Aus dem Institut für Pathologie,
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Non-coding RNA expression profiling of B-cell malignant (non-Hodgkin) lymphomas

zur Erlangung des akademischen Grades

Doctor rerum medicinalium (Dr. rer. medic.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Kebria Hezaveh

aus Iran-Arak

Datum der Promotion: 26.02.2016

Table of Contents

Figure Directory

Table Directory

Abstract

MicroRNAs are small (20-23nt in length), non-coding and highly conserved molecules, which are involved in several regulatory processes like cell growth, proliferation, differentiation, immune response and apoptosis, and play important roles in several diseases, including cancers like lymphoma. Germinal center (GC) derived B-cell lymphomas, including Burkitt lymphoma (BL), diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL), are the most frequent malignant lymphomas. Although clear distinctions on histologic and genetic grounds exist, there are also a large number of cases with intermediate features, not unequivocally attributable to one of these entities. The ICGC-MMML-Seq Consortium aims at fully characterizing a total of 250 GC derived B-cell lymphomas. Here we generated miRNA profiles from 56 patient samples including BL, DLBCL and FL using Illumina technology. Over the past decade, many studies have attempted to distinguish lymphoma subtypes using miRNAs profiling. However, available data is preliminary, as published profiles are not derived from large sample collections and also originate mostly from PCR-based approaches and microarrays. Yet, only sequencing-based approaches allow for an unbiased analysis and the discovery of novel miRNAs and small RNA classes. Our initial differential expression analyses comparing BL against non-BL showed eight miRNAs to be differentially expressed. In addition, we analyzed miRNA deregulation between non-BL subtypes including FL and DLBCL and also compared each of them separately to BL. A signature of 87, 98 and 108 miRNAs was obtained that differentiated FL from DLBCL, BL from DLBCL and BL from FL, respectively. Mutational analysis identified 17 mutations in 12 patients corresponding to eight distinct miRNAs. Among eight mutated miRNAs, *miR-142* with a total of seven different mutations in six patients was the most frequently mutated miRNA. Among predicted novel miRNAs, we successfully validated four candidates by Northern Blot experiments and we then tried to uncover their function in lymphomagenesis by performing further functional studies.

In addition, our data gave insight into the role of lncRNAs in GCB-lymphomas. We found the differential expression as well as differential methylation pattern of the lncRNA AP000251 in GCB-lymphoma subtypes. To investigate which miRNA-target pairs are more likely to display regulation in lymphoma, we performed AGO2-PAR-CLIP in four lymphoma (Raji, NAMALWA, SU-DHL-4 and SU-DHL-6) cell lines. We identified putative miRNA-targets

from each PAR-CLIP library which might represent a helpful tool to find potential therapeutic targets and prognostic markers in lymphoma.

Zusammenfassung

MicroRNAs sind kurze (20-23 Nukleotide lange), nicht-kodierende und hoch konservierte Moleküle, die an vielen regulatorischen Prozessen wie z.B. Zellwachstum, Proliferation, Differenzierung, Immunantwort und Apoptose beteiligt sind. Darüber hinaus spielen sie eine wichtige Rolle in zahlreichen Krankheiten, u.a. in verschiedenen Krebsarten wie z.B. Lymphomen.

Keimzentrums-B-Zell-Lymphome, zu denen Burkitt-Lymphome (BL), diffus großzellige B-Zell-Lymphome (DLBCL) und follikuläre Lymphome (FL) zählen, stellen die häufigsten aggressiven Lymphome dar. Und auch wenn prinzipiell eindeutige histologische und genetische Unterscheidungskriterien beschrieben wurden, existieren in der täglichen Praxis doch zahlreiche Fälle mit intermediärem Phänotyp, die sich nicht eindeutig in eine der o.g. Kategorien einordnen lassen.

Das ICGC-MMML-Seq Konsortium hat sich die vollständige Charakterisierung von 250 Keimzentrums-B-Zell-Lymphomen zum Ziel gesetzt. Im Rahmen dieser Arbeit wurden miRNA Profile von 56 Patientenproben (zusammengesetzt aus BL, DLBCL und FL) mit Illumina Technology sequenziert.

In den letzten zehn Jahren haben verschiedene Studien versucht, die Lymphomsubtypen an Hand von miRNA Profilen zu unterscheiden. Die in diesem Rahmen generierten Daten sind jedoch insoweit nur als vorläufig zu betrachten, da sie zum einen nicht aus großen Patientenkollektiven rekrutiert wurden und zum anderen auf PCR-basierenden Methoden zurückgehen. Nur eine auf den neuen Sequenziertechniken beruhende Herangehensweise erlaubt jedoch die unverzerrte Analyse und die Möglichkeit, sowohl neue miRNAs als auch neue Klassen kleiner RNAs zu entdecken.

Im ersten Schritt wurden Burkitt-Lymphome mit nicht-Burkitt-Lymphome verglichen, dort zeigte sich die differentielle Expression von acht miRNAs. Zusätzlich wurde die miRNA Deregulation zwischen FL und DLBCL sowie auch jeweils individuell gegen BL analysiert. So konnten Expressionssignaturen beschrieben werden, die FL von DLBCL (87 miRNAs), BL von DLBCL (98 miRNAs) und BL von FL (108 miRNAs) unterscheiden.

In der Mutationsanalyse basierend auf 12 Patientendatensätzen wurden 17 Mutationen, die acht verschiedene miRNAs betrafen, gefunden. Am häufigsten war miRNA-142 mit insgesamt sieben Mutationen in sechs Patienten betroffen.

Unter allen bioinformatisch vorhergesagten neuen miRNAs konnten vier Kandidaten durch Northern Blot Experimente validiert werden. Anschließend wurde damit begonnen, deren Rolle in der Lymphomentstehung durch funktionelle Studien näher zu charakterisieren.

Darüber hinaus gewährend die Sequenzierdaten Einblicke in die Rolle von langen nicht-kodierenden RNAs (lncRNAs). So konnte z.B. ein differentielles Methylierungs- und Expressionsmuster der lncRNA AP000251 beschrieben werden.

Um zu klären, welche miRNA-mRNA Regulationspaare eine relevante Aufgabe in Lymphomen übernehmen, wurden AGO2-PAR-CLIP Experimente in vier Lymphomzelllinien (Raji, NAMALWA, SU-DHL-4, SU-DHL-6) durchgeführt. Die identifizierten mRNAs, die eine validierte Regulation durch miRNAs zeigen, können nun als Ausgangspunkt dienen, um mögliche therapeutisch nutzbare Strukturen sowie prognostische Marker in Lymphomen zu beschreiben.

1 Introduction

1.1 Lymphoma

Lymphomas represent a heterogeneous group of malignancies which are characterized by neoplastic transformation of lymphocytes at different stages of lymphocyte development. Due to the heterogeneity of lymphomas many efforts have been made to classify lymphomas. The first systematic attempts to define distinct subtypes of lymphoma stem from the 1930s. Since then, sequential revisions in lymphoma classification schemes have been proposed with parallel developments taking place in genetics and molecular biology that improved our understanding of the molecular underpinnings and their assumed cells of origin (Sheehan and Rappaport 1970). The latest accepted definition of distinct subtypes of lymphoma is from the World Health Organization (WHO) classification, published in 2008 (Swerdlow, Campo et al. 2008) with the most recent update being released in 2008 (Vardiman, Thiele et al. 2009). According to the WHO classification system, lymphomas comprise more than 50 distinct subtypes (Swerdlow, Campo et al. 2008). This classification is based on the cell type of origin, histology, immunophenotype, clinical characteristics and genetic aberrations. Traditionally, Hodgkin lymphoma and non-Hodgkin lymphoma (NHL) are considered as the two main types of lymphoma, which account for 10% and 90% of all cases, respectively. They have distinct histological, immunophenotype and genetic features, different molecular pathogenesis, sites of primary manifestation of disease, distinct sites of primary manifestation, etc (Swerdlow, Campo et al. 2008).

1.1.1 Non-Hodgkin lymphoma (NHL)

NHL represents an etiologically and clinically heterogeneous group of lymphoid malignancies with around 50 entities recognized in the most recent WHO classification (Campo, Swerdlow et al. 2011). NHLs consist of three distinct lymphocyte types, B, T and NK cells; however most of them are of B-cell origin. B-cell NHLs arise at different stages of B-cell development and account for 85% of NHLs worldwide and 4% of all newly diagnosed cancers each year

(Muller, Ihorst et al. 2005, Alexander, Mink et al. 2007). The incidence of histopathologic subtypes of NHLs is highly age-dependent. Burkitt lymphoma and diffuse large B-cell (DLBCL) are the most common subtypes in children and young adults. Small lymphocytic and follicular lymphomas (FL) are most common in patients over the age of 60. The majority of NHLs occur in patients between 60 and 70 years old with slight male-to-female predominance. The most common types of NHL are diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL), which together account for around 60% of all NHLs (Anderson, Armitage et al. 1998, Muller, Ihorst et al. 2005, Swerdlow, Campo et al. 2008).

1.1.1.1 Diffuse large B-cell lymphoma (DLBCL)

The term DLBCL is applied to a group of clinically aggressive lymphomas that collectively represent 25% to 35% of all NHLs (Cultrera and Dalia 2012). DLBCL is composed of multiple clinical subtypes and morphologic patterns. Individual entities of DLBCL have similar clinical behavior, as well as similar treatment regimens. DLBCLs without specific clinical or pathologic features fall into the biologically heterogeneous category of diffuse large B-cell lymphoma, not otherwise specified (DLBCL-NOS). The median age of DLBCL-NOS is between 6th and 7th decade with a slight male predominance (Armitage and Weisenburger 1998, Said 2013). The other entities are T- cell/histiocyte rich large B-cell lymphoma, DLBCL associated with chronic inflammation, Epstein-Barr virus (EBV)+ DLBCL of the elderly, lymphomatoid granulomatosis, primary mediastinal large B-cell lymphoma, intravascular large B-cell lymphoma, primary cutaneous DLBCL, leg type, ALK+ large B-cell lymphoma, plasmablastic lymphoma, and primary effusion lymphoma (1997, Jaffe and Pittaluga 2011) some subtypes of DLBCL-NOS were categorized into one of three accepted variants, i.e., activated B-cell-like, germinal center B-cell-like and primary mediastinal large B-cell lymphomas; these types show differences in genetic alterations, molecular signaling pathways and clinical outcomes. Because gene expression profiling is not practical in routine use for DLBCL phenotypic classification, alternative techniques such as immunohistochemistry-based methods are required, though their accuracy is currently not very good (Carbone, Gloghini et al. 2014).

1.1.1.1.1 Clinical aspects

In general, DLBCL is a fast growing and aggressive form of NHL. DLBCLs usually arise as de novo malignancies, but in some cases they also present as a transformation of a previously

existing indolent lymphoma. All indolent lymphomas such as follicular lymphoma (FL), chronic lymphocytic leukemia (CLL) and marginal zone B cell lymphoma carry the risk of transformation, but the transformation is best investigated in FL (Bastion, Sebban et al. 1997, Montoto, Davies et al. 2007).

In some cases lymphomas initially occur in extra nodal sites such as the bone marrow, lung, kidneys and liver and usually present at an early stage of disease, but sometimes extranodal site involvement can also occur in primary nodal lymphoma; patients are then more likely to present with stage II disease (Lopez-Guillermo, Colomo et al. 2005).

The international prognostic index (IPI) is a commonly used tool to classify patients into risk groups using prognostic factors such as age, high serum lactate dehydrogenase (LDH), Ann Arbor stage II-IV, poor overall performance status, patient's performance status greater than 1 based on the eastern cooperative oncology group (ECOG) criteria and more than one extra-nodal site (1993, Ferreri, Blay et al. 2003).

1.1.1.1.2 Treatment

In 1976, cyclophosphamide, doxorubicin, vincristine and prednisolone (CHOP) was first described as the standard treatment for patients with DLBCLs. Since the early 2000s CHOP was improved by the addition of the CD20 monoclonal antibody rituximab (R-CHOP) and established as the current standard of care (Habermann, Weller et al. 2006).

1.1.1.1.3 Immunophenotype and genetics

In DLBCL tumor cells, typically one or more B-cell-associated markers are expressed such as CD19, CD20, CD22 and CD79a and approximately in 5-10% of de novo DLBCL CD5 (usually a T-cell-associated antigen), as a marker of T cells, is coexpressed (Falini, Pileri et al. 1990, Pileri, Ascani et al. 2000). Recent gene expression profiling analysis has confirmed at least three major distinct subgroups of DLBCL: GC-like DLBCL, non-GC-like DLBCL and type 3 or intermediate. Germinal center B-cell- like (GCB) DLBCLs present with the expression pattern of normal GC B-cells (GC signature genes) while activated B-cell-like (ABC) DLBCLs express gene characteristic of in vitro activated peripheral blood B-cells (activated B-cell signature genes) and some genes of plasma cell origin, suggesting a post germinal center origin. Type 3 does not express high levels of either GCB or ABC gene sets (Schneider, Pasqualucci et al. 2011).

Molecular cytogenetic studies of de novo DLBCLs have identified chromosomal translocations including C-MYC, BCL2 and BCL6 (Pileri, Ascani et al. 2000, Akyurek, Uner et al. 2012, Ott, Rosenwald et al. 2013), the latter being the most common translocation detected in 20 to 40% of all DLBCLs (Lo Coco, Ye et al. 1994). BCL2 translocations and MYC rearrangements occur in 20-30% and 10% of DLBCL cases, respectively (Weiss, Warnke et al. 1987, Min, Xia et al. 2005). GCB-DLBCL and non-GCB-DLBCL carry distinct chromosomal aberrations. For instance, the t(14; 18)(q32; q21) translocation involving the BCL2 gene and amplification of c-rel locus is strongly associated with GCB-DLBCL. On the other hand, high expression of nuclear factor kB (NF-kB) target genes has been detected in non-GCB-DLBCL (but not in GC-like DLBCL). The same is true for trisomy 3, gain of 18q21-q22, loss of 6q21-22 and deletion of the INK4A-ARF locus on chromosome 9 (Bea, Zettl et al. 2005, Lenz, Wright et al. 2008). Somatic hypermutation (SHM) of the immunoglobulin genes that are activated by activation-induced cytidine deaminase (AID) is related to non-GCB-DLBCL. Ongoing SHM is only observed in GCB-DLBCL, not in ABC-DLBCL, but the total amount of SHM is higher in ABC-DLBCL (Lossos, Alizadeh et al. 2000). In general, ABC-DLBCL has a poor prognosis compared to GCB-DLBCL (Alizadeh, Eisen et al. 2000).

1.1.1.2 Follicular lymphoma (FL)

FL is the most commonly occurring form of indolent B-cell NHL, and it comprises about 20-25% of all lymphomas. FLs arise following the malignant transformation of cells derived from germinal follicles of peripheral lymphoid tissues. These cells are called centrocytes (small cells) and centroblasts (large cells). There is much heterogeneity within the FL category. WHO divided FLs into 3 separate grades, defined by the presence and number of centroblasts within neoplastic follicles. In grades 1–2 the proportion of small cells is predominant. Clinically, they have an indolent course and are considered 'low grade', whereas grade 3 FLs show a high proportion of large cells and exhibit more aggressive clinical behaviour (Wahlin, Yri et al. 2012)

1.1.1.2.1 Clinical aspects

Peripheral lymphadenopathy is a common clinical presentation of FL. However, so-called B symptoms, such as fever, night sweats, and weight loss, are rarer in FL patients. Bone marrow involvement has been reported in 50 to 60 % of patients (Salles 2007). FL lymphoma is an

age-related disease which mostly occurs in the sixth decade of life. The majority of cases are diagnosed with advanced disease (stage 3 or 4), which present a poorer prognosis compared to stage 1 or 2 disease. Large (>5cm) peripheral lymph nodes, two or more extra-nodal sites of disease and extensive bone marrow involvement are also associated with poor outcome. Elevated serum LDH, ESR and ß -macroglobulin as well as anaemia and hypo-albuminaemia (O'Brien, Easterbrook et al. 1991) are negative predictors of outcome while a higher peripheral lymphocyte count is a predictor of increased survival (Siddiqui, Ristow et al. 2006). Four different clinical prognostic indices (IPI, ILI, FLIPI, and FLIPI2), have been developed to classify FLs; however, risk-adapted treatment is not yet available (Oken, Creech et al. 1982, Federico, Bellei et al. 2009).

Table 1.1 clinical prognostic features in FL. Comparison of clinical prognostic indices in FL (Oken, Creech et al. 1982, Federico, Bellei et al. 2009)

Subtype group	**IPI**		**ILI**		**FLIPI**		**FLIPI-2**	
Prognostic Factors	Age	> 60	Age	> 60	Age	≥ 60	Age	> 60
	Stage	3 to 4	Sex	Male	Stage	3 to 4	B2M	> ULN
	LDH(*iu)*	> ULN	LDH(*iu*)	> ULN	LDH(*iu*)	> ULN	Hb*(g/dl)*	< 12
	Extranodal	≥ 2	Extranodal	≥ 2	Hb*(g/dl)*	< 12	LN sites	> 6
					LN sites	≥ 5		
Study type	Retrospective single-center		Retrospective multi- center		Retrospective multi-center		Retrospective multi-center	

1.1.1.2.2 Immunophenotype and genetics

Immunohistological and molecular analyses indicate that follicular lymphomas showing atypical immunophenotypes and molecular profiles also exist (Marafioti, Copie-Bergman et al. 2013). Immunodiagnostic features of FL include CD10, BCL-6 and BCL-2 expression (DeWolf, Lange et al. 1981, Jacob, Kelsoe et al. 1991). Around 85% of FL cases carry the translocation t (14; 18) (q32; q21), which leads to the over-expression of the anti-apoptotic BCL2 protein. Many of the remaining cases have translocations between 3q27 (involving the BCL6 oncogene) and immunoglobulin gene loci (Otsuki, Yano et al. 1995). For the first time, in 1986, one study showed the presence of SHM in FL (Cleary, Meeker et al. 1986). SHM is responsible for the high intraclonal diversity of FLs (Carlotti, Wrench et al. 2009). Additionally, malignant clonal expansion occurring after the initiation of SHM demonstrates that FL is a GCB cell origin malignancy (Pasqualucci, Trifonov et al. 2011).

1.1.1.3 Burkitt lymphoma (BL)

BL is a highly aggressive lymphoma with an extremely high proliferation rate and rapid death in the absence of treatment (Nomura et al., 2008). This, in adults relatively uncommon, type of lymphoma was first described by Denis Burkitt as a sarcoma involving the jaws in African children (Boerma, Siebert et al. 2009).

1.1.1.3.1 Clinical aspects

In the WHO classification, BLs comprise three clinical variants; endemic, sporadic and immunodeficiency-associated BL. Endemic BL refers to those cases affecting children, usually 4–7 years old, in the geographic areas of endemic malaria falciparum in Africa. This type involves the bones of the jaw and other facial bones, as well as kidneys, gastrointestinal tract, ovaries, breast and other extranodal sites (Campo, Swerdlow et al. 2011). All endemic cases are characterized by Epstein-Barr virus (EBV) infection (Kelly, Stylianou et al. 2013).

Sporadic BL occurs worldwide and typically in industrialized countries. It account for 1-2% of all lymphomas (1%–2% in adults, up to 40% of lymphoma in children in U.S.A and Western Europe). The average age at disease onset in children is 5-10 years and in adults is 30 years of age (Blum, Lozanski et al. 2004). The abdominal organs and lymph nodes are the most common sites of involvement. An EBV infection of neoplastic cells can be detected in 15%-30% of sporadic BL cases (Burmeister, Schwartz et al. 2005).

Immunodeficiency-associated BL is most commonly seen in HIV-positive patients and less frequently in allograft recipients. BL accounts for 30%–40% of NHL in HIV positive patients. BL can be the initially presenting symptom of HIV infection in these patients (Gong, Stenzel et al. 2003, Xicoy, Ribera et al. 2003). EBV infection is seen in up to 40% of immunodeficiency-related BL cases.

Most subtypes of BLs respond well to therapy; the mainstay of treatment is intensive combination chemotherapy regimens such as CODOX-M/IVAC or hyper-CVAD (Thomas, Cortes et al. 1999, Mead, Sydes et al. 2002). Poor prognostic indicators, especially in children, are bone marrow and central nervous system involvement, unresected tumor larger than 10 cm in diameter and high LDH serum levels (Raphael, Gentilhomme et al. 1991).

1.1.1.3.2 Morphology and Immunophenotype

From a morphological point of view, the 2008 WHO classification describes two subtypes, i.e. BL with plasmacytoid differentiation and atypical BL/Burkitt-like lymphoma (Campo, Swerdlow et al. 2011).

1.1.1.3.3 Pathogenesis

A defining feature of BL is MYC translocations resulting in the overexpression of the gene. In 80% of BL cases the translocation partner for C-MYC is the IgH locus forming t (8:14) (q24:q32) and in the remaining 20%, it is the kappa or lambda light chain leading to formation of the t (2; 8) or t (8; 22), respectively. The other common lymphoma-associated translocations such as BCL-2 or BCL-6 translocation have so far not been detected in BLs. Although MYC translocation is a hallmark of BLs, it is not restricted to BLs. Additionally, around 10% of patients do not carry MYC translocations. While BLs implicate a high-level expression of MYC, the normal proliferating tissues, i.e., the normal germinal center B cells even with high division rates have very low levels of MYC expression. The proto-oncogene MYC is required for selection in the germinal center and cyclic re-entry. The other genetic event which has been shown in one third of BL cases is TP53 gene mutations, occurring in immunocompetent and immunosuppressed patients (Ferry 2006).

1.2 Non-coding RNAs

The term non-coding RNAs (ncRNAs) is commonly used for a group of RNAs that does not encode for a protein. NcRNAs are involved in multiple biological processes including tumorigenesis. According to their length, ncRNAs are divided into two categories. Transcripts that are shorter than 200 nucleotides are considered as small ncRNAs and greater than 200 nucleotides as long ncRNAs (lncRNAs). Small ncRNAs include small interfering RNAs (siRNA), small nucleolar RNAs (snoRNA), small nuclear RNAs (snRNA), piwi-interacting RNA (piRNA), and miRNAs (miRNA) (Gutschner and Diederichs 2012).

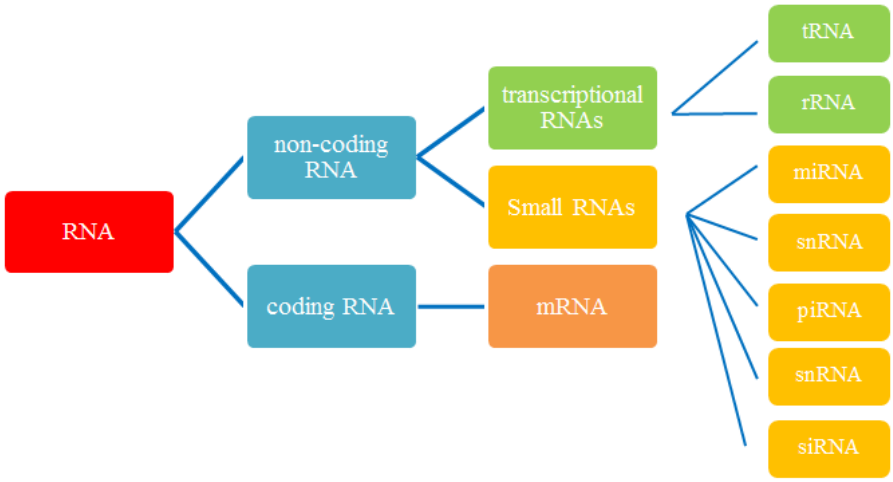


Figure 1.1 The RNA family and different types of non-coding RNAs (ncRNAs).

1.2.1 MiRNA

MiRNAs are a family of endogenous small RNAs (21-23 nucleotides in length) that were first discovered in 1993, when lin-4 gene, crucial for nematode Caenorhabditis elegans development, was found to not encode for a protein, but still negatively regulating lin-14 expression (Lee, Feinbaum et al. 1993). The second important miRNA, *Let 7*, was discovered in 2000 again in C. elegans. The identification of these two miRNAs stimulated an explosion in the field of genetics. Over the past decade, hundreds of miRNAs have been identified in vertebrates, flies, worms, plants and viruses. The miRNA sequence repository database miRBase, currently shows in 24,521 entries in its 2013 release (Kozomara and Griffiths-Jones 2011).

MiRNAs have emerged as crucial factors in basic biological processes such as growth, invasion, angiogenesis, proliferation, and differentiation via the posttranscriptional regulation of (as an estimation) more than one third of all human protein coding genes (Bartel 2004). MiRNA genes are dispersed through the human genome, covering 1 to 3% of the entire genome.

These genes can be classified as intergenic and intronic. Although miRNA genes are scattered in both intron or intergenic regions, most miRNA genes are found in intronic regions (Negrini, Nicoloso et al. 2009).

1.2.1.1 MiRNA biogenesis

Transcription is one of the most important regulatory steps in the biosynthesis of miRNAs. Although transcription of a miRNA gene is mediated by RNA polymerases II and III, most mammalian miRNAs are transcribed by RNA polymerase II (pol II), which produces a primary miRNA (pri-miRNA) transcript. Primary miRNAs consist of multiple stem loop/hairpin structures, 5' cap and a 3' poly-A-tail, being subsequently processed into more than one functional miRNA. Pri-miRNAs are cleaved by Drosha (nuclearRNase III) and Di George Syndrome critical region gene 8 (DGCR8) to form hairpin structures of about 60 to 100 nucleotides in the nucleus called miRNA precursor or pre-miRNAs. The pre-miRNAs are recognised by Exportin 5 and transported in a Ran-GTP dependent manner to the cytoplasm. In the cytosol, pre-miRNAs are processed by RNase III enzyme Dicer to form double stranded miRNAs: miRNA* duplex of about 15 to 22 nucleotides. The two miRNA strands are then separated by helicase enzyme into mature miRNA and passenger miRNA (miRNA*). The mature strand associates with an Argonaute (AGO) protein within the RNA-induced silencing complex (RISC) where it can bind to its target mRNAs. The miRNA strand, which gets incorporated into RISC, is called guide strand, while the opposite strand is known as passenger strand (the star (*)-strand), which is degraded.

However, recent deep sequencing data has revealed that some miRNAs* are not degraded and can play a functional role (Suzuki, Arase et al. 2011). Beside this classic miRNA biogenesis pathway, some alternative biogenesis pathways have recently been described. Mirtrons as alternative precursors for miRNA biogenesis are short hairpin introns which can resemble and mimic pre-miRNA hairpin structure. These mitrons bypass the Drosha processing step by using a splicing mechanism (Berezikov, Chung et al. 2007).

Dicer-independent miRNAs represent another alternative pathway so that pre-miRNA can be directly incorporated into the AGO complexes (Cheloufi, Dos Santos et al. 2010). Additionally, mature miRNA variants termed isomiRs can be generated by some post-transcriptional regulation processes. These isoforms are commonly reported in deep-sequencing studies, however their biological function still remains unclear (Cheloufi, Dos Santos et al. 2010).

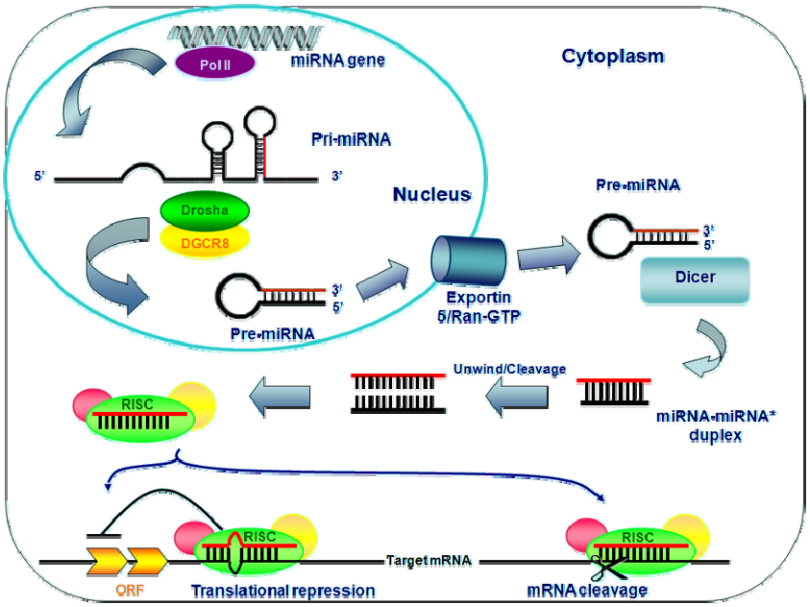


Figure 1.2 MiRNA biogenesis. Schematic representation of the known biogenic pathway of miRNA processing and maturation (Castellano and Stebbing 2013)

1.2.1.2 MiRNA function

The mechanism involved in miRNA-induced gene inhibition mainly depends on the level of complementarity between the miRNA and the mRNA targets leading to the inhibition of protein synthesis or accelerated mRNA degradation. Other factors include thermodynamic stability of miRNA and target interaction, target site accessibility and conservation of target sequences. Each miRNA can regulate up to several hundred target genes and each target gene may also be regulated by many different miRNAs (Bartel 2004, Lewis, Burge et al. 2005).

Functional studies showed that miRNAs interact with their target mRNA via a highly conserved sequence called seed sequence. This seed sequence contains nucleotides 2 to 7 at the 5' end; pairing to its specific target leads to translational repression or degradation of the target mRNA (Lewis, Shih et al. 2003, Eulalio, Huntzinger et al. 2008).

The process of translation can be divided into three steps, i.e., initiation, elongation and termination. Several studied show that miRNAs can trigger gene silencing at any of these steps (Filipowicz, Bhattacharyya et al. 2008).

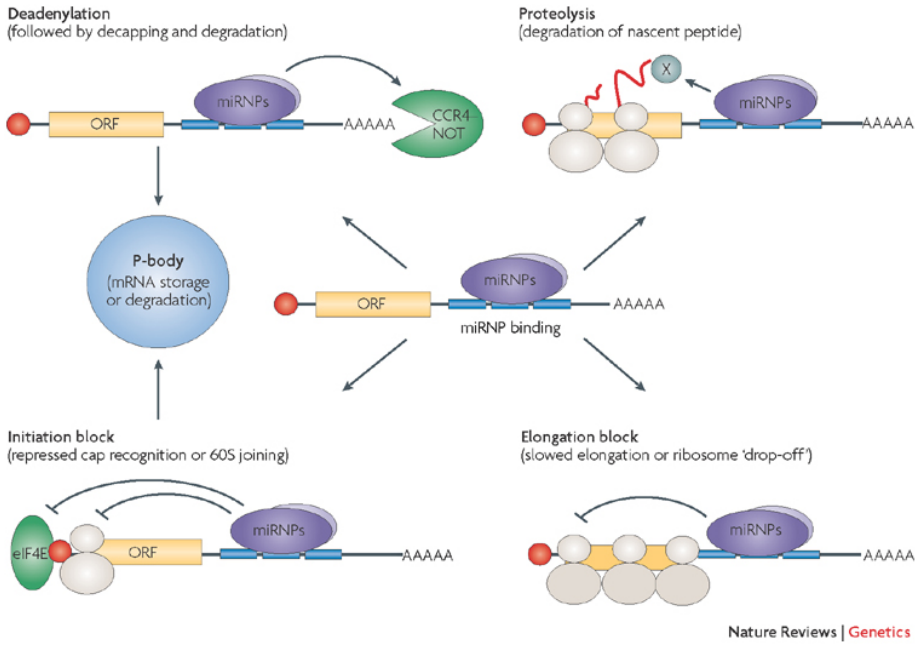


Figure 1.3 Mechanism of miRNA actions. Possible mechanisms of the miRNA-mediated post-transcriptional gene repression in animal cells (Filipowicz, Bhattacharyya et al. 2008)

MiRNAs are key regulators of multiple processes in higher eukaryotes. From an evolutionary point of view, most miRNAs are conserved and have similar functions. Both miRNAs, *lin-4 and let*-7, which were first discovered in *C. elegans*, have key roles in developmental timing. Multiple studies have demonstrated that several miRNAs are involved in a wide range of biological processes such as cell cycle control, homeostasis, hypoxia, circadian rhythm, neurogenesis, insulin secretion, aging, immune responses and viral replication (Bartel 2004). MiRNA expression profiles have shown that deregulation of miRNAs leads to the development of cancer and other types of disease, including diabetes, cardiovascular diseases, neurodegenerative diseases and autoimmune disorders (Kent and Mendell 2006, Bryant, Lutherborrow et al. 2009).

1.2.1.3 MiRNAs and cancer

Profiling studies have also revealed that several miRNAs exhibit oncogenic properties. The first report suggesting a potential involvement of miRNAs in cancer described *miR-15a* and *miR-16*, which were found to be located at chromosomal region 13q14, a genomic region which is deleted in most cases of chronic lymphocytic leukemia (CLL) resulting in the loss of expression of both *miR-15 and miR-16* (Calin, Dumitru et al. 2002).

A few years later the same group discovered that 52.5% of human miRNAs genes are located at cancer-associated genomic regions or in fragile sites suggesting that miRNAs have crucial

roles in human cancer pathogenesis (Calin, Sevignani et al. 2004). In a subsequent study, Zhang *et a.* demonstrated that a high proportion of genomic loci containing miRNA genes exhibit DNA copy number variations (CNVs) (Zhang, Huang et al. 2006). After these initial studies, many reports showed different miRNA expression in a variety of cancer types, including solid tumors and hematological malignancies (Lu, Getz et al. 2005, Calin and Croce 2006).

1.2.1.4 Role of miRNAs in B-cell NHL

MiRNAs are important elements in various steps of B-cell development; as a consequence, miRNA dysregulation of multiple genetic mechanisms leads to the activation of several oncogenic pathways, including the reprogramming of B-cells. Expression profiling studies have suggested that specific miRNAs can be used to characterize different subtypes of B-cell lymphomas (Malumbres, Sarosiek et al. 2009). *MiR-155* has been identified as one of the main oncomirs in solid tumors and hematological malignancies. Furthermore, *miR-155* is one of the first described oncogenic miRNAs in B-cell lymphoma. This miRNA was initially identified as a non-coding RNA termed bic (Rodriguez, Vigorito et al. 2007). Physiologically, *miR-155* is highly expressed during B-cell activation upon the germinal center reaction and has a central role in antibody class switching and plasma cell formation, both processes are thus impaired through *miR-155* deficiency (Rodriguez, Vigorito et al. 2007, Thai, Calado et al. 2007). The oncogenic role of *miR-155* in lymphomagenesis is still under investigation, but recently one study showed that *miR-155* can repress SH2-domain containing inositol-5-phosphatase 1 (SHIP-1), which has crucial functions during normal B-cell development (O'Connell, Chaudhuri et al. 2009). Additionally, SHIP-1 plays a vital role as the phosphatase which negatively regulates the AKT pathway. Overexpression of *mir-155* results in activation of AKT and subsequently induced B-cell proliferation (Havelange, Garzon et al. 2009). More evidence for a role of *miR-155* in lymphomagenesis has been obtained from experiments with mice transplanted with *miR-155* which display extensive myeloproliferation leading to B-cell malignancies. In addition to *SHIP-1*, other *miR-155* targets such as C/EBPß, PU.1 and CSFR were also validated (O'Connell, Chaudhuri et al. 2009). The second most common oncogenic miRNA cluster which represents an important player in lymphomagenesis is the *miR-17-92* cluster, consisting of six miRNAs (*miR17, miR-18a, miR-19b, miR-20a, miR-19b*, *miR-92). MiR-17-92* polycistron is located in the 13q31-32 region which is commonly amplified in B-cell lymphomas. Olive et al. showed that over-expression of *miR-17-92* in lymphocytes of the E(mu) MYC mouse model of B-cell lymphoma leads to accelerated tumor development

(Olive, Jiang et al. 2010). Although deletion of the whole *miR-17-92* cluster leads to lymphoma development, several individual members of this miRNA cluster have also been shown to play a critical role in tumor progression. In addition, some miRNAs, such as *miR-15a* and *miR-16* may function as tumor suppressors in B-cell lymphoma (Cimmino, Calin et al. 2005).

1.2.1.5 MiRNAs in DLBCL

MiRNA profiling of DLBCLs showed distinct expression patterns of miRNAs compared to other NHLs such as BL, CLL and FL (Zhang, Jima et al. 2009, Robertus, Kluiver et al. 2010). *MiR-150, miR-17-5p, miR-145*, and *miR-328* are upregulated in DLBCL compared to healthy lymph node samples and FL (Roehle, Hoefig et al. 2008). In addition, ABC-DLBCL and GCB DLBCL subtypes, which have different genetic aberrations and clinical outcomes, can be distinguished not only by gene expression profiles, but also by distinct miRNA expression profiles. In particular, GCB DLBCL shows a significant *miR-17-92* cluster up regulation, compared to ABC-DLBCL and B-cell controls. *MiR-155* is one of several miRNAs that has been identified to be significantly overexpressed in DLBCL especially in ABC-DLBCL (Kluiver, Haralambieva et al. 2006). Bone morphogenetic protein (BMP)-responsive transcriptional factor SMAD5 is directly targeted by *miR-155* so that *miR-155* overexpressing DLBCLs are less sensitive to growth inhibitory effects of bothTGF-ß1 and BMPs. The expression of specific miRNAs such as *miR-18a, miR-21, miR-155, miR-181* and *miR-222* can be useful for survival prediction in patients with DLBCL treated with (R-) CHOP (Li, Kim et al. 2009).

1.2.1.6 MiRNAs in FL

FLs are characterized by upregulated expression of *miR-20a/b* and *miR-194* which can target proliferation inhibitors such as CDKN1A and SOCS2, respectively (Wang, Corrigan-Cummins et al. 2012). Several miRNAs such as *miR-155, miR-210, miR-106a, miR-149,* and *miR-139*, are overexpressed in both FL and DLBCL compared to normal B-cells. A significant proportion of FL cases transform to aggressive DLBCL with poor outcome, this group has been identified with different miRNA expression signatures from de novo DLBCL cases. Furthermore, six miRNAs (*miR-223, 217, 222, 221, let-7i, 7b)* were found to be expressed differentially in transformed DLBCL and normal DLBCL. All of these data suggest

that miRNAs can be considered as prognostic and diagnostic markers in lymphomas (Lawrie, Chi et al. 2009).

