Doctoral dissertation

Identification of new cytogenetic subgroups in chronic lymphocytic leukemia by combining genomic, exomic and transcriptomic profiling techniques

Ana E. Rodríguez Vicente

With the approval of Salamanca University Faculty of Medicine, this thesis will be defended on July 23 2012, in the Lecture Hall, Centro de Investigación del Cáncer, Salamanca.

Supervisor: Prof. Dr. Jesús M. Hernández Rivas

VNiVERSiDAD D SALAMANCA

Table of contents

Preface

Chronic lymphocytic leukemia (CLL) is a heterogeneous disease with marked variability in the clinical course. The disease is also characterized by the presence of a genetical heterogeneity that is becoming apparent through studies of immunoglobulin heavy chain gene, chromosomal aberrations, microRNA deregulation and genetic abnormalities identified by whole genome sequencing. Furthermore, a strong relationship between specific genetic aberrations and the clinical course of CLL is observed.

Recent advances in genomics have transformed the research on hematologic malignancies by improving molecular approaches to gene networks. The recognition of novel molecular variables identified by the use of high-throughput molecular analytical techniques could contribute to a better knowledge of the pathogenesis of the disease and to the development of more accurate biological predictive factors. Thus microarrays have emerged as powerful tools for increasing the potential of standard methods through genome-wide biological studies. Array-CGH analysis has shown great promise as a tool for the analysis of genetic alterations in complex cancer genomes providing, in a single experiment, a general view of genomic gains and losses. Therefore the use of aCGH may detect new genetic lesions in CLL. Microarray expression profiling of CLL has been used to define the patterns of gene expression related to different clinical outcomes and chromosomal abnormalities. The global views of gene expression, when coupled with available knowledge about gene function, also provide the basis for formulating testable hypothesis about specific cellular pathways of CLL. Moreover, recently, the application of next generation sequencing technology has identified previously unknown recurrent mutations in CLL.

The general aim of this thesis was to analyze and to characterize the heterogeneity of CLL patients by applying high-throughput genomic technologies to gain knowledge in the molecular characteristics of CLL and their possible influence on the disease. The first part of this thesis contains a general introduction of the field, including the basics of the techniques applied for genomic profiling and next-generation sequencing, and different aspects of CLL. The second section specifies the goals of the thesis and shows the results and conclusions as original papers followed by a general discussion

Original papers

This thesis is based on the following papers, which in the text are referred to by their roman numerals:

I. Rodríguez AE, Robledo C, García JL, González M, Gutiérrez NC, Hernández JA, Sandoval V, García de Coca A, Recio I, Risueño A, Martín-Núñez G, García E, Fisac R, Conde J, de Las Rivas J, Hernández JM. **Identification of a novel recurrent gain on 20q13 in chronic lymphocytic leukemia by array CGH and gene expression profiling.** Ann Oncol. 2012 Jan 6. [Epub ahead of print].

II. Hernández JA*, Rodríguez AE*, González M, Benito R, Fontanillo C, Sandoval V, Romero M, Martín-Núñez G, de Coca AG, Fisac R, Galende J, Recio I, Ortuño F, García JL, de las Rivas J, Gutiérrez NC, San Miguel JF, Hernández JM. **A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia**. Haematologica. 2009 Mar;94(3):364-71.
*shared first authorship

III. Rodríguez AE, Hernández JA, Benito R, Gutiérrez NC, García JL, Hernández-Sánchez M, Risueño A, Sarasquete ME, Fermiñán E, Fisac R, García de Coca A, Martín-Núñez G, de las Heras N, Recio I, Gutiérrez O, de las Rivas J, González M, Hernández-Rivas JM. **Molecular characterization of chronic lymphocytic leukemia patients with a high number of losses in 13q14.** Submitted.

IV. Rodríguez AE, Quwaider D, Benito R, Misiewicz-Krzeminska I, Hernández-Sánchez M, Rojas-Ricardo E, Sarasquete ME, Gutiérrez NC, Hernández-Rivas JM. **A common polymorphism in 3´UTR deregulates HSP90B1 expression in chronic lymphocytic leukemia.** Manuscript in preparation.

Introduction

Genetic profiling

Since 1956, when Tjio and Levan made the landmark observation that human cells contain 46 chromosomes, the practice of cytogenetics has witnessed many innovations, including chromosome banding techniques, fluorescence in situ hybridization (FISH), spectral karyotyping, and comparative genomic hybridization (CGH).[1] Of these advances, G-banded karyotyping has been the most useful method in routine studies and has hence been the gold standard for analyzing chromosomes for the past several decades. In CLL, the use of Fluorescence "in situ" hybridization also provided new genetic markers and became an available tool in clinical setting. These advances helped to elucidate the causes of a range of genetic disorders and cancers, and the experts in the field formulated criteria for their use in clinical diagnostics. The most recent microarray-based CGH (array CGH), offered a method for whole-genome analysis at resolutions much greater than that possible with conventional karyotyping.[2] The first genomic microarray technique for detecting DNA alterations, array comparative genomic hybridization (aCGH), was developed from conventional CGH in the late 1990´s.[3,4] This methodology is based on the detection of CNAs due to a higher or lower ratio of inbound differently labeled sample DNA compared to a reference DNA (Figure 1). Instead of metaphase chromosomes acting as targets for sample DNA binding, aCGH uses thousands of short DNA probes, each representing a unique part of the genome, arrayed onto for example a glass slide. The fluorescence for each probe is recorded separately for the sample and reference channel using a laser scanner, and the formed ratio (sample/reference) can be plotted against the known genomic positions of each probe generating a high-resolution map of DNA copy-number alterations. With the exception of regions on centromeres and telomeres, characterized by the presence of highly repetitive sequences, aCGH is available for the whole genome screening by using bacterial artificial chromosome (BAC) arrays or oligonucleotide arrays.

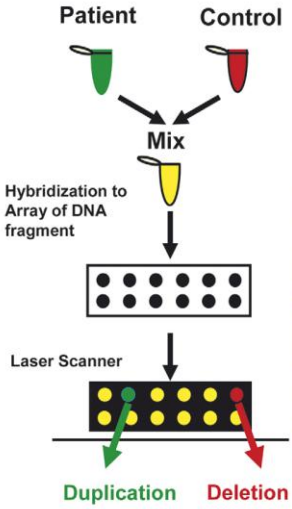

Figure 1. Principles of the aCGH technology. Whole genomic DNA from a control or reference (right) and genomic DNA from a test or patient (left) are differentially labeled with two different fluorescent dyes. The two genomic DNA samples are competitively cohybridized with large-insert clone DNA targets that have been robotically printed onto the microarray (middle). After hybridization of the DNA samples, the array is scanned to assess the relative fluorescence levels of each DNA for each target on the array. The ratio of fluorescent intensity between control and test DNA is used to identify copy number losses and gains on the chromosomes.[5]

The first genomic arrays were based on bacterial artificial chromosome clones mapped during the initial phases of the Human Genome Project. These arrays essentially represented multiple fluorescence in situ hybridization assays performed simultaneously. The first arrays, consisting of hundreds of bacterial artificial chromosome clones, mostly limited to genomic regions of known medical significance, were "targeted" formats, containing collections of clones from specific genomic regions, designed to detect genomic imbalances associated with known genetic disorders.[6,7] BAC probes are produced by insertions of DNA sequences to BAC (or PAC) clones in order to amplify the DNA sequences.[8-10] The DNA is extracted, fragmented and spotted to glass slides at specific positions, where each DNA probe will represent 100-200 kbp of the genome. Our group worked with a BAC array that contained 3528 bacterial artificial chromosomes (BACs) spanning the genome. The particular BAC and PAC set was distributed to academic institutions by the Welcome Trust Sanger Institute (Cambridge, United Kingdom) and contains targets spaced at ~1 Mb density over a full genome, a set of subtelomeric sequences for each chromosome arm, and a few hundred probes selected for their involvement in oncogenesis. The clone content is

available online in the ‘‘Cytoview’’ windows of the Sanger Institute mapping database site (available at: http://www.ensembl.org/ accessed August 2008). According to this database, clones were ordered along the chromosomes. Whole-genome arrays contained bacterial artificial chromosome clones from across the entire genome, which contained clones selected at regular intervals across the entire genome

More recently alternative designs based on oligonucleotide probes have been developed. These high-density whole-genome arrays have a resolution between 3 and 35 kb. The design of an oligonucleotide aCGH is somewhat different from BAC arrays. Variations include manufacturing method, sample throughput capacity and probe length (25–80 bases). All available oligonucleotide platforms are whole genome arrays, but they can be customized either to function as targeted arrays (similar to available targeted BAC arrays) or to substantially increase resolution in a specific genomic region of interest. This versatility will eventually prove useful as aCGH technology becomes widespread.

Next-Generation sequencing

Published by Sanger and colleagues in 1977, the dideoxynucleotide method for DNA sequencing remained the standard for the next 30 years. The advent of DNA sequencing has significantly accelerated biological research and discovery and the high demand for low-cost sequencing has driven the development of new sequencing technologies. Next-generation sequencing (NGS, also known as massively parallel sequencing) technologies have a higher throughput than traditional sequencing methods. It allows millions of sequencing reactions to happen in parallel, using different approaches, either by creating micro-reactors and/or attaching DNA molecules to solid surfaces or beads. Unlike previous methods, NGS generates millions of short reads (21-400 base pairs) and does not require amplification as sequencing could be performed from a single DNA molecule. The short reads can be quantified, allowing accurate copy number assessment. Moreover, with approaches that sequence both ends of a DNA molecule (paired end massively parallel sequencing), it has become possible to detect balanced and unbalanced somatic rearrangements (i.e. fusion genes) in a genome-wide fashion.

Recently developed targeted sequence enrichment, coupled with NGS, represents a beneficial strategy for enhancing data generation to answer questions in Medicine. This marriage of technologies offers to the researchers a simple method to isolate and analyze a few to hundreds, or even thousands, of genes or genomic regions from few to many samples in a relatively efficient and effective manner. Sequence capture is fully scalable, making it useful for experiments requiring few genes or genomic regions, as well as those involving entire exomes.

Targeted sequence enrichment refers to the suite of technologies designed to isolate a specific genomic fraction (e.g., genes, molecular markers, larger genomic regions) for subsequent NGS, ultimately resulting in an enriched pool of target sequences such that there is overall reduction in the genomic sequencing space, and hence greater sequence coverage for each targeted region. Several technologies exist for target enrichment, which can be classified by the mode of enrichment: (1) hybridization based sequence capture, (2) PCR-based amplification, and (3) molecular inversion probe-based amplification. The advantages of each of these have been discussed in depth.[11] Many protocols and technological alternatives are available for hybridization-based sequence capture. The two primary approaches involve hybridization of samples either to microarrays or to solution-based, pooled RNA-baits, both of which are complementary to the targeted genes. In both approaches, all targeted sequences are captured in one hybridization reaction.Since samples can be multiplexed using appropriate barcoding of NGS adapters, the potential to capture thousands of sequences simultaneously is proven. As with any technique, key elements must be taken into consideration when designing sequence capture experiments, including (1) the number of targets and the mean depth of coverage needed for each target, (2) the probe specificity, (3) the biological system studied, (4) the expected enrichment efficiency, and (5) the NGS technology used.

A final consideration in preparing for sequence capture is the type of NGS technology applied to the experiment. Four platforms for massively parallel DNA sequencing read production are used: the Roche/454 FLX[12] (http://www.454.com/enablingtechnology/the-system.asp), the Illumina/ Solexa Genome Analyzer (http://www.illumina.com/pages.ilmn?ID=203), the Applied Biosystems SOLiDTM System (http:// marketing.appliedbiosystems.com/images/

Product / Solid Knowledge / flash/ 102207/ solid.html) and the Ion Torrent system (http://www.iontorrent.com/ion-personal-genome-machine-sequencer/).

The Roche/454 FLX Pyrosequencer was the first to achieve commercial introduction (2004) and uses an alternative sequencing technology known as pyrosequencing. In pyrosequencing, each incorporation of a nucleotide by DNA polymerase results in the release of pyrophosphate, which initiates a series of downstream reactions that ultimately produce light by the firefly enzyme luciferase. The amount of light produced is proportional to the number of nucleotides incorporated (up to the point of detector saturation). The steps followed in this technology are summarized in Figure 2.

