UNIVERSITÀ DEL PIEMONTE ORIENTALE

School of Medicine

Department of Health Sciences

Master's degree in medical biotechnologies

Different immunoglobulin light chain genes rearrangements predict survival and Richter transformation in patients with chronic lymphocytic leukemia

Tutor: **Chiar.mo Prof. Gianluca GAIDANO**

Firmato digitalmente da Gianluca Gaidano
Data: 26.06.2024 13:51:18 CEST
Organizzazione: UNIVERSITA' DEGLI STUDI DEL PIEMONTE ORIENTALE/01943490027

Candidate: **Michel BADER**

Matricula Number: **20051533**

Academic Year 2023/2024

Table of Contents

1. SUMMARY

Chronic lymphocytic leukemia (CLL) is a heterogeneous disease with diverse clinical outcomes influenced by various molecular features. The mutational status of immunoglobulin heavy chain variable region (IGHV) has been a robust prognostic indicator, with mutated IGHV (M-IGHV) expressing patients associating to better survival compared to unmutated IGHV (UM-IGHV) patients. This study investigated the impact of light chain rearrangements on clinical outcomes, including overall survival (OS) and Richter transformation, in a prospective cohort of CLL patients. A total of 573 CLL patients were analysed, and their light chain gene rearrangements were identified through Sanger sequencing using leader and framework region 1 (FR1) primers. Statistical analysis was performed to evaluate the association of these rearrangements with OS and Richter transformation.

Among the 573 patients, 530 exhibited at least one productive light chain rearrangement. Kappa and lambda light chain genes were expressed in 57.7% and 31.3% of CLL patients, respectively. A total of 51 different productive IGKV (25) and IGLV (26) genes were identified. The most frequently rearranged kappa gene was IGKV4-1, present in 20.5% of patients with kappa rearrangements, while IGLV3-21 was the most common lambda gene, found in 19% of lambda expressing cases. The IGKV1-39 rearrangement was significantly associated with shorter OS, and it independently predicted poor prognosis. Additionally, specific light chain rearrangements, IGKV1-39, IGKV6-21, IGLV1-36, and IGLV8-61, were linked to a higher risk of Richter transformation. Multivariate analysis confirmed these rearrangements as independent risk factors for Richter transformation.

Our findings highlighted the prognostic significance of light chain gene rearrangements in CLL, with specific rearrangements correlating with poor survival and increased risk of Richter transformation. These insights underscored the potential of light chain rearrangements as biomarkers for risk stratification and targeted therapeutic strategies in CLL management.

2. INTRODUCTION

2.1 Epidemiology

In the 1970s, chronic lymphocytic leukemia (CLL) became firmly established as a clinical entity, building upon the groundbreaking work independently conducted by Dr. David Galton in London and Dr. William Dameshek in Boston during the 1950s and 1960s.[1]

CLL is the most prevalent adult leukemia in the Western world, representing approximately 25% of all adult leukemias and 25% of non-Hodgkin's lymphoma (NHL) cases. In the United States, the age-adjusted incidence of CLL is 4.5 cases per 100,000 individuals, with a median diagnosis age of 71 years.[2] Clinical manifestations of CLL encompass asymptomatic peripheral blood lymphocytosis or leukocytosis (mainly lymphocytosis), lymphadenopathy, hepatosplenomegaly, bone marrow failure, recurrent infections, and frequently autoimmune hemolytic anemia or autoimmune thrombocytopenia.[3]

2.2 Histogenesis

CLL is a B cell malignancy distinguished by the accumulation of mature clonal B cells expressing CD5. Its prevalence substantially rises with advancing age. Importantly, CLL cells display functional B cell receptors (BCRs) on their surfaces.[4] Somatic hypermutations occur during the normal B cell development process in germinal centres, particularly during the transition from naïve to memory B cells. CLL is categorized into two subgroups depending on the presence of somatic hypermutations within the variable regions of the immunoglobulin heavy chain gene (IGHV). Patients with CLL exhibiting mutated *IGHV* (*IGHV-M* CLL) generally have a more favourable prognosis compared to those with unmutated IGHV (*IGHV-UM* CLL).[5] Early investigations suggested separate origins for the two types of CLL, positing that *IGHV-UM* CLL arose from naïve B cells, while *IGHV-M* CLL originated from antigen-experienced B cells, such as memory B cells.[6]

However, immunological assessments of CLL-BCRs indicated that both types recognized self-antigens, at least under laboratory conditions, implying that CLL originates from self-reactive B cell precursors regardless of the IGHV mutation status.[7] The self-reactivity of B cell receptors (BCRs) in CLL has been extensively studied, representing one of the most investigated biological aspects of human CLL. Constitutively activated BCR signalling is a critical biological characteristic of CLL cells, providing rationale for BCR signalling-targeted therapies for CLL. Notably, approximately 1% of CLL cases exhibit a nearly identical amino

acid sequence in the heavy chain complementarity-determining region 3 (HCDR3) of their BCRs. Furthermore, in around 30% of CLL cases, the BCRs of CLL cells can be classified as specific stereotyped BCRs. These observations led to the hypothesis that some common self-antigens recognized by CLL-BCRs drive clonal expansion and play a significant role in CLL pathogenesis.[89]

The natural progression of many CLLs involves a transition towards a more malignant disease state. In most, if not all, cases, CLL is preceded by monoclonal B cell lymphocytosis (MBL), characterized by a very slow expansion of less than 5,000 monoclonal B cells in the peripheral blood.[10] MBL is detectable in approximately 5% of the elderly population and carries a risk of evolving into CLL of around 1% per year. Clinically, CLL exhibits significant heterogeneity, ranging from patients with highly stable disease and nearly normal life expectancy to those with rapidly progressive disease who face a shortened lifespan.[11] This variability in disease course is driven, at least in part, by heterogeneity in disease biology. Over time, a small fraction of CLL cases undergo transformation into a highly aggressive form known as Richter syndrome (RS), which morphologically resembles diffuse large B cell lymphoma (DLBCL).[12]

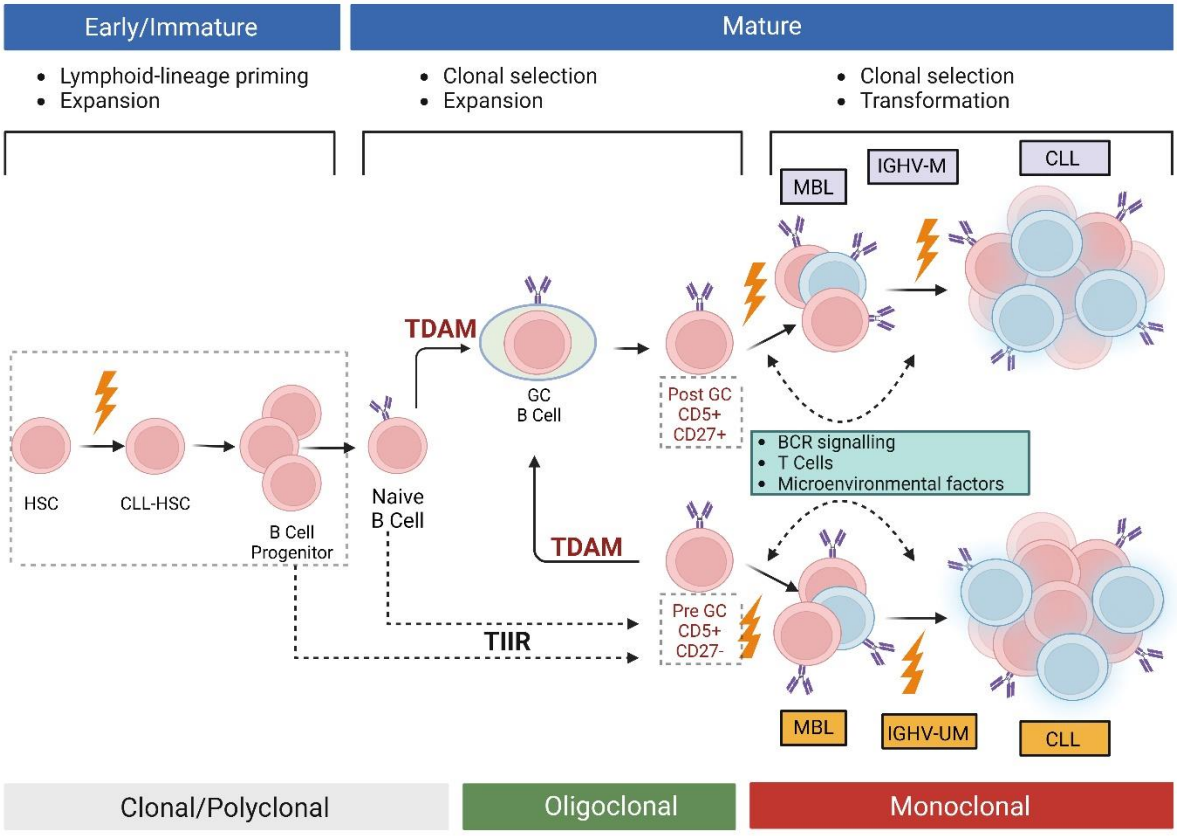

Figure 1 The cellular origin of chronic lymphocytic leukemia.

2.3 Pathogenesis

Chromosomal abnormalities in CLL have undergone extensive investigation and are utilized for risk assessments of the disease, alongside the genetic predisposition revealed by genome-wide association studies (GWAS). CLL cells exhibit immunoglobulin heavy chain (IGH)-related translocations at significantly lower frequencies compared to other types of mature B cell malignancies. This genetic feature serves to distinguish CLL from other mature B cell malignancies.[13]

The most encountered genetic abnormalities in CLL are deletions of *13q14* (*del13q14*), which are present in approximately 50%–60% of cases. These deletions are predominantly monoallelic and are more prevalent in IGHV-M CLL cases compared to *IGHV-UM* CLL cases. Although the clinical progression of CLL may be hastened in patients with extensive deletions involving the retinoblastoma gene (*RB1*).[14]

The long non-coding RNAs DLEU2 and DLEU1, along with the microRNA clusters *miR15A–miR16-1*, are situated within the minimal deleted region associated with *del13q14.*[15] Mice models engineered to harbor deletions of the corresponding murine locus developed clonal B cell lymphoproliferative disorders, indicating the significant involvement of microRNAs in CLL pathogenesis.[16] Moreover, the deletion of microRNAs *miR-15A–miR-16-1* led to the overexpression of BCL2, which offers a rationale for CLL therapeutic strategies aimed at targeting BCL2.[17]

The second most observed chromosomal abnormality in CLL is *trisomy 12*, detected in approximately 15% of patients. *Trisomy 12* represents a genetic lesion associated with intermediate risk.[18] The third most frequent chromosomal abnormality observed in CLL is deletions in the *11q22–q23* (*del11q*) chromosomal region, identified in approximately 15% of cases.[19] *Del11q* results in the loss of the tumour suppressor gene ataxia telangiectasia mutated (*ATM*), which encodes the DNA damage response (DDR) kinase *ATM.*[20] It's worth noting that approximately 25% of CLL patients with *del11q* deletions also carry mutations in the remaining *ATM* allele. The presence of both *del11q* and ATM mutation is linked to a poor prognosis.[21]

Deletions in the *17p13* chromosomal locus (*del17p*) are present in about 10% of CLL patients and are more commonly found in *IGHV-UM* CLL cases compared to *IGHV-M* CLL cases.[22] Del17p deletions typically encompass the entire short arm of chromosome *17*, resulting in the loss of the tumour suppressor gene *TP53.*[23] Missense mutations in the remaining *TP53* allele are detected in approximately 80% of CLL patients with *del17p*. This is in line with the

inactivation of *TP53*. Patients with *del17p* demonstrate high genomic complexity and generally have a worse overall prognosis compared to those with wild-type *TP53.* [24]

In addition to the extensive chromosomal abnormalities described above, advances in NGS technologies revealed recurrent driver mutations in CLL such as *SF3B1*, *ATM*, *TP53*, *NOTCH1*, *POT1*, *CHD2*, *XPO1*, *BIRC3*, *BRAF*, *MYD88*, *EGR2*, *MED 12*, *FBXW7*, *ASXL1*, *KRAS*, *NRAS*, *MAP2K1*, *NFKBIE*, *TRAF3*, *RPS15*, and *DDX3X*. [25]

The most mutated gene in CLL is *SF3B1*, identified in approximately 10%–15% of cases. Among *SF3B1* mutations, the *SF3B1 K700E* mutation is the most prevalent. [26] *SF3B1* mutations lead to alternative splicing in CLL cells, resulting in RNA changes that impact various CLL-associated pathways, including DNA damage response (DDR), telomere maintenance, and *NOTCH* signalling. Besides *SF3B1*, several other genes involved in RNA splicing, processing, and transport, such as *DDX3X* and *XPO1*, are mutated in CLL at lower frequencies. This indicates that deregulated RNA processing is a significant pathogenic pathway in CLL development.[27]

