The role of the transcription factor MYC in lymphomas

„Die Rolle des Transkriptionsfaktors MYC in Lymphomen“

DISSERTATION
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)
submitted to the
Department of Biology, Chemistry and Pharmacy
of Freie Universität Berlin

by
Karsten Kleo
from Cottbus
Student ID number: 4947549
February 2019
This dissertation was prepared
at the

University Medicine Berlin
Institute of Pathology
Department of Experimental Hematopathology
under the direction of
Prof. Dr. Michael Hummel

from July 2013 till February 2019

1st reviewer: Prof. Dr. Michael Hummel (Charité)
2nd reviewer: Prof. Dr. Rupert Mutzel (FU)

Day of thesis defend: 24.09.2019
1. Abstract

The transcription factor MYC, encoded by the MYC gene, plays a central role in many cellular processes, such as cell growth, apoptosis and cell communication. However, as an oncogene, MYC also plays a central role in initiation and progression of many different types of cancers, including malignant lymphomas, and is therefore the focus of many oncological studies.

Malignant lymphomas do not refer to a single disease entity, but describe a broad range of lymphatic neoplasias that derive from mature lymphoid cells. They can be subclassified into over 60 subtypes based on their differentiation, morphology and/or clinical course. Based on histology, malignant lymphomas can be generally distinguished into Hodgkin's lymphoma (HL) and Non-Hodgkin's lymphoma (NHL). NHL can be further sub-grouped according to their cell of origin into B- and T-cell NHL. Further subclassification exists based on additional histological, clinical and molecular criteria including chromosomal alterations and gene expression profiles. The present work deals with Burkitt's lymphoma (BL) and diffuse large B-cell lymphoma (DLBCL), which both belong to aggressive high-grade B-cell lymphoma. Here, we explore the role of the transcription factor MYC in the pathogenesis and clinical course of these lymphoma types.

One hallmark of BL is the t(8;14) chromosomal translocation, leading to an overexpression of MYC protein. MYC translocations are however, not restricted to BL, but can also occur in DLBCL, although at a much lower frequency. MYC break-positive BL differ significantly from the MYC break-positive DLBCL in their clinical course, with BL associated with a complete cure of the majority of cases.

To investigate the molecular differences between these two lymphoma subtypes, we performed a metabolic and proteomic study, identifying pyruvate as one of the discriminatory metabolites. This metabolic phenotype was further confirmed by proteomic studies of pyruvate metabolism-associated proteins (Schwarzfischer, et al., 2017). In a second study, the genome-wide MYC binding pattern of BL, MYC break-positive and -negative DLBCL was analyzed by chromatin immunoprecipitation followed by next generation sequencing (ChIP-Seq) and RNA-based next generation sequencing (RNA-Seq). Significant differences in the MYC DNA-binding patterns were identified, which were also mirrored in the different gene expression patterns. One of these differentially expressed genes code for the cell surface receptor CD97.
(ADGRE5), which is significantly over-expressed in BL, but absent in MYC break-positive and -negative DLBCL.

This finding was confirmed by independent validation experiments, including immunohistological staining of cell lines and primary patient samples (Kleo et al 2018 – submitted).

This doctoral thesis was complemented by an investigation of long noncoding RNAs (lncRNAs) and their role in modulating the MYC-driven cellular transcriptome. Using NGS, we identified 13 lncRNAs, which were differentially expressed between BL and DLBCLs, one of which was strongly regulated by MYC. This lncRNA was able to modulate MYC-induced cell cycle genes with a strong impact on cell cycle progression. We therefore called this lncRNA MINCR (MYC-induced non-coding RNA) (Doose, et al., 2015).

Taken together, this thesis provides additional evidence that MYC is not merely an on/off amplifier of gene activity but exerts specific actions on the gene expression program and – as a consequence – on cellular functions, a finding also true for aggressive lymphoma. Based on the MYC differences between BL and DLBCL, the identification of biomarkers for their distinction appears to be possible. Therefore it is justified to conclude that MYC plays an essential but diverse role in the pathogenesis of various lymphoma types, a finding which might be important for future treatment modalities.

The analyses of this work were partly conducted in cooperation with other research groups and led to three publications, which provide the scientific basis for this cumulative thesis.
2. Zusammenfassung


Um die molekularen Unterschiede zwischen diesen Lymphom-Subtypen zu untersuchen, führten wir zuerst eine metabolische Studie durch, in der Pyruvat als einer der diskriminierenden Metaboliten identifiziert wurde. Dieser metabolische Phänotyp wurde durch proteomische Untersuchungen der Pyruvatstoffwechselassozierten Proteine bestätigt (Schwarzfischer, et al., 2017). In einem zweiten Ansatz wurde das genomweite MYC-DNA-Bindungsmuster von MYC Bruch-positiven BL und

In einer weiteren Studie untersuchten wir die Rolle von langen nichtkodierenden RNAs (lncRNAs) und deren Einfluss auf das MYC-gesteuerte zelluläre Transkriptom. Unter den 13 differentiell exprimierten lncRNAs wurde eines stark durch MYC reguliert. Es ist in der Lage, MYC-induzierte Zellzyklusgene zu modulieren, was zu einer Kettenreaktion und letztlich zur Zellzyklusprogression führt. Daher nannten wir diese lncRNA MINCR (MYC-induzierte nicht-kodierende RNA) (Doose, et al., 2015).


Die Analysen dieser Arbeiten wurden in Kooperation mit anderen Forschungsgruppen durchgeführt und führten zu drei Publikationen, die für diese kumulative Arbeit verwendet werden.
3. Introduction

3.1 Lymphatic system

The lymphatic system is not a single organ, but rather distributed over the entire body with its functions and cells. It is closely connected to the body's defense system (immune system) and the blood-forming system in bone marrow (Milan, 1957, Olszewski, 1986, Padera, et al., 2016, Till, 1981). The lymphatic system consists of all lymphatics and the lymphatic organs, including the lymph nodes, spleen, lymphatic tissues in the gastrointestinal tract (e.g. the Peyer's plaques of the small intestine), the pharynx (tongues and palate almonds) as well as the thymus gland. The lymph nodes are small bean-shaped organs found in many parts of the body such as in the armpit, groin, pelvis, neck and the abdomen (Mahlke, et al., 2005, Suy, et al., 2016). They serve as filter stations for the lymph fluid, and produce and host cells that fight infections in the body. Via the lymph vessels, waste materials and disease promoters, e.g. bacteria or foreign body particles, are transported to the approx. 500-1000 lymph nodes (Grundmann, et al., 2012). The cells of the lymphatic system, the lymphocytes, originate from hematopoietic stem cells (HSCs) and differentiate into B-lymphocytes (B-cells) or T-lymphocytes (T-cells) (Delves, et al., 2017, Gutman and Weissman, 1972, Santambrogio, 2013, Tanaka and Iwakiri, 2016). They fulfill different tasks in the innate and adaptive immune response. B-lymphocytes are responsible for antibody-mediated immune defense, since they form immunoglobulins, which get in contact with a foreign substances (e.g. pathogens). T-lymphocytes directly have contact with foreign substances or virus-infected cells. They evoke the cellular immune response by presenting T-cell receptors (TCR) on their cell surface.
3.2. B-cell development

B-lymphocytes (or B-cells) are characterized by rearranged immunoglobulin heavy and light chain genes, which encode antibodies. Antibodies are membrane bound receptors in all B-cell subtypes, with the exception of plasma cells. Plasma cells secrete a huge amount of antibody molecules and build the basis of humoral immune response (Delves, et al., 2017, Murphy, et al., 2018). B-cells develop from hematopoietic stem cells through mature immune-competent B-cells, and then to antibody-secreting plasma cells (Shapiro-Shelef and Calame, 2005). The production of B-cells occurs through a series of processes. Clusters of Differentiation (CD) proteins (a group of cell surface markers) and membrane immunoglobulin expression (e.g. IgM and IgD) can be used to identify the different stages of B-cell development or activation (Bernard and Boumsell, 1984, Brisslert, et al., 2006, Gathings, et al., 1977). The primary cell population in the bone marrow, which is clearly attributed to the B-cell development, consists of the pro-B-cells (Audzevich, et al., 2017, Osmond, et al., 1992). They arise from the lymphoid precursor cells of the hematopoietic system and already express the B-cell marker CD19 and the CD45 isoform B220 (Osmond, et al., 1992). Other important markers are CD43 (Treasure, et al., 1992, Wells, et al., 1994, Wiken, et al., 1989) and CD25 (Brisslert, et al., 2006, Rolink, et al., 1994), which allow a distinction between different developmental stages (Benschop and Cambier, 1999).

The B-cell development takes place in two steps: first, an antigen-independent step that include the immunoglobulin gene rearrangement process for formation of a functional B-cell receptor (BCR) (also referred to antibody (Ab) at the protein level or Immunoglobulin (Ig) at the gene level) within the B-cells. This occurs in mammals initially in the fetal liver and later in the bone marrow. The second step involves foreign antigens, which initiate B-cell activation and differentiation in the periphery (Benschop and Cambier, 1999, Murphy, et al., 2018, Rink, et al., 2015).

The key element in both development stages is the BCR. Its functionality in combination with appropriate growth factors, is a prerequisite to achieve each development stage (Mauri and Bosma, 2012).

The modeling of the BCR is a complex process and plays a critical role for building the adaptive immune system by enabling the detection of a wide variety of antigens. Immunoglobulin genes encode the BCR. However, to generate the specificity and high variability of BCRs, gene rearrangement is necessary.
This immunoglobulin gene rearrangement processes take place during the B-cell development and concatenates, in several steps, noncontiguous variable (V), diversity (D), and joining (J) gene segments (Delves, et al., 2017, Murphy, et al., 2018, Rink, et al., 2015).


At the stage of late pro-B-cell, the cells begin to rearrange the gene of the immunoglobulin (Ig) heavy chain locus on chromosome 14. In the first step, a DH- is joined with a JH-segment followed by a fusion of VH-segment to the already generated DH/JH-segment rearrangements. When this process results in the formation of a functional reading frame for the Ig heavy chain, the cell expresses the pre-B-cell receptor (Pre-BCR) on the surface. This marks the first selection point in the B-cell development, and the cell reaches the pre-B-cell level (Melchers, 2005, Rink, et al., 2015, van Zelm, et al., 2007). The Pre-BCR transmits a positive signal to the cell, causing it to enter the cell cycle and undergo cell divisions. In the case of a non-functional immunoglobulin gene rearrangement, the second allele is used for a further round of recombination. If this also fails, the cell undergoes apoptosis (Mak and Saunders, 2011). If the pre-B-cell expresses a Pre-BCR onto the cell surface the rearrangement of the light chain locus on chromosome 2 (κ locus) or 22 (λ locus) takes place. The light chain is formed by recombination of VL- and JL-segments. However, no D-segment is involved for formation of a rearranged light chain gene (Ehlich, et al., 1993, Ghia, et al., 1996, Murphy, et al., 2018, Rink, et al., 2015). After a productive rearrangement of the light chain, the complete BCR (immunoglobulin) can be expressed on the surface of the immature B-cell. Further immunoglobulin gene modifications at later stages of the B-cell development include the class switching involving the constant region of the heavy chain gene region (Stavnezer, 1996). All five immunoglobulin classes (IgM, IgD, IgG, IgA and IgE) are coded by a constant (C) segments of heavy chains gene locus (μ, δ, γ, α and ε) and increase additionally the diversity of the immunoglobulin (Yuan, 1984, Yuan and Tucker, 1984).

The antigenic specificity of the immunoglobulins is mainly determined by the complementarity-determining regions (CDRs). The VH gene segments encode the CDR1 and CDR2, whereas, the V, D and J segments collectively contribute for CDR3 formation, which is the most critical determinant of antigenic specificity. The framework
regions (FR) between the CDRs are less variable and are important for antibody structure (Boudinot, et al., 2002, Kiyoi, et al., 1992, Warren, et al., 2013). The formation of BCR (antibodies) is not triggered by an antigen, rather, the B-cells form a large number of antibodies with various specificities. Each B-cell has a unique receptor specificity. Upon antigen contact, the B-cell is being activated. B-cell activation occurs under appropriate conditions (e.g. in presence of interleukins, T-helper cells) and via intracellular signal cascades (Delves, et al., 2017, Murphy, et al., 2018). The cell multiplies, leaves the secondary lymphatic organs and begin as plasma cells with the production of high-affinity antibodies.

B-cell activation by antigens can be T-cell independent (TI) or T-cell dependent (TD) (Lanzavecchia, 1985, Playfair and Purves, 1971). If T-cell interaction occurs, the antigen is internalized (endocytosis), processed (antigen processing) and presented as peptide fragments in major histocompatibility complex (MHC) class II molecules on the surface of B-cells (antigen presentation). The presented antigen (peptide fragment) on the B-cell surface, is recognized via a specific T-cell receptor (TCR) on the T-cell. Co-stimulated via CD40 (B-cell) and CD40-Ligand (T-cell) interaction, and in presence of cytokines, the B-cell becomes activated. The activated B-cell migrates together with T-cells into lymphoid follicles and form a so-called germinal center (Harwood and Batista, 2010, MacLennan, 1994). In the germinal center, the B-cell starts to proliferate, generate high-affinity B-cell receptors by somatic hypermutation and affinity maturation, and further differentiates into different B-cell types (e.g. plasma cell, memory cell).

B-cells, which undergo T-cell independent activation (e.g. TI-antigens such as poly- or lipopolysaccharides of bacteria), however, do not form a germinal center. They are not able to undergo affinity maturation or complete isotype switching. These B-cells become activated by TI antigen cross-linking, they start to proliferate and produce always antibodies of the IgM type. B-cells which are not activated by antigen binding within 3 days are subject to apoptosis (Mak and Saunders, 2011).
### 3.3. Classification of lymphomas

Lymphomas include all types of neoplasia that derive from lymphocytes of the immune system. A correct distinction and classification of lymphoma types may be difficult, but is very important for the prognosis and ultimately for the selection of therapy. Malignant lymphomas can be divided into two main groups: Hodgkin's lymphoma (HL) and non-Hodgkin's lymphoma (NHL) (Harris, et al., 2008, Jiang, et al., 2017, Swerdlow, et al., 2016).