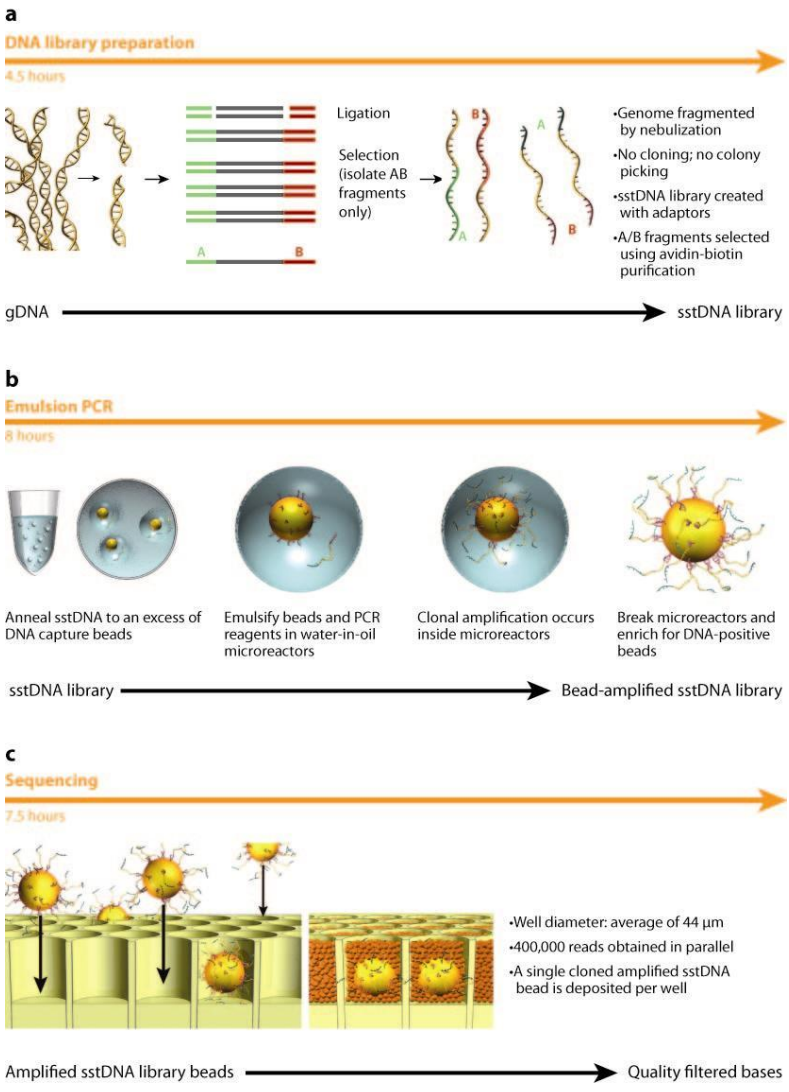

Figure 2. Work-flow of a NGS experiment with the 454-FLX technology.[13]

Chronic lymphocytic leukemia

Chronic lymphocytic leukemia (CLL) is the most common leukemia in Western countries mainly affecting individuals older than 50 years. Clinical diagnosis of chronic lymphocytic leukemia is defined by absolute lymphocytosis of at least 5×10^9/L, mature-appearing lymphocytes and an appropriate immunophenotype.[14] These characteristics distinguish CLL from mantle-cell lymphoma, splenic marginal-zone lymphoma and other chronic lymphoproliferative disorders, the diseases that could mimic CLL.[15]

CLL follows an extremely variable course, with survival ranging from months to decades. Recently, progress has been made in the identification of biological markers that could predict disease progression. Particularly, the expression of unmutated Ig genes, some cytogenetic abnormalities like 17p and 11q deletions and the expression of the zeta-associated protein 70 (ZAP-70) are associated to a poor prognosis. A major scientific goal is to find a biomolecular explanation for the CLL prognosis heterogeneity that can provide clues in the understanding of disease etiology and pathogenic mechanisms which promote the onset of the disease, as well as the progression and the evolution into more aggressive variants (Richter's lymphoma or prolymphocytic leukemia).[16]

Biology and pathophysiology

CLL is characterized as an abnormal proliferation of malignant B lymphocytes.[17] CLL was once thought to be a homogeneous disease, in which mature B cells accumulated largely due to a lack of normal cell death. Recently, however, CLL has been established as a heterogeneous disease of remarkable diversity. Differences in cell morphology, immunophenotype, cytogenetics, and molecular characteristics have been identified. This heterogeneity translates into varying clinical courses and responses to treatment.[18] Table 1 provides a historical overview of some of the changes in how CLL biology is viewed.

Table 1. Comparison of historical and current views of CLL biology[19]

Historical view		Current view
CLL is a clinically heterogeneous disease with a homogeneous celllar origin	→	CLL is a clinically heterogeneous disease originating from B lymphocytes that may differ in activation, maturity or cellular subgroup
CLL is a disease derived from naive B lymphocytes	→	CLL is a disease derived from antigen-experienced B lymphocytes that differ in the level of immunoglobulin variable gene mutations
Leukemic-cell accumulation occurs because of an inherent apoptotic defect involving the entire mass of leukemic	→	Leukemic cell accumulation occurs because of survival signals from the external enviroment
CLL is a disease of lymphocyte accumulation	→	CLL is a disease of lymphocyte accumulation with a higher associated level of proliferation than was previously recognized
Prognostic markers identify patients at various risk levels (low, intermediate or high in the Rai staging categories, and A, B or C in the Binet categories) with an acknowledged heterogeneity in clinical outcomes among patients in the low- and intermediate-risk categories	→	New molecular biomarkers are used in both diagnosis and prognosis to better assess patients

Malignant lymphocytes characteristically appear small and mature, although large atypical cells, cleaved cells, or prolymphocytic cells could also be observed.[20] In CLL, malignant, long-lived lymphocytes accumulate in the blood, bone marrow, spleen, and lymph nodes. The phenotype of CLL distinguishes it from other B-cell malignancies by the presence of B-cell markers (CD19, CD20, CD23, and CD43) along with CD5, an antigen normally found on T cells. Typically, CLL cells also express surface immunoglobulin (slg), CD79b, CD20, and CD22 at low density. The phenotypic features of CLL are not only used for initial diagnosis, but they also play a role in the assessment of minimal residual disease (MRD), an important prognostic parameter defined as <1 clonal B-cell in 10,000 leukocytes.[20]

Prognostic markers

The Rai and Binet staging systems are simple and reliable prognostic tools that are in the basis of the decision to treat most patients.[21,22] However, there is considerable variation in outcome with each stage. Therefore, additional factors are more commonly being considered in predicting individual patient prognosis and stratifying patient risks.

A quick review of the literature identifies over 35 different prognostic markers. Some of these are classified as “classical”, whereas others are classified as “biological” (Table 1). The classical prognostic markers tend to be those obtainable from routine history,

physical examination, and lab work. The biological prognostic factors tend to assess molecular aspects of the CLL cell themselves.

Table 2. Chronic lymphocytic leukemia: classical and biological prognostic markers.

Classical prognostic markers
Clinical stage
Blood lymphocyte count
Lymphocytes morphology in peripheral blood
Blood lymphocyte doubling time
Bone marrow infiltration degree (aspirate/biopsy)
Biological prognostic markers
Extensively studied
Serum markers
IGVH mutational status
V3-21 gene usage
Cytogenetics
CD38 expression
ZAP-70 expression
Requiring further studies
Chromosomal translocations
CLLU1 expression
microRNA signature
TCL-1 gene
Bcl2/Bax ratio
MDR1/MDR3 genes
Activation-induce cytidine deaminase (AID) mRNA
Lipoprotein lipase A expression
ADAM29 expression
VEGF
Thrombopoietin
Telomere lenght and telomerase activity
CD49d
CD69
FCRL
Treatment related
Response to therapy (minimal residual disease status after therapy)

Genomic abnormalities in CLL

Our knowledge of genomic aberrations in CLL has initially been based on chromosome banding studies using TPA as a mitogen.[23] These studies have been held back by the low mitotic activity of CLL cells. With this method, clonal chromosomal aberrations

were detected in only 40 to 50% of cases, the most common being trisomy 12 and abnormalities of chromosome band 13q14.[23] More recently, metaphase analysis has had a "revival" because the metaphase yield has been improved by stimulation of CLL cells with alternative methods like CD40 ligand expressing cells and IL-4 or the addition to the cultures of CpG-oligodeoxynucleotides and IL-2.[24] This approach revealed that translocations are present in about one third of CLL patients. After stimulation with CpG-oligodeoxynucleotides and IL-2, the observed aberration rate was comparable to the rate detected by parallel interphase FISH, and the method frequently detected balanced and unbalanced translocations.[25] These studies also demonstrated that additional cases of complex aberrations (more than 3 aberrations) can be detected. In a recent update on the results of metaphase cytogenetics after stimulation with CpG-oligonucleotide DSP30 and IL-2, a total of 500/506 (98.8%) cases were successfully analyzed.[26] Aberrations were detected in 415 of 500 (83.0%) cases by banding and only in 392 of 500 (78.4%) cases by FISH due to limitations of the probe set. Future studies will show if the metaphase cytogenetics after stimulation will result in a deeper understanding of the disease and prognostic subgroups.[27]

Fluorescence in-situ hybridization (FISH) allows the detection of chromosomal aberrations irrespective of the cell's ability to divide. [28] Genomic aberrations can be identified in about 80% of CLL cases by FISH with a disease specific probe set. The most common recurrent chromosomal abnormalities include trisomy 12 as well as deletions of 13q 11q, 17p and 6q.[29] A subdivision based on these aberrations is important, as it is associated with the rate of disease progression and the overall survival time of CLL. Five prognostic categories have been defined in a hierarchical model showing a shorter survival in patients with 17p deletion and 11q deletion (median survival of 32 and 79 months, respectively), while the survival is longer for patients with trisomy 12q normal karyotype, and deletion 13q as the sole abnormality (114, 111, and 133 months, respectively).[29] FISH is the current method of choice in clinical setting. However, FISH only "cover" a limited set of aberrations and the recent developments of comparative genome hybridization (CGH) and in particular high resolution SNP arrays have lead to a better understanding of the genetic profile of CLL.

IGVH mutational status

More than a decade ago the simultaneous publication of two papers seemed to explain the heterogeneity that clinicians had observed in the natural history of chronic lymphocytic leukemia (CLL). [30,31] Both studies mutually corroborated the finding that *IGVH* mutational status identifies two subsets of CLL, one with a median survival of 8 years and the other with a median survival of 25 years. Somatic mutations of the VH gene region of the heavy chain of immunoglobulins are present in about half of all CLL cases. Unmutated cases originate from cells in the pregerminal center and clearly are associated with a worse prognosis than mutated CLL cells arising from the postgerminal center. The definition of non-mutated *vs* mutated pattern resides in a cutoff point, defined arbitrarily as a homology greater than 98% (non-mutated) gene most similar to the germline.[32] Patients with CLL and an unmutated status have an unfavorable course and progress more rapidly, as opposed to patients exhibiting a mutated state, whose survival is much longer. Unmutated CLL patients have a greater tendency to acquire poor prognostic cytogenetic abnormalities. *IGVH* mutational status and cytogenetic abnormalities identified by FISH have a major impact on the survival of patients with CLL, but while cytogenetic changes during the course of the disease are relatively common, *IGVH* mutational status remains constant over time.

Gene expression profiling

The value of gene expression profiling (GEP) in the study of CLL is widely accepted. Such studies have identified new prognosis markers such as ZAP-70, LPL, PEG10 and CLLU1. Some of these are already well-established factors used in clinical practice, while the application of others is under study.

