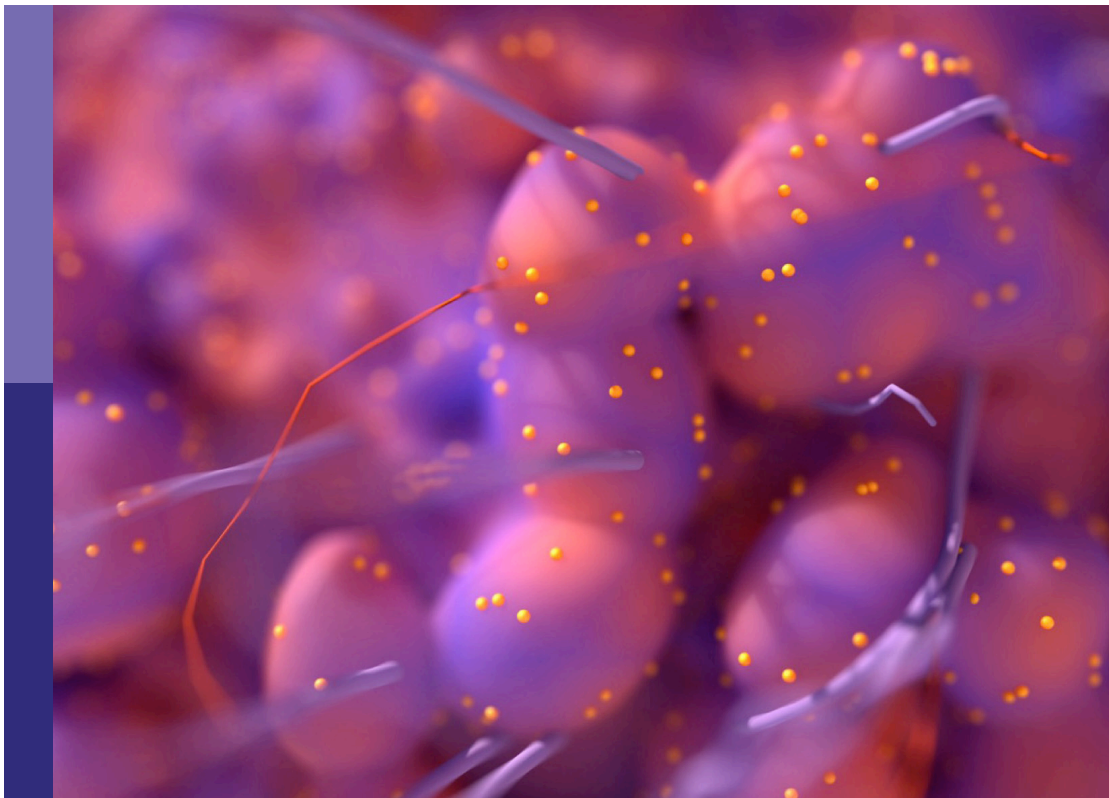
frontiers | Research Topics

Role of the antigen receptor in the pathogenesis of B-cell lymphoid malignancies

Edited by
Andreas Agathangelidis, Manlio Ferrarini and Franco Fais

Published in
Frontiers in Oncology



ISSN 1664-8714
ISBN 978-2-83252-114-4
DOI 10.3389/978-2-83252-114-4

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Role of the antigen receptor in the pathogenesis of B-cell lymphoid malignancies

Topic editors

Andreas Agathangelidis — National and Kapodistrian University of Athens, Greece
Manlio Ferrarini — University of Genoa, Italy
Franco Fais — University of Genoa, Italy

Citation

Agathangelidis, A., Ferrarini, M., Fais, F., eds. (2023). *Role of the antigen receptor in the pathogenesis of B-cell lymphoid malignancies*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83252-114-4

Table of contents

frontiers | Frontiers in Oncology

TYPE Editorial
PUBLISHED 23 March 2023
DOI 10.3389/fonc.2023.1188375

OPEN ACCESS

EDITED AND REVIEWED BY
Alessandro Isidori,
AORMN Hospital, Italy

*CORRESPONDENCE
Andreas Agathangelidis
agathan@biol.uoa.gr

SPECIALTY SECTION
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

RECEIVED 17 March 2023
ACCEPTED 20 March 2023
PUBLISHED 23 March 2023

CITATION
Agathangelidis A, Ferrarini M and Fais F (2023) Editorial: Role of the antigen receptor in the pathogenesis of B-cell lymphoid malignancies.
Front. Oncol. 13:1188375.
doi: 10.3389/fonc.2023.1188375

Editorial: Role of the antigen receptor in the pathogenesis of B-cell lymphoid malignancies

Andreas Agathangelidis[1]*, Manlio Ferrarini[2] and Franco Fais[2,3]

[1]Department of Biology, School of Science, National and Kapodistrian University of Athens, Athens, Greece, [2]Department of Experimental Medicine, University of Genoa, Genoa, Italy, [3]Molecular Pathology Unit, IRCCS Ospedale Policlinico San Martino, Genoa, Italy

KEYWORDS

B-cell lymphoid malignancies, B cell receptor, immunoglobulin, chronic lymphocytic leukemia (CLL), immunogenetics, monoclonal B cell lymphocytosis (MBL), ontogenesis, familial CLL

Editorial on the Research Topic
Role of the antigen receptor in the pathogenesis of B-cell lymphoid malignancies

This Research Topic contains nine manuscripts related to different scientific aspects of the *"Role of the Antigen Receptor in the Pathogenesis of B-Cell Lymphoid Malignancies"*. Each publication addresses pending issues, including (i) the role of the B cell receptor (BcR) in the ontogeny of these malignancies; (ii) the clinical relevance of particular properties of the clonotypic BcR immunoglobulin (IG) and other associated surface cell markers; and, finally, (iii) the benefits of targeting this crucial receptor for therapeutic purposes.

The notion that antigen selection of B cells through the BcR drives the pathogenesis of B-Cell Lymphoid Malignancies, such as CLL, is now well established. CLL is always preceded by monoclonal B-cell lymphocytosis (MBL), defined by a clonal B cell population of less than 5×10^9/L and no symptoms or signs of disease. In this context, the review by Galigalidou et al. contains valuable information on the role of microenvironmental interactions in MBL ontogenesis and its progression to CLL. More specifically, the study of immune cell (B and T cells) receptor repertoires revealed important differences between MBL and CLL, alluding to distinct selection forces, both in terms of the nature of the selective antigens as well as the persistence of these interactions. Furthermore, the study of residual B cells revealed an impaired B cell production in the bone marrow, already at the stage of MBL. Hence, the tumor microenvironment in MBL may be pivotal for understanding the initial steps of malignant transformation.

Along the same lines, Kolijn et al. provided relevant information about the ontogenesis of a specific type of CLL, defined as familial CLL. Of interest, all four affected siblings of one of the families included in the study carried BcR IG expressing the IGLV3-21 gene with the hallmark R110 mutation. The BcR IG in 2/4 siblings were assigned to either stereotyped subset #2 or its immunogenetic relative subset #169, both of which belong to the clinically aggressive IGLV3-21^{R110} CLL subgroup. Furthermore, the CLL clones within each family exhibited driver gene mutations previously associated with IGHV mutational status, cytogenetic aberrations and stereotyped subsets. Altogether, these findings underline the

notion that specific immunogenetic characteristics in combination with genetic aberrations drive CLL development, at least of the familial type.

In recent years, there has been accumulating evidence that the BcR expression levels and its functionality may associate directly or indirectly with other molecules. CD5 is considered among the most relevant ones in the context of CLL, which is located close to the BcR IG on the surface of the B cells and promotes cell survival and proliferation. In a brief research report, Maisano et al. described a novel peptide-based single-cell sorting methodology using the clonotypic BcR IG as bait. Hence, this study provided a proof of concept for the use of BcR IG ligands as probes for sorting and analyzing CLL clones. At the scientific level, transcriptomic analysis showed that the CD5 expression levels correlated with the expansion of the CLL clone, revealing a novel mechanism that could affect clonal expansion and persistence in CLL.

Over the years, immunogenetic studies in several B-Cell Lymphoid Malignancies support the theory for antigen drive by identifying distinct biases in the BcR IG gene repertoires. For example, the VH CDR3 sequences of the clonotypic BcR IG in CLL are characterized by length and amino acid composition restrictions. Rodriguez-Caballero et al. investigated the hydropathy index of the VH CDR3 in a large series of CLL patients and performed associations with other prognostic factors. Overall, two distinct subgroups of M-CLL patients emerged, displaying a neutral versus a negatively charged VH CDR3. Substantial differences were observed, with the M-CLL subgroup with neutral VH CDR3 being characterized by the predominance of the male gender, more advanced disease stage and a higher frequency of genetic aberrations, together with a higher rate of disease progression and shorter time-to-therapy (TTT). These findings further corroborate the relevance of the VH CDR3 in particular, and the BcR in general in CLL pathogenesis.

Another unique property of CLL is that a large fraction of clones (around 40%) are characterized by the expression of stereotyped BcR IG, which display distinct biological and clinical properties. Furthermore, CLL BcR IG have often been shown to carry autoreactive properties, alluding to a defect in immune tolerance in the respective patients. In two independent studies, Bagnara et al. and Vergani et al. performed high-throughput sequencing to explore the presence of stereotyped BcR IG in healthy donors. "CLL-like" stereotyped BcR IG were identified with no evidence of preferential accumulation in specific B-cell subpopulations (including $CD5^+$ B cells at this pre-leukemic phase), possibly because either the level of autoreactivity is not high enough to be considered as dangerous by tolerance mechanisms or due to editing of the clonotypic IG light chain genes.

The mutational load of the rearranged IGHV gene is considered one of the most accurate prognostic markers in CLL; in detail, M-CLL patients have better outcomes than patients with U-CLL, probably because somatic IGHV mutations may affect the BcR IG structure towards abolishing polyreactivity. Kaufman et al. tried to address the latter by comparing cases with different ratios of replacement (R) mutations that lead to non-conservative amino acid changes (Rnc) to the combined numbers of conservative (Rc) and silent (S) amino acid mutations. When comparing time-to-first-treatment (TTFT) of patients with (S+Rc)/Rnc ≤ 1 and >1, TTFTs were quite similar. As the authors proposed, the structure of the BcR IG may not be the most critical factor for dictating outcomes in CLL, yet one should keep in mind that SHMs, even those of non-conservative nature, do not always affect the BcR IG structure substantially.

Over the years, significant progress has been made in the therapeutic management of CLL as well as other B-cell malignancies; in detail, targets in the BcR signaling pathway, such as BTK and PI3Kδ, have emerged as a successful treatment strategy. Unfortunately, a proportion of patients still relapse, indicating the need to identify new therapeutic targets. In this context, Sana et al. provided a comprehensive review regarding the importance of studying and identifying new potential druggable targets, focusing on NFAT. These transcription factors are involved in inflammation and the development of both autoimmune and neoplastic diseases. In more detail, NFAT1 and NFAT2 were described to affect cell proliferation and cell death after BcR stimulation. Finally, targeting NFAT was beneficial in treating CLL and lymphoma in preclinical models, with ABC DLBCL cells being particularly dependent on the activation of the NFAT pathway.

Of interest, Arbel et al. demonstrated that the BcR pathway can be efficiently targeted in CLL cells using proteolysis targeting chimeras (PROTACs). More specifically, the reversible non-covalent compound (NC-1) could degrade BTK in CLL cells, leading to decreased baseline BTK phosphorylation. Furthermore, this led to lower levels of activation of BTK and other signaling molecules downstream of the BcR pathway, following IgM engagement. These effects were also found in samples from CLL patients with clinical resistance to ibrutinib and the BTK mutation C481Y.

Overall, this collection contains several new information, concepts, and ideas related to the role of BCR in lymphoproliferative diseases that can also be used as further insights for work in this field.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Oncology

REVIEW
published: 01 April 2021
doi: 10.3389/fonc.2021.651057

Role of NFAT in Chronic Lymphocytic Leukemia and Other B-Cell Malignancies

Ilenia Sana[1†], *Maria Elena Mantione*[1†], *Piera Angelillo*[1,2] *and Marta Muzio*[1*]

[1] Division of Experimental Oncology, San Raffaele Hospital IRCCS, Milano, Italy, [2] Lymphoma Unit, Department of Onco-Hematology, IRCCS San Raffaele Scientific Institute, Milan, Italy

In recent years significant progress has been made in the clinical management of chronic lymphocytic leukemia (CLL) as well as other B-cell malignancies; targeting proximal B-cell receptor signaling molecules such as Bruton Tyrosine Kinase (BTK) and Phosphoinositide 3-kinase (PI3Kδ) has emerged as a successful treatment strategy. Unfortunately, a proportion of patients are still not cured with available therapeutic options, thus efforts devoted to studying and identifying new potential druggable targets are warranted. B-cell receptor stimulation triggers a complex cascade of signaling events that eventually drives the activation of downstream transcription factors including Nuclear Factor of Activated T cells (NFAT). In this review, we summarize the literature on the expression and function of NFAT family members in CLL where NFAT is not only overexpressed but also constitutively activated; NFAT controls B-cell anergy and targeting this molecule using specific inhibitors impacts on CLL cell viability. Next, we extend our analysis on other mature B-cell lymphomas where a distinct pattern of expression and activation of NFAT is reported. We discuss the therapeutic potential of strategies aimed at targeting NFAT in B-cell malignancies not overlooking the fact that NFAT may play additional roles regulating the inflammatory microenvironment.

