UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1251

Exploring next-generation sequencing in chronic lymphocytic leukemia

VIKTOR LJUNGSTRÖM

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2016

ISSN 1651-6206
ISBN 978-91-554-9674-6
urn:nbn:se:uu:diva-302026

Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, Rudbecklaboratoriet, Dag Hammarskjölds v 20, Uppsala, Friday, 14 October 2016 at 09:15 for the degree of Doctor of Philosophy (Faculty of Medicine). The examination will be conducted in English. Faculty examiner: Associate Professor Catherine J. Wu (Department of Medical Oncology, Dana-Farber Cancer Institute).

Abstract
Ljungström, V. 2016. Exploring next-generation sequencing in chronic lymphocytic leukemia. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine* 1251. 61 pp. Uppsala: Acta Universitatis Upsaliensis. ISBN 978-91-554-9674-6.

Next-generation sequencing (NGS) techniques have led to major breakthroughs in the characterization of the chronic lymphocytic leukemia (CLL) genome with discovery of recurrent mutations of potential prognostic and/or predictive relevance. However, before NGS can be introduced into clinical practice, the precision of the techniques needs to be studied in better detail. Furthermore, much remains unknown about the genetic mechanisms leading to aggressive disease and resistance to treatment. Hence, in **Paper I**, the technical performance of a targeted deep sequencing panel including 9 genes was evaluated in 188 CLL patients. We were able to validate 143/155 (92%) selected mutations through Sanger sequencing and 77/82 mutations were concordant in a second targeted sequencing run, indicating that the technique can be introduced in clinical practice. In **Paper II** we screened 18 NF-κB pathway genes in 315 CLL patients through targeted deep sequencing which revealed a recurrent 4 base-pair deletion in the *NFKBIE* gene. Screening of *NFKBIE* in 377 additional cases identified the mutation in ~6% of all CLL patients. We demonstrate that the lesion lead to aberrant NF-κB signaling through impaired interaction with p65 and is associated with unfavorable clinical outcome. In **Paper III** we sought to delineate the genetic lesions that leads to relapse after fludarabine, cyclophosphamide, and rituximab treatment. Through whole-exome sequencing of pre-treatment and relapse samples from 41 cases we found evidence of frequent selection of subclones harboring driver mutations and subsequent clonal evolution following treatment. We also detected mutations in the ribosomal protein *RPS15* in 8 cases (19.5%) and characterization of the mutations through functional assays point to impaired p53 regulation in cells with mutated *RPS15*. **Paper IV** aimed at characterizing 70 patients assigned to three major subsets (#1, #2, and #4) through whole-genome sequencing. Besides recurrent exonic driver mutations, we report non-coding regions significantly enriched for mutations in subset #1 and #2 that may facilitate future molecular studies. Collectively, this thesis supports the potential of targeted sequencing for mutational screening of CLL in clinical practice, provides novel insight into the pathobiology of aggressive CLL, and demonstrates the clinical outcome and cellular effects of *NFKBIE* and *RPS15* mutations.

Keywords: CLL, next-generation sequencing, clonal evolution, stereotypy, RPS15, NFKBIE

Viktor Ljungström, Department of Immunology, Genetics and Pathology, Experimental and Clinical Oncology, Rudbecklaboratoriet, Uppsala University, SE-751 85 Uppsala, Sweden.

ISSN 1651-6206
ISBN 978-91-554-9674-6
urn:nbn:se:uu:diva-302026 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-302026)

No, no, you're not thinking; you're just being logical.

- Nils Bohr

List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals. Reprints were made with permission from the respective publishers.

I Sutton LA*, **Ljungström V***, Mansouri L, Young E, Cortese D, Navrkalova V, Malcikova J, Muggen AF, Trbusek M, Panagiotidis P, Davi F, Belessi C, Langerak AW, Ghia P, Pospisilova S, Stamatopoulos K, Rosenquist R. *Targeted next-generation sequencing in chronic lymphocytic leukemia: a high-throughput yet tailored approach will facilitate implementation in a clinical setting.* Haematologica. 2015 Mar;100(3):370-6.

II Mansouri L, Sutton LA, **Ljungström V**, Bondza S, Arngården L, Bhoi S, Larsson J, Cortese D, Kalushkova A, Plevova K, Young E, Gunnarsson R, Falk-Sörqvist E, Lönn P, Muggen AF, Yan XJ, Sander B, Enblad G, Smedby KE, Juliusson G, Belessi C, Rung J, Chiorazzi N, Strefford JC, Langerak AW, Pospisilova S, Davi F, Hellström M, Jernberg-Wiklund H, Ghia P, Söderberg O, Stamatopoulos K**, Nilsson M**, Rosenquist R**. *Functional loss of IκBε leads to NF-κB deregulation in aggressive chronic lymphocytic leukemia.* J Exp Med. 2015 Jun 1;212(6):833-43.

III **Ljungström V***, Cortese D*, Young E, Pandzic T, Mansouri L, Plevova K, Ntoufa S, Baliakas P, Clifford R, Sutton LA, Blakemore SJ, Stavroyianni N, Agathangelidis A, Rossi D, Höglund M, Kotaskova J, Juliusson G, Belessi C, Chiorazzi N, Panagiotidis P, Langerak AW, Smedby KE, Oscier D, Gaidano G, Schuh A, Davi F, Pott C, Strefford JC, Trentin L, Pospisilova S, Ghia P, Stamatopoulos K, Sjöblom T**, Rosenquist R**. *Whole-exome sequencing in relapsing chronic lymphocyticleukemia: clinical impact of recurrent RPS15 mutations.* Blood. 2016 Feb25;127(8):1007-16.

IV **Ljungström V***, Young E*, Agathangelidis A, Muggen AF, Plevova K, Rossi D, Davis Z, Sutton LA, Baliakas P, Ntoufa S, Davi F, Gaidano G, Oscier D, Pospisilova S, Langerak AW, Stamatopoulos K, Ghia P**, Rosenquist R**. *Whole-genome sequencing in subsets of chronic lymphocytic leukemia harboring stereotyped B-cell receptors.* Manuscript.

* Equal first authors
** Equal senior authors

Related papers published during the PhD period

1. Mansouri L, Sutton LA, **Ljungström V**, Sörqvist EF, Gunnarsson R, Smedby KE, Juliusson G, Stamatopoulos K, Nilsson M, Rosenquist R. *Feasibility of targeted next-generation sequencing of the TP53 and ATM genes in chronic lymphocytic leukemia.* Leukemia. 2014 Mar;28(3):694-6.

2. Parry M*, Rose-Zerilli MJ*, **Ljungström V***, Gibson J, Wang J, Walewska R, Parker H, Parker A, Davis Z, Gardiner A, McIver-Brown N, Kalpadakis C, Xochelli A, Anagnostopoulos A, Fazi C, Gonzalez de Castro D, Dearden C, Pratt G, Rosenquist R, Ashton-Key M, Forconi F, Collins A, Ghia P, Matutes E, Pangalis G, Stamatopoulos K, Oscier D, Strefford JC. *Genetics and Prognostication in Splenic Marginal Zone Lymphoma: Revelations from Deep Sequencing.* Clin Cancer Res. 2015 Sep15;21(18):4174-83.

3. Kiflemariam S, **Ljungström V**, Pontén F, Sjöblom T. *Tumor vessel up-regulation of INSR revealed by single-cell expression analysis of the tyrosine kinome and phosphatome in human cancers.* Am J Pathol. 2015 Jun;185(6):1600-9.

4. Malcovati L, Karimi M, Papaemmanuil E, Ambaglio I, Jädersten M, Jansson M, Elena C, Gallì A, Walldin G, Della Porta MG, Raaschou-Jensen K, Travaglino E, Kallenbach K, Pietra D, **Ljungström V**, Conte S, Boveri E, Invernizzi R, Rosenquist R, Campbell PJ, Cazzola M, Hellström Lindberg E. *SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts.* Blood. 2015 Jul 9;126(2):233-41.

5. Crona J, **Ljungström V**, Welin S, Walz MK, Hellman P, Björklund P. *Bioinformatic Challenges in Clinical Diagnostic Application of Targeted Next Generation Sequencing: Experience from Pheochromocytoma.* PLoS One. 2015 Jul 31;10(7):e0133210.

6. Moens LN, Falk-Sörqvist E, **Ljungström V**, Mattsson J, Sundström M, La Fleur L, Mathot L, Micke P, Nilsson M, Botling J. *HaloPlex Targeted Resequencing for Mutation Detection in Clinical Formalin-Fixed, Paraffin-Embedded Tumor Samples*. J Mol Diagn. 2015 Nov;17(6):729-39.

7. Pandzic T*, Larsson J*, He L, Kundu S, Ban K, Akhtar-Ali M, Hellström AR, Schuh A, Clifford R, Blakemore SJ, Strefford JC, Bauman TS, Lopez-Guillermo A, Campo E, **Ljungström V**, Mansouri L, Rosenquist R, Sjöblom T, Hellström M. *Transposon mutagenesis reveals fludarabine-resistance mechanisms in chronic lymphocytic leukemia.* Clin Cancer Res. 2016 Mar 8. pii: clincanres.2903.2015.

8. Bhoi S, **Ljungström V**, Baliakas P, Mattsson M, Smedby KE, Juliusson G, Rosenquist R, Mansouri L. *Prognostic impact of epigenetic classification in chronic lymphocytic leukemia: The case of subset #2. Epigenetics.* 2016 Jun 2;11(6):449-55.

9. Parker H, Rose-Zerilli MJ, Larrayoz M, Clifford R, Edelmann J, Blakemore S, Gibson J, Wang J, **Ljungström V**, Wojdacz TK, Chaplin T, Roghanian A, Davis Z, Parker A, Tausch E, Ntoufa S, Ramos S, Robbe P, Alsolami R, Steele AJ, Packham G, Rodríguez-Vicente AE, Brown L, McNicholl F, Forconi F, Pettitt A, Hillmen P, Dyer M, Cragg MS, Chelala C, Oakes CC, Rosenquist R, Stamatopoulos K, Stilgenbauer S, Knight S, Schuh A, Oscier DG, Strefford JC. *Genomic disruption of the histone methyltransferase SETD2 in chronic lymphocytic leukaemia.* Leukemia. 2016 Jun 10.

10. Navrkalova V, Young E, Baliakas P, Radova L, Sutton LA, Plevova K, Mansouri L, **Ljungström V**, Ntoufa S, Davies Z, Juliusson G, Smedby KE, Belessi C, Panagiotidis P, Touloumenidou T, Davi F, Langerak AW, Ghia P, Strefford JC, Oscier D, Mayer J, Stamatopoulos K, Pospisilova S, Rosenquist R, Trbusek M. *ATM mutations in major stereotyped subsets of chronic lymphocytic leukemia: enrichment in subset #2 is associated with markedly short telomeres*. Haematologica. 2016 Jun 16.

* Equal first authors

Contents

Abbreviations

AID	Activation induced cytidine deaminase
BcR	B cell receptor
BTK	Bruton's tyrosine kinase
CCF	Cancer cell fraction
CLL	Chronic lymphocytic leukemia
CSR	Class switch recombination
CDR	Complementarity-determining region
CR	Complete remission
D	Diversity
FCR	Fludarabine, cyclophosphamide, and rituximab
GC	Germinal center
HSC	Hematopoietic stem cell
IG	Immunoglobulin
IGH	Immunoglobulin heavy
iwCLL	International Workshop on CLL
J	Joining
MZ	Marginal zone
miRNA	MicroRNA
M-CLL	CLL with mutated IGHV genes
MDR	Minimal deleted region
MBL	Monoclonal B cell lymphocytosis
NGS	Next-generation sequencing
OS	Overall survival
PR	Partial remission
PFS	Progression-free survival
PD	Progressive disease
sCNA	Somatic copynumber aberration
SHM	Somatic Hypermutation
SD	Stable disease
TTFT	Time to first treatment
TLR	Toll-like receptor
U-CLL	CLL with unmutated IGHV genes
V	Variable
VAF	Variant allele frequency
WES	Whole-exome sequencing
WGS	Whole-genome sequencing

Introduction

Chronic lymphocytic leukemia (CLL) is a malignancy of the B cells characterized by the accumulation of clonal B cells in secondary lymphoid organs, bone marrow and peripheral blood[1]. CLL is the most common leukemia of the elderly in the Western population and around 500 cases are diagnosed annually in Sweden. The median age of diagnosis is 71 years although one third of patients with CLL are diagnosed at 65 years or younger[2]. The disease can present with lymphadenopathy, signs of impaired bone marrow function, i.e. anemia, bleeding tendency and infections, and systemic symptoms such as weight loss and fever. However, in the majority of cases (~85%), the patients are asymptomatic at the time of diagnosis, as the disease is often incidentally detected in a routine blood screen. For a CLL diagnosis the following criteria need to be met: blood lymphocytosis with clonal B cell count $\geq 5x10^9$/L, a typical immunophenotype ($CD5^+$, $CD19^+$, $CD23^+$, and CD20, CD22, and surface IG expressed at lower levels than normal B cells) as detected by flow cytometry, and a morphological picture with small, mature B lymphocytes[3]. The clinical course of the disease is highly variable with some patients surviving years, even decades, with little or no treatment, whereas others experience rapid disease progression and require several lines of treatment[4].

CLL is a malignancy and is as such driven by classic cancer-associated genetic events including oncogene activation and incativation of tumor suppressor genes[5]. However, a large body of research on CLL has also revealed that the malignant cells are heavily dependent on cues from the microenvironment and that signaling through various receptors including the B cell receptor (BcR) is vital for disease development[6-9]. In order to fully appreciate the molecular basis of CLL, the immunogenetic background of the disease and the evidence regarding disease ontogeny will be outlined before approaching the CLL genome.

Immunogenetics in CLL

The main physiological function of B cells is to produce immunoglobulins (IGs) that can recognize antigens and assist in the signal transmission necessary for elimination of potentially harmful agents. In order to recognize

the vast number of possible antigens, the immune system has to be able to produce a wide catalogue of antigen-specific B cells. The BcR is created during B cell development through a complex, programmed process that enables extensive antigen recognition capacity[10]. The first step involves the rearrangement of the immunoglobulin heavy chain (IGH) locus, which consists of variable (V), diversity (D), and joining (J) genes that are randomly recombined leading to combinatorial diversity[10]. During the recombination process, enzymatic processing can also lead to insertion or removal of nucleotides at the IGHV-IGHD and IGHD-IGHJ junctions, leading to increased variability of the IG antigen-binding site. There are three complementarity-determining regions (CDRs) that define the specificity of the antigen-binding site with the most variable being CDR3[11]. If the heavy chain rearrangement process is successful, IG light chain rearrangement will commence. This process is similar to the IGH rearrangement process, however, the light chain lacks D genes; hence there is only one recombination step[12].

In subsequent stages of B-cell development, if the BcR is stimulated through interaction with the corresponding epitope on an antigen, the IG molecule can undergo affinity maturation. This process involves somatic hypermutation (SHM) and class switch recombination (CSR) and takes place in secondary lymphoid organs[13]. SHM generates mutations in the variable domains of the IG molecule; B cells acquiring mutations that give their IG high affinity to the target antigen will be selected for and subject to clonal expansion. CSR is the process in which the isotype expression of the B cells is changed from IgM to IgA, IgE or IgG through alterations of the constant region of the IG molecule. If all steps in the IG maturation process are successful, the result is a functional IG with the highest achieved degree of antigen specificity[14].