NOTCH1 is the second most frequently mutated gene in CLL, present in approximately 10% of cases, with mutations more commonly observed in *IGHV-UM* CLL. Notably, around 40% of patients with *NOTCH1-mutated* CLL also harbour *trisomy 12*, suggesting a relationship between these two genetic abnormalities and CLL development.[28]

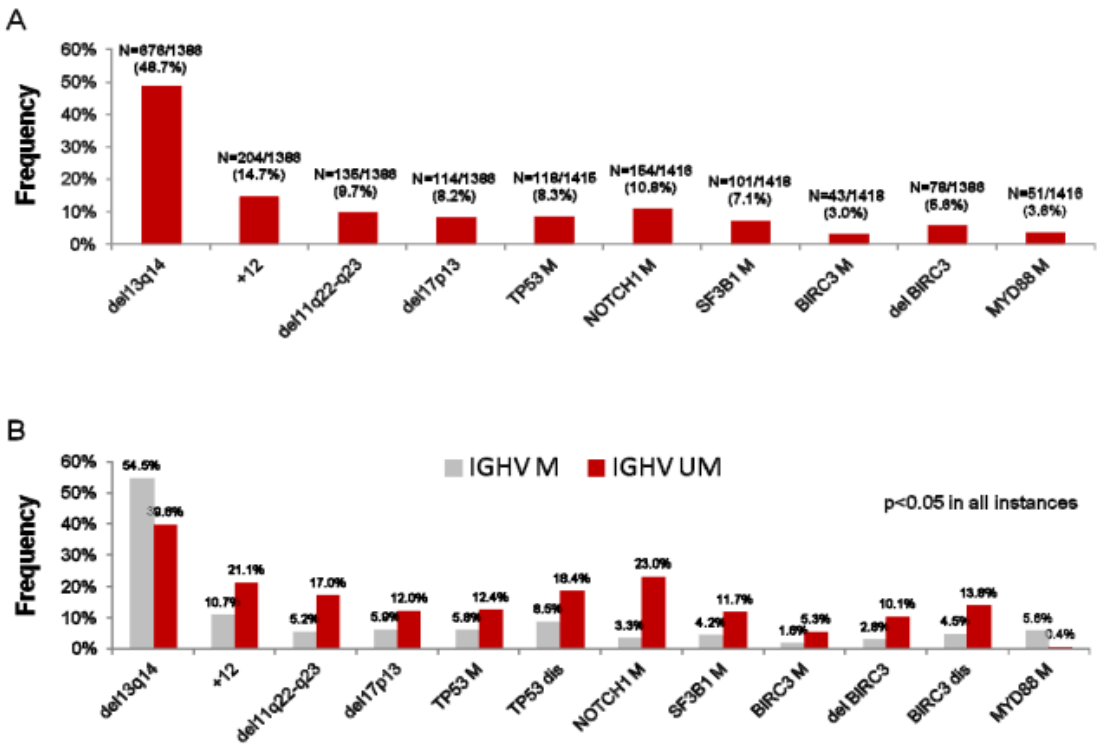

Figure 2 Genetic features in the study cohort. Panel A. Prevalence of the genetic features in the whole study cohort. Panel B. Prevalence of the genetic features in IGHV mutated (IGHV M) and IGHV unmutated (IGHV UM) cases. M, mutation; del, deletion; dis, disruption; p, p value from Fisher exact test corrected for multiple hypothesis testing [29]

2.4 Diagnosis and Prognosis Biomarkers

The diagnosis of CLL requires the presence of at least 5000 B-lymphocytes/μL in the peripheral blood for a minimum duration of 3 months. The clonality of these circulating B-lymphocytes must be verified through flow cytometry.[30] In CLL, cells co-express the surface antigen CD5 along with the B-cell antigens CD19, CD20, and CD23. Compared to normal B cells, CLL cells exhibit characteristically low levels of surface immunoglobulin, CD20, and CD79b. Each clone of leukemia cells is limited to expressing either kappa or lambda immunoglobulin light chains. It is important to note that CD5 expression can also be observed in other lymphoid malignancies, such as mantle cell lymphoma. A recent large-scale harmonization effort has confirmed that a panel consisting of CD19, CD5, CD20, CD23, kappa, and lambda is generally sufficient for establishing the diagnosis. In borderline cases, additional markers such as CD43, CD79b, CD81, CD200, CD10, or ROR1 may assist in refining the diagnosis.[31]

Mutations in specific genes have been reported to be associated with the etiology, progression, and treatment of CLL. Some of these mutations are known to affect the activity of relevant cellular signalling pathways. The B-cell surface receptor (BCR) encompasses the monomeric immunoglobulins IgM and IgD. The variable regions of these receptor molecules recognize foreign antigens, and Igα and Igβ are responsible for their intracellular signalling through phosphorylation of Src family kinases and subsequent activation of downstream effectors. Activation of the BCR signal transduction pathway is involved in the regulation of metabolism, survival, proliferation, differentiation, and migration of various cell types, including CLL cells.[32]

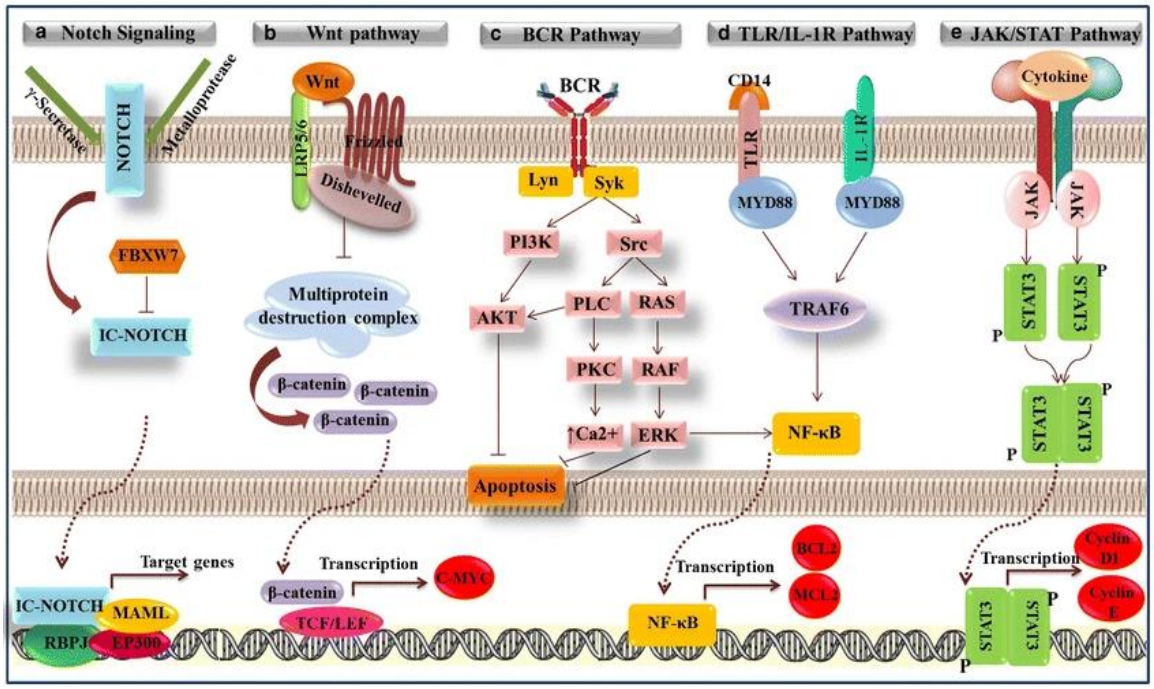

Figure 3 Signaling pathways in CLL cells. [32]

Approximately 60–65 % of CLL cases show mutated immunoglobulin heavy chain variable region *(*IGHV).[33]

IGHV gene mutations are markers for a favourable prognosis, with unmutated CLL (U-CLL) being clinically more aggressive than mutated CLL (M-CLL). The overall survival rate for U-CLL patients is 8–9 years, while for M-CLL patients it exceeds 24 years. Unmutated cases are associated with the expression of CD38 and ZAP-70, which are indicators of a poor prognosis.[34]

Gene mutation	Prevalence (%)	Association with	Functional role	Kind of mutation	Prognosis
IGHV	60–65 %	....	M-CLL typically has slower disease progression and longer survival.	....	Good
TP53	4–12 %	Del (17p13)	An essential role in apoptosis or cell cycle arrest after DNA damage.	inactivation	Poor
ATM	~12 %	Del (11q23)	Activates cell cycle checkpoints.	inactivation	Poor
			Induces apoptosis in response to DNA breaks.		
NOTCH1	~10 %	Trisomy 12	Regulates target genes, including *MYC*, *TP53* and molecules of the NFκB pathway.	activation	Poor
		UM- IGHV	Plays an important role in cell differentiation, proliferation, and apoptosis.		
SF3B1	5–10 %	Del (11q22-q23) ATM mutations, UM-IGHV, ZAP-70 expression	A core component of the spliceosome.	inactivation	Poor
BIRC3	~4 %	del(11q22-q23)	A negative regulator of the MAP3K14 (activator of NF-κB signaling).	inactivation	Poor
MYD88	3–5 %	M- IGHV and del (13q14)	Phosphorylated MYD88 resulting in activation the NF-κB pathway.	activation	....

Table 1 Most frequent somatic mutations in CLL.[32]

TP53 mutations

TP53 gene mutations have been observed in 4–12% of untreated CLL cases and in approximately 70% of cases with *17p* deletions. It has been well established that the presence of *TP53* mutations or deletions is associated with disease progression, a poor response to chemotherapy, and shorter survival in CLL patients.[35]

ATM mutations

ATM gene mutations have been reported in 12% of CLL patients and in about 30% of those with an *11q23* deletion. The absence of ATM protein expression leads to the loss of P53 function, chronic severe oxidative stress, rapid telomere shortening, loss of cell cycle checkpoint control, defective DNA repair, and consequently, disease progression.[36]

NOTCH1 mutations

NOTCH1 mutations in aggressive clinical phases of CLL have been linked to poor survival, resistance to treatment, disease progression, and an increased risk of progression to Richter's syndrome.[37] *NOTCH1* activating mutations have been observed in 10% of B-CLL cases. These mutations are associated with unmutated CLL (U-CLL) and *trisomy 12* and are linked to a poor prognosis.[38] The C.7544-7545del CT mutation is the most common *NOTCH1* mutation, occurring in approximately 80% of CLL cases.[39]

Other mutations reported in CLL include mutations in the *SF3B1*, *BIRC3*, *MYD88*, *KRAS*, *SMARCA2*, *NFKBIE* and *PRDK3* genes. [40] In chronic lymphocytic leukemia (CLL), the activation of proto-oncogenes via DNA promoter hypomethylation is a frequently observed phenomenon. Conversely, promoter hyper-methylation has been found to play a role in the silencing of tumour suppressor genes.[41]

Advancements in molecular understanding of the CLL genotype have significantly improved comprehension of the genetic underpinnings of chemotherapy refractoriness.[42]

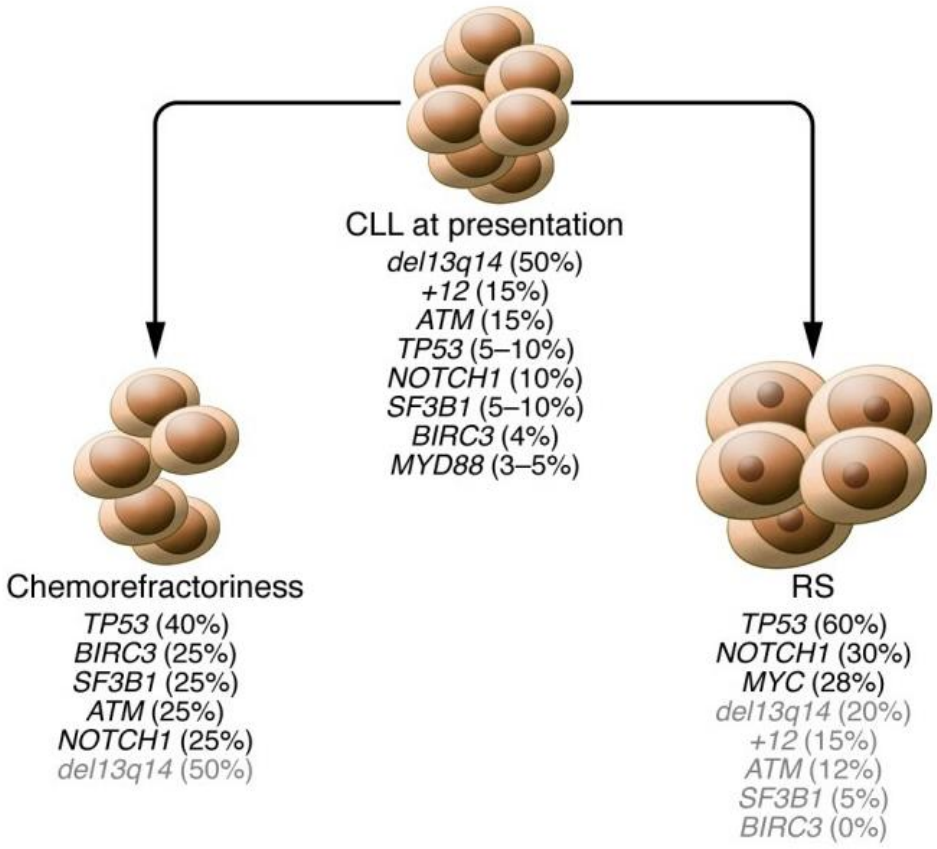

Figure 4 Genetic lesions of CLL at different phases of the disease. [42]