1.2.1.7 MiRNAs in BL

BL as a highly aggressive NHL is characterized by a high level of proliferation of tumor cells. Not only are their protein-coding gene expression profiles different from DLBCL, but BL tumor cells also display an miRNA signature distinct from DLBCLs. BLs express low levels of *miR-155* in comparison to normal B-cells. Recently, a signature of 38 miRNAs containing MYC-regulated and NF-kB pathway-associated miRNAs was published, which differentiated BL from DLBCL (Lenze, Leoncini et al. 2011). Sporadic subtype (sBL), Epstein Barr virus (EBV)-associated endemic subtype (eBL) and an HIV-associated subtype are three different subtypes of BL that be found with uniform biological aspect despite differences in their location of occurrence. MiRNA profiles demonstrate that sBL and eBL vary marginally only by six differentially expressed miRNA (Dave, Fu et al. 2006, Zhang, Nie et al. 2008)**.** Most BL tumors show translocation of the MYC gene on chromosome 8q24 to an immunoglobulin gene on chromosomes 14q32, 22q11 or 2p12, leading to dysregulation of the MYC gene. Studies have shown that MYC induces expression of the *miR-17-92* cluster in BL cases. In addition, MYC has been shown to be affected by miRNAs, *miR-17-5p* and *miR-20a*, two members of *miR-17-92* clusters that can regulate MYC by targeting E2F1 (O'Donnell, Wentzel et al. 2005).

Table 1.2 miRNA expression signatures of the most common NHLs (L Lim, Emilia et al. 2013)

Disease	Subtype	Upregulated miRNAs	Downregulated mRNAs
DLBCL	ABC and GCB	miR-150,miR-17-5p-miR-145, mir328,	-
	ABC vs GCB	mir155,miR-21,miR-221,miR-17-92	-
		cluster	-
		miR-146a, miR-146-b, miR-21, mir155,	-
		miR-500, miR-22, miR-363, miR-574,	
		miR-17, miR-19b, miR-20a, miR-29a,	-
	GCB	miR-92a, miR-106a, miR-720, miR-1260,	-
	Central nervous system	miR-1280	
	Testicular	miR-17-92 cluster	
		miR-17	
	Transformed from FL	miR-127, miR-223, miR-217, miR-222, let-7i, let-7b	
FL		miR-9, miR-301, miR-9*, miR-338, miR-213, miR-193a, miR-193b, miR-345,	miR-17, miR-30a, miR-33a, miR-106b, miR-141, mir202,

		miR-513b, miR-574, miR-54, miR-663, miR-1287, miR-1295, miR-1471	miR-205, miR-222, mir301b, miR-431, mir570
BL	eBL and sBL vs DLBCL	miR-371, miR-185, miR-93, miR-326, miR-497, miR-26a, miR-339, miR-485, miR-9, miR-193a, mir448, miR-429, miR-324, miR-340, mir105, miR-124	miR-221, miR-155, miR-146a, mir26b, mir23a, mir30d, miR-107, miR-103, miR-222, miR-26a, miR-30a, miR-142, miR-23b, miR-342, miR-29b, miR-34b

1.2.1.8 MiRNA detection

One of the best approaches to investigate the roles of miRNAs in cancers is to profile the mature miRNAs in specific tissue types of various cancers (Wark, Lee et al. 2008). Due to the short length of mature miRNAs, sequence similarities of miRNAs in the same sequence family and the fact that miRNAs only represent a small fraction (~0.01%) of the total RNA of a cell, miRNA detection faces several major technical challenges. Considerable effort was thus made to develop new approaches to miRNA profiling. Traditional strategies for detection of miRNA include Northern blotting, microarrays, and quantitative RT-PCR (qRT-PCR) with their relative strengths and weaknesses. Sequencing technologies have overcome previous technical limitations and provide promising methods for miRNA profiling. In 2006, appearance of next-generation sequencing (NGS) technology revolutionized the field of genomics. This method also has its advantages and disadvantages, the advantages include providing information about all RNA classes and the possibility of discovering novel miRNAs or other types of small RNAs, since there is no need to have prior sequence information. The disadvantage of NGS is the size of the data output, which is large and complex and therefore needs extensive bioinformatic analysis (Bernardo et al. 2012). Today, several NGS methods are available, including the Roche GS-FLX 454 Genome Sequencer (originally 454 sequencing), the Illumina HiSeq (originally Solexa technology), the ABI SOLiD analyzer, Polonator G.007 and the Helicos HeliScope platforms. Of these five platforms, the Illumina/Solexa HiSeq is currently the most widely used system (Zhang, Chiodini et al. 2011).

1.2.2 Long non-coding RNAs (LncRNAs)

LncRNAs (also referred to as lincRNA, for long intergenic ncRNA) are a large and heterogeneous group of RNA molecules with more than 200nts in length displaying different

molecular and cellular functions. According to the large-scale sequencing of full-length cDNA libraries, more than 3,300 human lncRNAs have so far been identified (Okazaki, Furuno et al. 2002, Guttman, Amit et al. 2009).

1.2.2.1 LncRNAs biogenesis

From the biogenesis point of view, lncRNAs exhibit a complicated biogenesis. Similar to the short-ncRNA, most lncRNAs are transcribed by RNA polymerase II (RNAP II), but some of them are also transcribed by RNAP III. Most lncRNAs (similar to protein-coding RNAs) are spliced, 5'-capped and polyadenylated. In addition, a majority of lncRNAs have proximal promoter and secondary RNA structures. LncRNAs originate from intronic, exonic, intergenic, intragenic, promoter regions, 3' and 5'UTRs. Additionally, most lncRNAs are transcribed from opposite strands of known protein-coding strands (sense) which are then called natural antisense transcripts (NATs), and recently were found to be generated from pseudogenes. NATs are subdivided into two groups: cis-NATs and trans-NATs, based on their function in cis or trans (Nie, Wu et al. 2012).

1.2.2.2 LncRNA function

Although cellular functions of many newly discovered lncRNAs still need to be clarified, it has been shown that a large proportion of lncRNAs play regulatory roles, including chromatin modification, genomic rearrangement, cell cycle regulation, genetic imprinting, transcription, splicing, mRNA decay and translation. Generally, lncRNAs regulate gene expression aspects in three steps: epigenetic, transcriptional and post-transcriptional (Clark and Blackshaw 2014).

1.2.2.3 The role of lncRNAs in cancer

Several studies have shown that the expression of lncRNAs is associated with human diseases, including cancers; however, the regulatory mechanisms of how lncRNAs contribute to cancer development are varied. The key role of lncRNAs is to interact with chromatin-modifying complexes, resulting in epigenetic changes (Mattick, Taft et al. 2010, Yang and Deng 2014). The lncRNAs ANRIL, XIST, HOTAIR and KCNQ1OT1 are able to reprogram the chromatin state in diverse cancers. Other lncRNAs like lincRNA-p21 are able to regulate protein signaling pathways associated with carcinogenesis. Some lncRNAs play a

role in RNA processing; the lncRNA MALAT1 acts as post-transcriptional regulator by controlling alternative splicing of pre-mRNAs (Tripathi, Ellis et al. 2010). Overexpression of MALAT1 leads to an increase in cell proliferation and migration in several cancers such as lung and colorectal cancer (Schmidt, Spieker et al. 2011, Xu, Yang et al. 2011). Some cancer cells recruit telomerases to achieve replicative immortality, lncRNA TERRA binds to telomerase and inhibits its function; in vitro downregulation of this lncRNAs has been linked to immortality of cancer cells by telomerase-mediated lengthening of telomeres (Redon, Reichenbach et al. 2010). Other lncRNAs can function as decoys; the tumor suppressor gene PTEN and its pseudogene PTENP1 are good examples of this mechanism. PTENP1 binds to miRNAs which inhibit PTEN expression and regulate cell proliferation (Poliseno, Salmena et al. 2010). Although knowledge about lncRNA functions in cancer has been growing, the function of most lncRNAs in many cancer types still remains undefined (Cheetham, Gruhl et al. 2013) (Figure 1.5).

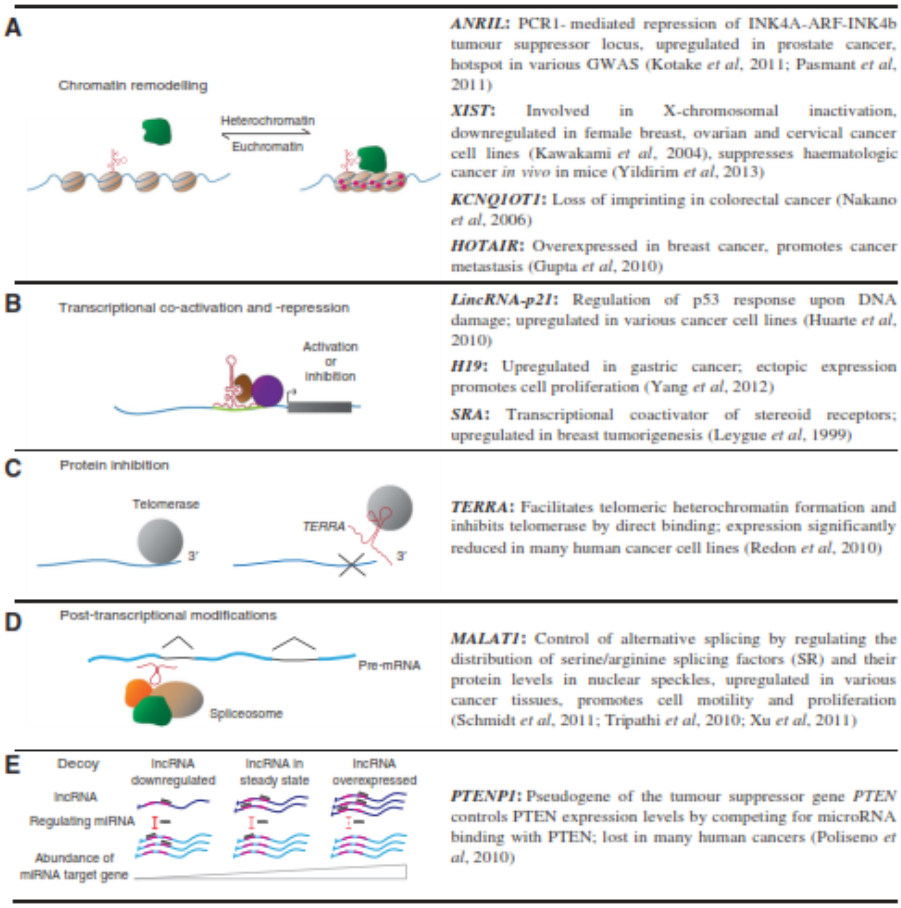


Figure 1.4 General mechanisms and examples of lncRNAs involved in cancer. LncRNAs act through a variety of mechanisms such as remodelling of chromatin (A), transcriptional co-activation or -repression (B), protein inhibition (C), as post-transcriptional modifiers (D) or decoy elements (E). Consequently, mis-expression of lncRNAs can lead to changed expression profiles of various target genes involved in different aspects of cellular homeostasis (Cheetham, Gruhl et al. 2013).

Compared to miRNAs, the number of known lncRNAs associated with blood cancers is much smaller, suggesting that much effort is still needed to expand our knowledge of lncRNAs and their roles in hematological malignancies, including lymphoma.

1.3 ICGC-MMML-Seq

The ICGC-MMML-Seq Consortium is the continuation of a long and outstanding tradition of interdisciplinary, multicenter collaboration in the field of lymphoma research in Germany. It is based on strongly interacting networks of reference pathologists, lymphoma scientists and clinical trial groups. These groups have joined-up in the ICGC-MMML-Seq with internationally renowned groups in the field of high-throughput sequencing and sequence analysis.

The ICGC-MMML-Seq Consortium aims to analyze 250 GCB-lymphomas according to the guidelines of the International Cancer Genome Consortium (ICGC). The core subtypes of GCB-lymphoma to be analyzed are FL, DLBCL and BL. The selection of these tumor types was based on their frequency, clinical relevance, socio-economic burden, the availability of material and overlapping biologic/genetic features.

The analysis pipeline of the ICGC-MMML-Seq contains three pre-analysis workpackages responsible for standardized sample acquisition through the clinical trial groups (WP1), reference pathology (WP2) and preparation of analytes (WP3). A special WPN provides samples from normal B-cells in order to address the germline problem inherent to B cells, since they are subject to somatic mutations even under physiological conditions. Sequencing covers genome (WP4), transcriptome (WP5), small RNAome (WP6) and DNA-methylome (WP7). Stepwise analysis of sequence data will be performed in two workpackages (WP8 and WP9), which will ensure integrative and comprehensive evaluation.

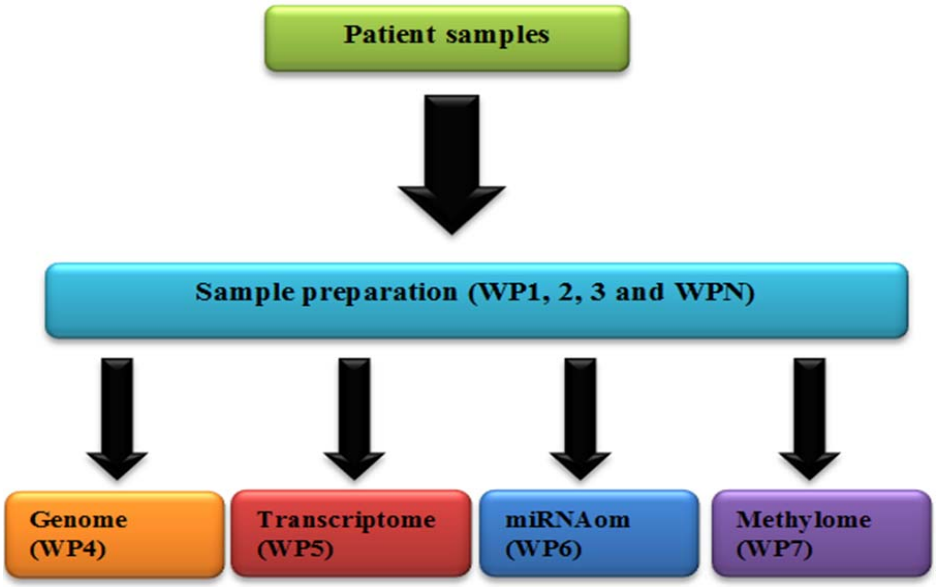


Figure 1.5 Sample workflow of the ICGC MMML-Seq Consortium.

1.4 Aim of the thesis

The use of NGS in small RNA profiling studies facilitates the identification of miRNA expression in many physiological and pathological processes. MiRNA expression signatures are deregulated in most human tumors and these aberrant miRNA expression patterns play crucial roles in tumorigenesis.

Here, NGS has been applied to analyze up to 250 GCB-lymphomas according to guidelines of ICGC-MMML-Seq. According to the WHO classification, there are two molecular subtypes of DLBCL, namely germinal center-derived (GCB) and activated B-cell (ABC). According to this definition, ABC cases do not directly derive from germinal center B-cells (lack of GC features) but display relatively high number of somatic immunoglobulin mutations. This identifies a clear relation of the ABC-type DLBCLs to s role of the GC in their pathogenesis. The ICGC-MML-Seq consortium followed this idea and decided to include ABC patients in its analysis of germinal center-related B-cell lymphomas. As my work is based on ICGC data, I am also following the latter approach in my thesis.

This was the first time that NGS was employed for miRNA studies in GCB-lymphomas. The aims of the study were first to investigate whether different GCB-lymphomas subtypes show specific miRNA expression, and to apply this knowledge to identifying potential biomarkers for distinguishing BL, FL and DLBCL. Deregulated miRNAs were identified by deep sequencing and were validated subsequently by RT-qPCR.

The second aim was to identify and validate potential novel miRNAs which may have roles in tumorigenesis of B-cell lymphomas.

The third aim was to characterize mRNA targeted by miRNAs in B-cell lymphoma cell lines by using AGO2-PAR-CLIP and the last objective of this study was to investigate the function of lncRNAs in lymphomas.

2 Materials

2.1 Human suspension cell lines

Cell Line: RAJI

DSMZ no.: ACC 319

Species: human (*Homo sapiens*)

Cell type: Burkitt lymphoma

Origin: established from the left maxilla of a 12-year-old African boy with Burkitt lymphoma in 1963; first continuous human hematopoietic cell line; classified as risk category 1 according to the German central commission for biological safety (ZKBS); cells are described to carry t (8;14) leading to MYC-IGH@ (MYC-IGH) fusion gene

Cell line: SU-DHL-4

DSMZ no.: ACC 495

Species: human (*Homo sapiens*)

Cell type: B-cell lymphoma

Origin: established from the peritoneal effusion of a 38-year-old man with B-NHL (diffuse large cell, cleaved cell type; originally described as "diffuse histiocytic lymphoma") in 1975; cell line carries EZH2 Y641S mutation; assigned to GCB-like lymphoma subtype (germinal center B-cell)

Cell line: SU-DHL-6

DSMZ no.: **ACC 572**

Species: human (*Homo sapiens*)

Cell type: B-cell lymphoma

Origin: established from the peritoneal effusion of a 43-year-old man with B-cell non-

Hodgkin lymphoma (B-NHL), described at the time as diffuse, mixed small and large cell type; cell line carries EZH2 Y641N mutation; assigned to GCB-like lymphoma subtype (germinal center B-cell)

Cell line: NAMALWA

DSMZ no.: ACC 69

Species: human (Homo sapiens)

Cell type: Burkitt lymphoma

Origin: established from the tumor mass of an African child with Burkitt lymphoma; this variant subline from the original NAMALWA (ACC 24) was described to be characterized by differential expression of immunoglobulin and MYC genes and of some cell surface antigens (see reference); cells were reported to harbor a retrovirus of simian origin, SMRV-H

2.2 Chemicals

2.2.1 General chemicals

Identifier	Company	Order number
1,4.dithio-DL-threitol (DTT)	Roth, Karlsruhe, Germany	6908.2
2-Propanol for Analysis	Merck, Darmstadt, Germany	67-63-0
Bis-Acrylamide 30% (19:1)	Bio-Rad, München, Germany	161-0154
Agarose	Biozym, Hessisch Odendorf, Germany	840004
Ammonium Persulfate	Sigma-Aldrich, St. Louis, MO, USA	A-3678
Bromphenol blue	Sigma-Aldrich, St. Louis, MO, USA	115-39-9
BSA 10 mg/ml	NEB, Frankfurt a. M., Germany	B90015
Chloroform 99.4%	Merck, Darmstadt, Germany	67-66-3
DMSO (Dimethyl sulfoxide)	Sigma-Aldrich, St. Louis, MO, USA	D2650
EDTA (Ethylenediamine-tetraacetic acid)	Sigma-Aldrich, St. Louis, MO, USA	60-00-4
Ethanol	Merck, Darmstadt, Germany	1.00983.2511
Ethidium bromide Solution	Sigma-Aldrich, St. Louis, MO, USA	E1510-10ML
Glycerine	Merck, Darmstadt, Germany	1.04094.2500
Glycine	Merck, Darmstadt, Germany	1.04201.1000
H_3BO_3 (Boric acid)	Merck, Darmstadt, Germany	1.00165.
HCl (Hydrochloric acid)	Merck, Darmstadt, Germany	1.09911.0001
KCl (Potassium chloride)	Merck, Darmstadt, Germany	1.04938.
KOH (Potassium hydroxide)	Merck, Darmstadt, Germany	105012
Magnesium (Mg_2)	Merck, Darmstadt, Germany	105815
Methanol	Merck, Darmstadt, Germany	1.06007.2500
Mg_2Cl	Merck, Darmstadt, Germany	1.05833.0250
Na_2EDTA	Sigma-Aldrich, St. Louis, MO, USA	E5134-500G

Na_2HPO_4 (di-Sodiumhydrogen-phosphat-Dihydrat)	Merck, Darmstadt, Germany	6580
NaCl (Sodium chloride)	Merck, Darmstadt, Germany	1.06404.1000
NaF (Sodium fluoride)	AppliChem, Darmstadt, Germany	A3904.0500
NaH_2PO_4 (Sodiumhydrogenphosphat Monohydrat	Merck, Darmstadt, Germany	6346.1000
NaOH (Sodiumhydroxide)	Merck, Darmstadt, Germany	1.06498.100
SDS 20%	Lifetechnologies, Germany	AM9820
50 x TAE (Tris/Acetic Acid/EDTA) Buffer	Bio-Rad, München, Germany	161-0743
TEMED (Tetramethylethylendiamin)	Merck, Darmstadt, Germany	1.10732.0100
Tris	Roth, Karlsruhe, Germany	5429.3
Tween 20	Merck, Darmstadt, Germany	8.22184.2500
Urea	Merck, Darmstadt, Germany	66612
β-Mercaptoethanol	Merck, Darmstadt, Germany	15433.0100

2.2.2 Specific chemicals

Identifier	Company	Order number
[γ-P32] Adenosine 5'-triphosphate (ATP)	Hartmann Analytic, Braunschweig, Germany	FP-301
4-Thioruidine 25 mg	Sigma-Aldrich, St. Louis, MO, USA	T4509
50 x Denhardt Solution	AppliChem, Darmstadt, Germany	A3792.0050
AccuPrime SuperMix I	Invitrogen, Carlsbad, CA, USA	12342-028
Complete, EDTA-free Protease Inhibitor Cocktail Tablets	Roche Diagnostics, Mannheim, Germany	11244800
GlycoBlue 300 µl (15 mg/ml)	Ambion, Huntingdon, Cambridgeshire, UK	AM9515
Isoamylalcohol	Merck, Darmstadt, Germany	100979
Phenol	Merck, Darmstadt, Germany	100206
Phenol acid	Sigma-Aldrich, St. Louis, MO, USA	P4682-400ML
Quick Start Bradford Dye Reagent	Bio-Rad, München, Germany	500-0205
Salmon Sperm DNA 1 ml (10 mg/ml)	Invitrogen, Carlsbad, CA, USA	15632-011
SYBR Green PCR Master Mix	ABI, Carlsbad, CA, USA	4309155
TaqMan Uni. PCR Master Mix, No Amperase UNG	ABI, Carlsbad, CA, USA	4324018
TRIzol Reagent	Invitrogen, Carlsbad, CA, USA	15596-018

2.3 Nucleic acids

2.3.1 Northern blot probes

Description	Company	Length	Sequence
NB-miR-1	MWG, Ebersberg, Germany	24 nt	AGATTCCTGTGCCCTCTGCCCTGG
NB-miR-2	MWG, Ebersberg, Germany	22 nt	CATTCGCCAGCTCACAATGGCT
NB-miR-3	MWG, Ebersberg, Germany	22 nt	CAGCCTCTGAGGATCCCAAGGT
NB-miR-4	MWG, Ebersberg, Germany	23 nt	CCTGCGGAGCAGAGACCAGCCAA
NB-miR-5	MWG, Ebersberg, Germany	23 nt	TCTCTCTCTCTGTGTCAGTGTGC

NB-miR-6	MWG, Ebersberg, Germany	23 nt	GCCTTGGAGGTAGACGGCCCTG
NB-miR-7	MWG, Ebersberg, Germany	23 nt	AGCAGGTACCACAGCAACTGCTT
NB-miR-8	MWG, Ebersberg, Germany	20 nt	CTGCCCGCTGCTGCTCTGTG
NB-miR-9	MWG, Ebersberg, Germany	22 nt	TGAGGGAGGAGCCACAGGCCTG
NB-miR-10	MWG, Ebersberg, Germany	24 nt	GGGAAGCTCTAGTGCCTGAGGAGG
NB-miR-11	MWG, Ebersberg, Germany	22 nt	TAATCTGACTAGGTTTGCCTTG
NB-miR-12	MWG, Ebersberg, Germany	26 nt	GACTGGGGAACAAGGGAGCAGGGGAC
NB-miR-13	MWG, Ebersberg, Germany	24 nt	TGGCCGCCATGCTCAGGTCTTGCG
NB-miR-14	MWG, Ebersberg, Germany	23 nt	GCTCGGCTCAGGCCCCAGCGCAG
NB-miR-15	MWG, Ebersberg, Germany	23 nt	GGCTCGGGCCTGGGGCTCTCCGG
NB-miR-16	MWG, Ebersberg, Germany	25 nt	GTCCAGACCCCGGGCTAGGCCCGGC
NB-miR-17	MWG, Ebersberg, Germany	23 nt	CGAGACATCTCGGCCCGAATGCT
NB-miR-18	MWG, Ebersberg, Germany	28 nt	AGGGTTGTAGTAGCCCGTAGGGGCCTAC
NB-miR-19	MWG, Ebersberg, Germany	28 nt	AGTAGCCCGTAGGGGCCTACAACGTTGG
NB-miR-20	MWG, Ebersberg, Germany	28 nt	TATTGCTAGGGCCCTGCAGTGA

2.3.2 Oligonucleotides

qRT-GAPDH-forward	MWG, Ebersberg, Germany	24 nt	ACCACTTTGTCAAGCTCATTTCCT
qRT-GAPDH-reverse	MWG, Ebersberg, Germany	23 nt	GTTGCTGTAGCCAAATTCGTTGT
LncRNA-AP000251-forward	MWG, Ebersberg, Germany	23 nt	GCAGTCGACGGGCTGATGAGCGA
LncRNA-AP000251-reverse	MWG, Ebersberg, Germany	20 nt	CAGGCGGCCGTGACTTTCAA
AP000251.2 gene specific primer	MWG, Ebersberg, Germany	26 nt	GAAAAGAGACCCAGAGTTGGGAGCGC

2.3.3 Locked Nucleic Acid (LNA™) - Exiqon

Description	Company	Length	Sequence
NB-miR-5	Exiqon, Denmark	18 nt	TCTCTCTGTGTCAGTGTG
NB-miR-19	Exiqon, Denmark	18 nt	TAGGGGCCTACAACGTTG

2.3.4 TaqMan miRNA assays

Identifier	Company	Order number	Assay ID	Description
miR-10b	ABI, Carlsbad, CA, USA	4427975	002218	TaqMan Micro RNA Assay
miR-141	ABI, Carlsbad, CA, USA	4427975	000463	TaqMan Micro RNA Assay
miR-142-5p	ABI, Carlsbad, CA, USA	4427975	002248	TaqMan Micro RNA Assay
miR-143	ABI, Carlsbad, CA, USA	4427975	002249	TaqMan Micro RNA Assay
miR-146a	ABI, Carlsbad, CA, USA	4427975	000468	TaqMan Micro RNA Assay
miR-150	ABI, Carlsbad, CA, USA	4427975	000473	TaqMan Micro RNA Assay
miR-21	ABI, Carlsbad, CA, USA	4427975	000397	TaqMan Micro RNA Assay
miR-22	ABI, Carlsbad, CA, USA	4427975	000398	TaqMan Micro RNA Assay
miR-28	ABI, Carlsbad, CA, USA	4427975	000411	TaqMan Micro RNA Assay

miR-92a	ABI, Carlsbad, CA, USA	4427975	000431	TaqMan Micro RNA Assay
miR-142-3p	ABI, Carlsbad, CA, USA	4427975	000464	TaqMan Micro RNA Assay
RNU24	ABI, Carlsbad, CA, USA	4427975	001001	TaqMan Micro RNA Assay

2.3.5 Other nucleic acids and nucleotides

Identifier	Company	Order number
Oligo dT Primers 100 µl (0.4 µg/µl)	Qiagen, Hilden, Germany	79237
Hexanucleotide Primers H0268-1UN	Sigma-Aldrich, St. Louis, MO, USA	067K6109
dATP, PCR Grade 100mM	Qiagen, Hilden, Germany	1039397
dCTP, PCR Grade 100mM	Qiagen, Hilden, Germany	1039396
dGTP, PCR Grade 100mM	Qiagen, Hilden, Germany	1039395
dTTP, PCR Grade 100mM	Qiagen, Hilden, Germany	1039394

2.3.6 Oligonucleotides for cloning of NB-miRNAs

(MWG, Ebersberg, Germany)

NB-miR-5-TopStrand oligo

5’-TGCTGGCACACTGACACAGAGAGAGAGAGTTTTGGCCATGACTGACTCTCTC TCTGTGTCAGTGTGC -3’

NB-miR-5-BottomStrand oligo

5’-CCTGGCACACTGACACAGAGAGAGAGTCAGTCATGGCCAAAACTCTCTCTCTCTGTGTCAGTGT GC C-3

NB-miR-19-TopStrand oligo

5’-TGCTGCCAACGTTGTAGGCCCCTACGGGCTACTGTTTTGGCCATGACTGACAGTAGCCCAGGGGCCTACA ACGTTGG -3’

NB-miR-19-TopStrand oligo

5’-CCTGCCAACGTTGTAGGCCCCTGGGCTACTGTCAGTCATGGCCAAAACAGTAGCCCGTAGGGGCCTACAA CGTTGG C-3’

2.4 Enzymes

Identifier	Company	Ordernumber
T4 Polynucleotide Kinase (10 U/µl)	NEB, Frankfurt a. M., Germany	M0236L
RNA Ligase 2 (Rnl2 (1-249) K227Q) 1 µg/µl	NEB, Frankfurt a. M., Germany	M0239L
SuperScriptIII Reverse Transcriptase 10,000 U (200 U/µl)	Invitrogen, Carlsbad, CA, USA	18080-044

2.5 Culture media, buffers and dilutions

2.5.1 Media and additives for cell cultivation

Identifier	Company	Ordernumber
Dulbecco' s Modified Eagle Medium 1x (DMEM)	Gibco, Invitrogen, Carlsbad, CA, USA	21969-035
RPMI 1640	Gibco, Invitrogen, Carlsbad, CA, USA	32404-014
Penicillin 10,000 U/ml	Gibco, Invitrogen, Carlsbad, CA, USA	15140-122
Streptomycin 10,000 µg/ml	Gibco, Invitrogen, Carlsbad, CA, USA	15140-122
L-Glutamine 100 x	Gibco, Invitrogen, Carlsbad, CA, USA	25030-024
Trypsin-EDTA 1 x	PAA, Pasching, Austria	L11-004
Dulbecco's PBS without Ca & Mg 1 x	PAA, Pasching, Austria	H15-002
Fetal Bovine Serum (FBS) heat inactivated	PAA, Pasching, Austria	A15-104
Geneticin G418	Gibco, Invitrogen, Carlsbad, CA, USA	10131035

2.5.2 PAR-CLIP buffers

1 M 4-Thiouridine stock solution

250 mg 4-thiouridine, 960.5 µl DMSO

NP40 lysis buffer

20 mM Tris-HCl (pH 7.5), 150 mM KCl, 2 mM EDTA, 1 mM NaF, 0.5% NP40, add fresh before use: 0.5 mM DTT, 1 x complete EDTA-free protease inhibitor cocktail (Roche Diagnostics)

Citrate-phosphate buffer, pH 5.0

4.7 g Citric acid monohydrate, 9.2 g Na2HPO4.7 H2O, H2O to 1 L

IP-wash buffer

50 mM HEPES-KOH, pH.7.5, 300 mM KCl, 0.05 % (v/v) NP40 substitute, 649 ml H2O, 0.5 mM DTT (add fresh), complete EDTA-free 2 protease inhibitor cocktail (add fresh)

10 X Dephosphorylation buffer

50 mM Tris-HCl, pH 7.9, 100 mM NaCl, 10 mM MgCl2.6H2O, 906.2 ml H2O, 1 mM DTT (add fresh)

Phosphatase wash buffer

50 mM Tris-HCl, pH 7.5, 20 mM EGTA-NaOH, pH 7.5, NP40 substitute 0.5 % (v/v)
905 ml H2O

Polynucleotide kinase (PNK) buffer without DTT

50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 10 mM MgCl2.6H2O, 923.3 ml H2O

PNK buffer with DTT

50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 10 mM MgCl2.6H2O, 923.3 ml H2O, 5 mM DTT (add fresh)

1 X SDS PAGE loading buffer

50 mM Tris-HCl, pH 6.8, EDTA-NaOH, pH 8.0 Glycerol, SDS DTT bromophenol blue, H2O to 10 ml

1 X MOPS running buffer

Dilute 1:20 from commercially available 20 X buffer (Invitrogen)

Proteinase K storage buffer

20 mg / ml Proteinase K, 50 mM Tris-HCl, pH 8, 30 mM CaCl2· 2H2O, 50 % Glycerol, H2O to 10 ml

2 X Proteinase K buffer

100 mM Tris-HCl, pH 7.5, 12.5 mM EDTA-NaOH, pH 8.0, 150 mM NaCl, SDS 2 % (v/v) H2O

Acidic phenol/chloroform/IAA (25:24:1)

Mix 25 ml acidic phenol, 24 ml chloroform and 1 ml isoamyl alcohol (overlay with 0.1 M citrate buffer, pH 4.3 ± 0.2)

50% DMSO

Mix 1 ml DMSO with 1 ml H2O

10 X RNA ligase buffer without ATP

0.5 M Tris-HCl, pH 7.6, 0.1 M 2-mercaptoethanol, acetylated BSA 1 mg/ml, 3.43 ml H2O

2 X Formamide loading dye

50 mM EDTA-NaOH, pH 8.0, 0.05 % (w/v) bromophenol blue formamide

10 X TBE

445 mM Tris base, 445 mM Boric acid, 10 mM EDTA-NaOH, pH 8.0, H2O to 1 L

0.4 M NaCl

0.4 M NaCl, 433.3 ml H2O

10 X RNA ligase buffer with ATP

0.5 M Tris-HCl, pH 7.6, 0.1 M MgCl2.6 H2O, 0.1 M 2-Mercaptoethanol Acetylated BSA, 1 mg/ml ATP H2O

150 mM KOH/20 mM Tris base

150 mM KOH, 20 mM Tris base H2O, 950 µl H2O

150 mM HCl

150 mM HCl, 987.6 µl H2O

5 X DNA loading dye

50 mM EDTA-NaOH, pH 8.0, 0.2 % (w/v) bromophenol blue, 20 % (w/v) ficoll type 400, H2O to 10 ml

3 M NaOAc (pH 5.2)

246.09 NaOAc, H2O to 1 L

2.5.3 Northern blot buffers

Sequa Gel 15%

10% Buffer, 3.33% Diluent, 1.66% Concentrate, 0.1% TEMED, 4% APS

10 X TBE Buffer (Tris-borate-EDTA)

890 mM Tris and 890 mM Boric acid, 20 mM Na2¯EDTA (pH 8)

2 X Bromphenol blue-Sample Buffer

8 M Urea, 50 mM EDTA, a pinch of bromphenol blue

Hybridization Solution Denhardt Solution

5 x SSC, 20 mM Na2HPO4 (pH 7.2), 1 % SDS, 5 x add before hybridization: 1% Salmon Sperm DNA (10 mg/ml)

Labeling reaction of 22 nt oligonucleotide (20 µM), γ-32P-ATP (150 to 3000 Ci/mmol), DNA probes 1 x T4 polynucleotide kinase (PNK) buffer, 2 U/µl T4 PNK

Wash solution I

100 ml 3M NaCl

20 ml 1M NaHPO4 (pH7.2)

20 ml 20% SDS

0.08 ml 0.5M EDTA

260 ml d.H2O

Wash solution II

6.67 ml 3M NaCl

10 ml 1MNaHPO4 (pH 7.2)

0.04 ml 0.5 M EDTA

173.3ml dH2O

2.5.4 Small RNA library preparation buffers

10 X TBE Buffer (Tris-borate-EDTA)

890mM Tris, 90 mM boric acid and 20 mM Na2EDTA (pH 8.0)

2.5.5 Cultivation of Raji, SU-DHL4, SU-DHL-6 and Namalwa

1 x DMEM 500 ml, 1% (v/v) Penicillin 10,000 U/ml, 1% (v/v) Streptomycin 10,000 µg/ml, 1 % 200 mM Glutamine, 10% FBS heat inactivated

2.5.6 Freezing medium for cryopreservation

90% culture medium and 10% DMSO

2.6 Kits, size markers and other materials

2.6.1 Kits

Kit Identifier	Company	Order number
TaqMan MiRNA Reverse Transcription Kit	ABI, Carlsbad, CA, USA	4366596
TruSeq Small RNA Sample Preparation Kit	Illumina, San Diego, CA,USA	RS-200-0012
TruSeq SR Cluster Kit v3-cBot-HS	Illumina, San Diego, CA,USA	GD-401-3001
TruSeq SBS Kit v3-HS	Illumina, San Diego, CA,USA	FC-401-3002
Corning®Costar®Spin-X®	Sigma Aldrich, Hamburg, Germany	CLS8162

2.6.2 Identifier

Identifier	Company	Order number
1kb Plus DNA Ladder 250 µg (1 µg/µl)	Invitrogen, Carlsbad, CA, USA	10787-018
Amersham Hybond-N+ Positively Charged Nylon Transfer Membrane	GE Healthcare, Freiburg, Germany	RPN203B
Cryo freezing container 500 ml	Nalgene, NY, USA	5100-0001
Culture Dish 150 x 25 mm	Corning, Amsterdam, Netherlands	430597
Dynabeads Protein G 30 mg/ml	Invitrogen, Carlsbad, CA, USA	100-03D
Extra Thick Blot Paper	Bio-Rad, München, Germany	170-3965
GeneChip Human Genome U133 A 2.0	Affymetrix, Santa Clara, CA, USA	900471
illustra MicroSpin G-25 Columns	GE Healthcare, Freiburg, Germany	27-5325-01
Microcentrifuge Tube 1.7ml prelubricated RNase/DNase free	Corning, Amsterdam, Netherlands	3207

2.5.1. Hardware

Hardware	Company
7900 HT Fast Real-Time PCR System	ABI, Carlsbad, CA, USA
Agilent 2100 Bioanalyzer Serial No. DE72905088	Agilent Technologies, Böblingen, Germany
Centrifuge 5403	Eppendorf, Hamburg, Germany

Centrifuge 5417R	Eppendorf, Hamburg, Germany
FLA-7000 Phosphor Imager	Fujifilm, Düsseldorf, Germany
Gene Amp PCR System 2700	ABI, Carlsbad, CA, USA
HeroLab UVT 2035 UV lamp	HeroLab, Wiesloch, Germany
Hybaid Hybridization Oven	Biometra, Göttingen, Germany
LAS-3000 mini 2UV Transilluminator	Fujifilm, Düsseldorf, Germany
Mastercycler gradient	Eppendorf, Hamburg, Germany
Milli-Q Integral 15 Serial No. FODA 15851D	Millipore, Billerica, MA, USA
NanoDrop Spectrophotometer ND-1000	PeqLab, Erlangen, Germany
Thermomixer Comfort	Eppendorf, Hamburg, Germany
Transfer Blot SD Semi-Dry Transfer Cell	Bio-Rad, München, Germany
Vortex2 Genie	Scientific Industries, NY, USA
HiSeq 2500	Illumina, San Diego, CA,USA

3 Methods

3.1 Cell Culture

3.1.1 Cultivation of human suspension cells

Four cell lines (SU-DHL-4, SU-DHL-6, Raji, Namalwa) were grown in DMEM (1 X Dulbecco's Modified Eagle Medium, Gibco Invitrogen) with 10% (v/v) FBS (PAA), 1% 200 mM L-glutamine (100 x, Gibco, Invitrogen) and 1% (v/v) 10,000 U/ml penicillin, 10,000 µg/ml streptomycin (Gibco, Invitrogen) until they reached a confluence of 90%. They were cultivated at 37°C in humid and saturated atmosphere with 5% CO_2. Cells were split 1:2 three times a week. The determination of cell numbers was performed by using the "Neubauer" counting chamber and 5 x 10^7 cells were used for further experimental analyses.