Keywords: nuclear factor of activated T cells, B-cell receptor, chronic lymphocytic leukemia, lymphoma, lymphoid malignancies

OPEN ACCESS

Edited by:
Franco Fais, University of Genoa, Italy

Reviewed by:
Magdalena Winiarska, Medical University of Warsaw, Poland
Albrecht Reichle, University Medical Center Regensburg, Germany

***Correspondence:**
Marta Muzio
muzio.marta@hsr.it

These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: *08 January 2021*
Accepted: *15 March 2021*
Published: *01 April 2021*

Citation:
Sana I, Mantione ME, Angelillo P and Muzio M (2021) Role of NFAT in Chronic Lymphocytic Leukemia and Other B-Cell Malignancies. Front. Oncol. 11:651057. doi: 10.3389/fonc.2021.651057

INTRODUCTION

Targeting B-Cell Receptor Signaling in Chronic Lymphocytic Leukemia

In recent years there have been significant improvements in the field of chronic lymphocytic leukemia (CLL) from both bench and bedside perspectives. CLL cells are addicted to different microenvironmental stimuli with a key role being played by the B-cell receptor (BCR) stimulation and/or constitutive cell autonomous BCR activation (1) leading to cell survival and proliferation.

Abbreviations: BCR, B cell receptor; BTK, Bruton Tyrosine Kinase; CLL, Chronic Lymphocytic Leukemia; DLBCL, Diffuse large B-cell lymphoma; NFAT, Nuclear Factor of Activated T cells; NFκB, Nuclear Factor kappa-light-chain-enhancer of activated B cells; PI3K, Phosphoinositide 3-kinases; TF, Transcription Factor.

On this scientific basis, several lines of research led to the development of small molecule inhibitors of the kinases that transmit the signals from the proximal BCR signaling complex to the downstream Transcription Factors (TFs) [i.e. Bruton Tyrosine Kinase (BTK) and Phosphoinositide 3-kinases delta (PI3Kδ)] (2, 3). CLL is a disease of the elderly, but it can also affect younger patients, who often are the most difficult to treat given the long disease history and the frequent need for multiple lines of treatment. To date, the only curative approach for CLL patients, unfortunately, is allogeneic bone marrow transplantation, a procedure burdened by a high rate of morbidity and mortality especially in older patients. Nowadays, one of the most powerful available tools for the treatment of CLL, is a novel small molecule that inhibits BTK, Ibrutinib. The drug was approved by the US Food and Drug Administration in 2014 for the treatment of relapsed refractory CLL and for CLL patients with the 17p deletion/mutation, known to frequently be chemo refractory. Ibrutinib was granted approval for first-line treatment of CLL in March 2016; since then, data from real world practice consistently shows a significant improvement in survival curves, in keeping with what investigators previously observed in clinical trials (4, 5). BTK is not the only target of novel non chemotherapeutic agents, as both BTK and PI3Kδ inhibitors are highly effective even for the treatment of refractory or relapsed disease. These drugs are designed to be administered until relapse/progression or unacceptable toxicity. Patients are therefore kept under follow-up during the administration for early detection of signs of clinical progression or toxicity. Several factors have been proposed as predictive markers for the emergence of resistance (i.e. prolonged lymphocytosis) (6, 7), which is currently an unmet clinical need. Novel therapeutic strategies are needed to cure refractory patients and to perhaps achieve deeper response with the intent of fully eradicating the disease.

In this context, we hypothesize that exploring other downstream signaling mediators including transcription factors may reveal novel vulnerabilities of malignant B cells, which could be of aid in treating CLL and other B cell malignancies. Given that resistance to targeted agents often occurs by mutation of the target kinase (8, 9), it is reasonable to hypothesize that blocking downstream signaling molecules could be a strategy to block the transmission of the survival signal to the nucleus. One of the transcription factors involved in B-cell antigen receptor signaling is NFAT, Nuclear factor of activated T-lymphocytes [others being Nuclear Factor kappa-light-chain-enhancer of activated B cells (NFκB), cMyc and activator protein 1 (AP1)]. Here, we briefly describe the biology of NFAT followed by a discussion on the expression pattern and functional role of NFAT transcription factors in CLL and other lymphoid malignancies.

The Nuclear Factor of Activated T-Cells Family of Transcription Factors

The NFAT family of TFs includes five members grouped by the presence of the REL homology region (RHR), a highly conserved DNA binding domain that confers a unique DNA binding specificity to these proteins (see **Figure 1** for a schematic representation). Several alternative names exist for each NFAT member, and they are all reported in **Table 1** together with essential information on each gene. Herein, we refer to the NFAT family members with their official names; NFAT1, NFAT2, NFAT3, NFAT4, and NFAT5.

Four NFAT proteins share sequence homology in the N-term regulatory region (NHR) responsible for the modulation by Calcium signaling, while NFAT5 is induced by osmotic stress (10). Briefly, calcium-dependent NFATs are normally retained in an inactive state into the cytoplasm of the cells by different kinases that phosphorylate the NHR domain (11). The stimulation of receptors such as the BCR in B-cells and

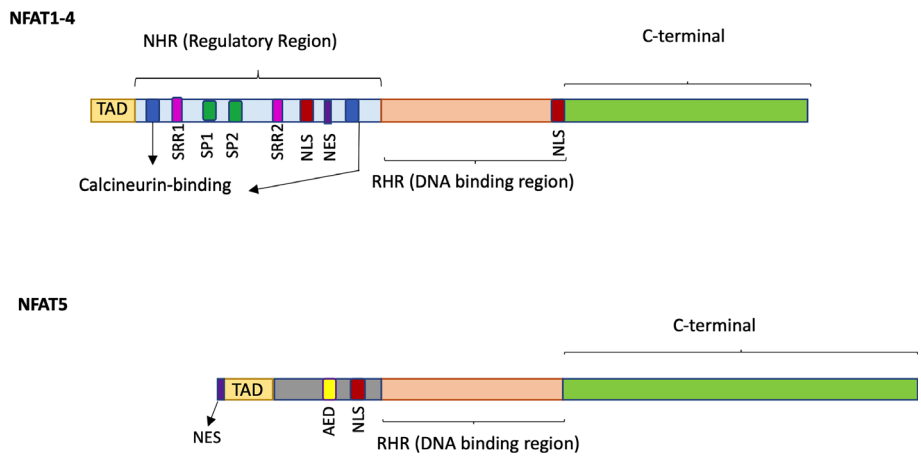


FIGURE 1 | Schematic representation of the structure of NFAT1-4 and NFAT5. NFAT proteins contain a REL homology region (RHR), the most conserved domain, which binds to DNA and is common to all five NFAT family members. The NFAT homology region (NHR) is the regulatory region conserved in NFAT1-4 members but not in NFAT5 and contains two calcineurin-binding sites, serine residues (SSR, SP), a Nuclear Localization Sequence (NLS) and a Nuclear Export Signal (NES). NFAT5 does not display the calcineurin binding site but owns an auxiliary export domain (AED) and a NLS and possess a nuclear export sequence (NES) located at the first 19 amino acids. All NFAT family members have a N-terminal Transactivation Domain (TAD) and a C-terminal domain which are the less conserved regions.

TABLE 1 | Summary of key features of human NFAT genes, mRNAs and proteins.

Name	Aliases	Chromosome	n° of splice variants (Uniprot)	n° of splice variants (Ensembl)	total n° of iso-forms	n° of different protein coding variants	UniprotKB entry	Ensembl entry	NCBI entry	RNA expression in tissues*	RNA expression in blood*
NFAT1	NFATc2, NFATp, NF-ATP	20q13.2	5	7	7	6	Q13469	ENSG00000101096	4773	Appendix, Lymph node, Tonsil	γδT cell, NK- cell, memory CD8 T-cell
NFAT2	NFATc1, NFATc, Nf-ATC, NF-ATc1.2	18q23	10 (+3)	18	19	11	O95644	ENSG00000131196	4772	Appendix, Tonsil, Lymph node	Non classical monocyte, Naive B cell, intermediate monocyte, memory B cell
NFAT3	NFATC4, NF-ATc4	14q12	24 (+3)	33	34	22	Q14934	ENSG00000100968	4776	Ovary, Cervix uterine, urinary bladder, gall bladder	not detected
NFAT4	NFATC3, NFATX	16q22.1	6 (+6)	22	25	5	Q12968	ENSG00000072736	4775	Thymus	γδT cell, memory CD8 T-cell, ^MAIT T cell, Naive CD8, NK-cells
NFAT5	TONEBP, NFATL1, KIAA0827, OREBP, NFATZ, NF-AT5	16q22.1	5 (+6)	13	14	4	O94916	ENSG00000102908	10725	Parathyroid gland, vagina, placenta, skin, bone marrow	Neutrophil, non classical monocyte

Tissues and cell types with relative high levels of mRNA expression are listed based on consensus data from Human protein atlas.

^MAIT T cell indicates Mucosal associated invariant T cell.

the T-cell receptor (TCR) in T-cells (12) generates a cascade that induces calcium mobilization through the activation of phospholipase C (PLCγ) that hydrolyzes phosphatidylinositol-3,4-bisphosphate (PIP_2) leading to the release of diacylglycerol (DAG) and inositol-1,4,5-trisphosphate (IP3). IP3 binds to the IP3 receptor and causes the release of Ca^{2+} from the endoplasmic reticulum and the consequent extracellular influx of Ca^{2+} in the cytosol by specific calcium channels (13, 14) (see **Figure 2** for a schematic representation). The presence of Ca^{2+} ions cause the binding of calmodulin (CaM) to the calcineurin phosphatase, leading to the dephosphorylation and activation of NFAT serine residues in the regulatory domain. Nuclear importing factors then mediate NFAT translocation into the nucleus where it binds the DNA alone or with other factors such as AP1, Stat3, GATA, c-Fos, c-Jun, and NFκB, to regulate gene expression either activating or silencing target genes; most of which are immune-related (15, 16). NFAT inactivation and the following relocation into the cytoplasm is operated by several kinases including glycogen-synthase kinase 3β [please refer to the following recent reviews for molecular details on NFAT proteins (16–19)].

The molecular activation of NFAT5 is more complex and still partially defined. Under isotonic conditions, there is a continuous shuttling of NFAT5 between the cytoplasm and the nucleus that can be regulated by tonicity stress. While hypotonic stress promotes nuclear export of the protein, hypertonic conditions induce transcription, translation and nuclear import of NFAT5 (20, 21). With its well-known tonicity-related regulatory role, it is emerging that several isotonic stimuli can promote NFAT5 activity; for instance, triggering of innate immunity receptors such as Toll-like receptors and consequent activation of reactive oxygen species (ROS) and mitogen activated protein kinases (MAPK) results in NFAT5 activation that, interestingly, shows distinctive features with respect to the osmotic activated response (22). However, the interplay between these two mechanisms of activation and the exact molecular pathways involved are still elusive and yet to be fully uncovered.

Studies on Different NFAT Deficient Mice Uncover Distinct Functional Roles in the B-Cell Lineage

To dissect the role of these TFs, several studies analyzed the phenotype of specific NFAT-deficient mice as well as combinations of two or more genetic deletions. For a comprehensive description of the murine lines please refer to online resources (e.g., the International Mouse Phenotyping Consortium: https://www.mousephenotype.org; MGI-Mouse Genome Informatics-: http://www.informatics.jax.org/). Herein, we focus on the results related to lymphocytes and the B-cell lineage as this may help in understanding and interpreting the specific role of individual NFAT proteins in B-cell malignancies.