Despite clinical and molecular differences, global gene expression profile analysis demonstrated that all CLL show a homogeneous gene expression profile irrespective of their IgV mutational status and differing from other lymphoid cancers, which suggests a common cellular precursor.[33,34] These analyses revealed that the gene expression profile of all CLL is related to that of antigen-experienced B cells, which in the human are defined by expression of the CD27 cell surface antigen, and that include classical memory B cells and marginal zone B cells which can be somatically mutated or

unmutated.[35,36] However, despite sharing a common signature, CLLs with mutated *IGVH* displayed a different expression of more than 100 genes. Among these, over-expression of genes encoding zeta-chain-associated protein 70 (ZAP-70), lipoprotein lipase (LPL), BCL-7A, dystrophin and gravin were observed in the unmutated cases, while the mutated cases over-express *Wnt3, CTLA-4, NRIP1* nuclear receptor gene, *ADAM29* and the transcription factor *TCF7.*[37] These results suggested that indolent mutated and aggressive unmutated CLLs constitute two variants of the same disease. The reasons accounting for these striking differences in clinical outcomes of these two variants remain unsolved. ZAP-70 appeared to be one of the most significant.[34] Subsequently, the correlation of ZAP-70 expression with the mutational status of the *IGVH* genes was assessed in larger series of CLL patients, where ZAP-70 was mostly found expressed in unmutated CLL.[38] Further clinical studies revealed that ZAP-70 was also an independent prognostic marker.[39] In addition, genes that are significantly more highly expressed are located in the corresponding aberrant chromosomal regions, indicating the existence of a genetic dosage effect, which may have a pathogenic role in CLL. Surprisingly, significant differences in gene expression according to sex were also found, which suggests that differences in molecular signatures relating to *IGVH* mutational status may be related to the sex of the patient.[40]

Kinetics of CLL

CLL had been considered as a disease of failed apoptosis and passive accumulation. This view is supported by the observation that the great majority (~98%) of peripheral circulating CLL cells are arrested in G0 or the early G1 phase and have overexpression of antiapoptotic proteins, such as BCL2. CLL B-cells are characterized by high expression of BCL-2 protein in the absence of specific translocations and by high expression of the p27kip protein, which blocks the progression into cell cycle. Given the key role of this protein in the cell cycle progression, the over-expression in CLL cells could account for the accumulation of B cells in early phases of the cell cycle. In addition, other members of the BCL-2 family such as anti-apoptotic proteins BCL-XL, BAG-1 and MCL-1 are overexpressed, while pro-apoptotic proteins like BAX and BCL-XS are under-expressed.[41,42]

The birth-rate of the leukemic cell clone was considered to be low and without influence on the disease. However, this view of CLL has been recently revised. Thus *in vivo* studies showed that there is a small but variable fraction of the monoclonal CLL cells that actually proliferate, and a fraction of cells that undergo apoptosis.[43] The same study showed that patients with accumulation of B cells showed a higher birth rate and a lower death rate of cells. In comparison, patients with stable B cell numbers displayed equilibrium between proliferation and apoptosis.

As opposed to *in vivo* results, apoptosis occurs after *in vitro* culture, which suggests a role of the microenvironment in CLL cell survival.[44] Therefore the host microenvironment of B cells determines the balance between resistance to apoptosis and active proliferation by providing different sets of signals and conditions. CLL B-cells communicate with cells in the microenvironment, such as T cells and stromal cells, and the interaction with these cells provide essential signals for survival and expansion of the CLL cells promoted by molecules such as IL-4 and CD40 ligation, whose effect also has been documented *in vitro*.[45,46]

MicroRNAs

MicroRNAs (miRNAs or miRs) are noncoding RNAs that have the capacity for simultaneous regulation of tens to hundreds of genes through direct targeting of untranslated regions (UTR)[47-49] . MiRNAs represent the best characterized member of a family of noncoding RNAs, many of which are currently being studied. MiRNA targeting results in either degradation of mRNA or inhibition of translation. In silico models suggest that miRNAs may be responsible for the regulation of more than 30% of the genome.[47] MiRNAs are important for biological processes, including proliferation, apoptosis, development, and cellular differentiation[50] and have the capacity to function as either oncogenes or tumor suppressors.[47] Deregulation in expression of miRNAs in some malignancies compared to normal tissue as well as their frequent location in fragile chromosomal regions supports their relevance to the development of malignancies.[51]

The first description of a miRNA associated to cancer was reported in CLL.[52] Calin et al showed that miR-15a/16-1 were located in 13q14.3 and were either deleted or downregulated in 68% of patients with CLL. Further studies found that these miRNAs

have as target the mRNA encoding the antiapoptotic protein B-cell CLL/lymphoma 2 (BCL2), the upregulation of which is critical for CLL cell survival.[53] Importantly, the direct interaction of *miR-15*/*miR-16* with BCL2 transcripts delayed protein translation, induced apoptosis, and reinforced the role of miRNAs as part of a new class of tumor suppressor genes.

Given the apparent link between cytogenetic abnormalities in CLL and miRNA deregulation, miRNAs have become attractive candidates as both biomarkers and therapeutic targets for CLL. The interaction between miRNAs, target genes, and pathways in CLL is clearly complex, as are the links between genotype and phenotype (Figure 3).

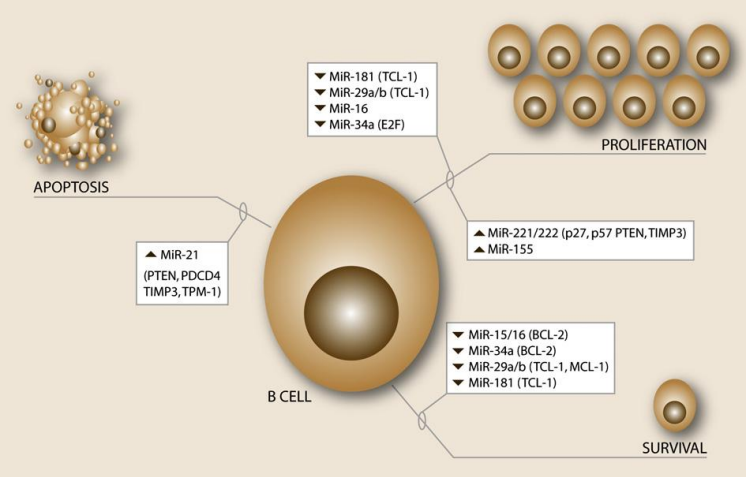

Figure 3. miRNAs implicated in apoptotic, survival, and proliferative pathways in CLL. Arrowheads represent expression patterns in CLL. Genes in parentheses represent validated targets for select miRNAs.[54]

Recently a miRNA analysis in a cohort of 61 treatment-naive CLL patients has been performed. The cohort represented several distinct karyotypes and the authors identified miRNAs that distinguished all karyotype groups, showing that miR-181a expression correlated with a shortened time to treatment as well as distinguished aggressive from indolent disease in 17p deletion cases[55]. Calin et al conducted miRNA expression profiling on a well-annotated series of 94 CLL patients and identified a panel of 13 genes that correlated with ZAP-70 expression and *IGVH* mutational status.[56] In addition, a subgroup of nine miRNAs (181b, 155, 146, 24-2, 23b, 23a, 222, 221, and 29c) differentiated patients with a short interval to therapy from patients

with a longer interval to therapy. The miR-34 family has been also implicated in CLL. Thus miR-34a expression partially correlated with p53 status and patients with p53 mutations or 17p- in general had lower miR-34a expression. However, the link between p53 mutational/deletional status and miR-34a has proven to be much more complex and further studies are required.

The *TCL1* gene, located at 14q32.1, is overexpressed in 25-35% of CLL cases.[57] *TCL1* overexpression is increased in CLL patients with unmutated *IGVH* and elevated ZAP-70, suggesting its role in aggressive forms of CLL. *TCL1* transgenic animals represent the best current animal model for the development of CLL.[58] Through the use of the *TCL1A* murine model, investigators have identified other miRNAs relevant to the pathogenesis of CLL such as miR-29 and miR-181.[59]

The rapid advancement of miRNA research in just the past few years suggests that the roles of many other miRNAs in CLL have yet to be discovered. Thus miRNAs remain one of the most exciting new avenues for CLL research.

Heat shock proteins in CLL

Over the last decade, research has implicated a group of molecular chaperones termed Heat Shock Proteins (HSPs) as major contributors to cancer progression and the development of chemo-resistance. In tumor cells, including CLL cells, HSPs are upregulated and may contribute to prolonged tumor cell survival via several mechanisms that remain to be fully revealed.[60-62]

Hsp90 is a molecular chaperone that catalyses the conformational maturation of a large number of signaling proteins in cancer that are collectively described as 'clients'. In advanced tumors, Hsp90 exists in an activated form complexed with other molecular chaperones, whereas in normal tissue Hsp90 is present in a latent, uncomplexed state.[63] Research into the involvement of *HSP90* in cancer progression has identified a wide variety of client proteins that may contribute to the progression of both solid tumors and hematological malignancies.

It has been demonstrated that ZAP-70 in CLL cells co-immunoprecipitates with HSP90, while ZAP-70 from normal T-cells does not. In addition, treatment of ZAP-70+ CLL cells with HSP90 inhibitors resulted in degradation of ZAP-70, while treatment of T-cells from CLL and control patients with HSP90 inhibitors did not affect the expression of

ZAP-70. Manipulation of ZAP-70- CLL cells to express ZAP-70 was shown to activate HSP90 and induce sensitivity to 17-AAG, an HSP90 inhibitor.[64]

HSP90 chaperones, a broad array of protein kinases involved in signal transduction pathways, including phosphorylated Akt, Lyn, B-Raf and IκK.[61,62,65] The interaction of HSP90 with these phosphorylated kinases prevents dephosphorylation of the kinase and its subsequent inactivation, therefore maintaining its activity. The over-activity of Akt,[66] Lyn[67] and implied elevation of IκK[68] in CLL cells has been previously observed and thought to contribute to cell survival and activation of downstream kinases. IkK over-activity has implications for a large number of NF-κB target genes including Bcl-2, X-IAP, c-FLIP and Mcl-1.[68] In fact, a number of these proteins have been implicated in the progression of CLL and have been found to correlate with poor prognosis which may suggest a clear link to over-activity of NF-κB[69,70]. Interestingly, Activation-Induced Cytidine Deaminase (AID), which has been shown to have prognostic significance in CLL,[71,72] is an HSP90 client and chemical inhibition of HSP90 activity results in destabilisation and proteosomal degradation of AID and reduced antibody diversification.[73] Furthermore, HSP90 inhibitor-induced cell death appears to be p53 independent suggesting that it may also be useful in the treatment of patients with a mutated p53 status.

It would appear that HSP90 inhibitor treatment of CLL cells targets a wide variety of proteins, many of which have been implicated in the progression of the disease. Thus, targeting HSPs is an attractive strategy in CLL.

Next generation sequencing studies in CLL

Whole genome sequencing of cases with CLL led to the discovery of several genes, previously unsuspected to be involved in this disease. Application of next generation sequencing technology has identified previously unknown recurrent mutations in *NOTCH1, XPO1, MYD88,* and *KLHL6* in 12.2%, 2.9%, 2.4%, and 1.8% of CLL cases, respectively.[74] *NOTCH1* and *XPO1* mutations are associated with *IGVH* unmutated CLL while *MYD88* and *KHL6* are associated with *IGVH* mutated CLL. Both *NOTCH1* and *MYD88* mutations appear to be activating mutations. *NOTCH1* mutated cases have been shown to overexpress *NOTCH1* pathway genes and are associated with unfavorable prognosis. Thus, in addition to a potential therapeutic target, this

mutation appears to confer prognostic information. Further studies on the coding genome of CLL have also revealed a recurrent mutation in a component of the spliceosome, *SF3B1*.[75,76] These somatic mutations were missense mutations that are clustered in hotspot codons 662, 666 and 700 and predicted poor prognosis. The mutations were found in 9.8% of CLL at diagnosis[75] and in 17% of fludarabine refractory cases.[76] This gene has also been shown to be mutated in myelodysplastic syndromes (refractory anemia with ring sideroblasts).[77] Therefore NGS appears to be a highly effective technique in identifying new genetic lesions and future studies are promising to contribute to an improved understanding of disease onset and evolution.

Objectives

The general aim of this thesis was to gain insight into the pathophysiological mechanisms of CLL in the different cytogenetic subgroups by applying a combined analysis of the genome and transcriptome of B clonal lymphocytes.

More specifically the aims were as follows:

- To screen and to identify new recurrent aberrations in a cohort of newly diagnosed CLL patients by genomic microarrays (aCGH).
- To analyze, by applying two microarray platforms, the gene expression profile of the cytogenetic CLL subgroups: normal FISH, 13q-, +12, 11q- and 17q-, as well as the new recurrent abnormalities identified by aCGH studies.
- To characterize the CLL patients with 13q deletion by gene expression profile and microRNA studies.
- To identify new variations in genes involved in CLL pathogenesis by applying next-generation sequencing techniques.
- To define the functional implication of the genes with altered expression and/or mutations in the pathogenesis of CLL.

