age in years (yrs) at diagnosis; sex, m for male and f for female; material/tissue is denoted as LN for lymph node, BM for bone marrow, cryo for cryopreserved tissue and FFPE for formalin-fixed, paraffin-embedded tissue; initial diagnosis denotes the diagnosis of the patients before classification as mnBLL: BL, Burkitt lymphoma; DLBCL, diffuse-large B-cell lymphoma; FL, follicular lymphoma; NHL, non-Hodgkin lymphoma; ALL, acute lymphoblastic leukemia; Techniques used for the analyses: 450k, analysis of the global DNA methylation using HumanMethylation 450k Bead Chip; OncoScan, analyses of chromosomal imbalances using the OncoScan DNA microarray; NGS, next generation sequencing; [3] whole-genome sequencing performed within the framework of the ICGC-MMML-Seq project; [4] targeted resequencing using the TruSight™ One Sequencing panel.

The cases which were not part of the initial cohort (cases 18-25) were initially analyzed for the existence of the specific 11q alteration using fluorescence in situ hybridization (FISH) diagnostic

before further analyses were performed. Some cases have been identified in routine diagnostics in the laboratories of the Institute of Human Genetics, Kiel, Germany and by Prof. Dr. Wolfram Klapper (Section Hematopathology, Kiel, Germany), others were supplied by Dr. Elaine Jaffe (Laboratory of Pathology, National Cancer Institute, Bethesda, USA), Dr. Itziar Salaverria (Department of Pathology, Hematopathology Unit, Barcelona, Spain) and Dr. Kristian Schafernak (Ann & Robert H. Lurie Children´s Hospital of Chicago, Department of Pathology, Chicago, USA).

2.1.1.3 Sera from patients with Burkitt lymphoma or Burkitt-like lymphoma

Serum of a patient with mnBLL (case 10, 2.1.1.2) was used to analyze the IL10 level in comparison to sera of six patients with BL (BL controls) of the peripheral blood (pB) or bone marrow (bm). Characteristics of patients are given in Table 3. Sera were kindly provided by Dr. Christine Damm-Welk (Department of Pediatric Hematology and Oncology, Justus-Liebig University, Giessen) in the framework of the NHL-BFM and MMML project. Material for FISH diagnostics was available for the BL controls 1, 3-4 and 6. Those cases were screened at our Institute for the existence of a *MYC*-break and 11q aberration as part of routine FISH work-up.

Table 3: Characteristics of patients analyzed for IL10 levels in serum.

Case	Initial diagnosis	Age (yrs)	Sex	Material	*MYC*-break	11q alteration
Case 10	Pediatric FL[1]	6	f	pB EDTA	negative	positive[2]
BL control 1	BL	14	m	pB Heparin	positive	negative
BL control 2	BL	3	m	pB EDTA	na	na
BL control 3	BL	9	m	bm EDTA	positive	negative
BL control 4	BL	9	m	pB Heparin	positive	negative
BL control 5	BL	2	m	bm EDTA	na	na
BL control 6	BL	11	m	pB EDTA	positive	negative

Age denotes patient age in years (yrs) at diagnosis; sex is denoted m for male and f for female; initial diagnosis is denoted BL for Burkitt lymphoma and FL for follicular lymphoma; material is denoted pB for peripheral blood and bm for bone marrow; na, not analyzed. [1] as described in [166] case 10 was classified as follicular lymphoma displaying a follicular and diffuse growth pattern but the cytological features were Burkitt-like. Hence it was included in the series of *MYC*-negative Burkitt-like lymphoma. [2] as published in [166].

2.1.2 Cell lines

The identity of all cell lines used within this thesis has been validated by STR-profiling using the Stem Elite ID Kit as described in 2.2.5.5.

2.1.2.1 Burkitt lymphoma cell lines

All cell lines were obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Culture (Braunschweig, Germany). The only exception was the BL-30 cell line which was kindly provided by Dr. R. Schmitz, National Cancer Institute, NIH, Bethesda, USA. Cell line characteristics and first description are listed in Table 4.

Table 4: Characteristics of Burkitt lymphoma cell lines.

Cell line	Diagnosis	Age (yrs)	Sex	*MYC* translocation	EBV Status	Reference
BALM-16	Burkitt leukemia	42	m	t(8;22)(q24;q11) [1]	negative	[171]
BALM-18	Burkitt leukemia	35	m	t(8;14)(q24;q32) [2]	negative	[172]
BL-2	BL	7	m	t(8;22)(q24;q11) [1]	negative	[173]
BL-30	BL	19	m	t(8;14)(q24;q32) [2]	negative	[174]
BL-41	BL	8	m	t(8;14)(q24;q32) [1]	negative	[175]
BL-70	BL	16	m	t(8;14)(q24;q32) [1]	negative	[175]
BLUE-1	BL	29	m	t(8;14)(q24;q32) [1]	negative	[176]
Ca46	BL	16	m	t(8;14)(q24;q32) [1]	negative	[177]
DAUDI	BL	16	m	t(8;14)(q24;q32) [1]	positive	[178]
DG-75	BL	10	m	t(8;14)(q24;q32) [1]	negative	[179]
EB-1	BL	9	f	t(8;14)(q24;q32) [2]	positive	[180]
EB-3	BL	3	m	t(8;14)(q24;q32) [3]	positive	[181]
Namalwa	BL	3	f	t(8;14)(q24;q32) [2]	positive	[178]
Raji	BL	12	m	t(8;14)(q24;q32) [1]	positive	[182]
Ramos	BL	3	m	t(8;14)(q24;q32) [1]	negative	not published
U-698-M	BL	7	m	t(8;14)(q24;q32) [4]	negative	[183]

Age denotes patient age in years (yrs) when cell line was established, sex is denoted m for male and f for female, reference refers to the first description, information of *MYC* translocation: [1]from the DSMZ (http://www.dsmz.de/home.html, 10/11/2014), [2] published in Murga-Penas *et al.* [184], [3] published in Berger *et al.* [181], [4] analyzed at our Institute.

2.1.2.2 Non-Burkitt lymphoma cell lines

The cell line MLMA was obtained from the Japanese Collection of Research Bioresources Cell Bank (JCRB). All other cell lines were obtained from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Culture (Braunschweig, Germany). Cell line characteristics and first description are listed in Table 5. The data of the hallmark chromosomal aberrations were taken from the homepage of the DSMZ (http://www.dsmz.de/home.html, 10/11/2014). The exception was the data of the MLMA cell line which are in-house data.

Table 5: Characteristics of non-Burkitt lymphoma cell lines.

Cell line	Diagnosis	Age (yrs)	Sex	Hallmark chromosomal aberration	Reference
Su-DHL-6	GCB-DLBCL[1]	43	m	t(14;18)(q32;q21) (IGH:BCL2)	[185]
Su-DHL-10	GCB-DLBCL[1]	25	m	der(14)t(8;14)(q24;q32) (IGH:MYC), der(18)t(14;18)(q32;q21) (IGH:BCL2)	[185]
Karpas422	GCB-DLBCL[1]	72	f	t(14;18)(q32;q21) (IGH:BCL2)	[186]
RIVA	ABC-DLBCL[1]	57	f	der(18)amp(18)(q21)dup(18)(q21)	[187]
Su-DHL-5	GCB-DLBCL[1]	17	f	+12, del(6)(q13)	[185]
HT	GCB-DLBCL[1]	70	m	dup(11)(?q23qter)	[188]
MLMA	malignant lymphoma[3]	32	f	der(11)dup(11)(q22q25)hsr(11)(q22~23)	not published
L-428	Hodgkin lymphoma[2]	37	f	complex hyperdiploid karyotype	[189]
MC-116	B-cell lymphoma[2]	n.d.	n.d.	t(8;14)(q24;q32) (IGH:MYC)	[177]
OCI-LY7	DLBCL[1]	48	m	t(8;14)(q24;q32) (IGH:MYC)	[190]
Hek293	Human embryonic kidney cell[2]	n.d.	n.d.	none	[191]

Diagnosis are denoted GCB-DLBCL for germinal center B-cell like diffuse-large B-cell lymphoma and ABC-DLBCL activated B-cell like diffuse-large B-cell lymphoma; age denotes patient age in years (yrs) when cell line was established; sex is denoted m for male and f for female; Reference refers to the first description of the cell lines; n.d., no data available. [1] diagnosis as defined by Schuetz *et al.* [192]; [2] diagnosis as defined on the homepage of the DSMZ (http://www.dsmz.de/home.html, 10/11/2014), [3] diagnosis as defined in this thesis as described in 3.2.2.1.

2.1.2.3 Characterization of a cell line panel used for functional studies

All functional experiments described in 3.2 were performed using a cell line panel consisting of six Burkitt lymphoma cell lines (BL-2, BL-41, BL-70, BLUE-1, Ca46, U-698-M), four diffuse-large B-cell lymphoma cell lines which can be further separated in three germinal center B-cell-like DLBCL (Su-DHL-6, Su-DHL-10, Karpas422) and one activated B-cell-like DLBCL (RIVA), as well as three in 3.2.2 described mnBLL cell line models (HT, Su-DHL-5, MLMA). All DLBCL and BL cell lines were chosen based on the criteria that those are cell line models frequently used and published by the scientific community. Moreover, the BL and DLBCL cell lines were analyzed for the presence of the mnBLL-typical gain and loss pattern on chromosome 11. To this end available SNP 6.0 array data from the COSMIC database and the available karyotype data from the DSMZ (http://www.dsmz.de/home.html, 10/11/2014) were mined for chromosomal aberrations on chromosome 11 as summarized in Table 6. Refer to the appendix 8.3 for the complete karyotypes of the cell lines. Furthermore, all BL and DLBCL cell lines were screened using a FISH assay specific for the gain and loss pattern on chromosome 11. Table 6 gives an overview on the characteristics of the reference cell line panel.

Table 6: Overview on characteristics of the BL and DLBCL cell line of the reference cell line panel.

Cell line	Diagnosis	Karyotype aberration involving chr11q23-q25	SNP 6.0 array chromosome 11q	11q gain/loss (FISH)[1]	Typical 11q alteration
BL-2	BL	del(11)(q24.2)	n.a.	subclonal del(11q25)	no
BL-41	BL	negative	negative	negative	no
BL-70	BL	negative	negative	negative	no
BLUE-1	BL	negative	n.a.	negative	no
Ca46	BL	negative	negative	negative	no
U-698-M	BL	dup(11)(q23q13)	chr11:108038477-111222713,gain	negative	no
Su-DHL-6	GCB-DLBCL	dup(11)(q24q25)	chr11:127435602-127770927,loss	11q25 loss	no
Su-DHL-10	GCB-DLBCL	der(11)t(Y;11)(q11;q25)	negative	negative	no
Karpas422	GCB-DLBCL	negative	chr11:110931477-117046760,gain chr11:125071926-125083705,loss	11q24 gain, 11q25 loss	yes
RIVA	ABC-DLBCL	negative	n.a.	negative	no

Karyotype data available at the DSMZ homepage at (http://www.dsmz.de/home.html, 10/11/2014); SNP 6.0 array data of chromosome 11, available at COSMIC database (http://cancer.sanger.ac.uk/cancergenome/ projects/cell_lines/, 10/11/2014); n.a., not available; 11q gain/loss (FISH), screening for 11q aberration pattern using FISH; positive, detection of 11q aberration; negative, no detection of 11q aberration. [1]The FISH analysis was performed as part of the routine FISH work-up at the Institute of Human Genetics, Kiel and the data were kindly provided by Dr. med. Susanne Bens.

2.1.3 Cell culture media

Table 7: Cell culture media used in this study.

Reagent	Manufacturer
D-MEM	GIBCO Life Technologies, Carlsband, USA
FBS, South American (CE)	GIBCO Life Technologies, Carlsband, USA
L-Alanyl-L-Glutamin (200 mM)	Biochrome AG, Berlin, Germany
RPMI-1640	GIBCO Life Technologies, Carlsband, USA

2.1.4 Antibodies

Table 8: Antibodies used in this study.

Antibody [clone]	Manufacturer
Primary antibodies:	
CDw210 [3F9]	BD Bioscience, Franklin Lake, New Jersey, USA
ETS1 [EPR546(2)]	Epitomics, California, USA
FLI1 [C19]	Santa Cruz Biotechnology, Texas, USA
GAPDH [GT239]	GeneTex, Irvine, CA, USA
Goat IgG Control-PE	R&D Systems, Minneapolis, USA
ID3 [C20]	Santa Cruz Biotechnology, Texas, USA
IL10 (ab34843)	Abcam, Cambridge, UK
IL10RA (T3431)	Epitomics, California, USA
IL-10Rα-Phycoerythrin (FAB6280P)	R&D Systems, Minneapolis, USA
Lamin B1 (ab16048)	Abcam, Cambridge, UK
PAFAH1B2 (M01A) [2FA-1C10]	Abnova, Taipei City, Taiwan
β-Actin [AC-15]	Sigma Aldrich, St. Louis, USA
Secondary antibodies:	
Donkey anti-Mouse IgG (H&L), HRP Conjugate	Immuno Reagents, Inc., Raleigh, USA
Donkey anti-Rabbit IgG (H&L), HRP Conjugate	Immuno Reagents, Inc., Raleigh, USA

2.1.5 Primers for Sanger sequencing

The initial step in the primer design was the download of the genomic sequence of the gene from the UCSC genome bioinformatics site (http://genome.ucsc.edu/). The primers were designed to have an optimal length of 18-20 bp, a GC content between 45-60 % and to harbor no repeats or run of a single bases. The uniqueness of the primers in the human genome was rechecked by BLAT analyses (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) against the NCBI build 37/hg19. Oligonucleotides were obtained from Biomers, Ulm, Germany as lyophilisates and resuspended in ddH_2O to a final concentration of 200 pmol/µl.

2.1.5.1 Primers for Sanger sequencing of *TP53*

The following primers were designed to screen B-cell lymphoma cell lines for mutations within exon 4-9 of the *TP53* gene on chromosome 17p13.1.

Table 9: Primer sequences for Sanger sequencing of *TP53*.

***TP53* chromosome 17p13.1**					
Primer name	**Exon**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp** / **End (hg19) in bp**	**Product Size**	**Annealing temp.**
TP53_Ex4_F	4	CGTTCTGGTAAGGACAAGGG	7,579,223	484 bp	60 °C
TP53_Ex4_R		GGAATCCCAAAGTTCCAAAC	7,579,706		
TP53_Ex5-6_F	5-6	TAGTGGGTTGCAGGAGGTG	7,578,076	594 bp	65 °C
TP53_Ex5-6_R		TCAAATAAGCAGCAGGAGAAAG	7,578,669		
TP53_Ex7_F	7	CCTGCTTGCCACAGGTC	7,577,384	288 bp	55 °C
TP53_Ex7_R		GAGGTGGATGGGTAGTAGTATGG	7,577,671		
TP53_Ex8-9_F	8-9	TGGTTGGGAGTAGATGGAGC	7,576,758	492 bp	60 °C
TP53_Ex8-9_R		GCCCCAATTGCAGGTAAAAC	7,577,249		

2.1.5.2 Primers for Sanger sequencing of *CCND3*

The following primers were designed to screen different B-cell lymphoma cell lines for mutations within the hot spot mutation region in exon 5 of the *CCND3* gene on chromosome 6p21.1.

Table 10: Primer sequences for Sanger sequencing of *CCND3*.

***CCND3* chromosome 6p21.1**					
Primer name	**Exon**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp** / **End (hg19) in bp**	**Product Size**	**Annealing temp.**
CCND3_5_F	5	CCATGTGTTGGGAGCTGTC	41,903,607	328 bp	65 °C
CCND3_5_R		CTGGAGGCAGGGAGGTG	41,903,934		

2.1.5.3 Primers for Sanger sequencing of *CDKN2A*

The following primers were designed to screen B-cell lymphoma cell lines for mutations in exon 3-5 of the *CDKN2A* gene on chromosome 9p21.3

Table 11: Primer sequences for Sanger sequencing of *CDKN2A*.

***CDKN2A* chromosome 9p21.3**					
Primer name	**Exon**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp** / **End (hg19) in bp**	**Product Size**	**Annealing temp.**
CDKN2A_Ex3_F	3	CACAAGCTTCCTTTCCGTCA	21,970,789	484 bp	55 °C
CDKN2A_Ex3_R		CTGAGGCAAGACCGGAGAC	21,971,272		
CDKN2A_Ex4_F	4	GCAAATGCTCCCTCAGGAAT	21,968,595	333 bp	63 °C
CDKN2A_Ex4_R		AAGATGTGGCCTTTCCCTTC	21,968,927		
CDKN2A_Ex5_F	5	TAGGGACGGCAAGAGAGGAG	21,968,162	166 bp	60 °C
CDKN2A_Ex5_R		CCTGTAGGACCTTCGGTGAC	21,968,327		

2.1.5.4 Primers for Sequencing of *ID3*

The following primers for the *ID3* gene on chromosome 1p36.12 were designed to screen B-cell lymphoma cell lines for mutations within this gene as has been published by Richter *et al.* [170]. The primer pair covers the complete coding region of the gene including parts of the 3'UTR and the complete 5'UTR as well as the complete intron 1.

Table 12: Primer Sequences for Sanger Sequencing of *ID3*.

***ID3* chromosome 1p36.12**				
Primer name	**Primer Sequence (5'… 3')**	**Start (hg19) in bp / End (hg19) in bp**	**Product Size**	**Annealing temp.**
ID3_F	TCCAGGCAGGCTCTATAAGTG	23,885,341	694 bp	65 °C
ID3_R	CCGAGTGAGTGGCAATTTTT	23,886,034		

2.1.5.5 Primers for Sanger sequencing of *CREBBP*

The following primers were designed to screen B-cell lymphoma cell lines for mutations in the *CREBBP* gene on chromosome 16p13.3.

Table 13: Primer sequences for Sanger sequencing of *CREBBP*.

***CREBBP* chromosome 16p13.3**					
Primer name	**Exon**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp / End (hg19) in bp**	**Product Size**	**Annealing temp.**
CREBBP_Ex1_F	1	GCTCCTCTCCCTCGCAG	3,929,763	253 bp	65 °C
CREBBP_Ex1_R		CCGGACGCTCTCTTTCAG	3,930,015		
CREBBP_Ex2a_F	2	CATAGAAACGTGGCAGTTGG	3,900,593	490 bp	60 °C
CREBBP_Ex2a_R		GTGGCAGGGCTGCTAGTC	3,901,082		
CREBBP_Ex2b_F		CCTAAACAGGCAGCCAGC	3,900,229	468 bp	60 °C
CREBBP_Ex2b_R		TTTACGCATTACTCGGAGGG	3,900,696		
CREBBP_Ex3_F	3	GAAACTGTGTGAGCATTTCCC	3,860,489	374 bp	55 °C
CREBBP_Ex3_R		CCTATCACCTACTGACACACTTTTAG	3,860,862		
CREBBP_Ex4_F	4	GTGGTCGGTATTATCCATCAGC	3,843,336	400 bp	65 °C
CREBBP_Ex4_R		GGCAAATTCTTCCTGACCTC	3,843,735		
CREBBP_Ex5_F	5	GGTCTGCCTATACTGTGTTATGG	3,841,838	390 bp	60 °C
CREBBP_Ex5_R		TGTACCTTGGGCTGCTGTC	3,842,227		
CREBBP_Ex6_F	6	AAATCATTCGTGGGCTTCTC	3,832,600	397 bp	60 °C
CREBBP_Ex6_R		CTGAAAACTGCCTTGGGTTC	3,832,996		
CREBBP_Ex7_F	7	ATGGTGGCATGTTGGTTATC	3,831,138	257 bp	60 °C
CREBBP_Ex7_R		CAGTTTTGTGTGGTTCTCAGTC	3,831,394		
CREBBP_Ex8_F	8	TGGTGGCAGAAGAACCTTAC	3,830,629	332 bp	60 °C
CREBBP_Ex8_R		AAGCACGTGACTTGTATAGGC	3,830,960		
CREBBP_Ex9_F	9	AGGTGATTCTCCCGCCTCAG	3,828,551	401 bp	65 °C
CREBBP_Ex9_R		ATCTGGGAAGTCTCCTTGGTC	3,828,951		
CREBBP_Ex10_F	10	GTTAGAACCTACAACACAGATCATTC	3,827,942	320 bp	60 °C
CREBBP_Ex10_R		ATACACCCCAAACACGAAGG	3,828,261		
CREBBP_Ex11_F	11	GCTTTTGGACCTATTGGCTG	3,827,410	378 bp	60 °C
CREBBP_Ex11_R		TTGGAAGATCTTTCCTATAATCTCC	3,827,787		
CREBBP_Ex12_F	12	TTCTGTTGCCTGTGCGTTC	3,824,487	320 bp	55 °C
CREBBP_Ex12_R		CAAGTGACATGAATTCTGCTGC	3,824,806		
CREBBP_Ex13_F	13	CATCCTCTGGGGTTGTGAAG	3,823,635	399 bp	55 °C
CREBBP_Ex13_R		CATGAAATGTGCATTCTGGA	3,824,033		
CREBBP_Ex14_F	14	TCCATTTCTGGTAGGGACAGGTGC	3,820,539	463 bp	60 °C
CREBBP_Ex14_R		GGCCCAAAAACAGCAGAGACAGA	3,821,001		
CREBBP_Ex15_F	15	TTGTAGGTTGCATGAGCAGC	3,819,081	356 bp	55 °C
CREBBP_Ex15_R		CAGGGATACCCATGGCAG	3,819,436		
CREBBP_Ex16_F	16	CTTCCCGTGAGGTTGCG	3,817,651	329 bp	55 °C
CREBBP_Ex16_R		TCCTCCACATGGAATCCTAAC	3,817,979		
CREBBP_Ex17_F	17	CTTAGGAGCATCGTGGCTG	3,808,746	342 bp	60 °C
CREBBP_Ex17_R		ACAATCTTCAAGGCAGGGG	3,809,087		
CREBBP_Ex18_F	18	TTGGGAATGGAAGTTATTTTGG	3,807,688	484 bp	60 °C
CREBBP_Ex18_R		CACCAGACAGCAGATTGCAC	3,808,171		
CREBBP_Ex19_F	19	TCACATGCTATCCCAAAATGTC	3,807,172	369 bp	60 °C
CREBBP_Ex19_R		GCCTGAAATTGGGCCAC	3,807,540		
CREBBP_Ex20_F	20	TTGCTTGGGTGGCTGTG	3,801,630	241 bp	60 °C
CREBBP_Ex20_R		ATGGCACCGGTACCTTCC	3,801,870		

Table 13 continued					
Primer name	**Exon**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp / End (hg19) in bp**	**Product Size**	**Annealing temp.**
CREBBP_Ex21_F	21	GTTACAAAATAACATTCCAGAGACC	3,799,547	227 bp	60 °C
CREBBP_Ex21_R		CCCACAACCCACTCCATAAG	3,799,773		
CREBBP_Ex22-23_F	22-23	GGACGCACACACAGACTTCTAC	3,794,816	621 bp	60 °C
CREBBP_Ex22-23_R		AACCAAAGAACAATGGGGAC	3,795,436		
CREBBP_Ex24_F	24	ATCCCAGAGTCCACCTTTCC	3,790,319	424 bp	60 °C
CREBBP_Ex24_R		CAAGAGCTTTGCAGAGAGCA	3,790,740		
CREBBP_Ex25_F	25	GGTGTGCAGAAGCACCTTG	3,789,484	306 bp	65 °C
CREBBP_Ex25_R		GAAGGCTCACAGGCTCCTC	3,789,789		
CREBBP_Ex26_F	26	AATGACAGAGCAAGACCCTG	3,788,475	315 bp	55 °C
CREBBP_Ex26_R		TTAAAATACCCATTATTTCACGG	3,788,788		
CREBBP_Ex27_F	27	TAACTCCTTAAAGGCAGGGC	3,786,584	300 bp	55 °C
CREBBP_Ex27_R		AAAAGGCACACAAATATCCTCC	3,786,883		
CREBBP_Ex28_F	28	CATGGGACTCTGCCACAC	3,785,931	388 bp	60 °C
CREBBP_Ex28_R		GACACCACCACAGGAAGGAC	3,786,318		
CREBBP_Ex29_F	29	TGACCTACTTTGGCCTGAGC	3,781,671	377 bp	65 °C
CREBBP_Ex29_R		ACTTCCCTCCCACCACAGAC	3,782,047		
CREBBP_Ex30_F	30	CTATTCTGCAGGCTGGGTG	3,781,127	442 bp	60 °C
CREBBP_Ex30_R		AAAGGGACAGGATGCTTCG	3,781,568		
CREBBP_Ex31.1_F	31	GCACAGACCCAGACTTAGC	3,779,394	580 bp	55 °C
CREBBP_Ex31.1_R		GGAGAAGGCAGACTCTGC	3,779,973		
CREBBP_Ex31.3_F		CCTGTACCGGGTGAACATCAAC	3,778,459	677 bp	60 °C
CREBBP_Ex31.3_R		GCTGCCTCCGTAACATTTCTCG	3,779,135		
CREBBP_Ex31.4_F		CCAAGTACGTGGCCAATCAG	3,778,015	717 bp	65 °C
CREBBP_Ex31.4_R		ACCGCACCTGGTTACTAAGG	3,778,731		
CREBBP_Ex31.5_F		CAACATCCAGCAAGCCC	3,777,646	548 bp	65 °C
CREBBP_Ex31.5_R		GAACCTAGATGCCTGGATTTTC	3,778,193		

2.1.5.6 Primers for Sanger sequencing of *PCBP1*

The following primers for the *PCBP1* gene on chromosome 2p13.3 were designed to validate mutations detected by whole-genome sequencing in BL and to screen for mutations in cell lines as is denoted in a manuscript accepted for publication in the journal *Genes, Chromosomes and Cancer*.

Table 14: Primer sequences for Sanger sequencing of *PCBP1*.

***PCBP1* chromosome 2p13.3**					
Primer name	**Exon**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp / End (hg19) in bp**	**Product Size**	**Annealing temp.**
PCBP1_part1_F	1	GCCAGCCGCCAAAGACTTGA	70,314,811	711 bp	65 °C
PCBP1_part1_R		CCTCCAGGTCATGGGTGGCATG	70,315,521		
PCBP1_part2_F		GGATATGCTGCCCAACTCCACCG	70,315,280	748 bp	60 °C
PCBP1_part2_R		CCTGGAATCACTGACCAG	70,316,027		

2.1.5.7 Primers for Sanger sequencing of *SMARCA4*

The following primers for *SMARCA4* on chromosome 19p13.2 have been published by Schneppenheim *et al.* [193] and were used for validation of mutations detected by whole genome sequencing and for screening of mutations in the helicase domain of *SMARCA4* in B-cell lymphoma cell lines.

Table 15: Primer sequences for Sanger sequencing of *SMARCA4*.

***SMARCA4* chromosome 19p13.2**					
Primer name	**Exon**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp** / **End (hg19) in bp**	**Product Size**	**Annealing temp.**
SMARCA4_Ex9_F	9	GCCTTGCGGGGAGATGTGTCCACCATGCTG	11,105,437	315 bp	65°C
SMARCA4_Ex9_R		GGGGAGTGACCCCTGGAGCCCGCAGTACC	11,105,751		
SMARCA4_Ex15_F	15	GTCAGGAGCCAGCACATTGTCACAGATAG	11,120,988	287 bp	65°C
SMARCA4_Ex15_R		CGCACCACCTGGGAACACCTGCACCGAGG	11,121,274		
SMARCA4_Ex16_F	16	AGGACCCTCTGGTGTCCGACCCGGCCTTC	11,123,565	311 bp	65°C
SMARCA4_Ex16_R		TTGTGGTATTCTACTGCGGCAAACTTAGG	11,123,875		
SMARCA4_Ex17_F	17	TTGCACAGTGAGCCATTGATGAGAGACCG	11,129,573	215 bp	65°C
SMARCA4_Ex17_R		TCACTGTCCAGAGGTATGTGTGGACGTC	11,129,787		
SMARCA4_Ex18_F	18	GTGCCTGTGCCCCTCTTGCCACCTGGCC	11,130,195	273 bp	65°C
SMARCA4_Ex18_R		AACTTGTAGGGGCTTTGGAGGAGACGGGC	11,130,467		
SMARCA4_Ex19_F	19	CTCCCCATGTGCCGGGCCACCTGCTGCCC	11,132,340	392 bp	65°C
SMARCA4_Ex19_R		CCAGCTGTAGCTGGTGCTCAACACGTTCC	11,132,731		
SMARCA4_Ex20_F	20	CCTTCTAGTGAGACCTCTGTCGCCCTCC	11,134,124	271 bp	65°C
SMARCA4_Ex20_R		TGGGGAGAGGCCCTGAGCACGCCCAGCCC	11,134,394		
SMARCA4_Ex21_F	21	GGGTTCGGATGGGGGGAGTCAGGCCTCAA	11,134,938	243 bp	65°C
SMARCA4_Ex21_R		CTGCCTGCCACGCTGCCGGCCTTGGACAC	11,135,180		
SMARCA4_Ex22_F	22	AGCCCACCCCACCCCAGGAGGGCAAGACC	11,136,039	227 bp	65°C
SMARCA4_Ex22_R		GAGCTGTCGAGGAGAAGCCAGCTCTGCC	11,136,265		
SMARCA4_Ex23_F	23	GGACCGCAGCGGGGCCCGGTGGCCTGCTC	11,136,911	189 bp	60°C
SMARCA4_Ex23_R		GCAATAAAGCCAACAAAACGACAGAAAAC	11,137,099		
SMARCA4_Ex24_F	24	CCTGCCTTACCTGCCTGCAGGGTTCCAGG	11,138,388	307 bp	60°C
SMARCA4_Ex24_R		GTGAGGAGCTTCTGTGGCAGCCACAACAAC	11,138,694		
SMARCA4_Ex25_F	25	TCCTTGGTGTCCCCACTCTACCCCTGAGG	11,141,340	299 bp	65°C
SMARCA4_Ex25_R		GGCCGTCTCCTCGAGGTTTTGCAGGCACC	11,141,638		
SMARCA4_Ex26_F	26	CAGAGGCCACCTTCCCTTTTATGACCTCC	11,143,899	357 bp	65°C
SMARCA4_Ex26_R		GAAAGCCGCTCACGCGTCCACCATTCACGC	11,144,255		
SMARCA4_Ex27_F	27	AACTGCTGGTGAAAGACGCCGGATTGACA	11,144,380	233 bp	65°C
SMARCA4_Ex27_R		GGCCCTTGCTGGCCGTCTCAGCCGAGAAG	11,144,612		
SMARCA4_Ex28_F	28	GCTCGGCCGCCGCCCACCCCGGCCCCTCC	11,144,739	234 bp	65°C
SMARCA4_Ex28_R		CTAGGGATACCACCATGGGCACTAGGACG	11,144,972		
SMARCA4_Ex30_F	30	CGGCCTCTGCTTGTCGACCTGGGTGCTGG	11,151,923	383 bp	65°C
SMARCA4_Ex30_R		GAGTGCAGATGCCAGGCCTGCTCCCACGG	11,152,305		
SMARCA4_Ex33_F	33	GGCCGGGCAGGCAGCCCTCCAGTCGGGCC	11,170,368	262 bp	65°C
SMARCA4_Ex33_R		AAAGCTGGGGCCTTGGGGGCTCTCGGGCC	11,170,629		

2.1.6 Primers for validation of TERT transcripts on cDNA level

The following primers were designed to validate the existence and expression of a new exon of the *TERT* gene (NM_198253.2) on chromosome 5p15.33, as well as for Sanger sequencing of the PCR product.

Table 16: Primer for validation of existence and expression of a new exon of the *TERT* gene.

***TERT* chromosome 5p15.33**					
Primer used for	**Primer name**	**Primer Sequence (5'… 3')**	**Start (hg19) in bp** / **End (hg19) in bp**	**Product Size**	**Annealing temp.**
sequencing of 5' end of new exon	TERT_Ex2_F	CTCCAGGCACAACGAACGC	1,287,089	176 bp	60 °C
	TERT_Ex new_R1	GATCAACACACACTCGGCAGG	1,293,564		
detection of *TERT* new exon and exon2V transcript	TERT_Ex2V2_F	CGAGTGGACACGGTGATCTC	1,285,577	296 bp	60 °C
	TERT_Ex new_R1	GGCAACCCAAGAGGTGGTGAG	1,294,834		
detection of splice variants of *TERT* new exon, sequencing of 3' end of new exon	TERT_Ex new_F	CCAGCTTTACCTGTGCTGG	1,282,669	447 bp, 222 bp, 181 bp	65 °C
	TERT_Ex3_R	CAGCCAGTGCAGGAACTTGG	1,286,390		
detection of basal TERT expression	TERT_Ex3_F	CGTCGTCGAGCTGCTCAGGTC	1,253,287	137 bp	60 °C
	TERT_Ex4_R	GCTGCACCCTCTTCAAGTGC	1,283,125		

2.1.7 Primers for quantitative real-time PCR (qPCR)

2.1.7.1 Self-designed qPCR primers

Primers were designed and the uniqueness in the human genome was rechecked by BLAT analyses (http://genome.ucsc.edu/cgi-bin/hgBlat?command=start) against the NCBI build 37/hg19. Oligonucleotides were obtained from Biomers, Ulm, Germany as lyophilisates and resuspended in ddH_2O to a final concentration of 200 pmol/µl. The primers were designed to have an optimal annealing temperature at 58°C.

Table 17: Self-designed primers for qPCR.

Gene	**Primer name**	**Detected Transcript**	**Primer Sequence (5'… 3')**	**Chr**	**Start (hg19) in bp** / **End (hg19) in bp**	**Product Size**
IL10	IL10_qPCR_F	NM_000572.2	CTACGGCGCTGTCATCGAT	1	206,940,948	110 bp
	IL10_qPCR_R		CTCATGGCTTTGTAGATGCC		206,945,839	
KMT2A	KMT2A_qPCR_F	NM_001197104.1 NM_005933.3	CCAGTAGTGGGCATGTAGAG	11	118,355,671	156 bp
	KMT2A_qPCR_R		GATGTTGCCTTCCACAAACG		118,359,464	

2.1.7.2 Pre-designed qPCR primers

Predesigned oligonucleotides were obtained from Qiagen, Hilden, Germany as lyophilisates and resuspended in ddH_2O according to manufacturer´s instruction.

Table 18: Predesigned qPCR primers obtained from Qiagen.

Gene	Detected transcript(s)	QuantiTect Primer Assay
ETS1	NM_005238	Hs_ETS1_1_SG
FLI1	NM_001167681, NM_002017	Hs_FLI1_1_SG
IL0RA	NM_001588	Hs_IL10RA_1_SG
PAFAH1B2	NM_002572	Hs_PAFAH1B2_1_SG
GUSB	NM_000181	Hs_GUSB_1_SG
HPRT1	NM_000194	Hs_HPRT1_1_SG

2.1.8 Plasmids

Commercially available vectors were obtained as DNA solutions. Table 19 lists plasmid DNA used in this study. Refer to 8.5 for the vector maps of the respective plasmids.

Table 19: List of plasmid DNA used in this study.

Plasmid DNA	Function	Manufacturer
ID3-pCMV6-AC-*GFP* (RG200583)	expression of ID3 (NM_002167.2) in B-cell lymphoma cell lines	OriGene Technologies, Rockville, USA
pMD.G	Packaging plasmids for generation of lentiviral vectors	Kindly provided by Dr. rer. nat. E. Murga Penas, Institute of Human Genetics, Kiel
pCMV-dR8.91	Packaging plasmids for generation of lentiviral vectors	Kindly provided by Dr. rer. nat. E. Murga Penas, Institute of Human Genetics, Kiel
pLKO_IPTG_3xLacO non-targeting control shRNA	Negative control for baseline cellular response to knock down experiments	Sigma Aldrich, St. Louis, USA
pLKO_IPTG_3xLacO PAFAH1B2 shRNA1 (TRCN0000218283)	Knock down of PAFAH1B2 expression in B-cell lymphoma cell lines	Sigma Aldrich, St. Louis, USA
pLKO_IPTG_3xLacO PAFAH1B2 shRNA2 (TRCN0000218878)	Knock down of PAFAH1B2 expression in B-cell lymphoma cell lines	Sigma Aldrich, St. Louis, USA
pLKO_IPTG_3xLacO PAFAH1B2 shRNA3 (TRCN0000230674)	Knock down of PAFAH1B2 expression in B-cell lymphoma cell lines	Sigma Aldrich, St. Louis, USA
pmax	Control of transfection efficiency, contains *GFP* gene	Amaxa/Lonza, Cologne, Germany

2.1.9 Special laboratory equipment

Laboratory equipment, which was specifically used for experiments in this study, is listed in Table 20. Standard laboratory equipment is not separately listed.

Table 20: Lab equipment used in this study.

Equipment	Manufacturer
Amersham Hyperfilm™ ECL (5x7 inch)	GE Healthcare, Munich, Germany
Criterion TGX Precast Gel Any kD	Bio-Rad, Munich, Germany
Criterion TGX Precast Gel 18 % Resolve	Bio-Rad, Munich, Germany
Hypercassette™	GE Healthcare, Munich, Germany
Immobilin-P PVDF membrane	Merck Millipore, Darmstadt, Germany
Magnetic beads	Agencourt, Beverly, MA, USA
Neubauer hematocytometer	Brand, Wertheim, Germany
Polystrene Round-Bottom Tube (FACS)	BD Bioscience, Franklin Lake, New Jersey, USA
Tissue culture flasks T25	Sarstedt, Newton, USA

Table 20 continued	
Equipment	**Manufacturer**
Tissue culture flasks T75	Sarstedt, Newton, USA
Tissue culture plate, 6-well	Sarstedt, Newton, USA
Tissue culture plate, 24-well	Sarstedt, Newton, USA
Tissue culture plate, 96-well	Sarstedt, Newton, USA
Whatman filter paper	GE Healthcare, Munich, Germany

2.1.10 Chemicals and reagents

The following Table 21 lists all chemicals and reagents used in this study and their manufacturers.

Standard chemicals are not separately listed.

Table 21: Chemicals and reagents used in this study.

Chemicals and reagents	Manufacturer
Acetic acid	J.T. Baker, Griesheim, Deutschland
Agar-Agar	Merck Millipore, Darmstadt, Germany
Agarose	Biozym, Hessisch Oldendorf, Germany
Aqua bidest (ddH_2O)	Fresenius Kabi AG, Bad Homburg, Germany
BKM120	Selleckchem, Houston, USA
Blotto, non-fat dry milk powder	Santa Cruz Biotechnology, Texas, USA
Bromphenol Blue	Serva, Heidelberg, Germany
Carbenicillin	Sigma Aldrich, St. Louis, USA
Cell Titer Blue reagent	Promega, Mannheim, Germany
cOmplete™ Protease Inhibition cocktail (PIC) Tablet	Roche, Rotkreuz, Switzerland
Denhardt´s solution	Sigma Aldrich, St. Louis, USA
DMSO	Sigma Aldrich, St. Louis, USA
dNTP	Roche, Rotkreuz, Switerland
EB-buffer	Qiagen, Hilden, Germany
Ethanol	J.T. Baker, Griesheim, Deutschland
Ethidiumbromide	Sigma Aldrich, St. Louis, USA
Glycerol	Sigma Aldrich, St. Louis, USA
Glycine	Sigma Aldrich, St. Louis, USA
Hering Sperm DNA	Promega, Mannheim, Germany
Human Cot-1	Life Technologies, Carlsband, USA
Human Total RNA Master Panel II	Clontech, Daint-Germain-en-Laye, France
HyperLadder I	Bioline, Taunton, USA
IGEPAL CA-630	Sigma Aldrich, St. Louis, USA
IPTG	Sigma Aldrich, St. Louis, USA
Isopropyl alcohol	Sigma Aldrich, St. Louis, USA
LB broth base	Life Technologies, Carlsband, USA
Luminata™ Forte Western HRP Substrate	Merck Millipore, Darmstadt, Germany
MES hydrate Sigma Ultra	Sigma Aldrich, St. Louis, USA
MES sodium salt	Sigma Aldrich, St. Louis, USA
Methyl alcohol	J.T. Baker, Griesheim, Deutschland
Methylated DNA	Merck Millipore, Darmstadt, Germany
Nsp I	New England Biolabs, Frankfurt a.M., Germany
NuPAGE LDS loading buffer (4x)	Life Technologies, Carlsband, USA
Paraformaldehyde	Sigma Aldrich, St. Louis, USA
PBS	GIBCO Life Technologies, Carlsband, USA
Polybrene	Merck Millipore, Darmstadt, Germany
Ponceau Red	Sigma Aldrich, St. Louis, USA
Precision PlusPotein Standard DualXtra	Bio-Rad, Munich, Germany
Puromycin	Sigma Aldrich, St. Louis, USA
RNase A	Qiagen, Hilden, Germany

Table 21 continued	
Reagents	**Manufacturer**
Roentoroll 25, part 1 and part2	Tetenal, Norderstedt, Germany
Roti®-Histol	Carl ROTH, Karlsruhe, Germany
SDS	Serva, Heidelberg, Germany
S.O.C. medium	Life Technologies, Carlsband, USA
Sodium cholorid (NaCl)	Merck Millipore, Darmstadt, Germany
Sodium deocycholat (Na-DOC)	Sigma Aldrich, St. Louis, USA
Streptavidin coated Sepahrose beads	GE Healthcare, Munich, Germany
Sty I	New England Biolabs, Frankfurt a.M., Germany
Superfix 25	Tetenal, Norderstedt, Germany
T4 DNA Ligase	New England Biolabs, Frankfurt a.M., Germany
10x TBE buffer	Life Technologies, Carlsband, USA
TITANIUM Taq Polymerase	Takara Clontech, Daint-Germain-en-Laye, France
TMACL	Sigma Aldrich, St. Louis, USA
TriDye ladder	New England Biolabs, Frankfurt/Main, Germany
TritonX-100	Sigma Aldrich, St. Louis, USA
TRIS	Merck Millipore, Darmstadt, Germany
TRIS-HCl	Sigma Aldrich, St. Louis, USA
Trypan blue solution (0.4 %)	Sigma Aldrich, St. Louis, USA
Trypsin EDTA (1x)	GE Healthcare, Munich, Germany
Tween®20	Sigma Aldrich, St. Louis, USA

2.1.11 Buffers and solutions

The following table enlists all buffers and solutions used in this study as well as their composition. If not other specified the buffers were used at room temperature.

Table 22: Buffers and solutions used in this study.

Buffers and solutions	Composition
Agarose gel loading buffer	1.9 mM Bromphenol blue 3.1 M glycerol add ddH_2O to 100 ml store at 4°C
Antibody diluent (WB)	1.5 g Blotto, non-fat milk powder TBS-T to 50 ml
Bjerrum Schäfer transfer buffer, 10x	480 mM TRIS 380 mM glycine ddH_2O to 1 l
Bjerrum Schäfer transfer buffer, 1x	100 ml 10x Bjerrum Schäfer transfer buffer 100 ml methyl alcohol ddH_2O to 1 l, store at 4 °C
Blocking Solution (WB)	3 g Blotto, non-fat milk powder TBS-T to 50 ml
FACS wash buffer	0.5 % FBS PBS to 50 ml
FACS wash buffer II	0.5 % FBS 0.5 % Tween®20 PBS to 50 ml
Freezing media	10 % DMSO FBS to 50 ml Store at 4 °C
LB medium	20 g LB ddH_2O to 1 l, autoclave and store at 4 °C

Table 22 continued	
Buffers and solutions	**Composition**
LB agar	10 g LB 15 g Agar-Agar ddH_2O to 0.5 l autoclave and store at 4 °C
Paraformaldehyde solution 1 % in PBS	1 g paraformaldehyde PBS to 100 ml Boil until dissolved and store at 4 °C
Permeabilization buffer	0.5 % Tween20 PBS to 50 ml
Ponceau Red Staining Solution	0.5 % Ponceau S 1 % acetic acid ddH_2O to 100 ml
RIPA	50 mM TRIS-HCl (pH8) 150 mM NaCl 1 % IGEPAL 0.5 % Na-DOC 0.1 % SDS 1 cOmplete™ PIC Tablet ddH_2O to 50 ml store at -20 °C
SDS-running buffer, 10x	250 mM TRIS 2 M glycine 1 % SDS ddH_2O to 1 l
SDS-running buffer, 1x	100 ml 10x SDS-running buffer ddH_2O to 1 l
TBE, 1x	100 ml 10x TBE buffer ddH_2O to 1 l
TBS, 10x	5 M NaCl 0.2 TRIS ddH_2O to 1 l
TBS, 1x	100 ml TBS ddH_2O to 1 l
TBS-T	0.1 % Tween®20 1 l 1x TBS

2.1.12 Kits

The kits which have been used in this study and their manufacturer are listed in Table 23.

Table 23: Name and manufacturer of the kits used in this study.

Kits	Manufacturer
AccuPrime Polymerase System	Life Technologies, Carlsband, USA
AmpliTaq Gold Polymerase System	Life Technologies, Carlsband, USA
BCA Macro Assay Kit	Serva, Heidelberg, Germany
Big Dye Terminator® v1.1 Cycle Sequencing Kit	Applied Biosystems, Foster City, USA
Cell Line Nucleofector® Kit V	Amaxa/Lonza, Cologne, Germany
EasyPure DNA Purification Kit	Biozym, Hessisch Oldendorf, Germany
Experion® RNA StdSens Analysis Kit	Bio-Rad, Munich, Germany
EZ-96 DNA Methylation™ Kit	Zymo Research, Freiburg, Germany
Genome wide human SNP NSP/STY Assay Kit 5.0/6.0	Affymetrix, Santa Clara, USA
Human IL-10 ELISA Kit	Thermo Scientific, Waltham, Massachusetts, USA
Human Methylation 450k Beadchip Kit	Illumina, San Diego, USA
jetPEI transfection reagent	Polyplus-transfection, Illkirch, France
Lenti-X™ GoStix™	Clontech, Daint-Germain-en-Laye, France
Montage 96 Well Sequencing Reaction Cleanup kit	Merck Millipore, Darmstadt, Germany
NucleoBond Xtra Maxi EF plasmid purification kit	Machery-Nagel, Düren, Germany
OncoScan™ FFPE Assay Kit	Affymetrix, Santa Clara, USA
One Shot® TOP10 Chemically Competent *E.coli*	Life Technologies, Carlsband, USA
PARIS™ kit	Life Technologies, Carlsband, USA
PUREGENE DNA Isolation Kit (Gentra Systems)	Qiagen, Hilden, Germany
RNeasy Mini kit	Qiagen, Hilden, Germany
QIAmp DNA FFPE tissue kit	Qiagen, Hilden, Germany
Qiagen MiniElute 96 PCR Purification Kit	Qiagen, Hilden, Germany
OncoScan™ FFPE Assay Kit	Affymetrix, Santa Clara, USA
QuantiTect®Reverse Transcription Kit	Qiagen, Hilden, Germany
QuantiTect®SYBR®Green PCR Kit	Qiagen, Hilden, Germany
Qubit® dsDNA BR assay	Life Technologies, Carlsband, USA
Qubit RNA HS assay	Life Technologies, Carlsband, USA
STEM ELITE ID	Promega, Mannheim, Germany
TruSightTM One Sequencing Panel	Illumina, San Diego, USA

2.1.13 Instruments

The following table lists the instruments used in this study. Standard instruments are not separately listed.

Table 24: Instruments used in this study.

Instruments	Model	Manufacturer
Autoclave	V-Autoklav	Webeco, Fridolfing, Germany
Centrifuge (Cell culture)	Centrifuge 5810	Eppendorf, Hamburg, Germany
Centrifuge (Cell culture)	Multifuge 1L-R (rotor: 75002000)	Heraeus, Hanau, Germany
Centrifuge (DNA, Eppendorf tube)	Biofuge primo R (rotor: 7593)	Heraeus, Hanau, Germany
Centrifuge (DNA, 15 ml falcon tube)	Megafuge 1.0 R (rotor: BS4402/A)	Heraeus, Hanau, Germany
Centrifuge (Methylation, plate)	Megafuge 40R (rotor: 75003607)	Heraeus, Hanau, Germany
Centrifuge (Eppendorf tubes)	Mikro 200 (rotor: 1195-A)	Hettich, Tuttlingen, Germany
Centrifuge (Protein, bacteria harvest)	Multifuge X1R (rotor: 75003602, 75003658)	Heraeus, Hanau, Germany
Centrifuge (Preparation of recombinant bacteria)	SORVALL SUPER T21 (rotor: SL-50T)	Thermo Scientific, Waltham, Massachusetts, USA

Table 24 continued		
Instruments	**Model**	**Manufacturer**
CO_2 incubator for cell culture	Galaxy B nunc™	RS Biotech, Irvine, UK
Electroporator	Nucleofector I device	Lonza, Cologne, Germany
ELISA plate reader	Synergy HT	BioTek, Luzern, Switzerland
Flow cytometer	AccuriC6	BD Bioscience, Franklin Lake, New Jersey, USA
Flow cytometer	FACS Calibur Analyzer	BD Bioscience, Franklin Lake, New Jersey, USA
Fluorescence microscope	Eclipse TS100	Nikon, Tokyo, Japan
Fluorometer	Qubit 2.0	Life Technologies, Carlsband, USA
Gene-Chip Washing Station	Gene Chip® Fluidics Station 450	Affymetrix, Santa Clara, USA
Gel chamber	Model 41-2026	Promega, Mannheim, Germany
Gel chamber	Hoefer HE 33 mini horizontal submarine unit	Amersham Bioscience, Munich, Germany
Gel chamber for SDS-Page	Criterion Cell	Bio-Rad, Munich, Germany
Gel documentation device	BioDoc Analyzer	Biometra, Göttingen, Germany
Hybridization Oven	GeneChip® Hybridization Oven 645	Affymetrix, Santa Clara, USA
Hybridization Oven (450k array)	Hybrid oven	Illumina, San Diego, USA
Incubator	Function Line	Heraeus, Hanau, Germany
Incubator with shaking device for bacteria	SM30	Edmund Bühler GmbH, Tübingen
Incubator for bacteria	B5025	Heraeus, Hanau, Germany
Incubator (450k array)	Hybex-microsample incubator	Illumina, San Diego, USA
Microfuidics system	Experion® Electrophoresis System	Bio-Rad, Munich, Germany
Microscope	Photo Zoom inverted microscope	Cambridge Instruments, Wetzlar, Germany
Microwave	M500	Philips, Hamburg, Germany
Plate shaker	MS 3 basic	IKA, Staufen, Germany
Plate shaker (450k array)	High-Speed Microplate Shaker	Illumina, San Diego, USA
Power Supply for western blot	Power Pac HC™	Bio-Rad, Munich, Germany
Power Supply for agarosegel	ST606	Life Technologies, Carlsband, USA
quantitative real-time PCR instrument	LightCycler® 480 II	Roche, Rotkreuz, Switzerland
Tube Roller Mixer	RMS V-30	Bibby Scientific limited group, Staffordshire, UK
Safety Cabinet	Nuaire Biological Safety Cabinet Class II	Zapf Instruments, Sarstedt, Germany
Scanner	GeneChip® Scanner	Affymetrix, Santa Clara, USA
Scanner (450k array)	iScan	Illumina, San Diego, USA
Sequencer, 4-capillary	ABI PRISM®3100/3130 Genetic Analyzer	Applied Biosystems, Foster City, USA
Sequencer	MiSeq desktop sequencer	Illumina, San Diego, USA
Tank for western blot	Criterion Blotter	Bio-Rad, Munich, Germany
Thermocycler	T3000 Thermocycler	Biometra, Göttingen, Germany
Thermomixer	Thermomixer Comfort	Eppendorf, Hamburg, Germany
Vacuum tool	Vacuum Prep Tool	Biotage, Uppsala, Sweden
Vortexer	REAX 2000	Heidolph, Schwabach, Germany
Weighing device	LC621P	Sartorius, Göttingen, Germany
X-ray film processor	Curix 60	Agfa HelthCare, Berlin, Germany

2.1.14 Software

The following table lists the software used in this study. Standard software is not enlisted in the following Table 25.

Table 25. Software used in this study.

Software	Manufacturer	used for
BD C Sampler Software	BD Bioscience, Franklin Lake, New Jersey, USA	Analyses of flow cytometry data
BioDoc Analyze	Applied Biosystems, Foster City, USA	Photodocumentation
Cell Quest Software	BD Bioscience, Franklin Lake, New Jersey, USA	Analyses of flow cytometry data
Command Console	Affymetrix, Santa Clara, USA	Analyses of Human SNP 6.0 array data
Experion® Software	Bio-Rad, Munich, Germany	Experion® RNA assay
Gen5™ Software	BioTek, Luzern, Switzerland	Control console for Synergy HT
GeneMapper 5 Software	Applied Biosystems, Foster City, USA	Sequence analyses
Genome Studio Software 2011.1	Illumina, San Diego, USA	Analyses 450k array data
Genotyping Console	Affymetrix, Santa Clara, USA	Analyses of Human SNP 6.0 arrays
Graphpad Prism 5	Graphpad Software, San Diego, CA, USA	Statistical analyses
Integrative Genomics Viewer 2.3	Broad Institute, Cambridge, Boston, USA	Viewing of BAM files
Lightcycler®480 SW 1.5.1	Roche, Rotkreuz, Switerland	Control console and analyses of qPCR data
Microsoft Office 2007	Microsoft, Unterschleißheim, Germany	Digital documentation
MiSeq Reporter	Illumina, San Diego, USA	Analyses of TRUE Sight Panel
Nexus Express for OncoScan 3.0	Affymetrix, Santa Clara, USA	Analyses of OncoScan copy number data
Sequencing Analyses Software 6.0	Applied Biosystems, Foster City, USA	Analyses of obtained sequences

2.1.15 Online databases and algorithms

Online databases and tools that were used in this study are listed in Table 26. As the databases have been continuously used, a single date of accession can sometimes not be given. If analyses or data extractions have been performed at a specific time point, this will be mentioned in the text.

Table 26: Online databases and algorithms used in this study.

Name	URL	Reference
BLAT	https://genome.ucsc.edu/cgi-bin/hgBlat?command=start	[194]
CCLE	http://www.broadinstitute.org/ccle/home	
COSMIC v69	http://cancer.sanger.ac.uk/cancergenome/projects/cosmic	
dbSNP (build 141)	http://www.ncbi.nlm.nih.gov/projects/SNP/	
DSMZ homepage	http://www.dsmz.de/home.html	
ExPASy translate tool	http://web.expasy.org/translate/	
GeneCards®	http://www.genecards.org/	
Human Genome Variation Society	http://www.hgvs.org/mutnomen/	[195]
ICGC data portal	https://dcc.icgc.org/	
LALIGN	http://www.ch.embnet.org/software/LALIGN_form.html	[196]
Mechismo	http://mechismo.russelllab.org/	[197]
NHLBI Exome Sequencing Project (ESP6500SI-V2)	http://evs.gs.washington.edu/EVS/	
Online STR Analyses	http://www.dsmz.de/STRanalyses	[198]
Pfam 27.0	http://pfam.xfam.org/	[199]
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/index.shtml	[200]
PubMed	http://www.ncbi.nlm.nih.gov/pubmed/	
SMART	http://smart.embl-heidelberg.de/	[201]
UCSC Genome Bioinformatics	http://genome.ucsc.edu/	[202]

2.2 Methods

2.2.1 Preparation and analyses of DNA

2.2.1.1 Extraction of genomic DNA

2.2.1.1.1 DNA extraction from FFPE-embedded tissue

DNA from formalin-fixed, paraffin-embedded (FFPE) tissue was extracted using the QIAmp DNA FFPE tissue kit. For preparation of DNA from FFPE blocks, 8-10 tissue sections of ~8 µm were used. For removal of the paraffin, Roti®-Histol was added to the sections and mixed thoroughly. The mixture was centrifuged and washed once more with Roti®-Histol. Afterwards the tissue was washed twice with absolute ethanol to remove any remaining Roti®-Histol. After drying the pellet at room temperature, the pellet was resuspended in ATL buffer and proteinase K. The lysate was incubated at 56 °C overnight (o/n). The next day, the lysate was incubated at 90 °C for 1 h to remove any modifications of the formaldehyde on the DNA. After adding AL buffer, the lysate was mixed and the DNA was precipitated using absolute ethanol. The lysate was then loaded onto a Mini Elute column and centrifuged. The column was washed with AW1 buffer and AW2 buffer by centrifugation at 8,000 rpm (Mikro 200). Finally, the DNA was eluted using ddH_2O and stored at 4 °C. The concentration was measured as described in 2.2.1.2.

2.2.1.1.2 DNA extraction from cell lines

DNA from cell lines in culture was extracted using the PUREGENE DNA Isolation Kit. The protocol is outlined in the following.

Cells from cultured cell lines were harvested by centrifuging for 3 min at 1500 rpm (Multifuge 1L-R) at room temperature. The supernatant was discarded and the pellet resuspended in Cell Lysis Solution corresponding to the amount of starting material. The lysate was incubated o/n at room temperature. The next day RNase A-solution was added to the lysate and incubated for another 30 min at room temperature. Then Protein Precipitation Solution was added to the lysate and mixed thoroughly for 1 min and afterwards incubated on ice for 15 min. The lysate was centrifuged at 4500 rpm for 30 min (Megafuge 1.0R). After transferring the supernatant to a fresh Eppendorf tube, the DNA was precipitated using the same volume of absolute isopropyl alcohol. The DNA became visible as a white thread and was transferred into a fresh Eppendorf tube. The DNA was washed once with 70 % ethanol and then dried at room temperature. Dependent on the DNA pellet size, ddH_2O was added and resolved o/n at room temperature. The DNA was stored afterwards at 4 °C and the concentration was measured as described in 2.2.1.2.

2.2.1.2 Measurement of DNA concentration using Qubit®dsDNA BR assay

The DNA concentration was measured using Qubit® dsDNA BR assay. The advantage of this assay is that it is a fluorescence-based technique which is highly selective for double stranded (ds) DNA over RNA. Therefore there is no bias of any leftover RNA when measuring DNA concentration as it might be the case using conventional spectrophotometric DNA measurement methods. The measurement was performed according to manufacturer´s protocol (Life Technologies, Carlsband, USA).

2.2.1.3 Plasmid propagation and transformation of *E. coli*

For each transformation one vial chemically competent TOP10 *E.coli* was thawed on ice. Up to 1 µg of plasmid DNA was added to each vial and tube content was mixed by gentle agitation. The transformation mixture was incubated for 30 min on ice, then at 42 °C for 30 s and afterwards immediately put on ice for at least 2 min. Afterwards, 250 µl of pre-warmed S.O.C. media was added to the transformation mix and incubated at 37 °C under constant shaking at 225 rpm in a horizontal shaking incubator for 1 h. 25 µl of the transformation mix was plated onto LB agar plates containing 100 µg/ml carbenicillin. Plates were incubated at 37 °C o/n. Single clones were picked and grown at 37 °C for at least 8 h in an initial solution of 5 ml LB media containing 100 µg/ml carbenicillin. To generate glycerol stocks, 0.5 ml glycerol was added to 0.5 ml of this bacteria suspension. Stocks were stored at -80 °C. The left-over of the bacteria suspension was used for the preparation of plasmid DNA (2.2.1.4).

2.2.1.4 Preparation of plasmid DNA from recombinant *E. coli*

Plasmid DNA was prepared using the NucleoBond® Xtra Maxi EF kit according to manufacturer´s instruction. After adding 4 ml of the mini initial solution (2.2.1.3) to 300 ml LB containing 50 µg/ml carbenicillin, the bacteria suspension was incubate o/n at 37 °C under vigorous shaking. The next day, the bacteria suspension was centrifuged at 5500 rpm (Multifuge X1R) for 15 min to pellet the cells. The supernatant was discarded and the pellet resuspended in RES-EF buffer + RNase A. LYS-EF buffer was added, mixed by inverting and incubated for 5 min at room temperature. NEU-EF buffer was added, mixed by inverting and loaded onto an equilibrated filter column. After the solution had passed through the filter completely, FIL-EF buffer was added to wash and run through. The filter was discarded and the column was washed with ENDO-EF buffer and WASH-EF buffer. Plasmid DNA was eluted with ELU buffer. To the elute half of the volume absolute isopropyl alcohol was added and mixed thoroughly to precipitate the DNA. The suspension was centrifuged at 18,500 g (SORVALL SUPER T21) for 45 min at 4 °C. Discarding the supernatant, the pellet was washed with 70 % ethanol and centrifuged at 15,000 g (Multifuge X1R) for 15 min at room temperature. The supernatant was

discarded carefully and the pellet dried at room temperature. After resuspension of the pellet in ddH_2O, plasmid DNA was dissolved o/n at 4 °C. The DNA concentration was determined as described in 2.2.1.2.

2.2.1.5 Amplification of DNA fragments by PCR

Amplification of fragments from genomic DNA (gDNA), plasmid DNA or complementary DNA (cDNA) was performed using either the AmpliTaq Gold Polymerase system or the Accuprime Polymerase system in case of GC-rich target DNA. Reactions were set up in volumes of 25 µl.

For a typical PCR of gDNA each 25 µl reaction mix contained:

AmpliTaq Gold Polymerase System		Accuprime Polymerase System	
1 µl	gDNA (100 ng/µl)	1 µl	gDNA (100 ng/µl)
0.5 µl	sense-primer (20 pmol/µl)	0.5 µl	sense-primer (20 pmol/µl)
0.5 µl	antisense-primer (20 pmol/µl)	0.5 µl	antisense-primer (20 pmol/µl)
1 µl	dNTP solution (5 mM per nucleotide)	2.5 µl	10x buffer
2.5 µl	10x buffer II	0.3 µl	Accuprime Polymerase
2 µl	$MgCl_2$ (25 mM)	20.2 µl	ddH_2O
0.2 µl	AmpliTaq Gold Polymerase		
17.3 µl	ddH_2O		

The amount of DNA, cDNA or plasmid DNA used for the PCR was dependent on the experimental setting. Thus, the amount of ddH_2O was adjusted not to exceed the final reaction volume of 25 µl.

In order to achieve PCR condition yielding the specific product, cycling was performed at the highest, most stringent annealing temperature which was determined for each primer pair by test reaction runs.

A typical PCR program was run as follows:

Denaturation	95 °C	5 min	
	95 °C	30 s	36 cycles
Annealing	55-65 °C	30 s	
Elongation	68 °C Accuprime Polymerase/ 72 °C AmpliTaq Gold Polymerase	30 s	
	68 °C or 72 °C	5 min	
Cooling	4 °C	on hold	

PCR products were checked by agarose gel electrophoresis (2.2.1.6).

2.2.1.6 Agarose gel electrophoresis

Agarose gel electrophoresis was performed depending on the size of the expected PCR product in agarose gels differing in their agarose concentration. 2 % gels were used in case of PCR products smaller than 1,000 bp in size, 1 % gels in case of larger fragments as in the case of plasmid DNA and 3 % gels in case of fragments about 100 bp in size. Agarose was boiled in TBE buffer in a microwave until the solution became transparent and then cooled down until lukewarm. Ethidiumbromide was added to a final concentration of ~1µM and mixed well. The agarose was poured into a horizontal gel chamber and a comb was placed into the gel to create wells. After becoming solid again the PCR products were loaded. 3.5 µl of gel loading buffer were added to 10 µl of PCR product. Depending on the size of the expected PCR product, the 100 bp TriDye ladder was used as site standard for products smaller 800 bp and the HyperLadder standard was used for PCR products bigger than 800 bp. 2-3 % agarose gels were run 10 V/cm and 1 % agarose gels 5 V/cm until single marker bands were clearly visible. Gels were documented using the BioDoc Analyzer. The size of the fragment was determined by comparison of bands relative to the standard.

2.2.1.7 Purification of PCR products from agarose gels

To analyze a biallelic involvement of *ID3* mutations in BL-70 cells, the ID3 cDNA was amplified as described in 2.2.1.5. The PCR yielded two bands on the agarose gel. The two bands were extracted using the EasyPure DNA Purification Kit according to manufacturer's instructions. The target PCR bands were cut out from the agarose gel. The cut-out gel band was weight and the volume of the gel was estimated at 100 µg gel responding to ~100 µl gel. MELT buffer - half of the gel volume- as well as SALT buffer - 4.5 of the gel volume- was added. The mixture was boiled at 55 °C until the agarose was completely dissolved. 10 µl BIND was added to the mixture and incubated for 5 min at room temperature. The mixture was centrifuged at 10,000 rpm for 5 s (Micro 200) and the supernatant was discarded. The pellet was resuspended in 1 ml WASH buffer and mixed well for 5-10 s using a vortexer. After centrifugation at 10,000 rpm for 5 s (Micro 200), the supernatant was discarded and the pellet was dried at room temperature. Finally, the pellet was resuspended in 20 µl ddH_2O and incubated for 5 min at room temperature. After an additional centrifugation step as described above, the supernatant containing the PCR product was transferred into a fresh tube. The sequence of the PCR product was analyzed by Sanger sequencing as described in 2.2.1.9.

2.2.1.8 Post reaction purification of PCR products

Prior to sequencing, amplified products were subjected to a cleanup step for removal of nucleotides, polymerase and salts. This step was processed in a 96-well plate using the Qiagen MiniElute 96 PCR

Purification Kit. A volume of 40 µl ddH_2O was added to the PCR product and mixed well. The mixture was loaded onto the 96-well plate and a vacuum was applied to draw the liquid through the membrane until completely dry. The plate was removed and put on a paper towel to remove any leftover liquid from underneath. For repeated washing, another 50 µl of ddH_2O were added to each well and drawn though the membrane as above until the membrane was dry. Another 20 µl of ddH_2O were added to each well. For DNA recovery, the plate was sealed with a plastic foil and put on a horizontal plate shaker for 5 min at 1,000 rpm. Purified PCR-products were transferred into a fresh tube.

2.2.1.9 DNA sequencing

The didesoxy-method described by Sanger *et al.* [203] was used for sequencing of PCR products or plasmid DNA. The determination of the nucleotide sequence was performed by capillary electrophoresis according to fragment length and detection of the fluorescent dye.

Sequencing reactions were set up using the Big Dye Terminator® v1.1 Cycle Sequencing Kit in volumes of 20 µl as follows:

1 µl	Big Dye Terminator
2 µl	5x Big Dye Terminator reaction buffer
1 µl	primer (20 pmol/µl)
1 µl	purified PCR product
15 µl	ddH_2O

The PCR reaction was run under the following conditions:

Denaturation	96 °C	2 min	
	96 °C	10 s	25 cycles
Annealing	57 °C	5 s	
Elongation	60 °C	4 min	
Cooling	4 °C	on hold	

Prior to electrophoretic separation, a cleanup step was performed using the Montage 96 well Sequencing Reaction Cleanup Kit to remove any leftover nucleotides, buffer and salts. A volume of 20 µl injection solution was added to the reaction and loaded onto the plate. A vacuum was applied to draw the liquid through the membrane until completely dry. The plate was removed and dried from underneath using a paper towel. To wash the membrane another 20 µl injection solution was loaded and again it was drawn through the membrane as described above. Finally 20 µl ddH_2O were loaded, the plate sealed with plastic foil and put on a horizontal plate shaker for 5 min at 1,000 rpm

for recovery of the sequencing products. Purified sequencing products were transferred into fresh tubes.

Sequencing was done on ABI PRISM® 3100 or 3130 Genetic Analyzer. The sequencing products were diluted between 3x-5x in ddH_2O. Injection time varied from 5 s to 15 s according to product concentration; run time varied between 18 min to 30 min according to fragment length. The analyses of the sequence raw data was performed using the Sequencing Analyses Software Version 6.0 and afterwards compared manually against references sequences of the NCBI build 37/hg19. Variances from the reference sequence were described according to HGVS mutation nomenclature recommendations [195]. Every detected single nucleotide variant (SNV) was screened for annotation as single nucleotide polymorphisms (SNP) in the dbSNP 141 database. The consequences of mutational changes on the protein code were derived from the DNA codon table which is attached in the appendix (8.4).

2.2.1.10 High-throughput sequencing

Using high-throughput sequencing it is possible to decipher the sequence of the whole-genome, exome, transcriptome and epigenome. Different sequencing techniques exist today depending on the provider, platform and applied kit. In this study exome sequencing and targeted-resequencing was performed for three mnBLL cases (cases 2, 3 and 8) and selected cell lines as listed in Table 27.

Table 27: Cell lines analyzed with high-throughput sequencing.

Cases	Enrichment	Sequencing platform	Custom/In-house	Analyses performed
BL-2 cells	Nimbl-SeqCapEZ-V2 (Roche)	HiSeq2000 (Illumina)	Cologne Center for Genomics (CCG), Cologne, Germany	Varbank (CCG)
MLMA cells				
Su-DHL-5 cells	TruSeq™ Exome Enrichment Kit (Illumina)	HiSeq2000 (Illumina)	AROS Applied Biotechnology, Arhus, Denmark	Dr. Matthias Schlesner (DKFZ)
HT cells				
mnBLL cases 2, 3 and 8	TruSight™ One Sequencing Panel (Illumina)	MiSeq (Illumina)	In-house	In-house
BL-70, BLUE-1, U-698-M cells				

All three enrichment techniques are based on hybrid capture. The differences of the enrichment panels and platforms, as well as the bioinformatic analyses are described in the following sections.

2.2.1.10.1 Custom exome sequencing

Custom exome sequencing of the BL-2 and MLMA cell lines was accomplished at the Cologne Center for Genomics (CCG). They used the Nimbl-SeqCapEZ-V2 kit for library preparation and enrichment which is a hybrid capture technique. This means that input DNA strands are hybridized to

complementary DNA fragments which carry a bead. With this panel more than 20,000 genes are covered including microRNA genes. The sequencing was performed on a HiSeq2000 platform which is based on the sequencing-by-synthesis method. The bioinformatic analyses were performed by CCG. The pipeline for the analyses is called varbank and the complex workflow can be seen on the homepage of varbank (http://varbank.ccg.uni-koeln.de/).

Custom exome sequencing of Su-DHL-5 and HT cell lines was accomplished at the AROS Applied Biotechnology Company. They used the TruSeq™ Exome Enrichment Kit for library preparation. This enrichment technique is based as well on hybrid capture. This panel includes 20,794 genes and targets more than 200,000 exons. In contrast to the Nimble-SeqCapEZ enrichment it targets also 5'- and 3'-UTR, as well as non-coding RNAs. The sequencing was also performed on a HiSeq2000. The bioinformatic analyses of this data set was performed by Dr. Matthias Schlesner, German Cancer Research Center (DKFZ), Division Theoretical Bioinformatics, Heidelberg, Germany in the framework of the ICGC MMML-Seq as already described [170].

2.2.1.10.2 Targeted-resequencing

Targeted-resequencing of the mnBLL cases 2, 3 and 8 as well as of the cell lines BL-70, BLUE-1 and U-698-M was performed at our Institute using the TruSight™ One Sequencing Panel. This panel targets 4,813 genes associated with human diseases based on information in the human gene mutation database and the online Mendelian inheritance in man catalog. M.Sc. Ulrike Paul of the high-throughput sequencing core facility of the Institute of Human Genetics, Kiel, performed the assay and provided the data for the further analyses. The applied protocol is explained briefly in the following. DNA samples were prepared by tagmentation using transposomes which cut the DNA strands into 300 bp fragments. In a PCR reaction oligonucleotides containing sequencing primers, indices unique to each sample and primers for binding to the flow cell were synthesized to the DNA fragments which were also amplified in parallel. In this panel three sample libraries were pooled. The enrichment of the target genes was achieved using a hybrid capture technique. In this panel biotinylated probes which are complementary to the target regions were hybridized to the samples and the enrichment was performed using streptavidin beads. After washing the samples to eliminate all non-bound DNA fragments, the enriched DNA was eluted from the beads. The prepared libraries were loaded onto a flow cell for sequencing on a MiSeq sequencer. The sequencing method is the same as for the HiSeq2000 (2.2.1.10.1). In the MiSeq workflow the removal of the indices, the alignment of the reads to the GRCh37/hg19 reference genome and production of so called BAM files was already included. The BAM files were used to produce the variance call files (VCF) using the MiSeq Reporter Software. The VCF files contain the information about the type of single nucleotide variant (SNV) as well as if it has already been annotated as a SNP in the dbSNP or a COSMIC SNP

databases. Additionally, the variants were selected according to the quality. Variants which (i) did not pass the filter, (ii) had a read depth lower than 20-fold and (iii) a mutated allele frequency (MAF) below 0.1 were excluded from further analyses. To eliminate false-positive variants from the filtered list, the BAM files were uploaded in the Integrative Genomics Viewer (version 2.3) which is a freeware program launched by the Broad Institute [204]. Each variant read was analyzed individually for artifacts, e.g. a read harboring many SNV is likely of bad quality especially when other samples show at the same position many SNV as well.

2.2.2 Preparation and analyses of RNA

2.2.2.1 Extraction of RNA from cell lines

Two different protocols for extraction of RNA from cell lines were used as described in the following.

2.2.2.1.1 RNA extraction with RNeasy Mini Kit

RNA extraction from cell lines was performed using the RNeasy Mini Kit according to manufacturer´s instructions. The principle is outlined in the following. Between $1x10^6$ - $3x10^6$ cells were pelleted and lysed using RLT Buffer supplemented with 10 % β-mercaptoethanol. Lysates were loaded onto a QIAshredder spin column for homogenization of the lysate. To the homogenate, 70 % ethanol was added and loaded onto RNeasy mini column. The column was washed once with RW1 buffer and then twice with RPE buffer, to wash out any leftover protein and DNA. The RNA was eluted using RNase-free water and stored at -80 °C. The RNA extracted with this method was used for the expression analysis of the alternative transcripts of the *TERT* gene.

2.2.2.1.2 RNA extraction with PARIS Kit

RNA from $1x10^6$ cells was extracted using the PARIS™ Kit. The extraction was performed according to the manufacturer´s protocol which is outlined in the following. Overall $1x10^6$ Cells were lysed in ice-cold Cell Disruption Buffer. The lysate was homogenized in Lysis/Binding Buffer. Afterwards, absolute ethanol was added to the mixture and loaded onto a column to purify the RNA. The column was washed once with wash solution 1 and then twice with wash solution 2/3, to wash out any leftover protein and DNA. RNA was eluted using elution buffer. The RNA was stored at -80 °C. The RNA extracted with this method was used for the expression analysis using qPCR.

2.2.2.2 Determination of RNA quality using Experion® RNA StdSens analysis kit

RNA samples were measured using the Experion® electrophoresis system and the Experion® RNA StdSens analyses kit according to manufacturer´s instructions. The purity and integrity of the RNA

was reported as a RNA quality indicator (RQI), which was calculated by the Experion software. A RQI value of 8-10 denotes RNA with a good quality. RNA with a RQI below 8 was not used for further experiments.

2.2.2.3 Determination of RNA concentration using Qubit® RNA HS Assay

For a quick control of the quantity of the RNA samples, they were measured using the Qubit® RNA HS Assay which is a fluorescence-based assay. It is a much faster method to determine the RNA concentration than using the Experion® electrophoresis system. An advantage of this fluorometric quantification of RNA in comparison to spectrophotometric approaches is the high selectivity of the fluorescent dye for RNA over DNA or free nucleotides. The measurement was performed according to manufacturer´s protocol.

2.2.2.4 Reverse transcription

Synthesis of cDNA from RNA was performed using the QuantiTect Reverse Transcription Kit according to manufacturer´s instructions. For each reaction 1 µg RNA extract was used. Prior to reverse transcription, RNA was subjected to a gDNA wipeout step to digest any left-over DNA. The cDNA synthesis was set up in a volume of 20 µl:

1 µl	RT Primer Mix (consisting of long random and oligo dT primer)
1 µl	Reverse Transcriptase (RT)
4 µl	5x Quantiscript RT buffer
14 µl	Template RNA diluted in dH_2O (after gDNA elimination)

The synthesis was performed at 42 °C for 30 min with a final inactivation step of the reverse transcriptase at 95 °C for 3 min. The cDNA was diluted 1/5 or 1/10 depending on the subsequent applied assay. cDNA was short-term stored at 4 °C or long-term at -20 °C.

2.2.2.5 Quantitative real-time PCR

Quantitative real-time PCR (qPCR) is a fluorescence-based PCR used to simultaneously amplify and quantify target DNA molecules. This technique is commonly used in gene expression analyses. The qPCR reaction mix contains the intercalating fluorophore SYBR® Green which is a cyanine dye. The maximum excitation wavelength of SYBR® Green is 497 nm and fluorescence emission of SYBR® Green I-DNA complex is at 520 nm wavelength which was detected using the Light Cycler ® 480 II.

The concentration of the cDNA was not measured, as after the cDNA synthesis, the reaction mix contains not only the cDNA but also RNAs and nucleotides. Therefore, 1 µl of a 1:10 diluted cDNA

was used per qPCR reaction. Thus, as 1 µg RNA was initially applied for the cDNA synthesis, each qPCR reaction contained about 5 ng RNA.

The qPCR reaction was set up using the QuantiTect® SYBR® Green PCR Kit. A typical qPCR was set up in a volume of 20 µl:

1 µl	1:10 diluted cDNA
10 µl	2x QuantiTect SYBR Green PCR Master Mix
2 µl	QuantiTect Primer Assay
7 µl	ddH_2O

In some cases self-designed primers obtained from Biomers were used instead of the QuantiTect Primer Assay:

1 µl	1:10 diluted cDNA
10 µl	2x QuantiTect SYBR Green PCR Master Mix
0.5 µl	sense-primer (20 pmol/µl)
0.5 µl	antisense-primer (20 pmol/µl)
8 µl	ddH_2O

The qPCR was run under the following conditions:

Denaturation	95 °C	15 min	
	95 °C	30 s	40-45 cycles
Annealing	58 °C	30 s	
Elongation	72 °C	30 s	
Melting curve	72-95 °C	continuous	

At the end of every qPCR reaction a melting curve analysis was performed as a quality control to detect unspecific PCR products. In the melting curve analysis the dissociation-characteristics of double stranded DNA is assessed which is dependent on the length and the GC-content of each DNA fragment. The dissociation is measured as a reduction of fluorescence signal. The temperature at which 50 % of the dsDNA is denatured is called the melting point (Tm). Thus, each qPCR product has a specific melting point. In case of unspecific amplification, several varying melting points would be detected in the melting curve analysis. The respective Cq values were excluded from further analyses.

To quantify the initial amount of target DNA, the so called quantification cycle (Cq) value was used, which has formerly been called crossing point (CP). In an ideal PCR reaction the efficiency is 2, meaning that with every cycle the DNA amount is doubled, which is reflected by a doubling of the fluorescence intensity. In reality the efficiency differs for every PCR reaction depending for example on the primers and target DNA sequence. Therefore, the efficiency of every primer set was determined using the slope of the log-linear portion of the calibration curve. Each calibration curve

was set up using a serial dilution of self-prepared BL-2 or MLMA cDNA from 0.5 ng to 500 ng in triplicates. For most of the assays cDNA from BL-2 was used, as the analyzed genes were not located in a region of chromosomal aberration in this cell line. The sole exception was the calibration curve for the IL10 transcript, for which the MLMA cDNA was used. The calibration curve qPCR was run at least two times. The efficiency of the qPCR (E) was calculated using the formula: $E=10^{-1/slope}$. The mean of the efficiencies was used for further quantification of the gene expression. An example of a calibration curve is given in Figure 7.

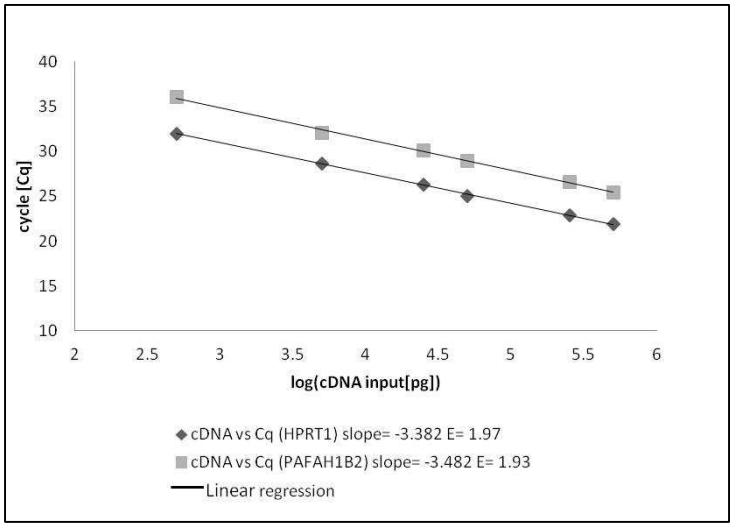


Figure 7: Calibration curve for calculation of PCR efficiency. The squares depict the calibration curve using the HPRT1 QuantiTect Primer Assay and the rhombuses depict the calibration curve for the PAFAH1B2 QuantiTect Primer Assay for one run. The slope and the efficiency (E) are given for this PCR reaction.

The gene expression of the target gene was normalized to the expression of the reference genes *GUSB* and *HPRT1* using the following formula: Ratio= $E_{Ref}^{Cq(Ref)}/E_{target}^{Cq(target)}$ [E_{Ref}: PCR efficiency of the reference gene; E_{target}: PCR efficiency of the target gene; Cq(Ref): Cq value of the reference gene; Cq(target): Cq value of the target gene]. The ratios of normalizations to either GUSB or HPRT1 were averaged (Ratio mRNA expression). The qPCR was repeated at least three times. Using the averaged ratios, the mean ratio and the standard deviation was calculated. The above described approach for gene expression quantification is recommended by the MIQE Guidelines for qPCR [205]. As the PCR efficiencies of the reference gene and target gene must be the same for the ΔΔCq method in order to generate an accurate comparison, which is often not warranted, the here described method is the means of choice.

2.2.3 Array-based techniques

2.2.3.1 Human SNP array 6.0 - DNA microarray

Using the genome-wide human SNP array 6.0 it is possible to analyze more than 900,000 SNPs, distributed throughout the whole genome, as well as the same number of copy number alterations (CNA) in a single experiment for one DNA sample. On the slide of the SNP array 25 mer oligonucleotides, which are called probes, are spotted. Each CNV locus is detected by one probe on the array. The SNP array loci are detected by allele-specific probes, meaning that one probe binds either to allele A or allele B of the respective SNP.

Using the human SNP array 6.0, the cell line MLMA was analyzed. The assay was performed by Claudia Becher, technician of the array-core facility at our Institute who supplied the data for the subsequent analyses which were performed in the framework of this thesis. The genome-wide human SNP array 6.0 analyses have been performed using the Genome wide human SNP NSP/STY assay kit 5.0/6.0 according to manufacturer´s instructions. The workflow is outlined shortly in the following. Prior to hybridization of the DNA samples to the array, the complexity of the genome was reduced by digestion using a restriction enzyme followed by an adapter-ligation and PCR-amplification. The DNA fragments were labeled with biotin and hybridized to the array where they bound to their respective probes. DNA fragments, which were not bound, were washed away. After staining of the biotin-labeled DNA fragments, the array was scanned using the GeneChip Scanner and the CN status was analyzed using the Genotyping Console software.

The analyses of CNA were performed using the following settings: only those CNAs were included which were larger than 100 kb, contained more than 20 markers and which overlapped with known copy number variations (CNV) less than 70 %.

2.2.3.2 OncoScan - DNA microarray

Using the OncoScan^{TM} FFPE Assay Kit it is possible to obtain whole-genome copy number and somatic mutation information from FFPE DNA samples. The method is based on the DNA microarray technique which covers SNPs and, thus, can extract genomic aberrations and loss-of-heterozygosity (LOH) information. The advantage of the OncoScan assay is that it is possible to use FFPE DNA which is often too degraded to obtain good results in conventional DNA microarray systems. The OncoScan assay is based on the molecular inversion probe (MIP) technology. Refer to Figure 8 for detailed description of the MIP technology. The assay was performed according to manufacturer´s instructions by Claudia Becher, technician of the array-core facility of our Institute. The data were supplied for the subsequent analysis of the mnBLL cases 18-25 which was performed in the framework of this thesis. The workflow is outlined shortly in the following.

The sample-DNA was diluted to a final concentration of 12 ng/µl and incubated o/n to anneal with the MIP probe. The following day the gap fill reaction was performed. For this the total volume was divided into two reactions to which either the AT- or the GC nucleotide mix was added. The gap was filled using special gap fill enzymes. Adding exonucleases, unligated probes which were not-gap filled, were removed. Afterwards the samples were cleaved using the cleavage enzyme to linearize the gap-filled, now circular MIP probes. Afterwards the gap-filled, linearized MIP probes were amplified to enrich the MIP product. After an additional cleavage step, the hybridization target was obtained which consists of the tag sequence and the primer sequence. This product is hybridized to the OncoScan™ array overnight (o/n). After washing the array at the Gene Chip® Fluidics Station 450 to remove unbound products, the array was stained using Streptavidin, R-Phycoerythrin Conjugate (SAPE). The fluorescence signals of the arrays were scanned using the Gene Chip® Scanner and so called CEL files were generated.