3.1.2 Cryopreservation of human cells

Cells were removed from the culture flask and pelleted for 5 min at 500 xg. Cells were washed with PBS, and 1 x 10^6 cells/ ml were resuspended in 1 ml freezing medium and transferred into cryotubes. The cryotubes were placed into a cryobox filled with isopropanol and frozen at -80°C for at least 24 h. Subsequently, the cryotubes were placed into the gas phase of liquid nitrogen. For recultivation of the cells, an aliquot was thawed in a water bath at 37°C. The freezing medium was removed by pelletizing the cells at 500 x g for 5 min and resuspending the cells into fresh medium.

3.2 Methods of molecular biology

3.2.1 Competent E.coli by Hanahan method

E. coli strain DH5α was streaked onto an LB plate and was grown overnight at 37°C. Single colonies were picked from the plate and incubated in 100 ml of SOB in a flask. E. coli strain DH5α was grown with shaking at 37°C until OD 600 reached 0.4 to 0.6. Cells were incubated on ice for 10 min, spun down at 2500 rpm for 10 min at 4°C, resuspended gently in 24 ml of cold FSB and then kept on ice for 10-15 min, then spun again at 2500 rpm for 10 min at 4°C and resuspended gently in 20 ml of cold FSB. Before freezing the cells, DMSO was added to a final concentration of 7% and the cell suspension was kept on ice for 10 min. The cell suspension was aliquoted into 100-200 µl batches, quickly frozen in liquid nitrogen and stored at -80°C.

3.2.2 Transformation of chemically competent E. coli

3 µl of e.g. the LR reaction was added into 100µl of competent cells, mixed gently by ticking against the tube and left on ice for 15-20 min. Then the sample was heat-shocked at 42°C for 1 min followed by 10 min recovery on ice. Afterwards, 400 µl of LB medium without antibiotic was added to the transformed cells and incubated at 37°C for 60 min with gentle shaking. The cells were plated onto LB agar plates containing appropriate selective antibiotics and incubated overnight at 37°C.

3.2.3 DNA mini-prep

For purification of small amounts of DNA (Mini-prep), 5-Prime Fast plasmid kit was used. According to the protocol, 3 ml LB-AMP-Medium (1 % (w/v) cultures were inoculated overnight. The next day, bacteria were harvested at 13,000 rpm for 1 min. The resulting pellet was lysed using solutions provided by the manufacturer (ICE-COLD Complete Lysis Solution) and the lysate was incubated for 3 min at room temperature. Subsequently, the lysate was applied to the spin column according to the manual.

3.2.4 DNA Maxi-prep

Preparation of lates amounts of purer plasmid DNA was performed using QIAGEN Plasmid Maxi Kitaccording to the manufacturer's instructions. The DNA yield was then determined using a spectrophotometer to measure absorbance at the wavelength of 260 nm.

3.2.5 RNA extraction (TRIzol reagent)

1 ml of TRIzol reagent was added to 1 x 10^7 cells, which were resuspended by pipetting. After this, 200 µl of chloroform was added to the mixture that was vortexed for 20 sec and incubated for 2 to 3 min at room temperature. The mixture was then centrifuged for 15 min at 4°C and 12,000 xg. The upper phase was mixed with P: C: I (Phenol: chloroform: isoamyl alcohol), vortexed and centrifuged again. 3 to 4 volumes of ethanol were added to the upper phase; both were vortexed and incubated at -20°C for at least 2 h. The RNA was pelletized by centrifugation for 15 min at 4°C and 12,000 x g. The ethanol supernatant was removed, and the RNA was air-dried. Finally, the RNA was resuspended in 25 µl nuclease free water.

3.2.6 Concentration and purity determination of nucleic acids

The concentration of nucleic acids was determined using the NanoDrop spectrophotometer measuring absorption at 260 nm. The ratio E260/E280 served as estimation for the purity of nucleic acids. For pure solutions, a value above 1.8 was expected.

3.2.7 Polyacrylamide gel electrophoresis of RNA

50 µg of total RNA was mixed with 2 x bromophenol blue sample buffer and loaded onto a 15 x 17 cm, 15% polyacrylamide gel, together with a molecular weight scale consisting 1 fmol, 2 fmol, 4 fmol and 20 fmol of synthetic novel miRNAs. The gel was started with 10 W for 10 min and 30 W for 1 h in 1 x TBE buffer. Subsequently, the gel was stained with ethidium bromide, and the RNA bands were detected using the FLA 7000 system (Fujifilm).

3.3 Polymerase chain reaction (PCR)

3.3.1 Standard PCR

After cDNA synthesis, the PCR was used to amplify certain gene regions, using specific PCR primers or as much gene regions as possible using random hexamer primers. The standard PCR was used only as control PCR for first detection of Argonaute-associated genes. For these PCRs, the AccuPrime Super Mix I was used. The mixture contains anti-Taq DNA polymerase antibodies, thermostable AccuPrime protein, Mg2++, deoxyribonucleotide triphosphates (dNTPs), and recombinant Taq DNA polymerase at concentrations sufficient to allow amplification during PCR. The PCR reaction was performed in 25 µl containing ~ 10% of cDNA mix (5 µl), 1 x AccuPrime Super Mix I, 0.25 µM of each PCR primer and H2O. The annealing temperature for gene specific primer was 55 to 58°C. For PCR, 0.33 µM random hexamer primers and an annealing temperature of 20°C were applied. A standard PCR looked like this: 94°C for 5 min, 94°C for 40 sec, 55°C for 40 sec, 68°C for 40 sec, repeating steps 2 to 4 for 29 times, and 68°C for 7 min. A standard PCR with hexamer primers looked like this: 94°C for 5 minutes, 94°C for 1 min, 20°C for 1 min and 25 sec, 68°C for 1 min, repeating steps 2 to 4 for 29 times, and 68°C for 7 min (Gene Amp PCR System 2700, Applied Biosystems).

3.3.2 Quantitative reverse transcription-real time-PCR (qRT-PCR)

For qRT-PCR, the SYBR Green PCR Master Mix (Applied Biosystems) was used. The reaction was performed in 25 µl containing cDNA mix (1.5 µl), 1 x SYBR Green PCR master mix, 10 µM of each PCR primer and H2O. PCR cycling conditions were: 50°C for 2 min, 95°C for 10 min, 40 cycles of: 95°C for 30 sec, 58°C for 30 sec, 72°C for 30 sec and 72°C for 7 min, hold at 4°C. Subsequently, the 7500HT Real-Time PCR System (Applied Biosystems) was used.

3.3.3 TaqMan miRNA assay

The TaqMan miRNA Assays (Applied Biosystems) were used for detection of individual miRNAs, using qrt-RT-PCR, and performed according to the manufacturers' instructions (see TaqMan® MicroRNA Assays Protocol; Applied Biosystems). The reverse transcription of miRNAs was modified and differed from the manufacturer's instructions. 20 ng of total RNA

was needed. The master mix for the cDNA synthesis contained 1 mM dNTPs, 3.3 U/µl MultiScribe Reverse Transcriptase, 1 X Reverse Transcription buffer, 0.252 U/µl RNase Inhibitor and 20% TaqMan miRNA Primer. The next steps were performed according to the TaqMan® MicroRNA assay protocol (Applied Biosystems).

3.4 Methods for assessing novel miRNAs

3.4.1 Northern blot analysis of novel miRNAs

After gel electrophoresis, the separated RNA was transferred with 0.5 x TBE buffer to a Hybond N+ membrane (GE Healthcare), using a Transfer Blot SD Semi-Dry Transfer Cell (Bio-Rad). The blotting was performed with 300 mA for 2 h. After wards, the membrane was dried and UV crosslinked at a wavelength of 366 nm for a few sec. EDC cross linking was performed by adding 0.753 g EDC to 24 ml of 1-methylimidazole buffer. The membrane was soaked in EDC solution and incubated at 60°C for 1 h. For blotting, the DNA probes were labeled with ATP γ - 32P. The labeling mix was incubated for 1 h at 37°C. 30 µl of 30 mM EDTA (pH 8.0) was then added to the mix to stop the reaction. The labeled DNA probe was purified using an equilibrated MicroSpin G-25 column. The remaining PNK (polynucleotide kinase) was inactivated by incubation at 95°C for 1 min. For preparation of prehybridization and hybridization solution, salmon sperm DNA was denatured at 100°C for 5 min before adding to the hybridization solution. The hybridization solution was pre-warmed to 40°C to 50°C in a rotator. For prehybridization, 15 ml of hybridization solution was added to the dried, UV and EDC cross-linked membrane, and incubated at 40°C to 50°C for 2 h under permanent rotation. Subsequently, the hybridization solution was replaced by 15 ml fresh hybridization solution, containing the labeled DNA probe followed by incubation at 40°C to 50°C overnight under permanent rotation. The next day, the hybridization solution was removed, and the membrane was washed for 10 min with 100 ml washing buffer I, and then again for 10 min with 100 ml washing buffer II at 50°C. The membrane was wrapped in cling film and placed into a developer cassette with a white screen onto the membrane for 4 to 24 h. Following, a picture was taken of the exposed screen using the FLA 7000 system (Fujifilm). For stripping the DNA probes off the membrane, the membrane was washed for 5 min in boiling water containing 1% SDS. Finally, the membrane was exposed on the screen for 3 h again to check if the DNA probes disappeared.

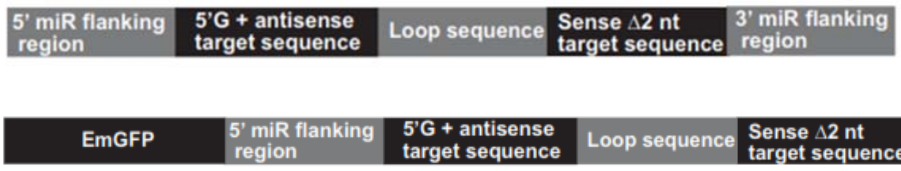
3.4.2 Knockdown of novel miRNAs

Raji and SU-DHL-4 cell lines were used for the downregulation of *NB-miR-5* and *NB-miR-19*. Downregulation was performed using miRCURY LNA™ miRNA Inhibitor that was obtained from Exiqon. 2 x 10^6 cells of each cell line were pelleted via centrifugation at 90 x g for 10 min at room temperature. Raji and SU-DHL-4 cells were transfected with the Cell Line Nucleofector® Kit V (Lonza), programs M-013 and M-13, respectively. Cells were resuspended in 100 µl of room-temperature Nucleofector®Solution and then nucleofected by using Amaxa Nucleofector II machine with a concentration of 30 nM of LNA oligonucleotide following the manufacturer's instructions. Cells were analyzed 24 h post nucleofection using fluorescence microscopy.

3.4.3 Cloning of novel miRNAs

Novel miRNAs (*NB-miR-5* and *NB-miR-19*) were cloned into pcDNA™6.2-GW/EmGFP-miR using the BLOCK-iT™ Pol II miR RNAi Expression Vector Kits. Single-stranded DNA oligos encoding the pre-miRNA of interest were designed according to manufacturer's instructions.

The engineered pre-miRNA sequence structure is based on the murine miR-155 sequence. We designed the ds-oligo encoding the engineered pre-miRNA according to the following structural features: a four nucleotide, 5'overhang (TGCT) complementary to the vector (required for directional cloning), a 5'G plus short 21 nucleotide antisense sequence (mature miRNA), followed by a short spacer of 19 nucleotides to form the terminal loop and a short sense target sequence with two nucleotides removed (Δ2) to create an internal loop. The pre-miRNA sequence and adjacent miR-155 flanking regions are denoted as the pre-miRNA expression cassette and are shown below (Figure 3.1)



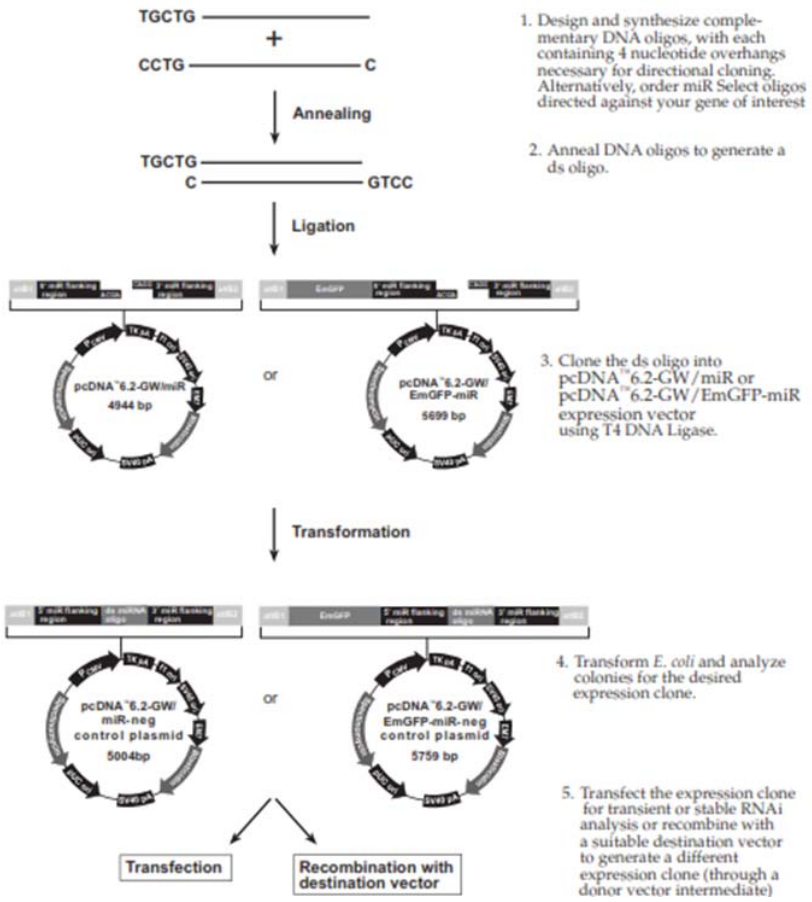


Figure 3.1 Cloning of novel miRNAs. The major steps necessary to produce a pcDNA6.2GW/miR or pcDNA 6.2-GW/EmGFP-miR expression clone using the BLOCKiT Pol II miR RNAi Expression Vector Kits

Once the engineered pre-miRNA expression cassette is introduced into the mammalian cells for expression, the pre-miRNA forms an intramolecular stem loop structure similar to the structure of endogenous pre-miRNA that is then processed by the endogenous Dicer enzyme into a 22 nucleotide mature miRNA. In order to determine the optimal transfection conditions, Raji (BL) and SU-DHL-4 (non-BL) were transfected with the destination vector.

3.4.3.1 Annealing the single-stranded oligos to generate ds-oligos

To generate a double-stranded oligos, the following annealing reaction was performed:

- 200 µM top strand oligo 5 µl
- 200 µM bottom strand oligo 5 µl
- 10 X oligo annealing Buffer 2 µl
- DNase/RNase-free water 8 µl
- total volume 20 µl

The reaction mixture was incubated at 95°C for 4 min and subsequently cooled at room temperature for 5-10 min. After spinning down the sample in a microcentrifuge for 5 sec, ds oligo mixture was diluted 5,000-fold by performing serial 100-fold and 50-fold dilutions; the first into DNase/RNase-free water and the second into 1X oligo annealing Buffer. Final concentration was 10 nM.

3.4.3.2 Clone the ds oligo into pcDNA™6.2-GW/EmGFP miR

The following ligation reaction was set up for cloning the ds oligo into pcDNA™6.2-GW/EmGFP miR

- 5X Ligation Buffer 4 µl
- pcDNA™6.2-GW/± EmGFP-miR (5 ng/µl), linearized 2 µl
- ds oligo (10 nM; 1:5,000 dilution) 4 µl
- DNase/RNase-Free water 9 µl
- T4 DNA Ligase (1 U/µl) 1 µl
- total volume 20 µl

After mixing the reaction, it was incubated for 5 min at room temperature

3.4.3.3 Transform One Shot® TOP10 Chemically Competent E. coli

2 µl of the ligation reaction was added to a vial of one Shot® TOP10 chemically competent E. coli and mixed gently. The reaction tube was incubated on ice for 5 to 30 min. Cells were heat-shocked for 30 sec at 42°C without shaking. Immediately afterwards, the tube was transferred to ice. Next, 250 µl of room temperature S.O.C. medium was added to the cells and incubated at 37°C for 1 h with shaking. 20-100 µl of bacterial culture was spread on a pre-warmed LB agar plate containing 50 µg/ ml spectinomycin and incubated overnight at 37°C.

3.5 Sample preparation for sequencing

3.5.1 Processing of frozen tumor samples

Frozen material of each DLBCL, BL and FL was collected. Fresh frozen tissue specimen of each single case was reviewed by means of histology and immunohistochemistry, employing

antibodies to determine its cellular composition (e.g. CD20, CD3, CD68, Ki-67). Frozen tissue sections were directly used for RNA and DNA extraction.

Total RNA of each sample was extracted from the specimens by the AMBION mirVana Kit. For cell lysis, the lysis/binding solution provided by AMBION, which stabilizes RNA and inactivates RNase, was used. After elution of the total RNA with nuclease-free water, the RNAs were stored at -80°C.

These steps were performed by the respective ICGC-MMML-Seq work package as specified above.

3.5.2 Quality control of the extracted RNA

The quality of the total RNA was analyzed by the Agilent BioAnalyzer 2100. Visual inspection of the BioAnalyzer profiles as well as consideration of the derived RIN values were used to evaluate RNA quality. These steps were performed by the respective ICGC-MMML-Seq work package as specified above.

3.5.3 MiRNA library preparation

Library preparation was performed using the TruSeq Small RNA Sample prep kit (Illumina, San Diego, CA, USA) with 1 µg of total RNA serving as an input.

3.5.4 3'-adapter ligation for cDNA library preparation

As a first step, 3'-adapter was ligated using T4 RNA ligase 2-truncated. For each library, 1 µg of RNA and 3'-adaptor was incubated at 70°C for 2 min and placed immediately on ice.

Reagent	Volume (µl)
RNA 3' Adapter (RA3)	1
1 µg Total RNA in Nuclease-free Water	5
Total volume	6

For 3'-ligation the following mix was prepared. 3'-ligation reaction was incubated on the pre-heated thermal cycler at 28°C for 1 h.

Reagent	Volume (µl)
5 X HM Ligation Buffer (HML)	2
RNase Inhibitor	1
T4 RNA Ligase 2, truncated	1
Total volume	6

With the reaction tube remaining on the thermal cycler, 1 µl of stop solution (STP) was added and the reaction incubated at 28°C for 15 min.

3.5.5 5'-adapter ligation for cDNA library preparation

In this step, the 5'-adapter is joined to the 3'-ligated RNA to enable the cDNA synthesis in the next step. Before starting the reaction, the 5'-adapter was incubated on the pre-heated thermal cycler at 70°C for 2 min and immediately placed on ice. Ligation of 5'-adapter was performed using T4 RNA ligase 1 and ATP. The reaction tube was incubated on the pre-heated thermal cycler at 28°C for 1 h and then placed on ice.

3.5.6 Reverse transcription and amplification

Reverse transcription followed by PCR was performed to create cDNA constructs based on the small RNAs ligated to 3' and 5' adapters. According to this protocol, small fragments that have adapter molecules on both ends were selectively enriched. First, RNA RT primer was added to the 5' and 3' adapter-ligated RNA, then the reaction tube was incubated at 70°C for 2 min and placed immediately on ice.

Reagent	Volume (µl)
5' and 3' adapter-ligated RNA	6
RNA RT Primer (RTP)	1
Total volume	7

After adding the primer, the following mixture was prepared for performing reverse transcription.

Reagent	Volume (µl)
5X First Strand Buffer	2
12.5 mM dNTP mix	0.5
100 mM DTT	1
RNase Inhibitor	1
Super Script II Reverse Transcriptase	1
Total volume per Sample	5.5

The reaction was incubated in the pre-heated thermal cycler at 50°C for 1 h and placed on ice.

3.5.7 PCR amplification

Reagent	Volume (µl)
Ultra-Pure Water	22.5
5X Phusion HF Buffer	10
RNA PCR Primer (RP1)	2
RNA PCR Primer Index (RPIX)	2
25 mM dNTP	0.5
Phusion DNA Polymerase	0.5
Total volume per Sample	37.5

For each reaction, only one of the 48 RNA PCR primer indices was used during the PCR step.

Index	Sequence	Index	Sequence	Index	Sequence
RPI1	ATCACG	**RPI17**	GTAGAG	**RPI33**	TCATTC
RPI2	CGATGT	**RPI18**	GTCCGC	**RPI34**	TCCCGA
RPI3	TTAGGC	**RPI19**	GTGAAA	**RPI35**	TCGAAG
RPI4	TGACCA	**RPI20**	GTGGCC	**RPI36**	TCGGCA
RPI5	ACAGTG	**RPI21**	GTTTCG	**RPI37**	GCCGCG
RPI6	GCCAAT	**RPI22**	CGTACG	**RPI38**	GCCTTA
RPI7	CAGATC	**RPI23**	GAGTGG	**RPI39**	GCTCCA
RPI8	ACTTGA	**RPI24**	GGTAGC	**RPI40**	GGCACA
RPI9	GATCAG	**RPI25**	CTAGCT	**RPI41**	GGCCTG
RPI10	TAGCTT	**RPI26**	CTAGCT	**RPI42**	TCTACC
RPI11	GGCTAC	**RPI27**	CTATAC	**RPI43**	TGAATG
RPI12	CTTGTA	**RPI28**	CTCAGA	**RPI44**	TGCCAT
RPI13	AGTCAA	**RPI29**	CTGCTG	**RPI45**	TGCTGG
RPI14	AGTTCC	**RPI30**	TAATCG	**RPI46**	TGGCGC
RPI15	ATGTCA	**RPI31**	TACAGC	**RPI47**	TTCGAA
RPI16	CCGTCC	**RPI32**	TATAAT	**RPI48**	TTCTCC

The reaction was run according to the following temperature program: 30 sec at 98°C, 15 cycles of: 10 sec at 98°C, 30 sec at 60°C, 15 sec at 72°C, 10 min at 72°C and hold at 4°C. For checking the library before gel fractionation, each sample was run on a high sensitivity DNA chip according to the manufacturer's instructions.

3.5.8 Purification of cDNA constructs

50 µl of PCR reaction was mixed with 5X loading dye and loaded onto a 6% Novex®TBE gel together with a 1 µl of High Resolution Ladder and 2 µl of Custom Ladder. Two wells were loaded with 30 µl each of amplified cDNA. 6% PAGE Gel was run at 145 V for 60 min in 1X TBE buffer.

Using a clean scalpel, bands corresponding to approximately the size of the adapter-ligated constructs derived from the 22 nt and 30 nt small RNA fragments (between 145bp and 160 bp of custom ladder) were cut out. The band of interest was placed into the 0.5 ml gel breaker

tube. The tube was centrifuged at 14,000 rpm in a microcentrifuge for 2 min at room temperature. After adding 300 µl of ultra-pure water to the gel pieces in the 2 ml tube, the tube was incubated at 4°C overnight. Next day, eluate and gel pieces were passed through Corning®Costar®Spin-X® tube filters (Sigma Aldrich). Precipitation was performed by the addition of 2 µl of Glycogen (20 mg/ml; Invitrogen), 30 µl of 3 M NaOAc and 975 µl of pre-chilled -15°C to -25°C 100% EtOH. After washing with 75% EtOH, the pellet was air dried at 25°C and dissolved in 20 µl EB buffer (Qiagen).

3.5.9 Library validation

The purified PCR products were quantified by loading 1 µl of the suspended construct on Agilent Technologies 2100 Bio analyzer using a DNA specific chip such as the DNA-1000 or High Sensitivity DNA chip.

3.5.10 Denaturing the template DNA

For denaturation, the template DNA with 0.1 N NaOH to the final DNA concentration of 20-pM, 10 µl of 2 nM template DNA added to the 10 µl of 0.1 N NaOH. The tube was first vortexed briefly to mix the template solution and then centrifuged at 280 x g for 1 min. Template solution was incubated at room temperature for 5 min. After the incubation time, 980 µl of pre-chilled HT1 (Hybridization Buffer) was added to the 20 µl of template solution.

3.5.11 cBot

To make a dilution of denatured DNA with pre-chilled HT1 to a total volume of 1,000 µl, the denatured DNA was diluted to the desired concentration using the following example:

Final Concentration (pM)	**6**	**7**	**10**	**12**	**15**	**18**	**20**
20 pM denatured DNA(µl)	300	350	500	600	750	900	1000
Pre-chilled HT1(µl)	700	650	500	400	250	100	0

120 µl of the 7 pM template DNA were dispensed into 0.2 ml eight-tube strips. After loading the cBot reagent, the cBot machine was run for 4h and S.R.Amp-Lin/Block/Hyb.V8 program was selected for loading the sample into the flowcell.

3.5.12 Deep sequencing

High-throughput sequencing was performed on a Hiseq 2500 for 51 cycles plus 7 cycles to read the indices.

3.6 PAR-CLIP

3.6.1 Thionucleoside labeling

Four lymphoma cell lines (SU-DHL4, SU-DHL6, Raji and NAMALWA) were grown in DMEM with 10% FBS. For each cell line, 6 (75cm) flasks were used. Stock solution of 4-thiouridine (4SU) (Sigma, T4509) was diluted in DMSO to 1M and kept at -20ºC. For labeling, 4SU was diluted in FBS and added to the cells to a final concentration of 100 µM. 4SU labeling was accomplished overnight.

3.6.2 Crosslinking, cell lysis and immunoprecipitation

After labeling, the medium was aspirated from the plates and the cells were crosslinked on ice with 365nm UV-lamps (energy setting 1500 µJ x 100/cm2). Cells were washed in cold PBS and pelleted by centrifugation. For cell lysis, the cell pellets were resuspended in 3 cell pellet volumes of NP40 lysis buffer and incubated on ice for 10 min. Cell lysates were cleared by centrifugation at 13,000 x g for 15 min at 4ºC; lysates were then cleared further by filtering through a 5 µm membrane syringe filter. Cleared lysates were partially digested with RNaseT1 (final concentration of 100 U/µl (Fermentas, 10,000 U/µl)) for 15 min at 22ºC in a water bath and then cooled on ice for 5 min. AGO2 protein complexes were immunoprecipitated from filtered cell lysates using antibodies conjugated to magnetic protein G dynabeads (Invitrogen). Anti-AGO2 antibody (C1.9E8.2, Millipore) was added to dynabeads to a final concentration of 0.25 mg/ml. Lysates were incubated with antibody coupled beads for 1 h at 4ºC. In PAR-CLIP experiments, the ionic strength of wash buffer is one of the important factors for the antibody-antigen interactions and it is variable based on the target in the experiment, e.g., the anti-FLAG M2 antibody will bind its antigen in the presence of 500 mM, NaCl, while the anti-AGO2 antibody should not be used with salt concentrations higher than 300 mM NaCl (Hafner, Landthaler et al. 2010). In order to ensure to not exceed the maximal salt concentration at which the antibody recognizes its antigen, both IP wash buffer and High salt wash buffer were substituted by NP40-lysis buffer, and the number of wash steps was reduced from three and five times to two and three times,

respectively. Beads were washed two times with NP40-lysis buffer and treated with RNase T1 in one volume of NP40-lysis buffer at a final concentration of 50 U/µl for 15 min at 22ºC. The beads were immediately washed two times with NP40-lysis buffer and resuspended in NEB buffer #3 containing 0.5 U/µl calf intestinal phosphatase (NEB). Dephosphorylation was performed at 37ºC in a thermomixer (Eppendorf) with shaking at 900 rpm for 1h. Beads were then washed twice with NP-40 lysis buffer and then labeled with [ɤ-32P]-ATP (Hartmann Analytic GmbH) in PNK buffer. Gamma-ATP was used at a final concentration of 0.1µCi/µl. The labeling reaction was performed at 37ºC with for 1 h. Regular ATP (Fermentas) was added to a final concentration of 100 µM and the reaction was further incubated for 5 min. The beads were washed 5 times with 800 µl PNK buffer without DTT and resuspended in 35 µl of 2 X SDS-PAGE loading buffer.

3.6.3 SDS-PAGE and electroelution of RNA

Beads in SDS loading buffer were boiled at 95ºC for 5 min and the supernatant was loaded onto an SDS gel (NuPAGE Novex 4-20% BT Gel, Invitrogen). Gels were exposed for 2-5 min onto a phosphorimaging screen and visualized on the FLA 7000 imager. The radioactive band corresponding to AGO2 (97 KD) was cut out of the gel. AGO2-RNA complexes were electroeluted from the gel using D-Tube Dyalyzer Kit MWCO 3.5kDa (Novagen) for 2 h at 100V in SDS running buffer (25mM Tris base, 192 mM glycine, 0.1% SDS). The electroeluate (~750 µl) was combined with 2 X proteinase K buffer (200 mM Tris-HCl pH 7.5, 150 mM NaCl, 12.5 mM EDTA, 2% SDS) containing proteinase K (Roche, 03115879001) to yield a final concentration of 2 mg/ml. The reaction was incubated at 55ºC for 30 min. Immunoprecipitated RNA was recovered by phenolchloroform extraction and ethanol precipitation, also using glycoblue (Ambion). The PAR-CLIP libraries were sequenced on an Illumina Hiseq2500 using the 50 cycles single read protocol.

3.6.4 RNA seq-Data analysis

After sequencing and in order to trim adapter sequences and low quality ends, we applied the cutadapt program (http://code.google.com/p/cutadapt/). Since very short reads were not informative, all reads bellow 18 bases in length were removed. Next, all reads were aligned against the reference sequence of the human genome 19 (hg19). This identifies most of the expressed transcripts of the sample with the read number as an approximation of the

expression of each transcript. This step was performed using the short read aligner TopHat, which is also aware of splice junctions that could be spanned by individual reads (Trapnell, Pachter et al. 2009). Using a combination of the two programs HTSeq and edgeR, expression values were counted, normalized appropriate to the library size and tested for differential expression compared to the reference mock samples of the same cell line.

4 Results

4.1 MiRNA expression profiling of GC lymphoma subtypes

4.1.1 Global miRNA expression analysis by deep sequencing

Sequencing based technologies provide an unbiased approach for studying of small non-coding RNAs such as miRNAs. To uncover global miRNAs expression as well as new miRNAs in GCB samples, we performed a large-scale miRNA expression profiling analysis by next generation sequencing on 56 samples, including 21 FL (18 grade 1/2, 1 grade 1/2/3A,1 grade 2/3A, 1grade 3A), 17 BL, 18 DLBCL (10 ABC, 8 GCB) specimens and five normal germinal center B-cell samples, which were isolated from surgically extracted tonsils and then purified by fluorescence-activated cell sorting (FACS) as being CD20+CD38+ (Supplement Table S. 1). Libraries were generated using Illumina TruSeq Small RNA Sample Prep Kit. Bar-coded libraries were mixed, pooled in equimolar (10 nmol/l) amounts and diluted to 6.5pmol/l for cluster formation on a single-flow cell lane.

In order to identify new classes of small-non coding RNAs, libraries were prepared with two distinct sized fractions, the “classic” miRNA size fraction of 18-35 nucleotides and a larger size fraction of 35-90 nucleotides, herein after referred to as small fragment and large fragment, respectively. Small fragment sequencing data from the Illumina HiSeq 2500 was processed and de-multiplexed using the CASAVA1.6 software (Illumina) pipeline to generate raw fastq reads.

To directly identify known miRNAs, the mapped reads were aligned to known miRNA sequences in the miRNA database (miRBase v19), using BamUtils from the NGSUtils-0.5.5 package ("multiple partial" option) (Breese and Liu 2013). The matched sequence reads were derived from data sets, and remaining unmatched reads were used later for identifying novel miRNAs. In the next step, we used the cutadapt program (http://code.google.com/p/cutadapt/) to remove the ligated adapter sequences. Furthermore, we filtered out reads that aligned against various other RNA species, including other non-coding small RNA species and RNA degradation products using the data from University of Santa Cruz (UCSC) Genome Browser.

After filtering and trimming of the reads for low quality and adaptors from the 56 classical size-selected sequencing libraries, we obtained a total of 1,340,000,000 high confidence single-end reads, equivalent to ~23,900,000 reads mapping to miRNAs per sample. For sample comparison we required 5 million or more total miRNA sequence reads per library and libraries with less than 5 million reads were resequenced.

Quality control checks on raw sequence data were carried out using the FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). As we expected, most of the reads were located in ncRNA, intronic and intergenic regions, Quality analysis showed that miRNAs typically constituted more than 50% of all non-coding RNAs (Figure 4.1.A).

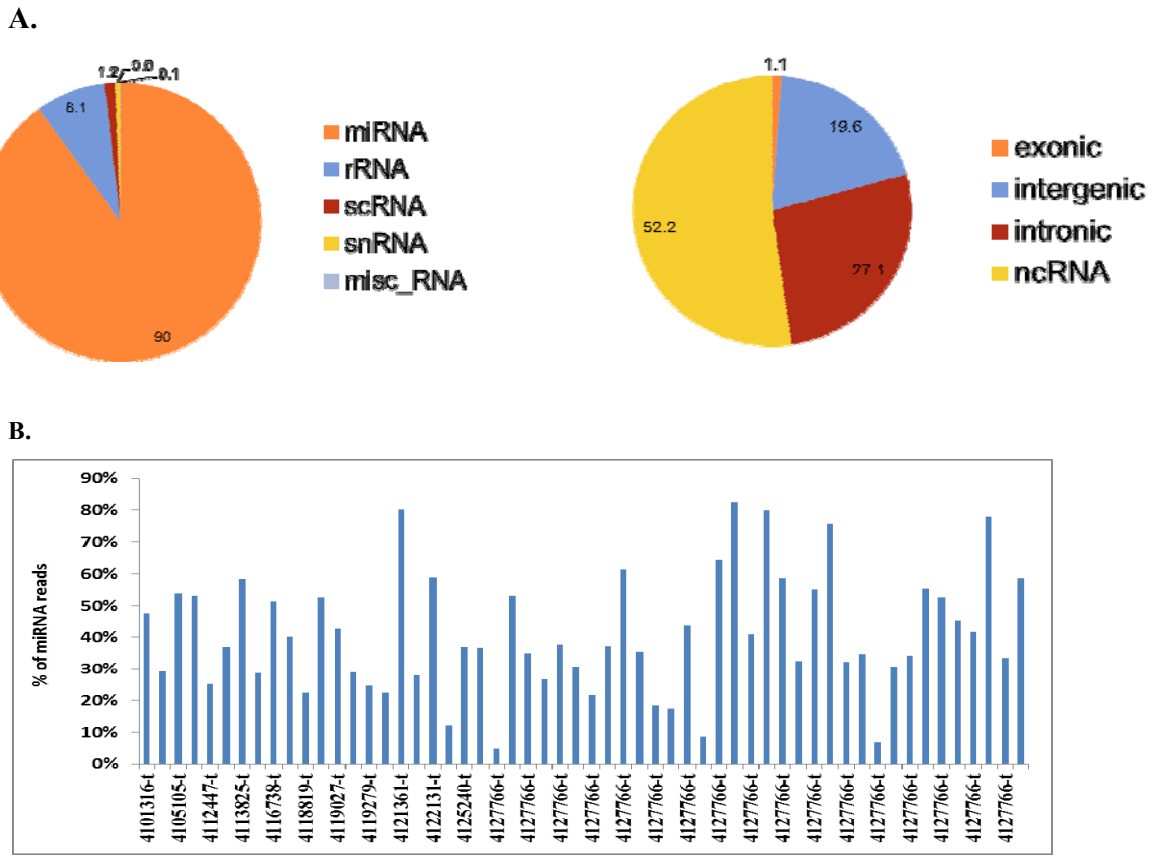


Figure 4.1 Quality analysis of miRNAs**.** (A) The left pie-chart represents percentages of the different RNA classes found in samples, the right pie-charts represent the mapped read distributions of protein-coding genes (CDS; orang), and intergenic regions (blue), interonic (red) and non-coding RNA (yellow). (B) Quality of all micrna reads in mapped reads (in percent) after filtering and other pre-alignment processing steps such as adapter clipping.

We identified the frequencies of reads by considering different classes of small RNAs or other genomic regions and calculated the average frequencies, comparing all 56 tumor samples. The

frequency of reads mapping to mature miRNAs ranged from 5 to 90% in the libraries and gave an average of 41% (Figure 4.2). Size distribution analysis showed that the matched sequences predominantly consisted of reads approximately 22-23nt in length. All of the high-quality clean reads larger than 18 nucleotides were mapped to the human genome. Figure 4.2.C shows size distribution of four patient samples.