The tumor cells of Hodgkin's lymphoma are mononucleated Hodgkin- and multinucleated Reed-Sternberg cells. They constitute a minority of the cells (0.1–10% of the cells) in affected lymph nodes (Amini and Enblad, 2003, Kuppers and Hansmann, 2005, Kuppers, et al., 1994). HL can be further sub-classified into two disease categories, the nodular lymphocyte predominant HL (NLPHL) and the classical HL (cHL) which can be further distinguished into four subtypes: nodular sclerosing (NScHL), mixed cellularity (MCcHL), lymphocyte depleted (LDcHL) and lymphocyte rich (LRcHL) (Amini and Enblad, 2003, Harris, et al., 2008, Jiang, et al., 2017, Swerdlow, et al., 2016).

The classification of NHL has evolved steadily throughout the last century. An early classification system for NHL proposed by Gall and Mallory (Gall and Mallory, 1942) was further refined in the 1950s by Hicks et al (Hicks, et al., 1956). Twenty years later, it was recognized that NHLs were tumors of the immune system and derive from T- or B-cells, which led to the immunologically based classifications (Kiel classification) (Fulle and Pribilla, 1980, Lennert and Stein, 1978, Lennert, et al., 1975, Lukes and Collins, 1974). Therefore NHL are divided into major categories: B-cell lymphomas (B-NHL; approx. 80-90% of cases) and T-cell lymphomas (T-NHL; approx. 10-20% of cases) (Ansell, 2015, Ansell and Armitage, 2005, Nogai, et al., 2011). A more clinically guided classification of B-NHL consists of indolent B-NHL or low-grade malignancy (e.g. follicular lymphomas, chronic lymphocytic leukemia, etc.) and aggressive B-NHL or high-grade malignancy (BL and DLBCL) (Cheson, 2008, Fulle and Pribilla, 1980, Jiang, et al., 2017, Weisenburger, et al., 1982). In 1994 the International Lymphoma Study Group (ILSG) published the "Revised European-American Classification of Lymphoid Neoplasms" (R.E.A.L. Classification) which summarized the current existing knowledge about neoplasms of the immune system and their clinical aspects to simplify cooperation between clinicians and scientists (Harris, 1995, Harris, et al., 1994,
Sander, et al., 1997). The increased understanding of the immune system, the genetic abnormalities associated with NHL, and the methodological developments have led to the identification of several previously unrecognized subtypes of lymphomas (Martin-Subero, et al., 2009).

Microarray gene expression analyses showed that DLBCL can be subdivided into 3 molecular different subgroups (GCB, ABC, type 3) (Alizadeh, et al., 2000, Rosenwald, et al., 2002). Furthermore, bases on gene-expression profiling analysis of BL by Hummel et al (Hummel, et al., 2006), a molecular definition of BL was outlined which sharply differentiates BL from DLBCL. The World Health Organization (WHO) classification of lymphoid neoplasms of 2008 (Harris, et al., 2008) and its revision in 2017 (Arber, et al., 2016, Campo, et al., 2011, Swerdlow, et al., 2016, Wang and He, 2016) considers in addition to histological and immunophenotypical features, many important molecular characteristics (Swerdlow, et al., 2017, Swerdlow, et al., 2008).

The incidence of HL in USA and Europe population of is approximately 2-3/100,000 per year. For NHL, the incidence is approximately 11-12/100,000 per year. Hence lymphomas are a common cancer diseases (Morton, et al., 2006, Torre, et al., 2015).

![Figure 2: Simplified classification of malignant lymphoma.](image)

**Figure 2: Simplified classification of malignant lymphoma.** Based on histological features, malignant lymphoma can be classified into Hodgkin - and non-Hodgkin lymphoma. Non-Hodgkin lymphomas (NHL) comprises a large group of morphologically, immunophenotypically or genetically distinct lymphomas derived from B- or T-cells. Depicted and highlighted are only those NHL entities, which are of interest for this thesis: MYC-break positive Burkitt lymphoma (BL) and diffuse large B-cell lymphoma (DLBCL) with MYC-break (DLBCLpos) and those without a MYC-break (DLBCLneg).
3.3.1 Burkitt lymphoma (BL)

BL is one of the fastest growing human tumors and was named after its discoverer, the tropical doctor, Dr. Denis Burkitt (Burkitt, 1958). BL can be divided into three different types: (i) **Endemic Burkitt Lymphoma**: It is the most common type of BL, occurring mainly in Central and East Africa as well as in South America. Mostly, children and adolescents between 2-20 years are affected, with boys being affected twice as often as girls. (ii) **HIV-associated Burkitt lymphoma**: The HIV infection and the resulting immune deficiency predispose to the development of malignant tumors, typically only in the advanced stage of immunodeficiency. Nevertheless, HIV-associated Burkitt lymphoma belongs to the second most common HIV-associated neoplasia after Kaposi’s sarcoma. (iii) **Sporadic Burkitt lymphoma**: It is the rarest type of BL and occurs worldwide in patients on an average, 10 years older than those with endemic lymphoma (Knowles, 2001, Robertson, 2012). Endemic BL usually occurs as a maxillary or facial bone tumor (50%-60% of eBL cases), with non-endemic BL usually affecting the abdominal organs (91%) (Dozzo, et al., 2017).


The cause of BL is largely unknown, but an association with viral infections such as Epstein-Barr virus or HIV is already known (Hensel, et al., 2011, Hoffmann, et al., 2015, Pannone, et al., 2014). Around 90-95% of BL carry translocations which juxtapose the MYC gene from chromosome 8 next to the immunoglobulin heavy gene locus on chromosome 14 (approx. 85%), or immunoglobulin κ light chain on chromosome 2, or the λ light chain on chromosome 22 (Bernheim, et al., 1981, Croce, et al., 1983, Dallafavera, et al., 1982, Hummel, et al., 2006, Manolov and Manolova, 1972, Taub, et al., 1982, Zech, et al., 1976). Very rare BL cases lacking an identifiable MYC rearrangement or with unknown translocations (non-IG) partners are also known and are associated with an adverse clinical outcome (Haralambieva, et al., 2004, Hummel, et al., 2006, Leucci, et al., 2008). Furthermore, recent NGS studies of BL identify somatic mutations in the transcription factor TCF3, or its negative regulator ID3 in approximately 70% of BL cases. This respectively affects the function of transcription

3.3.2. Diffuse large B-cell lymphoma (DLBCL)

30-40% of the diagnosed B-cell non-Hodgkin's lymphomas (B-NHL) are diffuse large B-cell lymphomas (Coiffier, 2001, Lossos, 2005, Nayak, et al., 2013). Thus, DLBCL is the most common malignant lymphoma worldwide. The highest probability of developing a DLBCL is between the ages of 40 and 80 years (Chan, et al., 1997). Comparable to BL, men are more likely to be affected than women (Savage, et al., 2009). DLBCL is clinically, molecular and morphologically very heterogeneous, and the delineation of BL from DLBCL is not always possible by morphology and immunohistology. Early classification systems such as the Kiel classification from 1974, subdivided DLBCL based on morphologic criteria into centroblastic, immunoblastic and anaplastic DLBCL subtypes (Fulle and Pribilla, 1980, Gerardma.R, et al., 1974, Lennert and Stein, 1978, Lukes and Collins, 1974, Stansfeld, et al., 1988). Centroblastic subtype is the most common and is associated with better survival when compared to immunoblastic and anaplastic DLBCL (Engelhard, et al., 1997, Federico, et al., 2005, Swerdlow, et al., 2016).

Molecular gene expression analysis revealed that DLBCL can be sub-divided based on the similarities with the transcriptional profiles of their presumed cell of origin (COO) into two molecular groups corresponding to germinal center B-cells (GCB) and activated B-cells (ABC) (Alizadeh, et al., 2000, Rosenwald, et al., 2002, Sehn and Gascoyne, 2015). Around 30% of DLBCL cannot be assigned to either the GCB or the ABC type and are summarized as intermediate type. Patients with GCB-type DLBCL show a more favorable clinical course than those of the ABC or intermediate type (Alizadeh, et al., 2000, Blenk, et al., 2007, Hans, et al., 2004, Monti, et al., 2005).

The knowledge of the chromosomal alterations in DLBCL allows for further subtyping, which can be correlated with prognosis. DLBCL patients with chromosomal translocations involving MYC and BCL2 or BCL6 gene (double-hit lymphoma – DHL) and less often, all three genes (triple-hit lymphoma – THL) represents about 5% of all cases of all DLBCL cases. Patients with MYC translocations with or without additional BCL2 or BCL6 translocation usually show an aggressive clinical course with a poor

3.4. Therapies of aggressive B-cell non-Hodgkin's lymphoma (B-NHL)

Untreated patients suffering from aggressive B-NHL die within a few months. On the other hand, aggressive B-NHL are very sensitive to radiation and chemotherapy. A first improvement in treatment of DLBCL patients was achieved by introduction of the CHOP treatment scheme, a polychemotherapy, consisting of cytostatics (drug treatment of cyclophosphamide, hydroxydaunububin, oncovine and prednisolone), almost 40 years ago (Child, et al., 1983, Heinz, et al., 1985).

BL are usually treated with CHOP therapy plus methotrexate, followed by ifosfamide / etoposide / cytosine arabinoside according to the treatment protocol applied to patients with B-cell acute lymphocytic leukemia (B-ALL). This therapy is well-suited and promising for children, but not for adult patients (Atra, et al., 1998, Kujawski, et al., 2002, Pohlen, et al., 2011, Schwarzbich, et al., 2016). Since BL is a very fast-growing tumor, chemo- or radiotherapy leads to very efficient killing of the tumor cells. Response rates and long-term relapse-free survival after treatment are correspondingly high (up to 80%), but in advanced stages or adult age, the prognosis is less favorable (Sweetenham, et al., 1996).

The survival rate of DLBCL patients varies considerably, reflecting the (molecular) heterogeneity of the individual tumors. This holds true despite the addition of Rituximab to the CHOP scheme (R-CHOP) which represents standard of care for DLBCL patients since approx. 15 years (Mounier, et al., 2003, Savage, et al., 2009, Winter, et al., 2006). Rituximab is a monoclonal antibody directed against the B-cell surface marker CD20. With this treatment, a significant improvement in overall survival (OS) can be reached. Unfortunately, dependent on patient age, about 30-60% of the DLBCL patients display no longtime treatment success (Coiffier, et al., 2002, Nogai, et al., 2011, Rosenwald, et al., 2002, Sehn and Gascoyne, 2015).

The impact of molecular diversity of DLBCL (regarding to gene expression signatures, genetic alterations and other molecular features) for the selection of eligible therapeutic options is still underdeveloped. Hence, new treatment options in combinations with
precise molecular and histopathological characterization are necessary to improve clinical outcome.

3.5. Transcription factor MYC

MYC is one of the most important and known transcription factors. It was first discovered in avian leukemia virus (ALV), which cause myelocytomatosis in chicken and therefore, carries the name v-MYC (viral MYC) (Duesberg and Vogt, 1979, Sheiness, et al., 1978). The cellular version of the MYC protein, known as c-MYC (cellular MYC, and later only designated as MYC), was discovered initially in BL (Colby, et al., 1983, Dallafavera, et al., 1982, Dominguez-Sola, et al., 2007). Subsequently, it was described in other mammals (Hall, et al., 1991, Ma and Erickson, 1988, Reuse, et al., 1990) and plants (Gong, et al., 1999). MYC is a nuclear phosphoprotein, which is able to bind DNA, to initiate and regulate the transcription of a plethora of genes. MYC is a major global transcription factor that regulates approximately 10-15% of all human genes (Fernandez, et al., 2003, Knoepfler, 2007, Nie, et al., 2012, Poole and van Riggelen, 2017). MYC binds sequence-specifically to the DNA, together with MYC-associated X-factor protein (MAX) (Nair and Burley, 2003). These MYC-MAX protein complexes can act as activators (Eberhardy and Farnham, 2001, Frank, et al., 2003, Liu, et al., 2003, Saunders, et al., 2006, Vervoorts, et al., 2003, Wu, et al., 1999) or suppressors of RNA polymerases and affect the expression of other genes (Brenner, et al., 2005, Gartel and Shchors, 2003, Gomez-Roman, et al., 2003, Latchman, 1997, Lee, et al., 1997, Mitchell and Tjian, 1989, Si, et al., 2010). Activation of MYC in noncancerous cells is triggered by growth factors and adequate nutrients, which bind to the cell surface and initialize via activation of histone acetyltransferases (TRRAP-related histone acetylation complexes and INI1-associated chromatin-modulating proteins), a genomic activation cascade, which is responsible for cell proliferation and differentiation (Liu, et al., 2003, Marcu, et al., 1992, Meichle, et al., 1992, Stine, et al., 2015). There are multiple levels of feedback loops and checkpoints of MYC activity especially because MYC is involved in many cellular processes such as cell growth, differentiation, cell cycle progression, apoptosis and cellular transformation (Amati, et al., 1998, Grandori, et al., 2000, Henriksson and Luscher, 1996, Levens, 2002, Stine, et al., 2015). In contrast, in cancer cells, an aberrant MYC activity is likely to promote tumorigenesis and is associated with approximately 60-70% of human cancers (Dang, et al., 2009, Dominguez-Sola and Gautier, 2014, Poole and van Riggelen, 2017,

