From THE DEPARTMENT OF LABORATORY MEDICINE
Karolinska Institutet, Stockholm, Sweden

GENETICS OF DIFFUSE LARGE B-CELL LYMPHOMA

Konstantinos Georgiou

Karolinska Institutet

Stockholm 2015

Published by Karolinska Institutet.
Printed by Eprint AB, Stockholm, Sweden

ISBN 978-91-7676-099-4

Genetics of diffuse large B-cell lymphoma

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Konstantinos Georgiou

Principal Supervisor:
Professor Qiang Pan-Hammarström
Karolinska Institutet
Department of Laboratory Medicine

Co-supervisors:
Professor Klas Wiman
Karolinska Institutet
Department of Oncology-Pathology

Dr. Noel Filipe da Cunha Carvalho de Miranda
Leiden University Medical Center
Department of Pathology

Opponent:
Dr. Harald Holte
Oslo University Hospital
Department of Oncology

Examination Board:
Professor Magnus Björkholm
Karolinska Institutet
Department of Medicine-Solna

Professor Sam Okret
Karolinska Institutet
Department of Biosciences and Nutrition

Associate Professor Tobias Sjöblom
Uppsala University
Department of Immunology

Αφιερωμένο στους γονείς μου, Γιώτα και Τάσο,
με αγάπη και ευγνωμοσύνη!

ABSTRACT

Diffuse large B-cell lymphoma (DLBCL) is one the most common forms of non-Hodgkin lymphoma and one of the most aggressive B-cell neoplasms. Although most patients respond to current standard treatments, a significant number of them relapse and become refractory to treatment. Hence, there is a need for new approaches in the management of DLBCL. In recent years, a number of studies using next-generation sequencing (NGS) have contributed to the characterization of the disease and have revealed a set of deregulated cellular mechanisms in DLBCL. However, DLBCL is a very heterogeneous disease and additional work is required in order to achieve a comprehensive understanding of the mechanisms underlying its lymphomagenesis.

The work described in this thesis aimed to further characterize the genome of DLBCL and identify novel genomic alterations by employing a number of NGS methods. Furthermore, this thesis sought to functionally describe the effect of some of these alterations and explore how they could be meaningful from a therapeutic perspective.

We first embarked on a targeted sequencing approach in order to assess the impact of mutations in a set of key DNA repair genes. DNA repair is critical in B-cell development as it ensures the regulation of antibody diversification processes. The results showed that half of the tumors analyzed carry mutations in DNA repair genes and the most frequently targeted mechanisms are mismatch repair, DNA damage response, homologous recombination and non-homologous end-joining. Moreover, functional analysis enabled the association of a number of those mutations to specific phenotypes.

Whole exome sequencing identified a number or previously unreported somatic mutation targets in DLBCL. In addition, it revealed that some genes known to be involved in DLBCL were mutated at a much higher frequency in Chinese patients as compared to patients from Western populations. The impact of mutations in *DTX1*, a gene encoding for a negative regulator of Notch signaling, was functionally assessed and a number of mutations were found to be deleterious.

Finally, whole genome sequencing enabled the identification of structural variations such as translocations in the genome of DLBCL. *IGH* translocations are a hallmark of DLBCL and the locus is often found juxtaposed to proto-oncogenes which leads to their upregulation. We identified *PD-L1*, a common culprit of immune suppression in cancer, as a novel *IGH* translocation partner in DLBCL. Translocations involving the *PD-L1* locus were found to impact patient survival and were associated with the more aggressive non-germinal center-like subtype of the disease.

LIST OF SCIENTIFIC PAPERS

I. de Miranda NF*, Peng R*, **Georgiou K**, Wu C, Falk Sörqvist E, Berglund M, Chen L, Gao Z, Lagerstedt K, Lisboa S, Roos F, van Wezel T, Teixeira MR, Rosenquist R, Sundström C, Enblad G, Nilsson M, Zeng Y, Kipling D, Pan-Hammarström Q.
*Equal contribution
DNA repair genes are selectively mutated in diffuse large B-cell lymphomas.
J Exp Med. 2013 Aug 26;210(9):1729-42.

II. de Miranda NF*, **Georgiou K***, Chen L*, Wu C, Gao Z, Zaravinos A, Lisboa S, Enblad G, Teixeira MR, Zeng Y, Peng R, Pan-Hammarström Q.
*Equal contribution
Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients.
Blood. 2014 Oct 16;124(16):2544-53.

III. **Georgiou K**, Chen L, Berglund M, de Miranda NF, Lisboa S, Fangazio M, Ren W, Zhu S, Hou Y, Wu K, Fang W, Wang X, Zhang H, Zhang L, Zeng Y, Bhagat G, Nordenskjöld M, Sundström C, Dalla-Favera R, Enblad G, Teixeira MR, Pasqualucci L, Peng R, Pan-Hammarström Q.
Genetic basis of *PD-L1* overexpression in diffuse large B-cell lymphomas.
Submitted for publication.

CONTENTS

LIST OF ABBREVIATIONS

ABC	Activated B-cell-like
A-EJ	Alternative end-joining
AID	Activation–induced cytidine deaminase
APC	Antigen presenting cell
BCR	B-cell receptor
BER	Base-excision repair
BL	Burkitt's lymphoma
BTK	Bruton's tyrosine kinase
BWA	Burrows Wheeler Aligner
C	Constant
CSR	Class-Switch recombination
D	Diversify segment
DDR	DNA damage response
DLBCL	Diffuse large B-cell lymphoma
DSB	DNA double strand break
EBV	Epstein-Barr virus
FISH	Fluorescence *in situ* hybridization
FL	Follicular lymphoma
GCB	Germinal center B-cell-like
HIV	Human immunodeficiency virus
HLA	Human leucocyte antigen
HR	Homologous recombination
IGH	Immunoglobulin heavy locus
IGV	Interactive genomic viewer
Indel	Insertions and deletions
J	Joining
MAF	Minor allele frequency
mDTX1	Mutated DTX1

MMR	Mismatch repair
MRN	The MRE11, RAD50, NBS1 complex
MSI	Microsatellite instability
MSS	Microsatellite stable
NCID1	NOTCH1 intracellular domain
NER	Nucleotide excision repair
NGS	Next-generation sequencing
NHEJ	Non-homologous end-joining
NHL	Non-Hodgkin lymphomas
NK	Natural killer cell
PCR	Polymerase chain-reaction
qPCR	Quantitative real-time polymerase chain-reaction
S	Switch
SHM	Somatic hypermutation
SNP	Single nucleotide polymorphism
SNV	Single nucleotide variation
STR	Short tandem repeat
STS	Sequence tagged site
SV	Structural variation
TCR	T-cell receptor
V	Variable
WES	Whole-exome sequencing
WGS	Whole-genome sequencing

1 INTRODUCTION

1.1 DNA REPAIR

1.1.1 DNA and mutagenic factors

The genome constitutes the genetic heritage and blueprint of an individual. In a complex being such as a human, the genomic sequence defining it is the result of a lengthy and selective evolutionary process. In order to maintain a well-functioning body and assure the perpetuation of the species, the genetic heritage contained within the DNA sequence must be guarded and preserved.

The approximately $3.2*10^9$ base-pairs that form the human genome in each nucleated cell are constantly exposed to mutational factors that can alter the DNA sequence. The nature of these mutation-inducing factors is variable, as they can be environmental such as chemical substances, ionizing radiation and infections, or intrinsic such as reactive oxygen species and errors occurring during the replication of DNA. The resulting alterations in the DNA can be nucleotide substitutions and DNA single or DNA double stand breaks (DSBs), leading to single nucleotide variations (SNVs), insertions/deletions (indels) and structural variations (SVs) (1, 2).

To counteract this, human cells possess a sophisticated machinery composed of an array of pathways and effector molecules that are collectively referred to as DNA damage response and repair factors. These constantly surveil the genome and ensure its repair when it is damaged by mutations or other forms of genetic alterations. Depending on their function, they are organized into mechanisms ensuring different but complementary and often overlapping tasks (3-5).

1.1.2 DNA damage response

The DNA damage response (DDR) mechanism is the first line of defense against DNA damage and acts as a surveillance mechanism. Upon the detection of lesions in the genome, it activates cell-cycle control checkpoints and triggers an appropriate DNA repair response (4, 6). This results in cell-cycle arrest, repair of the lesion or apoptosis of the cell and therefore ensure that the lesions are repaired; and even more importantly that they are not propagated to the next generation of daughter-cells (Figure 1) (6, 7). Notable sensor and effector molecules of the DDR pathway are p53, CHK1, CHK2, PARP1, DNA-PKcs (PRKDC), ATM and ATR (6, 8).

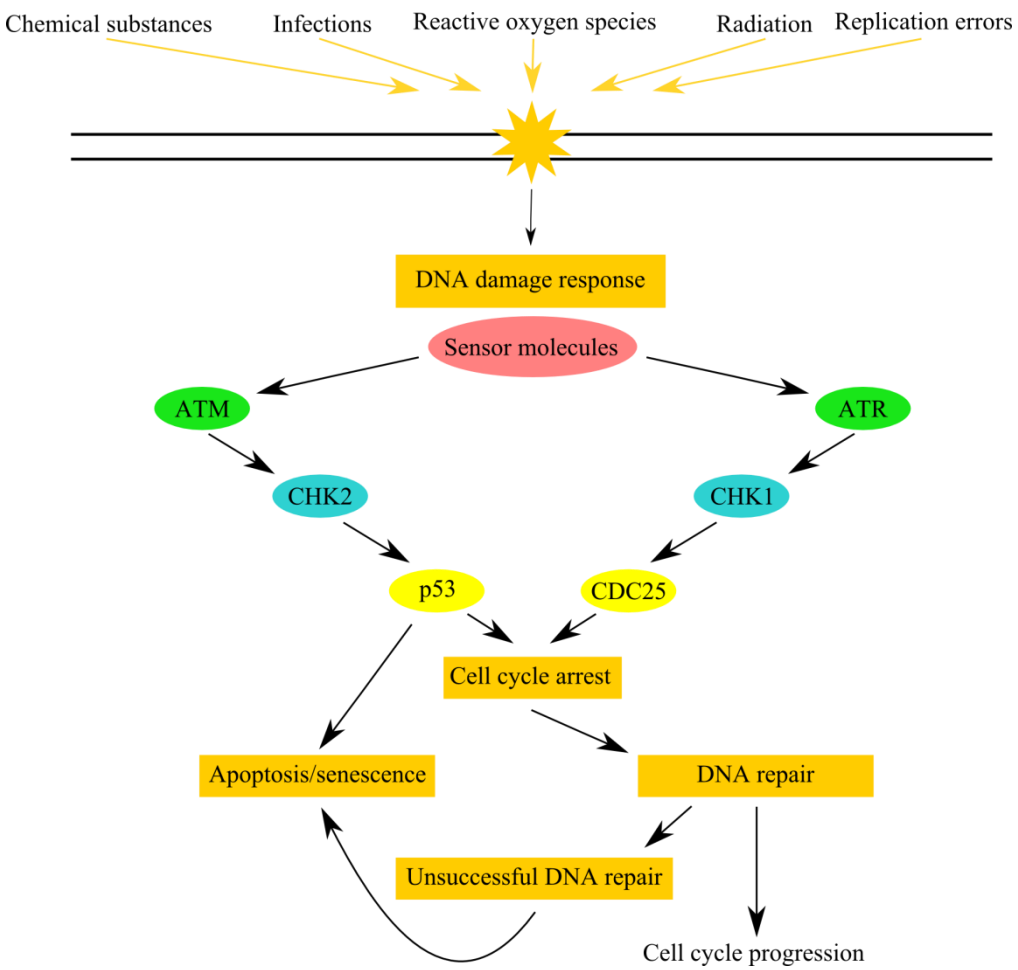

Figure 1: DNA damage response. Intrinsic and extrinsic factors may alter the DNA structure. In response to this damage, protective mechanisms are activated that lead to cell cycle arrest, apoptosis or repair of the damage. Apoptosis can be induced directly or subsequently to unsuccessful DNA repair.

1.1.3 DNA double strand breaks

Among the most dangerous forms of DNA damage are DSBs. When not properly repaired, they cause a cell to undergo apoptosis, a form of programmed cell-death. Moreover, unrepaired DSBs can lead to genomic instability (9). DSBs are resolved by homologous recombination (HR) or via the non-homologous end-joining (NHEJ) and alternative end-joining (A-EJ) pathways (10, 11). ATM, a Ser/Thr kinase plays an important role in the repair of DSBs by activating many of the factors involved (12).

1.1.3.1 Homologous recombination

HR, as the name suggests, uses sequence homology to repair DSB and for this to take place a template is required. BRCA1 accumulates at DSB sites and participates in the

choice of the HR pathway for the repair of a DSB (13). HR continues with the formation of the MRE11, RAD50 and NBS1 (MRN) complex at the site of the break. The exonuclease activity of EXO1 ensures the resection of nucleotides situated at the break. The resulting exposed single strands are coated by RPA in order to minimize damage and prevent the formation of secondary structures. BRCA2 promotes the assembly of RAD51/RAD52 to the single strands. RAD51 and RAD52 ensure the association of the damaged DNA to the sister chromatid in a process called strand-invasion. This leads to the formation of characteristic structures known as Holliday junctions. The sister chromatid then acts as a template for the extension and repair of the damaged strand (Figure 2) (14, 15).

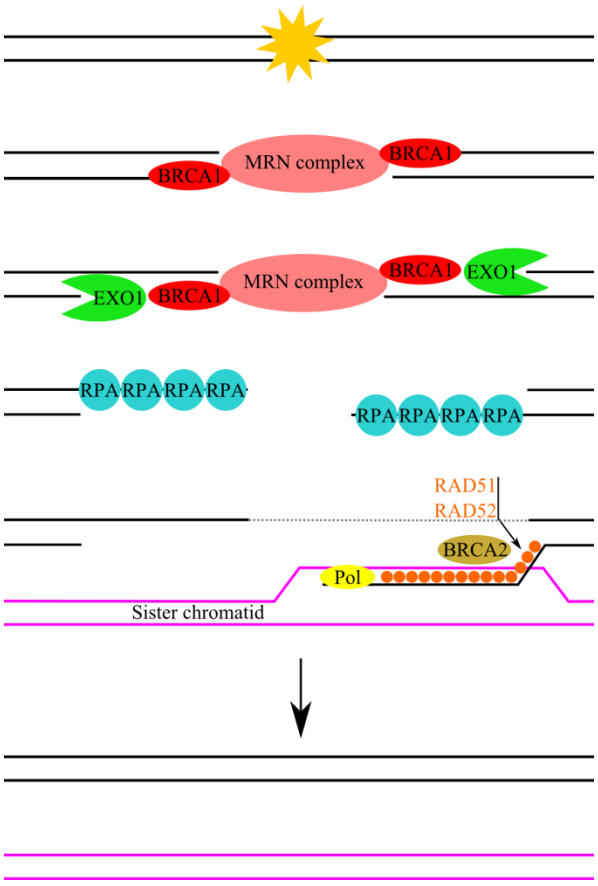

Figure 2: Homologous recombination. DSBs are detected by the MRN complex and the recruitment of BRCA1 to the site of the break promotes HR. The exonuclease activity of EXO1 results in the exposure of single strands that are coated with RPA. Following strand invasion, the damaged strands are repaired by using the sister chromatid as template.

1.1.3.2 Non-homologous end-joining and alternative end-joining

DSB can also be resolved by the NHEJ pathway. Although it shares a number of common molecular factors with HR, the two processes are fundamentally different. The NHEJ pathway repairs DNA breaks by joining the two exposed ends without making use of sequence homology (16). In the early stages of the process, ATM phosphorylates MDC1 and 53BP1 (*TP53BP1*) (17). This is followed by the recruitment of KU70 and KU80 to the site of the break and the formation of heterodimers (18). The dimer acts as a scaffold for the recruitment of other molecules

to the site (19). DNA-PKcs is believed to initiate the nuclease activity of Artemis (DCLRE1C) by phosphorylation, which along with polymerases μ and λ renders the ends of the breaks compatible for ligation (20-22). Finally, XRCC4 and Cernunnos promote the ligation of the two ends by DNA ligase IV (Figure 3) (10). A-EJ is the second end-joining repair pathway. In contrast with NHEJ, the A-EJ pathway makes use of sequence microhomology in order to repair DSBs, although this is not always the case (10, 23). In A-EJ, PARP1 associates to the broken DNA ends and not KU70/80 (24). A series of molecular factors follows PARP1 to the site, including XRCC1, polymerase θ and DNA ligase III (Figure 3) (25, 26).

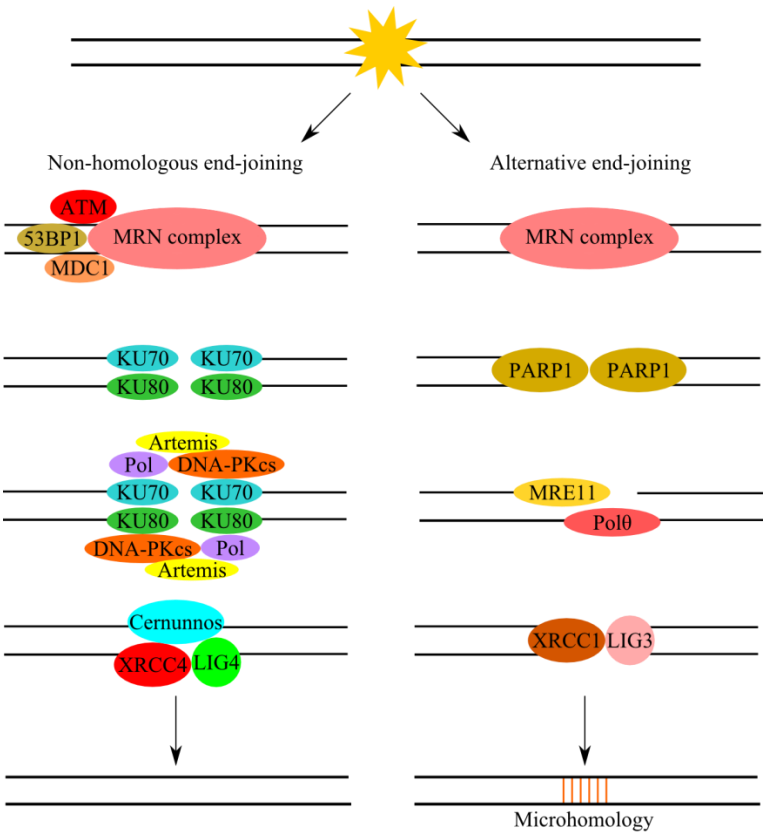

Figure 3: Non-homologous end-joining and alternative end-joining. Repair of DSB by making use of the NHEJ and A-EJ pathways. Assembly of the KU70/KU80 to the site leads to the formation of the NHEJ protein complex. Recruitment of PARP1 to the site inhibits the formation of the KU70/KU80 complex and as a result the DSB is repaired by the A-EJ pathway.

1.1.4 Single nucleotide lesion

Single nucleotide variations or SNVs in DNA consist of the replacement of one of the DNA bases adenine, thymine, guanine or cytosine with one of the other three. A single nucleotide alteration, if not correctly repaired, can have detrimental effects as it

might lead to the replacement of one amino-acid with another, the introduction of a stop codon and thus termination of the extension of the amino-acid chain or altered splicing of a transcript. Base-excision repair (BER), nucleotide-excision repair (NER) and mismatch repair (MMR) deal with single nucleotide alterations.

1.1.4.1 Base-excision repair

BER is responsible for removing damaged bases. DNA glycosylases such as UNG and OGG1 detect and remove the damaged bases, leaving an abasic DNA backbone at the site (27). The APE1 endonuclease cleaves the backbone and introduces a break, allowing polymerase β to incorporate the matching nucleotide (28). Finally, the break is sealed by DNA ligase III in complex with XRCC1 (Figure 4) (29, 30).