TP53 disruption, however, accounts for only 30%–40% of CLL cases refractory to fludarabine, the primary drug for treating this leukemia. In the remaining cases, the molecular basis of chemorefractoriness remains unknown. Recently identified genetic abnormalities in *NOTCH1*, *SF3B1*, and *BIRC3* have shed some light on this issue. These alterations are more prevalent in patients requiring therapy and resistant to fludarabine (*NOTCH1* mutations: 20%; *SF3B1* mutations: 17%; *BIRC3* disruption: 24%) compared to CLL at diagnosis or compared to cases requiring therapy but responsive to the drug.[43] Among chemorefractory patients, alterations in *NOTCH1*, *SF3B1*, and *BIRC3* seem to be predominantly mutually exclusive with *TP53* disruption, suggesting that multiple pathways can contribute to treatment failure.[44]

2.5 Light Chain Rearrangements

Currently, one of the most effective prognostic and therapeutic indicators in CLL, identified over 20 years ago, is the mutational status of IGHV. Survival analysis has demonstrated that patients with mutated BCR (M), characterized by somatic hypermutation (SHM) in the IGHV region resulting in less than 98% identity compared to the germline sequence, experience significantly longer survival compared to unmutated cases (UM). Additionally, multivariate analyses have confirmed that this mutational status serves as a robust independent poor prognostic marker in CLL.[45]

In the majority of CLL cases, the mutational statuses of heavy and light chains are consistent, except in the case of IGLV3-21. Among most IGLV3-21 rearrangements (24 out of 25), a similar LCDR3 (QVWDSSSDHPWV) is shared, accompanied by a single point mutation, specifically IGLV3-21^{R110}.[46]

Genome-wide methylation studies have revealed three epigenetic CLL subtypes, which are associated with IGHV mutational status and patient outcome. These subtypes include memory-like CLL (m-CLL; predominantly M-IGHV, with a favourable prognosis), intermediate CLL (i-CLL; characterized by a mixture of M- and U-IGHV, with an intermediate prognosis), and naïve-like CLL (n-CLL; predominantly U-IGHV, with a poor prognosis). The prognostic significance of this epigenetic classification has been validated in independent population-based studies and clinical trial cohorts.[47] These findings imply that *IGLV3-21^{R110}* may be overrepresented in i-CLL and could delineate a subgroup of patients with aggressive disease within this intermediate subtype. IGLV3-21^{R110} was detected in 38% of patients with i-CLL exhibiting poor clinical outcomes, resembling those seen in patients with n-CLL. Conversely, i-CLL lacking IGLV3-21^{R110} showed transcriptomic and clinical similarities to m-

CLL/M-IGHV tumors.[48] Compared to mutated-CLL without IGLV3-21^{R110}, CLL with IGLV3-21^{R110} showed a shorter overall survival (OS) and time to first treatment (TTFT) ($p = 0.05$, $p < 0.0001$, respectively), despite 75% (18/24) of these patients expressing mutated heavy chains. Overall, *IGLV3-21^{R110}* delineates a CLL subgroup with distinctive biological characteristics and an unfavourable prognosis, regardless of the *IGHV* mutational status, highlighting the significant value of the light chain.[46]

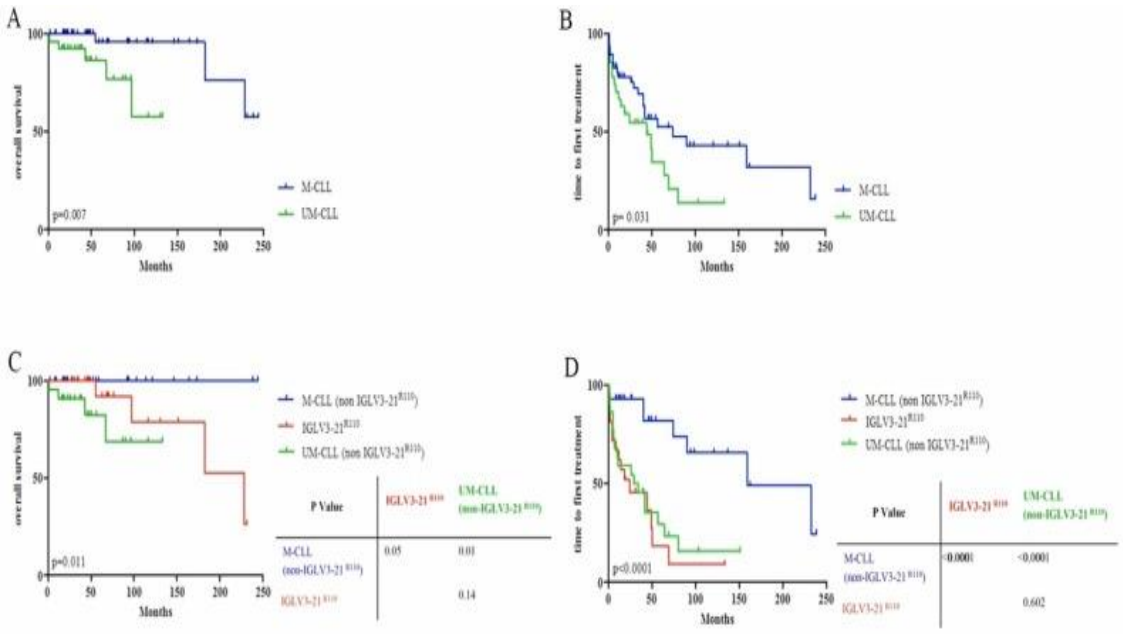

Figure 5 Survival analysis of the different groups. (A) Analysis of overall survival in M-CLL and UM-CLL patients ($p = 0.007$). (B) Analysis of time to first treatment in M-CLL and UM-CLL patients ($p = 0.031$). (C) Analysis of overall survival in in M-CLL non *IGLV3–21^{R110}*, UM-CLL non *IGLV3–21^{R110}* and *IGLV3–21^{R110}* ($p = 0.011$). (D) Analysis of time to first treatment in M-CLL non *IGLV3–21^{R110}*, UM-CLL non *IGLV3–21^{R110}* and *IGLV3–21^{R110}* ($p < 0.0001$). [49]

2.6 Richter Syndrome

Richter syndrome (RS) refers to the transformation of Chronic Lymphocytic Leukemia (CLL) into a high-grade aggressive lymphoma. It represents a grave complication of CLL, significantly worsening the disease prognosis due to both clinical and molecular alterations, exerting a profound impact on patient survival. RS occurs in approximately 2%–10% of CLL patients, with 95%–99% of cases characterized by a switch of the CLL clone to Diffuse Large B-cell Lymphoma (DLBCL-RT), while 1%–5% of cases undergo transformation into Hodgkin's variant (HV-RT).[50]

A subset of CLL cases progress to Richter syndrome (RS), marking the clinico-pathological transformation of CLL into DLBCL. The cumulative incidence of RS at 5 and 10 years after CLL diagnosis is approximately 10% and 15%, respectively.[51]

TP53 disruption (50%–60%), *NOTCH1* activation (30%), and *MYC* abnormalities (20%) have emerged as the most common recurrent genetic abnormalities in Richter syndrome (RS). These abnormalities are often acquired at the time of transformation, underscoring their pathogenetic significance in acquiring the DLBCL phenotype.[52]

Richter syndrome (RS) cases that are clonally unrelated to the paired CLL phase exhibit distinct biological and clinical characteristics compared to clonally related cases. As such, they should be regarded, and likely managed, as de novo DLBCL arising in the context of CLL. Rituximab-containing polychemotherapy serves as the cornerstone for induction treatment in patients with clonally related DLBCL transformation. Younger patients who respond to induction therapy should be considered for stem cell transplants to extend survival.[53]

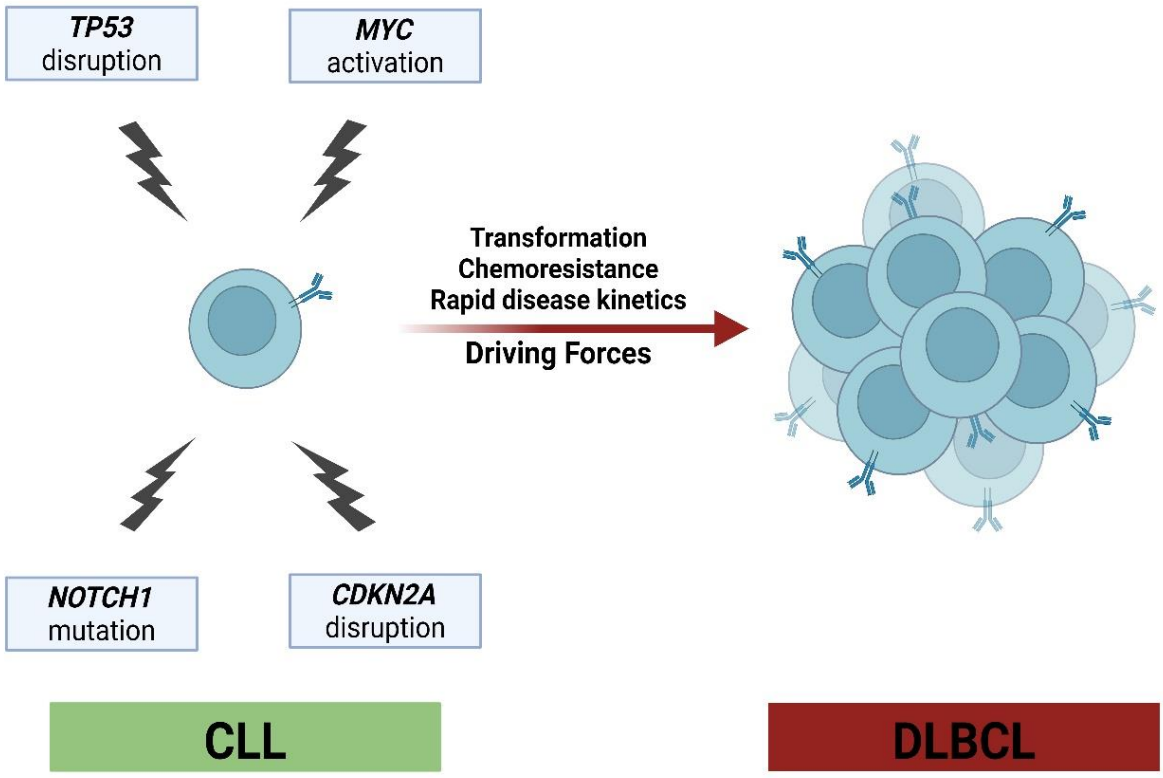

Figure 6 Genetic lesions contributing to chronic lymphocytic leukemia (CLL) transform into diffuse large B-cell lymphoma (DLBCL).

3. OBJECTIVES OF THE STUDY

Given the established importance of the IGHV mutational status and the role it plays in CLL prognosis, and due to the relative rarity of studies focusing on immunoglobulin light chains genes, more specifically their possible clinical association, this study aimed to:

1. Identify the light chain gene rearrangements repertoire in a large cohort of 573 CLL patients collected prospectively

2. Study these rearrangements and their possible correlation to overall survival and/or the risk of Richter transformation.

4. MATERIALS AND METHODS

4.1 Study population

The study included 573 CLL patients referring to our institution from which tumour genomic DNA was extracted from peripheral blood mononuclear cells. Samples were collected at the time of diagnosis and all cases were provided with the main clinical and biological data of the disease. Patients informed consents were provided in accordance with the local institutional review board requirements and the Declaration of Helsinki. The study was approved by the local ethical committee (study number CE 120/19).

4.2 Separation of mononuclear cells from peripheral blood

The separation of mononuclear cells was performed by density gradient centrifugation. Peripheral Blood (PB) samples were diluted 1:2 with physiological solution (NaCl 0.9%) and then centrifuged at the following program (1800 revolutions per minute (rpm) for 25 minutes) in a gradient differentiation Sigma Diagnostic Histopaque-1077 Cell Separation Medium (Sigma-Aldrich, St. Louis, MO, USA) solution that is used to separate blood cells into specific cell populations based on their densities. It separates white blood cells into granulocytes and mononuclear cells (monocytes and lymphocytes).