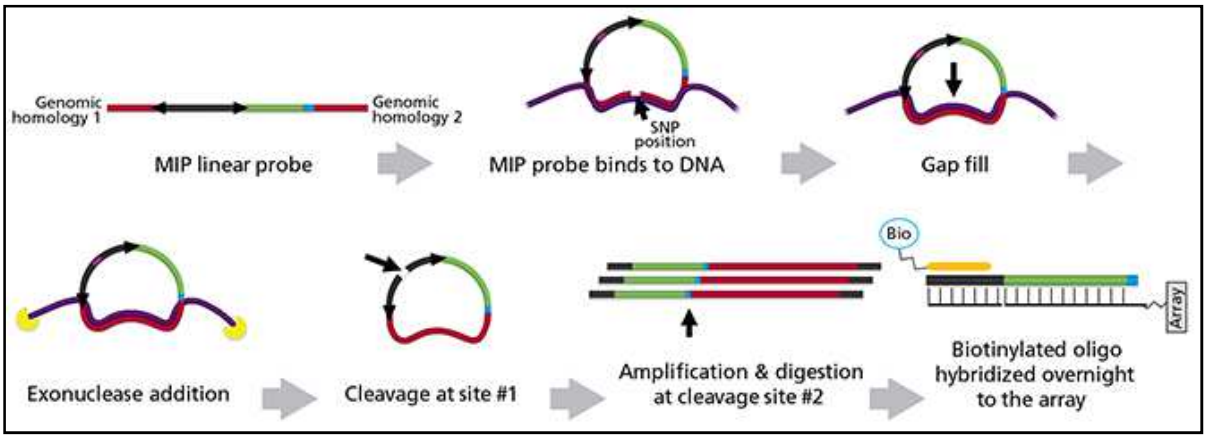


Figure 8: The MIP technology uses a MIP linear probe which contains homologous genomic regions adjacent to the target SNP as well as universal primer binding sites, cleavage sites and a tag sequence. This MIP probe binds to the homologous genomic region with the SNP position left out. The MIP probes form a circular structure. In a next step, the gap is filled-in. After subsequent ligation which can only take place if a nucleotide was incorporated, exonucleases are added which digest unligated probes. Afterwards the circular structure is linearized using an enzymatic digestion of one of the cleavage sites. The linearized products are amplified with the universal primer and cleaved a second time at the cleavage site 2. The product is hybridized to the array. During the hybridization step the digested product binds to the probes on the array and becomes simultaneously biotin-labeled. From http://www.affymetrix.com, 15/10/2014.

The CEL files were uploaded to the OncoScan™ Console which converted the CEL files to OSCHIP files. The converted files were uploaded to the Nexus Express for OncoScan 3.0 software to analyze the chromosomal aberrations. DNA samples derived from FFPE have generally a worse quality than DNA samples derived from cryo material or fresh tissues which is reflected by a higher number of CNA. To exclude false positive CNA, stringent filter criteria were applied. Thus, those CNA were excluded from further analyses which showed copy number gains and losses smaller than 100 kb, LOG2 ratios between -0.3 and 0.3 and overlap of more than 70 % with described CNV.

2.2.4 Analyses of global DNA methylation

The assays for analyses of DNA methylation were performed by Lorena Valles under the supervision of Dipl. Molmed. Julia Kolarova and Prof. Dr. rer. nat. Ole Ammerpohl within the epigenetic core facility of the Institute of Human Genetics, Kiel. They provided the data of the cases as listed in 2.1.1.2, 8.1 and 8.2 for the analyses described in this study. The protocols used are outlined in the following.

2.2.4.1 Bisulfite conversion of gDNA for HumanMethylation 450k Bead Chip

The first step to the analyses of DNA methylation is the conversion of unmethylated cytosine to uracil using sodium bisulfite. Methylated cytosine on the other hand becomes not deaminated by the sodium bisulfite treatment. The conversion is followed by an amplification reaction in which uracil is replaced by thymine. Thus, unmethylated cytosine is detected as thymine in subsequent sequencing reactions. The conversion was performed using the EZ-96 DNA Methylation™ Kit according to manufacturer´s instruction. Briefly, 1 µg of gDNA was diluted in 45 µl ddH_2O and 5 µl M-dilution buffer was added. The samples were incubated at 37 °C for 15 min in a thermocycler. 100 µl of the CT Conversion reagent were added and mixed well. The conversion reaction was performed in a thermocycler using the following program:

95 °C	30 s	} 16 cycles
50 °C	60 min	
4 °C	on hold	

400 µl binding buffer and converted sample were loaded onto the Zymo-spin columns and mixed by pipetting. After centrifugation at 10,000 g for 30 s, the supernatant was discarded. After applying 100 µl M-Wash buffer to the columns, they were centrifuged at 10,000 g for 30 s and the supernatant was discarded. 200 µl M-Desulphonation buffer was added and incubated at room temperature for 15-20 min. After centrifugation at 10,000 g (Mikro 200) for 30 s, the supernatant was discarded. 200 µl M-Wash buffer was added to every column and centrifuged at 10,000 g (Mikro 200) for 30 s. The washing step was repeated and the columns transferred to new reaction tubes. The converted DNA was eluted in 10 µl ddH_2O by centrifugation at 10,000 g (Mikro 200) for 30 s. The converted DNA was stored at -20 °C.

2.2.4.2 HumanMethylation 450k Bead Chip analysis

To determine the methylation landscape of mnBLL, they were analyzed using the HumanMethylation 450k Bead Chip. The results were compared to the available methylation data of BL and DLBCL, which have been already analyzed in the framework of the ICGC MMML-Seq and MMML projects. The quite

recently obtained data are still undergoing statistical analysis, but the data to determine the methylation status of the *IL10* locus were available for initial analysis. The principle of the HumanMethylation 450k Bead Chip is shortly outlined in the following:

The HumanMethylation 450k Bead Chip allows the parallel analyses of 485,577 different cytosine-guanine-dinucleotides (CpGs) corresponding to 21,231 genes. The principle of the array is based on two different bead-type assays: Infinium I assay and Infinium II assay. The Infinium I assay uses two different bead types which have probe sequences either matching the methylated or unmethylated CpG site. If a CpG target site matches with the respective probe, a single nucleotide is extended which is labeled with biotin or dinitrophenyl (DNP). The level of methylation for a CpG site is determined by the ratio from fluorescent signals from the methylated bead vs. the unmethylated bead. The Infinium II assay on the other hand has only one bead type corresponding to one CpG site. The 3'-end of the probe sequence stops one base before the query CpG target site. Thus, in a single-nucleotide extension either a biotin-labeled cytosine or guanine in case of a methylated CpG site or a DNP-labeled adenine or thymine in case of an unmethylated CpG site is incorporated. Thus, other than in the Infinium I assay the level of methylation for a CpG site is determined by the ratio of fluorescent signal from the methylated CpG vs. fluorescent signal from the unmethylated CpG.

The analyses of the DNA-methylation of mnBLL was performed using the "Illumina Infinium HD MethylationAssay" manual (state 2011) and the Human Methylation 450k Bead Chip Kit which is shortly outlined in the following. The bisulfite converted DNA was enzymatically fragmented. The DNA was precipitated using absolute isopropyl alcohol and subsequently resuspended in hybridization buffer. The samples were loaded onto the Human Methylation 450k Bead Chip and incubated in a Hybrid oven at 48 °C for 16-22 h. During this time the fragmented DNA binds to its complementary probes. DNA molecules which were not bound were removed by a washing step. During the single strand elongation a biotin-labeled cytosine or guanine or a DNP-labeled adenine or thymine was incorporated. Different fluorescence-labeled antibodies interact with the either biotin- or DNP-labeled nucleotide and the fluorescence signal was measured using the iScan. The methylation status was given as AVG-β values which correspond to the ratio of methylated signal to the sum of methylated and unmethylated signals at one CpG.

2.2.5 Cell biology methods

2.2.5.1 Cultivation of cells

The cultivation of eukaryotic cell lines was performed under sterile conditions using a laminar flow hood, sterile instruments and sterile growth media. RPMI-1640 medium supplemented with 10 % (v/v) Fetal Bovine Serum (FBS) and glutamine (2 mM) was used for cultivation of the cell lines BL-2,

BL-30, BL-41, BL-70, BLUE-1, Ca46, DAUDI, DG-75, EB-1, EB-3, Namalwa, Raji, Ramos, U698-M, Su-DHL-5, MLMA, Su-DHL-10 and RIVA. The percentage of FBS was raised to 20 % in the media for the Karpas422 and Su-DHL-6 cell lines and to 15 % for the HT cells. The cell line Hek293 was cultivated in D-MEM and 10 % FBS (v/v). The cell lines were cultivated in T25 flasks using an incubator with humidified atmosphere containing 5 % CO_2 at 37 °C. Cells were split every two or three days to prevent an overgrowth. To split the adherent Hek293 cells, they were washed once with PBS and then treated with 1 ml trypsin until the cells were detached from the bottom of the flask. The trypsin was inactivated with media containing 10 % FBS and the cells were pelleted at 1,300 rpm for 3 min (Multifuge 1L-R). The supernatant was discarded, the cells supplemented with their respective fresh media and a part of the cells were used for continued cultivation. Suspension cells were split by diluting them with fresh complete medium.

2.2.5.2 Freezing and thawing of cells

In order to freeze cells, they were pelleted at 1,300 rpm for 3 min (Multifuge 1L-R). The supernatant was discarded and the pellet resuspended in 1 ml freezing medium (90 % FBS, 10 % DMSO). The cell suspension was transferred into a cryo vial and frozen at -80°C o/n. The next day the vial was transferred into a -145 °C freezer.

To thaw cells the content of the cryo vial was diluted in 5 ml of the respective cell line cultivation media. The cells were pelleted at 1,300 rpm for 3 min (Multifuge 1L-R) and the supernatant was discarded. After resuspension of the cell pellet in fresh medium, the cells were transferred into a T25 flask and cultivated in an incubator.

2.2.5.3 Cell counting

The cell number was determined using a Neubauer hemocytometer. 10 µl cell suspension was mixed with 10 µl trypan blue solution and half of the mixture was applied to the hemocytometer. Four of the nine quadrants were counted. The number of cells in the growing culture was estimated using the following formula: number of cells per milliliter of culture= (number of cells in each quadrant/4) x 10,000 x 2 (to take into account the dilution factor due to the volume of trypan blue solution added). Based on this number, the volume of cell suspension for every experiment was calculated.

2.2.5.4 Mycoplasma test

Mycoplasmas are intracellular bacteria. Upon infection they can influence the cell physiology as inhibition of cell proliferation or alteration of gene expression. Therefore regular testing of the cell culture for Mycoplasma contamination is essential. To screen the cell lines for possible mycoplasma

contamination a PCR-based assay was performed. This assay detects the 16S rRNA gene of Mycoplasma genus which is well preserved in prokaryotes. Primer sequences were as follows: 5'-CACCATCTGTCACTCTGTTAACC-3' and 5'-GGAGCAAACAGGATTAGATACCC-3' published by [206]. The cell lines were cultured for at least two days in the same media and 1 ml cell supernatant was centrifuged at 250 g (Multifuge X1R) for 3 min to remove any remaining cells. Afterwards, the supernatant was transferred into a fresh tube and centrifuged again at 15,000 rpm (Multifuge X1R) for 10 min. After discarding the supernatant, the pellet was resuspended in 50 µl ddH_2O and heated to 95 °C for 3 min in a thermomixer. 5 µl of this solution were used in a PCR set up as described in 2.2.1.5. The PCR products were analyzed on an agarose gel. The presence of a 270 bp sized band indicated a Mycoplasma-positive sample.

2.2.5.5 Authentication of cell lines using STR profiles

Inter- and intraspecies cross-contamination of human and animal cells and misidentification are among the major problems in biological studies based on cell cultures and usually lead to the production of false data [207],[208]. Consequently, regular authentication of the cell lines in culture as well as when obtaining new cell lines is crucial. The authentication of the cell lines was performed using microsatellite loci in the human genome harboring short tandem repeat (STR) DNA markers. The usage of nine different STR loci allows the individualization of cell lines at the DNA level. The analyses were performed using the StemElite™ ID System according to manufacturer´s instructions. In this assay, nine STR loci are amplified in a multiplex PCR. The primers for the amplification of the nine loci are labeled either with fluorescein, carboxytetramethylrhodamine or 6-carboxy-4',5'-dichloro-2',7'-dimethoxy-fluorescein. Thus, based on the various lengths of the PCR products and fluorescence-labels the loci can be differentiated by capillary electrophoresis and three-fluorescence-color detection. The DNA of the cell lines was isolated as described before (2.2.1.1). Capillary electrophoresis was performed on an ABI PRISM® 3130 Genetic Analyzer. Allele analyses were done using the GeneMapper 5 Software. The resulting STR profiles were compared to the online STR Analyses database for cell lines (http://www.dsmz.de/STRanalyses). Only cell lines with correct STR profiles were used for further analyses.

2.2.5.6 Ectopic expression of genes of interest

Transfection is a process of introducing nucleic acids such as plasmid DNA into eukaryotic cells. In this study electroporation was used to transfect human B-cell lymphoma cell lines. Electroporation uses an electric field generated by an electroporator to disrupt the permeability of the cell membrane. If the electroporation is performed in presence of nucleic acids, these will enter the cells due to the

change in permeability. The pulse duration and shape of the electric field can be set manually, in order to obtain the ideal transfection condition. The best electroporation conditions are those that achieve a high transfection efficiency and high cell viability. The electroporation protocol needed to be optimized for each cell line. The optimization was performed by testing different transfection programs of the Nucleofector I device using the pmax vector included in the transfection kit, which harbors a gene encoding the green fluorescent protein (GFP). Three days after the electroporation, the transfection efficiency was determined. In total 100 cells were counted using a fluorescence microscope and the amount of GFP-expressing cells was determined. Additionally, cells were stained with trypan blue solution (2.2.5.3) to determine the number of viable cells.

In this study the BL cell lines BL-2 and DG-75 were transfected with the commercially available human *ID3*-pCMV6-AC-*GFP* plasmid (2.1.8). The nucleofection protocol is described in the following. At the day of the transfection cells were harvested and counted. Per transfection reaction $1x10^6$ cells were used. The appropriate number of cells was pelleted at 90 g (Multifuge 1L-R) for 5 min. The cell pellet was resuspended carefully in nucleofection solution V supplemented with 18 % nucleofection supplement. 5 µg of plasmid DNA were added to the cell suspension and mixed carefully. The cell-DNA suspension was transferred to nucleofection cuvettes using the supplied transfer pipette. The transfection program of the Nucleofector I device was set to C-09 for BL-2 and to R-13 for DG-75. After nucleofection cells were directly transferred into a well of a 6-well plate and supplemented with medium. To ensure proper expression of ID3-GFP*,* cells were cultivated for two days before further experiments were performed.

2.2.5.7 Knock down of target genes

In this study three different, commercially available short hairpin RNA (shRNA) constructs (pLKO_IPTG_3xLacO) were used, targeting PAFAH1B2 (PAFAH1B2 shRNA1-3) (2.1.8). ShRNAs use the endogenous cellular processing system of microRNAs (miRNAs) to form the effector complex for RNA interference.

A lentiviral vector system was used to transfer foreign DNA in human B-cell lymphoma cell lines. This gene delivery method is called transduction. To generate lentiviral particles, two different components are needed: packaging constructs which make up the structural components of the vector and a transgene construct containing the shRNA targeting the gene of interest as well as an resistance gene for an antibiotic. The structural components consist of the structural protein gag which forms the viral capsid and the nucleoprotein complex, the reverse transcriptase and integrase protein pol and the viral glycoprotein env which is displayed on the surface of the vector and defines the tropism. The packing constructs pCMV-dR8.91 [209] and pMD.G [210] were kindly provided by Dr. rer. nat. Eva Murga-Penas, Institute of Human Genetics, Kiel.

In the following the transduction protocol is explained in detail. The following experiments were performed according to S2 gene technology safety regulations at the respective S2-laboratory of the Department of Pediatrics, group of Dr. med. Denis Schewe, Christian-Albrechts University Kiel & University Hospital Schleswig Holstein, Kiel.

2.2.5.7.1 Production of lentiviral vectors

Due to its high transfection efficiency, Hek293 cells are commonly used for production of lentiviral particles. At day one, the Hek293 cells were seeded at a density of 3.5 million cells in a 10 cm dish. The following day the cells were transfected with 5 µg of each the packaging constructs pMD.G and pCMV-dR8.91 as well as 10 µg of the pLKO_IPTG_3xLacO construct carrying the shRNA sequence. The jetPEI DNA transfection kit was used for the transfection of the Hek293 cells in a 10 cm culture dish. For each transfection, all three plasmid constructs were diluted in 150 mM NaCl to a final volume of 250 µl in one mixture. In an additional mix, 40 µl of the cationic polymer of the jetPEI Kit were diluted in 150 mM NaCl to a final volume of 250 µl. The two solutions were mixed and incubated for 30 min at room temperature, to let the nucleic acids and polymer complex build up. The plasmid/polymer complexes were added drop wise to the Hek293 cells. The transfected cells were incubated for 24 h in the incubator before the medium was changed. 48 h after transfection the viral titer was described to be the highest and the viral vector was harvested. The supernatant was collected and centrifuged for 5 min at 1,100 rpm (Centrifuge 5810) to remove any remaining dead cells. The medium was sterile filtered using a 0.45 µm membrane to remove any cell debris. The filtered virus-containing medium was aliquoted and stored at -20°C. The transduction was performed shortly afterwards as the virus titer decreases during long time storage.

The presence of lentivirus in the supernatant was confirmed using the lentiviral titer test Lenti-X™ GoStix™ according to manufacturer´s instructions. This rapid strip test detects the lentiviral p24 capsid protein. The supernatant is applied to the GoStix cassette and Chase Buffer is added. If the sample contains sufficient levels of lentivirus, a band will appear in the cassette after 30-180 s. The lower detection limit of this assay is about $5x10^5$ infectious units/ml.

2.2.5.7.2 Transduction of target cells with lentiviral vectors

Prior to the transduction, Su-DHL-5 cells were seeded at a density of $2x10^5$ cells/well/1 mL in a 6-well plate. After at least 6 h, 2 µl polybrene (10 mg/ml) was added to the cells. Polybrene is a polycation which facilitates the interaction between the negatively charged viral particles and their target cells and therefore increases the transduction efficiency. Afterwards, 1 ml of either freshly produced or thawed medium containing the viral particles were added to the target cells. The transduced cells were incubated for 24 h in the incubator. The following day the medium was changed and fresh

medium added. Two days after transduction the selection of the Su-DHL-5 cells was started using medium supplemented with 2 µg/ml puromycin. The puromycin concentration, which effectively kills Su-DHL-5 cells, has been defined beforehand in a titration curve.
After one week of puromycin selection only those cells survived which had taken up the lentiviral shRNA construct.

2.2.5.7.3 Induction of shRNA expression using IPTG

The expression of the shRNA of interest from the pLKO_IPTG_3xLacO vector is controlled by a modified human U6 promoter construct containing LacO (operator) sequences. Furthermore, the vector contains a LacI (repressor) gene (refer to the appendix 8.5 for an overview on the vector map). The LacI protein is expressed from the lentiviral vector and in the absence of isopropyl-β-D-thiogalactoside (IPTG) it is bound to the LacO promoter region. Hence, it prohibits the expression of the shRNA. Binding of IPTG to the LacI leads to a conformation change of LacI and subsequent release from the LacO sequence. Hence, the RNA polymerase can bind to the U6 promoter and the shRNA is transcribed.
To determine the IPTG concentration giving the best knock down efficiency, the transduced cells were treated with various concentrations ranging from 10 µM to 1 mM IPTG diluted in sterile ddH_2O for three days. To ensure that IPTG had no impact on the gene expression of the target gene and/or cell viability, cells transduced with a non-targeting control (NTC) shRNA were included in the analyses. The knock down efficiency was analyzed using Western blot. The best knock down efficiency was obtained after treatment of Su-DHL-5 cells with 100 µM and 250 µM IPTG. These concentrations were used in subsequent experiments. To determine whether a knock down of PAFAH1B2 had an impact on cell viability, three days after induction cells were analyzed using a cell viability assay (2.2.5.10).

2.2.5.8 Treatment of cell lines with neutralizing antibodies

2.2.5.8.1 IL10 inhibition

To test if neutralizing IL10 in the cell culture supernatant has an impact on cell viability, cells were treated with a neutralizing IL10 antibody (ab34843). The analyses were performed in BL (BL-2, BL-70), mnBLL (HT, Su-DHL-5, MLMA) as well as in one non-BL (RIVA) cell lines. Cell lines were harvested, centrifuged at 1,100 rpm for 3 min (Multifuge 1L-R) and supplemented with fresh medium to remove any remaining IL10. Cells were counted and seeded at a density of 15,000 cells/well in a 96-well plate. The cell lines were treated with various concentrations of IL10 antibody ranging from 50 µg/ml to 1 mg/ml, which was diluted in PBS if necessary, in duplicates. For each cell line an untreated control was included in the experiment. After two days of treatment in an incubator, cells were

analyzed using the cell viability assay (2.2.5.10). The experiment was repeated three times as biological replicate.

2.2.5.8.2 IL10RA inhibition

The inhibition of IL10RA with a neutralizing antibody was performed using two different antibodies:

- CDw210 (3F9) has been published to neutralize IL10RA [211],[212]
- IL10RA (T3431) was used for IL10RA Western blot

The analyses were performed in a BL (BL-2) and a mnBLL (Su-DHL-5) cell line. Cell lines were harvested, centrifuged at 1,100 rpm for 3 min (Multifuge 1L-R) and supplemented with fresh medium. Cells were counted and seeded at a density of 15,000 cells/well in a 96-well plate. Cells were treated in duplicate with various concentration of antibody ranging from 2.5 µg/ml to 10 µg/ml diluted in PBS. The experiment was performed as biological triplicate. For each cell line an untreated control was included in the experiment. After two days of treatment in an incubator, cells were analyzed using the cell viability assay (2.2.5.10). The experiment was repeated three times as biological replicate.

2.2.5.9 Treatment of cell lines with BKM120

In order to analyze if the mnBLL cell lines depend on the downstream signaling of the IL10 receptor for survival, PI3K signaling was inhibited using the selective PI3K inhibitor BKM120, which is also termed Buparlisib. This compound has been recently published to successfully inhibit the proliferation of BL cell lines, which were shown to depend on the PI3K signaling for cell survival [72]. The analyses were performed in BL (BL-2, BL-41, BL-70), mnBLL (HT, Su-DHL-5, MLMA) as well as non-BL (Su-DHL-10, Karpas422) cell lines. Cells were seeded in duplicate at a density of 6,000 cells/well in a 96-well plate and treated with various concentration of BKM120 ranging from 0.15 µM to 2 µM BKM120 diluted in DMSO. For each cell line an untreated control was included in the experiment. After three days of treatment in an incubator, cell viability was analyzed using the cell viability assay (2.2.5.10). The experiment was repeated three times as biological replicate.

2.2.5.10 Cell viability assay

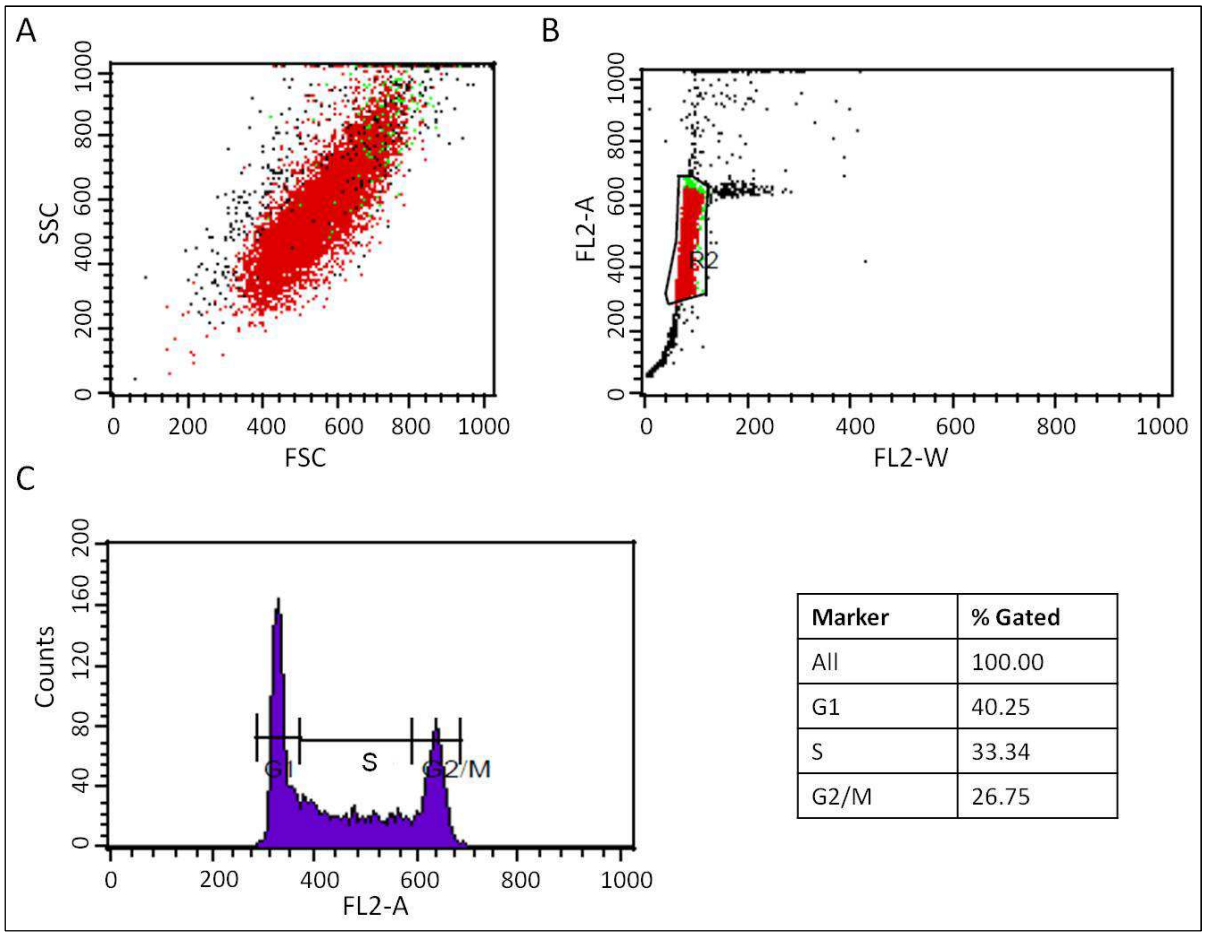
The determination of the amount of viable cells was performed using the Cell Titer Blue reagent. This assay is based on the reduction of resazurin to resorufin which is highly fluorescent ($579_{Ex}/590_{Em}$). The reduction becomes furthermore visible as a “blue shift” from blue to pink (absorbance maximum resazurin: 605 nm and resorufin: 573 nm). Only metabolically active cells are capable of this reaction which is an indicator for cell viability.

After treating cells in a 96-well plate using the above described assays, 20 µl of cell titer blue reagent was added to the cells. Depending on the cell type, the cells were incubated between one to three hours in an incubator until the reduction became visible as described above. The fluorescence of resorufin was measured at $540_{Ex}/590_{Em}$ using the Synergy HT ELISA plate reader. The fluorescence signals were displayed by the Gen5^{TM} Software and subsequent exported as an excel document. In each experiment a medium control (blank) was included, which was used to determine the background fluorescence of the medium. Initially the fluorescence signal of the blank was subtracted from every fluorescence signal. As every treatment condition was performed in duplicate, the mean of the fluorescence signals was calculated. The relative viability was calculated as ratio of fluorescence signal from treated samples to untreated control.

2.2.5.11 Cell cycle measurement

Cell cycle analyses were performed using propidium iodide according to a published protocol [213] and with the help of Prof. Dr. rer. nat. S. Adam, Institute of Immunology, Kiel. Propidium iodide is an intercalating agent which binds to DNA. When bound to DNA, propidium iodide has its fluorescence excitation maximum at 535 nm and the emission maximum at 617 nm. In cell cycle analyses the nuclear DNA content of a cell is quantitatively measured using flow cytometry, based on the principle that the incorporated amount of propidium iodide is proportional to the amount of DNA [214]. Thus, based on the detected fluorescence signal, the cells can be assigned to the various cell cycle phases (in cells without polyploidy= G0/1-phase: diploid chromosomal DNA content, S-phase: between 1x to 2x chromosomal DNA content, G2/M-phase: tetraploid chromosomal DNA content). For the analysis, $1x10^6$ cells were harvested and centrifuged at 1,100 rpm for 3 min (Multifuge 1L-R). The supernatant was discarded and the cells were washed with 1 ml of ice-cold PBS, followed by centrifugation at 1,100 rpm for 3 min (Multifuge 1L-R). This washing step was repeated once more. After discarding the supernatant the cells were resuspended in ice-cold 500 µl PBS. For fixation of the cells, the same volume of absolute ethanol was added dropwise under constant mixing on a vortex mixer. Afterwards, the cells were incubated for 30 min on ice. Cells were centrifuged at 1,100 rpm for 3 min (Multifuge 1L-R), the supernatant discarded and the cells resuspended in 200 µl PBS. To remove cellular RNA, 8 µl RNase A (1 mg/ml, diluted in ddH_2O) were added. Finally, 200 µl propidium iodide (50 µg/ml, diluted in ddH_2O) were added and cells were incubated for 30 min at room temperature in the dark. Cell cycle analysis was performed by flow cytometry using a FACS Calibur Analyzer and the Cell Quest software. The workflow of the analyses as depicted in Figure 9 is briefly described in the following. In each cell cycle analysis $1x10^4$ cells were analyzed. To exclude dead cells and debris from the following analysis, the cell population was initially analyzed in a two-parameter dot plot of the signals from the forward scatter (FSC) and side-scatter (SSC) (Figure 9A). The viable cell population

was gated and the emitted fluorescent light of propidium iodide which was recorded in the FL2 channel was analyzed. A common problem of the DNA analysis is the findings of doublets. Doublets are two cells with G1-phase DNA content which are recorded as one event and are hence, similar to cells with G2/M-phase DNA content. These doublets can be excluded by analyzing the pulse-width (FL2-W) of the fluorescent signal of propidium iodide (PI), as FL2-W increases with the diameter of the particle. Thus, G1-phase doublets have a higher FL2-W signal. Hence, in a dot plot graph depicting the FL2-W vs. the FL2-A (pulse area), the latter representing the relative DNA amount, a gate was set around the single cell population (Figure 9B). This gated cell populations were shown in a histogram depicting the FL2-A signal vs. the count of cells. Furthermore, the four distinct phases of a proliferating cell population can be identified (Figure 9C). Using the Cell Quest software the amount of cells in each cell cycle phase is calculated.



Marker	% Gated
All	100.00
G1	40.25
S	33.34
G2/M	26.75

Figure 9: Example of cell cycle analysis of BL-2 cell line. A: Dot blot graph of forward-scatter (FSC) vs. side-scatter (SSW) signals of the cell population. B: Depicts the two-parameter dot blot of FL2-W vs. FL2-A. The gate, depicted here as box with R2 label, contains solely single cells. C: Histogram of FL2-A, representing the relative DNA amount vs. the count of cells (counts). The cell cycle phases are indicated by the respective lines and labeled as G1 (G1-phase), S (S-phase) and G2/M (G2/M-phase). The table shows the amount of cells in the respective cell cycle phases. % Gated refers to the amount of cells gated in G1-, S- and G2/M-phases in the histogram. FL2-W, fluorescence channel 2 width; FL2-A, fluorescence channel 2 area; counts, number of cells.

2.2.6 Proteinbiochemical analyses

2.2.6.1 Protein extraction from cell lines

To obtain whole protein extracts from cell lines, cells were harvested and washed once with PBS. Depending on the size of the cell pellet, the cells were lysed in 40-100 µl RIPA buffer. The lysate was incubated on ice for 30 min and mixed from time to time. After centrifugation of the lysate at 14,000 rpm (Multifuge X1R) for 10 min at 4 °C, the supernatant was transferred into a new tube and stored at -80 °C.

2.2.6.2 Determination of protein concentration

The protein concentration was determined using the BCA Macro assay kit. The analyses were performed in a 96-well plate and each sample was measured in duplicate. In each measurement a standard curve according to manufacturer´s protocol was included. To obtain a 1/5 dilution, 10 µl water were added to 2.5 µl of the protein solution. 100 µl master mix, consisting of 1/50 dilution of solution A in solution B, was added, mixed and incubated for 30 min at 37 °C. The absorption was measured at 562 nm using the Synergy HT and the data were analyzed using the Gen5™ Software. Using the standard curve the protein concentrations of the samples were interpolated.

2.2.6.3 Western blot analyses

2.2.6.3.1 SDS-polyacrylamid gel electrophoresis

10-20 µg of each protein sample was diluted in 4x NUPAGE® LDS sample buffer and boiled at 95 °C for 5 min. Afterwards they were centrifuged shortly to collect the liquid at the bottom of the tube. Premade Criterion TGX Precast gels were inserted into the Criterion Cell gel chamber, which was filled with SDS-Running buffer according to manufacturer´s instructions. Depending on the size of the protein of interest a gradient gel, termed Criterion TGX Precast Any kD gel or the Criterion TGX Precast 18 % Resolving gel was used. The protein samples were loaded and the separation was performed at 100 V for 20 min until the electrophoretic front had entered the stacker gel completely. Afterwards the voltage was set to 150 V and separation completed through the running gel. The gel was taken out of the unit and the plastic case was broken to uncover the gel for blotting.

2.2.6.3.2 Gel Transfer

Proteins were blotted onto a PVDF-membrane using a tank blot cell transfer device according to manufacturer´s instructions. Prior to blotting, the PVDF-membrane was activated using absolute methyl alcohol for 30 s, afterwards, it was equilibrated together with Whatman filter papers in the

Bjerrum-Schäfer transfer buffer for at least 10 min. Blotting was performed at 100 V and 0.5 Amp for 1 h. To monitor a successful transfer, the membrane was stained with Ponceau S which binds reversible to the amino-residues of the proteins. For the staining, the membrane was incubated for at least 2 min at room temperature with Ponceau Red Staining Solution. Afterwards, the membrane was washed with H_2O until the protein bands were clearly visible. Subsequently, the membrane was washed with H_2O until the staining came off.

2.2.6.3.3 Immunoblot

The membrane was incubated in blocking solution for at least 1 h at room temperature or o/n at 4 °C. The specific antibody against the protein of interest was diluted in antibody diluent based on the manufacturer´s instructions or previous results. The antibody solution was added to the membrane and incubated at room temperature on a tube roller mixer for 1 h. The membrane was washed twice with antibody diluent for 5 min at room temperature. HRP-coupled secondary antibody was diluted 1/2,000 in antibody diluent and added to the blot. Incubation was performed for 1 h at room temperature on a tube roller mixer. The membrane was washed three times with TBS-T for 10 min each at room temperature and then in TBS for 5 min at room temperature for detergent removal.

2.2.6.3.4 Chemiluminescence

Luminata Forte was added to the membrane for 1-2 min to induce the horseradish peroxidase reaction. The liquid was drawn off until the membrane ran dry. The membrane was put into a X-ray film box. An Amersham Hyperfilm™ ECL was exposed for 1-60 s depending on signal intensity and developed using the Curix 60.

2.2.6.4 Enzyme-Linked Immunosorbent Assay for IL10 expression analyses

To determine the IL10 protein amount in cell culture supernatant, the cell lines BL-2, BL-41, BL-70, BLUE-1, Ca46, U-698-M, HT, Su-DHL-5, MLMA, Su-DHL-6, Su-DHL-10, Karpas422 and RIVA were seeded at a density of 250,000 cells/well in a 24-well plate. After two days, cell supernatant was harvested and stored at -80°C. Two different media conditions were tested: media supplemented with 5 % FBS or 10 % FBS. For detection of IL10 protein the human IL-10 Enzyme-linked Immunosorbent Assay (ELISA) Kit was used according to manufacturer´s instructions. The protocol is outlined in the following. 50 µl of the cell supernatant sample was added per well in duplicate. In the same plate different concentrations (0-600 pg/ml) of purified IL10 were also added in order to create a reference standard curve. 50 µl of biotinylated antibody reagent was added and incubated for 2 h

at room temperature. Solution was discarded and each well washed three times with 1x wash buffer at room temperature. 100 µl streptavidin-HRP solution was added and incubated for another 30 min at room temperature. The solution was discarded and washed again three times with 1x wash buffer. To each well 100 µl TMB substrate solution was added and incubated for 30 min at room temperature in the dark. After adding 100 µl of stop solution to each well, the color change was measured at 450 nm and at the reference wavelength 550 nm to correct for background signal using the Synergy HT. The measured absorbance was displayed by the Gen5^{TM} Software and subsequent reported as an excel data sheet. The average absorbance in nm obtained for each concentration of the standard curve was plotted on the vertical axis and the corresponding IL10 concentration in pg/ml was plotted on the horizontal axis. Using linear regression, a best-fit straight line was plotted through the data points. The equation of the linear regression was: y=mx+b [y: absorbance (nm); m: coefficient, rate and slope of line; x: concentration (pg/ml) and b: point at which standard curve crosses vertical-axis). This equation was used to determine the IL10 concentration of each sample. Initially, as the experiment was performed in duplicates, the mean of the measured absorbance from each sample was calculated. Using the average absorbance the IL10 concentration was interpolated from the above described equation.

The IL10 protein amount was analyzed also in sera from Burkitt lymphoma and MYC-negative Burkitt-like lymphoma patients using the same protocol as outlined above. Refer to 2.1.1.3 for an overview on the cases.

2.2.6.5 Expression analyses of IL10RA using flow cytometry

2.2.6.5.1 Cell surface staining of IL10RA

To analyze the surface IL10RA expression on cell lines, 5x10^5 cells per staining were used. All centrifugation steps were performed at 1,300 rpm for 3 min using the Multifuge X1R.

The counted cells were centrifuged. The supernatant was discarded and the cells resuspended in 500 µl FACS wash buffer. The cells were centrifuged again and the washing step was repeated once more. Finally, the cell pellet was resuspended in 15 µl FACS wash buffer and either 125 ng IL10Rα-PE antibody or goat IgG control-PE antibody were added. The cells were incubated for 35 min on ice in the dark. Afterwards, cells were washed twice with FACS wash buffer as described above. The cell pellet was resuspended in 250 µl 1 % PFA solution and stored at 4 °C in polystyrene round-bottom tubes until the measurement was performed. The amount of IL10RA-positive cells was determined by flow cytometry counting at least 1x10^4 cells using the Accuri C6 at the Department of Pediatrics, group of Dr. Denis Schewe, Christian-Albrechts University Kiel & University Hospital Schleswig-Holstein, Kiel. The workflow of the analyses using the BD C Sampler Software is briefly described in

the following. To exclude dead cells and debris from following analysis, the cell population was initially analyzed in a two-parameter dot plot of the signals from the forward scatter (FSC) and side-scatter (SSC). A gate was set on the viable cells. The viable cell population was subsequent analyzed using a two-parameter dot plot of the FSC vs. the FL2A in which the fluorescence signal of the PE-conjugated antibody was detected. Initially, the cells stained with the isotype control (Goat IgG control-PE antibody) were analyzed to determine the background fluorescence signal. Using these settings as a cut-off, the cells stained with the IL10Rα-PE antibody were analyzed. Those cells which showed higher fluorescence intensity than the isotype control were defined as IL10RA-positive. The amount of IL10RA-posivie cells was used for further analysis. Moreover, for visualization of the IL10RA-positive cells, the cell population stained with IL10Rα-PE antibody were displayed in the same histogram (FL2A vs. count of cells) as the cell population stained with the isotype control.
The analysis of the IL10RA cell surface expression was performed in duplicate. The number of IL10RA-positve cells of each experiment was averaged.

2.2.6.5.2 Intracellular staining of IL10RA

The analyses of the intracellular IL10RA expression in cell lines was performed using the same antibodies as described in 2.2.6.5.1. For each staining $5x10^5$ cells were used. All centrifugation steps were performed at 1,300 rpm for 3 min using the Multifuge X1R.
After the counted cells were centrifuged, the supernatant was discarded and the cells resuspended in 500 µl PBS. After an additional centrifugation step, the cell pellet was fixed in 1 ml 1 % paraformaldehyde solution for 20 min on ice. Cells were pelleted by centrifugation and the supernatant was discarded. Cells were washed once with PBS as described above and the cell membranes were permeabilized using 1 ml permeabilization buffer for 2 min on ice. Cells were washed twice as described above using 500 µl FACS wash buffer II. After resuspension of the cell pellet in 15 µl FACS wash buffer II, 125 ng IL10Rα-PE antibody or goat IgG control-PE antibody were added and incubated for 35 min on ice in the dark. Afterwards, cells were washed twice with FACS wash buffer II. The cell pellet was resuspended in 250 µl 1 % PFA solution and stored at 4 °C in polystyrene round-bottom tubes until the measurement was performed. The amount of IL10RA-positive cells was determined as described above (2.2.6.5.1) using the Accuri C6 at the Department of Pediatrics, group of Dr. Denis Schewe, Christian-Albrechts University Kiel & University Hospital Schleswig-Holstein, Kiel.
The analysis of the intracellular IL10RA expression was performed in duplicate. The number of IL10RA-positve cells of each experiment was averaged.

2.2.7 Bioinformatics

2.2.7.1 Modeling of open reading frames

Using the ExPASy translation tool the transcript sequences of *TERT* (NM_198253.2) including the different splice variants from the new exon in combination with exon 2 or exon 2V were uploaded to obtain the open reading frames (ORF). For each uploaded CDS three 5'-3' reading frames (frame 1-3) and the included ORFs were reported. Of note is that it is possible that one reading frame contains more than one ORF. Only ORFs longer than 100 amino acids were considered for further analysis. Furthermore, to determine if the ORFs contain domains similar to TERT or even the same domains, the predicted protein sequences were uploaded in the SMART tool which compares the protein sequence with a database of protein domain sequences. Afterwards, the predicted protein domains of each ORF were compared using the LALIGN tool to the protein sequence of the so called wildtype TERT protein to decipher similarities and differences.

2.2.7.2 Pluripotency plot

The pluripotency plot was provided by M.Sc. Michael Lenz, Joint Research Center for Computational Biomedicine, RWTH Aachen University. The pluripotency plot depicts the median expression level for each gene/tag of the U133A microarray of all ESC and all mBL samples. Every microarray tag is represented by a grey dot whereas the genes of the pluripotency signature are depicted by red-circled black dots. Furthermore, the log-fold change was calculated. The log-fold change of 1 and -1 was used as cut-off for definition of differentially expressed genes.

2.2.7.3 Statistical analyses

Based on the data of the experiments the mean and the standard deviation were calculated.

The significance of the differences in the cell cycle measurement (3.1.3.2) has been calculated using the unpaired, parametric t-test as published [170]. P-values were denoted as: not significant $p > 0.05$, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

The gene expression data and the gene expression analyses were performed and provided by M.Sc. Christian Kohler, Institute of Functional Genomics, Statistical Bioinformatics, University of Regensburg as has been published in [166]. In brief, the expression levels were compared by Student´s t-test using R package multitest. The false discovery rates (FDR) were calculated according to Benjamini and Hochberg and a FDR cut-off of 0.1 was applied.

The methylation value of the *IL10* locus (3.2.3.4.2) was analyzed in mnBLL in comparison to BL and DLBCL. The mean of each CpG for each entity as well as for the respective cell lines was calculated. The significance of the differences of the methylation was calculated using the unpaired, parametric t-test ($p>0.05$ not significant).

3 Results

The aim of this thesis was the identification and characterization of new functionally relevant candidate genes in pediatric B-cell lymphoma. The focus was on the most common subtype of pediatric B-cell lymphoma: *MYC*-positive Burkitt lymphoma, and its counterpart *MYC*-negative Burkitt-like lymphoma. With respect to the diverse questions raised in each lymphoma subtype, different techniques or combinations of those were used in the analyses. Figure 10 outlines the structure of findings presented in the results section.

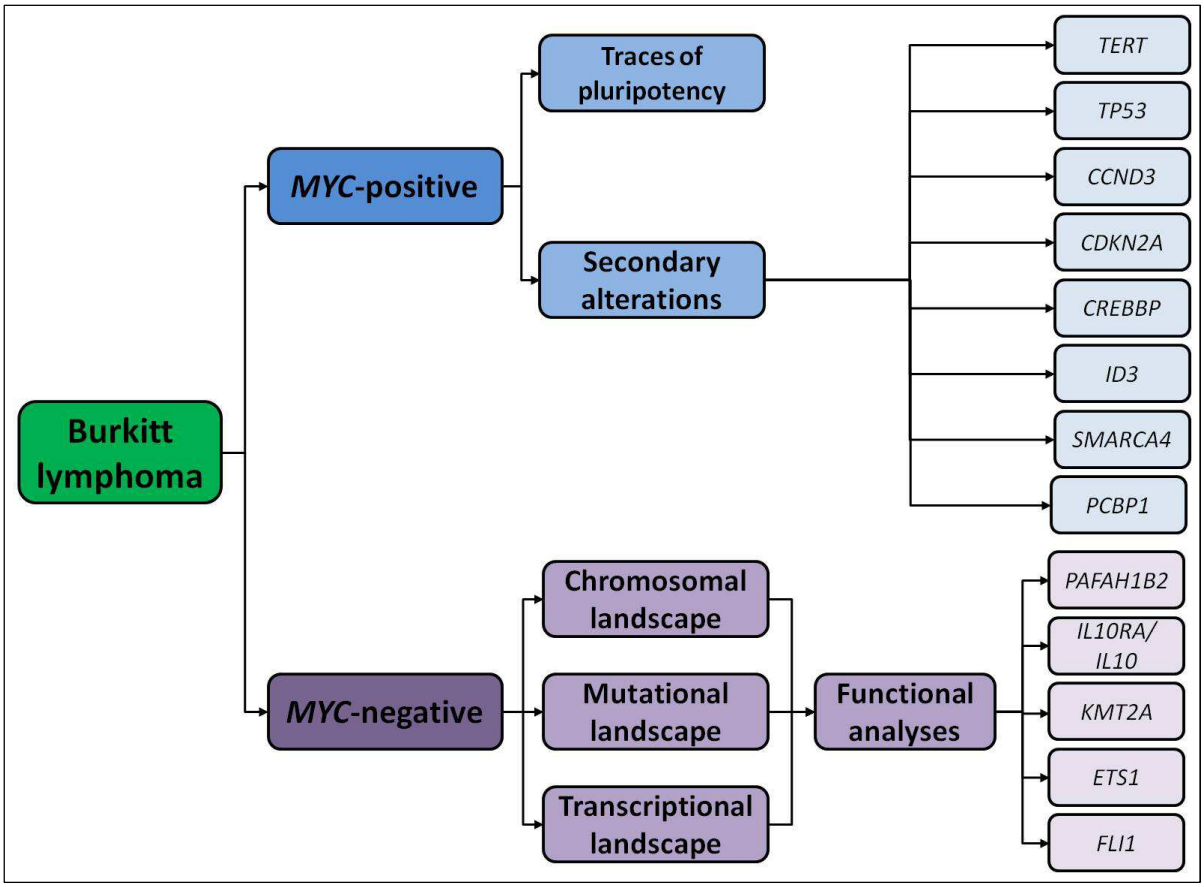


Figure 10: Overview on the herein conducted analyses of *MYC*-positive Burkitt lymphoma and *MYC*-negative Burkitt-like lymphoma.

3.1 Investigation and characterization of *MYC*-positive Burkitt lymphomas

3.1.1 Investigation of potential traces of pluripotency in Burkitt lymphomas

The hallmark of *MYC*-positive Burkitt lymphomas (BL) is the *IG-MYC* translocation leading to an overexpression of the MYC protein. Different hypotheses about the origin of the cell in which the *MYC* translocation is initiated have been proposed (1.3.3.1). One of these hypotheses suggests that MYC activation occurs within hematopoietic stem cells or cells with stem-cell like capacities which might be precursor cells of BL. This hypothesis is supported by the fact that the *MYC* gene is one of

the four reprogramming factors sufficient to transform somatic cells into inducible pluripotent stem cells. To test this hypothesis, BL were analyzed for potential traces of pluripotency in collaboration with the Centrum for Integrative Psychiatry, Kiel. Results of these analyses, which will be detailed below, are accepted for publication in the journal *Blood Cancer Journal* reflecting in a first authorship.

In order to identify traces of pluripotency in BL a recently published bioinformatic assay called PluriTest [215] was applied. Thus test is based on the analysis of the gene expression of a defined set of pluripotency signature genes. The bioinformatic procedures of this test were performed by M.Sc. Michael Lenz, Joint Research Center for Computational Biomedicine, RWTH Aachen University. Gene expression data of 221 mature aggressive B-cell lymphomas including molecular BL (mBL, n=44; *MYC*-positive n=38; *MYC*-negative n=4; non-*IG*-*MYC* positive n=1; not analyzed n=1), non-mBL (n=129) and lymphomas with a gene expression signature intermediate between mBL and non-mBL (intermediate lymphomas, n=48), published by the MMML project [108] were studied. The 177 non-mBL were classified into three groups according to their *MYC* translocation status: *MYC*-positive, non-*IG*-*MYC*-positive or *MYC*-negative. To improve the context of interpretation in the PluriTest, gene expression data of 173 primary germ cell tumors (GCT) [216],[217],[218] as well as of a large expression atlas consisting of a variety of somatic tissues [219] were also included in the analyses. Figure 11 gives an overview on the results of the analyses. The PluriTest is based on two classifiers termed pluripotency and novelty score. The pluripotency score describes to what extent a sample contains the pluripotency signature. The novelty score reports the extent of dissimilarity of the sample to the pluripotent sample models.

The B-cell lymphomas analyzed regardless of their diagnosis had a strikingly lower pluripotency score and higher novelty score than the GCT (Figure 11A). Thus, clearly showing a closer resemblance of most GCT expression profiles to human pluripotent stem cells (hPSC) than any of the analyzed B-cell lymphomas. To get a better overview on the single disease subgroups with regard to their *MYC* status, only the 44 mBL were initially compared to 115 *MYC*-negative non-mBL. As shown in Figure 11B, the mBL have a slightly higher pluripotency score and a lower novelty score than the *MYC*-negative non-mBL, leading to an effective separation of these tumor entities in two groups ($p<10^{-17}$, Mann-Whitney-U-test). Inclusion of all the other subgroups, using the mBL and *MYC*-negative non-mBL as empirical density map, showed that the *MYC*-positive intermediate and non-mBL do not have a higher pluripotency score than their *MYC*-negative counterpart but are intermingled between mBL and *MYC*-negative non-mBL (Figure 11C).

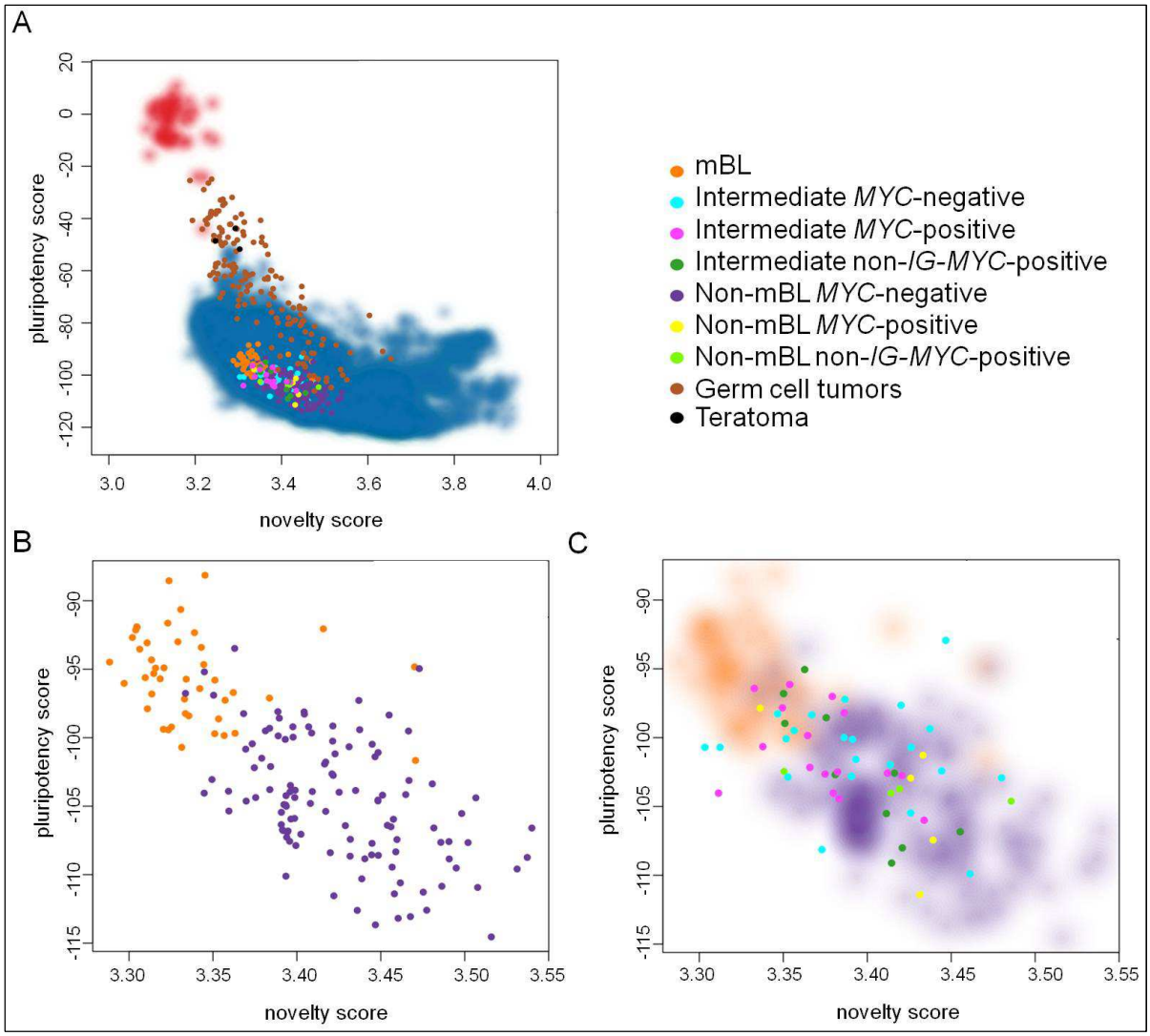


Figure 11: PluriTest analysis of *MYC*-positive and –negative B-cell lymphomas. A: PluriTest results of B-cell lymphomas in comparison to germ cell tumors and somatic tissues. The red and blue background encodes an empirical density map indicating the location of pluripotent (red) and somatic (blue) cells in the reference data set. B: Zoom into the PluriTest results, now only shown for the all mBL and *MYC*-negative non-mBL. C: The mBL and *MYC*-negative non-mBL were used as empirical density maps, into which the *MYC*-positive and *MYC*-negative intermediate B-cell lymphomas and non-mBL were integrated. Modified from Wagener *et al.*, *Blood Cancer Journal,* accepted for publication.

To determine which genes of the pluripotency signature differ between the embryonic stem cells (ESC) and the mBL, the gene expression was compared in a pluripotency plot (Figure 12). In total 181 genes (=279 oligonucleotide tags) build up the pluripotency signature [215]. The genes coding for more than one transcript can be covered by more than one oligonucleotide tag. Hence, it is possible that the transcripts of one gene might be differentially expressed. A total of 18 oligonucleotide tags (=13 genes, 6 %) were higher expressed in the mBL than in the ESC (log-fold change>1). Among those were genes related to increased cell cycle and proliferative activity like *BCL11A* [220] or *PLCG2* [221]. In contrast, 163 oligonucleotide tags (=114 genes, 58 %) were higher expressed in the ESC than in the mBL (log-fold change< -1). Among them were genes such as *NANOG, POU5F1,* and *SOX2* which are mechanistically linked to pluripotency [222]. A total of 98 tags (=81 genes, 35 %) were expressed at

comparable levels in mBL and ESC. Among those were genes which are related to cell cycle and proliferative activity as *CDC25A* [223] and *AURKB* [224].

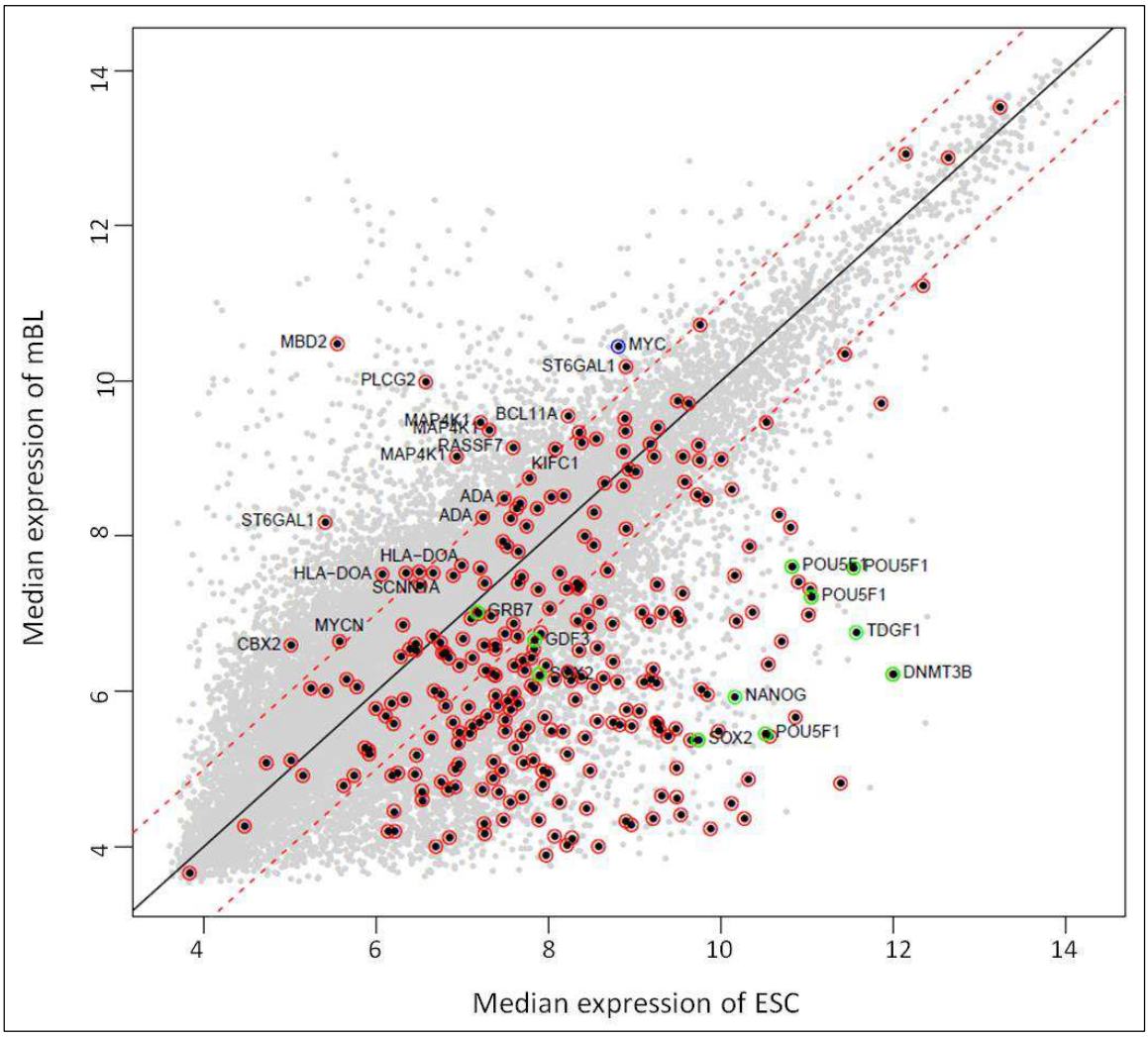


Figure 12: Pluripotency plot comparing the median gene expression of 44 mBL and 43 ESC with regard to the expression of pluripotency signature genes. The grey empirical density map is based on all genes analyzed on the respective gene expression platform. Red-circled dots mark genes of the pluripotency signature, green-circled dots mark genes mechanistically linked to pluripotency and the blue-circled dot depicts the *MYC* gene. The red dashed lines denote the log-fold change of ±1. Modified from Wagener *et al.*, *Blood Cancer Journal,* accepted for publication.

Taken together, the PluriTest algorithm showed that mBL map closer to human pluripotent stem cells (hPSC) than non-mBL and intermediate lymphomas. Nevertheless, based on data presented one can conclude that mBL do not resemble cells with pluripotent stem cell features.

3.1.2 Identification and characterization of a new exon in the *TERT* gene

As described in 1.3.4, TERT is part of the telomerase complex, which is highly active in embryonic cells and BL. A recent analysis confirmed the high TERT expression in BL using gene expression arrays, and hence, led to the inclusion of this gene in a newly defined molecular gene expression classifier for BL [108].

Within the ICGC MMML-Seq project, RNA-sequencing of primary BL was performed. Analysis of these RNA-sequencing data as well as of normal germinal center B-cells, performed by Dr. Stephan Bernhart from the Bioinformatics, LIFE Research Center for Civilization Diseases in Leipzig, again confirmed significantly higher expression of TERT in BL than in the GCB-controls (BL vs. GCB-controls log_2FC 2.21, adj. p-value 3.94^{-10}, t-test p-value corrected according to Benjamini and Hochberg). Strikingly, mining the transcriptome data of BL revealed reads within the second intron of the *TERT* gene mapping to chr5:1,286,281-1,287,202 bp (hg19). In contrast, these reads were barely present in the GCB-controls (Figure 13A). This suggests that, lying between exon 2 and exon 3, a new, so far not described exon exists. This exon, hereinafter called “new exon”, might be part of the coding sequence of an alternatively spliced transcript variant of TERT. As TERT contributes to the immortalization of BL (1.3.4), alternative splicing of *TERT* might be an event in Burkitt lymphomagenesis secondary to *MYC*-translocation which might influence TERT function. Thus, in the framework of this thesis the validation of the existence and further molecular analyses of this new exon were performed which are presented in the following.

To exclude that the reads of the new exon were misaligned repetitive elements, the region of the new exon was analyzed using the RepeatMasker track from the UCSC browser (Figure 13B). This showed that a large part of the new exon overlaps with short and long interspersed nuclear elements (SINE and LINE). Nevertheless, a BLAT search of the putative exon sequence derived from the transcriptome analysis confirmed the uniqueness of the sequence in the human genome.

The transcriptome data also revealed that the new exon has three different splice sites at the 3’-terminus leading to different exon sizes. The variants of the new exon are termed hereinafter V1-V3 with V1 being the shortest and V3 the longest variant (Figure 13B).

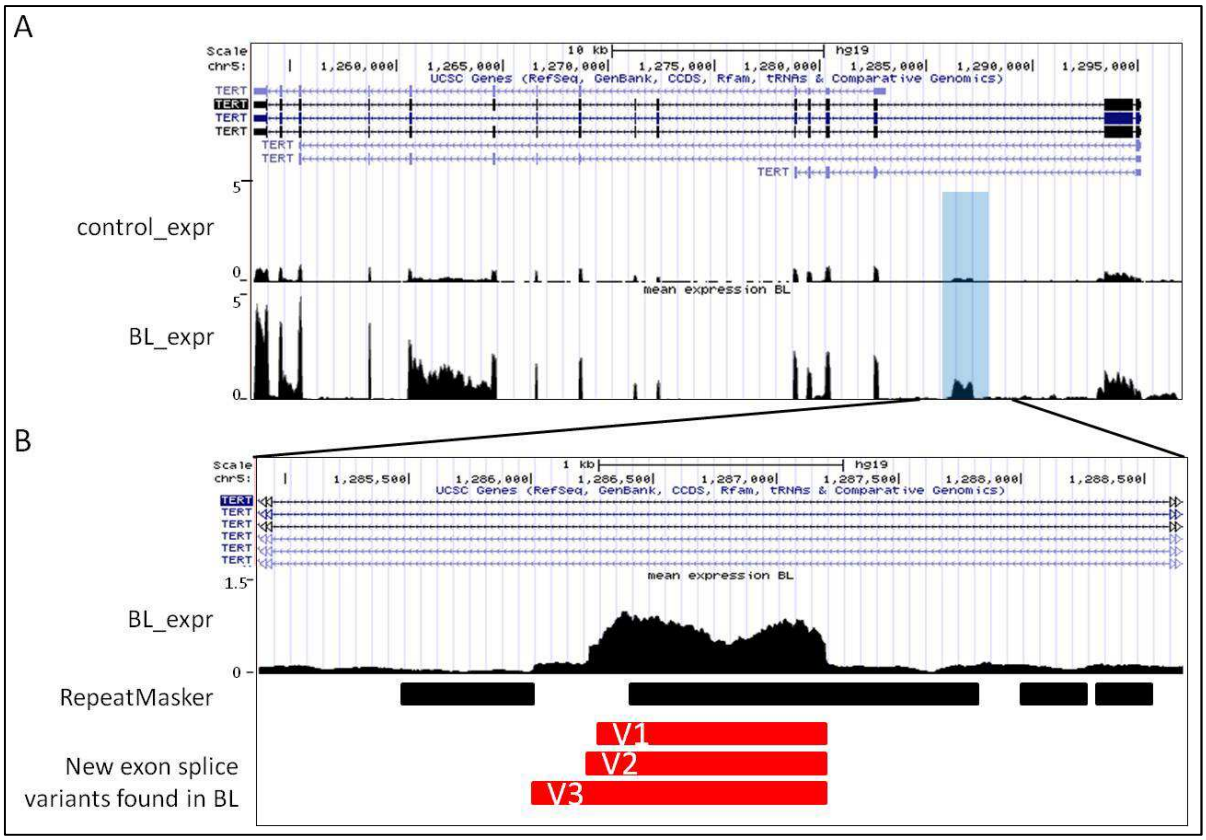


Figure 13: Differential expression of *TERT* gene on chromosome 5p15 (localized on the reverse strand). A: Depicted are from the top to the bottom: scale bar in kb, genome base position in bp, the UCSC gene track showing all transcripts of the *TERT* gene (reading direction exon 1 on the right to last exon on the left), and the mean read density obtained in the ICGC MMML-Seq project given as reads per million of the transcriptome data of all analyzed GCB-cell samples (control_expr) and BL samples (BL_expr). Highlighted in blue is the newly identified exon (chr5:1,286,281-1,287,202 bp, hg19) which is visible in the ICGC MMML-Seq samples but is not annotated yet in the UCSC gene track. B: Magnification of the region surrounding the new exon. Depicted are from the top to the bottom: scale bar in kb, genome base position in bp, the UCSC gene track showing all transcripts of the *TERT* gene (reading direction from the right (exon 1) to the left), the transcriptome track of BL (BL_expr), a track in which the black bars indicate position of repetitive elements (RepeatMasker) and a track in which the red bars depict the position and the length of the three splice variants (V1-V3) of the new exon identified in BL.

The existence of the new exon and its three variants, as well as the basal expression of TERT was analyzed in eleven BL cell lines as well as in eight primary BL on mRNA level. All BL cell lines and 7/8 primary BL expressed detectable amounts of TERT, as well as the three splice variants of the new exon (Figure 14), thus, confirming its existence.

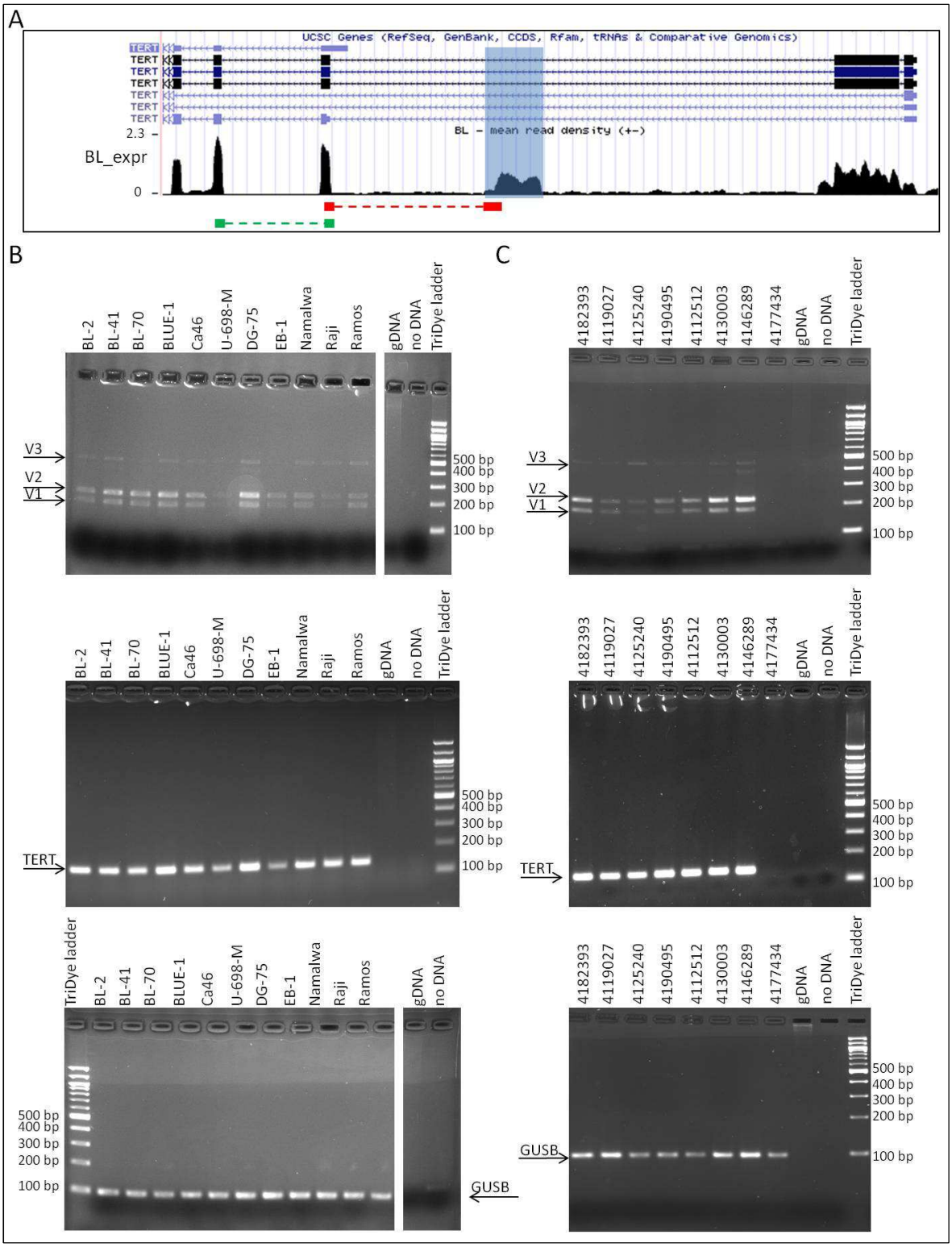


Figure 14: Basal expression of the *TERT* gene and validation of the new exon using PCR. (A) Depicted are from the top to the bottom: the UCSC gene track showing all transcripts of the *TERT* gene (reading direction from the right (exon 1) to the left), the transcriptome track of BL (BL_expr, in reads per million) in which the new exon is highlighted in blue as well as a schematic overview on the location of the PCR products with regard to the *TERT* gene. In latter track the red symbol depicts the PCR product detecting all three variants of the new exon in one PCR. The size of the PCR products were as follows: V1=181 bp, V2=222 bp and V3=447 bp. The green symbol represents the PCR product (137 bp) applied for analysis of basal TERT expression. Expression analysis was performed in eleven BL cell lines (B) and eight BL (C). Agarose gel analysis in the top panel depicts the expression of the three variants of the new exon, the middle panel shows basal TERT expression. Additionally, the integrity of the used cDNA was analyzed using the cDNA-specific primers for the reference gene *GUSB* (99 bp) (lower panel). Genomic DNA (gDNA) and water (no DNA) were used as negative controls. The arrows indicate the band corresponding to the size of the respective PCR products.

After confirmation of the expression of the new exon, its 5' and 3' exon borders were determined in the BL cell line DG-75 using Sanger sequencing (Figure 15). Based on the sequencing results, the exact coordinates of all three splice variants of the new exon were determined as follows: new exon V1 chr5:1,286,280-1,287,201 bp, new exon V2 chr5:1,286,239-1,287,201 bp and new exon V3 chr5:1,286,014-1,287,201 bp (referring to hg19). Refer to 8.6 for the sequences of the new exon variants.

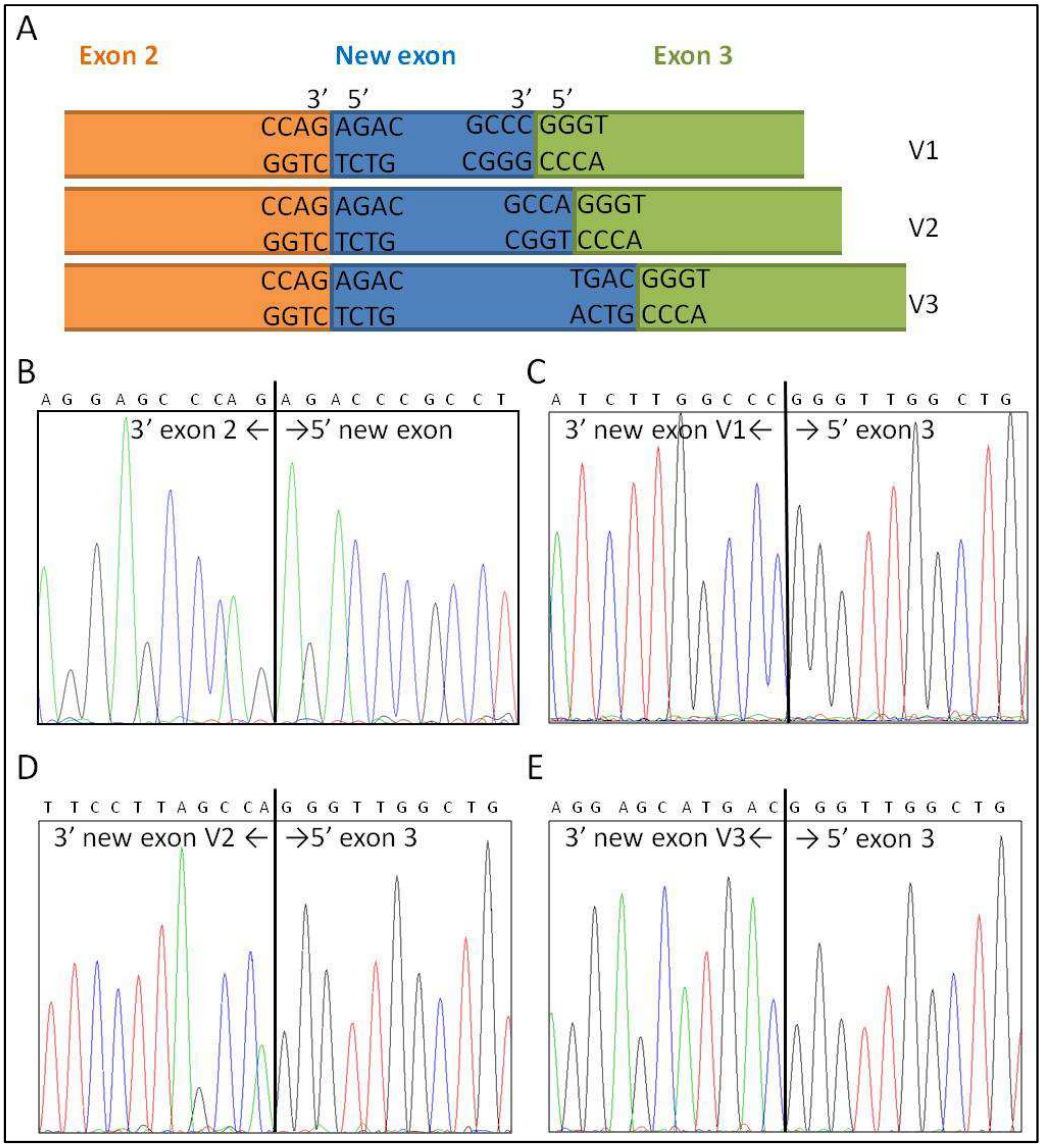


Figure 15: Validation of the 5' and 3' exon borders of the new exon of the *TERT* gene in the DG-75 cell line using Sanger sequencing. A: Schematic overview on the exon 2 - new exon and exon 3 -new exon junctions (note the inverse orientation as compared to Figure 13 and Figure 14 as focused on RNA). The letters indicate the four nucleotides adjacent to the junction as can be seen in the electropherograms of B-E. B: 3'end of exon 2 and 5'end of the new exon. C-E: 3' end of the new exon V1 (C), V2 (D) and V3 (E) and 5' end of exon 3.

In addition to the new exon, the transcriptome data of the BL from the ICGC MMML-Seq indicated the existence of a splice variant of exon 2, termed in the following exon 2V (Figure 16A). In particular, an alternative splice site seems to be used extending exon 2 into the second intron. To analyze if this

variant of exon 2 is transcribed together with the new exon PCR was performed on a panel of eleven BL cell lines. All BL cell lines expressed a TERT transcript harboring exon 2V and the new exon (Figure 16B/C). The 3'end of this exon 2 variant was analyzed using Sanger Sequencing (Figure 16D). The genomic coordinates of the exon 2V were determined as chr5:1,293,079-1,294,781 bp (hg19). Refer to 8.7 for the sequence of the splice site variant of exon 2.

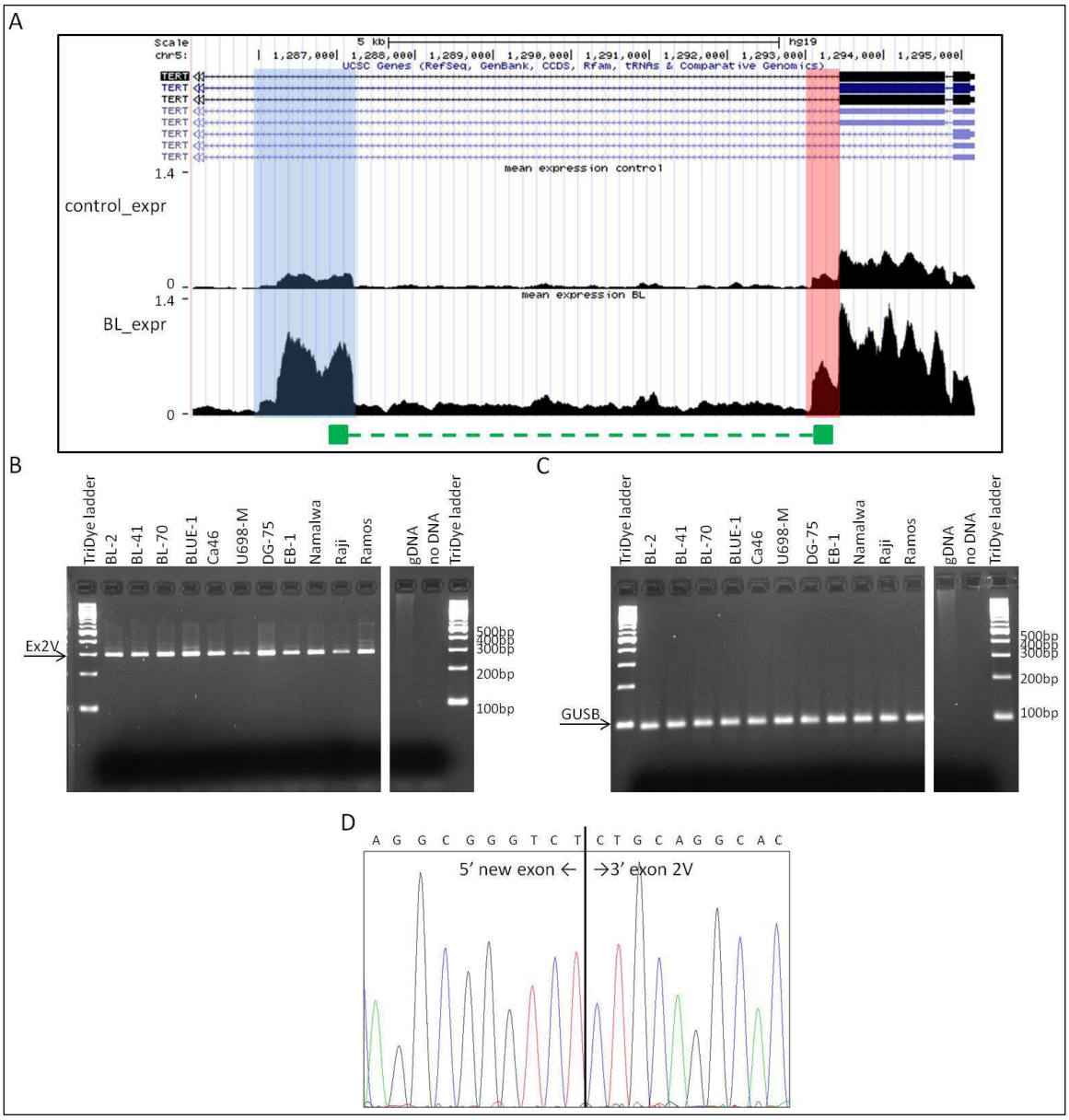


Figure 16: Validation of the expression of the variant of exon 2 together with the new exon. A: Depicted are from the top to the bottom: scale bar in kb, genome base position in bp, the UCSC gene track showing all transcripts of the *TERT* gene (reading direction from the right (exon 1) to the left), and the mean read density obtained in the ICGC MMML-Seq project given as reads per million of the transcriptome data of all analyzed GCB-cell samples (control_expr) and BL samples (BL_expr) as well as a schematic overview on the location of the PCR product with regard to the *TERT* gene. In latter track the green symbol represents the PCR product (296 bp) detecting expression of exon 2V together with the new exon. Highlighted in blue is the newly identified exon and highlighted in red is 3'end of exon 2V which is visible in the ICGC MMML-Seq samples but is not annotated yet in the UCSC gene track. B: Agarose gel analysis depicts expression of the exon 2V together with the new exon performed on eleven BL cell lines. C: Agarose gel analysis of the reference gene *GUSB* (99 bp PCR product) in the BL cell lines. D: Electropherogram showing exon borders of the 5' end of the new exon and 3' end of exon 2V. Genomic DNA (gDNA) and water (no DNA) were used as negative controls. The arrows indicate the band corresponding to the respective size of the PCR products.

As an additional exon as well as an alternative splice site might change the open reading frame (ORF) of the gene and, hence, altering the respective protein, the impact of the new exon variants on the TERT protein sequence was analyzed. Using the ExPASy translation tool the ORFs of the three 5'-3' reading frames of the coding sequence of the alternative TERT transcript were modeled (2.2.7.1). These included the alternative transcripts of the three different variants of the new exon in combination with either exon 2 or exon 2V. Table 28 summarizes for each of these six alternative TERT transcripts the predicted ORFs in the respective reading frames (frame 1-3). Interestingly, two ORFs were predicted in the first reading frame of the alternative transcripts including the new exon variants V2 and V3. Remarkably, overall only four different ORFs (ORF1-4) were identified in the six alternative transcripts as depicted in Figure 17.

Table 28: Overview on four predicted ORFs (ORF1-4) in the three frames of the respective alternative transcripts of TERT.

TERT transcripts	Frame 1	Frame 2	Frame 3
New exon V1	ORF1	ORF4	/
New exon V2	ORF1 + ORF3	/	/
New exon V3	ORF1 + ORF3	/	/
New exon V1 + ex2V	ORF2	/	ORF4
New exon V2 + ex2V	ORF2	ORF3	/
New exon V3 + ex2V	ORF2	ORF3	/

TERT transcripts: alternative TERT transcripts including the different splice variants of the new exon (new exon V1-V3) without or in combination with the alternative splice variant of exon 2 (ex2V); frame, reading frame; ORF, open reading frame.

The predicted protein sequences of all ORFs were compared to the ORF of the assumed wildtype (wt) TERT protein to identify altered regions. As mentioned above, overall four different ORFs were identified. The features of each of these ORFs as summarized in Figure 17 are presented in the following.

The first ORF (ORF1), 606 aa in length, is encoded in the first reading frame of the alternative transcripts including the new exon variants V1-V3. This ORF shares the same start codon with the wt TERT protein whereas a premature stop codon is localized within the new exon. The first 524 aa of this ORF1 match to the first 524 aa of the wt TERT protein. Overall, ORF1 lacks the C-terminal domain including the domains of the reverse transcriptase (RT), the C-terminal extension domain (CTE) as well as of the T-motif. Furthermore, only 46 aa of the TERT RNA binding domain (TRBD) domain are present.

The ORF2, encoded in the first reading frame of the transcripts including the new exon variants as well as the exon 2V, also lacks the same C-terminal domain as ORF1. Furthermore, as ORF1, ORF2 shares the start codon with wt TERT protein, whereas a premature stop codon maps within exon 2V. The first 525 aa match with the first 525 aa of the wt TERT protein. The remaining 49 aa at the C-terminal of ORF2 do not overlap with wt TERT protein.

In contrast, the ORF3 and ORF4 are predicted to lack the N-terminal domain. The start codon of ORF3 maps within the third exon of the transcript, whereas the stop codon is the same as in the wt TERT protein. The 584 aa of ORF3 match completely to the amino acids 549-1132 of the wt TERT protein. This ORF lacks the TERT essential N-terminal (TEN) domain, the QFP motif as well as the first 90 aa of the TRBD domain.

The start codon of ORF4 maps within the new exon of the transcript, whereas the stop codon is as in ORF3 the same as in the wt TERT protein. Overall 606 aa of the ORF4 match to amino acids 526-1132 of the wt TERT protein sequence whereas the first 15 aa do not match. ORF4 lacks the same N-terminal domains as ORF3.