Results

Paper I – Array-based genomic profiling in CLL

The advancement of genomics technologies has offered a range of experimental approaches to study the molecular variation underlying the biological and clinical heterogeneity of CLL. Microarray-based methodology provides genomic information of the whole genome in a single experiment. Thus both CGHarrays and expression arrays give a full scope of the cancer genome. Several CLL studies have applied microarrays for the detection of copy number abnormalities (CNAs), and have served/supplied/added/produced additional information on genome-wide alterations in this disease.[78-84] Furthermore expression microarrays have contributed to the molecular characterization of CLL.[33,34,40,85,86]

Our aim was to investigate genomic aberrations in CLL by a global analysis of the genome and the transcriptome and to study the copy number variations in CLL cytogenetic subgroups. The outline for the study is shown in Figure 4.

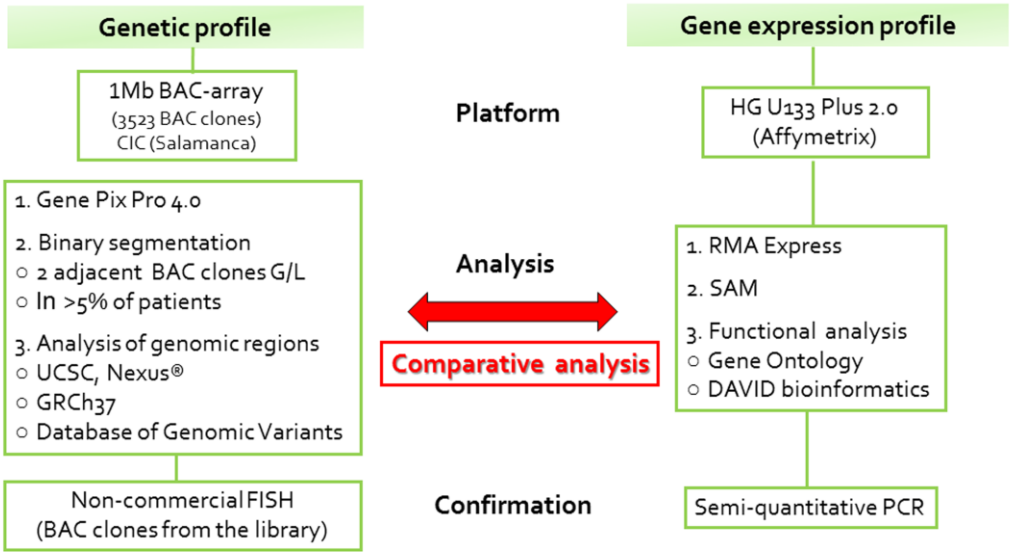

Figure 4. Outline for the genomic characterization of CLL

We have applied high density aCGH screening in 67 CLL samples. Our customized 1Mb-BAC array contained 3,523 clones, kindly provided by the Sanger Institute, Cambridge (UK), spanning the genome. The data analysis was carried out with GenePixPro 4.0. To

study the gains and losses a binary segmentation method was applied. In a selected group of patients, the results were confirmed by oligonucleotide arrays (NimbleGen) and FISH. The recurrent alterations were confirmed by FISH. Gene expression profile was analyzed by HU-133 Plus 2.0 array (Affymetrix). In all cases clinical data, FISH, IGH somatic mutations and immunophenotype were also available.

A total of 75% of patients displayed cytogenetic aberrations: 443 altered chromosomal regions were found, of which 237 (53%) were deletions. The median number of changes per patient was five (range 0–14). Overall array CGH results correlated with FISH analysis (+12, 11q- and 17p- cases), although the correlation was only 48% in 13q- subgroup. Besides the alterations reported by FISH (11q-, +12, 11q-, and 17p-), genomic arrays detected other recurrent changes: losses in 11q13.3 (16%), in 5q13.3–q14.1, 5q31.1 and 7q22 (6% each), as well as gains in 1q21.3–q22 (22%), 11q13.3 (21%), 16q23.2–q24.2 (21%), 6p21.31–p21.1 (19%) and 10q22.3 (7.5%). These aberrations can also be observed in patients with normal FISH. Thus our data confirms the heterogeneity of CLL.

We also defined a new recurrent gain in 20q in 19% (13/67) of the CLL patients. Genomic analysis identified the minimum region of gain on 20q13.12 of ~2.31 MB involving three clones at linear positions (42 188 467–44 495 323). To better characterize the 20q13.12 gain, FISH analysis was also carried out in a validation series of 58 patients and 17% patients showed this alteration in ≥4.5% cells (ranging from 4,5% to 12%). This gain was associated (P=0.002) with a high frequency of chromosomal changes as revealed by aCGH and overexpression of the genes located in this region.

Oligonucleotide expression microarrays correlated with the regions of loss or gain of genomic material, suggesting that the changes in gene expression are related to alterations in copy number, which confirms the gene dosage effect previously described in CLL.

Papers II and III – Characterization of CLL patients with a high number of 13q- cells.

Patients with loss on 13q show the better outcome in CLL. However, in some series with a long-term follow-up, patients with CLL and a normal karyotype showed a longer survival from 12 years on, as compared to patients with 13q-,[30] suggesting that 13q- subgroup could be heterogeneous. In addition, several studies have demonstrated that the percentage of cells displaying a particular cytogenetic abnormality (e.g. loss of *P53*)[87] or antigenic markers (e.g. CD38 or ZAP-70)[88] can be related to prognosis. For these reasons, an analysis of patients diagnosed with CLL and deletion in 13q14, as the sole cytogenetic abnormality, was carried out. The clinical features, including outcome, and the biological features of the patients displaying different degrees of infiltration by 13q- cells, were also assessed.

The study population performed in Paper II comprised 350 non-selected patients, from nine Spanish institutions, diagnosed as CLL. Clinical data were recorded by reviewing the clinical histories of patients included in the study. In most cases (283 patients; 81%) the FISH study was performed at the time of diagnosis. Progression was defined according to previously reported criteria: the presence of disease-related symptoms, massive or progressive organomegaly, bone marrow failure or recurrent infections[89]. At the time of diagnosis 109 patients showed 13q deletion as the sole FISH abnormality and the study was focused on this group of patients. A significantly longer survival was observed in the cohort of patients with losses in 13q in less than 80% of cells. Thus, in the subgroup of patients with 80% or more of cells with loss of 13q the overall survival was 56 months (95% CI: 39-73 months), while in the group of patients in whom less than 80% of cells showed losses in 13q, the overall survival had not been reached (95% CI: 163-254 months) (P<0.0001). A significantly shorter time to first therapy (TFT) was observed in the cohort of patients with 80% or more of cells showing losses in 13q (median of 38 months; 95% CI: 21-55 months) as compared to those cases with less than 80% of 13q- (median of 87 months; 95% CI: 21-153 months) (P=0.05). Thus, the percentage of cells displaying 13q deletion influences the outcome of CLL patients.

These results have been corroborated by two independents studies.[90,91] However, to the best of our knowledge, the molecular characteristics of 13q- CLL patients have not been so far analyzed in detail in order to better understand why these patients could

have a different outcome. Thus, as a next step toward the elucidation of biological differences within 13q- subgroup, a gene expression profile studies using two microarray platforms from *Affymetrix*: HG U133 Plus 2.0 and Human Exon 1.0 ST was carried out.

For the gene expression profile analysis two groups of patients with 13q- were compared: those in whom 80% or more of cells showed 13q- (13q-H CLL) and those in whom less than 80% of cells showed 13q losses (13q-L CLL). The gene function analysis performed in Paper II revealed that most of these genes were involved in apoptosis, cellular growth and proliferation, mitochondrial, endoplasmic reticulum (ER) and calcium mediated activity. In Paper III, the gene expression analysis was performed using the Affymetrix Human Exon arrays 1.0, which offer a more fine-grained view of gene expression than the former generation of chips. Thus, the data obtained provided great insights into the biological mechanisms underlying the clinical differences observed in this CLL subgroup. Overall, the correlation between both platforms was excellent. In Paper III we also performed a validation analysis on CD19+ immunomagnetically purified cells from an independent cohort of patients. The overview of the experimental strategy followed in Paper III is shown in Figure 5.

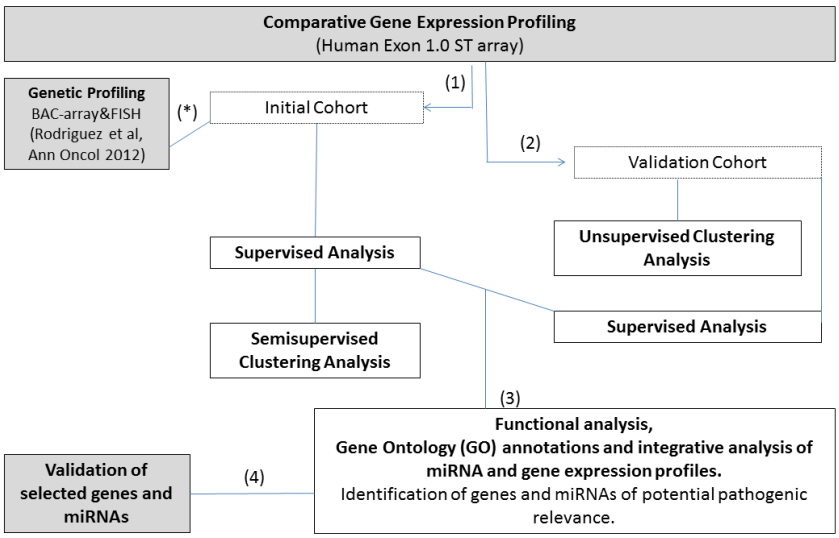

Figure 5. Overview of the experimental strategy followed in Paper III.

The transcriptome of 97 CLL patients and 5 healthy controls was analyzed. We split the samples in two cohorts: (1) Initial Cohort (n=65): PBMC cells, isolated from peripheral blood using Ficoll gradient and (2) Validation Cohort: CD19+ cells from 32 CLL patients and 5 healthy controls, immunomagnetically purified. The Initial Cohort was a subset of a group of CLL patients previously studied by our group by using genomic microarrays (Rodríguez et al, Ann Oncol 2012). Differentially expressed genes were identified using the SAM test in a supervised approach. To test whether the resulting lists of candidate genes were significantly associated with the CLL subgroups, a Global Test algorithm was used. In addition, in the Validation Cohort, we performed an unsupervised clustering analysis (Principal Component Analysis) to explore and represent the differences between the 13q- categories and controls. (3) The panels of differentially expressed genes were further functionally analyzed. An integrative analysis of the miRNA and gene expression profile was also performed (Supplementary Figure 2 in Paper III). (4) Gene expression changes for selected genes and miRNAs were tested using semi-quantitative SYBRgreen PCR and Taqman microRNA assays, respectively.

A total of 3,450 genes significantly distinguished 13q-H from 13q-L patients. These comprised 1,244 overexpressed genes and 2,206 underexpressed in the 13q-H group, defining the 13q-H signature. To assess the biological significance of the deregulated genes, a further analysis of the 3,450 deregulated genes characterizing the 13q-H CLL was carried out, revealing in this group of patients the involvement of several pathways, primarily related to cell proliferation, apoptosis and cell signaling. Moreover, 13q-H CLL patients were also characterized by a striking overrepresentation of deregulated miRNAs. A total of 15 miRNAs were deregulated in 13q-H relative to 13q-L patients. Hsa-miR-155 was the most highly upregulated miRNA (Rfold=3.70), while hsa-miR-223 was the most significantly downregulated (Rfold=0.10). The post-transcriptional regulatory network of miRNA and genes in CLL patients with more than 80% of 13q- cells was carried out by analyzing the miRNA-mRNA relationships and the pathway analysis demonstrated that B cell receptor signaling, PI3K signaling and NFkB signaling were among the most strongly affected pathways in 13q-H patients, highlighting the importance of miRNA regulation in CLL. We also analyzed the gene signature of CLL high risk cytogenetic subgroups in comparison with 13q- patients.

Surprisingly, our results suggest that some of the biological characteristics of 13q-H CLL patients were similar to those of high-risk cytogenetic subgroups, since they share the deregulation of several key signaling pathways. By contrast, 13q-L patients had similar gene expression to that of CLL with normal FISH.