The human MYC proto-oncogene (HGN-ID 7553) is located on chromosome 8q24 and consists of 3 exons which encode a 2168 bp long mRNA, 1320 bp of which are protein coding (Dallafavera, et al., 1982, Persson, et al., 1984). An open reading frame results in three distinct MYC protein isoforms: MYC-1 (454 aa), MYC-2 (439 aa) and a substantially shorter form called MYC-S (339 aa) (Hann, 1995, Hann, et al., 1994, Hann, et al., 1988, Spotts, et al., 1997) (Figure 3). The most frequently occurring AUG-initiated form, MYC-2 (75-90%), is heterogeneously expressed in all human tissues. It possesses a nucleus localization signal (NLS) and a basic helix-loop-helix (bHLH), as well as a leucine-zipper (LZ) structural motif, capable of binding DNA. The canonical binding sequence (E-Box) of MYC protein is CACGTG, but it is also able to bind non-canonical binding motives ((G|A)CA(A|C|G|T|N)(A|C|G|T|N)TG(G|A)) (Blackwell, et al., 1993, Chaudhary and Skinner, 1999, Nair and Burley, 2003, Zeller, et al., 2006). Transcriptional gene activation involves the recruitment of multiple coactivators and protein complexes (e.g. mediator complex, positive transcription elongation factor b (P-TEFb), the ATPases TIP48 and TIP49, and histone acetyltransferases such as CREB-binding protein (CBP) and p300, GCN5 and TIP60) to E-box elements (Adhikary and Eilers, 2005, Blackwood and Eisenman, 1991). Both MYC mRNA and MYC protein have very short half-lives (approximately 10 min and 25 min, respectively) (Ciechanover, et al., 1991, Dani, et al., 1984, Hann, 2006, Herrick and Ross, 1994, Rabbitts, et al., 1985). Normal fibroblasts in cell culture express a few thousand molecules of MYC protein per cell (Waters, et al., 1991), but this can be over two orders of magnitude higher in cancer cells and cell lines (Lin, et al., 2012, Maguire, et al., 1983, Taub, et al., 1982). The oncogenic potential of MYC in lymphomagenesis was first demonstrated in 1983 by Adams et al., who showed that the coupling of MYC gene with the immunoglobulin enhancer in transgenic mice resulted in the development of immature and mature B-cell neoplasms (Adams, et al., 1983, Adams, et al., 1985). The MYC proto-oncogene transforms into an oncogene often caused by a chromosomal translocation, insertional mutagenesis (via viruses) or through amplification of MYC gene (Meyer and Penn, 2008, Yokota, et al., 1986). MYC expression in various tumors is very heterogeneous, depending on the type of MYC dysregulation, and whether MYC overexpression is a timely or late event in the tumorigenesis (Shachaf, et al.,
2008). Because of its ubiquitous role, MYC is thought to be an attractive therapeutic target.

Figure 3: (A) Structure of MYC gene consisting of three exons, potential promotor areas (P0-3) and polyadenylation signals (pA1 and pA2) are indicate. (B) Structure of MYC isoforms (MYC-1, MYC-2 and MYC-S) including the high homologous conserved MYC Boxes (MB1, MB2) within the N terminal Domain (NTD), a nucleus localization signal (NLS) within the central peptide region, and bHLH-LZ motif within the C-terminal Domain (CTD). Stars indicates potential posttranslational modification sites. Figure adapted and modified from the UniProt Knowledgebase, 2017, human MYC # P01106, viewed 05.09.2017 <http://www.uniprot.org/uniprot/P01106> and literature (Chen, et al., 2014, Sarid, et al., 1987).

3.6. Identification of MYC alterations used to support diagnostics

In the case of lymphomas, specific chromosomal abnormalities can be used to support diagnostic process. Detection of a MYC translocation, in addition to other diagnostic criteria (e.g. B-cell marker and proliferation rate), is a hallmark of BL, but is also found in a fraction of DLCBL. Currently, there are three diagnostic methods available to detect MYC alterations in cell suspensions or tissue samples: conventional cytogenetics, fluorescent in situ hybridization (FISH) and immunohistochemistry (IHC) (Fernandez, et al., 2012, Nguyen, et al., 2017, Raess, et al., 2018).

Translocations of the MYC gene to other chromosomal locations can be identified by conventional cytogenetics or karyotype analysis. Unfortunately, this technique requires fresh tissue material and is time and labor intensive. It is thus not applicable in many diagnostic settings. An alternative method is the fluorescence in situ hybridization (FISH), using break-apart or fusion fluorochrome-labeled long DNA probes. A major advantage of the FISH method is that it can be performed on both fresh and formalin-fixed paraffin-embedded (FFPE) tissue sections. However, based on the design of
DNA probes, FISH might eventually produce false negative results. This is due to the large spatial distribution of the MYC breakpoints in the genomic region of up to 1000 Kb and the involvement of different break point regions (Haralambieva, et al., 2004, van Rijk, et al., 2008). The detection of MYC protein can be achieved by IHC on FFPE tissue section by antibodies targeting the N-terminus of MYC protein (Cattoretti, 2013, Gurel, et al., 2008). The IHC method can be carried out quickly and shows MYC deregulation by detection of increased MYC protein level, therefore applicable in most pathology laboratories. Finally, MYC (over-) expression can also be identifiable at the transcriptional level. Therefore, a variety of different methods are available ranging from real-time RT PCR, array-based approaches, NanoString Encounter methods to RNA-based NGS (Alidousty, et al., 2018, Jackstadt, et al., 2013). However, the detection of MYC (over-) expression at the transcriptional level is currently not the method of choice in the diagnostic setting.

4. Aim of thesis

The aim of this work is to investigate the various mode of action of MYC in two types of aggressive B-cell lymphoma, Burkitt lymphoma (BL) and diffuse large B-cell lymphoma (DLBCL). The understanding of the molecular mechanisms of MYC activity might help to identify new diagnostic or therapeutic options.

To shed more light on the role of MYC, cell lines derived from: (1) MYC-break positive BL, (2) MYC-break positive DLBCL and (3) MYC-break negative DLBCL were investigated by complex high-throughput experiments including (i) metabolomics, (ii) proteomics, (iii) DNA binding of MYC and (iv) global gene expression. In addition, the impact of long noncoding RNAs (lncRNA) on the ability to modulate the MYC driven cellular transcriptome were investigated.
5. Results

5.1. Publications

**Publication 1** (Kleo, et al., 2018, BMC Cancer, submitted)


The original article is online available at:
Published at BMC Cancer. 2019 Apr 5;19(1):322
PMID: 30953469

**Publication 2** (Schwarzfischer, et al., 2017)


Published at J Proteome Res. 2017 Mar 3;16(3):1105-1120.
The original article is online available at:
DOI: https://doi.org/10.1021/acs.jproteome.6b00164
PMID: 28161958

**Publication 3** (Doose, et al., 2015)


Published at Proc Natl Acad Sci U S A. 2015 Sep 22;112(38):E5261-70.
The original article is online available at:
DOI: https://doi.org/10.1073/pnas.1505753112
PMID: 26351698
5.2. Manuscript 1

Title: Identification of ADGRE5 as discriminating MYC target between Burkitt lymphoma and diffuse large B-cell lymphoma.


Journal: BMC Cancer, submitted 2018

The original article is online available at:
Published at BMC Cancer. 2019 Apr 5;19(1):322.
PMID: 30953469

5.2.1. Synopsis

MYC is a well described oncogenic transcription factor that plays a multifunctional role in many biological processes including cell proliferation and differentiation. MYC is also associated with many types of cancers including aggressive lymphoma such as BL and DLBCL. BL patients usually carry MYC translocation and show a better clinical course, whereas DLBCL patients with MYC translocation have a significantly worse clinical outcome. These opposite clinical courses raise the question regarding the role MYC plays in this distinct observations. In this work, three groups of cell lines derived from different aggressive non-Hodgkin B-cell lymphomas were employed: (i) Burkitt lymphoma (BL), (ii) MYC positive diffuse large B-cell lymphoma (DLBCLpos) – carrying MYC translocations and (iii) DLBCL without MYC translocation (DLBCLneg). Genome-wide MYC-DNA binding sites were determined by chromatin immunoprecipitation, followed by a high-throughput sequencing experiment (ChIP-Seq). As a control, the ChIP-Seq experiment was performed with a H3K4me3 antibody, indicating areas of transcriptionally active nearby genes. ChIP-Seq data was bioinformatically evaluated and differential binding analysis was performed to ascertain the MYC binding status and histone modification status within the investigated entities (BL vs. DLBCL; BL vs. DLBCLpos; BL vs. DLBCLneg and DLBCLneg vs. DLBCLpos). In addition, global RNA expression profiles were determined by RNA sequencing (RNA-Seq). Our study identified different genome-wide MYC-DNA binding sites and RNA expressions profiles in BL and DLBCL cell lines. Furthermore, we show that ADGRE5 (alias CD97) is a
MYC target gene, which was found to be specifically expressed in BL. These results suggest that ADGRE5 is promising marker for distinction between BL and DLBCL irrespective of the MYC translocation status of DLBCL.

5.2.2. Contribution

<table>
<thead>
<tr>
<th>I Concept</th>
<th>General study design</th>
<th>quota</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cell culture experiments design</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>ChIP-Seq experiments design</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>RNA-Seq experiments design</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Molecular analysis (RT-PCR, WB, IHC)</td>
<td>100%</td>
</tr>
</tbody>
</table>

| II Execution | Cell culture for ChIP-Seq & RNA-Seq | 100% |
|              | Protein, DNA & RNA Isolation | 100% |
|              | Wester-Blot (WB) | 100% |
|              | Realtime PCR (RT-PCR) | 100% |
|              | Chromatin Immunoprecipitation (ChIP) | 100% |
|              | Sequencing | 10% |
|              | Immunohistology (IHC) | 100% |
|              | Bioinformatics data analysis (Regensburg) | 30% |

| III Reporting | Preparation of “abstract & introduction” section | 95% |
|              | Preparation of “material/method” section | 85% |
|              | Preparation of “results” section (including figures) | 95% |
|              | Preparation of “discussion” section | 95% |

5.2.3. Manuscript Kleo et al.
Identification of ADGRE5 as discriminating MYC target between Burkitt lymphoma and diffuse large B-cell lymphoma.

Karsten Kleo1*, Lora Dimitrova1, Elisabeth Oker1, Nancy Tomaszewski1, Erika Berg1, Franziska Taruttis2, Julia C. Engelmann2,a, Philipp Schwarzfischer3, Jörg Reinders3, Rainer Spang2, Wolfram Gronwald3, Peter J. Oefner3, Michael Hummel1

1 Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Institute of Pathology, D-10117 Berlin, Germany
2 Statistical Bioinformatics and 3 Functional Genomics, Institute of Functional Genomics, University of Regensburg, D-93053 Regensburg, Germany
a Current address: Department of Marine Microbiology and Biogeochemistry, NIOZ Royal Netherlands Institute for Sea Research, 1790 AB Den Burg, The Netherlands

* Corresponding author
E-Mail: karsten.kleo@charite.de
Abstract

**Background:** MYC is a heterogeneously expressed transcription factor that plays a multifunctional role in many biological processes such as cell proliferation and differentiation. It is also associated with many types of cancer including the malignant lymphomas. There are two types of aggressive B-cell lymphoma, namely Burkitt lymphoma (BL) and a subgroup of diffuse large cell lymphoma (DLBCL), which both carry MYC translocations and overexpress MYC but both differ significantly in their clinical outcome. In DLBCL, MYC translocations are associated with an aggressive behavior and poor outcome, whereas MYC-positive BL show a superior outcome.

**Methods:** To shed light on this phenomenon, we investigated the different modes of actions of MYC in aggressive B-cell lymphoma cell lines subdivided into three groups: (i) MYC-positive BL, (ii) DLBCL with MYC translocation (DLBCLpos) and (iii) DLBCL without MYC translocation (DLBCLneg) for control. In order to identify genome-wide MYC-DNA binding sites a chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-Seq) was performed. In addition, ChIP-Seq for H3K4me3 was used for determination of genomic regions accessible for transcriptional activity. These data were supplemented with gene expression data derived from RNA-Seq.

**Results:** Bioinformatics integration of all data sets revealed different MYC-binding patterns and transcriptional profiles in MYC-positive BL and DLBCL cell lines indicating different functional roles of MYC for gene regulation in aggressive B-cell lymphomas. Based on this multi-omics analysis we identified ADGRE5 (alias CD97) - a member of the EGF-TM7 subfamily of adhesion G protein-coupled receptors - as a MYC target gene, which is specifically expressed in BL but not in DLBCL regardless of MYC translocation.

**Conclusion:** Our study describes a diverse genome-wide MYC-DNA binding pattern in BL and DLBCL cell lines with and without MYC translocations. Furthermore, we identified ADGRE5 as a MYC target gene able to discriminate between BL and DLBCL irrespectively of the presence of MYC breaks in DLBCL. Since ADGRE5 plays an important role in tumor cell formation, metastasis and invasion, it might also be instrumental to better understand the different pathobiology of BL and DLBCL and help to explain discrepant clinical characteristics of BL and DLBCL.

**Keywords:** ADGRE5, CD97, MYC, ChIP-Seq, RNA-Seq, lymphoma, BL, DLBCL
**Background**

The transcription factor MYC plays a multifunctional role in many cellular processes such as cell cycle progression, apoptosis and cellular transformation. Over-expression of MYC leads to an increased replication activity and is associated with different types of cancer. This holds also true for tumors of the immune system especially aggressive B-cell non-Hodgkin lymphomas (B-NHL) such as Burkitt lymphoma (BL) and diffuse large B-cell lymphoma (DLBCL). BL is an extremely fast growing tumor that carries immunoglobulin/MYC translocations in almost all cases. The tumor is predominantly found in male children but may also occur in adults especially with a compromised immune system. Treatment of BL is mainly based on high dose chemotherapy with usually favorable clinical outcome [1]. In contrast, DLBCL rarely carries MYC rearrangements, which may be associated with both immunoglobulin and non-immunoglobulin genes. Whereas DLBCL without MYC translocation reveals long-term survival of 60-70% of the patients treated with combined immune-chemotherapy, DLBCL with MYC translocation – regardless of its translocation partner – shows a very poor clinical outcome [2-8]. It is currently unclear why BL and DLBCL with MYC translocations display this very different clinical course. In addition, molecular features for a precise stratification of patients into BL and DLBCL with MYC translocation are lacking despite the need for different treatment modalities. To determine the potentially different role of MYC in BL and DLBCL, we aimed at identifying their molecular features by means of chromatin immunoprecipitation combined with high-throughput sequencing (ChIP-Seq) and whole transcriptome shotgun sequencing (RNA-Seq) employing B-cell lymphoma cell lines. Validation of the results was performed with primary lymphoma tissue samples.