With the notable exception of NFAT3, for which no specific studies analyzed the B-cell compartment, deletion of individual NFAT or double inactivation of NFAT genes *in vivo* revealed the importance of this TF family in the regulation of lymphocyte differentiation, proliferation, apoptosis, cytokine production, and

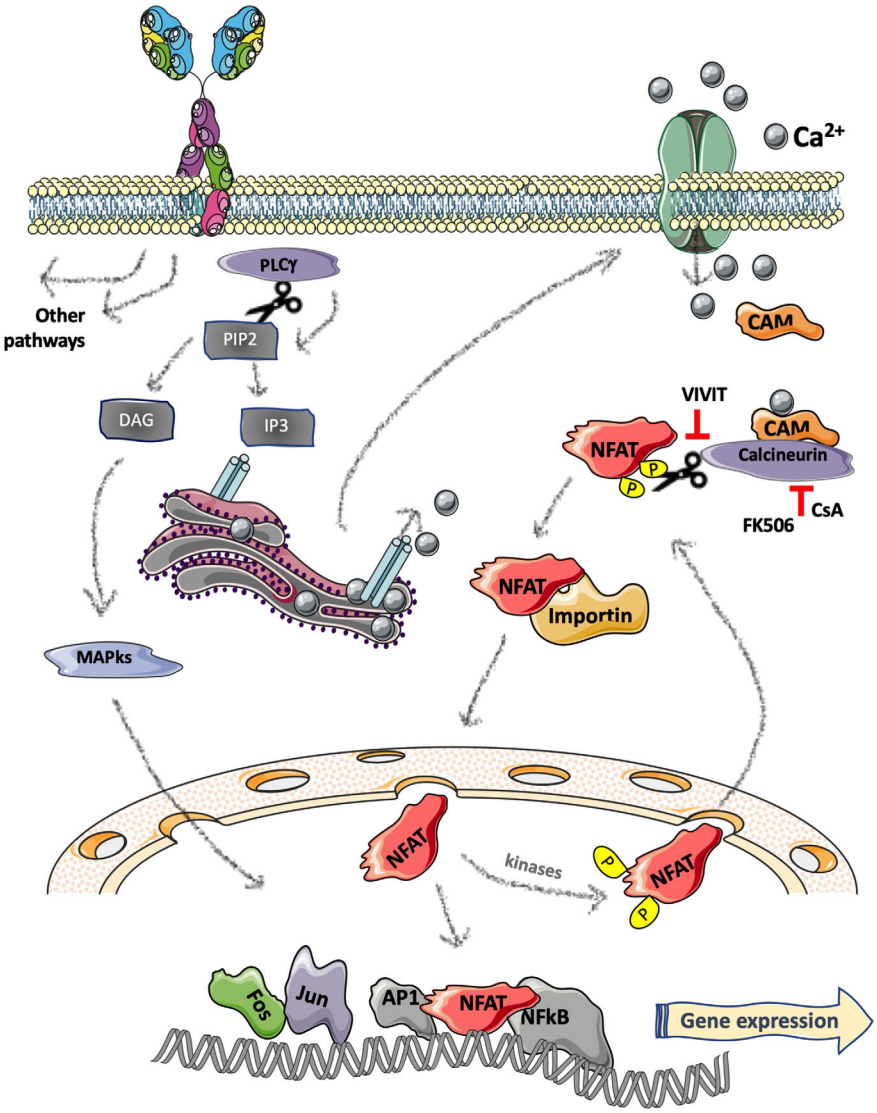


FIGURE 2 | Schematic representation of B-cell receptor induced NFAT pathway. Stimulation of the B-cell receptor (BCR) activates a cascade resulting in phospholipase C (PLCγ) activation and hydrolyzation of phosphatidylinositol-3,4-bisphosphate (PIP_2) generating the release of two second messengers: diacylglycerol (DAG) and inositol-1,4,5-trisphosphate (IP3). IP3 binds to the IP3 receptor located on the endoplasmic reticulum leading to release of Ca^{2+} from the ER and an extracellular influx of Ca^{2+} in the cytosol that causes the binding of calmodulin (CaM) to the calcineurin phosphatase and the consequent dephosphorylation and activation of NFAT. The nuclear translocation of NFAT is mediated by nuclear importing factors (e.g., importin). In the nucleus, NFAT binds to the DNA either alone or with other factors regulating gene expression. **Figure 2** was created using Servier Medical Art (https://smart.servier.com).

inflammation. The vast majority of the immunological studies, as well as the B cell-centered ones, focused on NFAT1 and NFAT2 that regulate different B-cell populations in diverse ways (23). Between the two, NFAT2 deficiency has a more severe impact; indeed, the global loss of NFAT2 determines prenatal lethality around day 14/15 upon gestation (24, 25). However, chimeric mice with lymphocyte-restrained NFAT2 loss showed defects in BCR-mediated proliferation of B cells (26), and a peculiar deficiency in peritoneal CD5+ B1a B-cells (27). Specifically, the defect in BCR-induced proliferation is determined by NFAT2-dependent expression of CD22, Rcan1, Tnfsf14, FasL and other key proteins of the BCR signaling pathway. In addition, the abrogation of NFAT2-mediated calcium flux response facilitates activation-induced cell death (AICD), which leads to the loss of CD5+ peritoneal B1a cells. Moreover, the lack of NFAT2-mediated repression of IL-10 production, impacts on IFN-γ production by CD4+ T-cells, impairing the capacity of B-cells to stimulate T cell proliferation (28). Mice lacking NFAT2 in pro-B cells have deficient expression of an essential TF determining B-cell lineage fate, EBF1, and a similar phenotype of EBF1-deficient

mice (29) with defective Immunoglobulin (Ig) gene rearrangement, and pre-BCR formation which impairs B cell development and leads to severe B-cell lymphopenia (30). The early developmental role of NFAT2 in peripheral B-cells was endorsed by a recent study that characterized lymphocyte dissemination in mice bearing deletion of NFAT2 in CD19-positive cells. The authors confirmed NFAT2-dependent deficiency in peritoneal CD5+ B1a B-cells that was accompanied by increased immature and mature follicular B cell populations (31). At later stages, NFAT2 loss of function causes functional defects only of mature B cells that promote mild clinical course of experimental autoimmune encephalomyelitis (28).

NFAT1-null mice have a normal development and a less severe phenotype. Nevertheless, after 6 months a proportion of the litters showed alterations of the immune system such as lymph node hyperplasia and splenomegaly accompanied by enlarged germinal centers and pronounced retardation in the involution of the thymus. NFAT1 deficient mice also displayed a hyperproliferative syndrome, higher B and T cell counts, dysregulated production of IL-4, and higher primary and secondary immune responses (32–34). Notably, NFAT1 plays a pivotal role in regulating the response of B-cells to self-antigen, balancing the processes of anergy and self-tolerance (35). Moreover, NFAT1 in B cells, controls and represses the expression of Cyclin E1 and E2, taking control of cell cycle progression and proliferation rates (36).

NFAT4 shows a more restricted role in regulating T cell activity; however, double deficient mice for NFAT1 and NFAT4 experience lymphadenopathy, splenomegaly, and a strongly activated phenotype with a substantial increase in serum IgE and IgG1 levels, similar to single knock-out mice (37). One of the origins of the lymphadenopathy was attributed to the observed resistance to apoptosis, due to decreased FasL expression and defective AICD induction (38). Moreover, the absence of both NFAT1 and NFAT4 drives naive CD4 T cells into Th2 cell differentiation even in the absence of endogenous IL-4, and boosts their responsiveness to TCR-mediated activation and secretion of Th2-type lymphokines (39). The elevated Th2 cytokine production also leads to hyperactivation of mature follicular B cells but not of marginal zone (MZ) B cells. This evidence indirectly links the loss of both NFAT1 and NFAT4 to the altered B cell phenotype of these mice, which have a lower representation of MZ B cells and a higher number of mature follicular B cells (40).

Focusing on NFAT5, the tonicity-responsive member of the family, its complete loss of function results in gestational lethality. Heterozygous animals show a phenotype marked by lymphoid hypocellularity, with thymus and spleen hypoplasia, defective antigen-specific antibody responses (in particular IgG secretion) and less mature CD4 and CD8 cells in the spleen and lymph nodes (41). These indications highlight the role for NFAT5-mediated adaptation to physiologic osmotic stress for lymphocyte-mediated immunity, with a putative B-cell centered role on T-cell dependent Ig response and proliferation, specifically under hypertonic conditions (42). Interestingly, NFAT5 showed a tonicity-independent role in the development and activation of macrophages where NFAT5 accumulation and the following increased expression of target genes such as TNF and IL6 can be mediated by Toll-like receptors and NFκB pathway activation (43). NFκB-mediated expression of NFAT5 also has a crucial role in pre–T-cell receptor thymocytes where it regulates the expression of the prosurvival factors A1 and Bcl2, and attenuates the proapoptotic p53/Noxa axis (44).

From all these observations it emerges that not only are NFAT1, NFAT2, and NFAT4 involved in the regulation and homeostasis of B-cells and BCR signaling, but NFAT5 also plays a crucial role which has yet to be fully characterized. On the contrary, no functional data are available for NFAT3. Nonetheless, both NFAT1 and NFAT2 are expressed by distinct B-cell malignancies as described in detail below, while less information is available on NFAT3, NFAT4, and NFAT5.

NFAT EXPRESSION AND ACTIVATION IN CLL AND OTHER LYMPHOID MALIGNANCIES

Expression and Function of NFAT in B-Cell Lymphomas

The category of non-Hodgkin lymphomas (NHL) comprises a large spectrum of entities ranging from indolent to highly aggressive diseases (45). Pathogenesis of most NHLs is unknown. For some subtypes, a chronic immune stimulation role has been suggested, thus, gaining insight on the multistep mechanism that leads to malignant transformation is key for the development of new treatments.

NFAT has neither been highlighted as a prominent B-cell lymphoma-associated molecule, nor a frequently mutated gene; however, distinct studies reported specific molecular and functional features of NFAT family members that may open up interesting novel therapeutic perspectives.

First, NFAT2 could be detected by IHC in lymphoid cells in routine biopsies of several hematologic malignancies, while nuclear NFAT2 was observed in a proportion of Burkitt and diffuse large B cell lymphoma (DLBCL) samples, suggesting an ongoing activation of the pathway in this type of lymphoma (46). In contrast, NFAT2, but not NFAT1, is downregulated by promoter methylation in Hodgkin's lymphoma cells (47). To note, the pattern of expression of NFAT does not fully reflect its role, as the activity of this TF is regulated by different mechanisms that eventually control the shuttling between the nucleus and the cytoplasm.

By using patients' derived cells and cell lines, Pham et al. not only reported constitutive activation of NFAT2 in large B-cell lymphoma, but also demonstrated that NFAT2 and NFκB cooperate to drive CD40L expression, which in turn triggers pro-survival signaling. NFAT siRNA inhibitors as well as drugs targeting NFAT activation blocked CD40L expression and induced apoptosis (48). The same group also demonstrated that both large B-cell lymphoma and mantle cell lymphoma constitutively express NFAT1 and NFAT2 that control BLyS expression and survival signaling again in cooperation with NFκB (49). Another interesting target of NFAT2 is c-Myc that

can be transcriptionally upregulated by NFAT2 through an epigenetic chromatin remodeling in DLBCL (50).

On the contrary, a tumor suppressive role of NFAT1 was proposed in DLBCL where its downregulation correlated with increased cyclin E expression; specifically, NFAT1 directly controlled cyclin E induction by binding to its promoter in lymphoma cell lines (36). Moreover, NFAT1 was implicated in mediating BCR-induced cell death in Hodgkin lymphoma cells, a phenomenon observed in selected cell lines; both Cyclosporin and FK506 inhibited apoptotic signaling, further supporting the functional involvement of NFAT (51).

Regarding the mechanisms responsible for NFAT overexpression and activation in lymphoma patients, while most of the studies suggest transcriptional and/or post-transcriptional mechanisms of regulation of NFAT in B-cell lymphoma, there is also evidence of NFAT gene amplification in a proportion of DLBCL of the ABC type (52). An increase of hexosamine biosynthetic pathway and O-GlcNAc metabolism plays a critical role in DLBCL cell proliferation, and is responsible for the observed NFκB and NFAT activation (53). Constitutive ongoing BCR signaling may also explain constitutive NFAT2 activation in DLBCL (54). However, a recent report suggested a BCR-independent, Calcium-dependent pathway towards NFAT2 activation in DLBCL (55, 56).

Finally, in addition to protein phosphorylation and splicing, distinct post-translational modifications modulate its activity including acetylation, SUMOylation and cleavage (57). However, no specific data on these mechanisms operating in B-cell malignancies are available.

Expression Pattern and Functional Role of NFAT in CLL

CLL involves mature clonal B-cells that accumulate in the peripheral blood and lymphoid organs where they receive supportive signals form the microenvironment. The BCR, the key protein for every B-lymphocyte, not surprisingly modulates CLL cells biology as well; several signaling molecules downstream of the BCR such as kinases and TFs are involved, and they have been recently targeted for therapeutic purposes (58, 59). Among the most relevant BCR-mediated TFs, we focus on NFAT, not overlooking the fact that NFκB as well as other TFs play a relevant role and may complement the activity of NFAT itself.

In 1996, Schuh et al. demonstrated for the first time that in contrast to normal B-cell, malignant cells isolated from the peripheral blood of patients with CLL show nuclear/active NFAT1 even in the absence of *in vitro* stimulation; in parallel, NFκB and AP1 activation were also observed (60). CLL cells show higher mRNA levels of expression of both NFAT1 and NFAT2 as compared to normal lymphocytes (61); hypomethylation of the NFAT2 promoter region as well the first intron region may explain higher levels of both mRNA and protein in CLL as compared to normal B-cells types (62). A comprehensive study on the epigenome and regulatory chromatic landscape of CLL highlighted that active chromatin regions were enriched for binding motifs of NFAT (as well as FOX and TCF/LEF transcription families) (63). When different groups of CLL patients with different clinic-biological characteristics were analyzed, NFAT promoter hypomethylation correlated with clinical staging (64). More recently, looking for markers of a specific subset of CLL, bearing trisomy of chromosome 12, Abruzzo at el. discovered that NFAT1, NFAT2, and NFAT4 mRNAs are significantly overexpressed (65).

Based on all these somewhat descriptive analyses, several authors suggested that BCR-mediated NFAT2 overexpression may be implicated in CLL pathobiology and may potentially be targeted for therapeutic purposes. Along this line of reasoning, several questions emerged: Which are the target genes of NFAT transcription factors in leukemic cells? Which are the functional consequences of NFAT hyperactivation in CLL? Which are the molecular mechanisms regulating NFAT activation in malignant cells?