Therefore, both clinical and biological studies demonstrate that CLL patients with 13q- are a heterogeneous group.

Paper IV – Mutational study of CLL by targeted resequencing.

The use of next generation sequencing (NGS) techniques has taken genomic research into a new era. The application of this technology has provided a better knowledge of the genetic complexity and heterogeneity of CLL. Technological innovation in DNA sequencing offers the promise of a more comprehensive, cost-effective and systematic ascertainment of genetic variation. A major bottleneck, however, is in isolating the target DNA to be sequenced. Based in literature and our previous results from GEP studies, we have applied targeted genome capture by using a NimbleGen array and a high-throughput FLX-454 sequencing to the genomic DNA of four CLL patients. By this targeted array-based sequence enrichment we have isolated the sequences of 93 selected genes (Supplementary Table S1) and two entire regions: 13q14.3 (50043128-50382849 bp) and 17p13.1 (7500000-7535000 bp).

The overview of the experimental design followed in this study is shown in Figure 6.

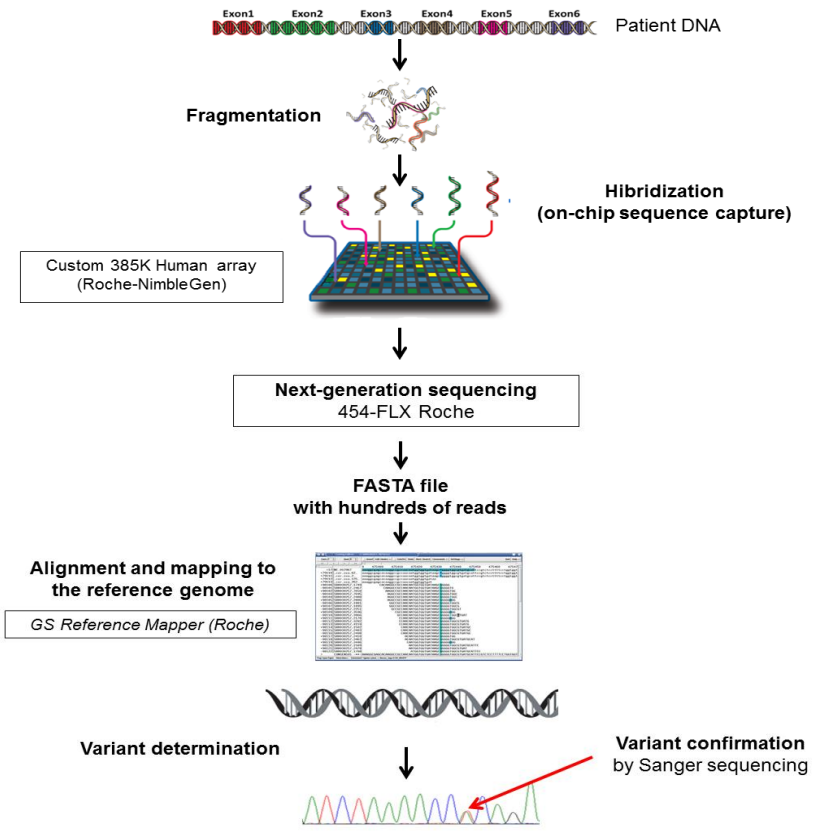

Figure 6. Targeted genome capture & next-generation sequencing

Briefly, genomic DNA was sonicated or nebulized to small fragment size (500bp). The fragment pool was denatured and hybridized to a NimbleGen Sequence Capture array designed to target the genomic regions of interest. Following hybridization, any unbound fraction was removed by washing and the enriched fragment pool was eluted and recovered from the array. The enriched fragment pool was amplified by ligation-mediated PCR. The end product was an enriched, amplified, target DNA pool ready for high-throughput sequencing. The software determines the sequence of more than 1,000,000 individual reads per 10-hour instrument run. The sequencing-data analysis consist of align the reads and map them to the reference genome. Basic raw data analysis was carried out using the GS Run Browser and GS Reference Mapper software version 2.0.01 (Roche Applied Science). Each sequence read was compared with the entire appropriate version of the human genome. Captured sequences mapped uniquely back to regions within the target regions were considered sequencing hits. We used a custom-made data analysis pipeline to annotate detected variants with various kinds of information, including known single-nucleotide polymorphisms (SNPs), amino acid consequences, genomic location and miRNA binding sites. To validate the variants identified Sanger sequencing was performed. A total of 4 patients with CLL and 4 patients with other hematological malignancies (used as controls) were analysed.

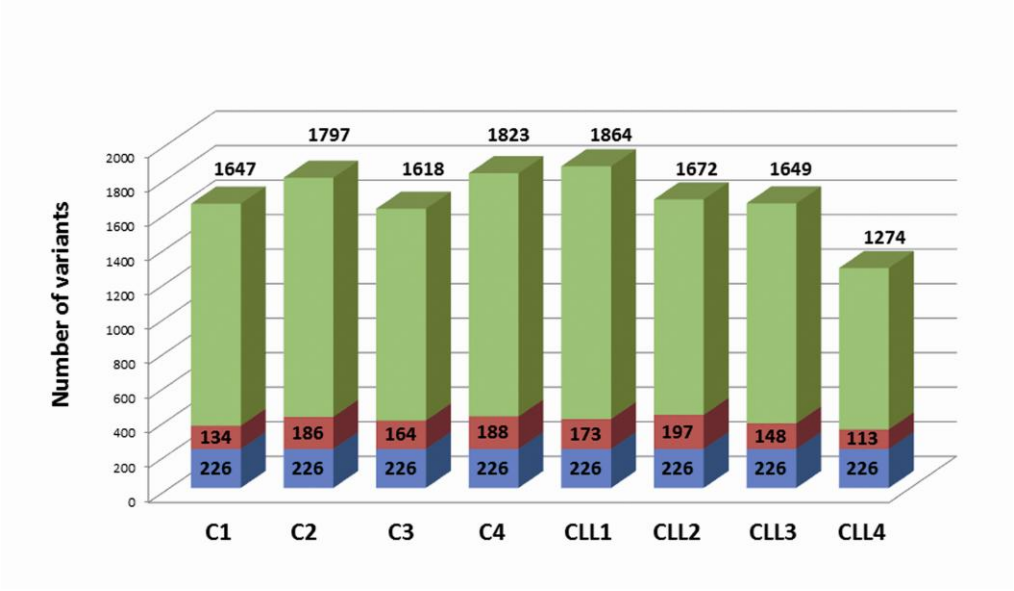

Figure 6. Number of variants identified in CLL and control samples.

The enrichment assay followed by NGS allowed the detection of over 1 600 variations/sample (median 1 721, range 1 618-1 823) (Figure 6). All putative variants were first compared with published single nucleotide polymorphism (SNP) data (dbSNP build 130; http://www.ncbi.nlm.nih.gov/projects/SNP). Most of the variants detected were identified as known SNPs and 226 variants were present in all the patients. Thus, altogether they were discarded. Overall, 10% of variants detected in each sample were variations not previously described. 73 were missense variations affecting 33 genes were detected. Most of the genes had one (70%) or two (12%) variations. Results are summarized in Table 3.

Number of variations/gen	Genes		
	n	%	Gene Symbol
1	23	70	ABCB6, ADAMTS1, ALAS2, BCL6, CAPRIN2, CDK6, CYP4F3, DICER1, DLEU7, FECH, GART, HSPD1, MLF1, MLL3, MRLPL27, MRLPL35, MRLPL37, MTFR1, PCCB, PIK3CA, RAPGEF2, SLC25A37, SOD1
2	4	12	ACAT1, CHFR, PHLPP2, TIMM50
3	1	3	FLT3
4	3	9	ATM, GCAT, GDF15
6	1	3	TET2
15	1	3	SYNE1

Table 3. **Variations detected in the targeted resequencing experiment.**

Interestingly, a 4bp insertion/deletion polymorphism (rs2307842) in the 3´UTR of *HSP90B1*, target site for miR-223 was detected (Figure 7). We hypothesized that this 'GACT' deletion disrupts the binding site for miR-223 thereby increasing the translation of HSP90B1. Both HSP90 and miR-223 have been previously implicated in CLL, thus we decided to gain insight in this field.

First we confirmed that miR-223 regulates HSP90B1 expression by 3´UTR reporter assays. The rationale for using this assay is that the binding of a given miRNA to its specific mRNA target site will repress reporter protein production (luciferase) thereby reducing activity/expression that can be measured and compared to a control (renilla). The experimental approach is to clone the 3´UTR of the target gene of interest immediately downstream of the luciferase open reading frame sequence contained in the reporter plasmid. The recombinant plasmid and the miRNA of interest are then

transiently transfected into a host cell, and luciferase activity is measured 24-48 hours after transfection.

HSP90B1 3´UTR (WT)

5´-TGTAAATTTGTACTATTT**AACTGAC**TATTCTTGATGTAAAATCTTGTC-3´

HSP90B1 3´UTR (MUT)

5´-TGTAAATTTGTACTATTT**AACT**ATTCTTGATGTAAAATCTTGTC-3´

Figure 7. HSP90B1 3´UTR wild type and mutated.
The wild-type (WT) and mutated (MUT) 3´UTR of HSP90B1, with the seed region (bold) and base deletion (bold and underlined) were sub-cloned into luciferase reporter and are shown.

The double-stranded oligonucleotides corresponding to the wild-type (WT-3'UTR) or mutant (MUT-3'UTR) miR-223 binding site in the UTR of HSP90B1 were synthesized and pmirGLO Vectors made up of an SV40 promoter, the Renilla-luciferase gene, and the 3´UTR of HSP90B1 were transfected into HEK293 cells. This assay confirmed the influence of the rs2307842 on HSP90B1 transcript stability, as the mutant reporters had more luciferase activity than their respective wild-type ones, which means that the 3´UTR polymorphism is affecting the binding of the miRNA (Figure 8).

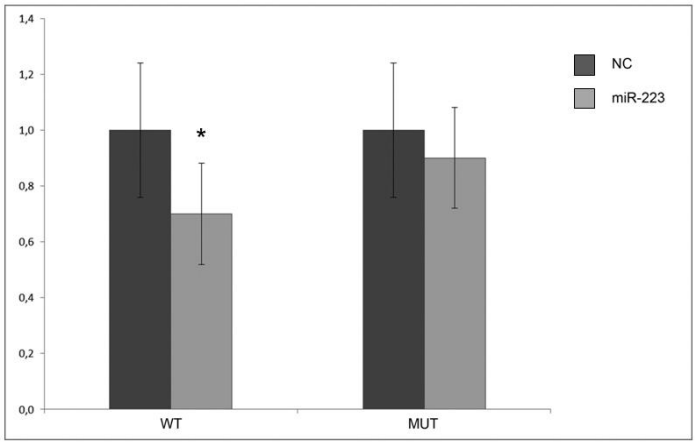

Figure 8. Luciferase reporter assays to confirm targeting of HSP90B1 3´UTR by miR-223.
Ectopic miR-223 expression inhibits wild-type but not mutant HSP90B1 3´UTR reporter activity in HEK293 cells. Cells were co-transfected with miR-223 precursor/negative control (NC) miRNA and with either WT or MUT HSP90B1 3´UTR reporter construct. Luciferase activity assay was performed at 24h after transfection. The columns represent normalized relative luciferase activity (RLU) by means with 95% confidence intervals from 4 independent experiments (Mann-Whitney test, *P<0.05).

We also validate this finding by transfecting two cell lines with synthetic miR-223 and measuring HSP90B1 expression by semi-quantitative PCR and Western blot. The cell lines were MM1S (mutant 3´UTR) and H929 (wild-type 3´UTR). All the experiments were done in triplicate and both semi-quantitative PCR and western blot analyses of HSP90B1 mRNA and protein showed that the rs2307842 determines an increased expression of the miR-223 target.

To evaluate the clinical impact of HSP90B1 3´UTR polymorphism, we expanded the study to 109 additional patients with CLL and 32 healthy controls. Sequencing of the HSP90B1 3´UTR region was performed by pyrosequencing (PyroMark Q24 system, Qiagen). rs2307842 was detected in 27/109 (25%) patients and 8/32 (25%) healthy controls, as expected. Overall, we did not find any significant relationship between rs2307842 and clinical characteristics of CLL patients. We also analyzed HSP90B1 expression in a series of 21 CLL patients (15 WT and 6 mutant) by semi-quantitative PCR and we observed a statistically significant HSP90B1 overexpression (P=0.039) in the mutant subgroup compared with WT-patients (Figure 9).