4.3 Genomic DNA extraction from PB mononuclear cells

DNA extraction was performed using the technique of "Salting out" which provides a rapid extraction of DNA by cell lysis, precipitation of proteins with salts and recovery of DNA with ethanol. The cells were suspended in a lysis buffer, consisting of (10 mM Tris-HCl, 10 mM NaCl and 2 mM EDTA), to which a surfactant with denaturing function, sodium dodecyl sulphate (SDS) was added to a final concentration of 0.5%, and Pronase E, at a final concentration of 400 µg/ml. The cell lysate was incubated for 16-20 hours at 37° C in a thermostatic stirrer. After the incubation period and after verifying the complete cell lysis, the proteins were eliminated by NaCl precipitation at a final concentration equal to 1.6 M and centrifugation at 3200 rpm for 20 minutes. The DNA was recovered in a jelly-like form of DNA which was seen after the addition of 100% ethanol. The DNA was extracted with Pasteur glass loops and subjected to three washes with 75% ethanol. Once the ethanol has evaporated, each sample was dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). This reagent has an alkaline pH and allows a correct dissolution and conservation of DNA, since the buffering

activity counteracts the acidification of the solution and the degradation of the DNA, while the calcium chelating activity inhibits the proliferation of microorganisms.

4.4 Quantification of DNA

DNA concentration was determined by the NanoDrop One Microvolume UV-Vis Spectrophotometer (Thermo Scientific, Massachusetts, USA). Nanodrop is a UV-visible spectrophotometer that uses small amounts of starting material exploiting a technology based on the surface tension that small volumes of DNA (1-2 µl) exert when they are placed between two neighbouring surfaces. The instrument allows to analyse samples with a concentration range between 2.0 and 27,500 ng/µL allowing the elimination of the pre-measurement dilution phase. The Blank solution was TE buffer. A drop of the sample placed on the podium is in direct contact with two optical fibers that allow a quick analysis in the order of a few seconds. The sample is read at a wavelength of 230 nm, 260 nm and 280 nm to evaluate, in addition to DNA concentration, the presence of contaminants such as ethanol or proteins. A ratio is calculated between the different readings. A ratio value between 1.9 and 2.1 is indicative of a good-quality DNA. The DNA concentration is calculated by the Lambert-Beer law, an empirical relation that correlates the quantity of light absorbed by a medium to its chemical nature and concentration.

4.5 Light chain gene rearrangements identification

In this cohort, light chain gene rearrangements were amplified with a polymerase chain reaction (PCR) method using GoTaq G2 hot start polymerase kit (Promega) and the reaction mixture was prepared as stated in (**Table 2**).

PCR Reaction			
Reagent	**Initial concentration**	**Final concentration in mix**	**Volume in mix**
GoTaq Flexi Buffer (Promega)	5X	1X	6 µl
Mgcl2	25 mM	1.5 mM	1.8 µl
dNTPs	10 mM	100 µM	0.3 µl
GoTaq (Promega)	5 U/µL	0.25 U/µL	0.05 µl
Forward Primer	5 µM	0.16 µM	1 µl
Reverse Primer	5 µM	0.16 µM	1 µl
DNA	50ng/µl	100 ng/µl	2 µl
H2O			17.85 µl
PCR cycling conditions			
Stage	**Temperature**	**Time**	**Cycles**

Stage 1.	95°C	3 min	
Stage 2. Step 1.	95°C	30 sec	Step 2. Temperature is lowered by 0.5 for each cycle
Step 2.	65°C	40 sec	
Step 3.	72°C	40 sec	(x 12 cycles)
Stage 3. Step 1.	95°C	30 sec	
Step 2.	59°C	40 sec	(x 35 cycles)
Step 3.	72°C	40 sec	
Stage 4.	72°C	10 min	

Table 2 PCR reaction and conditions

Leader primers were used for the amplification of the entire sequence of the rearranged light chain gene, thus enabling the true level of somatic hypermutation (SHM) to be determined. The sequences of leader primers are detailed in (**Table 3**) for kappa light chains and in (**Table 4**) for lambda light chains.

Primer Mix	Name	Sequence 5'→ 3'
Kappa light chain primers		
LVK1mix	LVK1a	CCTGCTCAGCTCCTGGGRCTCCTGC
	LVK1b	CCCGCTCAGCTCCTGGGGCTYCTGC
LVK2	LVK2	CTCCTGGGGCTGCTAATGCTCTGG
LVK3	LVK3	CTCTTCCTCCTGCTACTCTGGCTC
LVK4	LVK4	ATGGTGTTGCAGACCCAGGTCTT
LVK5-6 mix	LVK5	TTCCTCCTCCTTTGGATCTCTG
	LVK6	TTCTGCTSCTCTGGGTTCCAG
FR1VK1	FR1VK1	GACATCCRGWTGACCCAGTCTCCWTC
FR1VK2mix	FR1VK2a	CAGWCTCCACTCTCCCTGCCCGTC
	FR1VK2b	CAGACTCCACTCTCTCTGTCCGTC
	FR1VK2c	CAGACTCCACTCTCCTCACCTGTC
FR1VK3mix	FR1VK3a	TCTCCAGSCACCCTGTCTTTGTCTC
	FR1VK3b	TCTCCAGCCACCCTGTCTGTGTCTC
FR1VK4	FR1VK4	AGACTCCCTGGCTGTGTCTCTGGGC
FR1VK5-6mix	FR1VK5	CAGTCTCCAGCATTCATGTCAGCGA
	FR1K6a	TTTCAGTCTGTGACTCCAAAGGAG
	FR1K6b	TTCCTCTCTGTGACTCCAGGGGAG
JK outer mix	JK1-4 out	GTTTGATYTCCASYTTGGTCCC
	JK5 out	GTTTAATCTCCAGTCGTGTCCC

Table 3 Leader primers and Framework region (FR) 1 primers for kappa light chain genes

Primer Mix	Name	Sequence 5'→ 3'
Lambda light chain primers		
LVL1	LVL1	CTGGTCYCCTCTCYTCCTCAC
LVL2mix	LVL2a	CTCCTCAGCCTCCTCACTCAGG
	LVL2b	CTCCTCACYCTCCTCACTCAGG
LVL3amix	LVL3	CTCCTKCTCCCCCTCCTCACT
	LVL3e	CTCCTGCTCCCACTCCTCAAC
	LVL3h	GCCTGGACCGTTCTCCTCCTC
LVL3bmix	LVL3j	GCCTGGACCGCTCTCCTTCTG
	LVL3l	TCTGGCTCACTCTCCTCACTC
	LVL3r	CCTCGGCGTCCTTGCTTACTG
LVL4mix	LVL4a	CTCTTCCCTCTCCTCCTCCAC
	LVL4b	CCTCACCCTCCTCCTCCACTG
	LVL4c	CCTACTGCCCTTCATTTTCTC
LVL5-6mix	LVL5	TGYTCCTCTCTCACTGCACAG
	LVL6	CCTCCTCGCTCACTGCACAG
LVL7-10mix	LVL7	ATGGCCTGGACTCCTCTCTTT
	LVL8	ATGGCCTGGATGATGCTTCTC
	LVL9	TGCTCCTCACCCTCCTCAGTC
	LVL10	CTGGGCTCTGCTCCTCCTGAC
FR1VL1	FR1VL1	GGTCCTGGGCCCAGTCTGTG
FR1VL2kupp	FR1VL2kupp	CAGTCTGCCCTGACTCAGCCT
FR1VL3mix	FR1VL3a	CTCAGCCACCCTCAGTGTCCGT
	FR1VL3b	CTCAGCCACCCTCGGTGTCAGT
	FR1VL3c	CTCAGGACCCTGCTGTGTCTGT
FR1VL4-6mix	FR1VL4kupp	TTTCTTCTGAGCTGACTCAGGAC
	FR1VL5-9	CAGSCTGTGCTGACTCAGCC
	FR1VL6kupp	GAGTCTCCGGGGAAGACGGTA
FR1VL7-10mix	FR1VL7kupp	GTGGTGACTCAGGAGCCCTCAC
	FR1VL7-8	CTGTGGTGACYCAGGAGCC
	FR1VL9kupp	GCTGACTCAGCCACCTTCTGCA
	FR1VL10	CAGGCAGGGCWGACTCAGC
JL outer mix	JLa out	CTAGGACGGTSASCTTGGTCCC
	JL7 out	CGAGGACGGTCAGCTGGGTGCC
IGLV3-21^{R110} primers		
	IGLV3-21_F	GCTCTGTGACCTCCTATGTGC
	JC1 intron	GGCCCCATGGAGAAAGGTAA
	JC2/3 intron	CAAGCAAGGGTCTGAACAGG
	JC7 intron	AGGGGCAAAGATTCCAGACA

Table 4 Leader primers and Framework region (FR) 1 primers for lambda light chain genes including *IGLV3-21^{R110}* primers.

In cases of leader primers' inability to identify the rearranged gene, framework region 1 (FR1) primers were used as alternatives (forward primers) with the same reverse primers (JK, JL outer mix) mentioned above, and their sequences are listed in (**Table 3**). PCR products were then visualized by 2% agarose gel electrophoresis containing a DNA interlayer, in TBE 1X buffer (1M Tris, 1M boric acid and 20 mM EDTA, pH 8.0). The bands were visualized with

UV using the ChemiDoc MP imaging system instrument (Bio-Rad Laboratories S.R.L. Milan, Italy). The templates were then purified using MultiScreen PCRμ96 Filter Plate (Merck-Millipore, Damstadt, Germany), using a vacuum pump (Merck-Millipore), according to the company's protocol.

The purified DNA was subjected to bi-directional Sanger sequencing. The sequence reaction was performed using the BrilliantDye Terminator v1.1 (Nimagen, Gelderland, Netherlands). The Sanger reaction mixture can be seen in detail in (**Table 5**). The Sanger reaction protocol was set on the Thermal Cycler Veriti 96 well (Applied Biosystem, Forest City, USA) as shown in (**Table 5**).

Sanger Sequencing Reaction			
Reagent	**Initial concentration**	**Final concentration in mix**	**Volume in mix**
BrilliantDye Buffer v1.1 (Nimagen)	5X	0.4X	1.6 μL
BrilliantDye v1.1 RR Premix	2.5X	0.1X	0.8 μL
Forward Primer	5 μM	0.175 μM	0.7 μL
DNA	PCR product		1 μL
H2O			15.9 μL
Sanger cycling conditions			
Stage	**Temperature**	**Time**	**Cycles**
Stage 1.	98°C	3 min	
Stage 2.	96°C 60°C	10 sec 4 min	x 27 cycles

Table 5 Sanger sequencing reaction

Sequence reaction products were purified with the Montage SEQ 96 Sequencing reaction cleanup kit (Merck-Millipore, Burlington, USA). A volume of 5 μl of sequenced DNA was added to 10μl of formamide, and together they were denatured and subjected to capillary electrophoresis with the automated ABI PRISM 3130XL Genetic Analyzer (Applied Biosystem, Forest City, USA) and SeqStudio Genetic Analyzer for Human Identification (Applied Biosystem, Forest City, USA). The migration of DNA fragments was detected by fluorescence emission, and the data were analyzed and displayed as a nucleotide sequence in a chromatogram as in (**Figure 7**).

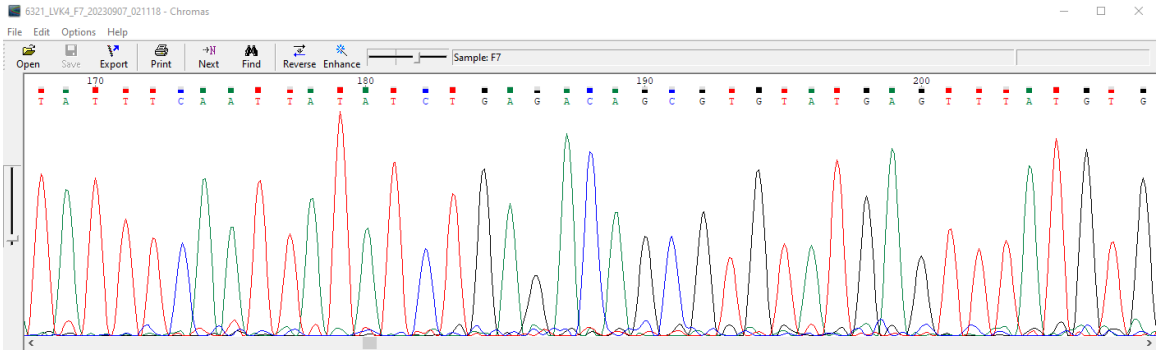

Figure 7 Sequence chromatogram illustrated by Chromas

4.6 Statistical analysis

The primary endpoints were: i) overall survival (OS), defined as the time between CLL diagnosis and the date of death from any causes (event), or last follow up (censoring); and ii) the risk of Richter Transformation defined as the time between CLL diagnosis and Richter transformation (event) or death/last follow up (censoring). Survival analysis was performed using the Kaplan-Meier method and compared between strata using the Log-rank test. A false discovery rate approach was used to account for multiple testing, and adjusted p-values were calculated using Bonferroni and Benjamini–Hochberg correction. Cox regression test was used in the univariate and the multivariate analysis. Pearson's correlation was used to compare the prevalence of rearrangements in different molecular and clinical subgroups of patients. The tests were performed using the Statistical Package for the Social Sciences (SPSS) software v.24.0 (Chicago, IL, USA) and R Studio Version 1.2.1335 2009-2019, Inc. Statistical significance was defined as p-value <0.05.