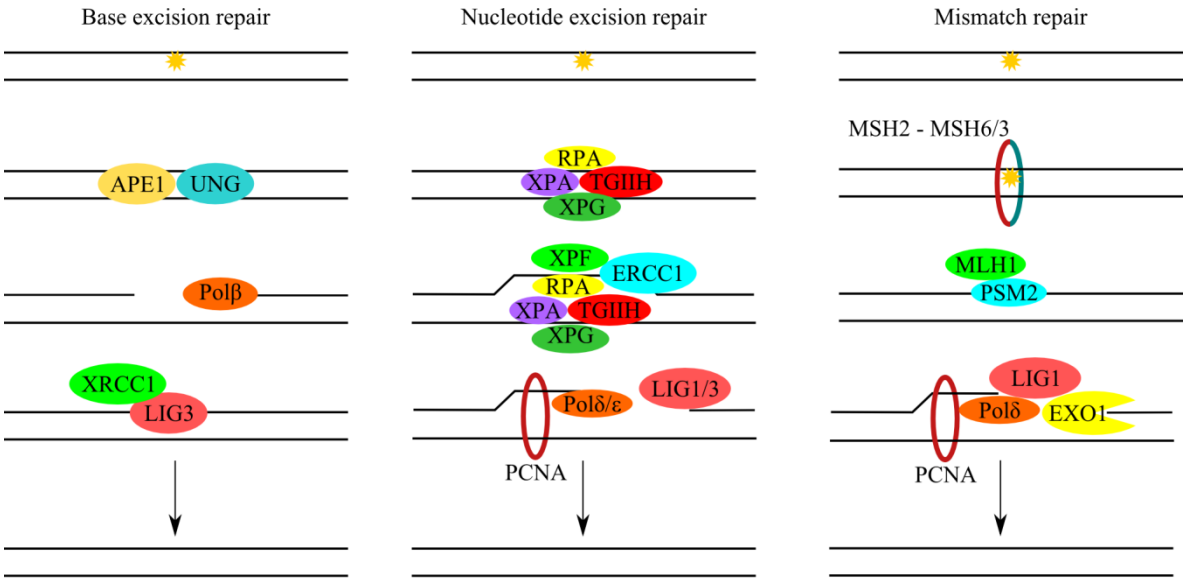

Figure 4: Repair of single nucleotide lesions. These three DNA repair pathways assure the repair of single nucleotide alterations. BER and NER replace damaged bases and nucleotides. MMR is responsible for the correction of replication errors that result in mismatches.

1.1.4.2 Nucleotide-excision repair

NER deals with bulky loops that are often caused by ultraviolet radiation (31). It requires the unwinding of the double strand on the site by the RPA/XPA/XPG and TGIIH complex and the excision of nucleotides by XPF associated to ERCC1. PCNA, polymerases δ and ε complete the repair along with DNA ligases I and III (Figure 4) (32).

1.1.4.3 Mismatch repair

MMR handles incorrectly incorporated nucleotides that result from erroneous insertions and deletions during DNA replication (33). The MSH2/MSH6 and MSH2/MSH3 heterodimers identify mismatches and abnormal loops in the structure of the genome. MLH1 associates to the genome in complex with PMS2 which introduces single strand breaks at the site. The exonuclease activity of EXO1 removes a number of nucleotides situated around the mismatch. Polymerase δ completes the repair process with the aid of PCNA and DNA ligase I (Figure 4) (34).

1.2 ADAPTIVE IMMUNITY AND ANTIBODY GENERATION

The term "immune system" encompasses a set of mechanisms and structures that protect the host from invading biological pathogens such as bacteria, viruses, fungi and parasites as well as tumors. Immunity is organized into two major systems, the innate and adaptive immune systems (35, 36).

1.2.1 Innate immunity

Innate immunity comprises a set of physical barriers that protect the body (including the skin and mucosa in contact with the external environment), the complement system of proteins and specialized cell types such as macrophages, Langerhans cells, dendritic cells, neutrophils, eosinophils, basophils, mast cells and natural killer (NK) cells. Cells of the innate immunity express surface receptors that recognize common structures found on different categories of pathogens. The innate immune system confers an immediate protection against disease that is not specific to the particular pathogen (35, 37, 38).

1.2.2 Adaptive immunity

Adaptive immunity, the other major branch of the immune system, is characterized by a slower but specific response to individual pathogens. Its principal components are T and B-lymphocytes which express receptors that specifically recognize epitopes, antigenic determinants of particular pathogens (36, 39). Development of this specificity requires time and therefore adaptive immunity arises much later than innate immunity during immune response (39). However, once a pathogen has been encountered by the adaptive system, an immunological memory is formed that makes it possible to mediate a very fast and efficient response during a future encounter with

the same pathogen (36). The two immune systems are not independent, and rather are strongly interconnected as the innate immunity is involved in antigen presentation and the activation of adaptive immunity (40).

1.2.2.1 T-lymphocytes

T-cells express the antigen-binding T-cell receptor (TCR) on their surface which recognizes epitopes mounted on human leucocyte antigen (HLA) molecules (39). The HLA class I molecules are expressed by almost every cell type and present to the immune system antigenic determinants from the cytoplasm. T-cells are activated upon the formation of a synapse between the TCR and an HLA class I molecule presenting a matching epitope. $CD8^+$ cytotoxic T-cells are able to recognize HLA class I signaling and respond by the secretion of perforin and granzyme B that cause the death of the target-cell (41, 42). Another class of T-lymphocytes, the $CD4^+$ helper T-cells are activated upon interaction with HLA class II molecules that are expressed by professional antigen presenting cells (APC). They respond with the secretion of cytokines that mediate the immune response by other cells, notably B-cells (41, 43).

1.2.2.2 B-lymphocytes

B-cells can also express epitope-binding receptors on their surface. The B-cell receptor (BCR) however can bind antigens directly and does not require them to be presented by HLA molecules. Engagement of the antigen initiates a signaling cascade that leads to the proliferation and differentiation into antibody-producing plasma cells and memory B-cells (44).

1.2.3 Antibody production in B-cells

B-cells secrete their epitope-binding receptor in the form of antibodies following differentiation into plasma cells. Antibody production is the main function of the B-cells. Secreted antibodies bind to their corresponding antigen and lead to a direct neutralization, opsonization of phagocytic cells or activate an antibody-dependent-cell-mediated cytotoxicity by NK cells (44-46). The generation of an immense library of antibodies specific to each epitope is achieved as a result of the processes of antibody diversification (47). This begins early on in B-cell development in the pro-B-cell stage, before the cells migrate from the bone marrow to the lymph nodes (45).

1.2.3.1 V(D)J recombination

In the pro-B-cell stage, the enzymes RAG1 and RAG2 introduce DSB into the immunoglobulin heavy-chain (*IGH)* locus on Chromosome 14. This initiates a process of *V*(*D*)*J* recombination where segments of the variable (*V*), diversify (*D*) and joining (*J*) regions of the *IGH* locus are recombined to assemble the variable region of the antibody gene. This recombination is ensured by the resolution of the DSBs by the NHEJ pathway. Each *V*(*D*)*J* recombination cycle produces a unique rearranged *IGH* from which a unique and antigen-specific BCR will arise (10, 48-51). Upon expression of the mature BCR, B-cells migrate to the peripheral lymphoid organs (lymph nodes, tonsils, spleen, Peyer's patches and mucosa associated lymphoid tissue) where they form germinal centers and undergo the final stages of their differentiation (Figure 5) (52).

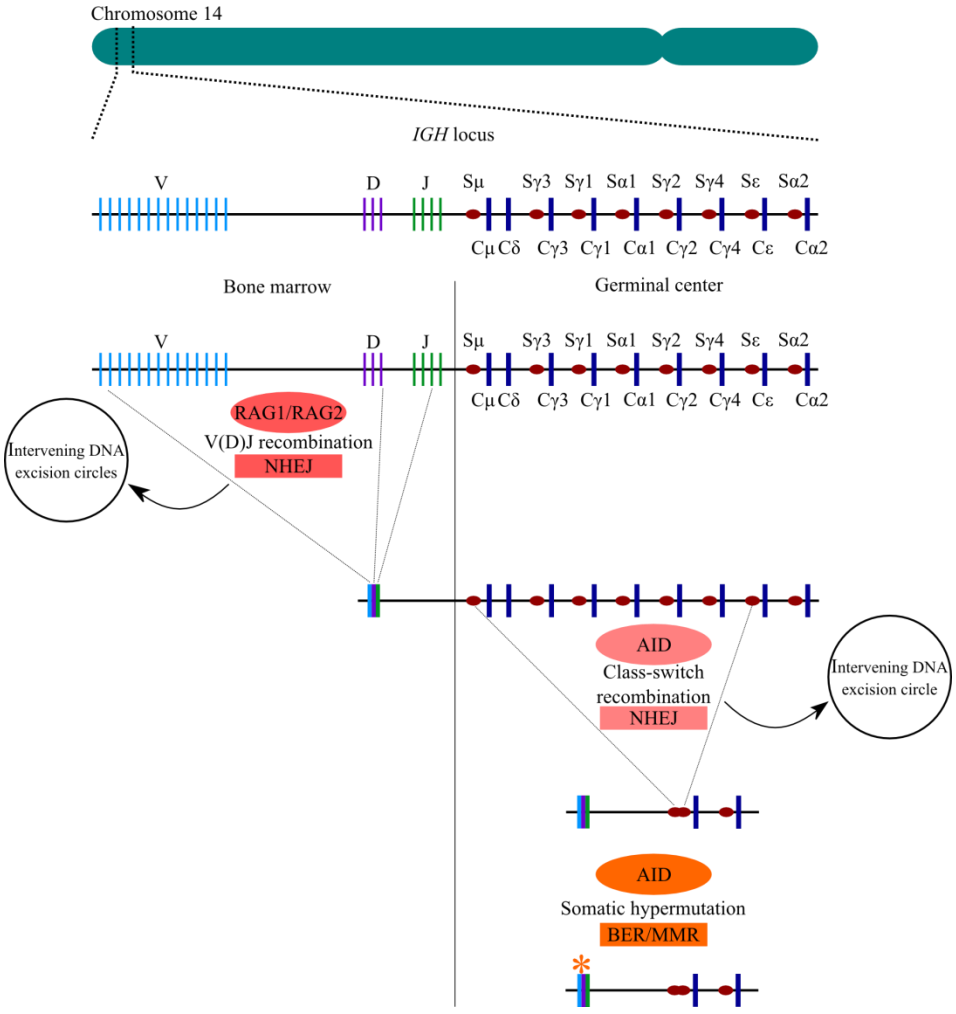

Figure 5: Antibody diversification processes. V(D)J recombination takes place in the bone marrow. RAG1/RAG2 introduce DSBs in the *IGH* locus that are resolved by NHEJ. The intervening DNA is discarded in the form of excision circles. CSR and SHM take place in the germinal centers. Mismatches resulting from AID activity in the S regions lead to DSBs that are resolved by NHEJ. AID-induced lesions in the V regions during SHM are processed by BER and MMR.

1.2.3.2 Class-switch recombination and somatic hypermutation

Recognition of the cognate antigen by the BCR activates B-cells and triggers the final stages of their differentiation. Class-switch recombination (CSR) is a process in which the gene coding for the constant (*C*) region the IgM (*Cμ*) antibody class is replaced by a downstream gene coding for another isotype (*Cα, Cγ* or *Cε*). Thus, CSR allows the cell to produce different classes of antibodies (*IgA, IgG*, *IgE* respectively). CSR is initiated by the deamination of cytidine residues in the switch (S) regions of the *IGH* locus by the activation-induced cytidine deaminase (AID) (52). The mismatches induced by the activity of AID on the *S* regions lead to DSBs that are repaired by the NHEJ pathway (17). Moreover, AID is also acting on the *V* regions during the process of somatic hypermutation (SHM). Mismatches introduced by AID here are resolved by the BER and MMR pathways. The purpose of SHM is to give rise to antibodies with enhanced affinities for the matching antigen. These processes finalize the differentiation into mature, antibody-expressing B-cells (Figure 5) (52).

Thus, antibody generation and diversification involves the formation of point mutations, DSBs and recombinations in the genome. These expose B-cells to particular mutagenic processes that need to be highly regulated and surveilled by DNA damage response and repair mechanisms.

1.3 CANCER

Cancers, also known as malignant neoplasms, are above all a “disease of the genome”. They arise due to alterations in the sequence and structure of DNA that lead to a loss of control of the cell’s proliferative rhythm (1). Such a transformation is not a single event but rather a series of sequential mutations over a long period of time that escape the surveillance of the cell’s DNA damage response and repair mechanisms (2, 53). Its most defining characteristic is the uncontrolled growth of the transformed cell population into an invading tumor that disrupts the normal function of the organism (54, 55). Tumors achieve an independence in their rate of growth by sustaining a constant proliferation signaling, becoming insensitive to apoptosis and growth suppressors, and by manipulating immune responses. As the disease progresses, they gain invasive capabilities and metastases occur throughout the body with new tumors developing in locations in the body that are distant to the initial transformation site. Although the malignant cells compose the bulk of the mass, we now know that tumors are complex biological entities and their growth is supported by tumor-associated stroma, vasculature and other types of non-malignant cells that are found infiltrated into the tumor, forming what is referred to as the tumor microenvironment (55).

1.3.1 Tumor suppressor genes and proto-oncogenes

In normal cells, proliferation, cell cycle progression and division are controlled by promoting or suppressing factors that form an equilibrium based on endocrine and exocrine signaling. From the early stages in the formation of a malignant tumor, this equilibrium is shifted towards perpetual growth. Tumor-suppressor genes that limit proliferation are inactivated by mutations whilst growth-promoting proto-oncogenes acquire activating mutations (56, 57). *TP53*, often referred to as the guardian of the genome, is a tumor suppressor gene with a central role in the maintenance of genome integrity and the progression of cell cycle control. It is found to be mutated in approximately 50% of all cancers making it the gene most often associated with cancer development (58-60). *RB1*, *TP16*, *BRCA1*, *BRCA2*, *CDK1*, *AKT1* and *PTEN* are other key tumor-suppressors frequently found to harbor inactivating mutations or copy number losses in cancers (61-66). Oncogenes however, whose activity promotes tumorigenesis, are found to be deregulated by different mechanisms. They are often targeted by mutations that lock them into a constitutively activated state of transmitting growth signals, or their expression levels are increased by promoter demethylation, copy number gains and structural rearrangements (56). *RAS*, *MYC*, *BCL2*, *BCL6* and *WNT* are some of the most commonly dysregulated proto-oncogenes in cancers (56, 67-70). Together, loss-of-function mutations in tumor-suppressors and gain-of-function mutations in proto-oncogenes are the primary factors that confer tumors with perpetually sustained growth signaling and resistance to growth suppression and apoptosis.

1.3.2 Angiogenesis

Angiogenesis is one of the most evident contributors to support tumor-growth by the microenvironment. Every cell, including those that are cancerous, requires a constant supply of nutrients and oxygen. As a tumor grows uncontrollably, cells are progressively distancing themselves from capillaries (71). This, however, is being counteracted by the production of angiogenic factors such as the vascular endothelial growth factor. This stimulates nearby vessels to grow into the tumor and provide access to the blood supply. In fact, the hypoxic environment created by the distancing of cells from the vessels is one of the main triggers of the angiogenic switch in a tumor (72). Simultaneously however, in order to adapt to their environment, tumor cells alter their metabolism and shift towards glycolysis even in the presence of oxygen. This phenomenon is known as the Warburg effect (73).

1.3.3 The immune compartment

Another key component of the tumor microenvironment that has received wide attention in recent years is the immune compartment. An intriguing and unexpected observation is that chronic inflammation can foster disease progression by promoting angiogenesis, providing growth signals and contributing to metastasis (55, 74).

Additionally, malignant cells appear to manipulate immunity in order to avoid detection and destruction. Cancerous cells acquire numerous mutations that give rise to neo-antigens that should, in theory, render them targetable by the immune system. In spite of that, tumors evolve to exploit a set of mechanisms that mask them from immune-recognition, block the activity of immune factors or induce the apoptosis of immune cells (75). These include the loss of expression of HLA molecules and the deregulation of immune-checkpoint pathways.

The CTLA-4/CD28 – B7 and PD-1 – PD-L1 immune-checkpoint pathways have received considerable attention recently as targeting these mechanisms demonstrated an impressive therapeutic potential in different malignancies. The two pathways work in similar ways to regulate the activity of T-cells that infiltrate the tumor microenvironment (76). Their precise mechanisms differ however as CTLA-4/CD28 interactions regulate early immune responses in lymphoid tissues, whereas the PD-1 – PD-L1 axis is mostly active in late responses in the peripheral tissues. Moreover, while CTLA-4 ligands are expressed by professional APCs, PD-L1 is expressed by many cell types as well as tumor cells (77).

1.3.3.1 CTLA-4/CD28 – B7

CTLA-4 and CD28 are expressed by $CD4^+$ helper T-cell and they compete for binding with B7 on the surface of the targeted cell. Binding to CD28 transmits an activation signal to T-cells whilst engagement of B7 by CTLA-4 triggers a negative signaling cascade that inhibits T-cell activity (78). Therapeutic antibodies targeting CTLA-4 such as Ipilimumab inhibit the interaction between B7 and CTLA-4 and block negative signaling, while indirectly promoting the binding of B7 to CD28. In the United States and the European Union, Ipilimumab has received approval as an therapy against melanoma and clinical trials assessing its efficacy against other cancers are underway (Figure 6) (79-82).

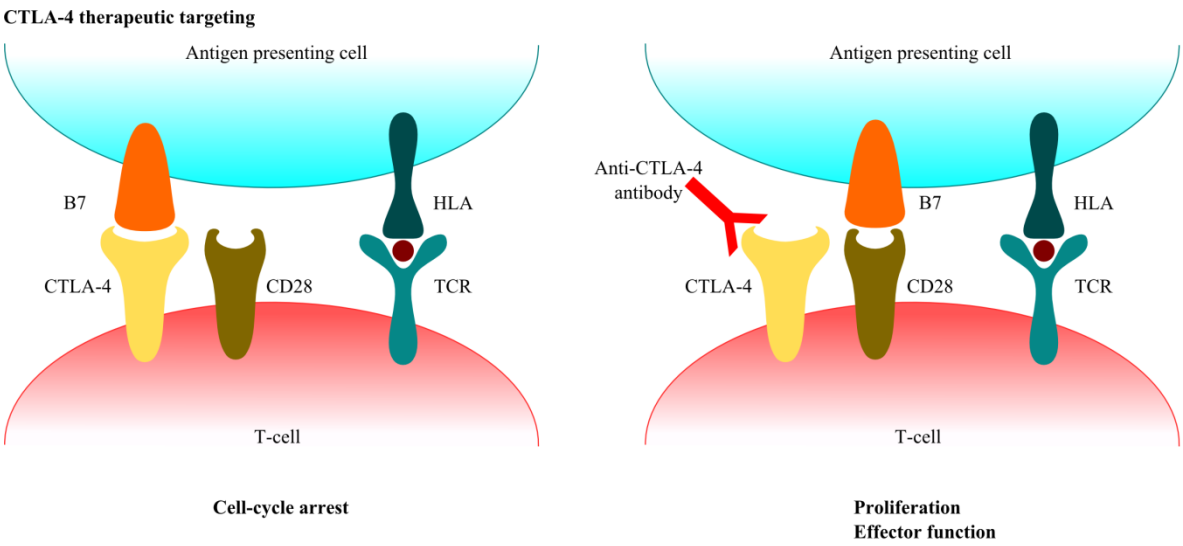

Figure 6: Anti-CTLA-4 treatment. Interaction between B7 and CTLA-4 leads to cell-cycle arrest in T-cells. CTLA-4 blockade favors the engagement of B7 by CD28 which leads to the proliferation and activation of T-cells thus promoting an effective immune response.

1.3.3.2 PD-1 – PD-L1/PD-L2

PD-L1 and PD-L2 expressed on antigen presenting cells (APC) and tumor cells interact with PD-1, their receptor on T-cells, leading to the functional exhaustion of T-cells and promoting their apoptosis. Under normal conditions, the function of this signaling pathway is to moderate the activity of effector cells and prevent excessive damage to the surrounding healthy tissue (83-85).

Different cancers, however, have been found to exploit this mechanism in order to shield themselves from immune attack by upregulating PD-L1 expression (86-89). Blocking of the interaction between PD-1 and PD-L1 with therapeutic antibodies in order to restore immune function has shown very promising results in clinical trials. Nivolumab, an anti-PD-1 monoclonal antibody has received approval from the Food and Drug Administration (FDA) in the United States for use in the treatment of aggressive melanoma and lung cancer (90). Currently, clinical trials indicate that patients affected by a wide spectrum of malignancies like renal carcinoma, bladder cancer, head and neck cancers, ovarian carcinoma and Hodgkin lymphoma can benefit from anti-PD-1 – PD-L1 treatments (91-94). In addition, the therapeutic effect is durable and longer-lasting as compared to CTLA-4-targeting treatments and toxicological studies show that it is better tolerated with fewer side-effects (Figure 7) (95).

In recent years, interest fueled by the increasing incidence of cancer and advancements in technology, notably in next-generation sequencing, has led to a

better understanding of the disease. The newly-acquired knowledge of the process of malignant transformation has contributed to the development of new therapeutic strategies (96, 97).