The IG gene repertoire in CLL

Since the IG processing steps involve numerous different combinations of genes, stochastic mutagenesis in the IG junctions, SHM, and CSR; the number of potential IG gene sequences is in the range of trillions[10]. CLL, as well as other malignancies, clonally expands from a progenitor cell and malignant cells will throughout disease development all carry almost identical IG molecules in their BcR. Analysis of the CLL IG configuration is therefore instrumental in order to understand the molecular timeline and to better define subgroups within the disease[7]. Even though the number of potential IG gene configurations is astronomic, it was observed in the early 1990s that CLL displays a remarkably restricted and biased IGHV gene repertoire with an overrepresentation of the IGHV1-69, IGHV3-7, IGHV3-21 and IGHV4-34 genes[8,15-17]. Another striking observation was that the

presence of SHM is not uniform across the gene subgroups, as genes such as IGHV3-7 and IGHV4-34 were associated with a high mutational load, whereas IGHV1-69 exhibited few or no mutations[15]. Subsequent studies also revealed biased IGHV-IGHD-IGHJ combinations and recurrent pairings of specific IG heavy and light chain genes[18]. Two seminal studies published back-to-back in 1999 revealed that approximately 50% of all CLL cases carry IGHV genes that have undergone SHM[19,20]. Even more intriguing was the discovery that patients with no or limited SHM (≥98% identity to the germline sequence) displayed an aggressive clinical behavior, whereas patients with mutated IGHV genes (<98% identity to the germline sequence) followed a significantly more indolent clinical course. The term unmutated CLL (U-CLL) describes the former group whereas mutated CLL (M-CLL) refers to the latter group[19,20].

The revelations that CLL presents with a restricted IGHV gene repertoire and that the degree of SHM is influenced by the IG gene usage led to further endeavors to characterize the IGHV gene configuration in the disease. It was soon revealed that unrelated CLL cases could express highly similar variable heavy (VH) CDR3 sequences based on their amino acid motifs[21]. In 2003, it was observed that 50% of IGHV3-21 CLL cases harbor almost identical VH CDR3 sequences together with biased usage of the lambda IGLV3-21 gene, and soon thereafter several other examples were presented with subgroups of patients harboring quasi-identical or 'stereotyped' BcRs, defined as similar IGHV-IGHD-IGHJ gene usage and >60% VH CDR3 amino acid identity[22-26]. The presence of highly homologous BcRs in CLL in combination with evidence of the CLL clone emerging from an antigen-experienced B cell (discussed below) led to the hypothesis that antigen stimulation is central for disease development[27]. This notion was corroborated by immunogenetic studies that revealed recurrent patterns of SHM within stereotyped subsets that indicate the recognition of a common discrete epitope. Today, there is a growing body of evidence supporting an antigen-driven model for CLL pathogenesis[28].

Major work has since been conducted in the area of BcR stereotypy and it is now accepted that approximately 30% of all CLL cases can be assigned to stereotyped subsets[29]. Around 200 distinct subsets have been identified and 19 major subsets make up 40% of all stereotyped cases[29]. Mounting evidence also points to homogeneous clinical and biological characteristics as well as similar clinical outcome within stereotyped subsets. In a recent study encompassing 8593 CLL cases, patients within the same subset followed the same clinical course with significantly different time-to-first-treatment (TTFT) compared to other stereotyped subsets. For example, two of the largest subsets, subsets #1 (IGHV1/5/7/IGKV1-39) and #2 (IGHV3-21/IGVL3-21) are associated with aggressive disease with poor outcome[30].

Subset #1 cases belong to U-CLL but subset #2 cases are a mix of U-CLL and M-CLL meaning that the assignment to subset #2 is a stronger prognostic indicator than IGHV status in this particular group[31]. In contrast, subset #4 patients (IGHV4-34/IGKV2-30) have a low age at diagnosis and present with a more indolent disease that rarely requires treatment[30]. As expected from their clinical profile, subset #4 cases have mutated IGHV molecules and accounts for 2% of all M-CLL cases[29]. Notably, in a study using IGs derived from subset #1 cells, it was revealed that the they effectively bind to malondialdehyde, a degradation product from the reaction between unsaturated lipids and reactive oxygen species[32]. In another study, it was established that IGs produced by cells from subset #6 binds specifically to non-muscle myosin heavy chain IIA[33]. Hence, autoantigens have been suggested as a potential culprit in CLL antigen stimulation[27]. However, binding to exoantigens such as microbial epitopes has also been demonstrated and there is an association with persistent Epstein-Barr infection and Cytomegalovirus infections and subset #4[32,34].

Table 1. Characteristics of subsets #1, #2, #4, and #8[26,29,30,35-37].

Subset	Size (%)	IGHV gene(s)	SHM status	Associated sCNAs	Associated mutations	TTFT (years)
#1	2.4	IGHV clan I	U-CLL	del(17p), del(11q)	*TP53* and *NOTCH1*	1.6
#2	2.8	IGHV3-21	U- and M-CLL	del(13q), del(11q)	*SF3B1*	1.9
#4	1.0	IGHV4-34	M-CLL	del(13q)		11.0
#8	0.5	IGHV4-39	U-CLL	trisomy12	*NOTCH1*	1.5

Normal counterparts of CLL

While commonly used in the literature, the term 'cell of origin' might not be adequate in this specific disease since further differentiation may take place between the initiating event and transformation, meaning that the actual cell of origin may be at an earlier stage of differentiation. Hence, the normal B cell counterpart of CLL could be a more accurate description. The IG configuration of the CLL clone in conjunction with the surface marker profile would theoretically provide a means for accurate delineation of the normal B cell counterpart. However, the complexity of the immune system confounds the issue and although there have been extensive efforts in defining the cell of origin of CLL for several decades, the topic still remain controversial. A reasonable hypothesis is that U-CLL and M-CLL originate from two distinct cell populations where the former stems from naïve B cells prior to SHM and the latter derives from B cells that have matured in a T cell dependent setting and acquired SHM in the germinal center (GC)[6]. Though practical, mounting evidence contests this 2-cell origin model. The prime

evidence came with gene-expression profiling studies that revealed a common gene expression signature in CLL with very few differences between U-CLL and M-CLL in contrast to thousands of differences compared to normal B cells. The normal B-cell population most resembling the signature of CLL was $CD27^+$ memory B cells, indicating a potential common origin from this population for both M-CLL and U-CLL[38,39]. It has also been demonstrated that B cells of marginal zone (MZ) origin can undergo CSR, with or without SHM, triggered by autoantigens or in response to bacterial polysaccharides in a T cell-independent manner and without GC formation. This phenomenon might be related to subgroups of U-CLL that have undergone class-switching. However, even though this model is mechanistically appealing, MZ B cells are $CD5^-CD23^-CD22^+$, a surface phenotype divergent from CLL cells[21,40].

The IGs expressed by CLL cells, and U-CLL in particular, are often polyreactive and able to bind to autoantigens and exoantigens, a behavior resembling that of natural antibodies[41,42]. These antibodies are detectable in normal individuals and are capable of recognizing epitopes highly conserved in different pathogens, commensal organisms, and endogenous intracellular structures[43]. The innate-like B1 lineage of B cells constitutes a major source of natural antibodies in mice[44]. This population is abundant in the peritoneal cavity and displays a restricted BcR IG repertoire with little or no SHM that largely resembles that of stereotyped BcR IGs in CLL[45,46]. The existence of human B cells with these exact features still remain to be fully characterized and the view that this potential population is the normal counterpart will remain contested until further evidence is presented.

In a recent study based on gene expression profiling, a novel memory $CD5^+CD27^+$ IGHV-M post-germinal center B cell subset was discovered that was suggested to derive from $CD5^+CD27^-$ B cells[47]. This model postulates that the former population would give rise to M-CLL whereas the latter would be the cell of origin for U-CLL. Interestingly, the authors also state that these B cell populations preferentially express stereotyped BcRs with SHM patterns corresponding to those observed in CLL[47].

Major efforts have been undertaken to characterize the CLL epigenome in recent years and the findings from these studies add to the complexity of the topic. In a large-scale study of the epigenome of CLL and physiological B cell populations, the authors suggest returning to the early view of U-CLL corresponding to naïve B cells and M-CLL to a memory B cell origin[48]. In another study with similar experimental design, the authors propose a less categorical view and instead hypothesize that the heterogeneity of the disease represents an origin from a continuous spectrum of maturation states represented by normal B cell development stages[49].

Collectively, this demonstrates that the normal B cell counterpart is a complex issue that likely will require integration of several analysis modalities in order to reach consensus[6].

Precursor stages of CLL

CLL is preceded by monoclonal B cell lymphocytosis (MBL), a condition defined by the presence of a clonal B cell population with fewer than 5 x 10^9/L B-cells in peripheral blood in an otherwise healthy individual[50]. Screening for MBL in healthy subjects using eight-color flow cytometry have revealed a prevalence of >10% in subjects older than 40 years and >20% in individuals older than 60 years[51-53]. The condition can further be subdivided into low-count (<0.5 x 10^9/L) and high-count (≥0.5 x 10^9/L) MBL. The former rarely gives rise to CLL whereas the latter progresses into CLL at a rate of 1% to 2% per year[54]. Low-count MBL presents a distinct IGHV gene repertoire compared to CLL whereas that of high-count MBL closely resembles early stage CLL[50,55,56]. CLL-associated genetic lesions can also be detected also in the MBL stage of the disease[54]. Chromosomal aberrations associated with a favorable prognosis are detected in similar frequencies in MBL and CLL, and recent studies have also shown that about half of all high-count MBL cases carry recurrent driver mutations such as mutations in *SF3B1*, *NOTCH1* and *BIRC3*[57,58].

In recent years, several studies have presented evidence suggesting that the earliest genetic and epigenetic events in CLL development might occur in pluripotent hematopoietic stem cells (HSCs). In 2014, a study identified cases that shared known CLL driver mutations (e.g. *SF3B1*, *NOTCH1* and *BRAF)* between the CLL clone and immature progenitor populations as well as monocyte populations. The authors suggest that the alterations were acquired in an early progenitor cell that later could contribute to both the lymphoid and myeloid lineages[59]. Xenografting of HSCs from patients with CLL in mice has also shown the ability of these populations to engraft and cause mono- or oligoclonal B cell proliferation. This would imply that early lesions are present in the HSC population leading to B cell lineage bias[60]. These observations were supported by recent studies of large cohorts without hematological malignancies where typical mutations associated with myeloid and lymphoid neoplasias were detected. The phenomenon has been named clonal hematopoiesis and is associated with increased risk of hematological malignancies, including CLL[61,62].

The CLL genome

Next-generation sequencing (NGS) has enabled the characterization of all genetic lesions present in the CLL genome at a single time point. Through whole-genome sequencing (WGS) and whole-exome sequencing (WES) it has been revealed that each CLL tumor carries on average 0.9 mutations per megabase and ~10-30 non-silent mutations per case. The total number of somatic mutations is higher in the M-CLL subgroup than in U-CLL (~2,800 vs ~2,000), likely reflective of off-target SHM in the former group. The average number of non-silent mutations is similar to that observed in acute myeloid leukemia (0.8 mutations per megabase) but significantly lower than many solid cancer forms such as lung cancer and melanoma that carry on average 10 mutations per megabase [63-66].

Analysis of somatic mutational signatures from WGS data has revealed three separate signatures in CLL[63,64]. First, an age-related signature characterized by C→T transitions at CpG sites, likely related to the elevated rate of spontaneous deamination of 5-methylcytosine. Second, a signature characterized by T:A→G:C transversions can be detected in M-CLL. Finally, a signature characterized by an enrichment of C→T and C→G mutations at TpCpN trinucleotides can be observed at IG loci and genes with high expression in the germinal center. This signature has been attributed to the activation-induced cytosine deaminase (AID) that mediates SHM and CSR[63,64].

Genomic aberrations

While CLL is not associated with a specific genetic aberration, more than 80% of CLL patients carry cytogenetically detectable recurrent chromosomal abnormalities at diagnosis[67]. The most frequently observed aberrations are 13q14 deletion [del(13q)], 11q22-23 deletion [del(11q)], trisomy 12, and 17p13 deletion [del(17p)][67-70].

13q14 deletion

The most frequent chromosomal aberration detected in CLL is del(13q) which affects ~55% of all cases[67]. The deletion is more frequently observed

in M-CLL and isolated del(13q) is associated with favorable prognosis. However, ~20% have large 13q deletions that spans the well-characterized tumor-suppressor gene *RB1* and these patients have shortened survival and accelerated disease progression[71,72]. del(13q) is also found in up to 44% of MBL cases, indicating that the lesion is an early event in CLL development but not associated with progression to advanced disease[57]. The deletion regularly encompasses the microRNA (miRNA) cluster *miR15a-miR16-1,* residing in an intron of *DLEU2*, and the *DLEU7* gene[73,74]. The oncogenic function of del(13q) have been associated with deregulation of the anti-apoptotic *BCL2* gene due to deletion of *miR15a-miR16-1,* and increased NF-κB signaling due to loss of *DLEU7*. In contrast to other recurrent somatic copy number aberrations (sCNAs) in CLL, biallelic deletions of 13q are observed in up to 25% of del(13q) patients[72,75].

11q22-23 deletion

Around 18% of CLL cases display del(11q) at diagnosis[67]. The deletion is almost exclusively monoallelic and the minimal deleted region (MDR) spans the tumor suppressor gene ataxia telangiectasia mutated (*ATM*), which encodes an important regulator of DNA repair and cell division that activates p53 in response to double-strand breaks[76]. Mutations in *ATM* are observed in approximately 20-40% of cases with del(11q) compared to ~12% of all CLL, indicating that disruption of the gene could be a driving mechanism behind the deletion[66,77,78]. Although this is an appealing mechanism, the fact that a majority of del(11q) cases do not carry biallelic defects in *ATM* could point to other genes that also have a role in the pathogenesis of the deletion. In fact, it has recently been reported that deletions spanning the *BIRC3* gene are associated with very unfavorable outcome[79] (discussed below). del(11q) is associated with lymphadenopathies and dismal prognosis and the lesion is primarily detected in patients with U-CLL[72].

Trisomy 12

Trisomy 12 is detected in 11-16% of all CLL and is associated with intermediate risk[67]. This aberration is enriched in patients that have undergone Richter's transformation and is also associated with a higher incidence of secondary tumors[80]. The underlying mechanism by which trisomy 12 contribute to disease development is still poorly understood[68].

17p13 deletion

The deletion of 17p is found in less than 10% of all CLL cases at diagnosis and is heavily associated with U-CLL[67]. The deletion spans the *TP53* gene that has a central role in cell cycle control and is recurrently inactivated in

many cancer forms[81]. Mutations in *TP53* were identified in CLL already in 1991 and are associated with advanced disease stages and very poor response to therapy[82-84]. *TP53* mutations are most frequently found in patients with del(17p) leading to functional disruption of both alleles [67,83,84].

Recurrently mutated genes

NGS-based sequencing studies enabled characterization of the genomic landscape of CLL with unprecedented resolution. In line with the knowledge from the pre-NGS era, somatic mutations were detected in the aforementioned tumor suppressor genes *TP53* and *ATM* that are associated with DNA-damage and cell cycle control. Recurrent mutations were also discovered in pathways associated with B cell biology such as the NF-κB pathway and in Notch signaling and subunits of the RNA splicing machinery[65,66,85,86].

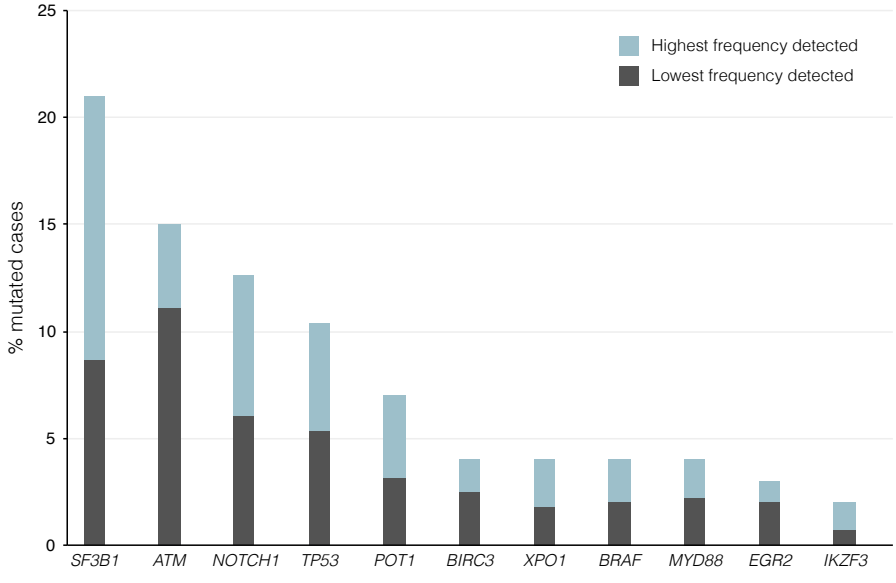

Figure 1. Frequencies of recurrently mutated genes in CLL[64,87,88].