5. RESULTS

5.1 Patient characteristics

A total of 573 patients were analyzed in this cohort and 530 harbored at least one productive rearrangement. Forty-two (7.9%) patients presented double productive rearrangements. Patient characteristics were consistent with a real-world cohort of unselected newly diagnosed CLL, as the median age at diagnosis was 70 years, 58.5% were male, the median lymphocyte count was 9900/ul, 36.8% had unmutated (UM)-IGHV, 46.5% had *del13q*, and 11.8% harbored *TP53* disruption (i.e. *TP53* mutation and/or 17p deletion). All patient characteristics are recorded in (Table 6).

Characteristics	Values
Median age at diagnosis (range)	70 (36 - 92)
Sex	
Male	310 (58.5%)
Female	219 (41.3%)
Median lymphocyte count	9900/µl
Median haemoglobin level	13.7 g/dl
Median platelet count	$200x10^3$/µl
Light chain	
Kappa	305 (57.6%)
Lambda	167 (31.5%)
Undefined	58 (10.9%)
IGHV	
Mutated	327 (63.1%)
Unmutated	191 (36.8%)
Binet staging system	
Binet A	414 (81.5%)
Binet B	53 (10.4%)
Binet C	41 (8.1%)
***TP53* disrupted**	62 (11.8%)
del17p	43 (8.7%)
TP53	50 (9.6%)
Trisomy 12	92 (18.7%)
NOTCH1	35 (11.4%)
del13q	229 (46.5%)
del11q	37 (7.5%)

MYD88	11 (3.6%)
SF3B1	21 (6.8%)
BIRC3	5 (1.6%)
Stereotyped cases	
Subset 1	11 (13.8%)
Subset 2	12 (15.0%)
Subset 4	6 (7.5%)
Subset 8	6 (7.5%)

Table 6 Patient characteristics

According to the Binet staging system, 414 patients (81.5%) were classified as Binet A, 53 patients (10.4%) as Binet B and 41 patients (8.1%) as Binet C. CLL stereotyped subsets were identified in 80/530 patients (15.1%), namely 11/80 patients (13.8%) belonged to subset #1, 12/80 (15%) to subset #2, 6/80 (7.5%) to subset #4 and 6/80 (7.5%) to subset #8. **Table 7** includes all CLL subsets identified in this study.

CLL Subset	No. (%)	CLL Subset	No. (%)
Subset#2	12 (15)	**Subset#66**	1 (1.3)
Subset#1	11 (13)	**Subset#7E**	1 (1.3)
Subset#8	6 (7.5)	**Subset#6**	1 (1.3)
Subset#4	6 (7.5)	**Subset#216**	1 (1.3)
Subset#59	3 (3.8)	**Subset#77**	1 (1.3)
Subset#68	2 (2.5)	**Subset#218**	1 (1.3)
Subset#204	3 (2.5)	**Subset#95**	1 (1.3)
Subset#10	4 (2.5)	**Subset#223**	1 (1.3)
Subset#16	5 (2.5)	**Subset#11**	1 (1.3)
Subset#7H	6 (2.5)	**Subset#246**	1 (1.3)
Subset#215	1 (1.3)	**Subset#5**	1 (1.3)
Subset#247	1 (1.3)	**Subset#248**	1 (1.3)
Subset#221	1 (1.3)	**Subset#23**	1 (1.3)
Subset#31	1 (1.3)	**Subset#250**	1 (1.3)
Subset#251	1 (1.3)	**Subset#206**	1 (1.3)
Subset#41	1 (1.3)	**Subset#7B**	1 (1.3)
Subset#217	1 (1.3)	**Subset#209**	1 (1.3)
Subset#9	1 (1.3)	**Subset#29**	1 (1.3)
Subset#226	1 (1.3)	**Subset#210**	1 (1.3)
Subset#63	1 (1.3)	**Subset#201**	1 (1.3)
Subset#249	1 (1.3)	**Subset#203**	1 (1.3)

Table 7 The frequency of CLL subsets in the present cohort

After a median follow-up of 11.6 years, the median overall survival for CLL patients was 11.7 years The 10-year overall survival (OS) for entire cohort was 61.0%. As expected, the well-known CLL prognostic biomarkers associated with poor outcomes. The 10-year overall survival (OS) for UM-IGHV patients was 48.1% compared to 67.9% for mutated (M)-IGHV patients (p<0.0001) (**Figure 8A**). Similarly, the 10-year OS for *TP53* disrupted cases was 36.7% compared to 64.2% for wild type *TP53* patients (p<0.0001) (**Figure 8B**).

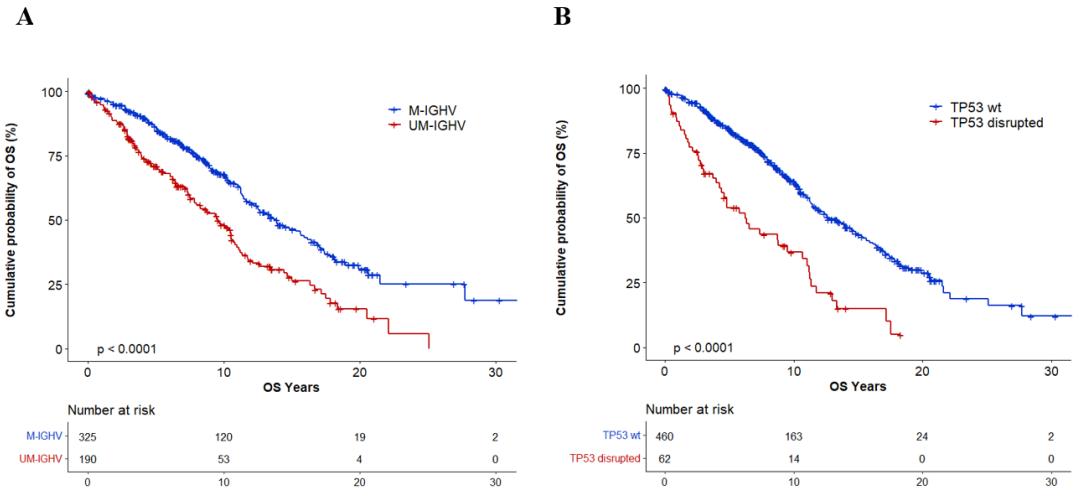

Figure 8 Kaplan-Meier curve in terms of overall survival of the well-known prognostic factors of CLL. (A) IGHV mutational status. The blue line represents patients with mutated heavy chains (M-IGHV), while the red line includes patients with unmutated heavy chains (UM-IGHV). (B) *TP53* disrupted (*TP53* mutation or del17p). Patients with disruptive aberrations of TP53 are indicated in red, while patients which harbour a wild type (wt)-*TP53* are indicated in blue. The log-rank statistics p values are indicated adjacent to the curves. The number of patients at risk in each group is indicated in the risk table under each graph.

5.2 Light chain rearrangements

According to flow-cytometry analysis, kappa light chains were expressed in 306 (57.7%) patients and lambda light chains in 166 (31.3%), while in 58 (10.9%) cases the expression of light chains was too low on the surface of the B cell that impaired determination of the CLL light chain restriction.

Overall, a total of 51 different productive IGKV (N=25) and IGLV (N=26) genes were identified. Leader primers were used to identify rearrangements in 404 patients (76.2%), while, in cases without productive rearrangements with the leader primers, the FR1 primers were used (N=126, 23.8%). For kappa light chains, the most frequently rearranged gene was IGKV4-1 in

84 patients, that comprised 20.5% of the total kappa rearrangements, followed by IGKV3(D)-20 in 60 patients (14.8%) and IGKV1-39 in 42 patients (10.4%) (**Figure 9A**). For lambda light chains, the most frequently rearranged gene was IGLV3-21 in 32 patients, that comprised 19% of the total lambda rearrangements (N=168), followed by IGLV2-14 in 27 patients (16.1%), and IGLV3-25 in 12 patients (7.1%). Among the 32 patients who expressed the lambda rearrangement IGLV3-21, nineteen of them (59.4%) harboured the IGLV3-21*04^{R110} configuration (**Figure 9B**).

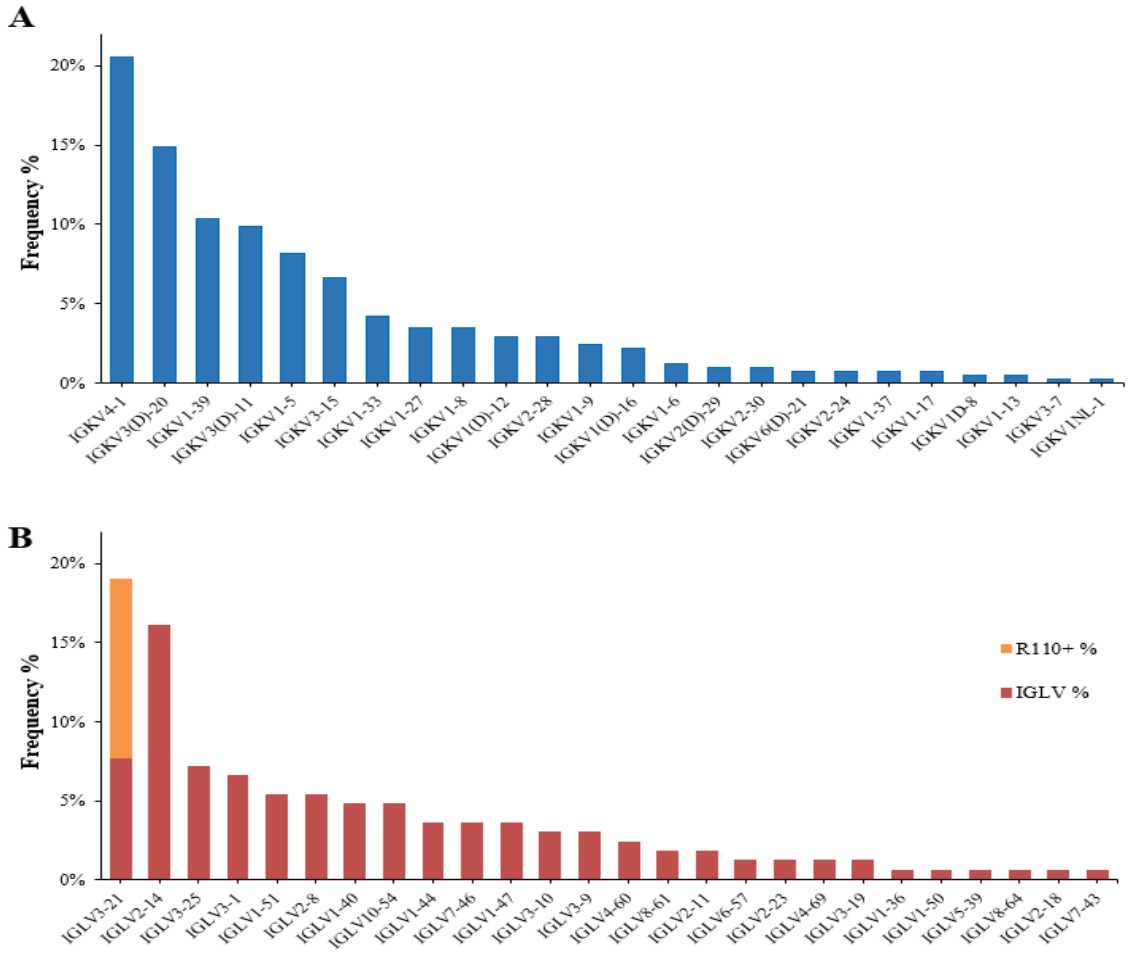

Figure 9 Frequency of rearranged light chain genes found in the cohort and their association to CLL molecular features. (A) Bar chart depicting the frequency of kappa light chain rearrangements (B) Bar chart depicting the frequency of lambda light chain rearrangements. IGLV3-21^{R110} rearrangements are colour coded in orange.

Interestingly, 57/58 cases whose light chain expression was undefined through flow-cytometry analysis, displayed kappa rearrangements through sequencing. However, they did not present similarities with regards to rearrangements or any other characteristics. One unique case presented with double rearrangements that included the two most frequent genes, *IGKV4-1*, and *IGLV3-21*.

By evaluating the association of light chain rearrangements with baseline CLL molecular features, among lambda rearrangements, both *IGLV3-21* and *IGLV3-21*^{R110} genes significantly associated with the usage of *IGHV3-21* genes, with IGHV borderline mutational status (defined as a variable region homology to the germline sequence of 97-97.99%), with subset#2, and with *SF3B1* mutations (all p<0.001). Among kappa rearrangements, *IGKV4-1* gene showed neither positive nor negative association with any of the CLL molecular features, whereas *IGKV1-39* gene significantly associated with *NOTCH1* mutations (p<0.0001), trisomy 12 (p=0.0002), UM-IGHV (p=0.0001), subset #1 (p<0.0001), and subset #8 (p<0.0001). Additionally, IGKV3-11 gene associated with IGHV4-34 genes (p=0.0086) and subset #4 (p<0.0001) (**Figure 10**).