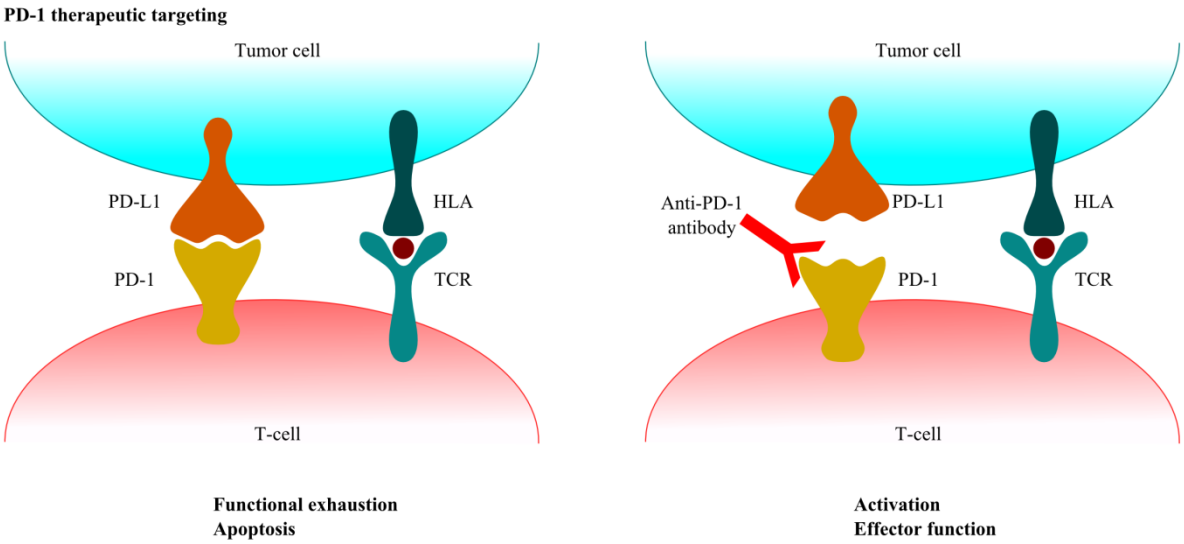

Figure 7: Anti-PD-1 treatment. The interaction of PD-1 and PD-L1 leads to the functional exhaustion of T-cells and promotes their apoptosis. Blocking this interaction by anti-PD-1 (or anti-PD-L1) therapeutic antibodies can lead to a more robust immune response by lifting inhibitory signals induced by the PD-1 – PD-L1 interaction.

1.4 B-CELL LYMPHOMAS

Lymphomas, myelomas and leukemia constitute a group of cancers called hematological malignancies. Lymphomas arise from cells of the lymphoid lineage and depending on the cell of origin they are classified as T-cell, B-cell or NK cell lymphomas (98, 99). The majority of lymphomas originate from B-cells. The genomic recombinations and somatic mutations that B-cells undergo during antibody diversification are believed to confer them an additional susceptibility for malignant transformation (100). Risk factors that have been associated with lymphomas include viral infections such as the Epstein-Barr virus (EBV) and the human immunodeficiency virus (HIV), immunosuppressive treatments following organ transplantation and germline mutations in key genes involved in apoptosis such as *BCL2L11*, *BAG5*, *BCLAF1*, and *CASP9*. (101-105).

1.4.1 Hodgkin lymphoma

Hodgkin lymphoma (HL) which represents 12% of lymphoma cases, is a particular subtype of B-cell lymphoma that is distinguished by the presence of large, multi-

nucleated, $CD30^+$ cells of characteristic morphology known as Reed-Sternberg cells (106, 107). Management of HL involves chemotherapy that generally yields very good results and the vast majority of patients survive their disease (108).

1.4.2 Non-Hodgkin lymphoma

Other types of B-cell lymphoma are collectively referred to as non-Hodgkin lymphomas (NHL). The two most common types of NHL are follicular lymphoma (FL) and diffuse large B-cell lymphoma (DLBCL) (109). FL accounts for 25% of NHL and is a mature B-cell lymphoma that arises from the germinal centers of the lymph nodes (110). It is considered to be an indolent malignancy, as many patients are asymptomatic over long periods of time, however, it remains an incurable disease (111). A subset of FL cases, about 35%, undergo histologic transformation into a high-grade and much more aggressive disease, DLBCL. The causes behind this are not yet clearly defined, but loss of the CDKN2A or CDKN2B-cell-cycle control kinases, loss of p53, deregulation of *MYC* expression, *BCL6* translocations, as well as *B2M* and *CD58* mutations have been implicated in the transformation (110, 112, 113). Evidence also suggests a link between aberrant SHM in FL and transformation into more aggressive NHL subtypes (114). Patients whose FL undergoes transformation have a poor prognosis even with the use of high-dose chemotherapy (115). This emphasizes the need for better understanding of the mechanisms of transformation as well as the genetics and pathophysiology of FL and DLBCL.

1.5 DIFFUSE LARGE B-CELL LYMPHOMA

DLBCL is the most frequent type of B-cell lymphoma, representing approximately one third of all new NHL diagnoses (99). It has an annual incidence rate of approximately 7 per 100 000 individuals in the United States and this has been increasing steadily over the past decades (107, 116). In those aged over 65 years, the incidence of DLBCL climbs to 68 per 100 000 persons (107). Although a number of cases are due to transformation from FL into DLBCL, the majority of diagnoses are *de novo* DLBCLs (117).

1.5.1 Disease subtypes and common genetic lesions

DLBCL originates from B-cells in germinal centers that undergo clonal expansion following cognate antigen recognition (118). The genetic background and clinical

manifestations of the disease vary considerably between cases, making DLBCL a very heterogeneous disease (119). Expression profiling has been the basis for classifying DLBCL cases into two subtypes, the germinal center B-cell-like (GCB) and the activated B-cell-like (ABC) subtypes (120). Molecular profiling has contributed to a better prediction of disease outcome and risk stratification for patients (121, 122). The ABC subtype is characterized by increased aggressiveness and has been associated with a less favorable disease outcome (123). GCB DLBCL is considered to be the malignant counterpart of germinal center B-cells. It is characterized by activating mutations in *EZH2*, amplifications in the *REL* proto-oncogene and translocations of *BCL2* to the *IGH* locus that lead to its increased expression. ABC DLBCL resembles more the post-germinal center activated B-cells where SHM activity has stopped. Its distinctive features are *BCL6* translocations and mutations in the regulator of hematopoietic differentiation *PRDM1*. Interestingly, *BCL6* translocations and loss-of-function of *PRDM1* both block terminal differentiation of B-cells and their occurrence is mutually exclusive. Another hallmark of ABC DLBCL is the chronically activated BCR signaling and NF-κB pathway. This is often the result of activating mutations in *CARD11*, *MYD88* and *CD79B* or inactivating mutations in *TNFAIP3* (*A20*). *BCL2* is also frequently upregulated in the ABC subtype, but in contrast to GCB, this is achieved not by genomic rearrangements, but by amplifications of the locus (118, 121, 123-130).

GCB	**ABC**
Mutations in gene coding for histone modifiers	
PI3K/Akt/mTOR activation	
MYC translocations	
BCL2 translocations	*BCL6* translocations
REL amplifications	*BCL2* amplifications
EZH2 mutations	*PRDM1* mutations
PTEN loss	BCR – NF-κB activation

Table 1: Oncogenic pathways in DLBCL.

Other common mutation targets in DLBCLs are genes implicated in the immune response such as *B2M*, *CD58* and *CD70*, the antiproliferative factors *TP53*, *BTG1*, *BTG2* and *SGK1*, the PI3K/Akt/mTOR pathway member *PTEN*, the histone modifiers *MLL2*, *CREBBP* and *EP300*, and the transcription factor *MEF2B (131-135)*. In addition to *BCL2* and *BCL6*, somatically acquired translocations involving the *IgH*

locus, *MYC*, *TP63* and *CIITA* have also been described in DLBCL. In fact, *IgH* locus rearrangements are considered as a hallmark of the disease (136-140).

1.5.2 The role of antibody diversification in DLBCL

The mutation load of DLBCL is relatively high when compared to other forms of B-cell malignancies (128). A contributor to that is aberrant targeting of non-immunoglobulin regions by AID during the germinal center stage of B-cell development in which DLBCL arises. A number of proto-oncogenes such as *BCL6*, *RHOH*, *MYC*, *PAX5* and *PIM1* have been identified as targets of illicit AID activity (141). Mismatches introduced by AID can also lead to DSB in the genome that can be recombined, resulting in oncogenic translocations between proto-oncogenes and the *IGH* locus (141, 142). Interestingly, studies on AID-deficient mouse models found that lack of AID prevents translocations between the *IGH* locus and *MYC (143, 144)*. Moreover, *in vitro* studies show that *BCL2* is cleaved by the RAG complex at its major breakpoint region which suggests a role of V(D)J recombination in the formation of *BCL2* rearrangements (145, 146). These findings demonstrate how physiological processes specific to B-cells contribute to the formation of B-cell lymphomas.

1.5.3 Immune system evasion

One of the most frequent mutation targets in DLBCL is *B2M* which is found to be affected by genetic alterations in approximately 30% of cases. Additional mechanisms that lead to loss of *B2M* expression might be active in DLBCL, as 60% of cases are *B2M*-deficient (147). As a result, over half of DLBCLs do not express HLA class I on their surface which compromises antigen presentation and masks them from $CD8^+$ cytotoxic T-cells (147, 148). Moreover, about 20% of DLBCLs carry mutations in *CD58*, a member of the immunoglobulin superfamily which is important for the activation of T-cells and NK cells (147, 149, 150). DLBCL derives from B-cells which are professional APCs expressing HLA class II. The loss of expression of HLA class II molecules has been described in DLBCL and was found to be associated with decreased immunosurveillance and poor disease outcome. In many cases this is due to the downregulation of *CIITA*, the master regulator of HLA class II expression, or the expression of *FOXP1* (151-153). In summary, DLBCLs have several different ways to escape surveillance and attack by the immune system.

1.5.4 Management of DLBCL and prospective targeted therapies

More than half of the patients suffering from DLBCL can be cured by currently available treatments (154). However, DLBCL is amongst the most aggressive forms of lymphoma and relapse after treatment occurs in a significant proportion of patients. These patients are refractory to treatment and usually succumb to their disease, which underlines the necessity for establishing new therapeutic options (99, 155).

1.5.4.1 CHOP

The first generation of chemotherapy combinations for DLBCL was introduced in the 1970's and was based on anthracycline regimens (156). Since the early 90's, the classical chemotherapeutic treatment for DLBCL has been a combination of Cyclophosphamide, Doxorubicin, Vincristine and Prednisone, commonly referred to as CHOP. This regimen provides similar results in terms of survival as the previously used anthracycline treatments but is better tolerated by patients (157).

1.5.4.2 Anti-CD20 antibodies

The next milestone in the management of DLBCL came with the supplementation of CHOP with Rituximab (R), an anti-CD20 chimeric therapeutic antibody and the first immunotherapeutic agent to be used for the management of DLBCL (158, 159). Rituximab contributed to a significant increase in the overall survival of patients of all ages and R-CHOP has become the gold standard for the treatment of DLBCL (160). Etoposide, a topoisomerase inhibitor is effective against primary mediastinal B-cell lymphoma and its overall utility as an adjunct to R-CHOP in DLBCL is currently being assessed (160, 161).

1.5.4.3 Salvage therapy

Autologous stem cell transplantation with $CD34^+$ cells and transplantation with autologous T-cells expressing anti-CD19 chimeric antigen receptors are used as salvage therapies after relapse following chemotherapy. Although complete remission and an increase in overall survival are achieved, outcomes can be further improved, hence the need for new approaches in the management of DLBCL (162-164).

1.5.4.4 New therapeutic approaches

Rituximab demonstrates an important "evolutionary step" in the way drugs are developed and selected. It is a targeted therapy, specifically directed against a molecular marker of B-cell lymphoma, the surface antigen CD20 (165). As such, it illustrates a shift from traditional chemotherapy towards advanced therapies aiming at specific targets on tumor cells. A number of molecules currently under evaluation are listed in Table 2.

Drug	**Molecule class**	**Targeting feature**	**Status (FDA)**
E7438	EZH2 inhibitors	*EZH2* activating mutations	Phase II trial
ABT-199	BCL2 inhibitors	*BCL2* overexpression	Phase I trial
Ibrutinib	BTK inhibitors	Constitutively active NF-κB pathway	Phase III trial
Entospletinib	SYK inhibitors	Constitutively active NF-κB pathway	Phase II trial
BKM120	PI3K blockade	PI3K/Akt/mTOR activation	Phase II trial
MK2206	Akt inhibitors	PI3K/Akt/mTOR activation	Phase II trial
Everolimus	mTOR inhibitors	PI3K/Akt/mTOR activation	Phase II trial
Veliparib	PARP1 inhibitors	PARP1 in tumors with HR defects	Phase II trial
SGN-CD70A	CD70 blockade	CD70 – CD27 signaling	Phase I trial
-	CD27 agonists	CD70 – CD27 signaling	Not investigated
Pembrolizumab, Nivolumab, Pidilizumab	PD-1 blockade	PD-1 – PD-L1/2 signaling	Phase II trial
Atezolizumab	PD-L1 blockade	PD-1 – PD-L1 signaling	Not investigated
RO4929097	γ-secretase inhibitor	Notch signaling	Phase I trial

Table 2: List of targeted therapies for DLBCL. Information derived from www.clinicaltrials.gov.

The histone-lysine *N*-methyltransferase EZH2 is frequently mutated in DLBCLs of the GCB subtype. Mutations altering the Tyr641 residue in the protein's SET domain are identified in approximately 22% of patients (125). Studies in DLBCL cell lines

found that these alterations confer a gain-of-function in *EZH2* which results in increased trimethylation of H3K27 (166). *In vitro* and *in vivo* studies have shown that as a consequence, the expression of important tumor suppressor genes is silenced and thus Tyr641 mutations in *EZH2* are important in promoting GCB lymphomagenesis (167, 168). EZH2 inhibitors have shown potential in studies as they promote apoptosis and cell-cycle arrest both *in vivo* and *in vitro* and are currently the subject of clinical trials for their effectiveness against lymphoma *(169-171).*

As mentioned earlier, *BCL2* and *BCL6* are often upregulated in DLBCL and inhibitors targeting these anti-apoptotic factors are being investigated (160). Although earlier studies of BCL2 inhibitors have shown toxicity due to BCL-XL blockade, newly developed agents show more promising results with less side effects (172, 173). Studies on BCL6 inhibitors have also provided a proof of principle in mouse studies where a reduction in the size of the tumors was observed (174, 175).

Chronically activated BCR signaling leads to the constitutive activation of the NF-κB pathway, a major characteristic of ABC DLBCL (176). The NF-κB pathway constitutes an important set of molecular factors involved in proliferation, invasion and metastasis and its targeting shows therapeutic potential (177). Loss of *A20* and mutations in *CD79A/B*, *MYD88* and *CARD11* are often associated with a constitutively active NF-κB in the ABC subtype of DLBCL (178). Targeting of the Bruton's tyrosine kinase (BTK) by Ibrutinib is highly toxic to ABC DLBCL with upstream mutations in *CD79A/B* but not in patients with downstream C*ARD11* or *MYD88* mutations alone. This observation demonstrates the principle of targeted therapies and how their outcome can be predicted (127, 179). Patients with mutations in other members of the BCR/NF-κB pathway upstream of BTK such as SYK and LYN could also potentially benefit from Ibrutinib treatment.

PI3K/Akt/mTOR signaling is important for various processes relevant to cancer development such as survival and proliferation and is linked to the BCR/NF-κB pathway through *SYK* and *CD79* (180, 181). The use of SYK inhibitors in DLBCL requires further investigation, but data suggests that this approach could be promising (182). Mutations of the *PTEN* tumor suppressor gene are one of the hallmarks of GCB DLBCL and they constitute another potent mechanism for the constitutive activation of the PI3K/Akt/mTOR pathway in DLBCL (183, 184). Additional classes of drugs targeting this pathway have been developed and are undergoing clinical investigation (185). These include molecules that block PI3K signaling, Akt inhibitors and mTOR inhibitors (186-188).

The targeting of DNA repair pathways, such as HR, has received growing attention in recent years. Although not directly a participant in HR, PARP1 associates to single

strand breaks during replication and in the absence of its activity the replication fork collapses. This however, is managed by HR in normal cells. In tumor cells with HR deficiencies, repair of collapsed replication forks is not possible and leads to cell death. PARP1 inhibitors could therefore be used in the context of synthetic lethality to treat tumors with HR deficiencies. (189, 190).

CD27 interaction with CD70 constitutes one of the immune-modulatory pathways regulating lymphocyte activity. Expression of CD70 by tumor infiltrating lymphocytes and subsequent exhaustion of effector T-cells has been observed in NHL (191). Moreover, the overexpression of CD70 by malignant cells has been described in B-cell lymphomas where it transmits a proliferative signal favoring tumor growth (192, 193). Studies with agonist anti-CD27 antibodies on transgenic mice suggest they can restore T-cell activity and promote anti-tumor immunity thereby suggesting another potential therapeutic avenue (194).

An immune checkpoint with a great therapeutic potential currently being investigated is the PD-1 – PD-L1/PD-L2 pathway. Up to 87% of HL patients with refractory disease have shown a response to treatment with Nivolumab, a PD-1 inhibitor, and *in vitro* studies on DLBCL have demonstrated the restoration of activity in infiltrating T-lymphocytes upon PD-1 blockade (91, 195). Histological evaluation of DLBCL patients has shown that a considerable proportion of patients express the PD-1 ligand PD-L1 (196). EBV infection accounts for the expression of PD-L1 in a number of tumors. However, the mechanisms of PD-L1 expression in DLBCL are not yet fully elucidated (197, 198). Nonetheless, expression of PD-L1 by DLBCL further suggests that targeting of PD-1 or its ligands could prove beneficial for patients suffering from aggressive and refractory DLBCL.

Such specific approaches to therapy require an understanding of the genetic background of the disease (199). Knowledge of the tumorigenic mechanisms particular to an individual case will help in predicting the outcome of therapies and make more precise choices of regimen for each patient (160, 200). The characterization of cancers, which is crucial to this goal, has been greatly empowered in recent years by the rapid development of sequencing technologies (201).

1.5.5 Next-generation sequencing of DLBCL

One of the main driving forces behind the advancements of biomedical research in the past 15 years has been the development of next generation sequencing (NGS) methods (96). The Sanger method was used during the first wave of large-scale sequencing in cancer genomics. This led to important discoveries that identified a

number of key genes involved in cancer. However, the method was too costly, slow and laborious to be implemented on a larger scale (202-204). NGS brought a rapid decline in the price per nucleotide sequenced and unprecedented progress in the speed and throughput of massive parallel sequencing (96, 205). A revolution ensued in genomics, transcriptomics and epigenetics as tremendous amounts of data could be generated at a manageable cost (206).

NGS methods have been employed to decipher the genome of DLBCL and much of the knowledge on the genetics of DLBCL mentioned earlier in this introduction is the result these studies (132, 133, 137, 184, 207). These studies have demonstrated the great level of heterogeneity characterizing the genetic landscape of DLBCL and prompted further investigation. The identification of clinically relevant biomarkers can form the basis for the development of new drugs, aid clinicians in choosing the appropriate therapies and establishing reliable prognostic models (208-210).

2 AIM OF THE STUDY

2.1 GENERAL AIM

The aim of this thesis is to explore the genome of DLBCL by employing a series of NGS methods. The data generated is used to identify novel mutation targets and structural alterations that contribute to the lymphomagenesis of DLBCL. Furthermore, this thesis aimed to functionally characterize the impact of selected NGS findings and explore how they could contribute to new approaches in the treatment of DLBCL.

2.2 SPECIFIC AIMS

Paper I: This project focused on assessing the role of mutations in DNA repair genes in the development of lymphomas. Moreover, it sought to investigate their relation to the processes of antibody diversification.

Paper II: Here, a whole exome sequencing (WES) approach was used to identify genes targeted by mutations in DLBCL. In addition, it sought to address experimentally the impact of mutations in a number of newly discovered mutational targets.

Paper III: This project investigated the role of genetic alterations as mechanisms of PD-1-ligand upregulation in DLBCL. Moreover, it aimed to describe the distribution of those alterations across the disease subtypes and their impact on patient survival.

3 MATERIALS AND METHODS

An extensive description of the reagents, patient material and methods used can be found in the scientific papers appended to this thesis. Here is an overview of selected methods to facilitate the interpretation of the results.

3.1 TARGET ENRICHMENT BY THE SELECTOR TECHNOLOGY

The Selector method, which preceded the current Agilent HaloPlex method, was used for the enrichment of the coding regions of 73 key DNA repair genes in 29 lymphoma samples, 22 DLBCL, 5 FL and 2 Burkitt's lymphomas (BL), all of which had matching blood controls (211). The *A20* gene was also included and served as an internal quality control. The genomic DNA was first fragmented by restriction nucleases and probes labeled with biotin were used to pull-down the 1158 exonic target regions with streptavidin-coated magnetic beads. The probe design causes a circularization of their target region by being complementary to both its extremities. Following the pull-down step, the circularized fragments were ligated and enriched by amplification with a Phi29 polymerase (211). Of note, each target region was covered by more than one probes as to minimize hybridization failure due to single-nucleotide polymorphisms (SNPs).