NOTCH1 mutations

NOTCH1 encodes a ligand-activated transcription factor involved in regulation of cell differentiation, proliferation and apoptosis and the gene was found recurrently mutated in the first NGS initiatives in CLL[65,86]. The frequency of *NOTCH1* mutations ranges from 5-15% and the mutation is observed in higher frequencies in more aggressive patient cohorts[86,88,89].

Compared to general patient cohorts, a higher concentration of *NOTCH1* mutations have recently been observed in the stereotyped subsets #1, #59, and #8 where the mutations were detected in 20-30% of the patients[36,37]. The C-terminus end of the protein contains a PEST domain that targets the protein for proteasome degradation via the FBXW7-SCF ubiquitin ligase complex and thereby limits the intensity and duration of Notch signaling[90]. A 2 base-pair frame-shift mutation affecting this domain is the most common *NOTCH1* lesion and accounts for 85-90% of all *NOTCH1* mutations in CLL[86]. The functional role of *NOTCH1* mutations in B-lymphocytes is not yet fully characterized, but initial results indicate that the truncating mutation generate a more stable and active isoform of the protein[65]. Cases with *NOTCH1* mutations have a more advanced Rai and Binet stage and more frequently carry unmutated IGHV genes. Trisomy 12 is significantly enriched in *NOTCH1*-mutated cases whereas del(13q) is seen less frequently[91,92].

SF3B1 mutations

Two large WES studies revealed that *SF3B1*, a gene that encodes an element of the splicing machinery, was mutated in 10-15% of the cases making it one of the most frequently mutated genes in CLL[66,85]. Interestingly, a strong association between *SF3B1* mutations and subset #2, and recently also subset #3, has been discovered[35-37]. In these subgroups, mutation frequencies of ~45% have been detected indicating that the lesions have a central role in the pathobiology of these subsets. The mutations cluster to a few evolutionarily conserved hotspots within six HEAT repeat domains. The exact driving mechanism behind *SF3B1* mutations has not yet been elucidated but comparative analysis of exon arrays revealed 184 genes with exons showing differential inclusion in *SF3B1*-mutated cases[85]. Transcriptome sequencing of *SF3B1* mutated CLL cases revealed the same levels of aberrant exon inclusion and furthermore 79 splice junctions specific to *SF3B1*-mutated cells[93]. Taken together, these findings indicate that *SF3B1* mutations may lead to incorrect splicing due to a defective spliceosome complex and that the oncogenic effect of the mutations could be associated with the expression of abnormal transcripts.

BIRC3 aberrations

The product encoded by *BIRC3* is part of a protein complex that negatively regulates MAP3K14, a central regulator of non-canonical NF-κB signaling. The gene can be disrupted either by deletions or truncating mutations that removes the C-terminal RING domain necessary for degradation of MAP3K14[79]. *BIRC3* lesions are rare at diagnosis (4%) but are enriched in

fludarabine-refractory CLL (24%) and display mutual exclusivity with *TP53* defects[79,88].

MYD88 mutations

MYD88 encodes a cytosolic adaptor protein involved in the Toll-like receptor (TLR) pathway and mutations in the gene have been detected in 2-5% of CLL patients. In contrast to the majority of the recurrently mutated genes in CLL, *MYD88* mutations are almost exclusively found in M-CLL[65,88,94].

Other recurrently mutated genes

Studies of more than 1000 CLL genomes[64,66,85,87] through WES or WGS have revealed the most frequently mutated genes in the disease, albeit with varying prevalence of the mutations depending on the characteristics of the study cohort. Studies based on larger-sized cohorts have also uncovered a very long 'tail' of less frequently mutated genes that might have implications on disease biology in specific subgroups of the disease[95].

Mutations in *POT1* are found in ~3% of all CLL and are associated with U-CLL and predominantly advanced disease[96]. The encoded protein is part of the shelterin complex that is involved in the protection of telomeres and mutations in *POT1* have been demonstrated to hinder effective binding to telomeric DNA, resulting in telomeric and chromosomal aberrations[96]. Mutations have also been observed in RNA export factor genes. The most frequently mutated gene within this entity is *XPO1,* but mutations in *NXF1*, *DDX3X* and *RANBP2* have also been described. *XPO1* mutations are isolated to U-CLL cases, suggesting that they might drive the development of more aggressive disease forms though the exact prognostic relevance of the mutations still remains to be elucidated[65,87,97]. Mutations in genes encoding B cell transcription factors have been detected in moderate frequencies in CLL. The prime examples are mutations in *EGR2* and *IKZF3* that are observed in 2.8% and 2% of general CLL, respectively[59,87]. *EGR2* mutations are significantly enriched in aggressive subgroups and in particular treatment refractory cases (Young et al., under review). Finally, several pathways including chromatin modification (e.g. *CHD2*, *ASXL1* and *HIST1H1B),* MYC-regulation (e.g. *MGA* and *PTPN11*), and RAS genes (e.g. *BRAF*, *KRAS*, and *NRAS*), are recurrently affected by mutations, however again at lower frequencies[64,87]. The already extensive list of recurrently mutated genes in CLL is still expected to expand as saturation analysis has estimated that ~2000 cases are required to delineate mutations present in 1-2% of the population[98].

Non-coding mutations

Most large-scale sequencing initiatives in CLL are based on exome sequencing and mutations in coding regions of the genome are both more likely to have a direct effect on cellular mechanisms, and more straightforward to distinguish from background mutations. Subsequently, the absolute majority of the recurrent mutations detected in CLL reside in the coding parts of the genome. A recent sequencing study including WGS of 150 CLL and MBL cases reported two non-coding genomic regions with recurrent somatic mutations[64]. First, mutations in the 3'UTR region of *NOTCH1* were detected in 4/150 patients, leading to similar functional consequences and clinical outcome as the common 2 base-pair deletions. Second, an intergenic region of chromosome 9p13 displayed a high frequency of mutations in M-CLL and the locus was identified as a *PAX5* enhancer by ChIP-seq analysis. In a number of cases the mutations were the only observed recurrent lesions and the authors suggest that they constitute a potential driving mechanism[64]. With the advent of more powerful sequencing platforms, WGS is now gradually becoming more available for genome-wide mutational screening of larger cohorts. This could potentially reveal novel recurrent mutations in non-coding regions explaining the disease phenotype in subgroups of CLL.

CLL tumor heterogeneity and clonal evolution

Cancer is thought to be a clonal process with step-wise accumulation of genetic lesions that leads to tumor heterogeneity and clonal evolution[99]. The latter refers to the Darwinian selection and expansion of subclones that have acquired genetic and epigenetic traits that make them the fittest in the current cellular ecology. The addition of treatment leads to evolutionary bottlenecks that increase the risk of selecting a fit subpopulation that have acquired resistance-mechanisms, thereby resulting in a more aggressive disease form[95,99,100]. The fact that genetic lesions can be present only in a minor population of the tumor was first demonstrated in CLL by karyotyping and fluorescence in situ hybridization and later by SNP arrays[101,102]. Array-based studies also revealed that minor clones harboring del(11q) and del(17p) can expand under the selective pressure of treatment and that the CLL genome can be altered at relapse compared to at diagnosis[101,103,104].

Through NGS-based studies, these concepts have now been refined and today clonal dynamics can be traced at base-pair level. In order to quantify the tumor population affected by a specific lesion Carter et al. developed an algorithm that infers tumor purity and tumor cell ploidy from sCNAs and point mutations and calculates the cancer cell fraction (CCF) affected by

each lesion[105]. A lesion with high CCF is likely to be clonal and correspond to either founder events or later events from a subclone that have outlasted a selective sweep that eliminated all other cancer cells not harboring the mutation. In contrast, mutations with lower CCF most likely represent events acquired later in the disease[105,106]. Two main patterns have been observed in studies of clonal evolution. Cases showing evidence of one main clone that over time has acquired driver events are interpreted to follow a linear evolution. If there are multiple, competing subclones that evolve over time, the evolution pattern is instead described as branched[95,107]. In 2012, Schuh et al published the results of longitudinal WGS analysis of three CLL patients subjected to multiple rounds of treatment. The temporal patterns after treatment were variable between the patients varying from interclonal equilibrium between several subpopulations to clonal evolution, where one minor subclone expanded to become the major clone[108]. Landau et al. have presented two seminal large-scale studies delineating the clonal dynamics of CLL in relation to treatment[87,97]. The first study was based on a cohort of 149 CLL patients and investigated clonal composition and clonal dynamics over time using WES combined with copy-number analysis. This allowed for the identification of genetic lesions acquired early in the disease such as *MYD88* mutations and trisomy 12 and later, predominantly subclonal, lesions such as *TP53* and *SF3B1* mutations. Presence of a subclonal driver mutation was also observed to predispose to disease progression. For 18 patients where consecutive samples were available, clonal dynamics over time were analyzed and it was found that clonal evolution more frequently occurred in patients that had received treatment suggesting that this selective pressure may accelerate the evolutionary process[97]. The second study had a similar experimental setup but with a larger cohort consisting of 538 CLL cases of which 278 were collected from the prospective CLL8 trial[87,109]. Longitudinal investigation of 59 cases with samples available from pre-treatment and relapse from treatment with fludarabine, cyclophosphamide and rituximab (FCR), revealed three types of driver patterns: i) drivers with predominantly stable CCF [del(13q), del(11q) and trisomy 12)], ii) drivers with increasing CCF (del(17p) and *TP53*), and iii) drivers with shifting CCF (*SF3B1*, *ATM* and *POT1*)[87]. The latter group indicates that clones harboring certain recurrent mutations might be the fittest at specific evolutionary time points but may be outcompeted at later stages. On the other hand, *TP53* aberrations seem to confer a resistance mechanism potent enough to withstand treatment and clonal competition at any stage of the disease.

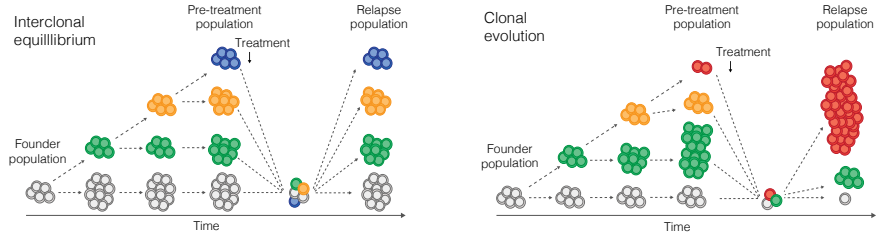

Figure 2. Two models of temporal dynamics in CLL relapsing after treatment as proposed by Schuh et al. and Landau et al.

The importance of subclonal lesions in *TP53* was further elucidated in a recent study where ultra-deep targeted NGS of the gene was applied to samples from 309 patients newly diagnosed with CLL[110]. This allowed for detection of variants with allelic frequencies down to ~0.3% that would not be detectable by Sanger sequencing. Small *TP53* mutated subclones were identified in 9% of untreated CLL and patients carrying subclonal *TP53* mutations showed the same poor survival as patients harboring clonal *TP53* lesions. Longitudinal analysis also revealed that these small subclones identified before treatment expanded under the selective pressure of treatment and became the dominant clone at the time of relapse[110]. Using a similar experimental design, Malcikova et al. were able to reproduce these results in an independent cohort of 330 patients[111]. The unequivocal results from these studies, together with the evidence from WES studies of inevitable selection for clones carrying *TP53* aberrations, clearly indicate that future clinical tests should be sensitive enough to detect minor populations harboring potent driver mutations.

The advent of drugs targeting the BcR pathway holds great promise for the treatment of CLL. However, early studies have demonstrated resistance against ibrutinib through mutations at the drug-binding site in *BTK* or in the downstream signaling molecule *PLCG2*[112,113]. Burger et al. recently presented the first data on clonal evolution in relation to BcR inhibition. WES and deep sequencing of five patients relapsing from ibrutinib treatment confirmed the role of *BTK* and *PLCG2* mutations, however only in two of these cases[114]. The remaining tumors displayed a variety of known and potential novel driver lesions, including del(8p), thus highlighting that the heterogeneity of CLL matters even with targeted treatment and that investigation of larger cohorts will be necessary in order to fully outline the resistance mechanisms to BcR inhibition[114].

Prognostic markers

The route to establishing an accurate diagnosis is usually straightforward in CLL patients. However, the process of deciding the best therapeutic option is complicated by the clinical heterogeneity of the disease, indicating the need for novel prognostic and predictive tools to aid in treatment decisions[4]. Two staging systems, Rai and Binet, are used in clinical practice today[115,116]. However, the majority of CLL cases are diagnosed at early clinical stages and the applicability in this patient category is questioned as these systems poorly discriminate aggressive from indolent disease in asymptomatic patients[4]. Many molecular studies have aimed at refining the prognostication of CLL resulting in a wealth of suggested clinical and biological markers[117]. Several of these markers have proven to carry independent prognostic value; however before a prognostic/predictive marker can be introduced into routine practice, it has to be validated in prospective clinical studies and each marker must undergo standardization as well as assessment of its feasibility for routine practice[118]. Out of all these potential markers, only a few selected are widely accepted and applied today and most of these are based on analysis of the tumor genome.

IGHV gene mutational status

As previously alluded to, the IGHV gene mutational status is one of the most robust prognostic markers in CLL as it is detectable at diagnosis, remains stable throughout disease, is independent of clinical stage and other biomarkers, and it is today regarded as the gold-standard prognostic marker[119-121]. It has been thoroughly demonstrated that U-CLL displays shorter TTFT and OS and moreover, more recent studies have revealed longer progression-free survival (PFS) from FCR treatment in M-CLL compared to U-CLL[122,123].

Genomic aberrations

The four recurrent cytogenetic aberrations detailed above are of prognostic significance and can stratify patients into risk groups according to the Döhner hierarchical model[67]. Cases with del(17p) have the worst prognosis followed by del(11q). In contrast, isolated del(13q) is associated with favorable outcome also compared to patients with normal karyotype.

Trisomy 12 is associated with shorter time to disease progression but this aberration does not appear to be associated with a worse overall outcome[67]. Certain genomic aberrations can also predict treatment response in CLL. In the German CLL8 study, when FCR was compared to FC, presence of del(17p) was still the strongest negative prognostic factor and associated with particularly poor outcome after chemoimmunotherapy. In contrast, patients harboring del(11q), del(13q) and trisomy 12 benefited greatly from FCR[109]. As previously mentioned, the tumor suppressor gene *TP53* is the main target of del(17p). In approximately 5% of all patients, *TP53* mutations are detected in the absence of del(17p) and patients harboring these mutations have equally adverse prognosis and response to treatment[83,84,124]. Hence, screening of del(17p) and *TP53* mutations are recommended before initiation of treatment[3,125].