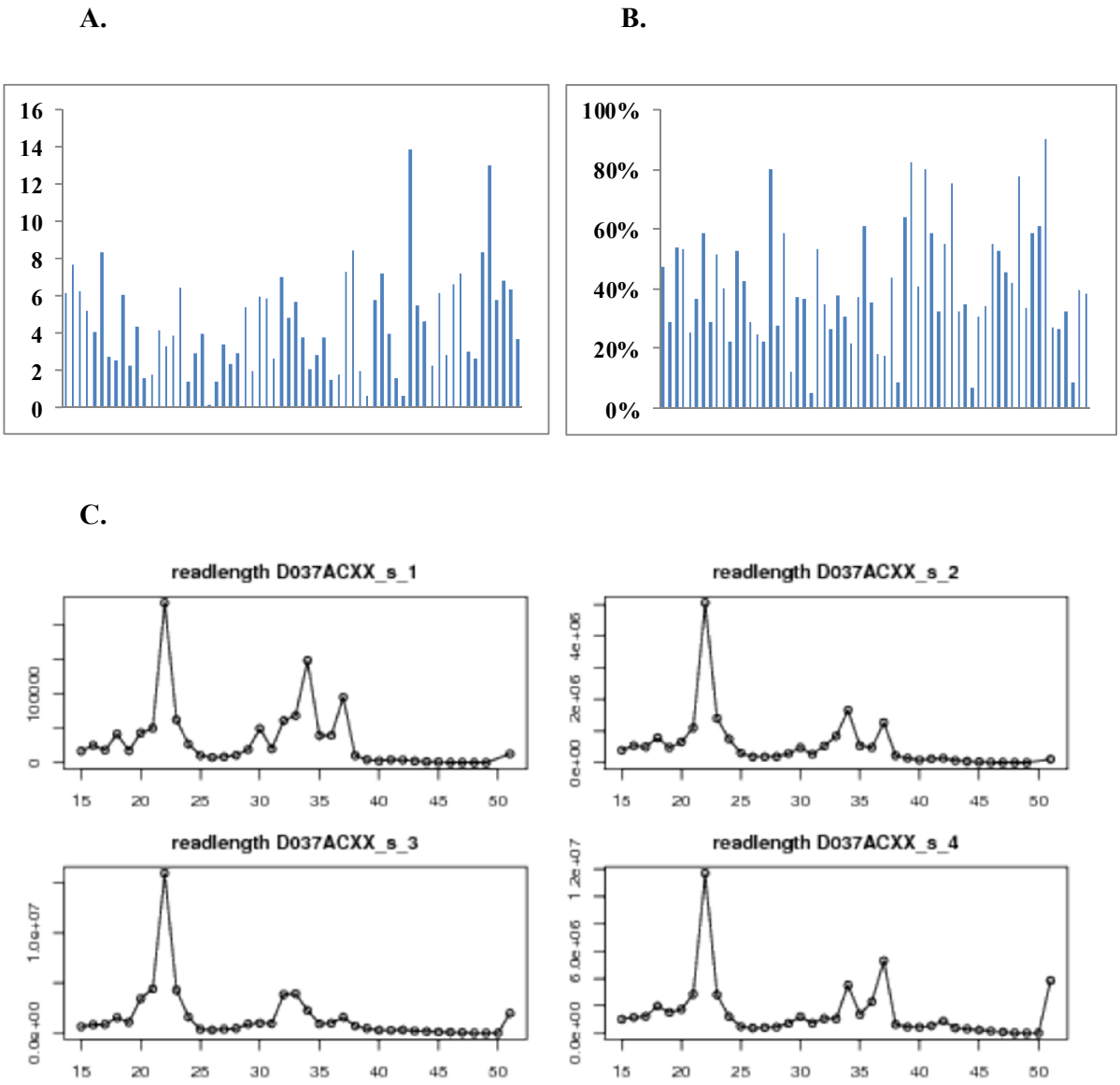


Figure 4.2 Overview of mapped reads, miRNAs and frequencies of RNA classes. (A) Number of reads ($x10^6$) mapped to the human genome (hg19) for all samples. These include all RNA species (premature miRNA, snoRNA, snRNA, miscRNA, rRNA, tRNA and mRNA). (B) Frequencies of reads mapped to annotated mature miRNAs (in percent) for all samples using the miRNA database. (C) Size distribution (bp) shown for sequence reads that matched known miRNAs. Examples of four individual libraries are given.

4.1.2 MiRNA expression differences between BL and non-BL

We performed unsupervised hierarchical clustering for all 56 samples (Figure 4.3). Most of the miRNAs showed very similar expression levels across subtypes of GCB lymphoma samples, therefore they did not cluster apart. Next, we performed statistical analyses to clarify differentially expressed miRNAs in the three GCB-lymphoma subtypes analyzed.

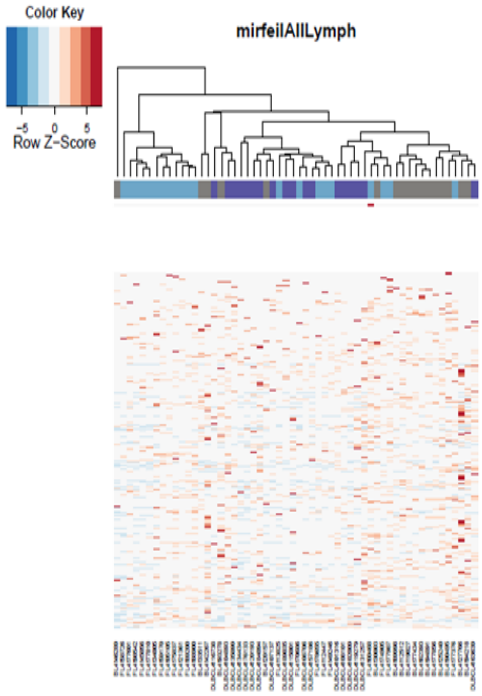


Figure 4.3 MiRNA expression profile across all subtypes

The first step of the analysis was to identify a set of miRNA with significant differences in expression level between BL and non-BL. After analysis, our result demonstrated eight miRNAs differentially expressed between BL and non-BL (Table 4.1). Three miRNAs (*miR-548ac, miR-150 and miR-211)* were expressed at a significantly higher level in non-BL than in BL cases, while expression of *miR-1275, miR-1258, and miR-1244-1/2/3* was reduced in non-BL entities. Our analysis of these data for *miR-150* is shown in Figure 4.4 and demonstrated that *miR-150* is highly expressed in non-BL cases compared to BL cases. Comparing DLBCL and FL with one another (Figure 4.4) showed *miR-150* strongly downregulated in DLBCL. Further differentially expressed miRNAs include e.g. *miR-211, miR-548ac* and *miR-1244-1/3*. Relatively little is known about those miRNAs and they represent novel de-regulated targets in lymphomagenesis.

Table 4.1 Top eight differentially expressed miRNAs in BL versus non-BL lymphomas. The list is ranked by p-value. Negative fold changes represent miRNAs that are downregulated in non-BL samples and positive values those which are upregulated compared to BL samples.

miRNA ID	log2 fold change	p-value	location	hematologic malignancy
miR-1275	-3.83	4.38e-05	chr6:33967749-33967828	not yet associated
miR-1258	-2.71	1.45e-04	chr2:180725563-180725635	not yet associated
miR-548ac	4.01	5.30e-04	chr1:117102646-117102733	not yet associated
miR-150	3.26	4.48e-04	chr19:50004042-50004125	Lymphoma, AML, MLL, leukemia
miR-211	4.24	2.64e-05	chr15:31357235-31357344	not yet associated
miR-1244-2	-3.86	1.26e-05	chr12:12264886-12264970	not yet associated
miR-1244-3	-3.8	2.02e-05	chr2:232578024-232578108	not yet associated
miR-1244-1	-4	6.47e-06	chr2:232578024-232578108	not yet associated

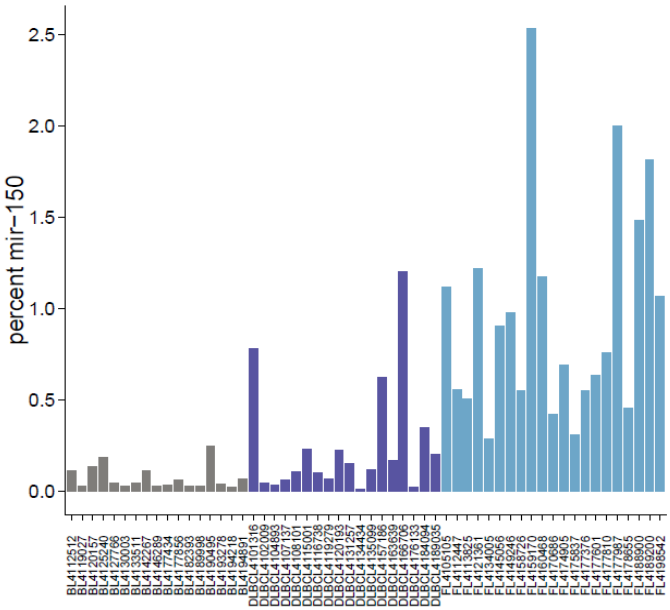


Figure 4.4 Expression of *miR-150* in FL, DLBCL and BL samples (in percent)

4.1.3 MiRNA expression differences between DLBCL and FL

In a second step of analysis, we assessed miRNA deregulation between non-BL subtypes including FL and DLBCL and also compared each of them separately to BL (Supplement Table S.2). For this purpose, we analyzed differential miRNA expression between DLBCL and FL. Our data showed that 87 ($p<0.001$) miRNAs were significantly dysregulated between these two lymphoma subtypes. FL was characterized mostly via downregulation of miRNAs (51 of 87 miRNAs). MiR-150 that was previously demonstrated to be differentially expressed in FL and DLBCL (Roehle, Hoefig et al. 2008) was shown to be overexpressed in FL.

Interestingly; we observed the overexpression of components of the *miR-17-92* cluster (*miR-17, miR-18a and miR-20a*) in DLBCL in comparison to FL. In addition, comparing FL to non-tumoral samples showed three miRNAs of the *miR-17-92* cluster (*miR-17, miR-18a and miR-92a*) and two paralog components (*miR-106a* and *miR-106b*) were downregulated in FL.

4.1.4 MiRNA expression differences between DLBCL and BL

98 miRNAs were found to be differentially expressed ($p<0.001$) between BL and DLBCL. Eight miRNAs (*miR-146b*, *miR-221, miR-222, miR-30d, miR-26a-2, miR-339, miR-9 and miR-193a*) were previously reported to be dysregulated between BLs and DLBCLs (Lenze, Leoncini et al. 2011). Among the others, *miR-9* and *miR-193* were identified to be upregulated in DBCL vs. BL while Lenze *et al.* reported these two miRNAs as being upregulated in BL. The other miRNAs that were previously reported included *miR-573, miR-17 and miR-92, miR-9, miR-26* and *miR-374.*

4.1.5 MiRNA expression differences between FL and BL

Based on a 2-fold or greater difference in mean values in FL samples compared to BL ($P<0.001$), we found that 108 miRNAs were significantly differentially expressed in these two subtypes, of which 65 miRNAs were downregulated and 42 miRNAs showed upregulation in FL.

4.1.6 High expression of *miR-143*

The most strongly upregulated miRNA compared to non-tumoral samples was *miR-143*, in germinal center or naive B-cells sample read counts for *miR-143* accounted for 0.05% of all sequenced reads, suggesting that *miR-143* is hardly expressed in normal samples. Expression of *miR-143* in lymphoma samples is shown in Figure 4.5; *miR-143* expression in tumoral samples was up to 68% of all sequenced reads.

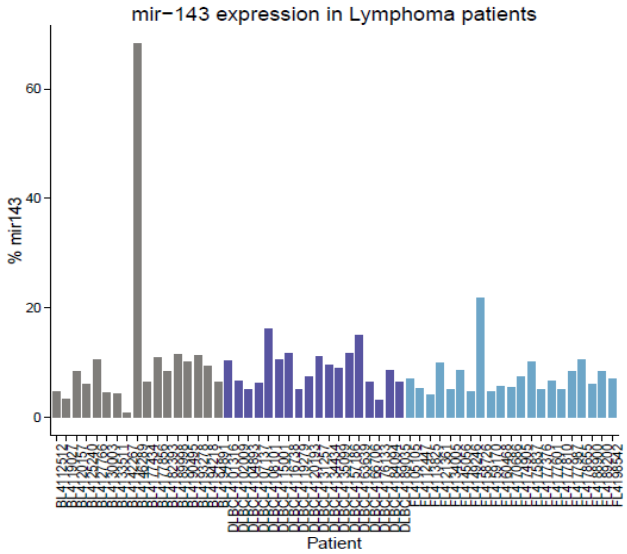


Figure 4.5 *miR-143* expression in three subtypes of GCB-lymphomas (in percent).

4.1.7 *MiR-142* is recurrently mutated in its mature sequence in DLBCL as well as FL

We focused our attention on mutations present in either the transcriptome and/or the miRNA datasets. First, we searched for those mutations which could be found on both DNA (whole genome sequencing) as well as RNA levels (RNAseq, miRNAseq). In total, we identified 17 mutations in 12 patients (6 mutations in 5 DLBCL patients, 4 mutations in 3 FL patients, 7 mutations in 4 BL patients) corresponding to eight distinct miRNAs (*miR-142, miR-612, miR-3655, miR-4322, miR-4537, miR-4538, miR-4539 and miR-5195*) (Table 4.2). In four cases, the mutations were within the mature sequence (*miR-142, miR-612, miR-4537 and miR-4538*). *MiR-142* was the most frequently mutated miRNA with a total of seven different mutations in six patients. Two of those mutations were located within the seed sequence. This corresponds to a mutation frequency of 15,4% in non-BL; looking at the subgroups, this breaks up into 22,2% in DLBCL and 9,5% in FL. *MiR-612* (only mutated in DLBCL; mutation frequency 5,6%), *miR-4538* (BL, mutation frequency 5,9%) and *miR-4537* (BL) showed one mutation each in the mature sequence. In the case of the latter, a second BL patient carried three additional mutations outside of the mature sequence. This corresponds to a mutation frequency for *miR-4537* of 11, 8% in BL patients. For four other miRNAs (*miR-3655, miR-4322, miR-4539, miR-5195*) we could only detect mutations outside the mature regions and in only one patient each.

Table 4.2 List of miRNAs with mutations and location of mutations.

Name	mature	chr	nt	type	pid	diagnosis	subtype	Mut freq mRNA	Mut freq miRNA	Rel cov miRNA	Rel cov mRNA
miR-142	n	Chr17	56408638	G->A	4101316	Non-BL	DLBCL	15,9%	6,3%	1760,59	0,67
miR-142	y	Chr17	56408624	C->T	4102009	Non-BL	DLBCL	44,4%	56,8%	2203,64	3,46
miR-142	y	Chr17	56408616	A->C	4112447	Non-BL	FL	23,0%	77,3%	2689,82	1,66
miR-142	n	Chr17	56408630	C->T	4120193	Non-BL	DLBCL	29,8%	15,6%	1711,43	2,29
miR-142	y	Chr17	56408620	A->T	4160468	Non-BL	FL	14,2%	0,0%	11968,60	0,69
miR-142	y	Chr17	56408621	A->G	4160468	Non-BL	FL	14,4%	0,1%	11968,60	0,69
miR-142	y	Chr17	56408612	A->T	4176133	Non-BL	DLBCL	38,7%	11,7%	1060,39	1,54
miR-3655	n	Chr5	140027478	A->G	4177376	Non-BL	FL	18,4%	0,0%	0,06	0,77
miR-4322	n	Chr19	10341109	C->T	4135099	Non-BL	DLBCL	16,7%	100,0%	0,00	0,06
miR-4537	n	Chr14	106325659	G->A	4189998	BL	BL	39,1%	22,2%	0,07	4,08
miR-4537	n	Chr14	106325674	C->G	4189998	BL	BL	41,1%	9,1%	0,07	4,08
miR-4537	n	Chr14	106325693	C->G	4189998	BL	BL	47,8%	100,0%	0,07	4,08
miR-4537	y	Chr14	106325609	G->A	4193278	BL	BL	84,3%	100,0%	0,02	4,11
miR-4538	y	Chr14	106324395	C->T	4146289	BL	BL	44,2%	0,0%	0,00	2,41
miR-4539	n	Chr14	106323713	C->G	4182393	BL	BL	29,5%	16,7%	0,08	1,78
miR-5195	n	Chr14	107259212	C->G	4182393	BL	BL	48,8%	90,9%	0,14	3,20
miR-612	y	Chr11	65211962	G->A	4135099	Non-BL	DLBCL	21,1%	85,7%	0,06	0,13

Next, we examined for mutations which were seen exclusively on the RNA (and not the DNA) level and which thus represent bona fide miRNA editing events. Starting with several thousand candidate SNVs (all SNVs sequenced in our miRNA data), we limited our search to those SNVs in the seed regions (within nucleotides 2 to 7, which are important in conferring binding-site specificity of the respective miRNAs to their target mRNAs) and excluded known SNVs (dbSNP) as well as rare (below 10%) variants. The last 40 candidates were manually evaluated (correct position of SNVs in sequence reads, sequencing quality of errors etc.) and the list was thus narrowed down to seven SNVs (Table 4.3). Of the two changes other than A-to-H (with H being A, C or T), the first one (chr11: 93466874) was only found in one sample while the second one being more common (chr14: 77732582) and was observed in all three subtypes. The five SNVs which represented A-to-H changes mapped to five miRNAs (*miR-1260b, miR-4772, miR-376a1, miR-376a2 and miR-376c*), with the three latter all belonging to the same genomic cluster. According to our data editing frequencies ranged from 31-86%.

Table 4.3 miRNA editing. The seven SNVs correspond to different miRNAs. Two of five SNVs show other than A-to-H (with H being A, C or T) changes and the rest show A-to-H changes.

chrom	location	miRNA	type	% alternative	BL	FL	DLBCL
chr2	103048801	miR-4772	A->C	31,0%	0	0	1
chr11	93466874	miR-1304	T->C	68,4%	0	0	1
chr11	96074619	miR-1260b	A->G	35,3%	0	3	0
chr14	101506074	miR-376c	A->G	45,2%	3	8	2
chr14	101506460	miR-376a2	A->G	86,2%	11	7	5
chr14	101507167	miR-376a1	A->G	86,2%	11	7	6
chr14	77732582	miR-1260a	T->G	98,7%	9	12	6

4.1.8 Validation of NGS data using qRT-PCR

Validation of our NGS data was performed by real-time RT-PCR analysis.We selected a subset of dysregulated miRNAs (*miRs-10b, -141a, -142b, -143, -146a, -150, -21,-22b-5p, -28, and -92a*). According to the NGS data, we prepared two lists of samples that showed the minimum and maximum expression levels of each selected miRNAs (Table 4.4 & Table 4.5).

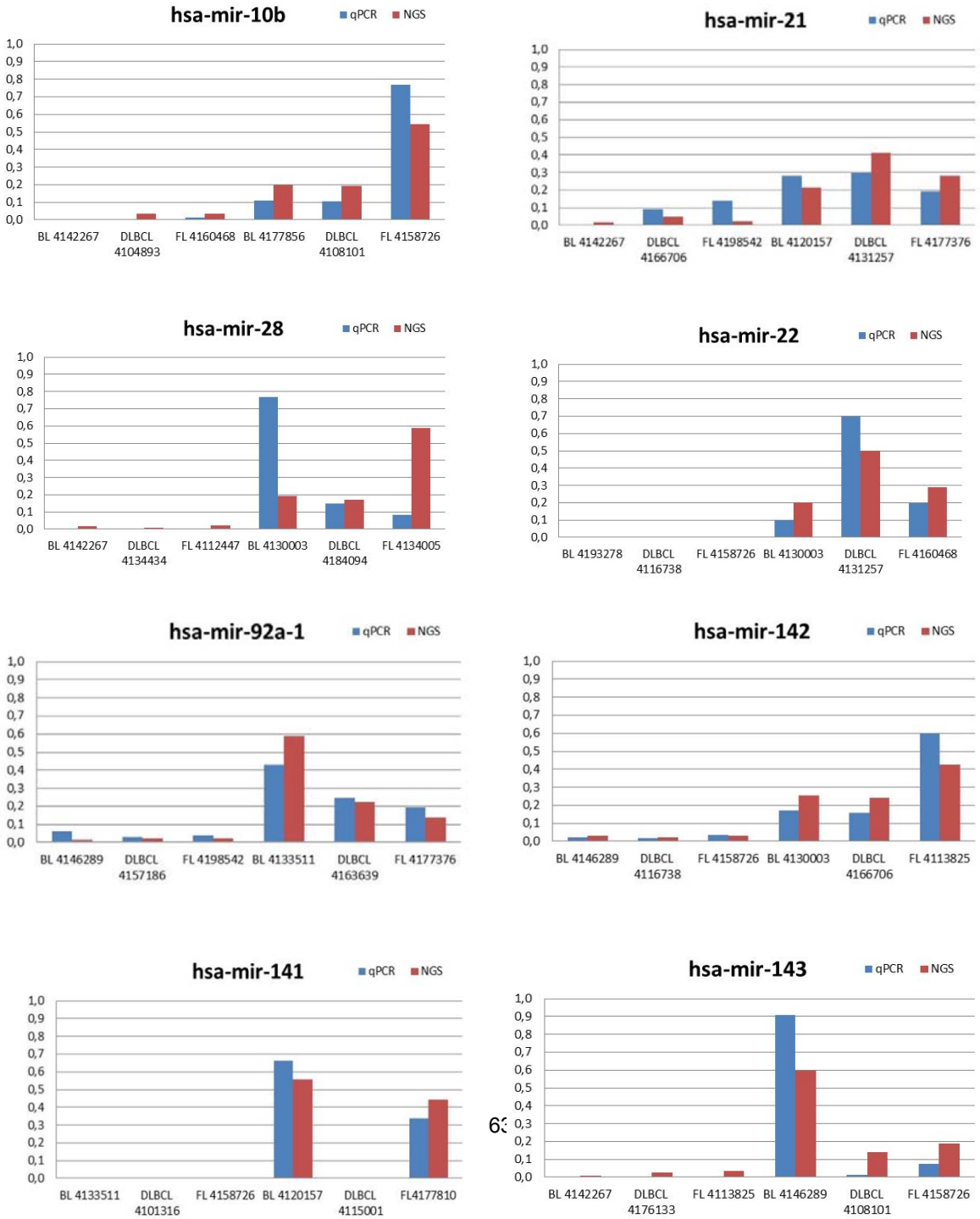
Table 4.4 Minimum expression of selected miRNA in the three subtypes of the GCB lymphoma.

miRNA	BL patient	BL expr.	DLBCL patient	DLBCL expr.	FL patient	FL expr.
miR-10a	BL 4142267	0.05	DLBCL 4104893	2.56	FL 4160468	2.58
miR-141	BL 4133511	0.00	DLBCL 4101316	0.00	FL 4158726	0.00
miR-142	BL 4146289	0.64	DLBCL 4116738	0.50	FL 4158726	0.70
miR-143	BL 4142267	0.84	DLBCL 4176133	3.11	FL 4113825	4.02
miR-146a	BL 4146289	0.02	DLBCL 4131257	0.51	FL 4134005	0.29
miR-150	BL 4194218	0.02	DLBCL 4134434	0.01	FL 4134005	0.29
miR-21	BL 4142267	0.60	DLBCL 4166706	2.28	FL 4198542	0.99
miR-22	BL 4193278	0.05	DLBCL 4116738	0.05	FL 4158726	0.04
miR-28	BL 4142267	0.05	DLBCL 4134434	0.04	FL 4112447	0.05
miR-92a-1	BL 4146289	0.09	DLBCL 4157186	0.05	FL 4198542	0.11

Table 4.5 Maximum expression of selected miRNA in three different subtypes of the GCB lymphoma samples.

miRNA	BL patient	BL expr.	DLBCL patient	DLBCL expr.	FL patient	FL expr.
miR-10a	BL 4177856	14.87	DLBCL 4108101	14.50	FL 4158726	41.35
miR-141	BL 4120157	0.56	DLBCL 4115001	7.46	FL4177810	0.45
miR-142	BL 4130003	5.70	DLBCL 4166706	5.46	FL 4113825	9.57
miR-143	BL 4146289	68.23	DLBCL 4108101	16.11	FL 4158726	21.68
miR-146a	BL 4120157	19.26	DLBCL 4134434	30.47	FL 4113825	14.65
miR-150	BL 4190495	0.25	DLBCL 4166706	1.20	FL 4159170	2.53
miR-21	BL 4120157	10.01	DLBCL 4131257	19.16	FL 4177376	13.11
miR-22	BL 4130003	2.63	DLBCL 4131257	6.59	FL 4160468	3.79
miR-28	BL 4130003	2.45	DLBCL 4184094	1.65	FL 4134005	1.03
miR-92a-1	BL 4133511	1.01	DLBCL 4163639	0.90	FL 4177376	3.05

We applied the TaqMan miRNA qRT-PCR assay (Applied Biosystems) utilizing 20ng of RNA and specific RT primers for individual miRNAs. To confirm consistent relative expression of each of these RNA species within individual samples, the Ct values for RNU24 (SNORD24 small nucleolar RNA, C/D box 24) were correlated across the ten GCB lymphoma samples. With the exception of *miR-28,* which showed a somewhat inconsistent expression pattern compared to NGS sequencing data, the other nine miRNAs (*miRs-10b, -141a, -142b, -143, -146a, -150, -21, -22b-5p and -92a*) fully confirmed the NGS sequencing data (Figure 4.6).



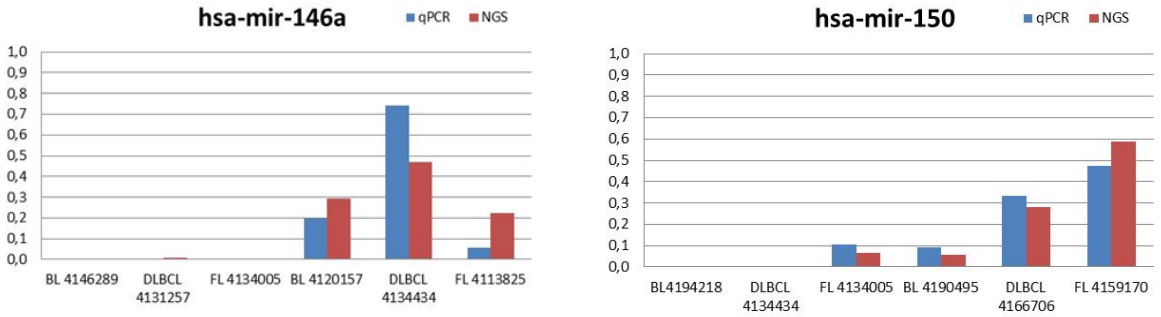


Figure 4.6 Validation of miRNA expression by quantitative Real-Time PCR**.** Represented are the comparisons between DLBCL, BL and FL. Blue bar: results of qRT-PCR, red bar: results obtained by NGS.

4.2 Novel miRNAs

The aim of the Düsseldorf workpackage within the ICGC MMML-Seq framework was to provide a comprehensive list of known small non-coding RNAs as well as novel miRNAs in GCB lymphomas. Therefore we tried to identify novel miRNA amongst the unclassified sequences in our libraries. We applied miRanalyzer for identification of novel miRNAs. Miranalyzer is a web server and stand-alone tool for the analysis of high throughput sequencing experiments (Hackenberg, Sturm et al. 2009). The fastq files were transformed into a read count file, which is filtered to keep only sequences from 17 to 26 bases. These reads were successively mapped to several databases in order to identify known miRNAs, discard messenger RNA contaminations and select sequences for the miRNA prediction step. Sequences that can be mapped to the reference genome, but cannot be assigned to any annotations (unclassified sequences), were used to predict novel miRNAs. 283 putative novel miRNAs were predicted by miRanalyzer (Supplement Table S. 3)

4.2.1 Experimental validation of novel miRNAs

To identify novel miRNAs in our libraries, the miRanalyzer output served as the starting point for the validation. Since novel miRNAs are often of low expression and generally have a sequence similar to known genomic repeats, we assumed that many of these predicted miRNAs were either randomly derived from fragments of various RNAs or regulatory small RNAs (Berezikov, Thuemmler et al. 2006). We addressed this issue by performing Northern blot assay to further confirm the presence of bona fide miRNAs. Northern blot as a classic low-throughput technique is the gold standard to validate the existence of computationally predicted novel miRNAs. Northern blot experiments were carried out according to a well-

established protocol (Landgraf, Rusu et al. 2007). Total RNA was derived from four lymphoma cell lines (NAMALWA (BL), Raji (BL), SU-DHL-4, (DLBCL) and SU-DHL-6 (FL)). The 20 most promising candidates were selected to be checked by Northern blot (Table 4.6)

Table 4.6 Top 20 novel predicted miRNAs in GCB lymphomas.

NB-miRNA	Genomic location	Sequence
NB-miR-1	chr1:43914204-43914292(+)	CCAGGGCAGAGGGCACAGGAATCT
NB-miR-2	chr20:43272717-43272813(+)	GCAGGGCCGTCTACCTCCAAGGC
NB-miR-3	chr3:156878062-156878153(+)	ACCTTGGGATCCTCAGAGGCTG
NB-miR-4	chr8:27290881-27290962(+)	TTGGCTGGTCTCTGCTCCGCAGG
NB-miR-5	chr10:50035510-50035603(+)	GCACACTGACACAGAGAGAGAGA
NB-miR-6	chr20:43272714-43272812(-)	CAGGGCCGTCTACCTCCAAGGC
NB-miR-7	chr10:75936514-75936605(+)	AAGCAGTTGCTGTGGTACCTGCT
NB-miR-8	chr5:72143946-72144041(+)	AGACGCTGGCGGCCGGGCTGCCA
NB-miR-9	chr1:43914204-43914292(+)	CAGGCCTGTGGCTCCTCCCTCA
NB-miR-10	chr1:32410164-32410309(+)	CCTCCTCAGGCACTAGAGCTTCCC
NB-miR-11	chr20:43272714-43272812(-)	CAAGGCAAACCTAGTCAGATTA
NB-miR-12	chr2:219206629-219206714(+)	GTCCCCTGCTCCCTTGTTCCCCAGTC
NB-miR-13	chr17:73008757-73008842(+)	CGCAAGACCTGAGCATGGCGGCCA
NB-miR-14	chr17:73008757-73008842(+)	CTGCGCTGGGGCCTGAGCCGAGC
NB-miR-15	chrX:118533303-118533392(+)	CCGGAGAGCCCCAGGCCCGAGCC
NB-miR-16	chr11:64863680-64863778(+)	GCCGGGCCTAGCCCGGGGTCTGGAC
NB-miR-17	chr15:45003719-45003814(+)	AGCATTCGGGCCGAGATGTCTCG
NB-miR-18	chrM:3363-3463(-)	GTAGGCCCCTACGGGCTACTACAACCCT
NB-miR-19	chrM:3363-3463(-)	CCAACGTTGTAGGCCCCTACGGGCTACT
NB-miR-20	chr12:52453530-52453613(+)	TCACTGCAGGGCCCTAGCAATA

The design of the probes was reverse complementary to the mature miRNA sequences (Table 4.7). Of the 20 candidates, four miRNAs (*NB-miR-1, NB-miR-5, NB-miR-19 and NB-miR-20*) showed mature processed fragments in Northern blot analysis.

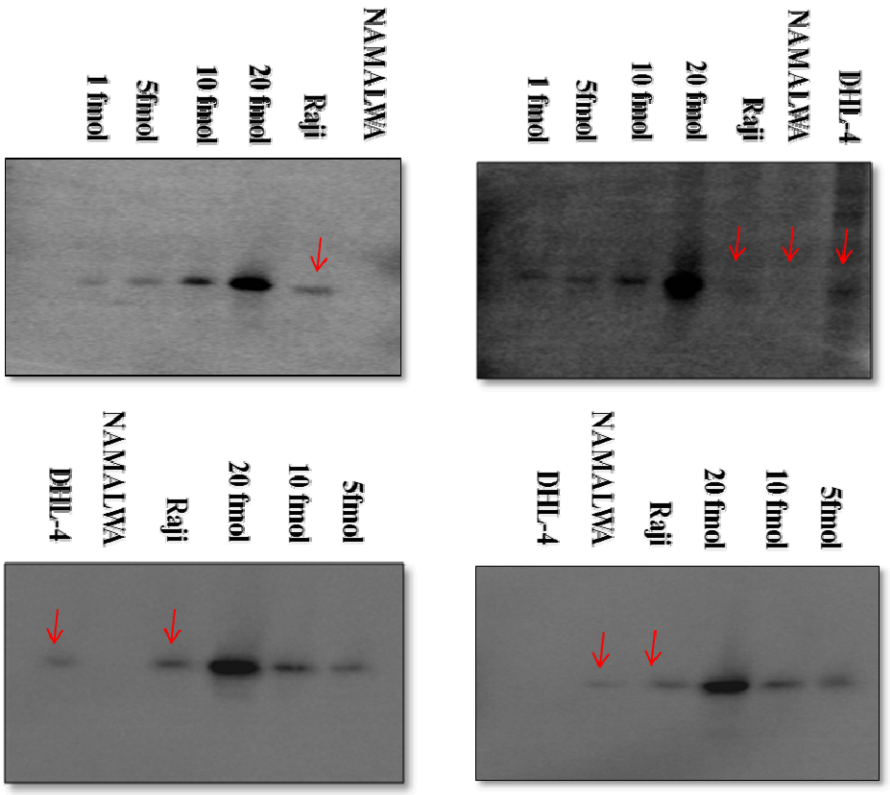
Table 4.7 Northern blot results of the top 20 novel predicted miRNAs in GCB lymphomas

NB-miRNA	Probe sequence	Result
NB-miR-1	**AGATTCCTGTGCCCTCTGCCCTGG**	**Positive**
NB-miR-2	GCCTTGGAGGTAGACGGCCCTGC	Negative
NB-miR-3	CAGCCTCTGAGGATCCCAAGGT	Negative
NB-miR-4	CCTGCGGAGCAGAGACCAGCCAA	Negative
NB-miR-5	**TCTCTCTCTCTGTGTCAGTGTGC**	**Positive**
NB-miR-6	GCCTTGGAGGTAGACGGCCCTG	Negative
NB-miR-7	AGCAGGTACCACAGCAACTGCTT	Negative
NB-miR-8	TGGCAGCCCGGCCGCCAGCGTCT	Negative
NB-miR-9	TGAGGGAGGAGCCACAGGCCTG	Negative
NB-miR-10	GGGAAGCTCTAGTGCCTGAGGAGG	Negative
NB-miR-11	TAATCTGACTAGGTTTGCCTTG	Negative
NB-miR-12	GACTGGGGAACAAGGGAGCAGGGGAC	Negative
NB-miR-13	TGGCCGCCATGCTCAGGTCTTGCG	Negative

NB-miR-14	GCTCGGCTCAGGCCCCAGCGCAG	Negative
NB-miR-15	GGCTCGGGCCTGGGGCTCTCCGG	Negative
NB-miR-16	GTCCAGACCCCGGGCTAGGCCCGGC	Negative
NB-miR-17	CGAGACATCTCGGCCCGAATGCT	Negative
NB-miR-18	AGGGTTGTAGTAGCCCGTAGGGGCCTAC	Negative
NB-miR-19	**AGTAGCCCGTAGGGGCCTACAACGTTGG**	**Positive**
NB-miR-20	**TATTGCTAGGGCCCTGCAGTGA**	**Positive**

In order to quantify the intensity of the bands obtained, we designed RNA standard (Eurofins) for validated novel miRNAs (*NB-miR-1, NB-miR-5, NB-miR-19 and NB-miR-20)*. The first novel miRNA (*NB-miR-1*) was detected only in the Raji cell line, the second novel miRNA (*NB-miR-5*) was moderately expressed in SU-DHL-4 cell line and weakly expressed in NAMALWA and Raji cell lines, the third and fourth novel miRNA (NB-miR-19/20) similarly observed in Raji, and also respectively expressed in SU-DHL-4 and NAMALWA (Figure 4.7.A). The secondary structures of the potential miRNAs were predicted by RNAfold (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi) (Figure 4.7.B).

A.



B.

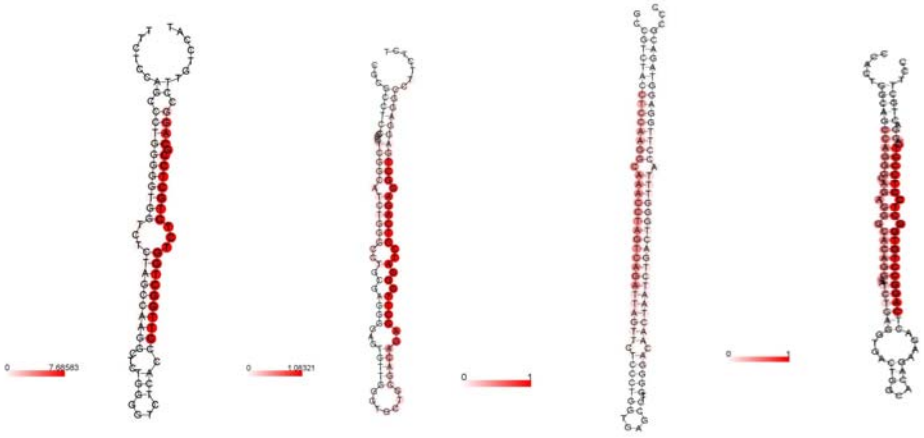


Figure 4.7 Validation of novel miRNA candidates by Northern Blot. (A) Northern Blot experiments with probe reverse complementary to candidate, 1/5/10/20: fmol size standard were used. Bands corresponding to mature miRNA are indicated by arrows. (B) Predicted folding structure (hairpin) of pre-miRNAs of candidates. Intensity of red shading correlates to the number of sequenced reads across a particular region across the pre-miRNA (the darker, the more reads detected)

4.2.2 Overexpression of novel miRNAs

The induction of miRNA expression in mammalian cell system is the initial step in different studies to identify miRNA function. Of the four novel miRNAs which we could successfully validate by Northern blot assay, we subjected two novel miRNAs (*NB-miR-5 and NB-miR-19*) to further functional analysis. Transient overexpression of novel miRNAs was achieved by transfection of double-stranded RNA molecules that mimic the Dicer cleavage product. To achieve that, *NB-miR-5* and *NB-miR-19* were cloned into pcDNA™6.2-GW/± EmGFP-miR (Invitrogen). pcDNA™6.2-GW/± EmGFP-miR vectors are designed to accept engineered pre-miRNA sequences. We set up the experiment with nucleofection using the Cell Line Nucleofector® Kit V (Lonza) and used 1 µg per 12-well plate for each cell line. Transfection efficiency was examined in two cell lines and found to be ~70%. In conclusion, we showed that synthetic pre-miRs can be delivered efficiently into Raji and SU-DHL-4 cell lines using nucleofection method. As negative control we employed empty destination vector.

4.2.3 Functionality tests of pre-miRs

To identify whether transfected pre-miRs can effectively increase intracellular levels of specific miRNAs, we conducted a quantification assay using stem-loop RT followed by

TaqMan PCR analysis. For these experiments, we selected Raji and SU-DHL-4 cell lines and transfected them with 1µg of destination vector. To optimize the time point after transfection that provides the maximum expression of NB-miRNAs, mature miRNAs were assayed by qRT-PCR 24h, 48h and 72h after transfection (Figure.3.10.). Relative miRNA expression changes were calculated by using a critical threshold (C (t)) method, also known as $2^{-\Delta\Delta CT}$ (Livak and Schmittgen 2001). Expression level of abundant small nuclear RNA (RNU24) was used as a reference for normalization throughout the experiments. The highest expression of novel miRNAs in both cell lines was achieved 24h after transfection. We detected significantly increased levels of both NB-miRNAs (around 11-fold in Raji cell line and around 20 -fold in SU-DHL-4 cell line) in the cells transfected with novel miRNAs compared to the mock transfected cells (Figure 4.8).