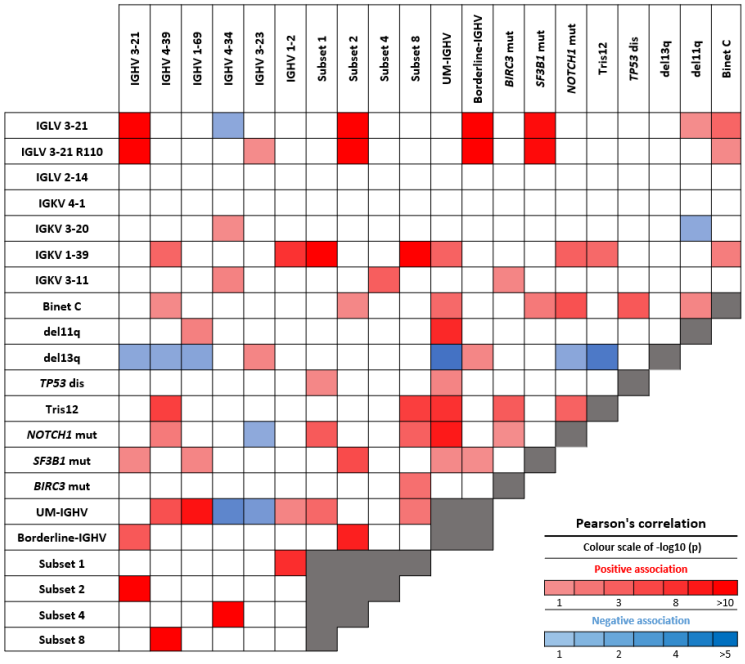

Figure 10 Correlation map comparing the most recurrent light chain rearrangements to the molecular features of CLL. The red color scale of -log$_{10}$(p) points to a co-occurrence between two variables. The blue color scale of -log$_{10}$(p) points to exclusivity between two variables. The intensity of the color corresponds to the strength of the correlation.

5.3 Association of light chain rearrangements with overall survival and Richter transformation

The 51 different kappa and lambda rearrangements identified in this cohort were studied for their association with overall survival (OS), and with the risk of Richter transformation. Bonferroni adjusted univariate analysis for OS (**Table 8**) revealed that only patients with *IGKV1-39* rearrangement (N=41) significantly associated with shorter OS (p<0.0001) (**Table 9**). Patients expressing the IGKV1-39 gene demonstrated a 10-year OS of 38.8% compared to 62.7% for patients harbouring other rearrangements (**Figure 11A**). Importantly, in multivariate analysis, IGKV1-39 gene (HR 1.77, 95% CI 1.17-2.69, p=0.007) independently associated with shorter OS when adjusted for UM-IGHV and *TP53* disruption (**Figure 11B**).

Rearrangement	HR	95% CI		P
K117	0.05	1.09E-04	18.61	0.3131
K127	2.23	1.14	4.36	0.0189
K133	1.37	0.73	2.59	0.3276
K137	0.68	0.10	4.87	0.7025
K139	2.21	1.48	3.30	0.0001
K15	0.63	0.37	1.07	0.0862
K16	0.42	0.06	2.98	0.3834
K19	1.00	0.41	2.42	0.9984
K1(D)12	1.09	0.92	4.38	0.9060
K1D13	0.05	3.16E-05	76.86	0.4222
K1(D)16	0.68	0.10	4.87	0.7030
K1(D)8	1.45	0.20	10.33	0.7130
K1NL1	0.05	2.41E-06	1.02E+03	0.5528
K224	3.26	0.45	23.48	0.2399
K228	1.54	0.68	3.46	0.2990
K230	0.34	0.05	2.42	0.2803
K2(D)29	6.50	1.60	26.47	0.0089
K315	0.87	0.50	1.53	0.6316
K37	3.81	0.53	27.29	0.1831
K3(D)11	1.00	0.64	1.56	1.0000
K3D15	0.05	5.50E-10	4.47E+06	0.7479
K3(D)20	1.86	0.08	3.83	0.5350
K41	0.60	0.37	0.99	0.0443
K621	4.47	1.11	18.09	0.0357
K6D21	51.77	6.63	404.40	0.00
L1054	0.45	0.11	1.80	0.2590
L136	4.49	0.63	32.22	0.1351
L140	1.03	0.42	2.49	0.9527

L144	0.98	0.24	3.93	0.9727
L147	3.37	1.25	9.09	0.0167
L150	43.01	5.59	330.75	0.0003
L151	3.07	1.26	7.50	0.0139
L211	1.65	0.53	5.16	0.3910
L214	0.46	0.20	1.02	0.0573
L218	21.13	2.86	156.20	0.0028
L223	0.05	9.17E-10	2.67E+06	0.7407
L28	1.87	0.83	4.20	0.1316
L31	1.45	0.60	3.52	0.4148
L310	0.86	0.21	3.47	0.8332
L319	0.05	1.59E-05	152.86	0.4632
L321	0.86	0.47	1.58	0.6301
L325	0.40	0.15	1.08	0.0701
L39	1.26	0.31	5.09	0.7428
L460	3.06	0.97	9.60	0.0556
L469	0.75	0.10	5.32	0.7696
L539	0.05	1.58E-07	1.55E+04	0.6416
L657	1.01	0.14	7.19	0.9950
L743	2.71	0.38	19.38	0.3210
L746	0.84	0.21	3.39	0.8098
L861	2.39	0.33	17.15	0.3853
L864	2.74	0.38	19.60	0.3153

Table 8 Univariate analysis of light chain rearrangements associated with overall survival

Rearrangement	HR	95% CI	P	P*
IGKV1-39	2.21	1.48-3.3	0.0001	0.006
IGKV1-27	2.23	1.14-4.36	0.019	0.95
IGKV2(D)-29	6.50	1.60-26.47	0.044	>0.999
IGKV4-1	0.60	0.37-0.98	0.017	0.85
IGLV1-47	3.37	1.25-9.09	0.014	0.7
IGLV1-51	3.07	1.26-7.50	0.014	0.71

Table 9 Univariate analysis of light chain rearrangements associated significantly with overall survival. P, P-value; P*, Bonferroni corrected p value; OS, overall survival; CI, confidence interval; HR, hazard ratio.

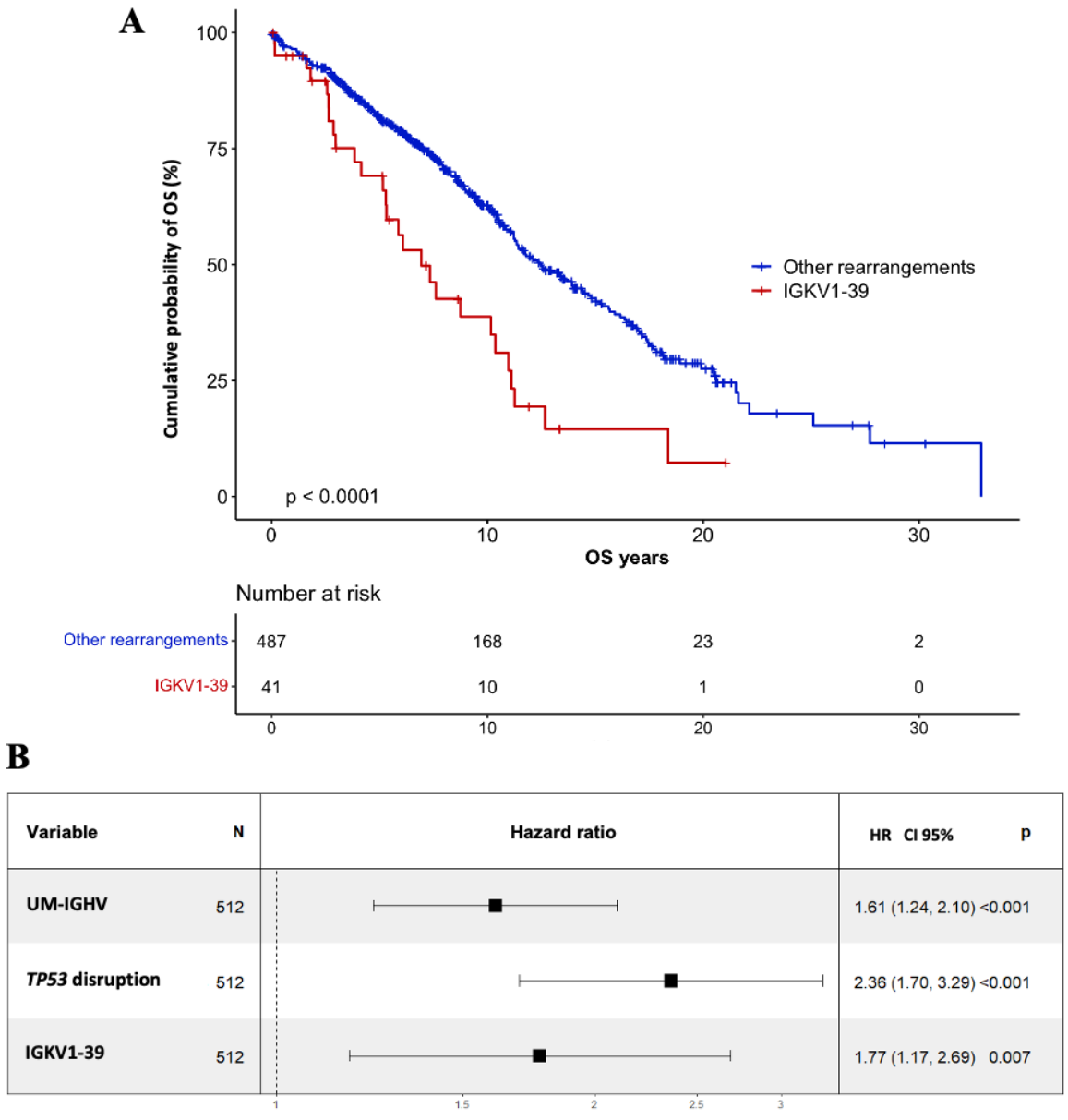

Figure 11 Clinical impact of different light chain rearrangements on overall survival. (A) Kaplan-Meier curve of overall survival for patients with IGKV1-39 rearrangement (red line) compared to other rearrangements (blue line). (B) Forest plot multivariate analysis in terms of overall survival of IGKV1-39 rearrangement and UM-IGHV and *TP53* disruption.

In addition, light chain rearrangements were correlated with the risk of Richter transformation and after univariate analysis (**Table 10**) adjusted for multiple comparisons the IGKV1-39 (N=41), IGKV6-21 (N=2), IGLV1-36 (N=1) and IGLV8-61 (N=3) rearrangements associated with a higher risk of Richter transformation (**Table 11**). By combining cases with the above-mentioned high-risk light chain rearrangements, a total of 47 CLL patients have been

obtained. The 10-year risk of Richter transformation in patients harbouring these high-risk rearrangements was 26.9% compared to 3.3% for the remaining cases (p<0.0001) (**Figure 12A**).

Rearrangement	HR	95% CI		P
K117	0.05	6E-16	4E+12	0.8544
K127	4.14	0.97	17.76	0.0555
K133	2.90	0.68	12.41	0.1514
K137	0.05	2E-11	1E+08	0.7839
K139	4.73	1.74	12.86	0.0024
K15	0.04	3E-05	61.47	0.3988
K16	5.34	0.7	39.73	0.1018
K19	2.32	0.3	17.22	0.4122
K1(D)12	0.00	0.64	1.56	0.9998
K1D13	0.05	1E-12	2E+09	0.8106
K1(D)16	0.00	3E-13	9E+09	0.8201
K1(D)8	0.00	4E-15	6E+11	0.8451
K1NL1	0.05	3E-17	8E+13	0.8663
K224	0.05	2E-19	1E+16	0.8827
K228	0.05	7E-08	3E+04	0.6579
K230	0.05	5E-10	5E+06	0.7481
K2(D)29	0.00	3E-19	1E+16	0.8824
K315	0.90	0.1	6.7	0.9157
K37	0.05	8E-22	3E+18	0.8973
K3(D)11	0.00	3E-14	9E+10	0.8346
K3D15	0.05	1E-22	3E+19	0.9018
K3(D)20	0.00	9E-13	3E+09	0.8117
K41	0.04	1E-04	14.9	0.2898
K621	30.45	3.8	241.3	0.0012
K6D21	0.00	1E-12	2E+09	0.9998
L1054	0.05	2E-08	1E+05	0.6851
L136	31.51	4.1	239.9	0.0009
L140	3.55	0.5	26.4	0.2160
L144	0.05	1E-09	2E+06	0.7339
L147	0.05	3E-11	9E+07	0.7820
L150	0.05	3E-45	1E+42	0.9529
L151	0.05	8E-10	3E+06	0.7407
L211	0.05	5E-12	5E+08	0.7973
L214	0.88	0.1	6.5	0.9001
L218	0.05	4E-37	6E+33	0.9420
L223	0.05	2E-20	1E+17	0.8889
L28	0.05	9E-09	3E+05	0.7017
L31	2.53	0.3	18.8	0.3645

L310	0.05	5E-11	5E+07	0.7755
L319	0.05	2E-13	1E+10	0.8234
L321	0.05	9E-06	223.8	0.4772
L325	1.56	0.2	11.6	0.6636
L39	0.05	2E-10	1E+07	0.7581
L460	0.05	2E-12	1E+09	0.8034
L469	0.05	1E-13	2E+10	0.8262
L539	0.05	5E-18	5E+14	0.8732
L657	0.05	2E-13	1E+10	0.8213
L743	0.05	9E-19	3E+15	0.8786
L746	0.05	9E-10	3E+06	0.7396
L861	17.19	2.3	130.5	0.0060
L864	0.05	9E-19	3E+15	0.8786

Table 10 Univariate analysis of light chain rearrangements associated with time-to-Richter transformation

Rearrangement	HR	95% CI	P	P*
IGKV1-39	4.73	1.74-12.86	0.002	0.0401
IGKV6-21	30.39	3.84-240.85	0.001	0.0306
IGLV1-36	31.51	4.14-239.94	0.0001	0.0306
IGLV8-61	17.19	2.26-130.49	0.006	0.049

Table 11 Univariate analysis of light chain rearrangements associated significantly with the risk of Richter transformation, P-value; P*, Benjamini-Hochberg correction; CI, confidence interval; HR, hazard ratio.