No specific NFAT ChIP-seq analysis was performed in CLL cells, thus hampering a broad view of all its target genes; nevertheless, independent studies demonstrated that distinct genes that are typically expressed by leukemic cells are directly regulated by NFAT family members in different cell types.

CD23 is a receptor for FcE, and a distinctive molecule expressed on the surface of CLL cells as well as released in the serum (66). Two different isoforms regulated by two different promoters exist, namely CD23a and CD23b. The CD23b promoter is specifically regulated by NFAT1 and NFAT2 in concert with STAT6 (67); in contrast, CD23a expression is regulated by Notch2 (68). CD23 expression can also be upregulated by BCR stimulation in CLL cells where blocking NFAT prevents CD23 induction (61).

CD5 expression is regulated by NFAT in normal B-cell populations (27, 69, 70); however, it is not known if the same occurs in CLL where it is distinctively expressed on the cell surface.

LCK was recently identified as a direct NFAT2 target gene in both human and mouse CLL samples (71). CD5 mediated IL10 production is regulated by NFAT in CLL cells; by cooperating with STAT3, NFAT2 binds to IL5 and IL13 promoters and the IL10 enhancer to upregulate their expression (72).

NFAT2 is constitutively active in approximately half of CLL cases, the same than that characterized by concomitant MAPK phosphorylation and anergy in terms of response to the BCR. Despite constitutive basal levels of NFAT2 activation, it can be further induced after BCR stimulation, at least in a group of responding cases that are characterized by an adverse clinical outcome (61). Not only is NFAT overexpressed and activated in CLL, but NFAT binding sites are hypomethylated in leukemic samples suggesting the overactivation of target genes that may be related to autoreactive BCR (73).

To assess the functional role of NFAT during disease development and/progression, different approaches were used including ablation of NFAT2 in a mouse model of leukemia, and the use of drugs targeting NFAT activation. The two approaches address different questions whilst also being complementary; while genetic inactivation results in complete inhibition of a single molecule *in vivo* over time, drug treatment suddenly

interrupts different NFAT family members as well as any additional off targets. Here, we focus on genetic approaches while describing potential drug targeting in a dedicated paragraph below for both CLL and lymphoma.

Overexpression of the TCL1 oncogene in the B-cell lineage results in the development of a malignant disease resembling human CLL, as characterized by the accumulation of CD5-positive clonal B-cells in the peripheral blood and lymphoid organs (74); this is a widely used and accepted mouse model of leukemia. Leukemic cells of this model show constitutive activation of NFAT2 and somewhat anergic features (71), as previously shown for a group of CLL cases (75). Genetic inactivation of NFAT2 in the B-cell types of these mice led to rapid acceleration of leukemia development and progression toward and aggressive disease resembling Richter transformation, occurring in a small proportion of CLL patients (71). The authors suggest that NFAT2 is a key regulator of anergy in CLL. To better understand if NFAT deletion impacted directly on different BCR downstream signaling or rather on clonal selection of different BCR recombination, Muller et al. analyzed clonal evolution in leukemic mice and found that NFAT2 signaling in CLL cells precipitates the oligoclonal selection of preferentially unmutated BCRs (76).

Overall, these data suggest that NFAT2, being implicated in the maintenance of anergy, may restrict leukemia development; however, at the same time, anergic signaling provides a survival advantage to the cells. It is important to evaluate the effect of NFAT2 inhibition after leukemia development, as data from primary patient samples suggest that it may have therapeutic activity by interrupting anergy (77).

Therapeutic Perspectives

Given the central role of NFAT in regulating the adaptive immune response, it has been thoroughly scrutinized as a "specific" drug target to achieve immunosuppression in the context of organ transplantation, or to dampen excessive autoimmune manifestations. The most widely used drugs targeting the NFAT pathways are Cyclosporine A (CsA) and Tacrolimus, both originally isolated from fungi, acting with slightly different mechanisms of action. Briefly, as schematically reported in **Figure 2**, they inhibit the activity of the phosphatase Calcineurin, thus augmenting the phosphorylation state of all its substrates including NFAT1-4 family members; to do so, CsA binds to Cyclophilin while FK-506 (an alternative name of Tacrolimus) binds to FKBP12 (78). Both drugs are widely used to prevent graft-rejection or to treat autoimmune diseases; however, several years ago, anecdotical reports suggested its potential positive effect in the context of CLL (79, 80).

Nowadays, accumulating preclinical data from many different labs suggest that targeting NFAT may represent a novel therapeutic approach to treat at least a subset of NFAT-positive B-cell malignancies. In particular, CsA, FK-506 as well as a short cell-permeable NFAT-specific inhibitory peptide (called VIVIT as according to its core aminoacidic sequence) have been tested *in vitro* and *in vivo* in mouse models of leukemia and lymphoma (81, 82). To note, VIVIT peptide directly binds to the NFAT-docking portion of Calcineurin, thus being more selective towards these transcription factors as compared to CsA.

Targeting CLL cells with the VIVIT peptide blocks BCR-mediated NFAT2 activation and target genes in responsive cells (61, 71); at the same time, VIVIT blocks constitutive NFAT activation in CLL, thus rescuing leukemic cells from anergy (77). More importantly, targeting spontaneous NFAT activation with VIVIT not only blocked related biochemical pathways but also induced cell death *in vitro* and delayed leukemia progression *in vivo* in mouse models (77). These observations suggest that NFAT controls survival signals in anergic cells, though, at the same time keeps the cells in an indolent state. In fact, as mentioned above, deleting the whole NFAT2 gene from cells before leukemia development, triggers faster accumulation of the disease and progression (71).

Cyclosporin and FK506 inhibited NFAT signaling leading to the abrogation of pro-survival effects present in CLL cells, even in the presence of supportive signaling given by stroma cells (64). Interestingly, the use of the BTK signaling inhibitor, a clinically approved drug targeting upstream BCR signaling, abrogated NFAT activity in leukemic cells suggesting that downstream NFAT activation may represent a novel therapeutic target in the cases where resistance to Ibrutinib arise due to mutations in upstream molecules of the signaling cascade (64).

Targeting the upstream BTK kinase with different inhibitors, including Ibrutinib, blocks NFAT2 together with its targets including IL10; to note, IL10 can control PDL1 expression preferentially in the ABC subset of DLBCL thus suggesting that targeting NFAT may impact on the expression of this molecule, therefore being relevant for anti-tumor immunity (54). Along this line, a recent paper confirmed chronic NFAT activation in ABC-DLBCL controlling IL10 production and demonstrated the efficacy of calcineurin inhibitors in blocking NFAT signaling and reducing proliferation. However, the authors proposed a BCR-independent mechanism of NFAT activation, yet a BCR-dependent expression of NFAT protein that is dependent on NFκB signaling. Interestingly, blocking calcineurin synergized with BCL2 and MCL1 inhibitors to kill lymphoma cells (55).

Based on all these observations, using CsA, FK-506, or novel formulations of VIVT appear to be a promising therapeutic perspective. However, the concomitant immunosuppressive activity of these drugs may potentially counterbalance the direct anti-tumor activity, and it should be carefully considered for any putative future clinical approach (83). Along this line, it will be crucial to design and engineer novel drugs targeting NFAT in the malignant cells only. Finally, several additional drugs targeting NFAT are emerging, and it will be important to test them either alone or in combination with signaling inhibitors using advanced preclinical models and primary tumor samples.

DISCUSSION

The B-cell receptor is the key molecule regulating the pathobiology of both normal and malignant B-lymphocytes;

accordingly, targeting the kinases proximal to the BCR (such as BTK and PI3Kδ), has emerged as a successful treatment strategy. Nevertheless, BCR targeting is in some cases not sufficient to achieve disease control and, eventually, tumor cells find alternative mechanisms to survive and proliferate. For this reason, efforts devoted to studying and identifying additional new potential druggable targets along the BCR signaling cascade are warranted. In this review, we explored the biology of a key family of transcription factors that are activated after BCR stimulation, namely NFATs.

First, to obtain insight into the intrinsic role of different NFATs in the B-cell context, we reported an overview of genetically modified mice where a dual role of NFAT1 and NFAT2 emerged, regulating both cell proliferation and cell death after BCR stimulation. Next, we detailed the expression, activation status, and functional role of different NFAT family members in CLL and other B-cell malignancies. NFAT1 and NFAT2 were described to be not only overexpressed but also functionally implicated in the regulation of malignant B-cell biology. In contrast, no information was available on NFAT3 and NFAT4 in this context, while a recent paper demonstrated that NFAT5 is overexpressed in CLL where it facilitates malignant cells survival and activation (84); yet, NFAT5 activation is not dependent on BCR but it is regulated by osmotic stress and inflammatory stimuli. Overall, we described a cell autonomous function of BCR-related NFAT activation in leukemia and lymphoma cells. Targeting this molecule using a specific inhibitor was shown to be beneficial in treating CLL and lymphoma in preclinical models; however, NFAT deletion in mouse models broke anergy with a paradoxical induction of leukemia progression *in vivo* (85).

It has been known for years now that the NFAT family of transcription factors are involved in several processes which are central for immune system function, inflammation, and the development of both autoimmune and neoplastic diseases (86). This is particularly important in light of the role that the tumor microenvironment has been recognized to have for tumor survival and progression. NFAT may exert additional effects in different cell types, including stroma cells, by regulating inflammation and inflammation-associated cancer, as previously reported by other reviews on this topic (86, 87). To note, inflammation is a hallmark of CLL as well as other lymphoid malignancies where infiltrating immune cells, stroma, and vessels contribute to shape a complex tumor microenvironment (88–90). With this in mind it is reasonable to hypothesize that inhibition of the NFAT pathway could be effective for the treatment of lymphoproliferative disease since it affects cell function and survival both on and off the tumor. Recently Bucher et al. have reported strong evidence showing that NFAT signaling is chronically activated in DLBCL regulating cell survival and inflammatory cytokines (55, 56). In particular ABC DLBCL cells seem to be particularly dependent on the activation of the NFAT pathway. Moreover, data shows that blockade of signals generated from BCR activation is not able to affect NFAT1/2 phosphorylation or translocation to the nucleus. These findings suggest that other mechanisms could be responsible for the pathway activation. In summary, evidence has accumulated showing that the NFAT family controls biological processes on and off the tumor, which should be carefully analyzed in the context of targeting for any future treatment of lymphoproliferative diseases.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

This work was supported by Ministero della Salute under Ricerca Finalizzata 2018–ID. RF-2018-12367072 project—P.I. MM and co-P.I. PA; AIRC under IG 2019–ID. 23088 project—P.I. MM.

REFERENCES

1. ten Hacken E, Burger JA. Microenvironment interactions and B-cell receptor signaling in Chronic Lymphocytic Leukemia: Implications for disease pathogenesis and treatment. *Biochim Biophys Acta - Mol Cell Res* (2016) 1863:401–13. doi: 10.1016/j.bbamcr.2015.07.009
2. Bosch F, Dalla-Favera R. Chronic lymphocytic leukaemia: from genetics to treatment. *Nat Rev Clin Oncol* (2019) 16:684–701. doi: 10.1038/s41571-019-0239-8
3. Schiattone L, Ghia P, Scarfò L. The evolving treatment landscape of chronic lymphocytic leukemia. *Curr Opin Oncol* (2019) 31:568–73. doi: 10.1097/CCO.0000000000000585
4. Shanafelt TD, Wang XV, Kay NE, Hanson CA, O'Brien S, Barrientos J, et al. Ibrutinib–Rituximab or Chemoimmunotherapy for Chronic Lymphocytic Leukemia. *N Engl J Med* (2019) 381:432–43. doi: 10.1056/nejmoa1817073
5. Moreno C, Greil R, Demirkan F, Tedeschi A, Anz B, Larratt L, et al. Ibrutinib plus obinutuzumab versus chlorambucil plus obinutuzumab in first-line treatment of chronic lymphocytic leukaemia (iLLUMINATE): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol* (2019) 20:43–56. doi: 10.1016/S1470-2045(18)30788-5
6. Rossi D, Gaidano G. Lymphocytosis and ibrutinib treatment of CLL. *Blood* (2014) 123(12):1772–4. doi: 10.1182/blood-2014-01-549493
7. Woyach JA, Furman RR, Liu T-M, Ozer HG, Zapatka M, Ruppert AS, et al. Resistance Mechanisms for the Bruton's Tyrosine Kinase Inhibitor Ibrutinib. *N Engl J Med* (2014) 370:2286–94. doi: 10.1056/nejmoa1400029
8. Kittai AS, Woyach JA. Resistance Mechanisms to Targeted Agents in Chronic Lymphocytic Leukemia. *Cancer J (United States)* (2019) 25:428–35. doi: 10.1097/PPO.0000000000000406
9. Sedlarikova L, Petrackova A, Papajik T, Turcsanyi P, Kriegova E. Resistance-associated mutations in chronic lymphocytic leukemia patients treated with novel agents. *Front Oncol* (2020) 10:894. doi: 10.3389/fonc.2020.00894
10. Müller MR, Rao A. NFAT, immunity and cancer: A transcription factor comes of age. *Nat Rev Immunol* (2010) 10:645–56. doi: 10.1038/nri2818
11. Okamura H, Aramburu J, García-Rodríguez C, Viola JPB, Raghavan A, Tahiliani M, et al. Concerted dephosphorylation of the transcription factor