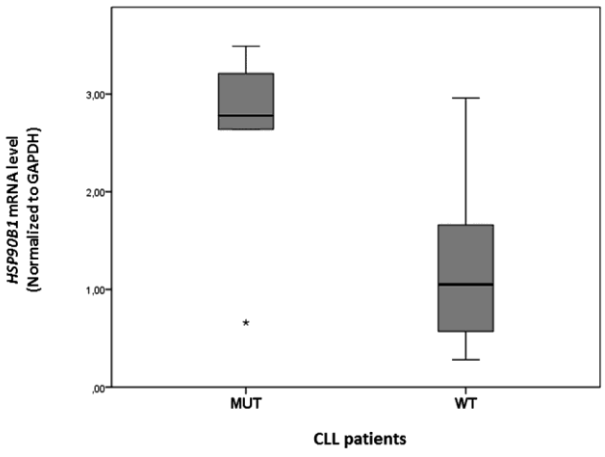

Figure 9. . Gene expression levels of HSP90B1 in CLL patients assessed by semi-quantitative PCR analysis. Box plots show the relative upregulation of HSP90B1 in CLL patients with rs2307842 (MUT) compared with wild type-CLL patients (WT). The thick line inside the box plot indicates the median expression levels and the box shows the 25th and 75th percentiles, while the whiskers show the maximum and minimum values. Outliers (extreme values falling out of the main distribution) are represented by open circles. Statistical significance was determined using the Mann-Whitney U test (P<0.05).

Thus, we demonstrate that, in CLL patients, the presence of rs2307842 in HSP90B1 3´UTR region leads to HSP90B1 overexpression due to the disruption of the miR-223 binding site.

Downregulation of miR-223 has been recently reported to have prognostic significance in CLL.[92] However, there is no evidence of the pathogenetic mechanism of this miRNA in CLL patients, and no target has been proposed or validated for miR-223 in CLL until date. Thus, this work provides novel information about how the downregulation of miR-223 can be determining the poor outcome of CLL patients, maybe through upregulation of HSP90B1 expression. These results could point out HSP90B1 as a new pathogenic mechanism in CLL, given its role in apoptosis resistance and cell survival promotion in several types of cancer. Moreover, HSP90 chaperones a number of important proteins with a key role in CLL such as ZAP-70, Akt and AID. HSP90 has been also involved in *BCL-2* regulation and apoptosis resistance in CLL. Thus, our results could be a first step towards elucidation of HSP90 role in CLL and further investigations in the field are needed.

General discussion

CLL has been established as a heterogeneous disease of remarkable diversity. No single genetic mutation or abnormality responsible for CLL development has been identified. The disease is characterized by a variety of chromosomal abnormalities, which have been shown to be important predictors of disease outcome. Considering the molecular heterogeneity of CLL, genome-wide studies and analysis of global pathways are powerful tools for the analysis of genetic alterations in CLL, rather than the study of single targets.

Comparative genomic hybridization arrays are reliable tools to investigate the genetic abnormalities in CLL with a higher resolution, compared with FISH and conventional cytogenetics. Several studies have demonstrated the presence of new, cytogenetically cryptic and recurrent chromosomal changes, such as gain of 2p and deletions of 8p, 18q and 22q. Our study identified a previously undescribed recurrent region of gain in CLL, located on 20q13 in 19% of CLL patients. This frequency is similar to other well characterized abnormalities in CLL (+12, 11q-, and 17p-). Gains in 20q13 in CLL did not occur as a single aberration, because all CLL patients with gains in this region also displayed additional genetic changes. In fact, gains on 20q were associated with genomic complexity. It is of note that genomic complexity has a significant impact on cancer prognosis and a number of studies have described the presence of several genomic changes as being predictors of disease progression and chemosensitivity in CLL.[93,94] The presence of a large number of genomic alterations in 20q13-gain patients suggest that this new genetic entity could be associated with a more advanced disease in CLL, as has been suggested in non-Hodgkin lymphomas.[95] The CLL patients with gains on 20q presented with high WBC and B-symptoms that also supports the idea that this abnormality could be a marker of evolution. However our study failed to demonstrate an association between diagnosis and progression status regarding 20q gain. Genomic instability could therefore be another molecular feature of CLL progression, as has been recently suggested.[96-98] Abnormalities of chromosome 20 are frequently observed aberrations in cancer.[99] In addition, the presence of gains on 20q has been associated with aggressive tumor behavior and poor clinical prognosis.[100,101] By contrast, deletions of the long arm of chromosome 20 are a common chromosomal abnormality associated with myeloid malignancies, and are rarely seen in lymphoid

malignancies.[99] The gain on 20q13 in CLL could be relevant to the pathogenesis and evolution of CLL because eleven protein-coding cancer-related genes have been identified in this region. It should be noted that all of these genes were upregulated in the CLL patients showing 20q gains in comparison with the other CLL cases. Therefore we suggest that 20q13.12 overexpressed genes may also be important in the evolution of CLL and warrant further studies.

Gene expression profiling is a powerful tool to better understand the biology, the clinical outcome and molecular mechanisms implicated in chronic lymphocytic leukemia (CLL). Such studies have identified new prognosis markers, being ZAP-70 the most used in the clinical practice. Gene expression profiling demonstrated a distinct pattern for CLL mutated and unmutated subtypes with ZAP-70 standing out as the gene that most stringently separated the subsets. Several studies have demonstrated that ZAP70 expression in CLL correlates strongly with IgVH mutation status and it is a powerful prognostic factor.[40,102] The study of the gene expression profile in CLL have given us insights into the molecular mechanisms involved in its pathogenesis by analyzing the impact of genomic aberrations on the expression of genes located on the corresponding loci.[90,91,103-106] . Our data revealed an association between altered transcription levels and genomic imbalances in the genetic subsets of CLL, indicating that gene dosage might have pathogenic effects in CLL. This gene dosage effect was also present in CLL displayed gain on 20q.Genomic arrays have also revealed new insights in CLL cytogenetic subgroups defined by FISH, such as 13q deletions. Although deletions of 13q are often cytogenetically criptic, they represent the most frequently observed FISH-aberration in CLL. Deletions at 13q14 are very heterogeneous in size and breakpoints. Classically, when present as a solitary aberration, the del(13q) implies a favorable prognosis, although this statement is now under discussion and 13q- CLL patients are a clinically heterogeneous group.[87,107] Several studies have reported that the extent of the deletion is associated with disease characteristics, for example del(13q) type II (long, involving RB1, related with disease progression) and del(13q) type I (short, not involving RB1, related with disease progression only when associated with other aberrations).[90]

The present study demonstrated that the number of malignant cells carrying this genetic lesion strongly correlates with the disease outcome. The clinical relevance of

the number of cells displaying a specific genetic abnormality has been recently demonstrated in CLL patients. Thus, the presence of more than 20% of cells with loss of *TP53* has been associated with an adverse prognosis, while patients with less than 20% of cells with loss of *TP53* had a prognosis similar to that of the global series.[100,101] Our results showed that, according to the percentage of cells with 13q-, two prognostic groups could be established: the patients with a high proportion (>80%) of 13q- cells had both a shorter overall survival than that of patients with <80% 13q- cells and a shorter time to first therapy. These results regarding the relationship between the number of 13q- cells and the outcome have been corroborated by two independent groups.[91,103] Moreover, to better understand the molecular mechanisms responsible for this different outcome, a molecular characterization of these patients by gene expression profiling and microRNAs studies, was carried out. We demonstrated that several pathways were upregulated in CLL patients with higher percentages of 13q- cells: BCR signaling, NFkB signaling and antiapoptotic pathways being of special interest in CLL. The BCR pathway is activated in poor prognosis CLL patients (IGHV unmutated), and the overexpression of several molecules involved in this pathway has been reported in advanced stages of the disease. This is consistent with the poor outcome observed in the CLL patients with a high load of 13q- cells. We also reported the overexpression of genes involved in promoting cell survival and antiapoptotic pathways, as well as the downregulation of several proapoptotic genes in these CLL patients. Therefore an imbalance between the proliferative and apoptotic signals could explain the lymphocytosis and the poor outcome described in these patients. The influence of other factors with prognostic relevance in CLL, such as IGVH mutational status, was discarded. Moreover, deregulation of several miRNAs was also observed. This is an interesting result, as several studies have recently identified miRNAs that could act as prognostic indicators in CLL, typically by correlating expression levels of these miRNA with established prognostic indicators.[108]

Surprisingly, our results suggest that some of the biological characteristics of CLL patients with higher percentages of 13q- cells are similar to those of high-risk cytogenetic subgroups (17p- and 11q-), since they share the deregulation of several key signaling pathways, which highlights the fact that CLL patients carrying higher percentages of 13q- cells have more aggressive clinical course. Whether a high load of

13q- cells has an independent prognostic value in CLL or it is a marker of progression remains unsolved.

In Paper I, both FISH and aCGH revealed that 13q- was a heterogeneous group in size of the deletion and percentage of cells displaying the abnormality. However, we could not evaluate the length of the deletion in this initial cohort despite using genomic arrays. Correlation between FISH and aCGH was poor in these patients, we think this is due to (1) the lack of BAC clones in the array that map 13q14 region and (2) genetic aberrations are more difficult to detect using aCGH when present in lower levels, and approximately half of our 13q- patients had lower percentages of 13q- by FISH. In fact, when aCGH failed to demonstrate the presence of 13q deletion, FISH data revealed that most of these cases had <30%. This could justify the lack of correlation between both techniques in this CLL subgroup. This was not the case for the rest of FISH subgroups (17p-, 11q- and +12), which showed excellent correlations. Thus, in our initial cohort of 13q- , we evaluated the size of the deletion by the FISH probes LSI-D13S319 probe (13q deletion) and LSI-RB probe (defining the larger 13q deletions) and analyzed the overall survival (Supplementary Figure 1, Paper II). This approach was similar to the one followed by Dal Bo et al.[90] Dal Bo et al hypothesized that 13q- cases with a high load of 13q- cells and worse prognosis have also larger 13q deletions. However our analysis showed that CLL patients with lower percentages of 13q- cells by FISH were commonly associated with shorter deletions (not involving the RB1 locus). In addition, a correlation between the presence of either a short deletion or the presence of a low number of 13q- cells and a better outcome was present. However the patients with a short deletion and a high number of 13q- cells did not show a better prognosis than patients with high number of 13q- cells showing that the presence of high number of 13q- cells is always associated worst shorter survival independently of the deleted size on 13q- (Supplementary Figure 1, Paper II).

Altogether, our data support the concept that patients with 13q- do not constitute homogeneous group and may have implications for the biology and the prognostic stratification of CLL.

In the last part of this thesis, we present a mutational study in CLL patients by NGS techniques. Recent results highlight the great impact of these massively parallel

platforms on genetics. The classical mutational studies in CLL involve immunoglobulin heavy chain, ATM and P53 genes. However, whole genome sequencing has provided new insights in the mutational status of this disease, involving several genes previously unrelated to CLL.[74,75] One of the most striking aspects of these studies is that, despite the high number of patients analyzed (xxx in total), the number of genes showing recurrent mutations (eight in the Spanish studies[74,75]) and the percentages of patients with these mutations are very low. This confirms the concept of CLL as a very heterogeneous disease.

We decided to follow a targeted resequencing strategy to capture the genes and regions of interest. Our selection was based on literature results (*ATM, P53, PHLPP, PI3K*) and from the results of the gene expression profiling studies (*HSP90B1, E2F1, RAPGEF2*). We detected a high number of variants, most of them annotated polymorphisms. Just 10% of detected variants had not been previously described. However, we did not reach a significant incidence, in spite of extending our series.