The high-risk rearrangements were subjected to a multivariate analysis which confirmed their independent association to higher risk of Richter transformation (HR=4.89, 95% CI 1.26-18.94, p=0.02) when adjusted to two main CLL features that predispose to Richter transformation, namely *NOTCH1* mutations and subset #8 (**Figure 12B**).

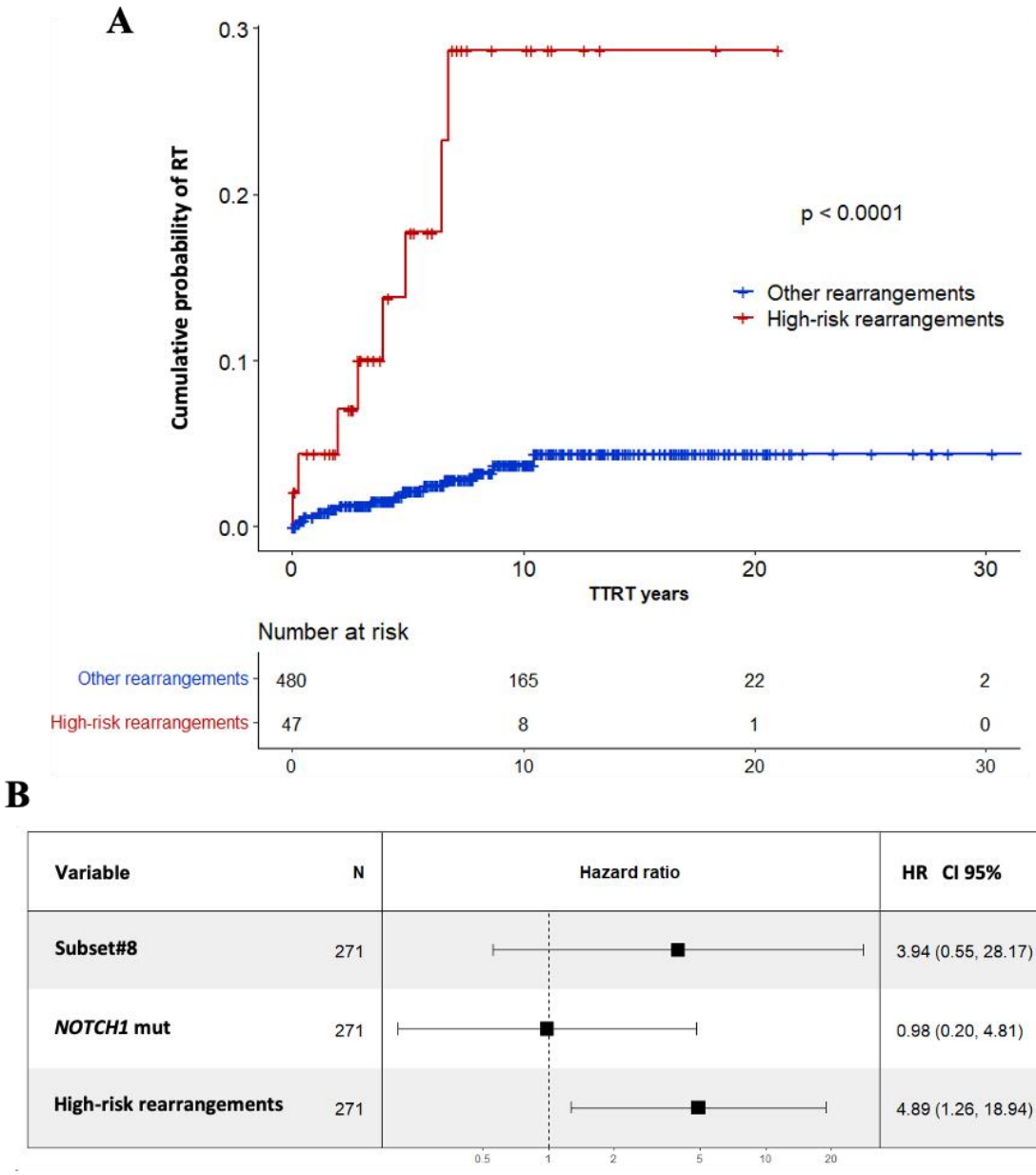

Figure 12 Clinical impact of different light chain rearrangements on the risk of Richter transformation. (A) Kaplan-Meier curve of the risk of Richter transformation (RT) for "high-risk rearrangements" (red line) compared to other rearrangements (blue line). (B) Forest plot multivariate analysis in terms of risk of Richter transformation.

6. DISCUSSION

Chronic lymphocytic leukemia (CLL) is a heterogeneous disease characterized by diverse clinical presentations and molecular features that significantly influence patient outcomes.

In this cohort, the identification of IGLV3-21 and its variant IGLV3-21^{R110} as frequently rearranged genes in our cohort merits attention. These rearrangements showed significant associations with specific CLL molecular subtypes, including IGHV3-21 usage, *SF3B1* mutations, and subsets #2 and #8. This highlights the intricate interplay between light chain rearrangements and underlying CLL biology, suggesting potential implications for disease pathogenesis and therapeutic strategies tailored to molecular subtypes.

In addition, we investigated the impact of light chain gene rearrangements on clinical outcomes, including overall survival and the risk of Richter transformation. The analysis revealed a notable diversity in light chain rearrangements, with 51 distinct kappa and lambda light chain rearrangements identified.

The median follow-up of the present cohort of more than 11 years may represent a suitable platform of evaluating overall survival and Richter transformation end points. Individual light chain genes were studied for their possible association to multiple clinical outcomes as this was the main focus of this study.

We observed that patients harbouring the IGKV1-39 rearrangement exhibited significantly shorter OS compared to those with any other rearrangement. This finding was robust in both univariate (even when corrected for Bonferroni) and multivariate analyses, where IGKV1-39 remained independently associated with poorer OS after adjusting for key prognostic factors such as unmutated IGHV status and *TP53* disruption. The mechanism underlying the adverse prognosis associated with IGKV1-39 calls for further investigation, as its identification could potentially guide risk stratification and therapeutic decisions in clinical practice.

Moreover, our study unveiled a novel insight into the association between specific light chain rearrangements and the risk of Richter. Notably, IGKV1-39, IGKV6-21, IGLV1-36, and IGLV8-61 rearrangements were identified as high-risk factors predisposing CLL patients to Richter transformation. Patients carrying these rearrangements had a substantially higher 10-

year risk of RT compared to those without these rearrangements, underscoring their potential utility as prognostic biomarkers for aggressive disease behaviour.

Furthermore, the limitations of our study include its reliance on single-centre data, which may limit the generalizability of findings across broader CLL populations. Future prospective studies with larger, multicenter cohorts are warranted to validate our findings and explore additional molecular mechanisms linking light chain rearrangements to CLL pathobiology and clinical outcomes. Additionally, integrating high-throughput sequencing technologies could enhance our understanding of clonal evolution and heterogeneity within CLL, offering deeper insights into the prognostic significance of light chain gene rearrangements. Finally, and due to the long follow-up of our cohort, most of patients were treated with chemo-immunotherapy and the prognostic impact in terms of overall survival of immunoglobulin light chain gene rearrangements in patients treated BTKi and BCL2 could not be established.

In conclusion, our study provides comprehensive insights into the role of light chain gene rearrangements as prognostic markers in CLL. By elucidating their associations with OS, Richter transformation, and molecular subtypes, we contribute to the growing body of evidence supporting personalized medicine approaches in CLL management. These findings underscore the potential clinical utility of light chain rearrangements in guiding risk stratification and therapeutic decisions, thereby enhancing patient outcomes in this complex hematological malignancy.

7. REFERENCES

1. Rai, K.R., and Jain, P. (2016). Chronic lymphocytic leukemia (CLL)—Then and now. Am J Hematol *91*, 330–340. https://doi.org/10.1002/AJH.24282.

2. Thandra, K.C., Barsouk, A., Saginala, K., Padala, S.A., Barsouk, A., and Rawla, P. (2021). Epidemiology of Non-Hodgkin's Lymphoma. Medical Sciences *9*. https://doi.org/10.3390/MEDSCI9010005.

3. Mukkamalla, S.K.R., Taneja, A., Malipeddi, D., and Master, S.R. (2023). Chronic Lymphocytic Leukemia. 10–13.

4. Chiorazzi, N., Rai, K.R., and Ferrarini, M. (2005). Chronic lymphocytic leukemia. N Engl J Med *352*, 804–815. https://doi.org/10.1056/NEJMRA041720.

5. Hamblin, T.J., Davis, Z., Gardiner, A., Oscier, D.G., and Stevenson, F.K. (1999). Unmutated Ig VH Genes Are Associated With a More Aggressive Form of Chronic Lymphocytic Leukemia. Blood *94*, 1848–1854. https://doi.org/10.1182/BLOOD.V94.6.1848.

6. Rosenwald, A., Alizadeh, A.A., Widhopf, G., Simon, R., Davis, R.E., Yu, X., Yang, L., Pickeral, O.K., Rassenti, L.Z., Powell, J., et al. (2001). Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. J Exp Med *194*, 1639–1647. https://doi.org/10.1084/JEM.194.11.1639.

7. Hervé, M., Xu, K., Ng, Y.S., Wardemann, H., Albesiano, E., Messmer, B.T., Chiorazzi, N., and Meffre, E. (2005). Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive B cell precursors despite expressing different antibody reactivity. J Clin Invest *115*, 1636–1643. https://doi.org/10.1172/JCI24387.

8. Gobessi, S., Laurenti, L., Longo, P.G., Carsetti, L., Berno, V., Sica, S., Leone, G., and Efremov, D.G. (2009). Inhibition of constitutive and BCR-induced Syk activation downregulates Mcl-1 and induces apoptosis in chronic lymphocytic leukemia B cells. Leukemia *23*, 686–697. https://doi.org/10.1038/LEU.2008.346.

9. Kikushige, Y. (2020). Pathogenesis of chronic lymphocytic leukemia and the development of novel therapeutic strategies. J Clin Exp Hematop *60*, 146. https://doi.org/10.3960/JSLRT.20036.

10. Marti, G.E., Rawstron, A.C., Ghia, P., Hillmen, P., Houlston, R.S., Kay, N., Schleinitz, T.A., and Caporaso, N. (2005). Diagnostic criteria for monoclonal B-cell lymphocytosis. Br J Haematol *130*, 325–332. https://doi.org/10.1111/J.1365-2141.2005.05550.X.

11. Rawstron, A.C. (2009). Monoclonal B-cell lymphocytosis. Hematology Am Soc Hematol Educ Program, 430–439. https://doi.org/10.1182/ASHEDUCATION-2009.1.430.

12. Gaidano, G., Foà, R., and Dalla-Favera, R. (2012). Molecular pathogenesis of chronic lymphocytic leukemia. J Clin Invest *122*, 3432–3438. https://doi.org/10.1172/JCI64101.

13. Döhner, H., Stilgenbauer, S., Benner, A., Leupolt, E., Kröber, A., Bullinger, L., Döhner, K., Bentz, M., and Lichter, P. (2000). Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med *343*, 1910–1916. https://doi.org/10.1056/NEJM200012283432602.

14. Ouillette, P., Collins, R., Shakhan, S., Li, J., Li, C., Shedden, K., and Malek, S.N. (2011). The prognostic significance of various 13q14 deletions in chronic lymphocytic leukemia. Clin Cancer Res *17*, 6778–6790. https://doi.org/10.1158/1078-0432.CCR-11-0785.

15. Palamarchuk, A., Efanov, A., Nazaryan, N., Santanam, U., Alder, H., Rassenti, L., Kipps, T., Croce, C.M., and Pekarsky, Y. (2010). 13q14 deletions in CLL involve cooperating tumor suppressors. Blood *115*, 3916–3922. https://doi.org/10.1182/BLOOD-2009-10-249367.