A

Characteristics	ORF1	ORF2	ORF3	ORF4
Start codon of ORF	Same as wt TERT	Same as wt TERT	Within exon 3	Within new exon
Stop codon of ORF	Within new exon	Within Ex2V	Same as wt TERT	Same as wt TERT
Length (aa)	606	574	584	622
Overlap with TERT wt protein sequence	1-524 aa	1-525 aa	549-1132 aa	526-1132 aa
Identity with TERT wt protein sequence	86.5 %	94.5 %	100 %	97.4 %
Retained domains	**N-terminal domain**: TEN, QFP, first part of TRBD		**C-terminal domain**: last part of TRBD, T-motif, RT, CTE	

B

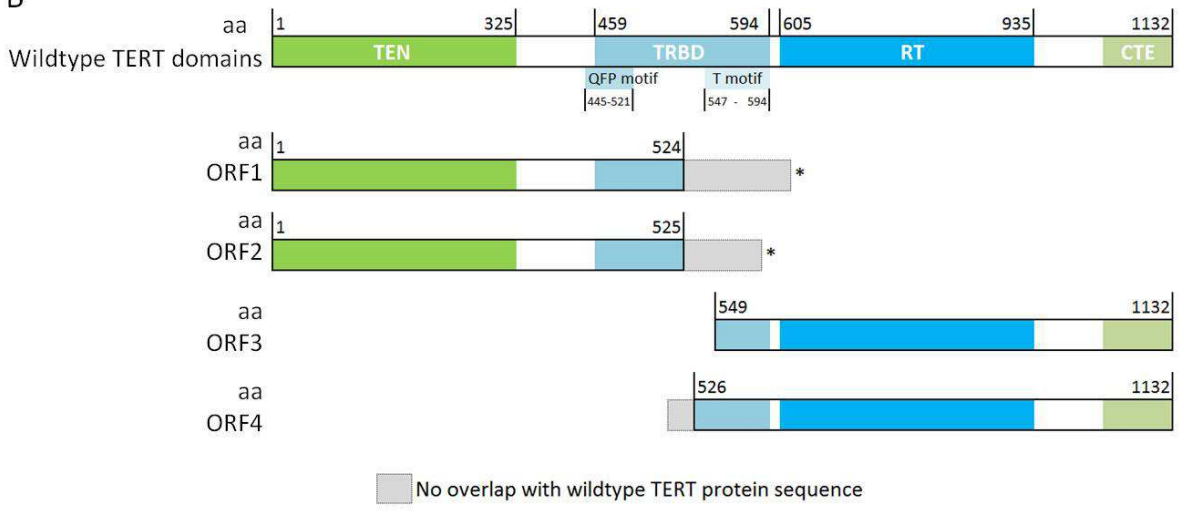


Figure 17: Overview on the four predicted ORFs (ORF1-4) of the six TERT transcripts including the three variants of the new exon and exon 2V. A: Summary of characteristics of the four ORFs: Position of the start and stop codon within the coding sequence of the ORF, length of ORF in aa, overlap with TERT wt protein sequence, overall identity to TERT wt protein sequence as well as overview on domains which are retained in the ORFs. B: On the top the domains of the wildtype TERT protein are depicted: TEN, TERT essential N-terminal; TRBD, TERT RNA binding domain; RT, reverse transcriptase and CTE, C-terminal extension. Within the TRBD two motifs are localized: QFP and T motif [225]. The lower part illustrates the four predicted ORFs (ORF1-4). The same color code in the ORFs as in the wt TERT protein indicates overlapping sequences whereas the grey colored boxes display amino acid sequences which differ from the wt TERT protein sequence. The numbers indicate the amino acids (aa). The * indicates a premature stop codon.

Taken together, analysis of the transcriptome data of the *TERT* gene in BL led to the identification of a new exon which used three different splice sites and, thus, had three different splice variants. Moreover, a new splice variant of exon 2 was detected. Thus, the novel findings presented herein led to the identification of six splice variants of the *TERT* gene. The expression of the various transcripts could be validated in BL cell lines as well as in primary BL. The ICGC MMML-Seq data show that these variants are also expressed in normal GCB-cells, though at lower levels as BL. Inclusion of the new splice variants in the reading frame of the TERT transcripts lead to a disruption of the wt ORF. Overall four different ORFs were predicted in the alternative transcripts, either lacking the N- or C-terminal domain.

3.1.3 The mutational landscape in Burkitt lymphoma

In the following, the results of the analyses of the mutational landscape in *MYC*-positive BL will be presented. This chapter is divided into (i) results derived from screening of BL cell lines for mutations within genes previously associated with lymphomagenesis of BL and/or DLBCL and (ii) results of the analyses of genes newly identified to be recurrently mutated in BL and BL cell lines.

3.1.3.1 Investigation of recurrently mutated genes in Burkitt lymphoma cell lines

Cell lines are widely used as models to study human cancers. A range of BL cell lines is worldwide available. Nevertheless, these BL cell lines differ like primary BL in their mutational landscape. Therefore, it is reasonable to characterize mutations in BL-related genes to be able to choose the appropriate cell line for further functional studies.

In the framework of this thesis, a panel of twelve cell lines (BL-2, BL-41, BL-70, BLUE-1, Ca46, DAUDI, DG-75, EB-1, Namalwa, Raji, Ramos, U-698-M) derived from children with BL was analyzed (2.1.2.1). These cell lines were screened by Sanger sequencing for mutations in genes described to be recurrently mutated in BL, i.e. *TP53* [91] and *CCND3* [133], as well as in genes described to be recurrently altered in DLBCL, i.e. *CREBBP* [226] and *CDKN2A* [227]. Figure 18 shows examples of the mutation screening using Sanger sequencing and Table 29 gives an overview on the results of this screening. The details will be presented in the following.

In 10/12 (83 %) BL cell lines overall 13 mutations were detected within the *TP53* gene. Of these, the majority were missense mutations (12/13) whereas the remaining was a nonsense mutation. All mutations were described to be probably damaging by PolyPhen. Despite this, six mutations (6/13, 46 %) were described as SNP in the dbSNP database.

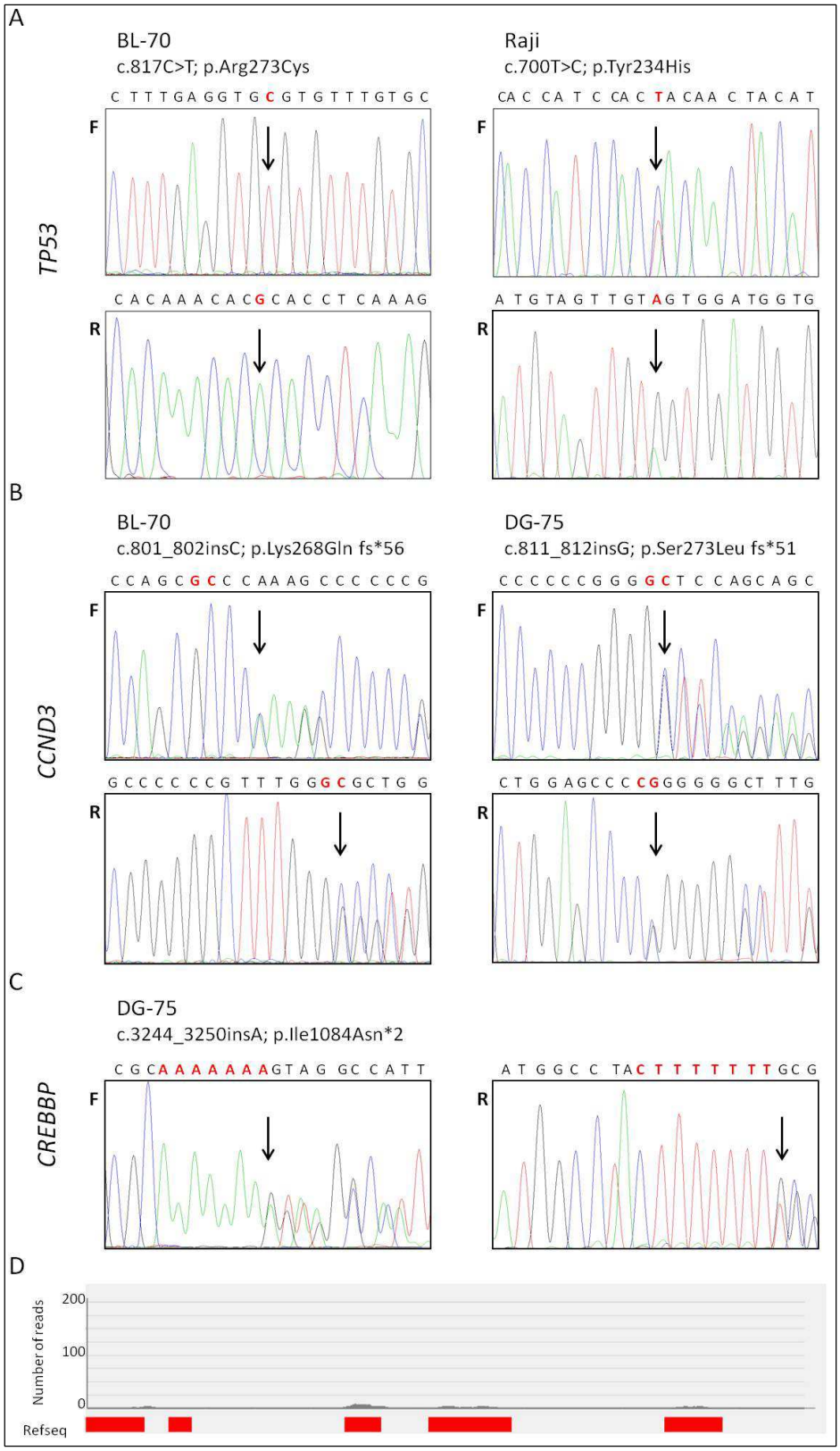


Figure 18: Overview on mutations detected in BL cell lines. A-C: Electropherograms showing the sequences around the mutational sites within the *TP53* (A), *CCND3* (B) and *CREBBP* (C) genes. The upper part of each panel shows the sequence derived from the forward primer (F) and the lower part of the reverse primer (R). The letter highlighted in red marks the nucleotide which is mutated. D: Plot showing the coverage of the *CDKN2A* gene in BL-2 as analyzed by whole-exome sequencing. The red bars of the Refseq track depict the various exons of those transcripts of *CDKN2A* with a consensus coding sequence or Refseq ID. The average exon coverage was between 0 to 5.46 reads per exon.

A total of 4/12 (33 %) BL cell lines harbored mutations within the fifth exon of the *CCND3* gene. Three of those mutations were insertions or deletions which lead to a frameshift and subsequent truncation of the protein. The remaining change was a missense mutation, which has been described as possibly damaging by PolyPhen. All mutations have not been reported as SNP in the dbSNP database.

Table 29: Overview on the mutations and alterations detected in 12 BL cell lines using Sanger sequencing.

Gene	Cell line	genomic position (hg19) in bp	mutation	consequence on protein	PolyPhen
TP53 NM_001126114 chr17p13.1	BL-41	chr17:7,577,538	c.743G>A	p.Arg248Gln[1]	damaging
	BL-70	chr17:7,577,121	c.817C>T	p.Arg273Cys[2]	damaging
	BLUE-1	chr17:7,577,120	c.818G>A	p.Arg273His[3]	damaging
	Ca46	chr17:7,577,538	c.743G>A	p.Arg248Gln[1]	damaging
	DAUDI	chr17:7,578,212	c.637C>T	p.Arg213*	n.p.
		chr17:7,577,141	c.797G>A	p.Gly266Glu	damaging
	DG-75	chr17:7,577,548	c.733G>A	p.Gly245Ser[3]	damaging
		chr17:7,577,090	c.848G>A	p.Arg283His	damaging
	Namalwa	chr17:7,577,538	c.743G>A	p.Arg248Gln[1]	damaging
	Raji	chr17:7,578,211	c.638G>A	p.Arg213Gln	damaging
		chr17:7,577,581	c.700T>C	p.Tyr234His	damaging
	Ramos	chr17:7,577,520-7,577,521	c.760_761AT>GA	p.Ile254Asp	damaging
	U-698-M	chr17:7,578,190	c.659A>G	p.Tyr220Cys	damaging
CCND3 NM_001760.4 chr6p21.1	BL-41	chr6:41,903,711-41,903,759	c.798_846del(#)	p.Pro267Leu fs*21	n.p.
	BL-70	chr6:41,903,788-41,903,789	c.801_802insC	p.Lys268Gln fs*56	n.p.
	BLUE-1	chr6:41,903,688	c.869T>G	p.Ile290Arg	possibly damaging
	DG-75	chr6:41,903,746-41,903,747	c.811_812insG	p.Ser273Leu fs*51	n.p.
CREBBP NM_004380 chr16p13.3	DG-75	chr16:3,817,721-3,817,727	c.3244_3250insA	p.Ile1084Asn*2	n.p.
CDKN2A ENSG00000147889 chr9p21.3	BL-2	chr9:21,967,753-21,995,301	homozygous deletion		n.p.

Sanger sequencing was performed on a panel of 12 BL cell lines (BL-2, BL-41, BL-70, BLUE-1, Ca46, DAUDI, DG-75, EB-1, Namalwa, Raji, Ramos, U-698-M). Genomic position (in bp) refers to the genomic build hg19; mutation, affected nucleotide in cDNA; (#), complex aberration; consequence on protein, amino acid change introduced by mutation; all mutations were revised for a SNP annotation in the dbSNP database (Build 141, 15/12/2014); six *TP53* mutations had a dbSNP annotation: [1], rs11540652, [2], rs121913343, [3], rs28934576; PolyPhen: n.p., not predictable with PolyPhen.

Only one BL cell line (DG-75) (1/12, 8 %) carried a mutation within the *CREBBP* gene. This mutation was an insertion leading to a frameshift and premature stop codon in exon 16 of the gene.

None of the cell lines carried a point mutation within the *CDKN2A* gene. Nevertheless, no PCR product was obtained for the BL-2 cell line indicating a homozygous loss or mutation of the primer binding sites in the gene. Further, analysis of this cell line by whole-exome sequencing failed to produce reads of the *CDKN2A* locus indicating a homozygous deletion of *CDKN2A* (Figure 18D). The

size of the homozygous deletion is unclear, as the exact breakpoints of the homozygous loss cannot be determined by whole-exome sequencing. But based on the read coverage of adjacent genes, the homozygous loss seems to encompass two more genes: *CDKN2B* (chr9:22,002,903-22,009,362 bp, hg19) and *DMRTA1* (chr9:22,446,840-22,452,472 bp, hg19) which lie downstream from the *CDKN2A* locus. Refer to 8.8 for an overview on the chromosomal region encompassing the homozygous loss.
Taken together, the BL cell lines showed a high frequency of mutations in *TP53* (83 %) and *CCND3* (33 %). *CDKN2A* and *CREBBP*, which are genes recurrently mutated in DLBCL, were each solely mutated in one of twelve BL cell lines (8 %).

3.1.3.2 Identification and characterization of recurrent *ID3* mutations in Burkitt lymphoma

In the framework of the ICGC MMML-Seq project, our group has recently described mutations within the *ID3 (*Inhibitor of DNA Binding 3, Dominant Negative Helix-Loop-Helix Protein) gene in 68 % (36/53) of BL [170]. The *in vitro* analyses of the functional consequences of the *ID3* mutations presented in the publication were performed in the framework of this thesis and are outlined below. This contribution led to a co-authorship in the manuscript published in the journal *Nature Genetics*.

Initially, a panel of 18 aggressive B-cell lymphoma cell lines was screened for *ID3* mutations using Sanger sequencing. This panel consisted of twelve BL- (BL-2, BL-41, BL-70, BLUE-1, Ca46, DAUDI, DG-75, EB-1, Namalwa, Raji, Ramos, U-698-M) and six non-BL (MC-116, HT, Karpas422, Su-DHL-5, Su-DHL-6, Su-DHL-10) cell lines. The results of the mutation analyses of *ID3* are summarized in Table 30.

Table 30: Mutation analyses of *ID3* in B-cell lymphoma cell lines by Sanger sequencing. Adapted from Richter *et al.*[170].

Cell line	Diagnosis	genomic position (hg19) in bp	mutation	consequence on protein	bi-allelic
BL-2	BL	chr1:23,885,618	c.300G>A (sm)	splicing	
		chr1:23,885,617	c.300+1G>C	splicing	
BL-41	BL	chr1:23,885,716	c.202G>C	p.Glu68*	Y
BL-70	BL	chr1:23,885,654-23,885,778	c.139_264del	p.Cys47Pro*32	Y
		chr1:23,885,617	c.300+1G>A	splicing	
BLUE-1	BL	chr1:23,855,678-23,885,682	c.236_240delACCTG	p.Asp79Alafs*13	
Ca46	BL	chr1:23,885,758	c.160C>G	p.Leu54Val	
		chr1:23,885,728	c.190C>T	p.Leu64Phe	
DAUDI	BL	chr1:23,885,758	c.160C>G	p.Leu54Val	
		chr1:23,885,697	c.241C>T	p.Gln81*	
Namalwa	BL	chr1:23,885,385-23,885,708	c.210_360+66del(#)	p.Ile74Val*26	
U-698-M	BL	chr1:23,885,752	c.166C>T	p.Pro56Ser	
		chr1:23,885,685	c.233T>C	p.Leu78Pro	
MC-116	B-cell lymphoma	chr1:23,885,618	c.300G>A(sm)	splicing	
		chr1:23,885,617	c.300+1G>A	splicing	

Diagnosis: BL, Burkitt lymphoma; mutation, affected nucleotide in cDNA; wt, wildtype; sm, silent mutation affecting last base of ID3 exon1; (#), complex aberration; consequence on protein, amino acid change introduced by mutation; y, biallelic status was confirmed. All consequences of the mutations on cDNA and protein level were described referring to the ID3 transcript NM_002167. The genomic positions (in bp) refer to the hg19 reference genome.

Overall, 8/12 (67 %) BL cell lines and one cell line classified as B-cell lymphoma (MC-116) harbored mutations within the *ID3* gene. The eight mutated BL cell lines carried in total 13 mutations within the *ID3* gene of which two cell lines harbored biallelic mutations. The 13 mutations in the BL cell lines were classified as missense mutations (5/13, 39 %), nonsense mutations (2/13, 15 %), frameshift deletions (3/13, 23 %) and splice site mutations (3/13, 23 %).

Remarkably, all three splice site mutations affected the end of exon 1, with the mutation c.300+1G>C occurring in two cell lines (BL-2, BL-70). Further analysis of the two *ID3* mutations in the BL-70 cells by RT-PCR and Sanger sequencing (2.2.1.7), revealed that the mutations were biallelic as two bands were detectable in the agarose gel analysis after the amplification. Both bands were analyzed by Sanger sequencing. Sequence analysis identified on the one allele the deletion but normal splicing of exon 1 on exon 2 (Figure 19A) whereas on the other allele a truncated exon 1 was spliced on exon 2 (Figure 19B). Thus, the splicing mutation led to an in-frame message with a loss of the 57 bp at the 3' end of exon 1.

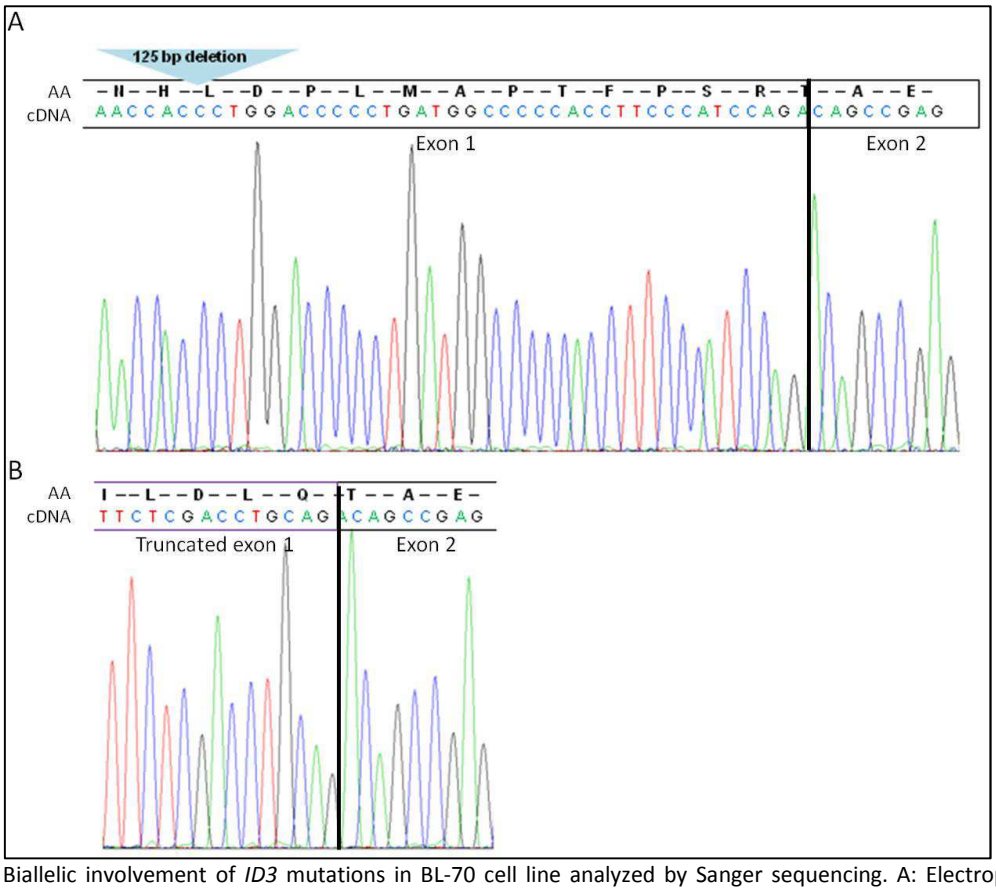


Figure 19: Biallelic involvement of *ID3* mutations in BL-70 cell line analyzed by Sanger sequencing. A: Electropherogram showing the cDNA sequence of the *ID3* allele harboring the deletion c.139_264del. The blue arrow indicates the position of the 125 bp deletion. The exon border is indicated by the black line confirming normal splicing. B: Electropherogram showing the cDNA sequence of the *ID3* allele with the splicing mutation (c.300+1G>A). This mutation leads to the usage of an alternative splice site within exon 1 and, thus, to an in-frame loss of 57 bp at the 3' end of the exon 1. Adapted from [170].

In summary, the screening of cell lines showed a predominant occurrence of *ID3* mutations in cell lines derived from BL of which two-third were affected, as well as a tendency to deleterious mutations.

To examine if the *ID3* mutations might lead to a decreased ID3 protein expression, Western blot analysis was performed using protein extracts of a panel of 14 B-cell lymphoma cell lines, consisting of ten BL- and four non-BL cell lines with known *ID3* mutation status. The results of the analysis are shown in Figure 20.

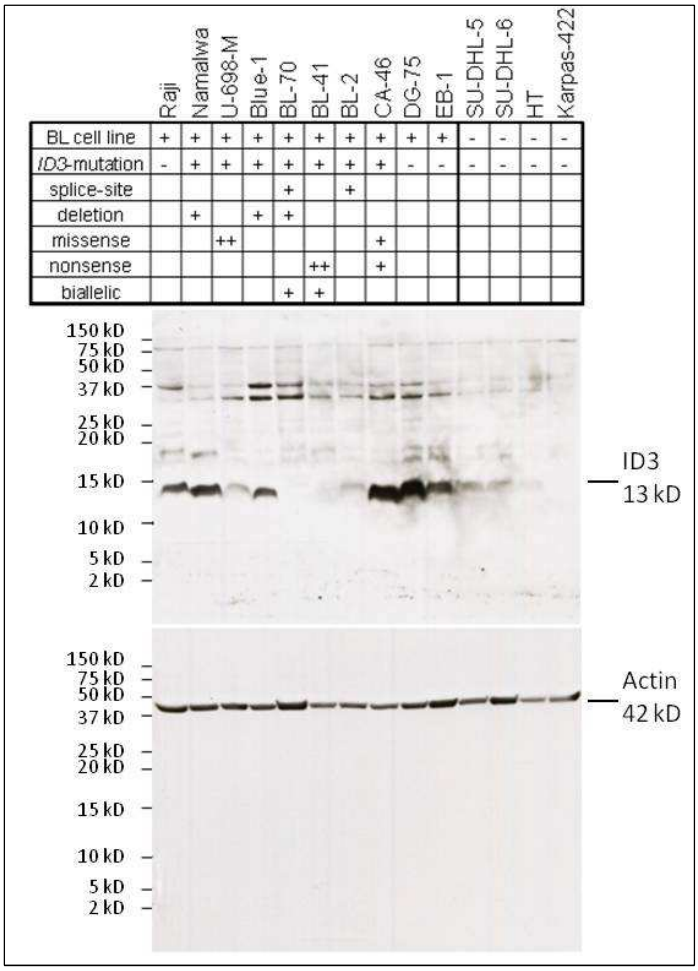


Figure 20: Western blot analyses of ID3 protein in B-cell lymphoma cell lines. The table at the top summarizes the mutation status of each cell line (+, one mutation, ++ two mutations) Depicted are the Immunoblots for the ID3 protein (13 kD) and the reference protein Actin (42 kD). Adapted from Richter *et al.* [170]. Although the ID3 antibody detected additional unspecific bands, it was determined as the best available antibody after screening of various commercially, available ID3 antibodies.

Based on the ID3 protein expression level detected by Western blot, the cell lines can be separated into different groups: cell lines with no ID3 expression (BL-2, BL-41, BL-70), cell lines (Ca-46, Namalwa) expressing ID3 at comparable levels as the cell lines with wildtype *ID3* (Raji, DG-75, EB-1) as well as cell lines with lower ID3 expression (BLUE-1, U-698-M) than the *ID3*-wildtype cell lines. The ID3 expression status correlated only partly with the *ID3* mutation status. For example, Namalwa

cells harbored a frameshift deletion but expressed ID3 at comparable levels as *ID3*-wildtype cell lines, whereas in BL-2 cells, harboring splice site mutations, no ID3 expression was detectable at all. The non-BL cell lines on the other hand showed a barely detectable or no ID3 expression at all. But it should be also taken into consideration that the detection of the expression depends on the presence of the epitope of the antibody. Thus, a here described loss of expression might be also attributable to a loss of the epitope due to an alteration of the ID3 protein sequence.

To test if a complete abolishing of ID3 function due to loss of expression or inactivating mutations might give a selective growth advantage, BL-2 cells with *ID3* mutation and complete loss of ID3 expression and DG-75 cells with wildtype-*ID3* and detectable ID3 expression were transfected with an ID3-GFP construct to re-/overexpress the ID3 protein.

The amount of ID3-GFP positive cells was determined using flow cytometry. Figure 21A and C shows that the overall ID3-GFP positivity was low for both cell lines (11.5-18 % for DG-75 and 4-5 % for BL-2 cells).

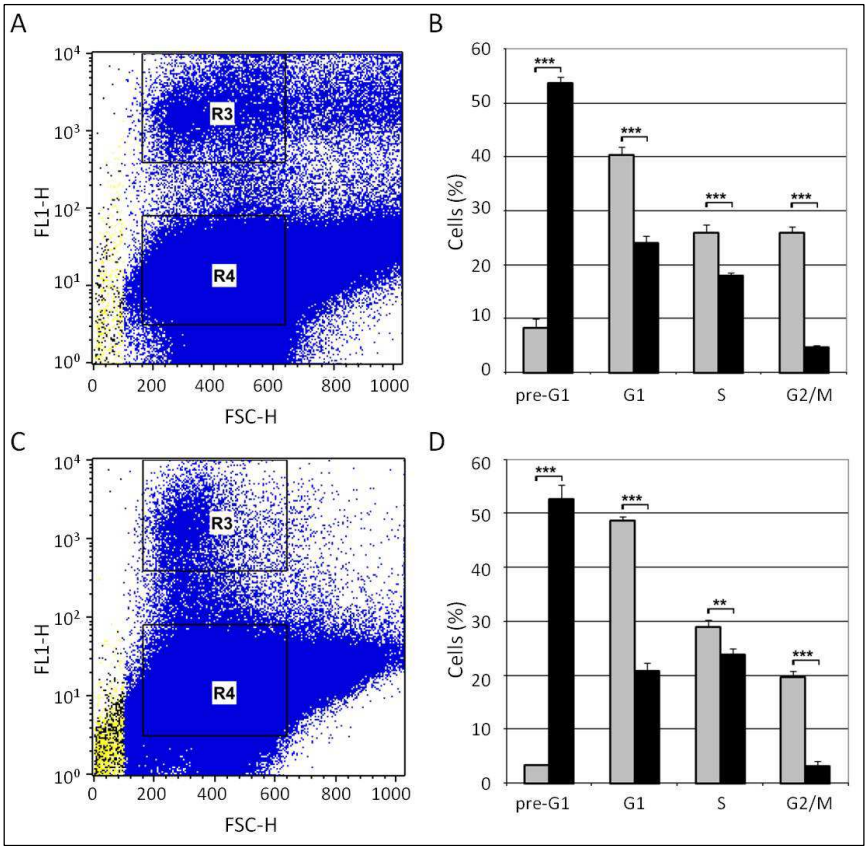


Figure 21: Cell cycle analyses of BL-cell lines transfected with an ID3-GFP construct. The *ID3*-wildtype DG-75 (A,B) and *ID3*-mutated BL-2 (C,D) cell lines were transfected with *ID3*-pCMV6-AC-*GFP*. A,C: Depict the results of the flow cytometry. The ID3-GFP-expressing (R3) and ID3-GFP-negative (R4) cells were gated into different populations. B,D: Depicts the cell cycle analyses of ID3-GFP-positive (black) and ID3-GFP-negative (grey) cell populations. The graphs depict the mean and the standard deviation. FL1-H, fluorescence channel 1 height; FSC-H, forward scatter height; G1, Gap 1 phase, S, synthesis phase, G2/M, Gap 2/Mitosis phase. Adapted from Richter *et al.* [170].

Analysis of the cell cycle (Figure 21B,D) showed that in both cell lines most of the ID3-GFP negative cell population was in G1 phase (DG-75 40.4 %, BL-2 ~48.6 %) and less than 10 % of the cells were in the apoptotic phase or pre-G1 phase. 52.6 % of BL-2 cells re-expressing the ID3 protein were in the pre-G1 phase, which is a significant increase as compared with ID3-GFP-negative cells (52.6 5 vs. 3.3 %, $p<0.001$, t-test). The other half of the ID3-GFP-positive BL-2 cells were in the G1 or S phase of the cell cycle. Only 3.2 % of those cells were in the G2/M phase which is a significant reduction in comparison to the ID3-GFP-negative BL-2 cells (3.2 % vs. 19.6 %, $p<0.001$, t-test). The majority of ID3-GFP-negative DG-75 cells were within the G1 phase (40.4 %) whereas in ID3-GFP-positive DG-75 cells 53.7 % were in the pre-G1 phase. This increase of cells in the pre-G1 phase in comparison to ID3-GFP-negative DG-75 cells is significant (53.7 % vs. 8.2 %, $p<0.001$, t-test).

Taken together, 67 % of the Burkitt lymphoma cell lines carried *ID3* mutations, from which the majority were deleterious and possibly inactivating. Some of the mutations led to a complete loss of the ID3 expression as shown by Western blot analysis. Re-expression or overexpression of ID3-GFP in BL cell lines with a mutated-*ID3* gene (BL-2) or wildtype *ID3* gene (DG-75) led to a significant increase of the number of cells in the pre-G1 phase of the cell cycle ($p<0.001$, t-test). Thus, *ID3* might be a novel tumor suppressor gene in *MYC*-positive Burkitt lymphoma.

3.1.3.3 Identification and characterization of recurrent *SMARCA4* mutations in Burkitt lymphoma

Further analysis of the mutational landscape of BL within the framework of the ICGC MMML-Seq project revealed the *SMARCA4 (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4)* gene to be mutated in 43 % of BL (9/21). The results of the identification and characterization of the *SMARCA4* mutations as presented below were embedded within a study on integrated analyses of the genome, methylome and epigenome in BL. The results of this study are currently under revision in the journal *Nature Genetics* and the data presented in this thesis lead to a co-authorship.

Initially, the focus was on the *SMARCA4* mutations identified in nine BL all analyzed within the ICGC MMML-Seq project. Based on the lack of these mutations within the respective germline, these mutations were described as somatic. The mutations were validated using Sanger sequencing as part of this thesis. After reviewing the nine mutations at the dbSNP141 and ESP6500 databases, they were confirmed as single nucleotide variants (SNV). Additionally, to elucidate a possible exclusiveness of *SMARCA4* mutation to BL, the whole-genome sequencing data of 99 non-BL of the ICGC MMML-Seq project were screened for mutations within the *SMARCA4* gene. Only 6 % of the non-BL (6/99) harbored *SMARCA4* mutations.

To identify additional cases with mutations and asses the frequency of *SMARCA4* mutations in independent cohorts, the published next generation sequencing data of BL from Schmitz *et al.* [72] and Love *et al.* [228] were mined. Schmitz *et al.* [72] have performed RNA-sequencing of 28 BL. Of those eleven BL (11/28, 39 %) were described to carry a *SMARCA4* mutation of which one mutation was described as SNP in the dbSNP141 database. Love *et al.* [228] performed whole-exome sequencing of 51 primary BL, of which eight (8/51, 16 %) carried *SMARCA4* mutations. Refer to Table 31 for an overview on the mutations.

Additionally, the whole-exome data of 13 BL cell lines available at the COSMIC database (http://cancer.sanger.ac.uk/cancergenome/projects/cell_lines/), published by Schmitz *et al.* [72] or by Love *et al.* [228] were reviewed for mutations within the *SMARCA4* gene. Using this approach, five BL cell lines (DAUDI, DG-75, EB-3, JIYOYEP and Ramos) were identified to harbor *SMARCA4* mutations. Two of the cell lines (Ramos, EB-3) harbored mutations which were reported in the dbSNP database as SNP and, hence, they were subsequently excluded. Using Sanger sequencing the *SMARCA4* mutation in DG-75 was confirmed. The *SMARCA4* mutation in DAUDI cells on the other hand could not be confirmed. The latter could be attributed to the fact that the whole-exome sequencing (data available at COSMIC cell line database, http://cancer.sanger.ac.uk/cell_lines/, 31/01/2014) and the in-house Sanger sequencing have been performed on DAUDI cells obtained from varying suppliers (American Type Culture Collection (ATCC) vs. DSMZ, respectively). Thus, DAUDI cells have likely been cultured for varying periods of time and the *SMARCA4* mutation reported in the COSMIC cell line database could represent a mutation acquired during the culturing of the cell line. As we do not possess the JIYOYEP cell line or DNA from this cell line, the validation of the mutation in this cell line was not possible.

In addition to the mining of available data, targeted resequencing using the TruSight One panel covering the *SMARCA4* gene was performed in three BL cell lines for which neither whole-exome nor whole-genome sequencing data were available (BL-70, BLUE-1, U-698-M). Only the cell line BLUE-1 carried a *SMARCA4* mutation but this was reported as SNP in the dbSNP database. Taken together, in two out of 16 BL cell lines (13 %) *SMARCA4* mutations could be identified and verified. Refer to Table 31 for a complete overview on the *SMARCA4* mutations. Figure 22A shows examples of the mutation validation using Sanger sequencing.

A total of 29 mutations were identified in *SMARCA4* of which all except one were missense mutations (28/29, 97 %). One mutation probably affects splicing (1/29, 3 %) (Table 31). All 28 missense mutations were described by PolyPhen as damaging.

Recurrent sites of substitutions were amino acids 973, 1192 and 1232 each altered in three cases, as well as the amino acid 1243 which was substituted in two cases. A total of 18 mutations affected the SNF2 family N-terminal domain (SNF2_N) and Helicase conserved C-terminal domain (Helicase_C),

(38 % and 24 %, respectively). Refer to Figure 22B for an overview on the distribution of the mutations in the *SMARCA4* encoded BRG1 protein which clearly depicts the accumulation of mutations within the SNF2_N and Helicase_C domains.

The Mechismo tool was used to analyze if the here described mutations might affect protein-protein, protein-DNA/RNA or protein-small molecule interaction. None of the 28 *SMARCA4* missense mutations was reported to alter any of the above mentioned interactions.

Table 31: Overview on *SMARCA4* (NM_001128844.1) mutations in primary BL and BL cell lines.

	PID	Genomic position (hg19) in bp	Mutation	Consequence on protein	Affected domain	Validated using Sanger sequencing	PolyPhen
ICGC cohort	4125240	chr19:11,105,679	c.1593+2T>C	Splicing	HSA	validated	n.p.
	4194218	chr19:11134,251	c.2917C>T	p.Arg973Trp	SNF2_N	validated	damaging
	4182393	chr19:11,134,252	c.2918G>A	p.Arg973Gln	SNF2_N	validated	damaging
	4194891	chr19:11,134,252	c.2918G>A	p.Arg973Gln	SNF2_N	validated	damaging
	4127766	chr19:11,134,254	c.2920C>T	p.Pro974Ser	SNF2_N	validated	damaging
	4146289	chr19:11,134,275	c.2942A>G	p.Lys981Glu	SNF2_N	validated	damaging
	4108627	chr19:11,141,498	c.3475G>A	p.Gly1159Arg	Helicase_C	validated	damaging
	4177856	chr19:11,144,113	c.3694G>A	p.Gly1232Ser		validated	damaging
	4112512	chr19:11,144,146	c.3727C>T	p.Arg1243Trp		validated	damaging
Schmitz *et al.*	SLN2448	chr19:11,094,973	c.146C>T	p.Pro49Leu		validated by [1]	damaging
	SLN2386	chr19:11,105,645	c.1561C>T	p.Arg521Trp	HSA	validated by [1]	damaging
	SLN2389	chr19:11,132,500	c.2716C>T	p.Arg906Cys	SNF2_N	validated by [1]	damaging
	SLN2402	chr19:11,132,516	c.2732G>A	p.Gly911Asp	SNF2_N	validated by [1]	damaging
	SLN2391	chr19:11,134,230	c.2896C>T	p.Arg966Trp	SNF2_N	validated by [1]	damaging
	SLN2540	chr19:11,141,508	c.3485G>T	p.Gly1162Val	Helicase_C	validated by [1]	damaging
	SLN2536	chr19:11,143,976	c.3556C>T	p.Ala1186Val	Helicase_C	validated by [1]	damaging
	SLN2528	chr19:11,143,999	c.3580G>A	p.Gly1194Arg	Helicase_C	validated by [1]	damaging
	SLN2448	chr19:11,144,113	c.3694G>A	p.Gly1232Ser		validated by [1]	damaging
	SLN2430	chr19:11,144,146	c.3727C>T	p.Arg1243Trp		validated by [1]	damaging
Love *et al.*	1060	chr19:11,132,428	c.2644G>A	p.Glu882Lys	SNF2_N	validated by [2]	damaging
	1064	chr19:11,132,513	c.2729C>T	p.Thr910Met	SNF2_N		damaging
	324	chr19:11,143,993	c.3574C>T	p.Arg1192Cys	Helicase_C		damaging
	510	chr19:11,143,993	c.3574C>T	p.Arg1192Cys	Helicase_C		damaging
	1090	chr19:11,143,993	c.3574C>T	p.Arg1192Cys	Helicase_C		damaging
	403	chr19:11,144,113	c.3694G>A	p.Gly1232Ser			damaging
	508	chr19:11,144,131	c.3712T>C	p.Ser1238Pro			possibly damaging
	1063	chr19:11,144,152	c.3733G>A	p.Ala1245Thr			possibly damaging
Cell lines	JIYOYEP[3]	chr19:11,132,426	c.2642A>G	p.Asp881Gly	SNF2_N		damaging
	DG-75[3]	chr19:11,152,098	c.4286G>A	p.Arg1429His		validated	damaging

PID: patient identification; Affected domain: HSA, Helicases and associated with SANT domains; SNF2_N, SNF2 family N-terminal domain; Helicase_C, Helicase conserved C-terminal domain; mutation, affected nucleotide in cDNA based on transcript NM_006196; consequence, amino acid change introduced by mutation The genomic positions (in bp) refer to the hg19 reference genome. If not otherwise indicated the validation using Sanger sequencing has been performed in the framework of this thesis. [1] Validation by Sanger sequencing by Schmitz *et al.* [72], [2] Validation by Sanger sequencing by Love *et al.* [228], [3] data from COSMIC Cell Line Project data base (http://cancer.sanger.ac.uk/cancergenome/projects/cell_lines/). PolyPhen: n.p., not predictable with PolyPhen.

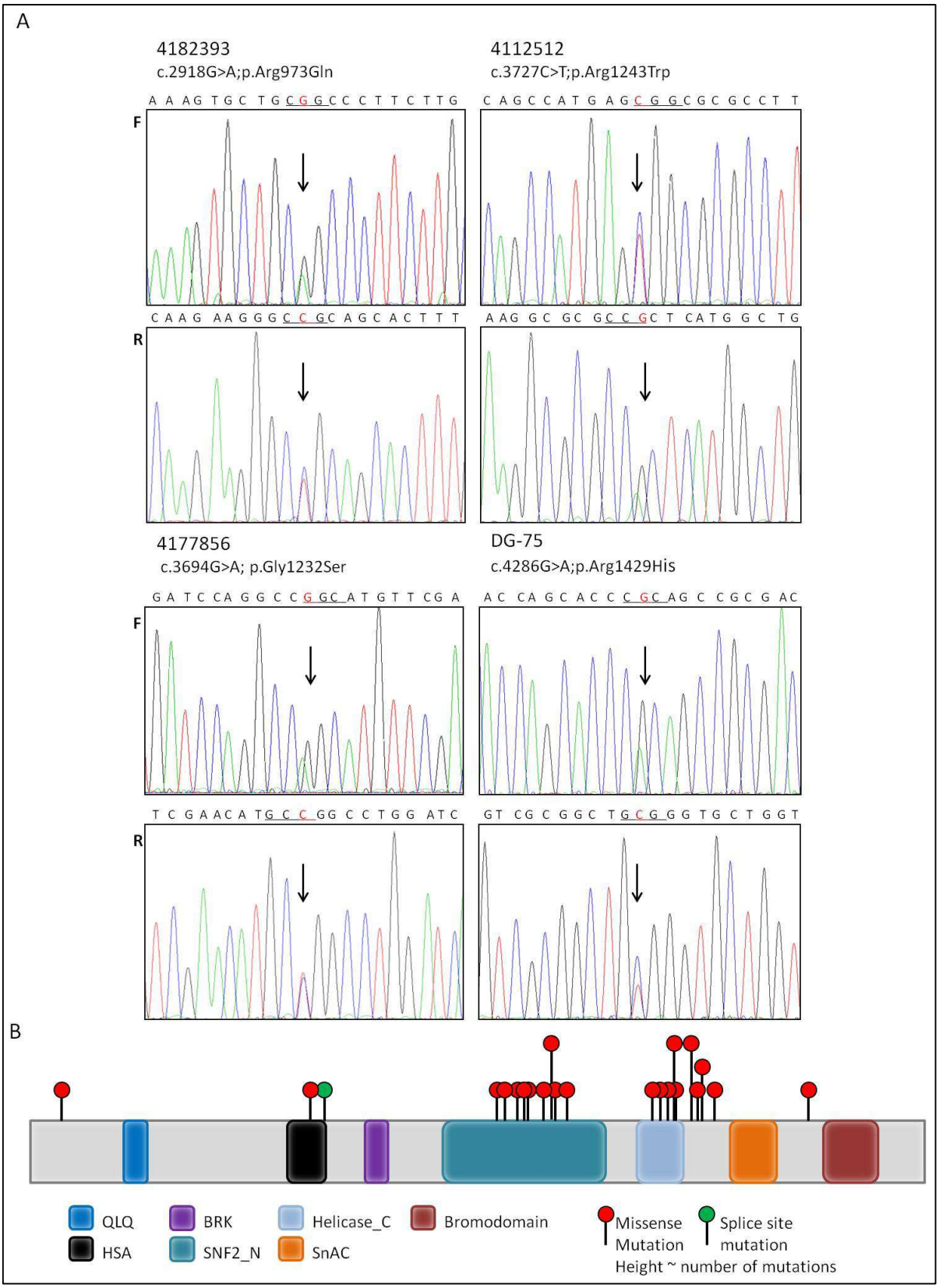


Figure 22: Overview on *SMARCA4* mutations. A: Electropherograms showing the sequences around the mutational sites within the *SMARCA4* gene in three BL and one BL cell line (DG-75). The upper panel shows the sequence derived from the forward primer (F) and the lower panel of the reverse primer (R). The site of mutation is indicated by the arrow. The letter highlighted in red marks the nucleotide which is mutated; the underlined letters mark the mutated codon. B: Overview on the mutation sites in BRG1 of all mutated BL (refer to Table 31) and BL cell lines. The protein model is based on Pfam domains.

Taken together, analyses of three independent BL-cohorts, with in total exactly 100 BL, led to the detection of *SMARCA4* mutations in 27 % of BL. The majority of these mutations were missense mutations of which 62 % affected either the SNF2_N domain or the Helicase_C domain.

3.1.3.4 Identification and characterization of recurrent *PCBP1* mutations in Burkitt lymphoma

In the framework of this thesis the *Poly(rC) binding protein 1* (*PCBP1*) gene was newly identified as recurrently mutated gene in BL. The results of the identification and characterization of the *PCBP1* mutations in BL as detailed below were summarized in a manuscript which is accepted for publication in the journal *Genes, Chromosomes and Cancer*.

Initially, analysis of whole-genome sequencing data of 17 BL from the ICGC MMML-Seq cohort revealed three BL to harbor mutations within the *PCBP1* gene (3/17, 18 %). The mutations were verified using Sanger sequencing. Moreover, by screening of the corresponding germline tissue of the patients the mutations were shown to be of somatic origin. All mutations were confirmed as SNV by comparison to the dbSNP141 and ESP6500 databases.

In order to identify additional cases with *PCBP1* mutations and to determine the frequency of mutations in this gene in an independent series, a cohort consisting of 28 *IG-MYC*-positive mBL from the MMML-network project [108] was screened using Sanger sequencing (MMML cohort). In this cohort, three BL carried *PCBP1* mutations (3/28, 11 %) which were also shown to be SNV. Thus, combining the discovery cohort of the ICGC MMML-Seq and the validation cohort of the MMML, we identified *PCBP1* mutations in 13 % (6/45) of BL.

Moreover, a panel consisting of 16 BL cell lines (BALM-16, BALM-18, BL-2, BL-30, BL-41, BL-70, BLUE-1, Ca46, DAUDI, DG-75, EB-1, EB-3, Namalwa, Raji, Ramos, U-698-M) was screened for alterations of *PCBP1*. Six BL cell lines carried mutations within *PCBP1* (6/16, 38 %) which were also shown to be SNV. Refer to Table 32 for an overview on the mutations. Figure 23A shows examples of electropherograms of identified mutations.

In total twelve mutations within the *PCBP1* gene were identified. The majority (7/12, 58 %) were missense mutations. The mutation p.Cys293Tyr occurred in two cases (MPI-017, BL-70 cell line). The other mutations were nonsense mutations (3/12, 25 %) as well as an insertion and deletion leading to a premature stop codon (2/12, 17 %). Two of the nonsense mutations affected adjacent amino acids (Gln184* and Tyr183*). The majority (6/7) of the seven missense mutations were defined by PolyPhen as possibly or even probably damaging. Only the mutation p.Val267Met was predicted to be benign.

Table 32: Overview on the mutations in *PCBP1* (NM_006196) identified in BL and BL cell lines.

	PID	Genomic position (hg19) in bp	Mutation	Consequence on protein	Affected domain	PolyPhen
ICGC cohort	4193278	chr2:70,314,913 -70,314,914	c.41_42delTC	p.Leu14fs*22	Truncating in KH I	n.p.
	4177434	chr2:70,315,476	c.601_602insT	p.Cys201Leufs*12	Truncating after KH II	n.p.
	4127766	chr2:70,315,425	c.550C>T	p.Gln184*	Truncating after KH II	n.p.
MMML cohort	MPI-017	chr2:70,315,753	c.878G>A	p.Cys293Tyr	KH III	possibly damaging
	MPI-005	chr2:70,315,759	c.884T>C	p.Ile295Thr	KH III	possibly damaging
	MPI-101	chr2:70,315,848	c.973A>G	p.Arg325Gly	KH III/NLS I	probably damaging
Cell lines	BL-30	chr2:70,315,779	c.907A>G	p.Asn303Asp	KH III	possibly damaging
	BL-41	chr2:70,315,674	c.799G>A	p.Val267Met	NLSI	benign
	BL-70	chr2:70,315,753	c.878G>A	p.Cys293Tyr	KH III	possibly damaging
	DAUDI	chr2:70,315,312	c.437T>G	p.Ile146Ser	KH II	probably damaging
	EB-3	chr2:70,315,424	c.549C>G	p.Tyr183*	Truncating after KH II	n.p.
	Namalwa	chr2:70,315,236	c.361A>T	p.Lys121*	Truncating after KH II	n.p.

PID: patient identification; mutation, affected nucleotide in cDNA based on transcript NM_006196; consequence, amino acid change introduced by mutation. Affected domains: KH, K-Homology domain; NLS, nuclear localization signal. PolyPhen: n.p., not predictable with PolyPhen. The genomic positions refer to the hg19 reference genome.

The PCBP1 protein consists of three K-homology (KH) domains (KH I-III) and two nuclear localization signals (NLS I-II). The majority of mutations affected the KH III domain (10/12, 83 %) due to a complete loss of the domain (5/10) or an amino acid substitution. Moreover, 58 % of the mutations affect the NLS I and/or NLS II due to a loss of both NLS or a missense mutation in one of them (Figure 23B). Five of the six missense mutations affected directly the KH III domain or lay next to it (p.Val267Met). The mutation p.Ile146Ser was the exception, as it lay within the KH II domain. The four mutations lying within the KH III domain were predicted by the Mechismo tool to have an effect on DNA/RNA interaction. The substitution p.Arg325Gly was predicted to have a strong disabling effect on oligonucleotide binding due to a loss of the positive charge while the other three mutations, p.Cys293Tyr, p.Ile295Thr and p.Asn303Asp, were predicted to have weaker effects.

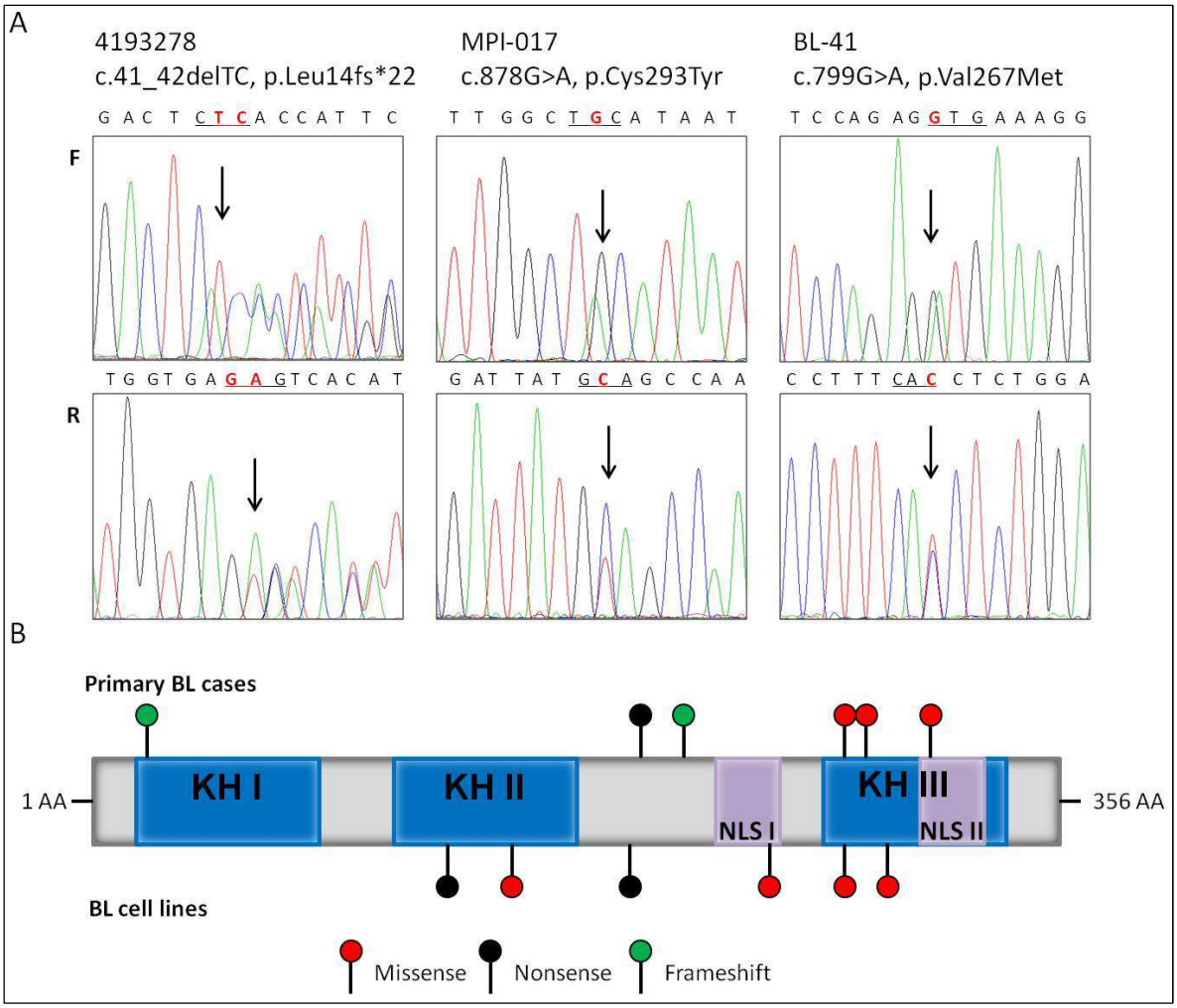


Figure 23: Overview on the *PCBP1* mutations. A: Electropherograms showing the sequence around the mutational sites within *PCBP1* gene in two primary BL cases and one BL cell line (BL-41). The site of mutation is indicated by the arrow. The upper panel shows the sequence derived from the forward primer (F) and the lower panel of the reverse primer (R). The letter highlighted in red marks the mutated nucleotide; the underlined letters mark the mutated codon. B: Overview on the mutational sites in PCBP1 with regard to its subdomains. KH, K-homology domain; NLS, nuclear localization signal. The numbers indicate the amino acid (aa). Modified from Wagener *et al.*, *Genes, Chromosomes and Cancer*, accepted for publication.

To analyze if the mutations have an impact on the PCBP1 expression, the transcriptome data available for the three cases of the ICGC MMML-Seq cohort were analyzed. This analysis revealed a balanced expression of the mutated and wildtype *PCBP1* alleles in cases 4177434 and 4127766 (relative expression of the mutated allele in case 4177434=0.47 and in case 4127766=0.55). Case 4193278 on the other hand, which carried a frameshift mutation leading to a truncation within the KH I domain, showed a more than five-fold lower expression of the mutated allele than of the wildtype allele (relative expression of the mutated allele=0.16). This points to a possible degradation of the mutated allele by nonsense mediated decay (NMD).

Analysis of the available clinico-pathological data of the BL with wildtype or mutated *PCBP1* of the ICGC MMML-Seq and MMML cohorts showed no striking differences between the two groups. The only significant difference was that the *PCBP1* mutated cases showed a higher frequency of

MUM1/IRF4 expression than the wildtype-*PCBP1* cases (6/6, 100 % vs. 12/32, 39 %, p=0.008, Fisher's exact test). Refer to 8.9 for an overview on immunohistochemical variables.

Taken together, 6 of 45 (13 %) investigated BL carried *PCBP1* mutations. The mutations predominantly affected the KH III domain of the PCBP1 protein (10/12, 83 %) either due to a complete loss or by an amino acid substitution. Four of the missense mutations might alter the RNA interaction efficiency of PCBP1. All *PCBP1* mutated cases expressed MUM1/IRF4.

3.1.3.5 Summary of the mutational landscape in *MYC*-positive Burkitt lymphoma cell lines

Figure 24 summarizes the results of the analyses of the mutational landscape in the BL cell line models. The most frequently mutated genes were *TP53* and *ID3* genes in 10/12 (83 %) and 8/12 (67 %) BL cell lines, respectively. 6/16 BL cell lines (38 %) carried *PCBP1* mutations and 4/12 (33 %) carried *CCND3* mutations. In 8 % of BL cell lines, corresponding to one BL cell line, mutations within *SMARCA4* (1/13), *CREBBP* (1/12) and *CDKN2A* (1/12) were detected. The overview on the mutational status of the respective genes shows that every cell line, besides the EB-1, BALM-16 and BALM-18, harbored at least one mutation in one of the above mentioned genes. Of note is, that only *PCBP1* mutation screening was performed for the BALM-16 and BALM-18 cell lines. Thus, the existence of mutations in the other six analyzed genes for these two cell lines cannot be excluded. Moreover, the frequency and distribution of *PCBP1* and *SMARCA4* mutations in the BL cell line models resembled that in primary BL.

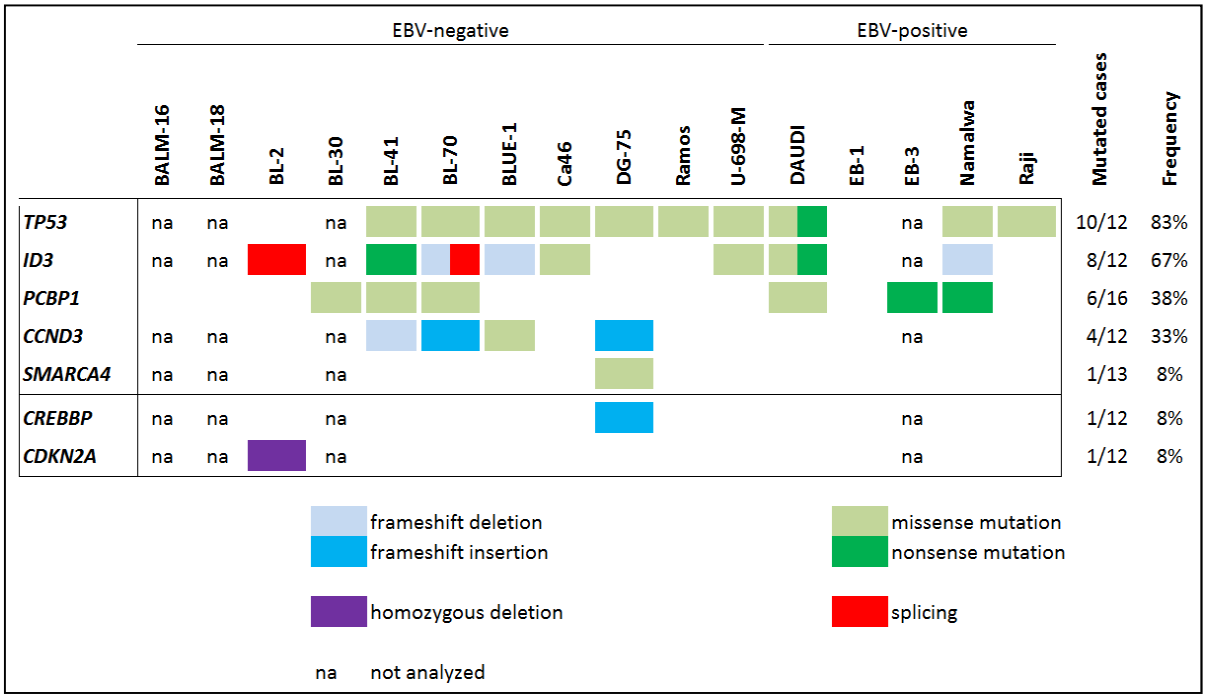


Figure 24: Overview on mutation status of the seven analyzed genes identified or verified by Sanger sequencing in overall 16 BL cell lines investigated as potential models for functional analyses. The five genes of the upper panel are frequently mutated in primary Burkitt lymphomas whereas the genes of the lower panel are more frequently mutated in primary DLBCL.

3.2 Identification and characterization of *MYC*-negative Burkitt-like lymphomas

In 2014, our group described the existence of a subtype of B-cell lymphomas termed *MYC*-negative high grade B-cell lymphoma resembling Burkitt lymphoma [166]. As already described in 1.4, these *MYC*-negative Burkitt-like lymphomas (mnBLL) were characterized by a common pattern of aberration on chromosome 11q consisting of interstitial gains and telomeric losses.

The aim of the analyses of mnBLL presented in the following was the characterization of the chromosomal, mutational and transcriptional landscape. These analyses should pave the way for the identification of genes which might contribute to the pathogenesis of these lymphomas. Moreover, cell line models for mnBLL were identified and characterized. These models were subsequently used for functional analyses of the potential candidate genes involved in the pathogenesis of this lymphoma subtype.

Part of the characterization of mnBLL was performed in the framework of this thesis leading to a shared first authorship in the above mentioned work published in the journal *Blood* [166]. The data presented below are extended follow-up analyses of this publication taking into account the data published in *Blood* [166] as well as data acquired afterwards.

3.2.1 Characterization of newly recruited *MYC*-negative Burkitt-like lymphomas

In addition to the 17 mnBLL already published in *Blood*, eight new mnBLL (cases 18-25) were identified in the routine tumor genetic diagnostics of the Institute of Human Genetics, Kiel, Germany or were referred to by other institutions for reference diagnostics. These cases were recognized by the referring pathologist to harbor features of BL. Furthermore, they were identified based on the cytogenetic analyses showing that they did not carry a break within the *MYC* gene, but harbored the gain and loss pattern on chromosome 11. All newly recruited cases were characterized as mnBLL based on the following published criteria:

Cytogenetic analysis:	Immunophenotype:
Presence of 11q gain/loss	Coexpression of CD20 and CD10
Lack of *MYC* break	Lack of BCL2 expression
Lack of *BCL2* break	>90 % Ki67 positivity

The information concerning the cytogenetic findings and immunophenotype were obtained from the clinical records and/or referring institutions. The results are summarized in Table 33. The *BCL2* break status was missing for three cases. Nevertheless, as the *BCL2* translocation is associated with an overexpression, the BCL2-negative immunophenotype of those cases suggested for a lack of a *BCL2* break.

Table 33: Cytogenetic features and immunophenotype of mnBLL cases investigated in this thesis but not reported in [166].

Case	Cytogenetics				Immunophenotype			
	11q gain/loss	*IGH-MYC* fusion	*MYC* break	*BCL2* break	CD20	CD10	BCL2	Ki67
Case 18	positive	n.d.	negative	n.d.	+	+	-	~ 100 %
Case 19	positive	negative	negative	negative	+	+	-/+	> 95 %
Case 20	positive	negative	negative	negative	+/-	+	-	> 95 %
Case 21	positive	negative	negative	n.d.	+	+	-	90 %
Case 22	positive	negative	negative	n.d.	n.d.	+	-	> 95 %
Case 23	positive	negative	negative	negative	n.d.	+	+/-	95 %
Case 24	positive	negative	negative	n.d.	+	+	-	100 %
Case 25	positive	negative	negative	negative	+	+	-	100 %

+, positive; -, negative; +/-, majority positive with negative fractions; -/+, majority negative with positive fractions; n.d., no data available.

After confirmation that the newly recruited cases harbor the typical cytogenetic and immunophenotypical features of the mnBLL, array-based analyses using the OncoScan DNA-microarray were performed to further characterize the alteration on chromosome 11 as well as the global chromosomal landscape. The results of these analyses are presented in the following.

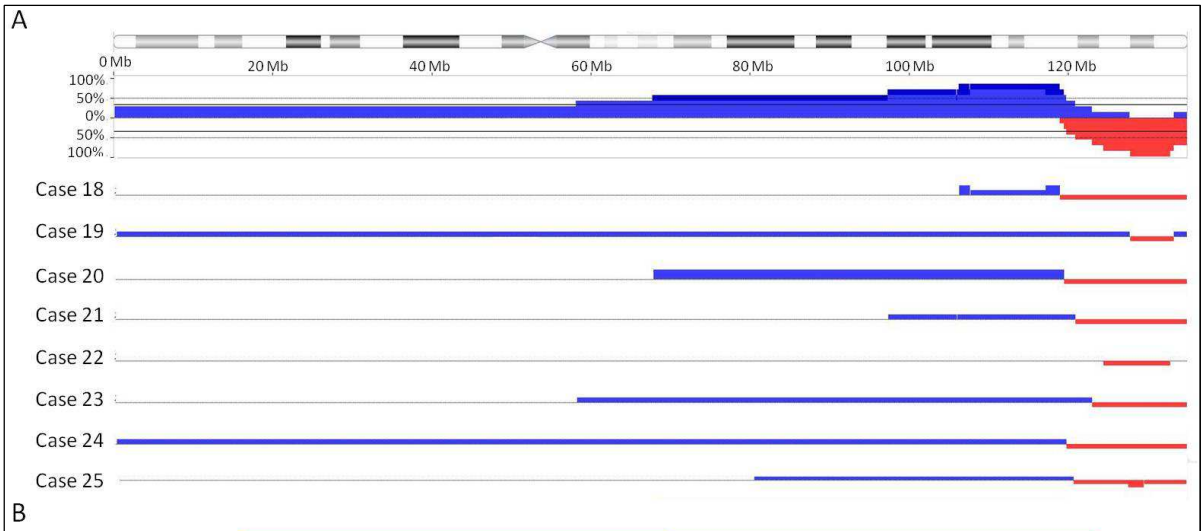
3.2.1.1 Array-based characterization of chromosome 11 aberrations in *MYC*-negative Burkitt-like lymphoma

In the following, the alterations on chromosome 11 observed in the eight novel mnBLL are described. Findings are summarized in Figure 25. For a complete list of all copy number alterations (CNA) in these eight mnBLL cases refer to appendix 8.10.

Seven of the eight mnBLL (cases 18-21 and 23-25) carried the typical copy number pattern of gain and loss on chromosome 11 (Figure 25A). The exception was case 22 which solely harbored a loss of 11q24.2-q25. Interestingly, cases 19 and 24 showed in contrast to all the other previously described mnBLL cases a complete gain of chromosome 11 with a telomeric loss in 11q25. Cases 18 and 20 carried in addition to the interstitial gains in 11q22.3-q23.3 a focal high-level amplification in 11q23.3.

Based on the breakpoints in chromosome 11 the minimal regions of gain and loss in the eight analyzed mnBLL cases were defined (Figure 25B). The minimal region of gain in 11q23.3 was defined as chr11:106,233,094-118,966,521 bp and the minimal region of amplification as chr11:117,127,159-118,966,521 bp. Telomeric losses of 11q affecting 11q24-qter were present in all eight samples. The minimal region of loss was defined as chr11: 127,799,447-132,877,670 bp. Moreover case 25 harbored a homozygous loss in 11q24.2-q24.3, spanning 1.95 Mb (chr11:127,621,020-129,572,876 bp).

A



B

	Copy number gain chr11 (in bp)		Copy number loss chr11 (in bp)	
PID	Start	End	Start	End
Case 18	**106,233,094**	**118,966,521**	118,971,251	134,938,847
Case 19	192,764	127,778,007	**127,799,447**	133,280,976
Case 20	67,767,160	119,460,505	119,488,812	134,938,847
Case 21	97,346,172	120,907,482	120,931,054	134,938,847
Case 22			124,440,617	**132,877,670**
Case 23	58,169,996	122,984,689	122,996,050	134,938,847
Case 24	192,764	119,754,895	119,771,843	134,938,847
Case 25	80,282,364	120,687,234	120,702,378	134,938,847

Figure 25: Chromosomal view of chromosome 11 analyzed by OncoScan DNA-microarray in eight newly recruited mnBLL. A: Depicted from top to the bottom: Ideogram of chromosome 11, diagram showing the cumulative percentage of cases sharing the copy number alteration and chromosome 11 profiles of cases 18-25. Blue color represents copy number gains, red color represents copy number losses. The height of the bars refers to the copy number state. B: Breakpoints of the alterations in chromosome 11. Highlighted breakpoints mark the minimal region of gain (blue) and loss (red). Start and end refer to the genomic position according to hg19 in bp.

Next, the minimal regions for the newly recruited mnBLL cases were compared to those of the published cases. Refer to Table 34 for an overview on the new and published minimal regions.

The minimal region of gain, defined in the newly recruited mnBLL, was 9 Mb larger than the already published minimal region. Therefore, the data from the newly recruited cases did not reduce the size of the already published minimal region of gain. In some of the published as well as in two of the newly identified mnBLL amplifications within the minimal gain region were detected. Thus, a minimal region of amplification was defined. The newly defined minimal region of amplification was ~1 Mb larger and the breakpoints lay outside of the already published minimal region. Thus, the minimal region of amplification did not alter either.

In contrast, on the basis of the results of the newly recruited mnBLL the size of the minimal region of loss could be reduced to chr11:127,799,447-132,877,670 bp. Thus, the size of the minimal region of loss was reduced by approximately 2.4 Mb now being ~ 5.1 Mb in size.

Intriguingly, in addition to the previously described case 9, a second mnBLL case (case 25) also harbored an homozygous loss in 11q. Based on those two cases the minimal homozygous loss region was defined as chr11:127,816,801-129,341,359 bp which is part of the minimal region of loss.

Table 34: Overview on minimal gain and loss regions in chromosome 11q.

Chromosome 11	Published breakpoints (bp)	New breakpoints (bp)	Final breakpoints (bp)
Minimal gain region	115,025,608-118,434,149	106,233,094-118,966,521	115,025,608-118,434,149
Minimal amplification region	117,602,151-118,434,149	117,127,159-118,966,521	117,602,151-118,434,149
Minimal loss region	127,471,805-134,940,727	127,799,447-132,877,670	127,799,447-132,877,670
Minimal homozygous loss region	127,816,801-129,341,359	127,621,020-129,572,876	127,816,801-129,341,359

Published breakpoints: Minimal regions already published by Salaverria *et al.* [166]; New breakpoints: Minimal regions defined based on the eight newly recruited mnBLL cases; Final breakpoints: Newly defined minimal regions based on the new and published regions. All genomic positions in bp refer to the hg19 reference genome. The underlined genomic positions define breakpoints of the final minimal regions.

Taken together, extension of the published cohort of 17 to a total of 25 mnBLL led to a re-definition of the minimal region of loss, which is now defined as chr11:127,799,447-132,877,670 bp (11q24.2-q25) which includes a minimal homozygous loss region at chr11:127,816,801-129,341,359 bp (11q24.3). The already described minimal regions of gain and amplification did not change.

3.2.1.2 Genome-wide array-based characterization of chromosomal imbalances in *MYC*-negative Burkitt-like lymphoma

After characterization of the aberration in chromosome 11, the OncoScan DNA-microarray data of the 25 mnBLL were mined for concomitant and recurrent CNA. The median number of CNA in the recently acquired cases (cases 18-25) was 6. The 17 cases already published have been analyzed with different platforms, thus, the median number of CNA differed. Cases 1-3 and 8 analyzed with SNP array 6.0 harbored a median of 18.5 CNA. Cases 4-7 and 9-12, whose DNA was derived from FFPE tumor material, were analyzed with array-CGH and harbored a median of 4 CNA. The MPI cases analyzed with array-CGH harbored a median of 11.5 CNA. Thus, the median number of CNA in the new cases lay in the range of described of a median 4-18.5 CNA.

Concomitant to the chromosome 11 alterations, the most abundant copy number alterations in all 25 mnBLL cases, defined as imbalances detected in at least five cases, were a trisomy 12, as well as losses in 6q14.3-q21 and gains in 13q31.3 and 18q21.2 (Table 35).

Table 35: Overview on concomitant CNA occurring in more than five of the in total 25 mnBLL cases.

	PID	Breakpoints in bp (hg19)	Minimal region of imbalances in bp
Copy number loss 6q14.3-q21	Case 1	chr6:74,937,191 - 118,113,917	**chr6:86,450,066 - 110,118,776**
	Case 3	chr6:**86,450,066** - 145,479,373	
	Case 22	chr6:67,759,432 - **110,118,776**	
	MPI-078	chr6:80,425,984 - 110,631,941	
	MPI-315	chr6:46,810,436 - 170,958,026	
	MPI-382	chr6:80,425,984 - 170,957,975	
Copy number gain 13q31.3	Case 1	chr13:77,877,525 - 106,193,991	**chr13:91,967,744 - 92,150,929**
	Case 2	chr13:**91,967,744** – 92,361,860	
	Case 18	chr13:91,359,736 - 92,328,701	
	Case 20	chr13:89,614,856 - **92,150,929**	
	MPI-382	chr13:51,430,518 - 114,914,704	
Copy number gain 18q21.2	Case 1	chr18:40,819,075 - 62,093,468	**chr18:51,129,489 - 53,141,537**
	Case 2	chr18:51,050,447 – **53,141,537**	
	Case 3	chr18:45,252,881 - 77,920,288	
	Case 20	chr18:**51,129,489** – 74,637,067	
	Case 24	chr18:18,554,307 – 53,950,927	
Trisomy 12	Case 5	Whole chromosome	
	Case 7		
	Case 18		
	Case 21		
	Case 23		
	Case 24		

PID: patient identification. All genomic positions in bp refer to the hg19 reference genome. Highlighted in bold letters are the breakpoints which define the minimal region of loss or gain.

The most frequent alteration was a copy number loss in 6q14.3-q21 which was identified in six out of 25 cases (24 %) (cases 1, 3, 22 and MPI-078, MPI-315, MPI-382). The minimal loss region was chr6:86,450,066-110,118,776 bp. Two tumor suppressor genes in lymphoma map within this 24 Mb large region of copy loss: *EPH Receptor A7* (*EPHA7)* [229] and *PR domain containing 1, with ZNF domain* (*PRDM1)* [230]. A total of 24 % (6/25) of the mnBLL harbored a complete trisomy 12 (cases 5, 7, 18, 21, 23 and 24). A copy number gain in 13q31.3 was identified in five out of 25 cases (20 %)(cases 1, 2, 18, 20 and MPI-382). Within the 0.2 Mb minimal region of gain in 13q31.31 (chr13:91,967,744-92,150,929 bp) maps the miR-17-92 cluster which has been for example described to be gained and overexpressed in pediatric BL [231]. Additionally, a recurrent copy number gain in 18q21.2 was detected which affected five out of 25 cases (20 %) (cases 1, 2, 3, 20 and 24). Within the minimal region of gain, which encompassed ~ 2 Mb (chr18:51,129,489-53,141,537 bp), map seven

genes including the *Transcription Factor 4* (*TCF4)*. This gene belongs, as its homologue *TCF3,* to the bHLH family of transcription factor and plays a major role in the Wnt/β-catenin signaling pathway in various cancer types [232],[233].

3.2.2 Cell line models for the functional study of *MYC*-negative Burkitt-like lymphoma

Cell line models are crucial tools for functional studies. Thus, to perform functional studies for mnBLL, such cell line models were indispensable. The two cell lines HT and Su-DHL-5 have already been identified as mnBLL cell line models [166]. In the following, the identification and characterization of new possible cell line models for mnBLL with the typical 11q alteration is described.

3.2.2.1 Identification and characterization of cell line models for *MYC*-negative Burkitt-like lymphoma

To identify further cell line models the available Affymetrix SNP array 6.0 data at the COSMIC database (http://cancer.sanger.ac.uk/cell_lines/, 11/09/2012) were mined for B-cell lymphoma cell lines harboring the peculiar alteration pattern on chromosome 11 characteristic for mnBLL. Using this approach, the two cell lines MLMA and Karpas422 were identified to carry a mnBLL-typical gain and loss pattern on chromosome 11.

Karpas422 has been described to be a germinal-center B-cell-like DLBCL cell line, harboring a translocation t(14;18) juxtaposing the *BCL2* gene to the *IGH* gene resulting in a BCL2 overexpression. In line, this cell line has been described to have a high BCL2 protein expression [234]. But as described in 3.2.1 the criteria for the diagnosis of mnBLL is a lack of a *BCL2* break as well as a lack of BCL2 expression. Thus, Karpas422 is not an appropriate cell line model for mnBLL.

The MLMA cell line has been described to have been established from a patient with human malignant lymphoma being IgD, IgM positive with hairy B-cells (http://cellbank.nibio.go.jp/english/, 15/12/2014). This differs from the diagnosis of mnBLL which have mostly been diagnosed as high grade B-cell lymphoma or atypical BL (2.1.1.2). To elucidate if the MLMA cell line has been established from hairy cell leukemia or from another B-cell malignancy and if it is indeed an appropriate mnBLL cell line model, the cells were characterized at our Institute by routine cytogenetic and molecular analyses. A hallmark of hairy cell leukemia is the mutation p.Val600Glu within the *BRAF* gene, which can be detected in nearly all typical cases [235]. MLMA cells did not carry this typical p.Val600Glu mutation. Thus, indicating that MLMA cell line has not been established from classical hairy cell leukemia. Immunophenotyping of MLMA cells was performed by Dr. Sebastian Böttcher at the II. Medical Clinic, Kiel. The immunophenotyp showed coexpression of

CD20, CD10, CD38, CD45 and CD43, whereas CD5, CD103 and CD138 were not expressed. Especially the lack of CD103 expression distinguished the cell line from the typical hairy cell leukemia [1]. Instead, the high coexpression of the other markers was indicative for a classification as mature B-cell non-Hodgkin-lymphoma. The CD10, CD20 and CD38 expression is well in line with a typical BL phenotype as also most mnBLL show.

Furthermore, cytogenetic analyses performed in routine tumor genetic diagnostics of the Institute of Human Genetics, Kiel revealed that the MLMA cell line did not harbor a *MYC* or a *BCL2* break which is in line with the established mnBLL criteria (3.2.1).

Our own SNP array 6.0 analyses of MLMA cells revealed extensive and complex rearrangements (overall 77 CNA). Refer to appendix 8.11 for a complete overview on the CNA. In agreement with the data in the COSMIC database, MLMA cells were shown to carry the typical gain and loss pattern on the long arm of chromosome 11 (Figure 26).

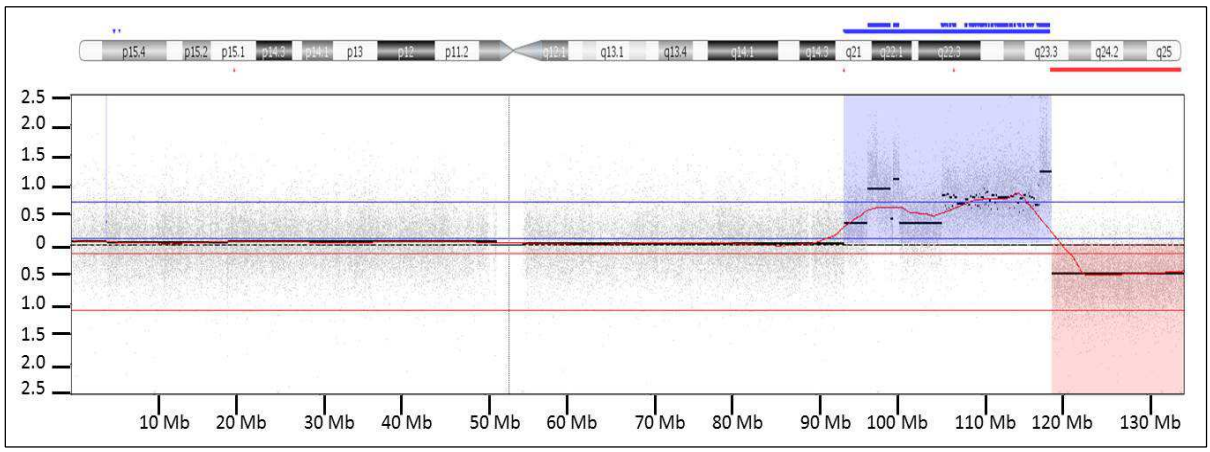


Figure 26: Overview on chromosome 11 analyzed by SNP array 6.0 in the MLMA cell line. The y-axis depicts the log2 ratio copy number and the x-axis the chromosomal positions in Mb. The blue highlighted region marks the copy number gain region, the red region marks the copy number loss regions.

The gain region covered chr11:93,756,820-118,962,816 bp (11q21-q23.3) which included three high level copy number gain regions 11q21.2-q22.1, 11q22.1 and 11q23.3. The gain region in MLMA cell line is 4-fold larger than the previously defined minimal region of gain, spanning 12 Mb. The terminal loss in MLMA cells spans chr11:118,970,202-134,944,770 bp (11q23.3-q25) which covers completely the minimal loss region (11q24.2-q25) defined for the primary mnBLL.

Since the analyses indicate that MLMA cell line was established from a mature B-cell non-Hodgkin-lymphoma and that it carries the typical chromosomal alteration on chromosome 11, it qualified as a novel mnBLL cell line model.

3.2.3 Identification and characterization of possible candidate genes contributing to the development of *MYC*-negative Burkitt-like lymphomas

After extensive chromosomal characterization of the primary mnBLL as well as the corresponding cell line models, the next aim was to identify possible candidate genes and pathogenetic mechanisms which might contribute or lead to the development of mnBLL. In the following the identification of possible candidate genes by global expression profiling, mutation analyses and text mining as well as subsequent functional characterization in the cell line models is described.

3.2.3.1 Identification of possible candidate genes on chromosome 11 in *MYC*-negative Burkitt-like lymphomas

We consider the 11q alteration as pathogenetic and likely primary event in the pathogenesis of mnBLL. Thus, to identify possible candidate genes involved in the pathogenesis of mnBLL, the analyses were focused on those genes mapping to the minimal regions of gain and loss on chromosome 11. These regions encompassed in total 54 protein coding genes and 8 genes of non-coding RNA (refer to Appendix 8.12 for the complete list of genes and to 8.13 for the complete list of non-coding RNAs mapping to the critical regions). Three different criteria have been applied to identify possible candidate genes: differential expression (3.2.3.1.1), recurrent mutations (3.2.3.1.2) and already published association with lymphomas (3.2.3.1.3). The results of the selections are summarized in the following.

3.2.3.1.1 Differentially expressed genes in *MYC*-negative Burkitt-like lymphomas

The first criterion applied to attenuate the list of possible candidate genes was the identification of those genes which map to the regions of alteration in 11q and are also differentially expressed at the mRNA level in mnBLL in comparison to *IG-MYC* positive, molecular BL (mBL) or DLBCL. For this analysis the available gene expression data of six mnBLL cases of the MMML cohort in comparison to 46 *IG-MYC* mBL and 198 DLBCL have been analyzed. Bioinformatic evaluations were supported by M.Sc. Christian Kohler, Institute of Functional Genomics, Statistical Bioinformatics, University of Regensburg.

The results of the analyses of the gene expression data of those genes which are mapping to the minimal region of gain are summarized in Figure 27.

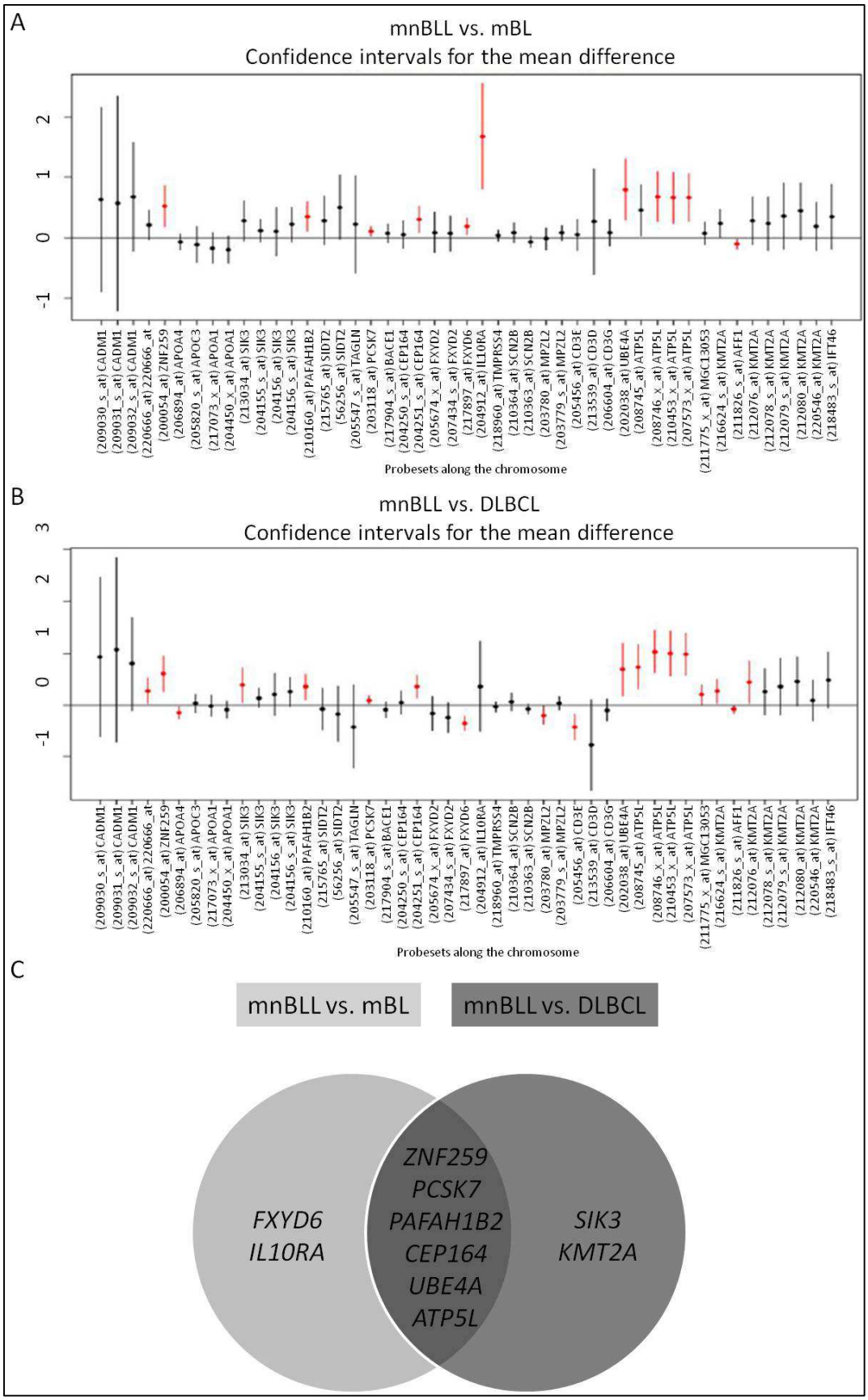


Figure 27: Expression analyses of the genes mapping to the minimal region of gain in 11q in mnBLL. Comparison of transcript expression in mnBLL vs. mBL (A) and vs. DLBCL (B). The mean expression is indicated by a dot and the vertical lines depict 95 % confidence intervals for the fold changes between the two compared groups. Significant differentially expressed genes (adj. p-value <0.1) are highlighted in red. Comparison of the differentially expressed genes of the minimal region of gain, revealed eight genes to be significantly higher expressed in mnBLL than in mBL and eight genes higher than in DLBCL. Six genes showed higher expression in mnBLL as compared to both contrasting groups. Modified from [166].