**Methods**

**Cell Culture**

Three MYC break positive BL cell lines (Blue-1 / ACC-594; BL-2 / ACC-625 and BL-41 / ACC-160), two MYC break positive (Carnaval / ACC-724; U2932-R2 / ACC-633) and two MYC break negative (Karpas-422 ACC-32, U2932-R1 / ACC-633) DLBCL cell lines (overview Fig 1A) were obtained in 2012 from the German Collection of Microorganisms and Cell Cultures (DSMZ). The sub-clones U2932-R1 and U2932-R2 were kindly provided by Dr. Quentmeier (DSMZ, Braunschweig, Germany) [9]. All cell lines were negatively tested for mycoplasma contamination prior to use and are currently not listed as cross-contaminated or misidentified cell lines according the International Cell Line Authentication Committee (ICLAC). All cell lines were cultivated in RPMI 1640 medium supplemented with GlutaMAX™-I (Gibco, Thermo Fisher Scientific) and containing 20% of heat inactivated fetal bovine serum (PAN Biotech, Aidenbach, Germany) under a humidified atmosphere with 5% CO₂ at 37°C. Cells were thawed and continuously split 3 times per week for a maximum period of three weeks. Cell counting was performed on a BD Accuri C6 Flow Cytometer (BD Biosciences, New Jersey, United States) and cell viability was determined by propidium iodide (PI) – staining (BD Bioscience,
Heidelberg, Germany) according to the manufacturer’s recommendations. Only cells, which exhibited more than 90% vitality, were used for further investigation.

**Western Blotting**

1.5 x 10^6 vital cells were washed three times with PBS and lysed with protease inhibitors containing RIPA buffer supported by sonication. After measuring protein concentration using the BCA protein assay kit (Pierce, Thermo Fisher Scientific), protein lysates were separated under denaturing conditions via gels electrophoresis using 16% sodium dodecyl sulfate polyacrylamide gels (Invitrogen, California, United States) and transferred to Hybond-ECL nitrocellulose membranes (Amersham Biosciences, New Jersey, United States) by electroblotting. Membranes were blocked with a PBST 5% dry milk solution for 1 hour followed by incubation with the respective primary antibody solution at 4°C overnight. Subsequently, membranes were washed three times with PBST and incubated for 1 hour with a secondary antibody conjugated with horseradish peroxidase (information on primary and secondary antibodies is available in Supplementary S1 Table). Chemiluminescence was detected using HRP substrate (Luminata Forte, Merck Chemicals GmbH, Darmstadt, Germany) and FusionCapt Advance analysis Software (Fusion device, Vilber Lurmat GmBH, Eberhardzell, Germany).

**Quantitative real-time PCR analysis**

Total RNA was isolated from 1 x 10^6 vital cells after washing with PBS employing NucleoSpin RNA Kit (MACHEORY-NAGEL GmbH & Co. KG, Düren, Germany). RNA fluorometric quantification was performed by means of the Qubit RNA quantification assay (Thermo Fisher Scientific). Total RNA was reverse transcribed into complementary DNA (cDNA) using TaqMan reverse transcription reagents. Real-time PCR analysis was realized using TaqMan Real-Time PCR Master Mix on a Step One Plus Real-Time PCR System (Thermo Fisher Scientific). All procedures were performed according to the manufacturer’s recommendations. RT-PCR Taq-Man probes are listed in Supplementary S1 Table. Relative RNA expression was calculated according to the comparative Ct method [10] using the average expression based on triplicates of two biological replicates of each cell line. For endogenous control b2-microglobulin (B2M) or succinate dehydrogenase complex, subunit A (SDHA) were used.

**ChIP-Seq experiments**

Chromatin immunoprecipitation (ChIP) was done according to published protocols [11, 12] with few modifications. Briefly, 2 x 10^7 vital cells were fixed for 10 min at 4°C in medium containing 1% formaldehyde. After blocking with 0.1 M glycine and washing four times with PBS, the cells were snap frozen and stored at -80°C. After thawing on ice each cell pellet was resuspended in 5 mL cold LB1 lysis buffer, incubated for 10 min at 4°C and for further 10 min ambient temperature in 5 mL LB2 lysis buffer before being finally dissolved in 3 mL LB3 buffer. Sonication was performed for 45 min [three cycles of 15 min each at high power in pulsed
mode (30 s on and 30 s off)] using titanium rods combined with a Bioruptor Sonicator (Diagenode, Seraing, Belgium). After addition of 300 μL 10% (vol/vol) Triton X-100 and centrifugation the supernatant was removed, 50 μL of which were stored as input DNA sample. 1.5 mL of the supernatant was incubated with 10 μg MYC antibody or 5 μg H3K4me3 antibody at 4°C overnight. For ChIP antibody information, refer to Supplementary S1 Table. For precipitation of DNA indirectly bound to the respective antibody, 30 μg Dynabeads coupled with Protein G (Thermo Fisher Scientific) were added for each μg antibody and incubated for 3 h at 4°C. Subsequently, the beads were washed and the immunoprecipitated (IP) DNA was eluted. Finally, the eluate (input DNA and IP DNA) was reverse cross-linked overnight at 65°C followed by digestion with RNase A and Proteinase K. The resulting DNA was phenol/chloroform extracted, precipitated and the DNA was resuspended in 30 μL 10 mM Tris-HCl, pH 8.0. DNA was subjected to fluorometric quantification by the Qubit DNA quantification assay (Thermo Fisher Scientific). Ten ng of chromatin-immunoprecipitated DNA sample were processed with NEBNext ChIP-Seq Library Prep Master Mix Set for Illumina library generation according to the manufacturer’s recommendations. All amplified libraries were analyzed with the DNA 1000 Kit on the 2100 Bioanalyzer (Agilent, California, United States).

Single-read NGS was done on an Illumina HiSeq 1500 system (50 cycles). Illumina adapters were trimmed from the raw sequence data and low quality bases and reads were removed with trimmomatic (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) [13]. Sequence data was aligned to the main chromosomes of the human reference genome (GRCh38) with bowtie version 0.12.7 (-e 70 -k 1 -m 1 -n 2 –best) [14]. H3K4 and MYC peaks were called with MACS2 [15] with a q-value cut-off of 0.1 and the peaks from the two replicate ChIP samples were summarized with IDR [16], keeping all peaks with an IDR < 0.1. Final peaks were annotated to the nearest transcription start site (TSS) using gene annotation from Ensembl release 77. Only peaks with a maximum distance of 2,000 bp to a TSS were kept. Artificial peaks were removed using the ENCODE blacklist (https://sites.google.com/site/anshulkundaje/projects/blacklists). Differential peaks between DLBCL with and without MYC break and BL were estimated using DiffBind [17] tool.

**RNA-Seq analysis**

Total RNA was isolated from 1 x 10^6 lymphoma cells, which were previously spiked in with 1 x 10^5 insect cells (Schneider cells) for data calibration [18]. The quality of the RNA was determined with an Agilent 4200 TapeStation and Software A.01.05 (Agilent, California, United States). 500 ng RNA per sample were processed using the Illumina TruSeq Stranded mRNA LT Sample Prep Kit following the manufacturer’s instructions to generate libraries for RNA sequencing. Samples were sequenced on a Hi-Seq 4000 (single read mode; length 150 bp) using the Illumina HiSeq 3000/4000 SBS 150 cycle kit. Sequence reads were aligned to a concatenated genome that consisted of the human (GRCh38) and the Drosophila melanogaster (BDGP5) reference genome, using STAR alignment tool [19] with default parameters. Gene annotation from Ensembl release 77 and feature Counts [20] with default parameters were used to assign read counts to human and *Drosophila* genes. Before differential gene expression analysis, we calculated DESeq2 sample sizeFactors [21] on the
*Drosophila* gene counts and applied them to the human sample data. This way, gene expression levels of the cell lines were calibrated to the number of sample cells. Then, gene expression levels were modeled with a generalized linear model assuming negative binomial distributed data and categorical variables for the lymphoma type (BL or DLBCL) and MYC status (MYC break positive or negative). Gene expression changes were tested for significance with the Wald test and fold changes with an associated False Discovery Rate (FDR) below 0.05 were considered significant differentially expressed.

### Proteomics

The SWATH-MS-based quantification of the proteins ADGRE5, BYSL and NPM1 was obtained from previously published data [22]. SWATH-MS measurements were carried out on a TripleTOF 5600+ (Sciex, Darmstadt, Germany) coupled to an Ultimate 3000 nano-HPLC-system (Dionex, Idstein, Germany) using an 88 min-binary gradient. The PeakView 2.1 software (Sciex, Darmstadt, Germany) was employed for quantification of the peptides based on an in-house library. Only peptides with FDR<1% and confidence >95% were considered for quantification. Peptide intensities were summed up and normalized to total protein intensity. Statistical tests were conducted using heteroskedastic 1-way ANOVA.

### Immunohistochemistry

Immunohistochemical staining was performed using sections derived from formalin-fixed paraffin-embedded cell line blocks (n = 12) and primary tissue samples (n = 38). The use of human primary tissue samples was approved by the Institutional Review Board of the Charité – Berlin (EA4/104/11). The immunostaining carried out using the Leica Bond-maX autostainer (Leica Biosystems, Illinois, United States) according to the manufacturer’s protocol. After heat-induced epitope retrieval, the sections were incubated with anti-c-myc and anti-CD97 (ADGRE5) rabbit antibodies, respectively (dilution 1:200). Horseradish peroxidase-labeled Anti-rabbit-IgG using the Bond Polymer Refine Detection Kit (Leica Biosystems, Illinois, United States) was employed to convert the chromogen substrate. Staining was performed with appropriate positive and negative controls.
Results

First, we determined MYC mRNA and MYC protein expression by qRT-PCR, Western blotting and immunohistochemistry, respectively, in cell lines derived from BL, DLBCLpos and DLBCLneg patients (Fig 1 B-D). With the exception of BL-41, all MYC break positive cell lines showed high expression of MYC mRNA. The level of MYC protein expression corresponded without exception with the presence of MYC breaks. The discrepant results between MYC RNA and MYC protein expression in BL-41 might reflect a longer half-life time of the MYC protein in BL-41 as compared to the other cell lines with MYC breaks [23-27]. Thus, less RNA is required to generate high amounts of MYC protein.

Fig 1. MYC expression in DLBCL and BL cell lines. (A) Cell lines categorized according to their genomic MYC status (MYC break). (B) Quantitative MYC RNA expression as determined by RT-PCR; endogenous control for normalization: B2M expression. (C) Western Blot analysis of MYC protein expression. (D) Immunohistochemical (IHC) staining for the cellular localization and distribution of MYC protein.

To investigate the MYC DNA-binding capabilities in BL and DLBCL, we performed MYC ChIP-Seq experiments to determine genome-wide MYC DNA-binding sites. Additional ChIP-Seq experiments for trimethylation of histone H3 at lysine 4 (H3K4me3) were carried out in order to locate genomic areas with open chromatin as indicators for potential transcriptional activity.
of nearby genes [28, 29]. To bioinformatically identify differential MYC DNA-binding sites the DiffBind package [17] was employed using a pairwise comparison of the cell groups (BL vs. DLBCL; BL vs. DLBCLpos; BL vs. DLBCLneg and DLBCLpos vs. DLBCLneg). Similar differential binding analysis was performed with H3K4me3 ChIP-Seq data to ascertain genome wide differential histone patterns and potential active transcriptional sites.

Detailed results of the bioinformatics analyses are available in supplementary S2 File (ChIP-Seq data), while Figs 2 and 3 depict aggregated data. The overall number of MYC DNA-binding sites was higher (approx. 2-fold) in MYC break positive (BL, DLBCLpos) than MYC break negative (DLBCLneg) cells (Fig 2A). Next, we explored whether genes associated with MYC-binding differed between the three groups of cell lines. Our data clearly indicate that there is not only a difference in the number of genes but in addition, that also different genes are targeted by MYC and/or H3K4 (Fig 2B). To identify differential MYC-binding genes we performed a differential peak analysis comparing four combinations: BL vs. DLBCL, BL vs. DLBCLneg, DLBCLpos vs. DLBCLneg and BL vs. DLBCLpos. Cell lines carrying MYC breaks have more genes located in the vicinity of MYC-binding sites which leads to a higher number of differential MYC-binding peaks in relation to MYC break negative cell lines (Fig 2C). Figure 3 highlights a list of twenty target genes selected that yielded the highest fold changes. The analysis of the MYC-binding motifs of MYC target genes showed an interesting distribution (Fig 2D) with a preference for non-canonical E-Box motives (approx. 45 %), while only 4% carried exclusively the classical canonic E-Box motif (CACGTG) and 19% both motifs. Strikingly, 32% of identified MYC targets genes displayed no known MYC-binding motifs. Non-canonical and/or canonical E-box was present in approx. 68 % of MYC target genes, thus corroborating previous studies of global mapping of MYC-binding sites [30]. However, the presence of E-box motives in the binding loci did not correlate with the regulation of associated genes [31, 32].
Fig 2. Differential binding patterns obtained by ChIP-Seq experiments. (A) Total gene counts identified by MYC-Chip, H3K4me3-Chip, and an overlay of MYC/H3K4me3-Chip peaks after MACS2 IDR peak calling. (B) Venn diagrams illustrate the number of identified targets after IDR peak calling of MYC and H3K4 ChIP, respectively, limited to within 2,000 bp from Origin of Replication (ORI). Each count presents a single Ensembl gene ID. (C) Differential binding analysis between different lymphoma entities. Each count presents a single Ensembl gene ID, limitation by 2,000 bp of ORI, IDR < 0.1 and p-value < 0.05. (D) Distribution of MYC E-Box binding motif within the identified genes with differential MYC-Chip peaks.
Fig 3. Selected differentially bound genes derived from MYC and H3K4me3 ChIP-Seq experiments. Twenty selected target genes with a highest Log Fold Change value obtained from differential binding analysis for each pairwise comparison (BL_vs_DLBCL; BL_vs_DLBCLpos; BL_vs_DLBCLneg; DLBCLpos_vs_DLBCLneg). Restriction to 2,000 bp upstream of transcriptional start site; IDR < 0.1 and p-value < 0.05.