NFAT1 induces a conformational switch that regulates transcriptional activity. *Mol Cell* (2000) 6:539–50. doi: 10.1016/S1097-2765(00)00053-8
12. Gwack Y, Feske S, Srikanth S, Hogan PG, Rao A. Signalling to transcription: Store-operated Ca2+ entry and NFAT activation in lymphocytes. *Cell Calcium* (2007) 42:145–56. doi: 10.1016/j.ceca.2007.03.007
13. Kar P, Nelson C, Parekh AB. Selective activation of the transcription factor NFAT1 by calcium microdomains near Ca2+ release-activated Ca2+ (CRAC) channels. *J Biol Chem* (2011) 286:14795–803. doi: 10.1074/jbc.M111.220582
14. Medyouf H, Ghysdael J. The calcineurin/NFAT signaling pathway: A novel therapeutic target in leukemia and solid tumors. *Cell Cycle* (2008) 7:297–303. doi: 10.4161/cc.7.3.5357
15. Hogan PG, Chen L, Nardone J, Rao A. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes Dev* (2003) 17:2205–32. doi: 10.1101/gad.1102703
16. Qin J-JJ, Nag S, Wang W, Zhou J, Zhang W-DD, Wang H, et al. NFAT as cancer target: Mission possible? *Elsevier* (2014) 1846(2):297–331. doi: 10.1016/j.bbcan.2014.07.009
17. Gwack Y, Sharma S, Nardone J, Tanasa B, Iuga A, Srikanth S, et al. A genome-wide Drosophila RNAi screen identifies DYRK-family kinases as regulators of NFAT. *Nature* (2006) 441:646–50. doi: 10.1038/nature04631
18. Okamura H, Garcia-Rodriguez C, Martinson H, Qin J, Virshup DM, Rao A. A Conserved Docking Motif for CK1 Binding Controls the Nuclear Localization of NFAT1. *Mol Cell Biol* (2004) 24:4184–95. doi: 10.1128/mcb.24.10.4184-4195.2004
19. Beals CR. Nuclear Export of NF-ATc Enhanced by Glycogen Synthase Kinase-3. *Science* (1997) 275:1930–3. doi: 10.1126/science.275.5308.1930
20. Miyakawa H, Woo SK, Dahl SC, Handler JS, Kwon HM. Tonicity-responsive enhancer binding protein, a Rel-like protein that stimulates transcription in response to hypertonicity. *Proc Natl Acad Sci USA* (1999) 96(5):2538–42. doi: 10.1073/pnas.96.5.2538
21. Woo SK, Lee S, Kwon HM. TonEBP transcriptional activator in the cellular response to increased osmolality. *Pflugers Arch Eur J Physiol* (2002) 444:579–85. doi: 10.1007/s00424-002-0849-2
22. Kim NH, Hong BK, Choi SY, Kwon HM, Cho CS, Yi EC, et al. Reactive oxygen species regulate context-dependent inhibition of NFAT5 target genes. *Exp Mol Med* (2013) 45:e32. doi: 10.1038/emm.2013.61
23. Peng SL, Gerth AJ, Ranger AM, Glimcher LH. NFATc1 and NFATc2 together control both T and B cell activation and differentiation. *Immunity* (2001) 14:13–20. doi: 10.1016/S1074-7613(01)00085-1
24. De La Pompa JL, Timmerman LA, Takimoto H, Yoshida H, Elia AJ, Samper E, et al. Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. *Nature* (1998) 392:182–6. doi: 10.1038/32419
25. Ranger AM, Grusby MJ, Hodge MR, Gravallese EM, De La Brousse FC, Hoey T, et al. The transcription factor NF-ATc is essential for cardiac valve formation. *Nature* (1998) 392:186–90. doi: 10.1038/32426
26. Yoshida H, Nishina H, Takimoto H, Marengère LEM, Wakeham AC, Bouchard D, et al. The transcription factor NF-ATc1 regulates lymphocyte proliferation and Th2 cytokine production. *Immunity* (1998) 8:115–24. doi: 10.1016/S1074-7613(00)80464-1
27. Berland R, Wortis HH. Normal B-1a cell development requires B cell-intrinsic NFATc1 activity. *Proc Natl Acad Sci USA* (2003) 100:13459–64. doi: 10.1073/pnas.2233620100
28. Bhattacharyya S, Deb J, Patra AK, Pham DAT, Chen W, Vaeth M, et al. NFATc1 affects mouse splenic B cell function by controlling the calcineurin-NFAT signaling network. *J Exp Med* (2011) 208:823–39. doi: 10.1084/jem.20100945
29. Lin H, Grosschedl R. Failure of B–cell differentiation in mice lacking the transcription factor EBF. *Nature* (1995) 376:263–7. doi: 10.1038/376263a0
30. Giampaolo S, Wójcik G, Klein-Hessling S, Serfling E, Patra AK. B cell development is critically dependent on NFATc1 activity. *Cell Mol Immunol* (2019) 16:508–20. doi: 10.1038/s41423-018-0052-9
31. Märklin M, Heitmann JS, Kauer J, Wirths S, Müller MR. Genetic loss of NFAT2 (NFATc1) impairs B cell development of B1 and B2 B cells. *Cell Immunol* (2020) 349:104048. doi: 10.1016/j.cellimm.2020.104048
32. Xanthoudakis S, Viola JPB, Shaw KTY, Luo C, Wallace JD, Bozza PT, et al. An enhanced imune respones in mice lackinh the transcription factor NFAT1. *Science* (1996) 272:892–5. doi: 10.1126/science.272.5263.892
33. Ranger AM, Hodge MR, Gravallese EM, Oukka M, Davidson L, Alt FW, et al. Delayed lymphoid repopulation with defects in IL-4-driven responses produced by inactivation of NF-ATc. *Immunity* (1998) 8:125–34. doi: 10.1016/S1074-7613(00)80465-3
34. Hodge MR, Ranger AM, De La Brousse FC, Hoey T, Grusby MJ, Glimcher LH. Hyperproliferation and dysregulation of IL-4 expression in NF-ATp-deficient mice. *Immunity* (1996) 4(4):397–405. doi: 10.1016/S1074-7613(00)80253-8
35. Barrington RA, Borde M, Rao A, Carroll MC. Involvement of NFAT1 in B cell self-tolerance. *J Immunol* (2006) 177:1510–5. doi: 10.4049/jimmunol.177.3.1510
36. Teixeira LK, Carrossini N, Sécca C, Kroll JE, DaCunha DC, Faget DV, et al. NFAT1 transcription factor regulates cell cycle progression and cyclin E expression in B lymphocytes. *Cell Cycle* (2016) 15:2346–59. doi: 10.1080/15384101.2016.1203485
37. Ranger AM, Oukka M, Rengarajan J, Glimcher LH. Inhibitory function of two NFAT family members in lymphoid homeostasis and Th2 development. *Immunity* (1998) 9:627–35. doi: 10.1016/S1074-7613(00)80660-3
38. Rengarajan J, Mittelstadt PR, Mages HW, Gerth AJ, Kroczek RA, Ashwell JD, et al. Sequential Involvement of NFAT and Egr Transcription Factors in FasL Regulation. *Immunity* (2000) 12:293–300. doi: 10.1016/S1074-7613(00)80182-X
39. Rengarajan J, Tang B, Glimcher LH. Nfatc2 and nfatc3 regulate th2 differentiation and modulate tcr-responsiveness of naïve th cells. *Nat Immunol* (2002) 3:48–54. doi: 10.1038/ni744
40. Samanta DN, Palmetshofer A, Marinkovic D, Wirth T, Serfling E, Nitschke L. B Cell Hyperresponsiveness and Expansion of Mature Follicular B Cells but Not of Marginal Zone B Cells in NFATc2/c3 Double-Deficient Mice. *J Immunol* (2005) 174:4797–802. doi: 10.4049/jimmunol.174.8.4797
41. Berga-Bolanos R, Alberdi M, Buxade M, Aramburu J, Lopez-Rodriguez C. NFAT5 induction by the pre-T-cell receptor serves as a selective survival signal in T-lymphocyte development. *Proc Natl Acad Sci* (2013) 110:16091–6. doi: 10.1073/pnas.1215934110
42. Go WY, Liu X, Roti MA, Liu F, Ho SN. NFATS/TonEBP mutant mice define osmotic stress as a critical feature of the lymphoid microenvironment. *Proc Natl Acad Sci USA* (2004) 101:10673–8. doi: 10.1073/pnas.0403139101
43. Buxadé M, Lunazzi G, Minguillón J, Iborra S, Berga-Bolaños R, del Val M, et al. Gene expression induced by Toll-like receptors in macrophages requires the transcription factor NFAT5. *J Exp Med* (2012) 209:379–93. doi: 10.1084/jem.20111569
44. Berga-Bolaños R, Drews-Elger K, Aramburu J, López-Rodríguez C. NFAT5 Regulates T Lymphocyte Homeostasis and CD24-Dependent T Cell Expansion under Pathologic Hypernatremia. *J Immunol* (2010) 185:6624–35. doi: 10.4049/jimmunol.1001232
45. Swerdlow SH, Campo E, Pileri SA, Lee Harris N, Stein H, Siebert R, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* (2016) 127:2375–90. doi: 10.1182/blood-2016-01-643569
46. Marafiot T, Pozzobon M, Hansmann M-L, Ventura R, Pileri SA, Roberton H, et al. The NFATc1 transcription factor is widely expressed in white cells and translocates from the cytoplasm to the nucleus in a subset of human lymphomas. *Br J Haematol* (2005) 128:333–42. doi: 10.1111/j.1365-2141.2004.05313.x
47. Akimzhanov A, Krenacs L, Schlegel T, Klein-Hessling S, Bagdi E, Stelkovics E, et al. Epigenetic changes and suppression of the nuclear factor of activated T cell 1 (NFATC1) promoter in human lymphomas with defects in immunoreceptor signaling. *Am J Pathol* (2008) 172:215–24. doi: 10.2353/ajpath.2008.070294
48. Pham LV, Tamayo AT, Yoshimura LC, Lin-Lee YC, Ford RJ. Constitutive NF-κB and NFAT activation in aggressive B-cell lymphomas synergistically activates the CD154 gene and maintains lymphoma cell survival. *Blood* (2005) 106:3940–7. doi: 10.1182/blood-2005-03-1167
49. Fu L, Lin-Lee YC, Pham LV, Tamayo A, Yoshimura L, Ford RJ. Constitutive NF-κB and NFAT activation leads to stimulation of the BLyS survival pathway in aggressive B-cell lymphomas. *Blood* (2006) 107:4540–8. doi: 10.1182/blood-2005-10-4042
50. Pham LV, Tamayo AT, Li C, Bueso-Ramos C, Ford RJ. An epigenetic chromatin remodeling role for NFATc1 in transcriptional regulation of growth and survival genes in diffuse large B-cell lymphomas. *Blood* (2010) 116:3899–906. doi: 10.1182/blood-2009-12-257378
51. Kondo E, Harashima A, Takabatake T, Takahashi H, Matsuo Y, Yoshino T, et al. NF-ATc2 induces apoptosis in Burkitt's lymphoma cells through