Overall, eight genes were significantly higher expressed in mnBLL than in mBL and eight genes were significantly higher expressed in mnBLL than in DLBCL (adj. *P*-values <0.1, t-test p-value corrected according to Benjamini and Hochberg). Of those genes, six genes (*ZNF259, PAFAH1B2, PCSK7, CEP164, UBE4A, ATP5L*) were higher expressed in mnBLL as compared to both contrasting groups. The *IL10RA* gene stood out in the gene expression comparison mnBLL vs. mBL, as it showed the highest mean gene expression difference from all genes of the 11q region (adj. p-value 5.51E-09, t-test p-value corrected according to Benjamini and Hochberg).

Figure 28 depicts the results of the analyses of the gene expression data of those genes which are mapping to the minimal region of loss. The comparison revealed six genes to be significantly lower expressed on mRNA level in mnBLL than in mBL and four genes to be lower expressed than in DLBLC (adj. *P*-values <0.1, t-test p-value corrected according to Benjamini and Hochberg). Two of those genes (*FLI1* and *SNX19*) were lower expressed in mnBLL as compared to both contrasting groups. Besides these two genes, the *PRDM10 gene* stood out in the analysis comparing mnBLL vs. mBL showing a significantly reduced expression comparable to *FLI1* (adj. p-value 0.013, t-test p-value corrected according to Benjamini and Hochberg).

Taken together, using transcriptional expression analyses, ten genes were identified to be differentially expressed at the mRNA level in mnBLL in comparison to BL and DLBCL. These genes were subsequently included in the list of possible candidate genes: *ZNF259, PAFAH1B2, PCSK7, CEP164, UBE4A, ATP5L, FLI1, SNX19, IL10RA* and *PRDM10.*

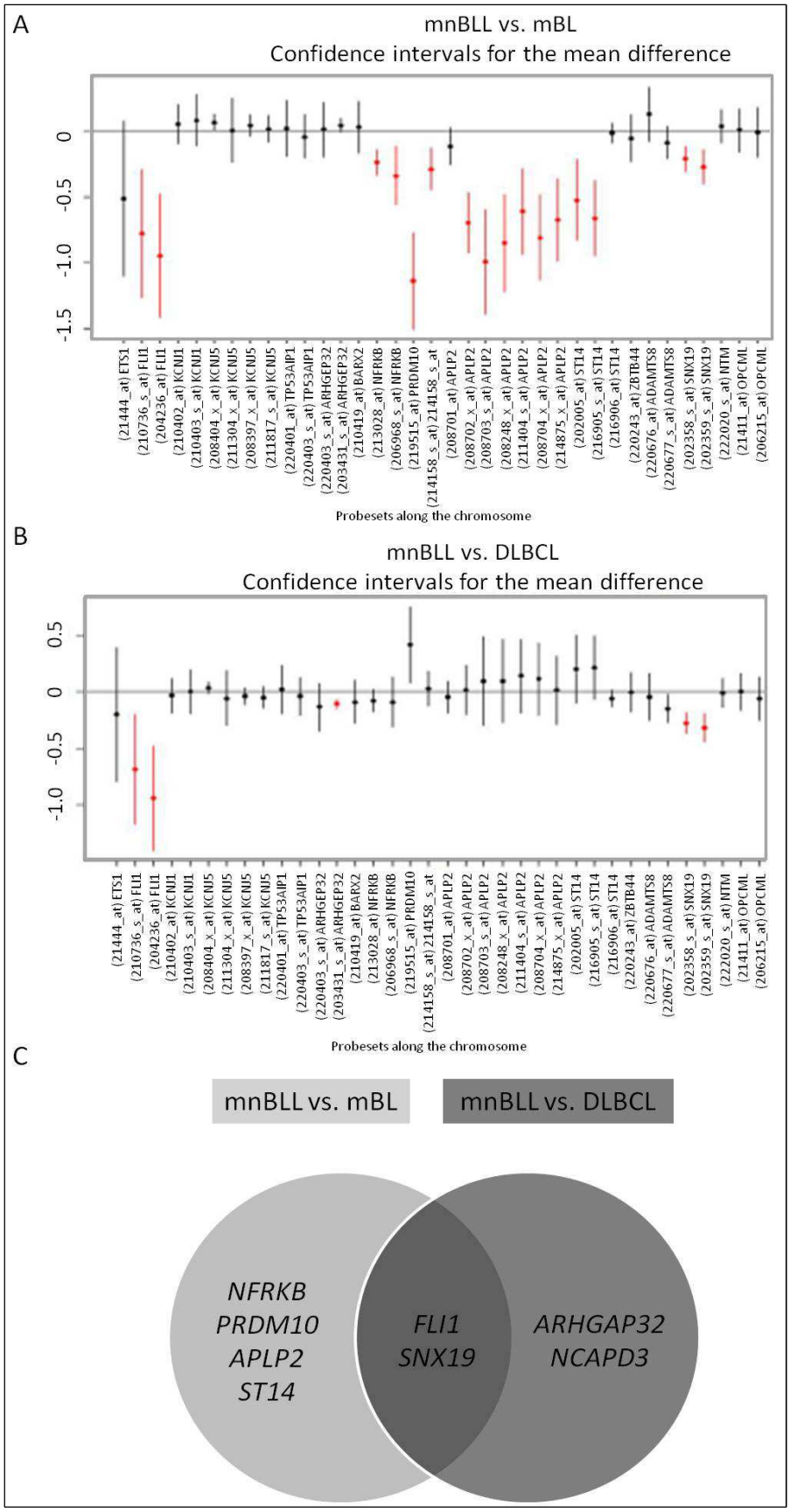


Figure 28: Expression analyses of the genes mapping to the minimal region of loss in 11q in mnBLL. Comparison of transcript expression in mnBLL vs. mBL (A) and vs. DLBCL (B). The mean expression is indicated by a dot and the vertical lines depict 95 % confidence intervals for the fold changes between the two compared groups. Significant differentially expressed genes (adj. p-value <0.1) are highlighted in red. Comparison of the differentially expressed genes of the minimal region of loss, revealed six genes to be significantly lower expressed in mnBLL than in mBL and four genes lower than in DLBCL. Two genes were lower expressed in mnBLL as compared to both contrasting groups. Modified from [166].

3.2.3.1.2 Identification of recurrently mutated genes in *MYC*-negative Burkitt-like lymphomas

To get an overview on the mutational landscape in the mnBLL, next generation sequencing data of four mnBLL (cases 1-3 and 8) and the three mnBLL cell lines (HT, Su-DHL-5, MLMA) were analyzed. Initially, the sequencing data were examined for recurrent mutations within the 54 genes of the minimal region of gain and loss on chromosome 11. Recurrency was defined as a gene being affected by protein changing mutations in at least two cases. To note is, that cases 2, 3 and 8 were analyzed by targeted resequencing using a gene panel in which only 22 of those 54 genes were covered. In the list of recurrently mutated genes were included (i) only nonsynonymous mutations affecting the coding sequence (CDS) of a gene (ii) mutations with a mutated allele frequency (MAF) higher than 0.1 and (iii) mutations which have no dbSNP annotation (build 141, 27/01/2015). According to these definitions none of the 54 genes of the minimal region of gain and loss was found to be recurrently mutated in the analyzed mnBLL and cell lines. Nevertheless, nine genes of the minimal regions were affected by protein coding mutations in one of the respective cases (summarized in Table 36).

Table 36: Overview on genes affected by protein coding mutations mapping to the minimal region of gain and loss on chromosome 11.

	PID	Gene	Genomic Position (hg19) in bp	Mutation	Consequence on protein	MAF	PolyPhen
Gain region	HT	*BUD13* NM_032725	chr11:116,633,542	c.763A>T	p.Thr255Ser	0.14	benign
	HT	*SIK3* NM_025164	chr11:116,730,273	c.2155A>T	p.Ile719Phe	0.24	benign
	Su-DHL-5	*TAGLN* NM_003186	chr11:117,074,158	c.316G>T	p.Val106Phe	0.19	damaging
	HT	*DSCAML1* NM_020693	chr11:117,299,170	c.4842C>A	p.Cys1614*	0.15	damaging
	HT	*TMPRSS13* NM_001077263	chr11:117,789,189	c.386C>T	p.Ala129Val	0.26	damaging
	HT	*IL10RA* NM_001558	chr11:117,860,284	c.316A>G	p.Ser106Gly	0.13	benign
	HT	*TMPRSS4* NM_019894	chr11:117,984,012	c.772G>A	p.Val258Met	0.16	damaging
Loss region	Case 8	*ETS1* NM_005238.3	chr11:128,355,983	c.532T>A	p.Tyr154*	0.56	damaging
	HT	*SNX19* NM_014758	chr11:130,785,119	c.716T>G	p.Val239Gly	1	benign

PID, patient identification; genomic position refers to the human reference genome hg19 in bp; MAF, mutated allele frequency. Gain region, minimal region of gain; loss region, minimal region of loss.

Among those nine mutated genes was the *V-Ets Avian Erythroblastosis Virus E26 Oncogene Homolog 1* (*ETS1)* gene which maps within the minimal region of homozygous loss in chromosome 11q in mnBLL. The mnBLL case 8 carried a mutation within the ETS1 gene, showing a MAF of 0.56,

which leads to a premature stop codon. This mutation was verified by Sanger sequencing. Sanger sequencing of *ETS1* in 15 additional mnBLL revealed six mutations in three of 15 cases (20 %) [166]. Three *ETS1* mutations detected in case 5 and case 7 were missense mutations, of which two were predicted by PolyPhen to be benign and one to be probably damaging. Case 3 harbored a splice site mutation. The other two mutations occurring in case 5 and case 7 were silent mutations. In total four of sixteen (25 %) analyzed mnBLL carried mutations within *ETS1*. Please refer to Table 37 for an overview on all *ETS1* mutations.

Table 37: Overview on *ETS1* mutations detected in 16 analyzed mnBLL cases.

***ETS1* (NM_005238.3)**						
PID	**Genomic Position (hg19) in bp**	**Mutation**	**Consequence on protein**	**MAF**	**PolyPhen**	**detected by**
Case 3	chr11:128,391,807	c.82+1G>A	splicing		n.p.	Sanger sequencing
Case 5	chr11:128,391,865	c.25C>G	p.Pro9Ala		probably damaging	Sanger sequencing
	chr11:128,333,521	c.993C>T	p.Gly331Gly		benign	Sanger sequencing
Case 7	chr11:128,391,826	c.64G>A	p.Glu22Lys		benign	Sanger sequencing
	chr11:128,350,277	c.800C>T	p.Ser267Phe		benign	Sanger sequencing
	chr11:128,354,796	c.653C>T	p.Thr218Ile		benign	Sanger sequencing
Case 8	chr11:128,355,983	c.462T>A	p.Tyr154*	0.56	n.p.	Targeted Resequencing, Sanger sequencing

PID, patient identification; genomic position refers to the human reference genome hg19; MAF, mutated allele frequency; n.p., not predictable with PolyPhen. Modified from [166].

The remaining eight genes were solely mutated within the mnBLL cell lines. The HT cell line carried seven and the Su-DHL-5 cell line one mutations within genes of the minimal gain and loss region. Seven of the eight mutations had a MAF lower than 0.3 suggesting for a subclonal origin and, thus, a minor role in lymphomagenesis. The mutation with a higher MAF was detected within the *SNX19* gene in HT cells, mapping to the minimal region of loss. Remarkably, as described previously, SNX19 was lower expressed in mnBLL than in mBL and DLBCL (3.2.3.1.1). The *SNX19* missense mutation c.716T>G had a MAF of 1 meaning that only one mutated copy of the gene was present in the cell line. Nevertheless, by using PolyPhen, the mutation was classified as benign indicating that the mutation has probably no major impact on the protein function.

To elucidate if genes outside the minimal region of gain and loss in chromosome 11 are targeted by recurrent mutations, the analyses were expanded to the whole-exome.

Initially, the whole-genome sequencing data from case 1 and the exome sequencing data from the three mnBLL cell lines were screened for genes which were mutated in more than two cases and which have not been described as SNP in the dbSNP141 database. This analysis resulted in the detection of 63 recurrently mutated genes which are enlisted in the appendix 8.14. Of those recurrently mutated genes only four were mutated in case 1. The mnBLL cell lines on the other hand

showed a high mutation frequency. About 90 % (57/63) of these recurrently mutated genes were mutated in the HT cells. Thus, the list of recurrently mutated genes might be biased by the high mutation load in the mnBLL cell lines. To further reduce the list of recurrently mutated genes, the targeted resequencing data of three mnBLL, were analyzed for mutations within those 63 genes. To note is, that of those only 22 were covered in the targeted resequencing panel. None of the three mnBLL cases harbored a mutation in those 22 genes.

To sum up, next generation sequencing analysis of four mnBLL and three mnBLL cell lines, did not lead to the identification of recurrently mutated genes in mnBLL other than the already described *ETS1* gene. *ETS1* was therefore added to the list of possible candidate genes.

3.2.3.1.3 Examination of candidate genes for published association with lymphomagenesis

All 54 genes of the minimal gain and loss region were examined for a published association or function with regard to lymphomagenesis using the PubMed library. In the following only those genes will be mentioned for which an association with lymphomas has been published.

Only for three of the ten possible candidate genes identified in the gene expression analyses (3.2.3.1.1) (*PAFAH1B2, IL10RA* and *FLI1*) a role with regard to lymphomagenesis has been published. *PAFAH1B2* has been described to be a recurrent target of chromosomal rearrangement in lymphoma [236]. The expression of the *IL10RA* gene has been published to be regulated by MYC in MYC-induced B-cell lymphomas in mice [237] and to be overexpressed in mantle cell lymphoma [238]. The *FLI1* gene has been described to be recurrently deregulated in DLBCL [239].

Additionally, the lysine (K)-specific methyltransferase 2A *(KMT2A)* gene, formerly known as *MLL,* and the *ETS1* gene were added to the list of possible candidate genes based on their published association with lymphoma. The *KMT2A* gene maps to the minimal region of amplification and is known to be frequently translocated in acute lymphoblastic leukemia and lymphoblastic lymphoma [240],[241]. Furthermore, translocations and complex alterations involving *KMT2A* have been identified in DLBCL [242],[239],[243],[244],[245]. Although the difference of KMT2A expression was not significant (adj. p-value >0.1) the mnBLL cases showed a slightly higher expression than the BL and DLBCL. The sole exception was one oligonucleotide tag covering a transcript variant of KMT2A (212076_at) which was significantly higher expressed in mnBLL than in DLBCL (adj. p-value 0.0018, t-test p-value corrected according to Benjamini and Hochberg). *ETS1* has been described to contribute to the regulation of plasmacytic differentiation in normal B-cells via negative regulation of BLIMP1 which is encoded by *PRDM1* [246]. BLIMP1 is a regulator of the plasmacytic differentiation in the germinal center [247]. ETS1 has been proposed to be a tumor suppressor gene in Hodgkin lymphoma [248] and an oncogene in a subset of DLBCL with terminal 11q gains [239]. Figure 29 gives an overview on the identified candidate genes after applying the three different filtering criteria.

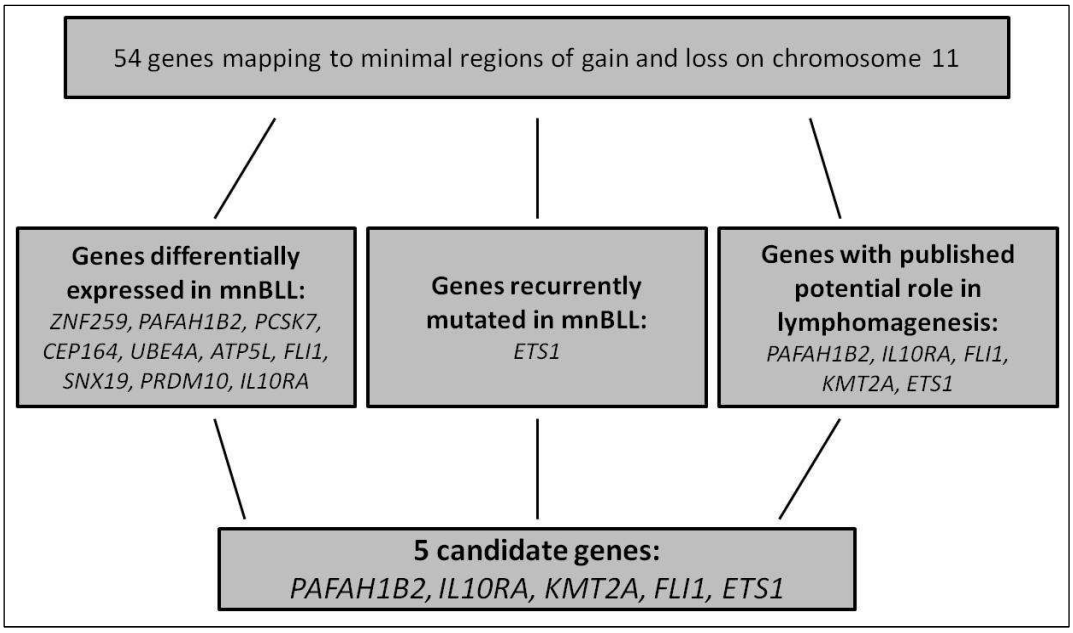


Figure 29: Overview on identified candidate genes after applying three different filter criteria.

After applying the three different criteria, the final list of candidate genes on chromosome 11 potentially involved in the pathogenesis of mnBLL encompassed five possible genes: *PAFAH1B2, IL10RA, FLI1, ETS1* and *KMT2A*. Table 38 summarizes the features of these candidate genes.

Table 38: Overview on the identified candidate genes in mnBLL.

	Gene	Function	Differential expression mnBLL		Mutated in	association with lymphoma (examples)
			vs. BL	vs. DLBCL		
Gain	*PAFAH1B2*	catalytical subunit of PAFAH, inactivation of platelet activating factor	significantly higher expressed	significantly higher expressed	0/16 cases 0/3 cell lines	recurrent chromosomal alterations in lymphoma [243], [244],[239],[245]
Amplification	*IL10RA*	Subunit of interleukin 10 receptor, mediation of immunosuppressive signals	significantly higher expressed	no difference	0/16 cases 1/3 cell lines	role in BL [237], overexpressed in mantle cell lymphoma [238]
	KMT2A	transcriptional coactivator, methyltransferase activity	higher expressed	higher expressed	0/16 cases 0/3 cell lines	recurrent translocation in lymphoblastic lymphoma [243]
Homozygous loss	*FLI1*	member of ETS family of transcription factors	significantly lower expressed	significantly lower expressed	0/16 cases 0/3 cell lines	deregulated in DLBCL [239]
	ETS1	member of ETS family of transcription factors	no difference	no difference	4/16 cases 0/3 cell lines	deleted and downregulated in Hodgkin lymphoma [248], amplified in DLBCL [239]

Gain, minimal region of gain; amplification, minimal region of amplification; homozygous loss, minimal region of homozygous loss

3.2.3.2 Expression of *MYC*-negative Burkitt-like lymphoma-related candidate genes in cell line models

To analyze the expression of the five candidate genes on mRNA and protein level in the mnBLL cell line models quantitative real-time PCR (qPCR) and Western blot were performed in three mnBLL, six BL- and five DLBCL cell lines (2.1.2.3). The results of the expression analyses of the three candidate genes mapping to the minimal region of gain are summarized in Figure 30.

The expression of PAFAH1B2 on mRNA level was higher in the mnBLL cell lines than in the BL and DLBCL cell lines (mean ratio mRNA expression 1.12 vs. 0.77 and 0.51, respectively). Notably, half of the analyzed BL cell lines (BLUE-1, Ca46, U-698-M) showed PAFAH1B2 expression levels comparable to the mnBLL cell lines HT and Su-DHL-5. In comparison to the other mnBLL cell lines, the MLMA cells exhibited a low expression (ratio mRNA expression 0.45). On protein level, the mnBLL cell lines exhibited a much higher PAFAH1B2 expression than the BL and DLBCL cell lines. The DLBCL cell lines particularly had a barely detectable PAFAH1B2 expression. Taken together, these findings are in agreement with the array-based gene expression analyses in the mnBLL showing a significant higher expression of PAFAH1B2 in mnBLL than mBL and DLBCL (3.2.3.1.1).

Quantification of the IL10RA mRNA level in the cell lines showed a higher expression in the mnBLL cell lines than in the BL and DLBCL cell lines. In latter the IL10RA mRNA expression was barely detectable (mean ratio mRNA expression 1.15 vs. 0.03 and 0.11, respectively). The mnBLL cell line MLMA exhibited the highest IL10RA mRNA expression level (ratio mRNA expression 2.33). On the protein level, IL10RA was not only highly expressed in the mnBLL cell lines, but also in four BL cell lines (BL-2, BL-41, BLUE-1, Ca46) and in two DLBCL cell lines (Karpas422, RIVA) which differs from the findings on mRNA level. Furthermore, IL10RA protein expression was not detectable in the BL cell line BL-70. Taken together the results in the cell lines on mRNA level confirm the array-based gene expression analyses showing an increased expression level in mnBLL (3.2.3.1.1). On the protein level the ten analyzed cell lines, irrespective of the lymphoma entity, exhibited a comparable IL10RA expression level.

The expression of KMT2A in the cell line panel was solely analyzed on mRNA level, as a suitable antibody for detection of the KMT2A protein was not available. The mnBLL cell lines exhibited a higher KMT2A expression than the BL and DLBCL cell lines on mRNA level, (mean ratio mRNA expression 9.22 vs. 2.53 and 1.85, respectively). The mnBLL cell line MLMA exhibited the highest KMT2A expression (ratio mRNA expression 18.1), showing a ~3.5-fold higher expression than the other mnBLL cell lines HT and Su-DHL-5 (ratio mRNA expression 4.3 and 5.3 respectively). This is in agreement with the array-based gene expression analysis of the mnBLL, showing a slightly, though not significant, higher expression level in comparison to BL and DLBCL (3.2.3.1.1).

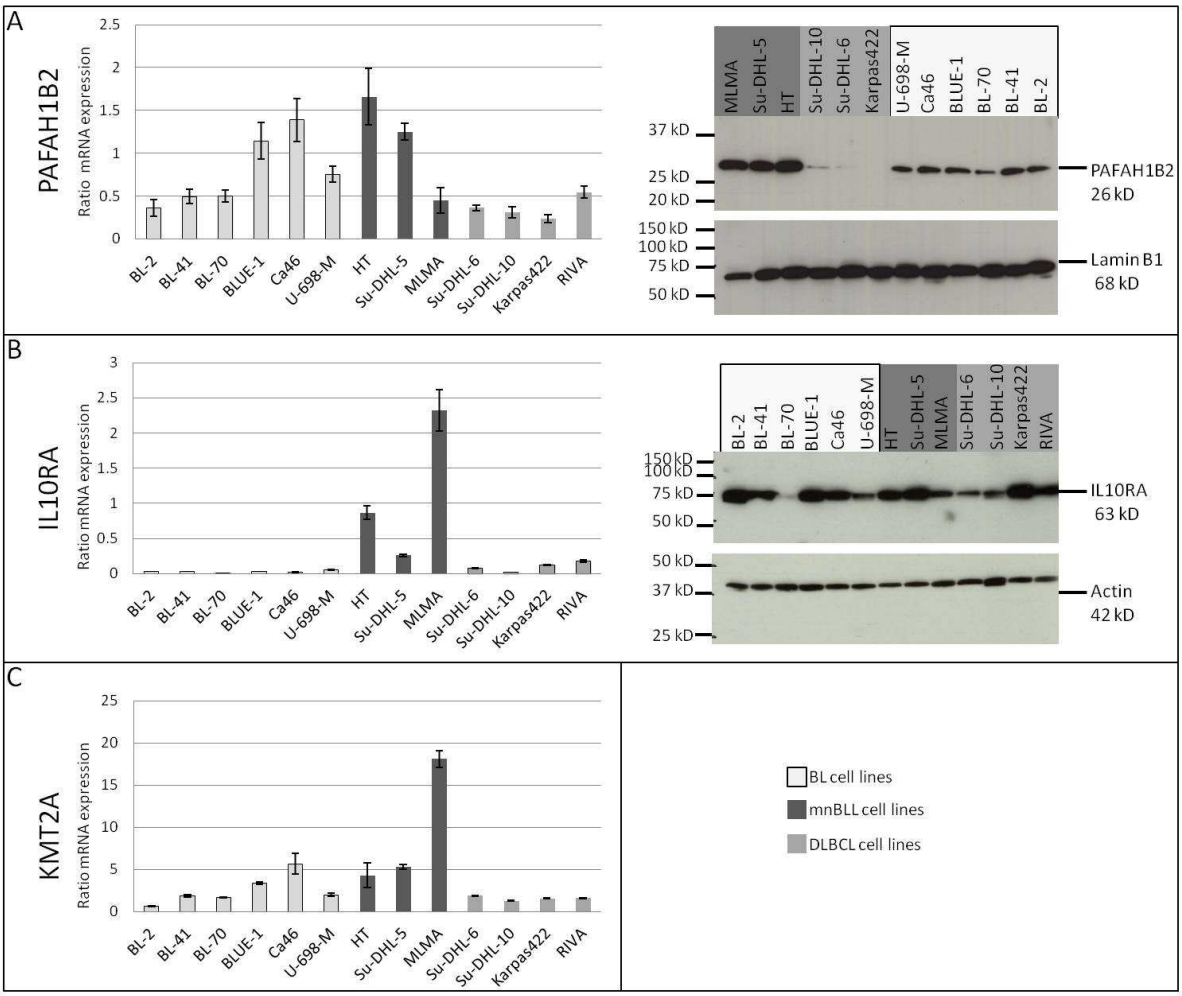


Figure 30: Expression analyses of candidate genes mapping to the minimal region of gain in mnBLL. The three panels depict expression analyses in B-cell lymphoma cell lines of PAFAH1B2 (A), IL10RA (B) and KMT2A (C) on mRNA level using qPCR and on protein level using Western blot. The qPCR graphs depict the mean of the relative expression and the standard deviation of three independent technical replicates. The Western blot which is displayed is an example of at least two performed technical and biological replicates. The complete Western blot is depicted in the appendix (8.15). The PAFAH1B2 Western blot figure has been adapted from Salaverria *et al.* [166].

Subsequently, the expression levels of the candidate genes *FLI1* and *ETS1* mapping to the minimal region of loss, more precisely into the minimal region of homozygous loss, were analyzed. The findings are documented in Figure 31.

The mean FLI1 expression on mRNA level was lower in the mnBLL than in the BL cell lines and slightly lower than in the DLBCL cell lines (mean mRNA expression ratio 3.25 vs. 5.15 and 3.44, respectively). Nevertheless, considerable heterogeneity in the expression level was observed. The MLMA cell line exhibited a much higher expression (ratio mRNA expression 6.3) than the other two mnBLL cell lines HT and Su-DHL-5 (ratio mRNA expression 1.2 and 2.2, respectively). Moreover, the DLBCL cell line Karpas422 exhibited a very low FLI1 expression (ratio mRNA expression 0.3) in comparison to the other DLBCL cell lines, which showed comparable FLI1 expression levels to the BL cell lines. To note is, that the FLI1 antibody used for protein expression analyses detects two protein isoforms of 48 kD and 51 kD size [249]. The BL cell lines expressed both isoforms at comparable levels, whereas in the

DLBCL cell lines the 48 kD isoform was higher expressed than the 51 kD isoform. The mnBLL cell lines HT and Su-DHL-5 seemed to have lost the expression of both isoforms whereas the MLMA cell line solely expressed the 48 kD isoform but at comparable levels to the DLBCL cell lines. Taken together, the analyses of FLI1 mRNA and protein expression revealed a lower expression in the mnBLL cell lines than in the BL and DLBCL cell lines. This is in agreement with the findings of the array-based gene expression analyses of the mnBLL which exhibited a significant lower FLI1 mRNA expression than the mBL and DLBCL (3.2.3.1.1).

Furthermore, the expression levels of ETS1 were analyzed. At mRNA level the mean ETS1 expression in the mnBLL cell lines was lower than in the BL cell lines but slightly higher than in the DLBCL cell lines (mean ratio mRNA expression of 13.39 vs. 21.36 and 10.96, respectively). At protein level the expression was comparable in all cell lines, whereas some of the BL cell lines (BL-2, Ca46, U-698-M) showed a slightly higher expression than the other cell lines. Overall the ETS1 expression differences in the different cell line entities were minor which is in line with the qPCR and array-based analyses (3.2.3.1.1).

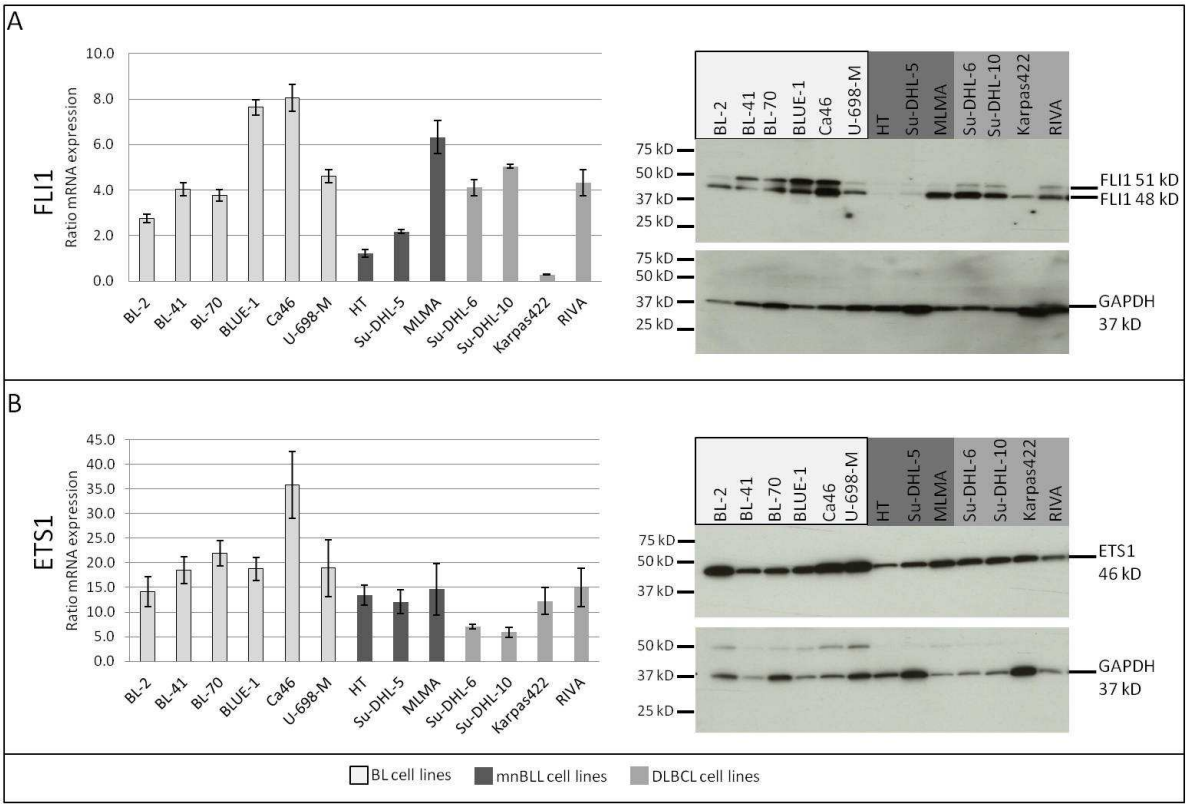


Figure 31: Expression analyses of candidate genes mapping to the minimal region of loss in mnBLL. The panel depicts expression analyses of FLI1 (A) and ETS1 (B) on mRNA level using qPCR and on protein level using Western blot in B-cell lymphoma cell lines. The qPCR graphs depict the mean and the standard deviation of three independent technical replicates. The Western blot shown is an example of at least two performed technical and biological replicates. The complete Western blot is depicted in the appendix (8.15).

Taken together, the expression analyses of the candidate genes in the mnBLL cell lines in comparison to BL and DLBCL cell lines confirmed PAFAH1B2 and IL10RA to be higher and FLI1 to be lower expressed.

3.2.3.3 Knock down study of the candidate gene *PAFAH1B2*

The first candidate gene chosen for the functional studies was the *PAFAH1B2* gene for which, as reported in 1.4, an oncogenic function has been assumed. *PAFAH1B2*, mapping to the minimal region of gain, was constitutively up-regulated on mRNA as well as on protein level in mnBLL. Moreover, it has been described to be an oncogenic target of a translocation involving the *IGH* locus in lymphoma [236]. *PAFAH1B2*, also called PAF-AH1b Alpha 2 subunit, encodes the beta subunit of the platelet-activating factor acetylhydrolase (PAF-AH) which inactivates the platelet activating factor (PAF). Among others, it has been described to have anti-apoptotic functions [250].

To elucidate if *PAFAH1B2* could be indeed a potential oncogene in mnBLL, knock down analyses of this gene were performed using lentiviral shRNA constructs. The analyses were initially based on the hypothesis that if *PAFAH1B2* plays a vital role in mnBLL, the induced loss of expression might lead to a loss of viability or growth suppression. The Su-DHL-5 cell line was used for the knock down analyses which were performed as described in 2.2.5.7. In total three different shRNAs targeting PAFAH1B2 (PAFAH1B2 shRNA1-3) as well as a non-targeting control (NTC) shRNA were delivered to the Su-DHL-5 cells using lentiviral vectors. Three days after induction of shRNA expression using IPTG the knock down efficiency of the PAFAH1B2 shRNA constructs in comparison to the non-targeting control shRNA was analyzed by Western blot as depicted in Figure 32A. All PAFAH1B2 shRNA constructs repressed the PAFAH1B2 expression almost completely. However, PAFAH1B2 shRNA1 seemed to be expressed even without IPTG induction, indicating a leakiness of the construct.

To determine if the knock down had an influence on the viability, the transduced cells were analyzed using a cell viability assay. This analysis showed that although the expression of PAFAH1B2 was almost completely reduced, the cell viability of the cells was not altered in comparison to the NTC control (Figure 32B).

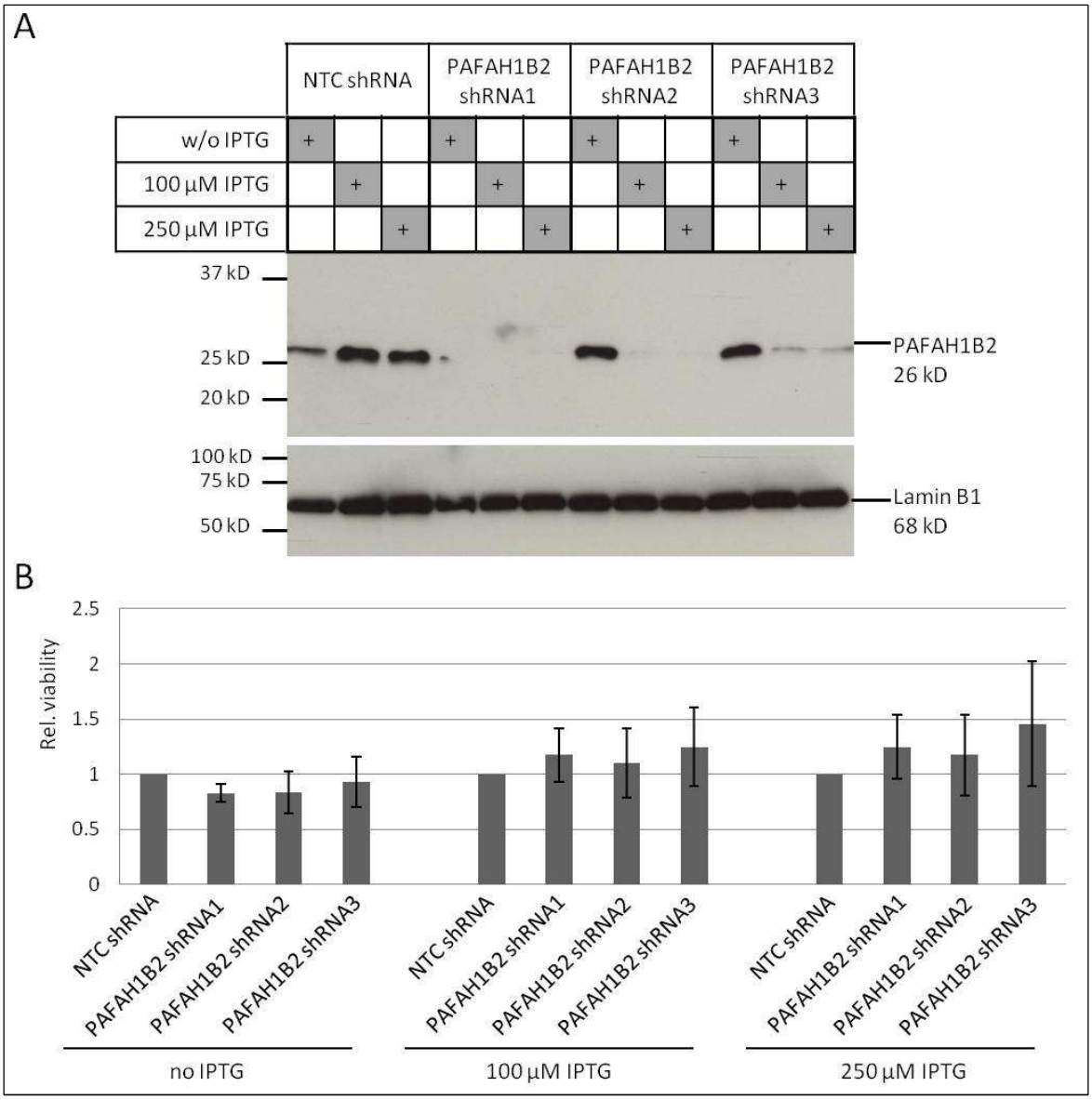


Figure 32: Knock down of PAFAH1B2 expression in Su-DHL-5 cell line. A: Example of knock down efficiency of PAFAH1B2 in the Su-DHL-5 cell line using Western blot. Cells have been treated with 100 µM IPTG or 250 µM IPTG and the expression of PAFAH1B2 was compared to cells not treated with IPTG (w/o IPTG) and to cells expressing the NTC shRNA. Treatment condition is indicated by + in a grey box. Detection of Lamin B1 protein was used as loading control. Refer to the appendix 8.15 for the complete Western blot figure. B: Analyses of cell viability after knock down of PAFAH1B2. Depicted are the mean and the standard deviation of three independent experiments.

The results of the PAFAH1B2 knock down experiments indicate that PAFAH1B2 expression and function is not essential for the viability of the mnBLL cell line Su-DHL-5 under the experimental conditions tested. This suggests that *PAFAH1B2* might not be on its own a driving oncogenic force in the tumorigenesis of mnBLL.

3.2.3.4 Analyses of the IL10R pathway in *MYC*-negative Burkitt-like lymphomas

The next candidate gene studied was *IL10RA*. *IL10RA*, mapping to the minimal region of amplification, was significantly higher expressed on mRNA level in mnBLL than in mBL, though at

protein level the expression was more homogenous in the cell lines. The IL10 receptor (IL10R) consists of an α-subunit which is encoded by *IL10RA* and a β-subunit which is encoded by *IL10RB*. The ligand of the IL10R is interleukin 10 (IL10). Upon binding of the ligand, the IL10 receptor activates among others the downstream PI3K signaling pathway [251]. PI3K signaling and its components have been introduced in 1.3.4 as an oncogenic signaling pathway in BL.

3.2.3.4.1 Expression analysis of IL10RA

As shown in Figure 30, mRNA expression of IL10RA was higher in the mnBLL cell lines than in BL and DLBCL cell lines. Only the overall IL10RA protein expression level was analyzed in the B-cell lymphoma cell lines by Western blot (3.2.3.2). Thus, since the main site of function of IL10R is on the cell surface, the cell surface expression of IL10RA was examined. To this end, cell surface and intracellular staining of IL10RA in a cell line panel consisting of three mnBLL (HT, Su-DHL-5, MLMA), three BL (BL-2, BL-41, BL-70) and three DLBCL (Su-DHL-6, Karpas422, RIVA) cell lines was performed and evaluated using flow cytometry (2.2.6.5). The results are summarized in Figure 33.

The highest IL10RA cell surface expression exhibited the BL cell lines showing a median of 22.5 % IL10RA-positive cells in comparison to 14.6 % in mnBLL cell lines and 5.9 % in DLBCL cell lines. The BL cell line BL-2 and the mnBLL cell line MLMA had the lowest amount of IL10RA positive cells (2.1 % and 3.3 %, respectively) (Figure 33A).

The intracellular IL10RA staining revealed on the other hand, that the mnBLL cell lines exhibited a higher median IL10RA-positivity of 73.7 % in comparison to 59.2 % in BL and 54.1 % in DLBCL cell lines (Figure 33A). The amount of IL10RA-positivity is influenced by the overall expression, which has been shown by Western blot to differ between the cell lines entities (Figure 30). Therefore, the ratio of IL10RA cell surface expression to the overall expression, which is reflected by the intracellular expression, was calculated. Figure 33B shows that the BL cell lines had the highest relative surface expression (median 0.41) whereas the mnBLL and DLBCL cell lines exhibited solely half of the relative surface expression (0.21 and 0.11 respectively). To note is that the BL cell line BL-70 and the DLBCL cell line RIVA exhibited the highest surface expression ratio with 0.63 and 0.54 respectively. Oddly, the BL-70 cell line showed in the Western blot analyses the lowest IL10RA expression (Figure 30).

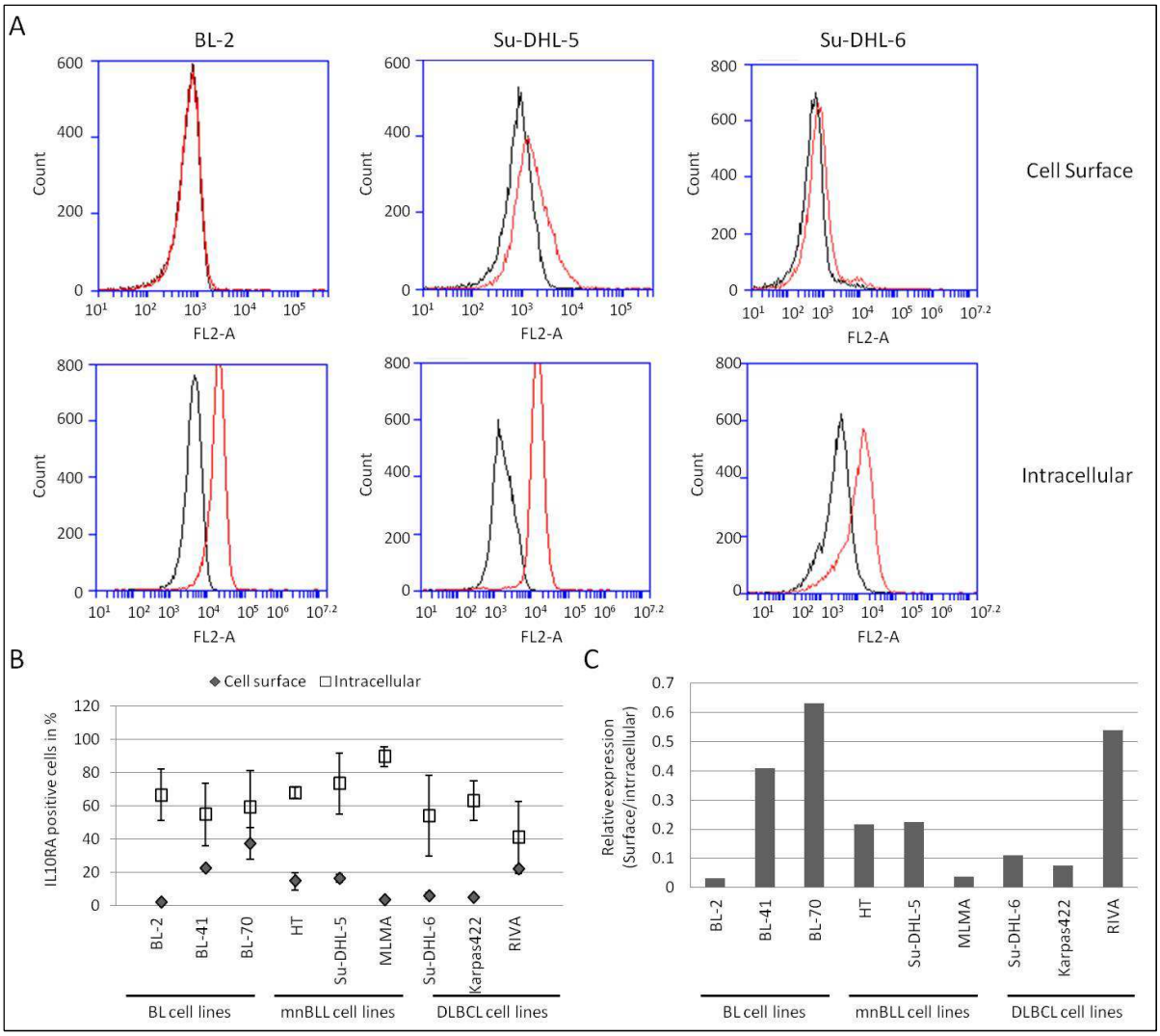


Figure 33: Analyses of IL10RA cell surface and intracellular expression using flow cytometry. A: Representative histograms of the IL10RA expression analysis with FACS in a BL (BL-2), a mnBLL (Su-DHL-5) and a DLBCL (Su-DHL-6) cell line. The upper and the lower panel depict the cell surface and the intracellular expression analysis, respectively. The black peaks represent the isotype control measurement and the red peaks the IL10RA measurement. FL2-A, fluorescence channel 2 area; count, number of cells. B: Percentages of cells expressing IL10RA on the cell surface or intracellular. The measurements have been repeated at least two times. Depicted are the mean and the standard deviation. C: Ratio of mean IL10RA cell surface expression to intracellular expression. The higher the ratio the higher the relative amount cell surface expression.

Taken together, the flow cytometry analyses of the IL10RA cell surface expression did not indicate that the mnBLL cell lines have a higher IL10RA cell surface expression than the BL and DLBCL cell lines. Nevertheless, mnBLL cell lines tended to have a higher intracellular IL10RA expression which is in line with the results of the qPCR analyses (3.2.3.2).

3.2.3.4.2 Expression analysis of IL10

Next, the differential expression of the IL10R ligand, IL10, was analyzed in the mnBLL as well as in the mnBLL cell lines. The gene expression data of six of the mnBLL were provided by the MMML project. The data and statistical analysis were provided by M.Sc. Christian Kohler, Institute of Functional Genomics, Statistical Bioinformatics, University of Regensburg. Strikingly, IL10 which maps to

chromosome 1q32.1 showed a significantly higher expression in the mnBLL than in the mBL analyzed based on U133A gene expression array data (adjusted p-value 0.003, t-test p-value corrected according to Benjamini and Hochberg) (Figure 34A).

Analyses of the mRNA level of IL10 in the mnBLL cell lines in comparison to BL and DLBCL cell lines showed an extraordinary high expression of IL10 (mean ratio mRNA expression 0.372) whereas within the BL and the DLBCL cell lines its expression was barely detectable (mean ratio mRNA expression 0.002 and 0.0004, respectively). Of note is that the mnBLL cell lines Su-DHL-5 and MLMA showed a 20-90 fold higher IL10 expression than the other cell lines (Figure 34B) whereas the mnBLL cell line HT exhibited a barely detectable expression.

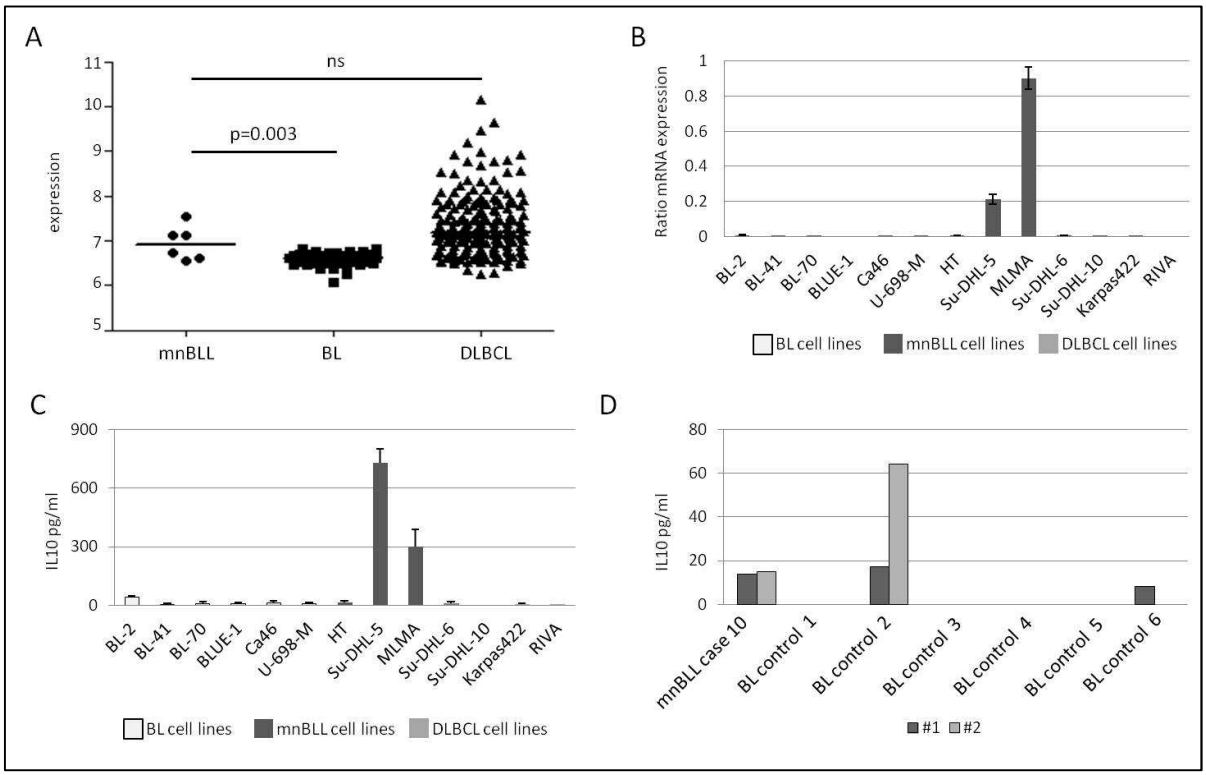


Figure 34: IL10 expression in B-cell lymphoma cell lines and sera of patients with B-cell lymphoma. A: U133A gene expression microarray data of 6 mnBLL, 46 BL and 198 DLBCL. The mnBLL show a significant higher IL10 mRNA than the BL (adjusted p-value 0.003). Lines indicate the median. B: Mean and standard deviation of three independent technical replicates of the IL10 mRNA expression analyses in the cell lines by qPCR. C: Mean and standard deviation of three independent measurements of IL10 level in the cell culture supernatant of different B-cell lymphoma cell lines by ELISA. D: Results of two measurements of IL10 level (#1, #2) by ELISA which have been performed on the serum from initial diagnosis of one patient with mnBLL (case 10) and six patients with BL (BL control 1-6).

The extracellular IL10 protein expression level was analyzed using ELISA (2.2.6.4). The cell lines with the highest IL10 concentration in the cell supernatant were the mnBLL cell lines Su-DHL-5 (730 pg/ml) and MLMA (299 pg/ml). Besides the BL cell line BL-2 (41 pg/ml), the other cell lines showed only a barely detectable or no expression of IL10 protein in the supernatant at all (Figure 34C). Thus, the results of the IL10 protein expression analysis in the cell lines were in line with the findings of the IL10 qPCR.

Furthermore, the IL10 protein level was analyzed in the serum of one mnBLL patient (case 10), of whom material was available, in comparison to sera of six patients with primary BL from initial diagnosis (Figure 34D). The IL10 level within the serum from the patient with mnBLL presentation was about ~14 pg/ml. In four of the BL control sera, no IL10 was detectable at all. The other two BL control sera (BL control 2 and 6) exhibited comparable IL10 level to the mnBLL.

To sum up, two of the three mnBLL cell lines exhibited a high IL10 concentration in the cell culture supernatant whereas in the BL and DLBCL cell lines IL10 was not expressed or below the detection limit. Within the serum of the patient with mnBLL, IL10 was detectable at comparable levels to that of two of the four sera from patients with BL. These findings suggest that both IL10 and IL10RA are regularly expressed in the mnBLL.

The finding that the *IL10RA* gene is overexpressed in mnBLL can be explained by the recurrent chromosomal gain on 11q. To unravel the reason for the high IL10 expression in mnBLL, several approaches were used: (i) analyses of chromosomal alterations of *IL10* locus; (ii) identification of mutations within *IL10* gene and (iii) analyses of methylation status of *IL10* locus.

By mining the CNA data, only two mnBLL were identified to exhibit chromosomal alterations affecting 1q32.1 including the *IL10* locus: MPI-315 carried a chromosomal loss (1q32.1-qter) and MPI-382 carried a chromosomal gain (1q21.1-qter). Hence, solely for the mnBLL case MPI-382 a higher IL10 expression might be explainable by a copy number gain covering the *IL10* locus (Figure 35A). Thus, analyses of the CNA render chromosomal alterations as a cause for the high IL10 expression in mnBLL unlikely.

The already described whole-genome, exome and targeted re-sequencing data of the mnBLL and the mnBLL cell lines exhibited no mutations affecting *IL10* rendering also oncogenic mutations unlikely.

Thus, using the HumanMethylation 450k Bead Chip array data the *IL10* gene locus was analyzed for differential methylation comparing eight mnBLL with 18 BL and seven DLBCL, as well as the corresponding cell line subtypes. Ten CpGs at the *IL10* gene locus including CpGs ~20 kb up- and downstream of the gene, were covered on the 450k array. The mnBLL showed a mean methylation of the complete *IL10* locus of 0.27 and the BL and DLBCL of 0.25 and 0.35, respectively. The mean methylation of the respective cell lines models was 0.23, 0.47 and 0.45 in mnBLL, BL and DLBCL cell lines, respectively. Overall, the methylation status of none of the CpGs linked to the *IL10* gene differed significantly in mnBLL compared to BL or DLBCL ($p > 0.05$, t-test) (Figure 35B).

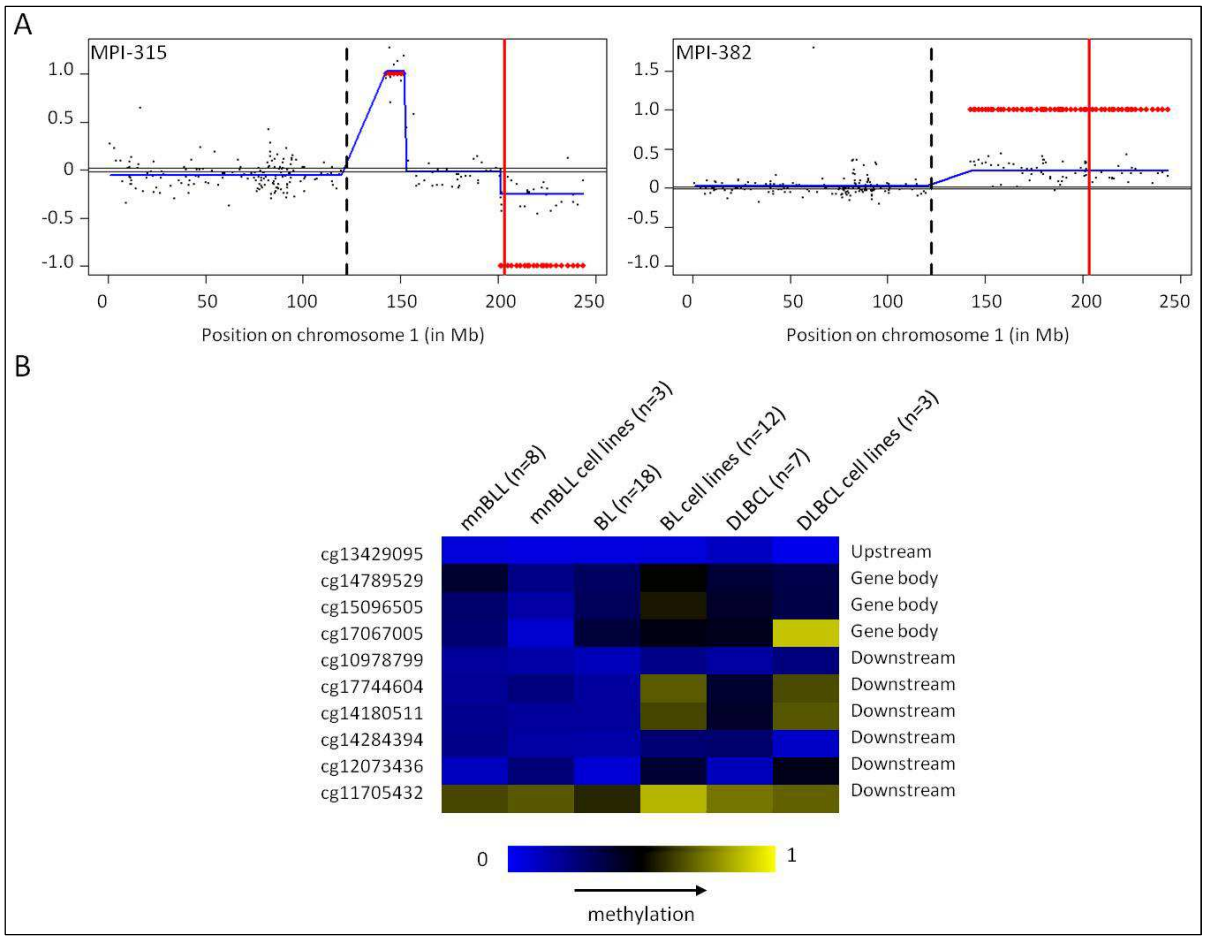


Figure 35: Chromosomal and methylation landscape of the *IL10* gene locus in mnBLL. A: Plots showing the alterations on chromosome 1 for the mnBLL cases MPI-315 and MPI-382. The x-axis depicts the position on the chromosome in Mb and the y-axis the normalized log2Ratio. The blue line outlines the chromosomal profile. The red line indicates the position of the *IL10* gene locus. B: The heatmap shows the methylation values (mean AVG-β value) of the single CpGs at the *IL10* gene locus of eight mnBLL, three mnBLL cell lines, 18 BL, twelve BL cell lines, seven DLBCL and three DLBCL cell lines. The blue color indicates low methylation and the yellow color high methylation. The localization of each CpG is given as "upstream" of the gene, "gene body" within the gene and "downstream" of the gene.

Taken together, the increased expression of IL10 in the mnBLL in comparison to the BL and DLBCL can yet neither be explained by chromosomal or mutational alterations nor by changes at the methylation level.

3.2.3.4.3 Functional analyses of the IL10 signaling pathway

The overexpression of IL10RA and IL10 might indicate a possible autocrine signaling via the IL10 receptor pathway leading to a subsequent self-stimulation of growth and survival. To test this hypothesis, single components of this putative autocrine signaling pathway were inhibited: (i) the IL10 receptor subunit IL10RA, (ii) the ligand IL10 and (iii) the IL10R-downstream PI3K signaling.

IL10R inhibition was performed in the mnBLL cell line Su-DHL-5 and in the BL cell line BL-2 as described in 2.2.5.8. Two different IL10RA antibodies were used: T3431, which has been also used for IL10RA Western blot analyses and 3F9 for which neutralizing properties have already been described

[211],[212]. Treatment with the T3431 antibody led to a loss of viability in both cell lines. The highest concentration applied of 10 µg/ml antibody led to 80 % reduction of cell viability in BL-2 cells and ~90 % in Su-DHL-5 cells (Figure 36A). Treatment with 3F9 antibody on the other hand, had no impact at all on the cell viability of BL-2 and Su-DHL-5 cells (Figure 36B).

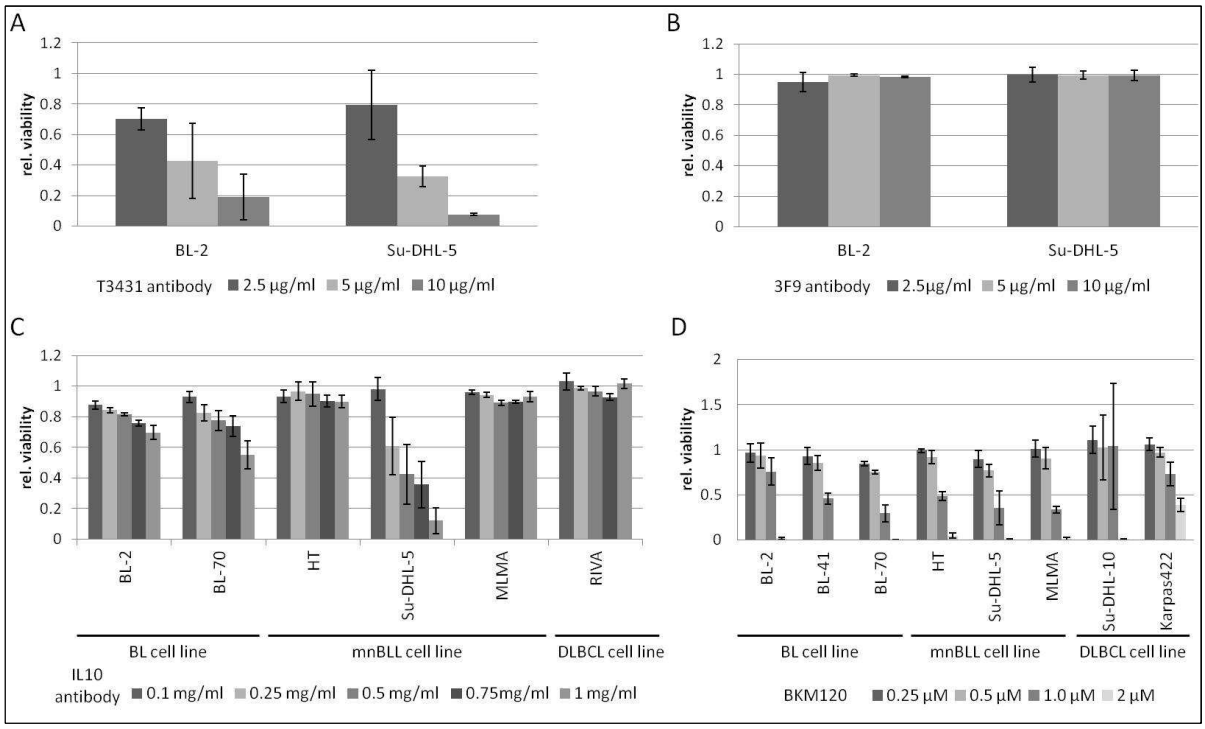


Figure 36: Treatment of different components of the IL10 signaling pathway. A,B: Treatment of BL-2 (BL cell line) and Su-DHL-5 (mnBLL cell line) cells with varying concentration of two different IL10RA antibodies (T3431 and 3F9). C: Treatment of BL cell lines (BL-2, BL-70), mnBLL cell lines (HT, Su-DHL-5, MLMA) and DLBCL cell line (RIVA) with various concentrations of neutralizing IL10 antibody (ab34843). D: Impact of cell line treatment with BKM120 on cell viability. Studied were three BL cell lines (BL-2, BL-41, BL-70), three mnBLL cell lines (HT, Su-DHL-5, MLMA) and two DLBCL cell lines (Su-DHL-10, Karpas422). The graphs depict the mean and the standard deviation of three experiments.

Inhibition of IL10 was performed using various concentrations (0.1 mg/ml-1 mg/ml) of neutralizing IL10 antibody (ab34843) on two BL cell lines (BL-2, BL-70), three mnBLL cell lines (HT, Su-DHL-5, MLMA) and one DLBCL cell line (RIVA) as outlined in 2.2.5.8. After two days of treatment the impact on cell viability was analyzed (2.2.5.10). The only cell line showing clearly reduced cell viability was the mnBLL cell line Su-DHL-5 (Figure 36C). Treatment of Su-DHL-5 cells with 0.25 mg/ml IL10 antibody reduced the viability about 40 %. The highest applied concentration of 1 mg/ml IL10 antibody reduced the viability by about 90 %. The BL cell lines BL-2 and BL-70 had a reduced cell viability of about 30 % and 45 %, respectively, after treatment with 1 mg/ml IL10 antibody. Otherwise, the IL10 antibody treatment had no impact on cell viability on all the other cell lines.

Finally, the IL10 receptor downstream PI3K signaling pathway has been inhibited. The compound Buparlisib (BKM120) has been reported to be a potent inhibitor of the PI3K class IA in different clinical trials [252],[253]. Moreover, Schmitz *et al.* [72] showed that BKM120 is toxic to BL cell lines. Using their treatment protocol as outlined in 2.2.5.9 the impact of BKM120 on the mnBLL cell lines

was analyzed. The BKM120 treatment was performed in three BL cell lines (BL-2, BL-41 and BL-70), two DLBCL cell lines (Su-DHL-10 and Karpas422) and the three mnBLL cell lines (HT, Su-DHL-5 and MLMA). Figure 36D summarizes the results of three independent experiments. The highest applied concentration of 2 µM BKM120, led to a complete loss of cell viability in the BL cell lines, the DLBCL cell line Su-DHL-10 and the mnBLL cell lines. Lower BKM120 concentrations as 1 µM reduced the cell viability of the same cell lines about 50 %. These results indicate that like the BL cell lines, the cell survival of mnBLL cell lines seems to depend among others on PI3K signaling.

Taken together, the inhibition of IL10RA reduced the cell viability in the tested mnBLL cell line but as well in the BL cell line. Treatment with the inhibiting IL10 antibody lead to a nearly complete loss of cell viability within the mnBLL cell line Su-DHL-5. Inhibition of the IL10R downstream PI3K signaling pathway led to the complete loss of cell viability of all mnBLL cell lines. Thus, the findings indicate a role of the IL10 signaling in the pathogenesis of mnBLL.

3.2.4 Integrative analysis of the findings in *MYC*-negative Burkitt-like lymphomas

In the following, as well as in Figure 37 the results of the analysis of the genomic, genetic and expressional landscape of the mnBLL are summarized.

The hallmark characteristics of the mnBLL is the chromosomal alteration on chromosome 11q consisting of a copy number gain (11q23.3) including a minimal region of amplification and a telomeric copy number loss (11q24.3-q25) including a minimal region of homozygous loss. These alterations were present in 23 of 25 analyzed mnBLL (92 %). Two cases solely harbored the telomeric loss. Additionally, recurrent copy number alteration as copy loss in 6q14.3-q21 involving the *EPHA7* and *PRDM1* genes, copy gains affecting 13q31.3 and 18q21 involving *miR-17-92* gene cluster and *TCF4* gene respectively, might cooperate with the 11q alteration in the pathogenesis of mnBLL.

No recurrently mutated gene could be identified by gene-wide sequencing approaches. But recurrent mutation of *ETS1* mapping to the minimal region of homozygous loss in mnBLL was observed by Sanger sequencing.

Expression analysis on mRNA and protein level combined with data mining approaches led to the identification of *PAFAH1B2, IL10RA* and *KMT2A* as oncogene candidates and of *FLI1* and ETS1 as tumor suppressor gene candidates in mnBLL.

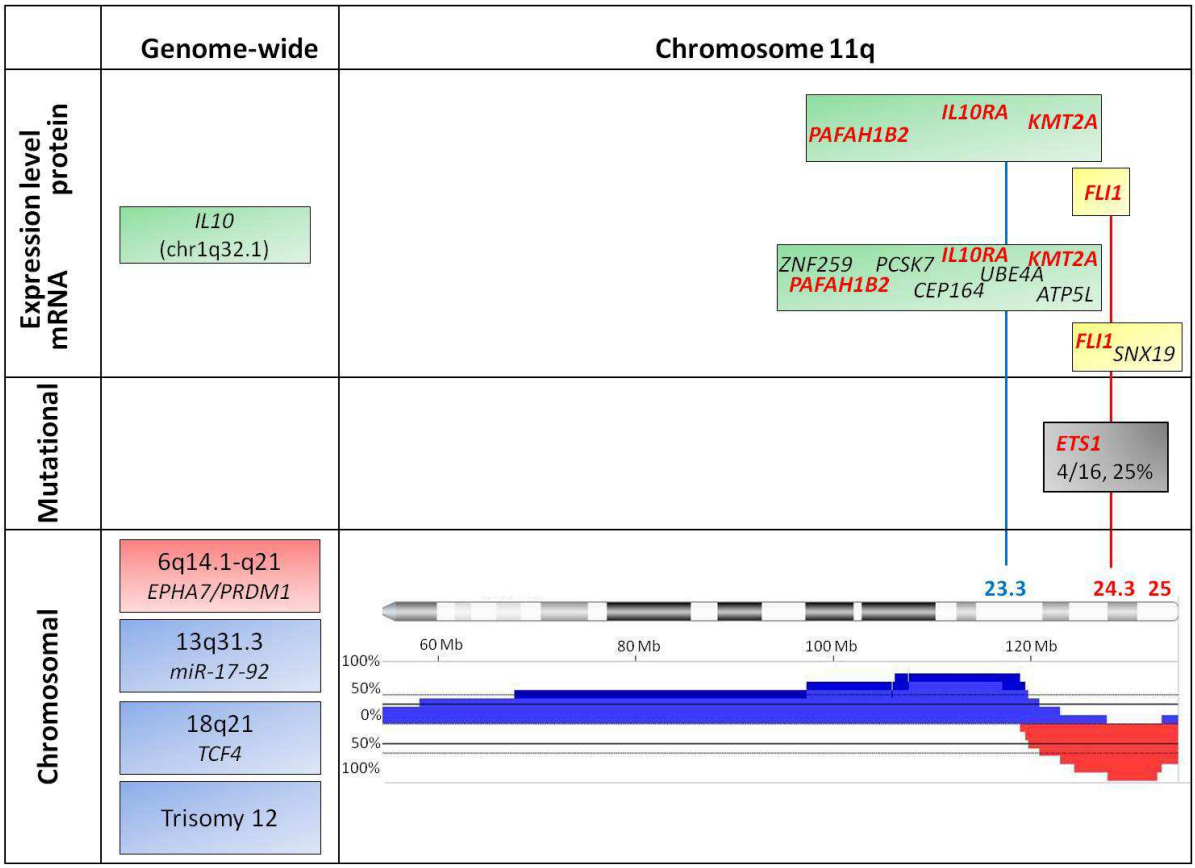


Figure 37: Summary of recurrent genomic and genetic alterations in mnBLL, as well as differential gene expression on mRNA and protein level in mnBLL in comparison to BL and DLBCL. Blue color indicates copy number gains, red color indicates copy number losses, green color indicates genes significantly higher expressed whereas yellow color indicates genes significantly lower expressed in mnBLL than in comparison to BL and DLBCL. Gene names highlighted in red represent identified possible candidate genes contributing to the lymphomagenesis of the mnBLL.

Thus, based on these comprehensive analyses, five possible candidate genes contributing to the lymphomagenesis of mnBLL were described: *PAFAH1B2, IL10RA, KMT2A, FLI1* and *ETS1*.

Functional studies of *PAFAH1B2* in a mnBLL cell line suggested that this gene plays a minor role in the lymphomagenesis of this tumor, whereas functional studies of *IL10RA* and its related ligand (IL10) as well as the downstream signaling pathway points toward a specific role in mnBLL.

4 Discussion

The analyses of this thesis focused on two mature aggressive B-cell lymphomas of childhood: *MYC*-positive Burkitt lymphoma (BL), and its *MYC*-negative counterpart, *MYC*-negative Burkitt-like lymphoma (mnBLL).

The aims of the analyses of BL were to decipher if BL possesses traces of pluripotency and to study the expression of a newly identified alternative spliced isoforms of *TERT* in BL. Moreover, BL should be investigated for recurrently mutated genes and their role in Burkitt lymphomagenesis studied. Furthermore, the aim of this thesis was to analyze the chromosomal, mutational and transcriptional landscape of the mnBLL in order to identify potential candidate genes which might have a tumorigenic role in their lymphomagenesis.

The following discussion will in the first part focus on the findings in BL and in the second part on those in mnBLL. Thereafter, the characteristics of both lymphoma subtypes will be compared.

4.1 *MYC*-positive Burkitt lymphoma

The discussion of the *MYC*-positive Burkitt lymphoma can be roughly divided into two parts: analyses based on gene expression data and analyses of the mutational landscape which led to the identification of secondary alterations occurring concomitant to the MYC deregulation in BL. The major findings of this thesis in BL are:

- The PluriTest algorithm showed a higher similarity of BL to hPSC than of non-BL. Nevertheless, an origin from or resemblance to cells with PSC features for the BL could not be established.
- The existence of a new exon in the *TERT* gene could be verified. Modeling of the effects of the splice variant including this new exon predicted (i) translation of proteins with either a loss of the N-terminal or C-terminal domain or (ii) due to the introduction of a premature stop codon the degradation of the mRNA by NMD.
- The BL cell lines showed a high frequency of mutations in genes known to be recurrently mutated in BL (*TP53, CCND3*), whereas genes recurrently mutated in DLCBL (*CREBBP, CDKN2A*) were less frequently mutated.
- Three genes were identified to be recurrently mutated in BL and BL cell lines:
  - *ID3* was identified to harbor inactivating mutations and subsequent functional analyses point towards a tumor suppressive function in BL.
  - *SMARCA4* missense mutations recurrently affecting its catalytic domain were identified in BL.
  - *PCBP1* was identified to be recurrently mutated in BL likely leading to a loss-of-function.

In the following these major results are discussed.

4.1.1 Traces of pluripotency in Burkitt lymphoma

As already mentioned in 1.3.3, the *MYC*-translocation is the initiating event in the pathogenesis of BL and different theories exist at which stage of the B-cell development transformation might occur. One hypothesis has been that pediatric BL might derive from or resemble cells with – potential- stem cell features. To test this hypothesis, BL were analyzed herein for traces of pluripotency using the PluriTest.

The PluriTest showed that the B-cell lymphomas, regardless of their subtype, had in comparison to germ cell tumors (GCT) or in particular teratomas, which have pluripotent capacities [254], a strikingly lower pluripotency score and higher novelty score. But taking the *MYC*-status into account, mBL could be efficiently separated from *MYC*-negative non-mBL showing a higher pluripotency score and novelty score. The *MYC*-positive and –negative intermediate B-cell lymphomas and *MYC*-positive non-mBL intermingled between both indicating that the *MYC*-status alone does not explain the shift of the mBL towards hPSC. This might be due to the fact that in non-mBL the *MYC* translocation is a secondary event occurring during the progression of the tumor [255],[256].

Comparison of the expression of the 181 genes, which make up the pluripotency signature, between mBL and ESC revealed that about 41 % of these genes were expressed at similar or higher levels in mBL. hPSC showed just as mBL a rapid proliferation rate driven by certain genes [162],[257] probably leading to the described shared gene expression pattern between mBL and ESC which is in agreement with recent findings by Ben-Porath *et al.* [258]. Moreover, Kim *et al.* [259] have shown that the similarities of hPSC and some tumor-associated signatures are contributed by a regulatory network of the *MYC* gene. The findings that the major pluripotency genes like *NANOG* or *POU5F1* [260],[261] were lower expressed in mBL whereas genes related to increased cell cycle and proliferation were expressed at comparable levels or even higher in mBL, further indicates that the high similarity is based on a shared proliferation signature in the mBL and hPSC. Thus, it seems as if the pluripotency score and the novelty score of mBL is shifted towards hPSC due to a *MYC*-driven proliferation signature. Overall the results suggest that the cell in which the *IG-MYC* translocation is initiated, seems not derive from a B-lymphoid progenitor cell with stem-cell like features. This is further supported by the findings of the *IGH* breakpoint analyses in the BL variants, which suggest either a origin of the translocation in bone marrow precursor B-cells during VDJ rearrangement particularly in endemic BL [106],[107] or in mature germinal center B-cells during class switch recombination [104] or during somatic hypermutation in sporadic BL [262] as described in 1.3.3.1.

Taken together, the results summarized in this thesis show that mBL resemble more hPSC than non-mBL and intermediate B-cell lymphomas. Nevertheless, this finding could not be associated to an origin or resemblance of mBL to B-lymphoid progenitors with stem cell features.

4.1.2 Possible implications of the new exon of *TERT* gene for Burkitt lymphoma

Analysis of the transcriptome data of primary BL within the ICGC MMML-Seq project revealed a new, so far undescribed exon within the *TERT* gene lying between the second and third exon. The 5' and 3' exon borders as well as the expression of the three splice variants of this new exon were herein verified in a panel of eleven BL cell lines, as well as in seven primary BL. Moreover, the expression of a new, so far undescribed splice variant of exon 2 (exon 2V) could be confirmed. The possible implications of the expression of the alternative TERT transcripts for BL will be discussed in the following.

The *TERT* gene maps to chromosome 5p15 and harbors diverse functions. It is the catalytic subunit of the telomerase complex [140],[141]. In addition, the telomerase consists of a telomerase RNA (TERC) which directs the telomere synthesis [142]. The canonical function of the telomerase is the maintenance of the telomere ends [140],[141]. Furthermore, non-telomeric functions have been described in the last years which include amongst others regulation of gene expression [263],[264],[265],[266], regulation of the cell cycle [267] and inhibition of apoptosis [268],[269],[270]. With regard to these various cellular functions, oncogene-like features have been attributed to TERT likely contributing to malignant phenotypes [271]. This is supported by the findings that the vast majority of cancers express high levels of TERT [272],[273],[274] which is correlated with high telomerase activity [139]. In line with this, BL have been shown to have the highest telomerase activity among malignant B-cell lymphoma [139] which is reflected by a likely MYC-induced TERT expression [275],[276],[277]. Due to its high expression it is one of the genes defining the gene expression signature of BL [108]. In line with these findings, TERT expression was detectable in all BL cell lines as well as in all but one primary BL. Reviewing the available karyotype and SNP array data as well as the whole-genome sequencing data which were available within the ICGC MMML-Seq project failed to explain the lack of TERT expression in this single BL case.

As described in 3.1.2 overall six alternative TERT transcripts might occur due to the inclusion of the three variants of the new exon in combination with either exon 2 or exon 2V in the reading frame of the wildtype TERT transcripts. Prediction of the open reading frames (ORFs) showed that overall four different ORFs might occur, encoding proteins of which either two lack the N-terminal or the C-terminal domain of the wildtype TERT protein. But it should be noted that based on this modeling, it is not clear if these proteins are translated at all. The ORFs encoding proteins lacking the C-terminal domain harbor premature stop codons. Alternative splicing of TERT has been already reported n several studies leading to the detection of more than 20 splice variants [278],[279],[280],[281],[282]. Remarkably, most of them have a premature stop codon [283] making the transcripts more prone to NMD [284]. This is thought to be one of the mechanisms regulating TERT activity [285],[286]. In line with this, a recent study has shown that the TERT upregulation during lymphomagenesis is counter-

regulated by a switch of alternative splicing towards transcripts with premature stop codon subjected to NMD [287]. Thus, the inclusion of the new exon could also be a counter-regulating mechanism and the predicted ORFs lacking the C-terminal domain might be not translated at all. However, if transcripts with premature stop codons are dedicate to NMD, their expression is normally difficult to detect without inhibition of the NMD machinery as has been reported for other transcripts [288]. Thus, as the alternative TERT transcripts can be easily detected in the here studied cases, it is also conceivable that the particular high TERT expression in BL might overrun this negative feedback regulation. Alternatively, the alternative transcripts might not be subjected to NMD. In line with this, Wang *et al.* [289] have reported that high MYC expression inhibits NMD in B-cells and, hence, MYC-target genes were stabilized and upregulated. Hence, due to the high MYC expression in the BL cells, NMD might be suppressed and the alternative transcripts harboring a premature stop codon might be translated.

Thus, assuming that the N-terminal and the C-terminal truncated proteins are translated, different functions are conceivable which are shortly discussed in the following.

The C-terminal domain of TERT, which is lost in ORF1 and ORF2, contains two important domains: (i) the reverse transcriptase (RT) domain which is responsible for telomerase activity [290] and (ii) the TERT C-terminal extension (CTE) domain which participates in protein-protein or protein-DNA interactions and regulates the enzyme location [291],[292]. Furthermore, the T-motif is lost which is important for maintenance of enzymatic activity [293] by mediating TERT-TERC binding interaction and, thus, influences the elongation of the telomeres. Thus, TERT proteins lacking these domains have lost their catalytic activity which is needed for its diverse functions [294]. But the N-terminal TEN domain which is also called RNA interaction domain-1 (RID1) [295] is retained. This domain has been reported to bind single-stranded telomeric DNA [296],[297]. Various studies have analyzed the properties of TERT proteins lacking the C-terminus. Some have postulated a dominant-negative effect on wildtype-TERT function for example by competing for TERC binding [278],[298],[299]. Other studies have shown that TERT variants lacking the catalytic domain retain the ability to induce proliferation [279] or anti-apoptotic effects [278]. Thus, the C-terminal truncated proteins encoded by the ORF1 and 2 might counteract the function of the physiological TERT in a dominant-negative manner or might even have alternative functions.

The N-terminal truncated TERT proteins encoded by ORF3 and ORF4 retain the functionally important catalytic RT domain. But the above described TEN domain is lost and only half of the sequence of the TRBD domain is retained in those N-terminal truncated proteins. The TRBD domain has a high affinity for TERC and, hence, is important for the TERT-TERC interaction promoting a stable assembly of the telomerase complex [300],[301]. Thus, as the truncated protein cannot be directed to the telomeric DNA and the assembly with TERC is also likely impaired, it might have no function.

Nevertheless, to elucidate the function of the alternative TERT proteins in BL, further functional studies need to be performed.

Taken together, based on the existing data it is yet not possible to conclude if a truncated TERT protein is expressed. Furthermore, the function of these truncated proteins remain unclear although it is conceivable that the expression of the alternative transcript including the variants of the new exon might (i) be a mechanism to negatively regulate the TERT expression (NMD) or function (dominant negative effect) or (ii) contribute to wildtype TERT function mediating additional functions. Thus, further functional studies need to be conducted to unravel if the alternative transcripts are indeed translated and what kind of function they have in BL cells.

4.1.3 Analysis of the mutational landscape in Burkitt lymphoma

The first step in the analysis of the mutational landscape in BL was the identification of appropriate *in vitro* models for BL. Thus, a panel of twelve BL cell lines was screened for mutations using Sanger sequencing of two genes described to be recurrently mutated in BL (*TP53, CCND3*) as well as of two genes described to be recurrently mutated in DLBCL (*CREBBP and CDKN2A*). The results of these mutational analyses will be discussed in the following.

The transcription factor *TP53*, which is a tumor suppressor, has the highest mutation frequency of all analyzed genes by whole-genome sequencing among the 50 cancers analyzed within the ICGC project (https://dcc.icgc.org/, 15/01/2015). Inactivating mutations in *TP53*, which are in the majority missense mutations, have been published to occur in about 33 % of primary BL [72],[101],[170],[228]. Screening of the twelve BL cell lines revealed that 83 % carried *TP53* mutations of which the majority (12/13) were missense mutations. Various reasons for the discrepant *TP53* mutation frequencies between BL and BL cell lines are conceivable: (i) high passage time of cell lines might lead to acquisition of growth-advantage mutations [302]. (ii) Within the processing of the NGS data annotated SNPs as well as germline SNV were excluded in the primary BL. The latter was not possible in the cell lines due to the lack of germline material. (iii) DNA from primary BL were derived from pretreatment tissue whereas cell lines may derive from tumor tissues after treatment which might induce *TP53* mutations [303],[304] as in the case of DAUDI cells [178]. (iv) The acquisition of *TP53* mutation poses a growth advantage to the cells which might lead to a positive selection for culturing [125],[305]. Interestingly, the recurrent *TP53* mutations at positions 245, 248 and 273, described as damaging SNPs by PolyPhen and the dbSNP 141 database, correspond to six already defined “hot spot” residues which have been defined by Freed-Pastor and Prives [306] based on *TP53* missense mutation data of 25,902 human cancer patients. These hot spot mutations, as well as the other seven identified mutations all map within a conserved DNA-binding

domain of the TP53 protein [307]. The mutations affecting this domain have been shown to lead to a prolonged half-life of the protein [308],[309],[310]. TP53 is activated upon cellular stress like e.g. DNA damage leading to a negative regulation of the cell cycle [311] and induction of apoptosis [312],[313],[314]. Thus, a stable mutated-TP53 protein can exert a dominant-negative effect leading to an abrogation of the wildtype TP53 function in the cells [315]. The selective acquisition of *TP53* mutations is important for *MYC*-positive cancers like BL, since cells with a MYC overexpression, which represents already a cellular stress factor [316], are more prone to TP53-dependent apoptosis [317],[318].