MYC-binding is not the sole factor for activating of gene expression and associated functional consequences. In order to gain a deeper insight into the transcriptional landscape and the impact of the various MYC-binding patterns, we performed RNA-Seq and correlated the results with the presence of MYC breaks and with the MYC and H3K4 binding patterns. In Fig 4A the number of genes differentially expressed among the three cell line groups is given. The highest number of differentially expressed genes was found between BL and DLBCL in general, whereby the expression difference between BL and DLBCLpos was the lowest. This demonstrates that MYC has a major impact through activation of the same gene set, which constitutes a significant proportion of the entire transcriptome. In harmony with this notion, the comparison of the RNA-Seq data between BL and DLBCLneg, and DLBCLpos and DLBCLneg revealed very similar numbers of differentially expressed genes. This reinforces the similarity in the gene expression profiles of both types of MYC break positive cell lines. Lists of differentially expressed genes are given in Supplementary S3 File (RNA-Seq data). In Fig 4B some differentially expressed genes (from Supplementary S3 File) are functionally grouped into clustering of differentiation (CD) molecules (B1), integrin molecules (B2) or MYC-related molecules (B3) and visualized as heat maps. Most of the identified CD molecules seem to be upregulated in MYC break positive (BL, DLBCLpos) cell lines compared to MYC break negative (DLBCLneg) cell lines.
To validate the data derived from genome-wide DNA-binding and gene expression, we performed additional gene-specific ChIP (MYC and H3K4me3) and RT-PCR experiments (Fig 4C). The selection criteria for the target genes were MYC DNA-binding according to ChIP-Seq and differential expression according to RNA-Seq. Among the identified genes, ZAP-70, ADGRE5, CDK20, GPAM, SMAD1 and TERT were the most interesting. Genes lacking differentially expression such as LARS, FARSA and already described as MYC target genes like BYSL [33] and NMP1 [34] were selected as positive control. By independent validation assays we were able to confirm the data derived from our ChIP-Seq and RNA-Seq analyses and demonstrate that a set of genes is able to reliably differentiate between lymphoma entities.
A

B

C1

C2

C3

Myc ChIP
H3K4me3 ChIP
mRNA Expression

B1

B2

B3

Delta Ct-value

Log2FoldChange

p < 0.05

p < 0.05
**Fig 4. RNA-Seq and validation of selected targets.** (A) Overview of total counts of identified RNA-Seq targets after differential expression analysis between pairs of lymphoma (sub-) entities. (B) Exemplary heatmaps of differentially expressed targets grouped for (B1) clustering of differentiation molecules, (B2) integrin molecules, and (B3) MYC-related molecules. (C) Summary of the validation experiment for selected targets via additional MYC / H3K4me3 ChIP enrichment analysis (C1) and additional TAQ-MAN RT-PCR analyses (C2) in BL (Blue-1; BL-2, BL-41), DLBCLneg (Karpas-422; U2932-R1) and DLBCLpos (U2932-R2; Carnaval) cell lines (n = 2 biological replicates).

Interestingly **ADGRE5** (previously designated as **CD97** and marked by an arrow in Fig 4, B1) shows strong selective expression in BL cell lines. On the other hand, some integrin molecules known to be binding partners of ADGRE5 are downregulated in BL as compared to DLBCL.

To validate this interesting outcome we quantify the proteins of ADGRE5 and already known homogeneous expressed MYC targets like BYSL and NPM1, obtained from previously published proteomic data [22] (Fig 5 A) and western blot analysis (Fig 5 C). Finally, we demonstrated the discriminating character of ADGRE5 between BL and DLBCL in additional immunostainings of cell lines and FFPE tissue samples (Fig 5 D).

**Fig 5. ADGRE5 (alias CD97) protein expression in cell lines and patient tumor samples.** (A) Proteomic analysis of ADGRE5, BYSL and NPM1 level in BL (i), DLBCLneg (ii) and DLBCLpos (iii) cell lines. (B) Schematic model of largest ADGRE5 isoform EGF (1,2,3,4,5), Arg-Gly-Asp integrin-binding motif (RGD), GPCR-autoproteolysis-inducing domain (GAIN), epidermal growth factor domain (EGF), and nucleoside position of potential N-glycosylation sites are indicated. (C) Western Blot analysis of expression of ADGRE5, BYSL, NMP1 and endogenous control ß-actin and GAPDH in selected cell lines. (D) Immunohistochemical (IHC) staining of ADGRE5 protein in BL, DLBCLpos and DLBCLneg cell lines and FFPE tumor tissue samples. (E) Overview of ADGRE5 staining FFPE tumor tissue samples and supplementary information.
Discussion

Our RNA-Seq and ChIP-Seq data showed a significant overrepresentation of ADGRE5 in BL as compared to DLBCL regardless of the MYC break status of the latter. This finding was confirmed by independent additional target-specific ChIP experiments, RT-PCR and re-analysis of published proteomic data [22] (see Fig 4C1-2 and 5A). For further exploration, we selected ADGRE5 as an interesting candidate gene. ADGRE5 previously designated as CD97 [35] belongs to the adhesion G protein-coupled receptor (GPCR) subfamily E and was the first receptor of this gene family found to be associated with cancer [36]. ADGRE5 is a member of the EGF-TM7 (seven-span transmembrane protein containing epidermal growth factor domains; Fig. 5 B) protein family and is constitutively expressed in granulocytes, monocytes as well as in subsets of T- and B-cells [37-42]. An increased ADGRE5 expression is found in some types of leukemia [43-46]. Interestingly, differential expression of ADGRE5 has also been described for several solid cancers such as lung, thyroid and colorectal carcinomas, indicating a tumor and/or tissue-specific expression pattern [47-49]. Upregulation of ADGRE5 is often observed at the invading tumor front as well as in advanced tumor stages. Furthermore, ADGRE5 presents an unfavorable prognostic factor [50-54]. Depending on the cell type and tumor grade, ADGRE5 protein exists in three isoforms resulting from alternative splicing [55]. ADGRE5 protein is cleaved by self-catalytic proteolysis into a large extracellular subunit, which contains three (EGF1,2,5), four (EGF1,2,3,5) or five (EGF1,2,3,4,5) extracellular N-terminal epidermal growth factor (EGF)-like domains, that are coupled to the seven-span transmembrane subunit (TM7) via an extended spacer region [56-58]. As a surface receptor molecule, ADGRE5 has the ability to bind ligands of the cellular and extracellular matrix, enhances proteolytic activity of matrix metalloproteinases (MMPs) and thus triggers secretion of chemokines [59]. Finally, ADGRE5 interacts with CD55 [60-63], the glycosaminoglycan chondroitin sulfate [64, 65], integrin [66] or CD90 [67] in an isoform-specific manner. Initial functional studies suggest that ADGRE5 is relevant for cell adhesion, migration and invasion [53, 59, 68].

To determine ADGRE5 isotype distribution in aggressive lymphoma, we performed Western blot analyses (Fig 5C) and found that the short isoform (EGF1,2,5) is homogeneously expressed in all cell lines, while the largest isoform (EGF1,2,3,4,5) was preferentially present in BL cell lines. This is a very striking finding, as the EGF-like repeat 4, which has been reported to interact with chondroitin sulfate, is only found in the largest ADGRE5 isoform [64-66]. The interaction of chondroitin sulfate and ADGRE5 (EGF4) mediates cell adhesion and angiogenesis and plays an important role in the interaction of activated T-cells, dendritic cells and macrophages. This observation fits very nicely to features of BL, especially angiogenesis and macrophage attraction.

In order to determine the expression profile of ADGRE5 in primary patient specimens, we performed immunohistochemical staining (IHC). Figure 5D illustrates the higher expression of ADGRE5 on the cell surface of BL cell lines and primary BL patient specimens. In contrast, primary tissue specimens obtained from DLBCL patients and cell lines largely lacks ADGRE5 protein expression irrespectively of the presence of MYC breaks. (Additional IHC staining results are shown in Supplementary S4 Fig). Table 5E summarizes the ADGRE5 IHC results
obtained for 38 patients suffering from BL and DLBCL, respectively. Overall, ADGRE5 is significantly more frequently positive in BL patients (88%) as compared to DLBCLs patients that were mainly negative (80%). Thus, the data derived from our cell line experiments are nicely reflected in primary patient specimens.

**Conclusion**

Here we describe the impact of MYC in three types of aggressive B-cell non-Hodgkin lymphomas: BL and DLBCL with and without MYC break (DLBCLneg and DLBCLpos, respectively). More MYC-binding sites were found by MYC ChIP-Seq in BL and DLBCLpos as compared to DLBCLneg. Interestingly, MYC was found to be bound to different target genes in BL and DLBCLpos, which is also reflected by their gene expression differences. By combined analyses, *ADGRE5* (*CD97*) was identified as an interesting differentially expressed MYC target gene. Its expression, in particular that of the largest *ADGRE5* isoform (EGF1,2,3,4,5), was significantly higher in BL than DLBCL. This observation was confirmed by additional validation experiments and by immunohistochemistry of primary FFPE patient samples. Based on the reported function of the EGF4 repeat as a receptor for chondroitin sulfate, we hypothesize that this might contribute to some peculiar features of BL, namely macrophage attraction and angiogenesis. In addition, we suggest ADGRE5 as a marker to discriminate between BL and DLBCL (regardless of the presence of MYC breaks) in patient stratification.

**Abbreviations**

- **BL**: Burkitt lymphoma; **B-NHL**: B-cell non-Hodgkin lymphomas; **ChIP**: Chromatin immunoprecipitation; **DLBCL**: diffuse large B-cell lymphoma; **DLBCLneg**: DLBCL without MYC break; **DLBCLpos**: DLBCL with MYC break; **EGF-TM7**: seven-span transmembrane protein containing epidermal growth factor domains; **FDR**: False Discovery Rate; **GPCR**: G protein-coupled receptor; **IDR**: Irreproducible Discovery Rate; **ORI**: Origin of Replication; **TSS**: Transcription start site

**Declarations**

**Ethics approval and consent to participate**

The use of human primary tissue samples for the Immunohistostaining was approved by the Institutional Review Board of the Charité – Berlin (EA4/104/11).

**Consent for publication**

Not applicable.

**Availability of data and material**

All data generated or analyzed during this study are included in this published article and its supplementary information files.
Competing interests

The authors declare that they have no competing interests.

Funding

This research was financially supported by the BMBF (Federal Ministry for Education and Research, Germany) in the context of the E:bio Consortium Molecular Mechanisms in Malignant Lymphoma with MYC-Deregulation (MMML-MYC-SYS), grant numbers 0316166H and 0316166G. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors’ contributions

KK, LD, FT, JCE, PS, RS, WG, PJO and MH design and conception the study. KK, EO, NT, EB and PS performed the acquisition of data. Analysis of data: KK, LD, FT, JCE, PS, RS, WG, PJO, MH; interpretation of results: KK, LD, MH; writing and revising of manuscript: KK, FT, JCE, PS, JR, RS, WG, PJO and MH. All authors read and approved the final manuscript (KK, LD, EO, NT, EB, FT, JCE, PS, JR, RS, WG, PJO and MH).

Acknowledgements

The authors would like to thank Edda von der Wall, Anke Sommerfeld and Hedwig Lammert for their excellent technical assistance and Dr. Hilmar Quentmeier for providing U-2932 subclones. Special thanks go to Ulrike Krüger for Sequencing at the NGS core facility of the Berlin-Brandenburg Center for Regenerative Therapies.

List of additional files

Fig 1. MYC expression in DLBCL and BL cell lines.

Fig 2. Differential binding patterns obtained by ChIP-Seq experiments.

Fig 3. Selected differentially bound genes derived from MYC and H3K4me3 ChIP-Seq experiments.

Fig 4. RNA-Seq and validation of selected targets.

Fig 5. ADGRE5 (alias CD97) protein expression in cell lines and patient tumor samples.

S1 Table. Compilation of antibodies used for Chromatin Immunoprecipitation (ChIP), Western Blot (WB) or Immunohistochemistry (IHC) and list of TaqMan Assays used for RT-PCR analysis.

S2 File. ChIP-Seq data of differentially bound MYC and H3K4 target genes derived from comparison of BL vs. DLBCL, BL vs. DLBCLneg, DLBCLneg vs. DLBCLpos and BL vs. DLBCLpos.

S3 File. RNA-Seq data of differentially expressed genes derived from comparison of BL vs. DLBCL, BL vs. DLBCLneg, DLBCLneg vs. DLBCLpos and BL vs. DLBCLpos.

S4 Fig. ADGRE5 IHC staining of cell line and patient tissue samples.

Supporting information files
S1 Table. Compilation of antibodies used for Chromatin Immunoprecipitation (ChIP), Western Blot (WB) or Immunohistochemistry (IHC) and list of TaqMan Assays used for RT-PCR analysis. *indicated endogenous control.