signaling via the B cell antigen receptor. *Eur J Immunol* (2003) 33:1–11. doi: 10.1002/immu.200390000
52. Lenz G, Wright GW, Emre NCT, Kohlhammer H, Dave SS, Davis RE, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc Natl Acad Sci USA* (2008) 105:13520–5. doi: 10.1073/pnas.0804295105
53. Pham LV, Bryant JL, Mendez R, Chen J, Tamayo AT, Xu-Monette ZY, et al. Targeting the hexosamine biosynthetic pathway and O-linked N-acetylglucosamine cycling for therapeutic and imaging capabilities in diffuse large B-cell lymphoma. *Oncotarget* (2016) 7:80599–611. doi: 10.18632/oncotarget.12413
54. Li L, Zhang J, Chen J, Xu-Monette ZY, Miao Y, Xiao M, et al. B-cell receptor–mediated NFATc1 activation induces IL-10/STAT3/PD-L1 signaling in diffuse large B-cell lymphoma. *Blood* (2018) 132:1805–17. doi: 10.1182/blood-2018-03-841015
55. Bucher P, Erdmann T, Grondona P, Xu W, Schmitt A, Schürch C, et al. Targeting chronic NFAT activation with calcineurin inhibitors in diffuse large B-cell lymphoma. *Blood* (2020) 135:121–32. doi: 10.1182/blood.2019001866
56. Muppidi JR. A role for NFAT signaling in ABC-DLBCL. *Blood* (2020) 135:81. doi: 10.1182/blood.2019004199
57. Lee N, Kim D, Kim W-U. Role of NFAT5 in the Immune System and Pathogenesis of Autoimmune Diseases. *Front Immunol* (2019) 10:270. doi: 10.3389/fimmu.2019.00270
58. Fabbri G, Dalla-Favera R. The molecular pathogenesis of chronic lymphocytic leukaemia. *Nat Rev Cancer* (2016) 16:145–62. doi: 10.1038/nrc.2016.8
59. Arnason JE, Brown JR. Targeting B Cell Signaling in Chronic Lymphocytic Leukemia. *Curr Oncol Rep* (2017) 19:1–13. doi: 10.1007/s11912-017-0620-7
60. Schuh K, Avots A, Tony HP, Serfling E, Kneitz C. Nuclear NF-ATp is a hallmark of unstimulated B cells from B-CLL patients. *Leuk Lymphoma* (1996) 23:583–92. doi: 10.3109/10428199609054868
61. Le Roy C, Deglesne PA, Chevallier N, Beitar T, Eclache V, Quettier M, et al. The degree of BCR and NFAT activation predicts clinical outcomes in chronic lymphocytic leukemia. *Blood* (2012) 120:356–65. doi: 10.1182/blood-2011-12-397158
62. Pei L, Choi JH, Liu J, Lee EJ, McCarthy B, Wilson JM, et al. Genome-wide DNA methylation analysis reveals novel epigenetic changes in chronic lymphocytic leukemia. *Epigenetics* (2012) 7:567–78. doi: 10.4161/epi.20237
63. Beekman R, Chapaprieta V, Russiñol N, Vilarrasa-Blasi R, Verdaguer-Dot N, Martens JHA, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. *Nat Med* (2018) 24:868–80. doi: 10.1038/s41591-018-0028-4
64. Wolf C, Garding A, Filarsky K, Bahlo J, Robrecht S, Becker N, et al. NFATC1 activation by DNA hypomethylation in chronic lymphocytic leukemia correlates with clinical staging and can be inhibited by ibrutinib. *Int J Cancer* (2018) 142:322–33. doi: 10.1002/ijc.31057
65. Abruzzo LV, Herling CD, Calin GA, Oakes C, Barron LL, Banks HE, et al. Trisomy 12 chronic lymphocytic leukemia expresses a unique set of activated and targetable pathways. *Haematologica* (2018) 103:2069–78. doi: 10.3324/haematol.2018.190132
66. Rawstron AC, Kreuzer KA, Soosapilla A, Spacek M, Stehlikova O, Gambell P, et al. Reproducible diagnosis of chronic lymphocytic leukemia by flow cytometry: An European Research Initiative on CLL (ERIC) & European Society for Clinical Cell Analysis (ESCCA) Harmonisation project. *Cytom Part B - Clin Cytom* (2018) 94:121–8. doi: 10.1002/cyto.b.21595
67. Kneitz C, Goller M, Tony HP, Simon A, Stibbe C, König T, et al. The CD23b promoter is a target for NF-AT transcription factors in B-CLL cells. *Biochim Biophys Acta - Mol Basis Dis* (2002) 1588:41–7. doi: 10.1016/S0925-4439(02)00114-X
68. Hubmann R, Schwarzmeier JD, Shehata M, Hilgarth M, Duechler M, Dettke M, et al. Notch2 is involved in the overexpression of CD23 in B-cell chronic lymphocytic leukemia. *Blood* (2002) 99:3742–7. doi: 10.1182/blood.v99.10.3742
69. Teutsch M, Higer M, Wang D, Wortis HW. Induction of CD5 on B and T cells is suppressed by cyclosporin A, FK-520 and rapamycin. *Int Immunol* (1995) 7:381–92. doi: 10.1093/intimm/7.3.381
70. Berland R, Wortis HH. An NFAT-dependent enhancer is necessary for anti-IgM-mediated induction of murine CD5 expression in primary splenic B cells. *J Immunol* (1998) 161:277–85.
71. Märklin M, Heitmann JS, Fuchs AR, Truckenmüller FM, Gutknecht M, Bugl S, et al. NFAT2 is a critical regulator of the anergic phenotype in chronic lymphocytic leukaemia. *Nat Commun* (2017) 8:1–14. doi: 10.1038/s41467-017-00830-y
72. Garaud S, Morva A, Lemoine S, Hillion S, Bordron A, Pers J-O, et al. CD5 Promotes IL-10 Production in Chronic Lymphocytic Leukemia B Cells through STAT3 and NFAT2 Activation. *J Immunol* (2011) 186:4835–44. doi: 10.4049/jimmunol.1003050
73. Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowitz M, Ruppert AS, et al. DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia. *Nat Genet* (2016) 48:253–64. doi: 10.1038/ng.3488
74. Bichi R, Shinton SA, Martin ES, Koval A, Calin GA, Cesari R, et al. Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression. *Proc Natl Acad Sci USA* (2002) 99:6955–60. doi: 10.1073/pnas.102181599
75. Muzio M, Apollonio B, Scielzo C, Frenquelli M, Vandoni I, Boussiotis V, et al. Constitutive activation of distinct BCR-signaling pathways in a subset of CLL patients: a molecular signature of anergy. *Blood* (2008) 112:188–95. doi: 10.1182/blood-2007-09-111344
76. Müller DJ, Wirths S, Fuchs AR, Märklin M, Heitmann JS, Sturm M, et al. Loss of NFAT2 expression results in the acceleration of clonal evolution in chronic lymphocytic leukemia. *J Leukoc Biol* (2019) 105:531–8. doi: 10.1002/JLB.2AB0218-076RR
77. Apollonio B, Scielzo C, Bertilaccio MTS, Ten Hacken E, Scarfò L, Ranghetti P, et al. Targeting B-cell anergy in chronic lymphocytic leukemia. *Blood* (2013) 121:3879–88. doi: 10.1182/blood-2012-12-474718
78. Lee JU, Kim LK, Choi JM. Revisiting the concept of targeting NFAT to control T cell immunity and autoimmune diseases. *Front Immunol* (2018) 9:2747. doi: 10.3389/fimmu.2018.02747
79. Wright SJ, Keating MJ. Cyclosporine a in chronic lymphocytic leukemia: Dual anti-leukemic and immunosuppressive role? *Leuk Lymphoma* (1995) 20:131–6. doi: 10.3109/10428199509054764
80. Kyasa MJ, Parrish RS, Schichman SA, Zent CS. Autoimmune cytopenia does not predict poor prognosis in chronic lymphocytic eukemia/small lymphocytic lymphoma. *Am J Hematol* (2003) 74:1–8. doi: 10.1002/ajh.10369
81. Yu H, Van Berkel TJC, Biessen EAL. Therapeutic potential of VIVIT, a selective peptide inhibitor of nuclear factor of activated T cells, in cardiovascular disorders. *Cardiovasc Drug Rev* (2007) 25:175–87. doi: 10.1111/j.1527-3466.2007.00011.x
82. Aramburu J, Yaffe MB, López-Rodríguez C, Cantley LC, Hogan PG, Rao A. Affinity-driven peptide selection of an NFAT inhibitor more selective than cyclosporin A. *Science* (1999) 285:2129–33. doi: 10.1126/science.285.5436.2129
83. Nishii K, Katayama N, Mitani H, Matsumoto T, Miwa H, Kita K, et al. Effects of cyclosporin A on refractory B-cell chronic lymphocytic leukemia. *Int J Hematol* (2000) 71:59–65.
84. Chen BL, Li Y, Xu S, Nie Y, Zhang J. NFAT5 Regulated by STUB1, Facilitates Malignant Cell Survival and p38 MAPK Activation by Upregulating AQP5 in Chronic Lymphocytic Leukemia. *Biochem Genet* (2021), 123456789. doi: 10.1007/s10528-021-10040-3
85. Bachireddy P, Wu CJ. Arresting the Inflammatory Drive of Chronic Lymphocytic Leukemia with Ibrutinib. *Clin Cancer Res* (2016) 22:1547–9. doi: 10.1158/1078-0432.CCR-15-3106
86. Pan M-G, Xiong Y, Chen F. NFAT gene family in inflammation and cancer. *Curr Mol Med* (2013) 13:543–54. doi: 10.2174/1566524011313040007
87. Zanoni I, Granucci F. Regulation and dysregulation of innate immunity by NFAT signaling downstream of pattern recognition receptors (PRRs). *Eur J Immunol* (2012) 42:1924–31. doi: 10.1002/eji.201242580
88. Caligaris-Cappio F. Inflammation, the microenvironment and chronic lymphocytic leukemia. *Haematologica* (2011) 96:353–5. doi: 10.3324/haematol.2010.039446
89. Schulz A, Toedt G, Zenz T, Stilgenbauer S, Lichter P, Seiffert M. Inflammatory cytokines and signaling pathways are associated with survival of primary chronic lymphocytic leukemia cells in vitro: a dominant role of CCL2. *Haematologica* (2011) 96:408–16. doi: 10.3324/haematol.2010.031377

90. Carbone A, Tripodo C, Carlo-Stella C, Santoro A, Gloghini A. The Role of Inflammation in Lymphoma. *Adv Exp Med Biol* (2014) 816:315–33. doi: 10.1007/978-3-0348-0837-8_12

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

ORIGINAL RESEARCH
published: 13 May 2021
doi: 10.3389/fonc.2021.646971

Proteolysis Targeting Chimeras for BTK Efficiently Inhibit B-Cell Receptor Signaling and Can Overcome Ibrutinib Resistance in CLL Cells

Yamit Shorer Arbel[1†], Ben-Zion Katz[1,2†], Ronen Gabizon[3], Amit Shraga[3], Yotam Bronstein[1], Talia Kamdjou[2], Anat Globerson Levin[4], Chava Perry[1,2], Irit Avivi[1,2], Nir London[3] and Yair Herishanu[1,2]*

[1] Sackler Faculty of Medicine, Tel Aviv University, Tel-Aviv, Israel, [2] Department of Hematology, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, [3] Department of Organic Chemistry, The Weizmann Institute of Science, Rehovot, Israel, [4] Immunology Research Laboratory, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel

OPEN ACCESS

Edited by:
Franco Fais,
University of Genoa, Italy

Reviewed by:
Varsha Gandhi,
University of Texas MD Anderson Cancer Center, United States
Tanja Nicole Hartmann,
University of Freiburg Medical Center, Germany

***Correspondence:**
Yair Herishanu
yairh@tlvmc.gov.il

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: 28 December 2020
Accepted: 30 March 2021
Published: 13 May 2021

Citation:
Shorer Arbel Y, Katz B-Z, Gabizon R, Shraga A, Bronstein Y, Kamdjou T, Globerson Levin A, Perry C, Avivi I, London N and Herishanu Y (2021) Proteolysis Targeting Chimeras for BTK Efficiently Inhibit B-Cell Receptor Signaling and Can Overcome Ibrutinib Resistance in CLL Cells. Front. Oncol. 11:646971. doi: 10.3389/fonc.2021.646971

Proteolysis targeting chimeras (PROTACs) are small molecules that form ternary complexes between their target and E3 ligase, resulting in ubiquitination and proteasomal degradation of the target protein. Using our own designed Bruton's tyrosine kinase (BTK) PROTAC compounds, we show herein efficient BTK degradation in chronic lymphocytic leukemia (CLL) cells. The reversible non-covalent compound (NC-1) was the most potent and therefore we focused on this PROTAC to investigate its subsequent effects on the BCR pathway. NC-1 decreased baseline BTK phosphorylation as well as activation of BTK and other signaling molecules downstream of the BCR pathway, following IgM engagement. These effects were also obtained in samples from CLL patients with clinical resistance to ibrutinib and mutations at C481. NC-1 treatment further decreased baseline CD69 surface levels, completely abrogated its upregulation following IgM activation, decreased CLL cells migration toward SDF-1 and overcame stromal anti-apoptotic protection. In conclusion, our results indicate that targeting BTK using the PROTAC strategy could be a potential novel therapeutic approach for CLL.

Keywords: PROTACs (proteolysis targeting chimeras), BTK - Bruton's tyrosine kinase, CLL (Chronic Lymphocytic Leukemia), ibrutinib, BCR (B cell receptor) signaling

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is the most common leukemia in the western world (1). Although in recent years there has been considerable progress in the treatment options for CLL, a definitive cure is still only achievable with allogeneic stem cell transplantation (2). When considering the essential role of BCR signaling in CLL pathogenesis (3), this pathway has become a target for anti-CLL therapy. Small molecules directed against kinases of the BCR pathway show impressive clinical activity. One of these molecules is ibrutinib, an irreversible inhibitor of Bruton's tyrosine kinase (BTK) which has a critical role in the amplification of the BCR

signal. Nevertheless, due to side effects and resistance that emerges over-time, there is a need to develop novel approaches to target the BCR pathway (4–7).

Proteolysis-targeting chimeras (PROTACs), a novel strategy that utilizes the intracellular ubiquitin-proteasome system to induce targeted protein degradation, are receiving much attention in the field of targeted therapies. This technology is based on using hetero-bifunctional molecules that direct a ligand to bind with the target protein, linked with another ligand to recruit an E3 ubiquitin ligase. When the ternary complex (target-PROTAC-E3) is formed, the recruited E3 employs an E2 ubiquitin-conjugating enzyme to transfer ubiquitin to the surface of the targeted protein. This process leads to proteasomal degradation of the target protein (8–12).