Prognostic value of novel mutations

Studies have revealed that CLL patients harboring *NOTCH1* mutations have shorter PFS and OS. Cases carrying *NOTCH1* mutations require treatment more frequently and earlier than unmutated cases, but TTFT does not seem to be independent of IGHV mutations status[88]. Treatment refractoriness is seen more frequently in *NOTCH1*-mutated cases and they do not reach negative minimal residual disease to the same extent as *NOTCH1*-unmutated cases. The ratio of complete or partial response to treatment does however not seem to be dependent on *NOTCH1* mutations[91,126]. The mutations may carry predictive value as it has been demonstrated that *NOTCH1* mutated patients had decreased benefit of addition of the CD20 antibody rituximab in the CLL8 trial[127]. This phenomenon could be explained by the recent discovery that *NOTCH1* mutated cells have decreased CD20 expression levels[128]. *SF3B1* mutations are associated with unmutated IGHV genes and are detected in a higher frequency in patients with advanced clinical stage[88,129]. *SF3B1* mutations predict for shorter time to disease progression and OS, and seem to be related to refractoriness to fludarabine-based therapy as they were more frequently found in refractory cases (17-30%) than newly diagnosed cases (5-10%)[91,130]. *BIRC3* lesions have been associated with chemorefractoriness and aggressive disease and a prognostic impact similar to *TP53* aberrations have been observed[79,129]. However, in the UK LRF CLL4 study, *BIRC3* deletion and/or mutation did not have an impact on OS and PFS in del(11q) patients treated with first-line therapy[131]. Hence, as with several novel mutations in CLL, the exact prognostic relevance remains to be fully elucidated. Finally, it has been suggested that *MYD88* mutations are associated with younger age at diagnosis and good outcome[94]. However, these results have not been successfully replicated in independent cohorts and the prognostic relevance of *MYD88* mutations remains ambiguous[132].

Prognostic indices

The rapidly expanding number of biomarkers with potential prognostic relevance in CLL has led to a great interest in the development of prognostic indices with the ability of combining multiple variables into efficient tools for predicting OS and/or TTFT. There have been numerous suggestions of novel prognostic schemes in recent years and two of these will be discussed below.

Rossi et al. presented an algorithm based on gene mutations in combination with cytogenetic aberrations that suggested a hierarchical model with four risk groups: i) a high risk group with patients displaying *TP53* and/or *BIRC3* aberrations, ii) intermediate risk patients that carried *NOTCH1* and/or *SF3B1* mutations and/or del(11q), iii) low risk patients with trisomy 12 or no chromosomal aberrations, and iv) very low risk patients with isolated del(13q). Through sequential analysis, the authors also demonstrated that the model retained prognostic relevance over time regardless of clonal evolution[129]. However, the model has not been successfully validated in independent studies as overlapping OS and TTFT were observed between the different risk groups[88,133].

The International CLL-IPI Study Group has recently proposed a novel prognostic index that identifies five independent prognostic factors, namely: presence of *TP53* aberrations, IGHV mutational status, high or low serum β2-microglobulin concentration, Binet A/Rai 0 vs Binet B-C/Rai I-IV, and age >65 years. A scoring method was subsequently developed based on the independent factors allowing separation into four risk-groups that predict OS[134]. The index was successfully validated in two independent cohorts and its ability to also predict TTFT in early-stage patients was recently demonstrated[135]. However, prospective validation in large patient cohorts, including balanced number of patients with indolent disease, is required before the index can be suggested for introduction in clinical practice.

A prognostic index that integrates biomarkers from different modalities will most likely be necessary in order to stratify CLL patients in the most efficient way. However, the construction of the 'perfect' prognostic index will require detailed studies of very large patient cohorts that evenly represent all stages and subgroups of the disease.

Treatment of CLL

The vast majority (~85%) of CLL patients are asymptomatic upon diagnosis and will not be treated as long as they do not present any symptoms; approximately 40% of all CLL patients will never require treatment[4]. One of the major clinical challenges is to, as early as possible, accurately define patients that are in need of treatment and to offer them the most suitable treatment regimen available.

According to current best clinical practice, treatment should only be initiated in patients presenting with active disease according to the criteria defined by the International Workshop on Chronic Lymphocytic Leukemia[3] (iwCLL). The criteria include evidence of progressive bone marrow failure, bulky disease, non-responsive autoimmune anemia and/or thrombocytopenia, constitutional symptoms, and progressive lymphocytosis. For patients in need of treatment, the choice of treatment regimen is based on the physical condition of the patient and the status of the tumor suppressor gene *TP53*. FCR is the gold-standard first-line regimen for physically fit patients below 65 years of age and with no *TP53* aberration. Patients in the same category but with *TP53* aberration present are instead recommended treatment with BcR signaling inhibitors (e.g. ibrutinib) and anti-CD20 antibodies[136,137]. This group of patients can also be eligible for allogeneic stem cell transplantation. Chlorambucil is the main recommendation for elderly and/or frail patients and combination with CD20 antibodies is recommended if tolerated[3].

Assessment of response to treatment is also conducted according to the 2008 iwCLL guidelines[3]. The main treatment response categories are complete remission (CR; peripheral blood lymphocyte count $\leq 4 \times 10^9$/L, absence of constitutional symptoms, absence of significant lymphadenopathy, normalized blood counts, and no hepato- or splenomegaly, for at least 2 months after completion of therapy), partial remission (PR; 50% reduction in peripheral blood lymphocyte count, reduction in lymphadenopathy, improved blood counts, 50% reduction in hepato- or splenomegaly if present pretreatment, for a duration of at least 2 months after completions of therapy), progressive disease (PD; 50% increase in peripheral blood lymphocytes, appearance of a new lesion such as enlarged lymph node or hepato- or splenomegaly, transformation into a more aggressive disease form, or occurrence of cytopenia). The classification stable disease (SD)

defines cases where CR or PR is not achieved but without evidence of PD. The 2008 iwCLL guidelines defines disease relapse as a patient who has previously achieved CR or PR but exhibits evidence of PD after a period of 6 or more months[3].

As previously mentioned, the current first-line therapy for fit patients is FCR. This chemoimmunotherapy approach where purine analogues, alkylating agents, and monoclonal antibodies are combined, have achieved response rates of about 90% but most patients still inevitably relapse[109,138,139]. Recently, novel drugs targeting BcR signaling such as ibrutinib, an oral, selective and irreversible inhibitor of Bruton's tyrosine kinase (BTK), and idelalisib, an inhibitor of the PI3K catalytic subunit-δ, have been proven highly effective in relapsed as well as in treatment-naive CLL, including high-risk cases with *TP53* aberrations[136,137,140]. However, due to an increased rate of adverse events associated with idelalisib treatment, the drug is currently only approved as second-line treatment in high-risk groups of CLL. The first results from the second-generation BTK inhibitor acalabrutinib are promising with 95% overall response rates in relapsed CLL[141]. Other novel therapeutic options include the second-generation CD20 antibody ofatumumab and the selective BCL2-inhibitor venetoclax[142,143].

Next-generation sequencing technologies

Studies of cancer over the last decades have revealed that one of the most fundamental mechanisms of the disease is the stepwise accumulation of genetic lesions[144]. Because of this, the ability to detect somatically acquired genetic alterations specific to the tumor cells is important for improving our understanding of the disease biology but also to improve diagnostic and prognostic/predictive tools for cancer patients[145]. Sanger sequencing has long been considered the gold standard for mutation detection and while it is robust and yields highly reliable results, the throughput and genomic coverage is limited and variants with low allelic frequencies can usually not be detected[146].

NGS technologies offer much greater throughput and sensitivity than previously available techniques, enabling the analysis of complete cancer genomes, exomes, or transcriptomes in one sequencing run[147]. As the name implies, WGS offers genomic coverage of all sequencable regions in the genome, however usually with depth of coverage of only approximately 30x, hence mutations with low allelic frequencies are challenging to detect. On the other side of the NGS technique spectrum is targeted deep-sequencing where selected genomic regions are captured and sequenced to a high depth, usually over 1000x. This allows for multiplexing of more samples and superior detection of subclonal mutations but relies on *a priori* knowledge of which regions to include in the analysis. Finally, WES has emerged as one of the most popular library techniques since it offers a combination of the two aforementioned methods. This technique is based on target capture of all known coding regions of the genome and renders superior depth of coverage compared to WGS but with much broader genomic coverage than targeted deep sequencing[147].

Table 2. Comparison of WGS, WES, and target enrichment

Technique	Target coverage	Target size	Sequencing required
WGS	30X	~3300Mb	~100Gb
WES	100X	~60Mb	~10Gb
Targeted deep seq	1000X	< 0.5Mb	< 0.5Gb

The most fundamental characteristic of NGS techniques is massive parallelization, meaning the ability of sequencing of millions of genomic fragments in parallel and computationally aligning the reads to a pre-constructed reference or creating full length contigs de-novo. This enables reading each nucleotide of interest with a high degree of redundancy that in turn allows for the detection of mutations with low allelic frequencies. There are several commercially available NGS platforms today, but many large-scale sequencing efforts in cancer have utilized the Illumina sequencing technique (Illumina, San Diego, CA)[148].

NGS techniques are now being introduced into clinical practice for detection of somatic mutations of prognostic relevance in cancer and for facilitating diagnostics concerning inherited diseases[149-151]. Although the techniques are the same as those used in research projects, implementation of NGS in clinical tests requires several additional criteria to be fulfilled. First, library preparation and sequencing should be fully standardized to ensure that data quality and yield is sufficient for high-confidence variant calling. Second, the minimal requirements to report the status of a genomic position should be well defined and these requirements should account for disease biology (e.g. subclonality) and sampling issues (e.g. impurity of the tumor sample). Finally, all steps of the method need to be robust and reproducible, both regarding laboratory procedures and the computational analysis.

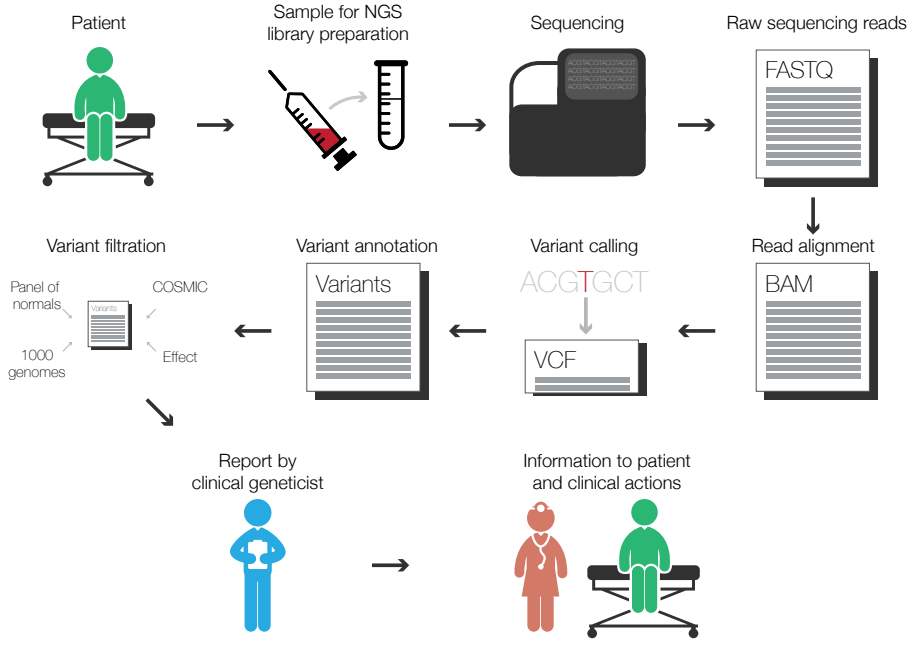

Figure 3. Principal workflow of clinical NGS tests.

Computational analysis of NGS data

One of the major hurdles in interpreting NGS results is the amount of data that is produced. Compared to Sanger sequencing results that can be analyzed manually using specialized software, NGS data analysis involves steps that would not be feasible without computational power. Depending on the input entity (e.g. DNA or RNA) and the type of sequencing data generated (WGS, targeted sequencing or transcriptome sequencing), different processing steps are needed and there is no universal analysis pipeline suitable for all types of data. However, some processing steps are required in all analysis pipelines for DNA sequencing, namely I) base calling and read demultiplexing which can be handled internally in many sequencing machines today, II) quality control to ensure sufficient quality and yield of the data for each run and sample, III) removal of sequencing adapters, IV) alignment of sequencing reads to a reference sequence, V) variant calling, and VI) variant annotation. All processing steps can to some extent affect the end results but alignment, variant calling and variant annotation have the largest impact[152].

High quality alignment is fundamental for good downstream results and the challenge is often to find a balance between alignment stringency and the yield of aligned reads. Pseudogenes and repetitive genomic regions can lead to false positive variants if too promiscuous read alignments are allowed and conversely, too stringent settings can discard variants that affect a large proportion of the read, such as large insertions and deletions. There are many variant calling algorithms available with different strategies to acquire high sensitivity and specificity but they can broadly be divided by their ability to detect germline or somatic variants. The previous are built for direct comparison of sequencing data to a known reference and reporting all positions where a mutation is found, whereas the latter compares a tumor sequencing file to a matched normal sequencing file and a known reference to detect somatically acquired mutations. Using a matched high-quality normal sample for variant filtering would be the ideal scenario for detection of cancer mutations in clinical practice, but because of the increased cost this is not always feasible. Finally, mutations are annotated to identify the genes and transcripts they affect and to predict the potential change in the amino acid sequence. Variant information from previous large scale sequencing studies of healthy individuals can also be added to filter out germline variants that most likely are of no clinical significance. Private variants and variants specific to minor ethnical populations might not be detected by this technique however, and caution is required when interpreting the results without germline DNA[152,153].

Present investigations

Thesis aims

The aim of this thesis was to explore NGS as a novel strategy for accurate screening of recurrent mutations in CLL, and for the discovery of novel genetic mechanisms associated with disease evolution and resistance to FCR treatment. The specific aims of each paper were:

I To evaluate the technical performance of targeted NGS for mutational screening in CLL and to assess its concordance with Sanger sequencing.

II To apply targeted NGS to investigate the genetic background of the constitutional NF-κB activation in CLL.

III To characterize the genetic mechanisms underlying relapse after treatment with FCR using WES on longitudinal samples.

IV To delineate the genomic landscape in the stereotyped subsets #1, #2 and #4 using WGS.

Materials and methods

Patient material

All patients included in the studies were diagnosed according to iwCLL criteria and displayed a typical CLL immunophenotype. Informed consent was collected according to the Helsinki Declaration and ethical approval was granted by local ethical review committees.

Paper I encompassed 188 CLL patients with poor-prognostic features (73% IGHV unmutated, 27% IGHV3-21 subset #2) that are part of an international multicenter cohort. **Paper II** included 692 CLL cases that were collected from collaborating institutions in Sweden, Greece, Italy, France, Czech Republic, the Netherlands, the USA, and the UK. This study also included 136 mantle cell lymphoma cases, 66 diffuse large B cell lymphoma (DLBCL) cases and 170 splenic marginal zone lymphoma cases, all diagnosed according to the WHO classification. **Paper III** included a discovery cohort of 41 CLL cases that had received FCR treatment and had either obtained a complete remission (CR; n = 32) or partial remission (n = 9) and from which pre-treatment and relapse samples were available. Seven collaborating institutions in Sweden, Greece, Italy, France, the Czech Republic, the United Kingdom, and Germany provided the samples. Two independent validation cohorts were also selected. First, 790 CLL cases and 30 cases with Richter transformation were provided by the aforementioned institutions and from additional institutions in Greece, The Netherlands, and the United states. Second, samples from 329 untreated patients from the UK LRF CLL4 clinical trial were investigated. **Paper IV** comprised 76 patients from collaborating institutions in the Netherlands, the Czech Republic, the United Kingdom, Italy, Sweden, Greece and France were selected based on the expression of stereotyped BcR IGs and assignment to one the major subsets #1 (n=27), #2 (n=28) or #4 (n=21) as previously defined.

Methods

Library preparation and sequencing

The HaloPlex (Agilent Technologies, Santa Clara, CA) target enrichment system[154] was used for mutational screening in **Paper I and Paper II**. Target probes were designed using the SureDesign service (https://earray.chem.agilent.com/suredesign/home.htm). Capture of the target regions was performed using the HaloPlex kit and following the manufacturer's instructions. In **Paper III**, exome capture and library generation was accomplished using the TruSeq Exome Enrichment Kit (Illumina, San Diego, CA). Finally, in **Paper IV**, WGS libraries were constructed using the TruSeq Nano Kit (Illumina, San Diego, CA). Paired-end sequencing was performed for all libraries on the Illumina HiSeq 2000 (**Paper I-III**) or the Illumina HiSeq 2500 or HiSeq X Ten system (**Paper IV**).