3.2 SOLID SEQUENCING AND DATA ANALYSIS

The enriched targets were barcoded and a sequencing library was constructed using the standard SOLiD method. The barcoded, adaptor-ligated fragments were attached to beads and underwent clonal amplification in an emulsion polymerase chain reaction (PCR). The beads were then sequenced with the SOLiD "sequencing-by-ligation" method that employs fluorescently-labeled di-base probes. In brief, four fluorophores are used in the SOLiD method to label the probes and therefore each color represents 4 of the 16 possible di-nucleotide combinations. A primer is hybridized to the adaptor followed by the first fluorescent probe. The probes and primer are being ligated by a DNA ligase present in the reaction mix. The fluorescent light is measured and the next probe is incorporated and ligated. Following a cycle of 7 reactions, a 35 base pair read is generated. The length of individual reads can be extended by additional cycles. The strand is removed by denaturation and a new primer is introduced into the mixture, offset by one base as compared to the previous one. The probe ligation cycle is repeated five times and the target sequence is resolved.

SNPmania, an in-house software, was used for the calling of SNVs as well as for small indels. Comparing the minor-allele frequency (MAF) between tumor and blood control allowed for the determination of the somatic status of alterations. The hg18 human genome assembly was used as a reference. An *in silico* prediction of the functional impact of mutations was performed with the SIFT and Polyphen-2 algorithms (212, 213). Comparison of the allelic ratios of heterozygous positions between paired tumor and blood controls enabled the detection of allelic imbalances.

3.3 ILLUMINA HISEQ SEQUENCING

3.3.1 Whole-exome sequencing

DNA extracted from DLBCL tumors and their paired blood controls was processed into 150-200 bp-long fragments prior to WES. Adaptors were ligated to both ends and the fragments were purified and amplified using a ligation-mediated PCR method. The exonic regions were enriched by hybridization to the SureSelect Biotinylated RNA library. The enriched libraries were sequenced on an Illumina Hiseq 2000 sequencer which utilizes a sequencing-by-synthesis method. In brief, the libraries are loaded on a flow cell with primers fixed on its surface. The individual fragments of the library hybridize to the primers due to sequence complementarity between the fragments' adaptors and the primers. The fragments are amplified by a solid-phase bridge amplification step. Following denaturation of the double strands, reversible chain-terminator nucleotides labeled with fluorophores are introduced into the reaction mix. A polymerase incorporates the first nucleotide to the single strands. Each of the four nucleotides is labeled with a different fluorophore and thus the incorporated nucleotide can be identified by a camera. Further on, the chain-terminator nucleotide is modified in order to render further polymerization of the chain possible. The cycle is repeated and the individual reads are produced.

The raw data was processed to remove the adaptor sequences and filter-out low quality reads. The sequencing reads were aligned using the Burrows Wheeler Aligner (BWA) (214). Varscan was used for the detection of SNVs and GATK for calling of indels (215, 216). Predictions on the impact of missense mutations was performed with SIFT and Polyphen-2 as mentioned earlier. The Interactive Genomic Viewer (IGV) allowed for visual inspection of the mutations (217).

3.3.2 Whole-genome sequencing

DNA from tumors and paired blood controls was fragmented with a Covaris E-210 ultrasonicator. The shearing process was adjusted to yield fragments of approximately 500 bp. The fragments were subsequently purified, end-blunted, A-tailed and ligated to adaptors prior to amplification by 10 to 12 PCR cycles. The libraries were then quantified by quantitative real-time PCR (qPCR) and a bioanalyzer. Whole-genome sequencing (WGS) was performed by the Illumina Hiseq 2000 platform as described earlier. A number of additional samples were processed by the Illumina Hiseq X10 platform (218).

Mutation analysis was performed similarly to the WES and in addition, SegSeq was used for detecting copy number changes (219). Moreover, SVs such as translocations were detected with the in-house SeekSV software.

3.3.3 RNA sequencing

RNA was extracted from tumors and fragmented with the NEBNext Magnesium RNA Fragmentation Module. The fragmented RNA was ligated to adaptors and hybridized to reverse transcription primers before synthesis of cDNA by the SuperScript III reverse transcriptase. The obtained cDNA was amplified by PCR and subsequently underwent size selection by AMPure beads. Sequencing of the cDNA libraries was performed on the Illumina Hiseq 2000 platform (220).

The expression levels were calculated with the “fragments per kilobase of transcript per million mapped read” (FPKM) method and fusion transcripts were detected with the SOAPfuse software (221).

3.4 454 GS FLX+ TITANIUM SEQUENCING

Selected genes were sequenced in extended cohorts using this method. The target regions were amplified in microfluidic PCR reactions on a Fluidigm Access array using the FastStart High Fidelity PCR system by Roche. The primers used carried CS1/CS2 tail adaptors. The DNA templates from different samples were pooled together and added to the array along with barcoded sequencing primers containing adaptors complementary to the amplification primer tails. The target regions were thus enriched and barcoded in order to be able to distinguish the amplicons derived from each individual sample. Following the enrichment step, the libraries were analyzed on a 2100 Bioanalyzer by Agilent technologies and purified by gel extraction. Finally,

sequencing was performed on a GS FLX+ System and the data was analyzed by the GS Amplicon Variant Analyzer.

3.5 ALLELIC IMBALANCE AND MICROSATELLITE INSTABILITY

The presence of allelic imbalances in selected loci was assessed by making use of polymorphic sequence-tagged site (STS) markers. The latter consist of repetitive motifs within the genome where the number of repetitions varies across different alleles. In an individual heterozygous for a particular STS marker, the distribution of the two alleles is normally 1:1. This ratio is altered in the case of allelic imbalances such as loss or gain of one of the alleles (Figure 8). Therefore, by measuring the size and ratio of STS markers in a certain position, the presence of allelic imbalances in a locus can be detected. To this end, heterozygous STS markers in the vicinity of the locus of interest were amplified by PCR using forward primers labeled with 6-carboxyfluorescein (6-FAM) fluorescent dyes.

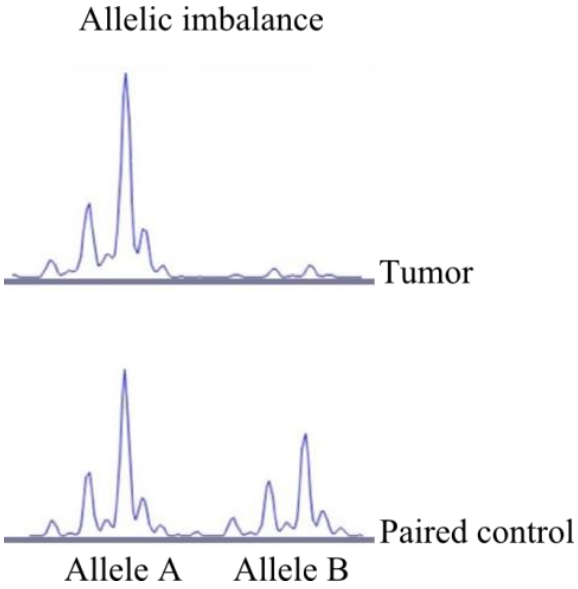

Figure 8: Example of allelic imbalance. Each of the two alleles is represented by a major peak surrounded by minor peaks. In the control two alleles can be observed whilst in the tumor only one is detected.

The amplicons were then loaded onto a capillary DNA analyzer in order to measure the fragment's size and intensity of fluorescence from each of the two alleles. The fluorescent signal is proportional to the presence of the individual alleles. The allelic distribution of STS markers of tumor samples and paired normal tissues were analyzed in order to obtain an interpretable result. An allelic imbalance was called when the distribution of alleles between the normal control and the tumor exceeded 50%. The loci of *EXO1*, *PARP1*, *MDC1* and *RPA1* were assessed for allelic imbalances. Analysis of the *TP53BP1* locus was attempted as well but proven to be

difficult due to a high degree of homozygosity of the STS markers assessed for that locus. Two STS markers from each site of the loci were analyzed to ensure the accuracy and reproducibility of the result. The markers where selected from the UCSC genome browser and the Ensembl genome browser. The primers necessary for their amplification were either retrieved from the same databases or designed using Primer 3. High-performance liquid chromatography (HPLC) was used for the purification of the fluorescently labeled oligonucleotides. The DNA analyzer used was a 48-capillary 3730 DNA Analyzer and the generated data was processed with the Peak Scanner software by Applied Biosystems.

The presence of microsatellite instability (MSI) was evaluated in a similar fashion. Short tandem repeats (STRs) are repetitive sequences within the genome and can serve as markers for MSI. The size of an STR can vary among individuals due to the number or repetitions. However, they constitute a genetic fingerprint and the number of repetitions remains unaltered under normal circumstances (222). Instability in the size of STRs is observed when MMR is impaired (Figure 9) (34). The markers used for the detection of MSI were the mononucleotide markers BAT-25 and BAT26, as well as the dinucleotide markers S2S123 and D10S197. The marker amplification and process by the DNA analyzer was performed in a manner similar to that used for the detection of allelic imbalances. However, the analysis consisted of identifying changes in the size of the STR markers between the normal control and the tumor sample.

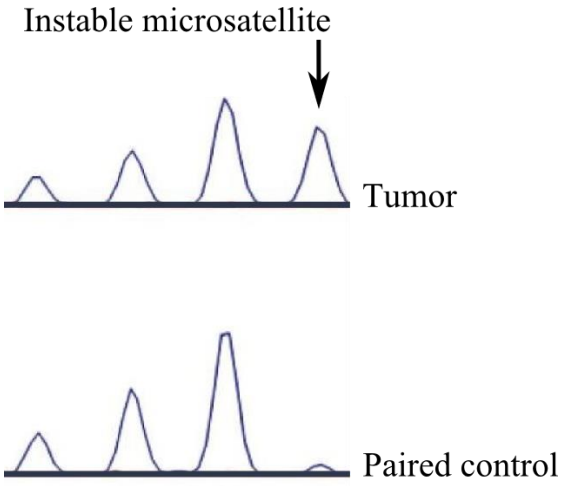

Figure 9: Example of microsatellite instability. Fragment analysis detects an additional peak in the tumor sample. This indicates instability in the size of the microsatellite due to an increase in the number of repetitions of the motif.

3.6 NOTCH DUAL LUCIFERASE REPORTER ASSAY

DTX1 is a negative regulator of the NOTCH signaling pathway. NOTCH1 is a transmembrane receptor. Upon engagement by a ligand, the NOTCH1 intracellular

domain (NCID1) is translocated to the nucleus where it interacts with the CSL transcription factor and drives the expression of its target genes. DTX1 represses the expression of those target genes by interacting with NCID1 and inhibiting the recruitment of coactivators (223).

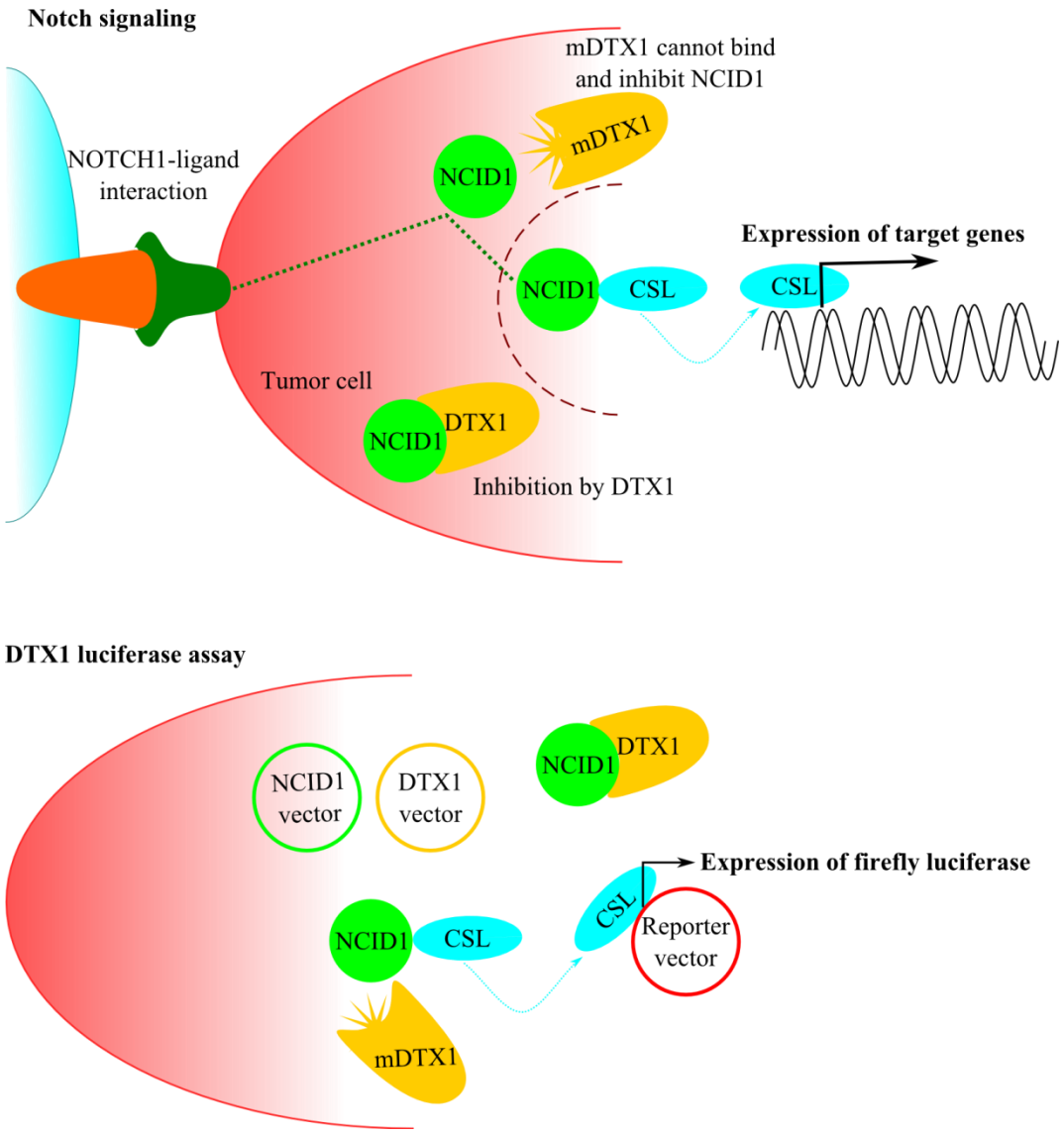

Figure 10: Notch luciferase reporter assay. Interaction of Notch1 with its ligand leads to the liberation of NCID1 into the cytoplasm. NCID1 then translocates to the nucleus where it interacts with co-activators and CSL and drives the expression of its target genes. DTX1 acts as a negative regulator by binding to NCID1 and inhibiting downstream interactions. In our system, WT DTX1 could bind to NCID1 thus inhibiting the expression of firefly luciferase. Some DTX1 mutants (mDTX1) did not inhibit the expression of the reporter thus indicating a loss of the ability to bind NCID1.

This interaction between DTX1 and NCID1 was used in order to assess the impact of mutations identified on *DTX1* by means of a dual luciferase reporter assay (Figure

10). cDNA was used as a template for the amplification of the coding region of the wild type *DTX1* and *NCID1*. The amplification primers were designed in such a way as to confer the ends of the amplicons with restriction sites that allowed their cloning into the pcDNA 3.1 mammalian expression vector by Life technologies. Site-directed mutagenesis was used for the generation of 14 *DTX1* mutants discovered in the DLBCL patients by DNA sequencing. The reporter gene, firefly luciferase, was cloned into a vector. The expression of firefly luciferase was coupled to the transcription element *CSL* whose activation depends on binding by NCID1. A pRL-TK plasmid expressing *Renilla* luciferase served as internal control. The plasmids were co-transfected into the U2OS osteosarcoma cell-line using the Turbofect reagent by Fermentas. After 48 hours of incubation, the cells were lysed and the firefly luciferase levels were measured with the Dual-luciferase reporter assay by Promega on a Thermo Scientific Varioskan Flash microplate luminometer. The *Renilla* luciferase levels were also measured and used for normalization.

4 RESULTS AND DISCUSSION

4.1 PAPER I

DNA repair genes are selectively mutated in diffuse large B-cell lymphomas.

This project focused on assessing the role of DNA repair genes in the development of lymphomas, especially in relation to the processes of antibody diversification. The results showed that 50% of the DLBCLs involved in the study carried mutations in at least one of the screened genes. Functional studies also associated defects in specific DNA repair mechanisms to unique phenotypes.

4.1.1 Mutations identified in DNA repair genes

The targeted sequencing of the coding regions of 73 key DNA repair genes across 29 mature B-cell lymphomas identified a number of somatic, novel germline and rare germline (MAF<0.01) mutations. Approximately half of the 442 SNVs identified by the SOLiD method passed visual inspection by IGV and 124 resulted in non-synonymous alterations. Sixty-nine of those were validated by Sanger sequencing. This includes all 33 novel and 21 rare germline SNVs as well as 15 of the 21 somatic SNVs. The majority of those mutations, including all the somatic mutations, were found in the 22 DLBCL tumors despite the comparable sequencing performance in the FL and BL tumors. BFAST enabled the discovery of four somatic deletions in four different DLBCL tumors. Collectively, the SOLiD sequencing method revealed 73 non-synonymous mutations in the cohort (Table 3).

4.1.2 Distribution of mutations across DNA repair pathways

Nineteen somatic and 45 germline mutations were distributed in 22 DLBCL tumors. The pathways more frequently affected by mutations were the DDR, MMR, HR and NHEJ pathways.

In the DDR pathway, mutations were found in *ATM*, *MDC1*, *RAD50*, *RPA1*, *TP53*, and *TP53BP1*, as well as in two novel mutation targets in DLBCL, *CHEK2* and *PARP1*. The *EXO1*, *MLH3, MSH2, MSH3*, *MSH6, PMS1* and *PMS2* MMR factors were found to be mutated in our cohort. The HR factors affected by mutations were

BLM, *BRCA1*, *BRCA2*, *FANCA*, *FANCG*, *RAD51B* and *RAD54B*. NHEJ was targeted by mutations in the *DCLRE1C*, *PNKP*, *PRKDC*, *XRCC5* and *XRCC6* genes.

Pathway	Gene	Somatic	Novel germline	Rare germline (MAF<0.01)
DDR	*ATM*	p.L1206V; p.Y1215H	-	p.R629W
	CHEK2	p.Q336L	-	p.E528K
	MDC1	-	p.H35L; p.N280D; p.P965L	-
	PARP1	p.K637Sfs*13	p.S776G	-
	RAD50	-	p.V400E	p.L1125F
	RPA1	-	p.G160R	-
	TP53	p.R273C; p.F134V; p.R273H	-	p.V31I
	TP53BP1	-	-	p.T519A (x2); p.A1714S (x2)
MMR	*EXO1*	p.K471R	-	-
	MLH3	-	p.L1111F; p.C40Y	p.I988M; p.S946F
	MSH2	p.S540G; p.L833V	p.G820D	-
	MSH3	-	p.P657S; p.R1061G (x2)	-
	MSH6	p.K543R	p.T563N	-
	PMS1	-	p.L369P; p.D397E	p.R919C
	PMS2	-	p.M362K	p.S128L
HR	*BLM*	-	p.S580P	p.V765I
	BRCA1	-	-	p.Y856H
	BRCA2	p.V849A; p.M3322I	p.S1744I	p.C315S; p.I1929V
	FANCA	-	p.R1321H; p.T620I	-
	FANCG	-	p.M431R	-
	RAD51B	-	p.E346D	-
	RAD54B	-	p.P55L	p.I778V
NHEJ	*DCLRE1C*	p.I543Rfs*12	-	-
	PNKP	-	p.P112R	-
	PRKDC	p.F3418L	p.D566N; p.K1984N	-
	XRCC5	p.T283A	-	-
	XRCC6	p.I267Kfs*2	-	-
BER	*LIG3*	-	p.R343Q	-
	POLB	-	p.R333Q	-
	UNG	-	p.E121Q	-
	XRCC1	-	-	p.A121T
NER	*DDB1*	p.A9G	-	p.Y517C
	ERCC5	-	p.D387N	-
	LIG1	-	p.T311M	-
Other	*A20*	p.C627Ffs*44	p.P714S	-
	DNTT	-	p.R335W	p.V37I; p.V371I

Table 3: Somatic, novel germline and rare germline mutations found in DLBCLs by the SOLiD sequencing approach. The genes are arranged in pathways; however, some genes belong to more than one pathway. *LIG1* and *LIG3* are involved in BER and NER, and *PNKP* is involved in BER and NHEJ. As a member of the MRN complex, *RAD50* participates in all DSB repair pathways. Adapted from *J Exp Med. 2013 Aug 26;210(9):1729-42*.