PROTACs have a number of advantages over standard chemical inhibitors, including increased selectivity, inhibition of all target protein functions (13–16), longer lasting effects due to the need for a new synthesis of target (17) and induction of degradation by sub-stoichiometric concentrations of PROTAC (18). BTK is an established target for non-covalent PROTACs (19–24). Our previous results showed efficient BTK degradation with reversible covalent PROTACs, as well as their non-covalent and irreversible counterparts (25).

In this study we examined the effect of the PROTACs on BCR signaling, activation, migration and apoptosis in CLL cells. We show efficient inhibition of the BCR signaling pathway while using PROTACs in both wild-type (WT) and BTK mutated CLL cells. These results provide a basis for further preclinical study of BTK PROTACs as a novel strategy for CLL therapy.

MATERIALS AND METHODS

Patients and Samples

Cells were obtained from peripheral blood samples donated by patients fulfilling the standard criteria for CLL after signing informed consent approved by the Tel-Aviv Sourasky Medical's Institutional Review Board according to the Helsinki Accords (**Table 1**). Peripheral blood mononuclear cells (PBMC) were isolated by Ficoll density-gradient centrifugation. Viable frozen cells were kept in FCS containing 10% DMSO and stored in liquid nitrogen. Before use, frozen cells were thawed and cultured at 37°C, 5% CO2, in RPMI medium supplemented with 10% FCS, penicillin, streptomycin, and L-glutamine. The samples used contained more than 90% CLL cells.

Antibodies and Reagents

ERK1/2, Phospho-ERK1/2 (Thr202/Tyr204), Akt (pan), phospho-Akt (S473),PLC γ2, pPLC γ2(Tyr1217),BTK, Phospho-BTK (Tyr223),Lyn(5G2), cleaved PARP (Asp214), CD79a, phospho-CD79a (Tyr182), Syk,phospho-Syk (Tyr525/526), SHIP1 and phospho-SHIP1 (Tyr1020) antibodies were from Cell Signaling Technology (Beverly, MA). Anti-SRC family (phospho Y418)-Phospho-Lyn (Y396) was obtained from Abcam (Cambridge,UK).Purified anti-human actin antibody was obtained from MP Biomedicals (Illkirch,France). Goat anti Rabbit IgG (H+L)-HRP conjugate and Goat anti Mouse IgG (H+L)-HRP conjugate and Goat F(ab')2 anti-human IgM or IgG were from Jackson Immunoresearch Laboratories, (West Grove, PA). Dynabeads Human T-

TABLE 1 | Patient characteristics.

Patient	Gender/ Age (y)	Binet stage	ALC(x10^9/L)	IGHV mutational status	FISH/TP53
CLL_01	M/67	A	102	M-IGHV	Del13q
CLL_02	M/72	B	203	UM-IGHV	Del11q
CLL_03	F/61	B	193	UM-IGHV	Del13q
CLL_04	F/64	C	194	M-IGHV	Del13q
CLL_05	M/76	C	117	M-IGHV	Del17p
CLL_06	M/71	A	70	UM-IGHV	Del17p/TP53mut
CLL_07	M/56	B	76	UM-IGHV	Del17p/TP53mut
CLL_08	M/43	A	120	UM-IGHV	Del17p
CLL_09	M/57	B	66	UM-IGHV	Del11q
CLL_10	M/57	B	144	UM-IGHV	Del11q
CLL_11	M/70	B	87	UM-IGHV	Del11q
CLL_12	F/71	B	121	UM-IGHV	Del13q
CLL_13	M/52	A	146	UM-IGHV	Del11q
CLL_14	F/72	A	190	M-IGHV	Del13q
CLL_15	M/54	B	61	M-IGHV	Normal
CLL_16	M/76	A	59	M-IGHV	Del13q
CLL_17	M/67	B	88	UM-IGHV	Normal
CLL_18	M/60	B	219	UM-IGHV	Del11q
CLL_19	M/64	C	33	UM-IGHV	Del11q/TP53mut
C481Y-mutation					
CLL_20	M/60	C	71	UM-IGHV	Del17p/TP53mut
C481S-mutation					
CLL_21	F/54	C	40	UM-IGHV	Del17p/TP53mut
C481F-mutation					
CLL_22	F/70	C	70	ND	ND

M, male; F, female; y, years; ALC, absolute lymphocyte count; M-IGHV, mutated IGHV; UM-IGHV-IGHV, unmutated IGHV; TP53mut, TP53 gene mutated; ND-no data.

Activator CD3/CD28 were obtained from Thermo Scientific (Rockford, IL). All antibodies utilized in the study were used in concentrations according to the manufacturer's instructions. Ficoll-Paque PLUS from GE healthcare (Uppsala, Sweden), dimethyl sulfoxide (DMSO) from Merck (Darmstadt, Germany), RPMI, fetal calf serum (FCS), Dulbecco's phosphate buffered saline (PBS), L-glutamine and penicillin-streptomycin were from Biological Industries (Beit-Haemek, Israel). BTK PROTACs RC-2, IR-2 and NC-1 (**Supplementary Figure 1**) were designed as previously described (25).

Targeting BTK in CLL Cells

CLL cells were incubated with PROTACs (RC-2, IR-2 and NC-1) at the indicated doses and time intervals at 37°C. The PROTACs were dissolved in DMSO, and controls were treated with DMSO accordingly.

Western Blotting

CLL cells were lysed in RIPA lysis buffer (Cell Signaling Technology, Beverly, MA) containing phosphatase inhibitor cocktail 2 and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO). Extract from cell lysates were separated on 4–15% Criterion™ TGX™ Precast Midi Protein Gel (Bio-Rad Laboratories) and transferred electrophoretically to nitrocellulose membrane (Bio-Rad Laboratories). The membranes were incubated with the designated antibodies and HRP conjugated secondary antibodies according to the manufacturer's instructions. Bands were detected using MyECL Imager (Thermo Scientific, Rockford, IL).

Flow Cytometry

For activation marker analysis, CLL and normal B-cells ($5x10^6$ cells/mL, 500,000 cells per tube) were stained with APC Mouse Anti-human CD19 and PE Mouse Anti-human CD69 (BD Biosciences, CA, USA) and incubated for 30 minutes on ice. For normal T cell activation analysis, peripheral blood mononuclear cells were stained with APC Mouse anti-human CD3 and PE Mouse anti-human CD69 (BD Biosciences, CA, USA). Isotype controls were APC Mouse IgG1, κ and PE Mouse IgG1, κ (BD Biosciences, CA, USA). Cells were washed and then suspended in 0.5 mL PBS/1% FCS. For cell viability and apoptosis analysis, CLL cells ($5x10^6$ cells/mL, 500,000 cells per tube) were stained with the Annexin V/propidium Iodide MEBCYTO® Apoptosis Kit (MBL, Nagoya, Japan), according to the manufacturer's instructions. In both assays samples were acquired by BD FACSCanto II and analyzed using BD FACSDiva software.

Migration Assay

Peripheral blood CLL cells were cultured in 6-well dishes ($5x10^6$ cells/mL in RPMI 10% FCS) and incubated with 100 nM NC-1 BTK PROTAC for 18 hours.DMSO treated cells served as controls. A total of 100 µL, containing 5×10^5 cells, was added to the top chamber of a 6.5-mm diameter Transwell culture inserts (Costar, Cambridge, MA) with a pore size of 5 µm. Filters then were transferred to wells containing medium with 100 ng/mL SDF-1α (Merck). Wells containing medium without SDF-1α served as a negative control. The chambers were incubated for 2 hours at 37°C in 5% CO2. After this incubation, the cells in the lower chamber were suspended and divided into aliquots for counting. The experiments were performed in triplicates.

Co-Culture Assay

CLL cells ($10x10^6$ cells/mL) were co-cultured with HS-5 (ATCC® CRL-11882™) in 20:1 ratio in 6-well dishes. The medium for co-culture was DMEM (ATCC® 30-2002™) supplemented with 5% FCS and 1% penicillin-streptomycin. The cells were incubated with 100nM NC-1 PROTAC for 48 hours at 37°C. Proteins were extracted from cells treated with PROTAC and controls and analyzed by Western blotting.

Statistical Analysis

In order to compare between the two paired groups, within each type of experiment, the Student's t test was applied to compare the means of normal distributed dependent variables and the Wilcoxon Signed-rank test was applied in order to compare the distribution of non-parametric dependent variables. All statistical analyses were performed using GraphPad Prism 8.0 software (GraphPad Software, San Diego, CA, USA). A P-value of <0.05 was considered as statistically significant.

RESULTS

Efficient PROTAC-Mediated BTK Degradation in CLL Cells

We evaluated the ability of reversible covalent (RC-2), irreversible covalent (IR-2) and reversible non-covalent (NC-1) compounds (**Supplementary Figure 1**) to induce BTK degradation in CLL cell. The cells were treated for 18 hours based on our previous results using the three compounds in Ramos cells (25), as well as a time course experiment in CLL cells treated with 100 nM NC-1 PROTAC. Treatment with NC-1 decreased BTK in a time dependent manner and completely abolished BTK after 18 hours of incubation (**Supplementary Figure 2**). We also measured the expression levels of other proteins of the BCR pathway to test off-target effects. Proteins were extracted from cells treated with PROTACs and controls and analyzed by Western blotting. In our previous report (25), we showed that the three compounds decreased BTK levels, an effect that was more pronounced at the higher doses of 100-1000 nM. The NC-1 compound completely abolished BTK at ≥100 nM, and had a higher degradation potency compared to RC-2 and IR-2. In this work, total Lyn and Syk levels were also evaluated to confirm specificity of the PROTACs to BTK. Consistent with our previous report in the Ramos cell line, NC-1 and IR-2 also reduced Lyn levels, which is compatible with the known off-target effect of ibrutinib (26) (**Figures 1A, C**). Then, the effect of the most potent PROTAC, NC-1, at 100 nM was evaluated in 8 patients with different chromosomal abnormalities and/or *TP53* mutation. In all samples, the NC-1 compound led to BTK as well as Lyn degradation. Total Syk and

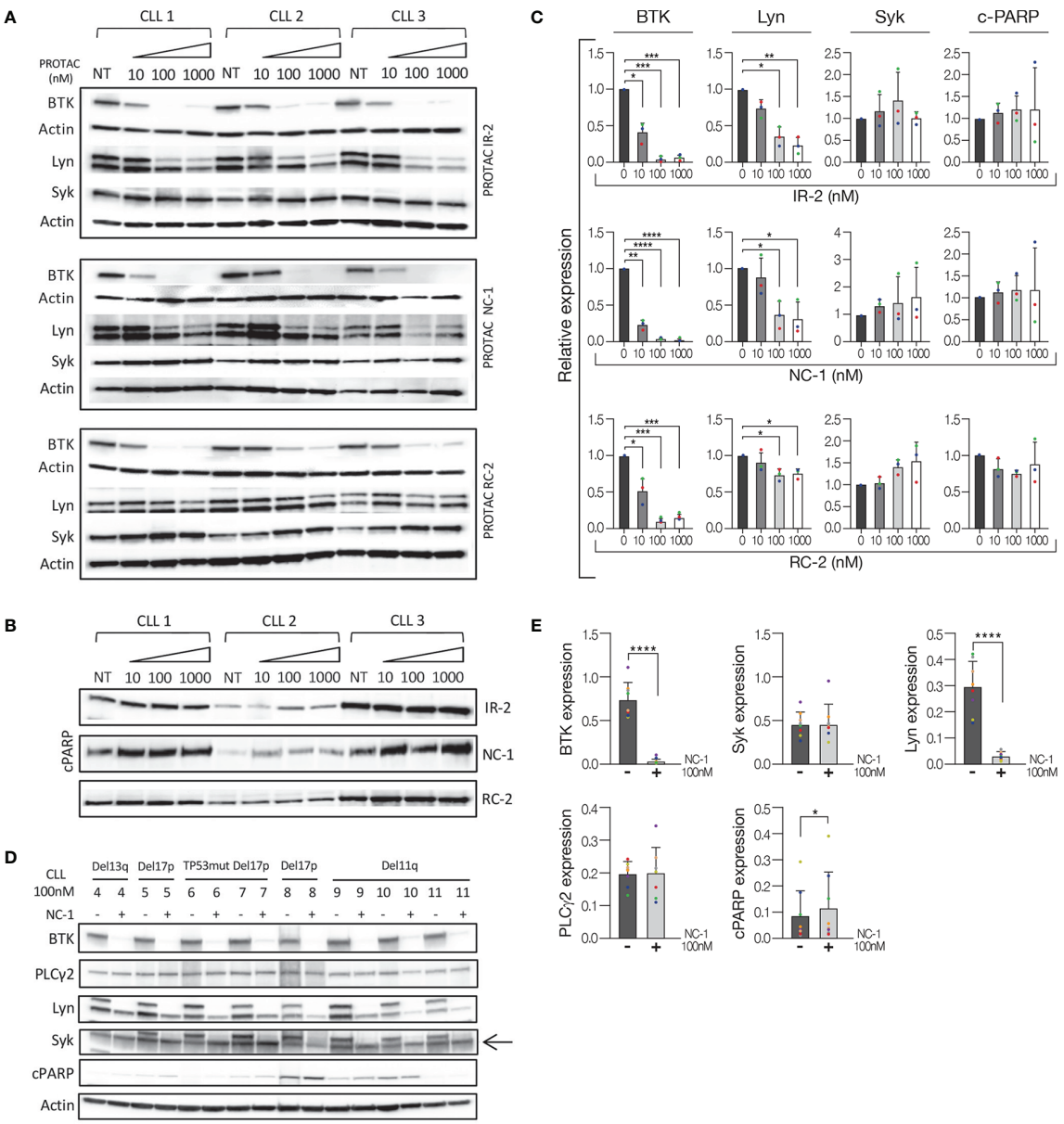


FIGURE 1 | Efficient BTK degradation in CLL cells. Peripheral blood CLL cells of three patients were cultured in 6-well dishes (20x10^6 cells/mL in RPMI 10% FCS) and incubated with BTK PROTACs in different concentrations (1000 nM, 100 nM and 10 nM) for 18 hours at 37° in a humidified 5% CO2 atmosphere. DMSO treated cells served as controls. Then, proteins were extracted and analyzed by Western blot. **(A)** A Western blot analysis showing total BTK, Lyn and Syk levels. Actin was used to verify equal loading. **(B)** Cleaved PARP levels in cells treated with different types and concentrations of PROTAC as a marker to measure apoptosis levels. **(C)** Quantification of BTK, Lyn, Syk and cleaved PARP levels in A and B by normalization to actin using myImageAnalysis^{TM} Software (n=3). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001. **(D, E)** Peripheral blood CLL cells of 8 patients with various genetic abnormalities were cultured in 6-well dishes (20x10^6 cells/mL in RPMI 10% FCS) and incubated with 100nM NC-1 for 18 hours at 37° in a humidified 5% CO2 atmosphere. DMSO treated cells served as controls. Then, proteins were extracted and analyzed by Western blot analysis. Levels of BTK, PLCγ2, Lyn, Syk and cleaved PARP (cPARP) are indicated (n=8). *p<0.05, ****p<0.0001.