NGS data analysis

In **Paper I and II**, sequencing adaptors were trimmed from all raw sequencing reads using CutAdapt and aligned to the Hg19 human genome reference using the MOSAIK aligner[155]. Variants were identified using either Varscan2[156] or an in-house, purpose-built variant caller (SNPmania) and annotated using ANNOVAR[157]. Variant allele frequency (VAF) cutoff for mutational screening was set to 10%. In **Paper III** the sequencing reads were processed through the bcbio-nextgen framework. BWA-mem[158] (v0.7.10) was used for read alignment to the Hg19 human reference genome, realigned around indels using GATK (v3.2) and optical and PCR duplicates were marked using Sambamba (v0.4.7). Small mutations were detected using VarScan2 (v2.3.6) with a 10% VAF cutoff and sCNAs were derived from WES data using EXCAVATOR[159]. SciClone was used for clustering of the VAFs enabling analysis of clonal composition and dynamics at the time point before initiation of treatment and at relapse[160]. In **Paper IV**, the raw sequencing reads were either processed in Piper, a pipeline built on top of GATK queue or the WGS analysis pipeline provided by the Beijing Genomics Institute. The reads were aligned to the Grch37 or Hg19 reference genome using BWA v0.7.12 and deduplicated and realigned using GATK v3.3.0. SNVs and small indels were detected using VarScan2 with a 10% VAF cutoff. Computation intensive analyses were executed on the Uppmax high-performance cluster Milou.

Sanger sequencing and GeneScan analysis

In **Paper I-III**, screening and validation of small mutations was performed through Sanger sequencing. Selected variants were bi-directionally sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit and an ABI 3730 DNA Analyzer (Life Technologies, Carlsbad, CA). In **Paper II**, a GeneScan assay was designed for detection of the 4bp recurrent deletion in *NFKBIE*. Briefly, genomic DNA was amplified by hot-start PCR with Platinum-Taq DNA Polymerase (Invitrogen) and 60°C as annealing temperature. The fragment length of the PCR products was assessed by capillary electrophoresis with ABI3730XL DNA Analyzer (Applied Biosystems) and analyzed with Peak Scanner Software v1.0 (Applied Biosystems).

Statistical analysis

Differences in frequencies were evaluated using descriptive statistics. In Paper II and III, Paired Student's *t* test was used to assess differences between subgroups with at least four patients, and the Chi-square test was used to assess independence of association between categorical variables. In Paper II, the differences in *NFKBIE* mutation frequency among CLL subset cases were assessed using Friedman ANOVA. OS was calculated from the date of diagnosis until the date of last follow-up or death. Survival curves were constructed using the Kaplan-Meier method, and the log-rank test was used to determine differences between survival proportions. For all comparisons, p-values were two-sided and a significance level of $p<0.05$ was set. All statistical analyses were performed using Statistica software version 10 (Paper III) or version 12 (Paper II; Stat Soft).

Results and Discussion

Paper I - Targeted next-generation sequencing in CLL

Since CLL is recognized as a clinically heterogeneous disease, even though the diagnosis usually is straightforward, the need for precise markers to support therapy decisions and prognostication is still significant. Beyond mutations in *TP53*, whose prognostic relevance already have been established[83,84,124], large-scale sequencing studies and subsequent validation of large and well-characterized cohorts of CLL patients have revealed a number of recurrently mutated genes with prognostic potential. Mutations in *ATM*, *NOTCH1*, *SF3B1*, *MYD88* and *BIRC3* are currently under evaluation but appear to have prognostic significance in CLL[64-66,85,87,88,91,97,126,161]. *KLHL6*, *POT1*, and *XPO1* have been found recurrently mutated in CLL but their prognostic relevance is not yet established[64,65,87,97,133]. Mutational screening in clinical routine today is based on Sanger sequencing and while highly reliable, the low throughput makes analysis of several genes or genes with many exons unfeasible. Hence, targeted NGS is emerging as a much more efficient means to identify mutations. However, before NGS can be introduced in any clinical test, the analytical accuracy and reproducibility of the assay needs to be thoroughly evaluated.

To investigate if targeted NGS can be integrated in clinical routine practice, we designed a HaloPlex panel targeting all coding exons of the nine aforementioned genes and selected 188 CLL patients with poor-prognostic features for mutational screening. The HaloPlex libraries were sequenced on two lanes on the HiSeq2000 instrument (Illumina, San Diego, CA). The procedure was repeated in a new HaloPlex panel including six of the genes (*BIRC3*, *NOTCH1*, *POT1*, *SF3B1*, *TP53* and *XPO1*) and 63 of the CLL cases to estimate technical precision and reproducibility.

Sequencing of the libraries yielded an average on-target coverage of ~1500 reads/base. However, the coverage of targeted deep-sequencing is not evenly distributed and a detailed analysis of the coverage distribution revealed that ≥80% of the bases in the regions of interest were covered at least 100x in 180/188 samples. Adequate depth of coverage is also fundamental for accurate and confident variant calling and for clinical applications it should be ensured that each reported position has reached a set threshold.

Variant calling and filtering revealed mutation rates in line with previous screening efforts of aggressive CLL patient cohorts; *SF3B1* (37/180; 21%), *ATM* (35/180; 19%), *NOTCH1* (33/180; 18%), *TP53* (25/180; 14%), *XPO1* (14/180, 8%), *POT1* (9/180; 5%), *BIRC3* (4/180; 2%), *MYD88* (2/180; 1%) and *KLHL6* (1/180; 0.6%). We selected 155 mutations with VAF ranging from 10-99% for validation with Sanger sequencing and 143/155 (92%) could be confirmed. One explanation for these discrepancies is that most of the discordant variants had allelic frequencies at the detection border of Sanger sequencing (average, 16%; range, 11-27%) and dropout of the variant allele may have led to a homozygous reference call. When we compared the data from the second HaloPlex panel we found that 77/82 variants were concordant and that no additional variants could be detected. This suggests that the false-positive rate of the assay is good (0%) but that some refinement is needed to perfect the true-positive rate (94%).

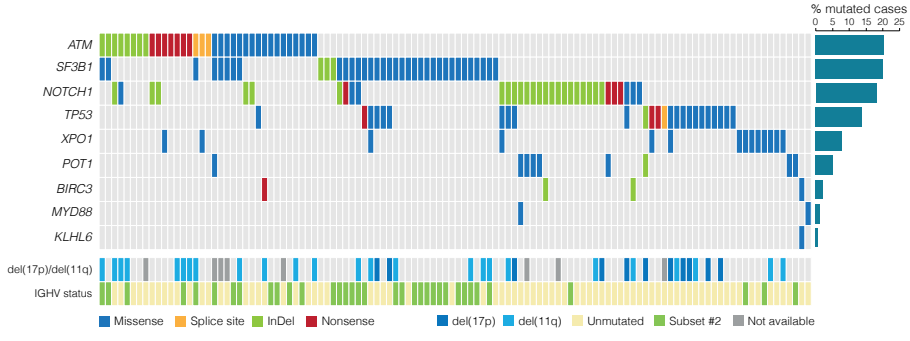

Figure 4. Brick plot depicting the mutations detected by targeted deep sequencing.

As previously discussed, depth of coverage is vital for the sensitivity of the assay and need to be optimal to prevent false-negative calls. The coverage yield is dependent on several pre-analytical conditions and thorough quality controls are important to ensure that the amount and quality of the input DNA is sufficient and that all library preparation steps are successful. Poor capture of certain genomic regions can affect the coverage and extra sets of probes can be added to the design to overcome this. Finally, the different thresholds in the bioinformatics workflow need to be balanced with regards to sensitivity and specificity and must respect both technical issues such as normal tissue contamination and biological phenomena such as intraclonal heterogeneity.

In conclusion, we demonstrate that targeted NGS can be applied for routine mutational screening in CLL. The future introduction into clinical routine will require stringent protocols for all analytical steps and during an initial phase all positive results will most likely require validation with Sanger

sequencing. However, as the technique matures we foresee that it will replace low throughput techniques and lead to more cost effective and sensitive mutational screening.

Paper II – Molecular basis of NF-κB deregulation in CLL

Nuclear factor-κB (NF-κB) signaling pathway is essential for several fundamental cellular processes such as cell survival, inflammation, and proliferation and aberrant NF-κB activation is a recurrent phenomenon in most mature B cell malignancies[162-165]. Even though the phenomenon is widely accepted, the genetic background explaining the deregulation in CLL is largely unknown. The only recurrent genetic lesions that have been shown to contribute to the activation are *BIRC3* aberrations in the noncanonical NF-κB pathway, and *MYD88* mutations the TLR signaling pathway[65,66,79,166]. In 2014, Damm et al. reported that ~11% of patients in a cohort of patients with advanced stage CLL carried mutations in the *NFKBIE* gene, which encodes IκBε, a negative regulator of NF-κB in B cells, but the functional and clinicobiological consequences of these mutations remained to be determined[59].

With this in mind, we sought to achieve a detailed characterization of the NF-κB signaling pathway in CLL through high-sensitive genomic techniques and functional assays. To this end, we designed a HaloPlex panel targeting 18 genes in the NF-κB core complex and performed targeted-deep sequencing in a discovery cohort comprising 124 CLL cases and a subsequent validation cohort of 191 CLL cases enriched for stereotyped subsets associated with poor prognosis. The most frequently mutated gene was *NFKBIE,* in which mutations were detected in 21/315 cases, and strikingly, 13/21 *NFKBIE* mutations constituted an identical 4bp deletion. Prompted by this finding, we next screened 377 CLL cases from both aggressive and more indolent disease subgroups using a GeneScan assay targeting the deletion and detected 22 additional *NFKBIE* deletions. Collectively, 43/692 (6.2%) patients carried aberrations in *NFKBIE* and the mutations were significantly enriched in poor-prognostic subgroups; for example, 37/43 (86%) *NFKBIE* aberrant cases concerned U-CLL and 17/112 (15%) of cases assigned to the poor-prognostic subset #1 carried *NFKBIE* aberrations. As expected from this enrichment, the clinical profile of *NFKBIE* aberrations was associated with significantly shorter TTFT, similar to U-CLL and del(17p). Longitudinal samples were available from 14 treated CLL cases and in 6/14 cases we observed an increasing VAF of *NFKBIE* mutations at relapse. Although the relevance of *NFKBIE* in relationship to

clonal evolution needs to be studied in better detail, this could indicate that the mutations provide clonal advantage, at least at certain disease stages.

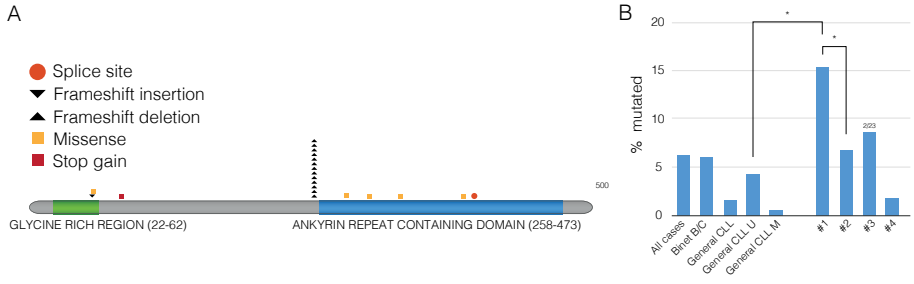

Figure 5. A) Locations of mutations detected in *NFKBIE*. B) Mutation frequencies in the study cohorts.

The physiological function of IκBε is to limit nuclear migration of NF-κB dimers (e.g. p65 and REL) through binding via the ankyrin repeats region and thereby providing negative regulation of NF-κB activation[162,167]. Since the 4bp *NFKBIE* deletion leads to truncation of the ankyrin repeat region and the majority of the other *NFKBIE* mutations clustered to this region we hypothesized that the mutations would lead to aberrant protein levels and NF-κB dimer interaction. Indeed, Western blot analysis revealed significantly lower IκBε levels and higher p65 levels in *NFKBIE* deleted patients compared to wild type (WT) patients. The p65 binding properties of WT and deleted IκBε were investigated through co-immunoprecipitation and the proximity ligation assay[168], both indicating a reduced interaction for the mutant protein. Finally, analysis of nuclear fractions of p50 and p65 revealed increased levels in *NFKBIE* deleted cases vs WT patients. Collectively, this points to a deactivating function of the mutations leading to increased levels of nuclear p50/p65 and subsequent NF-κB activation.

In summary, this study provides evidence of the genetic alterations leading to constitutional NF-κB activation in CLL with the prime finding being the recurrent 4bp *NFKBIE* deletion. We also describe the functional consequences of the deletion with the major effect being reduced IκBε expression and IκBε-p65 interaction, which in turn lead to increased phosphorylated p65. Collectively, lesions within the NF-κB may emerge as prognostic markers and components in the pathway could potentially constitute suitable targets for novel therapies in CLL.

Paper III – WES in relapsing CLL

FCR is the gold-standard first-line regimen in medically fit patients with CLL; however, despite good response rates most patients will eventually relapse, usually also acquiring a more aggressive disease phenotype. Until recently, the only molecular mechanism associated with FCR relapse and refractoriness was *TP53* aberrations, which were observed in up to 40% of fludarabine-refractory CLL[169,170]. NGS and targeted re-sequencing studies have revealed that mutations in *NOTCH1*, *SF3B1* and *BIRC3* are enriched in fludarabine refractory patients (44,53,54,76) and that approximately 10% of the cases harbor mutations in *FAT1*[79,86,130,171,172]. However, these studies have either analyzed biologically and clinically heterogeneous patient groups or only analyzed samples at the time of refractoriness. Very recently, Landau et al. presented the first large-scale genomic study including a high number of cases uniformly treated with FC or FCR. Longitudinal samples from pre-treatment and relapse were available for 59 cases and allowed characterization of temporal dynamics as previously discussed. The study also reported novel putative driver mutations such as in *RPS15,* although the functional impact of these mutations was not explored[87].

In order to characterize mutations enriched in patients relapsing from FCR treatment and to track clonal dynamics across treatment relapse, we collected pre-treatment and relapse samples from 41 patients receiving FCR with either partial response (PR, with ≥4 cycles of treatment required) or complete response (CR, ≥1 cycle of treatment required). The 28 samples with matched germline DNA allowed for detection of all somatic mutations within coding regions and revealed on average 15.2 non-silent mutations in the pre-treatment samples and 17.6 in the relapse samples. This indicates that chemoimmunotherapy do not increase the mutational load, which is in line with previous studies. However, closer inspection of the mutation classes revealed that the relapse-specific mutations were biased towards the C•G→A•T type, indicating that treatment may have an effect on the mutational spectrum. As expected, mutations associated with adverse prognosis in CLL were detected at high frequencies at relapse; *TP53* (n=8; 19.5%), *NOTCH1* (n=8; 19.5%), *ATM* (n=7; 17%), *SF3B1* (n=6; 14.6%), *NFKBIE* (n=4; 9.8%), *EGR2* (n=4; 9.8%) and *BIRC3* (n=3; 7.3%). Notably, we observed that *RPS15*, a gene only recently associated with CLL pathobiology, was mutated in a large proportion of the cases (n=8; 19.5%). The mutations in *RPS15* followed a distinct pattern with single nucleotide variants residing within a 7 amino-acid long evolutionary conserved region.

RPS15 encodes a component of the 40S ribosomal subunit and has, besides its role in protein translation, been shown to stabilize p53 by interacting with the MDM2-p53-MDMX network, thereby inhibiting MDM2-mediated p53

degradation[173]. Transcripts containing two recurrent *RPS15* mutations (p.P131S, p.G132A) were constructed using site directed mutagenesis. Transient expression of either WT or mutant *RPS15* in the HCT116 colorectal cancer cell line revealed impaired ability of $RPS15^{G132A}$ in regulating endogenous p53 with reduced p53 stability over time compared to WT transcript. Accordingly, both mutants revealed higher level of p53 ubiquitination compared to WT RPS15. Further functional analyses are required in order to reveal the complete functional impact of *RPS15* mutations as the lesions also could affect the ribosomal machinery. However, our results suggest that at least part of the oncogenic effect of the mutations could stem from impaired binding of mutant RPS15 to MDM2 or MDMX compared to WT protein leading to more pronounced p53 degradation.