The *CCND3* gene mapping to chromosome 6p21 has been identified to be recurrently mutated in primary BL, showing mutation frequencies between 24 % and 38 % [133],[170]. Those studies have shown that the mutations cluster in a region which is encoded by the fifth exon of *CCND3*. Four of twelve screened BL cell lines (33 %) carried *CCND3* mutations within this exon. The majority of these mutations led to a frameshift and a loss of the C-terminal domain of the protein which includes the phosphorylation motif. This is in concordance with the findings in primary BL [133],[170]. The only missense mutation detected within the BL cell lines affected the residue Ile290 which has been described to be part of the phosphorylation motif that regulates among others the protein stability [319],[320]. Thus, all four mutations identified in BL cell lines affected the phosphorylation site of the protein. *CCND3* encodes cyclin D3, which regulates the G1-S cell cycle transition by binding to the cyclin-dependent kinases CDK4 and CDK6 [321] in germinal center B-cells [322],[323]. In functional studies, Schmitz *et al.* [72],[135] have shown that *CCND3* mutations led to a stabilization of the protein and consequently to a longer half-life. Based on those findings they suggested that BL cell lines as well as primary BL likely rely on CCND3 signaling which is augmented upon activating CCND3 mutations [72]. Thus, the here identified *CCND3* mutations in BL cell lines harbor the same kind of activating mutations as the published primary BL.

CREBBP, mapping to chromosome 16p13.3, is recurrently mutated or deleted in DLBCL with frequencies between 8-41 % [324],[325],[326]. In contrast, *CREBBP* is barely (4 %) or not at all mutated in BL [72],[170]. Furthermore, screening of the CONAN database for copy number analyses in cell lines revealed that none of the herein studied BL cell lines carried CNA affecting the *CREBBP* locus (http://cancer.sanger.ac.uk/cell_lines/conan/search, 03/04/2015). Based on this, *CREBBP* mutations are a possible criterion to differentiate between BL and DLBCL. Among the BL cell lines screened for *CREBBP* mutations, only the DG-75 cell line (1/12, 8 %) carried a mutation (p.Ile1084Asn*2). The main function of CREEBP is the acetylation of histones [327],[328] which is important for the regulation of transcription [329],[330]. Half of the mutations identified in DLBCL affected the histone acetyltransferase (HAT) domain which has been described to reduce the activity of the protein [324]. The *CREBBP* mutation identified in DG-75 led to a loss of the functional HAT

domain of the protein. Anyhow, this does not implicate that DG-75 cells are DLBCL-like. *CREBBP* is a large gene with 31 exons spanning 155 Mb, hence, the probability for a mutational event increases. The probability is even higher in a cell line as DG-75 which has been shown to harbor microsatellite instability (http://cancer.sanger.ac.uk/cell_lines/, 03/04/2015). This is in line with the findings by Wagener *et al.* [331] who showed that the cells show a high mutation load which is associated with DNA MMR deficiency (Figure 38).

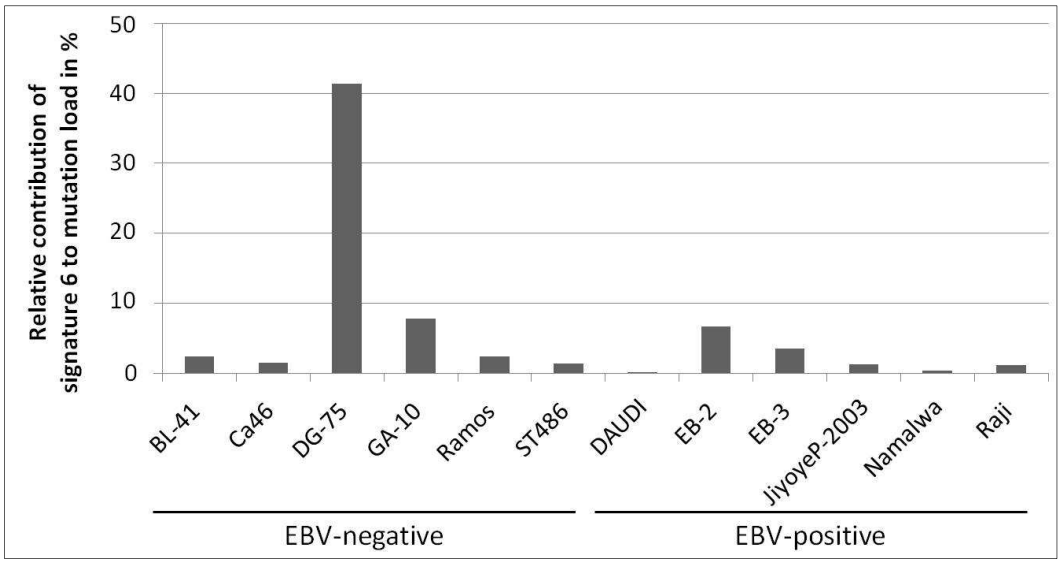


Figure 38: Relative contribution of signature 6 (DNA MMR deficiency associated) to the mutation load in 12 BL cell lines. The DG-75 cells have the highest mutation load associated with signature 6 in comparison to the other BL cell lines. Data from [331].

The *CDKN2A* gene maps to a tumor suppressor gene locus on chromosome 9p21 [332]. *CDKN2A* encodes two alternative proteins: (i) $p16^{INK4a}$ has been described to inhibit CDK4 and CDK6 [333],[334] which interact with CCND3 [321] and (ii) $p14^{ARF}$ which has been described to inhibit MDM2 which destabilizes the TP53 protein [335]. The gene is frequently altered in malignant lymphomas [336], especially in DLBCL [227] and FL [337]. Three different *CDKN2A* inactivating mechanisms have been described: (i) deletion, (ii) methylational silencing and (iii) inactivating mutations [338],[339]. Although the overwhelming majority of inactivating alterations in DLBCL and FL are attributed to deletion in 9p21 including *CDKN2A*, a small fraction of cases harbors inactivating mutations [72],[227],[337],[339]. In Burkitt lymphomas the major mechanism of *CDKN2A* inactivation has been attributed to epigenetic silencing [134],[340] whereas deletions and inactivating mutations occur at a much lower frequencies [72],[341]. In line with these findings, none of the here analyzed BL cell lines carried a *CDKN2A* mutation. Of note is that the BL-2 cell line carried a homozygous deletion of the *CDKN2A* locus as detected by exome sequencing. After screening of the above described CONAN database none of the other herein studied BL cell lines harbored CNA affecting the *CDKN2A* locus (http://cancer.sanger.ac.uk/cell_lines/conan/search, 03/04/2015).

Figure 39 gives a mechanistic overview on the impact of the mutated genes on the physiology of the BL.

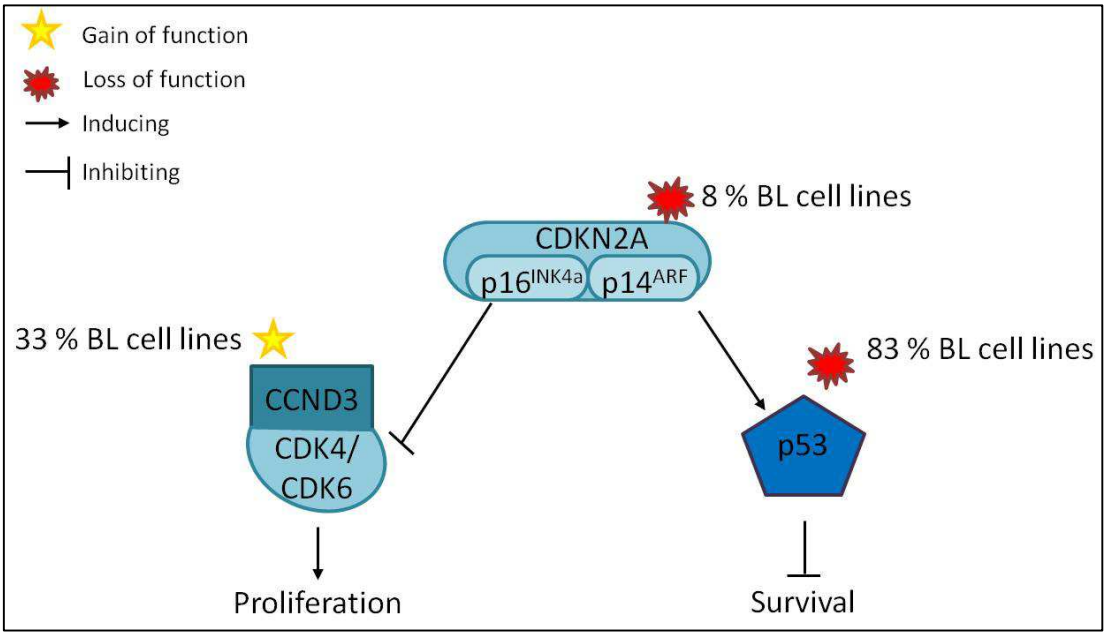


Figure 39: Overview on recurrently mutated *CCND3, TP53* and *CDKN2A* in Burkitt lymphoma and their impact on cellular pathways. Inactivating mutations or deletions in *CDKN2A* were detected in 8 % of BL cell lines. The gene encodes two alternative proteins: $p16^{INK4a}$ and $p14^{ARF}$. Inactivation of $p16^{INK4a}$ leads to a loss of CDK4 and CDK6 inhibition, which are bound to CCND3, thus, increasing the CCND3-CDK induced proliferation of the cell. This pro-proliferative pathway is further enhanced by activating *CCND3* mutations (33 % of BL cell lines) leading to a stabilization of the protein. On the other hand, loss of $p14^{ARF}$ leads to a decrease in TP53 protein stability reducing the TP53-induced apoptosis. Furthermore, BL and BL cell lines harbor recurrent inactivating mutations within *TP53* (83 % of BL cell lines).

The TP53 gene was the gene with the highest mutation frequency in BL cell lines (83 %). These mutations lead to an inactivation of the protein and, thus, a loss of TP53-induced apoptosis. The inactivation of TP53 can be supported by inactivation of the *CDKN2A*-encoded protein $p14^{ARF}$ which regulates via MDM2 the stabilization of TP53 protein. Sánchez-Beato *et al.* [341] have reported that *CDKN2A* and *TP53* alterations might occur mutually exclusive. These holds true for the BL cell lines investigated herein. Furthermore, loss of the *CDKN2A*-encoded protein $p16^{INK4a}$ leads to an increased activation of CCND3-CDK4/6 and subsequently higher proliferation activity. This effect is enhanced by activating mutations in *CCND3*, increasing the stability of the protein and fostering progression through the cell cycle.

To sum up, the mutational landscapes of the BL cell lines and primary BL with regard to the mutation frequencies and types within *TP53, CCND3, CREBBP* and *CDKN2A* resemble each other. Thus, the selected BL cell lines are appropriate *in vitro* models for further functional analyses of BL. Overall, the mutations in the analyzed genes, lead to an increase in proliferation and an inhibition of anti-survival signals via TP53.

4.1.3.1 Recurrency of *ID3* mutations in Burkitt lymphoma

The analyses of the frequency and the functional impact of *ID3* mutations on BL cell lines performed in this thesis were embedded within the published analysis of primary BL within the ICGC MMML-Seq project [170].

A total of 18 B-cell lymphoma cell lines were screened for *ID3* mutations. Of those, only BL cell lines carried *ID3* mutations (67 %, 8/12). This is in agreement with the findings in primary *MYC*-positive BL. *ID3* mutations were detected in 68 % of the Burkitt lymphomas of those [170]. The latter findings of our group were confined in two independent NGS studies in which 58 % [72] and 34 % [228] of BL were described to carry *ID3* mutations. In addition to the BL cell lines, only the MC-116 cell line of the B-cell lymphoma cell line panel (1/6, 16 %) which carried an *ID3* mutation. Remarkably, this cell line carries a *MYC*-translocation [177]. In line with this finding in the cell line, also six primary *MYC*-positive non-BL were identified to carry *ID3* mutations (13 %, 6/45), suggesting an association between *MYC*-translocation status and *ID3* mutation.

The *ID3* gene, consisting of two coding exons, is located on chromosome 1p36.12 and encodes the Inhibitor of DNA Binding 3, Dominant Negative Helix-Loop-Helix protein. This protein belongs to the Inhibitor of Differentiation (ID) family of helix-loop-helix (HLH) proteins [342] which consists of four members in mammals (ID1-4). Their main function is the regulation of proliferation and differentiation [343],[344],[345]. Some ID3 functions have been attributed to cell cycle regulation. Thus, ID3 has been described to induce growth arrest in B-lymphocyte progenitors [346]. One of its main functions is the antagonization of the DNA-binding capacity of the basic HLH (bHLH) proteins TCF3 and TCF4 through heterodimerization via its HLH domain [347],[348],[349]. *TCF3*, also known as *E2A*, plays an important role during B-cell lymphocyte development [350],[351] and is involved in the germinal center reaction (1.2.3). As far as yet known, *TCF4* (E2-2) on the other hand plays in comparison to *TCF3* a minor role in B-cell development [352]. Nevertheless, it is rather important for plasmacytoid dendritic cell development [353] and neurodevelopment [354],[355]. Generally the ID proteins are ubiquitously expressed [356]. However, within the hematopoietic tissue ID3 has been described to show a lymphoid-specific expression [357]. Within the tonsils, ID3 is higher expressed within the centroblasts of the dark zone than in the centrocytes of the light zone [135]. This pattern resembles that of TCF3 and suggests an important role in germinal center B-cell regulation. Remarkably, ID3 deficiency in a mouse model led to an induction of γδ T-cell lymphoma [358], indicating a tumor suppressive role for ID3.

ID3 mutations have yet been described in the literature mainly within B-cell lymphoma [359],[360]. But overexpression of ID3 has been observed in different tumor types like prostate cancer [361] and medulloblastoma [362]. In these tumors, ID3 overexpression was associated with poor prognosis. Strikingly, *ID3* is not only recurrently mutated in BL but is among those genes which are particularly

highly expressed in BL [108]. This strong expression might be induced by the high expression of the *MYC* oncogene in BL. Indeed, the ID3 promoter harbors a MYC binding site [363],[364]. In line with this, the ID3 protein expression was lower or non-detectable in non-BL cell lines which were *MYC*-negative; whereas the *IG-MYC* positive non-BL cell line Su-DHL-6 exhibited a low ID3 expression.

A total of 13 mutations were detected in eight BL cell lines. The majority of those mutations (8/13, 65 %) skewed towards deleterious changes caused by nonsense mutations, deletions or splice site mutations. The type of *ID3* mutations in primary BL described in the literature differed between the study cohorts: Richter *et al.* [170] have described an overall deleterious character of the mutations, whereas Schmitz *et al.* [72] and Love *et al.* [228] have described that the majority of *ID3* mutations in BL were missense mutations. This discrepancy can be in part explained by the fact that Schmitz *et al.* used RNA-sequencing to identify SNV. Thus, transcripts harboring a nonsense mutations or frameshift deletion/insertion leading to a premature stop codon are probably not detected with this technique due to nonsense mediated decay (NMD).

The ID3 protein expression analysis in the BL cell lines gave further evidence for the deleterious character of the *ID3* mutations. In five of the eight *ID3*-mutated BL cell lines (63 %) ID3 protein expression was lower in comparison to ID3-wildtype BL cell lines or completely absent. Two of the BL cell lines with complete loss of ID3 expression carried biallelic *ID3* mutations. Hence, it can be speculated that although the ID3 mRNA levels are increased in primary BL [108], the ID3 protein expression is also reduced in primary BL with biallelic *ID3* mutations.

Two BL cell lines (BL-2, BL-70) carried the same splice site mutation which leads to an aberrant splicing as shown for BL-70 cells. In line with this, aberrant splicing due to a splicing mutation could be also shown for one of the primary BL samples [170]. Protein expression analysis showed a loss of ID3 expression in BL-70 cells, pointing towards NMD of the aberrantly spliced mRNA. Thus, these findings are in line with the overall deleterious character of the mutations.

In addition to the deleterious mutations, five missense mutations were identified in the BL cell lines. All of them were located within the HLH domain of the ID3 protein. The HLH domain mediates the protein homo- and heterodimerization of ID proteins [356]. The accumulation of missense mutations in the HLH domain has been also observed in the primary BL [72],[170], [228]. Furthermore, Richter *et al.* have shown that several of those mutations were predicted to affect the interaction of ID3 with TCF3 and/or TCF4 [170]. This might lead to an increased activity of TCF3 and maybe TCF4, giving the BL a selective growth and survival advantage. This hypothesis was corroborated by functional experiments in this thesis. These showed that overexpression of wildtype-ID3 in *ID3*-unmutated (DG-75) and *ID3*-mutated (BL-2) BL cell lines led to a significant increase of apoptosis. This is in line with the findings by Love *et al.* [228] who showed that an overexpression of wildtype-ID3 in BL cell lines

reduced the cell viability whereas overexpression of mutant-ID3 did not affect the viability. These findings further support a possible tumor-suppressor function for ID3 in BL.

However, although *ID3* mutations have a high frequency of 34-68 % in BL, not all primary BL and BL cell lines carry mutations in this gene. Interestingly, Schmitz *et al.* have shown that 11 % of BL harbor activating mutations within *TCF3* [72]. Mutations of *TCF3* have meanwhile also been detected in follow up analysis of more primary BL (23 %, 3/13) within the ICGC MMML-Seq project (https://dcc.icgc.org/, 15/01/2015). Thus, up to 70 % of BL harbor either inactivating *ID3* or activating *TCF3* mutations [72]. In line, the BL cell lines DG-75 and Ramos, lacking *ID3* mutations, carried *TCF3* mutations (p.Asp561Val and p.Asp561Glu, respectively) [72]. Remarkably, in both cell lines the mutation affected a residue in the bHLH domain which Schmitz *et al.* described as evolutionary conserved [72]. Moreover, Richter *et al.* [170] described that 28 % of BL without an *ID3* mutation harbored a copy number gain in 18q. A minimal region of gain in chr18:51,894,728–54,354,319 bp (hg19) including the *TCF4* gene could be defined for those cases which was, furthermore, associated with a higher expression of the *TCF4* gene [170]. As TCF4 has been described to share targets with TCF3 [365], it is conceivable that the overexpression of TCF4 is another activating mechanism in BL.

In further experimental studies, Schmitz *et al.* have shown that loss of TCF3 inactivation by ID3 and/or activating *TCF3* mutations led to a TCF3-dependent activation of the tonic BCR signalling and subsequently activation of the PI3K pathway, increasing the proliferation rate of the cells [72]. As mentioned in the introduction (1.3.4), the PI3K pathway plays an important role as survival determinant in BL [133]. This view has been further confirmed by specific inhibition of the PI3K in BL cell lines with the clinical trial compound BKM120 [72]. The loss of viability of BL cell lines due to BKM120 treatment was confirmed in functional experiments on the same BL cell lines used by Schmitz *et al.* [72] in the framework of this thesis (Figure 36D).

To sum up, a high frequency of inactivating mutations within *ID3* is specific for BL and points towards a tumor-suppressor function of this gene. The loss of ID3 function increases TCF3 activity and, thus, PI3K signaling which ultimately leads to an enforced proliferation of the cells. This is schematically summarized in Figure 40.

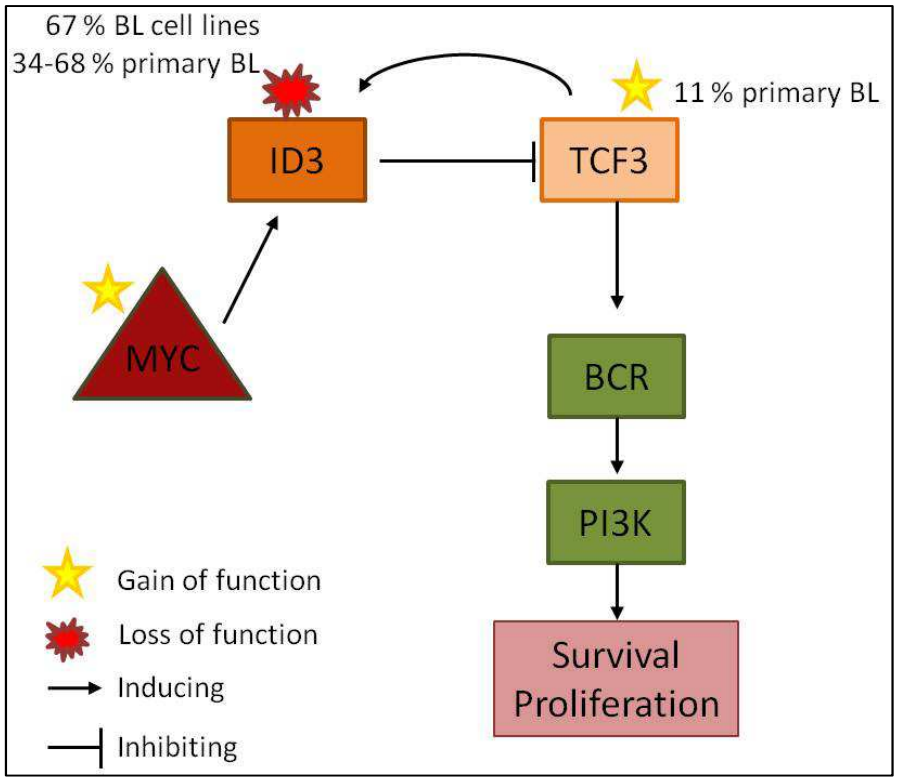


Figure 40: Schematic overview on the aberrations in the ID3-TCF3-PI3K signaling pathway in BL. ID3 is particularly highly expressed in BL which is likely induced by the high MYC oncogene expression in BL. Two mutational events lead to an increase of TCF3 activity: (i) inactivating mutations of *ID3* as detected in 34-68 % of primary BL and 67 % of BL cell lines ablate its negative control of TCF3 and (ii) activating *TCF3* mutations were detected in 11 % of primary BL [72]. Higher activity of TCF3 augments BCR signaling which subsequently enforces PI3K signaling, leading to increased survival and proliferation of the BL cells. Adapted from Ott *et al.* [71].

4.1.3.2 Recurrency of *SMARCA4* mutations in Burkitt lymphoma

In the following the results of the *SMARCA4* mutation analysis in BL, which were embedded within the ICGC MMML-Seq analyses of primary BL, will be discussed. SMARCA4 is of special interest in BL as it is, just as ID3 (4.1.3.1), among those genes which are particularly highly expressed in BL [108]. This strong expression might be induced by the MYC oncogene [366].

The aim of this study was to analyze the functional impact of the *SMARCA4* mutations since so far only the fact that these mutations recurrently occur in BL has been described in literature [72],[170],[228].

SMARCA4 maps to chromosome 19p13.2. Mutations in *SMARCA4* have been observed in 43 % (9/21) of primary BL within our ICGC MMML-Seq cohort and in 13 % of BL cell lines (2/16). To enlarge the cohort of BL, exome and RNA-sequencing data of two independent cohorts [72],[228] were included for the characterization of the *SMARCA4* mutations. The overall *SMARCA4* mutation frequency in BL in all three cohorts was 27 % (27/100). The mutation frequencies differed between the cohorts (ICGC MMML-Seq: 43 %, Schmitz: 39 %, Love: 16 %) which might be due to the different sequencing methods applied (whole-genome sequencing vs. RNA-sequencing vs. exome sequencing). Another reason might be the selection of the cohorts: all analyzed primary BL of the ICGC cohort were derived

from children (median age 6 years at diagnosis), whereas in the studies of Love *et al.* and Schmitz *et al.* also BL from adults were included [72],[228]. Furthermore, the ICGC cohort consists of molecular defined BL hereby excluding lymphomas with morphological features of BL but not representing BL at the molecular level like e.g. some double-hit lymphomas [367],[368].

SMARCA4 has been described as one of 15 genes being frequently mutated in a set of over 1,000 pediatric cancers [369]. Furthermore, *SMARCA4* mutations have been described in Atypical Teratoid/Rhabdoid Tumor (AT/RT) which primarily affect young children [193],[370] and in small cell carcinoma of the ovary, hypercalcemic type (SCCOHT) occurring in woman under 40 years of age [371]. The *SMARCA4* mutations of both tumor types were predominantly nonsense mutations or deletions leading to a loss of the C-terminal domains and SMARCA4 expression [370],[371]. Moreover, *SMARCA4* germline mutations have been reported in 4/7 families with SCCOHT and in 6/7 children with AT/RT [370],[371]. On the contrary, the *SMARCA4* mutations identified in BL were all somatic and in the overwhelming majority missense mutations (97 %). Thus, although those tumors share the high frequency of *SMARCA4* mutations, the functional impact likely varies due to the different types of mutations.

The *SMARCA4* gene encodes the Brahma-related gene-1 (BRG1) protein which is the catalytic subunit of the ATP-dependent chromatin remodeler called SWI/SNF complex [372]. Thus, it confers the DNA-dependent ATPase activity of the complex. Binding of SWI/SNF to the nucleosome changes the conformation at the transcription start site in such a manner, that the RNA polymerase II can bind and initiate the transcription of the gene [373],[374]. In addition to BRG-1, the SWI/SNF complex contains 8-10 subunits which are referred to as BRG-1 associated factors (BAF) [375],[376]. Among them are e.g. ARID1A (BAF250A) and ARID1B (BAF250B). The BAF subunits contain DNA-binding and protein binding motifs which influence the selectivity and activity of the ATPase subunits [377]. BRG1 itself contains a Bromo domain which binds acetylated histones [378] and is necessary for the stable association of the complex with the chromatin [379]. Furthermore, interaction between the SWI/SNF complex and gene-specific transcription factors directs the complex to the target genes [380],[381],[382]. The BRG1 SWI/SNF complex is a master regulator of gene expression as it has been shown to bind to ~51,000 regions in embryonic mouse tissues [383] as well as about 15,000 islands across the genome in murine embryonic stem cells [384]. In line with this, BRG1 has been described to associate for example with the *p21* promoter in a p53-dependent manner [385],[386] and, thus, plays an important role in cell cycle control. Furthermore, BRG1 has been reported to modulate TERT expression [387]. A further function of BRG1 is the regulation of differentiation of diverse tissues and at various developmental stages [388],[389],[390]. BRG1 has been also linked to B-cell development and function. Hence, as it has been described to play a role in the V(D)J recombination and BCR

assembly [391],[392]. Moreover, it has been recorded to be required for B-cell proliferation upon activation and, thus, to increase the efficiency of BCR signaling [393].

Furthermore, different studies have shown a tumor-suppressor function for various subunits of the SWI/SNF complex including BRG1 [385],[394]. Hence, BRG1 has been shown to bind to the retinoblastoma tumor suppressor protein (RB) and consequently control the cell cycle progression [394]. Moreover, BRG1 re-expression in BRG1-mutant or deficient cell lines inhibits their growth [385]. To elucidate the impact of the *SMARCA4* mutations on its functions, the sites of mutations were analyzed on protein level as discussed in the following.

Strikingly, 62 % of the missense mutations lay within the SNF_N or Helicase domains which together build the ATPase domain of BRG1 [372]. Different studies have revealed that mutations within the ATPase domain might impair the function of the protein [395]. For example, Bultman *et al.* [395] have shown that the mutation Glu1083Gly in the murine *Brg1* gene does not alter the stability of the protein, its capability to form the SWI/SNF complex or the ATPase activity. Instead, the mutation led to a reduced transcription of the target gene β-globulin, as well as to a methylation of the CpGs. In contrast, in BL the genomic targets of SMARCA4 are hypomethylated but an increase of SMARCA4 target genes transcription was not observed (data not shown, manuscript under revision at *Nature Genetics*). In another study it has been described that a missense mutation within the ATPase domain of BRG1 leads to a dominant negative inhibition of transcription due to the formation of a nonfunctional complex [372]. Based on this, it is conceivable that the *SMARCA4* mutations identified in the BL ablate the ATPase activity and, thus, its function, but the ability of the complex to interact with the DNA is still present. The protein modeling of SMARCA4 and its mutations further corroborates this (Figure 41). The results suggest that the helicase function of the protein is impaired either due to direct interference with ATP-binding or obstruction of the interaction of the N-and C-terminal helicase domains (data provided by Dr. rer. nat. Robert B. Russell, Cell Networks, Bioquant, University of Heidelberg, Heidelberg, Germany).

In comparison to *ID3*, the mutation frequency of *SMARCA4* is lower though a significant amount of cases is affected. Nevertheless, also other members of the SWI/SNF complex have been described to be recurrently mutated in BL. Giulino-Roth *et al.* [91] have shown that 17 % (5/29) of primary pediatric BL carry truncating *ARID1A* mutations and one BL harbored a mutation within *SMARCB1* (BAF47). In line with this, eight of 12 primary BL of the ICGC MMML-Seq cohort without *SMARCA4* mutation carried mutations within subunits of the SWI/SNF complex as in *ARID1A* (17 %, 2/12), *ARID1B* (8 %, 1/12) and *SMARCB1* (8 %, 1/12) or in newly identified subunits of SWI/SNF complex [396] as *BCL7A* (17 %, 2/12) and *BCL11B* (8 %, 1/12). Thus, 81 % of BL (17/21) carried mutations within the SWI/SNF complex (data not shown, manuscript under revision at *Nature Genetics*). This

suggests that the inactivation of this complex might confer a selective pathogenic advantage for the BL cells.

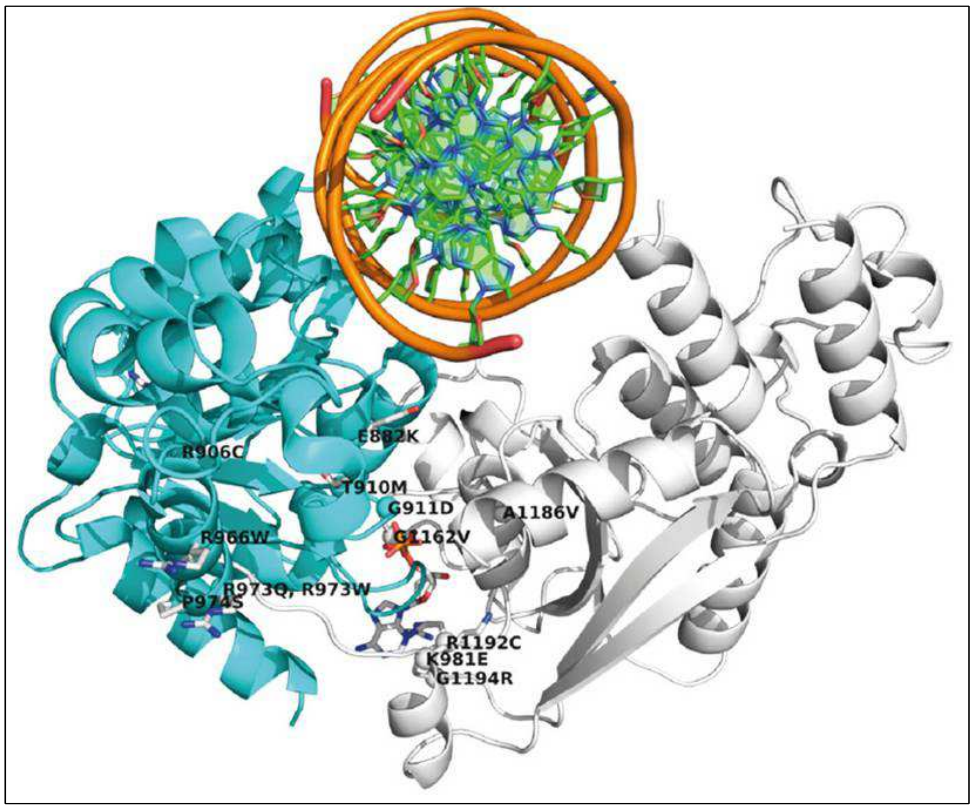


Figure 41: Protein model of SMARCA4 showing sites of mutations in BL. The N-terminal helicase domain (SNF2_N) is shown in cyan and the conserved C-terminal helicase domain (Helicase_C) is shown in grey. DNA is depicted as a cartoon and ATP with sticks, both colored by atom types. Residues mutated in BL are shown as spheres for C-alpha atoms, with side-chains as sticks colored by atom types. The model was provided by Dr. rer. nat. Robert Russell Cell Networks, Bioquant, University of Heidelberg, Heidelberg, Germany and is also included in Kretzmer *et al.*, *Nature Genetics*, under revision.

Taken together, 43 % of primary BL were identified using whole-genome sequencing and verified using Sanger sequencing to carry *SMARCA4* mutations. The overwhelming majority of mutations were missense mutations which likely lead to an impaired ATPase function and, thus, might have a dominant negative effect on the transcription of the SMARCA4 target genes. Hence, the mutations further corroborate the published function of SMARCA4 within the SWI/SNF complex as tumor suppressor.

4.1.3.3 Recurrency of *PCBP1* mutations in Burkitt lymphoma

Using whole-genome sequencing in a cohort of pediatric Burkitt lymphoma several genes were identified which are mutated at lower frequencies (9 and 20 %) than those discussed before. Among those genes is *PCBP1* which was mutated in 3 of 17 (18 %) whole-genome sequenced BL.

After verification of the mutations and their somatic origin using Sanger sequencing, an independent cohort of 28 primary BL, as well as 16 BL cell lines was screened for mutations using Sanger Sequencing. In total, *PCBP1* mutations were identified in 13 % (6/45) of primary BL and in 38 % (6/16)

of BL cell lines. Of note is, that five of those cell lines (BL-30, BL-70, DAUDI, EB-3, Namalwa) have been already described in the literature to carry *PCBP1* mutations [72],[228], thus proving the correctness of the here presented data. Mining of the data of the two published mutation screens in BL revealed that *PCBP1* mutations have been observed at lower frequencies by Schmitz *et al.* [72] (2/28, 7 %) and by Love *et al.* [228] (2/51, 4 %) but were not further discussed in those studies. The lower frequency might be again explained by the different sequencing methods used (RNA- vs. whole-genome- sequencing) as well as by the composition of the cohorts.

In comparison to the more frequently mutated genes like *ID3* or *CCND3*, the significance of genes mutated at a lower frequency is unclear. Nevertheless, a recent meta-analysis of pan-cancer exome data showed that most cancer related genes are mutated at frequencies between 2 and 20 % [397]. It is conceivable that genes with a low mutation frequency contribute to the pathogenesis of the tumor by interplay with other mutations deregulating the same pathway. In line with this, *TCF3* has been reported to be mutated in only 11 % of BL. Nevertheless, together with the above described *ID3* mutations, deregulation of the ID3-TCF3-PI3K pathway occurs in about 70 % BL [72]. Furthermore, *SMARCA4* mutations occur in 43 % of BL but overall 81 % of BL carry mutations within subunits of the SWI/SNF complex (4.1.3.2). Moreover, the *ras homology family member A* (*RHOA*) gene has been described to be mutated in 8.5 % of pediatric BL [398]. Recently, the Gα13-dependent pathway signaling which also includes RHOA has been reported by Muppidi *et al.* [399] to be deregulated in BL. Frequent mutations in components of this pathway as *GNA13* (7/31, 22.6 %) and *P2RY8* (10/31, 32.3 %) have been reported [399]. The Gα13 pathway inhibits the cell growth and the migration of GC B-cells [399]. Its deregulation has been reported to induce lymphoma formation in a mouse model [400]. Thus, as the above presented examples demonstrate, it is conceivable that *PCBP1* mutations also contribute to the lymphomagenesis of BL by interplay with other mutations.

PCBP1, also called heterogeneous nuclear ribonucleoprotein E1 (hnRNP-E1) or Alpha-CP1, is a member of the poly(C)-binding proteins. It has a high and specific affinity for binding C-rich oligonucleotide sequences. The protein consists of three KH domains (KH), which act as independent RNA/DNA binding units [401],[402], as well as of two nuclear localization signals (NLS) [403]. The first NLS lies between the second and third KH, whereas the second NLS lies within the KH III. The major functions of PCBP1 can be attributed to transcriptional regulation of gene expression, translational control and modulation or pre-mRNA splicing. PCBP1 controls translation by binding to mRNA templates containing differentiation control elements (DICE) [404]. The silencing of pseudoexon splicing as part of the pre-spliceosomal complex [405] as well as the transcriptional activation of the eukaryotic translation initiation factor 4E (eIF4E) [404] were also attributed to PCBP1. Furthermore, PCBP1 has been shown to regulate the expression of DNA polymerase eta (POLH) by stabilizing its mRNA [406].

PCBP1 protein is localized in the cytoplasm as well as in the nucleus [404]. In agreement with PCBP1's function in mRNA splicing, it has been also described to accumulate in nuclear speckles [403],[407] in which splicing factors are generally concentrated [408].

In this study a total of 12 *PCBP1* mutations were observed. Ten of those affected the KH III domain either due to a complete loss of the domain (5/10) or due to a substitution (5/10). Moreover, seven mutations affected the NLS I and/or NLS II due to a complete loss of both (5/7) or a substitution (2/7) in one of the NLS. The functional impact of the mutations might vary. Nevertheless, the nonsense and frameshift mutations leading to loss of functional domains are probably more damaging than the substitutions. For example the nonsense mutation in case 4177434, leading to a loss of KH III domain and both NLS, probably completely impairs the nuclear trafficking. In contrast, the missense mutation of BL-41 in NLS I might slightly decrease the PCBP1 expression within the nucleus. This is supported by the findings by Chkheidze and Liebhaber [403] showing that the depletion of only one NLS signal leads to the retention of some of the protein in the cytoplasm whereas the depletion of both NLS leads to a loss of nuclear trafficking.

Additionally, four missense mutations were predicted by Mechismo [197] to have an effect on the oligonucleotide interaction. Hence, the mutation p.Arg325Gly was predicted to have a strong disabling effect on oligonucleotides binding. The substitutions p.Cys293Tyr, p.Ile295Thr and p.Asn303Asp on the other hand, were predicted to have weaker effects. Thus, the missense mutations might also lead to a loss-of-function. Strikingly, the mutation p.Cys293Tyr, which was found in MPI-017 as well as in the BL-70 cell line, has been also described in each one BL sample by Schmitz *et al.* [72] and by Love *et al.* [228].

Furthermore, the available clinico-pathological and molecular cytogenetic data, which were provided by the ICGC MMML-Seq and the MMML-project, were analyzed for differences in the *PCBP1* mutated and wildtype BL groups. Interestingly, all six mutated cases expressed IRF4, whereas only 39 % of the *PCBP1* wildtype cases expressed IRF4 (*PCBP1* mutated vs. wildtype p=0.008). In the literature, IRF4 expression is detected in 20-40 % of BL [409],[410]. IRF4, also known as MUM1, belongs to the interferon regulating factor (IRF) family of transcription factors. The expression of IRF4 is restricted to the immune system [411],[412] and it is a crucial regulatory factor in B-cell development [413]. Low levels of IRF4 expression have been described in the germinal center B-cells [414] whereas within the terminal differentiated plasma cells higher expression levels have been described [415]. The IRF4/MUM1 positive immunoprofile of the BL might reflect a pathological intermediate GC stage of differentiation [409]. This is further supported by the finding that those cases are also positive for the GC marker BCL6 (8.9). The implications of *PCBP1* mutations in MUM1 expressing Burkitt lymphoma are unclear at present and the findings need further confirmation in independent cohorts.

To sum up, *PCBP1* was found to be mutated in 13 % of primary BL. The mutations primarily affect the KH III domain and/or the NLS. The consequences of the mutations point to a reduction or even loss of function. Furthermore, *PCBP1* mutations might accumulate in IRF4/MUM1 positive BL.

4.1.3.4 Conclusions from the analyses of the mutational landscape in Burkitt lymphoma

The aim of the analyses of the mutational landscape in BL was to identify recurrently mutated genes in primary BL and to determine the consequences of the mutated genes for Burkitt lymphomagenesis. The mutation screening was also performed in BL cell lines to identify appropriate cell line models for BL for further functional studies. The following figure summarizes the comparison of the mutation frequencies of the here analyzed genes in BL cell lines to the mutation frequencies determined in the framework of this thesis or published elsewhere for primary BL (Figure 42).

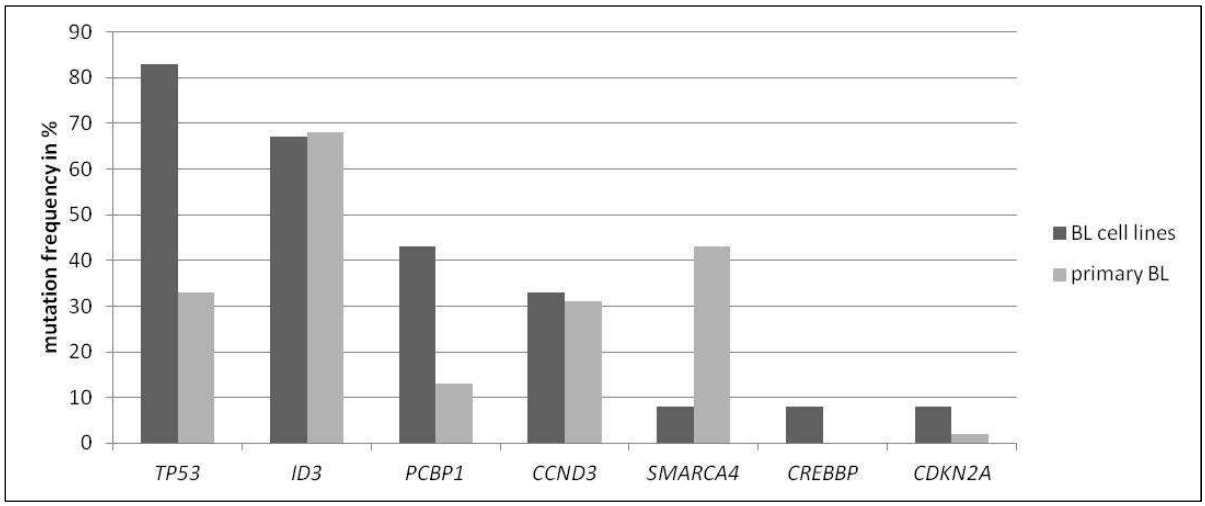


Figure 42: Comparison of mutation frequencies in seven genes detected in BL cell lines and primary BL. The mutation frequency of *ID3*, *PCBP1*, *SMARCA4* of the primary BL were determined within this thesis, whereas the mutation frequencies of *TP53* [72],[101],[170],[228], *CCND3* [133], [170], *CREBBP* [72],[170] and *CDKN2A* [72],[341] have been already published.

Overall, all genes mutated in primary BL are also mutated in the BL cell lines. Nevertheless, the mutation frequencies of *TP53* and *PCBP1* are higher in the BL cell lines than in the BL whereas the mutation frequency of *SMARCA4* is lower in the BL cell lines. A higher mutation frequency in the BL cell lines, as already discussed in 4.1.3, might be related to (i) the high passage number of the cell lines favoring acquisition of mutations which give an growth advantage as for example inactivation of *TP53* (ii) treatment-induced mutations in BL cell lines and (iii) as in the case of *TP53* inclusion of mutations designated as SNP in BL cell lines which were excluded in the processing of the whole-genome data. In case of *TP53* mutations, exclusion of the mutations annotated as SNPs, decreases the frequency to 42 % which is comparable to the mutation frequency observed in primary BL. But it is important to include these mutations annotated as SNPs as they have been described to have dominant negative effects as described in 4.1.3. Furthermore, the type of identified mutations is

consistent between the cell lines and the primary BL. Hence, taken together, the selected cell lines seem to be appropriate *in vitro* models for BL.

Remarkably, some of those genes which were identified to be recurrently mutated in BL have been described to interact with each other leading to a mechanistic model of BL lymphomagenesis. Figure 43 illustrates the interaction of the recurrently mutated genes in BL and the involved oncogenic pathways.

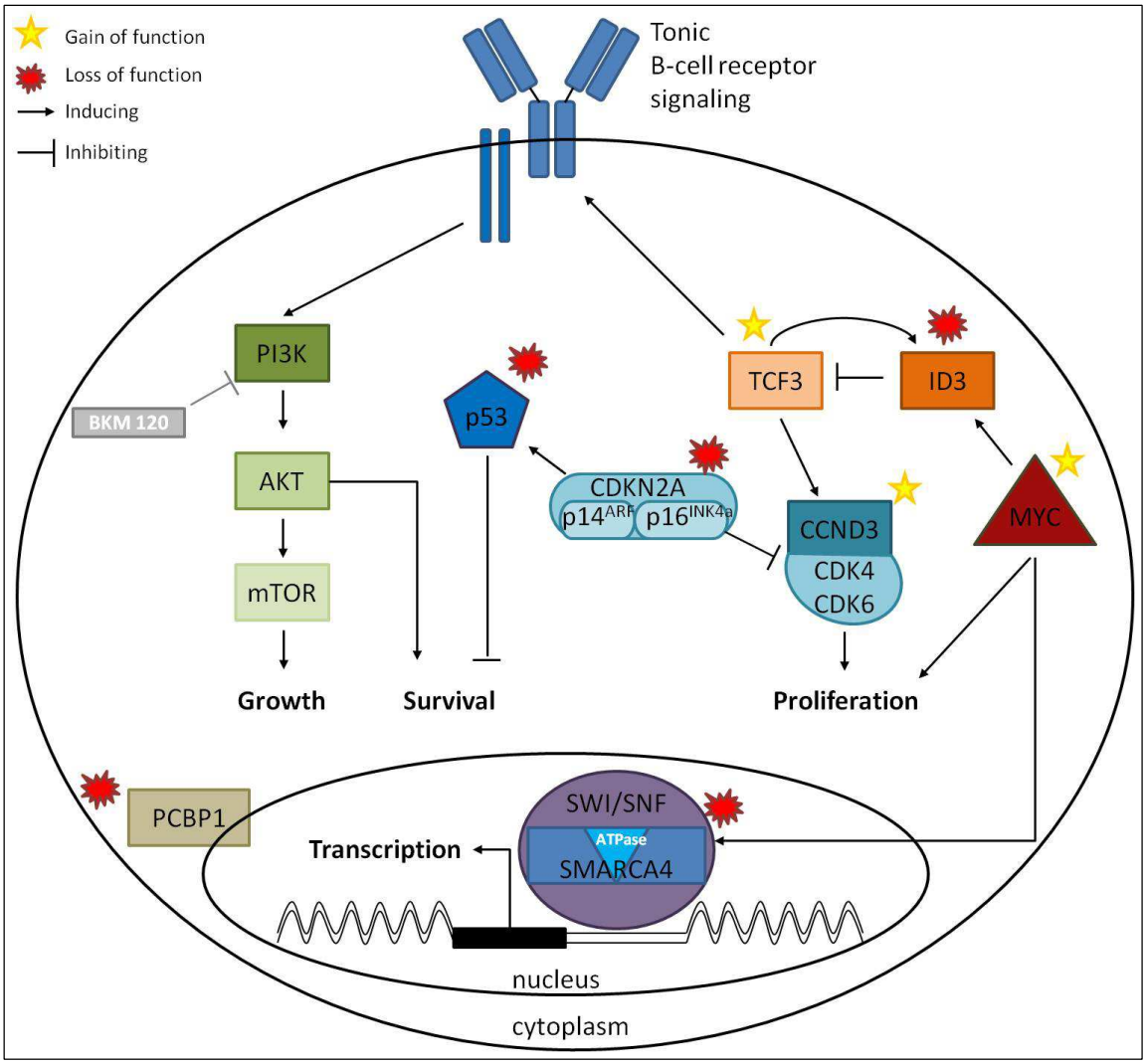


Figure 43: Overview on recurrently altered oncogenic pathways in Burkitt lymphoma. Four oncogenic pathways can be defined: 1. Inactivation of *TP53* leads to a decrease of TP53-dependent apoptosis, which is further augmented by the finding of recurrent inactivation of *CDKN2A* encoding $p14^{ARF}$ which regulates the stability of the TP53 protein. 2. Activation of CCND3 mediated proliferation via CDK4 and CDK6 is enhanced due to *CCND3* activating mutations. Furthermore *CDKN2A* also encodes the CDK4 and CDK6 regulating protein $p16^{INK4a}$ and its inactivation leads to an augmentation of the pro-proliferative pathway. 3. Inactivating mutations in *ID3* lead to a loss of negative regulation of TCF3 which has been furthermore described to carry activating mutations [72]. TCF3 has been described to activate CCND3, thus, contributing to the proliferative pathway. Furthermore, it enhances BCR signaling via various mechanisms, which leads to increased PI3K/AKT signaling and, hence, to the survival and growth of the BL cells. 4. Inactivation of the MYC induced SWI/SNF catalytic subunit SMARCA4 due to mutations within the ATPase domain likely lead to a dominant negative effect with a subsequent decrease of SMARCA4 target gene transcription. The effect of the likely inactivating *PCBP1* mutations is unclear to the present. But since the NLS domain of the protein is a major target of the mutations, an impaired trafficking to the nucleus is conceivable.

Strikingly, based on the analysis of the mutational landscape four major oncogenic pathways could be identified, which constitute promising targets for BL therapy. Among those targets is PI3K A specific inhibitor for PI3K is the compound BKM120 which is recently applied in several clinical trials [252],[416]. Schmitz *et al.* [72] could show that BL cell lines treated with BKM120 undergo cell death. This effect on BL cell lines could be reproduced in functional studies conducted within this thesis as described in 3.2.3.4.3. Hence, BKM120 is probably a portent BL therapeutic. A further promising target for BL therapy might be the cyclin-dependent kinases CDK4 and CDK6 which mediate proliferation. A potent inhibitor of the cyclin dependent kinases is the compound PD0332991 which has been shown by Schmitz *et al.* [72] to effectively block the G_1-S-phase transition in BL cell lines. Moreover, this inhibitor has proven in several clinical trials its clinical benefit in the treatment of refractory NHL and multiple myeloma [417],[418]. Moreover, the Akt-downstream component, mTOR, might be a further target for BL therapy. A potent inhibitor of the mTOR complex is rapamycin and its derivates which have been shown in several clinical trials to be potent lymphoma therapeutics [419]. In line with this, treatment with rapamycin has been described to be toxic to most of the BL cell lines [72] or to decrease the tumor growth in a transgenic mouse model of BL [420].

4.2 *MYC*-negative high-grade lymphoma resembling Burkitt lymphoma

The existence of rare *MYC*-negative Burkitt lymphoma or Burkitt-like lymphoma has been widely discussed for a long time [1],[164],[169]. Recently, by using a comprehensive genomics and gene expression approach our group identified a series of mBL and intermediate B-cell lymphoma lacking a *MYC*-translocation but sharing a peculiar pattern of chromosomal gain in 11q23.3 and loss in 11q24.2-q25 [166]. The aims of this thesis were the analyses of the chromosomal, mutational and transcriptional landscape of these *MYC*-negative Burkitt like lymphoma (mnBLL) in order to identify potential candidate genes which might contribute to its lymphomagenesis.

Major results of these analyses were:

- The minimal region of loss and the included minimal region of homozygous loss were redefined by analysis of newly recruited mnBLL. Moreover, recurrent secondary alterations, present in more than 20 % of mnBLL affecting 6q, 13q and 18q as well as a trisomy 12 were identified.
- *ETS1* is recurrently mutated in 4/16 (25 %) mnBLL.
- Ten genes of the altered 11q region were identified to be differentially expressed in mnBLL in comparison to BL and DLBCL.

- Combining the results of the analysis of the chromosomal, mutational and transcriptional landscape with literature search led to the identification of *PAFAH1B2, IL10RA* and *KMT2A* as potential oncogenes and of *FLI1* and ETS1 as potential tumor suppressor genes in mnBLL.
- Expression analysis of the five candidate genes on mRNA and protein level in mnBLL cell lines were in agreement with the expression profiles in the mnBLL.
- Analysis of the IL10RA cell surface and extracellular IL10 expression in combination with inhibition of IL10, IL10RA and PI3K revealed the dependency of the mnBLL cell lines on PI3K signaling and suggests an oncogenic role for the IL10 signaling pathway.

Hereinafter, the results of the follow up analyses with regard to the already published data in *Blood* [166] will be discussed.

4.2.1 The chromosomal landscape of *MYC*-negative Burkitt-like lymphomas

We described 17 *MYC*-negative Burkitt-like lymphomas sharing a special pattern of alteration on chromosome 11 consisting of a copy number gain (11q23.3) and a telomeric copy number loss (11q24.2-qter) [166]. In this thesis eight additional cases (case 18-25) were characterized. Moreover, in addition to the two already described cell lines with the typical mnBLL features (HT, Su-DHL-5) [166], the MLMA cell line was identified as an additional mnBLL cell line model. In order to identify potential pathogenetically relevant candidate genes, the aim of the analysis of the chromosomal landscape of mnBLL was the definition of the minimal gain and loss region in 11q as well as the identification of recurrent secondary alterations for all 25 mnBLL cases.

Figure 44 gives an overview on the alterations of chromosome 11 in the mnBLL as well as in three mnBLL cell lines. Combining the breakpoint data in 11q of all 25 cases led to a re-definition of the minimal region of loss and homozygous loss, whereas the already described minimal regions of gain and amplification remained unchanged. The occurrence of imbalances on chromosome 11 as 11q23 duplications or der(11q) have been described in *MYC*-negative lymphomas with Burkitt morphology [168],[169]. Furthermore, alterations on chromosome 11q23-qter have been described in a variety of hematological disorders, especially translocations involving the *KMT2A* gene on 11q23 [421],[422]. In DLBCL a gain in 11q22.3-qter or 11q24.3 is among the most frequent chromosomal rearrangements [239],[245]. But in contrast to the mnBLL these DLBCL harbor solely a copy number gain but nor loss in 11q.

As all analyzed mnBLL harbor at least the deletion on 11q24.3, this region might contain a tumor suppressor gene(s) important for its pathogenesis. Remarkably, two mnBLL (case 9 and case 25) harbor a homozygous deletion of 11q24.3 including seven genes (8.12) and one ncRNA (8.13). Strikingly, among those genes are two transcription factors of the ETS-family, *ETS1* and *FLI1*, which are known to play important roles in the pathogenesis of various cancers [423],[424],[425]. In

contrast to the here assumed tumor suppressive function, an oncogenic role has been described for those two genes in a subset of DLBCL [239]. In this subset of DLBCL, *FLI1* and *ETS1* map to a minimal region of gain [239]. A tumor suppressive function for ETS1 has been already suggested for Hodgkin lymphoma [248],[426]. This points towards an ambivalent role for ETS1 in tumors, having an oncogenic role in DLBCL but a tumor suppressive in mnBLL. The potential impact of those two genes on the pathogenesis of mnBLL will be discussed in further detail in 4.2.4.3 and 4.2.4.4.

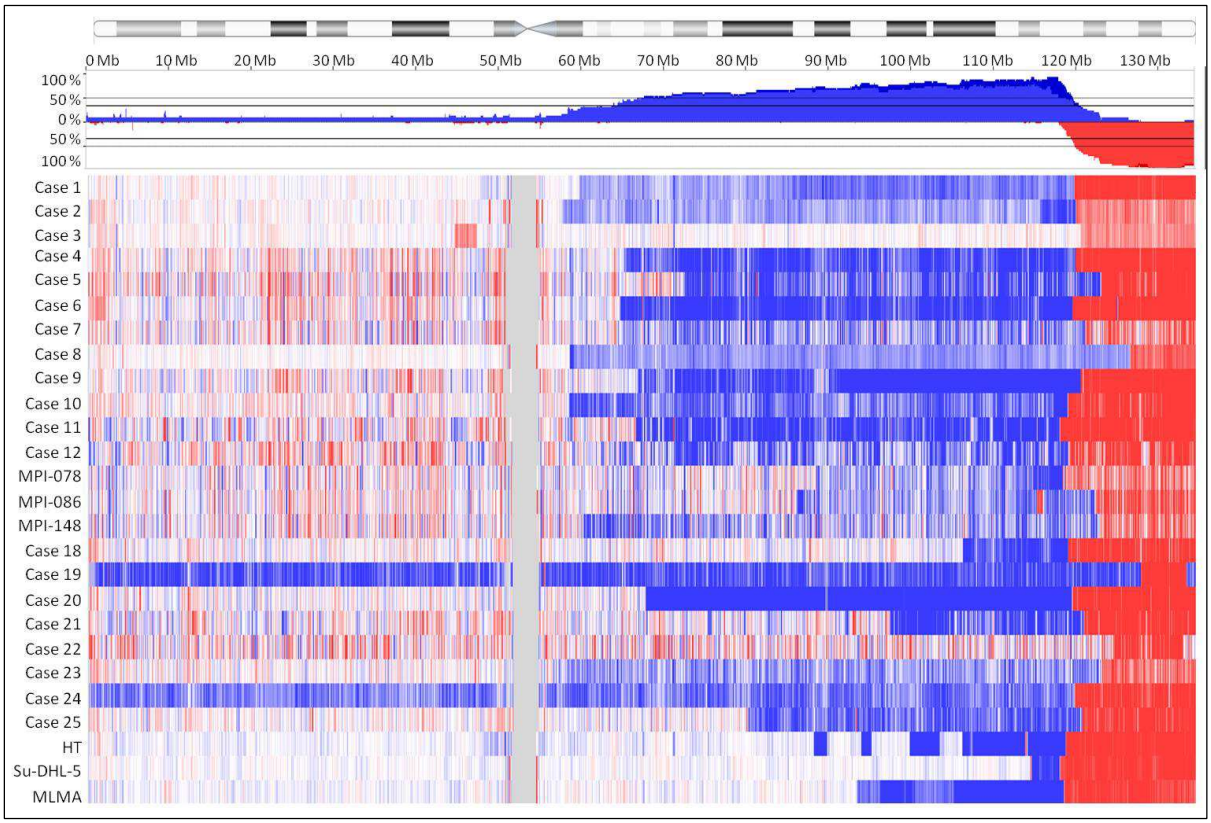


Figure 44: Genomic profile of chromosome 11 analyzed by SNP 6.0 array and OncoScan in mnBLL (n=23) and mnBLL cell lines (n=3). Depicted are the ideogram of chromosome 11, the cumulative percentage of cases sharing this copy number alteration and the analyzed cases: case 1-12 and 18-25, MPI-cases and cell lines HT, Su-DHL-5 and MLMA. Missing are three MPI cases, which have been analyzed using an array-CGH platform. Blue color represents copy number gains, and red color represents copy number losses.

A major hallmark of BL is the low genomic complexity harboring less than five CNA in addition to the *MYC*-translocation [170],[427],[428]. Hence, the genomic complexity of the mnBLL was used to determine if mnBLL resemble BL also at chromosomal level. As already described in 3.2.1.2 the genomic complexity of mnBLL varied strongly among the cases with a range of 2 – 28 CNA per case. The big difference in the amount of CNA might be attributed to the different platforms used for the CN analyses or to the different fixation methods of the tissue of which DNA was extracted. It is known that DNA derived from FFPE material has a lower quality due to degradation [429], leading to a higher background in the CN analyses. Thus, more stringent criteria needed to be applied which might lead to reduced detection sensitivity for smaller alterations. Therefore, it is only possible to

compare the CNA analyzed with the same platform and using DNA derived from the same fixation procedure. This has been already performed for the initial mnBLL cohort showing that the mnBLL have a high amount of CNA in comparison to BL (11.5 vs. 4) and DLBCL (11.5 vs. 6) [166]. Thus, the higher genomic complexity is not in agreement with the Burkitt-like phenotype. Hence, as the MYC deregulation in BL is already such a crucial alteration for the malignant transformation, the low genomic complexity in mnBLL might be an indicator that the mnBLL do not harbor such a potent alteration. And, thus, more alterations are necessary to induce full blown lymphomagenesis. This is in line with the observation that the absence of *MYC*-rearrangements in B-cell lymphoma has been described to be associated with higher genomic complexity [169].

Thus, the chromosomal landscape of the mnBLL was analyzed for secondary recurrent alterations concomitant to the 11q alteration. Recurrency of secondary alterations was defined as alterations occurring in more than 5 of the 25 (20 %) mnBLL. Hence, following secondary alterations were identified: deletion on 6q14.3-q21 and gains in 13q31.3, 18q21 as well as trisomy 12.

Trisomy 12 was identified in 20 % of mnBLL. A partial or a complete trisomy 12 has been described to be common in BL [131],[132], DLBCL [430] and CLL. In the latter, trisomy 12 has been associated with a favorable overall survival [431]. Based on cytogenetic studies in B-cell lymphoproliferative malignancies a “minimal duplication region” 12q13-q22 has been identified in cases with partial trisomy 12 [432]. Five mnBLL (cases 1, 8 and MPI-078, MPI-086, MPI-315) as well as the mnBLL cell lines Su-DHL-5 and MLMA harbor copy number gains in this minimal gain region on chromosome 12. Hence, a total of eleven mnBLL (44 %) harbor a duplication of 12q13-q22 due to a complete or partial trisomy 12. Thus, partial trisomy 12 can be considered as the second most frequent aberration in the mnBLL. The tumorigenic impact of the trisomy 12 is not clear. Overexpression of the oncogenic genes *CDK4* and *MDM2* located within the minimal region of gain on chromosome 12 has been described for CLL harboring a partial or complete trisomy 12 an [433],[434]. Hence, it is conceivable that these genes may also play a role in mnBLL. The oncogenic properties of those two genes for BL in interplay with *CCND3* activating mutations and *CDKN2A* inactivation have already been discussed in 4.1.3.

Chromosome 6q was identified to be deleted in six mnBLL cases and in the Su-DHL-5 cell line. The minimal deleted region was defined as 6q14.3-q21. Chromosomal deletions affecting 6q have been described to occur in 25 to 70 % of mature B-cell lymphoma [229],[435],[436],[437]. The 6q deletion has been associated with a progression from low to high grade FL [438]. Moreover, Offit *et al.* [436] have identified a minimal region of molecular deletion in 6q21 which was mainly seen in high-grade NHL. Two known tumor suppressor genes are located within the minimal region defined for the mnBLL: *EPHA7* and *PRDM1*. The ephrin receptor A7 (EPHA7) gene belongs to the family of receptor tyrosine kinases which control several physiological processes as adhesion or migration [439]. The tumor suppressive property of a truncated, soluble form of *EPHA7* ($EPHA7^{TR}$) has been recently

shown for FL [229],[440]. The knock down of EPHA7^{TR} accelerated the development of lymphoma in a murine model of FL [229]. Upon binding of EPHA7^{TR} to the EPHA2 receptor it acts as a dominant, soluble inhibitor of its oncogenic signaling in lymphoma cells [229],[441]. PRDM1 on the other hand, plays an important role in terminal B-cell differentiation as a transcriptional repressor of genes including AID and ID3 required for B-cell proliferation and affinity maturation [442]. Frequent inactivation of *PRDM1* has been described in about 25 % of activated B-cell like DLBCL [443]. Pasqualucci *et al.* [443] suggested that the loss of *PRDM1* contributes to the lymphomagenesis by blocking the terminal B-cell differentiation towards plasma cells.

In 20 % of the mnBLL cases and in the Su-DHL-5 cell line a 0.2 Mb sized minimal region of gain was identified in 13q31.31 which encompassed the *miR-17-92* locus. This locus is frequently amplified in BL [132],[231],[444] and associated with a higher expression of this oncomir [231],[445],[446]. The *miR-17-92*, encoding six individual miRNAs [447], has been described to have oncogenic functions in survival, proliferation and differentiation [448],[449],[450]. In a murine model overexpression of miR-17-92 together with MYC led to an increased tumor development [451] supporting its role in BL lymphomagenesis. Furthermore, Olive *et al.* [452] have shown that overexpression of miR-17-92 reduced the expression of PTEN which in turn increased PI3K/Akt signaling. As already discussed in 4.1.3.1, the PI3K signaling pathway is one of the main oncogenic pathways deregulated in BL. Furthermore, the mnBLL have been shown to have a comparable PI3K activity index as BL [166] and based on the experiments conducted herein, the survival of the mnBLL cell lines seems also to depend on the PI3K signaling (4.2.4.5). Hence, it is conceivable, that the gain of miR-17-92 in mnBLL might contribute to the activation of the PI3K signaling.

Finally, a chromosomal gain on 18q21.1 including the *TCF4* gene was identified in 20 % of mnBLL and in the MLMA cell line. TCF4 is part of the Wnt/β-catenin signaling [453],[454] which has been described to play a role in various solid tumors as breast cancer or colorectal carcinoma [232],[455],[456]. TCF4 has been described to play a minor role in B-cell development in comparison to the other bHLH family member TCF3 [352]. Nevertheless, it has been shown that they share targets including pro-B-cell genes [365]. As already mentioned in 4.1.3.1, BL without *ID3* mutation carried 18q gains, including the *TCF4* locus, which was associated with a higher TCF4 expression [170]. Therefore, it is reasonable to speculate that TCF4 overexpression due to the 18q gains could constitute an alternative mechanism to activate for example PI3K signaling in mnBLL [72],[170] and, thus, contribute to its lymphomagenesis.

To sum up, together with the hallmark alteration on 11q, several recurrent CNA have been identified in mnBLL with known tumor suppressor genes, as *EPHA7* and *PRDM1,* as well as oncogenes, as *CDK4*, *MDM2* and *miR-17-92,* mapping to these regions. Some of those CNA also occur in BL, hence, the deregulation of those genes might contribute to the pathogenesis of mnBLL by activating

tumorigenic mechanisms similar to those in BL. This might explain to some extent the BL-like phenotype.

4.2.2 The mutational landscape of *MYC*-negative Burkitt-like lymphomas

The analyses of the mutational landscape were dedicated to: (i) identification of recurrently mutated genes within the altered 11q region probably indicating roles as oncogenes or tumor suppressor genes and (ii) identification of recurrently mutated genes throughout the exome potentially cooperating with the 11q changes in the lymphomagenesis of mnBLL. The results of these analyses are discussed in the following.

Screening of the 54 genes of the altered 11q region by next generation sequencing did not lead to the identification of recurrently mutated genes. Overall, only nine genes of this region were found sporadically mutated including *ETS1* which maps to the focal homozygous deletion. Screening of the mnBLL cohort for *ETS1* mutations by Sanger sequencing led to the identification of mutations in 25 % (4/16) of cases [166]. Moreover, some of the *ETS1* mutations might be inactivating, although a modeling of the functional consequences yielded no conclusive results [166]. Case 8 harboring an *ETS1* mutation has been also analyzed by targeted-resequencing. The MAF of this mutation is 0.56. As this case has only one remaining *ETS1* allele due to the heterozygous deletion of the other, this MAF indicates a subclonal mutational event occurring during tumor progression. Furthermore, a germline mutation cannot be excluded. Thus, besides the mnBLL cases with homozygous deletion in 11q, the cases with heterozygous deletion in 11q do not seem, at least on mutational level, an inactivation of the remaining allele.

As *FLI1* maps also to the focal homozygous deletion of the mnBLL, the remaining allele has been screened for inactivating mutations using Sanger sequencing [166]. But none of the mnBLL harbored a mutation within this gene which does not lend support to a tumor-suppressive function though haploinsufficiency cannot be excluded.

The analyses of the complete exome revealed no recurrently mutated genes. This might be due to the fact that only whole-genome sequencing data were available for one case whereas the other cases were analyzed by targeted resequencing of solely a fraction of the genes of the complete exome. Thus, genes with recurrent mutations might be simply not yet identified in this analysis, leaving *ETS1* as the sole recurrently mutated gene identified in the mnBLL up to now.

4.2.3 The transcriptional landscape of *MYC*-negative Burkitt-like lymphomas

To determine an expression signature specific for mnBLL, the gene expression data provided by the MMML project were compared to those of BL and DLBCL separately, focusing on the 54 genes of the

aberrant 11q region. By this approach, combining the data of both comparisons, eight genes were identified to be either higher or lower expressed in mnBLL. Furthermore, two genes (*IL10RA*, *PRDM1*) showed a particular higher differential expression in mnBLL than in BL. Thus, overall ten genes were differentially expressed in mnBLL relative to BL and DLBCL, however, only three (*PAFAH1B2, IL10RA, FLI1*) were described in the literature to be associated with lymphoma (3.2.3.1.3). The role of these three candidate genes for the pathogenesis of mnBLL will be discussed in 4.2.4.

In the publication of the initial cohort of mnBLL, a global gene expression analysis has been performed and the differentially expressed genes between mnBLL and BL, as well as DLBCL were determined [166]. The six candidate genes (*EPHA7*, *PRDM1*, *MDM2*, *CDK4*, *TCF4*, *mir-17-92*) mapping to the regions of secondary alterations (4.2.1) were analyzed for differential expression between mnBLL and BL was well as mnBLL and DLBCL. Of note is that the *mir-17-92* gene cluster is not covered by the U133A gene expression array, hence, its expression was not analyzable. The analysis of the remaining five genes revealed *MDM2*, located on chromosome 12q15, to be lower expressed in mnBLL than in BL (adj. p-value 0.045, t-test p-value corrected according to Benjamini and Hochberg). This finding seems contradictory as 24 % of mnBLL harbor a complete trisomy 12 and, thus, a higher expression of *MDM2* in mnBLL in comparison to BL might be expected. But of note is that the gene expression analysis has been performed for six mnBLL (case 1, MPI-078, MPI-086, MPI-148, MPI-315, MPI-382). These six mnBLL cases do not harbor a trisomy 12 or other alterations affecting the *MDM2* locus. Hence, the detected lower MDM2 expression in mnBLL might simply reflect the fact that those cases do not harbor a chromosomal gain of the *MDM2* locus. This might also hold true for the other four candidate genes (*EPHA7*, *PRDM1*, *CDK4*, *TCF4*) as not all mnBLL harbor all secondary CNA. Thus, for a proper conclusion, the gene expression of the candidate genes needs to be analyzed in those cases which harbor the respective secondary alteration. Furthermore, these secondary alterations detected in mnBLL have been described to be also frequently altered in BL as well as DLBCL (4.2.1). Hence, this might lead to false negative or false positive results in the gene expression analysis.

Taken together, the transcriptome data are a further criterion to narrow down the list of potent candidate genes identified in the genomic and genetic analysis. But as the analyses are based on data from six mnBLL (6/25, 24 %), the possibility of false negative and false positive results needs to be considered.

4.2.4 Possible candidate genes in *MYC*-negative Burkitt-like lymphomas

In the following the role of the five possible candidate genes (*KMT2A*, *PAFAH1B2, ETS1, FLI1* and *IL10RA)* for the pathogenesis of mnBLL are discussed. These genes were identified based on four different criteria: (1) localization within the minimal aberrant region in 11q, (2) differential expression, (3) recurrent mutations and (4) published association with lymphomagenesis.

4.2.4.1 Oncogenic role of *KMT2A* in *MYC*-negative Burkitt-like lymphomas

The *KMT2A* gene maps to the minimal region of amplification in the mnBLL. In line with this, recurrent amplification of 11q23 including the *KMT2A* gene have been described in B-cell lymphomas like DLBCL [239],[245]. Moreover, up to 10 % of AML [457] or up to 13 % of ALL [458] carry a *KMT2A* rearrangement. As gene amplification is one of the mechanisms activating proto-oncogenes, it is conceivable that *KMT2A* has oncogenic properties in the mnBLL.

KMT2A, former known as MLL, is a methyltransferase which is part of a multiprotein complex [459],[460]. It mediates chromatin remodeling and transcriptional activation via chromatin modifications like trimethylation of histone 3 lysine 4 (H3K4me3) [461]. Other functions for KMT2A have been described in the regulation of transcriptional initiation and elongation via the RNA polymerase II [462] or regulation of the S-phase checkpoint in the cell cycle [463]. KMT2A has been shown to be critical for proper hematopoiesis, as severe defects like reduced numbers of hematopoietic stem cells were observed in homozygous knockout mice, as well as in *KMT2A*-depleted zebrafish [464],[465],[466].

KMT2A is slightly higher expressed in mnBLL in comparison to BL and DLBCL, but this difference is not significant (adj. p-value >0.1) [166]. In line with this, the KMT2A expression analysis in the respective cell line subtypes showed a higher KMT2A expression in the mnBLL cell lines. However, this might be biased by the particular high expression in MLMA cells, which was four times higher than in the other two mnBLL cell lines.

A higher KMT2A expression has been shown to be associated with an increased expression of downstream target genes [467],[468] indicating an aberrant transcription regulation [469]. To clarify if this might be also the case in mnBLL, the gene expression data need to be analyzed with regard to the expression of the KMT2A target genes in comparison to for example BL or DLBCL.

In conclusion, the amplification of *KMT2A* is not strongly associated with a stark overexpression in mnBLL, which does not support the likeliness of this gene to be an essential oncogene in the mnBLL.

4.2.4.2 Oncogenic role of *PAFAH1B2* in *MYC*-negative Burkitt-like lymphomas

PAFAH1B2 maps to the minimal region of gain of mnBLL. A disruption of this gene due to a translocation t(11;14)(q23;q32) has been described in a case with B-cell lymphoma, placing the coding sequence of *PAFAH1B2* under the influence of *IGH* regulatory elements [236],[470]. This translocation exhibits similarities to the t(8;14)(q24;q32) in BL leading to a deregulation of the *MYC* gene, which is the hallmark and the driving force of this lymphoma. As the amplification of *PAFAH1B2* in mnBLL led to an increased expression on mRNA and protein level, it is conceivable that PAFAH1B2 might be a driving force of mnBLL.

PAFAH1B2 is a catalytic subunit of the platelet activating factor acetylhydrolase (PAF-AH) [471] which hydrolyses the phospholipid mediator PAF [472]. The synthesis of PAF is induced upon inflammatory stimuli [473] and it mediates upon binding to the extracellular PAF receptor (PAFR) or by PAFR-independent mechanisms a wide range of cellular functions including proliferation, differentiation and transformation [474],[475],[476]. Furthermore, regulatory activities for PAF in B-lymphocytes [477] have been described as treatment of B-cells or an BL cell line with PAF leads to a decreased proliferation [478]. In line with this, patients with lymphoid or non-lymphoid malignancies have been shown to harbor decreased levels of PAF [479]. Due to the higher PAFAH1B2 expression in the mnBLL it might be possible, that the processing and, hence, the level of PAF is decreased. Thus, this might contribute to the proliferative properties of the cells. In line with this, an anti-apoptotic function has been described for PAFAH1B2 due to the hydrolysis of PAF which has been described to induce by a PAFR-independent mechanism caspase-3-dependent apoptosis [250].

Another oncogenic pathway of PAFAH1B2 might be via its interaction with ZFP36L1 which is frequently deregulated in other B-cell malignancies [325],[480],[481]. ZFP36L1 belongs to the family of tristetraprolins [482] which play a role in mRNA degradation and have been associated with tumor suppressive mechanisms [483]. ZFP36L has been described to locate next to the *IGH* locus due to a interstitial deletion in e.g. chronic lymphatic leukemia [480],[481] and to be hypermutated in DLBCL [325]. Thus, the interaction of PAFAH1B2 might lead to a negative regulation of ZFP36L1.

Despite these published evidences, using knock down experiments in the mnBLL cell line Su-DHL-5 an essential role for this gene could not be confirmed. The loss of PAFAH1B2 did not reduce the cell viability under the tested experimental conditions. But as its function might be to protect from the negative influence of PAF, which was not part of the experimental setting, its true role *in vivo* might be missed under the applied experimental conditions.

Taken together, a role for PAFAH1B2 as a driving force in the lymphomagenesis of mnBLL could not be confirmed. Nevertheless, it is conceivable that PAFAH1B2 contributes to the progression of the tumor by degradation of PAF and, hence, might maintain the proliferation and survival of the cells.

4.2.4.3 Tumor suppressive role of *ETS1* in *MYC*-negative Burkitt-like lymphomas

ETS1 belongs to the family of ETS transcription factors [484]. The hallmark of these transcription factors is a conserved DNA-binding domain, called ETS domain, which interacts with a specific nucleotide sequence in the target genes containing a GGAA/T core sequence [485],[486].

Functions for *ETS1* in B-cell maturation and survival have been described as regulation of class-switch recombination [487] and BCR-signaling dependent S-phase entry [488]. Furthermore, ETS1 has been described to regulate the expression of PRDM1 [246] which inhibits PAX5. As PRDM1 is important for the plasmacytic differentiation [247], its inhibition mediated by ETS1 might contribute to the

regulation of B-cell maturation. Furthermore, ETS1 has been shown to be important for the transcriptional regulation of diverse cytokines like IL2 and IL10 in various lymphoid cells [489],[490],[491].

As already mentioned in 4.2.1, *ETS1* is one of the few genes localized in the focal homozygous deletion, indicating a role as tumor suppressor gene in mnBLL. A tumor suppressive function for ETS1 has been already proposed in Hodgkin lymphoma which were shown to harbor recurrent deletions of the *ETS1* locus as well as frequent *ETS1* mutations [248],[426]. According to Knudson´s two hit theory as described in 1.3.2 both alleles of a tumor suppressor need to be inactivated. Only two of the 25 analyzed mnBLL harbored a homozygous loss of the *ETS1* locus, whereas the majority showed a heterozygous loss. In addition, although mutations of *ETS1* were detected in 25 % of mnBLL, those mutations were subclonal and modeling of a tumorigenic impact of the mutations was not conclusive [166]. Analyses of the mRNA and protein expression indicated that the second allele might be still active, as the ETS1 expression was comparable between mnBLL and BL as well as DLBCL, and in the respective cell line models. Hence, in contrast to the two mnBLL cases with the homozygous loss in 11q, the other mnBLL cases as well as the mnBLL cell lines do not seem to have lost the ETS1 expression and, thus, function. Consequently, an essential tumor suppressive role for *ETS1* in mnBLL is unlikely. But it might be that co-deregulated genes of the 11q region might influence directly or indirectly the regulation of ETS1 function and contribute in this manner to the tumorigenicity of ETS1 in mnBLL.

4.2.4.4 Tumor suppressive role of *FLI1* in *MYC*-negative Burkitt-like lymphomas

FLI1 belonging to the family of ETS transcription factors [492] maps also to the focal homozygous deletion in 11q24.3. The *FLI1* gene is localized 106 kb telomeric from *ETS1* and it is thought that both genes arose from a common ancestor by gene duplication [492]. In line with this, both genes have been shown to target the same core consensus sequence within the target genes [493]. FLI1 is mainly expressed in the hematopoietic lineages including lymphocytes but also in a variety of other tissues [494]. It has been shown in a mouse model that loss of FLI1 expression leads to impaired B-cell development likely due to downregulation of PAX5 and TCF3, and overexpression of ID1 and ID2 [495]. In line with this, FLI1 has been shown to build a complex with PAX5 and ETS1 in a murine B-cell line and, hence, regulate the transcription of CD79 by binding to its promoter [496]. Furthermore, different studies in erythroid cells have shown that FLI1 is a transcriptional activator of BCL2 [497], MDM2 [498] and miR-17-92 [499] expression, hence contributing to the survival and proliferation of these cells. In monocytes, FLI1 is a positive regulator of the expression of the IL10 cytokine [500]. Taken together, the expression of FLI1 in B-cells as well as in erythroid cells is important for the maturation and survival of the respective cells.

Since *FLI1* maps to the focal homozygous deletion it might have tumor suppressive properties in mnBLL. On transcriptional level, mnBLL showed a significant lower FLI1 expression than BL and DLBCL. This holds also true for the respective cell lines on mRNA as well as on protein level. Interestingly, the DLBCL cell line Karpas422 which was dismissed as mnBLL cell line model based on the presence of the *BCL2*-translocation showed a reduced FLI1 expression. Analysis of the protein expression revealed, that the two FLI1 isoforms were differentially expressed between the lymphoma subtypes, indicating different mechanism of FLI1 translational control. In the mnBLL cell lines no FLI1 expression (HT and Su-DHL-5) or solely expression of the small isoforms (MLMA) was detectable. The BL cell lines on the other hand expressed both isoforms and the DLBCL cell lines expressed rather the smaller isoforms. Up to now it is still discussed if both isoforms have diverse functions [249],[501]. Thus, whether the different expression pattern of the FLI1 isoforms reflects different functions of FLI1 in the lymphoma subtypes is not clear.

Remarkably, the mRNA expression of FLI1 was reduced in the mnBLL cell lines, whereas on protein level the expression was lost. Whether the primary mnBLL express FLI1 protein or not is not known. Nevertheless, it is conceivable that the FLI1 protein expression is lost in the primary mnBLL as in the mnBLL cell lines. The reason for the lack of FLI1 expression is yet unclear As the mnBLL cell lines as well as 23/25 mnBLL solely harbor a heterozygous loss of the *FLI1* locus and the coding sequence of *FLI1* is not frequently mutated,. A possible explanation might be an epigenetic regulation of the FLI1 expression. A differential methylation of *FLI1* in mnBLL in comparison to BL as well as DLBCL could not be detected (data not shown). The chromatin state of Su-DHL-5 cells and six other lymphoma cell lines including BL and DLBCL cell lines have been analyzed by Dr. med. Anke Bergmann, Institute of Human Genetics, Kiel within the European BLUEPRINT consortium and were kindly provided. Figure 45 gives an overview on the chromatin states of the *FLI1* locus in the cell lines. In general, the chromatin states were similar besides a block of repressed heterochromatin within the first intron in those cell lines with lack FLI1 expression (Su-DHL-5, Karpas422). This finding might indicate an epigenetic regulation of the remaining *FLI1* allele in the mnBLL cell line.

So far, FLI1 has been shown to harbor oncogenic functions in different malignancies, as in Ewing sarcoma which harbor a recurrent FLI1-EWS fusion [502] or in DLBCL which harbor recurrent *FLI1* amplification leading to an increased cell viability [239]. The mnBLL represent the first so far described malignancy in which FLI1 might have a tumor-suppressive function. This is in contrast to the functions of FLI1 in B-cell maturation and proliferation which point towards an oncogenic role in those cells. But as ETS1 and FLI1 share the consensus sequence of the target genes [493] and as they share some functions in the B-cell maturation and proliferation, it is conceivable that ETS1 might substitute for some of the FLI1 functions which are essential for the cell survival whereas those functions which are essential to be lost for the pathogenesis of the mnBLL are not substituted.

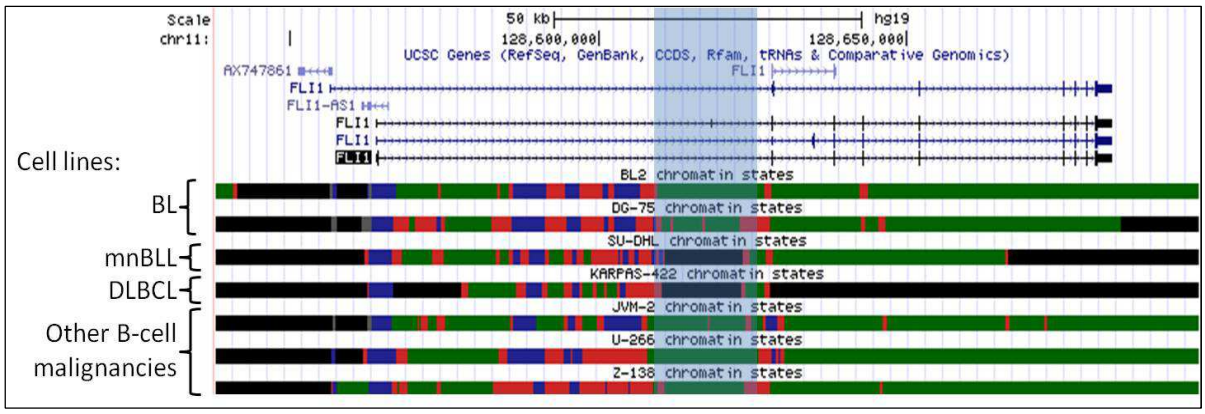


Figure 45: Chromatin states of the *FLI1* locus in mnBLL (Su-DHL-5), BL (BL-2 and DG-75), DLBCL (Karpas422) cell lines and cell lines from three other B-cell malignancies (JCM-2, U-266, Z-138). Depicted are from the top to the bottom: scale bar in kb, genome base position in bp, the UCSC gene track showing all transcripts of the *FLI1* gene (reading direction exon 1 on the left to last exon on the right) as well as the chromatin states of the respective cell lines. The chromatin states are represented by the following color code: black, repressed heterochromatin; grey, repressed promoter; blue, active promoter; red, regulatory element and green, transcribed region. Highlighted in blue is a block of repressed chromatin in the mnBLL cell line Su-DHL-5 and the DLBCL cell line Karpas422 in the first intron of FLI1 which might be associated with the loss of FLI1 expression.

Taken together, *FLI1* maps to a focal deletion and its expression is reduced in the mnBLL indicating a tumor suppressive role in the mnBLL. It is conceivable that due to the loss of FLI1 the transcription of some of its target genes becomes deregulated and, thus, contribute to an increased proliferation and survival of mnBLL cells.

4.2.4.5 Contribution of *IL10RA* and IL10 signaling to pathogenesis of *MYC*-negative Burkitt-like lymphomas

IL10RA was considered as a possible candidate gene, as it maps to the minimal region of amplification and is significantly higher expressed on mRNA level in mnBLL than in BL.

IL10RA encodes the α-subunit of the IL10 receptor (IL10R). The receptor is formed upon binding of the cytokine IL10 to IL10RA which leads to a recruitment of IL10RB and, thus, establishment of the active receptor [503]. IL10R belongs to the class II cytokine receptor family consisting of 6 different receptor chains forming heterodimers [504]. IL10RA has been described to be constitutively expressed in most hematopoietic cells [505],[506], but also in non-hematopoietic cells. In the latter, its expression needs to be usually induced [507],[508]. IL10RB on the other hand is constitutively expressed in most cells and tissues [509]. So far, there has been no indication for an activation-induced expression of IL10RB in immune cells [510].