4.1.3 Mutation frequency

WES data was available for 17 of the 22 DLBCL tumors involved in the study. This allowed the comparison of the mutation frequency between DNA repair genes and the entire coding genome. The mean frequency of somatic mutations in DNA repair genes was 4.21 mutations/Mb versus 3.15 mutations/Mb for the entire coding genome. Of note, the mutation frequency in DNA repair genes identified by the SOLiD method in the 22 DLBCL samples was 4.16 mutations/Mb which is comparable to the WES result.

Even though DNA repair genes do not appear to be significantly more targeted by somatic mutations as compared to the rest of the coding genome, approximately half of the DLBCL cohort was affected by somatic mutations in at least one DNA repair gene.

4.1.4 Cohort expansion for selected genes

Seven of the genes harboring mutations identified by the SOLiD method were selected for sequencing in expanded cohorts. The Sanger method was used for the sequencing of *CHEK2*, *PARP1* and MSH2 in 235, 130 and 94 DLBCLs respectively. This included the 22 patients that underwent SOLiD sequencing. The other four genes, *DDB1*, *XRCC5*, *DCLRE1C* and *MLH3* were sequenced in 48 DLBCL tumors using the 454 GS FLX+ Titanium sequencing platform.

In addition to the two variants identified by the SOLiD method, seventeen tumors that did not undergo SOLiD sequencing were found to harbor *CHEK2* mutations. These were novel mutations and rare germline variants. Among the variants, two were splicing mutations, one was a frameshift deletion and one was a premature stop codon. The rest were SNVs, most of which were predicted by the SIFT and Polyphen-2 algorithms to have a damaging effect on the protein function. In summary, 8% of the cohort had *CHEK2* mutations. Mutations in *CHEK2* may contribute to carcinogenesis by impairing the signal transmission from DDR sensor proteins to p53. Therefore, *CHEK2* mutations could compromise cell-cycle checkpoints and the induction of apoptosis following DNA damage (224). The sequencing of an additional 90 DNA samples derived from healthy blood donors did not find any *CHEK2* mutations other than two germline variants that have previously been found by genotyping. This further supports the association of *CHEK2* variants with the development of DLBCL.

The Sanger sequencing of *PARP1* in the extended cohort identified four additional mutations. Those were two rare germline SNVs, one novel premature stop codon and one novel SNV. In total, 5% of the cohort was carrying *PARP1* mutations.

MSH2 cohort expansion by the Sanger method revealed a novel SNV and a previously reported rare SNV. Overall, *MSH2* was mutated in 2% of our cohort.

The 454 GS FLX+ Titanium platforms identified two novel mutations in *DDB1*, one novel and one rare variant in both *MLH3* and *XRCC5* and one rare variant in *DCLRE1C*. In summary, the combined sequencing results found *DDB1*, *MLH3*, *XRCC5* and *DCLRE1C* to be mutated in 6%, 13% 6% and 4% of the tumors respectively.

4.1.5 Allelic imbalances

SOLiD sequencing data also allowed for the comparison of allelic ratios of heterozygous germline SNVs between tumors and paired normal controls. This approach found imbalances in the loci of *EXO1* (6 tumors), *PARP1* and *RPA1* (3 tumors each), *TP53BP1*, *ERCC6* and *ATR* (2 tumors each) and *MDC1* (1 tumor).

Analysis of STS markers in the loci of *EXO1*, *PARP1*, *MDC1* and *RPA1* across 28 DLBCL tumors and paired normal controls confirmed all but one of the allelic imbalances found by the SOLiD method. Moreover, it identified additional cases of allelic imbalances in these loci that were not detected through the analysis of the SOLiD data. Overall, the *RPA1*, *EXO1*, *MDC1* and *PARP1* loci were affected by allelic imbalances in 25%, 21%, 18% and 14% of the tumors respectively.

An attempt to assess any impact on the expression levels of these genes, found that *PARP1* is expressed at significantly lower levels in tumors with allelic imbalances in the *PARP1* locus. This decrease in expression might reflect a loss in copy numbers in the locus. This is in conjunction with the observation that high PARP1 activity resulting from high levels of DNA damage, is associated with increased cell death (225). Therefore, a lower expression of *PARP1* might represent a survival mechanism in these tumors. The expression levels of *RPA1*, *EXO1* and *MDC1* did not appear altered by the presence of allelic imbalance in their loci.

4.1.6 Microsatellite instability and mutational load

MSI analysis of four STR markers was performed in order to detect genetic instability induced by the presence of MMR defects. The results showed MSI in four of the 28

tumors analyzed. Interestingly, all four tumors harbored mutations in at least one MMR gene. In addition, in one sample with an *MSH2* mutation from which a paired normal control was not available, MSI analysis was performed using Promega's MSI analysis system employing five mononucleotide markers. MSI was detected in this sample as well bringing the total number of MMR-mutated samples affected by MSI to five. Mutations in the MMR pathway have previously been associated with MSI in other malignancies, but never before in DLBCL (34).

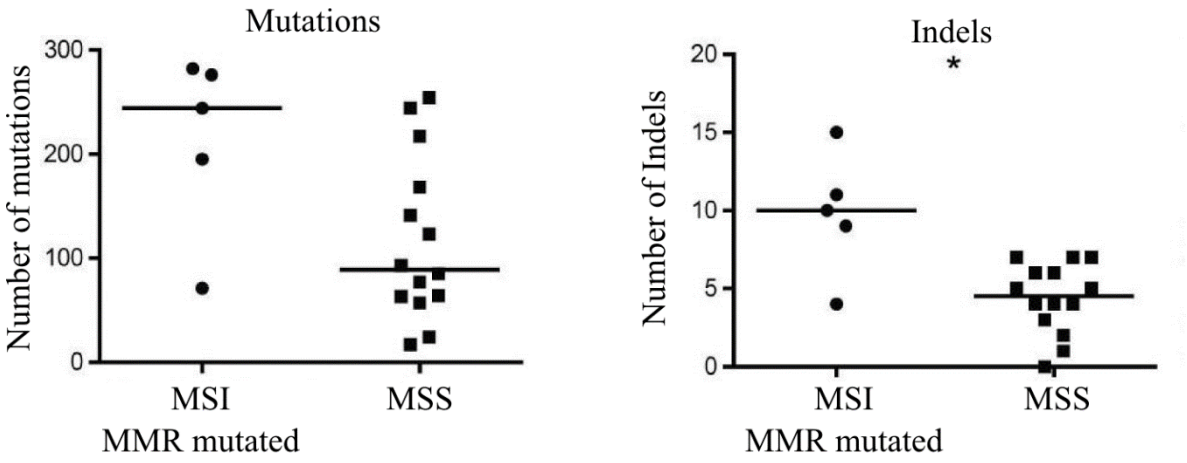

Figure 11: Increased mutation load in tumors with MSI/MMR mutations. Mutations in the MMR pathway were associated with MSI. The same tumors exhibited a higher mutational load as compared with the rest of the cohort, especially in the case of insertions and deletions. * P<0.05.

Moreover, MMR deficiencies have been associated with higher mutation loads in other human malignancies such as gastric carcinoma (226). WES data from the five MSI samples, 14 microsatellite stable (MSS) samples and their paired controls was used in order to investigate for associations between MMR defects and mutational load or mutation signatures. The MSI tumors were characterized by a heavier mutational load as compared to the MSS tumors even though statistical significance was marginally not attained (Mann-Whitney P=0.05). Interestingly however, the presence of indels was significantly higher in the MSI tumors (P=0.02) and three out of the four deletions identified by the SOLiD method were found in MMR-deficient tumors displaying MSI (Figure 11).

Furthermore, the MMR-mutated MSI tumors associated with a particular mutational signature as C:G>A:T transversions were significantly more common in these tumors. In B-cells, MMR is involved in the processing of mismatches induced by the mutagenic activity of AID during CSR and SHM and protects the genome from off-target AID activity (34, 52, 227-230). Therefore, the enrichment in C:G>A:T

transversions in these tumors might reflect a compromised surveillance of the AID activity due to MMR deficiencies. Accordingly, the presence of G:C>A:T transversions has previously been associated with MMR-deficiencies in mice, further supporting a role for MMR-deficiency in the etiology of DLBCL (231).

4.1.7 NHEJ defects and chromosomal translocations

Tissue sections were available from 13 DLBCL samples including three of the four tumors presented with NHEJ defects. Fluorescence *in situ* hybridization (FISH) was used in order to investigate whether these mutations associate with breakpoints in the locus of *IGH*. Dual-color break-apart probes were used to assess the status of the *IGH* and *BCL6* loci that are often found rearranged in DLBCL. The results found two samples with breaks in the *IGH* locus, both of which were harboring NHEJ mutations (*XRCC5* and *DCLRE1C*). None of the tumors with wild type NHEJ genes were affected by breaks in the *IGH* locus. This finding indicates that NHEJ defects might be involved in the formation of translocations between the *IGH* locus and proto-oncogenes in DLBCL. Of note, breaks in the *BCL6* locus were also observed in the two samples with *IGH* breaks; however, there was no further evidence supporting that the two loci were indeed juxtaposed to each other.

4.2 PAPER II

Exome sequencing reveals novel mutation targets in diffuse large B-cell lymphomas derived from Chinese patients.

In this paper, a WES approach was used to identify genes targeted by mutations in DLBCL and describe the frequency at which they are affected. The results revealed a number of novel and driver mutation targets for Chinese patients and underscored the genetic heterogeneity of DLBCL. Moreover, the functional implications of mutations in two of the most interesting gene targets, *DTX1* and *LYN*, were addressed experimentally.

4.2.1 Whole exome sequencing results

The coding genome of 31 Chinese DLBCL tumors and their normal blood controls were analyzed by WES. All variants reported below passed visual inspection by IGV. The number of somatic non-synonymous mutations identified in each tumor varied from 8 to 183 (median=54). In agreement with the findings of Paper I, the five tumors with somatic mutations in MMR genes presented heavier mutational loads than the remaining 26 tumors.

Gene	This study	Morin et al. (2011)	Pasqualucci et al. (2011)	Lohr et al. (2012)	Zhang et al. (2013)
CD70	23%	7%	-	10%	-
B2M	23%	8%	13%	10%	4%
MYD88	19%	13%	8%	12%	15%
TMSB4X	19%	-	-	-	-
TP53	16%	12%	17%	25%	5%
BTG2	16%	13%	-	2%	-
UBE2A	13%	-	17%	6%	-
SLITRK3	13%	-	-	-	-
IGLL5	13%	-	-	-	-
DCDC5	13%	-	-	-	-
DTX1	13%	-	-	6%	4%
LYN	10%	-	-	-	-
CD79B	10%	8%	12%	16%	8%
A20	10%	3%	15%	2%	10%
MLL2	6%	22%	23%	29%	-
BCL2	3%	28%	9%	23%	4%
EZH2	3%	22%	6%	14%	1%
CARD11	3%	10%	10%	20%	11%
MEF2B	-	15%	8%	18%	3%

Table 4: Gene mutation frequencies across different DLBCL NGS studies. Adapted from *Blood. 2014 Oct 16;124(16):2544-53.*

Non-synonymous mutations in at least 4 of the 31 tumors (~13%) were identified in 28 genes. These genes included common and previously described mutational targets in DLBCL, as well as the novel recurrently mutated genes *TMSB4X*, *DCDC5*, *IGLL5* and *SLITRK3*. In addition, a number of known targets in DLBCL were mutated at higher frequencies (*CD70*, *B2M*, *BTG2* and *DTX1)* or lower frequencies (*BCL2*, *EZH2*, *MEF2B* and *MLL2*) in our cohort as compared to similar studies focused on Western cohorts (Table 4) (132-134, 207). Differences in the mutation frequencies between our results and those of previous studies might reflect the exposure to

different environmental mutagens or a different germline background. *LYN* was found mutated in 3 of the 31 samples (10%) and was an additional target of interest due to its involvement in the BCR/NF-κB pathway's regulation. The BCR/NF-κB pathway is often affected by mutations in *MYD88*, *CARD11* and *CD79B* that lead to its constitutive activation in DLBCL. *LYN* however has not before been described as recurrently mutated in DLBCL.

The Molecular Signatures Database allowed for the assignment of genes carrying non-synonymous somatic mutations into canonical pathways. The most frequently affected processes included antigen processing and immune recognition, apoptosis, the BCR/NFkB pathway, chemokine and toll-like receptors, JAK/STAT signaling, cell cycle control, and Notch signaling.

4.2.2 Cohort expansion

Sanger sequencing was used for the sequencing of the coding regions of selected mutation targets in an expanded cohort of DLBCLs. *TMSB4X*, *CD70*, *B2M*, *DTX1*, *LYN* and *UBE2A* were analyzed across 136 Chinese DLBCL tumors including the 31 that underwent WES. *A20* and *TP53* were analyzed in 91 samples. The cohort expansion found the *CD70*, *TMSB4X*, *B2M*, *A20*, *DTX1*, *UBE2A*, *TP53* and *LYN* to be mutated in 22%, 16%, 13%, 14%, 12%, 10%, 9% and 8% of the cohort respectively.

4.2.3 *CD70* mutations

CD70 is a surface signaling molecule involved in the regulation of immunity by binding to CD27, its receptor on T-cells (193). In our cohort, *CD70* was the most frequently mutated gene. Remarkably, approximately 50% of the *CD70* mutations identified here resulted in premature stop codons or were located on splice sites, suggesting a deleterious effect for those variants. Moreover, WGS data from 10 samples identified copy number losses in the *CD70* locus in three samples. Taken together, this data suggests that *CD70* mutations contribute to the lymphomagenesis of DLBCL by disrupting anti-tumor T-cell activity.

4.2.4 *DTX1* mutations

DTX1 acts as negative regulator of the Notch signaling pathway (223). In a subset of DLBCL, Notch signaling is activated by gain-of-function mutations in *NOTCH1* and *NOTCH2* (232). However, in our WES sequencing effort only one *NOTCH1* and two

NOTCH2 mutations were identified. Screening of mutational hotspots of *NOTCH1* and *NOTCH2* in 91 DLBCLs found them to be mutated in 6% and 7% of the cohort respectively. *DTX1* was mutated in 12% of the tumors and was therefore the most frequently targeted component of the Notch signaling pathway in our cohort. Interestingly, *DTX1* and *NOTCH1* or *NOTCH2* mutations were in general not found together in the same patients (Table 5). This further suggests that *DTX1* mutations are an additional mechanism of Notch signaling activation in DLBCL and highlights the importance of the pathway in lymphomagenesis. Screening for *DTX1* mutations in a cohort of Swedish DLBCLs found alterations in 3 out of the 35 tumors (9%), a result comparable to that of the Chinese cohort.

Gene	Sample	Mutations
DTX1	DL19	p.G58D
	DL27	p.M10L; p.F82L
	DL29	p.H34R; p.A56T; p.G58D
	DL30	p.F17Y
	DL34	p.G7D; p.W30*
	DL42	p.E33D
	DL44	p.P11S; p.W30*; p.[T43_C45delinsS]; p.L69Q
	DL49	p.P18L; p.W30C; p.W30*; p.H34N; p.H34Y; p.Q77E
	DL52	p.P11fs*2
	DL64	p.W28*; p.I73V; p.I74fs*112; + splice site mutation
	DL72	p.G58D
	DL95	p.Q68H; p.Q81L
	DL103	p.P39R
	DL119	p.W30*
	DL121	p.S78C
	DL137	p.E54K
NOTCH1	DL40	p.H2241L
	DL41	p.Q2501*
	DL69	p.N2528K
	DL80	p.E2515*
	DL87	p.L1703Q
NOTCH2	DL14	p.Q2409*
	DL20	p.E2420K
	DL29	p.Q2285*
	DL62	p.D2163N
	DL93	p.R2400*
	DL98	p.P2157T

Table 5: *DTX1, NOTCH1* and *NOTCH2* non-synonymous mutations. Mutations in *DTX1* and *NOTCH1* or *NOTCH2* are generally mutually exclusive, with the sole exception of sample DL29. Adapted from *Blood. 2014 Oct 16;124(16):2544-53.*

All of the 43 *DTX1* mutations identified in the Chinese cohort were located in exon 1. Interestingly, many of them were G:C>A:T transversions found within WRC/GYW motifs which suggests aberrant AID targeting. However, FISH analysis in 35 tumors with *DTX1* mutations revealed no breaks in the locus of *DTX1* that would indicate translocations.

The impact of 14 of the mutants was assessed by means of a dual luciferase assay as described in the methods section. Four of the mutants demonstrated a significantly decreased inhibition of expression of the luciferase reporter gene, suggesting a functional consequence for those mutations. Importantly, those four mutations were present in 40% of the samples in which non-synonymous *DTX1* alterations were identified.

4.2.5 *LYN* mutations

LYN acts as a negative regulator of BCR signaling. It is a tyrosine kinase implicated in signal transduction through the phosphorylation of factors associated with the BCR complex and CD22 (233). Interestingly, studies in mice have shown an association between *LYN* deficiencies and autoimmune and myeloproliferative diseases (234, 235). Therefore, deleterious *LYN* mutations might interfere with the downregulation of BCR signaling and contribute to the constitutive activation of the NF-κB pathway (Figure 12).

The mutations identified in our cohort were located mostly in the protein's kinase domain which is responsible for the phosphorylation of downstream factors (234). Furthermore, approximately one third of the non-synonymous mutations led to premature stop-codons and truncated peptides.

To analyze the effect of 7 of these mutations at the protein level, U2OS cells were transfected with a mammalian expression vector carrying the wild type *LYN* or one of the 7 mutations. The cells were cultured for 48 hours before being analyzed by Western blot. As expected, no protein was detected in the cells transfected with *LYN* mutants harboring premature stop-codon mutations. In addition, reduction in the protein levels was observed in one of the missense mutants indicating that this particular mutation might lead to a less stable peptide. Further functional characterization of the missense mutations in the kinase domain could elucidate their impact on LYN's function as a regulator of BCR signaling.

Of note, *LYN* was also sequenced in 96 Swedish DLBCLs and the mutation rate in this cohort was only 2%, indicating a population bias.

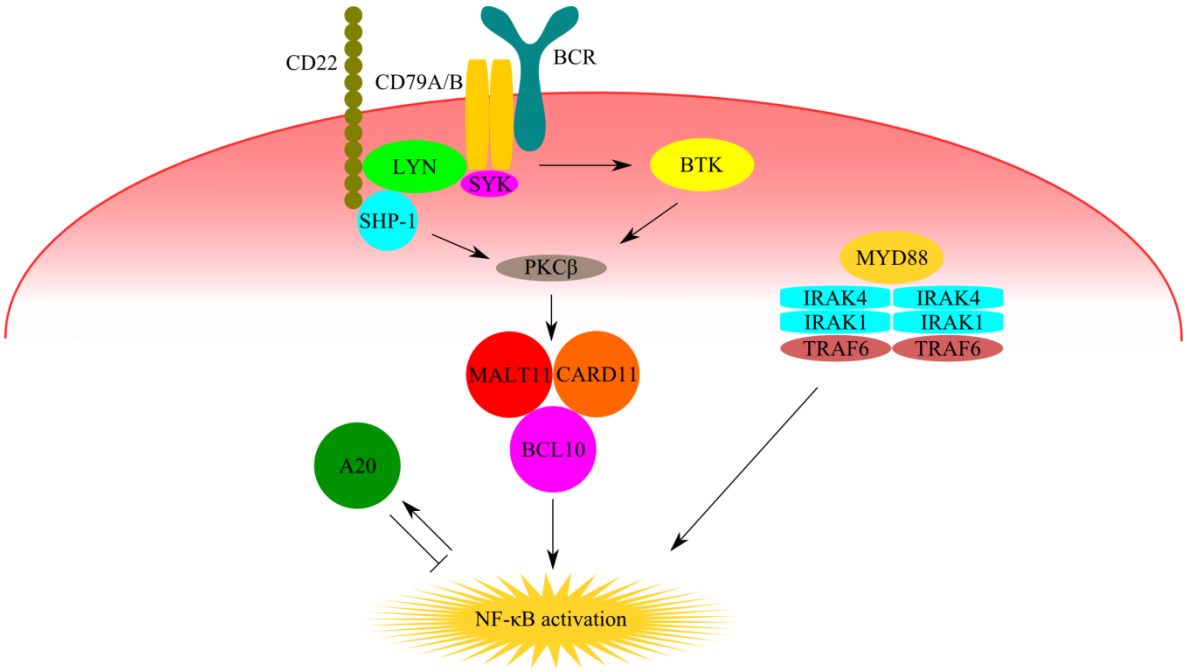

Figure 12: BCR/NF-κB signaling. Deleterious mutations in *CD79B*, *CARD11*, *MYD88* and *A20* are recurrent in DLBCL. *LYN* mutations could represent an additional mechanism contributing to the constitutive activation of the NF-κB pathway in DLBCL. Interestingly, as LYN is found upstream of BTK, targeting with BTK inhibitors could be envisioned as a treatment modality for these patients.