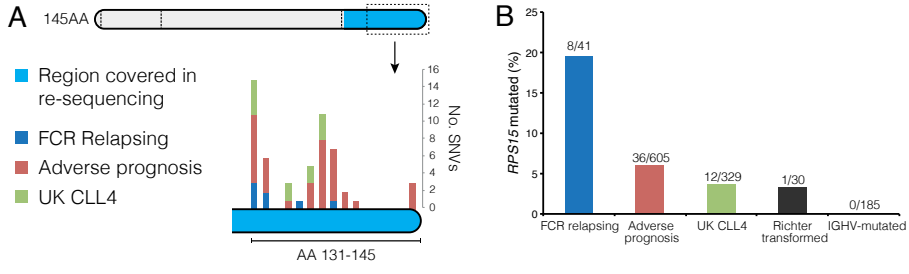

Figure 6. A) Locations of mutations detected in *RPS15* by WES and targeted re-sequencing. B) Frequencies of *RPS15* mutations in the study cohorts.

To assess the temporal dynamics of the detected mutations, 25 of the cases with matched germline DNA were subjected to analysis of clonal composition using SciClone. Each case harbored between 3 to 7 (sub)clones and relapse-specific mutations were detected in all samples. Relapse-specific subclones had expanded to become the dominant clone in 14 cases, indicating that they harbored mutations that made them evolutionarily superior. Inspection of the recurrently mutated genes revealed that mutations in i) *RPS15* remained stable at clonal levels over time, ii) *TP53*, *EGR2*, *NOTCH1* and *BIRC3* expanded or remained stable, and iii) *SF3B1* and *ATM* exhibited a mixed pattern.

In summary, this study provides novel insight into the heterogeneous genetic landscape of CLL relapsing after FCR treatment. We demonstrate that the selective pressure of treatment leads to marked shifts in the (sub)clonal populations, often leading to the expansion of more aggressive subclones. Furthermore, our finding of recurrent mutations in *RPS15* and the evidence of their functional consequences in vitro points to a novel mechanism for p53 dysregulation in CLL.

Paper IV – WGS in stereotyped subsets of CLL

Although the coding regions of the CLL genome are well studied, some CLL cases do not exhibit any recurrent mutations, potentially alluding to the existence of driver mutations in non-coding parts of the genome. WGS studies are intrinsically challenging due to the vast amounts of data generated and the lack of efficient methods for prioritizing variants in non-coding regions. CLL adds to the complexity as the tumor cells, in addition to cancer-associated mutations, will carry physiological somatic alterations stemming from IG rearrangements and SHM. Puente et al. provided novel insight into the non-coding regions of the CLL genome in their study of 150 CLL and MBL genomes that resulted in the discovery of recurrent *NOTCH1* 3'UTR mutations and *PAX5* enhancer mutations. In this study, patients were sub grouped either based on disease stage (MBL vs CLL) or by their IGHV mutational status (U-CLL vs M-CLL)[64]. While these are widely accepted entities in the disease, we hypothesized that more detailed group classifications might be required in order to overcome the molecular heterogeneity. Studies of CLL cases carrying stereotyped BcRs have revealed similar clinicobiological profiles and also enrichment of recurrent mutations in certain subsets with the prime examples being *SF3B1* mutations in up to 40% of patients in subset #2 and *NOTCH1* mutations in up to 30% of subset #1 cases[30,35-37]. With this in mind, we selected 70 cases from the two adverse-prognosis subsets #1 (n=25) and #2 (n=26) and the indolent subset #4 (n=19) for WGS analysis.

Previous studies have shown that M-CLL cases have higher total mutation burden than U-CLL[64] and our findings support this as the unmutated subset #1 had on average ~1800 mutations per case and the mutated subset #4 had an average of ~2400 mutations per case ($p < 0.005$). As expected, Subset #2 that is a mix of U-CLL and M-CLL had an intermediate mutation burden of ~2100 mutations per case on average. However, the mean number of non-synonymous was similar across the subsets (14.3 in subset #1, 15.1 in subset #2, and 11.9 in subset #4) indicating that a high proportion of the mutational load in M-CLL cases likely stems from off target SHM and constitutes passenger mutations. Analysis of all SNVs enabled the detection of two mutational signatures that to a high degree resembled the signatures previously observed in CLL. Signature 1, with high similarity to the age-related signature involving C→T transitions at CpG sites, was mainly contributing in subset #1 cases and signature 2, defined by T→G transversions and resembling the signature observed in M-CLL, primarily concerned subset #4 cases whereas subset #2 showed evidence of both signatures. In signature based clustering analysis, subset #1 and subset #4 clustered separately from each other whereas subset #2 showed intermediate

behavior, once again indicating that the mutational processes leading to SHM is the likely contributor to the signatures.

Analysis of exonic mutations corroborated the previously observed subset-specific enrichment of recurrent mutations. *SF3B1* mutations were detected in 9 (13%) cases of which 7 belonged to subset #2. Four cases (6%) carried *NOTCH1* mutations and 3 of these belonged to subset #1. We also detected *TP53* mutations in 4 cases (6%; 2 in subset #1, 1 in subset #2, and 1 in subset #4) and *ATM* mutations in 3 cases (4%; 1 in subset #1 and 2 in subset #2). Three mutations were detected in the recently discovered driver gene *RPS15,* of which all cases belonged to subset #1, potentially pointing to a subset enrichment of these mutations. Finally, three subset #2 cases carried mutations in *CTCFL*, a gene not previously associated with CLL. The gene product is suggested to be involved in regulation of methylation but only low levels of *CTCFL* transcripts have been detected in gene expression profiling studies of B cells. Thirty-eight cases (19/25 subset #1, 17/26 subset #2, and 2/19 subset #4 cases) carried at least one exonic putative driver mutation suggested in two large-scale sequencing initiatives[64,87].

We therefore turned to mutations in non-coding regions to investigate if these could constitute alternative driving mechanisms. After filtration, we detected a set of recurrently affected gene regions with enrichment of mutations in subset #1 and #2. Intronic mutations in *MSI2* were detected in 9 patients of whom 6 belonged to subset #2. Interestingly, a novel splice-isoform of this gene has been reported in an RNAseq study of subset patients[174]. Intronic mutations in *CTBP2* were detected in 7 patients of whom 5 were assigned to subset #1 and intriguingly, the intronic mutations in the paralog *CTBP1* were detected in 4 patients, all assigned to subset #2. Taking putative exonic drivers and the 50 gene regions into account, all subset #1 and #2 cases were represented by at least one mutation.

In conclusion, we here present the first WGS initiative in patients assigned to stereotyped subsets. In line with previous reports, we acknowledge the influence of the SHM machinery in the mutational load and characteristics. Furthermore, we present novel candidate mutations in non-coding regions based on their enrichment in aggressive subsets that may be incorporated in future studies in order to characterize their potential functional impact.

Concluding remarks

The first descriptions of cancer can be found in papyrus rolls dating back to 1600 BC and the disease group has both scared and fascinated us for as long as we have known it. Driven by the wish to understand and also fight cancer, we have developed tools that have enabled investigation with gradually higher resolution, from the invention of the light microscope to array based techniques. As this thesis hopefully has conveyed, NGS techniques have led to major breakthroughs in the last decade and several key pieces have been added to the pathogenesis puzzle. With the case of CLL, we have learnt about pathways recurrently affected by genetic lesions, acquired important knowledge from the transcriptomic landscape, and seen examples of what epigenetic studies can teach us regarding disease ontogeny and clinical outcome.

In this thesis, we aimed at adding some pieces to the puzzle by investigating genetic mechanisms associated with well-defined subgroups of CLL. First, our findings from the NF-κB pathway show how recurrent genetic lesions can lead to constitutional pathway activation and explain the unfavorable clinical behavior associated with the mutations. Second, the results from the molecular landscape of CLL relapsing after FCR treatment were not unexpectedly both heterogeneous and complex. The revelation that aberrations in ribosomal proteins might play a role in the disease biology and in treatment resistance motivates further functional studies, as the mutations may also affect ribosomal function and translation stringency. Finally, the first WGS results from stereotyped subsets of CLL illustrate the complexity of the CLL genome as it contains a combination of cancer-associated and physiological somatic events. However, well-defined subgroups such as aggressive and indolent subset cases seem to hold potential in discriminating disease-related events.

The heterogeneity of the disease also constitutes a challenge for efficient prognostication and treatment decisions in clinical practice. The discovery of recurrent mutations with prognostic and/or predictive potential holds great promise for improving clinical decision-making. In order to perform stringent and sensitive mutational screening in clinical practice, well-established and reproducible techniques are needed. In this thesis we demonstrate that the performance of targeted deep sequencing is sufficient

for incorporation into clinical practice for mutational screening of CLL. However, even though many recurrent mutations have been detected, much work still remains in defining the exact catalogue of disease related mutations in CLL.

Insight from molecular studies and development of novel treatments shows that we are slowly honing in on the mechanisms central to CLL development and the scientific community has recently begun discussing if the first cases of cured CLL have been observed. Even though much work remains before a cure is truly available for all CLL cases, it is reassuring that the drug repertoire and the toolbox for clinical decision-making are rapidly increasing.

Acknowledgments

This work was carried out at the Department of Immunology, Genetics and Pathology, Rudbeck laboratory, Uppsala University.

I have had the benefit of working closely with two very skilled supervisors. I begin by expressing my gratitude to my main supervisor **Tobias Sjöblom**. Thank you for taking me in and for the valuable lessons about how research actually works, that one should never cheat with experimental design, and for helping me develop as a scientist. To my main co-supervisor **Richard Rosenquist,** thank you for the never-ending support throughout the last four years, for letting me work independently and develop my skills in genetics as well as bioinformatics, and for teaching me the values of strong collaborations. Thanks to both of you for believing in me - your mentoring skills complement each other perfectly.

To my co-supervisor **Larry Mansouri**, thank you for the great hands-on lessons on lab techniques (that I then fled from as quickly as I could), for teaching me how to survive in Rudbeck, and for being an inspiring colleague. Last but not least, to my co-supervisor **Mats Nilsson**: Thank you for always being excited by new findings and for always having ideas about new techniques that have led us forward.

To the past and present members of the MolHem group: **Diego** – thank you for these years spent in parallel, for almost constant confusion, and for great scientific ideas. **Emma** – thank you for putting up with our chaos and for your *realistic* view of the world. **Lesley** – thank you for keeping track of every major project and for guiding me through the first publication. **Panos** – your clinical knowledge has been invaluable, thank you for your patience with never-ending editing of lists and statistics. **Sujata** – thank you for the support throughout our years in the lab. **Tanja** – your honesty and knowledge is invaluable, thank you for all the input during these years. **Aron** – It won't be long until you school me in bioinformatics and I'm looking forward to that day. To our clinical branch: **Mattias** and **Karin**, it has been nice getting to know you, see you at Ackis soon. To the extension of the MolHem group: **Kostas** – thank you for always being frank in every discussion, it made me develop as a scientist. **Aliki, Andreas, Nikos, Anastasia and Stavroula** – it was great working with you, email and Skype

goes a long way. **Paolo and Jon** – thank you for involving me in interesting projects and for providing precious samples.

A big thanks to past and present members of the MolCan group: **Lucy** – for valuable discussions and for battling the Comdia project together. **Verónica** – for your support and company during late nights at the lab. **Chatarina, Ivaylo, Karin, Snehangshu, and Tom** – for valuable input on projects and great kick-off sessions.

To all the fantastic people at Klinisk Genetik and Clinical Sequencing: **Lucia, Elin, Lotte, Claes, Malin, Patrik, Johan, Britt-Inger** and **Eva** – thank you for eternal patience with delivering sequencing data and helping out with every thinkable NGS related question.

Thank you to all the staff working at IGP and in particular the head of department **Karin Forsberg Nilsson** for your support of the PhD students, and **Christina** and **Helene** for always helping us PhD students in times of confusion. Thank you also to the **PhD student council** for arranging conferences and symposia (and also for putting up with me). To all the **friends from the lab**, thanks for beer-clubs, game-nights and alternative journal clubs!

My **old and new friends from medical school and the clinic**, I want to mention you all but the pages are running out. Thank you for always being there and constantly making sure that I one day will come back to you in the clinic. Thank you for listening to my never-ending talks about new mutations in CLL. Extra thanks to the heroes that remained in Uppsala and provided moral support and dinners in times of need. The **Enköping-gang**: Thanks for being part of the football madness and for concert trips across the country.

To my family: Thank you **Mom** and **Dad** for all the love and support you have given me. Thank you for giving me the courage to dive into science and for teaching me that different is ok. To **Sara**, **Filip**, **Jocke** and **Nellie** – Thank you for your support and curiosity. I'm grateful to have you in Uppsala. **Folke** och **Louise**, jag är oändligt glad att jag har er. Utan er hade jag aldrig varit den jag är idag.

References

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

2 Chiorazzi, N., Rai, K. R. & Ferrarini, M. Chronic lymphocytic leukemia. *The New England journal of medicine* **352**, 804-815, doi:10.1056/NEJMra041720 (2005).

3 Hallek, M. *et al.* Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. *Blood* **111**, 5446-5456, doi:10.1182/blood-2007-06-093906 (2008).

4 Baliakas, P., Mattsson, M., Stamatopoulos, K. & Rosenquist, R. Prognostic indices in chronic lymphocytic leukaemia: where do we stand how do we proceed? *J Intern Med* **279**, 347-357, doi:10.1111/joim.12455 (2016).

5 Fabbri, G. & Dalla-Favera, R. The molecular pathogenesis of chronic lymphocytic leukaemia. *Nature reviews. Cancer* **16**, 145-162, doi:10.1038/nrc.2016.8 (2016).

6 Chiorazzi, N. & Ferrarini, M. Cellular origin(s) of chronic lymphocytic leukemia: cautionary notes and additional considerations and possibilities. *Blood* **117**, 1781-1791, doi:10.1182/blood-2010-07-155663 (2011).

7 Stevenson, F. K., Krysov, S., Davies, A. J., Steele, A. J. & Packham, G. B-cell receptor signaling in chronic lymphocytic leukemia. *Blood* **118**, 4313-4320, doi:10.1182/blood-2011-06-338855 (2011).

8 Kostareli, E., Gounari, M., Agathangelidis, A. & Stamatopoulos, K. Immunoglobulin gene repertoire in chronic lymphocytic leukemia: insight into antigen selection and microenvironmental interactions. *Mediterr J Hematol Infect Dis* **4**, e2012052, doi:10.4084/MJHID.2012.052 (2012).

9 Sutton, L. A. & Rosenquist, R. The complex interplay between cell-intrinsic and cell-extrinsic factors driving the evolution of chronic lymphocytic leukemia. *Seminars in cancer biology* **34**, 22-35, doi:10.1016/j.semcancer.2015.04.009 (2015).

10 Tonegawa, S. Somatic generation of immune diversity. *Biosci Rep* **8**, 3-26 (1988).

11 Weigert, M. *et al.* The joining of V and J gene segments creates antibody diversity. *Nature* **283**, 497-499 (1980).

12 Jung, D. & Alt, F. W. Unraveling V(D)J recombination; insights into gene regulation. *Cell* **116**, 299-311 (2004).

13 Wagner, S. D. & Neuberger, M. S. Somatic hypermutation of immunoglobulin genes. *Annu Rev Immunol* **14**, 441-457, doi:10.1146/annurev.immunol.14.1.441 (1996).

14 Klein, U. & Dalla-Favera, R. Germinal centres: role in B-cell physiology and malignancy. *Nat Rev Immunol* **8**, 22-33, doi:10.1038/nri2217 (2008).