S2 File. ChIP-Seq data of differentially bound MYC and H3K4 target genes derived from comparison of BL vs. DLBCL, BL vs. DLBCLneg, DLBCLneg vs. DLBCLpos and BL vs. DLBCLpos.

S3 File. RNA-Seq data of differentially expressed genes derived from comparison of BL vs. DLBCL, BL vs. DLBCLneg, DLBCLneg vs. DLBCLpos and BL vs. DLBCLpos.

S4 Fig. ADGRE5 IHC staining of cell line and patient tissue samples. (A) Burkitt Lymphoma (BL): Cell lines (1-5: Blue-1 Bl-41, BL-2, DG-75, CA-46) and primary tumor tissues (7-14) are mainly ADGRE5 positive. (B) Diffuse large B cell lymphoma without MYC break (DLBCLneg): Cell lines (1-4: Karpass-422, U2932-R1, HT, WSU-DLCL2) and primary tumor tissues (5-19). (C) DLBCLpos: Cell lines (1-3: Carnaval, U2932-R2, SU-DHL-10) and primary tumor tissues (4-6). DLBCLpos and DLBCLneg are mainly negative for ADGRE5. Strong positive staining in some tissue sections results from macrophages or T-cells.

References


49. Liu Y, Chen L, Peng SY, Chen ZX, Hoang-Vu C: Role of CD97(stalk) and CD55 as molecular markers for prognosis and therapy of gastric carcinoma patients. J Zhejiang Univ Sci B 2005, 6(9):913-918.


5.3. Manuscript 2

Title: Comprehensive metaboproteomics of Burkitt’s and diffuse large B-cell lymphoma cell lines and primary tumor tissues reveals distinct differences in pyruvate content and metabolism.


Published at J Proteome Res. 2017 Mar 3;16(3):1105-1120.
The original article is online available at: DOI: https://doi.org/10.1021/acs.jproteome.6b00164
PMID: 28161958

5.3.1. Synopsis

The distinction of BL and DLBCL can be difficult at the transcriptional, immunophenotypical and genomic level. To contribute to this distinction and to obtain a better understanding of the different biological mechanisms between both subtypes, we developed a metabolomic and proteomic approach to study both established cell lines, as well as cryopreserved and formalin-fixed paraffin-embedded (FFPE) tissue sections of BL and DLBCL. BL and DLBCL cell lines were cultured, and metabolites were characterized from cell pellets and from cell culture supernatants. For metabolite analysis, multidimensional nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) were used to identify metabolites, which differ in abundance between BL and DLBCL.

Proteins were isolated from the cell pellets and then subjected to label-free nano LC-SWATH-MS. Combined metaboproteomic analysis was also performed on frozen lymphoma tissue specimens (n=11). In addition, proteomics was applied to formalin-fixed, paraffin-embedded (FFPE) tissue specimens (n=14). NMR analyses revealed that DLCBL cell lines produce and secrete significantly more pyruvic acid compared to BL cell lines. This observation was confirmed by materials from two independent laboratories in which the cell lines were cultivated under slightly different conditions. Investigation of single proteins did not provide a comprehensive view of the complex biological mechanisms of tumorigenesis. Therefore, additional comprehensive proteomic analysis was performed by using nano LC-SWATH-MS.
technology to investigate protein profiles of BL and DLBCL. The metabolic phenotype was mirrored at the protein level leading to numerous differentially expressed proteins in the two disease entities.

The proteomic analysis of the FFPE and frozen BL and DLBCL tissue specimens revealed that with this approach, one could clearly distinguish between BL and DLBCL. The concordance between the cryopreserved and the FFPE samples was excellent. Combined metabolic and proteomic analysis of BL and DLBCL showed that DLBCL cell lines and primary lymphoma tissues produce and secrete more pyruvic acid than BL. Pyruvic acid plays an important role in cellular redox metabolism and is also believed to be an angiogenic factor. Therefore, metabolic and proteomic analysis is a useful tool to discriminate between BL and DLBCL. Furthermore, it helps to identify differential biological processes which could be an explanation of the heterogeneity and clinical behavior observed within these lymphoma entities.

5.3.2. Contribution

| I Conception | Cell culture experiments design (Laboratory 1) | 100% |
| II Execution | Cell culture optimization | 100% |
| | Cell culture vitality measurements (FACS) | 100% |
| | Cell culture sample preparation | 100% |
| | Selection of patient material (frozen & FFPE tissue) | 100% |
| | Preparation & provision of patient material | 95% |
| III Reporting | Cell culture experimental (Lab.1) material/method section | 100% |

5.3.3. Manuscript Schwarzfischer et al.
5.4. Manuscript 3

**Title:** MINCR is a MYC-induced IncRNA able to modulate MYC’s transcriptional network in Burkitt lymphoma cells.


**Journal:** Proc Natl Acad Sci U S A. 2015 Sep 22;112(38):E5261-70.

Published at Proc Natl Acad Sci U S A. 2015 Sep 22;112(38):E5261-70. The original article is online available at: DOI: [https://doi.org/10.1073/pnas.1505753112](https://doi.org/10.1073/pnas.1505753112) PMID: 26351698

**5.4.1. Synopsis**

In this publication, we described the impact of the transcription factor MYC and long non-coding RNAs (IncRNAs) to the cellular transcriptome. Based on RNA-Seq data from two MYC-inducible cell lines and a cohort of 91 B-cell lymphomas, we identified 13 differentially expressed IncRNAs in IG-MYC-positive BL, which are regulated in the same direction by MYC in the model cell lines. One of the IncRNAs was investigated more intensively and designated MYC-induced long noncoding RNA (MINCR). This MINCR showed a strong correlation with MYC expression in MYC-positive lymphomas. To clarify the cellular role of MINCR, a RNA interference (RNAi) experiment followed by differential gene expression analysis was performed. The results revealed significant clustering of cell cycle genes next to genes down-regulated after MINCR knockdown. Interestingly, these cell cycle genes were also enriched for MYC binding sites in their promoter regions. MINCR knockdown led to reduction of this MYC binding to the promoter regions of these genes. This suggests that MINCR can act as a modulator of the MYC transcriptional program and furthermore, leads to a reduction of MYC activity during cell cycle progression. The down-regulation of Aurora kinases A/B, chromatin licensing and DNA replication factor 1 (CDT1) explains the reduction of cellular proliferation observed in the MINCR knockdown model. Therefore, MINCR can
be regarded as a new interactor and potential therapeutic target of MYC regulatory transcription network via controlling the expression of cell cycle genes.

5.4.2. Contribution

<table>
<thead>
<tr>
<th>I Conception</th>
<th>ChIP experimental design</th>
<th>quota</th>
</tr>
</thead>
<tbody>
<tr>
<td>II Execution</td>
<td>ChIP experimental protocol optimization</td>
<td>70%</td>
</tr>
<tr>
<td></td>
<td>ChIP experimental provision</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>ChIP experimental evaluation</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>ChIP experimental interpretation</td>
<td>60%</td>
</tr>
<tr>
<td>III Reporting</td>
<td>ChIP experimental material/method section</td>
<td>80%</td>
</tr>
</tbody>
</table>

5.4.3. Manuscript Doose et al.
6. Discussion

MYC is one of the most described transcription factors which is involved in a wide range of cellular processes such as apoptosis, cell cycle, metabolism, cell growth, adhesion, angiogenesis and differentiation. It regulates the expression of a variety of protein-encoding and non-coding genes by activating or suppressing their respective expression. Due to this global role of MYC in many cellular processes, it is obvious that aberrant regulation of MYC also has far-reaching consequences for the cell fate. This oncogenic character ultimately leads to the fact that MYC is associated with well over 70% of all cancers diseases (Dang, et al., 2009, Dominguez-Sola and Gautier, 2014, Poole and van Riggelen, 2017, Sheiness, et al., 1978). Although MYC is important for the renewal of normal proliferating cells (stem cell compartment and immune cells), many studies have focused on MYC as the target of cancer therapeutics because of its predominant role in tumor development (Chen, et al., 2014, Chen, et al., 2018, Ponzielli, et al., 2005). Inhibition of MYC transcription via bromodomain proteins (eg, BRD2, BRD3, BDR4 and BDRT) using BET (Bromodomain and Extra-Terminal motif) inhibitors is well established in model systems, but a broad clinical application is lacking. BET inhibitors lead to cell growth inhibition especially in BL cell lines (Delmore, et al., 2011, Mertz, et al., 2011). Other inhibitor strategies target MYC mRNA translation (via mTOR, CPEB), MYC stability (via USO28, PLK1, AURKA) or MYC-MAX dimerization (via Mycro1 and Mycro2) (Clausen, et al., 2010, Follis, et al., 2009, Follis, et al., 2008, Hammoudeh, et al., 2009, Huang, et al., 2006, Kiessling, et al., 2006, Mustata, et al., 2009, Park, et al., 2004, Park, et al., 2004, Prochownik and Vogt, 2010). Despite the success in model systems employing direct MYC inhibitors, strategies that do not focus on MYC itself but its target genes are more promising alternatives. However, it is of particular importance to first identify these MYC target genes and then better understand the role of MYC in those respective cancer types.

„Identification of ADGRE5 as discriminating MYC target between Burkitt lymphoma and diffuse large B-cell lymphoma”

(Kleo et al.)

In the publication of Kleo et al. (submitted to BMC Cancer, 2018), we first looked closely at the role of MYC with respect to its target genes in two types of aggressive lymphomas, namely BL and DLBCL. To this end, we identified MYC and H3K4me3
DNA binding sites in BL and DLBCL cell lines via ChIP-Seq experiments. For better characterization of MYC-related subgroups we subdivided the BL and DLBCL cell lines into three groups (MYCpos BL, MYCpos DLBCL and MYCneg DLBCL) as defined by the presence or absence of a MYC translocation. In parallel, the cellular RNA expression was investigated by the help of RNA-Seq experiments. The methods of choice (ChIP-Seq and RNA-Seq) are powerful tools to identify genome-wide genomic MYC binding sites and RNA expressions profiles in the investigated cell lines. ChIP-Seq offers a number of advantages such as high resolution, low background and complete genome coverage (Park, 2009). A critical element using ChIP technology is the quality of the antibody (here antibody against MYC protein), which have influence to the number of enriched MYC target sites in relation to the background (input DNA). Furthermore, it is important to note that a small number of cell lines was used for this work (n=7), as these cell lines were the only ones available, which had the required characteristics for this study.

We identified a large number of genomic MYC binding sites which were different among the investigated molecular lymphoma groups (BL and DLBCLs).

The number of identified MYC DNA binding sites was almost twice in cell lines with MYC translocation (BL, DLBCLpos) as compared to cell lines without MYC translocation (DLBCLneg). This indicates that the number of MYC-binding target sites is related to the increased MYC protein level in cell lines with a chromosomal MYC rearrangement.

Previous studies show that gene profiles transcriptionally regulated by MYC vary in different cell types and show only a little overlap into distinct basic functional groups (such as metabolism, cell-cycle or DNA replication) (Li, et al., 2003, Zeller, et al., 2003). This observation also holds true for B-cell Lymphoma, where significant number of MYC DNA binding sites have been identified to occur in common for all three groups. Further emphasis was also placed on the MYC DNA binding sites that are different between the investigated groups (BL; DLBCLpos and DLBCLneg). The observation that MYC can bind to different target genes in B-cell lymphomas (here BL, DLBCL) implies that MYC alone could not explain the differences in gene expression per se. Other transcription regulatory elements may also have effects, and MYC may function here as a universal enhancer of individual cellular gene expression.

A universal transcriptional enhancer activity of MYC was described by Lie and Nie in 2012 (Lin, et al., 2012, Nie, et al., 2012), and we intended to test this hypothesis in the
B-cell lymphoma entities - BL and DLBCL - as well. The B-cell transcriptome is specific to the particular lymphoma entity and the identification of different MYC-DNA binding pattern results in a diagnostic potential for a better differentiation between the individual lymphoma entities.

The binding of MYC to the promoter region is important, but not sufficient for the transcription of respective genes. The observation that only about half of the identified MYC DNA binding sites match those of H3K4me3 DNA binding sites (indicator for transcription activity) suggests that MYC can act as a transcriptional activator as well as transcriptional repressor. This property of MYC has already been described in other studies (Dominguez-Sola, et al., 2007, Herkert and Eilers, 2010, Kretzner, et al., 1992).

To correlate MYC binding and corresponding transcription, we performed RNA-Seq and determined specific differential expression profiles for the respective lymphoma entities. The number of differently expressed transcripts between MYC-break positive and negative cell lines was much higher when compared to MYC-break positive BL and DLBCL cell lines, where the differences were not all related to MYC binding. Therefore we decided to focus on MYC target genes which are different based on Chip-Seq as well by RNA-Seq between the investigated groups. For validation, we selected 4 MYC target genes (NPM1, LARS, BYSL and FARSA) which are equally expressed in all three groups and 6 MYC targets (ZAP70, ADGRE5, CDK20, GPAM, SMAD1 and TERT) which are characterized by differential expression. For further validation, we focused on ADGRE5 due to its biological function and due to the availability of specific antibodies, which could be applied in further experiments. These experiments, comprising additional cell lines and primary patient tumor samples, confirmed ChIP-Seq and RNA-Seq findings, and demonstrated that ADGRE5 is significantly highly expressed in BL than in DLBCL independent of the MYC break status of the latter.