IL10 signaling has been described to have immunstimulatory and anti-inflammatory functions in immune cells. It regulates the immunomodulation by downregulation of pro-inflammatory cytokines [511],[512] and upregulation of cytokine antagonists like the interleukin 1 receptor antagonist [513],[514]. Moreover, it inhibits MHC II expression in antigen presenting cells [515]. Among the immunstimulatory functions is the induction of proliferation in B-cells [516] as well as the induction

of maturation and isotype switching [517]. A role for IL10 signaling by stimulation of growth and progression in B-cell lymphoma has been suggested in various studies [237],[518],[519]. Furthermore, studies of *in vivo* murine lymphoma models with inducible MYC expression have shown that high MYC expression down-regulates IL10RA. But when MYC is turned off again, IL10RA expression recovers and induces cell proliferation [520]. In other experiments it was demonstrated that lymphomas expressing high IL10RA levels have selective growth advantage over lymphomas expressing low IL10RA levels likely due to upregulation of the anti-apoptotic MCL1 protein [237]. Thus, IL10 signaling might represent an oncogenic stimulus in the mnBLL due to autocrine signaling as has been proposed for other B-cell lymphoma [521],[522].

To get a better overview on the components of the IL10 signaling, this pathway, depicted in Figure 46, is shortly introduced in the following.

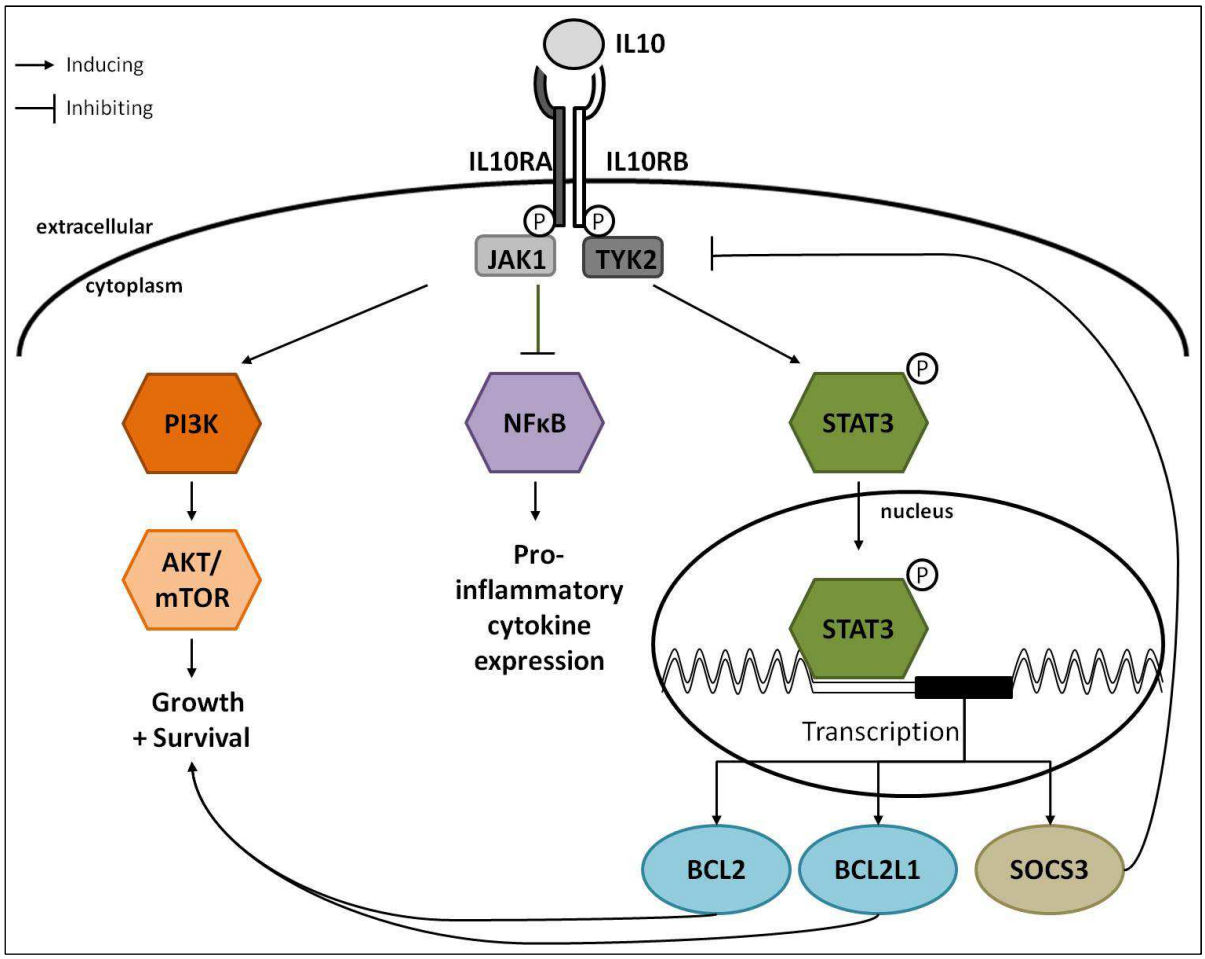


Figure 46: Schematic overview on IL10 signaling. Upon binding of IL10 to its receptor, the tyrosine kinases JAK1 and TYK2 become phosphorylated and activate STAT3, which translocates into the nucleus where it induces the expression of several genes including the anti-apoptotic genes *BCL2* and *BCL2L1* as well as *SOCS3* which is a negative regulator of the IL10 signaling. IL10 signaling inhibits NFκB leading subsequently to downregulation of pro-inflammatory cytokines. Furthermore, the PI3K-Akt pathway becomes activated inducing the growth and survival of the cells.

Initially, IL10 binds to IL10RA which subsequent interacts with IL10RB to form the active IL10R [503]. This leads to the phosphorylation of the tyrosine kinases TYK2 and JAK1 which are bound to IL10RB

and IL10RA, respectively. Different pathways have been described to be induced afterwards. Of note is that most of the published analyses of IL10 downstream signaling pathways have been performed in monocytes and T-cells, but only rudimentarily in B-cells. Thus, it is not clear if all pathways are active in B-cells. Upon activation of JAK1 and TYk2, the transcription factor STAT3 becomes activated [523] and translocates into the nucleus where it regulates the expression of several genes including the anti-apoptotic genes *BCL2* [524] and *BCL2L1* [525]. Moreover, it regulates SOCS3 expression which is important for the negative feedback regulation of IL10 signaling [526]. Furthermore, IL10 signaling has been described to inhibit NFκB activation [527] by either blocking of the nuclear translocation or inhibition of the DNA binding activity [528], hence regulating the expression of inflammatory cytokines [529]. A third pathway activated by IL10 signaling via activation of the Insulin receptor substrate 2 [530] is the PI3K-Akt pathway which is important for the proliferative but not for the inflammatory effects [251].

An initial step to determine an oncogenic role of the IL10 pathway in mnBLL in this thesis was the expression analyses of IL10RA and IL10 in the respective cell line models. On mRNA level, the IL10RA expression was highest in the mnBLL cell lines, whereas on protein level also four of the BL as well as two of the DLBCL cell lines showed expression levels comparable to the mnBLL cell lines. This suggests a post-transcriptional regulation of the IL10RA levels. Whether this acts via mRNA degradation or enhanced translation is not clear. Screening of the cell lines for IL10RA cell surface expression revealed a higher surface expression in BL cell lines than in mnBLL cell lines. Corinti *et al.* [531] have shown that mature dendritic cells selectively upregulated IL10RA mRNA but did not exhibit a higher surface expression. They attributed this effect to possible posttranslational modifications as negative feedback regulation to prevent abnormal IL10RA expression levels [531]. Although these studies have been performed in dendritic cells, it might be that in the mnBLL a similar mechanism is active. Alternatively, due to a higher IL10 abundance and consequently signaling, a higher rate of IL10R internalization might occur [532]. Another possibility is that due to receptor shedding IL10RA becomes soluble. So far, a soluble form of IL10RA *in vitro* has been not described. Nevertheless, the IL22 receptor, which also belongs to the class II cytokine receptor family of IL10, exists as a soluble form (IL22BP) [533]. Furthermore, it cannot be definitely excluded that the results of the protein expression analyses are prone to technical artifacts. The antibody used for the FACS analysis was raised against a recombinant soluble form of IL10RA whereas the antibody for the Western blot was raised against a central epitope within IL10RA. Thus, the efficiency and specificity of the antibodies might differ. In line with this, the analysis of the IL10RA expression in the BL-70 cell line lacked detectable expression in the Western blot whereas in the FACS analysis the cells exhibited the highest relative surface expression. This cell line had no mutations or structural alteration within the epitope recognized by the antibody used for Western blot analysis. Thus, the marked difference

in both experiments might be a result of an unspecific binding of the FACS antibody. Despite of these technical issues it can be stated that the expression of IL10RA was detectable in the mnBLL which is the crucial prerequisite for the IL10 signaling [534].

An elevated expression of IL10 was detectable on mRNA and protein level in two of the three mnBLL cell lines but not in the BL and DLBCL cell lines as well as in the mnBLL cell line HT. The lack of IL10 expression in latter cell line might be due to intercellular viability. To analyze the IL10 levels in mnBLL serum was only available for case 10. Within the serum of mnBLL case 10 an average IL10 level of 14 pg/ml was detected. Comparable levels of IL10 were detected in the serum of two patients with BL. In line with this, similar IL10 levels have been detected in various studies analyzing serum from patients with NHL and HL [535],[536],[537]. These studies showed an average IL10 level of 15-21 pg/ml. These levels were elevated in comparison to those of healthy donors which showed average IL10 level of 10 pg/ml in serum [535],[537]. Although the applied IL10 ELISA assays in those studies differed from the one used in this thesis, the IL10 level of 10 pg/ml in healthy donors might be a good reference value for the mnBLL case. Hence, the mnBLL case showed also a slightly higher IL10 level. Interestingly, the increased IL10 levels in the NHL and HL patients were associated with shorter failure-free survival [537],[538].

The high IL10 expression in mnBLL cannot be explained by genomic, genetic or epigenetic alterations. Nevertheless, its expression is also regulated by a variety of transcription factors [539],[540]. Moreover, in different cell types the regulating network possibly varies [541]. As already mentioned, the transcription factors ETS1 and FLI1 have been shown in various studies to regulate IL10 expression [500],[542],[543]. In those studies IL10 was negatively regulated by ETS1 in T cells but positively in monocytes. FLI1 was described as a positive regulator of IL10 expression in monocytes. A further indication for an ETS family regulated IL10 expression in B-cells comes from a study by Bonetti *et al.* [239]. They have shown that knock down of FLI1 led to increased expression of IL10 in a DLBCL cell line. Thus, the combination of gain and loss in 11q might contribute directly in *cis* to the upregulation of the IL10R subunit IL10RA and indirectly in *trans* to the upregulation of IL10

To further stratify a possible role for IL10 signaling in mnBLL, single components of the pathway were inhibited herein. Figure 47 gives an overview on the possibility of therapeutic interaction and possible oncogenic mechanism conferred by IL10 signaling in mnBLL as outlined in the following.

Blocking of IL10RA using two neutralizing antibodies resulted in opposing effects. One antibody reduced the cell viability not only of the mnBLL cell line (Su-DHL-5) but also of the BL cell line (BL-2). This might reflect either that the antibody has unspecific side effects or that both cell lines need the IL10 signaling for survival. Nevertheless, the latter is contradicted by the finding that the BL-2 cells do not show reduced cell viability upon treatment with the neutralizing IL10 antibody. Thus, the observed effects of the IL10RA blocking in BL-2 cells are more likely side effects.

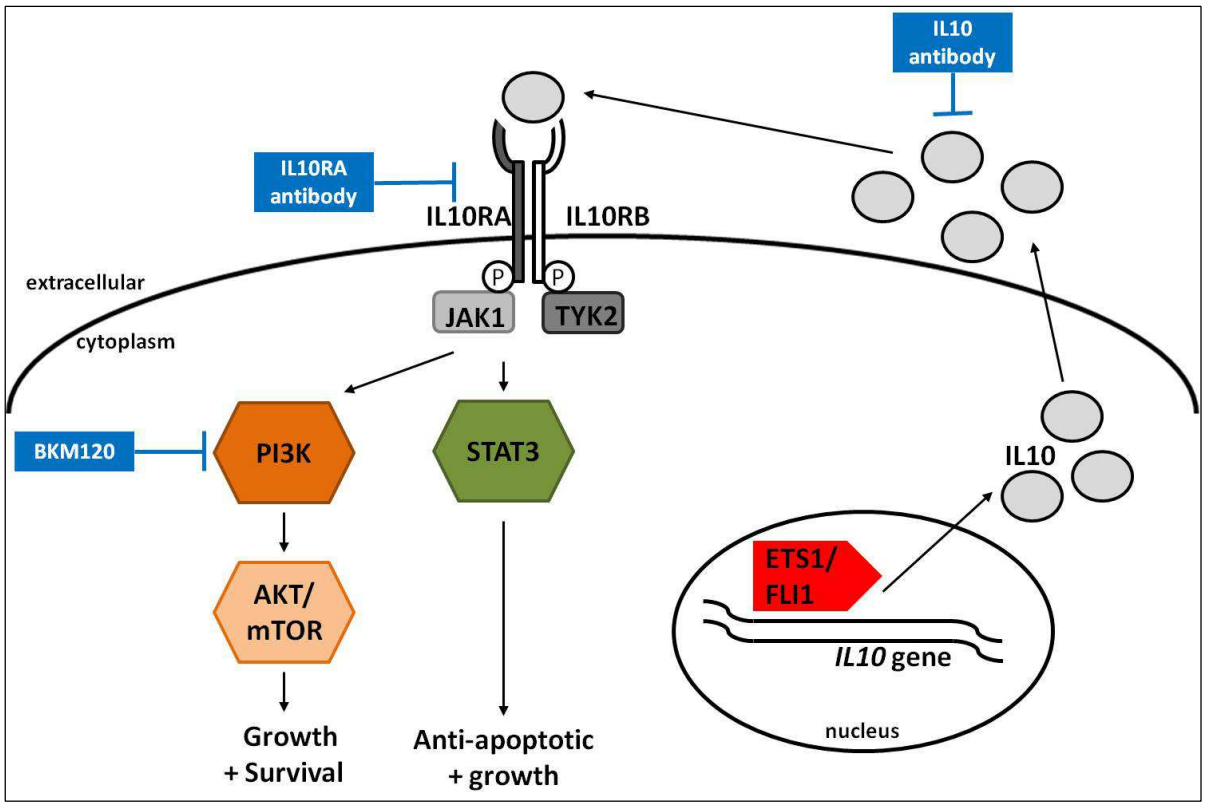


Figure 47: Overview on the possibility of therapeutic interaction and possible oncogenic mechanism conferred by IL10 signaling. The alpha subunit of the IL10 receptor (IL10RA) is overexpressed in the mnBLL which might lead to an overrepresentation on the cell surface. The expression of the IL10R ligand IL10 might be negatively regulated by ETS1 or FLI1 which map to the minimal region of homozygous loss. In mnBLL the expression of FLI1 is lost, indicating a loss of transcriptional regulation of IL10. A higher expression of IL10 might lead to an autocrine activation of the IL10 receptor. Activation of the receptor leads to the phosphorylation of the kinases JAK1 and TYK2 which subsequently activate survival and growth pathways via the STAT3 and PI3K signaling. Different components of the pathway were inhibited to analyze the dependency of mnBLL cell lines on IL10 signaling which are indicated by blue boxes.

Treatment of the BL, DLBCL and mnBLL cell lines with IL10 neutralizing antibody points towards a dependency of the mnBLL cell line Su-DHL-5 on IL10 signaling. The viability of Su-DHL-5 cells was reduced to 10 % after treatment with high doses whereas BL and DLBCL cell lines showed no or only slight reduction. Oddly, treatment of the other two mnBLL cell lines HT and MLMA with the neutralizing IL10 antibody did not reduce their cell viability, although MLMA cells expressed IL10 at high levels. Again this might point towards intercellular variability and hence, the survival of the MLMA cells might not solely depend on IL10 signaling as the Su-DHL-5 cells. Furthermore, this indicates that IL10 signaling might not be the sole unifying mechanism in the mnBLL.

IL10 signaling can activate the PI3K pathway [251]. Indeed, in mnBLL the PI3K signaling activity index was shown to be similar to that in *MYC*-positive BL [166]. As already described the activation of this pathway is important for the survival of BL [72],[133]. Therefore in this thesis, the PI3K pathway was inhibited *in vitro* using the PI3K class IA specific inhibitor called BKM120. This inhibitor has already been used in several clinical trials [252],[416]. As already mentioned in 4.1.3.1 the BL cell lines showed reduced cell viability which was in line with the findings reported by Schmitz *et al.* [72]. From the technical side of view this indicates appropriate experimental conditions upon treatment with

BKM120. All three mnBLL cell lines also showed reduced cell viability after treatment with BKM120 suggesting that the PI3K pathway plays also an important role in mnBLL. Nevertheless, as the PI3K pathway is activated by various stimuli including BCR signaling as summarized in 1.3.4, these results are not a clear indicator for an IL10 dependent activation of PI3K in mnBLL. In spite of this, the finding that the mnBLL seem to dependent as the BL on PI3K activity might be a further cause for the Burkitt-like phenotype of mnBLL. Moreover, independent of the underlying mechanisms of activation of PI3K signaling, the responsiveness of the mnBLL cell lines to BKM120 suggests that this compound might be valuable in an anti-tumoral therapy approached in patients with mnBLL.

During the writing of this thesis Béguelin *et al.* [544] described a deregulated IL10 signaling in DLBCL due to frequent genomic amplifications and overexpression of the IL10R subunits as well as of IL10. They suggested that, due to the deregulation of IL10R and IL10, an IL10-dependent auto-stimulatory loop is activated and, thus, by IL10 induced STAT3 signaling the proliferation and survival of the DLBCL cell is induced [544]. As described above, phosphorylated STAT3 regulates, after translocation into the nucleus, the expression of several target genes. Moreover, recent studies have shown that also unphosphorylated STAT3 can translocate into the nucleus and activate transcription of its target genes [545],[546]. Remarkably, STAT3 can induce its own expression upon binding to its own promoter [547]. Investigation of the available U133A gene expression data revealed that the STAT3 mRNA levels were significant higher in mnBLL than in comparison to BL (adjusted p-value $7.02x10^{6}$, t-test p-value corrected according to Benjamini and Hochberg) whereas STAT3 was significant lower expressed in mnBLL than in DLBCL (adjusted p-value 0.03, t-test p-value corrected according to Benjamini and Hochberg). The latter might reflect the finding that activated B-cell like DLBCL have been reported to harbor elevated levels of STAT3 contributing to the pathogenesis of this lymphoma [548],[549]. Nevertheless, the mnBLL showed in comparison to BL elevated STAT3 levels which might reflect an IL10 signaling induced STAT3 activation and, thus, STAT3 expression. Whether STAT3 signaling is activated in mnBLL and, hence, confers oncogenic signals to the mnBLL cells needs to be further investigated.

Taken together, based on the functional analyses performed in the mnBLL an oncogenic role for IL10 signaling is conceivable as (i) IL10RA is higher expressed in mnBLL than in BL, (ii) IL10 is highly expressed in mnBLL indicating a possible autocrine signaling, (iii) the PI3K pathway activity index is in mnBLL similar to BL which highly dependent on its signaling for survival, (iv) inhibition of PI3K signaling *in vitro* led to a complete loss of cell viability in mnBLL cell lines. Further experiments need to be performed to investigate the mechanisms and its potential to targeted intervention.

4.2.5 Summary of the analyses of *MYC*-negative BL-like lymphomas

By a combination of genomic, genetic, expression and functional studies a set of candidate genes possibly harboring a tumorigenic role in mnBLL pathogenesis was identified. The complex alteration on 11q is the hallmark of mnBLL. Several of the genes mapping to this region are deregulated and might contribute to the pathogenesis of mnBLL by cooperating with other oncogenes and/or tumor suppressor genes of the same region or of the other recurrently altered region. Figure 48 gives a schematic overview on the candidate genes and their potential tumorigenic functions in mnBLL.

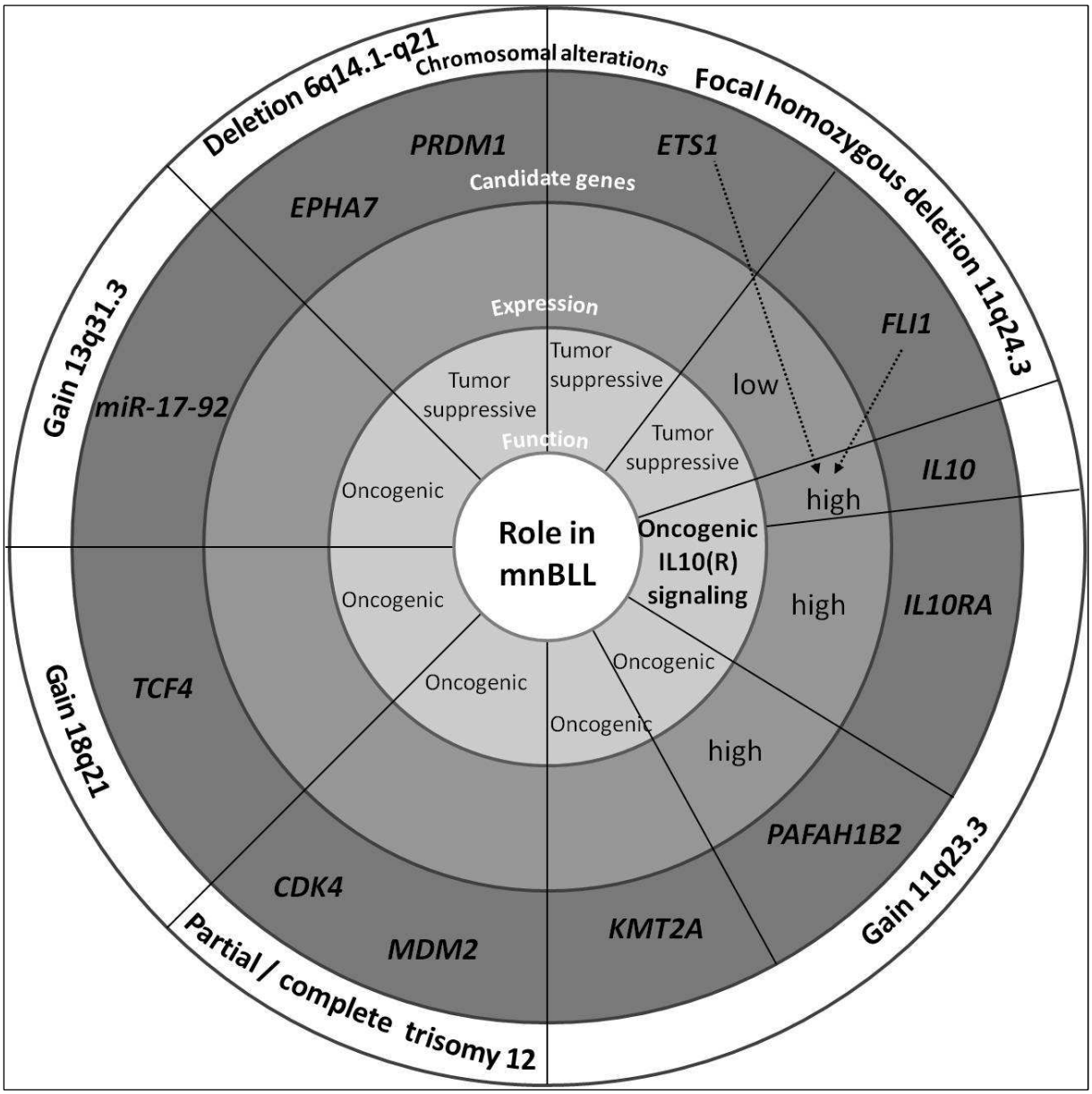


Figure 48: Schematic overview on potential candidate genes and their potential roles in mnBLL. Based on recurrent genomic alteration (white outer circle) various candidate genes have been identified (dark grey circle). Some of this candidate genes show a differential expression in the gene expression analysis which was also confirmed in the mnBLL cell line panels (middle grey circle). Based on the function of the gene, published association with B-cell lymphomas or even functional experiments performed within this thesis, a possible tumorigenic function of the candidate genes was suggested (light grey circle). Based on the study of the literature some of the candidate genes might interact or regulate each other as indicated by dashed lines.

4.3 Comparison of *MYC*-positive Burkitt lymphoma and *MYC*-negative Burkitt-like lymphoma

Based on the pathological and molecular features mnBLL resemble BL rather than DLBCL This led to their naming as “High grade B-cell lymphoma with features of BL, MYC-, with 11q-gain/loss pattern” [166]. In line with these findings, initial analyses of the methylation landscape confirmed that the mnBLL are more similar to BL than to DLBCL (Wagener and Ammerpohl, unpublished data). Different hypotheses exist to explain this similarity: (i) the altered 11q region harbors a gene with MYC-like properties, (ii) similar (signaling) pathways are deregulated or (iii) both lymphoma subtypes derive from the same cell of origin. In the following, based on all available data the three hypotheses will be discussed.

None of the genes mapping to the altered 11q region have been described to be such a potent, general amplifier of the active transcriptional program of a cell like MYC. The finding that the mnBLL do not harbor such a potent deregulated *MYC*-like gene is further supported by the higher genomic complexity in contrast to BL. Thus, mnBLL harbor up to 24 CNA in addition to the 11q alteration. This indicates the necessity of mnBLL to acquire more alterations to induce malignant transformation in contrast to primary BL. In BL the deregulation of MYC is already such a crucial alteration that only 1-2 additional CNA [170],[427] are required for full blown lymphomagenesis. Of note is that 64 % of the mnBLL harbored at least one of the four recurrent alterations as defined in 3.2.1.2. Remarkably, the spectrum of CNA identified in mnBLL is similar to that found in BL [131],[132] suggesting for similar tumorigenic mechanisms. In line with the chromosomal alterations, analysis of the mutational landscape does also suggest for more complex transformation mechanism in mnBLL than in BL. BL harbor frequent mutations within a core set genes like *ID3, CCND3, TP53* and *SMARCA4* leading subsequent to a deregulation of pathways already described in 4.1.3.4. In contrast, none of the analyzed mnBLL harbored a mutation within those core genes. Thus, the lack of a deregulated *MYC*-like gene in mnBLL might be reflected by the necessity to acquire more alterations for full blown lymphomagenesis.

It is highly probable that the Burkitt-like phenotype of mnBLL is due to the shared deregulation of the PI3K signaling which has been shown to be essential for BL (4.1.3.1). In the publication of the initial cohort, a PI3K activity index in mnBLL comparable to that in BL has been described [166]. This could be linked in subsequent analyses performed in this thesis to a deregulated IL10 signaling in mnBLL. Further experiments need to be performed to confirm the association of increased IL10 signaling with a higher PI3K signaling and, hence, as the essential oncogenic pathway in mnBLL.

MnBLL harbor a germinal center B-cell like gene expression signature [166]. The germinal center can be divided into the dark and light zone. Based on gene expression signatures derived from cells of these zones, BL and DLBCL were shown to harbor similarities with the cells of the dark and light zone, respectively [68]. To unravel from which zone the mnBLL might derive, we applied the same

approach as published by Victora *et al.* [68]. The analyses performed by Dr. Maciej Rosolowski, Institute for Medical Informatics, Statistics and Epidemiology, Leipzig, indicate that the mnBLL might also derive from dark zone B-cells. This further supports an association between the Burkitt-like phenotype in mnBLL and the shared cell of origin with BL.

Taken together, on genomic and genetic level the mnBLL show marked differences which are likely attributed to a more complex pathogenic mechanism. The findings that the mnBLL do also depend on PI3K signaling for survival and that they probably have a common cell of origin might cause the Burkitt-like phenotype of mnBLL.

4.4 Conclusions

The analyses of this thesis focused on two mature aggressive B-cell lymphoma subtypes: *MYC*-positive Burkitt lymphoma and *MYC*-negative Burkitt-like lymphoma. In the following the major observations of this thesis are summarized for each lymphoma subtype.

Conclusion from the analyses of *MYC*-positive Burkitt lymphoma:

- The PluriTest algorithm showed a higher similarity of BL to hPSC than of non-BL. Nevertheless, an origin of BL from or resemblance to cells with PSC features could not be established.
- The transcription of the newly identified exon of the *TERT* gene could either be a mechanism to negatively regulate the TERT expression or function, or the truncated proteins might contribute to the wildtype TERT function.
- Analyses of the mutational landscape of the BL cell lines demonstrated the resemblance of the BL cell line models with the primary BL. Thus, the cell lines seem to be appropriate *in vitro* models for functional analyses
- Analyses of the mutational landscape in primary BL and BL cell lines led to the identification of recurrently mutated candidate genes potentially contributing to the lymphomagenesis:
  - Inactivation of the *ID3* tumor suppressor contributes to subsequent deregulation of the TCF3-BCR-PI3K pathway. This deregulation in combination with *IG-MYC* translocation is crucial for the pathogenesis of BL.
  - The frequent mutations in the catalytic domain of SMARCA4 likely have a dominant negative effect. Hence, SMARCA4 might also be a potential tumor suppressor in BL.
  - *PCBP1* harbors frequent mutations affecting the KH III domain and/or NLS likely leading to a reduction or even loss of function.

Conclusion from the analyses of *MYC*-negative Burkitt-like lymphoma:

- Characterization of the altered 11q region of all 25 mnBLL led to a redefinition of the minimal region of loss and the included minimal region of homozygous loss. Moreover, recurrent secondary alterations affecting 6q, 13q and 18q as well as a complete trisomy 12 were identified. Each region harbored potential tumorigenic candidate genes.
- Integrative analysis of the chromosomal, mutational and transcriptional landscape of the altered 11q region as well as a comprehensive literature search led to the identification of three potential oncogenes (*IL10RA, KMT2A* and *PAFAH1B2*) and two potential tumor suppressor genes.
- By interplay of IL10RA and possibly FLI1/ETS1 induced IL10 deregulation, an autocrine IL10 signaling loop might be activated which subsequently induces the oncogenic PI3K pathway. Thus, the deregulation of IL10 signaling might be a key oncogenic pathway in mnBLL.

4.5 Perspectives

After identification of several potential candidate genes in *MYC*-positive Burkitt lymphoma and *MYC*-negative Burkitt-like lymphoma within this thesis, the next step should be the extensive investigation of their mechanisms of action using functional analyses in the here identified cell line models.

For instance, to elucidate if the mutations within the ATPase domain of *SMARCA4* have indeed a dominant negative effect on the transcription of its target genes, a CRISPR-Cas system [550] could be used to induce side-directed mutations within *SMARCA4*. By subsequent chromatin immunoprecipitation (ChIP) the binding of the protein encoded by either wildtype or mutated *SMARCA4* to its target genes could be determined.

The *PCBP1* mutations likely lead to a loss of the KH III domain and/or NLS domain which might impair its function through cytoplasmatic retention. To investigate this in more detail the cellular expression of PCBP1 in *PCBP1* mutated and wildtype cell lines as well as the primary BL could be analyzed using immunefluorescence.

The transcription of the new exon of TERT might have different impacts on the protein function either leading to translation of an altered protein or NMD. To elucidate if a truncated protein is indeed translated from the alternative transcript, the translation of this transcript could be analyzed using *in vitro* translation assays. Furthermore, to elucidate if the transcript is expressed *in vivo* a specific antibody needs to be engineered to detect particularly the alternative protein e.g. in Western blot. In a next step the function of this alternative transcript could be analyzed using e.g. RNA interference assays. To study whether the alternative transcript is subjected to NMD, inhibition of the NMD machinery using specific compounds like caffeine or emetine [551],[552] in cell lines expressing this transcript could be performed. Afterwards, the abundance of the transcript could be quantified using qPCR.

The description of the features of the new lymphoma subtype mnBLL is based on overall 25 cases. Thus, the aim of future studies needs to be the recruitment of further cases to get an even better overview on the chromosomal and mutation landscape as has already been started within this thesis. Furthermore, analyses of the methylation landscape might help to identify potential pathomechanisms contributing to its lymphomagenesis as has been published for other B-cell lymphomas [553]. The deregulated IL10 signaling pathway has been identified as a potential oncogenic pathway in mnBLL. To further investigate its role in mnBLL, the activation by phosphorylation of downstream components like the IL10 receptor associated tyrosine kinases JAK1 and TYK2 using specific phospho-antibodies could be analyzed. Furthermore, in addition to the PI3K inhibition, further components of the pathway upstream of PI3K like e.g. JAK1 and TYK2 need to be specifically inhibited. The basis for the autocrine signaling is the expression of the ligand, IL10, which might be regulated by FLI1 and/or ETS1. Thus, the physical interaction of those transcription factors

with the *IL10* gene locus could be analyzed using ChIP. Depending on the results, two different functional strategies might be applied to analyze if the IL10 expression is indeed regulated by those two transcription factors: (i) re-expression of the FLI1/ETS1 in the mnBLL cell lines and/or (ii) downregulation of FLI1/ETS1 in a cell line by RNA interference or usage of specific compounds as have been described for FLI1 [554], followed by quantification of the IL10 levels using IL10 ELISA.

Taken together, after identification and description of potential functional relevant candidate genes in this thesis, the subsequent elucidation of their tumorigenic role in the lymphomagenesis of BL and mnBLL might help to understand their relevance for these lymphomas and, thus, their future suitability as potential specific therapeutic targets for the respective lymphoma subtypes.

5 Summary

Burkitt lymphomas (BL) account for about 25 % of the mature B-cell lymphomas in childhood. The hallmark of BL is the chromosomal translocation t(8;14)(q24;q32), juxtaposing the *MYC* gene to the *immunoglobulin* (*IG*) heavy chain locus, or one of its light chain variants. These translocations result in overexpression of the oncogenic transcription factor MYC. However, the *IG-MYC* translocation alone is not sufficient to drive lymphomagenesis. Additional alterations are necessary for malignant transformation like mutations within *CCND3* or *TP53*. In addition to *MYC*-positive BL, the existence of a subset of *MYC*-negative B-cell lymphoma resembling BL has been recently described. This lymphoma subset has been termed *MYC*-negative Burkitt-like lymphoma (mnBLL) and is characterized by a recurrent pattern of centromeric gain and telomeric loss on chromosome 11.

The aims of this thesis were the identification and characterization of secondary genetic alterations in BL contributing to its lymphomagenesis. Moreover, BL should be investigated for traces of pluripotency as MYC is one of four reprogramming factors able to induce pluripotency in somatic cells. With regard to mnBLL, the aim of this thesis was the thorough description of the chromosomal, mutational and transcriptional landscape of these lymphomas in order to identify potential candidate genes contributing to its lymphomagenesis, which further should be functionally characterized.

The analyses of BL were performed using 51 primary BL of the MMML cohort, which contains more than 800 molecularly characterized B-cell lymphomas, and of the ICGC MMML-Seq cohort. From the latter cohort, data from whole-genome and RNA-sequencing were available. Moreover, 25 mnBLL were analyzed of which 8 were newly recruited within this thesis. In addition, a panel of 23 B-cell lymphoma cell lines (16 BL, 3 mnBLL and 6 diffuse large B-cell lymphoma cell lines) was used for mutational screening and functional analyses. A variety of techniques was used to study the chromosomal (OncoScan DNA array, SNP 6.0 array), mutational (whole-genome, whole-exome and targeted re-sequencing, Sanger sequencing) and transcriptional (qPCR, Western blot, flow cytometry, IL10 ELISA, PluriTest) landscapes of BL and mnBLL as well as of the respective cell line models. Furthermore, the cell lines were used for various functional analyses including cell cycle and viability assays, gene knock down using lentiviral shRNA vectors or inhibition of IL10 signaling components using neutralizing IL10RA and IL10 antibodies as well as the PI3K inhibitor BKM120.

Application of the PluriTest algorithm on gene expression data of 221 mature aggressive B-cell lymphomas demonstrated that BL showed a higher similarity to pluripotent stem cells (PSC) than non-BL. Nevertheless, an origin from or resemblance to cells with PSC features could not be established for BL. Analyses of the transcriptome of the BL led to the identification of a new, so far not described exon within the *TERT* gene. The expression of this new exon was verified in 11/11 BL cell lines and in 7/8 primary BL. Modeling of the impact of the new exon on the TERT reading frame revealed that either (i) N- or C-terminal truncated proteins might be translated or (ii) the transcript

might be subjected to nonsense-mediated decay due to the introduction of a premature stop codon. Analyses of whole-genome sequencing data of BL and mutational screening of BL cell lines by Sanger sequencing led to the identification of frequently mutated genes so far not described in association with Burkitt lymphomagenesis. Hence, *ID3* mutations were identified in 35/53 (68 %) BL and 8/12 (67 %) BL cell lines. Re-expression or overexpression of *ID3* in BL cell lines induced the amount of cells in the pre-G1-phase as analyzed by cell cycle analysis, hence, further supporting the tumor suppressive function of ID3 in BL. Furthermore, 9/21 (43 %) BL and 2/16 (13 %) BL cell lines were shown to carry *SMARCA4* mutations. 62 % of these mutations affected the catalytic domain. Lastly, in 6/45 (13 %) BL and 6/16 (38 %) BL cell lines *PCBP1* mutations were identified which might lead to a cytoplasmatic retention and, hence, reduced function.

Analysis of the copy number data determined in all 25 mnBLL led to a redefinition of the minimal region of loss in 11q which includes a minimal region of focal homozygous loss of 1.5 Mb. Moreover, recurrent, secondary alterations in mnBLL affected 6q, 13q, 18q and 12p-q. Each of these altered regions harbored potential pathogenically relevant candidate genes. Furthermore, mutations of *ETS1* mapping to the region of loss in 11q were identified in 4/16 (25 %) mnBLL. Combining the results of the analyses of the chromosomal and mutational landscape with transcriptome data of 6 mnBLL provided by the MMML network project, and literature search led to the identification of 3 potential oncogenes (*IL10RA, PAFAH1B2, KMT2A*) and two potential tumor suppressor genes (*ETS1, FLI1*) in the altered 11q region. The expression profiles of these candidate genes were also analyzed in the mnBLL cell line models and were in agreement with those in the primary cases. Knock down analysis of PAFAH1B2 in a mnBLL cell line did not indicate an essential role of the PAFAH1B2 protein for the viability of the cells under the applied conditions. However, expression analysis of IL10 and IL10RA as well as inhibition of the IL10 signaling pathway components IL10, IL10RA and PI3K in B-cell lymphoma cell lines led to the identification of the IL10 signaling as a potential oncogenic signaling pathway in mnBLL. Hence, due to the interplay of IL10RA and FLI1/ETS1-induced IL10 deregulation, an autocrine IL10 signaling loop might be activated inducing PI3K signaling. The latter has been described as an oncogenic signaling pathway in BL.

Taken together, the analyses performed in this thesis led to the identification of secondary alterations likely contributing to the lymphomagenesis of BL. Moreover, the investigations led to a detailed characterization of *MYC*-negative mnBLL on chromosomal, mutational and transcriptional level resulting in the identification of potential candidate genes. Thus, the results of this thesis provide novel insights into the pathogenetic mechanisms contributing to lymphomagenesis and, hence, may help to improve diagnostics and therapy of these mature aggressive B-cell lymphomas of childhood.

6 Zusammenfassung

Ungefähr 25 % der reifzelligen B-Zelllymphome bei Kindern sind Burkitt Lymphome (BL). Charakteristisch für BL ist die chromosomale Translokation t(8;14)(q24;q32) oder eine ihrer Varianten. Durch diese gelangt das MYC Onkogen unter die Kontrolle der regulatorischen Einheiten des *Immunglobulin* Schwereketten oder eines der Leichtketten Gene, was zur Deregulation des Transkriptionsfaktors MYC führt. Die *MYC* Translokation reicht jedoch für die maligne Transformation nicht aus. Vielmehr sind sekundäre Veränderungen für die Entstehung eines BL notwendig, wie Mutationen in den Genen *CCND3* oder *TP53*. Kürzlich konnten auch BL-ähnliche *MYC*-negative B-Zelllymphome (mnBLL) identifiziert werden. Charakterisiert sind diese mnBLL durch einen Zugewinn des zentromeren und einen Verlust des telomeren Bereichs des langen Arms von Chromosom 11.

Ziele dieser Arbeit waren die Identifizierung und Charakterisierung von sekundären genetischen Veränderungen, die zu der Entstehung des BL beitragen. Außerdem sollten BL auf Anzeichen von Pluripotenz untersucht werden, da MYC einer von vier Faktoren ist, die Pluripotenz in somatischen Zellen induzieren können. Zudem sollten mnBLL umfassend im Bezug auf chromosomale, genetische und transkriptionelle Veränderungen charakterisiert werden, um potentielle Kandidatengene zu identifizieren, die zur Entstehung dieser Lymphome beitragen. Diese Kandidatengene sollten funktionell charakterisiert werden.

Für die Analyse der BL standen insgesamt 51 primäre BL aus den ICGC MMML-Seq und MMML Kohorten zur Verfügung, wobei letztere insgesamt mehr als 800 molekular charakterisierte B-Zell-Lymphome umfasst. Für einige Fälle der ICGC MMML-Seq Kohorte waren zudem Gesamtgenom- und RNA-Sequenzierungsdaten verfügbar. Für die Charakterisierung der mnBLL wurden 25 Fälle untersucht, von denen 8 im Rahmen dieser Arbeit identifiziert wurden. Mutations- und funktionelle Analysen zu Kandidatengenen wurden an bis zu 23 B-Zell-Lymphom Zelllinien (ZL) (ZL: 16 BL, 3 mnBLL und 6 diffus-großzellige B-Zell-Lymphome) durchgeführt. Mit Hilfe verschiedener Methoden wurden chromosomale (DNA-Array: OncoScan, SNP 6.0), genetische (Gesamtgenom-, Gesamtexom-, Re- und Sanger Sequenzierung) und transkriptionelle (qPCR, Western Blot, IL10 ELISA, PluriTest, Durchfluss-zytometrie) Veränderungen in BL, mnBLL und den entsprechenden ZL untersucht. Zudem wurden funktionelle Analysen an ZL durchgeführt, wie z.B. Zellzyklus- und Zellviabilitätsuntersuchungen, *knock down* mittels lentiviraler shRNA Vektoren sowie Inhibition von Komponenten des IL10-Signalwegs mit neutralisierenden IL10- und IL10RA- Antikörpern und dem PI3K-Inhibitor BKM120.

Für die Pluripotenz-Analyse der BL wurden Genexpressionsdaten von 221 reifzelligen, aggressiven B-Zelllymphomen mittels PluriTest ausgewertet. Dabei wurde gezeigt, dass BL den pluripotenten Stammzellen (PSZ) ähnlicher sind als Nicht-BL. Eine Abstammung bzw. Ähnlichkeit der BL mit PSZ-ähnlichen Zellen konnte jedoch nicht nachgewiesen werden. Im Rahmen der Analyse der Transkriptomdaten der BL konnte ein neues, bisher nicht beschriebenes Exon des *TERT* Gens

identifiziert werden. Dessen Expression wurde in 11/11 BL ZL und in 7/8 primären BL validiert. Vorhersagen über den Einfluss des neuen Exons auf das Leseraster des *TERT* Gens ergaben, dass entweder (i) C- oder N-terminal trunkierte Proteine translatiert oder (ii) aufgrund eines vorzeitigen Stopcodons Transkripte durch *nonsense-mediated decay* degradiert werden. Durch die Auswertung der Gesamtgenomsequenzierung der BL sowie durch gezielte Sanger Sequenzierung konnten häufig mutierter Gene identifiziert werden, die bisher nicht in Zusammenhang mit BL beschrieben wurden. So konnten *ID3* Mutationen in 35/53 (68 %) BL sowie in 8/12 (67 %) BL ZL nachgewiesen werden. In Übereinstimmung mit einer tumorsuppressiven Rolle von ID3 wurden in Zellzyklusanalysen in BL ZL nach ID3 Re- bzw. Über-Expression vermehrt Zellen in der pre-G1 Phase nachgewiesen. Mutationen im *SMARCA4* Gen wurden in 9/21 (43 %) BL und in 2/16 (13 %) BL ZL identifiziert. Bei 62 % der Mutationen ist die katalytische Domäne des Proteins betroffen. Mutationen in *PCBP1* konnten in 6/45 (13 %) BL und 6/16 (38 %) BL ZL identifiziert werden. Diese könnte zu einer zytoplasmatische Retention führen und somit zu einer eingeschränkten Funktion.

Nach Analyse der chromosomalen Veränderungen aller 25 mnBLL wurde die Größe der minimalen Verlustregion in 11q und einer darin enthaltenen 1,5 Mb großen fokalen homozygoten Deletion eingegrenzt. Außerdem wurden wiederkehrende, sekundäre Imbalancen in den Regionen 6q, 13q, 18q und 12p-q und die darin enthaltenen potentiell pathogen relevanten Kandidatengene identifiziert. In 4/16 (25 %) mnBLL wurden Mutationen des *ETS1* Gens in der Verlustregion in 11q detektiert. Durch eine kombinierte Auswertung der Expressionsdaten von sechs mnBLL, der chromosomalen und genetischen Veränderungen der 11q Region sowie eine Literaturrecherche konnten drei potentielle Onkogene (*IL10RA, PAFAH1B2, KMT2A*) und zwei potentielle Tumorsuppressorgene (*ETS1, FLI1*) identifiziert werden. Deren Expressionsprofile in mnBLL ZL waren vergleichbar mit denen der primären mnBLL. Für *PAFAH1B2* konnte durch *knock down* Experimente keine essentielle Rolle für die Viabilität von mnBLL Zellen nachgewiesen werden. Jedoch zeigen Expressionsanalysen von IL10 und IL10RA, sowie die Inhibition von IL10RA, IL10 und PI3K, dass durch das Zusammenspiel von IL10RA- und FLI1/ETS1-induzierter IL10-Überexpression ein autokriner IL10 Signalweg aktiviert werden könnte, der den bereits bei BL als onkogen bekannten PI3K Signalweg aktiviert.

Zusammengefasst haben die hier durchgeführten Analysen zur Identifizierung von sekundären Veränderungen geführt, die möglicherweise zur Entstehung der BL beitragen. Außerdem wurden mnBLL umfassend auf chromosomaler, genetischer und transkriptioneller Ebene charakterisiert, wodurch potentielle Kandidatengene für deren Pathogenese identifiziert wurden. Somit tragen die Ergebnisse dieser Arbeit zu einem breiteren Verständnis der pathogenen Mechanismen bei, die zur Entstehung dieser B-Zelllymphome führen. Diese Ergebnisse können in Zukunft zur Verbesserung der Diagnostik und Therapie dieser Lymphome beitragen.

7 References

1. Swerdlow SH, International Agency for Research on Cancer, World Health Organization. WHO classification of tumours of haematopoietic and lymphoid tissues. Lyon, France: International Agency for Research on Cancer; 2008.

2. A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project. Blood. 1997;89: 3909–3918.

3. Armitage JO, Weisenburger DD. New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project. J Clin Oncol Off J Am Soc Clin Oncol. 1998;16: 2780–2795.

4. Küppers R. Mechanisms of B-cell lymphoma pathogenesis. Nat Rev Cancer. 2005;5: 251–262. doi:10.1038/nrc1589

5. Anderson JR, Armitage JO, Weisenburger DD. Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project. Ann Oncol Off J Eur Soc Med Oncol ESMO. 1998;9: 717–720.

6. Morton LM, Wang SS, Devesa SS, Hartge P, Weisenburger DD, Linet MS. Lymphoma incidence patterns by WHO subtype in the United States, 1992-2001. Blood. 2006;107: 265–276. doi:10.1182/blood-2005-06-2508

7. Miller RW, Young JL, Novakovic B. Childhood cancer. Cancer. 1995;75: 395–405.

8. Downing JR, Wilson RK, Zhang J, Mardis ER, Pui C-H, Ding L, et al. The Pediatric Cancer Genome Project. Nat Genet. 2012;44: 619–622. doi:10.1038/ng.2287

9. Scotting PJ, Walker DA, Perilongo G. Childhood solid tumours: a developmental disorder. Nat Rev Cancer. 2005;5: 481–488. doi:10.1038/nrc1633

10. Kaatsch P, Spix, C. German Childhood Cancer Registry, Annual Report 2013/2014 (1980-2013). Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI) at the University Medical Center of the Johannes Gutenberg University Mainz; 2014.

11. Oschlies I, Klapper W. Maligne Lymphome im Kindes- und Jugendalter: Praktisches Wissen für den diagnostischen Alltag. Pathol. 2014;35: 383–398. doi:10.1007/s00292-014-1891-5

12. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin. 2009;59: 225–249. doi:10.3322/caac.20006

13. Magrath I. Epidemiology: clues to the pathogenesis of Burkitt lymphoma. Br J Haematol. 2012;156: 744–756. doi:10.1111/j.1365-2141.2011.09013.x

14. Rochford R, Cannon MJ, Moormann AM. Endemic Burkitt's lymphoma: a polymicrobial disease? Nat Rev Microbiol. 2005;3: 182–187. doi:10.1038/nrmicro1089

15. Orem J, Mbidde EK, Lambert B, de Sanjose S, Weiderpass E. Burkitt's lymphoma in Africa, a review of the epidemiology and etiology. Afr Health Sci. 2007;7: 166–175. doi:10.5555/afhs.2007.7.3.166

16. Tao Q, Robertson KD, Manns A, Hildesheim A, Ambinder RF. Epstein-Barr virus (EBV) in endemic Burkitt's lymphoma: molecular analysis of primary tumor tissue. Blood. 1998;91: 1373–1381.

17. Roithmann S, Toledano M, Tourani JM, Raphael M, Gentilini M, Gastaut JA, et al. HIV-associated non-Hodgkin's lymphomas: clinical characteristics and outcome. The experience of the French Registry of HIV-associated tumors. Ann Oncol Off J Eur Soc Med Oncol ESMO. 1991;2: 289–295.

18. Weintraub J, Warnke RA. Lymphoma in cardiac allotransplant recipients. Clinical and histological features and immunological phenotype. Transplantation. 1982;33: 347–351.

19. Wilkinson AH, Smith JL, Hunsicker LG, Tobacman J, Kapelanski DP, Johnson M, et al. Increased frequency of posttransplant lymphomas in patients treated with cyclosporine, azathioprine, and prednisone. Transplantation. 1989;47: 293–296.

20. Jaglowski SM, Linden E, Termuhlen AM, Flynn JM. Lymphoma in adolescents and young adults. Semin Oncol. 2009;36: 381–418. doi:10.1053/j.seminoncol.2009.07.009

21. Boerma EG, van Imhoff GW, Appel IM, Veeger NJGM, Kluin PM, Kluin-Nelemans JC. Gender and age-related differences in Burkitt lymphoma--epidemiological and clinical data from The Netherlands. Eur J Cancer Oxf Engl 1990. 2004;40: 2781–2787. doi:10.1016/j.ejca.2004.09.004

22. Burkhardt B, Oschlies I, Klapper W, Zimmermann M, Woessmann W, Meinhardt A, et al. Non-Hodgkin's lymphoma in adolescents: experiences in 378 adolescent NHL patients treated according to pediatric NHL-BFM protocols. Leukemia. 2011;25: 153–160. doi:10.1038/leu.2010.245

23. Mbulaiteye SM, Anderson WF, Ferlay J, Bhatia K, Chang C, Rosenberg PS, et al. Pediatric, elderly, and emerging adult-onset peaks in Burkitt's lymphoma incidence diagnosed in four continents, excluding Africa. Am J Hematol. 2012;87: 573–578. doi:10.1002/ajh.23187

24. Klapper W, Szczepanowski M, Burkhardt B, Berger H, Rosolowski M, Bentink S, et al. Molecular profiling of pediatric mature B-cell lymphoma treated in population-based prospective clinical trials. Blood. 2008;112: 1374–1381. doi:10.1182/blood-2008-01-136465

25. Diviné M, Casassus P, Koscielny S, Bosq J, Sebban C, Le Maignan C, et al. Burkitt lymphoma in adults: a prospective study of 72 patients treated with an adapted pediatric LMB protocol. Ann Oncol Off J Eur Soc Med Oncol ESMO. 2005;16: 1928–1935. doi:10.1093/annonc/mdi403

26. Hoelzer D, Ludwig WD, Thiel E, Gassmann W, Löffler H, Fonatsch C, et al. Improved outcome in adult B-cell acute lymphoblastic leukemia. Blood. 1996;87: 495–508.

27. Blum KA, Lozanski G, Byrd JC. Adult Burkitt leukemia and lymphoma. Blood. 2004;104: 3009–3020. doi:10.1182/blood-2004-02-0405

28. Mead GM, Sydes MR, Walewski J, Grigg A, Hatton CS, Pescosta N, et al. An international evaluation of CODOX-M and CODOX-M alternating with IVAC in adult Burkitt's lymphoma: results of United Kingdom Lymphoma Group LY06 study. Ann Oncol Off J Eur Soc Med Oncol ESMO. 2002;13: 1264–1274.

29. Beck K, Peak MM, Ota T, Nemazee D, Murre C. Distinct roles for E12 and E47 in B cell specification and the sequential rearrangement of immunoglobulin light chain loci. J Exp Med. 2009;206: 2271–2284. doi:10.1084/jem.20090756

30. Dias S, Silva H, Cumano A, Vieira P. Interleukin-7 is necessary to maintain the B cell potential in common lymphoid progenitors. J Exp Med. 2005;201: 971–979. doi:10.1084/jem.20042393

31. Corfe SA, Paige CJ. The many roles of IL-7 in B cell development; mediator of survival, proliferation and differentiation. Semin Immunol. 2012;24: 198–208. doi:10.1016/j.smim.2012.02.001

32. Grimaldi CM, Hicks R, Diamond B. B cell selection and susceptibility to autoimmunity. J Immunol Baltim Md 1950. 2005;174: 1775–1781.

33. Gatto D, Brink R. The germinal center reaction. J Allergy Clin Immunol. 2010;126: 898–907; quiz 908–909. doi:10.1016/j.jaci.2010.09.007

34. Wabl M, Steinberg C. Affinity maturation and class switching. Curr Opin Immunol. 1996;8: 89–92.

35. Dominguez-Sola D, Victora GD, Ying CY, Phan RT, Saito M, Nussenzweig MC, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. Nat Immunol. 2012;13: 1083–1091. doi:10.1038/ni.2428

36. Calado DP, Sasaki Y, Godinho SA, Pellerin A, Köchert K, Sleckman BP, et al. The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. Nat Immunol. 2012;13: 1092–1100. doi:10.1038/ni.2418

37. Calame KL. Plasma cells: finding new light at the end of B cell development. Nat Immunol. 2001;2: 1103–1108. doi:10.1038/ni1201-1103

38. Teng Y, Takahashi Y, Yamada M, Kurosu T, Koyama T, Miura O, et al. IRF4 negatively regulates proliferation of germinal center B cell-derived Burkitt's lymphoma cell lines and induces differentiation toward plasma cells. Eur J Cell Biol. 2007;86: 581–589. doi:10.1016/j.ejcb.2007.05.006

39. Murphy K, Travers P, Walport M, Janeway C. Janeway's immunobiology. New York: Garland Science; 2012.

40. Paramithiotis E, Cooper MD. Memory B lymphocytes migrate to bone marrow in humans. Proc Natl Acad Sci U S A. 1997;94: 208–212.

41. Mamani-Matsuda M, Cosma A, Weller S, Faili A, Staib C, Garçon L, et al. The human spleen is a major reservoir for long-lived vaccinia virus-specific memory B cells. Blood. 2008;111: 4653–4659. doi:10.1182/blood-2007-11-123844

42. Ahmed R, Gray D. Immunological memory and protective immunity: understanding their relation. Science. 1996;272: 54–60.

43. Schroeder HW, Cavacini L. Structure and function of immunoglobulins. J Allergy Clin Immunol. 2010;125: S41–52. doi:10.1016/j.jaci.2009.09.046

44. De Groot C, Kapsenberg ML, Leene W. Observations on transmembrane structures of surface immunoglobulin in the plasma membrane of B lymphocytes. Biochim Biophys Acta. 1982;689: 275–282.

45. Kehry M, Ewald S, Douglas R, Sibley C, Raschke W, Fambrough D, et al. The immunoglobulin mu chains of membrane-bound and secreted IgM molecules differ in their C-terminal segments. Cell. 1980;21: 393–406.

46. Market E, Papavasiliou FN. V(D)J recombination and the evolution of the adaptive immune system. PLoS Biol. 2003;1: E16. doi:10.1371/journal.pbio.0000016

47. Lieber MR, Hesse JE, Mizuuchi K, Gellert M. Lymphoid V(D)J recombination: nucleotide insertion at signal joints as well as coding joints. Proc Natl Acad Sci U S A. 1988;85: 8588–8592.

48. Levy NS, Malipiero UV, Lebecque SG, Gearhart PJ. Early onset of somatic mutation in immunoglobulin VH genes during the primary immune response. J Exp Med. 1989;169: 2007–2019.

49. Muramatsu M, Sankaranand VS, Anant S, Sugai M, Kinoshita K, Davidson NO, et al. Specific expression of activation-induced cytidine deaminase (AID), a novel member of the RNA-editing deaminase family in germinal center B cells. J Biol Chem. 1999;274: 18470–18476.

50. Pasqualucci L, Guglielmino R, Houldsworth J, Mohr J, Aoufouchi S, Polakiewicz R, et al. Expression of the AID protein in normal and neoplastic B cells. Blood. 2004;104: 3318–3325. doi:10.1182/blood-2004-04-1558

51. Di Noia JM, Neuberger MS. Molecular mechanisms of antibody somatic hypermutation. Annu Rev Biochem. 2007;76: 1–22. doi:10.1146/annurev.biochem.76.061705.090740

52. Lieber MR, Yu K, Raghavan SC. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. DNA Repair. 2006;5: 1234–1245. doi:10.1016/j.dnarep.2006.05.013

53. Yan CT, Boboila C, Souza EK, Franco S, Hickernell TR, Murphy M, et al. IgH class switching and translocations use a robust non-classical end-joining pathway. Nature. 2007;449: 478–482. doi:10.1038/nature06020

54. Honjo T, Kinoshita K, Muramatsu M. Molecular mechanism of class switch recombination: linkage with somatic hypermutation. Annu Rev Immunol. 2002;20: 165–196. doi:10.1146/annurev.immunol.20.090501.112049

55. Stavnezer J, Schrader CE. IgH chain class switch recombination: mechanism and regulation. J Immunol Baltim Md 1950. 2014;193: 5370–5378. doi:10.4049/jimmunol.1401849

56. Küppers R, Klein U, Hansmann ML, Rajewsky K. Cellular origin of human B-cell lymphomas. N Engl J Med. 1999;341: 1520–1529. doi:10.1056/NEJM199911113412007

57. Tsujimoto Y, Gorham J, Cossman J, Jaffe E, Croce CM. The t(14;18) chromosome translocations involved in B-cell neoplasms result from mistakes in VDJ joining. Science. 1985;229: 1390–1393.

58. Jäger U, Böcskör S, Le T, Mitterbauer G, Bolz I, Chott A, et al. Follicular lymphomas' BCL-2/IgH junctions contain templated nucleotide insertions: novel insights into the mechanism of t(14;18) translocation. Blood. 2000;95: 3520–3529.

59. Klein U, Dalla-Favera R. Germinal centres: role in B-cell physiology and malignancy. Nat Rev Immunol. 2008;8: 22–33. doi:10.1038/nri2217

60. Weller S, Faili A, Garcia C, Braun MC, Le Deist F F, de Saint Basile G G, et al. CD40-CD40L independent Ig gene hypermutation suggests a second B cell diversification pathway in humans. Proc Natl Acad Sci U S A. 2001;98: 1166–1170. doi:10.1073/pnas.98.3.1166

61. Shaffer AL, Rosenwald A, Staudt LM. Lymphoid malignancies: the dark side of B-cell differentiation. Nat Rev Immunol. 2002;2: 920–932. doi:10.1038/nri953

62. Hardianti MS, Tatsumi E, Syampurnawati M, Furuta K, Saigo K, Nakamachi Y, et al. Activation-induced cytidine deaminase expression in follicular lymphoma: association between AID expression and ongoing mutation in FL. Leukemia. 2004;18: 826–831. doi:10.1038/sj.leu.2403323

63. Lossos IS, Alizadeh AA, Eisen MB, Chan WC, Brown PO, Botstein D, et al. Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas. Proc Natl Acad Sci U S A. 2000;97: 10209–10213. doi:10.1073/pnas.180316097

64. Rolink A, Melchers F. Molecular and cellular origins of B lymphocyte diversity. Cell. 1991;66: 1081–1094.

65. Klein U, Tu Y, Stolovitzky GA, Keller JL, Haddad J, Miljkovic V, et al. Transcriptional analysis of the B cell germinal center reaction. Proc Natl Acad Sci U S A. 2003;100: 2639–2644. doi:10.1073/pnas.0437996100

66. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403: 503–511. doi:10.1038/35000501

67. Dave SS, Fu K, Wright GW, Lam LT, Kluin P, Boerma E-J, et al. Molecular diagnosis of Burkitt's lymphoma. N Engl J Med. 2006;354: 2431–2442. doi:10.1056/NEJMoa055759

68. Victora GD, Dominguez-Sola D, Holmes AB, Deroubaix S, Dalla-Favera R, Nussenzweig MC. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. Blood. 2012;120: 2240–2248. doi:10.1182/blood-2012-03-415380

69. Robbiani DF, Bothmer A, Callen E, Reina-San-Martin B, Dorsett Y, Difilippantonio S, et al. AID is required for the chromosomal breaks in c-myc that lead to c-myc/IgH translocations. Cell. 2008;135: 1028–1038. doi:10.1016/j.cell.2008.09.062

70. Ramiro AR, Jankovic M, Eisenreich T, Difilippantonio S, Chen-Kiang S, Muramatsu M, et al. AID is required for c-myc/IgH chromosome translocations in vivo. Cell. 2004;118: 431–438. doi:10.1016/j.cell.2004.08.006

71. Ott G, Rosenwald A, Campo E. Understanding MYC-driven aggressive B-cell lymphomas: pathogenesis and classification. Hematol Educ Program Am Soc Hematol Am Soc Hematol Educ Program. 2013;2013: 575–583. doi:10.1182/asheducation-2013.1.575

72. Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M, et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature. 2012;490: 116–120. doi:10.1038/nature11378

73. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Kinzler KW. Cancer genome landscapes. Science. 2013;339: 1546–1558. doi:10.1126/science.1235122

74. Adamson ED. Oncogenes in development. Dev Camb Engl. 1987;99: 449–471.

75. Lodish HF. Molecular cell biology. New York: W.H. Freeman; 2000.

76. Santarius T, Shipley J, Brewer D, Stratton MR, Cooper CS. A census of amplified and overexpressed human cancer genes. Nat Rev Cancer. 2010;10: 59–64. doi:10.1038/nrc2771

77. Hagemeijer A, van der Plas DC, Soekarman D, van Denderen J, Grosveld G. The Philadelphia translocation in CML and ALL: recent investigations, new detection methods. Nouv Rev Fr Hématologie. 1990;32: 83–86.

78. Rajan L, Broussard D, Lozano M, Lee CG, Kozak CA, Dudley JP. The c-myc locus is a common integration site in type B retrovirus-induced T-cell lymphomas. J Virol. 2000;74: 2466–2471.

79. Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, et al. Translocation of the c-myc gene into the immunoglobulin heavy chain locus in human Burkitt lymphoma and murine plasmacytoma cells. Proc Natl Acad Sci U S A. 1982;79: 7837–7841.

80. Ladanyi M, Offit K, Jhanwar SC, Filippa DA, Chaganti RS. MYC rearrangement and translocations involving band 8q24 in diffuse large cell lymphomas. Blood. 1991;77: 1057–1063.

81. Bakhshi A, Jensen JP, Goldman P, Wright JJ, McBride OW, Epstein AL, et al. Cloning the chromosomal breakpoint of t(14;18) human lymphomas: clustering around JH on chromosome 14 and near a transcriptional unit on 18. Cell. 1985;41: 899–906.

82. Cleary ML, Sklar J. Nucleotide sequence of a t(14;18) chromosomal breakpoint in follicular lymphoma and demonstration of a breakpoint-cluster region near a transcriptionally active locus on chromosome 18. Proc Natl Acad Sci U S A. 1985;82: 7439–7443.

83. Rosenberg CL, Wong E, Petty EM, Bale AE, Tsujimoto Y, Harris NL, et al. PRAD1, a candidate BCL1 oncogene: mapping and expression in centrocytic lymphoma. Proc Natl Acad Sci U S A. 1991;88: 9638–9642.

84. Bosch F, Jares P, Campo E, Lopez-Guillermo A, Piris MA, Villamor N, et al. PRAD-1/cyclin D1 gene overexpression in chronic lymphoproliferative disorders: a highly specific marker of mantle cell lymphoma. Blood. 1994;84: 2726–2732.

85. Martín-Subero JI, Gesk S, Harder L, Sonoki T, Tucker PW, Schlegelberger B, et al. Recurrent involvement of the REL and BCL11A loci in classical Hodgkin lymphoma. Blood. 2002;99: 1474–1477.

86. Lowe SW, Cepero E, Evan G. Intrinsic tumour suppression. Nature. 2004;432: 307–315. doi:10.1038/nature03098

87. Motoyama N, Naka K. DNA damage tumor suppressor genes and genomic instability. Curr Opin Genet Dev. 2004;14: 11–16. doi:10.1016/j.gde.2003.12.003

88. Rocco JW, Sidransky D. p16(MTS-1/CDKN2/INK4a) in cancer progression. Exp Cell Res. 2001;264: 42–55. doi:10.1006/excr.2000.5149

89. O'Connor PM, Jackman J, Jondle D, Bhatia K, Magrath I, Kohn KW. Role of the p53 tumor suppressor gene in cell cycle arrest and radiosensitivity of Burkitt's lymphoma cell lines. Cancer Res. 1993;53: 4776–4780.

90. Huang J, Zhao Y-L, Li Y, Fletcher JA, Xiao S. Genomic and functional evidence for an ARID1A tumor suppressor role. Genes Chromosomes Cancer. 2007;46: 745–750. doi:10.1002/gcc.20459

91. Giulino-Roth L, Wang K, MacDonald TY, Mathew S, Tam Y, Cronin MT, et al. Targeted genomic sequencing of pediatric Burkitt lymphoma identifies recurrent alterations in antiapoptotic and chromatin-remodeling genes. Blood. 2012;120: 5181–5184. doi:10.1182/blood-2012-06-437624

92. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A. 1971;68: 820–823.

93. Cook WD, McCaw BJ. Accommodating haploinsufficient tumor suppressor genes in Knudson's model. Oncogene. 2000;19: 3434–3438. doi:10.1038/sj.onc.1203653

94. Fero ML, Randel E, Gurley KE, Roberts JM, Kemp CJ. The murine gene p27Kip1 is haplo-insufficient for tumour suppression. Nature. 1998;396: 177–180. doi:10.1038/24179

95. Kwabi-Addo B, Giri D, Schmidt K, Podsypanina K, Parsons R, Greenberg N, et al. Haploinsufficiency of the Pten tumor suppressor gene promotes prostate cancer progression. Proc Natl Acad Sci U S A. 2001;98: 11563–11568. doi:10.1073/pnas.201167798

96. Seidman JG, Seidman C. Transcription factor haploinsufficiency: when half a loaf is not enough. J Clin Invest. 2002;109: 451–455. doi:10.1172/JCI15043

97. Mestre-Escorihuela C, Rubio-Moscardo F, Richter JA, Siebert R, Climent J, Fresquet V, et al. Homozygous deletions localize novel tumor suppressor genes in B-cell lymphomas. Blood. 2007;109: 271–280. doi:10.1182/blood-2006-06-026500

98. Lindström MS, Klangby U, Wiman KG. p14ARF homozygous deletion or MDM2 overexpression in Burkitt lymphoma lines carrying wild type p53. Oncogene. 2001;20: 2171–2177. doi:10.1038/sj.onc.1204303

99. Paglia LL, Laugé A, Weber J, Champ J, Cavaciuti E, Russo A, et al. ATM germline mutations in women with familial breast cancer and a relative with haematological malignancy. Breast Cancer Res Treat. 2010;119: 443–452. doi:10.1007/s10549-009-0396-z

100. Gumy Pause F, Wacker P, Maillet P, Betts D, Sappino A-P. ATM gene alterations in childhood acute lymphoblastic leukemias. Hum Mutat. 2003;21: 554. doi:10.1002/humu.9140

101. Gaidano G, Ballerini P, Gong JZ, Inghirami G, Neri A, Newcomb EW, et al. p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. Proc Natl Acad Sci U S A. 1991;88: 5413–5417.

102. Møller MB, Ino Y, Gerdes AM, Skjødt K, Louis DN, Pedersen NT. Aberrations of the p53 pathway components p53, MDM2 and CDKN2A appear independent in diffuse large B cell lymphoma. Leukemia. 1999;13: 453–459.

103. Grønbaek K, Worm J, Ralfkiaer E, Ahrenkiel V, Hokland P, Guldberg P. ATM mutations are associated with inactivation of the ARF-TP53 tumor suppressor pathway in diffuse large B-cell lymphoma. Blood. 2002;100: 1430–1437. doi:10.1182/blood-2002-02-0382

104. Hecht JL, Aster JC. Molecular biology of Burkitt's lymphoma. J Clin Oncol Off J Am Soc Clin Oncol. 2000;18: 3707–3721.

105. Boxer LM, Dang CV. Translocations involving c-myc and c-myc function. Oncogene. 2001;20: 5595–5610. doi:10.1038/sj.onc.1204595

106. Haluska FG, Finver S, Tsujimoto Y, Croce CM. The t(8; 14) chromosomal translocation occurring in B-cell malignancies results from mistakes in V-D-J joining. Nature. 1986;324: 158–161. doi:10.1038/324158a0

107. Haluska FG, Tsujimoto Y, Croce CM. The t(8;14) chromosome translocation of the Burkitt lymphoma cell line Daudi occurred during immunoglobulin gene rearrangement and involved the heavy chain diversity region. Proc Natl Acad Sci U S A. 1987;84: 6835–6839.

108. Hummel M, Bentink S, Berger H, Klapper W, Wessendorf S, Barth TFE, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med. 2006;354: 2419–2430. doi:10.1056/NEJMoa055351

109. Hayday AC, Gillies SD, Saito H, Wood C, Wiman K, Hayward WS, et al. Activation of a translocated human c-myc gene by an enhancer in the immunoglobulin heavy-chain locus. Nature. 1984;307: 334–340.

110. ar-Rushdi A, Nishikura K, Erikson J, Watt R, Rovera G, Croce CM. Differential expression of the translocated and the untranslocated c-myc oncogene in Burkitt lymphoma. Science. 1983;222: 390–393.

111. Bemark M, Neuberger MS. The c-MYC allele that is translocated into the IgH locus undergoes constitutive hypermutation in a Burkitt's lymphoma line. Oncogene. 2000;19: 3404–3410. doi:10.1038/sj.onc.1203686

112. Dominguez-Sola D, Victora GD, Ying CY, Phan RT, Saito M, Nussenzweig MC, et al. The proto-oncogene MYC is required for selection in the germinal center and cyclic reentry. Nat Immunol. 2012;13: 1083–1091. doi:10.1038/ni.2428

113. Nahar R, Ramezani-Rad P, Mossner M, Duy C, Cerchietti L, Geng H, et al. Pre-B cell receptor-mediated activation of BCL6 induces pre-B cell quiescence through transcriptional repression of MYC. Blood. 2011;118: 4174–4178. doi:10.1182/blood-2011-01-331181

114. Gregory MA, Qi Y, Hann SR. Phosphorylation by glycogen synthase kinase-3 controls c-myc proteolysis and subnuclear localization. J Biol Chem. 2003;278: 51606–51612. doi:10.1074/jbc.M310722200

115. Lüscher B, Eisenman RN. New light on Myc and Myb. Part I. Myc. Genes Dev. 1990;4: 2025–2035.

116. Alt FW, DePinho R, Zimmerman K, Legouy E, Hatton K, Ferrier P, et al. The human myc gene family. Cold Spring Harb Symp Quant Biol. 1986;51 Pt 2: 931–941.

117. Watt R, Nishikura K, Sorrentino J, ar-Rushdi A, Croce CM, Rovera G. The structure and nucleotide sequence of the 5' end of the human c-myc oncogene. Proc Natl Acad Sci U S A. 1983;80: 6307–6311.

118. Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, et al. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell. 2012;151: 68–79. doi:10.1016/j.cell.2012.08.033

119. Lin CY, Lovén J, Rahl PB, Paranal RM, Burge CB, Bradner JE, et al. Transcriptional amplification in tumor cells with elevated c-Myc. Cell. 2012;151: 56–67. doi:10.1016/j.cell.2012.08.026

120. Chandriani S, Frengen E, Cowling VH, Pendergrass SA, Perou CM, Whitfield ML, et al. A core MYC gene expression signature is prominent in basal-like breast cancer but only partially overlaps the core serum response. PloS One. 2009;4: e6693. doi:10.1371/journal.pone.0006693

121. Sander S, Bullinger L, Klapproth K, Fiedler K, Kestler HA, Barth TFE, et al. MYC stimulates EZH2 expression by repression of its negative regulator miR-26a. Blood. 2008;112: 4202–4212. doi:10.1182/blood-2008-03-147645

122. Coller HA, Grandori C, Tamayo P, Colbert T, Lander ES, Eisenman RN, et al. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. Proc Natl Acad Sci U S A. 2000;97: 3260–3265.

123. Greasley PJ, Bonnard C, Amati B. Myc induces the nucleolin and BN51 genes: possible implications in ribosome biogenesis. Nucleic Acids Res. 2000;28: 446–453.

124. Reisman D, Elkind NB, Roy B, Beamon J, Rotter V. c-Myc trans-activates the p53 promoter through a required downstream CACGTG motif. Cell Growth Differ Mol Biol J Am Assoc Cancer Res. 1993;4: 57–65.

125. Zindy F, Eischen CM, Randle DH, Kamijo T, Cleveland JL, Sherr CJ, et al. Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. Genes Dev. 1998;12: 2424–2433.

126. Eischen CM, Weber JD, Roussel MF, Sherr CJ, Cleveland JL. Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis. Genes Dev. 1999;13: 2658–2669.

127. Adams JM, Harris AW, Pinkert CA, Corcoran LM, Alexander WS, Cory S, et al. The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. Nature. 1985;318: 533–538.

128. Harris AW, Pinkert CA, Crawford M, Langdon WY, Brinster RL, Adams JM. The E mu-myc transgenic mouse. A model for high-incidence spontaneous lymphoma and leukemia of early B cells. J Exp Med. 1988;167: 353–371.

129. Park SS, Kim JS, Tessarollo L, Owens JD, Peng L, Han SS, et al. Insertion of c-Myc into Igh induces B-cell and plasma-cell neoplasms in mice. Cancer Res. 2005;65: 1306–1315. doi:10.1158/0008-5472.CAN-04-0268

130. Martins CP, Brown-Swigart L, Evan GI. Modeling the therapeutic efficacy of p53 restoration in tumors. Cell. 2006;127: 1323–1334. doi:10.1016/j.cell.2006.12.007

131. Boerma EG, Siebert R, Kluin PM, Baudis M. Translocations involving 8q24 in Burkitt lymphoma and other malignant lymphomas: a historical review of cytogenetics in the light of todays knowledge. Leukemia. 2009;23: 225–234. doi:10.1038/leu.2008.281

132. Salaverria I, Zettl A, Beà S, Hartmann EM, Dave SS, Wright GW, et al. Chromosomal alterations detected by comparative genomic hybridization in subgroups of gene expression-defined Burkitt's lymphoma. Haematologica. 2008;93: 1327–1334. doi:10.3324/haematol.13071

133. Sander S, Calado DP, Srinivasan L, Köchert K, Zhang B, Rosolowski M, et al. Synergy between PI3K signaling and MYC in Burkitt lymphomagenesis. Cancer Cell. 2012;22: 167–179. doi:10.1016/j.ccr.2012.06.012

134. Klangby U, Okan I, Magnusson KP, Wendland M, Lind P, Wiman KG. p16/INK4a and p15/INK4b gene methylation and absence of p16/INK4a mRNA and protein expression in Burkitt's lymphoma. Blood. 1998;91: 1680–1687.

135. Schmitz R, Ceribelli M, Pittaluga S, Wright G, Staudt LM. Oncogenic mechanisms in Burkitt lymphoma. Cold Spring Harb Perspect Med. 2014;4. doi:10.1101/cshperspect.a014282

136. Shammas MA. Telomeres, lifestyle, cancer, and aging. Curr Opin Clin Nutr Metab Care. 2011;14: 28–34. doi:10.1097/MCO.0b013e32834121b1

137. Hu BT, Lee SC, Marin E, Ryan DH, Insel RA. Telomerase is up-regulated in human germinal center B cells in vivo and can be re-expressed in memory B cells activated in vitro. J Immunol Baltim Md 1950. 1997;159: 1068–1071.

138. Weng NP, Granger L, Hodes RJ. Telomere lengthening and telomerase activation during human B cell differentiation. Proc Natl Acad Sci U S A. 1997;94: 10827–10832.

139. Klapper W, Krams M, Qian W, Janssen D, Parwaresch R. Telomerase activity in B-cell non-Hodgkin lymphomas is regulated by hTERT transcription and correlated with telomere-binding protein expression but uncoupled from proliferation. Br J Cancer. 2003;89: 713–719. doi:10.1038/sj.bjc.6601112

140. Blackburn EH, Greider CW, Henderson E, Lee MS, Shampay J, Shippen-Lentz D. Recognition and elongation of telomeres by telomerase. Genome Natl Res Counc Can Génome Cons Natl Rech Can. 1989;31: 553–560.

141. Greider CW, Blackburn EH. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. Cell. 1985;43: 405–413.

142. Feng J, Funk WD, Wang SS, Weinrich SL, Avilion AA, Chiu CP, et al. The RNA component of human telomerase. Science. 1995;269: 1236–1241.

143. Greenberg RA, O'Hagan RC, Deng H, Xiao Q, Hann SR, Adams RR, et al. Telomerase reverse transcriptase gene is a direct target of c-Myc but is not functionally equivalent in cellular transformation. Oncogene. 1999;18: 1219–1226. doi:10.1038/sj.onc.1202669

144. Wang J, Xie LY, Allan S, Beach D, Hannon GJ. Myc activates telomerase. Genes Dev. 1998;12: 1769–1774.

145. Pope JH, Horne MK, Scott W. Transformation of foetal human keukocytes in vitro by filtrates of a human leukaemic cell line containing herpes-like virus. Int J Cancer J Int Cancer. 1968;3: 857–866.

146. Kang M-S, Soni V, Bronson R, Kieff E. Epstein-Barr virus nuclear antigen 1 does not cause lymphoma in C57BL/6J mice. J Virol. 2008;82: 4180–4183. doi:10.1128/JVI.02596-07

147. Kelly GL, Milner AE, Baldwin GS, Bell AI, Rickinson AB. Three restricted forms of Epstein-Barr virus latency counteracting apoptosis in c-myc-expressing Burkitt lymphoma cells. Proc Natl Acad Sci U S A. 2006;103: 14935–14940. doi:10.1073/pnas.0509988103

148. Kelly GL, Milner AE, Tierney RJ, Croom-Carter DSG, Altmann M, Hammerschmidt W, et al. Epstein-Barr virus nuclear antigen 2 (EBNA2) gene deletion is consistently linked with EBNA3A, -3B, and -3C expression in Burkitt's lymphoma cells and with increased resistance to apoptosis. J Virol. 2005;79: 10709–10717. doi:10.1128/JVI.79.16.10709-10717.2005

149. Nanbo A, Inoue K, Adachi-Takasawa K, Takada K. Epstein-Barr virus RNA confers resistance to interferon-alpha-induced apoptosis in Burkitt's lymphoma. EMBO J. 2002;21: 954–965. doi:10.1093/emboj/21.5.954

150. Rohn JL, Hueber AO, McCarthy NJ, Lyon D, Navarro P, Burgering BM, et al. The opposing roles of the Akt and c-Myc signalling pathways in survival from CD95-mediated apoptosis. Oncogene. 1998;17: 2811–2818. doi:10.1038/sj.onc.1202393

151. Han S-S, Yun H, Son D-J, Tompkins VS, Peng L, Chung S-T, et al. NF-kappaB/STAT3/PI3K signaling crosstalk in iMyc E mu B lymphoma. Mol Cancer. 2010;9: 97. doi:10.1186/1476-4598-9-97

152. Bouchard C, Marquardt J, Brás A, Medema RH, Eilers M. Myc-induced proliferation and transformation require Akt-mediated phosphorylation of FoxO proteins. EMBO J. 2004;23: 2830–2840. doi:10.1038/sj.emboj.7600279

153. Curnock AP, Knox KA. LY294002-mediated inhibition of phosphatidylinositol 3-kinase activity triggers growth inhibition and apoptosis in CD40-triggered Ramos-Burkitt lymphoma B cells. Cell Immunol. 1998;187: 77–87. doi:10.1006/cimm.1998.1335

154. Tuveson DA, Carter RH, Soltoff SP, Fearon DT. CD19 of B cells as a surrogate kinase insert region to bind phosphatidylinositol 3-kinase. Science. 1993;260: 986–989.

155. Vivanco I, Sawyers CL. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. Nat Rev Cancer. 2002;2: 489–501. doi:10.1038/nrc839

156. McConechy MK, Ding J, Cheang MCU, Wiegand KC, Senz J, Tone AA, et al. Use of mutation profiles to refine the classification of endometrial carcinomas. J Pathol. 2012;228: 20–30. doi:10.1002/path.4056

157. Yahiaoui OI, Nunès JA, Castanier C, Devillier R, Broussais F, Fabre AJ, et al. Constitutive AKT activation in follicular lymphoma. BMC Cancer. 2014;14: 565. doi:10.1186/1471-2407-14-565

158. Okkenhaug K, Vanhaesebroeck B. PI3K in lymphocyte development, differentiation and activation. Nat Rev Immunol. 2003;3: 317–330. doi:10.1038/nri1056

159. Welcker M, Orian A, Jin J, Grim JE, Grim JA, Harper JW, et al. The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. Proc Natl Acad Sci U S A. 2004;101: 9085–9090. doi:10.1073/pnas.0402770101

160. International Cancer Genome Consortium, Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, et al. International network of cancer genome projects. Nature. 2010;464: 993–998. doi:10.1038/nature08987

161. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell. 2006;126: 663–676. doi:10.1016/j.cell.2006.07.024

162. Thomson JA, Itskovitz-Eldor J, Shapiro SS, Waknitz MA, Swiergiel JJ, Marshall VS, et al. Embryonic stem cell lines derived from human blastocysts. Science. 1998;282: 1145–1147.

163. Neiman P, Wolf C, Enrietto PJ, Cooper GM. A retroviral myc gene induces preneoplastic transformation of lymphocytes in a bursal transplantation assay. Proc Natl Acad Sci U S A. 1985;82: 222–226.

164. Leucci E, Cocco M, Onnis A, De Falco G, van Cleef P, Bellan C, et al. MYC translocation-negative classical Burkitt lymphoma cases: an alternative pathogenetic mechanism involving miRNA deregulation. J Pathol. 2008;216: 440–450. doi:10.1002/path.2410

165. Haralambieva E, Schuuring E, Rosati S, van Noesel C, Jansen P, Appel I, et al. Interphase fluorescence in situ hybridization for detection of 8q24/MYC breakpoints on routine histologic sections: validation in Burkitt lymphomas from three geographic regions. Genes Chromosomes Cancer. 2004;40: 10–18. doi:10.1002/gcc.20009

166. Salaverria I, Martin-Guerrero I, Wagener R, Kreuz M, Kohler CW, Richter J, et al. A recurrent 11q aberration pattern characterizes a subset of MYC-negative high-grade B-cell lymphomas resembling Burkitt lymphoma. Blood. 2014;123: 1187–1198. doi:10.1182/blood-2013-06-507996

167. Bentink S, Wessendorf S, Schwaenen C, Rosolowski M, Klapper W, Rosenwald A, et al. Pathway activation patterns in diffuse large B-cell lymphomas. Leukemia. 2008;22: 1746–1754. doi:10.1038/leu.2008.166

168. Pienkowska-Grela B, Rymkiewicz G, Grygalewicz B, Woroniecka R, Krawczyk P, Czyz-Domanska K, et al. Partial trisomy 11, dup(11)(q23q13), as a defect characterizing lymphomas with Burkitt pathomorphology without MYC gene rearrangement. Med Oncol Northwood Lond Engl. 2011;28: 1589–1595. doi:10.1007/s12032-010-9614-0

169. Poirel HA, Cairo MS, Heerema NA, Swansbury J, Aupérin A, Launay E, et al. Specific cytogenetic abnormalities are associated with a significantly inferior outcome in children and adolescents with mature B-cell non-Hodgkin's lymphoma: results of the FAB/LMB 96 international study. Leukemia. 2009;23: 323–331. doi:10.1038/leu.2008.312

170. Richter J, Schlesner M, Hoffmann S, Kreuz M, Leich E, Burkhardt B, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat Genet. 2012;44: 1316–1320. doi:10.1038/ng.2469

171. Matsuo Y, MacLeod RA, Kojima K, Kuwahara K, Sakata A, Drexler HG, et al. A novel ALL-L3 cell line, BALM-16, lacking expression of immunoglobulin chains derived from a patient with hypercalcemia. Leukemia. 1997;11: 2168–2174.