Our custom-made data analysis allowed us to determine if any of the detected variants were related to microRNAs. Thus, we have detected the 4 bp deletion polymorphism (rs rs2307842) in HSP90B1 3´UTR, more specifically in the miR-223 binding site. miR-223 is part of the microRNA signature described by Calin et al which differentiates *IGVH*mutated/ZAP70+ and *IGVH*/ZAP70- CLL patients.[56] Moreover, miR-223 downregulation has been associated with disease progression and shorter overall survival and TFT in CLL.[55,86,92,109] Despite the proven implication of miR-223 expression in CLL prognosis, little is known on the molecular mechanisms which could be responsible for the poor outcome of these patients. In Paper IV we demonstrate that rs2307842 polymorphism alters the regulation of HSP90B1 expression by disrupting of the miR-223 binding site, leadind to the overexpression of this gene. Our date validate HSP90B1 as a target of miR-223. Moreover, our results show that CLL patients with rs2307842 have overexpression of HSP90B1. Interestingly, these patients usually have *IGVH* unmutated. In addition, CLL patients showing overexpression of HSP90B1 (both wt and with the rs2304872) share poor prognostic factors such as *IGVH* unmutated status or ZAP70 expression. HSP90 is a heat shock protein that stabilizes several important enzymes and antiapoptotic proteins with relevant functions in CLL REFs. HSP90 is commonly overexpressed in cancer, both solid and hematological types.

Therefore HSP90 inhibitors have been proposed as a novel therapeutic option of CLL, particularly as their mechanism of action appears independent of mutations of ATM or TP53. These agents have been reported to be effective in refractory CLL cases, alone or in combination with classical therapies by inducing apoptosis in vitro.[64,110-115] Thus our results open new avenues for exploring innovative pathways in CLL microRNA regulation, pathogenesis and therapy.

Concluding remarks

1. The use of genomic microarrays shows chromosomal alterations in 75% patients with CLL. Besides FISH-detected alterations, genomic microarrays identify new recurrent aberrations such as losses in 11q13.3, 5q13.3–q14.1, 5q31.1 and 7q22, as well as gains in 1q21.3–q22, 11q13.3, 16q23.2–q24.2, 6p21.31–p21.1, 20q13.12 and 10q22.3. These data confirm the heterogeneity of CLL.

2. Gene expression profiling in CLL patients shows that a significant number of the differentially expressed genes cluster in chromosomal regions affected by the respective genomic imbalances. This observation suggests that changes in gene expression are related with copy-number alterations, which confirms the gene dosage effect previously described in cancer.

3. A combined analysis of the genome and transcriptome of CLL patients reveals the presence of a new recurrent gain in 20q. Genomic analysis identifies the minimal region of gain on 20q13.12, with a size of ~2.31 Mb and present in 19% CLL patients. This gain is characterized by the overexpression of the genes located in it and a high frequency of chromosomal changes.

4. CLL patients showing 13q deletion as a sole abnormality are a heterogeneous group and the number of cells displaying this aberration is related to the outcome of these patients. CLL patients with a higher load of 13q- cells (>80%) are characterized by both a short overall survival and time to first therapy.

5. Clonal B lymphocytes from CLL patients with a high number of 13q- losses have more proliferation and less apoptosis than clonal B lymphocytes from CLL patients with a low number of 13q- losses, due to the deregulation of several genes involved in these pathways, as well as microRNAS such as miR-15a, miR-29a, miR-34a, miR-155 and miR-181b. The gene expression profile of CLL patients with a high number of losses in 13q is similar to the observed in high risk CLL patients (17p- and 11q-). Altogether, these results could justify the clinical differences observed between both 13q- CLL subgroups.

6. The presence of the polymorphism rs2307842 in the 3´UTR of *HSP90B1* disrupts the binding site for miR-223 and interferes with miR-223 function. This leads to the overexpression of *HSP90B1* only in B lymphocytes of CLL patients.

7. Hsa-miR-223 modulates *HSP90B1* expression. These results could provide a plausible explanation as to why CLL patients harboring miR-223 downregulatio are associated with a poor outcome. A correlation between *HSP90B1* and *BCL2* overexpression in CLL patients was observed. This could determine a higher resistance to apoptosis in CLL patients with *HSP90B1* overexpression.

References

1. Trask BJ. Human cytogenetics: 46 chromosomes, 46 years and counting. Nat Rev Genet 2002;3:769-778.

2. Speicher MR, Carter NP. The new cytogenetics: blurring the boundaries with molecular biology. Nat Rev Genet 2005;6:782-792.

3. Pinkel D, Segraves R, Sudar D et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 1998;20:207-211.

4. Solinas-Toldo S, Lampel S, Stilgenbauer S et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer 1997;20:399-407.

5. Shinawi M, Cheung SW. The array CGH and its clinical applications. Drug Discov Today 2008;13:760-770.

6. Cheung SW, Shaw CA, Yu W et al. Development and validation of a CGH microarray for clinical cytogenetic diagnosis. Genet Med 2005;7:422-432.

7. Shaffer LG, Kashork CD, Saleki R et al. Targeted genomic microarray analysis for identification of chromosome abnormalities in 1500 consecutive clinical cases. J Pediatr 2006;149:98-102.

8. Buiting K, Gross S, Lich C et al. Epimutations in Prader-Willi and Angelman syndromes: a molecular study of 136 patients with an imprinting defect. Am J Hum Genet 2003;72:571-577.

9. Redon R, Ishikawa S, Fitch KR et al. Global variation in copy number in the human genome. Nature 2006;444:444-454.

10. Peiffer DA, Le JM, Steemers FJ et al. High-resolution genomic profiling of chromosomal aberrations using Infinium whole-genome genotyping. Genome Res 2006;16:1136-1148.

11. Mamanova L, Coffey AJ, Scott CE et al. Target-enrichment strategies for next-generation sequencing. Nat Methods 2010;7:111-118.

12. Margulies M, Egholm M, Altman WE et al. Genome sequencing in microfabricated high-density picolitre reactors. Nature 2005;437:376-380.

13. Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 2008;9:387-402.

14. Dighiero G, Hamblin TJ. Chronic lymphocytic leukaemia. Lancet 2008;371:1017-1029.

15. Matutes E, Owusu-Ankomah K, Morilla R et al. The immunological profile of B-cell disorders and proposal of a scoring system for the diagnosis of CLL. Leukemia 1994;8:1640-1645.

16. Rossi D, Gaidano G. Richter syndrome: molecular insights and clinical perspectives. Hematol Oncol 2009;27:1-10.

17. Watson L, Wyld P, Catovsky D. Disease burden of chronic lymphocytic leukaemia within the European Union. Eur J Haematol 2008;81:253-258.

18. Wierda WG, O'Brien S, Wang X et al. Prognostic nomogram and index for overall survival in previously untreated patients with chronic lymphocytic leukemia. Blood 2007;109:4679-4685.

19. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med 2005;352:804-815.

20. Hallek M, Cheson BD, Catovsky D 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 2008;111:5446-5456.

21. Rai KR, Sawitsky A, Cronkite EP et al. Clinical staging of chronic lymphocytic leukemia. Blood 1975;46:219-234.

22. Binet JL, Auquier A, Dighiero G et al. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer 1981;48:198-206.

23. Juliusson G, Oscier DG, Fitchett M et al. Prognostic subgroups in B-cell chronic lymphocytic leukemia defined by specific chromosomal abnormalities. N Engl J Med 1990;323:720-724.

24. Mayr C, Speicher MR, Kofler DM et al. Chromosomal translocations are associated with poor prognosis in chronic lymphocytic leukemia. Blood 2006;107:742-751.

25. Dicker F, Schnittger S, Haferlach T, Kern W, Schoch C. Immunostimulatory oligonucleotide-induced metaphase cytogenetics detect chromosomal aberrations in 80% of CLL patients: A study of 132 CLL cases with correlation to FISH, IgVH status, and CD38 expression. Blood 2006;108:3152-3160.

26. Haferlach C, Dicker F, Schnittger S, Kern W, Haferlach T. Comprehensive genetic characterization of CLL: a study on 506 cases analysed with chromosome banding analysis, interphase FISH, IgV(H) status and immunophenotyping. Leukemia 2007;21:2442-2451.

27. Zenz T, Mertens D, Kuppers R, Dohner H, Stilgenbauer S. From pathogenesis to treatment of chronic lymphocytic leukaemia. Nat Rev Cancer 2010;10:37-50.

28. Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum Genet 1988;80:224-234.

29. Dohner H, Stilgenbauer S, Benner A et al. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med 2000;343:1910-1916.

30. Damle RN, Wasil T, Fais F et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood 1999;94:1840-1847.

31. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. Blood 1999;94:1848-1854.

32. Schroeder HW, Jr., Dighiero G. The pathogenesis of chronic lymphocytic leukemia: analysis of the antibody repertoire. Immunol Today 1994;15:288-294.

33. Klein U, Tu Y, Stolovitzky GA et al. Gene expression profiling of B cell chronic lymphocytic leukemia reveals a homogeneous phenotype related to memory B cells. J Exp Med 2001;194:1625-1638.

34. Rosenwald A, Alizadeh AA, Widhopf G et al. Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. J Exp Med 2001;194:1639-1647.

35. Klein U, Goossens T, Fischer M et al. Somatic hypermutation in normal and transformed human B cells. Immunol Rev 1998;162:261-280.

36. Tangye SG, Liu YJ, Aversa G, Phillips JH, de Vries JE. Identification of functional human splenic memory B cells by expression of CD148 and CD27. J Exp Med 1998;188:1691-1703.

37. Vasconcelos Y, De VJ, Vallat L et al. Gene expression profiling of chronic lymphocytic leukemia can discriminate cases with stable disease and mutated Ig genes from those with progressive disease and unmutated Ig genes. Leukemia 2005;19:2002-2005.

38. Van BF, Verhasselt B, Philippe J. Prognostic markers in chronic lymphocytic leukemia: a comprehensive review. Blood Rev 2009;23:25-47.

39. Bosch F, Muntanola A, Gine E et al. Clinical implications of ZAP-70 expressionin chronic lymphocytic leukemia. Cytometry B Clin Cytom 2006;70:214-217.

40. Haslinger C, Schweifer N, Stilgenbauer S et al. Microarray gene expression profiling of B-cell chronic lymphocytic leukemia subgroups defined by genomic aberrations and VH mutation status. J Clin Oncol 2004;22:3937-3949.

41. Dyer MJ, Zani VJ, Lu WZ et al. BCL2 translocations in leukemias of mature B cells. Blood 1994;83:3682-3688.

42. Vrhovac R, Delmer A, Tang R et al. Prognostic significance of the cell cycle inhibitor p27Kip1 in chronic B-cell lymphocytic leukemia. Blood 1998;91:4694-4700.

43. Messmer BT, Messmer D, Allen SL et al. In vivo measurements document the dynamic cellular kinetics of chronic lymphocytic leukemia B cells. J Clin Invest 2005;115:755-764.

44. Lagneaux L, Delforge A, Dorval C, Bron D, Stryckmans P. Excessive production of transforming growth factor-beta by bone marrow stromal cells in B-cell chronic

lymphocytic leukemia inhibits growth of hematopoietic precursors and interleukin-6 production. Blood 1993;82:2379-2385.

45. Buske C, Gogowski G, Schreiber K et al. Stimulation of B-chronic lymphocytic leukemia cells by murine fibroblasts, IL-4, anti-CD40 antibodies, and the soluble CD40 ligand. Exp Hematol 1997;25:329-337.

46. Granziero L, Ghia P, Circosta P et al. Survivin is expressed on CD40 stimulation and interfaces proliferation and apoptosis in B-cell chronic lymphocytic leukemia. Blood 2001;97:2777-2783.

47. Croce CM. Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet 2009;10:704-714.

48. Grimson A, Farh KK, Johnston WK et al. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 2007;27:91-105.

49. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281-297.

50. Di LG, Calin GA, Croce CM. MicroRNAs: fundamental facts and involvement in human diseases. Birth Defects Res C Embryo Today 2006;78:180-189.

51. Calin GA, Sevignani C, Dumitru CD et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 2004;101:2999-3004.

52. Calin GA, Dumitru CD, Shimizu M et al. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002;99:15524-15529.

53. Cimmino A, Calin GA, Fabbri M et al. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 2005;102:13944-13949.

54. Nana-Sinkam SP, Croce CM. MicroRNA in chronic lymphocytic leukemia: transitioning from laboratory-based investigation to clinical application. Cancer Genet Cytogenet 2010;203:127-133.

55. Visone R, Rassenti LZ, Veronese A et al. Karyotype-specific microRNA signature in chronic lymphocytic leukemia. Blood 2009;114:3872-3879.

56. Calin GA, Ferracin M, Cimmino A et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353:1793-1801.

57. Nicoloso MS, Kipps TJ, Croce CM, Calin GA. MicroRNAs in the pathogeny of chronic lymphocytic leukaemia. Br J Haematol 2007;139:709-716.