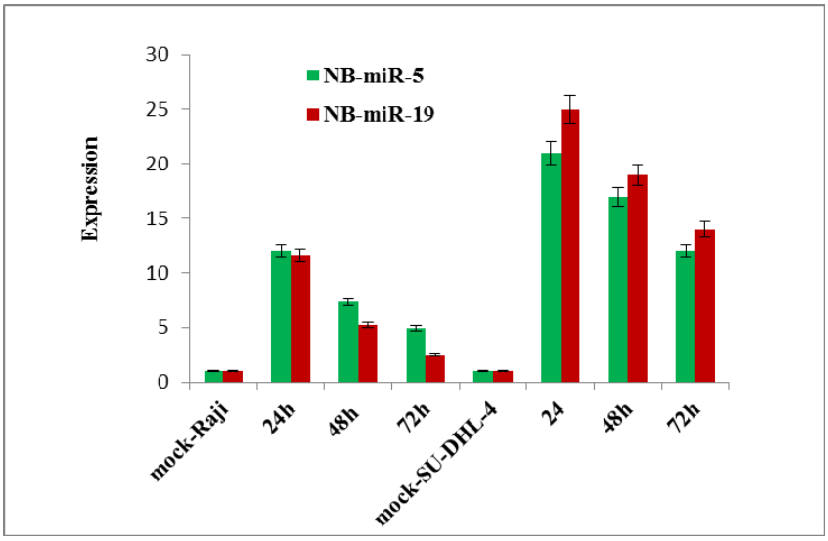


Figure 4.8 Expression levels of NB-miRNA-5 and NB-miRNA-19. Representative time point qRT-PCR experiment in the cells transfected with pre-NB-miRNAs. Raji and SU-DHL-4 were transfected with pre-NB-miRNAs and qRT-PCR was performed 24h, 48h, 72h after transfection. Expression levels of NB-miRNA-5 and NB-miRNA-19 in mock-transfected cells (only transfection reagent) were adjusted to 1 and used as a reference to normalize miRNA expression changes in the cells transfected. Results are shown as normalized expression mean values from three independent transfection experiments (n=3) ± SD.

4.2.4 Knockdown of novel miRNAs

Since miRNA overexpression experiments are prone to a certain degree of false positive results (Bracken, Gregory et al. 2008), we also applied miRNA silencing to the lymphoma cell lines. The small size of miRNAs makes their detection problematic; therefore miRNAs knockdown requires a highly specific and potent tool (Krutzfeldt, Poy et al. 2006). We employed miRNA knockdown application, based on using Locked Nucleic Acid (LNA™)-modified probes. Since LNA has high affinity for short RNA targets and a high mismatch

discrimination ability, it is considered an effective tool for functional analysis of miRNAs (You, Moreira et al. 2006). LNA probes for both novel miRNAs were designed by Exiqon. Similar to the overexpression experiment, Raji and SU-DHL-4 were selected for anti-miR transfection using the Cell Line Nucleofector® Kit V (Lonza). Usually, miRCURY LNA™ micRNA inhibitors display potent activity at final concentrations of 1-25 nM, but an extended range of 1-100 nM may be appropriate for optimization purposes. We transfeced the cells with different concentration (30 nm, 50 nm and 100 nm) and harvested the cells 24h after transfection. Quantification analysis revealed that anti-novel miRNAs transfection significantly reduced both NB-miRNAs levels compared to mock transfection (Figure 4.9) confirming that 100nm of anti-novel miRNAs had best overall knockout efficiency. Different concentrations of LNA in SU-DHL-4 cell line had little or no effect on *NB-miR-19* expression; and we could not obtain knockdown of *NB-miR-19* in SU-DHL-4. Even after increasing incubation time to 48h and 72h we obtained similar result (results not shown).

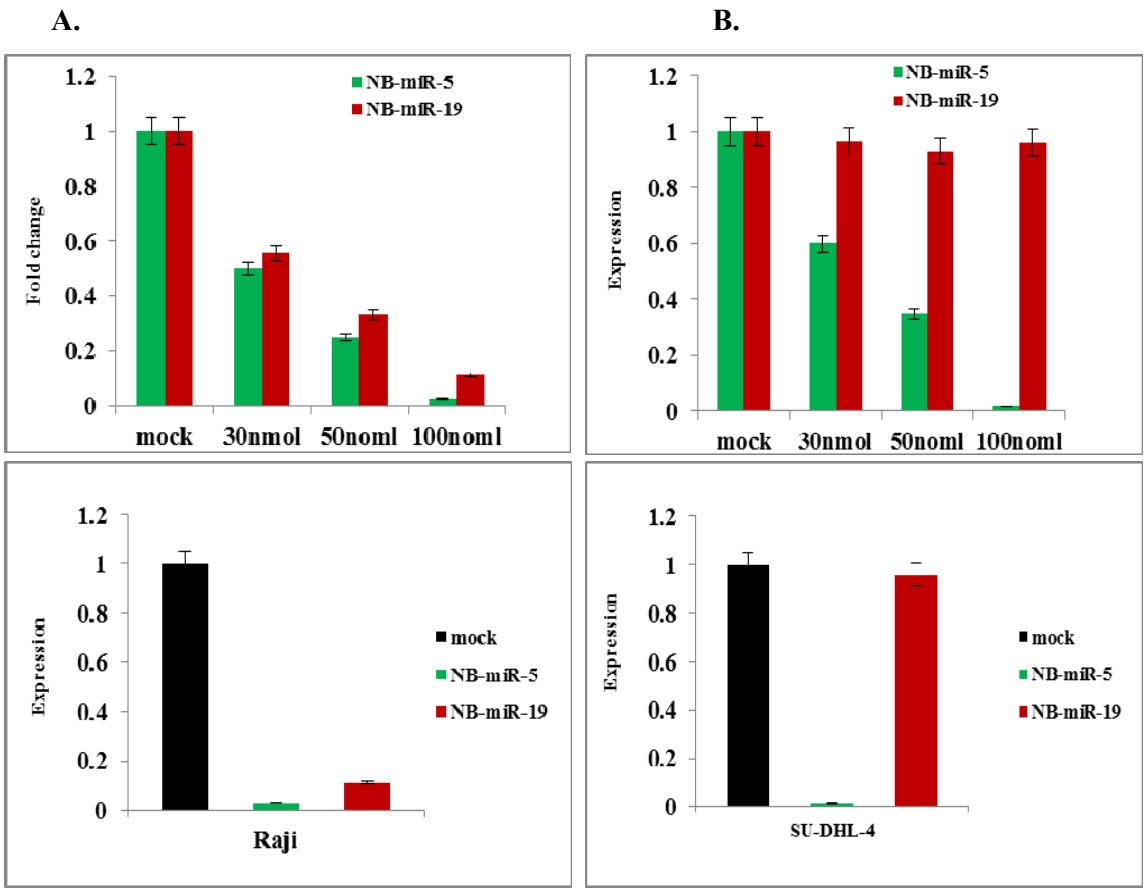


Figure 4.9 Different concentrations of LNA knockdown probes targeting NB-miRNAs**.** QRT-PCR was performed to quantify changes in the miRNA expression levels after transfection with a different concentration of LNA probes. Raji (A) and SU-DHL-4(B) were transfected with the indicated concentration of LNA probes and qRT-PCR was performed 24h after transfection. Expression levels of NB-miRNA-5 and NB-miRNA-19 in

mock-transfected cells (only transfection reagent) were adjusted to 1 and used as a reference to normalize miRNA expression changes in the transfected cells. Results are shown as normalized expression mean values from three independent transfection experiments (n=3) ± SD

4.2.5 RNA Seq-data of novel miRNAs

Since miRNAs regulate gene expression by translational repression and mRNA cleavage (Du and Zamore 2007), the effect of an miRNA modulation should also be assayed at the mRNA level. Therefore, after overexpression and knockdown of NB-miRNAs, we set out to perform RNAs-Seq experiments to identify their regulated mRNAs. We performed overexpression and knockdown experiment in Raji and SU-DHL-4 cell lines in triplicate. 24h after transfection cells were collected and total RNA was extracted by using Trizol. Prior to preparation of libraries, we verified increased and decreased of novel miRNAs expression by qRT-PCR. RNAseq libraries were generated using the TruSeq RNA sample prep kit (Illumina, San Diego, CA, USA), which generates libraries from 200 ng of total RNA. Each library was barcoded, pooled and sequenced using Illumina HiSeq 2500. We chose single-end 100-bp sequencing to maximize read depth and detection of small fold changes. Following mapping of the sequencing reads to the reference genome with TopHat, a total of 241 dysregulated genes were detected after knockdown and overexpression of *NB-miR-5* in SU-DHL-4, as result showed 208 and 34 genes were identified from overexpression and knockdown samples respectively, which shared 34 genes in common (Table 4.8). We obtained fewer genes after overexpression and knockdown of *NB-miR-5* in Raji compared to SU-DHL-4.

Since we could not get knockdown of *NB-miR-9* in SU-DHL-4, we obtained 2,188 different expressed genes after overexpression of this miRNA in SU-DHL-4. Our data indicated overexpression and knockdown of *NB-miR-19* in Raji changed mRNA levels of 174 genes, compared to mock (Table 4.9, Supplement Table S. 5).

Table 4.8 List of shared target genes after overexpression and knockdown of *NB-miR-5* in SU-DHL-4

list of overlapping target gene		
ACAP2	GNA13	PSMD8
AKAP11	GRK6	RASSF2
ATM	HELZ	RFC3
BCL2L11	LPIN1	RPN1
BRWD1	MCM2	RUVBL1
CDCA7	MCM5	FANCI
CLCN3	MCM7	FNBP4
CMTM6	MEF2A	NME1
CYBA	MGEA5	PIK3AP1
EFHD2	MRPL12	NIPBL
EP300	MYBL2	FAM214A

Table 4.9 Gene expression summary of both novel miRNAs overexpression and knockdown in Raji and SU-DHL-4 cell lines. Knockdown of *NB-miR-9* was not obtained.

Samples	NB-miR-5-KD	NB-miR-5-OE	NB-miR-19-KD	NB-miR-19-OE
Total No. of genes in SU-DHL-4	34	208	-	2,188
Upregulated in SU-DHL4	18	95	-	1,335
Downregulated in SU-DHL-4	16	113	-	853
Total No. of genes in Raji	12	28	75	99
Upregulated in Raji	10	8	51	21
Downregulated in Raji	2	20	24	78

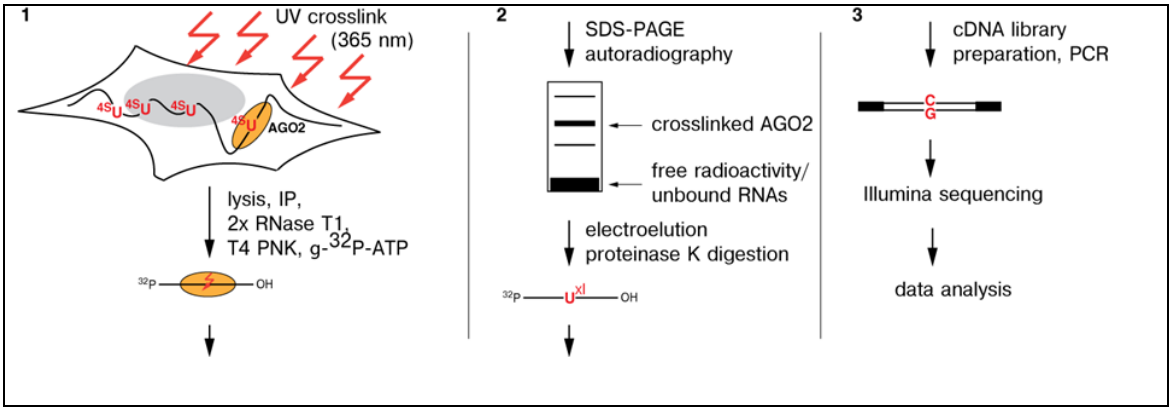
4.3 AGO2-PAR-CLIP

4.3.1 Photoactivatable nucleoside facilitates RNA-RBP crosslinking in cultured cells

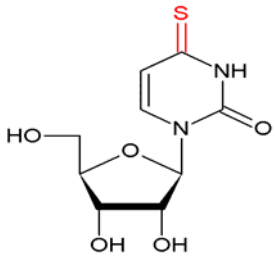
To uncover those mRNAs which are targeted by miRNAs in Argonaute-miRNA-mRNA interactions, we performed pull-down experiments of AGO2 complexes after photoactivated ribonucleoside-enhanced cross-linking followed by NGS sequencing of associated mRNAs (PAR-CLIP) (Hafner, Landthaler et al. 2010) (Figure 4.10.A). For immunoprecipitations (IPs) of endogenous AGO2 proteins, we utilized monoclonal anti-AGO2 antibody (C1.9E8.2, Millipore) to isolate AGO2-associated RNAs. PAR-CLIP experiments were performed in all four lymphoma cell lines (NAMALWA, Raji, SU-DHL-4 and SU-DHL-6) and AGO2-overexpressing HeLa cell line as positive control. Random or site-specific incorporation of photoactivatable nucleoside analogs into RNA in vitro has been used to probe RBP (RNA-binding protein) and RNP (ribonucleoprotein) -RNA interactions (Meisenheimer and Koch 1997, Kirino and Mourelatos 2008). Several of these photoactivatable nucleosides are readily taken up by cells without apparent toxicity and they have been used for in vitro crosslinking (Favre, Moreno et al. 1986). Here, we applied 4SU to cultured cells expressing the AGO2-

protein, followed by UV 365 nm irradiation (Hafner, Landthaler et al. 2010). 4SU (Figure 4.10.B) is incorporated into nascent transcript subsequently resulting in thymidine to cytidine (T-to-C) conversions at the cross-linked sites. The T-to-C conversions are used as markers to identify RNAs truly associated with AGO2 from false positive RNAs (Hafner, Landthaler et al. 2010, Spitzer, Hafner et al. 2014). The crosslinked RNA-protein complexes were isolated by IP, and the covalently bound RNAs were partially digested with RNase T1 and radiolabeled. After performing IPs, the bands that corresponded to the expected size of AGO2 were cut out (Figure 4.10.C).

A.

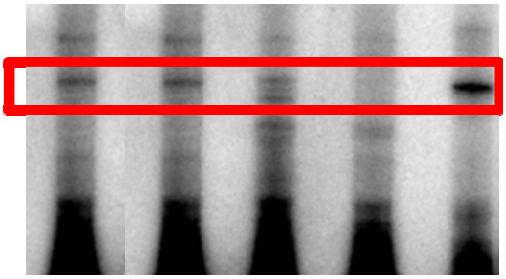


B.



4-thiouridine (4SU)

C.



SU-DHL-4 SU-DHL-6 Namalwa Raji Hela

Figure 4.10 PAR-CLIP methodology. (A) Illustration of PAR-CLIP, 4SU-labeled transcripts were crosslinked to AGO2 and partially RNase-digested RNA-protein complexes were immunopurified and size-fractionated. RNA molecules were recovered and converted into a cDNA library and deep sequenced. (B) Structure of 4-thiouridine (4SU). (C) Phosphor images of SDS-gels that resolved 5'-32P-labeled RNA-AGO2 immunoprecipitates (IPs) prepared from lysates from cells that were cultured in media in the presence of 4SU and crosslinked with UV 365 nm.

Barcoded libraries from four cell lines were mixed and pooled in equimolar (10 nmol/L) amounts and diluted to 7pmol/L for cluster formation on a single-flow cell lane. After filtering and trimming of the reads for low quality and adaptors from the four sequencing libraries, we obtained a total of 148,040,337 single-end reads (Table 4.10). Next, all remaining reads were aligned against the human reference genome sequence hg 19 using BWA allowing for two mismatches between a read sequence and the reference sequence. The identification of target regions that are bound by the protein of interest is performed by a hierarchical clustering algorithm implemented in our in-house pipeline. Each of the resulting 'clusters' of reads represents a target region of the protein. In this study, we got 21.397 clusters formed by four PAR-CLIP libraries (Supplement Table S. 4). To discard false-positives, we applied filters to identify the most confident clusters, e.g. filtering for a T-to-C conversion rate above 25% for each cluster and for at least 10 reads per cluster. Among four libraries, the majorities of reads were achieved from non-BL cell lines (SU-DHL-4 and SU-DHL-6) (Figure 4.11.A). Most of the reads were located in the CDS (coding DNA sequence) and 3'UTR, only few reads were located in 5'UTR (Figure 4.11.B). In the Raji cell line, we achieved reads neither in the 3' UTR nor in the 5' UTR region, but only in the CDS region.

Table 4.10 Summary of read alignment statistics for four cell lines (number of reads is shown in millions of reads)

Cell line	Number of reads after adaptor trimming	Number of aligned reads
Raji	10.37 M (7%)	1.31M (12.6%)
NAMALWA	16.25 M (10%)	1.32M (8.19%)
SU-DHL-4	17.01 M (12.6%)	1.14M (6.12%)
SU-DHL-6	18.66 M (11.6%)	1.38M (8.15%)

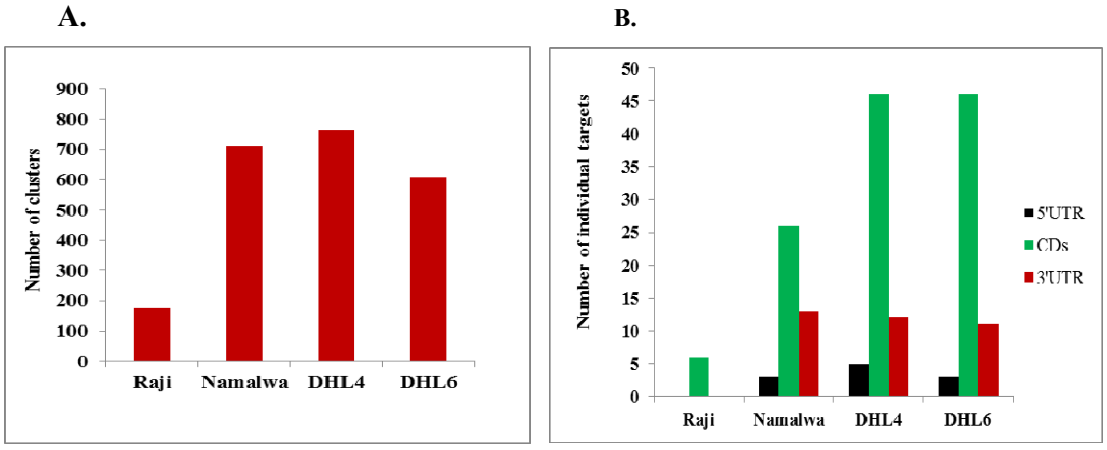


Figure 4.11 AGO2-PAR-CLIP summary of four cell lines. (A) Number of clusters achieved after read alignment and pile-up from each cell line. (B) Genomic location of PAR-CLIP isolated mRNAs and distribution of T-to-C sites in transcript regions.

After read alignment and pile-up we achieved 1962, 5937, 6302 and 7196 clusters from Raji, NAMALWA, SU-DHL-4 and SU-DHL-6, respectively (Supplement Table S.5). Non-specific bound clusters were excluded with following criteria: more than 10 reads per cluster and more than 25% T-C mutation. After filtering non-specific clusters, 176, 713, 763 and 609 mRNA clusters were detected in AGO2-PAR-CLIPs from Raji, NAMALWA, SU-DHL-4 and SU-DHL-6, respectively (Figure 4.11.A). The AGO2-PAR-CLIP mRNAs from the two non-BL cell lines exhibited six shared mRNAs (SH2B2, HERC2, SVEP1, SIPA1L3, CANT1, and DHX33). Our data did not show any common target in the two BL cell lines. We also captured nine targets (GSG2, SETD2, PHKB, MYBL1, APC2, DHX30, GLB1, CDH23, KIAA1274 and B2M) that were common in three out of the four cell lines. Only one target (GSG2) was found in all four cell lines (Table 4.11).

Table 4.11 Top six expressed miRNA targets in non-BL cell lines. MiRNA targets that were found in three out of four cell lines. Only GSG2 was found in all four libraries.

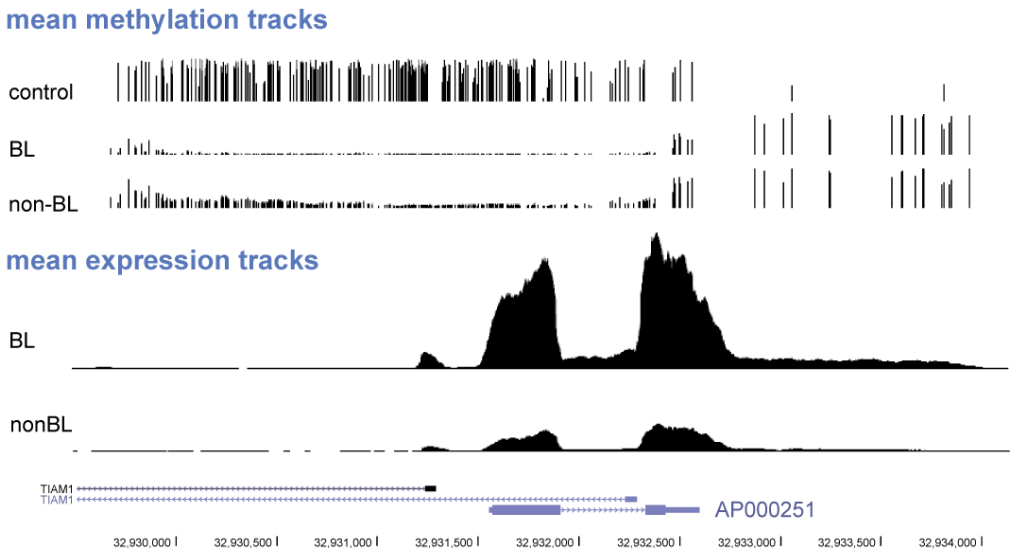
Common mRNA targets in non-BL cell lines	Common mRNA target in three cell lines
SH2B2	**GSG2**
HERC2	SETD2
SVEP1	PHKB
SIPA1L3	MYBL1
CANT1	APC2
DHX33	DHX30
	GLB1
	CDH23
	KIAA1274
	B2M

4.4 LncRNAs

4.4.1 Differentially expressed and methylated lncRNA AP000251

High-throughput RNA sequencing has promoted the study of lncRNAs. Recent transcriptomic studies in mammals have revealed an abundance of long non-coding RNAs (lncRNAs) that are interspersed with coding genes (Engstrom, Suzuki et al. 2006). We mapped all of the sequenced reads derived from GCB lymphoma and from healthy controls against the largest currently available lncRNA database. Searching for the lncRNAs showing the highest differential methylation and expression between BL and non-BL patient samples, we identified lncRNA AP000251 (Figure 4.12.A).

A.



B.

CATEGORY	GENE SYMBOL	ID	UCSC BROWSER	STRAND	EXONS NUMBER
lincRNA	AP000251.2	ENST00000433071.1	chr21: 32931536 - 32932800	+	2

Figure 4.12 Differentially expressed and methylated lncRNA AP000251. (A) Expression tracks are scaled to RPM (read per million) 0.7. (B) Number of exons; LncRNA AP000251 spans 1.263 bp in the genome (32,931,537 - 32,932,800) and consists of two exons (371 bp, 470 bp).

The lncRNA AP000251 lies downstream of TIAM1 (T-cell lymphoma invasion and metastasis 1) on chromosome 21. It spans 1.263 bp in the genome (32,931,537 - 32,932,800) and consists of two exons (371 bp, 470 bp) (Figure 4.12.B). The function of lncRNA AP000251 is currently unknown, but dysregulation in our tumor samples suggests that it has a critical biological function. To assess its function, we performed overexpression as well as knockdown experiments followed by qRT-PCR analysis to check for altered mRNA levels following AP000251 RNA level changes. Next, we performed RNAseq experiments on overexpressed and knockdown lncRNA AP000251 samples to identify those transcripts which are regulated by this lncRNAs.

4.4.2 LncRNA AP000251 validation

To assess the accuracy of next-generation sequencing data, we measured AP000251 expression patterns in lymphoma cell lines by qRT-PCR analysis. For this purpose, lncRNA AP000251 expression was confirmed using qRT-PCR in the set consisting of BL, non-BL and normal B cells. Two cell lines of each BL and non-BL subtypes were selected to compare expression levels of lncRNA AP000251. To prevent false amplification of contaminated genomic DNA, we designed primers spanning the exon-exon junction (Eurofins). GAPDH (Glyceraldehyde-3-phosphate dehydrogenase) was used as internal control. QRT-PCR analysis showed an expression pattern of lncRNA AP000251 in the cell lines similar to that observed in NGS sequencing data (Figure 4.13).

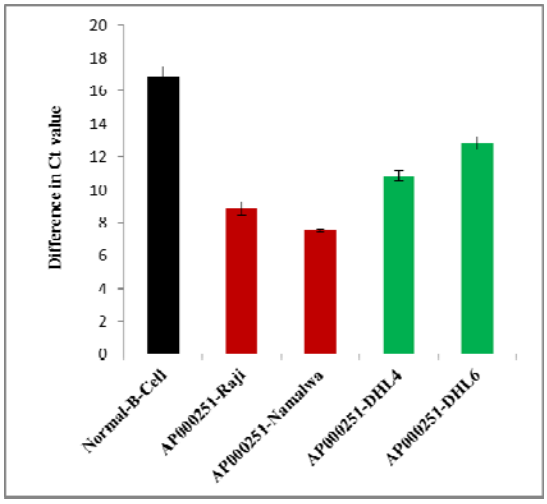


Figure 4.13 Expression levels of LncRNA AP000251 in lymphoma cell lines. LncRNA AP000251 RNA levels were measured by qRT-PCR in four different lymphoma cell lines, higher delta Ct value shows lower expression, GAPDH mRNA levels were used as an internal normalization control. Values represent the average of three technical replicates.

4.4.3 Overexpression of lncRNA Ap000251

Based on the different expression of lncRNA AP000251 in BL and non-BL, we selected Raji and SU-DHL-4 cell lines for the overexpression study. To make overexpression (OE) constructs, PCR products were digested with SalI and NotI, purified, and subcloned into the Gateway vector, pcDNA™6.2/EmGFP-Bsd/V5-DEST (Invitrogen). Because expression of the lncRNA is not detectable by Western blot experiment, this vector is suitable as it contains a separate transcriptional unit which allows high-level constitutive expression of Emerald Green Fluorescent Protein (EmGFP) fused to the blasticidin resistance gene. This way, judging transfection efficiency by fluorescence microscopy is easily achieved.

Overexpression constructs were sequenced prior to transfection. As we already set up electroporation-based transfection techniques for Raji and SU-DHL-4 cell line, destination vector was transfected into Raji and SU-DHL-4 cells by using nucleofection kit. Empty vector was used as negative control expression vector. According to the manufacturer's protocol, 2 x 10^6 cells of each cell line were transfected with 2 µg plasmid DNA. We examined three different time points post transfection to achieve a high level of efficiency. After 24, 48 and 72h, cells were analyzed for GFP expression by fluorescence microscopy. Cells were harvested 24-72h post-transfection and cellular RNA was extracted and reverse transcribed to cDNA. Next we applied qRT-PCR to measure expression of AP000251 at different time points. We found the best time point with maximum expression of lncRNA AP000251in both SU-DHL-4 and Raji cell lines was 24h after transfection (Figure 4.14).

A. **B.**

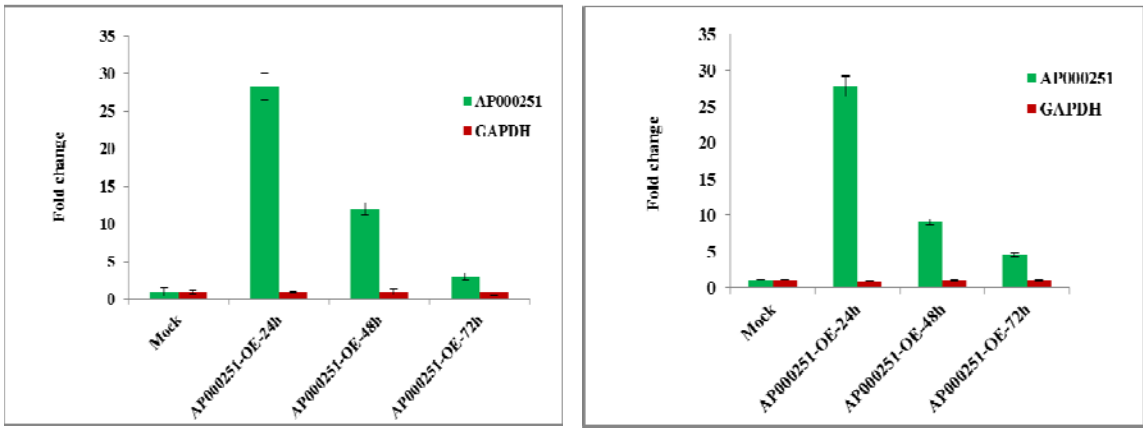


Figure 4.14 APl000251 fold change 24h after AP000251 transfection in (A) Raji and (B) SU-DHL4 cell lines. Values represent average and bars represent standard deviations of three independent experiments. Fold changes of AP000251 levels were calculated using $2^{\Delta\Delta Ct}$ method. (OE; Overexpression). Data for the 48 hour time point is not shown.

4.4.4 Knockdown of lncRNA AP000251

Short-hairpin RNA interference (shRNA)-encoding constructs directed against AP000251 were cloned into the pcDNA™6.2-GW/EmGFP-miR vector as previously described. Similar to overexpression, Raji and SU-DHL-4 cell line were transfected with the constructs using nucleofection. According to the GFP expression, we found the best time point for harvesting the cells was 24h after transfection. In Raji cell line, qRT-PCR analysis revealed transfection of three shRNAs significantly reduces endogenous levels of AP000251 relative to nonsilencing control, suggesting that this lncRNA is susceptible to Dicer-mediated

suppression (Figure 4.15.A). The expression of AP000251 in SU-DHL-4 cell line after transfection of three shRNAs remained unchanged (Figure 4.15.B). We repeated this experiment with a different concentration of shRNAs but were not able to achieve a significant reduction of AP000251 in transfected cells compared to the control (Figure 4.15.C).

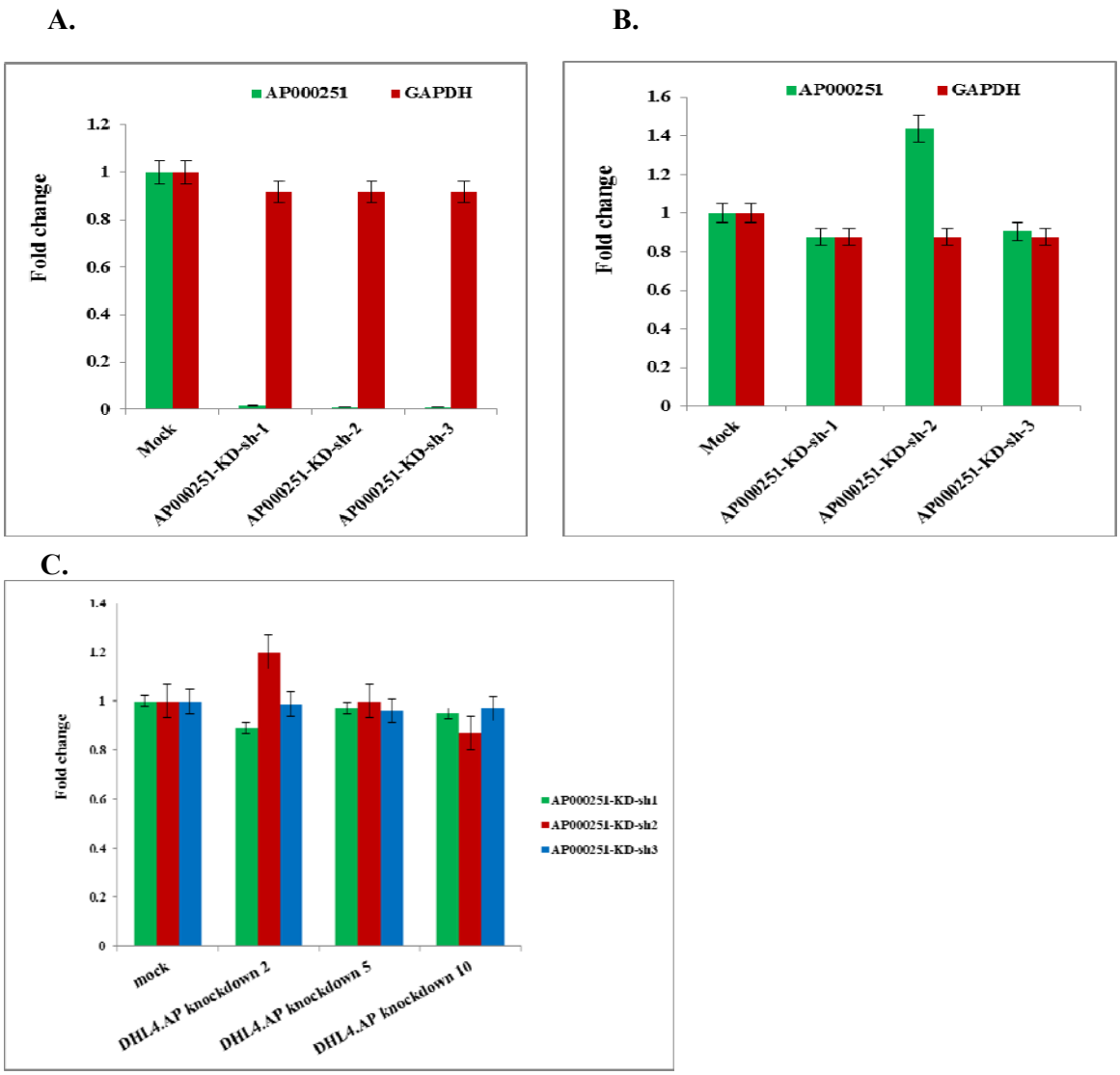


Figure 4.15 APl000251 fold changes 24h after transfection of three shRNAs in (A) Raji cell line, and (B) SU-DHL-4 cell line. (C) SU-DHL-4 cell line was also transfected with different concentrations (2, 5 and 10 µg) of three shRNAs against lncRNA AP000251.

4.4.5 AP000251 does not regulate TIAM gene expression

The lncRNA AP000251 lies downstream of TIAM1 (T-cell lymphoma invasion and metastasis 1) on chromosome 21. TIAM1 is a metastasis-related gene of T-cell lymphoma

that is also involved in the metastasis of a variety of other cancers and TIAM1/Rac signaling affects cell migration, invasion and tumor metastasis (Ellenbroek, Iden et al. 2012). To assess the function of lncRNA AP000251, we reasoned that AP000251 might act to regulate the gene expression of its neighboring gene. To assess this hypothesis, we again used three shRNAs directed against our lncRNA in order to identify the functional role following its depletion. We selected Raji cells for this experiment because of the ease of functional AP000251 knockdown and the detectable amounts of AP000251 and TIAM-1 in this cell line (Figure 4.16.A). We measured the expression of AP000251 and TIAM1 using qRT-PCR. After reduction of AP000251 in Raji, qRT-PCR analysis exhibits no significant expression change of TIAM1 (Figure 4.16.B). Next we checked the expression of TIAM1 after overexpression of AP000251 in Raji, but similar to the knockdown experiment, could not detect altered levels of TIAM1 (data not shown).

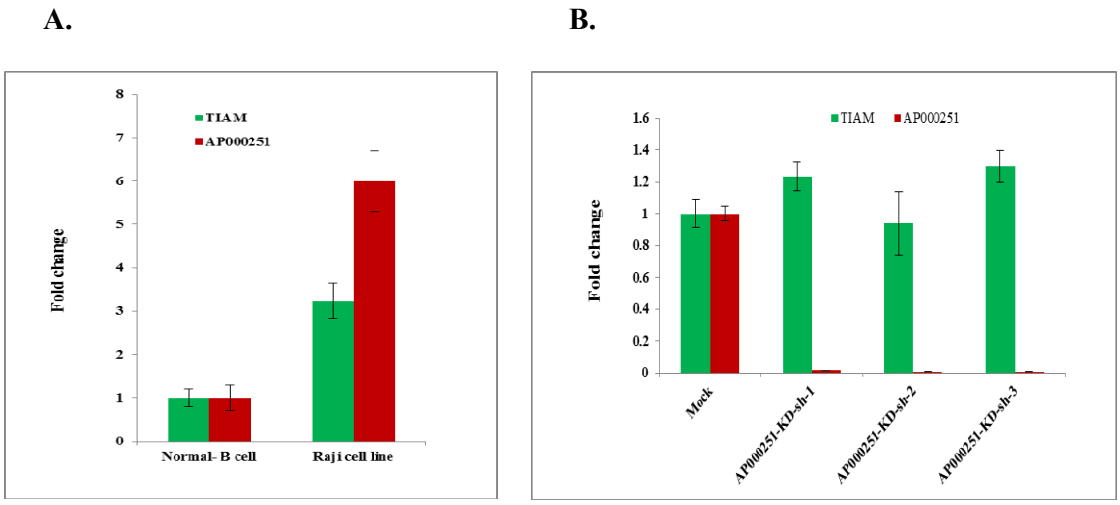


Figure 4.16 AP000251 does not regulate TIAM gene expression. (A) RNA levels of AP000251 and TIAM in normal B-cells and Raji cell line, were determined by Real-time RT-PCR and normalized to GAPDH levels.(B) Expression of TIAM in Raji cell line did not show any significant change in transfected cells with shRNAs against AP000251. Fold changes of AP000251 and TIAM levels were calculated using $2^{\Delta\Delta Ct}$ method.

To directly determine whether TIAM1 plays a role in expression and function of AP000251, we reduced endogenous TIAM1 RNA levels in Raji cells using two different siRNAs. 24h after transfection of siRNAs (50pmol), we harvested and examined AP000251 expression by qRT-PCR. However, reduced TIAM1 levels did not affect AP000251 expression (Figure 4.17).

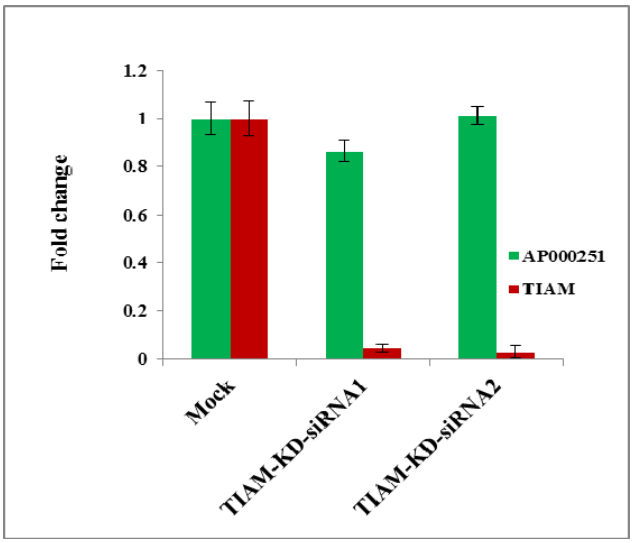


Figure 4.17 Knockdown of TIAM gene. (A) RNA levels of AP000251 and TIAM in Raji cell line were determined by qRT-PCR and normalized to GAPDH levels. Expression of AP000251 in Raji cell line did not show any significant change in transfected cells with siRNAs against TIAM. Fold changes of AP000251 and TIAM levels have been calculated using $2^{\Delta\Delta Ct}$ method.