15 Fais, F. *et al.* Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. *J Clin Invest* **102**, 1515-1525, doi:10.1172/JCI3009 (1998).
16 Schroeder, H. W., Jr. & Dighiero, G. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. *Immunol Today* **15**, 288-294, doi:10.1016/0167-5699(94)90009-4 (1994).
17 Kipps, T. J. *et al.* Developmentally restricted immunoglobulin heavy chain variable region gene expressed at high frequency in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* **86**, 5913-5917 (1989).
18 Dono, M. *et al.* Evidence for progenitors of chronic lymphocytic leukemia B cells that undergo intraclonal differentiation and diversification. *Blood* **87**, 1586-1594 (1996).
19 Hamblin, T. J., Davis, Z., Gardiner, A., Oscier, D. G. & Stevenson, F. K. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **94**, 1848-1854 (1999).
20 Damle, R. N. *et al.* Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* **94**, 1840-1847 (1999).
21 Chiorazzi, N. & Ferrarini, M. B cell chronic lymphocytic leukemia: lessons learned from studies of the B cell antigen receptor. *Annu Rev Immunol* **21**, 841-894, doi:10.1146/annurev.immunol.21.120601.141018 (2003).
22 Tobin, G. *et al.* Chronic lymphocytic leukemias utilizing the VH3-21 gene display highly restricted Vlambda2-14 gene use and homologous CDR3s: implicating recognition of a common antigen epitope. *Blood* **101**, 4952-4957, doi:10.1182/blood-2002-11-3485 (2003).
23 Messmer, B. T. *et al.* Multiple distinct sets of stereotyped antigen receptors indicate a role for antigen in promoting chronic lymphocytic leukemia. *J Exp Med* **200**, 519-525, doi:10.1084/jem.20040544 (2004).
24 Tobin, G. *et al.* Somatically mutated Ig V(H)3-21 genes characterize a new subset of chronic lymphocytic leukemia. *Blood* **99**, 2262-2264 (2002).
25 Ghiotto, F. *et al.* Remarkably similar antigen receptors among a subset of patients with chronic lymphocytic leukemia. *J Clin Invest* **113**, 1008-1016, doi:10.1172/JCI19399 (2004).
26 Stamatopoulos, K. *et al.* Over 20% of patients with chronic lymphocytic leukemia carry stereotyped receptors: Pathogenetic implications and clinical correlations. *Blood* **109**, 259-270, doi:10.1182/blood-2006-03-012948 (2007).
27 Rosen, A., Murray, F., Evaldsson, C. & Rosenquist, R. Antigens in chronic lymphocytic leukemia--implications for cell origin and leukemogenesis. *Seminars in cancer biology* **20**, 400-409, doi:10.1016/j.semcancer.2010.09.004 (2010).
28 Murray, F. *et al.* Stereotyped patterns of somatic hypermutation in subsets of patients with chronic lymphocytic leukemia: implications for the role of antigen selection in leukemogenesis. *Blood* **111**, 1524-1533, doi:10.1182/blood-2007-07-099564 (2008).
29 Agathangelidis, A. *et al.* Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. *Blood* **119**, 4467-4475, doi:10.1182/blood-2011-11-393694 (2012).
30 Baliakas, P. *et al.* Clinical effect of stereotyped B-cell receptor immunoglobulins in chronic lymphocytic leukaemia: a retrospective multicentre study. *Lancet Haematol* **1**, E74-E84 (2014).

31 Baliakas, P. *et al.* Not all IGHV3-21 chronic lymphocytic leukemias are equal: prognostic considerations. *Blood* **125**, 856-859, doi:10.1182/blood-2014-09-600874 (2015).
32 Lanemo Myhrinder, A. *et al.* A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. *Blood* **111**, 3838-3848, doi:10.1182/blood-2007-11-125450 (2008).
33 Chu, C. C. *et al.* Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood* **112**, 5122-5129, doi:10.1182/blood-2008-06-162024 (2008).
34 Kostareli, E. *et al.* Molecular evidence for EBV and CMV persistence in a subset of patients with chronic lymphocytic leukemia expressing stereotyped IGHV4-34 B-cell receptors. *Leukemia* **23**, 919-924, doi:10.1038/leu.2008.379 (2009).
35 Strefford, J. C. *et al.* Distinct patterns of novel gene mutations in poor-prognostic stereotyped subsets of chronic lymphocytic leukemia: the case of SF3B1 and subset #2. *Leukemia* **27**, 2196-2199, doi:10.1038/leu.2013.98 (2013).
36 Sutton, L. A. *et al.* Different spectra of recurrent gene mutations in subsets of chronic lymphocytic leukemia harboring stereotyped B-cell receptors. *Haematologica* **101**, 959-967, doi:10.3324/haematol.2016.141812 (2016).
37 Rossi, D. *et al.* Association between molecular lesions and specific B-cell receptor subsets in chronic lymphocytic leukemia. *Blood* **121**, 4902-4905, doi:10.1182/blood-2013-02-486209 (2013).
38 Rosenwald, A. *et al.* Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J Exp Med* **194**, 1639-1647 (2001).
39 Klein, U. *et al.* Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. *J Exp Med* **194**, 1625-1638 (2001).
40 Weill, J. C., Weller, S. & Reynaud, C. A. Human marginal zone B cells. *Annu Rev Immunol* **27**, 267-285, doi:10.1146/annurev.immunol.021908.132607 (2009).
41 Herve, M. *et al.* Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. *J Clin Invest* **115**, 1636-1643, doi:10.1172/JCI24387 (2005).
42 Catera, R. *et al.* Chronic lymphocytic leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. *Mol Med* **14**, 665-674, doi:10.2119/2008-00102.Catera (2008).
43 Baumgarth, N., Tung, J. W. & Herzenberg, L. A. Inherent specificities in natural antibodies: a key to immune defense against pathogen invasion. *Springer Semin Immunopathol* **26**, 347-362, doi:10.1007/s00281-004-0182-2 (2005).
44 Hardy, R. R. B-1 B cell development. *J Immunol* **177**, 2749-2754 (2006).
45 Caligaris-Cappio, F., Gobbi, M., Bofill, M. & Janossy, G. Infrequent normal B lymphocytes express features of B-chronic lymphocytic leukemia. *J Exp Med* **155**, 623-628 (1982).
46 Stall, A. M. *et al.* Ly-1 B-cell clones similar to human chronic lymphocytic leukemias routinely develop in older normal mice and young autoimmune (New Zealand Black-related) animals. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 7312-7316 (1988).

47 Seifert, M. *et al.* Cellular origin and pathophysiology of chronic lymphocytic leukemia. *J Exp Med* **209**, 2183-2198, doi:10.1084/jem.20120833 (2012).
48 Kulis, M. *et al.* Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. *Nature genetics* **44**, 1236-1242, doi:10.1038/ng.2443 (2012).
49 Oakes, C. C. *et al.* DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nature genetics* **48**, 253-264, doi:10.1038/ng.3488 (2016).
50 Strati, P. & Shanafelt, T. D. Monoclonal B-cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification. *Blood* **126**, 454-462, doi:10.1182/blood-2015-02-585059 (2015).
51 Nieto, W. G. *et al.* Increased frequency (12%) of circulating chronic lymphocytic leukemia-like B-cell clones in healthy subjects using a highly sensitive multicolor flow cytometry approach. *Blood* **114**, 33-37, doi:10.1182/blood-2009-01-197368 (2009).
52 Scarfo, L. & Ghia, P. What does it mean I have a monoclonal B-cell lymphocytosis?: Recent insights and new challenges. *Semin Oncol* **43**, 201-208, doi:10.1053/j.seminoncol.2016.02.013 (2016).
53 Ghia, P. *et al.* Monoclonal CD5+ and CD5- B-lymphocyte expansions are frequent in the peripheral blood of the elderly. *Blood* **103**, 2337-2342, doi:10.1182/blood-2003-09-3277 (2004).
54 Rawstron, A. C. *et al.* Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. *The New England journal of medicine* **359**, 575-583, doi:10.1056/NEJMoa075290 (2008).
55 Vardi, A. *et al.* Immunogenetics shows that not all MBL are equal: the larger the clone, the more similar to CLL. *Blood* **121**, 4521-4528, doi:10.1182/blood-2012-12-471698 (2013).
56 Dagklis, A. *et al.* The immunoglobulin gene repertoire of low-count chronic lymphocytic leukemia (CLL)-like monoclonal B lymphocytosis is different from CLL: diagnostic implications for clinical monitoring. *Blood* **114**, 26-32, doi:10.1182/blood-2008-09-176933 (2009).
57 Kalpadakis, C. *et al.* New insights into monoclonal B-cell lymphocytosis. *Biomed Res Int* **2014**, 258917, doi:10.1155/2014/258917 (2014).
58 Barrio, S. *et al.* Genomic characterization of high-count MBL cases indicates that early detection of driver mutations and subclonal expansion are predictors of adverse clinical outcome. *Leukemia*, doi:10.1038/leu.2016.172 (2016).
59 Damm, F. *et al.* Acquired initiating mutations in early hematopoietic cells of CLL patients. *Cancer Discov* **4**, 1088-1101, doi:10.1158/2159-8290.CD-14-0104 (2014).
60 Kikushige, Y. *et al.* Self-renewing hematopoietic stem cell is the primary target in pathogenesis of human chronic lymphocytic leukemia. *Cancer Cell* **20**, 246-259, doi:10.1016/j.ccr.2011.06.029 (2011).
61 Genovese, G. *et al.* Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *The New England journal of medicine* **371**, 2477-2487, doi:10.1056/NEJMoa1409405 (2014).
62 Jaiswal, S. *et al.* Age-related clonal hematopoiesis associated with adverse outcomes. *The New England journal of medicine* **371**, 2488-2498, doi:10.1056/NEJMoa1408617 (2014).
63 Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415-421, doi:10.1038/nature12477 (2013).

64 Puente, X. S. *et al.* Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526**, 519-524, doi:10.1038/nature14666 (2015).
65 Puente, X. S. *et al.* Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475**, 101-105, doi:10.1038/nature10113 (2011).
66 Wang, L. *et al.* SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *The New England journal of medicine* **365**, 2497-2506, doi:10.1056/NEJMoa1109016 (2011).
67 Dohner, H. *et al.* Genomic aberrations and survival in chronic lymphocytic leukemia. *The New England journal of medicine* **343**, 1910-1916, doi:10.1056/NEJM200012283432602 (2000).
68 Juliusson, G. *et al.* Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. *The New England journal of medicine* **323**, 720-724, doi:10.1056/NEJM199009133231105 (1990).
69 Zech, L. & Mellstedt, H. Chromosome 13--a new marker for B-cell chronic lymphocytic leukemia. *Hereditas* **108**, 77-84 (1988).
70 Gahrton, G., Robert, K. H., Friberg, K., Zech, L. & Bird, A. G. Extra chromosome 12 in chronic lymphocytic leukaemia. *Lancet* **1**, 146-147 (1980).
71 Ouillette, P. *et al.* The prognostic significance of various 13q14 deletions in chronic lymphocytic leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research* **17**, 6778-6790, doi:10.1158/1078-0432.CCR-11-0785 (2011).
72 Puiggros, A., Blanco, G. & Espinet, B. Genetic abnormalities in chronic lymphocytic leukemia: where we are and where we go. *Biomed Res Int* **2014**, 435983, doi:10.1155/2014/435983 (2014).
73 Calin, G. A. *et al.* A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *The New England journal of medicine* **353**, 1793-1801, doi:10.1056/NEJMoa050995 (2005).
74 Cimmino, A. *et al.* miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 13944-13949, doi:10.1073/pnas.0506654102 (2005).
75 Reddy, K. S. Chronic lymphocytic leukaemia profiled for prognosis using a fluorescence in situ hybridisation panel. *Br J Haematol* **132**, 705-722, doi:10.1111/j.1365-2141.2005.05919.x (2006).
76 Shiloh, Y. ATM and related protein kinases: safeguarding genome integrity. *Nature reviews. Cancer* **3**, 155-168, doi:10.1038/nrc1011 (2003).
77 Austen, B. *et al.* Mutations in the ATM gene lead to impaired overall and treatment-free survival that is independent of IGVH mutation status in patients with B-CLL. *Blood* **106**, 3175-3182, doi:10.1182/blood-2004-11-4516 (2005).
78 Schaffner, C., Stilgenbauer, S., Rappold, G. A., Dohner, H. & Lichter, P. Somatic ATM mutations indicate a pathogenic role of ATM in B-cell chronic lymphocytic leukemia. *Blood* **94**, 748-753 (1999).
79 Rossi, D. *et al.* Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. *Blood* **119**, 2854-2862, doi:10.1182/blood-2011-12-395673 (2012).
80 Strati, P. *et al.* Second cancers and Richter transformation are the leading causes of death in patients with trisomy 12 chronic lymphocytic leukemia. *Clin Lymphoma Myeloma Leuk* **15**, 420-427, doi:10.1016/j.clml.2015.02.001 (2015).

81 Bieging, K. T., Mello, S. S. & Attardi, L. D. Unravelling mechanisms of p53-mediated tumour suppression. *Nature reviews. Cancer* **14**, 359-370, doi:10.1038/nrc3711 (2014).
82 Gaidano, G. *et al.* p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* **88**, 5413-5417 (1991).
83 Zenz, T. *et al.* Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. *Blood* **112**, 3322-3329, doi:10.1182/blood-2008-04-154070 (2008).
84 Malcikova, J. *et al.* Monoallelic and biallelic inactivation of TP53 gene in chronic lymphocytic leukemia: selection, impact on survival, and response to DNA damage. *Blood* **114**, 5307-5314, doi:10.1182/blood-2009-07-234708 (2009).
85 Quesada, V. *et al.* Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature genetics* **44**, 47-52, doi:10.1038/ng.1032 (2012).
86 Fabbri, G. *et al.* Analysis of the chronic lymphocytic leukemia coding genome: role of NOTCH1 mutational activation. *J Exp Med* **208**, 1389-1401, doi:10.1084/jem.20110921 (2011).
87 Landau, D. A. *et al.* Mutations driving CLL and their evolution in progression and relapse. *Nature* **526**, 525-530, doi:10.1038/nature15395 (2015).
88 Baliakas, P. *et al.* Recurrent mutations refine prognosis in chronic lymphocytic leukemia. *Leukemia* **29**, 329-336, doi:10.1038/leu.2014.196 (2015).
89 Cortese, D. *et al.* On the way towards a 'CLL prognostic index': focus on TP53, BIRC3, SF3B1, NOTCH1 and MYD88 in a population-based cohort. *Leukemia* **28**, 710-713, doi:10.1038/leu.2013.333 (2014).
90 Weng, A. P. *et al.* Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* **306**, 269-271, doi:10.1126/science.1102160 (2004).
91 Oscier, D. G. *et al.* The clinical significance of NOTCH1 and SF3B1 mutations in the UK LRF CLL4 trial. *Blood* **121**, 468-475, doi:10.1182/blood-2012-05-429282 (2013).
92 Balatti, V. *et al.* NOTCH1 mutations in CLL associated with trisomy 12. *Blood* **119**, 329-331, doi:10.1182/blood-2011-10-386144 (2012).
93 Ferreira, P. G. *et al.* Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. *Genome Res* **24**, 212-226, doi:10.1101/gr.152132.112 (2014).
94 Martinez-Trillos, A. *et al.* Mutations in TLR/MYD88 pathway identify a subset of young chronic lymphocytic leukemia patients with favorable outcome. *Blood* **123**, 3790-3796, doi:10.1182/blood-2013-12-543306 (2014).
95 Guieze, R. & Wu, C. J. Genomic and epigenomic heterogeneity in chronic lymphocytic leukemia. *Blood* **126**, 445-453, doi:10.1182/blood-2015-02-585042 (2015).
96 Ramsay, A. J. *et al.* POT1 mutations cause telomere dysfunction in chronic lymphocytic leukemia. *Nature genetics* **45**, 526-530, doi:10.1038/ng.2584 (2013).
97 Landau, D. A. *et al.* Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **152**, 714-726, doi:10.1016/j.cell.2013.01.019 (2013).