58. Bichi R, Shinton SA, Martin ES et al. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. Proc Natl Acad Sci U S A 2002;99:6955-6960.

59. Pekarsky Y, Santanam U, Cimmino A et al. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. Cancer Res 2006;66:11590-11593.

60. Nylandsted J, Brand K, Jaattela M. Heat shock protein 70 is required for the survival of cancer cells. Ann N Y Acad Sci 2000;926:122-125.

61. Broemer M, Krappmann D, Scheidereit C. Requirement of Hsp90 activity for IkappaB kinase (IKK) biosynthesis and for constitutive and inducible IKK and NF-kappaB activation. Oncogene 2004;23:5378-5386.

62. Sato S, Fujita N, Tsuruo T. Modulation of Akt kinase activity by binding to Hsp90. Proc Natl Acad Sci U S A 2000;97:10832-10837.

63. Kamal A, Thao L, Sensintaffar J et al. A high-affinity conformation of Hsp90 confers tumour selectivity on Hsp90 inhibitors. Nature 2003;425:407-410.

64. Castro JE, Prada CE, Loria O et al. ZAP-70 is a novel conditional heat shock protein 90 (Hsp90) client: inhibition of Hsp90 leads to ZAP-70 degradation, apoptosis, and impaired signaling in chronic lymphocytic leukemia. Blood 2005;106:2506-2512.

65. Trentin L, Frasson M, Donella-Deana A et al. Geldanamycin-induced Lyn dissociation from aberrant Hsp90-stabilized cytosolic complex is an early event in apoptotic mechanisms in B-chronic lymphocytic leukemia. Blood 2008;112:4665-4674.

66. Ringshausen I, Schneller F, Bogner C et al. Constitutively activated phosphatidylinositol-3 kinase (PI-3K) is involved in the defect of apoptosis in B-CLL: association with protein kinase Cdelta. Blood 2002;100:3741-3748.

67. Contri A, Brunati AM, Trentin L et al. Chronic lymphocytic leukemia B cells contain anomalous Lyn tyrosine kinase, a putative contribution to defective apoptosis. J Clin Invest 2005;115:369-378.

68. Hertlein E, Wagner AJ, Jones J et al. 17-DMAG targets the nuclear factor-kappaB family of proteins to induce apoptosis in chronic lymphocytic leukemia: clinical implications of HSP90 inhibition. Blood 2010;116:45-53.

69. Pepper C, Lin TT, Pratt G et al. Mcl-1 expression has in vitro and in vivo significance in chronic lymphocytic leukemia and is associated with other poor prognostic markers. Blood 2008;112:3807-3817.

70. Pepper C, Hewamana S, Brennan P, Fegan C. NF-kappaB as a prognostic marker and therapeutic target in chronic lymphocytic leukemia. Future Oncol 2009;5:1027-1037.

71. Heintel D, Kroemer E, Kienle D et al. High expression of activation-induced cytidine deaminase (AID) mRNA is associated with unmutated IGVH gene status and unfavourable cytogenetic aberrations in patients with chronic lymphocytic leukaemia. Leukemia 2004;18:756-762.

72. Palacios F, Moreno P, Morande P et al. High expression of AID and active class switch recombination might account for a more aggressive disease in unmutated CLL patients: link with an activated microenvironment in CLL disease. Blood 2010;115:4488-4496.

73. Orthwein A, Patenaude AM, Affar eB et al. Regulation of activation-induced deaminase stability and antibody gene diversification by Hsp90. J Exp Med 2010;207:2751-2765.

74. Puente XS, Pinyol M, Quesada V et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature 2011;475:101-105.

75. Quesada V, Conde L, Villamor N et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet 2011;44:47-52.

76. Wang L, Lawrence MS, Wan Y et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med 2011;365:2497-2506.

77. Papaemmanuil E, Cazzola M, Boultwood J et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med 2011;365:1384-1395.

78. Pfeifer D, Pantic M, Skatulla I et al. Genome-wide analysis of DNA copy number changes and LOH in CLL using high-density SNP arrays. Blood 2007;109:1202-1210.

79. Lehmann S, Ogawa S, Raynaud SD et al. Molecular allelokaryotyping of early-stage, untreated chronic lymphocytic leukemia. Cancer 2008;112:1296-1305.

80. Gunn SR, Bolla AR, Barron LL et al. Array CGH analysis of chronic lymphocytic leukemia reveals frequent cryptic monoallelic and biallelic deletions of chromosome 22q11 that include the PRAME gene. Leuk Res 2009;33:1276-1281.

81. Patel A, Kang SH, Lennon PA et al. Validation of a targeted DNA microarray for the clinical evaluation of recurrent abnormalities in chronic lymphocytic leukemia. Am J Hematol 2008;83:540-546.

82. Schwaenen C, Nessling M, Wessendorf S et al. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. Proc Natl Acad Sci U S A 2004;101:1039-1044.

83. Forconi F, Rinaldi A, Kwee I et al. Genome-wide DNA analysis identifies recurrent imbalances predicting outcome in chronic lymphocytic leukaemia with 17p deletion. Br J Haematol 2008;143:532-536.

84. Grubor V, Krasnitz A, Troge JE et al. Novel genomic alterations and clonal evolution in chronic lymphocytic leukemia revealed by representational oligonucleotide microarray analysis (ROMA). Blood 2009;113:1294-1303.

85. Aalto Y, El-Rifa W, Vilpo L et al. Distinct gene expression profiling in chronic lymphocytic leukemia with 11q23 deletion. Leukemia 2001;15:1721-1728.

86. Li S, Moffett HF, Lu J et al. MicroRNA expression profiling identifies activated B cell status in chronic lymphocytic leukemia cells. PLoS One 2011;6:e16956.

87. Catovsky D, Richards S, Matutes E et al. Assessment of fludarabine plus cyclophosphamide for patients with chronic lymphocytic leukaemia (the LRF CLL4 Trial): a randomised controlled trial. Lancet 2007;370:230-239.

88. Krober A, Seiler T, Benner A et al. V(H) mutation status, CD38 expression level, genomic aberrations, and survival in chronic lymphocytic leukemia. Blood 2002;100:1410-1416.

89. Cheson BD, Bennett JM, Grever M et al. National Cancer Institute-sponsored Working Group guidelines for chronic lymphocytic leukemia: revised guidelines for diagnosis and treatment. Blood 1996;87:4990-4997.

90. Dal BM, Rossi FM, Rossi D et al. 13q14 Deletion size and number of deleted cells both influence prognosis in chronic lymphocytic leukemia. Genes Chromosomes Cancer 2011;50:633-643.

91. Van Dyke DL, Shanafelt TD, Call TG et al. A comprehensive evaluation of the prognostic significance of 13q deletions in patients with B-chronic lymphocytic leukaemia. Br J Haematol 2010;148:544-550.

92. Stamatopoulos B, Meuleman N, Haibe-Kains B et al. microRNA-29c and microRNA-223 down-regulation has in vivo significance in chronic lymphocytic leukemia and improves disease risk stratification. Blood 2009;113:5237-5245.

93. Kujawski L, Ouillette P, Erba H et al. Genomic complexity identifies patients with aggressive chronic lymphocytic leukemia. Blood 2008;112:1993-2003.

94. Kay NE, Eckel-Passow JE, Braggio E et al. Progressive but previously untreated CLL patients with greater array CGH complexity exhibit a less durable response to chemoimmunotherapy. Cancer Genet Cytogenet 2010;203:161-168.

95. Stephens PJ, Greenman CD, Fu B et al. Massive Genomic Rearrangement Acquired in a Single Catastrophic Event during Cancer Development. Cell 2011;144:27-40.

96. Yang SH, Seo MY, Jeong HJ et al. Gene copy number change events at chromosome 20 and their association with recurrence in gastric cancer patients. Clin Cancer Res 2005;11:612-620.

97. Zhu H, Lam DC, Han KC et al. High resolution analysis of genomic aberrations by metaphase and array comparative genomic hybridization identifies candidate tumour genes in lung cancer cell lines. Cancer Lett 2007;245:303-314.

98. Lassmann S, Weis R, Makowiec F et al. Array CGH identifies distinct DNA copy number profiles of oncogenes and tumor suppressor genes in chromosomal- and microsatellite-unstable sporadic colorectal carcinomas. J Mol Med 2007;85:293-304.

99. Bench AJ, Nacheva EP, Hood TL et al. Chromosome 20 deletions in myeloid malignancies: reduction of the common deleted region, generation of a PAC/BAC contig and identification of candidate genes. UK Cancer Cytogenetics Group (UKCCG). Oncogene 2000;19:3902-3913.

100. Guarini A, Chiaretti S, Tavolaro S et al. BCR ligation induced by IgM stimulation results in gene expression and functional changes only in IgV H unmutated chronic lymphocytic leukemia (CLL) cells. Blood 2008;112:782-792.

101. Rodriguez A, Villuendas R, Yanez L et al. Molecular heterogeneity in chronic lymphocytic leukemia is dependent on BCR signaling: clinical correlation. Leukemia 2007;21:1984-1991.

102. Dickinson JD, Joshi A, Iqbal J et al. Genomic abnormalities in chronic lymphocytic leukemia influence gene expression by a gene dosage effect. Int J Mol Med 2006;17:769-778.

103. Hernandez JA, Rodriguez AE, Gonzalez M et al. A high number of losses in 13q14 chromosome band is associated with a worse outcome and biological differences in patients with B-cell chronic lymphoid leukemia. Haematologica 2009;94:364-371.

104. Parker H, Rose-Zerilli MJ, Parker A et al. 13q deletion anatomy and disease progression in patients with chronic lymphocytic leukemia. Leukemia 2011;25:489-497.

105. Ouillette P, Erba H, Kujawski L et al. Integrated genomic profiling of chronic lymphocytic leukemia identifies subtypes of deletion 13q14
1. Cancer Res 2008;68:1012-1021.

106. Mosca L, Fabris S, Lionetti M et al. Integrative genomics analyses reveal molecularly distinct subgroups of B-cell chronic lymphocytic leukemia patients with 13q14 deletion. Clin Cancer Res 2010;16:5641-5653.

107. Thornton PD, Gruszka-Westwood AM, Hamoudi RA et al. Characterisation of TP53 abnormalities in chronic lymphocytic leukaemia. Hematol J 2004;5:47-54.

108. Ward BP, Tsongalis GJ, Kaur P. MicroRNAs in chronic lymphocytic leukemia. Exp Mol Pathol 2011;90:173-178.

109. Zhou K, Yi S, Yu Z et al. MicroRNA-223 expression is uniformly down-regulated in B cell lymphoproliferative disorders and is associated with poor survival in patients with chronic lymphocytic leukemia. Leuk Lymphoma 2012;53:1155-1161.

110. Best OG, Che Y, Singh N et al. The Hsp90 inhibitor SNX-7081 synergizes with and restores sensitivity to fludarabine in chronic lymphocytic leukemia cells with lesions in the TP53 pathway: a potential treatment strategy for fludarabine refractory disease. Leuk Lymphoma 2012

111. Best OG, Singh N, Forsyth C, Mulligan SP. The novel Hsp-90 inhibitor SNX7081 is significantly more potent than 17-AAG against primary CLL cells and a range of haematological cell lines, irrespective of lesions in the TP53 pathway. Br J Haematol 2010;151:185-188.

112. McCaig AM, Cosimo E, Leach MT, Michie AM. Dasatinib inhibits B cell receptor signalling in chronic lymphocytic leukaemia but novel combination approaches are required to overcome additional pro-survival microenvironmental signals. Br J Haematol 2011;153:199-211.

113. Hertlein E, Wagner AJ, Jones J et al. 17-DMAG targets the nuclear factor-kappaB family of proteins to induce apoptosis in chronic lymphocytic leukemia: clinical implications of HSP90 inhibition. Blood 2010;116:45-53.

114. Johnson AJ, Wagner AJ, Cheney CM et al. Rituximab and 17-allylamino-17-demethoxygeldanamycin induce synergistic apoptosis in B-cell chronic lymphocytic leukaemia. Br J Haematol 2007;139:837-844.

115. Jones DT, Addison E, North JM et al. Geldanamycin and herbimycin A induce apoptotic killing of B chronic lymphocytic leukemia cells and augment the cells' sensitivity to cytotoxic drugs. Blood 2004;103:1855-1861.