16. Klein, U., Lia, M., Crespo, M., Siegel, R., Shen, Q., Mo, T., Ambesi-Impiombato, A., Califano, A., Migliazza, A., Bhagat, G., et al. (2010). The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. Cancer Cell *17*, 28–40. https://doi.org/10.1016/J.CCR.2009.11.019.

17. Cimmino, A., Calin, G.A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., Wojcik, S.E., Aqeilan, R.I., Zupo, S., Dono, M., et al. (2005). miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A *102*, 13944–13949. https://doi.org/10.1073/PNAS.0506654102.

18. Landau, D.A., Tausch, E., Taylor-Weiner, A.N., Stewart, C., Reiter, J.G., Bahlo, J., Kluth, S., Bozic, I., Lawrence, M., Böttcher, S., et al. (2015). Mutations driving CLL and their evolution in progression and relapse. Nature *526*, 525–530. https://doi.org/10.1038/NATURE15395.

19. Döhner, H., Stilgenbauer, S., Benner, A., Leupolt, E., Kröber, A., Bullinger, L., Döhner, K., Bentz, M., and Lichter, P. (2000). Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med *343*, 1910–1916. https://doi.org/10.1056/NEJM200012283432602.

20. Shiloh, Y., and Ziv, Y. (2013). The ATM protein kinase: regulating the cellular response to genotoxic stress, and more. Nat Rev Mol Cell Biol *14*, 197–210. https://doi.org/10.1038/NRM3546.

21. Skowronska, A., Parker, A., Ahmed, G., Oldreive, C., Davis, Z., Richards, S., Dyer, M., Matutes, E., Gonzalez, D., Taylor, A.M.R., et al. (2012). Biallelic ATM inactivation significantly reduces survival in patients treated on the United Kingdom Leukemia Research Fund Chronic Lymphocytic Leukemia 4 trial. J Clin Oncol *30*, 4524–4532. https://doi.org/10.1200/JCO.2011.41.0852.

22. Wierda, W.G., O'Brien, S., Wang, X., Faderl, S., Ferrajoli, A., Do, K.A., Garcia-Manero, G., Cortes, J., Thomas, D., Koller, C.A., et al. (2011). Multivariable model for time to first

treatment in patients with chronic lymphocytic leukemia. J Clin Oncol *29*, 4088–4095. https://doi.org/10.1200/JCO.2010.33.9002.

23. Döhner, H., Fischer, K., Bentz, M., Hansen, K., Benner, A., Cabot, G., Diehl, D., Schlenk, R., Coy, J., Stilgenbauer, S., et al. (1995). p53 Gene Deletion Predicts for Poor Survival and Non-Response to Therapy With Purine Analogs in Chronic B-Cell Leukemias. Blood *85*, 1580–1589. https://doi.org/10.1182/BLOOD.V85.6.1580.BLOODJOURNAL8561580.

24. Zenz, T., Kröber, A., Scherer, K., Häbe, S., Bühler, A., Benner, A., Denzel, T., Winkler, D., Edelmann, J., Schwänen, C., et al. (2008). Monoallelic TP53 inactivation is associated with poor prognosis in chronic lymphocytic leukemia: results from a detailed genetic characterization with long-term follow-up. Blood *112*, 3322–3329. https://doi.org/10.1182/BLOOD-2008-04-154070.

25. Wang, L., Lawrence, M.S., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D.S., Zhang, L., et al. (2011). SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med *365*, 2497–2506. https://doi.org/10.1056/NEJMOA1109016.

26. Puente, X.S., Pinyol, M., Quesada, V., Conde, L., Ordóñez, G.R., Villamor, N., Escaramis, G., Jares, P., Beá, S., González-Díaz, M., et al. (2011). Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature *475*, 101–105. https://doi.org/10.1038/NATURE10113.

27. Wang, L., Brooks, A.N., Fan, J., Wan, Y., Gambe, R., Li, S., Hergert, S., Yin, S., Freeman, S.S., Levin, J.Z., et al. (2016). Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. Cancer Cell *30*, 750–763. https://doi.org/10.1016/J.CCELL.2016.10.005.

28. Del Giudice, I., Rossi, D., Chiaretti, S., Marinelli, M., Tavolaro, S., Gabrielli, S., Laurenti, L., Marasca, R., Rasi, S., Fangazio, M., et al. (2012). NOTCH1 mutations in +12 chronic lymphocytic leukemia (CLL) confer an unfavorable prognosis, induce a distinctive transcriptional profiling and refine the intermediate prognosis of +12 CLL. Haematologica *97*, 437–441. https://doi.org/10.3324/HAEMATOL.2011.060129.

29. Rossi, D., Spina, V., Bomben, R., Rasi, S., Dal-Bo, M., Bruscaggin, A., Rossi, F.M., Monti, S., Degan, M., Ciardullo, C., et al. (2013). Association between molecular lesions and specific B-cell receptor subsets in chronic lymphocytic leukemia. Blood *121*, 4902–4905. https://doi.org/10.1182/BLOOD-2013-02-486209.

30. Melo, J. V., Catovsky, D., and Galton, D.A.G. (1986). The relationship between chronic lymphocytic leukaemia and prolymphocytic leukaemia: II. PATTERNS OF EVOLUTION OF'PROLYMPHOCYTOID'TRANSFORMATION. Br J Haematol *64*, 77–86. https://doi.org/10.1111/j.1365-2141.1986.tb07575.x.

31. Hallek, M. (2019). Chronic lymphocytic leukemia: 2020 update on diagnosis, risk stratification and treatment. Am J Hematol *94*, 1266–1287. https://doi.org/10.1002/AJH.25595.

32. Shahjahani, M., Mohammadiasl, J., Noroozi, F., Seghatoleslami, M., Shahrabi, S., Saba, F., and Saki, N. (2015). Molecular basis of chronic lymphocytic leukemia diagnosis and prognosis. Cellular Oncology *38*, 93–109. https://doi.org/10.1007/S13402-014-0215-3/FIGURES/1.

33. Gaidano, G., Foà, R., and Dalla-Favera, R. (2012). Molecular pathogenesis of chronic lymphocytic leukemia. J Clin Invest *122*, 3432–3438. https://doi.org/10.1172/JCI64101.

34. Damle, R.N., Wasil, T., Fais, F., Ghiotto, F., Valetto, A., Allen, S.L., Buchbinder, A., Budman, D., Dittmar, K., Kolitz, J., et al. (1999). Ig V Gene Mutation Status and CD38 Expression As Novel Prognostic Indicators in Chronic Lymphocytic Leukemia. Blood *94*, 1840–1847. https://doi.org/10.1182/BLOOD.V94.6.1840.

35. Tam, C.S., Shanafelt, T.D., Wierda, W.G., Abruzzo, L. V., Van Dyke, D.L., O'Brien, S., Ferrajoli, A., Lerner, S.A., Lynn, A., Kay, N.E., et al. (2009). De novo deletion 17p13.1 chronic lymphocytic leukemia shows significant clinical heterogeneity: the M. D. Anderson and Mayo Clinic experience. Blood *114*, 957–964. https://doi.org/10.1182/BLOOD-2009-03-210591.

36. Rossi, D., and Gaidano, G. (2012). ATM and chronic lymphocytic leukemia: mutations, and not only deletions, matter. Haematologica *97*, 5–8. https://doi.org/10.3324/HAEMATOL.2011.057109.

37. Mansouri, L., Cahill, N., Gunnarsson, R., Smedby, K.E., Tjönnfjord, E., Hjalgrim, H., Juliusson, G., Geisler, C., and Rosenquist, R. (2013). NOTCH1 and SF3B1 mutations can be added to the hierarchical prognostic classification in chronic lymphocytic leukemia. Leukemia *27*, 512–514. https://doi.org/10.1038/LEU.2012.307.

38. Balatti, V., Bottoni, A., Palamarchuk, A., Alder, H., Rassenti, L.Z., Kipps, T.J., Pekarsky, Y., and Croce, C.M. (2012). NOTCH1 mutations in CLL associated with trisomy 12. Blood *119*, 329–331. https://doi.org/10.1182/BLOOD-2011-10-386144.

39. Foà, R., Del Giudice, I., Guarini, A., Rossi, D., and Gaidano, G. (2013). Clinical implications of the molecular genetics of chronic lymphocytic leukemia. Haematologica *98*, 675–685. https://doi.org/10.3324/HAEMATOL.2012.069369.

40. Rodríguez-Vicente, A.E., Díaz, M.G., and Hernández-Rivas, J.M. (2013). Chronic lymphocytic leukemia: a clinical and molecular heterogenous disease. Cancer Genet *206*, 49–62. https://doi.org/10.1016/J.CANCERGEN.2013.01.003.

41. Kulis, M., Heath, S., Bibikova, M., Queirós, A.C., Navarro, A., Clot, G., Martínez-Trillos, A., Castellano, G., Brun-Heath, I., Pinyol, M., et al. (2012). Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet *44*, 1236–1242. https://doi.org/10.1038/NG.2443.

42. Gaidano, G., Foà, R., and Dalla-Favera, R. (2012). Molecular pathogenesis of chronic lymphocytic leukemia. J Clin Invest *122*, 3432. https://doi.org/10.1172/JCI64101.

43. Rossi, D., Fangazio, M., Rasi, S., Vaisitti, T., Monti, S., Cresta, S., Chiaretti, S., Del Giudice, I., Fabbri, G., Bruscaggin, A., et al. (2012). Disruption of BIRC3 associates with fludarabine chemorefractoriness in TP53 wild-type chronic lymphocytic leukemia. Blood *119*, 2854–2862. https://doi.org/10.1182/BLOOD-2011-12-395673.

44. Rossi, D., Bruscaggin, A., Spina, V., Rasi, S., Khiabanian, H., Messina, M., Fangazio, M., Vaisitti, T., Monti, S., Chiaretti, S., et al. (2011). Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood *118*, 6904–6908. https://doi.org/10.1182/BLOOD-2011-08-373159.

45. Hamblin, T.J., Davis, Z., Gardiner, A., Oscier, D.G., and Stevenson, F.K. (1999). Unmutated Ig VH Genes Are Associated With a More Aggressive Form of Chronic Lymphocytic Leukemia. Blood *94*, 1848–1854. https://doi.org/10.1182/BLOOD.V94.6.1848.

46. Lan, X., Ruminy, P., Bohers, E., Rainville, V., Viennot, M., Viailly, P.J., Etancelin, P., Tilly, H., Mihailescu, S., Bouclet, F., et al. (2022). 5' Rapid amplification of cDNA ends (5'RACE): A simpler method to analyze immunoglobulin genes and discover the value of the light chain in chronic lymphocytic leukemia. Leuk Res *123*. https://doi.org/10.1016/J.LEUKRES.2022.106952.

47. Kulis, M., Heath, S., Bibikova, M., Queirós, A.C., Navarro, A., Clot, G., Martínez-Trillos, A., Castellano, G., Brun-Heath, I., Pinyol, M., et al. (2012). Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet *44*, 1236–1242. https://doi.org/10.1038/NG.2443.

48. Nadeu, F., Royo, R., Clot, G., Duran-Ferrer, M., Navarro, A., Martín, S., Lu, J., Zenz, T., Baumann, T., Jares, P., et al. (2021). IGLV3-21R110 identifies an aggressive biological subtype of chronic lymphocytic leukemia with intermediate epigenetics. Blood *137*, 2935–2946. https://doi.org/10.1182/BLOOD.2020008311.

49. Lan, X., Ruminy, P., Bohers, E., Rainville, V., Viennot, M., Viailly, P.J., Etancelin, P., Tilly, H., Mihailescu, S., Bouclet, F., et al. (2022). 5′ Rapid amplification of cDNA ends (5′RACE): A simpler method to analyze immunoglobulin genes and discover the value of the light chain in chronic lymphocytic leukemia. Leuk Res *123*, 106952. https://doi.org/10.1016/J.LEUKRES.2022.106952.

50. Innocenti, I., Benintende, G., Tomasso, A., Fresa, A., Autore, F., Larocca, L.M., and Laurenti, L. (2023). Richter transformation in Chronic Lymphocytic Leukemia. Hematol Oncol *41*, 293–300. https://doi.org/10.1002/HON.3106.

51. Rossi, D., Cerri, M., Capello, D., Deambrogi, C., Rossi, F.M., Zucchetto, A., De Paoli, L., Cresta, S., Rasi, S., Spina, V., et al. (2008). Biological and clinical risk factors of chronic

lymphocytic leukaemia transformation to Richter syndrome. Br J Haematol *142*, 202–215. https://doi.org/10.1111/J.1365-2141.2008.07166.X.

52. Rossi, D., Spina, V., Deambrogi, C., Rasi, S., Laurenti, L., Stamatopoulos, K., Arcaini, L., Lucioni, M., Rocque, G.B., Xu-Monette, Z.Y., et al. (2011). The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. Blood *117*, 3391–3401. https://doi.org/10.1182/BLOOD-2010-09-302174.

53. Rossi, D., and Gaidano, G. (2013). Richter syndrome. Adv Exp Med Biol *792*, 173–191. https://doi.org/10.1007/978-1-4614-8051-8_8/COVER.