98 Lawrence, M. S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495-501, doi:10.1038/nature12912 (2014).
99 Nowell, P. C. The clonal evolution of tumor cell populations. *Science* **194**, 23-28 (1976).
100 Greaves, M. & Maley, C. C. Clonal evolution in cancer. *Nature* **481**, 306-313, doi:10.1038/nature10762 (2012).
101 Gunnarsson, R. *et al.* Array-based genomic screening at diagnosis and during follow-up in chronic lymphocytic leukemia. *Haematologica* **96**, 1161-1169, doi:10.3324/haematol.2010.039768 (2011).
102 Han, T. *et al.* Cytogenetic evidence for clonal evolution in B-cell chronic lymphocytic leukemia. *Cancer Genet Cytogenet* **23**, 321-328 (1986).
103 Grubor, V. *et al.* Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). *Blood* **113**, 1294-1303, doi:10.1182/blood-2008-05-158865 (2009).
104 Ouillette, P. *et al.* Clonal evolution, genomic drivers, and effects of therapy in chronic lymphocytic leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research* **19**, 2893-2904, doi:10.1158/1078-0432.CCR-13-0138 (2013).
105 Carter, S. L. *et al.* Absolute quantification of somatic DNA alterations in human cancer. *Nature biotechnology* **30**, 413-421, doi:10.1038/nbt.2203 (2012).
106 Landau, D. A. & Wu, C. J. Chronic lymphocytic leukemia: molecular heterogeneity revealed by high-throughput genomics. *Genome Med* **5**, 47, doi:10.1186/gm451 (2013).
107 Yates, L. R. & Campbell, P. J. Evolution of the cancer genome. *Nature reviews. Genetics* **13**, 795-806, doi:10.1038/nrg3317 (2012).
108 Schuh, A. *et al.* Monitoring chronic lymphocytic leukemia progression by whole genome sequencing reveals heterogeneous clonal evolution patterns. *Blood* **120**, 4191-4196, doi:10.1182/blood-2012-05-433540 (2012).
109 Hallek, M. *et al.* Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. *Lancet* **376**, 1164-1174, doi:10.1016/S0140-6736(10)61381-5 (2010).
110 Rossi, D. *et al.* Clinical impact of small TP53 mutated subclones in chronic lymphocytic leukemia. *Blood* **123**, 2139-2147, doi:10.1182/blood-2013-11-539726 (2014).
111 Malcikova, J. *et al.* Detailed analysis of therapy-driven clonal evolution of TP53 mutations in chronic lymphocytic leukemia. *Leukemia* **29**, 877-885, doi:10.1038/leu.2014.297 (2015).
112 Woyach, J. A. *et al.* Resistance mechanisms for the Bruton's tyrosine kinase inhibitor ibrutinib. *The New England journal of medicine* **370**, 2286-2294, doi:10.1056/NEJMoa1400029 (2014).
113 Furman, R. R. *et al.* Ibrutinib resistance in chronic lymphocytic leukemia. *The New England journal of medicine* **370**, 2352-2354, doi:10.1056/NEJMc1402716 (2014).
114 Burger, J. A. *et al.* Clonal evolution in patients with chronic lymphocytic leukaemia developing resistance to BTK inhibition. *Nat Commun* **7**, 11589, doi:10.1038/ncomms11589 (2016).
115 Rai, K. R. *et al.* Clinical staging of chronic lymphocytic leukemia. *Blood* **46**, 219-234 (1975).

116 Binet, J. L. *et al.* A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. *Cancer* **48**, 198-206 (1981).
117 Cramer, P. & Hallek, M. Prognostic factors in chronic lymphocytic leukemia-what do we need to know? *Nature reviews. Clinical oncology* **8**, 38-47, doi:10.1038/nrclinonc.2010.167 (2011).
118 Binet, J. L. *et al.* Perspectives on the use of new diagnostic tools in the treatment of chronic lymphocytic leukemia. *Blood* **107**, 859-861, doi:10.1182/blood-2005-04-1677 (2006).
119 Oscier, D. G. *et al.* Multivariate analysis of prognostic factors in CLL: clinical stage, IGVH gene mutational status, and loss or mutation of the p53 gene are independent prognostic factors. *Blood* **100**, 1177-1184 (2002).
120 Tobin, G. *et al.* Patients with chronic lymphocytic leukemia with mutated VH genes presenting with Binet stage B or C form a subgroup with a poor outcome. *Haematologica* **90**, 465-469 (2005).
121 Ghia, P. *et al.* ERIC recommendations on IGHV gene mutational status analysis in chronic lymphocytic leukemia. *Leukemia* **21**, 1-3, doi:10.1038/sj.leu.2404457 (2007).
122 Lin, K. I. *et al.* Relevance of the immunoglobulin VH somatic mutation status in patients with chronic lymphocytic leukemia treated with fludarabine, cyclophosphamide, and rituximab (FCR) or related chemoimmunotherapy regimens. *Blood* **113**, 3168-3171, doi:10.1182/blood-2008-10-184853 (2009).
123 Fischer, K. *et al.* Long-term remissions after FCR chemoimmunotherapy in previously untreated patients with CLL: updated results of the CLL8 trial. *Blood* **127**, 208-215, doi:10.1182/blood-2015-06-651125 (2016).
124 Rossi, D. *et al.* The prognostic value of TP53 mutations in chronic lymphocytic leukemia is independent of Del17p13: implications for overall survival and chemorefractoriness. *Clinical cancer research : an official journal of the American Association for Cancer Research* **15**, 995-1004, doi:10.1158/1078-0432.CCR-08-1630 (2009).
125 Pospisilova, S. *et al.* ERIC recommendations on TP53 mutation analysis in chronic lymphocytic leukemia. *Leukemia* **26**, 1458-1461, doi:10.1038/leu.2012.25 (2012).
126 Villamor, N. *et al.* NOTCH1 mutations identify a genetic subgroup of chronic lymphocytic leukemia patients with high risk of transformation and poor outcome. *Leukemia* **27**, 1100-1106, doi:10.1038/leu.2012.357 (2013).
127 Stilgenbauer, S. *et al.* Gene mutations and treatment outcome in chronic lymphocytic leukemia: results from the CLL8 trial. *Blood* **123**, 3247-3254, doi:10.1182/blood-2014-01-546150 (2014).
128 Pozzo, F. *et al.* NOTCH1 mutations associate with low CD20 level in chronic lymphocytic leukemia: evidence for a NOTCH1 mutation-driven epigenetic dysregulation. *Leukemia* **30**, 182-189, doi:10.1038/leu.2015.182 (2016).
129 Rossi, D. *et al.* Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. *Blood* **121**, 1403-1412, doi:10.1182/blood-2012-09-458265 (2013).
130 Rossi, D. *et al.* Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood* **118**, 6904-6908, doi:10.1182/blood-2011-08-373159 (2011).
131 Rose-Zerilli, M. J. *et al.* ATM mutation rather than BIRC3 deletion and/or mutation predicts reduced survival in 11q-deleted chronic lymphocytic leukemia: data from the UK LRF CLL4 trial. *Haematologica* **99**, 736-742, doi:10.3324/haematol.2013.098574 (2014).

132 Baliakas, P. *et al.* Prognostic relevance of MYD88 mutations in CLL: the jury is still out. *Blood* **126**, 1043-1044, doi:10.1182/blood-2015-05-648634 (2015).
133 Jeromin, S. *et al.* SF3B1 mutations correlated to cytogenetics and mutations in NOTCH1, FBXW7, MYD88, XPO1 and TP53 in 1160 untreated CLL patients. *Leukemia* **28**, 108-117, doi:10.1038/leu.2013.263 (2014).
134 International, C. L. L. I. P. I. w. g. An international prognostic index for patients with chronic lymphocytic leukaemia (CLL-IPI): a meta-analysis of individual patient data. *Lancet Oncol* **17**, 779-790, doi:10.1016/S1470-2045(16)30029-8 (2016).
135 Molica, S. *et al.* The chronic lymphocytic leukemia international prognostic index predicts time to first treatment in early CLL: Independent validation in a prospective cohort of early stage patients. *American journal of hematology*, doi:10.1002/ajh.24493 (2016).
136 Byrd, J. C. *et al.* Targeting BTK with ibrutinib in relapsed chronic lymphocytic leukemia. *The New England journal of medicine* **369**, 32-42, doi:10.1056/NEJMoa1215637 (2013).
137 Furman, R. R. *et al.* Idelalisib and rituximab in relapsed chronic lymphocytic leukemia. *The New England journal of medicine* **370**, 997-1007, doi:10.1056/NEJMoa1315226 (2014).
138 Badoux, X. C. *et al.* Fludarabine, cyclophosphamide, and rituximab chemoimmunotherapy is highly effective treatment for relapsed patients with CLL. *Blood* **117**, 3016-3024, doi:10.1182/blood-2010-08-304683 (2011).
139 Robak, T. *et al.* Rituximab plus fludarabine and cyclophosphamide prolongs progression-free survival compared with fludarabine and cyclophosphamide alone in previously treated chronic lymphocytic leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 1756-1765, doi:10.1200/JCO.2009.26.4556 (2010).
140 Advani, R. H. *et al.* Bruton tyrosine kinase inhibitor ibrutinib (PCI-32765) has significant activity in patients with relapsed/refractory B-cell malignancies. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **31**, 88-94, doi:10.1200/JCO.2012.42.7906 (2013).
141 Byrd, J. C. *et al.* Acalabrutinib (ACP-196) in Relapsed Chronic Lymphocytic Leukemia. *The New England journal of medicine* **374**, 323-332, doi:10.1056/NEJMoa1509981 (2016).
142 Coiffier, B. *et al.* Safety and efficacy of ofatumumab, a fully human monoclonal anti-CD20 antibody, in patients with relapsed or refractory B-cell chronic lymphocytic leukemia: a phase 1-2 study. *Blood* **111**, 1094-1100, doi:10.1182/blood-2007-09-111781 (2008).
143 Souers, A. J. *et al.* ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nat Med* **19**, 202-208, doi:10.1038/nm.3048 (2013).
144 Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546-1558, doi:10.1126/science.1235122 (2013).
145 Shen, T., Pajaro-Van de Stadt, S. H., Yeat, N. C. & Lin, J. C. Clinical applications of next generation sequencing in cancer: from panels, to exomes, to genomes. *Front Genet* **6**, 215, doi:10.3389/fgene.2015.00215 (2015).
146 Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nature biotechnology* **26**, 1135-1145, doi:10.1038/nbt1486 (2008).
147 Rizzo, J. M. & Buck, M. J. Key principles and clinical applications of "next-generation" DNA sequencing. *Cancer prevention research* **5**, 887-900, doi:10.1158/1940-6207.CAPR-11-0432 (2012).

148 Goodwin, S., McPherson, J. D. & McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nature reviews. Genetics* **17**, 333-351, doi:10.1038/nrg.2016.49 (2016).
149 Kohlmann, A. *et al.* Next-generation sequencing technology reveals a characteristic pattern of molecular mutations in 72.8% of chronic myelomonocytic leukemia by detecting frequent alterations in TET2, CBL, RAS, and RUNX1. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 3858-3865, doi:10.1200/JCO.2009.27.1361 (2010).
150 Kohlmann, A. *et al.* The Interlaboratory RObustness of Next-generation sequencing (IRON) study: a deep sequencing investigation of TET2, CBL and KRAS mutations by an international consortium involving 10 laboratories. *Leukemia* **25**, 1840-1848, doi:10.1038/leu.2011.155 (2011).
151 Yang, Y. *et al.* Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *The New England journal of medicine* **369**, 1502-1511, doi:10.1056/NEJMoa1306555 (2013).
152 Bao, R. *et al.* Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Inform* **13**, 67-82, doi:10.4137/CIN.S13779 (2014).
153 Genomes Project, C. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56-65, doi:10.1038/nature11632 (2012).
154 Johansson, H. *et al.* Targeted resequencing of candidate genes using selector probes. *Nucleic Acids Res* **39**, e8, doi:10.1093/nar/gkq1005 (2011).
155 Lee, W. P. *et al.* MOSAIK: a hash-based algorithm for accurate next-generation sequencing short-read mapping. *PloS one* **9**, e90581, doi:10.1371/journal.pone.0090581 (2014).
156 Koboldt, D. C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* **22**, 568-576, doi:10.1101/gr.129684.111 (2012).
157 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164, doi:10.1093/nar/gkq603 (2010).
158 Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589-595, doi:10.1093/bioinformatics/btp698 (2010).
159 Magi, A. *et al.* EXCAVATOR: detecting copy number variants from whole-exome sequencing data. *Genome Biol* **14**, R120, doi:10.1186/gb-2013-14-10-r120 (2013).
160 Miller, C. A. *et al.* SciClone: inferring clonal architecture and tracking the spatial and temporal patterns of tumor evolution. *PLoS computational biology* **10**, e1003665, doi:10.1371/journal.pcbi.1003665 (2014).
161 Skowronska, A. *et al.* Biallelic ATM inactivation significantly reduces survival in patients treated on the United Kingdom Leukemia Research Fund Chronic Lymphocytic Leukemia 4 trial. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **30**, 4524-4532, doi:10.1200/JCO.2011.41.0852 (2012).
162 Bonizzi, G. & Karin, M. The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends in immunology* **25**, 280-288, doi:10.1016/j.it.2004.03.008 (2004).
163 Compagno, M. *et al.* Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* **459**, 717-721, doi:10.1038/nature07968 (2009).

164 Staudt, L. M. Oncogenic activation of NF-kappaB. *Cold Spring Harbor perspectives in biology* **2**, a000109, doi:10.1101/cshperspect.a000109 (2010).
165 Rossi, D., Ciardullo, C. & Gaidano, G. Genetic aberrations of signaling pathways in lymphomagenesis: Revelations from next generation sequencing studies. *Seminars in cancer biology*, doi:10.1016/j.semcancer.2013.04.002 (2013).
166 Mansouri, L., Papakonstantinou, N., Ntoufa, S., Stamatopoulos, K. & Rosenquist, R. NF-kappaB activation in chronic lymphocytic leukemia: A point of convergence of external triggers and intrinsic lesions. *Seminars in cancer biology*, doi:10.1016/j.semcancer.2016.07.005 (2016).
167 Alves, B. N. *et al.* IkappaBepsilon is a key regulator of B cell expansion by providing negative feedback on cRel and RelA in a stimulus-specific manner. *J Immunol* **192**, 3121-3132, doi:10.4049/jimmunol.1302351 (2014).
168 Soderberg, O. *et al.* Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nature methods* **3**, 995-1000, doi:10.1038/nmeth947 (2006).
169 Stilgenbauer, S. & Zenz, T. Understanding and managing ultra high-risk chronic lymphocytic leukemia. *Hematology / the Education Program of the American Society of Hematology. American Society of Hematology. Education Program* **2010**, 481-488, doi:10.1182/asheducation-2010.1.481 (2010).
170 Zenz, T. *et al.* TP53 mutation and survival in chronic lymphocytic leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **28**, 4473-4479, doi:10.1200/JCO.2009.27.8762 (2010).
171 Schnaiter, A. *et al.* NOTCH1, SF3B1, and TP53 mutations in fludarabine-refractory CLL patients treated with alemtuzumab: results from the CLL2H trial of the GCLLSG. *Blood* **122**, 1266-1270, doi:10.1182/blood-2013-03-488197 (2013).
172 Messina, M. *et al.* Genetic lesions associated with chronic lymphocytic leukemia chemo-refractoriness. *Blood* **123**, 2378-2388, doi:10.1182/blood-2013-10-534271 (2014).
173 Daftuar, L., Zhu, Y., Jacq, X. & Prives, C. Ribosomal proteins RPL37, RPS15 and RPS20 regulate the Mdm2-p53-MdmX network. *PloS one* **8**, e68667, doi:10.1371/journal.pone.0068667 (2013).
174 Mansouri, L. *et al.* Next generation RNA-sequencing in prognostic subsets of chronic lymphocytic leukemia. *American journal of hematology* **87**, 737-740, doi:10.1002/ajh.23227 (2012).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine 1251

Editor: The Dean of the Faculty of Medicine

A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title "Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine".)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-302026

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2016