ADGRE5 (alias CD97) belongs to the adhesion G protein-coupled receptor (GPCR) subfamily E, and has already been found to be related to several cancer entities (Aust, et al., 1997, Hamann, et al., 2015). As a member of the EGF-TM7 (seven-span transmembrane protein containing epidermal growth factor domains) protein family, it is constitutively expressed in granulocytes, monocytes as well as in subsets of T- and B-cells (Ancuta, et al., 2009, Eichler, et al., 1997, Gasz, et al., 2005, Jaspars, et al., 2001, Kop, et al., 2009, Veninga, et al., 2008). An increased ADGRE5 expression is also found in some types of leukemia (Bonardi, et al., 2013, Coustan-Smith, et al., 2011, Maiga, et al., 2016, Mirkowska, et al., 2013). The differential expression of
ADGRE5 between BL and DLBCL is well in line with other observations from solid cancers that ADGRE5 displays a tumor and/or tissue-specific expression pattern (Aust, et al., 2016, Boltze, et al., 2002, Liu, et al., 2005). Furthermore, some publications describe ADGRE5 as an unfavorable prognostic factor (He, et al., 2015, Safaee, et al., 2013, Si, et al., 2010, Steinert, et al., 2002, Wu, et al., 2012). As a surface receptor, ADGRE5 has the ability to bind ligands of the cellular and extracellular matrix, enhance proteolytic activity of matrix metalloproteinases (MMPs), and thus triggers secretion of chemokines (Galle, et al., 2006). Furthermore, initial functional studies suggest that ADGRE5 is relevant for cell adhesion, migration and invasion (Galle, et al., 2006, Kobayashi, et al., 2013, Steinert, et al., 2002). ADGRE5 interacts in an isoform-specific manner with CD55 (Abbott, et al., 2007, Chiu, et al., 2008, Hamann, et al., 1996, Toomey, et al., 2014), the glycosaminoglycan chondroitin sulfate (Kwakkenbos, et al., 2005, Stacey, et al., 2003), integrin (Wang, et al., 2005) or CD90 (Wandel, et al., 2012). Interestingly, we found by Western blotting that ADGRE5 isotype distribution is different in aggressive lymphoma, with the largest isoform (EGF1,2,3,4,5) preferentially present in BL cell lines. It was found that an interaction of chondroitin sulfate and ADGRE5 (via EGF-like repeat 4) mediates cell adhesion and angiogenesis, and plays an important role in the interaction of activated T-cells, dendritic cells and macrophages (Kwakkenbos, et al., 2005, Stacey, et al., 2003, Wang, et al., 2005). This observation fits nicely to features of BL such as (neo-)angiogenesis and macrophage attraction (Pham, et al., 2018, Ribatti, et al., 2013). Finally, the observed higher ADGRE5 expression in BL cell lines was also confirmed in additional BL cell lines and in primary BL patient tumor tissue.

In conclusion, this study demonstrates the different gene binding activity of MYC in different types of aggressive B-cell lymphoma, which might help to explain their different clinical course. The identified differential MYC binding patterns and the resulting expression of MYC target genes may provide new starting points for better diagnostic stratification of lymphoma patients. Furthermore, it might contribute to the development of new cancer therapeutics. Although the genomic and transcriptional state of a lymphoma cell provides fascinating insights into cellular processes, it constitutes only a partial aspect and a snapshot. It does not reflect the complexity and the interaction between individual proteins and the impact of metabolites. Therefore, the work of Schwarzfischer et al. 2017, aims elaborating the metabolome and proteasome of BL and DLBCL.
“Comprehensive metaboproteomics of Burkitt's and diffuse large B-cell lymphoma cell lines and primary tumor tissues reveals distinct differences in pyruvate content and metabolism”  
(Schwarzfischer et al. 2017)

Metabolomic and proteomic methods were used to perform mass spectroscopy in selected cell lines as well as in patient tissues (FFPE and frozen tumor tissue) derived from BL and DLBCL patients (Schwarzfischer, et al., 2017). In mass spectroscopy, ionized molecules are analyzed according to their mass and ion charge ratio. The resulting spectra are specific for the respective molecules, and targets are ultimately identified by comparison to known spectra available in public databases. Thus, mass spectroscopy technology allows a large number of biomolecules (metabolites and proteins) to be identified at the same time, providing a comprehensive insight into the underlying biological processes (Houk, et al., 1980, O'Reilly and Rappsilber, 2018, Zenobi, 2013).

The metabolomic analysis of cell culture supernatant from BL and DLBCL cell lines initially identified pyruvate as a discriminating metabolite. Pyruvate is produced and secreted in significantly higher levels in DLBCL as compared to BL. This interesting observation could be shown not only extracellularly (cell culture supernatant), but also intracellularly in the cell pellet as well as in frozen tumor materials. Pyruvate is an important intermediate in the breakdown of glucose during cellular energy production (glycolysis) (Berg, et al., 2017). It can also be formed by the metabolism of various amino acids such as alanine (transamination), serine (oxidative deamination) or cysteine (desulfuration). Pyruvate itself can then be further metabolized in the cellular metabolic cycle to lactate (lactate metabolism), to L-alanine (amino metabolism), to acetyl-coenzyme A (citrate cycle) or to oxaloacetate (glycogenesis) (Berg, et al., 2017). Furthermore, pyruvate is described as an angiogenesis factor in tumorigenesis (Jung, et al., 2011, Lee, et al., 2001), and is in harmony with the angiogenic pathway and the increased microvessel density (MVD) observed in DLBCLs (Passalidou, et al., 2003).

To identify differences at the protein level beyond pyruvate itself, several representative proteins of the associated KEGG pathways (glycolysis, PPP, oxidative phosphorylation and carbon metabolism) were identified in a subsequent proteomic investigation.
It could be shown that BL express significantly more LDHA (LDHA: lactate dehydrogenase A) which is responsible for the inter-conversion of L-lactate and pyruvate, with concomitant inter-conversion of NADH and NAD+ (Berg, et al., 2017). It is already known that LDHA transactivation is triggered by MYC (Dang, 2012, Shim, et al., 1997), and this could explain the observation that besides pyruvate, lactate is much more frequently accumulated in DLBCL as compared to BL cell lines and frozen tumor tissue.

The inter-conversion of L-Lactate and pyruvate is impaired by the reduced LDHA expression in DLBCLs, and the metabolites (L-Lactate and pyruvate) accumulate intra- and extracellularly (Berg, et al., 2017). Furthermore, the increased concentration of glycolytic enzymes, such as hexokinase 1 (HXK1), glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (also known as G3P), phosphoglycerate kinase 1 (PGK1) and triosephosphate isomerase (TPIS), observed in DLBCL also conclusively explains the increased pyruvate concentration because, such glycolytic enzymes catalyze the conversion of glucose stepwise to pyruvate (Berg, et al., 2017). These and others proteomic differences between BL and DLBCL reflect not only the metabolic phenotype but also could explain the higher proliferation observed in BL. Although a direct link to MYC deregulation has not been inferred, our data suggests that proteins of the one-carbon metabolism pathway like deoxyuridine 5′-triphosphate nucleotide hydrolase (DUT), thymidylate synthase (TYSY), dihydrofolate reductase (DYR), phosphoserine aminotransferase (SERC), serine hydroxymethyltransferase, mitochondrial (GLYM); D-3-phosphoglycerate dehydrogenase (SERA) and nucleoside diphosphate kinase A (NDKA) are more abundant in the BL.

The observation of increased expression of enzymes related to the one-carbon metabolism is in line with published data describing that MYC could activate the serine and glycine synthesis pathway (Shim, et al., 1997, Sun, et al., 2015). This could explain the increased utilization of glucose for the production of glycine in BL whereas, in DLBCL, the pyruvate synthesis is the preferred pathway. Ultimately, the glycine cleavage refuels the one-carbon metabolism which plays an essential role in biosynthesis of lipids, nucleotides and proteins, the maintenance of redox status and the substrates for methylation reactions (Locasale, 2013, Molyneux, et al., 2012, Vazquez, et al., 2013). This can be regarded as an additional argument for the role of MYC in cancers by its impact on specific metabolic pathways (Dang, 2012, Dang, 2013, Stine, et al., 2015).
To conclude, in Schwarzfischer et al. 2018, we investigated metabolic and proteomic differences between BL and DLBCL, and discovered differential intra- and extracellular concentrations of pyruvic acid between BL and DLBCL cell lines and tumor tissue samples. This metabolic phenotype is also reflected by proteomic differences in pyruvate pathway-related enzymes.

Complex cellular physiological and metabolomic processes are widely interconnected, and are controlled by different regulatory mechanisms. Next Generation Sequencing experiments have shown that up to 70% of the human genome is actively transcribed (Carninci, et al., 2005), but only 2% of these are ultimately protein coding (Djebali, et al., 2012, Mattick and Makunin, 2006). The functional roles of this huge amount of non-coding transcripts are still not yet fully understood, but more and more reports suggest their effects on cellular physiology and an abnormal expression in several diseases (Flockhart, et al., 2012, Holdt, et al., 2013, Huarte, et al., 2010, Prensner, et al., 2011, Sauvageau, et al., 2013). One class of non-coding transcripts are represented by long non-coding RNAs (lncRNA), which are involved in regulation of post-transcriptional, epigenetic and gene transcription processes (Ma, et al., 2015, Orom, et al., 2010, Orom and Shiekhattar, 2013, Perkel, 2013).

As a transcription factor, MYC underlies also a tight regulation and feedback control by co-factors and suppressors (Vervoorts, et al., 2003, Yuan, et al., 2009). One of the major ways to control MYC levels in cells is via targeted degradation by the ubiquitin–proteasome system (UPS) (Farrell and Sears, 2014, Popov, et al., 2007) and another regulatory element like lncRNA (Hamilton, et al., 2015, Iaccarino, 2017). However, despite the established role of transcription factor MYC in cancer therapy, little is known about the effects of long non-coding RNAs for the modulation of MYC expression an activity.

“MINCR is a MYC-induced IncRNA able to modulate MYC's transcriptional network in Burkitt lymphoma cells”

(Doose et al. 2015)

In the work from Doose et al 2015, we performed RNA-Seq derived from two MYC-inducible cell lines and a cohort of 91 B-cell lymphomas to identify MYC-regulated IncRNAs. Furthermore, the functional role of a group of differentially expressed IncRNAs with correlation to MYC expression was investigated using RNA interference (RNAi) and Chromatin Immunoprecipitation (ChIP).
RNA-Seq data obtained from two cell lines (P493-6 and hT-RPE-MycER) and germinal center (GC)-derived B-cell lymphomas samples revealed 143 IncRNAs. These IncRNAs were significantly differentially expressed by the activation of MYC in the P493-6 cells and the hT-RPE-MycER cells, respectively. Parallel analysis of 16 BL tumor samples compared to the normal GC B-cells revealed 367 IncRNAs significantly up- and down-regulated in BL, respectively. 13 IncRNAs differentially expressed in IG-MYC-positive BL and regulated by MYC in the inducible cell lines were identified (Doose et. al 2015 / Fig 1C), suggesting that MYC could be either directly or indirectly involved in transcriptional regulation of those RNAs (Doose, et al., 2015). In comparison to published MYC ChIP-Seq data from inducible P493-6 cells (Sabo, et al., 2014) and BL cell lines (Seitz, et al., 2011), we confirm that 7 out of 13 IncRNAs were bound by MYC under MYC-high conditions in the region around the transcriptional start site (TSS). This finding suggests that some of the identified IncRNAs may be directly targeted by MYC. Furthermore, at least 10 out of 13 IncRNAs were either transcribed regions or active promoters based on the analysis of DNA methylation and presence of DNase I hypersensitive sites (DHS) at the TSS site. Among the 13 IncRNA, the IncRNA with highest significance (Ensembl ID: ENSG00000253716) was regarded as a MYC-regulated transcript. LncRNA ENSG00000253716 had all characteristics associated with open chromatin and active transcription. Furthermore, the expression of ENSG00000253716 was reduced after MYC knockdown in BL cell lines, leading to the designation of this transcript as MYC-induced long non-coding RNA (MINCR). The MINCR gene is located on chromosome 8q24.3 and conserved throughout primates, but not in other vertebrates. MINCR also has a RefSeq entry defined as uncharacterized LOC100507316. ENCODE annotates at least six different isoforms transcribed from the MINCR gene locus, with a long isoform (MINCR_L) composed of three exons and all others containing two exons (short isoform: MINCR_S). A qRT-PCR analysis performed using isoform-specific sets of primers confirmed that the short isoform (MINCR_S) is more expressed in BL. Detailed characterizations of MINCR expression suggests that it is enriched in the nuclear fraction and is ubiquitously expressed among different tissues. Analyzing the expression of MINCR in the entire ICGC MMML-Seq cohort shows that the expression of the IncRNA MINCR strongly correlates with MYC expression, not only in BL, but also in other GC-derived MYC-positive lymphomas and other cancer types (Richter, et al., 2012, Richter, et al., 2012). To determine the functional role of MINCR, RNAi
knockdown was performed and the effect of a possible change in cell proliferation/viability was investigated. The data generated provides evidence that MINCR RNAi results in a time-dependent decrease in cellular proliferation, independent of the MYC activation status. However, MINCR knockdown led to a significant reduction in the percentage of cells with a DNA structure corresponding to cells in S phase and G2/M phase. MINCR knockdown had also effects on cellular viability and cell cycle progression. Performing RNA-seq of hT-RPE-MycER cells 48 h after transfection with the two MINCR siRNAs showed no significant changes in expression of the most proximal genes upstream and downstream from MINCR. Overall, our data demonstrates that the reduction of proliferation observed in hT-RPE-MycER cells after MINCR knockdown depends on the down-regulation of a set of genes important for cell cycle progression. Particularly, cell cycle genes coding for Aurora kinase A (AURKA), AURKB, and chromatin licensing and DNA replication factor 1 (CDT1) were also reproducibly down-regulated in BL-2 cells after MINCR knockdown. Additional MYC ChIP experiments indicate that some of MINCR-regulated genes are also MYC targets. Additional knockdown experiments of AURKA, AURKB and CDT1 led to reduction in cellular viability in hT-RPE-MycER cells. Furthermore, the effect was even more pronounced after MYC activation. These data imply that down-regulation of cell cycle genes (like AURKA, AURKB, and CDT1) can contribute to reduced basal proliferation observed after MINCR knockdown. Interestingly, AURKA and AURKB, just like CDT1, have already been shown to play a crucial role in MYC induced oncogenic transformation and lymphoma development (Seo, et al., 2005, Yang, et al., 2010).