4.3 PAPER III

Genetic basis of PD-L1 overexpression in diffuse large B-cell lymphomas.

This project investigated the role of cytogenetic alterations as mechanisms of PD-1-ligand upregulation in DLBCL. Moreover, it described the distribution of those alterations across the disease subtypes and their impact on survival. Finally, it discussed how this could influence the treatment of DLBCL by therapies targeting the PD-1 – PD-L1/PD-L2 pathway.

4.3.1 Cytogenetic alterations, disease subtype and patient survival

The genes coding for the PD-1 ligands, PD-L1 and PD-L2, are located adjacent to each other on Chromosome 9-p.1. FISH probes targeting Chromosome 9-p.1 were used in order to assess the cytogenetic status of the locus across a cohort of 170 DLBCLs. The cohort was composed of 80 Chinese, 64 Swedish and 26 American patients. Twenty-two samples were presented with Chromosome 9-p.1 copy gains, six with locus breaks and two with locus amplification representing 13%, 4% and 1% of the cohort respectively. Locus amplification was defined as the presence of at least five copies of the locus per cell. Intriguingly, one of the samples presented with a break in the *PD-L1/PD-L2* locus was already known to be deficient for the NHEJ gene *DCLRE1C* and harbored breaks in the *IGH* locus. This was known from our previous work presented in Paper I. Collectively this information suggests that the translocation partner of *IGH* in this sample may be the *PD-L1/PD-L2* locus.

The molecular subtype of the disease was made known by the Han's algorithm for 122 of the samples. Forty-nine samples (40%) belonged to the GCB subtype whilst 73 (60%) belonged to the non-GCB subtype. The results showed that cytogenetic alterations affecting the *PD-L1/PD-L2* locus were more frequent in the non-GCB samples. Specifically, only 4 of the 49 GCB samples were found to have alterations in the locus vs 19 of the non-GCB samples. It is notable that all the locus breaks and amplifications were exclusively found in samples belonging to the non-GCB subtype. Data from previous studies suggests that *PD-L1* is expressed mostly by non-GCB DLBCLs (88). This prompts speculation that the higher expression of *PD-L1* in non-GCB DLBCL might, in part, be due to cytogenetic alterations.

A trend towards decreased survival was observed in patients harboring alterations in the *PD-L1/PD-L2* locus as compared with the remaining patients. More importantly however, the presence of breaks in the *PD-L1/PD-L2* locus showed a strong association with a rapid disease progression and a sharp decrease in patient survival. Statistical significance was observed even when the survival analysis was restricted to the non-GCB subtype. Therefore this decrease in survival is independent of the disease subtype.

4.3.2 Identification of translocation partners by WGS

To confirm and further characterize the locus breaks and amplifications, WGS was performed on five of the affected tumors from which DNA was available. Four tumors harbored locus breaks and one was presented with locus amplification. In three of the four samples with locus breaks, SV analysis revealed the exact location of the

breakpoints and identified the translocation partners. One tumor with a locus break was left unresolved probably due to DNA material degradation. The *PD-L1/PD-L2* locus was juxtaposed to the *Sμ* region of the *IGH* locus in two of the samples including the one with a *DCLRE1C* deficiency. Of note, the translocation in this particular sample was also demonstrated by a 4-color FISH experiment (Figure 13). In addition, breaks in the *PD-L1/PD-L2* and *IGH* loci were found in the unresolved case suggesting that the two loci may be juxtaposed in this sample as well. This data shows that the *PD-L1/PD-L2* locus is a recurrent *IGH* translocation partner in DLBCL. In the 4th sample, the identified translocation partner was *PIM1*, a gene recurrently mutated in DLBCL and an aberrant SHM target (236) (Figure 14). Finally, copy number variation analysis confirmed the amplification identified by FISH.

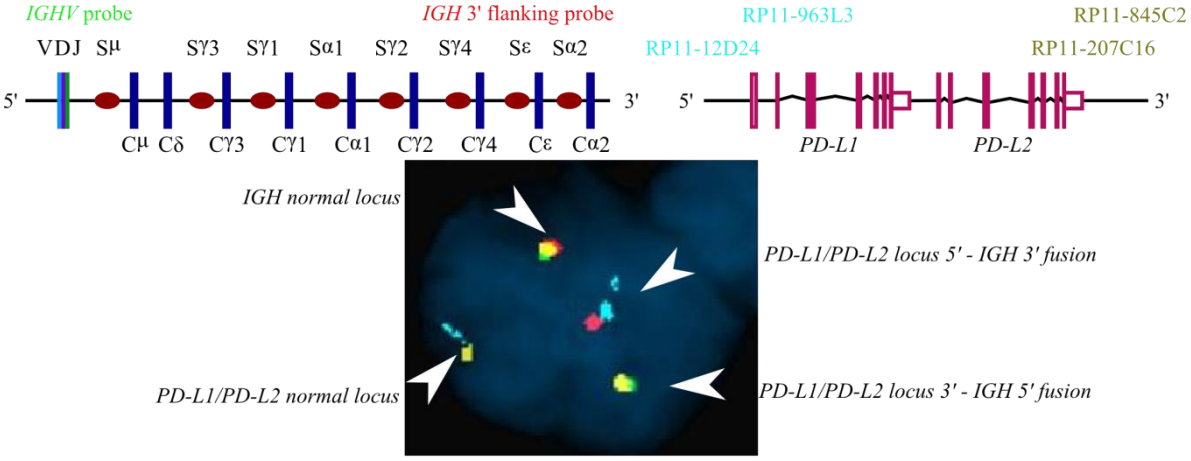

Figure 13: FISH analysis utilizing 4 colors. The rearrangement between the *PD-L1/PD-L2* and *IGH* loci is demonstrated here. The probes and the fluorophore's color used to label each probe are shown above the illustration of the loci.

4.3.3 *PD-L1* and *PD-L2* RNA expression analysis

RNA sequencing was available for 18 of the samples that underwent cytogenetic analysis by FISH. The median expression levels of PD-L1 were 9.4 FPKM. However, in two of the samples with suspected translocations between the *IGH* and *PD-L1/PD-L2* loci, the expression levels exceeded 40 FPKM suggesting that translocation to the *IGH* locus may lead to an upregulation or *PD-L1* expression. Collectively, a trend toward higher expression of *PD-L1* was observed in the samples harboring cytogenetic alterations, although, the results fell short of statistical significance probably due to the limited number of samples available. qPCR was performed in a larger number of samples (n=29) and found the expression of *PD-L1* to be

significantly higher in samples with alterations in the locus. Of note, no association was found between cytogenetic alteration and the expression levels of *PD-L2*.

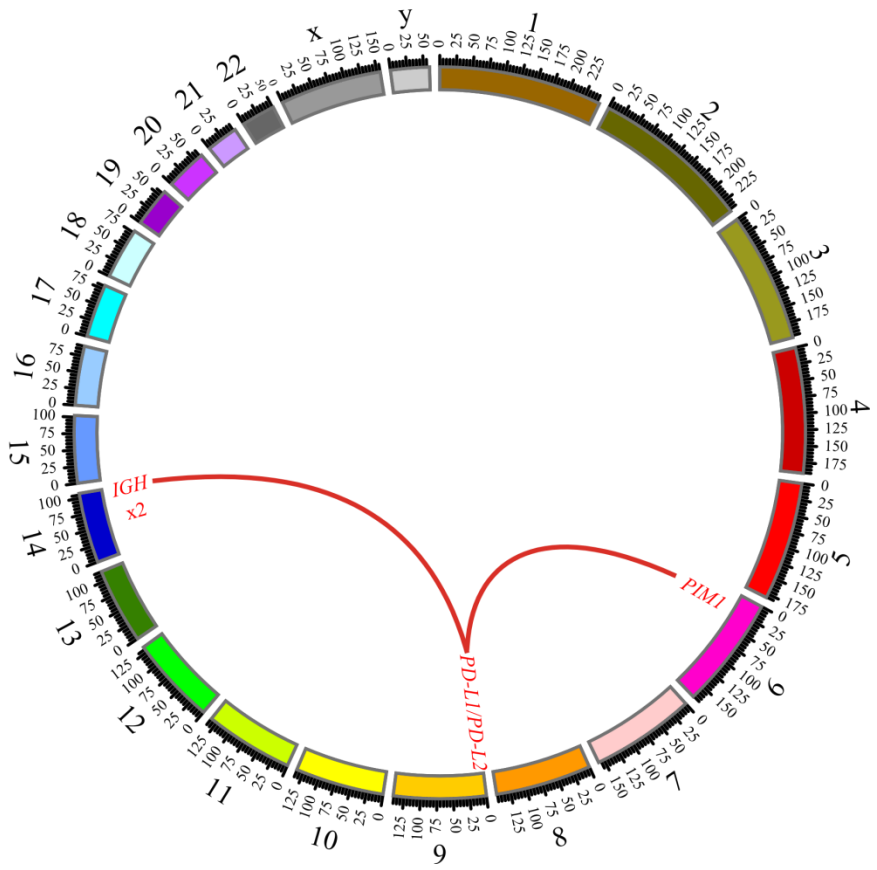

Figure 14: Circus plot showing the *PD-L1/PD-L2* translocation partners. The plot shows the locations of the concerned loci within the genome.

4.3.4 PD-L1 protein expression

PD-L1 immunostaining was performed on 167 of the samples and found 43 of them (26%) to be positive for PD-L1 protein expression. A strong association to cytogenetic alterations was found as 56% of the samples with copy gains, locus breaks or amplifications were expressing PD-L1 versus only 20% in the remaining samples. Notably, all samples with breaks/translocations and amplifications in the *PD-L1/PD-L2* locus were found to be positive for PD-L1 protein expression. This data provides further evidence that cytogenetic alterations, especially translocations and amplifications, constitute mechanisms of PD-L1 overexpression in DLBCL. Moreover, along with the decreased survival observed in patients with translocations, it suggests that a number of patients with aggressive DLBCL disease could benefit from therapies targeting the PD-1 – PD-L1/PD-L2 axis.

4.3.5 HLA class I expression

HLA class I is essential for antigenic presentation and recognition by cytotoxic $CD8^+$ T-cells. In most cancers, loss of HLA class I expression masks tumors from cytotoxic T-cells regardless of the expression status of PD-L1. The expression of the HLA class I complex was assessed in 90 tumors from the cohort and 39% were found to be negative. Contrary to expectation, almost all of the samples with breaks or amplifications in the *PD-L1/PD-L2* locus were negative for HLA class I. This could perhaps be attributed to the particular physiological function of B-cells. As professional APCs, B-cells express HLA class II which presents antigenic determinants to $CD4^+$ helper T-cells. Therefore, overexpression of PD-L1 in DLBCL, even in the absence of HLA class I expression, might confer malignant cells with further protection from immunity via an HLA class II dependent mechanism (Figure 15).

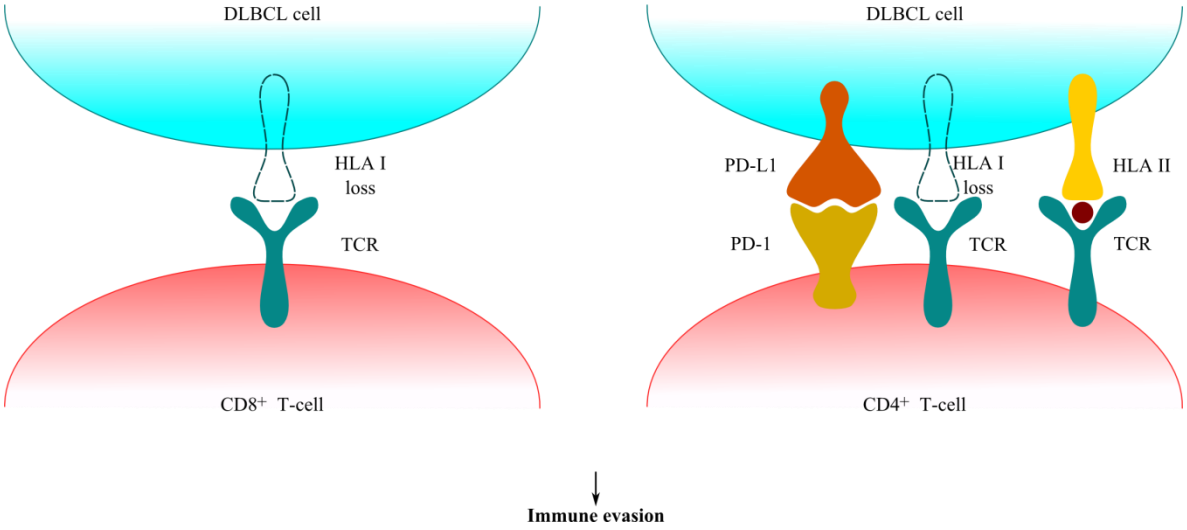

Figure 15: Immune evasion. Loss of HLA I suffice for immune escape from $CD8^+$ T-cells. However, as DLBCLs express HLA class II, expression of *PD-L1* can shield them from $CD4^+$ T-cell-mediated immune responses.

5 CONCLUDING REMARKS AND PERSPECTIVES

This thesis aimed to characterize the genome of DLBCL and describe novel mutation targets and structural alterations that contribute to the lymphomagenesis of DLBCL. Furthermore, this thesis attempted to explore how the knowledge generated by sequencing can contribute to new approaches in the treatment of DLBCL.

The key findings of this work are:

- Genes participating in DNA repair and antibody diversification are commonly affected by somatic and germline mutations in DLBCLs. In summary, 50% of the tumors in our cohort carry somatically acquired mutations in at least one DNA repair gene. Moreover, mutations in certain pathways can be associated with specific phenotypes. More specifically, MMR deficiencies translate into MSI and an increased mutational load, and NHEJ defects may predispose B-cells to chromosomal translocations involving the *IGH* locus due to an improper resolution of DSB.

- DLBCL is a heterogeneous disease characterized by a diverse mutational landscape and the incidence of mutations in certain genes varies across populations. A notable example is *CD70* which was found to be mutated at a much higher frequency in our cohort of Chinese patients, as compared to previous studies focusing on DLBCLs from Western populations.

- *DTX1* is recurrently targeted by mutations in DLBCLs and is the most commonly mutated member of the Notch pathway in Chinese DLBCLs. These mutations might be the result of illicit AID activity as they are mostly G:C>A:T transversions concentrated in WRC/GYW motifs. Importantly, a number of these mutations appear to have a deleterious effect on the function of *DTX1* as a negative regulator of Notch.

- *LYN* is a novel mutational target in DLBCL. Most of the mutations identified in our cohort are located in the protein's kinase domain, thereby suggesting a deleterious effect on protein function. Moreover, they often lead to premature stop-codons. Therefore, *LYN* deficiencies might represent an additional mechanism for constitutive activation of the BCR/NF-κB pathway.

- Cytogenetic alterations involving the *PD-L1/PD-L2* locus, especially translocations, constitute an intrinsic mechanism for PD-L1 expression in DLBCLs. These structural variations were mostly found in the more aggressive non-GCB subtype. A sharp decrease in survival was observed in patients with *PD-L1/PD*-L2 translocations and this was independent from the disease subtype. Finally, the *PD-L1/PD-L2* locus is described as a novel and recurrent *IGH* translocation partner in DLBCL.

Collectively, the results of this thesis show that a considerable proportion of the DLBCL patients are affected by deficiencies in pathways that are potentially targetable by therapies that are currently available or in development (Table 6). Therefore, a careful evaluation of patients and molecular characterization of each case can optimize therapy. This, in conjunction with the new developments in pharmacogenetics, could greatly improve disease outcomes.

Molecule class	**Target**	**Findings in the study cohort**
PD-1/PD-L1 therapeutic antibodies	PD-1 – PD-L1 pathway	PD-L1 overexpression
CD27 agonists	CD70 – CD27 pathway	*CD70* mutations & copy losses
Notch signaling factor inhibitors	Notch signaling	*DTX1* mutations
BTK inhibitors	BCR/NF-κB pathway	*LYN* mutations
PARP1 inhibitors	HR deficiencies	Mutations in HR factors
NEDD8 activation inhibitors	Ubiquitin-proteasome system	*UBE2A* mutations
EGFR/HER2 blockade	EGFR tyrosine kinases	*BTG2* mutations

Table 6: Defects identified in DLBCL tumors and potentially appropriate therapies that could benefit patients.

The expression of PD-L1 primarily by the non-GCB DLBCLs indicates that patients could be stratified and therapies targeting PD-1 or PD-L1 could be directed to patients with aggressive disease. Recently, a number of biomarkers that predict the response rate to therapies targeting the PD-1 – PD-L1 interaction have been identified. Expression of PD-L1 by the tumor cells or the tumor microenvironment has been associated to better responses. Intriguingly, however, even PD-L1-negative tumors have been found to respond to anti-PD-1 or anti-PD-L1 blockade to some extent (77, 237). In addition, better responses have been observed in tumors infiltrated by $CD8^+$ T-lymphocytes (238).

Interestingly, one of the DLBCLs enrolled in this study harbored a *PD-L1 – IGH* translocation and was found to be mutated for the NHEJ factor *DCLRE1C* and the MMR factor *MSH2*. MMR is involved in the processing of AID-induced mismatches in CSR, and NHEJ resolves the ensuing DSB (17, 227). It is therefore possible that the *PD-L1 – IGH* translocation in this sample may be the result of this combination of NHEJ and MMR mutations. Moreover, this is further supported by the fact that the breakpoint within the *PD-L1* gene is located within a WRC/GYW motif. In addition, this tumor, along with others harboring MMR deficiencies was presented with a heavier mutational load. Literature suggests that a higher number of mutations leads to

a rise in the number of neo-antigens rendering tumors more immunogenic (75). Studies have also suggested that as a consequence, tumors are more likely to respond to immunotherapy (239, 240). Taken together, this indicates that MMR mutations could constitute a biomarker for the selection of patients likely to respond to therapies targeting the PD-1 – PD-L1 axis.

The high prevalence of *CD70* mutations in our cohort indicates that immunotherapies targeting the CD70 – CD27 axis may also benefit DLBCL patients. Loss of CD70 might confer the tumor with an immune-suppressing activity by impeding CD27-mediated T-cell activation. Therefore, the use of CD27-activating antibodies may restore T-cell function in CD70 deficient patients (194).

LYN, a member of the BCR signaling pathway was found to be mutated in a considerable subset of the enrolled patients. The consequences of *LYN* mutations should be carefully investigated as they could contribute to the chronic activation of the NF-κB pathway. Follow-up studies could investigate the impact of *LYN* mutations on its kinase activity and how *LYN* deficiencies affect the expression of the NF-κB pathway factors. Moreover, the use of inhibitors of downstream effector molecules such as BTK could be envisioned for those patients.

DTX1 was mutated in 12% of our cohort. Together with other studies reporting mutations in *NOTCH1* and *NOTCH2*, they support that therapies targeting the Notch signaling pathway could also prove beneficial for DLBCL patients.

Additionally, *PARP1* was found mutated or expressed at lower levels in a subset of our cohort. Targeting of PARP1 in tumors deficient for HR leads to synthetic lethality and cell death. However, as a number of DLBCLs are already deficient for PARP1, it is likely that they will be insensitive to such therapies and patients should therefore be carefully evaluated prior to treatment.

The considerable efforts that have been made in recent years to characterize the genome of DLBCL have identified the most commonly mutated genes. The work conducted for this thesis reported a number of previously unknown mutational targets and underlined the heterogeneous genetic landscape of DLBCL. However, the collective results from this and previous studies prompt further investigation in order to achieve a comprehensive understanding of the DLBCL genome.

Down-sampling analysis of available data suggested that the genome of 2000 tumors must be sequenced in order to reach gene discovery saturation in DLBCL. It is estimated that this will identify 90% of the genes that are mutated in as low as 2% of

the patients (241). This is instrumental to the transition towards personalized medicine and a targeted therapy approach for the treatment of cancer. A profound knowledge of the disease landscape can guide the efforts for the development of novel drug molecules. Furthermore, in conjunction with patient screening upon diagnosis, it can provide individual patients with a treatment tailored to the genetic alterations characterizing their disease.

Our group is currently working towards this goal in collaboration with the Beijing Genomics Institute in order to develop a patient screening chip. The purpose of this effort is to provide a standardized and rapid diagnostic tool at a much lower cost than the currently available sequencing methods. It will enable the screening of newly-diagnosed or relapsed patients for defects in common mutational targets and assess the sensitivity of their disease to currently available treatments. Therefore, it will enable a more careful evaluation of each patient, provide a more accurate prediction of the disease outcome and direct a personalized selection of treatment. Patients could thus benefit from targeted therapies that are currently difficult to implement.