4.4.6 RNAseq analysis of following level alterations of AP000251

Although lncRNAs constitute a large portion of the mammalian transcriptome, most of their biological functions have remained elusive. In order to determine lncRNA AP000251 function, we decided to capture target genes correlated with altered AP000251 expression in lymphoma cell lines. We assumed these target genes might offer clues to AP000251 function. Therefore, we first performed knockdown and overexpression of AP000251 in lymphoma cell lines, next sequenced replicate samples using the Illumina HiSeq system. Since we previously failed to knockdown AP000251 in SU-DHL-4, we performed the knockdown experiment in Raji and overexpression was performed in SU-DHL-4. Triplicate independent experiments were set up for each cell line, 24h after transfection cells were collected and total RNA was extracted using Trizol. We generated libraries again using the TruSeq RNA sample prep kit (Illumina, San Diego, CA, USA). 200 ng of total RNA was used for library preparation. Libraries were barcoded, pooled and sequenced using Illumina HiSeq 2500. We chose single-end 100-bp sequencing to maximize read depth and detection of small fold changes. Using a combination of the two programs HTSeq and edgeR, expression values were counted, normalized according to the library size and tested for differential expression compared to the reference mock samples (transfected with empty vector) of the same cell line. After overexpression of AP000251 in SU-DHL-4, we observed 120 dysregulated genes compared to mock samples, including 70 downregulated and 50 upregulated genes. In addition, the

knockdown of our lncRNA in Raji resulted in 668 dysregulated genes, of which 230 were downregulated and 438 upregulated (Supplement Table S. 6).

Table 4.12 List of top dysregulated genes after knockdown (Raji) and overexpression (SU-DHL-4) of lncRNA AP00025.

SU-DHL-4 cell line			**Raji cell line**		
Gene	Log2FC	p-value	Gene	Log2FC	p-value
ESPL1	1.346064	2.28E-06	**HSPE1**	1.663837	9.8E-35
CENPE	1.286474	2.51E-05	**ATP5J2**	1.425549	7.02E-27
HMMR	1.282953	3.22E-08	**TYMS**	1.270261	3.22E-21
CENPF	1.215119	0.000117	**TOP2A**	1.18483	7.98E-15
ASPM	1.174282	0.000101	**PPIB**	1.042085	1.28E-15
NCAPD2	1.155437	1.83E-05	**NME1**	1.001574	2.72E-11
MKI67	1.13374	0.000727	**UQCRQ**	-1.0148	6.55E-08
E2F2	1.107877	1.35E-06	**NCR3LG1**	-1.03295	3.96E-14
BRD4	1.084665	4.38E-05	**NCOR2**	-1.06128	4.9E-12
FOXM1	1.059382	2.56E-06	**SYNE2**	-1.10225	1.01E-08
POLQ	1.048703	0.000116	**SREBF1**	-1.11664	1.13E-15
RBM3	-1.00491	7.24E-05	**DOT1L**	-1.11795	3.49E-08
RPL37	-1.00972	0.003245	**TXNIP**	-1.14556	4.48E-14
RPL27A	-1.06396	0.0063	**SRCAP**	-1.15507	0.001545
CD53	-1.08436	1.92E-06	**FCRL3**	-1.24273	7.56E-15
EVL	-1.09967	3.47E-06	**RNF213**	-1.24946	2.18E-07
LIMD2	-1.11688	5.74E-05	**MOB3A**	-1.25244	4.85E-15
NCF1	-1.12396	6.54E-06	**KMT2D**	-1.25323	5.49E-19
CD52	-1.17388	0.000413	**MYLIP**	-1.62058	1.11E-05
TMSB4X	-1.24453	0.000447	**DDIT4**	-2.72851	6.41E-06

5 Discussion

5.1 MiRNA expression in GCB lymphoma

5.1.1 MiRNA expression differences between BL and non-BL

B-cell-NHLs are a group of lymphoproliferative malignancies that are characterized by distinctive chromosomal abnormalities and aberrant gene expression patterns, which have been identified to be relevant to their pathogenesis. Although clear distinctions on histologic and genetic grounds exist, there are also a large number of cases with intermediate features, not unequivocally attributable to one of these entities (Campo, Swerdlow et al. 2011). Some studies have suggested that treatment for each NHL subtype should be different based on its genetic and molecular characteristics (Macintyre, Willerford et al. 2000). Despite the complexities of miRNAs function in tumorigensis, dysexpressed miRNAs have shown significant potential as biomarkers for detection, diagnosis, classification and treatment of cancer. Some studies have confirmed the diagnostic impact of miRNAs in human pathology (Baffa, Fassan et al. 2009, Fassan, Croce et al. 2011). Their competency as biomarkers arises from their widespread dysregulation and distinctive expression profiles. Furthermore, miRNAs are more stable than mRNAs and permit analysis of samples with different sources such as urine, blood, FFPE tissues (Fassan, Croce et al. 2011). Identification of miRNAs differentially expressed among lymphoma subtypes might further advance our understanding of lymphoma pathogenesis. MiRNA profiles of NHL subtypes have been determined in several efforts either by microarray analysis (Lenze, Leoncini et al. 2011) or qRT-PCR (Wang, Corrigan-Cummins et al. 2012). Although the data obtained with these methods are reliable, the expression of a given miRNA cannot be quantified exactly in relation to the total amount of miRNAs present in a sample and in addition, unknown miRNAs will not be detected on either microarray or qRT-PCR analysis (Szczyrba, Loprich et al. 2010).

MiRNA expression of many cancer types is now well established (Croce 2008). However the role of miRNAs in lymphoma lags behind (Lawrie 2007). Therefore we set out to perform a comprehensive inquiry of miRNA expression in a unique sample collection of germinal center

(GC) derived B-cell lymphomas including FL, DLBCL and BL. Total RNA samples for cDNA library preparation were retrieved from ICGC-MMML-Seq. The ICGC-MMML-Seq Consortium has aimed at fully characterizing a total of 250 GC-derived B-cell lymphoma samples over a period of five years. Since this sequencing effort is ongoing, this thesis is based on the analysis of the first 56 samples, for which full datasets (genome, transcriptome, small RNAome, methylome) exist.

Unsupervised analysis by clustering miRNA profiles of all cases revealed no significant distinction of DLBCL, BL and FL from each other and many miRNAs showed similar expression levels across subtypes of GCB lymphoma. As unsupervised approaches are unable to identify more subtle molecular differences, we applied statistical analyses to identify differentially expressed miRNAs in GCB-lymphoma subtypes. In the first step we categorized all samples into two main groups including BL and non-BLs.

In order to show that miRNAs are responsible for the difference between BL and non-BL, the miRNA profiles of 17 BL were compared to 39 of non-BL (FL and DLBCL). Eight miRNAs (*miR-150, miR-211, miR-548ac, miR-1275, miR-1258 and miR-1244-1/2/3)* were differentially expressed between BL and other NHL samples. Of the miRNAs downregulated in the BL cases, *miR-150* downregulation has previously been reported in BL (Wang, Yang et al. 2014). *MiR-150* plays an important role in BL by targeting MYB, which has an essential role in hematopoietic and lymphoid development and apoptosis (Xiao, Calado et al. 2007). Re-expression of *miR-150* reduced the proliferation of lymphoma cells, which suggests it to be a promising novel strategy for tumor treatment (Wang, Yang et al. 2014). Furthermore, *miR-150* has also been found to have a role in the adaptive immune response (Zhong, Zhang et al. 2012) and was shown to be upregulated in cutaneous marginal zone B-cell lymphomas (Monsalvez, Montes-Moreno et al. 2013). Further dysregulated miRNAs include *miR-211*, miR-*548ac and miR-1244-1/3*, which represent novel deregulated targets in lymphomagenesis. Seven of these eight miRNAs have not been associated with hematologic malignancies yet, which offers a unique starting point for future analysis. *MiR-17-92* is located at 13q31.3; amplification of this region has been identified in several hematopoietic malignancies including DLBCLs, FLs and BLs (Ota, Tagawa et al. 2004). The *miR-17-92* cluster consists of six miRNAs (*miR-17, miR-18a, miR-19a, miR-20a, miR-19b and miR-92a*) and two paralogs (the *miR-106b-25* cluster and the *miR-106a-363* cluster) (Petrocca, Vecchione et al. 2008, Okada, Kohanbash et al. 2010).

Despite the significant role of *miR-17-92* cluster, no study has previously evaluated its diagnostic value in differentiating DLBCL from BL. Here we observed high expression of

miR-17-92 in BL compared to DLBCL. Among the other miRNAs which were expressed differentially in DLBCL and BL cases, we found the *let-7* family highly expressed in DLBCL. In addition, a total of 198 miRNAs (150 upregulated, 48 downregulated) were deregulated in BL compared to non-tumoral samples. Among the downregulated miRNAs, we found the *let-7* family miRNAs (*let-7d* and *let-7g*) and *miR-155* that are lost in different neoplasia (Sampson, Rong et al. 2007) and BLs (Robertus, Kluiver et al. 2010) respectively. Loss of *miR-155* has been shown to increase levels of AID and to result in a higher frequency of MYC translocations (Dorsett et al, 2008). Other MYC-regulated miRNAs that we found to be upregulated in BL include *miR-23* targeting glutaminase, *miR-125b* targeting IRF4 and PRDM1/BLIMP1 (Robertus, Kluiver et al. 2010). Finally, we observed upregulation of *miR-34b* in BL which has been shown to control C-MYC expression in BL as the more common aberrant control exercised by the immunoglobulin enhancer locus (Leucci, Cocco et al. 2008).

5.1.2 MiRNA expression differences between DLBCL and BL

We compared miRNA expression between DLBCL and BL samples. 98 miRNAs were differentially expressed ($p<0.001$) between DLBCL and BL. Several studies have shown that miR-155 expression can differentiate BL from DLBCL (Gibcus, Tan et al. 2009, Lenze, Leoncini et al. 2011, Fassina, Marino et al. 2012). Lenze and coworkers, by microarray profiling of patient samples, reported expression of 38 miRNAs containing MYC regulated and nuclear factor-kB pathway-associated miRNAs differentially expressed between BL and DLBCL (Lenze, Leoncini et al. 2011). Eight of these miRNAs were included in our list of 98 most discriminatory miRNAs (*miR-146b, miR-221, miR-222, miR-30d, miR-26a-2, miR-339, miR-9, miR-193*). Recently Di Lisio and colleagues reported expression of 470 miRNAs in 43 DLBCL, and 28 BL cases as well as 15 controls (Di Lisio, Sanchez-Beato et al. 2012). In this report, comparative microarray analysis data identified 19 miRNAs discriminating BL and DLBCL. In comparison to our data, seven miRNAs were included in our list (*miR-573, miR-17, miR-92, miR-9, miR-26* and *miR-374)*. *MiR-26a* was previously identified as being highly expressed in DLBCL cases (Lenze, Leoncini et al. 2011); our data also confirmed overexpression of *miR-26a* in DLBCL vs. BL.

5.1.3 MiRNA expression differences between DLBC and FL

Next, we focussed on miRNAs that showed differential regulation in DLBCL and FL**.** DLBCL can arise as de novo or through the transformation of other types of lymphomas, such as FL. Although morphology as well the presence of follicular dendritic cells are essential in the routine differentiation between these two entities, the morphological differentiation between germinal center-derived DLBCL (GC-DLBCL) and high-grade (grade 3) FL can be challenging. Therefore, differentiation between DLBCL and FL can be problematic and needs new diagnostic tools. Expression profiling studies reported miRNAs might be useful tools in differentiating DLBCL and FL. Recently, several miRNAs expression profiling reports focused on DLBCL and FL (Roehle, Hoefig et al. 2008, Lawrie, Chi et al. 2009, Fassina, Marino et al. 2012).

Despite the important role of miRNA in differentiating DLBCL and FL, no study has previously assessed differentially expressed miRNA in these subtypes by using deep sequencing technology. In this study, when comparing DLBCL to FL profiles, 87 miRNAs were differentially expressed. FL was characterized mostly via downregulation of miRNAs (51 of 87). Some of these miRNAs have already been found to play an important role in tumor development. For instance, *miR-150* that was previously demonstrated as being differentially expressed in FL and DLBCL (Roehle, Hoefig et al. 2008, Lawrie, Chi et al. 2009), has been reported to control B-cell development and is significantly upregulated in patients with CLL (Fulci, Chiaretti et al. 2007, Zhou, He et al. 2010). Components of *miR-17-92* cluster have been previously reported as most dysregulated miRNAs in both DLBCL and FL (Roehle, Hoefig et al. 2008, Lawrie, Chi et al. 2009). MYC oncogene has been reported to upregulate the expression of *miR-17-92* and is considered as a key regulator of miRNAs in DLBCL (Li, Kim et al. 2009). After checking *miR-17-92* cluster expression in DLBCL and FL, as we expected, our data showed overexpression of three *miR-17-92* miRNAs (*miR-17, miR-18a, miR-20a*) in DLBCL in comparison to FL.

5.1.4 MiRNA expression differences between BL and FL

Here, for the first time, we report miRNA profiles comparing BL vs. FL. In a previous study, the comparison of MYC-high BLs to all other MYC-low lymphomas (MCL, FL and CLL) revealed 122 differentially expressed miRNAs, including 39 of 50 evaluable known targets of MYC (Robertus, Kluiver et al. 2010). Based on a 2-fold or greater difference in mean values in FL samples compared to BL ($P<0.001$), we found that 108 miRNAs were significantly

differentially expressed in these two subtypes, of which 65 miRNAs were down-regulated and 42 miRNAs were upregulated in FL. Here, our data showed that expression of *miR-9* in FL is higher than in BL. Upregulation of *miR-9* in FL was previously described in different studies (Roehle, Hoefig et al. 2008, Di Lisio, Sanchez-Beato et al. 2012). *MiR-9*, which is activated by MYC, has been shown to be involved in NF-kB pathway (Guo, Pu et al. 2009). Moreover, *miR-9,* by targeting PRDM1/Blimp-1 complex, which is a crucial complex in lymphoma, plays an important role in lymphomagenesis (Lawrie, Chi et al. 2009). Consistent with a previous report (Robertus, Kluiver et al. 2010), we found high expression of the *miR-17-92* cluster in BL compared to FL.

5.1.5 Overexpression of *miR-143* in lymphoma

MiR-143 has been suggested to be a tumor suppressor and was down-regulated in many types of cancer such as colorectal carcinoma (Michael, SM et al. 2003) and some lymphomas (Akao, Nakagawa et al. 2007). However, a recent study examined expression of 464 miRNAs in 80 DLBCL and 18 FL cases and showed that *miR-143* was upregulated in DLBCL and FL samples (Lawrie, Chi et al. 2009). In our study, we also showed that expression levels of *miR-143* were significantly upregulated in DLBCL, including ABC cases. *MiR-143* as a nuclear factor-kB (NF-KB) transactivational target promotes tumor cell invasion and metastasis (Zhang, Liu et al. 2009). High expression of NF-kB target genes has been shown to be a hallmark of ABC-DLBCL and was associated with a poor prognosis in these patients (Pavan, Spina et al. 2008). Taken together, *miR-143* might play an oncogenic role in B-cell lymphoma.

To conclude, this study shows the set of differentially expressed miRNAs that increases the range of potential diagnostic markers for lymphoma, especially when differential diagnosis of BL, DLBCL and FL is required. Our data supports and extends previous studies showing miRNAs as novel biomarkers to separate different subtypes of lymphoma. One explanation for the differences between our data and previous data is that available data are preliminary, as published profiles were not derived from large sample collections and also originate mostly from PCR-based approaches and microarrays.

5.1.6 Mutated *MiR-142* is found in DLBCL as well as FL

We identified 17 mutations in 12 patients corresponding to eight distinct miRNAs (*miR-142, miR-612, miR-3655, miR-4322, miR-4537, miR-4538, miR-4539 and miR-5195*). Of these eight mutated miRNAs, *miR-142* with a total of seven different mutations in six patients, was the most frequently mutated miRNA. Different studies indicated that *miR-142* functions as a tumor suppressor in many types of cancer, for instance by targeting CD133, ABCG2, and Lgr5 in colon cancer cells (Shen, Zeng et al. 2013) or by targeting MLL-AF4 oncogene in human acute lymphoblastic leukemia (Dou, Li et al. 2013). Furthermore, *miR-142* has been shown to function as a potential tumor suppressor in human osteosarcoma by targeting HMGA1 (Xu, Wang et al. 2014). Kwanhian et al. reported a mutation in the seed sequence of *miR-142* in approximately 20% of DLBCL patients; in addition, their data suggest that the mutations within *miR-142* result in a loss of function that might contribute to cell growth stimulation and induction of DLBCL (Kwanhian, Lenze et al. 2012). Our data therefore confirms the mutation frequency in DLBCL and this is the first report showing similar mutations also in FL.

Recently, much interest has been focused on RNA editing, increasing the possibility that some of the sequence variations are actually RNA editing sites rather than hereditary SNPs. We explored this possibility that mutations which were observed only on the RNA level represent bona fide miRNA editing events.

RNA editing is a process in which members of the double-stranded-RNA (dsRNA) specific ADAR (adenosine deaminase acting on RNA) family perform the hydrolytic deamination of adenosine to inosine (A-to-I editing) (Polson, Crain et al. 1991). RNA editing is crucial for the development and normal life of both invertebrates and vertebrates (Higuchi, Maas et al. 2000, Palladino, Keegan et al. 2000), but there is still a lack of comprehensive understanding of the role of RNA editing. When an RNA molecule contains an inosine, any one of 3 bases will be incorporated in the strand that is created by next generation sequencing: A, C or T (Illumina, personal communication), the most commonly sequenced change being A-to-G. In this study, after excluding known SNVs (dbSNP) as well as rare (below 10%) variants, we obtained seven SNVs. The five SNVs which represented A-to-H changes mapped to five miRNAs (*miR-160b, miR-4772, miR-376a1, miR-376a2 and miR-376c*). Although not yet in lymphoma, the miR-376 family has previously been shown to be subject to miRNA editing in different cancer types, resulting in an altered mRNA target profile (Mizuguchi, Mishima et al. 2011, Choudhury, Tay et al. 2012). The identification of edited RNA molecules, including miRNAs, which play crucial roles in lymphomagenesis in a sequence-dependent manner, will

increase our understanding of the role of A-to-I editing under normal conditions and the consequent impact of its disruption in cancer.

5.2 Novel miRNAs in GCB lymphoma

One of the important advantages of small RNA sequencing (in comparison to array based-approaches) is that it allows the bioinformatic prediction of novel miRNAs (Creighton, Reid et al. 2009). Our data also supports this potential of next generation sequencing to uncover novel miRNAs. After bioinformatic analysis, we found 283 predicted novel miRNAs. This is a striking advance compared to recent array-based studies of lymphoma which could not identify novel miRNAs (Lenze, Leoncini et al. 2011). To further validate our data, we subjected 20 miRNA candidates to Northern blot analysis. Northern blot is one of the early methods for the detection of single miRNA molecules. This technique has the advantage of target detection without any need for chemical or enzymatic modification of the target molecules (Wark, Lee et al. 2008). Furthermore, compared to other methods such as qRT-PCR, this method is able to detect both pre-miRNAs and mature miRNAs. Among 20 novel miRNA candidates, we successfully validated four of them. Those candidates not identified by Northern blotting may be below detection threshold and have low-level expression, therefore, only detectable by deep sequencing. In this case, Northern blot analysis may be not sensitive enough to detect their expression.

5.2.1 Application of pre-miRs and anti-miRs for novel miRNAs

The Northern blot result provides experimental evidence to confirm that these four novel miRNAs (*NB-miR-1, NB-miR-5, NB-miR-19* and *NB-miR-20*) are bona fide novel miRNAs expressed in lymphoma samples. Two novel miRNAs (*NB-miR-5 and NB-miR-19*), which were detected in both BL and non-BL cell lines, were selected for further functional analysis. To modulate miRNA activity, efficient and reliable approaches are required. Different technological innovations have been developed to facilitate the understanding of miRNA function. Overexpression and/or inhibition assays are the most often used approaches to investigate miRNA effects on gene expression. Anti-miR mechanisms block the mature miRNA and thus separate it from participating in the miRISC-associated inhibition of target mRNAs. RNA molecules that mimic mature miRNA have been used to increase the expression of specific miRNAs (Cheng, Byrom et al. 2005, Davis, Lollo et al. 2006,

Borgdorff, Lleonart et al. 2010). Here, we assessed the potency of miRNA mimicking and antisense oligonucleotides to increase and inhibit the activity of specific miRNAs in cell culture, respectively.

Several things should be considered before using anti-mir and pre-mir, the first is to protect synthetic anti-miRs and pre-miRs from cellular nucleases when applied to cell culture, the second is to improve affinity for the target miRNA (in case of anti-miRs) or mRNA (in case of pre-miRs) and enhanced RNase H or other enzyme activity that is involved in miRNA function (Esau 2008). To enhance novel miRNAs, instead of using mimic miRNA, we cloned pre-miRNAs into a destination vector. In order to improve affinity for the target miRNAs, we applied LNA probes. The use of LNA probes for the knockdown of miRNA is particularly effective because these probes have high affinity for their short RNA targets and allow the discrimination of closely related miRNA sequences. Next, we performed qRT-PCR to evaluate functionality of pre-miRs and anti-miRs in cell culture. MiRNA expression profiling by qRT-PCR shows that transfection with pre-miRs significantly increased the expression of both novel miRNAs in SU- DHL-4 and Raji cell lines. Except for *NB-miR-19* in SU-DHL-4 cell line, we observed a significant reduction of both novel miRNAs levels in cells transfected with LNA probes. It is worth nothing that, sometimes inhibition of miRNAs is achieved without detectable miRNA degradation (Elmen, Lindow et al. 2008, Davis, Propp et al. 2009), suggesting that qRT-PCR might not always be an appropriate method to measure inhibition. Furthermore, in some cases, high affinity chemical modifications make the anti-miR:miRNA complex stable and interfere with miRNA detection by qRT-PCR (Davis, Propp et al. 2009). However, compared to prior studies (Davis, Lollo et al. 2006, Ovcharenko, Kelnar et al. 2007, Lennox and Behlke 2010), we confirmed that synthetic pre-miR and anti-miR are functional methods to, respectively, increase or inhibit the activity of endogenous miRNAs in cultured cells.

5.2.2 RNAseq data of novel miRNAs

Various methods have been suggested for miRNA target identification. Measurements of mRNA expression by RNAseq method have revealed mRNAs with different expression in the presence or absence of individual miRNAs. To identify novel miRNA targets we performed RNAseq after overexpression and knockdown of both *NB-miR-5* and *NB-miR-19.* Overexpression and knockdown of both novel miRNAs caused upregulation and

downregulation of a large number of genes involved in functions related to cell proliferation, such as DNA replication, cell cycle and mitosis.

After knockdown and overexpression of *NB-miR-5* in SU-DHL-4, we detected 242 genes, and 34 genes were found to be shared in both experiments. We identified protein kinase-related genes in this group, including ATM, GRK6. ATM (ataxia telangiectasia mutated) gene encodes serine-threonine kinase that belongs to the PI3/PI4-kinase family (Shiloh 2001). This enzyme is an important cell cycle checkpoint kinase, playing a crucial role in signaling pathways activated by DNA damage (Lowndes and Murguia 2000). Many studies showed ATM has different targets; one of the important targets is p53 that participates in the ATM regulation of the G1/S checkpoint activated by DNA damage. In mantle cell lymphoma (MCL), a specific subtype of NHLs, ATM is an interesting target as a tumor suppressor gene with pathogenic function (Camacho, Hernandez et al. 2002). In addition, ATM aberrations in MCL are independent of p53 gene status; here we also observed no altered p53 gene expression. The other protein kinase is GRK6; this gene encodes a member of the guanine nucleotide-binding protein (G protein)-coupled receptor kinase and is known as one of the CXCR4 expression regulators (Fong, Premont et al. 2002). CXCR4 plays a crucial role in the segregation of dark and light zones in GCs (Allen, Ansel et al. 2004). In addition, CXCR4 has been known as a coreceptor for HIV infection and as a key molecule for cancer metastasis (Feng, Broder et al. 1996, Tsutsumi, Tanaka et al. 2007). Our data showed that overexpression of *NB-miR-5* in SU-DHL-4 activates maintenance complex component 2 (MCM2) genes, which prepare chromatin for DNA replication and as a result, progression of the cell through the S-phase (Dimitrova, Todorov et al. 1999). Among the transcription factors, cell division cycle-associated protein (CDCA7), which is known to participate in the regulation of cell proliferation, was upregulated in response to *NB-miR-5* overexpression in both Raji and SU-DHL-4. This gene has been identified as a C-MYC target gene and is frequently overexpressed in human cancers such as NHL (Osthus, Karim et al. 2005). After determining a list of genes, the next step would be to map these genes to known pathways/Gene Ontology terms and identify pathways which are overrepresented in a given set of genes.

5.2.2.1 *NB-miR-19* meets NF-KB and JAK/STAT pathway

Our data showed that overexpression of *NB-miR-19* in SU-DHL-4 led to altered expression of NF-KB members such as RelB, NF-kB1 and NF-kB2. Recently several studies implicated that aberrant NF-KB activation is a hallmark of several lymphoid malignancies and is directly

linked to advanced disease (Jost and Ruland 2007, Staudt 2010). Among NF-KB target genes, we found overexpression of BCL-2 family members which previously had been shown in the ABC-DLBCL subtype (Davis, Brown et al. 2001). We found 2.5-fold upregulation of CARD11, a membrane-associated guanylate kinase (MAGUK) family member, after *NB-miR19* overexpression. In ABC-DLBCL, survival of the malignant cells is dependent on constitutive activation of the NF-KB signaling pathway. Lenz et al. showed that mutation of CARD11 in DLBCL resulted in constitutive NF-KB activation and enhanced NF-KB activity upon antigen receptor stimulation (Lenz, Davis et al. 2008). Taken together, CARD11 contributes to lymphomagenesis and can be considered as a target for DLBCL therapy. In addition, our data connects expression of *NB-miR-19* to the JAK/STAT (Janus associated kinase-signal transducer and activator of transcription) signalling pathway. The JAK/STAT pathway as an important intracellular signaling cascades consists of a variety of cytokines and growth factors (Furqan, Mukhi et al. 2013). JAK/STAT pathways have also been identified to be critical players in the pathogenesis of leukemia and lymphoma (Kopp, Ralfkiaer et al. 2013). We found all four cytoplasmic tyrosine kinases member of JAK family (JAK1, 2, 3 and Tyk2) were upregulated around 3-fold in overexpressed *NB-miR-19* samples. JAK2 has been shown to upregulate the expression of about 2,000 genes in cancer cells, including well-defined oncogenes like MYC and NF-KB pathway genes, such as IRF4 and CD40 (Vainchenker and Constantinescu 2013). Interestingly, we found around 2-fold upregulation of CD40 in SU-DHL-4 after *NB-miR-19* overexpression. JAK2 mutations have been identified in a significant number of myeloproliferative disorders, therefore the JAK2/STAT signaling pathway is considered as a new druggable pathway (Levine, Pardanani et al. 2007). JAK1 plays an exclusive role in lymphohematopoiesis; Flex et al. implicated several acquired JAK1 missense mutations in adult lymphoblastic leukemia, which are associated with advanced age and poor prognosis (Flex, Petrangeli et al. 2008). Furthermore overexpression of *NB-miR-19* increased more than 4-fold the expression of STAT5 that was previously identified as a malefactor in the pathogenesis of lymphoma as well as several other cancers (Kopp, Ralfkiaer et al. 2013).

In conclusion, these results revealed that *NB-miR-19* might regulate the JAK/STAT pathway but the specific role of *NB-miR-19* in JAK/STAT signaling remains to be investigated and needs further functional analysis.

5.3 AGO2-PAR-CLIP reveals miRNA targets in lymphoma cell lines

MiRNAs are small, highly conserved small RNA molecules that play important roles in post-transcriptional gene regulation. MiRNAs are part of the RNA-induced silencing complex (RISC) and guide this complex to specific miRNA recognition elements (MREs) on the mRNA target molecules resulting in mRNA degradation and translational repression (Bartel 2009). Most of the MREs have been identified in 3'UTRs of protein coding genes (Maragkakis, Alexiou et al. 2009). However some reports indicated that MREs can also be located in CDSs of target genes (Tay, Zhang et al. 2008). For instance, *miR-148* represses specific splice variants of DNA methyltransferase 3b (Dnmt3b) by targeting sequences in CDS region (Duursma, Kedde et al. 2008). Elcheva et al. showed that the coding region of beta-transducin repeat containing protein 1 is regulated by *miR-183* (Elcheva, Goswami et al. 2009).

The most often used approach to investigate miRNA targets has been to transfect cells with pre-mir and/or anti-mir and measure resulting changes in mRNA or protein levels (Grimson, Farh et al. 2007, Baek, Villen et al. 2008). This approach is challenging because results may arise from changes in the levels of transcription factors that are targeted by the miRNA, rather than being direct consequences of miRNA-mRNA interaction (Cloonan, Brown et al. 2008). CLIP-based techniques such as PAR-CLIP allow for the direct identification and localization of MREs. Hafner et al., using AGO2-PAR-CLIP, showed that miRNAs bind to target mRNAs in approximately equal proportions on the 3'UTR as well as on the CDS (Hafner, Landthaler et al. 2010). Hundreds of miRNAs function through one of the four human Argonaute proteins (AGO) and mediate transcriptional inhibition of target mRNAs (Bartel 2009). The four human AGO genes are co-expressed and bind to miRNAs irrespective of their sequence. AGO2, in contrast to the other members, retains an active RNase H domain and is thus able to directly cleave target mRNAs with extensive complementarity to the bound miRNAs (Hutvagner and Simard 2008). In order to capture miRNA targets and describe their location within 3′ UTRs and CDS in lymphoma cell lines, we performed AGO2 PAR-CLIP experiments (Hafner, Landthaler et al. 2010, Spitzer, Hafner et al. 2014). Recently Farazi et al. revealed potential therapeutic targets and prognostic markers in breast cancer using AGO2 PAR-CLIP (Farazi, Ten Hoeve et al. 2014). To compare miRNAs targets in non-BL and BL subtypes, we cultured two cell lines of each subtypes in the presence of 100 µM 4SU. Among four PAR-CLIP libraries, the Raji sample showed lowest reads. Six PAR-CLIP targets

(SH2B2, HERC2, SVEP1, SIPA1L3, CANT1, and DHX3) were identified in both non-BL cell lines (SU-DH-4 & SU-DHL-6). SH2B2 (SH2B adaptor protein 2), also known as APS (Adaptor protein containing PH and SH2 domains), is exclusively expressed in B cells and tyrosine phosphorylated in response to BCR stimulation, suggesting its involvement in immunoreceptor signaling (Yokouchi, Suzuki et al. 1997). HERC2 is considered as a novel regulator of p53 signaling. Cubillos et al. demonstrated that HERC2 depletion reduces the transcriptional activity of p53 without affecting its stability and suggested HERC2 that might be associated with cancer in humans (Cubillos-Rojas, Amair-Pinedo et al. 2014). CANT1 (calcium activated nucleotidase 1) is an androgen-regulated protein, and is overexpressed in prostate cancer, reduction of CANT1 expression reduces prostate cell proliferation (Patel, Halling-Brown et al. 2013). Among the miRNA targets which were found in three cell lines, we identified MYBL1, also known as A-MYB. This is a gene member of MYB transcription factors family, including C-MYB and B-MYB. MYBL1 which was found in NAMALWA, SU-DHL-4 and SU-DHL-6 PAR-CLIP libraries, has been shown to be involved in DLBCL, BL and CLL pathogenesis (Golay, Luppi et al. 1996). The only target found in all four libraries was GSG2 (germ cell associated 2), also known as Haspin, which has been indicated as an important kinase involved in mitosis. Inhibition of Haspin has been proposed as a novel approach in anti-mitotic cancer therapeutics (Dai, Sullivan et al. 2006).

Frequently, activated pathways in NHLs are regulated by miRNA-triggered gene-silencing, in which AGO2, associated mRNAs could be silenced by any miRNA that is associated in the same Argonaute complex. Therefore, it remains a great challenge to confirm all interactions between miRNAs and their target-mRNAs in activated signaling pathways in lymphoma, in order to get more insights into lymphomagenesis. Recently, several approaches have been developed for elucidating of individual miRNA-mRNA interactions. For further analysis of single transcripts that are altered in their expression, dual-luciferase reporter assays could be performed to prove the connection between the overexpressed or inhibited miRNA and the selected target-mRNA. Thus, distinct regulatory mechanisms potentially involved in GCB lymphoma pathogenesis could be validated.

5.4 LncRNA and GCB lymphomas

5.4.1 Differentially expressed and methylated lncRNA AP000251

LncRNAs have been characterized as transcripts of around 200 nucleotides in length without protein coding capacities. Although lncRNAs were initially proposed to be transcriptional noise, current studies have indicated that lncRNAs, like miRNAs, play important roles as gene regulators through transcriptional, post-transcriptional and epigenetic mechanisms. Dysregulation of lncRNA expression is involved in many human diseases and in some cases, it has been observed as a primary feature of human cancers (Faghihi, Modarresi et al. 2008, Gupta, Shah et al. 2010, Lai, Yang et al. 2012). For instance, upregulation of MALAT-1 in bladder cancer triggers the Wnt pathway to promote epithelial mesenchymal transition resulting in metastasis (Ying, Chen et al. 2012). It is worth noting that genome-wide expression and functional significance of lncRNAs in NHLs still remains unresolved. High-throughput RNA sequencing has promoted the study to discover previously uncharacterized lncRNAs. In this study, we evaluated the lncRNA expression profile in the 56 GCB-lymphoma samples by deep sequencing to reveal the potential role of lncRNAs in the pathogenesis of NHLs. NGS sequencing data revealed a significantly differentially expressed lncRNA AP000251 in BL and non-BL samples compared to the control. For further validation of NGS results, qRT-PCR was performed to evaluate the expression patterns of AP000251 in BL and non-BL lymphoma cell lines.

The qRT-PCR result is in agreement with the data from the NGS. We demonstrated for the first time the differential expression of the lncRNA in GCB-lymphoma subtypes, suggesting AP000251 as a biomarker to distinguish BL from non-BL cases. In addition to differential expression of AP000251 in GCB subtypes, we also describe its differential methylation pattern in normal B-cells as well as BL and non-BL. Methylation changes in CpG islands (CGI) and CpG shores (low CpG density areas ~2 kb close to CGI) affect gene expression (Ji, Ehrlich et al. 2010). In general, DNA methylation is associated with loss of gene expression (Bird 1986). Several studies found that the expression of lncRNAs is regulated by epigenetic mechanism such as DNA methylation (Lujambio, Portela et al. 2010, Stadtfeld, Apostolou et al. 2010). Altered methylation patterns are thought to be an initial event in some cancers (Lujambio, Portela et al. 2010) and tumor specific methylation patterns can be applied for the detection and classification of cancer. Taken together, we assume that the overexpression of AP000251 might be caused by hypomethylation of lncRNA AP000251, suggesting a potential oncogenic role for AP000251.

5.4.2 Overexpression and knockdown of AP000251 in lymphoma cell lines

In contrast to miRNAs or mRNAs, the functions of lncRNAs cannot be determind according to their sequence or primary structure alone (Mercer, Dinger et al. 2009). Moreover, unlike miRNAs, which exhibit strong conservation across diverse species, the lack of strong conservation of lncRNAs compounds the difficulties (Kondrashov, Kiefmann et al. 2005). A number of lncRNAs termed ncRNA-activating (ncRNA-a) regulate transcription of neighbouring genes on the same chromosome by using a cis-mediated mechanism (Wang and Chang 2011, Lai, Orom et al. 2013, Melo, Drost et al. 2013). This group of lncRNAs act close to their site of synthesis in an RNA-dependent manner or by RNA-independent mechanisms. In the first mechanism, they facilitate looping onto the promoter regions of their transcriptional targets, while in the second mechanism they locally change the chromatin status. TIAM1 (a Rac-specific GEF) was originally characterized by its ability to promote T-cell lymphoma invasion and has been implicated to be overexpressed in some cancers such as breast cancer (Ellenbroek, Iden et al. 2012). To assess the function of lncRNA AP000251, we proposed AP000251 might regulate TIAM1 gene expression which lies upstream of AP000251 on chromosome 21. However, our qRT-PCR data showed no significant differences in TIAMI expression upon knockdown or overexpression of AP000251. Although a number of lncRNAs regulate transcription of neighbouring genes by the cis-acting mechanism, many lncRNAs have also been proposed to regulate gene expression using a trans-mechanism (Guttman, Amit et al. 2009, Hung, Wang et al. 2011). In order to capture target genes which are regulated by AP000251, we first performed knockdown and forced expression of AP000251 in a lymphoma cell line, next we sequenced replicate samples using the Illumina HiSeq system.

5.4.3 Knockdown of lncRNA AP000251 in Raji cell line

We found 668 dysregulated genes upon AP000251 knockdown in our profiling experiments, of which 230 genes and 438 genes have been shown to be downregulated and upregulated, respectively. Among the downregulated genes, we found genes that are involved in cell cycle regulation (CCNB1), mitosis (HMMR, BUB1B, TPX2 and KIF2C) (Hu, Wu et al. 2012) and DNA modulation (TOP2A, HMGB1) (Bjorck, Ek et al. 2005). CCNB1 forms a complex with CDC2, which is a crucial complex in the G2/M phase of the cell cycle. Since CCNB1 is a key

regulator of the G2/M transition of the cell cycle, this together with the other genes that have significant impact on survival and proliferation of cells (Bourne, Watson et al. 1996), might indicate a role for AP000251 in the proliferation phase of cells.