To conclude, the oncogenic transcription factor MYC is known to play a crucial role in the pathogenesis of many neoplasms by controlling the expression of many genes. The role of IncRNAs during MYC-induced oncogenic transformation was investigated in Doose et al. 2015, where we identified a new MYC-regulated IncRNA named MYC-induced long noncoding RNA (MINCR). MINCR is functional, and controls cell cycle progression by influencing the expression of MYC-regulated cell cycle genes. Therefore, MINCR is a novel player in the MYC transcriptional network, and represents a potential therapeutic target in the fight against malignant lymphoma and other cancers that rely on MYC expression.
7. Conclusion and Perspectives

MYC is a common oncogene identified as affected by chromosomal translocation in BL and in a subgroup of DLBCL (Ott, et al., 2013). The oncogenic events leading to aberrant MYC expression is usually associated with an aggressive clinical behavior in B-cell neoplasms. A MYC translocation – alone or accompanied with BCL2 and/or BCL6 translocations - confers an increase in aggressive behavior (Kramer, et al., 1998). However, MYC chromosomal translocation is also present at low frequencies in healthy individuals (Janz, et al., 2003), and therefore it cannot alone be the reason for lymphomagenesis.

Although BL and DLBCL are well-defined lymphoma entities, molecular-pathological studies have identified several subgroups with intermediate features, which are often associated with an aggressive cancer behavior. There are still challenges for routine diagnostics to identify these subgroups by immunohistochemistry or genetic analysis. Gene profiles transcriptionally regulated by MYC vary in different cell types and show only a little overlap between distinct functional groups (such as metabolism, cell-cycle or DNA replication) (Li, et al., 2003, Zeller, et al., 2003). A potential explanation for this differential MYC-related transcriptional activity is given by studies from Ni et al 2012 and Lin et al 2012. These studies suggest that MYC functions as an enhancer of generally transcribed genes in the cell (Lin, et al., 2012, Nie, et al., 2012). Therefore, the different MYC target gene signatures, observed between DLBCL and BL, may also reflect the different gene expression profiles, observed due to the cell origin in different germinal center compartments (Alizadeh, et al., 2000, Hans, et al., 2004, Hummel, et al., 2006, Victora, et al., 2012).

The present work compiles the data derived from three studies (Kleo et al., Schwarzfischer et al. and Doose et al.) to investigate various aspects of MYC-driven aggressive B-cell lymphomas. Several approaches such as ChIP-Seq, RNA-Seq, mass spectroscopy and RNA interference were employed to disclose differences between BL and DLBCLs, and to gain insights into their complex cell physiology. Besides the identification of differential genomic and transcriptional patterns we also discovered different metabolic and proteomic profiles, which expand our knowledge about the complex cellular physiological processes of MYC-driven B-cell lymphomas. Our already published and submitted data identified new differential target molecules and metabolites (such as ADGRE5 and pyruvate), which possess the potential to
support precise diagnostic distinction among the investigated lymphoma types. Furthermore, the impact of lncRNAs for MYC-associated B-cell lymphoma was investigated, leading to the identification of MINCR as a novel modifier of the MYC transcriptional network. The generated data sets represent an excellent basis for additional investigations to further explore aggressive B-cell lymphoma and to support the development of alternative therapeutic strategies to combat MYC-driven aggressive B-cell lymphomas.
8. Acknowledgment / Danksagung

An erster Stelle gilt mein Dank meinem Doktorvater Herr Prof. Dr. Michael Hummel für seine wissenschaftliche und konstruktive Unterstützung während der gesamten Bearbeitungsphase meiner Dissertation.

Besonders danke ich auch Dr. rer. nat. Lora Dimitrova, Elisabeth Oker und Nancy Tomaszewski für die technische und praktische Unterstützung und vielen lieben Worte während der gesamten Bearbeitungszeit meiner Dissertation.

Ein besonderer Dank gilt meinen Eltern Ulrich und Ilona und meinen Geschwistern Simone und Mario, die mich auf meinem akademischen Weg stets begleitet haben und ihn letztlich erst ermöglichten.

Danken möchte ich außerdem meinen Freunden Arthur Fandrich, Carolin Nitzold, David Ritzmann, Gabriele Weiss, Manja Stelter und Stev Töpfer, die mich in den vergangenen Jahren und darüber hinaus begleitet und moralisch unterstützt haben.

Allen Mitarbeitern der Arbeitsgruppe von Prof. Hummel, sowie allen Beteiligten meiner Arbeit insbesondere den Kooperationspartnern bin ich sehr dankbar für die gute und zahlreiche Unterstützung sowie die konstruktive und angenehme Zusammenarbeit.

“Ich habe gelernt, dass der Weg des Fortschritts weder kurz noch unbeschwerlich ist.“

Zitat: Marie Curie 1867 - 1934
9. References


Atra, A., M. Gerrard, R. Hobson, J. D. Imeson, S. Ashley and C. R. Pinkerton. (1998), 'Improved Cure Rate in Children with B-Cell Acute Lymphoblastic Leukaemia (B-All) and Stage Iv B-Cell Non-Hodgkin's Lymphoma (B-Nhl)--Results of the Ukccsg 9003 Protocol', Br J Cancer Vol. 77, No. 12, pp. 2281-5.


Lennert, K. and H. Stein. (1978), 'Personal Points of View on the Kiel Classification', *Recent Results Cancer Res* Vol. 64, pp. 31-7.


Diffuse Large B-Cell Lymphoma Identifies Robust Subtypes Including One Characterized by Host Inflammatory Response', *Blood* Vol. 105, No. 5, pp. 1851-61.


of Ebv in the Pathogenesis of Burkitt's Lymphoma: An Italian Hospital Based Survey', *Infect Agent Cancer* Vol. 9, No. 1, pp. 34.


Poole, C. J. and J. van Riggelen. (2017), 'Myc-Master Regulator of the Cancer Epigenome and Transcriptome', *Genes (Basel)* Vol. 8, No. 5.


Richter, J., M. Schlesner, S. Hoffmann, M. Kreuz, E. Leich, B. Burkhardt, M. Rosolowski, O.
Ammerpohl, R. Wagener, S. H. Bernhart, D. Lenze, M. Szczepanowski, M. Paulsen, S. Lipinski,
Korbel, D. Kube, D. Langenberger, C. Lawerenz, J. Lisfeld, K. Meyer, S. Picelli, J. Pischimarov,
B. Radlwimmer, T. Rausch, M. Rohde, M. Schilhabel, R. Scholtysik, R. Spang, H. Trautmann, T.
Zenz, A. Borkhardt, H. G. Drexler, P. Moller, R. A. MacLeod, C. Pott, S. Schreiber, L. Trumper,
M. Loeffler, P. F. Stadler, P. Lichter, R. Eils, R. Kuppers, M. Hummel, W. Klapper, P. Rosenstiel,
A. Rosenwald, B. Brors, R. Siebert and I. M. S. Project. (2012), 'Recurrent Mutation of the Id3
Gene in Burkitt Lymphoma Identified by Integrated Genome, Exome and Transcriptome

Richter, J., M. Schlesner, S. Hoffmann, M. Kreuz, E. Leich, B. Burkhardt, M. Rosolowski, O.
Ammerpohl, R. Wagener, S. H. Bernhart, D. Lenze, M. Szczepanowski, M. Paulsen, S. Lipinski,
Korbel, D. Kube, D. Langenberger, C. Lawerenz, J. Lisfeld, K. Meyer, S. Picelli, J. Pischimarov,
B. Radlwimmer, T. Rausch, M. Rohde, M. Schilhabel, R. Scholtysik, R. Spang, H. Trautmann, T.
Zenz, A. Borkhardt, H. G. Drexler, P. Moller, R. A. F. MacLeod, C. Pott, S. Schreiber, L. Trumper,
M. Loeffer, P. F. Stadler, P. Lichter, R. Eils, R. Kuppers, M. Hummel, W. Klapper, P. Rosenstiel,
A. Rosenwald, B. Brors, R. Siebert and I. M. S. Project. (2012), 'Sequencing of Pediatric Burkitt Lymphoma within the


Robertson, E. S. (2012), Burkitt’s Lymphoma, Springer New York.

Chain (Cd25, Tac) Expression Defines a Crucial Stage in Pre-B Cell Development', Int Immunol
Vol. 6, No. 8, pp. 1257-64.

Muller-Hermelink, E. B. Smeland, L. M. Staudt and L. L. M. P. Pr. (2002), 'The Use of
Molecular Profiling to Predict Survival after Chemotherapy for Diffuse Large-B-Cell

Verrecchia, C. Tonelli, G. Faga, V. Bianchi, A. Ronchi, D. Low, H. Muller, E. Guccione, S.
Campaner and B. Amati. (2014), 'Selective Transcriptional Regulation by Myc in Cellular

Safaee, M., A. J. Clark, M. C. Oh, M. E. Ivan, O. Bloch, G. Kaur, M. Z. Sun, J. M. Kim, T. Oh, M. S. Berger
and A. T. Parsa. (2013), 'Overexpression of Cd97 Confers an Invasive Phenotype in
Glioblastoma Cells and Is Associated with Decreased Survival of Glioblastoma Patients', PLoS
One Vol. 8, No. 4.

Sander, C. A., P. Kind, P. Kaudewitz, M. Raffeld and E. S. Jaffe. (1997), 'The Revised European-
American Classification of Lymphoid Neoplasms (Real): A New Perspective for the


10. Abbreviations

aa  aminoacid
Ab  antibody
ADGRE5  adhesion G protein-coupled receptor E5
BCR  B-cell receptor
bHLH  basic helix-loop-helix
BL  Burkitt lymphoma
B-NHL  B-Cell Non-Hodgkin lymphoma
bp  base pair
BYSL  Bystin Like
CD  cluster of differentiation
CDK20  Cyclin Dependent Kinase 20
CDRs  complementarity-determining regions
ChIP  Chromatin-Immunoprecipitation
GPAM  Glycerol-3-Phosphate Acyltransferase, Mitochondrial
DLBCL  diffuse large B-cell lymphoma
DLBCLneg  diffuse large B-cell lymphoma negative for MYC rearrangement
DLBCLpos  diffuse large B-cell lymphoma positive for MYC rearrangement
DNA  deoxyribonucleic acid
FARSA  Phenylalanyl-TRNA Synthetase Subunit Alpha
FFPE  formalin-fixed paraffin-embedded
FISH  fluorescent in situ hybridization
H3K4me3  histone H3 containing the trimethylated lysine 4
HL  Hodgkin lymphoma
HSCs  hematopoietic stem cells
Ig  immunoglobulin
IHC  immunohistochemistry
LC-SWATH-MS  liquid chromatography – sequential window acquisition of all theoretical fragment ion spectra – mass spectrometry
lncRNA  long non-coding RNA
LARS  Leucyl-TRNA Synthetase
LZ  leucine-zipper
Max  MYC-associated X-factor protein
MHC  major histocompatibility complex
MINCR  Myc-Induced long non-coding RNA
NHL  Non-Hodgkin lymphoma
NLS  nucleus localization signal
NPM1  Nucleophosmin 1
RNA  ribonucleic acid
SMAD1  SMAD Family Member 1
TCR  T-cell receptor
TD  T-cell dependent
TERT  Telomerase Reverse Transcriptase
TI  T-cell independent
WHO  World Health Organization
ZAP70  Zeta Chain Of T Cell Receptor Associated Protein Kinase 70
11. Selbständigkeitserklärung (independence declaration)


Karsten Kleo

Berlin, 28.02.2019
12. Curriculum vitae

For reasons of privacy protection, personal data are not included in the electronic version of the thesis.
Scientific contributions
Publication

2019

2018
Kleo K, Lora Dimitrova, Elisabeth Oker, Nancy Tomaszewski, Erika Berg, Franziska Taruttis, Julia C. Engelmann, Philipp Schwarzfischer, Jörg Reinders, Rainer Spang, Wolfram Gronwald, Peter J. Oefner, Michael Hummel (2018), 'Identification of ADGRE5 as discriminating MYC target between Burkitt lymphoma and diffuse large B-cell lymphoma' – *Submitted to BMC Cancer*

2017

2015

2012

2011
2018
DGVS & DGAV Kongress - Viszeralmedizin 2018 (München)
**Title:** Her2 als prädiktiver Marker in der neoadjuvanten Therapie bei Adenokarzinomen des Magens und Ösophagus. (talk)

102. Jahrestagung der Deutschen Gesellschaft für Pathologie e.V. (Berlin) **Title:** Immunprofiling mittels NanoString nCounter Technologie zur Identifizierung prädiktiver Marker bei Adenokarzinomen des ösophagogastralen Übergangs und Magens. (poster)

2017
101. Jahrestagung der Deutschen Gesellschaft für Pathologie e.V. (Erlangen) **Title:** Metaboproteomics of Burkitt and diffuse large B-cell lymphoma cell lines and primary tumor tissues. (poster)

2012
Biosensorweltkongress 2012 (Mexico) **Title:** Immunological detection of *Francisella tularensis* using QCM. (poster)

IMCS 2012 - The 14th International Meeting on Chemical Sensors (Nürnberg) **Title:** Fast detection of pathogenic bacteria by using different sensor techniques. (poster)

2011
7. Deutsches BioSensor Symposium (Heiligenstadt) **Title:** Detection of vaccinia virus DNA by Quartz Crystal Microbalance. (poster)

The 10th workshop on biosensors and bioanalytical microtechniques (Weimar) **Title:** Indirect detection of vaccinia virus by DNA using QCM. (poster)

2nd International Conference on Bio-Sensing Technology (Amsterdam) **Title:** Label free detection of DNA for virus analysis. (poster)

10. Dresdner Sensor Symposium (Dresden) **Title:** Label-freier Nachweis von DNA für die Virusedetektion. (talk)

- end -