172. Matsuo Y, Sugimoto A, Harashima A, Nishizaki C, Ishimaru F, Kondo E, et al. Establishment and characterization of a novel ALL-L3 cell line (BALM-18): induction of apoptosis by anti-IgM and inhibition of apoptosis by bone marrow stroma cells. Leuk Res. 1999;23: 559–568.

173. Bertrand S, Berger R, Philip T, Bernheim A, Bryon PA, Bertoglio J, et al. Variant translocation in a non endemic case of Burkitt's lymphoma: t (8;22) in an Epstein--Barr virus negative tumour and in a derived cell line. Eur J Cancer. 1981;17: 577–584.

174. Philip I, Philip T, Favrot M, Vuillaume M, Fontaniere B, Chamard D, et al. Establishment of lymphomatous cell lines from bone marrow samples from patients with Burkitt's lymphoma. J Natl Cancer Inst. 1984;73: 835–840.

175. Lenoir GM, Vuillaume M, Bonnardel C. The use of lymphomatous and lymphoblastoid cell lines in the study of Burkitt's lymphoma. IARC Sci Publ. 1985; 309–318.

176. Burmeister T, Macleod RAF, Reinhardt R, Mansmann V, Loddenkemper C, Marinets O, et al. A novel sporadic Burkitt lymphoma cell line (BLUE-1) with a unique t(6;20)(q15;q11.2) rearrangement. Leuk Res. 2006;30: 1417–1423. doi:10.1016/j.leukres.2006.03.026

177. Magrath IT, Pizzo PA, Whang-Peng J, Douglass EC, Alabaster O, Gerber P, et al. Characterization of lymphoma-derived cell lines: comparison of cell lines positive and negative for Epstein-Barr virus nuclear antigen. I. Physical, cytogenetic, and growth characteristics. J Natl Cancer Inst. 1980;64: 465–476.

178. Nadkarni JS, Nadkarni JJ, Clifford P, Manolov G, Fenyö EM, Klein E. Characteristics of new cell lines derived from Burkitt lymphomas. Cancer. 1969;23: 64–79.

179. Ben-Bassat H, Goldblum N, Mitrani S, Goldblum T, Yoffey JM, Cohen MM, et al. Establishment in continuous culture of a new type of lymphocyte from a “Burkitt like” malignant lymphoma (line D.G.-75). Int J Cancer J Int Cancer. 1977;19: 27–33.

180. Epstein MA, Barr YM. CULTIVATION IN VITRO OF HUMAN LYMPHOBLASTS FROM BURKITT’S MALIGNANT LYMPHOMA. Lancet. 1964;1: 252–253.

181. Berger R, Bernheim A. Cytogenetics of Burkitt’s lymphoma-leukaemia: a review. IARC Sci Publ. 1985; 65–80.

182. Pulvertaft JV. CYTOLOGY OF BURKITT’S TUMOUR (AFRICAN LYMPHOMA). Lancet. 1964;1: 238–240.

183. Nilsson K, Sundström C. Establishment and characteristics of two unique cell lines from patients with lymphosarcoma. Int J Cancer J Int Cancer. 1974;13: 808–823.

184. Maria Murga Penas E, Schilling G, Behrmann P, Klokow M, Vettorazzi E, Bokemeyer C, et al. Comprehensive cytogenetic and molecular cytogenetic analysis of 44 Burkitt lymphoma cell lines: secondary chromosomal changes characterization, karyotypic evolution, and comparison with primary samples. Genes Chromosomes Cancer. 2014;53: 497–515. doi:10.1002/gcc.22161

185. Epstein AL, Levy R, Kim H, Henle W, Henle G, Kaplan HS. Biology of the human malignant lymphomas. IV. Functional characterization of ten diffuse histiocytic lymphoma cell lines. Cancer. 1978;42: 2379–2391.

186. Dyer MJ, Fischer P, Nacheva E, Labastide W, Karpas A. A new human B-cell non-Hodgkin’s lymphoma cell line (Karpas 422) exhibiting both t (14;18) and t(4;11) chromosomal translocations. Blood. 1990;75: 709–714.

187. Th’ng KH, Garewal G, Kearney L, Rassool F, Melo JV, White H, et al. Establishment and characterization of three new malignant lymphoid cell lines. Int J Cancer J Int Cancer. 1987;39: 89–93.

188. Beckwith M, Longo DL, O’Connell CD, Moratz CM, Urba WJ. Phorbol ester-induced, cell-cycle-specific, growth inhibition of human B-lymphoma cell lines. J Natl Cancer Inst. 1990;82: 501–509.

189. Schaadt M, Diehl V, Stein H, Fonatsch C, Kirchner HH. Two neoplastic cell lines with unique features derived from Hodgkin’s disease. Int J Cancer J Int Cancer. 1980;26: 723–731.

190. Tweeddale ME, Lim B, Jamal N, Robinson J, Zalcberg J, Lockwood G, et al. The presence of clonogenic cells in high-grade malignant lymphoma: a prognostic factor. Blood. 1987;69: 1307–1314.

191. Graham FL, Smiley J, Russell WC, Nairn R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol. 1977;36: 59–74.

192. Schuetz JM, Johnson NA, Morin RD, Scott DW, Tan K, Ben-Nierah S, et al. BCL2 mutations in diffuse large B-cell lymphoma. Leukemia. 2012;26: 1383–1390. doi:10.1038/leu.2011.378

193. Schneppenheim R, Frühwald MC, Gesk S, Hasselblatt M, Jeibmann A, Kordes U, et al. Germline nonsense mutation and somatic inactivation of SMARCA4/BRG1 in a family with rhabdoid tumor predisposition syndrome. Am J Hum Genet. 2010;86: 279–284. doi:10.1016/j.ajhg.2010.01.013

194. Kent WJ. BLAT--the BLAST-like alignment tool. Genome Res. 2002;12: 656–664. doi:10.1101/gr.229202. Article published online before March 2002

195. Den Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat. 2000;15: 7–12. doi:10.1002/(SICI)1098-1004(200001)15:1<7::AID-HUMU4>3.0.CO;2-N

196. Huang X, Miller W. A time-efficient, linear-space local similarity algorithm. Adv Appl Math. 1991;12: 337–357. doi:10.1016/0196-8858(91)90017-D

197. Betts MJ, Lu Q, Jiang Y, Drusko A, Wichmann O, Utz M, et al. Mechismo: predicting the mechanistic impact of mutations and modifications on molecular interactions. Nucleic Acids Res. 2015;43: e10. doi:10.1093/nar/gku1094

198. Dirks WG, MacLeod RAF, Nakamura Y, Kohara A, Reid Y, Milch H, et al. Cell line cross-contamination initiative: an interactive reference database of STR profiles covering common cancer cell lines. Int J Cancer J Int Cancer. 2010;126: 303–304. doi:10.1002/ijc.24999

199. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014;42: D222–230. doi:10.1093/nar/gkt1223

200. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7: 248–249. doi:10.1038/nmeth0410-248

201. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. Proc Natl Acad Sci U S A. 1998;95: 5857–5864.

202. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002;12: 996–1006. doi:10.1101/gr.229102. Article published online before print in May 2002

203. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. 1977. Biotechnol Read Mass. 1992;24: 104–108.

204. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14: 178–192. doi:10.1093/bib/bbs017

205. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, et al. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem. 2009;55: 611–622. doi:10.1373/clinchem.2008.112797

206. Deng S, Hiruki C. Amplification of 16S rRNA genes from culturable and nonculturable Mollicutes. J Microbiol Methods. 1991;14: 53–61. doi:10.1016/0167-7012(91)90007-D

207. Drexler HG, Dirks WG, Matsuo Y, MacLeod R a. F. False leukemia-lymphoma cell lines: an update on over 500 cell lines. Leukemia. 2003;17: 416–426. doi:10.1038/sj.leu.2402799

208. Capes-Davis A, Theodosopoulos G, Atkin I, Drexler HG, Kohara A, MacLeod RAF, et al. Check your cultures! A list of cross-contaminated or misidentified cell lines. Int J Cancer J Int Cancer. 2010;127: 1–8. doi:10.1002/ijc.25242

209. Zufferey R, Nagy D, Mandel RJ, Naldini L, Trono D. Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. Nat Biotechnol. 1997;15: 871–875. doi:10.1038/nbt0997-871

210. Naldini L, Blömer U, Gallay P, Ory D, Mulligan R, Gage FH, et al. In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science. 1996;272: 263–267.

211. Levings MK, Sangregorio R, Sartirana C, Moschin AL, Battaglia M, Orban PC, et al. Human CD25+CD4+ T suppressor cell clones produce transforming growth factor beta, but not interleukin 10, and are distinct from type 1 T regulatory cells. J Exp Med. 2002;196: 1335–1346.

212. Srisodsai A, Kurotani R, Chiba Y, Sheikh F, Young HA, Donnelly RP, et al. Interleukin-10 induces uteroglobin-related protein (UGRP) 1 gene expression in lung epithelial cells through homeodomain transcription factor T/EBP/NKX2.1. J Biol Chem. 2004;279: 54358–54368. doi:10.1074/jbc.M405331200

213. Lüschen S, Ussat S, Scherer G, Kabelitz D, Adam-Klages S. Sensitization to death receptor cytotoxicity by inhibition of fas-associated death domain protein (FADD)/caspase signaling. Requirement of cell cycle progression. J Biol Chem. 2000;275: 24670–24678. doi:10.1074/jbc.M003280200

214. Nunez R. DNA measurement and cell cycle analysis by flow cytometry. Curr Issues Mol Biol. 2001;3: 67–70.

215. Müller F-J, Schuldt BM, Williams R, Mason D, Altun G, Papapetrou EP, et al. A bioinformatic assay for pluripotency in human cells. Nat Methods. 2011;8: 315–317. doi:10.1038/nmeth.1580

216. Korkola JE, Houldsworth J, Feldman DR, Olshen AB, Qin L-X, Patil S, et al. Identification and validation of a gene expression signature that predicts outcome in adult men with germ cell tumors. J Clin Oncol Off J Am Soc Clin Oncol. 2009;27: 5240–5247. doi:10.1200/JCO.2008.20.0386

217. Kushwaha R, Jagadish N, Kustagi M, Tomishima MJ, Mendiratta G, Bansal M, et al. Interrogation of a context-specific transcription factor network identifies novel regulators of pluripotency. Stem Cells Dayt Ohio. 2014; doi:10.1002/stem.1870

218. Palmer RD, Murray MJ, Saini HK, van Dongen S, Abreu-Goodger C, Muralidhar B, et al. Malignant germ cell tumors display common microRNA profiles resulting in global changes in expression of messenger RNA targets. Cancer Res. 2010;70: 2911–2923. doi:10.1158/0008-5472.CAN-09-3301

219. Lukk M, Kapushesky M, Nikkilä J, Parkinson H, Goncalves A, Huber W, et al. A global map of human gene expression. Nat Biotechnol. 2010;28: 322–324. doi:10.1038/nbt0410-322

220. Yu Y, Wang J, Khaled W, Burke S, Li P, Chen X, et al. Bcl11a is essential for lymphoid development and negatively regulates p53. J Exp Med. 2012;209: 2467–2483. doi:10.1084/jem.20121846

221. Hikida M, Casola S, Takahashi N, Kaji T, Takemori T, Rajewsky K, et al. PLC-gamma2 is essential for formation and maintenance of memory B cells. J Exp Med. 2009;206: 681–689. doi:10.1084/jem.20082100

222. Na J, Plews J, Li J, Wongtrakoongate P, Tuuri T, Feki A, et al. Molecular mechanisms of pluripotency and reprogramming. Stem Cell Res Ther. 2010;1: 33. doi:10.1186/scrt33

223. Hernández S, Hernández L, Beà S, Cazorla M, Fernández PL, Nadal A, et al. cdc25 cell cycle-activating phosphatases and c-myc expression in human non-Hodgkin's lymphomas. Cancer Res. 1998;58: 1762–1767.

224. Den Hollander J, Rimpi S, Doherty JR, Rudelius M, Buck A, Hoellein A, et al. Aurora kinases A and B are up-regulated by Myc and are essential for maintenance of the malignant state. Blood. 2010;116: 1498–1505. doi:10.1182/blood-2009-11-251074

225. Xia J, Peng Y, Mian IS, Lue NF. Identification of functionally important domains in the N-terminal region of telomerase reverse transcriptase. Mol Cell Biol. 2000;20: 5196–5207.

226. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet. 2011;43: 830–837. doi:10.1038/ng.892

227. Monti S, Chapuy B, Takeyama K, Rodig SJ, Hao Y, Yeda KT, et al. Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B cell lymphoma. Cancer Cell. 2012;22: 359–372. doi:10.1016/j.ccr.2012.07.014

228. Love C, Sun Z, Jima D, Li G, Zhang J, Miles R, et al. The genetic landscape of mutations in Burkitt lymphoma. Nat Genet. 2012;44: 1321–1325. doi:10.1038/ng.2468

229. Oricchio E, Nanjangud G, Wolfe AL, Schatz JH, Mavrakis KJ, Jiang M, et al. The Eph-receptor A7 is a soluble tumor suppressor for follicular lymphoma. Cell. 2011;147: 554–564. doi:10.1016/j.cell.2011.09.035

230. Mandelbaum J, Bhagat G, Tang H, Mo T, Brahmachary M, Shen Q, et al. BLIMP1 is a tumor suppressor gene frequently disrupted in activated B cell-like diffuse large B cell lymphoma. Cancer Cell. 2010;18: 568–579. doi:10.1016/j.ccr.2010.10.030

231. Schiffman JD, Lorimer PD, Rodic V, Jahromi MS, Downie JM, Bayerl MG, et al. Genome wide copy number analysis of paediatric Burkitt lymphoma using formalin-fixed tissues reveals a subset with gain of chromosome 13q and corresponding miRNA over expression. Br J Haematol. 2011;155: 477–486. doi:10.1111/j.1365-2141.2011.08883.x

232. Xu S, Tong M, Huang J, Zhang Y, Qiao Y, Weng W, et al. TRIB2 inhibits Wnt/β-Catenin/TCF4 signaling through its associated ubiquitin E3 ligases, β-TrCP, COP1 and Smurf1, in liver cancer cells. FEBS Lett. 2014;588: 4334–4341. doi:10.1016/j.febslet.2014.09.042

233. Cha N, Liu W, Yang N, Xie S, Gao Y, Chen X, et al. Oncogenicity of LHX4 in colorectal cancer through Wnt/β-catenin/TCF4 cascade. Tumour Biol J Int Soc Oncodevelopmental Biol Med. 2014;35: 10319–10324. doi:10.1007/s13277-014-2210-8

234. Masir N, Campbell LJ, Jones M, Mason DY. Pseudonegative BCL2 protein expression in a t(14;18) translocation positive lymphoma cell line: a need for an alternative BCL2 antibody. Pathology (Phila). 2010;42: 212–216. doi:10.3109/00313021003631296

235. Jain P, Pemmaraju N, Ravandi F. Update on the biology and treatment options for hairy cell leukemia. Curr Treat Options Oncol. 2014;15: 187–209. doi:10.1007/s11864-014-0285-5

236. Lecointe N, Meerabux J, Ebihara M, Hill A, Young BD. Molecular analysis of an unstable genomic region at chromosome band 11q23 reveals a disruption of the gene encoding the alpha2 subunit of platelet-activating factor acetylhydrolase (Pafah1a2) in human lymphoma. Oncogene. 1999;18: 2852–2859. doi:10.1038/sj.onc.1202645

237. Yu D, Cozma D, Park A, Thomas-Tikhonenko A. Functional validation of genes implicated in lymphomagenesis: an in vivo selection assay using a Myc-induced B-cell tumor. Ann N Y Acad Sci. 2005;1059: 145–159. doi:10.1196/annals.1339.047

238. Martínez N, Camacho FI, Algara P, Rodríguez A, Dopazo A, Ruíz-Ballesteros E, et al. The molecular signature of mantle cell lymphoma reveals multiple signals favoring cell survival. Cancer Res. 2003;63: 8226–8232.

239. Bonetti P, Testoni M, Scandurra M, Ponzoni M, Piva R, Mensah AA, et al. Deregulation of ETS1 and FLI1 contributes to the pathogenesis of diffuse large B-cell lymphoma. Blood. 2013;122: 2233–2241. doi:10.1182/blood-2013-01-475772

240. Chen CS, Sorensen PH, Domer PH, Reaman GH, Korsmeyer SJ, Heerema NA, et al. Molecular rearrangements on chromosome 11q23 predominate in infant acute lymphoblastic leukemia and are associated with specific biologic variables and poor outcome. Blood. 1993;81: 2386–2393.

241. Cimino G, Rapanotti MC, Rivolta A, Lo Coco F, D'Arcangelo E, Rondelli R, et al. Prognostic relevance of ALL-1 gene rearrangement in infant acute leukemias. Leukemia. 1995;9: 391–395.

242. Gindin T, Murty V, Alobeid B, Bhagat G. MLL/KMT2A translocations in diffuse large B-cell lymphomas. Hematol Oncol. 2014; doi:10.1002/hon.2158

243. Ahlmann M, Meyer C, Marschalek R, Burkhardt B, Koehler G, Klapper W, et al. Complex MLL rearrangement in non-infiltrated bone marrow in an infant with stage II precursor B-lymphoblastic lymphoma. Eur J Haematol. 2014;93: 349–353. doi:10.1111/ejh.12314

244. Yamamoto K, Yakushijin K, Okamura A, Hayashi Y, Matsuoka H, Minami H. Gain of 11q by an additional ring chromosome 11 and trisomy 18 in CD5-positive intravascular large B-cell lymphoma. J Clin Exp Hematop JCEH. 2013;53: 161–165.

245. Salaverria I, Martin-Guerrero I, Burkhardt B, Kreuz M, Zenz T, Oschlies I, et al. High resolution copy number analysis of IRF4 translocation-positive diffuse large B-cell and follicular lymphomas. Genes Chromosomes Cancer. 2013;52: 150–155. doi:10.1002/gcc.22014

246. John SA, Clements JL, Russell LM, Garrett-Sinha LA. Ets-1 regulates plasma cell differentiation by interfering with the activity of the transcription factor Blimp-1. J Biol Chem. 2008;283: 951–962. doi:10.1074/jbc.M705262200

247. Lin K-I, Angelin-Duclos C, Kuo TC, Calame K. Blimp-1-dependent repression of Pax-5 is required for differentiation of B cells to immunoglobulin M-secreting plasma cells. Mol Cell Biol. 2002;22: 4771–4780.

248. Overbeck BM, Martin-Subero JI, Ammerpohl O, Klapper W, Siebert R, Giefing M. ETS1 encoding a transcription factor involved in B-cell differentiation is recurrently deleted and down-

regulated in classical Hodgkin's lymphoma. Haematologica. 2012;97: 1612–1614. doi:10.3324/haematol.2012.061770

249. Sarrazin S, Starck J, Gonnet C, Doubeikovski A, Melet F, Morle F. Negative and translation termination-dependent positive control of FLI-1 protein synthesis by conserved overlapping 5' upstream open reading frames in Fli-1 mRNA. Mol Cell Biol. 2000;20: 2959–2969.

250. Bonin F, Ryan SD, Migahed L, Mo F, Lallier J, Franks DJ, et al. Anti-apoptotic actions of the platelet-activating factor acetylhydrolase I alpha2 catalytic subunit. J Biol Chem. 2004;279: 52425–52436. doi:10.1074/jbc.M410967200

251. Crawley JB, Williams LM, Mander T, Brennan FM, Foxwell BM. Interleukin-10 stimulation of phosphatidylinositol 3-kinase and p70 S6 kinase is required for the proliferative but not the antiinflammatory effects of the cytokine. J Biol Chem. 1996;271: 16357–16362.

252. Ando Y, Inada-Inoue M, Mitsuma A, Yoshino T, Ohtsu A, Suenaga N, et al. Phase I dose-escalation study of buparlisib (BKM120), an oral pan-class I PI3K inhibitor, in Japanese patients with advanced solid tumors. Cancer Sci. 2014;105: 347–353. doi:10.1111/cas.12350

253. Bendell JC, Rodon J, Burris HA, de Jonge M, Verweij J, Birle D, et al. Phase I, dose-escalation study of BKM120, an oral pan-Class I PI3K inhibitor, in patients with advanced solid tumors. J Clin Oncol Off J Am Soc Clin Oncol. 2012;30: 282–290. doi:10.1200/JCO.2011.36.1360

254. Bustamante-Marín X, Garness JA, Capel B. Testicular teratomas: an intersection of pluripotency, differentiation and cancer biology. Int J Dev Biol. 2013;57: 201–210. doi:10.1387/ijdb.130136bc

255. Aukema SM, Siebert R, Schuuring E, van Imhoff GW, Kluin-Nelemans HC, Boerma E-J, et al. Double-hit B-cell lymphomas. Blood. 2011;117: 2319–2331. doi:10.1182/blood-2010-09-297879

256. Barrans S, Crouch S, Smith A, Turner K, Owen R, Patmore R, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. J Clin Oncol Off J Am Soc Clin Oncol. 2010;28: 3360–3365. doi:10.1200/JCO.2009.26.3947

257. Kondo M, Wagers AJ, Manz MG, Prohaska SS, Scherer DC, Beilhack GF, et al. Biology of hematopoietic stem cells and progenitors: implications for clinical application. Annu Rev Immunol. 2003;21: 759–806. doi:10.1146/annurev.immunol.21.120601.141007

258. Ben-Porath I, Thomson MW, Carey VJ, Ge R, Bell GW, Regev A, et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat Genet. 2008;40: 499–507. doi:10.1038/ng.127

259. Kim J, Woo AJ, Chu J, Snow JW, Fujiwara Y, Kim CG, et al. A Myc network accounts for similarities between embryonic stem and cancer cell transcription programs. Cell. 2010;143: 313–324. doi:10.1016/j.cell.2010.09.010

260. Li Y-Q. Master stem cell transcription factors and signaling regulation. Cell Reprogramming. 2010;12: 3–13. doi:10.1089/cell.2009.0033

261. Medvedev SP, Shevchenko AI, Mazurok NA, Zakiian SM. [OCT4 and NANOG are the key genes in the system of pluripotency maintenance in mammalian cells]. Genetika. 2008;44: 1589–1608.

262. Küppers R, Dalla-Favera R. Mechanisms of chromosomal translocations in B cell lymphomas. Oncogene. 2001;20: 5580–5594. doi:10.1038/sj.onc.1204640

263. Kirkpatrick KL, Newbold RF, Mokbel K. The mRNA expression of hTERT in human breast carcinomas correlates with VEGF expression. J Carcinog. 2004;3: 1. doi:10.1186/1477-3163-3-1

264. Smith LL, Coller HA, Roberts JM. Telomerase modulates expression of growth-controlling genes and enhances cell proliferation. Nat Cell Biol. 2003;5: 474–479. doi:10.1038/ncb985

265. Choi J, Southworth LK, Sarin KY, Venteicher AS, Ma W, Chang W, et al. TERT promotes epithelial proliferation through transcriptional control of a Myc- and Wnt-related developmental program. PLoS Genet. 2008;4: e10. doi:10.1371/journal.pgen.0040010

266. Gonzalez OG, Assfalg R, Koch S, Schelling A, Meena JK, Kraus J, et al. Telomerase stimulates ribosomal DNA transcription under hyperproliferative conditions. Nat Commun. 2014;5: 4599. doi:10.1038/ncomms5599

267. Xiang H, Wang J, Mao Y, Liu M, Reddy VN, Li DW-C. Human telomerase accelerates growth of lens epithelial cells through regulation of the genes mediating RB/E2F pathway. Oncogene. 2002;21: 3784–3791. doi:10.1038/sj.onc.1205455

268. Lee J, Sung YH, Cheong C, Choi YS, Jeon HK, Sun W, et al. TERT promotes cellular and organismal survival independently of telomerase activity. Oncogene. 2008;27: 3754–3760. doi:10.1038/sj.onc.1211037

269. Massard C, Zermati Y, Pauleau A-L, Larochette N, Métivier D, Sabatier L, et al. hTERT: a novel endogenous inhibitor of the mitochondrial cell death pathway. Oncogene. 2006;25: 4505–4514. doi:10.1038/sj.onc.1209487

270. Rahman R, Latonen L, Wiman KG. hTERT antagonizes p53-induced apoptosis independently of telomerase activity. Oncogene. 2005;24: 1320–1327. doi:10.1038/sj.onc.1208232

271. Krupp G, Klapper W, Parwaresch R. Cell proliferation, carcinogenesis and diverse mechanisms of telomerase regulation. Cell Mol Life Sci CMLS. 2000;57: 464–486.

272. Kim NW, Piatyszek MA, Prowse KR, Harley CB, West MD, Ho PL, et al. Specific association of human telomerase activity with immortal cells and cancer. Science. 1994;266: 2011–2015.

273. Shay JW, Bacchetti S. A survey of telomerase activity in human cancer. Eur J Cancer Oxf Engl 1990. 1997;33: 787–791. doi:10.1016/S0959-8049(97)00062-2

274. Klingelhutz AJ. Telomerase activation and cancer. J Mol Med Berl Ger. 1997;75: 45–49.

275. Horikawa I, Cable PL, Afshari C, Barrett JC. Cloning and characterization of the promoter region of human telomerase reverse transcriptase gene. Cancer Res. 1999;59: 826–830.

276. Takakura M, Kyo S, Kanaya T, Hirano H, Takeda J, Yutsudo M, et al. Cloning of human telomerase catalytic subunit (hTERT) gene promoter and identification of proximal core promoter sequences essential for transcriptional activation in immortalized and cancer cells. Cancer Res. 1999;59: 551–557.

277. Wu KJ, Grandori C, Amacker M, Simon-Vermot N, Polack A, Lingner J, et al. Direct activation of TERT transcription by c-MYC. Nat Genet. 1999;21: 220–224. doi:10.1038/6010

278. Listerman I, Sun J, Gazzaniga FS, Lukas JL, Blackburn EH. The major reverse transcriptase-incompetent splice variant of the human telomerase protein inhibits telomerase activity but

protects from apoptosis. Cancer Res. 2013;73: 2817–2828. doi:10.1158/0008-5472.CAN-12-3082

279. Hrdlicková R, Nehyba J, Bose HR. Alternatively spliced telomerase reverse transcriptase variants lacking telomerase activity stimulate cell proliferation. Mol Cell Biol. 2012;32: 4283–4296. doi:10.1128/MCB.00550-12

280. Kilian A, Bowtell DD, Abud HE, Hime GR, Venter DJ, Keese PK, et al. Isolation of a candidate human telomerase catalytic subunit gene, which reveals complex splicing patterns in different cell types. Hum Mol Genet. 1997;6: 2011–2019.

281. Hisatomi H, Ohyashiki K, Ohyashiki JH, Nagao K, Kanamaru T, Hirata H, et al. Expression profile of a gamma-deletion variant of the human telomerase reverse transcriptase gene. Neoplasia N Y N. 2003;5: 193–197. doi:NO_DOI

282. Withers JB, Ashvetiya T, Beemon KL. Exclusion of exon 2 is a common mRNA splice variant of primate telomerase reverse transcriptases. PloS One. 2012;7: e48016. doi:10.1371/journal.pone.0048016

283. Saebøe-Larssen S, Fossberg E, Gaudernack G. Characterization of novel alternative splicing sites in human telomerase reverse transcriptase (hTERT): analysis of expression and mutual correlation in mRNA isoforms from normal and tumour tissues. BMC Mol Biol. 2006;7: 26. doi:10.1186/1471-2199-7-26

284. Nicholson P, Yepiskoposyan H, Metze S, Zamudio Orozco R, Kleinschmidt N, Mühlemann O. Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. Cell Mol Life Sci CMLS. 2010;67: 677–700. doi:10.1007/s00018-009-0177-1

285. Cong Y-S, Wright WE, Shay JW. Human telomerase and its regulation. Microbiol Mol Biol Rev MMBR. 2002;66: 407–425, table of contents.

286. Wong MS, Chen L, Foster C, Kainthla R, Shay JW, Wright WE. Regulation of telomerase alternative splicing: a target for chemotherapy. Cell Rep. 2013;3: 1028–1035. doi:10.1016/j.celrep.2013.03.011

287. Amor S, Remy S, Dambrine G, Le Vern Y, Rasschaert D, Laurent S. Alternative splicing and nonsense-mediated decay regulate telomerase reverse transcriptase (TERT) expression during virus-induced lymphomagenesis in vivo. BMC Cancer. 2010;10: 571. doi:10.1186/1471-2407-10-571

288. Andreutti-Zaugg C, Scott RJ, Iggo R. Inhibition of nonsense-mediated messenger RNA decay in clinical samples facilitates detection of human MSH2 mutations with an in vivo fusion protein assay and conventional techniques. Cancer Res. 1997;57: 3288–3293.

289. Wang D, Wengrod J, Gardner LB. Overexpression of the c-myc oncogene inhibits nonsense-mediated RNA decay in B lymphocytes. J Biol Chem. 2011;286: 40038–40043. doi:10.1074/jbc.M111.266361

290. Bachand F, Autexier C. Functional regions of human telomerase reverse transcriptase and human telomerase RNA required for telomerase activity and RNA-protein interactions. Mol Cell Biol. 2001;21: 1888–1897. doi:10.1128/MCB.21.5.1888-1897.2001

291. Beattie TL, Zhou W, Robinson MO, Harrington L. Functional multimerization of the human telomerase reverse transcriptase. Mol Cell Biol. 2001;21: 6151–6160.

292. Autexier C, Lue NF. The structure and function of telomerase reverse transcriptase. Annu Rev Biochem. 2006;75: 493–517. doi:10.1146/annurev.biochem.75.103004.142412

293. Weinrich SL, Pruzan R, Ma L, Ouellette M, Tesmer VM, Holt SE, et al. Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. Nat Genet. 1997;17: 498–502. doi:10.1038/ng1297-498

294. Mukherjee S, Firpo EJ, Wang Y, Roberts JM. Separation of telomerase functions by reverse genetics. Proc Natl Acad Sci U S A. 2011;108: E1363–1371. doi:10.1073/pnas.1112414108

295. Moriarty TJ, Huard S, Dupuis S, Autexier C. Functional multimerization of human telomerase requires an RNA interaction domain in the N terminus of the catalytic subunit. Mol Cell Biol. 2002;22: 1253–1265.

296. Jacobs SA, Podell ER, Cech TR. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. Nat Struct Mol Biol. 2006;13: 218–225. doi:10.1038/nsmb1054

297. Lue NF. A physical and functional constituent of telomerase anchor site. J Biol Chem. 2005;280: 26586–26591. doi:10.1074/jbc.M503028200

298. Colgin LM, Wilkinson C, Englezou A, Kilian A, Robinson MO, Reddel RR. The hTERTalpha splice variant is a dominant negative inhibitor of telomerase activity. Neoplasia N Y N. 2000;2: 426–432.

299. Yi X, White DM, Aisner DL, Baur JA, Wright WE, Shay JW. An alternate splicing variant of the human telomerase catalytic subunit inhibits telomerase activity. Neoplasia N Y N. 2000;2: 433–440.

300. Bosoy D, Peng Y, Mian IS, Lue NF. Conserved N-terminal motifs of telomerase reverse transcriptase required for ribonucleoprotein assembly in vivo. J Biol Chem. 2003;278: 3882–3890. doi:10.1074/jbc.M210645200

301. Lai CK, Mitchell JR, Collins K. RNA binding domain of telomerase reverse transcriptase. Mol Cell Biol. 2001;21: 990–1000. doi:10.1128/MCB.21.4.990-1000.2001

302. Anker L, Ohgaki H, Ludeke BI, Herrmann HD, Kleihues P, Westphal M. p53 protein accumulation and gene mutations in human glioma cell lines. Int J Cancer J Int Cancer. 1993;55: 982–987.

303. Tweddle DA, Malcolm AJ, Bown N, Pearson AD, Lunec J. Evidence for the development of p53 mutations after cytotoxic therapy in a neuroblastoma cell line. Cancer Res. 2001;61: 8–13.

304. Christiansen DH, Andersen MK, Pedersen-Bjergaard J. Mutations with loss of heterozygosity of p53 are common in therapy-related myelodysplasia and acute myeloid leukemia after exposure to alkylating agents and significantly associated with deletion or loss of 5q, a complex karyotype, and a poor prognosis. J Clin Oncol Off J Am Soc Clin Oncol. 2001;19: 1405–1413.

305. Thome KC, Radfar A, Rosenberg N. Mutation of Tp53 contributes to the malignant phenotype of Abelson virus-transformed lymphoid cells. J Virol. 1997;71: 8149–8156.

306. Freed-Pastor WA, Prives C. Mutant p53: one name, many proteins. Genes Dev. 2012;26: 1268–1286. doi:10.1101/gad.190678.112

307. Cho Y, Gorina S, Jeffrey PD, Pavletich NP. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. Science. 1994;265: 346–355.

308. Buschmann T, Minamoto T, Wagle N, Fuchs SY, Adler V, Mai M, et al. Analysis of JNK, Mdm2 and p14(ARF) contribution to the regulation of mutant p53 stability. J Mol Biol. 2000;295: 1009–1021. doi:10.1006/jmbi.1999.3387

309. Peng Y, Chen L, Li C, Lu W, Agrawal S, Chen J. Stabilization of the MDM2 oncoprotein by mutant p53. J Biol Chem. 2001;276: 6874–6878. doi:10.1074/jbc.C000781200

310. Frum RA, Grossman SR. Mechanisms of mutant p53 stabilization in cancer. Subcell Biochem. 2014;85: 187–197. doi:10.1007/978-94-017-9211-0_10

311. Bunz F, Dutriaux A, Lengauer C, Waldman T, Zhou S, Brown JP, et al. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. Science. 1998;282: 1497–1501.

312. Lowe SW, Schmitt EM, Smith SW, Osborne BA, Jacks T. p53 is required for radiation-induced apoptosis in mouse thymocytes. Nature. 1993;362: 847–849. doi:10.1038/362847a0

313. Clarke AR, Purdie CA, Harrison DJ, Morris RG, Bird CC, Hooper ML, et al. Thymocyte apoptosis induced by p53-dependent and independent pathways. Nature. 1993;362: 849–852. doi:10.1038/362849a0

314. Polyak K, Xia Y, Zweier JL, Kinzler KW, Vogelstein B. A model for p53-induced apoptosis. Nature. 1997;389: 300–305. doi:10.1038/38525

315. Willis A, Jung EJ, Wakefield T, Chen X. Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. Oncogene. 2004;23: 2330–2338. doi:10.1038/sj.onc.1207396

316. Maya-Mendoza A, Ostrakova J, Kosar M, Hall A, Duskova P, Mistrik M, et al. Myc and Ras oncogenes engage different energy metabolism programs and evoke distinct patterns of oxidative and DNA replication stress. Mol Oncol. 2015;9: 601–616. doi:10.1016/j.molonc.2014.11.001

317. Evan GI, Wyllie AH, Gilbert CS, Littlewood TD, Land H, Brooks M, et al. Induction of apoptosis in fibroblasts by c-myc protein. Cell. 1992;69: 119–128.

318. Sakamuro D, Eviner V, Elliott KJ, Showe L, White E, Prendergast GC. c-Myc induces apoptosis in epithelial cells by both p53-dependent and p53-independent mechanisms. Oncogene. 1995;11: 2411–2418.

319. Diehl JA, Zindy F, Sherr CJ. Inhibition of cyclin D1 phosphorylation on threonine-286 prevents its rapid degradation via the ubiquitin-proteasome pathway. Genes Dev. 1997;11: 957–972.

320. Casanovas O, Jaumot M, Paules A-B, Agell N, Bachs O. P38SAPK2 phosphorylates cyclin D3 at Thr-283 and targets it for proteasomal degradation. Oncogene. 2004;23: 7537–7544. doi:10.1038/sj.onc.1208040

321. Sherr CJ. Mammalian G1 cyclins. Cell. 1993;73: 1059–1065.

322. Cato MH, Chintalapati SK, Yau IW, Omori SA, Rickert RC. Cyclin D3 is selectively required for proliferative expansion of germinal center B cells. Mol Cell Biol. 2011;31: 127–137. doi:10.1128/MCB.00650-10

323. Peled JU, Yu JJ, Venkatesh J, Bi E, Ding BB, Krupski-Downs M, et al. Requirement for cyclin D3 in germinal center formation and function. Cell Res. 2010;20: 631–646. doi:10.1038/cr.2010.55

324. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature. 2011;471: 189–195. doi:10.1038/nature09730

325. Morin RD, Mungall K, Pleasance E, Mungall AJ, Goya R, Huff RD, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood. 2013;122: 1256–1265. doi:10.1182/blood-2013-02-483727

326. Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C, 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: 3879–3884. doi:10.1073/pnas.1121343109

327. Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell. 1996;87: 953–959.

328. Bannister AJ, Kouzarides T. The CBP co-activator is a histone acetyltransferase. Nature. 1996;384: 641–643. doi:10.1038/384641a0

329. Struhl K. Histone acetylation and transcriptional regulatory mechanisms. Genes Dev. 1998;12: 599–606.

330. Grunstein M. Histone acetylation in chromatin structure and transcription. Nature. 1997;389: 349–352. doi:10.1038/38664

331. Wagener R, Alexandrov LB, Montesinos-Rongen M, Schlesner M, Haake A, Drexler HG, et al. Analysis of mutational signatures in exomes from B-cell lymphoma cell lines suggest APOBEC3 family members to be involved in the pathogenesis of primary effusion lymphoma. Leukemia. 2015; doi:10.1038/leu.2015.22

332. Kamb A, Gruis NA, Weaver-Feldhaus J, Liu Q, Harshman K, Tavtigian SV, et al. A cell cycle regulator potentially involved in genesis of many tumor types. Science. 1994;264: 436–440.

333. Chan FK, Zhang J, Cheng L, Shapiro DN, Winoto A. Identification of human and mouse p19, a novel CDK4 and CDK6 inhibitor with homology to p16ink4. Mol Cell Biol. 1995;15: 2682–2688.

334. Hirai H, Roussel MF, Kato JY, Ashmun RA, Sherr CJ. Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. Mol Cell Biol. 1995;15: 2672–2681.

335. Zhang Y, Xiong Y, Yarbrough WG. ARF promotes MDM2 degradation and stabilizes p53: ARF-INK4a locus deletion impairs both the Rb and p53 tumor suppression pathways. Cell. 1998;92: 725–734.

336. Pinyol M, Cobo F, Bea S, Jares P, Nayach I, Fernandez PL, et al. p16(INK4a) gene inactivation by deletions, mutations, and hypermethylation is associated with transformed and aggressive variants of non-Hodgkin's lymphomas. Blood. 1998;91: 2977–2984.

337. Alhejaily A, Day AG, Feilotter HE, Baetz T, Lebrun DP. Inactivation of the CDKN2A tumor-suppressor gene by deletion or methylation is common at diagnosis in follicular lymphoma and associated with poor clinical outcome. Clin Cancer Res Off J Am Assoc Cancer Res. 2014;20: 1676–1686. doi:10.1158/1078-0432.CCR-13-2175

338. Soufir N, Queille S, Liboutet M, Thibaudeau O, Bachelier F, Delestaing G, et al. Inactivation of the CDKN2A and the p53 tumour suppressor genes in external genital carcinomas and their precursors. Br J Dermatol. 2007;156: 448–453. doi:10.1111/j.1365-2133.2006.07604.x

339. Guney S, Jardin F, Bertrand P, Mareschal S, Parmentier F, Picquenot J-M, et al. Several mechanisms lead to the inactivation of the CDKN2A (P16), P14ARF, or CDKN2B (P15) genes in the GCB and ABC molecular DLBCL subtypes. Genes Chromosomes Cancer. 2012;51: 858–867. doi:10.1002/gcc.21970

340. Baur AS, Shaw P, Burri N, Delacrétaz F, Bosman FT, Chaubert P. Frequent methylation silencing of p15(INK4b) (MTS2) and p16(INK4a) (MTS1) in B-cell and T-cell lymphomas. Blood. 1999;94: 1773–1781.

341. Sánchez-Beato M, Sáez AI, Navas IC, Algara P, Sol Mateo M, Villuendas R, et al. Overall survival in aggressive B-cell lymphomas is dependent on the accumulation of alterations in p53, p16, and p27. Am J Pathol. 2001;159: 205–213. doi:10.1016/S0002-9440(10)61686-0

342. Perk J, Iavarone A, Benezra R. Id family of helix-loop-helix proteins in cancer. Nat Rev Cancer. 2005;5: 603–614. doi:10.1038/nrc1673

343. Norton JD. ID helix-loop-helix proteins in cell growth, differentiation and tumorigenesis. J Cell Sci. 2000;113 (Pt 22): 3897–3905.

344. D'Cruz LM, Stradner MH, Yang CY, Goldrath AW. E and Id proteins influence invariant NKT cell sublineage differentiation and proliferation. J Immunol Baltim Md 1950. 2014;192: 2227–2236. doi:10.4049/jimmunol.1302904

345. Liu C, Jin R, Wang H-C, Tang H, Liu Y-F, Qian X-P, et al. Id1 expression promotes peripheral CD4+ T cell proliferation and survival upon TCR activation without co-stimulation. Biochem Biophys Res Commun. 2013;436: 47–52. doi:10.1016/j.bbrc.2013.05.054

346. Kee BL. Id3 induces growth arrest and caspase-2-dependent apoptosis in B lymphocyte progenitors. J Immunol Baltim Md 1950. 2005;175: 4518–4527.

347. Benezra R, Davis RL, Lockshon D, Turner DL, Weintraub H. The protein Id: a negative regulator of helix-loop-helix DNA binding proteins. Cell. 1990;61: 49–59.

348. Loveys DA, Streiff MB, Kato GJ. E2A basic-helix-loop-helix transcription factors are negatively regulated by serum growth factors and by the Id3 protein. Nucleic Acids Res. 1996;24: 2813–2820.

349. Littlewood TD, Evan GI. Transcription factors 2: helix-loop-helix. Protein Profile. 1995;2: 621–702.

350. Bain G, Maandag EC, Izon DJ, Amsen D, Kruisbeek AM, Weintraub BC, et al. E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell. 1994;79: 885–892.

351. Zhuang Y, Soriano P, Weintraub H. The helix-loop-helix gene E2A is required for B cell formation. Cell. 1994;79: 875–884.

352. Zhuang Y, Cheng P, Weintraub H. B-lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2-2, and HEB. Mol Cell Biol. 1996;16: 2898–2905.

353. Cisse B, Caton ML, Lehner M, Maeda T, Scheu S, Locksley R, et al. Transcription factor E2-2 is an essential and specific regulator of plasmacytoid dendritic cell development. Cell. 2008;135: 37–48. doi:10.1016/j.cell.2008.09.016

354. Powell LM, Jarman AP. Context dependence of proneural bHLH proteins. Curr Opin Genet Dev. 2008;18: 411–417. doi:10.1016/j.gde.2008.07.012

355. Flora A, Garcia JJ, Thaller C, Zoghbi HY. The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors. Proc Natl Acad Sci U S A. 2007;104: 15382–15387. doi:10.1073/pnas.0707456104

356. Riechmann V, van Crüchten I, Sablitzky F. The expression pattern of Id4, a novel dominant negative helix-loop-helix protein, is distinct from Id1, Id2 and Id3. Nucleic Acids Res. 1994;22: 749–755.

357. Deed RW, Jasiok M, Norton JD. Lymphoid-specific expression of the Id3 gene in hematopoietic cells. Selective antagonism of E2A basic helix-loop-helix protein associated with Id3-induced differentiation of erythroleukemia cells. J Biol Chem. 1998;273: 8278–8286.

358. Li J, Maruyama T, Zhang P, Konkel JE, Hoffman V, Zamarron B, et al. Mutation of inhibitory helix-loop-helix protein Id3 causes γδ T-cell lymphoma in mice. Blood. 2010;116: 5615–5621. doi:10.1182/blood-2010-03-274506

359. Gebauer N, Bernard V, Feller AC, Merz H. ID3 mutations are recurrent events in double-hit B-cell lymphomas. Anticancer Res. 2013;33: 4771–4778.

360. Spender LC, Inman GJ. Developments in Burkitt's lymphoma: novel cooperations in oncogenic MYC signaling. Cancer Manag Res. 2014;6: 27–38. doi:10.2147/CMAR.S37745

361. Sharma P, Patel D, Chaudhary J. Id1 and Id3 expression is associated with increasing grade of prostate cancer: Id3 preferentially regulates CDKN1B. Cancer Med. 2012;1: 187–197. doi:10.1002/cam4.19

362. Phi JH, Choi SA, Lim S-H, Lee J, Wang K-C, Park S-H, et al. ID3 contributes to cerebrospinal fluid seeding and poor prognosis in medulloblastoma. BMC Cancer. 2013;13: 291. doi:10.1186/1471-2407-13-291

363. Light W, Vernon AE, Lasorella A, Iavarone A, LaBonne C. Xenopus Id3 is required downstream of Myc for the formation of multipotent neural crest progenitor cells. Dev Camb Engl. 2005;132: 1831–1841. doi:10.1242/dev.01734

364. Seitz V, Butzhammer P, Hirsch B, Hecht J, Gütgemann I, Ehlers A, et al. Deep sequencing of MYC DNA-binding sites in Burkitt lymphoma. PloS One. 2011;6: e26837. doi:10.1371/journal.pone.0026837

365. Seet CS, Brumbaugh RL, Kee BL. Early B cell factor promotes B lymphopoiesis with reduced interleukin 7 responsiveness in the absence of E2A. J Exp Med. 2004;199: 1689–1700. doi:10.1084/jem.20032202

366. Schuldiner O, Benvenisty N. A DNA microarray screen for genes involved in c-MYC and N-MYC oncogenesis in human tumors. Oncogene. 2001;20: 4984–4994. doi:10.1038/sj.onc.1204459

367. Aukema SM, Kreuz M, Kohler CW, Rosolowski M, Hasenclever D, Hummel M, et al. Biological characterization of adult MYC-translocation-positive mature B-cell lymphomas other than molecular Burkitt lymphoma. Haematologica. 2014;99: 726–735. doi:10.3324/haematol.2013.091827

368. Aukema SM, Siebert R, Schuuring E, van Imhoff GW, Kluin-Nelemans HC, Boerma E-J, et al. Double-hit B-cell lymphomas. Blood. 2011;117: 2319–2331. doi:10.1182/blood-2010-09-297879

369. Huether R, Dong L, Chen X, Wu G, Parker M, Wei L, et al. The landscape of somatic mutations in epigenetic regulators across 1,000 paediatric cancer genomes. Nat Commun. 2014;5: 3630. doi:10.1038/ncomms4630

370. Hasselblatt M, Nagel I, Oyen F, Bartelheim K, Russell RB, Schüller U, et al. SMARCA4-mutated atypical teratoid/rhabdoid tumors are associated with inherited germline alterations and poor prognosis. Acta Neuropathol (Berl). 2014;128: 453–456. doi:10.1007/s00401-014-1323-x

371. Witkowski L, Carrot-Zhang J, Albrecht S, Fahiminiya S, Hamel N, Tomiak E, et al. Germline and somatic SMARCA4 mutations characterize small cell carcinoma of the ovary, hypercalcemic type. Nat Genet. 2014;46: 438–443. doi:10.1038/ng.2931

372. Khavari PA, Peterson CL, Tamkun JW, Mendel DB, Crabtree GR. BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. Nature. 1993;366: 170–174. doi:10.1038/366170a0

373. Neish AS, Anderson SF, Schlegel BP, Wei W, Parvin JD. Factors associated with the mammalian RNA polymerase II holoenzyme. Nucleic Acids Res. 1998;26: 847–853.

374. Schwabish MA, Struhl K. The Swi/Snf complex is important for histone eviction during transcriptional activation and RNA polymerase II elongation in vivo. Mol Cell Biol. 2007;27: 6987–6995. doi:10.1128/MCB.00717-07

375. Wang W, Côté J, Xue Y, Zhou S, Khavari PA, Biggar SR, et al. Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. EMBO J. 1996;15: 5370–5382.

376. Wang W, Xue Y, Zhou S, Kuo A, Cairns BR, Crabtree GR. Diversity and specialization of mammalian SWI/SNF complexes. Genes Dev. 1996;10: 2117–2130.

377. Patsialou A, Wilsker D, Moran E. DNA-binding properties of ARID family proteins. Nucleic Acids Res. 2005;33: 66–80. doi:10.1093/nar/gki145

378. Winston F, Allis CD. The bromodomain: a chromatin-targeting module? Nat Struct Biol. 1999;6: 601–604. doi:10.1038/10640

379. Martens JA, Winston F. Recent advances in understanding chromatin remodeling by Swi/Snf complexes. Curr Opin Genet Dev. 2003;13: 136–142.

380. Sena JA, Wang L, Hu C-J. BRG1 and BRM chromatin-remodeling complexes regulate the hypoxia response by acting as coactivators for a subset of hypoxia-inducible transcription factor target genes. Mol Cell Biol. 2013;33: 3849–3863. doi:10.1128/MCB.00731-13

381. Leung JY, Nevins JR. E2F6 associates with BRG1 in transcriptional regulation. PloS One. 2012;7: e47967. doi:10.1371/journal.pone.0047967

382. Bakshi R, Hassan MQ, Pratap J, Lian JB, Montecino MA, van Wijnen AJ, et al. The human SWI/SNF complex associates with RUNX1 to control transcription of hematopoietic target genes. J Cell Physiol. 2010;225: 569–576. doi:10.1002/jcp.22240

383. Attanasio C, Nord AS, Zhu Y, Blow MJ, Biddie SC, Mendenhall EM, et al. Tissue-specific SMARCA4 binding at active and repressed regulatory elements during embryogenesis. Genome Res. 2014;24: 920–929. doi:10.1101/gr.168930.113

384. Ho L, Miller EL, Ronan JL, Ho WQ, Jothi R, Crabtree GR. esBAF facilitates pluripotency by conditioning the genome for LIF/STAT3 signalling and by regulating polycomb function. Nat Cell Biol. 2011;13: 903–913. doi:10.1038/ncb2285

385. Hendricks KB, Shanahan F, Lees E. Role for BRG1 in cell cycle control and tumor suppression. Mol Cell Biol. 2004;24: 362–376.

386. Lee D, Kim JW, Seo T, Hwang SG, Choi E-J, Choe J. SWI/SNF complex interacts with tumor suppressor p53 and is necessary for the activation of p53-mediated transcription. J Biol Chem. 2002;277: 22330–22337. doi:10.1074/jbc.M111987200

387. Wu S, Ge Y, Huang L, Liu H, Xue Y, Zhao Y. BRG1, the ATPase subunit of SWI/SNF chromatin remodeling complex, interacts with HDAC2 to modulate telomerase expression in human cancer cells. Cell Cycle Georget Tex. 2014;13: 2869–2878. doi:10.4161/15384101.2014.946834

388. Seo S, Richardson GA, Kroll KL. The SWI/SNF chromatin remodeling protein Brg1 is required for vertebrate neurogenesis and mediates transactivation of Ngn and NeuroD. Dev Camb Engl. 2005;132: 105–115. doi:10.1242/dev.01548

389. Bultman S, Gebuhr T, Yee D, La Mantia C, Nicholson J, Gilliam A, et al. A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. Mol Cell. 2000;6: 1287–1295.

390. De la Serna IL, Carlson KA, Imbalzano AN. Mammalian SWI/SNF complexes promote MyoD-mediated muscle differentiation. Nat Genet. 2001;27: 187–190. doi:10.1038/84826

391. Patenge N, Elkin SK, Oettinger MA. ATP-dependent remodeling by SWI/SNF and ISWI proteins stimulates V(D)J cleavage of 5 S arrays. J Biol Chem. 2004;279: 35360–35367. doi:10.1074/jbc.M405790200

392. Morshead KB, Ciccone DN, Taverna SD, Allis CD, Oettinger MA. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. Proc Natl Acad Sci U S A. 2003;100: 11577–11582. doi:10.1073/pnas.1932643100

393. Holley DW, Groh BS, Wozniak G, Donohoe DR, Sun W, Godfrey V, et al. The BRG1 chromatin remodeler regulates widespread changes in gene expression and cell proliferation during B cell activation. J Cell Physiol. 2014;229: 44–52. doi:10.1002/jcp.24414

394. Dunaief JL, Strober BE, Guha S, Khavari PA, Alin K, Luban J, et al. The retinoblastoma protein and BRG1 form a complex and cooperate to induce cell cycle arrest. Cell. 1994;79: 119–130.

395. Bultman SJ, Gebuhr TC, Magnuson T. A Brg1 mutation that uncouples ATPase activity from chromatin remodeling reveals an essential role for SWI/SNF-related complexes in beta-globin expression and erythroid development. Genes Dev. 2005;19: 2849–2861. doi:10.1101/gad.1364105

396. Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J, et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nat Genet. 2013;45: 592–601. doi:10.1038/ng.2628

397. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature. 2014;505: 495–501. doi:10.1038/nature12912

398. Rohde M, Richter J, Schlesner M, Betts MJ, Claviez A, Bonn BR, et al. Recurrent RHOA mutations in pediatric Burkitt lymphoma treated according to the NHL-BFM protocols. Genes Chromosomes Cancer. 2014;53: 911–916. doi:10.1002/gcc.22202

399. Muppidi JR, Schmitz R, Green JA, Xiao W, Larsen AB, Braun SE, et al. Loss of signalling via Gα13 in germinal centre B-cell-derived lymphoma. Nature. 2014;516: 254–258. doi:10.1038/nature13765

400. Cattoretti G, Mandelbaum J, Lee N, Chaves AH, Mahler AM, Chadburn A, et al. Targeted disruption of the S1P2 sphingosine 1-phosphate receptor gene leads to diffuse large B-cell lymphoma formation. Cancer Res. 2009;69: 8686–8692. doi:10.1158/0008-5472.CAN-09-1110

401. Silvera D, Gamarnik AV, Andino R. The N-terminal K homology domain of the poly(rC)-binding protein is a major determinant for binding to the poliovirus 5'-untranslated region and acts as an inhibitor of viral translation. J Biol Chem. 1999;274: 38163–38170.

402. Thisted T, Lyakhov DL, Liebhaber SA. Optimized RNA targets of two closely related triple KH domain proteins, heterogeneous nuclear ribonucleoprotein K and alphaCP-2KL, suggest Distinct modes of RNA recognition. J Biol Chem. 2001;276: 17484–17496. doi:10.1074/jbc.M010594200

403. Chkheidze AN, Liebhaber SA. A novel set of nuclear localization signals determine distributions of the alphaCP RNA-binding proteins. Mol Cell Biol. 2003;23: 8405–8415.

404. Meng Q, Rayala SK, Gururaj AE, Talukder AH, O'Malley BW, Kumar R. Signaling-dependent and coordinated regulation of transcription, splicing, and translation resides in a single coregulator, PCBP1. Proc Natl Acad Sci U S A. 2007;104: 5866–5871. doi:10.1073/pnas.0701065104

405. Chaudhury A, Chander P, Howe PH. Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. RNA N Y N. 2010;16: 1449–1462. doi:10.1261/rna.2254110

406. Ren C, Cho S-J, Jung Y-S, Chen X. DNA polymerase η is regulated by poly(rC)-binding protein 1 via mRNA stability. Biochem J. 2014;464: 377–386. doi:10.1042/BJ20141164

407. Berry AM, Flock KE, Loh HH, Ko JL. Molecular basis of cellular localization of poly C binding protein 1 in neuronal cells. Biochem Biophys Res Commun. 2006;349: 1378–1386. doi:10.1016/j.bbrc.2006.09.012

408. Misteli T, Cáceres JF, Spector DL. The dynamics of a pre-mRNA splicing factor in living cells. Nature. 1997;387: 523–527. doi:10.1038/387523a0

409. Gualco G, Queiroga EM, Weiss LM, Klumb CEN, Harrington WJ, Bacchi CE. Frequent expression of multiple myeloma 1/interferon regulatory factor 4 in Burkitt lymphoma. Hum Pathol. 2009;40: 565–571. doi:10.1016/j.humpath.2008.07.021

410. Chuang S-S, Ye H, Du M-Q, Lu C-L, Dogan A, Hsieh P-P, et al. Histopathology and immunohistochemistry in distinguishing Burkitt lymphoma from diffuse large B-cell lymphoma with very high proliferation index and with or without a starry-sky pattern: a comparative study with EBER and FISH. Am J Clin Pathol. 2007;128: 558–564. doi:10.1309/EQJR3D3V0CCQGP04

411. Mittrücker HW, Matsuyama T, Grossman A, Kündig TM, Potter J, Shahinian A, et al. Requirement for the transcription factor LSIRF/IRF4 for mature B and T lymphocyte function. Science. 1997;275: 540–543.

412. Natkunam Y, Warnke RA, Montgomery K, Falini B, van De Rijn M. Analysis of MUM1/IRF4 protein expression using tissue microarrays and immunohistochemistry. Mod Pathol Off J U S Can Acad Pathol Inc. 2001;14: 686–694. doi:10.1038/modpathol.3880373

413. Lu R. Interferon regulatory factor 4 and 8 in B-cell development. Trends Immunol. 2008;29: 487–492. doi:10.1016/j.it.2008.07.006

414. Shaffer AL, Rosenwald A, Hurt EM, Giltnane JM, Lam LT, Pickeral OK, et al. Signatures of the immune response. Immunity. 2001;15: 375–385.

415. Sciammas R, Shaffer AL, Schatz JH, Zhao H, Staudt LM, Singh H. Graded expression of interferon regulatory factor-4 coordinates isotype switching with plasma cell differentiation. Immunity. 2006;25: 225–236. doi:10.1016/j.immuni.2006.07.009

416. Saura C, Bendell J, Jerusalem G, Su S, Ru Q, De Buck S, et al. Phase Ib study of Buparlisib plus Trastuzumab in patients with HER2-positive advanced or metastatic breast cancer that has progressed on Trastuzumab-based therapy. Clin Cancer Res Off J Am Assoc Cancer Res. 2014;20: 1935–1945. doi:10.1158/1078-0432.CCR-13-1070

417. Schwartz GK, LoRusso PM, Dickson MA, Randolph SS, Shaik MN, Wilner KD, et al. Phase I study of PD 0332991, a cyclin-dependent kinase inhibitor, administered in 3-week cycles (Schedule 2/1). Br J Cancer. 2011;104: 1862–1868. doi:10.1038/bjc.2011.177

418. Niesvizky R, Badros AZ, Costa LJ, Ely SA, Singhal SB, Stadtmauer EA, et al. Phase 1/2 study of CDK4/6 inhibitor palbociclib (PD-0332991) with bortezomib and dexamethasone in relapsed/refractory multiple myeloma. Leuk Lymphoma. 2015; 1–21. doi:10.3109/10428194.2015.1030641

419. Johnston PB, Yuan R, Cavalli F, Witzig TE. Targeted therapy in lymphoma. J Hematol OncolJ Hematol Oncol. 2010;3: 45. doi:10.1186/1756-8722-3-45

420. Cen O, Longnecker R. Rapamycin reverses splenomegaly and inhibits tumor development in a transgenic model of Epstein-Barr virus-related Burkitt's lymphoma. Mol Cancer Ther. 2011;10: 679–686. doi:10.1158/1535-7163.MCT-10-0833

421. Bernard OA, Berger R. Molecular basis of 11q23 rearrangements in hematopoietic malignant proliferations. Genes Chromosomes Cancer. 1995;13: 75–85.

422. Kobayashi H, Espinosa R, Thirman MJ, Gill HJ, Fernald AA, Diaz MO, et al. Heterogeneity of breakpoints of 11q23 rearrangements in hematologic malignancies identified with fluorescence in situ hybridization. Blood. 1993;82: 547–551.

423. Furlan A, Vercamer C, Bouali F, Damour I, Chotteau-Lelievre A, Wernert N, et al. Ets-1 controls breast cancer cell balance between invasion and growth. Int J Cancer J Int Cancer. 2014;135: 2317–2328. doi:10.1002/ijc.28881

424. Verschoor ML, Verschoor CP, Singh G. Ets-1 global gene expression profile reveals associations with metabolism and oxidative stress in ovarian and breast cancers. Cancer Metab. 2013;1: 17. doi:10.1186/2049-3002-1-17

425. Pallai R, Bhaskar A, Sodi V, Rice LM. Ets1 and Elk1 transcription factors regulate cancerous inhibitor of protein phosphatase 2A expression in cervical and endometrial carcinoma cells. Transcription. 2012;3: 323–335. doi:10.4161/trns.22518

426. Schwering I, Bräuninger A, Klein U, Jungnickel B, Tinguely M, Diehl V, et al. Loss of the B-lineage-specific gene expression program in Hodgkin and Reed-Sternberg cells of Hodgkin lymphoma. Blood. 2003;101: 1505–1512. doi:10.1182/blood-2002-03-0839

427. Aukema SM, Kreuz M, Kohler CW, Rosolowski M, Hasenclever D, Hummel M, et al. Biological characterization of adult MYC-translocation-positive mature B-cell lymphomas other than molecular Burkitt lymphoma. Haematologica. 2014;99: 726–735. doi:10.3324/haematol.2013.091827

428. Berger R, Bernheim A. Cytogenetic studies on Burkitt's lymphoma-leukemia. Cancer Genet Cytogenet. 1982;7: 231–244.

429. Koshiba M, Ogawa K, Hamazaki S, Sugiyama T, Ogawa O, Kitajima T. The effect of formalin fixation on DNA and the extraction of high-molecular-weight DNA from fixed and embedded tissues. Pathol Res Pract. 1993;189: 66–72. doi:10.1016/S0344-0338(11)80118-4

430. Scholtysik R, Kreuz M, Hummel M, Rosolowski M, Szczepanowski M, Klapper W, et al. Characterization of genomic imbalances in diffuse large B-cell lymphoma by detailed SNP-chip analysis. Int J Cancer J Int Cancer. 2015;136: 1033–1042. doi:10.1002/ijc.29072

431. Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000;343: 1910–1916. doi:10.1056/NEJM200012283432602

432. Dierlamm J, Wlodarska I, Michaux L, Vermeesch JR, Meeus P, Stul M, et al. FISH identifies different types of duplications with 12q13-15 as the commonly involved segment in B-cell lymphoproliferative malignancies characterized by partial trisomy 12. Genes Chromosomes Cancer. 1997;20: 155–166.

433. Kienle DL, Korz C, Hosch B, Benner A, Mertens D, Habermann A, et al. Evidence for distinct pathomechanisms in genetic subgroups of chronic lymphocytic leukemia revealed by quantitative expression analysis of cell cycle, activation, and apoptosis-associated genes. J Clin Oncol Off J Am Soc Clin Oncol. 2005;23: 3780–3792. doi:10.1200/JCO.2005.02.568

434. Winkler D, Schneider C, Kröber A, Pasqualucci L, Lichter P, Döhner H, et al. Protein expression analysis of chromosome 12 candidate genes in chronic lymphocytic leukemia (CLL). Leukemia. 2005;19: 1211–1215. doi:10.1038/sj.leu.2403778

435. Gaidano G, Hauptschein RS, Parsa NZ, Offit K, Rao PH, Lenoir G, et al. Deletions involving two distinct regions of 6q in B-cell non-Hodgkin lymphoma. Blood. 1992;80: 1781–1787.

436. Offit K, Parsa NZ, Gaidano G, Filippa DA, Louie D, Pan D, et al. 6q deletions define distinct clinico-pathologic subsets of non-Hodgkin's lymphoma. Blood. 1993;82: 2157–2162.

437. Taborelli M, Tibiletti MG, Martin V, Pozzi B, Bertoni F, Capella C. Chromosome band 6q deletion pattern in malignant lymphomas. Cancer Genet Cytogenet. 2006;165: 106–113. doi:10.1016/j.cancergencyto.2005.06.025

438. Nanjangud G, Rao PH, Teruya-Feldstein J, Donnelly G, Qin J, Mehra S, et al. Molecular cytogenetic analysis of follicular lymphoma (FL) provides detailed characterization of chromosomal instability associated with the t(14;18)(q32;q21) positive and negative subsets and histologic progression. Cytogenet Genome Res. 2007;118: 337–344. doi:10.1159/000108318

439. Pasquale EB. Eph receptors and ephrins in cancer: bidirectional signalling and beyond. Nat Rev Cancer. 2010;10: 165–180. doi:10.1038/nrc2806

440. Oricchio E, Wendel H-G. Mining the cancer genome uncovers therapeutic activity of EphA7 against lymphoma. Cell Cycle Georget Tex. 2012;11: 1076–1080. doi:10.4161/cc.11.6.19451

441. Nakamoto M, Bergemann AD. Diverse roles for the Eph family of receptor tyrosine kinases in carcinogenesis. Microsc Res Tech. 2002;59: 58–67. doi:10.1002/jemt.10177

442. Shaffer AL, Lin KI, Kuo TC, Yu X, Hurt EM, Rosenwald A, et al. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. Immunity. 2002;17: 51–62.

443. Pasqualucci L, Compagno M, Houldsworth J, Monti S, Grunn A, Nandula SV, et al. Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. J Exp Med. 2006;203: 311–317. doi:10.1084/jem.20052204

444. Berger R, Le Coniat M, Derré J, Vecchione D. Secondary nonrandom chromosomal abnormalities of band 13q34 in Burkitt lymphoma-leukemia. Genes Chromosomes Cancer. 1989;1: 115–118.

445. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. Nature. 2005;435: 834–838. doi:10.1038/nature03702

446. Tagawa H, Karube K, Tsuzuki S, Ohshima K, Seto M. Synergistic action of the microRNA-17 polycistron and Myc in aggressive cancer development. Cancer Sci. 2007;98: 1482–1490. doi:10.1111/j.1349-7006.2007.00531.x

447. Tanzer A, Stadler PF. Molecular evolution of a microRNA cluster. J Mol Biol. 2004;339: 327–335. doi:10.1016/j.jmb.2004.03.065

448. Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, et al. A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res. 2005;65: 9628–9632. doi:10.1158/0008-5472.CAN-05-2352

449. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT. c-Myc-regulated microRNAs modulate E2F1 expression. Nature. 2005;435: 839–843. doi:10.1038/nature03677

450. Ventura A, Young AG, Winslow MM, Lintault L, Meissner A, Erkeland SJ, et al. Targeted deletion reveals essential and overlapping functions of the miR-17 through 92 family of miRNA clusters. Cell. 2008;132: 875–886. doi:10.1016/j.cell.2008.02.019

451. He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, et al. A microRNA polycistron as a potential human oncogene. Nature. 2005;435: 828–833. doi:10.1038/nature03552

452. Olive V, Bennett MJ, Walker JC, Ma C, Jiang I, Cordon-Cardo C, et al. miR-19 is a key oncogenic component of mir-17-92. Genes Dev. 2009;23: 2839–2849. doi:10.1101/gad.1861409

453. Graham TA, Ferkey DM, Mao F, Kimelman D, Xu W. Tcf4 can specifically recognize beta-catenin using alternative conformations. Nat Struct Biol. 2001;8: 1048–1052. doi:10.1038/nsb718

454. Hecht A, Stemmler MP. Identification of a promoter-specific transcriptional activation domain at the C terminus of the Wnt effector protein T-cell factor 4. J Biol Chem. 2003;278: 3776–3785. doi:10.1074/jbc.M210081200

455. Wang S-H, Li N, Wei Y, Li Q-R, Yu Z-P. β-catenin deacetylation is essential for WNT-induced proliferation of breast cancer cells. Mol Med Rep. 2014;9: 973–978. doi:10.3892/mmr.2014.1889

456. Zhou D, Bai F, Zhang X, Hu M, Zhao G, Zhao Z, et al. SOX10 is a novel oncogene in hepatocellular carcinoma through Wnt/β-catenin/TCF4 cascade. Tumour Biol J Int Soc Oncodevelopmental Biol Med. 2014;35: 9935–9940. doi:10.1007/s13277-014-1893-1

457. Schoch C, Schnittger S, Klaus M, Kern W, Hiddemann W, Haferlach T. AML with 11q23/MLL abnormalities as defined by the WHO classification: incidence, partner chromosomes, FAB subtype, age distribution, and prognostic impact in an unselected series of 1897 cytogenetically analyzed AML cases. Blood. 2003;102: 2395–2402. doi:10.1182/blood-2003-02-0434

458. Armstrong SA, Look AT. Molecular genetics of acute lymphoblastic leukemia. J Clin Oncol Off J Am Soc Clin Oncol. 2005;23: 6306–6315. doi:10.1200/JCO.2005.05.047

459. Nakamura T, Mori T, Tada S, Krajewski W, Rozovskaia T, Wassell R, et al. ALL-1 is a histone methyltransferase that assembles a supercomplex of proteins involved in transcriptional regulation. Mol Cell. 2002;10: 1119–1128.

460. Dou Y, Milne TA, Tackett AJ, Smith ER, Fukuda A, Wysocka J, et al. Physical association and coordinate function of the H3 K4 methyltransferase MLL1 and the H4 K16 acetyltransferase MOF. Cell. 2005;121: 873–885. doi:10.1016/j.cell.2005.04.031

461. Slany RK. The molecular biology of mixed lineage leukemia. Haematologica. 2009;94: 984–993. doi:10.3324/haematol.2008.002436

462. Milne TA, Dou Y, Martin ME, Brock HW, Roeder RG, Hess JL. MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci U S A. 2005;102: 14765–14770. doi:10.1073/pnas.0503630102

463. Liu H, Takeda S, Kumar R, Westergard TD, Brown EJ, Pandita TK, et al. Phosphorylation of MLL by ATR is required for execution of mammalian S-phase checkpoint. Nature. 2010;467: 343–346. doi:10.1038/nature09350

464. McMahon KA, Hiew SY-L, Hadjur S, Veiga-Fernandes H, Menzel U, Price AJ, et al. Mll has a critical role in fetal and adult hematopoietic stem cell self-renewal. Cell Stem Cell. 2007;1: 338–345. doi:10.1016/j.stem.2007.07.002

465. Jude CD, Climer L, Xu D, Artinger E, Fisher JK, Ernst P. Unique and independent roles for MLL in adult hematopoietic stem cells and progenitors. Cell Stem Cell. 2007;1: 324–337. doi:10.1016/j.stem.2007.05.019

466. Wan X, Hu B, Liu J, Feng X, Xiao W. Zebrafish mll gene is essential for hematopoiesis. J Biol Chem. 2011;286: 33345–33357. doi:10.1074/jbc.M111.253252

467. Poppe B, Vandesompele J, Schoch C, Lindvall C, Mrozek K, Bloomfield CD, et al. Expression analyses identify MLL as a prominent target of 11q23 amplification and support an etiologic role for MLL gain of function in myeloid malignancies. Blood. 2004;103: 229–235. doi:10.1182/blood-2003-06-2163

468. Faber J, Krivtsov AV, Stubbs MC, Wright R, Davis TN, van den Heuvel-Eibrink M, et al. HOXA9 is required for survival in human MLL-rearranged acute leukemias. Blood. 2009;113: 2375–2385. doi:10.1182/blood-2007-09-113597

469. Yip BH, So CWE. Mixed lineage leukemia protein in normal and leukemic stem cells. Exp Biol Med Maywood NJ. 2013;238: 315–323. doi:10.1177/1535370213480717

470. Meerabux J, Yaspo ML, Roebroek AJ, Van de Ven WJ, Lister TA, Young BD. A new member of the proprotein convertase gene family (LPC) is located at a chromosome translocation breakpoint in lymphomas. Cancer Res. 1996;56: 448–451.

471. Ho YS, Swenson L, Derewenda U, Serre L, Wei Y, Dauter Z, et al. Brain acetylhydrolase that inactivates platelet-activating factor is a G-protein-like trimer. Nature. 1997;385: 89–93. doi:10.1038/385089a0

472. Stafforini DM, McIntyre TM, Zimmerman GA, Prescott SM. Platelet-activating factor acetylhydrolases. J Biol Chem. 1997;272: 17895–17898.

473. Zimmerman GA, McIntyre TM, Prescott SM, Stafforini DM. The platelet-activating factor signaling system and its regulators in syndromes of inflammation and thrombosis. Crit Care Med. 2002;30: S294–301.

474. Roth M, Nauck M, Yousefi S, Tamm M, Blaser K, Perruchoud AP, et al. Platelet-activating factor exerts mitogenic activity and stimulates expression of interleukin 6 and interleukin 8 in human lung fibroblasts via binding to its functional receptor. J Exp Med. 1996;184: 191–201.

475. Bennett SA, Birnboim HC. Receptor-mediated and protein kinase-dependent growth enhancement of primary human fibroblasts by platelet activating factor. Mol Carcinog. 1997;20: 366–375.

476. Bennett SA, Leite LC, Birnboim HC. Platelet activating factor, an endogenous mediator of inflammation, induces phenotypic transformation of rat embryo cells. Carcinogenesis. 1993;14: 1289–1296.

477. Denizot Y, Guglielmi L, Donnard M, Trimoreau F. Platelet-activating factor and normal or leukaemic haematopoiesis. Leuk Lymphoma. 2003;44: 775–782. doi:10.1080/1042819031000067549

478. Leprince C, Vivier E, Treton D, Galanaud P, Benveniste J, Richard Y, et al. Immunoregulatory functions of paf-acether. VI. Dual effect on human B cell proliferation. Lipids. 1991;26: 1204–1208.

479. Denizot Y, Dupuis F, Trimoreau F, Praloran V, Liozon E. Decreased levels of platelet-activating factor in blood of patients with lymphoid and nonlymphoid hematologic malignancies. Blood. 1995;85: 2992–2993.

480. Pospisilova H, Baens M, Michaux L, Stul M, Van Hummelen P, Van Loo P, et al. Interstitial del(14)(q) involving IGH: a novel recurrent aberration in B-NHL. Leukemia. 2007;21: 2079–2083. doi:10.1038/sj.leu.2404739

481. Nagel I, Bug S, Tönnies H, Ammerpohl O, Richter J, Vater I, et al. Biallelic inactivation of TRAF3 in a subset of B-cell lymphomas with interstitial del(14)(q24.1q32.33). Leukemia. 2009;23: 2153–2155. doi:10.1038/leu.2009.149

482. Stoecklin G, Colombi M, Raineri I, Leuenberger S, Mallaun M, Schmidlin M, et al. Functional cloning of BRF1, a regulator of ARE-dependent mRNA turnover. EMBO J. 2002;21: 4709–4718.

483. Stoecklin G, Gross B, Ming X-F, Moroni C. A novel mechanism of tumor suppression by destabilizing AU-rich growth factor mRNA. Oncogene. 2003;22: 3554–3561. doi:10.1038/sj.onc.1206418

484. Sharrocks AD. The ETS-domain transcription factor family. Nat Rev Mol Cell Biol. 2001;2: 827–837. doi:10.1038/35099076

485. Nye JA, Petersen JM, Gunther CV, Jonsen MD, Graves BJ. Interaction of murine ets-1 with GGA-binding sites establishes the ETS domain as a new DNA-binding motif. Genes Dev. 1992;6: 975–990.

486. Karim FD, Urness LD, Thummel CS, Klemsz MJ, McKercher SR, Celada A, et al. The ETS-domain: a new DNA-binding motif that recognizes a purine-rich core DNA sequence. Genes Dev. 1990;4: 1451–1453.

487. Nguyen HV, Mouly E, Chemin K, Luinaud R, Despres R, Fermand J-P, et al. The Ets-1 transcription factor is required for Stat1-mediated T-bet expression and IgG2a class switching in mouse B cells. Blood. 2012;119: 4174–4181. doi:10.1182/blood-2011-09-378182

488. Raghunandan R, Frissora FW, Muthusamy N. Modulation of Ets-1 expression in B lymphocytes is dependent on the antigen receptor-mediated activation signals and cell cycle status. Scand J Immunol. 2013;77: 75–83. doi:10.1111/sji.12012

489. Lee C-G, Kwon H-K, Sahoo A, Hwang W, So J-S, Hwang J-S, et al. Interaction of Ets-1 with HDAC1 represses IL-10 expression in Th1 cells. J Immunol Baltim Md 1950. 2012;188: 2244–2253. doi:10.4049/jimmunol.1101614

490. Tsao H-W, Tai T-S, Tseng W, Chang H-H, Grenningloh R, Miaw S-C, et al. Ets-1 facilitates nuclear entry of NFAT proteins and their recruitment to the IL-2 promoter. Proc Natl Acad Sci U S A. 2013;110: 15776–15781. doi:10.1073/pnas.1304343110

491. Russell L, Garrett-Sinha LA. Transcription factor Ets-1 in cytokine and chemokine gene regulation. Cytokine. 2010;51: 217–226. doi:10.1016/j.cyto.2010.03.006

492. Ben-David Y, Giddens EB, Letwin K, Bernstein A. Erythroleukemia induction by Friend murine leukemia virus: insertional activation of a new member of the ets gene family, Fli-1, closely linked to c-ets-1. Genes Dev. 1991;5: 908–918.

493. Watson DK, Smyth FE, Thompson DM, Cheng JQ, Testa JR, Papas TS, et al. The ERGB/Fli-1 gene: isolation and characterization of a new member of the family of human ETS transcription factors. Cell Growth Differ Mol Biol J Am Assoc Cancer Res. 1992;3: 705–713.

494. Pusztaszeri MP, Seelentag W, Bosman FT. Immunohistochemical expression of endothelial markers CD31, CD34, von Willebrand factor, and Fli-1 in normal human tissues. J Histochem Cytochem Off J Histochem Soc. 2006;54: 385–395. doi:10.1369/jhc.4A6514.2005

495. Zhang XK, Moussa O, LaRue A, Bradshaw S, Molano I, Spyropoulos DD, et al. The transcription factor Fli-1 modulates marginal zone and follicular B cell development in mice. J Immunol Baltim Md 1950. 2008;181: 1644–1654.

496. Maier H, Ostraat R, Parenti S, Fitzsimmons D, Abraham LJ, Garvie CW, et al. Requirements for selective recruitment of Ets proteins and activation of mb-1/Ig-alpha gene transcription by Pax-5 (BSAP). Nucleic Acids Res. 2003;31: 5483–5489.

497. Lesault I, Quang CT, Frampton J, Ghysdael J. Direct regulation of BCL-2 by FLI-1 is involved in the survival of FLI-1-transformed erythroblasts. EMBO J. 2002;21: 694–703.

498. Truong AHL, Cervi D, Lee J, Ben-David Y. Direct transcriptional regulation of MDM2 by Fli-1. Oncogene. 2005;24: 962–969. doi:10.1038/sj.onc.1208323

499. Kayali S, Giraud G, Morlé F, Guyot B. Spi-1, Fli-1 and Fli-3 (miR-17-92) oncogenes contribute to a single oncogenic network controlling cell proliferation in friend erythroleukemia. PloS One. 2012;7: e46799. doi:10.1371/journal.pone.0046799

500. Ho HH, Ivashkiv LB. Downregulation of Friend leukemia virus integration 1 as a feedback mechanism that restrains lipopolysaccharide induction of matrix metalloproteases and interleukin-10 in human macrophages. J Interferon Cytokine Res Off J Int Soc Interferon Cytokine Res. 2010;30: 893–900. doi:10.1089/jir.2010.0046

501. Mélet F, Motro B, Rossi DJ, Zhang L, Bernstein A. Generation of a novel Fli-1 protein by gene targeting leads to a defect in thymus development and a delay in Friend virus-induced erythroleukemia. Mol Cell Biol. 1996;16: 2708–2718.

502. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M, et al. Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. Nature. 1992;359: 162–165. doi:10.1038/359162a0

503. Kotenko SV, Krause CD, Izotova LS, Pollack BP, Wu W, Pestka S. Identification and functional characterization of a second chain of the interleukin-10 receptor complex. EMBO J. 1997;16: 5894–5903. doi:10.1093/emboj/16.19.5894

504. Josephson K, Logsdon NJ, Walter MR. Crystal structure of the IL-10/IL-10R1 complex reveals a shared receptor binding site. Immunity. 2001;15: 35–46.

505. Liu Y, Wei SH, Ho AS, de Waal Malefyt R, Moore KW. Expression cloning and characterization of a human IL-10 receptor. J Immunol Baltim Md 1950. 1994;152: 1821–1829.

506. Liu Y, de Waal Malefyt R, Briere F, Parham C, Bridon JM, Banchereau J, et al. The EBV IL-10 homologue is a selective agonist with impaired binding to the IL-10 receptor. J Immunol Baltim Md 1950. 1997;158: 604–613.

507. Weber-Nordt RM, Meraz MA, Schreiber RD. Lipopolysaccharide-dependent induction of IL-10 receptor expression on murine fibroblasts. J Immunol Baltim Md 1950. 1994;153: 3734–3744.

508. Roth I, Fisher SJ. IL-10 is an autocrine inhibitor of human placental cytotrophoblast MMP-9 production and invasion. Dev Biol. 1999;205: 194–204. doi:10.1006/dbio.1998.9122

509. Gibbs VC, Pennica D. CRF2-4: isolation of cDNA clones encoding the human and mouse proteins. Gene. 1997;186: 97–101.

510. Crepaldi L, Gasperini S, Lapinet JA, Calzetti F, Pinardi C, Liu Y, et al. Up-regulation of IL-10R1 expression is required to render human neutrophils fully responsive to IL-10. J Immunol Baltim Md 1950. 2001;167: 2312–2322.

511. De Waal Malefyt R, Abrams J, Bennett B, Figdor CG, de Vries JE. Interleukin 10(IL-10) inhibits cytokine synthesis by human monocytes: an autoregulatory role of IL-10 produced by monocytes. J Exp Med. 1991;174: 1209–1220.

512. D'Andrea A, Aste-Amezaga M, Valiante NM, Ma X, Kubin M, Trinchieri G. Interleukin 10 (IL-10) inhibits human lymphocyte interferon gamma-production by suppressing natural killer cell stimulatory factor/IL-12 synthesis in accessory cells. J Exp Med. 1993;178: 1041–1048.

513. Jenkins JK, Malyak M, Arend WP. The effects of interleukin-10 on interleukin-1 receptor antagonist and interleukin-1 beta production in human monocytes and neutrophils. Lymphokine Cytokine Res. 1994;13: 47–54.

514. Dickensheets HL, Freeman SL, Smith MF, Donnelly RP. Interleukin-10 upregulates tumor necrosis factor receptor type-II (p75) gene expression in endotoxin-stimulated human monocytes. Blood. 1997;90: 4162–4171.

515. Go NF, Castle BE, Barrett R, Kastelein R, Dang W, Mosmann TR, et al. Interleukin 10, a novel B cell stimulatory factor: unresponsiveness of X chromosome-linked immunodeficiency B cells. J Exp Med. 1990;172: 1625–1631.

516. Saeland S, Duvert V, Moreau I, Banchereau J. Human B cell precursors proliferate and express CD23 after CD40 ligation. J Exp Med. 1993;178: 113–120.

517. Burdin N, Van Kooten C, Galibert L, Abrams JS, Wijdenes J, Banchereau J, et al. Endogenous IL-6 and IL-10 contribute to the differentiation of CD40-activated human B lymphocytes. J Immunol Baltim Md 1950. 1995;154: 2533–2544.

518. Czarneski J, Lin YC, Chong S, McCarthy B, Fernandes H, Parker G, et al. Studies in NZB IL-10 knockout mice of the requirement of IL-10 for progression of B-cell lymphoma. Leukemia. 2004;18: 597–606. doi:10.1038/sj.leu.2403244

519. Alas S, Bonavida B. Rituximab inactivates signal transducer and activation of transcription 3 (STAT3) activity in B-non-Hodgkin's lymphoma through inhibition of the interleukin 10 autocrine/paracrine loop and results in down-regulation of Bcl-2 and sensitization to cytotoxic drugs. Cancer Res. 2001;61: 5137–5144.

520. Yu D, Dews M, Park A, Tobias JW, Thomas-Tikhonenko A. Inactivation of Myc in murine two-hit B lymphomas causes dormancy with elevated levels of interleukin 10 receptor and CD20: implications for adjuvant therapies. Cancer Res. 2005;65: 5454–5461. doi:10.1158/0008-5472.CAN-04-4197

521. Masood R, Zhang Y, Bond MW, Scadden DT, Moudgil T, Law RE, et al. Interleukin-10 is an autocrine growth factor for acquired immunodeficiency syndrome-related B-cell lymphoma. Blood. 1995;85: 3423–3430.

522. Beatty PR, Krams SM, Martinez OM. Involvement of IL-10 in the autonomous growth of EBV-transformed B cell lines. J Immunol Baltim Md 1950. 1997;158: 4045–4051.

523. Weber-Nordt RM, Riley JK, Greenlund AC, Moore KW, Darnell JE, Schreiber RD. Stat3 recruitment by two distinct ligand-induced, tyrosine-phosphorylated docking sites in the interleukin-10 receptor intracellular domain. J Biol Chem. 1996;271: 27954–27961.

524. Levy Y, Brouet JC. Interleukin-10 prevents spontaneous death of germinal center B cells by induction of the bcl-2 protein. J Clin Invest. 1994;93: 424–428. doi:10.1172/JCI116977

525. Li L, Krajewski S, Reed JC, Choi YS. The apoptosis and proliferation of SAC-activated B cells by IL-10 are associated with changes in Bcl-2, Bcl-xL, and Mcl-1 expression. Cell Immunol. 1997;178: 33–41. doi:10.1006/cimm.1997.1129

526. Williams L, Bradley L, Smith A, Foxwell B. Signal transducer and activator of transcription 3 is the dominant mediator of the anti-inflammatory effects of IL-10 in human macrophages. J Immunol Baltim Md 1950. 2004;172: 567–576.

527. Wang P, Wu P, Siegel MI, Egan RW, Billah MM. Interleukin (IL)-10 inhibits nuclear factor kappa B (NF kappa B) activation in human monocytes. IL-10 and IL-4 suppress cytokine synthesis by different mechanisms. J Biol Chem. 1995;270: 9558–9563.