AP000251 knockdown led to consistent level changes of PCNT (46,324,069-46,445,768) and SOD (31,659,621- 31,668,930) which are both located on chromosome 21. This might suggest a cis-acting regulatory role of AP000251 in lymphoma cell lines. PCNT (Pericentrin) is expressed in acute myeloid leukemia (AML) and high levels of PCNT correlate with aneuploidy and centrosome aberrations in AML patients. Importantly, the knockdown of AP000251 was associated with increased transcription of MYC, suggesting that AP000251 operates to restrict the activity of MYC. Among the other overexpressed genes, we found transcription factor TCF3 (E2A) that has been found mutated in BL cases and is involved in PI3K signaling pathway. In addition to an association with BL, E2A/PI3K has essential functions in B-cell fate and GC biology (Kwon, Hutter et al. 2008, Lin, Jhunjhunwala et al. 2010).

5.4.4 Overexpression of lncRNA AP000251 in SU-DHL-4 cell line

After overexpression of AP000251 in SU-DHL-4, 120 differentially expressed genes have been defined. We observed reduction of some interesting genes (E2F2, TOP2A, FOXM1, NUSAP1, BRCA1 and PLK1) upon overexpression of AP000251 in the SU-DHL-4 cell line. E2F2 is an E2Fs family member which is involved in the regulation of genes that are essential for proliferation, differentiation, development and apoptosis (Muller, Bracken et al. 2001). While the other members of this family (E2F1 and E2F3) show a broad expression, E2F2 expression is restricted to the hematopoietic cell lineage (Li, Zhu et al. 2003). Rempel et al. implicated a role for E2F2 activity in affecting MYC-induced oncogenesis (Rempel, Mori et al. 2009). NUSAP1 is a nucleolar-spindle-associated protein that plays a role in spindle microtubule organization (Raemaekers, Ribbeck et al. 2003). Previous reports have indicated that high expression of NUSAP1is consistent with microtubule bundling and cell cycle arrest at the G2/M checkpoint (Vanden Bosch, Raemaekers et al. 2010) and has been observed in many cancers including prostate, melanoma, glioblastoma, and hepatocellular carcinoma (Bogunovic, O'Neill et al. 2009, Satow, Shitashige et al. 2010). Breast cancer-associated gene 1 (BRCA1) is a specific tumor suppressor in breast and ovarian cancer (Miki, Swensen et al. 1994, Huen, Sy et al. 2010). It is worth noting that in this study we found downregulation of both NUSAP1and BRCA1upon overexpression of AP000251.

One report indicated that NUSAP1 and BRCA1 are implicated in the same pathways of DNA repair and centrosome duplication and NUSAP1 is able to control BRCA1 protein abundance (Kotian, Banerjee et al. 2014). FOXM1 (forkhead box transcriptional factor) expression is elevated in numerous types of human malignancies in liver, prostate, breast, colon, neural and other tissues (Kalin, Wang et al. 2006, Wang, Park et al. 2011). FOXM1 facilitates development of cancers by transcriptional regulation of genes involved in cell proliferation and promotes progression through G/S and G/M phases of the cell cycle (Wang, Chen et al. 2005). Finally, Polo-like kinase 1 (PLK1) plays a role in several aspects of the cell cycle progression, especially in the G2/M phase (Lowery, Lim et al. 2005). It is not expressed in most normal differentiated tissues; however, it is overexpressed in a wide variety of cancers and its expression is often associated with poor outcome (Liu, Zhang et al. 2008). Liu et al. showed overexpression of PLK1 in NHL patients and in addition there is a link between overexpression and clinical feature in this disease (Liu, Zhang et al. 2008).

Taken together, most genes affected upon dysregulation of AP000251 are cell cycle regulators; therefore it can be assumed that AP000251 might interfere with cell cycle regulation. However, further studies are required to determine whether lncRNA AP000251 may serve as a new therapeutic target and/or diagnostic biomarker in NHLs.

References

References(1993). "A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project." N Engl J Med **329**(14): 987-994.

(1997). "A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. The Non-Hodgkin's Lymphoma Classification Project." Blood **89**(11): 3909-3918.

Akao, Y., Y. Nakagawa, Y. Kitade, T. Kinoshita and T. Naoe (2007). "Downregulation of microRNAs-143 and -145 in B-cell malignancies." Cancer Sci **98**(12): 1914-1920.

Akyurek, N., A. Uner, M. Benekli and I. Barista (2012). "Prognostic significance of MYC, BCL2, and BCL6 rearrangements in patients with diffuse large B-cell lymphoma treated with cyclophosphamide, doxorubicin, vincristine, and prednisone plus rituximab." Cancer **118**(17): 4173-4183.

Alexander, D. D., P. J. Mink, H. O. Adami, E. T. Chang, P. Cole, J. S. Mandel and D. Trichopoulos (2007). "The non-Hodgkin lymphomas: a review of the epidemiologic literature." Int J Cancer **120 Suppl 12**: 1-39.

Alizadeh, A. A., M. B. Eisen, R. E. Davis, C. Ma, I. S. Lossos, A. Rosenwald, J. C. Boldrick, H. Sabet, T. Tran, X. Yu, J. I. Powell, L. Yang, G. E. Marti, T. Moore, J. Hudson, Jr., L. Lu, D. B. Lewis, R. Tibshirani, G. Sherlock, W. C. Chan, T. C. Greiner, D. D. Weisenburger, J. O. Armitage, R. Warnke, R. Levy, W. Wilson, M. R. Grever, J. C. Byrd, D. Botstein, P. O. Brown and L. M. Staudt (2000). "Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling." Nature **403**(6769): 503-511.

Allen, C. D., K. M. Ansel, C. Low, R. Lesley, H. Tamamura, N. Fujii and J. G. Cyster (2004). "Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5." Nat Immunol **5**(9): 943-952.

Anderson, J. R., J. O. Armitage and D. D. Weisenburger (1998). "Epidemiology of the non-Hodgkin's lymphomas: distributions of the major subtypes differ by geographic locations. Non-Hodgkin's Lymphoma Classification Project." Ann Oncol **9**(7): 717-720.

Armitage, J. O. and D. D. Weisenburger (1998). "New approach to classifying non-Hodgkin's lymphomas: clinical features of the major histologic subtypes. Non-Hodgkin's Lymphoma Classification Project." J Clin Oncol **16**(8): 2780-2795.

Baek, D., J. Villen, C. Shin, F. D. Camargo, S. P. Gygi and D. P. Bartel (2008). "The impact of microRNAs on protein output." Nature **455**(7209): 64-71.

Baffa, R., M. Fassan, S. Volinia, B. O'Hara, C. G. Liu, J. P. Palazzo, M. Gardiman, M. Rugge, L. G. Gomella, C. M. Croce and A. Rosenberg (2009). "MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets." J Pathol **219**(2): 214-221.

Bartel, D. P. (2004). "MicroRNAs: genomics, biogenesis, mechanism, and function." Cell **116**(2): 281-297.

Bartel, D. P. (2009). "MicroRNAs: target recognition and regulatory functions." Cell **136**(2): 215-233.

Bastion, Y., C. Sebban, F. Berger, P. Felman, G. Salles, C. Dumontet, P. A. Bryon and B. Coiffier (1997). "Incidence, predictive factors, and outcome of lymphoma transformation in follicular lymphoma patients." J Clin Oncol **15**(4): 1587-1594.

Bea, S., A. Zettl, G. Wright, I. Salaverria, P. Jehn, V. Moreno, C. Burek, G. Ott, X. Puig, L. Yang, A. Lopez-Guillermo, W. C. Chan, T. C. Greiner, D. D. Weisenburger, J. O. Armitage, R. D. Gascoyne, J. M. Connors, T. M. Grogan, R. Braziel, R. I. Fisher, E. B. Smeland, S. Kvaloy, H. Holte, J. Delabie, R. Simon, J. Powell, W. H.

Wilson, E. S. Jaffe, E. Montserrat, H. K. Muller-Hermelink, L. M. Staudt, E. Campo, A. Rosenwald and P. Lymphoma/Leukemia Molecular Profiling (2005). "Diffuse large B-cell lymphoma subgroups have distinct genetic profiles that influence tumor biology and improve gene-expression-based survival prediction." Blood **106**(9): 3183-3190.

Berezikov, E., W. J. Chung, J. Willis, E. Cuppen and E. C. Lai (2007). "Mammalian mirtron genes." Mol Cell **28**(2): 328-336.

Berezikov, E., F. Thuemmler, L. W. van Laake, I. Kondova, R. Bontrop, E. Cuppen and R. H. Plasterk (2006). "Diversity of microRNAs in human and chimpanzee brain." Nat Genet **38**(12): 1375-1377.

Bird, A. P. (1986). "CpG-rich islands and the function of DNA methylation." Nature **321**(6067): 209-213.

Bjorck, E., S. Ek, O. Landgren, M. Jerkeman, M. Ehinger, M. Bjorkholm, C. A. Borrebaeck, A. Porwit-MacDonald and M. Nordenskjold (2005). "High expression of cyclin B1 predicts a favorable outcome in patients with follicular lymphoma." Blood **105**(7): 2908-2915.

Blum, K. A., G. Lozanski and J. C. Byrd (2004). "Adult Burkitt leukemia and lymphoma." Blood **104**(10): 3009-3020.

Boerma, E. G., R. Siebert, P. M. Kluin and M. Baudis (2009). "Translocations involving 8q24 in Burkitt lymphoma and other malignant lymphomas: a historical review of cytogenetics in the light of todays knowledge." Leukemia **23**(2): 225-234.

Bogunovic, D., D. W. O'Neill, I. Belitskaya-Levy, V. Vacic, Y. L. Yu, S. Adams, F. Darvishian, R. Berman, R. Shapiro, A. C. Pavlick, S. Lonardi, J. Zavadil, I. Osman and N. Bhardwaj (2009). "Immune profile and mitotic index of metastatic melanoma lesions enhance clinical staging in predicting patient survival." Proc Natl Acad Sci U S A **106**(48): 20429-20434.

Borgdorff, V., M. E. Lleonart, C. L. Bishop, D. Fessart, A. H. Bergin, M. G. Overhoff and D. H. Beach (2010). "Multiple microRNAs rescue from Ras-induced senescence by inhibiting p21(Waf1/Cip1)." Oncogene **29**(15): 2262-2271.

Bourne, Y., M. H. Watson, M. J. Hickey, W. Holmes, W. Rocque, S. I. Reed and J. A. Tainer (1996). "Crystal structure and mutational analysis of the human CDK2 kinase complex with cell cycle-regulatory protein CksHs1." Cell **84**(6): 863-874.

Bracken, C. P., P. A. Gregory, N. Kolesnikoff, A. G. Bert, J. Wang, M. F. Shannon and G. J. Goodall (2008). "A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition." Cancer Res **68**(19): 7846-7854.

Breese, M. R. and Y. Liu (2013). "NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets." Bioinformatics **29**(4): 494-496.

Bryant, A., M. Lutherborrow and D. Ma (2009). "The clinicopathological relevance of microRNA in normal and malignant haematopoiesis." Pathology **41**(3): 204-213.

Burmeister, T., S. Schwartz, H. A. Horst, H. Rieder, N. Gokbuget, D. Hoelzer and E. Thiel (2005). "Molecular heterogeneity of sporadic adult Burkitt-type leukemia/lymphoma as revealed by PCR and cytogenetics: correlation with morphology, immunology and clinical features." Leukemia **19**(8): 1391-1398.

Calin, G. A. and C. M. Croce (2006). "MicroRNA signatures in human cancers." Nat Rev Cancer **6**(11): 857-866.

Calin, G. A., C. Sevignani, C. D. Dumitru, T. Hyslop, E. Noch, S. Yendamuri, M. Shimizu, S. Rattan, F. Bullrich, M. Negrini and C. M. Croce (2004). "Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers." Proc Natl Acad Sci U S A **101**(9): 2999-3004.

Camacho, E., L. Hernandez, S. Hernandez, F. Tort, B. Bellosillo, S. Bea, F. Bosch, E. Montserrat, A. Cardesa, P. L. Fernandez and E. Campo (2002). "ATM gene inactivation in mantle cell lymphoma mainly occurs by truncating mutations and missense mutations involving the phosphatidylinositol-3 kinase domain and is associated with increasing numbers of chromosomal imbalances." Blood **99**(1): 238-244.

Campo, E., S. H. Swerdlow, N. L. Harris, S. Pileri, H. Stein and E. S. Jaffe (2011). "The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications." Blood **117**(19): 5019-5032.

Carbone, A., A. Gloghini, Y. L. Kwong and A. Younes (2014). "Diffuse large B cell lymphoma: using pathologic and molecular biomarkers to define subgroups for novel therapy." Ann Hematol **93**(8): 1263-1277.

Carlotti, E., D. Wrench, J. Matthews, S. Iqbal, A. Davies, A. Norton, J. Hart, R. Lai, S. Montoto, J. G. Gribben, T. A. Lister and J. Fitzgibbon (2009). "Transformation of follicular lymphoma to diffuse large B-cell lymphoma may occur by divergent evolution from a common progenitor cell or by direct evolution from the follicular lymphoma clone." Blood **113**(15): 3553-3557.

Castellano, L. and J. Stebbing (2013). "Deep sequencing of small RNAs identifies canonical and non-canonical miRNA and endogenous siRNAs in mammalian somatic tissues." Nucleic Acids Res **41**(5): 3339-3351.

Cheetham, S. W., F. Gruhl, J. S. Mattick and M. E. Dinger (2013). "Long noncoding RNAs and the genetics of cancer." Br J Cancer **108**(12): 2419-2425.

Cheloufi, S., C. O. Dos Santos, M. M. Chong and G. J. Hannon (2010). "A dicer-independent miRNA biogenesis pathway that requires Ago catalysis." Nature **465**(7298): 584-589.

Cheng, A. M., M. W. Byrom, J. Shelton and L. P. Ford (2005). "Antisense inhibition of human miRNAs and indications for an involvement of miRNA in cell growth and apoptosis." Nucleic Acids Res **33**(4): 1290-1297.

Cimmino, A., G. A. Calin, M. Fabbri, M. V. Iorio, M. Ferracin, M. Shimizu, S. E. Wojcik, R. I. Aqeilan, S. Zupo, M. Dono, L. Rassenti, H. Alder, S. Volinia, C. G. Liu, T. J. Kipps, M. Negrini and C. M. Croce (2005). "miR-15 and miR-16 induce apoptosis by targeting BCL2." Proc Natl Acad Sci U S A **102**(39): 13944-13949.

Clark, B. S. and S. Blackshaw (2014). "Long non-coding RNA-dependent transcriptional regulation in neuronal development and disease." Front Genet **5**: 164.

Cleary, M. L., T. C. Meeker, S. Levy, E. Lee, M. Trela, J. Sklar and R. Levy (1986). "Clustering of extensive somatic mutations in the variable region of an immunoglobulin heavy chain gene from a human B cell lymphoma." Cell **44**(1): 97-106.

Cloonan, N., M. K. Brown, A. L. Steptoe, S. Wani, W. L. Chan, A. R. Forrest, G. Kolle, B. Gabrielli and S. M. Grimmond (2008). "The miR-17-5p microRNA is a key regulator of the G1/S phase cell cycle transition." Genome Biol **9**(8): R127.

Creighton, C. J., J. G. Reid and P. H. Gunaratne (2009). "Expression profiling of microRNAs by deep sequencing." Brief Bioinform **10**(5): 490-497.

Croce, C. M. (2008). "Oncogenes and cancer." N Engl J Med **358**(5): 502-511.

Cubillos-Rojas, M., F. Amair-Pinedo, R. Peiro-Jordan, R. Bartrons, F. Ventura and J. L. Rosa (2014). "The E3 ubiquitin protein ligase HERC2 modulates the activity of tumor protein p53 by regulating its oligomerization." J Biol Chem **289**(21): 14782-14795.

Cultrera, J. L. and S. M. Dalia (2012). "Diffuse large B-cell lymphoma: current strategies and future directions." Cancer Control **19**(3): 204-213.

Cyster, J. G. (2010). "Shining a light on germinal center B cells." Cell **143**(4): 503-505.

Dai, J., B. A. Sullivan and J. M. Higgins (2006). "Regulation of mitotic chromosome cohesion by Haspin and Aurora B." Dev Cell **11**(5): 741-750.

Dave, S. S., K. Fu, G. W. Wright, L. T. Lam, P. Kluin, E. J. Boerma, T. C. Greiner, D. D. Weisenburger, A. Rosenwald, G. Ott, H. K. Muller-Hermelink, R. D. Gascoyne, J. Delabie, L. M. Rimsza, R. M. Braziel, T. M. Grogan, E. Campo, E. S. Jaffe, B. J. Dave, W. Sanger, M. Bast, J. M. Vose, J. O. Armitage, J. M. Connors, E. B. Smeland, S. Kvaloy, H. Holte, R. I. Fisher, T. P. Miller, E. Montserrat, W. H. Wilson, M. Bahl, H. Zhao, L. Yang, J. Powell, R. Simon, W. C. Chan, L. M. Staudt and P. Lymphoma/Leukemia Molecular Profiling (2006). "Molecular diagnosis of Burkitt's lymphoma." N Engl J Med **354**(23): 2431-2442.

Davis, R. E., K. D. Brown, U. Siebenlist and L. M. Staudt (2001). "Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells." J Exp Med **194**(12): 1861-1874.

Davis, S., B. Lollo, S. Freier and C. Esau (2006). "Improved targeting of miRNA with antisense oligonucleotides." Nucleic Acids Res **34**(8): 2294-2304.

Davis, S., S. Propp, S. M. Freier, L. E. Jones, M. J. Serra, G. Kinberger, B. Bhat, E. E. Swayze, C. F. Bennett and C. Esau (2009). "Potent inhibition of microRNA in vivo without degradation." Nucleic Acids Res **37**(1): 70-77.

DeWolf, W. C., P. H. Lange, R. Shepherd, S. Martin-Alosco and E. J. Yunis (1981). "Association of HLA and renal cell carcinoma." Hum Immunol **2**(1): 41-44.

Di Lisio, L., M. Sanchez-Beato, G. Gomez-Lopez, M. E. Rodriguez, S. Montes-Moreno, M. Mollejo, J. Menarguez, M. A. Martinez, F. J. Alves, D. G. Pisano, M. A. Piris and N. Martinez (2012). "MicroRNA signatures in B-cell lymphomas." Blood Cancer J **2**(2): e57.

Dimitrova, D. S., I. T. Todorov, T. Melendy and D. M. Gilbert (1999). "Mcm2, but not RPA, is a component of the mammalian early G1-phase prereplication complex." J Cell Biol **146**(4): 709-722.

Du, T. and P. D. Zamore (2007). "Beginning to understand microRNA function." Cell Res **17**(8): 661-663.

Duursma, A. M., M. Kedde, M. Schrier, C. le Sage and R. Agami (2008). "miR-148 targets human DNMT3b protein coding region." RNA **14**(5): 872-877.

Elcheva, I., S. Goswami, F. K. Noubissi and V. S. Spiegelman (2009). "CRD-BP protects the coding region of betaTrCP1 mRNA from miR-183-mediated degradation." Mol Cell **35**(2): 240-246.

Ellenbroek, S. I., S. Iden and J. G. Collard (2012). "The Rac activator Tiam1 is required for polarized protrusional outgrowth of primary astrocytes by affecting the organization of the microtubule network." Small GTPases **3**(1): 4-14.

Elmen, J., M. Lindow, S. Schutz, M. Lawrence, A. Petri, S. Obad, M. Lindholm, M. Hedtjarn, H. F. Hansen, U. Berger, S. Gullans, P. Kearney, P. Sarnow, E. M. Straarup and S. Kauppinen (2008). "LNA-mediated microRNA silencing in non-human primates." Nature **452**(7189): 896-899.

Engstrom, P. G., H. Suzuki, N. Ninomiya, A. Akalin, L. Sessa, G. Lavorgna, A. Brozzi, L. Luzi, S. L. Tan, L. Yang, G. Kunarso, E. L. Ng, S. Batalov, C. Wahlestedt, C. Kai, J. Kawai, P. Carninci, Y. Hayashizaki, C. Wells, V. B. Bajic, V. Orlando, J. F. Reid, B. Lenhard and L. Lipovich (2006). "Complex Loci in human and mouse genomes." PLoS Genet **2**(4): e47.

Esau, C. C. (2008). "Inhibition of microRNA with antisense oligonucleotides." Methods **44**(1): 55-60.

Eulalio, A., E. Huntzinger and E. Izaurralde (2008). "Getting to the root of miRNA-mediated gene silencing." Cell **132**(1): 9-14.

Faghihi, M. A., F. Modarresi, A. M. Khalil, D. E. Wood, B. G. Sahagan, T. E. Morgan, C. E. Finch, G. St Laurent, 3rd, P. J. Kenny and C. Wahlestedt (2008). "Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase." Nat Med **14**(7): 723-730.

Falini, B., S. Pileri, H. Stein, D. Dieneman, F. Dallenbach, G. Delsol, O. Minelli, S. Poggi, M. F. Martelli, G. Pallesen and et al. (1990). "Variable expression of leucocyte-common (CD45) antigen in CD30 (Ki1)-positive anaplastic large-cell lymphoma: implications for the differential diagnosis between lymphoid and nonlymphoid malignancies." Hum Pathol **21**(6): 624-629.

Farazi, T. A., J. J. Ten Hoeve, M. Brown, A. Mihailovic, H. M. Horlings, M. J. van de Vijver, T. Tuschl and L. F. Wessels (2014). "Identification of distinct miRNA target regulation between breast cancer molecular subtypes using AGO2-PAR-CLIP and patient datasets." Genome Biol **15**(1): R9.

Fassan, M., C. M. Croce and M. Rugge (2011). "miRNAs in precancerous lesions of the gastrointestinal tract." World J Gastroenterol **17**(48): 5231-5239.

Fassan, M., K. Sachsenmeir, M. Rugge and R. Baffa (2011). "Role of miRNA in distinguishing primary brain tumors from secondary tumors metastatic to the brain." Front Biosci (Schol Ed) **3**: 970-979.

Fassina, A., F. Marino, M. Siri, R. Zambello, L. Ventura, M. Fassan, F. Simonato and R. Cappellesso (2012). "The miR-17-92 microRNA cluster: a novel diagnostic tool in large B-cell malignancies." Lab Invest **92**(11): 1574-1582.

Favre, A., G. Moreno, M. O. Blondel, J. Kliber, F. Vinzens and C. Salet (1986). "4-Thiouridine photosensitized RNA-protein crosslinking in mammalian cells." Biochem Biophys Res Commun **141**(2): 847-854.

Federico, M., M. Bellei, L. Marcheselli, S. Luminari, A. Lopez-Guillermo, U. Vitolo, B. Pro, S. Pileri, A. Pulsoni, P. Soubeyran, S. Cortelazzo, G. Martinelli, M. Martelli, L. Rigacci, L. Arcaini, F. Di Raimondo, F. Merli, E. Sabattini, P. McLaughlin and P. Solal-Celigny (2009). "Follicular lymphoma international prognostic index 2: a new prognostic index for follicular lymphoma developed by the international follicular lymphoma prognostic factor project." J Clin Oncol **27**(27): 4555-4562.

Feng, Y., C. C. Broder, P. E. Kennedy and E. A. Berger (1996). "HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor." Science **272**(5263): 872-877.

Ferreri, A. J., J. Y. Blay, M. Reni, F. Pasini, M. Spina, A. Ambrosetti, A. Calderoni, A. Rossi, V. Vavassori, A. Conconi, L. Devizzi, F. Berger, M. Ponzoni, B. Borisch, M. Tinguely, M. Cerati, M. Milani, E. Orvieto, J. Sanchez, C. Chevreau, S. Dell'Oro, E. Zucca and F. Cavalli (2003). "Prognostic scoring system for primary CNS lymphomas: the International Extranodal Lymphoma Study Group experience." J Clin Oncol **21**(2): 266-272.

Ferry, J. A. (2006). "Burkitt's lymphoma: clinicopathologic features and differential diagnosis." Oncologist **11**(4): 375-383.

Filipowicz, W., S. N. Bhattacharyya and N. Sonenberg (2008). "Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?" Nat Rev Genet **9**(2): 102-114.

Flex, E., V. Petrangeli, L. Stella, S. Chiaretti, T. Hornakova, L. Knoops, C. Ariola, V. Fodale, E. Clappier, F. Paoloni, S. Martinelli, A. Fragale, M. Sanchez, S. Tavolaro, M. Messina, G. Cazzaniga, A. Camera, G. Pizzolo, A. Tornesello, M. Vignetti, A. Battistini, H. Cave, B. D. Gelb, J. C. Renauld, A. Biondi, S. N. Constantinescu, R.

Foa and M. Tartaglia (2008). "Somatically acquired JAK1 mutations in adult acute lymphoblastic leukemia." J Exp Med **205**(4): 751-758.

Fong, A. M., R. T. Premont, R. M. Richardson, Y. R. Yu, R. J. Lefkowitz and D. D. Patel (2002). "Defective lymphocyte chemotaxis in beta-arrestin2- and GRK6-deficient mice." Proc Natl Acad Sci U S A **99**(11): 7478-7483.

Fulci, V., S. Chiaretti, M. Goldoni, G. Azzalin, N. Carucci, S. Tavolaro, L. Castellano, A. Magrelli, F. Citarella, M. Messina, R. Maggio, N. Peragine, S. Santangelo, F. R. Mauro, P. Landgraf, T. Tuschl, D. B. Weir, M. Chien, J. J. Russo, J. Ju, R. Sheridan, C. Sander, M. Zavolan, A. Guarini, R. Foa and G. Macino (2007). "Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia." Blood **109**(11): 4944-4951.

Furqan, M., N. Mukhi, B. Lee and D. Liu (2013). "Dysregulation of JAK-STAT pathway in hematological malignancies and JAK inhibitors for clinical application." Biomark Res **1**(1): 5.

Gascoyne, R. D., S. A. Adomat, S. Krajewski, M. Krajewska, D. E. Horsman, A. W. Tolcher, S. E. O'Reilly, P. Hoskins, A. J. Coldman, J. C. Reed and J. M. Connors (1997). "Prognostic significance of Bcl-2 protein expression and Bcl-2 gene rearrangement in diffuse aggressive non-Hodgkin's lymphoma." Blood **90**(1): 244-251.

Gibcus, J. H., L. P. Tan, G. Harms, R. N. Schakel, D. de Jong, T. Blokzijl, P. Moller, S. Poppema, B. J. Kroesen and A. van den Berg (2009). "Hodgkin lymphoma cell lines are characterized by a specific miRNA expression profile." Neoplasia **11**(2): 167-176.

Golay, J., M. Luppi, S. Songia, C. Palvarini, L. Lombardi, A. Aiello, D. Delia, K. Lam, D. H. Crawford, A. Biondi, T. Barbui, A. Rambaldi and M. Introna (1996). "Expression of A-myb, but not c-myb and B-myb, is restricted to Burkitt's lymphoma, sIg+ B-acute lymphoblastic leukemia, and a subset of chronic lymphocytic leukemias." Blood **87**(5): 1900-1911.

Gong, J. Z., T. T. Stenzel, E. R. Bennett, A. S. Lagoo, C. H. Dunphy, J. O. Moore, D. A. Rizzieri, J. H. Tepperberg, P. Papenhausen and P. J. Buckley (2003). "Burkitt lymphoma arising in organ transplant recipients: a clinicopathologic study of five cases." Am J Surg Pathol **27**(6): 818-827.

Grimson, A., K. K. Farh, W. K. Johnston, P. Garrett-Engele, L. P. Lim and D. P. Bartel (2007). "MicroRNA targeting specificity in mammals: determinants beyond seed pairing." Mol Cell **27**(1): 91-105.

Guo, L. M., Y. Pu, Z. Han, T. Liu, Y. X. Li, M. Liu, X. Li and H. Tang (2009). "MicroRNA-9 inhibits ovarian cancer cell growth through regulation of NF-kappaB1." FEBS J **276**(19): 5537-5546.

Gupta, R. A., N. Shah, K. C. Wang, J. Kim, H. M. Horlings, D. J. Wong, M. C. Tsai, T. Hung, P. Argani, J. L. Rinn, Y. Wang, P. Brzoska, B. Kong, R. Li, R. B. West, M. J. van de Vijver, S. Sukumar and H. Y. Chang (2010). "Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis." Nature **464**(7291): 1071-1076.

Gutschner, T. and S. Diederichs (2012). "The hallmarks of cancer: a long non-coding RNA point of view." RNA Biol **9**(6): 703-719.

Guttman, M., I. Amit, M. Garber, C. French, M. F. Lin, D. Feldser, M. Huarte, O. Zuk, B. W. Carey, J. P. Cassady, M. N. Cabili, R. Jaenisch, T. S. Mikkelsen, T. Jacks, N. Hacohen, B. E. Bernstein, M. Kellis, A. Regev, J. L. Rinn and E. S. Lander (2009). "Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals." Nature **458**(7235): 223-227.

Habermann, T. M., E. A. Weller, V. A. Morrison, R. D. Gascoyne, P. A. Cassileth, J. B. Cohn, S. R. Dakhil, B. Woda, R. I. Fisher, B. A. Peterson and S. J. Horning (2006). "Rituximab-CHOP versus CHOP alone or with maintenance rituximab in older patients with diffuse large B-cell lymphoma." J Clin Oncol **24**(19): 3121-3127.

Hackenberg, M., M. Sturm, D. Langenberger, J. M. Falcon-Perez and A. M. Aransay (2009). "miRanalyzer: a microRNA detection and analysis tool for next-generation sequencing experiments." Nucleic Acids Res **37**(Web Server issue): W68-76.

Hafner, M., M. Landthaler, L. Burger, M. Khorshid, J. Hausser, P. Berninger, A. Rothballer, M. Ascano, Jr., A. C. Jungkamp, M. Munschauer, A. Ulrich, G. S. Wardle, S. Dewell, M. Zavolan and T. Tuschl (2010). "Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP." Cell **141**(1): 129-141.

Havelange, V., R. Garzon and C. M. Croce (2009). "MicroRNAs: new players in acute myeloid leukaemia." Br J Cancer **101**(5): 743-748.

Hoeller, S., M. P. Bihl, D. Zihler, S. Cogliatti, M. Ponzoni, A. Zettl, P. Went, A. Foerster, P. Hirschmann, A. Tzankov and S. Dirnhofer (2012). "Molecular and immunohistochemical characterization of B-cell lymphoma-2-negative follicular lymphomas." Hum Pathol **43**(3): 405-412.

Hu, Y., G. Wu, M. Rusch, L. Lukes, K. H. Buetow, J. Zhang and K. W. Hunter (2012). "Integrated cross-species transcriptional network analysis of metastatic susceptibility." Proc Natl Acad Sci U S A **109**(8): 3184-3189.

Huen, M. S., S. M. Sy and J. Chen (2010). "BRCA1 and its toolbox for the maintenance of genome integrity." Nat Rev Mol Cell Biol **11**(2): 138-148.

Hung, T., Y. Wang, M. F. Lin, A. K. Koegel, Y. Kotake, G. D. Grant, H. M. Horlings, N. Shah, C. Umbricht, P. Wang, Y. Wang, B. Kong, A. Langerod, A. L. Borresen-Dale, S. K. Kim, M. van de Vijver, S. Sukumar, M. L. Whitfield, M. Kellis, Y. Xiong, D. J. Wong and H. Y. Chang (2011). "Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters." Nat Genet **43**(7): 621-629.

Hutvagner, G. and M. J. Simard (2008). "Argonaute proteins: key players in RNA silencing." Nat Rev Mol Cell Biol **9**(1): 22-32.

Jacob, J., G. Kelsoe, K. Rajewsky and U. Weiss (1991). "Intraclonal generation of antibody mutants in germinal centres." Nature **354**(6352): 389-392.

Jaffe, E. S. and S. Pittaluga (2011). "Aggressive B-cell lymphomas: a review of new and old entities in the WHO classification." Hematology Am Soc Hematol Educ Program **2011**: 506-514.

Ji, H., L. I. Ehrlich, J. Seita, P. Murakami, A. Doi, P. Lindau, H. Lee, M. J. Aryee, R. A. Irizarry, K. Kim, D. J. Rossi, M. A. Inlay, T. Serwold, H. Karsunky, L. Ho, G. Q. Daley, I. L. Weissman and A. P. Feinberg (2010). "Comprehensive methylome map of lineage commitment from haematopoietic progenitors." Nature **467**(7313): 338-342.

Jost, P. J. and J. Ruland (2007). "Aberrant NF-kappaB signaling in lymphoma: mechanisms, consequences, and therapeutic implications." Blood **109**(7): 2700-2707.

Kalin, T. V., I. C. Wang, T. J. Ackerson, M. L. Major, C. J. Detrisac, V. V. Kalinichenko, A. Lyubimov and R. H. Costa (2006). "Increased levels of the FoxM1 transcription factor accelerate development and progression of prostate carcinomas in both TRAMP and LADY transgenic mice." Cancer Res **66**(3): 1712-1720.

Kelly, G. L., J. Stylianou, J. Rasaiyaah, W. Wei, W. Thomas, D. Croom-Carter, C. Kohler, R. Spang, C. Woodman, P. Kellam, A. B. Rickinson and A. I. Bell (2013). "Different patterns of Epstein-Barr virus latency in

endemic Burkitt lymphoma (BL) lead to distinct variants within the BL-associated gene expression signature." J Virol **87**(5): 2882-2894.

Kent, O. A. and J. T. Mendell (2006). "A small piece in the cancer puzzle: microRNAs as tumor suppressors and oncogenes." Oncogene **25**(46): 6188-6196.

Kirino, Y. and Z. Mourelatos (2008). "Site-specific crosslinking of human microRNPs to RNA targets." RNA **14**(10): 2254-2259.

Kluiver, J., E. Haralambieva, D. de Jong, T. Blokzijl, S. Jacobs, B. J. Kroesen, S. Poppema and A. van den Berg (2006). "Lack of BIC and microRNA miR-155 expression in primary cases of Burkitt lymphoma." Genes Chromosomes Cancer **45**(2): 147-153.

Kondrashov, A. V., M. Kiefmann, K. Ebnet, T. Khanam, R. S. Muddashetty and J. Brosius (2005). "Inhibitory effect of naked neural BC1 RNA or BC200 RNA on eukaryotic in vitro translation systems is reversed by poly(A)-binding protein (PABP)." J Mol Biol **353**(1): 88-103.

Kopp, K. L., U. Ralfkiaer, L. M. Gjerdrum, R. Helvad, I. H. Pedersen, T. Litman, L. Jonson, P. H. Hagedorn, T. Krejsgaard, R. Gniadecki, C. M. Bonefeld, L. Skov, C. Geisler, M. A. Wasik, E. Ralfkiaer, N. Odum and A. Woetmann (2013). "STAT5-mediated expression of oncogenic miR-155 in cutaneous T-cell lymphoma." Cell Cycle **12**(12): 1939-1947.

Kotian, S., T. Banerjee, A. Lockhart, K. Huang, U. V. Catalyurek and J. D. Parvin (2014). "NUSAP1 influences the DNA damage response by controlling BRCA1 protein levels." Cancer Biol Ther **15**(5): 533-543.

Kozomara, A. and S. Griffiths-Jones (2011). "miRBase: integrating microRNA annotation and deep-sequencing data." Nucleic Acids Res **39**(Database issue): D152-157.

Krutzfeldt, J., M. N. Poy and M. Stoffel (2006). "Strategies to determine the biological function of microRNAs." Nat Genet **38 Suppl**: S14-19.

Kwon, K., C. Hutter, Q. Sun, I. Bilic, C. Cobaleda, S. Malin and M. Busslinger (2008). "Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development." Immunity **28**(6): 751-762.

Lai, F., U. A. Orom, M. Cesaroni, M. Beringer, D. J. Taatjes, G. A. Blobel and R. Shiekhattar (2013). "Activating RNAs associate with Mediator to enhance chromatin architecture and transcription." Nature **494**(7438): 497-501.

Lai, M. C., Z. Yang, L. Zhou, Q. Q. Zhu, H. Y. Xie, F. Zhang, L. M. Wu, L. M. Chen and S. S. Zheng (2012). "Long non-coding RNA MALAT-1 overexpression predicts tumor recurrence of hepatocellular carcinoma after liver transplantation." Med Oncol **29**(3): 1810-1816.

Landgraf, P., M. Rusu, R. Sheridan, A. Sewer, N. Iovino, A. Aravin, S. Pfeffer, A. Rice, A. O. Kamphorst, M. Landthaler, C. Lin, N. D. Socci, L. Hermida, V. Fulci, S. Chiaretti, R. Foa, J. Schliwka, U. Fuchs, A. Novosel, R. U. Muller, B. Schermer, U. Bissels, J. Inman, Q. Phan, M. Chien, D. B. Weir, R. Choksi, G. De Vita, D. Frezzetti, H. I. Trompeter, V. Hornung, G. Teng, G. Hartmann, M. Palkovits, R. Di Lauro, P. Wernet, G. Macino, C. E. Rogler, J. W. Nagle, J. Ju, F. N. Papavasiliou, T. Benzing, P. Lichter, W. Tam, M. J. Brownstein, A. Bosio, A. Borkhardt, J. J. Russo, C. Sander, M. Zavolan and T. Tuschl (2007). "A mammalian microRNA expression atlas based on small RNA library sequencing." Cell **129**(7): 1401-1414.

Lawrie, C. H. (2007). "MicroRNA expression in lymphoma." Expert Opin Biol Ther **7**(9): 1363-1374.

Lawrie, C. H., J. Chi, S. Taylor, D. Tramonti, E. Ballabio, S. Palazzo, N. J. Saunders, F. Pezzella, J. Boultwood, J. S. Wainscoat and C. S. Hatton (2009). "Expression of microRNAs in diffuse large B cell lymphoma is

associated with immunophenotype, survival and transformation from follicular lymphoma." J Cell Mol Med **13**(7): 1248-1260.

Lee, R. C., R. L. Feinbaum and V. Ambros (1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell **75**(5): 843-854.

Lennox, K. A. and M. A. Behlke (2010). "A direct comparison of anti-microRNA oligonucleotide potency." Pharm Res **27**(9): 1788-1799.

Lenz, G., R. E. Davis, V. N. Ngo, L. Lam, T. C. George, G. W. Wright, S. S. Dave, H. Zhao, W. Xu, A. Rosenwald, G. Ott, H. K. Muller-Hermelink, R. D. Gascoyne, J. M. Connors, L. M. Rimsza, E. Campo, E. S. Jaffe, J. Delabie, E. B. Smeland, R. I. Fisher, W. C. Chan and L. M. Staudt (2008). "Oncogenic CARD11 mutations in human diffuse large B cell lymphoma." Science **319**(5870): 1676-1679.