6 ACKNOWLEDGEMENTS

During this long journey toward a PhD, I came across many people I want to acknowledge for their contributions and help. In particular,

I would like to thank my main supervisor, Qiang Pan-Hammarström, for offering me a position in her lab, for closely following my progress and for providing all the technical and financial means necessary for the completion of this. Pan, I would also like to thank you for the supervision, your dedication to our work, the discussions and the valuable life-lessons.

My friend and co-supervisor Noel deMiranda for his support, supervision and guidance. I look up to you and admire you for your scientific talent, knowledge and clarity of spirit. I am also grateful for all the good memories, the long night discussions, the dinners and bottles of wine we shared together.

My co-supervisor Klas Wiman for his scientific advice and for always being available whenever I needed him.

Our divisional administrators Moa Thorin and Pernilla Klyve for doing such an excellent job and for having the answer to all the questions I ever had!

Rozina Caridha and Kerstin Bergman, our two amazing lab managers for making my work in the lab so much easier and well organized!

Apostolos Zaravinos, Weicheng Ren and Chenglin Wu for all their help in the lab and for being excellent coworkers in the cancer project.

I would like to thank Longyun Chen for the bioinformatics support and all the data analysis, without whom this work would not be possible.

The head of division, Lennart Hammarström, for his valuable comments and for helping me revise an ethical permit application.

A big thank you to Mattias Berglund for the great collaboration, the nice discussions and all the good work!

Susana Lisboa for welcoming me at the Portuguese Institute of Oncology and for introducing me to confocal microscopy and the FISH method.

Andrej Weintraub, Arja Kramsu and Marita Ward for making administrative issues as easy as they could ever be!

All through these years, I was lucky to be surrounded by good friends and colleagues. I am thankful for all the good times, the laughs, the activities, the meals and drinks we shared. Georgia Kokaraki, Kiki Liadaki, Åsa Wiktorsson, Amin Otmani, Marcel Franckowiak, Jesabel Varadé, Linkun Du, Yin Lin, Mahya Dezfouli, Annica Lindkvist, Omar Alkhairy, Che Kang Lim, Sofia Appelberg, Kasper Krogh-Andersen, Harold Marcotte, Xiaogan Wang, Fanglei Zuo, Renée Enqvist, Hassan Abolhassani, Radhika Kamdar, and Crónan Morrison, you really made my days brighter!

I am grateful to Rosa Romano for reviewing my thesis, showing me the best pizzeria in Stockholm (she is a true Napolitan) and for her choice of backdrop music in the lab.

My neighbor in the office, Jovanka King, thank you for proofreading my thesis and for the Vegemite experience!

Many thanks to my good friends and colleagues Ning Wang, Gökçe Günaydin, Margarita Bartish and Andrea Björkman for offering me all the help I ever asked during these years and for being such excellent people to be with both in the lab and out in the city.

My good old friend Thomas, our friendship is many years strong and even international. We created great memories during our two years in the amphitheaters of the University of Montpellier and the four years in the corridors of Huddinge Sjukhus. I am thankful for all the support you offered to me during my PhD, the good times, the coffees, the cigarettes and other intoxicating substances (all of them legal) but above everything I am happy…Correction, I am honored…. That you are my friend.

My thanks to you Thanos for being one of the sweetest people I ever met. Many times during my studies you had to encourage me and you always succeeded. Certainly, you did a lot for me….but I like to believe that our friendship is mutually beneficial since I can take credit for my scissors ridding you of that pony-tail hair (hallelujah)!!!

Chen and Linus, the dynamic-duo that I shared many Saturday wine-evenings with. I am particularly grateful to you for the unforgettable times in the streets, city-walls and cafeterias of Xi'an. We have many nights to remember together and I am sure there will be more to follow.

Special thanks to you Mr. Velzevoulis for being an ever so special person and for showing me some of the most arresting landscapes I have ever seen. You transformed my short PhD breaks into experiences of a lifetime. Also thank you for your contributions in saving my time by keeping me away from social media in the lab. With the kind of posts you make it is never a safe decision to go on Facebook while my screen is visible in public.

Tzela and Kosta, I want to thank you for being two truly inspirational persons in my life. Tzela, your pragmatic approach in life and your strong personality were examples to me. Kosta, thank you for infusing me with scientific curiosity since I was a little kid. You first introduced me to the wonders of the natural sciences, from the tiny cells to the huge galaxies. The books, the museum visits, the observatories and the long discussions under the starry skies of Greece, they all played their part. Perhaps unwillingly, you truly shaped my life.

Fireball!!! I thank you for reviewing my thesis…But above all I am grateful to you for your intangible contributions; all the jolly times in Stockholm and abroad, the times that you were here to lift my spirit, listen to me voice my frustrations and help me through them… Assisted at times by a glass of wine or two. I am so happy that you are part of my life, and in fact that you are central to it. I can't tell you how thrilled I am as we set off together, following the cheerful call of the Pet-Shop Boys!

7 REFERENCES

1. Stratton MR, Campbell PJ, Futreal PA (2009) The cancer genome. *Nature* 458(7239):719-724.
2. Helleday T, Eshtad S, Nik-Zainal S (2014) Mechanisms underlying mutational signatures in human cancers. *Nat Rev Genet* 15(9):585-598.
3. Friedberg EC, *et al.* (2006) DNA repair: from molecular mechanism to human disease. *DNA Repair* 5(8):986-996.
4. Harper JW, Elledge SJ (2007) The DNA damage response: ten years after. *Mol Cell* 28(5):739-745.
5. Ciccia A, Elledge SJ (2010) The DNA damage response: making it safe to play with knives. *Mol Cell* 40(2):179-204.
6. Jackson SP, Bartek J (2009) The DNA-damage response in human biology and disease. *Nature* 461(7267):1071-1078.
7. Campisi J, d'Adda di Fagagna F (2007) Cellular senescence: when bad things happen to good cells. *Nat Rev Mol Cell Biol* 8(9):729-740.
8. Shiloh Y (2003) ATM and related protein kinases: safeguarding genome integrity. *Nat Rev Cancer* 3(3):155-168.
9. Ferguson DO, Alt FW (2001) DNA double strand break repair and chromosomal translocation: lessons from animal models. *Oncogene* 20(40):5572-5579.
10. Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 79:181-211.
11. van den Bosch M, Lohman PH, Pastink A (2002) DNA double-strand break repair by homologous recombination. *Biol Chem* 383(6):873-892.
12. Shiloh Y, Ziv Y (2013) The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. *Nat Rev Mol Cell Biol* 14(4):197-210.
13. Escribano-Diaz C, *et al.* (2013) A cell cycle-dependent regulatory circuit composed of 53BP1-RIF1 and BRCA1-CtIP controls DNA repair pathway choice. *Mol Cell* 49(5):872-883.
14. Li X, Heyer WD (2008) Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res* 18(1):99-113.
15. Khanna KK, Jackson SP (2001) DNA double-strand breaks: signaling, repair and the cancer connection. *Nat Genet* 27(3):247-254.
16. Dudas A, Chovanec M (2004) DNA double-strand break repair by homologous recombination. *Mutat Res* 566(2):131-167.

17. Kotnis A, Du L, Liu C, Popov SW, Pan-Hammarstrom Q (2009) Non-homologous end joining in class switch recombination: the beginning of the end. *Philos Trans R Soc Lond B Biol Sci* 364(1517):653-665.

18. Lieber MR, Lu H, Gu J, Schwarz K (2008) Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. *Cell Res* 18(1):125-133.

19. Rooney S, Chaudhuri J, Alt FW (2004) The role of the non-homologous end-joining pathway in lymphocyte development. *Immunol Rev* 200:115-131.

20. Franco S, *et al.* (2008) DNA-PKcs and Artemis function in the end-joining phase of immunoglobulin heavy chain class switch recombination. *J Exp Med* 205(3):557-564.

21. Du L, *et al.* (2008) Involvement of Artemis in nonhomologous end-joining during immunoglobulin class switch recombination. *J Exp Med* 205(13):3031-3040.

22. Lieber MR, Gu J, Lu H, Shimazaki N, Tsai AG (2010) Nonhomologous DNA end joining (NHEJ) and chromosomal translocations in humans. *Subcell Biochem* 50:279-296.

23. Boboila C, *et al.* (2010) Alternative end-joining catalyzes class switch recombination in the absence of both Ku70 and DNA ligase 4. *J Exp Med* 207(2):417-427.

24. Wang M, *et al.* (2006) PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res* 34(21):6170-6182.

25. Frit P, Barboule N, Yuan Y, Gomez D, Calsou P (2014) Alternative end-joining pathway(s): bricolage at DNA breaks. *DNA Repair* 17:81-97.

26. Mateos-Gomez PA, *et al.* (2015) Mammalian polymerase theta promotes alternative NHEJ and suppresses recombination. *Nature* 518(7538):254-257.

27. Krokan HE, Standal R, Slupphaug G (1997) DNA glycosylases in the base excision repair of DNA. *Biochem J* 325 (Pt 1):1-16.

28. Demple B, Sung JS (2005) Molecular and biological roles of Ape1 protein in mammalian base excision repair. *DNA Repair* 4(12):1442-1449.

29. Ellenberger T, Tomkinson AE (2008) Eukaryotic DNA ligases: structural and functional insights. *Annu Rev Biochem* 77:313-338.

30. Nazarkina ZK, Khodyreva SN, Marsin S, Lavrik OI, Radicella JP (2007) XRCC1 interactions with base excision repair DNA intermediates. *DNA Repair* 6(2):254-264.

31. Huang JC, Hsu DS, Kazantsev A, Sancar A (1994) Substrate spectrum of human excinuclease: repair of abasic sites, methylated bases, mismatches, and bulky adducts. *Proc Natl Acad Sci U S A* 91(25):12213-12217.

32. Rouillon C, White MF (2011) The evolution and mechanisms of nucleotide excision repair proteins. *Res Microbiol* 162(1):19-26.

33. Kolodner RD, Marsischky GT (1999) Eukaryotic DNA mismatch repair. *Curr Opin Genet Dev* 9(1):89-96.

34. Jiricny J (2006) The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* 7(5):335-346.

35. Rodriguez RM, Lopez-Vazquez A, Lopez-Larrea C (2012) Immune systems evolution. *Adv Exp Med Biol* 739:237-251.

36. Travis J (2009) Origins. On the origin of the immune system. *Science* 324(5927):580-582.

37. Ricklin D, Hajishengallis G, Yang K, Lambris JD (2010) Complement: a key system for immune surveillance and homeostasis. *Nat Immunol* 11(9):785-797.

38. Hoffmann J, Akira S (2013) Innate immunity. *Curr Opin Immunol* 25(1):1-3.

39. Litman GW, Rast JP, Fugmann SD (2010) The origins of vertebrate adaptive immunity. *Nat Rev Immunol* 10(8):543-553.

40. Iwasaki A, Medzhitov R (2015) Control of adaptive immunity by the innate immune system. *Nat Immunol* 16(4):343-353.

41. Miller JF (2011) The golden anniversary of the thymus. *Nat Rev Immunol* 11(7):489-495.

42. Zhang N, Bevan MJ (2011) CD8(+) T cells: foot soldiers of the immune system. *Immunity* 35(2):161-168.

43. Luckheeram RV, Zhou R, Verma AD, Xia B (2012) CD4(+)T cells: differentiation and functions. *Clin Dev Immunol* 2012:925135.

44. LeBien TW, Tedder TF (2008) B lymphocytes: how they develop and function. *Blood* 112(5):1570-1580.

45. Cooper MD (2015) The early history of B cells. *Nat Rev Immunol* 15(3):191-197.

46. Ravetch JV, Bolland S (2001) IgG Fc receptors. *Annu Rev Immunol* 19:275-290.

47. Silverstein AM (2003) Splitting the difference: the germline-somatic mutation debate on generating antibody diversity. *Nat Immunol* 4(9):829-833.

48. Oettinger MA, Schatz DG, Gorka C, Baltimore D (1990) RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248(4962):1517-1523.

49. Helmink BA, Sleckman BP (2012) The response to and repair of RAG-mediated DNA double-strand breaks. *Annu Rev Immunol* 30:175-202.

50. Malu S, Malshetty V, Francis D, Cortes P (2012) Role of non-homologous end joining in V(D)J recombination. *Immunol Res* 54(1-3):233-246.

51. Jung D, Giallourakis C, Mostoslavsky R, Alt FW (2006) Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* 24:541-570.

52. Di Noia JM, Neuberger MS (2007) Molecular mechanisms of antibody somatic hypermutation. *Annu Rev Biochem* 76:1-22.

53. Lord CJ, Ashworth A (2012) The DNA damage response and cancer therapy. *Nature* 481(7381):287-294.

54. Sahai E (2005) Mechanisms of cancer cell invasion. *Curr Opin Genet Dev* 15(1):87-96.

55. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* 144(5):646-674.

56. Croce CM (2008) Oncogenes and cancer. *N Engl J Med* 358(5):502-511.

57. Sun W, Yang J (2010) Functional mechanisms for human tumor suppressors. *J Cancer* 1:136-140.

58. Hollstein M, Sidransky D, Vogelstein B, Harris CC (1991) p53 mutations in human cancers. *Science* 253(5015):49-53.

59. Levine AJ, Momand J, Finlay CA (1991) The p53 tumour suppressor gene. *Nature* 351(6326):453-456.

60. Muller PA, Vousden KH (2013) p53 mutations in cancer. *Nat Cell Biol* 15(1):2-8.

61. Sherr CJ, McCormick F (2002) The RB and p53 pathways in cancer. *Cancer Cell* 2(2):103-112.

62. Liggett WH, Jr., Sidransky D (1998) Role of the p16 tumor suppressor gene in cancer. *J Clin Oncol* 16(3):1197-1206.

63. Silver DP, Livingston DM (2012) Mechanisms of BRCA1 tumor suppression. *Cancer Discov* 2(8):679-684.

64. Anonymous (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* 474(7353):609-615.

65. Deshpande A, Sicinski P, Hinds PW (2005) Cyclins and cdks in development and cancer: a perspective. *Oncogene* 24(17):2909-2915.

66. Yuan TL, Cantley LC (2008) PI3K pathway alterations in cancer: variations on a theme. *Oncogene* 27(41):5497-5510.

67. Adams JM, Cory S (2007) The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* 26(9):1324-1337.

68. Polakis P (2012) Wnt signaling in cancer. *Cold Spring Harb Perspect Biol* 4(5).

69. Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D (2011) RAS oncogenes: weaving a tumorigenic web. *Nat Rev Cancer* 11(11):761-774.

70. Dang CV (2012) MYC on the path to cancer. *Cell* 149(1):22-35.

71. Hanahan D, Folkman J (1996) Patterns and emerging mechanisms of the angiogenic switch during tumorigenesis. *Cell* 86(3):353-364.

72. Carmeliet P (2005) VEGF as a key mediator of angiogenesis in cancer. *Oncology* 69 Suppl 3:4-10.

73. Vander Heiden MG, Cantley LC, Thompson CB (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* 324(5930):1029-1033.

74. Bondar T, Medzhitov R (2013) The origins of tumor-promoting inflammation. *Cancer Cell* 24(2):143-144.

75. Yadav M, *et al.* (2014) Predicting immunogenic tumour mutations by combining mass spectrometry and exome sequencing. *Nature* 515(7528):572-576.

76. Pico de Coana Y, Choudhury A, Kiessling R (2015) Checkpoint blockade for cancer therapy: revitalizing a suppressed immune system. *Trends Mol Med* 21(8):482-491.

77. Buchbinder EI, Desai A (2015) CTLA-4 and PD-1 Pathways: Similarities, Differences, and Implications of Their Inhibition. *Am J Clin Oncol*.

78. Alegre ML, Frauwirth KA, Thompson CB (2001) T-cell regulation by CD28 and CTLA-4. *Nat Rev Immunol* 1(3):220-228.

79. Hodi FS, *et al.* (2010) Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med* 363(8):711-723.

80. Sondak VK, Smalley KS, Kudchadkar R, Grippon S, Kirkpatrick P (2011) Ipilimumab. *Nat Rev Drug Discov* 10(6):411-412.

81. Ascierto PA, *et al.* (2014) The role of immunotherapy in solid tumors: report from the Campania Society of Oncology Immunotherapy (SCITO) meeting, Naples 2014. *J Transl Med* 12:291.

82. Carosella ED, Ploussard G, LeMaoult J, Desgrandchamps F (2015) A Systematic Review of Immunotherapy in Urologic Cancer: Evolving Roles for Targeting of CTLA-4, PD-1/PD-L1, and HLA-G. *Eur Urol* 68(2):267-279.

83. Keir ME, Francisco LM, Sharpe AH (2007) PD-1 and its ligands in T-cell immunity. *Curr Opin Immunol* 19(3):309-314.

84. Freeman GJ, *et al.* (2000) Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. *J Exp Med* 192(7):1027-1034.

85. Keir ME, Butte MJ, Freeman GJ, Sharpe AH (2008) PD-1 and its ligands in tolerance and immunity. *Annu Rev Immunol* 26:677-704.

86. Hamanishi J, *et al.* (2007) Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer. *Proc Natl Acad Sci U S A* 104(9):3360-3365.

87. Gadiot J, *et al.* (2011) Overall survival and PD-L1 expression in metastasized malignant melanoma. *Cancer* 117(10):2192-2201.

88. Andorsky DJ, *et al.* (2011) Programmed death ligand 1 is expressed by non-hodgkin lymphomas and inhibits the activity of tumor-associated T cells. *Clin Cancer Res* 17(13):4232-4244.

89. Velcheti V, *et al.* (2014) Programmed death ligand-1 expression in non-small cell lung cancer. *Lab Invest* 94(1):107-116.

90. Brahmer JR, *et al.* (2012) Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med* 366(26):2455-2465.

91. Ansell SM, *et al.* (2015) PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma. *N Engl J Med* 372(4):311-319.

92. Hamanishi J, *et al.* (2015) Safety and Antitumor Activity of Anti-PD-1 Antibody, Nivolumab, in Patients With Platinum-Resistant Ovarian Cancer. *J Clin Oncol*.

93. Zandberg DP, Strome SE (2014) The role of the PD-L1:PD-1 pathway in squamous cell carcinoma of the head and neck. *Oral Oncol* 50(7):627-632.

94. Tykodi SS (2014) PD-1 as an emerging therapeutic target in renal cell carcinoma: current evidence. *Onco Targets Ther* 7:1349-1359.

95. Luke JJ, Ott PA (2015) PD-1 pathway inhibitors: the next generation of immunotherapy for advanced melanoma. *Oncotarget* 6(6):3479-3492.

96. Mardis ER (2011) A decade's perspective on DNA sequencing technology. *Nature* 470(7333):198-203.

97. Gagan J, Van Allen EM (2015) Next-generation sequencing to guide cancer therapy. *Genome Med* 7(1):80.

98. Campo E, *et al.* (2011) The 2008 WHO classification of lymphoid neoplasms and beyond: evolving concepts and practical applications. *Blood* 117(19):5019-5032.

99. Sabattini E, Bacci F, Sagramoso C, Pileri SA (2010) WHO classification of tumours of haematopoietic and lymphoid tissues in 2008: an overview. *Pathologica* 102(3):83-87.

100. Kuppers R (2005) Mechanisms of B-cell lymphoma pathogenesis. *Nat Rev Cancer* 5(4):251-262.

101. Carbone A, Gloghini A, Dotti G (2008) EBV-associated lymphoproliferative disorders: classification and treatment. *Oncologist* 13(5):577-585.

102. Beral V, Peterman T, Berkelman R, Jaffe H (1991) AIDS-associated non-Hodgkin lymphoma. *Lancet* 337(8745):805-809.

103. Singavi AK, Harrington AM, Fenske TS (2015) Post-transplant lymphoproliferative disorders. *Cancer Treat Res* 165:305-327.

104. Kelly JL, *et al.* (2010) Germline variation in apoptosis pathway genes and risk of non-Hodgkin's lymphoma. *Cancer Epidemiol Biomarkers Prev* 19(11):2847-2858.

105. Straus SE, *et al.* (2001) The development of lymphomas in families with autoimmune lymphoproliferative syndrome with germline Fas mutations and defective lymphocyte apoptosis. *Blood* 98(1):194-200.

106. Kuppers R, Hansmann ML (2005) The Hodgkin and Reed/Sternberg cell. *Int J Biochem Cell Biol* 37(3):511-517.

107. Fisher SG, Fisher RI (2004) The epidemiology of non-Hodgkin's lymphoma. *Oncogene* 23(38):6524-6534.

108. Connors JM (2015) I. Hodgkin lymphoma: special challenges and solutions. *Hematol Oncol* 33 Suppl 1:21-24.

109. Anderson JR, Armitage JO, Weisenburger DD (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.

110. Pasqualucci L, *et al.* (2014) Genetics of follicular lymphoma transformation. *Cell Rep* 6(1):130-140.

111. Hiddemann W, Cheson BD (2014) How we manage follicular lymphoma. *Leukemia* 28(7):1388-1395.

112. Cullen MH, Lister TA, Brearley RI, Shand WS, Stansfeld AG (1979) Histological transformation of non-Hodgkin's lymphoma: a prospective study. *Cancer* 44(2):645-651.