528. Schottelius AJ, Mayo MW, Sartor RB, Baldwin AS. Interleukin-10 signaling blocks inhibitor of kappaB kinase activity and nuclear factor kappaB DNA binding. J Biol Chem. 1999;274: 31868–31874.

529. Asadullah K, Sterry W, Volk HD. Interleukin-10 therapy--review of a new approach. Pharmacol Rev. 2003;55: 241–269. doi:10.1124/pr.55.2.4

530. Zhou JH, Broussard SR, Strle K, Freund GG, Johnson RW, Dantzer R, et al. IL-10 inhibits apoptosis of promyeloid cells by activating insulin receptor substrate-2 and phosphatidylinositol 3'-kinase. J Immunol Baltim Md 1950. 2001;167: 4436–4442.

531. Corinti S, Albanesi C, la Sala A, Pastore S, Girolomoni G. Regulatory activity of autocrine IL-10 on dendritic cell functions. J Immunol Baltim Md 1950. 2001;166: 4312–4318.

532. Elbim C, Reglier H, Fay M, Delarche C, Andrieu V, El Benna J, et al. Intracellular pool of IL-10 receptors in specific granules of human neutrophils: differential mobilization by proinflammatory mediators. J Immunol Baltim Md 1950. 2001;166: 5201–5207.

533. Kotenko SV, Izotova LS, Mirochnitchenko OV, Esterova E, Dickensheets H, Donnelly RP, et al. Identification, cloning, and characterization of a novel soluble receptor that binds IL-22 and neutralizes its activity. J Immunol Baltim Md 1950. 2001;166: 7096–7103.

534. Ding Y, Qin L, Zamarin D, Kotenko SV, Pestka S, Moore KW, et al. Differential IL-10R1 expression plays a critical role in IL-10-mediated immune regulation. J Immunol Baltim Md 1950. 2001;167: 6884–6892.

535. Rautert R, Schinköthe T, Franklin J, Weihrauch M, Boll B, Pogge E, et al. Elevated pretreatment interleukin-10 serum level is an International Prognostic Score (IPS)-independent risk factor for early treatment failure in advanced stage Hodgkin lymphoma. Leuk Lymphoma. 2008;49: 2091–2098. doi:10.1080/10428190802441339

536. Fabre-Guillevin E, Tabrizi R, Coulon V, Monnereau A, Eghbali H, Soubeyran I, et al. Aggressive non-Hodgkin's lymphoma: concomitant evaluation of interleukin-2, soluble interleukin-2 receptor, interleukin-4, interleukin-6, interleukin-10 and correlation with outcome. Leuk Lymphoma. 2006;47: 603–611. doi:10.1080/10428190500361029

537. Sarris AH, Kliche KO, Pethambaram P, Preti A, Tucker S, Jackow C, et al. Interleukin-10 levels are often elevated in serum of adults with Hodgkin's disease and are associated with inferior failure-free survival. Ann Oncol Off J Eur Soc Med Oncol ESMO. 1999;10: 433–440.

538. Güler N, Kelkitli E, Atay H, Erdem D, Alaçam H, Bek Y, et al. The Relationship of T Helper-2 Pathway Components Interleukin-4, Interleukin-10, Immunoglobulin E, and Eosinophils with Prognostic Markers in Non-Hodgkin Lymphoma: A Case-Control Study. Turk J Haematol Off J Turk Soc Haematol. 2014;31: 381–387. doi:10.4274/tjh.2013.0328

539. Cao S, Liu J, Song L, Ma X. The protooncogene c-Maf is an essential transcription factor for IL-10 gene expression in macrophages. J Immunol Baltim Md 1950. 2005;174: 3484–3492.

540. Ziegler-Heitbrock L, Lötzerich M, Schaefer A, Werner T, Frankenberger M, Benkhart E. IFN-alpha induces the human IL-10 gene by recruiting both IFN regulatory factor 1 and Stat3. J Immunol Baltim Md 1950. 2003;171: 285–290.

541. Iyer SS, Cheng G. Role of interleukin 10 transcriptional regulation in inflammation and autoimmune disease. Crit Rev Immunol. 2012;32: 23–63.

542. Grenningloh R, Kang BY, Ho I-C. Ets-1, a functional cofactor of T-bet, is essential for Th1 inflammatory responses. J Exp Med. 2005;201: 615–626. doi:10.1084/jem.20041330

543. Li JCB, Lau ASY. A role for mitogen-activated protein kinase and Ets-1 in the induction of interleukin-10 transcription by human immunodeficiency virus-1 Tat. Immunology. 2007;121: 337–348. doi:10.1111/j.1365-2567.2007.02580.x

544. Béguelin W, Sawh S, Chambwe N, Chun Chan F, Jiang Y, Choo J-W, et al. IL10 receptor is a novel therapeutic target in DLBCLs. Leukemia. 2015; doi:10.1038/leu.2015.57

545. Yang J, Liao X, Agarwal MK, Barnes L, Auron PE, Stark GR. Unphosphorylated STAT3 accumulates in response to IL-6 and activates transcription by binding to NFkappaB. Genes Dev. 2007;21: 1396–1408. doi:10.1101/gad.1553707

546. Yoshida Y, Kumar A, Koyama Y, Peng H, Arman A, Boch JA, et al. Interleukin 1 activates STAT3/nuclear factor-kappaB cross-talk via a unique TRAF6- and p65-dependent mechanism. J Biol Chem. 2004;279: 1768–1776. doi:10.1074/jbc.M311498200

547. Narimatsu M, Maeda H, Itoh S, Atsumi T, Ohtani T, Nishida K, et al. Tissue-specific autoregulation of the stat3 gene and its role in interleukin-6-induced survival signals in T cells. Mol Cell Biol. 2001;21: 6615–6625.

548. Huang X, Meng B, Iqbal J, Ding BB, Perry AM, Cao W, et al. Activation of the STAT3 signaling pathway is associated with poor survival in diffuse large B-cell lymphoma treated with R-CHOP. J Clin Oncol Off J Am Soc Clin Oncol. 2013;31: 4520–4528. doi:10.1200/JCO.2012.45.6004

549. Scuto A, Kujawski M, Kowolik C, Krymskaya L, Wang L, Weiss LM, et al. STAT3 inhibition is a therapeutic strategy for ABC-like diffuse large B-cell lymphoma. Cancer Res. 2011;71: 3182–3188. doi:10.1158/0008-5472.CAN-10-2380

550. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. 2014;157: 1262–1278. doi:10.1016/j.cell.2014.05.010

551. Martin L, Grigoryan A, Wang D, Wang J, Breda L, Rivella S, et al. Identification and characterization of small molecules that inhibit nonsense-mediated RNA decay and suppress nonsense p53 mutations. Cancer Res. 2014;74: 3104–3113. doi:10.1158/0008-5472.CAN-13-2235

552. Johnson JK, Waddell N, kConFab Investigators, Chenevix-Trench G. The application of nonsense-mediated mRNA decay inhibition to the identification of breast cancer susceptibility genes. BMC Cancer. 2012;12: 246. doi:10.1186/1471-2407-12-246

553. De S, Shaknovich R, Riester M, Elemento O, Geng H, Kormaksson M, et al. Aberration in DNA methylation in B-cell lymphomas has a complex origin and increases with disease severity. PLoS Genet. 2013;9: e1003137. doi:10.1371/journal.pgen.1003137

554. Li Y-J, Zhao X, Vecchiarelli-Federico LM, Li Y, Datti A, Cheng Y, et al. Drug-mediated inhibition of Fli-1 for the treatment of leukemia. Blood Cancer J. 2012;2: e54. doi:10.1038/bcj.2011.52

8 Appendices

8.1 Characteristics of B-cell lymphomas cases of the ICGC MMML-Seq cohort

Table S 1 gives on overview on characteristics of the patients with BL, Burkitt leukemia or DLBCL of the ICGC MMML-Seq cohort as well as on the respective analyses they were subjected to.

S 1: Overview on BL and DLBCL cases of the ICGC MMML-Seq cohort used in this thesis for different analysis.

Case	Age (yrs)	Sex	Diagnosis	MYC-break	*SMARCA4* mutation[1]	*PCBP1* mutation[2]	TERT expression	450k
4177434	16	f	BL	positive	x	x	x	
4193278	17	m	BL	positive	x	x		x
4127766	8	m	BL	positive	x	x		
4190495	15	m	BL	positive	x	x	x	x
4182393	10	m	BL	positive	x	x	x	x
4119027	12	m	BL	positive	x	x	x	x
4125240	4	m	BL	positive	x	x	x	x
4112512	18	f	BL	positive	x	x	x	x
4130003	6	m	BL	positive	x	x	x	
4146289	14	m	BL	positive	x	x	x	
4108627	5	f	Burkitt leukemia	positive	x			
4133511	5	m	BL	positive	x	x		x
4142267	5	m	Burkitt leukemia	positive	x	x		
4161696	5	m	Burkitt leukemia	positive	x			
4177856	10	m	BL	positive	x	x		x
4178310	10	m	Burkitt leukemia	positive	x			
4178345	10	m	Burkitt leukemia	positive	x	x		
4189998	13	m	BL	positive	x	x		x
4190784	4	m	Burkitt leukemia	negative	x			
4194218	4	m	BL	positive	x	x		x
4194891	4	m	BL	positive	x	x		x
4116738	15	m	DLBCL	negative				x
4163639	75	f	DLBCL	negative				x
4135099	49	m	DLBCL	negative				x
4108101	66	m	DLBCL	negative				x
4101316	74	f	DLBCL	negative				x
4104893	16	m	DLBCL	negative				x

Age denotes patient age in years (yrs) at diagnosis; sex is denoted m for male and f for female; BL is denoted for Burkitt lymphoma, DLBCL for diffuse-large B-cell lymphoma; [1] results of the *SMARCA4* mutation analysis are under revision in the journal *Nature Genetics*; 450k, analysis of the global DNA methylation using HumanMethylation 450k Bead Chip; [2] results of the *PCBP1* mutations analysis are accepted for publication in the journal *Genes, Chromosomes and Cancer*.

8.2 Characteristics of B-cell lymphoma cases of the MMML cohort

Table S 2 gives on overview on characteristics of the patients with BL and DLBCL of the MMML cohort as well as on the respective analyses they were subjected to.

S 2: Overview on BL and DLBCL cases of the MMML cohort used in this thesis for different analyses.

Case	Age (yrs)	Sex	Diagnosis	MYC-break	*PCBP1* mutation[1]	450k
MPI-005	5	m	Atypical BL	positive	x	x
MPI-016	44	f	Atypical BL	positive	x	
MPI-017	40	m	BL	positive	x	x
MPI-027	63	f	Atypical BL	positive	x	
MPI-033	4	m	BL	positive	x	
MPI-043	5	m	Atypical BL	positive	x	
MPI-044	13	m	Atypical BL	positive	x	
MPI-048	13	f	BL	positive	x	x
MPI-055	12	m	Atypical BL	positive	x	
MPI-071	54	m	Atypical BL	positive	x	x
MPI-079	2	f	BL	positive	x	
MPI-080	2	m	BL	positive	x	
MPI-082	4	m	Atypical BL	positive	x	
MPI-084	57	m	BL	positive	x	x
MPI-085	26	f	BL	positive	x	x
MPI-089	7	m	BL	positive	x	
MPI-101	76	m	Atypical BL	positive	x	x
MPI-114	8	m	BL	positive	x	
MPI-123	24	m	Atypical BL	positive	x	
MPI-139	46	m	BL	positive	x	
MPI-144	33	m	BL	positive	x	x
MPI-171	24	f	DLBCL	positive		x
MPI-323	40	m	Atypical BL	positive	x	x
MPI-527	8	f	BL	positive	x	
MPI-529	10	f	BL	positive	x	
MPI-532	5	m	Atypical BL	positive	x	
MPI-579	5	f	BL	positive	x	
MPI-580	9	m	BL	positive	x	
MPI-590	17	m	Atypical BL	positive	x	
MPI-625	12	m	BL	positive	x	
MPI-680	43	m	Atypical BL	positive		x

Age denotes patient age in years (yrs) at diagnosis; sex is denoted m for male and f for female; BL is denoted for Burkitt lymphoma, DLBCL for diffuse-large B-cell lymphoma; 450k, analysis of the global DNA methylation using HumanMethylation 450k Bead Chip; [1] results of the *PCBP1* mutations analysis are accepted for publication in the journal *Genes, Chromosomes and Cancer*.

8.3 Karyotypes of cell lines studied

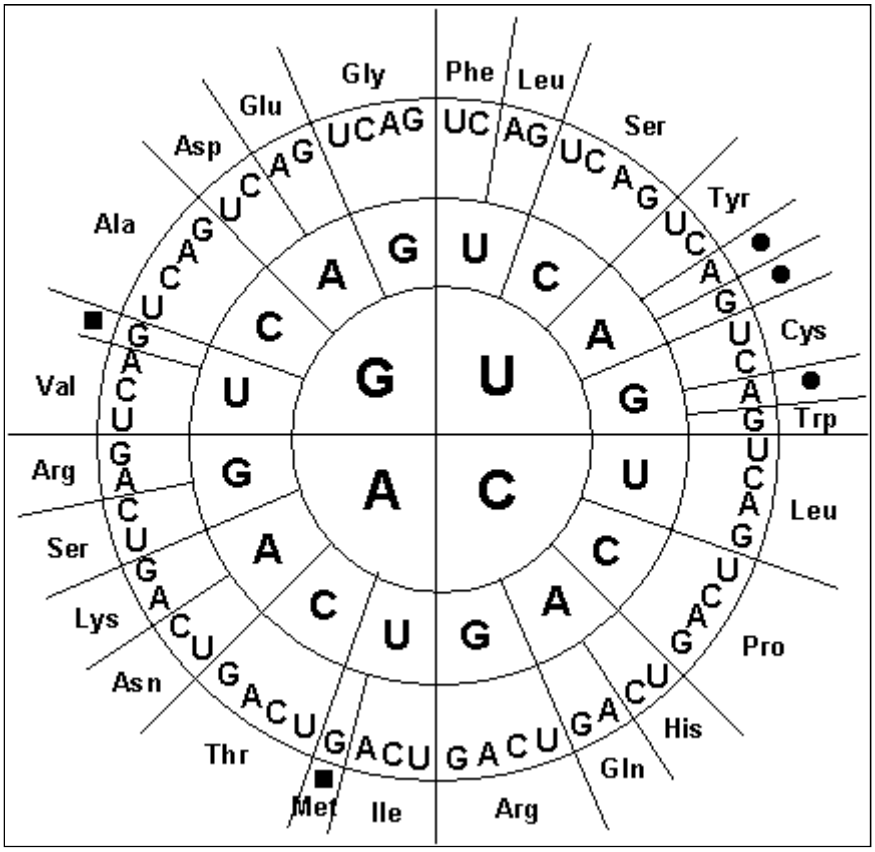
All but two cell line karyotypes were available and downloaded at the homepage of the DSMZ (http://www.dsmz.de/home.html, 05/10/2014). The exceptions were the karyotypes of U-698-M and MLMA cells which were obtained in-house.

S 3: Karyotypes of cell line studied.

Cell line	Diagnosis	Karyotype
BL-2	BL	44-47<2n>XY, der(1)t(1;7)(q32;q11.2),der(6)t(1;6)(q21;q25),t(8;22)(q24;q11.2)
BL-41	BL	48(42-49)<2n>XY, +7, -13, 2mar, add(8)(q24), t(8;14)(q24;q32), der(15)t(13;15)(q13;p11), add(17)(p12), subclonal rearrangements at 1q23, 7p22, 11q13
BL-70	BL	47(42-49)<2n>XY, +7, inv(1)(p21q21), del(2)(q33), t(8;14)(q24;q32), t(12;22)(q21.1;q13.2)
BLUE-1	BL	53(49-53)<2n>XY, +6, +13, +16, +20, +20, +20, +21, t(6;20)(q15;q11.2), der(6)t(6;20)(q15;q11.2), t(8;14)(q24;q32), der(20)del(20)(p12.2p13.2)t(6;20)(q15;q11.2)t(6;11)(q16;p13)x2
Ca46	BL	46(45-48)<2n>X/XY, dup(1)(q21q32), dup(7)(q12q22), t(8;14)(q24;q32)
U-698-M	BL	49(44-50)<2n>XY, +3, +7, -14, +mar, dup(1)(q43q21.2), der(2)t(2;3)(p16;p11), add(3)(p11), del(6)(q15q22), del(9)(p22), dup(11)(q23q13), add(13)(p12), add(16)(q24) +t(8;14)(q24;q23), In-house data
HT	mnBLL[1]	46(42-46)<2n>XY, +2, der(2;4)(p10;q10), der(2)(del)(2)(p11?q21), dup(10)(q11q22-23), dup(11)(?q23qter)
MLMA	mnBLL[1]	44~46,X,del(X)(q13q21),-2,-3,der(3;15)(q10;p10),-4,add(4)(p16),del(5)(q23q35),der(7)(?), -8,der(11)dup(11)(q22q25)hsr(11)(q22~23),-12,der(15)t(?8;15)(q23;q15),add(17)(p11), +18,add(18)(q21),der(19)t(12;19)(q14;q13.3)+3~5mar[cp8], In-house data
Su-DHL-5	mnBLL[1]	47(41-48)<2n>XX, +12, del(6)(q13), del(12)(q13) - sideline with del(6)x2
Su-DHL-6	GCB-DLBCL	47(42-48)<2n>X, -Y, +6, +7, del(4)(q23), del(6)(p21.3p22.2), i(6p), del(7)(q?22q?32), der(8)t(8;9)(q24;p13), der(9)t(8;19;9)(q24;q13;p13), dup(11)(q24q25), t(14;18)(q32;q21), der(22)t(?7;22)(?q32;p11)
Su-DHL-10	GCB-DLBCL	47(43-48)<2n>XY, +7, der(8)t(X;8)(q25;p23)t(8;X)(q24;q26)t(X;14)(q28;q32), del(10)(q22q24),der(11)t(Y;11)(q11;q25),der(14)t(8;14)(q24;q32),der(18)t(14;18)(q32;q21)
Karpas422	GCB-DLBCL	hyperdiploid, 10% polyploidy; 47(44-48)<2n>XX, +14, t(2;10)(p23;q22), t(4;11)(q21;q24), t(4;16)(q21;p13), der(14)t(14;18)(q32;q21)x2 *FISH: 4,7 t(4;16;10)(q21;p13;q23), no t(4;11)(q21;q23), t(14;18)(q32;q21), +der(14)t(14;18)(q32;q21)*
RIVA	ABC-DLBCL	48-51<2n>XX, +7, +18, +19, +mar, t(1;22;16)(p35;p12;q12.2), del(3)(q11-12), trp(3)(q11-12q28-29), t(4;8)(q22;q24), del(6)(q13), der(15)t(13;15) (q11;p1?1), der(18)amp(18)(q21) dup(18)(q21q23), der(19)t(6;19) (p21;p13)add(19(q13.3)

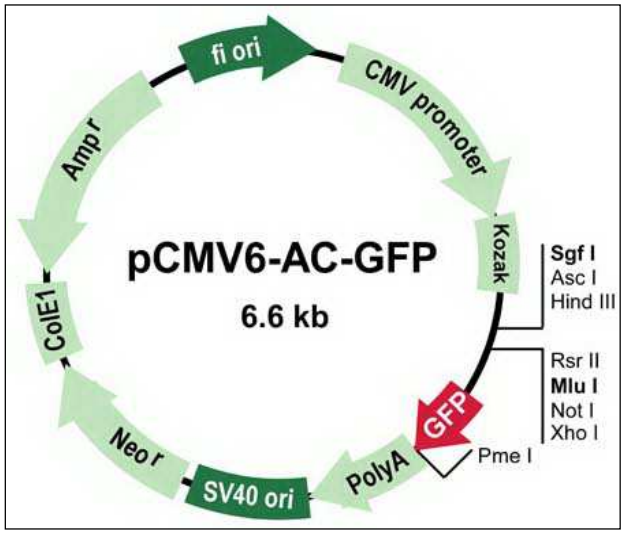
Diagnosis: mnBLL, *MYC*-negative Burkitt-like lymphoma; BL, Burkitt lymphoma, GCB DLBCL, germinal center B-cell like diffuse large B-cell lymphoma; ABC DLBCL, activated B-cell like diffuse large B-cell lymphoma. [1] Diagnosis as defined in this thesis (3.2.2).

8.4 Abbreviations and codons for amino acids and translation stop

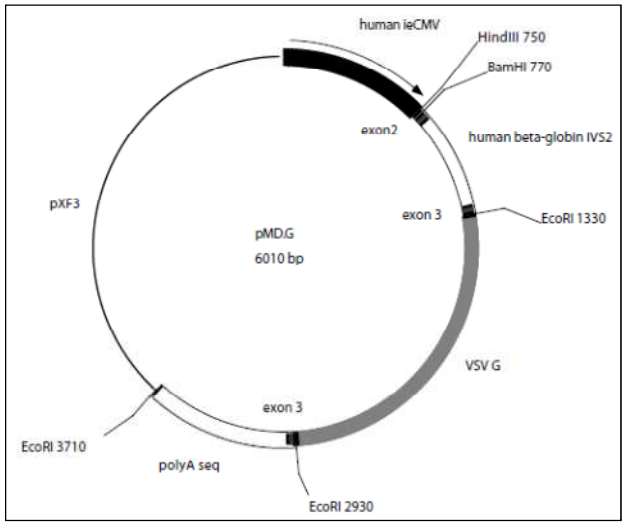


S 4: Overview on abbreviations and codons for amino acids and translational stop (•/*)
(from http://www.operon.com/products/custom_oligos/images/geneticcode.gif, 14/04/2015).

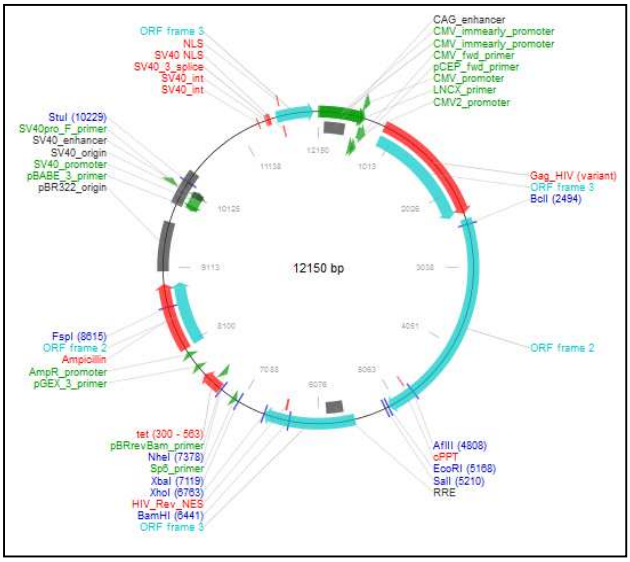
8.5 Vector maps of plasmids used in this study



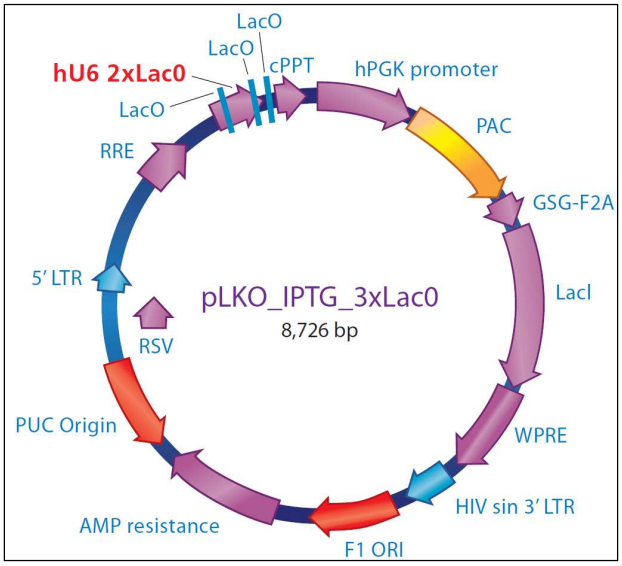
S 5: Vector map of pCMV6-AC-GFP which was obtained from OriGene Technologies (http://www.origene.com/destination_vector/PS100010.aspx, 25/03/2015).



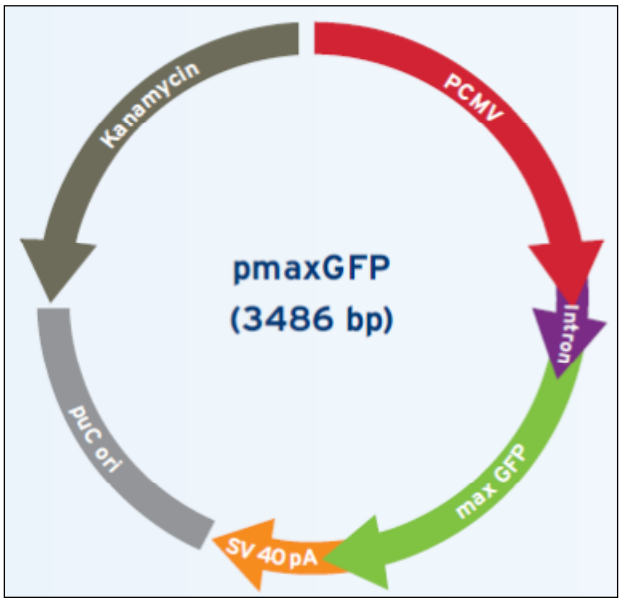
S 6: Vector map of pMD.G which was provided by Dr.rer. nat. E. Murga Penas, Institute of Human Genetics, Kiel (http://web.mit.edu/jacks-lab/protocols/pMDGmap.pdf, 25/03/2015).



S 7: Vector map of pCMV-dR8.91 which was provided by Dr. rer. nat. E. Murga Penas, Institute of Human Genetics, Kiel (http://www.addgene.org/vector-database/2221/, 25/03/2015).



S 8: Vector map of pLKO_IPTG_3xLacO which was obtained from Sigma Aldrich (http://www.sigmaaldrich.com/life-science/functional-genomics-and-rnai/shrna/library-information/vector-map.html#inducible, 25/03/2015).



S 9: Vector map of pmax which was obtained from Amaxa/Lonza (http://bio.lonza.com/fileadmin/ groups/ FAQs /public/Technology_Flyer.pdf, 25/03/2015).

8.6 Sequences of the new exon variants of the *TERT* gene

Sequence of new exon variant 1, chr5:1,286,280-1,287,201 bp (hg19):

5′AGACCCGCCTGGTGCACTCTGATTCTCCACTTGCCTGTTGCATGTCCTCGTTCCCTTGTTTCTCACCACCTCTTGGGTTGCCATGTGCGT
TTCCTGCCGAGTGTGTGTTGATCCTCTCGTTGCCTCCTGGTCACTGGGCATTTGCTTTTATTTCTCTTTGCTTAGTGTTACCCCCTGATCTT
TTTATTGTCGTTGTTTGCTTTTGTTTATTGAGACAGTCTCACTCTGTCACCCAGGCTGGAGTGTAATGGCACAATCTCGGCTCACTGCAAC
CTCTGCCTCCTCGGTTCAAGCAGTTCTCATTCCTCAACCTCATGAGTAGCTGGGATTACAGGCGCCCACCACCACGCCTGGCTAATTTTT
GTATTTTTAGTAGAGATAGGCTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCTGCCCGCCTTGGCCTCCCAC
AGTGCTGGGATTACAGGTGCAAGCCACCGTGCCCGGCATACCTTGATCTTTTAAAATGAAGTCTGAAACATTGCTACCCTTGTCCTGAG
CAATAAGACCCTTAGTGTATTTTAGCTCTGGCCACCCCCCAGCCTGTGTGCTGTTTTCCCTGCTGACTTAGTTCTATCTCAGGCATCTTGA
CACCCCCACAAGCTAAGCATTATTAATATTGTTTTCCGTGTTGAGTGTTTCTGTAGCTTTGCCCCCGCCCTGCTTTTCCTCCTTTGTTCCCC
GTCTGTCTTCTGTCTCAGGCCCGCCGTCTGGGGTCCCCTTCCTTGTCCTTTGCGTGGTTCTTCTGTCTTGTTATTGCTGGTAAACCCCAGC
TTTACCTGTGCTGGCCTCCATGGCATCTAGCGACGTCCGGGGACCTCTGCTTATGATGCACAGATGAAGATGTGGAGACTCACGAGGA
GGGCGGTCATCTTGGCCC3′

Sequence of new exon variant 2, chr5:1,286,239-1,287,201 bp (hg19):

5′AGACCCGCCTGGTGCACTCTGATTCTCCACTTGCCTGTTGCATGTCCTCGTTCCCTTGTTTCTCACCACCTCTTGGGTTGCCATGTGCGT
TTCCTGCCGAGTGTGTGTTGATCCTCTCGTTGCCTCCTGGTCACTGGGCATTTGCTTTTATTTCTCTTTGCTTAGTGTTACCCCCTGATCTT
TTTATTGTCGTTGTTTGCTTTTGTTTATTGAGACAGTCTCACTCTGTCACCCAGGCTGGAGTGTAATGGCACAATCTCGGCTCACTGCAAC
CTCTGCCTCCTCGGTTCAAGCAGTTCTCATTCCTCAACCTCATGAGTAGCTGGGATTACAGGCGCCCACCACCACGCCTGGCTAATTTTT
GTATTTTTAGTAGAGATAGGCTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCTGCCCGCCTTGGCCTCCCAC
AGTGCTGGGATTACAGGTGCAAGCCACCGTGCCCGGCATACCTTGATCTTTTAAAATGAAGTCTGAAACATTGCTACCCTTGTCCTGAG
CAATAAGACCCTTAGTGTATTTTAGCTCTGGCCACCCCCCAGCCTGTGTGCTGTTTTCCCTGCTGACTTAGTTCTATCTCAGGCATCTTGA
CACCCCCACAAGCTAAGCATTATTAATATTGTTTTCCGTGTTGAGTGTTTCTGTAGCTTTGCCCCCGCCCTGCTTTTCCTCCTTTGTTCCCC
GTCTGTCTTCTGTCTCAGGCCCGCCGTCTGGGGTCCCCTTCCTTGTCCTTTGCGTGGTTCTTCTGTCTTGTTATTGCTGGTAAACCCCAGC
TTTACCTGTGCTGGCCTCCATGGCATCTAGCGACGTCCGGGGACCTCTGCTTATGATGCACAGATGAAGATGTGGAGACTCACGAGGA
GGGCGGTCATCTTGGCCCGTGAGTGTCTGGAGCACCACGTGGCCAGCGTTCCTTAGCCA3′

Sequence of new exon variant 3, chr5:1,286,014-1,287,201 bp (hg19):

5′AGACCCGCCTGGTGCACTCTGATTCTCCACTTGCCTGTTGCATGTCCTCGTTCCCTTGTTTCTCACCACCTCTTGGGTTGCCATGTGCGT
TTCCTGCCGAGTGTGTGTTGATCCTCTCGTTGCCTCCTGGTCACTGGGCATTTGCTTTTATTTCTCTTTGCTTAGTGTTACCCCCTGATCTT
TTTATTGTCGTTGTTTGCTTTTGTTTATTGAGACAGTCTCACTCTGTCACCCAGGCTGGAGTGTAATGGCACAATCTCGGCTCACTGCAAC
CTCTGCCTCCTCGGTTCAAGCAGTTCTCATTCCTCAACCTCATGAGTAGCTGGGATTACAGGCGCCCACCACCACGCCTGGCTAATTTTT
GTATTTTTAGTAGAGATAGGCTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTGACCTCAAGTGATCTGCCCGCCTTGGCCTCCCAC
AGTGCTGGGATTACAGGTGCAAGCCACCGTGCCCGGCATACCTTGATCTTTTAAAATGAAGTCTGAAACATTGCTACCCTTGTCCTGAG
CAATAAGACCCTTAGTGTATTTTAGCTCTGGCCACCCCCCAGCCTGTGTGCTGTTTTCCCTGCTGACTTAGTTCTATCTCAGGCATCTTGA
CACCCCCACAAGCTAAGCATTATTAATATTGTTTTCCGTGTTGAGTGTTTCTGTAGCTTTGCCCCCGCCCTGCTTTTCCTCCTTTGTTCCCC
GTCTGTCTTCTGTCTCAGGCCCGCCGTCTGGGGTCCCCTTCCTTGTCCTTTGCGTGGTTCTTCTGTCTTGTTATTGCTGGTAAACCCCAGC
TTTACCTGTGCTGGCCTCCATGGCATCTAGCGACGTCCGGGGACCTCTGCTTATGATGCACAGATGAAGATGTGGAGACTCACGAGGA
GGGCGGTCATCTTGGCCCGTGAGTGTCTGGAGCACCACGTGGCCAGCGTTCCTTAGCCAGTGAGTGACAGCAACGTCCGCTCGGCCTG
GGTTCAGCCTGGAAAACCCCAGGCATGTCGGGGTCTGGTGGCTCCGCGGTGTCGAGTTTGAAATCGCGCAAACCTGCGGTGTGGCGC
CAGCTCTGACGGTGCTGCCTGGCGGGGGAGTGTCTGCTTCCTCCCTTCTGCTTGGGAACCAGGACAAAGGATGAGGCTCCGAGCCGTT
GTCGCCCAACAGGAGCATGAC3′

8.7 Sequence of the splice variant of exon 2 of the *TERT* gene

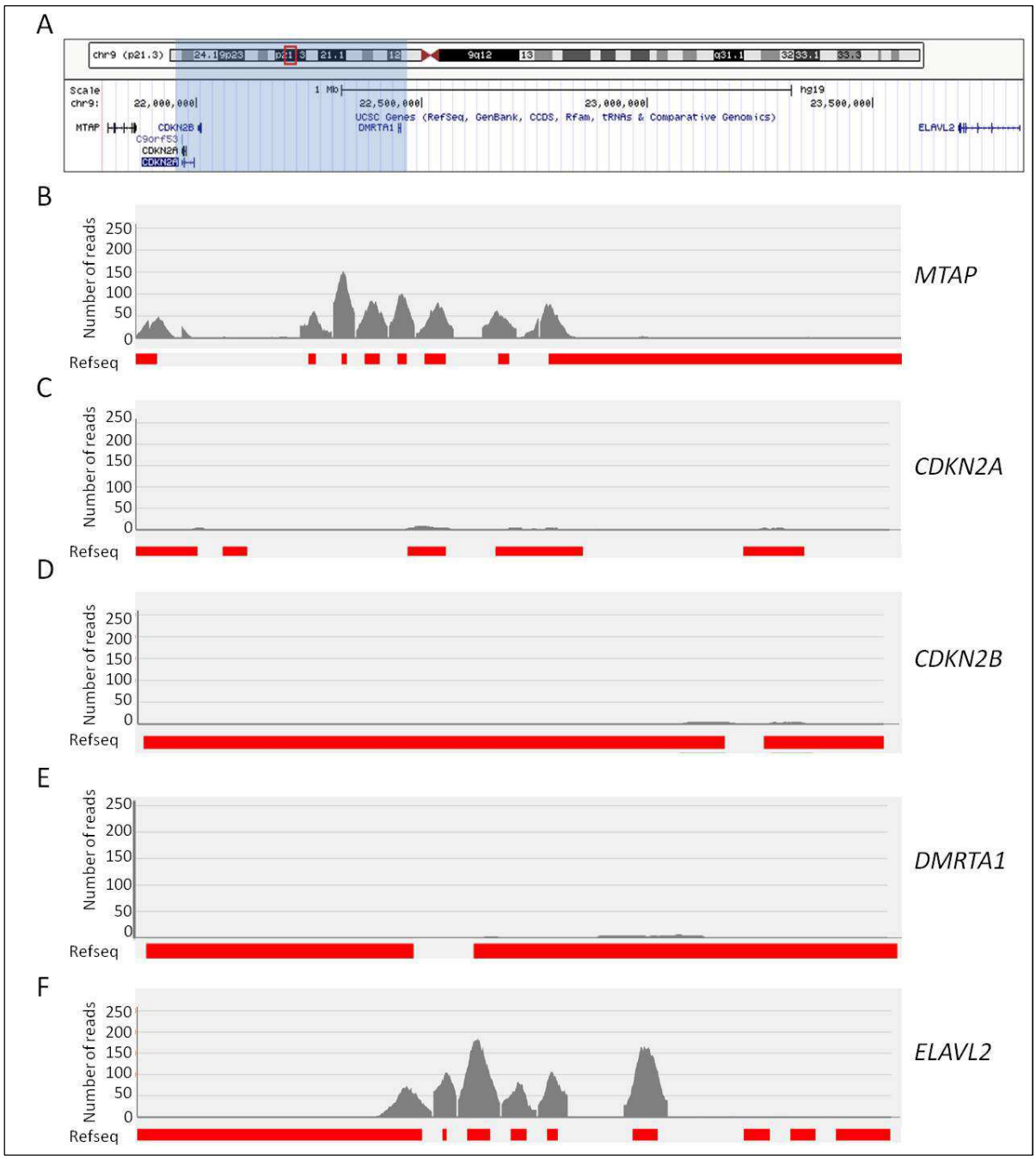
Sequence of splice variant of exon 2, chr5:1,293,079-1,294,781 bp (hg19):

5′GTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCGCGAAGAACGTGCTGGCCTTCGGCTTCGC
GCTGCTGGACGGGGCCCGCGGGGGCCCCCCCGAGGCCTTCACCACCAGCGTGCGCAGCTACCTGCCCAACACGGTGACCGACGCACT
GCGGGGGAGCGGGGCGTGGGGGCTGCTGCTGCGCCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGCACGCTGCGCGCTCTTTGT
GCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCCGCTGTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACAC
GCTAGTGGACCCCGAAGGCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGTCCCCCTGGGCCTGCCAGC
CCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCCAAGAGGCCCAGGCGTGGCGCTGCCCCTGAGCCGG
AGCGGACGCCCGTTGGGCAGGGGTCCTGGGCCCACCCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTTCTGTGTGGTGTCACCTG
CCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGCTCTCTGGCACGCGCCACTCCCACCCATCCGTGGGCCGCCAGCACCACGC
GGGCCCCCCATCCACATCGCGGCCACCACGTCCCTGGGACACGCCTTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTC
AGGCGACAAGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCTGACTGGCGCTCGGAGGCTCGTGGAGACCATC
TTTCTGGGTTCCAGGCCCTGGATGCCAGGGACTCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGCCCCTGTTTCT
GGAGCTGCTTGGGAACCACGCGCAGTGCCCCTACGGGGTGCTCCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTCACCCCAGCAGCC
GGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTCTGTGGCGGCCCCCGAGGAGGAGGACACAGACCCCCGTCGCCTGGTGCAGCTGCTC
CGCCAGCACAGCAGCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGCCTGCGCCGGCTGGTGCCCCCAGGCCTCTGGGGCTCCAGG
CACAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGCCAAGCTCTCGCTGCAGGAGCTGACGTGGA
AGATGAGCGTGCGGGACTGCGCTTGGCTGCGCAGGAGCCCAGGTGAGGAGGTGGTGGCCGTCGAGGGCCCAGGCCCCAGAGCTGAA
TGCAGTAGGGGCTCAGAAAAGGGGGCAGGCAGAGCCCTGGTCCTCCTGTCTCCATCGTCACGTGGGCACACGTGGCTTTTCGCTCAGG
ACGTCGAGTGGACACGGTGATCTCTGCCTCTGCTCTCCCTCCTGTCCAGTTTGCATAAACTTACGAGGTTCACCTTCACGTTTTGATGGA
CACGCGGTTTCCAGGCGCCGAGGCCAGAGCAGTGAACAGAGGAGGCTGGGCGCGGCAGTGGAGCCGGGTTGCCGGCAATGGGGAG
AAGTGTCTGGAAGCACAGACGCTCTGGCGAGGGTGCCTGCAG3′

The underlined sequence is the sequence which is in addition to the sequence of the wildtype exon 2 and, hence, defines the 3’ borders of the here newly identified splice variant.

8.8 Read coverage of *CDKN2A* locus in BL-2 cell line

A homozygous loss was detected in the *CDKN2A* locus on chromosome 9p21.3 by whole-exome sequencing of the BL-2 cell line. This homozygous loss encompassed the adjacent genes *CDKN2B* and *DMRTA1*. The *MTAP* gene upstream and the *ELAVL2* downstream to the *CDKN2A* locus were not affected by the deletion.



S 10: Read coverage of the chromosomal region chr9:21,795,637-23,842,777 bp (hg19) including the *CDKN2A* locus in BL-2 cell lines as analyzed by whole-exome sequencing. A: Depicted are from the top to the bottom: ideogram of chromosome 9, scale bar in kb, genome base position in bp, the UCSC gene track showing the genes of the region. Highlighted in blue is the region of homozygous loss. B-F: Plots of read coverage of *MTAP* (B), *CDKN2A* (C), *CDKN2B* (C), *DMRTA1* (D) *ELAVL2* (F). The red bars of the Refseq track depict all exons of transcripts with a consensus coding sequence or Refseq ID.

8.9 Immunohistochemical variables in *PCBP1* mutated and wildtype Burkitt lymphomas

In the following table S 11 the immunohistochemical variables of *PCBP1* mutated and wildtype patients are summarized. The data are part of a manuscript which is accepted for publication in the journal *Genes, Chromosomes and Cancer*. The data were provided within the ICGC MMML-Seq as well as MMML project and the differences in categorical variables between groups were analyzed by Dr. Markus Kreuz, Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig using Fisher's exact test. Of note is that immunohistochemistry data were not available of all BL cases.

S 11: Immunohistochemical variables in *PCBP1* mutated and wildtype BL.

	ICGC MMML-Seq cohort				MMML cohort				Combined cohorts				P-value[1]
	mut	(%)	wt	(%)	mut	(%)	wt	(%)	mut	(%)	wt	(%)	
Sex													
F	1	33	1	7	0	0	9	36	1	17	10	26	1
M	2	67	13	93	3	100	16	64	5	83	29	74	
Immunohistochemistry													
CD10													
neg	0	0	0	0	0	0	1	4	0	0	1	3	1
pos	2	100	12	100	3	100	24	96	5	100	36	97	
CD5													
neg	0	0	3	100	3	100	20	83	3	100	23	85	1
pos	0	0	0	0	0	0	4	17	0	0	4	15	
BCL2													
neg	2	67	11	92	3	100	19	83	5	83	30	86	1
pos	1	33	1	8	0	0	4	17	1	17	5	14	
BCL6													
neg	0	0	0	0	1	33	0	0	1	17	0	0	0.15
pos	3	100	10	100	2	67	24	100	5	83	34	100	
IRF4/MUM1													
neg	0	0	7	70	0	0	12	57	0	0	19	61	0.008
pos	3	100	3	30	3	100	9	43	6	100	12	39	
Ki67[2]													
neg	0	0	0	0	0	0	1	4	0	0	1	3	1
pos	3	100	11	100	3	100	24	96	6	100	35	97	

Percentages refer to analyzed cases. Of note is that immunohistochemistry data were not available of all BL cases. M, male; F, female; neg, negative; pos, positive; mut, *PCBP1* mutated; wt, *PCBP1* wildtype

[1] P-value (Fisher's exact test) refers to comparison of mutated vs. wildtype in combined cohorts, p-value < 0.05 was defined as significant.

[2] cut-off ≥90%

8.10 Overview on chromosome 11 aberrations in *MYC*-negative Burkitt-like lymphomas

S 12: Overview on the chromosome 11 aberrations identified in mnBLL cases 18-25.

Case	Aberration	Cytoband	Genomic position (hg19) in bp	Size (kb)
Case 18	High level gain	11q22.3	106,233,094-107,651,609	1,418.5
	CN Gain	11q22.3 - q23.3	107,670,228-117,119,552	9,449.3
	High level gain	11q23.3	117,127,159-118,966,521	1,839.4
	CN Loss	11q23.3 - q25	118,971,251-134,938,847	15,967.6
	CN Gain whole chr.	12p13.33 - q24.33	0-133,851,895	13,385.2
	CN Gain	13q31.3	91,359,736-92,328,701	969.0
	CN Gain whole chr.	19p13.3 - q13.43	0-59,128,983	59,129.0
Case 19	CN Gain	11p15.5 - q24.2	192,764-127,778,007	27,585.2
	CN Loss	11q24.2 - q25	127,799,447-133,280,976	5,481.5
	CN Gain	11q25	133,305,660-134,938,847	1,633.2
	CN Loss	14q31.3	86,454,666-88,383,985	1,929.3
Case 20	CN Gain	8q21.12 - q23.1	80,082,078-107,358,145	27,276.1
	CN Gain	8q23.3 - q24.23	112,501,602-136,517,735	24,016.1
	CN Gain	8q24.3	140,127,523-146,292,734	6,165.2
	High level gain	11q13.2 - q23.3	67,767,160-119,460,505	51,693.4
	CN Loss	11q23.3 - q25	119,488,812-134,938,847	15,450.0
	CN Gain	13q31.1 - q31.2	85,287,625-89,609,906	4,322.3
	High level gain	13q31.2 - q31.3	89,614,856-92,150,929	2,536.1
	CN Loss	13q31.3 - q34	92,165,452-115,103,150	22,937.7
	CN Gain	14q21.3 - q32.33	48,465,864-106,325,013	57,859.2
	High level gain	18q21.2 - q21.32	51,129,489-56,719,436	5,589.9
	CN Gain	18q22.1 - q23	62,979,859-74,637,067	11,657.2
	CN Loss	18q23	74,643,905-78,007,784	3,363.9
	CN Gain whole chr.	20p13 - p11.22	0-21,420,492	21,420.5
Case 21	CN Gain	11q22.1 - q22.3	97,346,172-104,405,081	7,058.9
	CN Gain	11q22.3	104,423,008-105,983,140	1,560.1
	CN Gain	11q22.3 - q23.3	106,005,901-120,907,482	14,901.6
	CN Loss	11q23.3 - q25	120,931,054-134,938,847	14,007.8
	CN Gain whole chr.	12p13.33 - q24.33	0-133,851,895	133,851.9
Case 22	CN Loss	6q12 - q21	67,759,432-110,118,776	42,359.4
	CN Loss	11q24.2 - q25	124,440,617-132,877,670	8,437.1
Case 23	CN Gain	8q22.3 - q24.3	104,078,514-146,292,734	42,214.2
	CN Gain	11q12.1 - q13.3	58,169,996-69,439,860	11,269.9
	CN Gain	11q13.3 - q24.1	69,502,677-122,984,689	53,482.0
	CN Loss	11q24.1 - q25	122,996,050-134,938,847	11,942.8
	CN Gain whole chr.	12p13.33 - q24.33	0-133,851,895	133,851.9
Case 24	CN Gain	1p36.33 – p36	754,192-20,740,016	19,985.8
	CN Loss	3p14.2	60,206,501-60,508,340	301.8
	CN Loss	4q13.2 – q35.2	67,042,255-190,915,650	123,873.4
	CN Gain	7q36.2 – q36.6	154,812,950-155,777,228	964.3
	CN Gain whole chr.	8p23.3 - q24.3	0-146,364,022	146,364.0
	CN Gain	11p15.5 - q23.3	192,764-119,754,895	119,562.1
	CN Loss	11q23.3 - q25	119,771,843-134,938,847	15,167.1
	CN Gain whole chr.	12p13.33 - q24.33	0-133,818,115	133,818.1
	CN Gain	18q11.1 - q21.31	18,554,307-53,950,927	35,396.6
	CN Loss	18q21.31 – q23	53,962,817-78,007,784	24,044.9

Table S 12 continued				
Case	**Aberration**	**Cytoband**	**Genomic position (hg19) in bp**	**Size (kb)**
Case 25	CN Loss	2q14.3	124,730,172-125,054,416	324.3
	CN Loss	4q21.21 - q22.1	80,280,705-91,015,463	10,734.8
	CN Gain	7q21.3 - q36.3	90,789,353-159,118,443	68,329.1
	CN Loss	7q31.1	110,934,682-111,195,204	260.5
	CN Gain	11q14.1 - q23.3	80,282,364-120,687,234	40,404.9
	CN Loss	11q23.3 - q25	120,702,378-134,938,847	14,236.5
	Homozygous Loss	11q24.2 - q24.3	127,621,020-129,572,876	1,951.9
	High level gain	16p13.3	83,887-536,686	452.8

CN, copy number

8.11 Overview on the chromosomal imbalances in the MLMA cell line

S 13: Overview on chromosomal imbalances identified in the MLMA cell line.

Chromosome	CNA	Cytoband	Start in bp (hg19)	End in bp (hg19)	Size (Mb)
2	Loss	p25.3-p25.1	12,772	12,003,659	11.99
	Loss	p22.3-p22.1	35,769,789	38,723,973	2.95
	Gain	p22.1-p16.3	38,727,427	50,813,047	12.09
	High level gain	p16.3	49,631,385	50,813,047	1.18
	High level gain	p16.1-p15	60,828,527	62,003,895	1.18
	Loss	p14-p13.3	68,224,520	69,089,184	0.87
	Loss	q22.3-q24.1	146,877,806	157,412,458	10.54
3	Gain	p26.3-p26.1	60,333	6,778,609	6.72
	Loss	p26.1-p25.3	6,785,721	10,270,528	3.48
	Loss	p25.3-p25.1	10,587,649	14,161,890	3.11
	Loss	p24.3	18,648,481	20,421,820	1.77
	Gain	p24.2-p22.3	26,223,735	33,753,419	7.53
	Loss	p22.3-p22.2	33,763,296	37,652,985	3.89
	Gain	p22.1	39,873,605	41,929,725	2.06
	Gain	p22.1	42,757,852	42,896,787	0.14
	Loss	p22.1-p21.31	42,968,294	44,316,466	1.35
	Loss	p21.31-p12.3	49,358,503	86,823,544	37.47
	Homoz. Loss	p14.3	54,692,206	57,604,882	2.91
	Loss	p14.2-p12.1	59,906,110	86,823,544	26.92
	Gain	p12.1	86,917,763	87,124,205	0.21
	Gain	p11.2	87,805,170	94,262,650	6.46
	High level gain	p11.2	94,262,653	94,660,166	0.4
	Gain	q11.2	97,068,564	98,801,390	1.73
	High level gain	q12.1	98,801,400	99,799,699	0.99
	Loss	q12.2-q13.32	99,799,700	118,260,457	18.46
	Gain	q25.32-q25.32	158,136,848	158,646,181	0.51
4	Loss	p16.3-p11	12,269	49,089,362	49.08
	Gain	p11-q13.1	52,344,170	60,132,170	7.79
	Gain	q24-q33	102,831,285	170,778,068	67.95
	Gain	q34.1-q35.2	172,956,807	19,1020,138	18.06
5	Loss	q12.1-q12.1	59,668,386	59,899,235	0.23
	Loss	q21.3-q23.2	105,071,875	126,622,890	21.55
	Gain	q23.2-q31.1	126,627,875	133,753,600	0.71
	High level gain	q31.1	133,753,689	135,164,389	1.41
7	High level gain	q21.11-q22.1	82,851,692	102,102,164	19.25
	Gain	q22.1-q31.1	102,340,594	110,224,593	7.88
	Loss	q31.1	110,225,748	110,445,011	0.22
	Gain	q31.1-q31.33	110,448,308	124,881,515	14.43
	Gain	q31.33	125,876,948	126,485,276	0.61
	Loss	q31.33	126,493,232	126,799,388	0.31
	Gain	q31.33-q33	126,801,929	137,101,515	10.30

Table S13 continued					
Chromosome	**CNA**	**Cytoband**	**Start in bp (hg19)**	**End in bp (hg19)**	**Size (Mb)**
7	Loss	q33	137,106,596	137,336,572	0.23
	Gain	q33-q34	137,345,877	140,060,653	2.71
	Loss	q34	140,065,886	140,324,890	0.26
	Gain	q34	140,330,182	140,813,472	0.48
	Loss	q34-q36.3	140,815,840	159,119,708	18.30
8	Loss	p23.3-q21.13	1	83,352,762	83.35
	Homoz. loss	q21.13-q21.2	83,389,273	86,565,730	3.18
	Loss	q21.2-q21.3	86,821,307	89,304,055	2.48
	Homoz. loss	q21.3	89,322,596	89,430,301	0.11
	Loss	q21.3	89,433,842	90,064,236	0.63
10	Gain	q26.3	132,072,271	135,506,692	3.43
11	Gain	q21	93,756,820	96,723,470	2.97
	High level gain	q21-q22.1	96,725,882	99,519,286	2.79
	Gain	q22.1	99,528,055	99,791,125	0.26
	High level gain	q22.1	99,799,894	100,545,859	0.75
	Gain	q22.1-q23.3	100,554,028	117,513,420	16.96
	High level gain	q23.3	117,530,600	118,962,816	1.43
	Loss	q23.3-q25	118,970,202	134,944,770	15.98
12	High level gain	q13.13-q15	54,893,027	69,935,275	15.04
13	Loss	q14.11	41,719,828	42,878,121	1.12
15	Gain	q11.2-q13.1	24,886,177	29,813,757	4.93
	High level gain	q13.1-q14	29,813,760	34,763,616	4.95
	Gain	q22.2-q22.31	59,533,114	66,252,621	6.72
	High level gain	q22.31-q25.2	66,252,625	83,649,934	17.4
	Loss	q26.1-q26.2	92,589,340	96,373,328	3.78
17	Loss	p13.3-p11.2	514	21,547,725	21.55
18	High level gain	q21.2-q21.33	49878,620	61,386,375	4.72
	Loss	q21.33-q23	61,388,177	78,015,057	16.63
	Homoz. Loss	q22.1-q23	62,941,778	74,502,552	11.56
19	Loss	p13.2	10,151,244	10,586,984	0.44
	Loss	q13.42-q13.43	56,124,049	59,097,842	2.97
22	Gain	q13.2-q13.33	42,025,057	51,188,494	9.16
X	Gain	p22.33-q11.2	1	63,921,440	63.92
	Loss	q11.2-q21.31	63,862,549	86,991,835	23.13
	Gain	q21.31-q25	87,103,000	125,431,093	38.33
	High level gain	q25-q28	125,419,892	154,929,486	29.51

Homoz. Loss: homozygous loss; start and end of copy number alteration in bp; size in mega bases (Mb).

8.12 Protein coding genes in the minimal regions of gain and loss on chromosome 11

The following table S 14 enlists all genes which map to the minimal region of gain or amplification, as well as minimal region of loss or homozygous loss on chromosome 11 as identified in mnBLL.

S 14: Complete list of genes mapping to the minimal regions of alteration on chromosome 11 in mnBLL

Gene	Full Name	Region Chr11 (hg19) in bp	Minimal region of
CADM1	cell adhesion molecule 1	chr11:115,044,345-115,375,241	Gain
BUD13	*BUD13* homolog (S. cerevisiae)	chr11:116,618,886-116,643,714	Gain
ZNF259	zinc finger protein 259	chr11:116,649,276-116,658,739	Gain
APOA5	apolipoprotein A-V	chr11:116,660,086-116,663,136	Gain
APOA4	apolipoprotein A-IV	chr11:116,691,418-116,694,011	Gain
APOC3	apolipoprotein C-III	chr11:116,700,624-116,703,787	Gain
APOA1	apolipoprotein A-I	chr11:116,706,469-116,708,338	Gain
SIK3	SIK family kinase 3	chr11:116,714,118-116,968,993	Gain
PAFAH1B2	platelet-activating factor acetylhydrolase 1b, catalytic subunit 2	chr11:117,015,000-117,041,761	Gain
SIDT2	SID1 transmembrane family, member 2	chr11:117,049,939-117,068,161	Gain
TAGLN	transgelin	chr11:117,070,040-117,075,508	Gain
PCSK7	proprotein convertase subtilisin/kexin type 7	chr11:117,075,788-117,102,811	Gain
CEP164	centrosomal protein 164kDa	chr11:117,198,571-117,283,982	Gain
RNF214	ring finger protein 214	chr11:117,103,404-117,156,404	Gain
BACE1	beta-site APP-cleaving enzyme 1	chr11:117,156,402-117,186,972	Gain
DSCAML1	Down syndrome cell adhesion molecule like 1	chr11:117,298,489-117,667,976	Amplification
FXYD2	Sodium/potassium-transporting ATPase subunit gamma	chr11:117,690,790-117,695,459	Amplification
FXYD6	FXYD domain-containing ion transport regulator 6 precursor	chr11:117,707,691-117,747,746	Amplification
TMPRSS13	transmembrane protease, serine 13	chr11:117,771,356-117,800,168	Amplification
IL10RA	interleukin 10 receptor, alpha	chr11:117,857,106-117,872,199	Amplification
TMPRSS4	Transmembrane protease serine 4	chr11:117,947,727-117,990,556	Amplification
SCN4B	Sodium channel subunit beta-4 precursor	chr11:118,004,092-118,023,630	Amplification
SCN2B	Sodium channel subunit beta-2 precursor	chr11:118,033,519-118,047,337	Amplification
AMICA1	adhesion molecule, interacts with CXADR antigen 1	chr11:118,064,442-118,095,809	Amplification
MPZL2	myelin protein zero-like 2	chr11:118,124,131-118,135,251	Amplification
MPZL3	Myelin protein zero-like protein 3 precursor	chr11:118,100,336-118,123,011	Amplification
CD3G	CD3g molecule, gamma (CD3-TCR complex)	chr11:118,215,059-118,224,497	Amplification
CD3E	CD3e molecule, epsilon (CD3-TCR complex)	chr11:118,175,295-118,186,890	Amplification
CD3D	CD3d molecule, delta (CD3-TCR complex)	chr11:118,209,789-118,213,459	Amplification
ATP5L	ATP Synthase, H+ Transporting, Mitochondrial Fo Complex, Subunit G	chr11:118,272,104-118,280,562	Amplification
UBE4A	ubiquitination factor E4A	chr11:118,230,296-118,269,926	Amplification
KMT2A	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila)	chr11:118,307,205-118,397,539	Amplification
TTC36	tetratricopeptide repeat domain 36	chr11:118,398,210-118,401,740	Amplification
THEM25	Homo sapiens transmembrane protein 25	chr11:118,401,803-118,406,550	Amplification
ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1 (avian)	chr11:128,328,656-128,457,453	Homozygous loss
FLI1	Friend leukemia virus integration 1	chr11:128,563,811-128,683,162	Homozygous loss

Table S 14 continued			
Gene	**Full Name**	**Region Chr11 (hg19) in bp**	**Minimal region of**
KCNJ1	potassium inwardly-rectifying channel, subfamily J, member 1	chr11:128,707,915-128,712,363	Homozygous loss
C11orf45	chromosome 11 open reading frame 45	chr11:128,769,460-128,775,964	Homozygous loss
KCNJ5	potassium inwardly-rectifying channel, subfamily J, member 5	chr11:128,761,313-128,787,951	Homozygous loss
TP53AIP1	tumor protein p53 regulated apoptosis inducing protein 1	chr11:128,769,460-128,775,592	Homozygous loss
ARHGAP32	Rho GTPase activating protein 32	chr11:128,769,460-128,775,592	Homozygous loss
BARX2	BARX homeobox 2	chr11:129,245,881-129,322,174	Homozygous loss
TMEM45B	transmembrane protein 45B	chr11:129,685,741-129,729,898	Loss
NFRKB	nuclear factor related to kappaB binding protein	chr11:129,733,670-129,762,904	Loss
PRDM10	PR domain containing 10	chr11:129,769,601-129,872,730	Loss
APLP2	amyloid beta (A4) precursor-like protein 2	chr11:129,940,463-130,014,706	Loss
ST14	suppression of tumorigenicity 14 (colon carcinoma)	chr11:130,029,682-130,080,257	Loss
ZBTB44	zinc finger and BTB domain containing 44	chr11:130,029,682-130,080,257	Loss
ADAMTS8	ADAM metallopeptidase with thrombospondin type 1 motif, 8	chr11:130,274,818-130,298,539	Loss
ADAMTS15	ADAM metallopeptidase with thrombospondin type 1 motif, 15	chr11:130,318,869-130,346,539	Loss
SNX19	sorting nexin 19	chr11:130,745,766-130,786,382	Loss
NTM	neurotrimin	chr11:131,780,712-132,206,716	Loss
OPCML	opioid binding protein/cell adhesion molecule-like	chr11:132,284,875-132,813,037	Loss

8.13 Non-coding RNAs in the minimal regions of gain and loss on chromosome11

Following table S 15 enlists all genes of non-coding RNAs mapping to the minimal region of gain or amplification, as well as minimal region of loss or homozygous loss on chromosome 11 in mnBLL.

S 15: Complete list of non-coding RNAs mapping to the minimal regions of alteration on chromosome 11 in mnBLL.

Gene	Full Name	Region chr11 (hg19) in bp	Minimal region of
LOC283143	uncharacterized LOC283143	chr11:115,626,051-115,630,918	Gain
mir-652		chr11:116,349,824-116,349,914	Gain
mir-548		chr11:116,822,173-116,822,256	Gain
LOC100652768	uncharacterized LOC100652768	chr11:117,066,329-117,072,630	Gain
piR-43793		chr11:117,158,614-117,158,647	Gain
LOC100526771	uncharacterized LOC100526771	chr11:117,886,487-117,957,508	Amplification
FLI-AS1	FLI1 antisense transcript	chr11:128,561,567-128,565,918	Homozygous loss
LINC00167	long intergenic non-protein coding RNA 167	chr11:129,872,519-129,875,381	Loss

8.14 List of recurrently mutated genes in *MYC*-negative Burkitt-like lymphoma

Combination of whole-genome data of mnBLL case 1 with whole-exome data of three mnBLL cell lines (HT, Su-DHL-5, MLMA) led to the identification of 63 recurrently mutated genes (mutated in at least two analyzed samples) which are enlisted in S 16. Three additional mnBLL cases (cases 2, 3 and 8) were analyzed using the TruSight™ One sequencing panel which covers 22 of the 63 genes. None of those 22 genes was found to be mutated in those three mnBLL. Highlighted in bold letters are those genes which were mutated in mnBLL case 1.

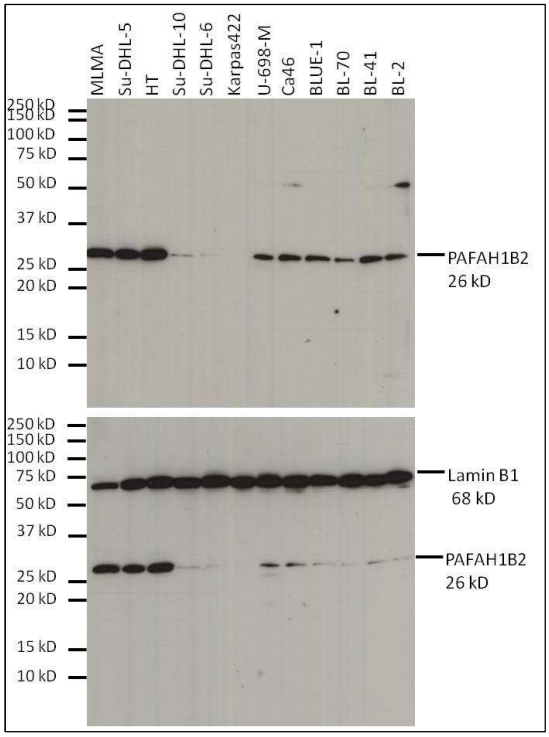
S 16: Overview on genes recurrently mutated in mnBLL and mnBLL cell lines.

Gene	Chromosome	TruSight™ Panel	Case 1	HT	Su-DHL-5	MLMA
ABP1	7			x		x
ADAMTS16	5	x		x	x	
ADCY6	12	x		x		x
AHNAK2	14			x	x	x
ALG10	12				x	x
ANGPTL2	9			x	x	
ARHGEF1	19			x	x	
ARID1B	6	x		x		x
ARMC3	10			x		x
ATRNL1	10	x		x	x	
BIRC3	11			x	x	
CD36	7	x			x	x
CERK	22				x	x
CLTCL1	22	x	x	x		
COL4A1	13	x		x		x
DDX3X	X			x	x	
DDX41	5			x	x	
DDX60	4		x			x
DLAT	11	x		x	x	
DPY19L2	12	x		x	x	
DYNC1H1	14	x		x		x
EIF4A1	17			x	x	
FAM186A	12			x	x	
FAM83H	8	x		x		x
FAT1	4			x	x	
FBN2	5	x		x		x
FBXL13	7			x		x
FBXO5	6			x		x
FHL5	6		x		x	
GNA13	17		x		x	
GNAS	20	x		x	x	
ITGA2	5	x		x		x
ITGAX	16			x	x	
KIAA1045	9			x	x	
KIAA1755	20			x		x
KRT3	12	x		x		x
LAMA3	18	x		x		x
TEKT4P2	21			x	x	
MAPK4	18			x	x	
MKI67	10			x		x
MYH14	19	x		x		x
NLRC5	16			x	x	
NOTCH3	19	x		x	x	
NOTUM	17			x		x
NUP188	9			x	x	
OR7G3	19			x		x

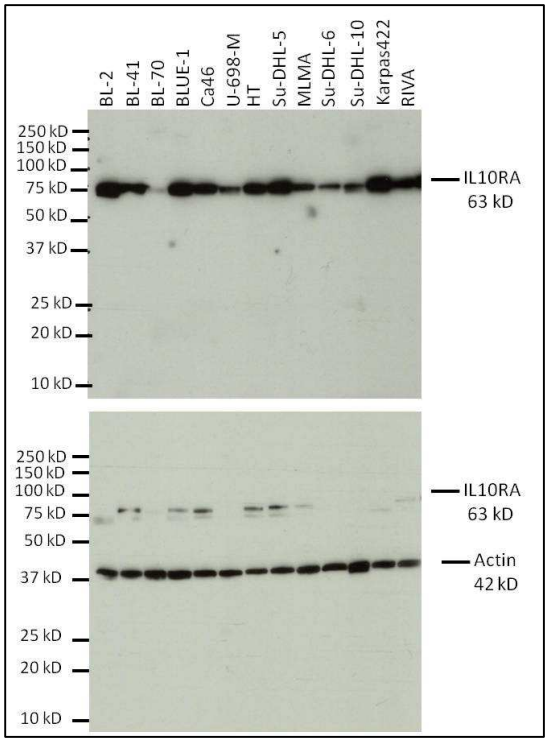
Table S 16 continued						
Gene	**Chromosome**	**TruSight™ Panel**	**Case 1**	**HT**	**Su-DHL-5**	**MLMA**
PABPC1	8			x	x	
PAPLN	14			x		x
PCDHB7	5			x	x	
PTX4	16			x	x	
RBFOX3	17			x		x
RPL35	9			x	x	
SENP3	17			x	x	
SGCD	5	x		x	x	
SPTY2D1	11			x		x
ST5	11	x		x		x
SVEP1	9			x		x
SYNRG	17			x	x	
SYT4	18			x	x	
TNXB	6	x		x		x
TP53	17	x		x		x
ULK2	17			x	x	
ZNF208	19			x	x	

8.15 Complete Western blot figures

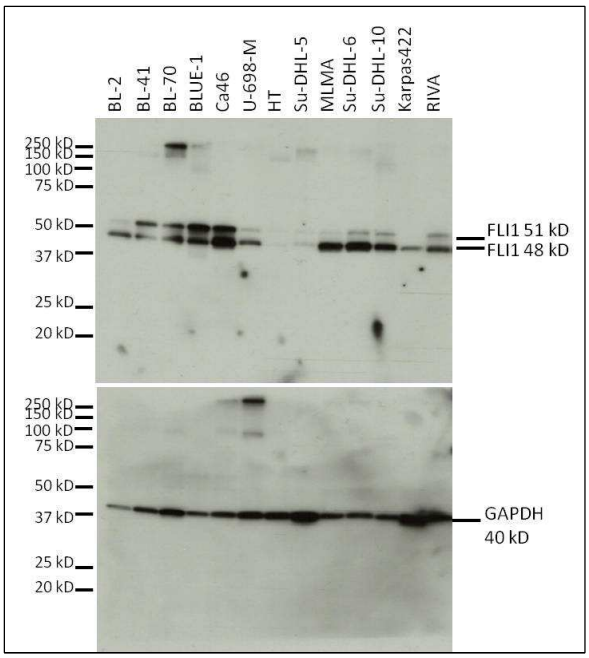
Following Western blot figures show the complete membrane.



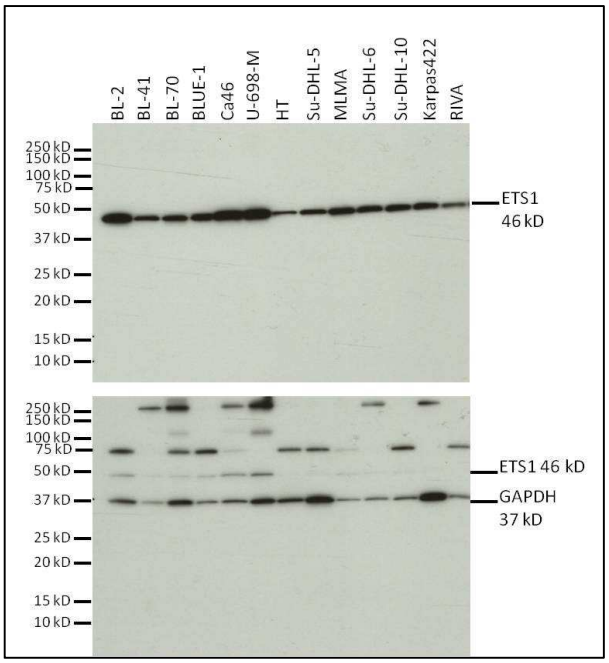
S 17: Western blot of PAFAH1B2.



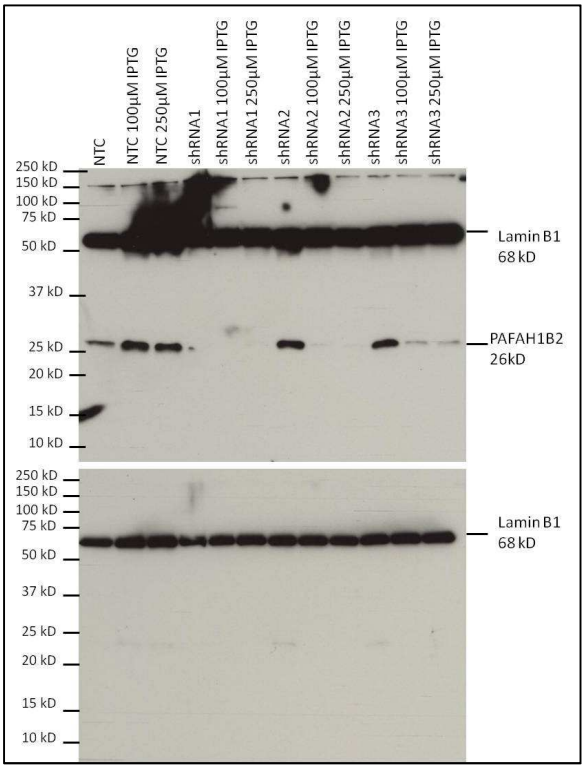
S 18: Western blot of IL10RA.



S 19: Western blot of FLI1.



S 20: Western blot of ETS1.



S 21: Control of knock down of PAFAH1Bs expression in Su-DHL-5 cell line using Western blot.

9 Danksagung

Ich möchte mich herzlichste bei allen bedanken, die mich während meiner Doktorarbeit unterstützt haben:

Bei Frau Prof. Dr. rer. nat. Manuela Dittmar möchte ich mich herzlichst für die Übernahme der Betreuung und des Hauptreferats meiner Doktorarbeit bedanken.
Ein ganz besonderer Dank gilt Prof. Dr. med. Reiner Siebert für die Bereitstellung des interessanten Dissertationsthemas, die Betreuung und die Unterstützung während der Anfertigung dieser Arbeit. Danke vor allem für die immerwährende Diskussionsbereitschaft und für die kritische Durchsicht dieser Arbeit.

Ich danke der KinderKrebsInitiative Buchholz/Holm-Seppensen für das entgegengebrachte Vertrauen und für die Möglichkeit Forschungsarbeiten im Rahmen des Christoph-Schubert-Stipendiums durchzuführen.

Mein Dank geht an alle Kooperationspartnern vor allem Prof. Dr. med. Wolfram Klapper, Dr. rer. nat. Monika Sczepanowski und Charlotte Botz-von Drathen für die Bereitstellung von Tumormaterialien; an Dr. med. Christine Damm-Welk und Prof. Dr. med. Brigitte Burkhardt für die Bereitstellung der Serumproben; an Dr. med. Denis Schewe für die Bereitstellung des FACS Geräts und die Möglichkeit die S2 Räumlichkeiten zu nutzen; an Prof. Dr. rer. nat. Sabine Adam für die Einweisung in die FACS Analysen; an M.Sc. Michael Lenz für die Unterstützung bei dem PluriTest und an das Team der MMML und ICGC MMML-Seq Projekte, vor allem hier an Dr. rer. nat Matthias Schlesner, Dr. rer. nat. Stephan H. Bernhart, M.Sc. Christian Kohler, Dr. rer. nat. Maciej Rosolowski und Dr. rer. nat. Markus Kreuz für die Bereitstellung von Daten und Unterstützung bei bioinformatischen Fragestellungen.

Des Weiteren möchte ich mich bei dem gesamten Team des Institutes für Humangenetik bedanken, insbesondere bei Prof. Dr. rer. nat. Ole Ammerpohl, Dr. rer. nat. Ingram Iaccarino und Dr. rer. nat. Julia Richter für ihre stets offene Tür und Unterstützung bei Fragen,
bei Dr. rer. nat. Andrea Haake, Dr. med Susanne Bens, Dr. rer. nat. Itziar Salaverria, Dr. rer. nat. Inga Nagel, Dr.med. Anke Bergmann, Dr. rer. nat. Eva M. Murga-Penas, und Dr. rer. nat. Inga Vater für die konstruktive und gute Zusammenarbeit,
bei dem Team der Molekulargenetik, Claudia Becher, Dorit Schuster, Gabi Riesen, Lorena Valles, Ute Jacobsen, für Ihre Hilfsbereitschaft, fachliche Unterstützung bei der Durchführung der experimentellen Arbeiten und für die harmonische Zusammenarbeit,
bei meinen Doktorandenmitstreitern Sietse M. Aukema, Ulrike Paul, Shaymaa El Gaafary und besonders Kirstin Hoff und Julia Kolarova für die Unterstützung, konstruktiven Gespräche und Zusammenhalt auch weit über die Arbeit hinaus.

10 Curriculum vitae

Persönliche Daten

Name	Rabea Wagener
Akademischer Grad	Diplom Molekulare Biomedizinerin (Dipl. Mol. Biomed.)
Anschrift	Weißenburgstraße 4
	24116 Kiel
	Deutschland
Geburtstag	27.05.1987
Geburtsort	Attendorn, Deutschland
Staatsangehörigkeit	deutsch
Familienstand	ledig
Email	rabeawagener@gmx.de

Schulausbildung und beruflicher Werdegang

1993-1997	Franziskus Grundschule, Lennestadt-Meggen
1997-2006	Gymnasium der Stadt Lennestadt, Lennestadt-Altenhundem
Juni 2006	Erlangung der Allgemeinen Hochschulreife
Oktober 2006-August 2011	Studium der Molekularen Biomedizin an der Rheinischen Friedrich-Wilhelms-Universität zu Bonn, Abschluss Diplom Molekulare Biomedizin
Oktober 2010- August 2011	Diplomarbeit am Institut für Humangenetik der Rheinischen Friedrich-Wilhelms-Universität zu Bonn, Titel: „Untersuchung des tumorigenen Potentials von *CHD1L* in Glioblastomen“
September 2011	Studentische Hilfskraft am Institut für Humangenetik der Rheinischen Friedrich-Wilhelms-Universität zu Bonn
seit Oktober 2011	Wissenschaftliche Angestellte am Institut für Humangenetik der Christian-Albrechts-Universität zu Kiel mit Zielrichtung Promotion

11 Publikationen

Originalarbeiten die Teil der vorliegenden Arbeit sind

Wagener R[1], Aukema SM[1], Schlesner M, Haake A, Burkhardt B, Claviez A, Drexler HG, Hummel M, Kreuz M, Loeffler M, Rosolowski M, López C, Möller P, Richter J, Rohde M, Betts MJ, Russell RB, Bernhart SH, Hoffmann S, Rosenstiel P, Schilhabel M, Szczepanowski M, Trümper L, Klapper W, Siebert R, on behalf of the ICGC MMML-Seq-Project and the "Molecular Mechanisms in Malignant Lymphomas" Network Project of the Deutsche Krebshilfe, The PCBP1 gene encoding Poly(rC) binding protein 1 is recurrently mutated in Burkitt lymphoma, *Genes, Chromosomes and Cancer*, zur Publikation akzeptiert, 2015
[1] shared first authorship

Wagener R, Lenz M, Schuldt B, Lenz I, Schuppert A, Siebert R, Müller FJ, Investigation of potential traces of pluripotency in germinal-center derived B-cell lymphomas driven by MYC., *Blood Cancer Journal*, zur Publikation akzeptiert 2015

Salaverria I[1], Martin-Guerrero I[1], **Wagener R**[1], Kreuz M[1], Kohler CK[1], Richter J., Pienkowska-Grela B., Adam P, Burkhardt B, Claviez A, Damm-Welk C, Drexler HG, Hummel M, Jaffe ES, Küppers R, Lefebvre C, Lisfeld J, Löffler M, Macleod RAF, Nagel I, Oschlies I, Rosolowski M, Russell R.B., Rymkiewicz G., Schindler D, Schlesner M., Scholtysik R, Schwaenen C, Spang R, Szczepanowski M, Trümper L, Vater I, Wessendorf S, Klapper W, and Siebert R, for the Molecular Mechanisms in Malignant Lymphoma Network Project and Berlin-Frankfurt-Münster Non-Hodgkin Lymphoma Group, A recurrent 11q aberration pattern characterizes a subset of MYC-negative high-grade B-cell lymphomas resembling Burkitt lymphoma, *Blood*. 2014; 123(8):1187-98
[1] shared first authorship

Richter J, Schlesner M, Hoffmann S, Kreuz M, Leich E, Burkhardt B, Rosolowski M, Ammerpohl O, **Wagener R**, Bernhart SH, Lenze D, Szczepanowski M, Paulsen M, Lipinski S, Russell RB, Adam-Klages S, Apic G, Claviez A, Hasenclever D, Hovestadt V, Hornig N, Korbel JO, Kube D, Langenberger D, Lawerenz C, Lisfeld J, Meyer K, Picelli S, Pischimarov J, Radlwimmer B, Rausch T, Rohde M, Schilhabel M, Scholtysik R, Spang R, Trautmann H, Zenz T, Borkhardt A, Drexler HG, Moller P, Macleod RA, Pott C, Schreiber S, Trumper L, Loeffler M, Stadler PF, Lichter P, Eils R, Kuppers R, Hummel M, Klapper W, Rosenstiel P, Rosenwald A, Brors B, Siebert R, Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing, *Nat Genet*. 2012; 44(12):1316-20

Originalarbeiten die nicht Teil der vorliegenden Arbeit sind

Wagener R, Alexandrov LB, Montesinos-Rogen M, Schlesner M, Haake A, Drexler HG, Richter J, Bignell GR, McDermott U, Siebert R, Analysis of mutational signatures in exomes from B-cell lymphoma cell lines suggest APOBEC3 family members to be involved in the pathogenesis of Primary Effusion Lymphoma, *Leukemia*. 2015; doi: 10.1038/leu.2015.22

Bens S, Zichner T, Stütz AM, Caliebe A, **Wagener R**, Hoff K, Korbel JO, von Bismarck P, Siebert R, SPAG7 is a candidate gene for the periodic fever, aphthous stomatitis, pharyngitis and adenopathy (PFAPA) syndrome., *Genes Immun*. 2014; 15(3):190-4

Veröffentlichte Tagungsbeiträge

Wagener R[1], Aukema SM[1], Schlesner M, Haake A, Burkhardt B, Drexler HG, Hummel M, Kreuz M, Löffler M, López C, Möller P, Richter J, Russell RB, Szczepanowski M, Trümper L, Klapper W, Siebert R, on behalf of the German ICGC MMML-Seq-Project and the MMML-Study Group, The PCBP1 gene encoding Poly(rC) binding protein 1 is recurrently mutated in Burkitt lymphoma, 1. International Conference on New Concepts in B Cell Malignancies: From molecular pathogenesis to personalized treatment, 14.-16. November 2014, Thessaloniki, Griechenland
[1] shared first authorship

Wagener R, Richter J, Burkhardt B, Schlesner M, Damm-Welk C, Woessmann W, Vater I, Siebert R, Mutational landscape of ALK-positive diffuse large B-cell lymphoma, 25. Jahrestagung der deutschen Gesellschaft für Humangenetik, 19.-21. März 2014, Essen, Deutschland

Bens S, Zichner T, Stütz AM, Caliebe A, **Wagener R**, Hoff K, Korbel JO, von Bismarck P, Siebert R, SPAG7 is a candidate gene for the periodic fever, aphthous stomatitis, pharyngitis, and adenopathy (PFAPA) syndrome, 25. Jahrestagung der deutschen Gesellschaft für Humangenetik, 19.-21. März 2014, Essen, Deutschland

Iaccarino I, **Wagener R**, Doose G, Hoffmann S, Höll J, Borkhardt A, Siebert R, Identification of MYC-regulated long intergenic non-coding RNAs, The Non-Coding Genome EMBO/EMBL Symposium, 9.-12. Oktober 2013, Heidelberg, Deutschland

Wagener R, Salaverria I, Martin Guerrero I, Iaccarino I, Ammerpohl O, Drexler HG, MacLeod RAF, Siebert R, Deutsche Krebshilfe Network “Molecular Mechanisms in Malignant Lymphomas”, Expression analyses of genes within a recurrent chromosome 11q aberration in a subset of MYC-negative mature aggressive B-cell lymphomas with features of Burkitt lymphoma, 6. Mildred Scheel Cancer Conference, 5.-.7. Juni 2013, Königswinter, Deutschland

Wagener R, Ammerpohl O, Kolarova J, Gutwein J, Murga Penas EM, Richter J, Drexler HG, MacLeod RAF, Siebert R, Significant differences in the DNA-methylation profile between EBV-positive and EBV-negative Burkitt Lymphoma cell lines, 24. Jahrestagung der deutschen Gesellschaft für Humangenetik, 20.-22.3 März 2013, Dresden, Deutschland

Wagener R, Ammerpohl O, Kolarova J, Gutwein J, Murga Penas EM, Richter J, Drexler HG, MacLeod RAF, Siebert R, EBV status affects differential DNA-methylation of Burkitt Lymphoma cell lines, 23. Jahrestagung der Gesellschaft für Virologie (GfV), 6.-9. März 2013, Kiel, Deutschland

12 Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation eigenständig und nur mit den von mir angegebenen Quellen und Hilfsmitteln bzw. mit der wissenschaftlichen Beratung meiner akademischen Betreuer angefertigt habe. Die vorliegende Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden und wurde bisher in gleicher oder ähnlicher Form an keiner anderen Universität zur Begutachtung im Rahmen eines Prüfungsverfahrens vorgelegt. Weiterhin bestätige ich, dass ich zuvor noch keine Promotionsversuche unternommen habe.

Folgende Teile der vorliegenden Arbeit wurden in Form eines Artikels in einer internationalen Fachzeitschrift veröffentlicht bzw. zur Publikation akzeptiert oder auf einem Kongress präsentiert:

Wagener R, Aukema SM, Schlesner M, Haake A, Burkhardt B, Claviez A, *et al.*, The PCBP1 gene encoding Poly(rC) binding protein 1 is recurrently mutated in Burkitt lymphoma, *Genes, Chromosomes and Cancer*, zur Publikation akzeptiert, 2015

Wagener R, Lenz M, Schuldt B, Lenz I, Schuppert A, Siebert R, Müller FJ, Investigation of potential traces of pluripotency in germinal-center derived B-cell lymphomas driven by MYC., *Blood Cancer Journal*, zur Publikation akzeptiert, 2015

Salaverria I, Martin-Guerrero I, Wagener R, Kreuz M, Kohler CW, Richter J, *et al.* A recurrent 11q aberration pattern characterizes a subset of MYC-negative high-grade B-cell lymphomas resembling Burkitt lymphoma. *Blood*. 2014;123: 1187–1198.

Richter J, Schlesner M, Hoffmann S, Kreuz M, Leich E, Burkhardt B *et al.* Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing, *Nat Genet*. 2012; 44(12):1316-20

Wagener R, Aukema SM, Schlesner M, Haake A, Burkhardt B, Drexler HG, Hummel M, Kreuz M, Löffler M, López C, Möller P, Richter J, Russell RB, Szczepanowski M, Trümper L, Klapper W, Siebert R, on behalf of the German ICGC MMML-Seq-Project and the MMML-Study Group, The *PCBP1* gene encoding Poly(rC) binding protein 1 is recurrently mutated in Burkitt lymphoma, 1. International Conference on New Concepts in B Cell Malignancies: From molecular pathogenesis to personalized treatment, 14.-16. November 2014, Thessaloniki, Griechenland

Wagener R, Salaverria I, Martin Guerrero I, Iaccarino I, Ammerpohl O, Drexler HG, MacLeod RAF, Siebert R, Deutsche Krebshilfe Network “Molecular Mechanisms in Malignant Lymphomas”, Expression analyses of genes within a recurrent chromosome 11q aberration in a subset of MYC-negative mature aggressive B-cell lymphomas with features of Burkitt lymphoma, 6. Mildred Scheel Cancer Conference, 5.-.7. Juni 2013, Königswinter, Deutschland

Kiel, den …….

Rabea Wagener