Lenz, G., G. W. Wright, N. C. Emre, H. Kohlhammer, S. S. Dave, R. E. Davis, S. Carty, L. T. Lam, A. L. Shaffer, W. Xiao, J. Powell, A. Rosenwald, G. Ott, H. K. Muller-Hermelink, R. D. Gascoyne, J. M. Connors, E. Campo, E. S. Jaffe, J. Delabie, E. B. Smeland, L. M. Rimsza, R. I. Fisher, D. D. Weisenburger, W. C. Chan and L. M. Staudt (2008). "Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways." Proc Natl Acad Sci U S A **105**(36): 13520-13525.

Lenze, D., L. Leoncini, M. Hummel, S. Volinia, C. G. Liu, T. Amato, G. De Falco, J. Githanga, H. Horn, J. Nyagol, G. Ott, J. Palatini, M. Pfreundschuh, E. Rogena, A. Rosenwald, R. Siebert, C. M. Croce and H. Stein (2011). "The different epidemiologic subtypes of Burkitt lymphoma share a homogenous micro RNA profile distinct from diffuse large B-cell lymphoma." Leukemia **25**(12): 1869-1876.

Leucci, E., M. Cocco, A. Onnis, G. De Falco, P. van Cleef, C. Bellan, A. van Rijk, J. Nyagol, B. Byakika, S. Lazzi, P. Tosi, H. van Krieken and L. Leoncini (2008). "MYC translocation-negative classical Burkitt lymphoma cases: an alternative pathogenetic mechanism involving miRNA deregulation." J Pathol **216**(4): 440-450.

Levine, R. L., A. Pardanani, A. Tefferi and D. G. Gilliland (2007). "Role of JAK2 in the pathogenesis and therapy of myeloproliferative disorders." Nat Rev Cancer **7**(9): 673-683.

Lewis, B. P., C. B. Burge and D. P. Bartel (2005). "Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets." Cell **120**(1): 15-20.

Lewis, B. P., I. H. Shih, M. W. Jones-Rhoades, D. P. Bartel and C. B. Burge (2003). "Prediction of mammalian microRNA targets." Cell **115**(7): 787-798.

Li, C., S. W. Kim, D. Rai, A. R. Bolla, S. Adhvaryu, M. C. Kinney, R. S. Robetorye and R. C. Aguiar (2009). "Copy number abnormalities, MYC activity, and the genetic fingerprint of normal B cells mechanistically define the microRNA profile of diffuse large B-cell lymphoma." Blood **113**(26): 6681-6690.

Li, F. X., J. W. Zhu, C. J. Hogan and J. DeGregori (2003). "Defective gene expression, S phase progression, and maturation during hematopoiesis in E2F1/E2F2 mutant mice." Mol Cell Biol **23**(10): 3607-3622.

Lin, Y. C., S. Jhunjhunwala, C. Benner, S. Heinz, E. Welinder, R. Mansson, M. Sigvardsson, J. Hagman, C. A. Espinoza, J. Dutkowski, T. Ideker, C. K. Glass and C. Murre (2010). "A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate." Nat Immunol **11**(7): 635-643.

Liu, L., M. Zhang and P. Zou (2008). "Expression of PLK1 and survivin in non-Hodgkin's lymphoma treated with CHOP." Acta Pharmacol Sin **29**(3): 371-375.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." Methods **25**(4): 402-408.

Lo Coco, F., B. H. Ye, F. Lista, P. Corradini, K. Offit, D. M. Knowles, R. S. Chaganti and R. Dalla-Favera (1994). "Rearrangements of the BCL6 gene in diffuse large cell non-Hodgkin's lymphoma." Blood **83**(7): 1757-1759.

Lopez-Guillermo, A., L. Colomo, M. Jimenez, F. Bosch, N. Villamor, L. Arenillas, A. Muntanola, S. Montoto, E. Gine, D. Colomer, S. Bea, E. Campo and E. Montserrat (2005). "Diffuse large B-cell lymphoma: clinical and biological characterization and outcome according to the nodal or extranodal primary origin." J Clin Oncol **23**(12): 2797-2804.

Lossos, I. S., A. A. Alizadeh, M. B. Eisen, W. C. Chan, P. O. Brown, D. Botstein, L. M. Staudt and R. Levy (2000). "Ongoing immunoglobulin somatic mutation in germinal center B cell-like but not in activated B cell-like diffuse large cell lymphomas." Proc Natl Acad Sci U S A **97**(18): 10209-10213.

Lowery, D. M., D. Lim and M. B. Yaffe (2005). "Structure and function of Polo-like kinases." Oncogene **24**(2): 248-259.

Lowndes, N. F. and J. R. Murguia (2000). "Sensing and responding to DNA damage." Curr Opin Genet Dev **10**(1): 17-25.

Lu, J., G. Getz, E. A. Miska, E. Alvarez-Saavedra, J. Lamb, D. Peck, A. Sweet-Cordero, B. L. Ebert, R. H. Mak, A. A. Ferrando, J. R. Downing, T. Jacks, H. R. Horvitz and T. R. Golub (2005). "MicroRNA expression profiles classify human cancers." Nature **435**(7043): 834-838.

Lujambio, A., A. Portela, J. Liz, S. A. Melo, S. Rossi, R. Spizzo, C. M. Croce, G. A. Calin and M. Esteller (2010). "CpG island hypermethylation-associated silencing of non-coding RNAs transcribed from ultraconserved regions in human cancer." Oncogene **29**(48): 6390-6401.

Macintyre, E., D. Willerford and S. W. Morris (2000). "Non-Hodgkin's Lymphoma: Molecular Features of B Cell Lymphoma." Hematology Am Soc Hematol Educ Program: 180-204.

Malumbres, R., K. A. Sarosiek, E. Cubedo, J. W. Ruiz, X. Jiang, R. D. Gascoyne, R. Tibshirani and I. S. Lossos (2009). "Differentiation stage-specific expression of microRNAs in B lymphocytes and diffuse large B-cell lymphomas." Blood **113**(16): 3754-3764.

Marafioti, T., C. Copie-Bergman, M. Calaminici, J. C. Paterson, V. H. Shende, H. Liu, M. Baia, A. D. Ramsay, C. Agostinelli, J. Briere, A. Clear, M. Q. Du, P. P. Piccaluga, N. Masir, E. P. Nacheva, P. Sujobert, K. Shanmugam, T. M. Grogan, S. P. Brooks, A. Khwaja, K. Ardeshna, W. Townsend, S. A. Pileri, C. Haioun, D. Linch, J. G. Gribben, P. Gaulard and P. G. Isaacson (2013). "Another look at follicular lymphoma: immunophenotypic and molecular analyses identify distinct follicular lymphoma subgroups." Histopathology **62**(6): 860-875.

Maragkakis, M., P. Alexiou, G. L. Papadopoulos, M. Reczko, T. Dalamagas, G. Giannopoulos, G. Goumas, E. Koukis, K. Kourtis, V. A. Simossis, P. Sethupathy, T. Vergoulis, N. Koziris, T. Sellis, P. Tsanakas and A. G. Hatzigeorgiou (2009). "Accurate microRNA target prediction correlates with protein repression levels." BMC Bioinformatics **10**: 295.

Mattick, J. S., R. J. Taft and G. J. Faulkner (2010). "A global view of genomic information--moving beyond the gene and the master regulator." Trends Genet **26**(1): 21-28.

McClure, R. F., E. D. Remstein, W. R. Macon, G. W. Dewald, T. M. Habermann, A. Hoering and P. J. Kurtin (2005). "Adult B-cell lymphomas with burkitt-like morphology are phenotypically and genotypically heterogeneous with aggressive clinical behavior." Am J Surg Pathol **29**(12): 1652-1660.

Mead, G. M., M. R. Sydes, J. Walewski, A. Grigg, C. S. Hatton, N. Pescosta, C. Guarnaccia, M. S. Lewis, J. McKendrick, S. P. Stenning, D. Wright and U. L. collaborators (2002). "An international evaluation of CODOX-M and CODOX-M alternating with IVAC in adult Burkitt's lymphoma: results of United Kingdom Lymphoma Group LY06 study." Ann Oncol **13**(8): 1264-1274.

Meisenheimer, K. M. and T. H. Koch (1997). "Photocross-linking of nucleic acids to associated proteins." Crit Rev Biochem Mol Biol **32**(2): 101-140.

Melo, C. A., J. Drost, P. J. Wijchers, H. van de Werken, E. de Wit, J. A. Oude Vrielink, R. Elkon, S. A. Melo, N. Leveille, R. Kalluri, W. de Laat and R. Agami (2013). "eRNAs are required for p53-dependent enhancer activity and gene transcription." Mol Cell **49**(3): 524-535.

Mercer, T. R., M. E. Dinger and J. S. Mattick (2009). "Long non-coding RNAs: insights into functions." Nat Rev Genet **10**(3): 155-159.

Michael, M. Z., O. C. SM, N. G. van Holst Pellekaan, G. P. Young and R. J. James (2003). "Reduced accumulation of specific microRNAs in colorectal neoplasia." Mol Cancer Res **1**(12): 882-891.

Miki, Y., J. Swensen, D. Shattuck-Eidens, P. A. Futreal, K. Harshman, S. Tavtigian, Q. Liu, C. Cochran, L. M. Bennett, W. Ding and et al. (1994). "A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1." Science **266**(5182): 66-71.

Min, D. L., H. L. Xia, X. Y. Zhou, M. H. Sun, W. T. Yang, T. M. Zhang, A. H. Zheng and D. R. Shi (2005). "[Analysis of bcl-6 protein expression and gene rearrangement in diffuse large B-cell lymphoma]." Zhonghua Bing Li Xue Za Zhi **34**(6): 327-331.

Monsalvez, V., S. Montes-Moreno, M. J. Artiga, M. E. Rodriguez, B. Sanchez-Espiridion, M. Lozano, R. Fernandez-de-Misa, J. L. Rodriguez-Peralto, M. A. Piris and P. L. Ortiz-Romero (2013). "MicroRNAs as prognostic markers in indolent primary cutaneous B-cell lymphoma." Mod Pathol **26**(2): 171-181.

Montoto, S., A. J. Davies, J. Matthews, M. Calaminici, A. J. Norton, J. Amess, S. Vinnicombe, R. Waters, A. Z. Rohatiner and T. A. Lister (2007). "Risk and clinical implications of transformation of follicular lymphoma to diffuse large B-cell lymphoma." J Clin Oncol **25**(17): 2426-2433.

Muller, A. M., G. Ihorst, R. Mertelsmann and M. Engelhardt (2005). "Epidemiology of non-Hodgkin's lymphoma (NHL): trends, geographic distribution, and etiology." Ann Hematol **84**(1): 1-12.

Muller, H., A. P. Bracken, R. Vernell, M. C. Moroni, F. Christians, E. Grassilli, E. Prosperini, E. Vigo, J. D. Oliner and K. Helin (2001). "E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis." Genes Dev **15**(3): 267-285.

Negrini, M., M. S. Nicoloso and G. A. Calin (2009). "MicroRNAs and cancer--new paradigms in molecular oncology." Curr Opin Cell Biol **21**(3): 470-479.

Nie, L., H. J. Wu, J. M. Hsu, S. S. Chang, A. M. Labaff, C. W. Li, Y. Wang, J. L. Hsu and M. C. Hung (2012). "Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer." Am J Transl Res **4**(2): 127-150.

O'Brien, M. E., P. Easterbrook, J. Powell, G. R. Blackledge, L. Jones, I. C. MacLennan and R. C. Leonard (1991). "The natural history of low grade non-Hodgkin's lymphoma and the impact of a no initial treatment policy on survival." Q J Med **80**(292): 651-660.

O'Connell, R. M., A. A. Chaudhuri, D. S. Rao and D. Baltimore (2009). "Inositol phosphatase SHIP1 is a primary target of miR-155." Proc Natl Acad Sci U S A **106**(17): 7113-7118.

O'Donnell, K. A., E. A. Wentzel, K. I. Zeller, C. V. Dang and J. T. Mendell (2005). "c-Myc-regulated microRNAs modulate E2F1 expression." Nature **435**(7043): 839-843.

Okada, H., G. Kohanbash and M. T. Lotze (2010). "MicroRNAs in immune regulation--opportunities for cancer immunotherapy." Int J Biochem Cell Biol **42**(8): 1256-1261.

Okazaki, Y., M. Furuno, T. Kasukawa, J. Adachi, H. Bono, S. Kondo, I. Nikaido, N. Osato, R. Saito, H. Suzuki, I. Yamanaka, H. Kiyosawa, K. Yagi, Y. Tomaru, Y. Hasegawa, A. Nogami, C. Schonbach, T. Gojobori, R. Baldarelli, D. P. Hill, C. Bult, D. A. Hume, J. Quackenbush, L. M. Schriml, A. Kanapin, H. Matsuda, S. Batalov, K. W. Beisel, J. A. Blake, D. Bradt, V. Brusic, C. Chothia, L. E. Corbani, S. Cousins, E. Dalla, T. A. Dragani, C. F. Fletcher, A. Forrest, K. S. Frazer, T. Gaasterland, M. Gariboldi, C. Gissi, A. Godzik, J. Gough, S. Grimmond, S. Gustincich, N. Hirokawa, I. J. Jackson, E. D. Jarvis, A. Kanai, H. Kawaji, Y. Kawasawa, R. M. Kedzierski, B. L. King, A. Konagaya, I. V. Kurochkin, Y. Lee, B. Lenhard, P. A. Lyons, D. R. Maglott, L. Maltais, L. Marchionni, L. McKenzie, H. Miki, T. Nagashima, K. Numata, T. Okido, W. J. Pavan, G. Pertea, G. Pesole, N. Petrovsky, R. Pillai, J. U. Pontius, D. Qi, S. Ramachandran, T. Ravasi, J. C. Reed, D. J. Reed, J. Reid, B. Z. Ring, M. Ringwald, A. Sandelin, C. Schneider, C. A. Semple, M. Setou, K. Shimada, R. Sultana, Y. Takenaka, M. S. Taylor, R. D. Teasdale, M. Tomita, R. Verardo, L. Wagner, C. Wahlestedt, Y. Wang, Y. Watanabe, C. Wells, L. G. Wilming, A. Wynshaw-Boris, M. Yanagisawa, I. Yang, L. Yang, Z. Yuan, M. Zavolan, Y. Zhu, A. Zimmer, P. Carninci, N. Hayatsu, T. Hirozane-Kishikawa, H. Konno, M. Nakamura, N. Sakazume, K. Sato, T. Shiraki, K. Waki, J. Kawai, K. Aizawa, T. Arakawa, S. Fukuda, A. Hara, W. Hashizume, K. Imotani, Y. Ishii, M. Itoh, I. Kagawa, A. Miyazaki, K. Sakai, D. Sasaki, K. Shibata, A. Shinagawa, A. Yasunishi, M. Yoshino, R. Waterston, E. S. Lander, J. Rogers, E. Birney, Y. Hayashizaki, F. Consortium, R. G. E. R. G. P. I and I. I. Team (2002). "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs." Nature **420**(6915): 563-573.

Oken, M. M., R. H. Creech, D. C. Tormey, J. Horton, T. E. Davis, E. T. McFadden and P. P. Carbone (1982). "Toxicity and response criteria of the Eastern Cooperative Oncology Group." Am J Clin Oncol **5**(6): 649-655.

Olive, V., I. Jiang and L. He (2010). "mir-17-92, a cluster of miRNAs in the midst of the cancer network." Int J Biochem Cell Biol **42**(8): 1348-1354.

Osthus, R. C., B. Karim, J. E. Prescott, B. D. Smith, M. McDevitt, D. L. Huso and C. V. Dang (2005). "The Myc target gene JPO1/CDCA7 is frequently overexpressed in human tumors and has limited transforming activity in vivo." Cancer Res **65**(13): 5620-5627.

Ota, A., H. Tagawa, S. Karnan, S. Tsuzuki, A. Karpas, S. Kira, Y. Yoshida and M. Seto (2004). "Identification and characterization of a novel gene, C13orf25, as a target for 13q31-q32 amplification in malignant lymphoma." Cancer Res **64**(9): 3087-3095.

Otsuki, T., T. Yano, H. M. Clark, C. Bastard, J. P. Kerckaert, E. S. Jaffe and M. Raffeld (1995). "Analysis of LAZ3 (BCL-6) status in B-cell non-Hodgkin's lymphomas: results of rearrangement and gene expression studies and a mutational analysis of coding region sequences." Blood **85**(10): 2877-2884.

Ott, G., A. Rosenwald and E. Campo (2013). "Understanding MYC-driven aggressive B-cell lymphomas: pathogenesis and classification." Hematology Am Soc Hematol Educ Program **2013**: 575-583.

Ovcharenko, D., K. Kelnar, C. Johnson, N. Leng and D. Brown (2007). "Genome-scale microRNA and small interfering RNA screens identify small RNA modulators of TRAIL-induced apoptosis pathway." Cancer Res **67**(22): 10782-10788.

Pasqualucci, L., V. Trifonov, G. Fabbri, J. Ma, D. Rossi, A. Chiarenza, V. A. Wells, A. Grunn, M. Messina, O. Elliot, J. Chan, G. Bhagat, A. Chadburn, G. Gaidano, C. G. Mullighan, R. Rabadan and R. Dalla-Favera (2011). "Analysis of the coding genome of diffuse large B-cell lymphoma." Nat Genet **43**(9): 830-837.

Patel, M. N., M. D. Halling-Brown, J. E. Tym, P. Workman and B. Al-Lazikani (2013). "Objective assessment of cancer genes for drug discovery." Nat Rev Drug Discov **12**(1): 35-50.

Pavan, A., M. Spina, V. Canzonieri, S. Sansonno, G. Toffoli and V. De Re (2008). "Recent prognostic factors in diffuse large B-cell lymphoma indicate NF-kappaB pathway as a target for new therapeutic strategies." Leuk Lymphoma **49**(11): 2048-2058.

Petrocca, F., A. Vecchione and C. M. Croce (2008). "Emerging role of miR-106b-25/miR-17-92 clusters in the control of transforming growth factor beta signaling." Cancer Res **68**(20): 8191-8194.

Pileri, S. A., S. Ascani, E. Sabattini, G. Fraternali-Orcioni, S. Poggi, M. Piccioli, P. P. Piccaluga, B. Gamberi, P. L. Zinzani, L. Leoncini and B. Falini (2000). "The pathologist's view point. Part II --aggressive lymphomas." Haematologica **85**(12): 1308-1321.

Poliseno, L., L. Salmena, J. Zhang, B. Carver, W. J. Haveman and P. P. Pandolfi (2010). "A coding-independent function of gene and pseudogene mRNAs regulates tumour biology." Nature **465**(7301): 1033-1038.

Raemaekers, T., K. Ribbeck, J. Beaudouin, W. Annaert, M. Van Camp, I. Stockmans, N. Smets, R. Bouillon, J. Ellenberg and G. Carmeliet (2003). "NuSAP, a novel microtubule-associated protein involved in mitotic spindle organization." J Cell Biol **162**(6): 1017-1029.

Raphael, M., O. Gentilhomme, M. Tulliez, P. A. Byron and J. Diebold (1991). "Histopathologic features of high-grade non-Hodgkin's lymphomas in acquired immunodeficiency syndrome. The French Study Group of Pathology for Human Immunodeficiency Virus-Associated Tumors." Arch Pathol Lab Med **115**(1): 15-20.

Redon, S., P. Reichenbach and J. Lingner (2010). "The non-coding RNA TERRA is a natural ligand and direct inhibitor of human telomerase." Nucleic Acids Res **38**(17): 5797-5806.

Rempel, R. E., S. Mori, M. Gasparetto, M. A. Glozak, E. R. Andrechek, S. B. Adler, N. M. Laakso, A. S. Lagoo, R. Storms, C. Smith and J. R. Nevins (2009). "A role for E2F activities in determining the fate of Myc-induced lymphomagenesis." PLoS Genet **5**(9): e1000640.

Robertus, J. L., J. Kluiver, C. Weggemans, G. Harms, R. M. Reijmers, Y. Swart, K. Kok, S. Rosati, E. Schuuring, G. van Imhoff, S. T. Pals, P. Kluin and A. van den Berg (2010). "MiRNA profiling in B non-Hodgkin lymphoma: a MYC-related miRNA profile characterizes Burkitt lymphoma." Br J Haematol **149**(6): 896-899.

Rodriguez, A., E. Vigorito, S. Clare, M. V. Warren, P. Couttet, D. R. Soond, S. van Dongen, R. J. Grocock, P. P. Das, E. A. Miska, D. Vetrie, K. Okkenhaug, A. J. Enright, G. Dougan, M. Turner and A. Bradley (2007). "Requirement of bic/microRNA-155 for normal immune function." Science **316**(5824): 608-611.

Roehle, A., K. P. Hoefig, D. Repsilber, C. Thorns, M. Ziepert, K. O. Wesche, M. Thiere, M. Loeffler, W. Klapper, M. Pfreundschuh, A. Matolcsy, H. W. Bernd, L. Reiniger, H. Merz and A. C. Feller (2008). "MicroRNA signatures characterize diffuse large B-cell lymphomas and follicular lymphomas." Br J Haematol **142**(5): 732-744.

Rowley, J. D. (1988). "Chromosome studies in the non-Hodgkin's lymphomas: the role of the 14;18 translocation." J Clin Oncol **6**(5): 919-925.

Said, J. W. (2013). "Aggressive B-cell lymphomas: how many categories do we need?" Mod Pathol **26 Suppl 1**: S42-56.

Salles, G. A. (2007). "Clinical features, prognosis and treatment of follicular lymphoma." Hematology Am Soc Hematol Educ Program: 216-225.

Sampson, V. B., N. H. Rong, J. Han, Q. Yang, V. Aris, P. Soteropoulos, N. J. Petrelli, S. P. Dunn and L. J. Krueger (2007). "MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells." Cancer Res **67**(20): 9762-9770.

Satow, R., M. Shitashige, Y. Kanai, F. Takeshita, H. Ojima, T. Jigami, K. Honda, T. Kosuge, T. Ochiya, S. Hirohashi and T. Yamada (2010). "Combined functional genome survey of therapeutic targets for hepatocellular carcinoma." Clin Cancer Res **16**(9): 2518-2528.

Schmidt, L. H., T. Spieker, S. Koschmieder, S. Schaffers, J. Humberg, D. Jungen, E. Bulk, A. Hascher, D. Wittmer, A. Marra, L. Hillejan, K. Wiebe, W. E. Berdel, R. Wiewrodt and C. Muller-Tidow (2011). "The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth." J Thorac Oncol **6**(12): 1984-1992.

Schneider, C., L. Pasqualucci and R. Dalla-Favera (2011). "Molecular pathogenesis of diffuse large B-cell lymphoma." Semin Diagn Pathol **28**(2): 167-177.

Sheehan, W. W. and H. Rappaport (1970). "Morphological criteria in the classification of the malignant lymphomas." Proc Natl Cancer Conf **6**: 59-71.

Shiloh, Y. (2001). "ATM and ATR: networking cellular responses to DNA damage." Curr Opin Genet Dev **11**(1): 71-77.

Siddiqui, M., K. Ristow, S. N. Markovic, T. E. Witzig, T. M. Habermann, J. P. Colgan, D. J. Inwards, W. L. White, S. M. Ansell, I. N. Micallef, P. B. Johnston, T. G. Call and L. F. Porrata (2006). "Absolute lymphocyte count predicts overall survival in follicular lymphomas." Br J Haematol **134**(6): 596-601.

Spitzer, J., M. Hafner, M. Landthaler, M. Ascano, T. Farazi, G. Wardle, J. Nusbaum, M. Khorshid, L. Burger, M. Zavolan and T. Tuschl (2014). "PAR-CLIP (Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins." Methods Enzymol **539**: 113-161.

Stadtfeld, M., E. Apostolou, H. Akutsu, A. Fukuda, P. Follett, S. Natesan, T. Kono, T. Shioda and K. Hochedlinger (2010). "Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells." Nature **465**(7295): 175-181.

Staudt, L. M. (2010). "Oncogenic activation of NF-kappaB." Cold Spring Harb Perspect Biol **2**(6): a000109.

Suzuki, H. I., M. Arase, H. Matsuyama, Y. L. Choi, T. Ueno, H. Mano, K. Sugimoto and K. Miyazono (2011). "MCPIP1 ribonuclease antagonizes dicer and terminates microRNA biogenesis through precursor microRNA degradation." Mol Cell **44**(3): 424-436.

Swerdlow, S. H., E. Campo, N. L. Harris, E. S. Jaffe, S. A. Pileri, H. Stein, J. Thiele and J. W. Vardiman (2008). WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, Fourth Edition.

Szczyrba, J., E. Loprich, S. Wach, V. Jung, G. Unteregger, S. Barth, R. Grobholz, W. Wieland, R. Stohr, A. Hartmann, B. Wullich and F. Grasser (2010). "The microRNA profile of prostate carcinoma obtained by deep sequencing." Mol Cancer Res **8**(4): 529-538.

Tay, Y., J. Zhang, A. M. Thomson, B. Lim and I. Rigoutsos (2008). "MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation." Nature **455**(7216): 1124-1128.

Thai, T. H., D. P. Calado, S. Casola, K. M. Ansel, C. Xiao, Y. Xue, A. Murphy, D. Frendewey, D. Valenzuela, J. L. Kutok, M. Schmidt-Supprian, N. Rajewsky, G. Yancopoulos, A. Rao and K. Rajewsky (2007). "Regulation of the germinal center response by microRNA-155." Science **316**(5824): 604-608.

Thomas, D. A., J. Cortes, S. O'Brien, S. Pierce, S. Faderl, M. Albitar, F. B. Hagemeister, F. F. Cabanillas, S. Murphy, M. J. Keating and H. Kantarjian (1999). "Hyper-CVAD program in Burkitt's-type adult acute lymphoblastic leukemia." J Clin Oncol **17**(8): 2461-2470.

Trapnell, C., L. Pachter and S. L. Salzberg (2009). "TopHat: discovering splice junctions with RNA-Seq." Bioinformatics **25**(9): 1105-1111.

Tripathi, V., J. D. Ellis, Z. Shen, D. Y. Song, Q. Pan, A. T. Watt, S. M. Freier, C. F. Bennett, A. Sharma, P. A. Bubulya, B. J. Blencowe, S. G. Prasanth and K. V. Prasanth (2010). "The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation." Mol Cell **39**(6): 925-938.

Tsutsumi, H., T. Tanaka, N. Ohashi, H. Masuno, H. Tamamura, K. Hiramatsu, T. Araki, S. Ueda, S. Oishi and N. Fujii (2007). "Therapeutic potential of the chemokine receptor CXCR4 antagonists as multifunctional agents." Biopolymers **88**(2): 279-289.

Vainchenker, W. and S. N. Constantinescu (2013). "JAK/STAT signaling in hematological malignancies." Oncogene **32**(21): 2601-2613.

Vanden Bosch, A., T. Raemaekers, S. Denayer, S. Torrekens, N. Smets, K. Moermans, M. Dewerchin, P. Carmeliet and G. Carmeliet (2010). "NuSAP is essential for chromatin-induced spindle formation during early embryogenesis." J Cell Sci **123**(Pt 19): 3244-3255.

Vardiman, J. W., J. Thiele, D. A. Arber, R. D. Brunning, M. J. Borowitz, A. Porwit, N. L. Harris, M. M. Le Beau, E. Hellstrom-Lindberg, A. Tefferi and C. D. Bloomfield (2009). "The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes." Blood **114**(5): 937-951.

Wahlin, B. E., O. E. Yri, E. Kimby, H. Holte, J. Delabie, E. B. Smeland, C. Sundstrom, B. Christensson and B. Sander (2012). "Clinical significance of the WHO grades of follicular lymphoma in a population-based cohort of 505 patients with long follow-up times." Br J Haematol **156**(2): 225-233.

Wang, I. C., Y. J. Chen, D. Hughes, V. Petrovic, M. L. Major, H. J. Park, Y. Tan, T. Ackerson and R. H. Costa (2005). "Forkhead box M1 regulates the transcriptional network of genes essential for mitotic progression and genes encoding the SCF (Skp2-Cks1) ubiquitin ligase." Mol Cell Biol **25**(24): 10875-10894.

Wang, K. C. and H. Y. Chang (2011). "Molecular mechanisms of long noncoding RNAs." Mol Cell **43**(6): 904-914.

Wang, M., W. Yang, M. Li and Y. Li (2014). "Low expression of miR-150 in pediatric intestinal Burkitt lymphoma." Exp Mol Pathol **96**(2): 261-266.

Wang, W., M. Corrigan-Cummins, J. Hudson, I. Maric, O. Simakova, S. S. Neelapu, L. W. Kwak, J. E. Janik, B. Gause, E. S. Jaffe and K. R. Calvo (2012). "MicroRNA profiling of follicular lymphoma identifies microRNAs related to cell proliferation and tumor response." Haematologica **97**(4): 586-594.

Wang, Z., H. J. Park, J. R. Carr, Y. J. Chen, Y. Zheng, J. Li, A. L. Tyner, R. H. Costa, S. Bagchi and P. Raychaudhuri (2011). "FoxM1 in tumorigenicity of the neuroblastoma cells and renewal of the neural progenitors." Cancer Res **71**(12): 4292-4302.

Wark, A. W., H. J. Lee and R. M. Corn (2008). "Multiplexed detection methods for profiling microRNA expression in biological samples." Angew Chem Int Ed Engl **47**(4): 644-652.

Weiss, L. M., R. A. Warnke, J. Sklar and M. L. Cleary (1987). "Molecular analysis of the t(14;18) chromosomal translocation in malignant lymphomas." N Engl J Med **317**(19): 1185-1189.

Xiao, C., D. P. Calado, G. Galler, T. H. Thai, H. C. Patterson, J. Wang, N. Rajewsky, T. P. Bender and K. Rajewsky (2007). "MiR-150 controls B cell differentiation by targeting the transcription factor c-Myb." Cell **131**(1): 146-159.

Xicoy, B., J. M. Ribera, J. Esteve, S. Brunet, M. A. Sanz, P. Fernandez-Abellan and E. Feliu (2003). "Post-transplant Burkitt's leukemia or lymphoma. Study of five cases treated with specific intensive therapy (PETHEMA ALL-3/97 trial)." Leuk Lymphoma **44**(9): 1541-1543.

Xu, C., M. Yang, J. Tian, X. Wang and Z. Li (2011). "MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis." Int J Oncol **39**(1): 169-175.

Yang, Q. Q. and Y. F. Deng (2014). "Long non-coding RNAs as novel biomarkers and therapeutic targets in head and neck cancers." Int J Clin Exp Pathol **7**(4): 1286-1292.

Ying, L., Q. Chen, Y. Wang, Z. Zhou, Y. Huang and F. Qiu (2012). "Upregulated MALAT-1 contributes to bladder cancer cell migration by inducing epithelial-to-mesenchymal transition." Mol Biosyst **8**(9): 2289-2294.

Yokouchi, M., R. Suzuki, M. Masuhara, S. Komiya, A. Inoue and A. Yoshimura (1997). "Cloning and characterization of APS, an adaptor molecule containing PH and SH2 domains that is tyrosine phosphorylated upon B-cell receptor stimulation." Oncogene **15**(1): 7-15.

You, Y., B. G. Moreira, M. A. Behlke and R. Owczarzy (2006). "Design of LNA probes that improve mismatch discrimination." Nucleic Acids Res **34**(8): e60.

Zhang, J., R. Chiodini, A. Badr and G. Zhang (2011). "The impact of next-generation sequencing on genomics." J Genet Genomics **38**(3): 95-109.

Zhang, J., D. D. Jima, C. Jacobs, R. Fischer, E. Gottwein, G. Huang, P. L. Lugar, A. S. Lagoo, D. A. Rizzieri, D. R. Friedman, J. B. Weinberg, P. E. Lipsky and S. S. Dave (2009). "Patterns of microRNA expression characterize stages of human B-cell differentiation." Blood **113**(19): 4586-4594.

Zhang, L., J. Huang, N. Yang, J. Greshock, M. S. Megraw, A. Giannakakis, S. Liang, T. L. Naylor, A. Barchetti, M. R. Ward, G. Yao, A. Medina, A. O'Brien-Jenkins, D. Katsaros, A. Hatzigeorgiou, P. A. Gimotty, B. L. Weber and G. Coukos (2006). "microRNAs exhibit high frequency genomic alterations in human cancer." Proc Natl Acad Sci U S A **103**(24): 9136-9141.

Zhang, T., K. Nie and W. Tam (2008). "BIC is processed efficiently to microRNA-155 in Burkitt lymphoma cells." Leukemia **22**(9): 1795-1797.

Zhang, X., S. Liu, T. Hu, S. Liu, Y. He and S. Sun (2009). "Up-regulated microRNA-143 transcribed by nuclear factor kappa B enhances hepatocarcinoma metastasis by repressing fibronectin expression." Hepatology **50**(2): 490-499.

Zhong, S., S. Zhang, E. Bair, S. Nares and A. A. Khan (2012). "Differential expression of microRNAs in normal and inflamed human pulps." J Endod **38**(6): 746-752.

Zhou, B., Y. He, X. Zhang, J. Xu, Y. Luo, Y. Wang, S. G. Franzblau, Z. Yang, R. J. Chan, Y. Liu, J. Zheng and Z. Y. Zhang (2010). "Targeting mycobacterium protein tyrosine phosphatase B for antituberculosis agents." Proc Natl Acad Sci U S A **107**(10): 4573-4578.

Abbreviations

4SU	4-thiouridine
ABC	Activated B-cell like
ADAR	Adenosine deaminase acting on RNA
AGO	Argonaute protein
AID	Activation-induced cytidine deaminize
AML	Acute myeloid leukemia
ATM	Ataxia telangiectasia mutated
BL	Burkitt lymphoma
BMP	Bone morphogenetic protein
bp	base pair
BRCA1	Breast cancer-associated gene 1
BSA	bovine serum albumin
CANT1	Calcium activated nucleotidase 1
CDCA7	Cell division cycle-associated protein
CGI	CpG islands
CHOP	Cyclophosphamide, doxorubicin, vincristine and prednisolone
CLIP	Crosslinking immunoprecipitation
CLL	Chronic lymphocytic leukemia
CNV	Copy number variation
Ct	Cycle threshold
DLBCL	Diffuse large B-cell lymphoma
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsRNA	Double-stranded-RNA
eBL	Associated endemic subtype
EBV	Epstein Barr virus
ECOG	Eastern Cooperative Oncology Group
FL	Follicular lymphoma
FOXM1	Forkhead box transcriptional factor
GC	Germinal center
GCB	Germinal center B-cell like
HT1	Hybridization Buffer
ICGC	International Cancer Genome Consortium
IP	Immunoprecipitation
JAK/STAT	Janus associated kinase-signal transducer and activator of transcription
KD	Knockdown
kDa	Kilodalton
LDH	Lactate dehydrogenase
LncRNA	Long noncoding RNA
MCL	Mantle cell lymphoma
MCM2	Maintenance complex component 2
miRNA	microRNA
miRNA*	microRNA star

MREs	miRNA recognition elements
mRNA	Messenger ribonucleic acid
NATs	natural antisense transcripts
ncRNA	non-coding RNA
NF-KB	Nuclear factor kB
NGS	Next-generation sequencing
NHL	Non-Hodgkin lymphoma
nt	Nucleotide
OE	Overexpression
P:C:I	phenol:chloroform:isoamyl alcohol
PAR-CLIP	Photoactivatable-Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation
PCNT	Pericentrin
PCR	Polymerase chain reaction
piRNA	piwi-interacting RNA
PLK1	Polo-like kinase 1
PNK	Polynucleotide kinase
pri-miRNA	primary miRNA
qRT-PCR	Quantitative reverse transcription-real time-PCR
RBP	RNA binding protein
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNAP II	RNA polymerase II
sBL	Sporadic BL
SDS	Sodium dodecyl sulfate
SHIP-1	SH2-domain containing inositol-5-phosphatase 1
SHM	Somatic hypermutation
shRNA	Short-hairpin RNA interference
siRNA	Small interfering RNA
snoRNA	Small nucleolar RNA
snRNA	Small nuclear RNA
STP	Stop solution
TIAM1	T-cell lymphoma invasion and metastasis 1
UTR	Untranslated region
UV	Ultraviolet
WHO	World Health Organization
WP	Work package

Supplement

1. **Supplemental Table S. 1 Lymphoma patient characteristics**
2. **Supplemental Table S. 2 MiRNA expression differences between GCB lymphoma subtypes**
3. **Supplemental Table S.3 Novel miRNAs**
4. **Supplemental Table S. 4 RNA Seq-data of novel miRNAs**
5. **Supplemental Table S. 5 AGO2-PAR-CLIP**
6. **Supplemental Table S. 6 RNAseq analysis for AP000251**

Acknowledgements

On this occasion, I wish to thank all persons involved in this project for the great support during my PhD thesis. I would like to, first and foremost, thank my supervisor Dr. Jessica Höll for providing me the great opportunity for an exciting research; for always finding the time to discuss my questions and for unconditionally supporting me in my professional interests. A special thanks goes to Prof. Dr. Arndt Borkhardt for giving me the opportunity to successfully complete my PhD thesis in the Department of Pediatric Oncology, Hematology and Immunology, Düsseldorf and jointly at the Institute of Pathology, Berlin. In this regard, I would also like to thank Prof. Dr. Michael Hummel for the evaluation of my thesis as first professor and the great thought-provoking impulses. Furthermore, I would like to thank all of my colleagues for their dedicated teamwork and the ever existing atmosphere of genuine warmth and friendship. Last but never least; I would like to thank my amazing family for the love, support, and constant encouragement that I have gotten over the years. In particular, I would like to thank my mother, all the support she has provided me over the years was the greatest gift and I would like to address a special word of thanks to my sister Samira for loving and supporting me.

Curriculum vitae

"My curriculum vita does not appear in the electronic version of my paper for reasons of data protection."

Affidavit

"I, Kebria Hezaveh certify under penalty of perjury by my own signature that I have submitted the thesis on the topic **"Non-coding RNA expression profiling of germinal-center derived B-cell malignant (non-Hodgkin) lymphomas (GCB-lymphomas)".** I wrote this thesis independently and without assistance from third parties, I used no other aids than the listed sources and resources.

All points based literally or in spirit on publications or presentations of other authors are, as such, in proper citations (see "uniform requirements for manuscripts (URM)" the ICMJE www.icmje.org) indicated. The sections on methodology (in particular practical work, laboratory requirements, statistical processing) and results (in particular images, graphics and tables) correspond to the URM (s.o) and are answered by me. My interests in any publications to this dissertation correspond to those that are specified in the following joint declaration with the responsible person and supervisor. All publications resulting from this thesis and which I am author correspond to the URM (see above) and I am solely responsible.

The importance of this affidavit and the criminal consequences of a false affidavit (section 156,161 of the Criminal Code) are known to me and I understand the rights and responsibilities stated therein.

Kebria Hezaveh