113. Casulo C, Burack WR, Friedberg JW (2015) Transformed follicular non-Hodgkin lymphoma. *Blood* 125(1):40-47.

114. Rossi D, *et al.* (2006) Aberrant somatic hypermutation in transformation of follicular lymphoma and chronic lymphocytic leukemia to diffuse large B-cell lymphoma. *Haematologica* 91(10):1405-1409.

115. Davies AJ, *et al.* (2007) Transformation of follicular lymphoma to diffuse large B-cell lymphoma proceeds by distinct oncogenic mechanisms. *Br J Haematol* 136(2):286-293.

116. Greenlee RT, Hill-Harmon MB, Murray T, Thun M (2001) Cancer statistics, 2001. *CA Cancer J Clin* 51(1):15-36.

117. Lossos IS, *et al.* (2002) Transformation of follicular lymphoma to diffuse large-cell lymphoma: alternative patterns with increased or decreased expression of c-myc and its regulated genes. *Proc Natl Acad Sci U S A* 99(13):8886-8891.

118. Pasqualucci L, Dalla-Favera R (2015) The genetic landscape of diffuse large B-cell lymphoma. *Semin Hematol* 52(2):67-76.

119. Gurbuxani S, Anastasi J, Hyjek E (2009) Diffuse large B-cell lymphoma--more than a diffuse collection of large B cells: an entity in search of a meaningful classification. *Arch Pathol Lab Med* 133(7):1121-1134.

120. Alizadeh AA, *et al.* (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403(6769):503-511.

121. Rosenwald A, *et al.* (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* 346(25):1937-1947.

122. Fiskvik I, *et al.* (2015) Combining MYC, BCL2 and TP53 gene and protein expression alterations improves risk stratification in diffuse large B-cell lymphoma. *Leuk Lymphoma* 56(6):1742-1749.

123. Rosenwald A, Staudt LM (2003) Gene expression profiling of diffuse large B-cell lymphoma. *Leuk Lymphoma* 44 Suppl 3:S41-47.

124. De Paepe P, De Wolf-Peeters C (2007) Diffuse large B-cell lymphoma: a heterogeneous group of non-Hodgkin lymphomas comprising several distinct clinicopathological entities. *Leukemia* 21(1):37-43.

125. Morin RD, *et al.* (2010) Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat Genet* 42(2):181-185.

126. Lenz G, *et al.* (2008) Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. *Science* 319(5870):1676-1679.

127. Davis RE, *et al.* (2010) Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* 463(7277):88-92.

128. Ngo VN, *et al.* (2011) Oncogenically active MYD88 mutations in human lymphoma. *Nature* 470(7332):115-119.

129. Pasqualucci L, *et al.* (2006) Inactivation of the PRDM1/BLIMP1 gene in diffuse large B cell lymphoma. *J Exp Med* 203(2):311-317.

130. Lossos IS, *et al.* (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.

131. Pasqualucci L, *et al.* (2011) Inactivating mutations of acetyltransferase genes in B-cell lymphoma. *Nature* 471(7337):189-195.

132. Lohr JG, *et al.* (2012) Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A* 109(10):3879-3884.

133. Morin RD, *et al.* (2011) Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* 476(7360):298-303.

134. Zhang J, *et al.* (2013) Genetic heterogeneity of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* 110(4):1398-1403.

135. Xu ZZ, *et al.* (2013) Activation of the PI3K/AKT/mTOR pathway in diffuse large B cell lymphoma: clinical significance and inhibitory effect of rituximab. *Ann Hematol* 92(10):1351-1358.

136. Scott DW, *et al.* (2012) TBL1XR1/TP63: a novel recurrent gene fusion in B-cell non-Hodgkin lymphoma. *Blood* 119(21):4949-4952.

137. Morin RD, *et al.* (2013) Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* 122(7):1256-1265.

138. Steidl C, *et al.* (2011) MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. *Nature* 471(7338):377-381.

139. Dalla-Favera R, Martinotti S, Gallo RC, Erikson J, Croce CM (1983) Translocation and rearrangements of the c-myc oncogene locus in human undifferentiated B-cell lymphomas. *Science* 219(4587):963-967.

140. Lenz G, *et al.* (2007) Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell-like diffuse large B cell lymphoma. *J Exp Med* 204(3):633-643.

141. Khodabakhshi AH, *et al.* (2012) Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget* 3(11):1308-1319.

142. Kuppers R, Dalla-Favera R (2001) Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* 20(40):5580-5594.

143. Ramiro AR, *et al.* (2004) AID is required for c-myc/IgH chromosome translocations in vivo. *Cell* 118(4):431-438.

144. Pasqualucci L, *et al.* (2008) AID is required for germinal center-derived lymphomagenesis. *Nat Genet* 40(1):108-112.

145. Raghavan SC, Swanson PC, Ma Y, Lieber MR (2005) Double-strand break formation by the RAG complex at the bcl-2 major breakpoint region and at other non-B DNA structures in vitro. *Mol Cell Biol* 25(14):5904-5919.

146. Raghavan SC, Swanson PC, Wu X, Hsieh CL, Lieber MR (2004) A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex. *Nature* 428(6978):88-93.

147. Challa-Malladi M, *et al.* (2011) Combined genetic inactivation of beta2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. *Cancer Cell* 20(6):728-740.

148. Gussow D, *et al.* (1987) The human beta 2-microglobulin gene. Primary structure and definition of the transcriptional unit. *J Immunol* 139(9):3132-3138.

149. Bolhuis RL, Roozemond RC, van de Griend RJ (1986) Induction and blocking of cytolysis in CD2+, CD3- NK and CD2+, CD3+ cytotoxic T lymphocytes via CD2 50 KD sheep erythrocyte receptor. *J Immunol* 136(11):3939-3944.

150. Wang JH, *et al.* (1999) Structure of a heterophilic adhesion complex between the human CD2 and CD58 (LFA-3) counterreceptors. *Cell* 97(6):791-803.

151. Rimsza LM, *et al.* (2004) Loss of MHC class II gene and protein expression in diffuse large B-cell lymphoma is related to decreased tumor immunosurveillance and poor patient survival regardless of other prognostic factors: a follow-up study from the Leukemia and Lymphoma Molecular Profiling Project. *Blood* 103(11):4251-4258.

152. Riemersma SA, *et al.* (2000) Extensive genetic alterations of the HLA region, including homozygous deletions of HLA class II genes in B-cell lymphomas arising in immune-privileged sites. *Blood* 96(10):3569-3577.

153. Brown PJ, *et al.* (2015) FOXP1 suppresses immune response signatures and MHC class II expression in activated B-cell-like diffuse large B-cell lymphomas. *Leukemia*.

154. Armitage JO (2012) My treatment approach to patients with diffuse large B-cell lymphoma. *Mayo Clin Proc* 87(2):161-171.

155. Sud R, Friedberg JW (2008) Salvage therapy for relapsed or refractory diffuse large B-cell lymphoma: impact of prior rituximab. *Haematologica* 93(12):1776-1780.

156. DeVita VT, Jr., *et al.* (1975) Advanced diffuse histiocytic lymphoma, a potentially curable disease. *Lancet* 1(7901):248-250.

157. Fisher RI, *et al.* (1993) Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. *N Engl J Med* 328(14):1002-1006.

158. Coiffier B, *et al.* (2002) CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *N Engl J Med* 346(4):235-242.

159. Pfreundschuh M, *et al.* (2011) CHOP-like chemotherapy with or without rituximab in young patients with good-prognosis diffuse large-B-cell lymphoma: 6-year results of an open-label randomised study of the MabThera International Trial (MInT) Group. *Lancet Oncol* 12(11):1013-1022.

160. Roschewski M, Staudt LM, Wilson WH (2014) Diffuse large B-cell lymphoma-treatment approaches in the molecular era. *Nat Rev Clin Oncol* 11(1):12-23.

161. Dunleavy K, *et al.* (2013) Dose-adjusted EPOCH-rituximab therapy in primary mediastinal B-cell lymphoma. *N Engl J Med* 368(15):1408-1416.

162. Gisselbrecht C, *et al.* (2010) Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. *J Clin Oncol* 28(27):4184-4190.

163. Eide MB, *et al.* (2011) High dose chemotherapy with autologous stem cell support for patients with histologically transformed B-cell non-Hodgkin lymphomas. A Norwegian multi centre phase II study. *Br J Haematol* 152(5):600-610.

164. Kochenderfer JN, *et al.* (2015) Chemotherapy-refractory diffuse large B-cell lymphoma and indolent B-cell malignancies can be effectively treated with autologous T cells expressing an anti-CD19 chimeric antigen receptor. *J Clin Oncol* 33(6):540-549.

165. Mian M, *et al.* (2014) A success story: how a single targeted-therapy molecule impacted on treatment and outcome of diffuse large B-cell lymphoma. *Anticancer Res* 34(5):2559-2564.

166. Yap DB, *et al.* (2011) Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. *Blood* 117(8):2451-2459.

167. Beguelin W, *et al.* (2013) EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. *Cancer Cell* 23(5):677-692.

168. Velichutina I, *et al.* (2010) EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. *Blood* 116(24):5247-5255.

169. McCabe MT, *et al.* (2012) EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. *Nature* 492(7427):108-112.

170. Qi W, *et al.* (2012) Selective inhibition of Ezh2 by a small molecule inhibitor blocks tumor cells proliferation. *Proc Natl Acad Sci U S A* 109(52):21360-21365.

171. Momparler RL, Cote S (2015) Targeting of cancer stem cells by inhibitors of DNA and histone methylation. *Expert Opin Investig Drugs* 24(8):1031-1043.

172. Wilson WH, *et al.* (2010) Navitoclax, a targeted high-affinity inhibitor of BCL-2, in lymphoid malignancies: a phase 1 dose-escalation study of safety, pharmacokinetics, pharmacodynamics, and antitumour activity. *Lancet Oncol* 11(12):1149-1159.

173. Souers AJ, *et al.* (2013) ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med* 19(2):202-208.

174. Cerchietti LC, *et al.* (2009) A peptomimetic inhibitor of BCL6 with potent antilymphoma effects in vitro and in vivo. *Blood* 113(15):3397-3405.

175. Cerchietti LC, *et al.* (2010) A small-molecule inhibitor of BCL6 kills DLBCL cells in vitro and in vivo. *Cancer Cell* 17(4):400-411.

176. Davis RE, Brown KD, Siebenlist U, Staudt LM (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.

177. Lim KH, Yang Y, Staudt LM (2012) Pathogenetic importance and therapeutic implications of NF-kappaB in lymphoid malignancies. *Immunol Rev* 246(1):359-378.

178. Compagno M, *et al.* (2009) Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* 459(7247):717-721.

179. Younes A, *et al.* (2014) Combination of ibrutinib with rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone (R-CHOP) for treatment-naive patients with CD20-positive B-cell non-Hodgkin lymphoma: a non-randomised, phase 1b study. *Lancet Oncol* 15(9):1019-1026.

180. LoPiccolo J, Blumenthal GM, Bernstein WB, Dennis PA (2008) Targeting the PI3K/Akt/mTOR pathway: effective combinations and clinical considerations. *Drug Resist Updat* 11(1-2):32-50.

181. Chen L, *et al.* (2013) SYK inhibition modulates distinct PI3K/AKT- dependent survival pathways and cholesterol biosynthesis in diffuse large B cell lymphomas. *Cancer Cell* 23(6):826-838.

182. Cheng S, *et al.* (2011) SYK inhibition and response prediction in diffuse large B-cell lymphoma. *Blood* 118(24):6342-6352.

183. Pfeifer M, *et al.* (2013) PTEN loss defines a PI3K/AKT pathway-dependent germinal center subtype of diffuse large B-cell lymphoma. *Proc Natl Acad Sci U S A* 110(30):12420-12425.

184. Lenz G, *et al.* (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.

185. Rodon J, Dienstmann R, Serra V, Tabernero J (2013) Development of PI3K inhibitors: lessons learned from early clinical trials. *Nat Rev Clin Oncol* 10(3):143-153.

186. Lannutti BJ, *et al.* (2011) CAL-101, a p110delta selective phosphatidylinositol-3-kinase inhibitor for the treatment of B-cell malignancies, inhibits PI3K signaling and cellular viability. *Blood* 117(2):591-594.

187. Petrich AM, *et al.* (2012) Akt inhibitors MK-2206 and nelfinavir overcome mTOR inhibitor resistance in diffuse large B-cell lymphoma. *Clin Cancer Res* 18(9):2534-2544.

188. Witzig TE, *et al.* (2011) A phase II trial of the oral mTOR inhibitor everolimus in relapsed aggressive lymphoma. *Leukemia* 25(2):341-347.

189. Helleday T, Bryant HE, Schultz N (2005) Poly(ADP-ribose) polymerase (PARP-1) in homologous recombination and as a target for cancer therapy. *Cell Cycle* 4(9):1176-1178.

190. Evers B, Helleday T, Jonkers J (2010) Targeting homologous recombination repair defects in cancer. *Trends Pharmacol Sci* 31(8):372-380.

191. Yang ZZ, *et al.* (2014) TGF-beta upregulates CD70 expression and induces exhaustion of effector memory T cells in B-cell non-Hodgkin's lymphoma. *Leukemia* 28(9):1872-1884.

192. Lens SM, *et al.* (1999) Aberrant expression and reverse signalling of CD70 on malignant B cells. *Br J Haematol* 106(2):491-503.

193. Nolte MA, van Olffen RW, van Gisbergen KP, van Lier RA (2009) Timing and tuning of CD27-CD70 interactions: the impact of signal strength in setting

the balance between adaptive responses and immunopathology. *Immunol Rev* 229(1):216-231.

194. He LZ, *et al.* (2013) Agonist anti-human CD27 monoclonal antibody induces T cell activation and tumor immunity in human CD27-transgenic mice. *J Immunol* 191(8):4174-4183.

195. Quan L, *et al.* (2015) PD-1 Blockade Can Restore Functions of T-Cells in Epstein-Barr Virus-Positive Diffuse Large B-Cell Lymphoma In Vitro. *PLoS One* 10(9):e0136476.

196. Kiyasu J, *et al.* (2015) Expression of programmed cell death ligand 1 is associated with poor overall survival in patients with diffuse large B-cell lymphoma. *Blood*.

197. Chen BJ, *et al.* (2013) PD-L1 expression is characteristic of a subset of aggressive B-cell lymphomas and virus-associated malignancies. *Clin Cancer Res* 19(13):3462-3473.

198. Green MR, *et al.* (2012) Constitutive AP-1 activity and EBV infection induce PD-L1 in Hodgkin lymphomas and posttransplant lymphoproliferative disorders: implications for targeted therapy. *Clin Cancer Res* 18(6):1611-1618.

199. Jardin F (2014) Next generation sequencing and the management of diffuse large B-cell lymphoma: from whole exome analysis to targeted therapy. *Discov Med* 18(97):51-65.

200. Steinhardt JJ, Gartenhaus RB (2012) Promising personalized therapeutic options for diffuse large B-cell Lymphoma Subtypes with oncogene addictions. *Clin Cancer Res* 18(17):4538-4548.

201. Simon R, Roychowdhury S (2013) Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov* 12(5):358-369.

202. Pao W, *et al.* (2004) EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proc Natl Acad Sci U S A* 101(36):13306-13311.

203. Sjoblom T, *et al.* (2006) The consensus coding sequences of human breast and colorectal cancers. *Science* 314(5797):268-274.

204. Wood LD, *et al.* (2007) The genomic landscapes of human breast and colorectal cancers. *Science* 318(5853):1108-1113.

205. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74(12):5463-5467.

206. Casey G, Conti D, Haile R, Duggan D (2013) Next generation sequencing and a new era of medicine. *Gut* 62(6):920-932.

207. Pasqualucci L, *et al.* (2011) Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* 43(9):830-837.

208. Woollard PM, *et al.* (2011) The application of next-generation sequencing technologies to drug discovery and development. *Drug Discov Today* 16(11-12):512-519.

209. Staton AD, Koff JL, Chen Q, Ayer T, Flowers CR (2015) Next-generation prognostic assessment for diffuse large B-cell lymphoma. *Future Oncol* 11(17):2443-2457.

210. Younes A, Berry DA (2012) From drug discovery to biomarker-driven clinical trials in lymphoma. *Nat Rev Clin Oncol* 9(11):643-653.

211. Johansson H, *et al.* (2011) Targeted resequencing of candidate genes using selector probes. *Nucleic Acids Res* 39(2):e8.

212. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 4(7):1073-1081.

213. Adzhubei IA, *et al.* (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7(4):248-249.

214. Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25(14):1754-1760.

215. Koboldt DC, *et al.* (2009) VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics* 25(17):2283-2285.

216. Li R, *et al.* (2009) SNP detection for massively parallel whole-genome resequencing. *Genome Res* 19(6):1124-1132.

217. Thorvaldsdottir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 14(2):178-192.

218. Lou H, *et al.* (2015) A 3.4-kb Copy-Number Deletion near EPAS1 Is Significantly Enriched in High-Altitude Tibetans but Absent from the Denisovan Sequence. *Am J Hum Genet* 97(1):54-66.

219. Chiang DY, *et al.* (2009) High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nat Methods* 6(1):99-103.

220. Podnar J, Deiderick H, Huerta G, Hunicke-Smith S (2014) Next-Generation Sequencing RNA-Seq Library Construction. *Curr Protoc Mol Biol* 106:4 21 21-24 21 19.

221. Jia W, *et al.* (2013) SOAPfuse: an algorithm for identifying fusion transcripts from paired-end RNA-Seq data. *Genome Biol* 14(2):R12.

222. Ellegren H (2004) Microsatellites: simple sequences with complex evolution. *Nat Rev Genet* 5(6):435-445.

223. Izon DJ, *et al.* (2002) Deltex1 redirects lymphoid progenitors to the B cell lineage by antagonizing Notch1. *Immunity* 16(2):231-243.

224. Ahn J, Urist M, Prives C (2004) The Chk2 protein kinase. *DNA Repair* 3(8-9):1039-1047.

225. Krishnakumar R, Kraus WL (2010) The PARP side of the nucleus: molecular actions, physiological outcomes, and clinical targets. *Mol Cell* 39(1):8-24.

226. Nagarajan N, *et al.* (2012) Whole-genome reconstruction and mutational signatures in gastric cancer. *Genome Biol* 13(12):R115.

227. Martin A, *et al.* (2003) Msh2 ATPase activity is essential for somatic hypermutation at a-T basepairs and for efficient class switch recombination. *J Exp Med* 198(8):1171-1178.

228. Bardwell PD, *et al.* (2004) Altered somatic hypermutation and reduced class-switch recombination in exonuclease 1-mutant mice. *Nat Immunol* 5(2):224-229.

229. Martomo SA, Yang WW, Gearhart PJ (2004) A role for Msh6 but not Msh3 in somatic hypermutation and class switch recombination. *J Exp Med* 200(1):61-68.

230. Liu M, *et al.* (2008) Two levels of protection for the B cell genome during somatic hypermutation. *Nature* 451(7180):841-845.

231. Zhang S, Lloyd R, Bowden G, Glickman BW, de Boer JG (2001) Msh2 DNA mismatch repair gene deficiency and the food-borne mutagen 2-amino-1-methy1-6-phenolimidazo [4,5-b] pyridine (PhIP) synergistically affect mutagenesis in mouse colon. *Oncogene* 20(42):6066-6072.

232. South AP, Cho RJ, Aster JC (2012) The double-edged sword of Notch signaling in cancer. *Semin Cell Dev Biol* 23(4):458-464.

233. Chan VW, Lowell CA, DeFranco AL (1998) Defective negative regulation of antigen receptor signaling in Lyn-deficient B lymphocytes. *Curr Biol* 8(10):545-553.

234. Scapini P, Pereira S, Zhang H, Lowell CA (2009) Multiple roles of Lyn kinase in myeloid cell signaling and function. *Immunol Rev* 228(1):23-40.

235. Chan VW, Meng F, Soriano P, DeFranco AL, Lowell CA (1997) Characterization of the B lymphocyte populations in Lyn-deficient mice and the role of Lyn in signal initiation and down-regulation. *Immunity* 7(1):69-81.

236. Pasqualucci L, *et al.* (2001) Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 412(6844):341-346.

237. Taube JM, *et al.* (2014) Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res* 20(19):5064-5074.

238. Tumeh PC, *et al.* (2014) PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* 515(7528):568-571.

239. Kelderman S, Schumacher TN, Kvistborg P (2015) Mismatch Repair-Deficient Cancers Are Targets for Anti-PD-1 Therapy. *Cancer Cell* 28(1):11-13.

240. Le DT, *et al.* (2015) PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med* 372(26):2509-2520.

241. Lawrence MS, *et al.* (2014) Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* 505(7484):495-501.