PLCγ2 levels were not affected by this treatment (**Figures 1D, E**). In order to analyze the effect of the PROTACs on cell viability we tested cleaved PARP levels as a measure of apoptosis. Cleaved-PARP levels show a mild increase in apoptosis after 18 hours of PROTAC treatment (**Figures 1B–E**).

In order to validate the mechanism of PROTAC-mediated BTK degradation, CLL cells were treated with bortezomib, a proteasome inhibitor (27), for 1 hour before treatment with 100 nM NC-1. We analyzed BTK levels after additional 4 hours. PROTAC treatment was assessed for 4 hours due to the

concomitant presence of bortezomib, which could cause cell death. Compatible with our previous results in Mino cells (25), bortezomib significantly inhibited degradation, suggesting proteasome-mediated degradation (**Supplementary Figure 3**).

BTK PROTAC Leads to BCR Signaling Pathway Inhibition

After demonstrating that the NC-1 PROTAC leads to effective BTK degradation, its effect on downstream elements in the BCR signaling pathway was tested. Peripheral blood CLL cells were treated with NC-1 or ibrutinib, and then activated with goat F (ab')2 anti-human IgM. Cells that were not treated or stimulated served as controls. Western blot analysis revealed that pre-treatment with NC-1 resulted in abolishment of BTK phosphorylation, and partial inhibition of phosphorylation increase in Akt, ERK and PLCγ2 following IgM cross-linking (**Figures 2A, B**). The inhibitory effect of ibrutinib on the BCR signaling was less prominent compared to that of NC-1 (**Figures 2A, B**). Since we have previously shown that Lyn is an off target of NC-1, we also analyzed the effect of this PROTAC on phosphorylation of proximal BCR elements. Treatment with 100nM NC-1 compound decreased Lyn levels, as well as phosphorylation of this protein and of CD79a after BCR activation (**Figures 2C, D**). The effect of NC-1 on Syk and Ship1 phosphorylation was heterogeneous and statistically non-significant, however in some samples (n=4) treatment with NC-1 clearly inhibited Syk phosphorylation after BCR activation (**Figures 2C, D**). The effect of ibrutinib on phosphorylation of the upstream signaling molecules of BCR pathway was not statistically significant (**Figures 2C, D**).

Efficient BCR Signaling Pathway Inhibition in Ibrutinib Resistant CLL Cells

In the next experiments, we analyzed the effect of the PROTACs on the BCR signaling pathway in ibrutinib resistant cells. Ibrutinib binds BTK at the cysteine 481 residues and mutations at this position have been identified as the most frequent mutations in patients with CLL who develop clinical resistance. Peripheral blood CLL cells from patients with resistance to ibrutinib were treated with PROTACs or ibrutinib and then activated with goat F(ab')2 anti-human IgM. Cells that were not treated or stimulated served as controls. Western blot analysis shows a substantial decrease in BTK levels in a dose-dependent manner in ibrutinib resistant cells (C481Y) treated with PROTACs (**Figure 3A**). In order to prove ibrutinib resistance, we analyzed its effect on phosphorylation levels of BTK in cells before and after acquiring the BTK C481Y mutation. It is important to emphasize that the mutations in BTK are heterozygous and in all experiments the cells from resistant patients were collected while patients continued ibrutinib treatment, and before further line of treatment. Because BTK mutations at C481 disrupt the covalent binding of ibrutinib to BTK, the cells treated with ibrutinib were washed before activation. As shown in **Figure 3B**, the relative decrease in pBTK levels after the cells were treated with ibrutinib was more prominent before the cells acquired the BTK C481Y mutation. Treatment with PROTACs led to a decrease in BTK and pBTK levels in cells resistant to ibrutinib (**Figure 3B**). We also show that there is no significant difference in the levels of BTK and pBTK between cells treated with PROTACs and washed before activation and those that were not washed (**Figure 3B**). The effect of the PROTAC NC-1 on BTK levels was also evaluated in cells with other BTK mutations, C481S and C481F. **Figure 3C** demonstrates BTK degradation in these ibrutinib resistant cells. Next, we analyzed the effect of NC-1 on downstream BCR signaling elements in cells from patients with ibrutinib clinical resistance. A representative Western blot analysis shows inhibition of the phosphorylation of BTK, Akt and ERK in cells treated with the PROTAC, an effect that was not observed when cells were treated with ibrutinib (**Figures 3D, E**). These results demonstrate that the PROTAC appears to be effective in degradation of BTK protein as well as inhibiting the downstream elements of the BCR signaling in BTK mutated cells.

BTK PROTAC Blocks CLL Cell Activation in Response to BCR Engagement

B-cell activation is generally accompanied by upregulation of cell surface expression of certain functional molecules. CLL cells present the phenotype of activated B cells based on the overexpression of the activation markers such as CD69, an early activation marker (28). CLL cells were treated with 100 nM NC-1, and then surface CD69 levels were determined by flow cytometry. As we expected, activation of the B cells with F(ab') 2 anti-human IgM caused an increase in CD69 expression (**Figures 4A, B**). In cells treated with NC-1 there was a decrease in CD69 expression both at baseline and after activation (**Figures 4A, B**). NC-1 PROTAC treatment also decreased CD69 expression in BTK C481S mutated cells, whereas ibrutinib did not alter CD69 expression at baseline or after activation (**Figure 4C**). Our results show that PROTAC treatment blocks BCR-mediated activation in both WT and mutated BTK CLL cells. The effect of NC-1 compound on activation marker expression was also tested in normal B and T cells. The PROTAC decreased CD69 surface levels in both normal B and T cells after activation. However, there was no significant effect on this marker in resting cells treated with 100 nM NC-1 compound (**Figures 4D, E**).

BTK PROTAC Induces Apoptosis, Inhibits Migration and Overcomes the Protective Effect of Stromal Cells on CLL Cells

After demonstrating the efficacy of PROTAC in suppressing elements in the BCR pathway and decreasing activation marker expression, we analyzed its effect on apoptosis. For this purpose, CLL cells were treated with 100 nM NC-1 PROTAC or left untreated. The concentrations of PROTAC used in the experiments are equivalent to the plasma therapeutic levels of ibrutinib in patients with CLL (29). Cells viability was determined by flow cytometry immediately after thawing, after 48 and 96 hours of incubation, using an Annexin V/PI apoptosis detection kit. As we expected, there was a decrease in cell viability over time. This effect was approximately 10% higher in cells that were treated with NC-1 BTK PROTAC (**Figures 5A, B**). The differences in cell viability between the treated cells and the

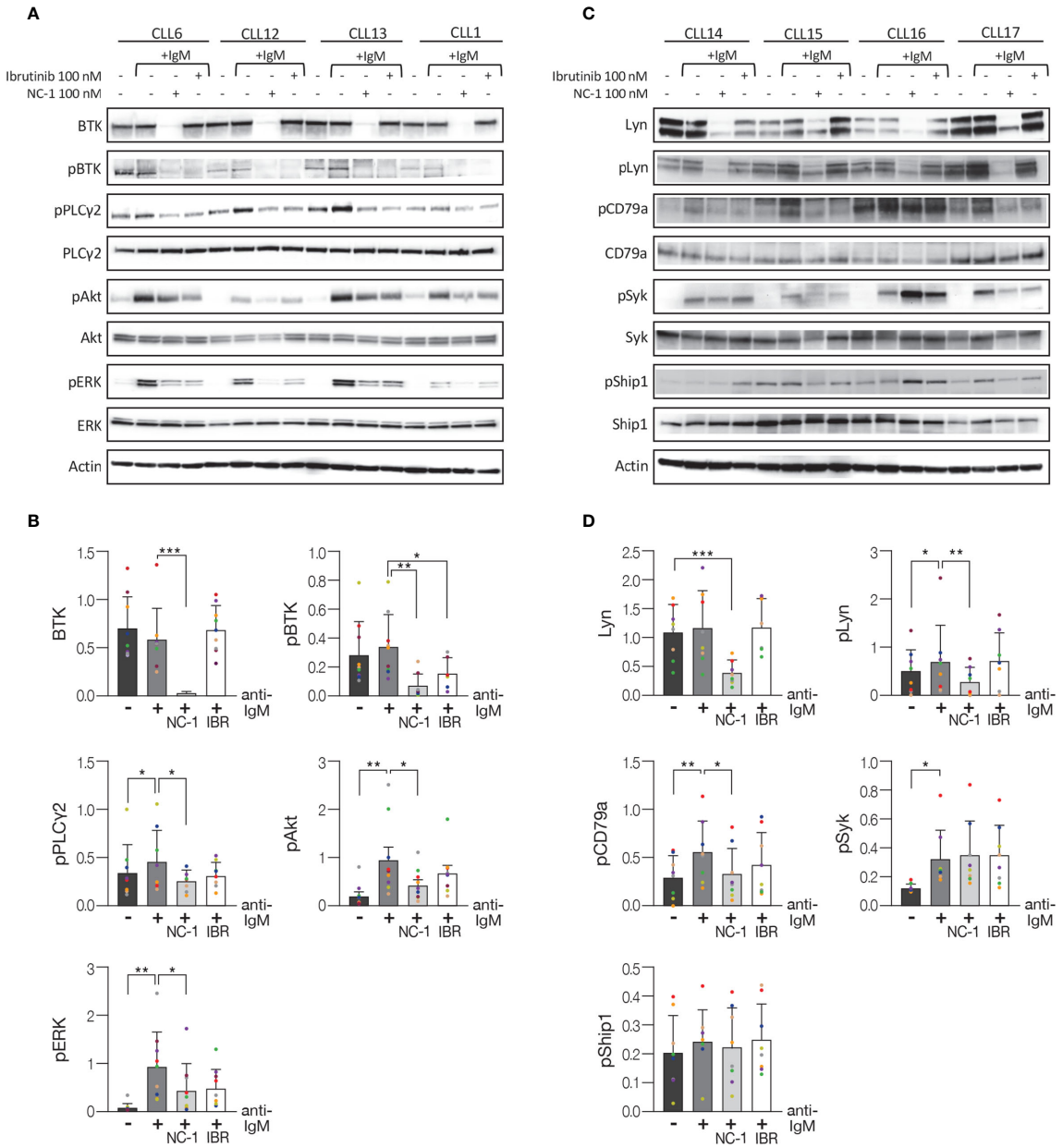


FIGURE 2 | BTK PROTAC leads to BCR signaling pathway inhibition. Peripheral blood CLL cells were cultured in 6-well dishes (20x10^6 cells/mL in RPMI 10% FCS) and incubated with 100 nM NC-1 for 18 hours or with ibrutinib (IBR) for 1 hour, at 37° in a humidified 5% CO2 atmosphere. DMSO treated cells served as controls. Following treatment, cells were incubated with goat F (ab') 2 anti-human IgM (10 μg/mL) for 15 minutes or left unstimulated. Cells treated with ibrutinib were washed before activation. Then, proteins were extracted and analyzed by Western blot analysis. **(A)** A representative Western blot showing BTK (Tyr223), PLCγ2 (Tyr1217), Akt (S473) and ERK (T202/Y204) phosphorylation, as well as total amount of these proteins. Actin was used to verify equal loading. **(B)** Quantification of BTK, pBTK, pPLCγ2, pAkt and pERK levels in A by normalization to actin using myImageAnalysis™ Software (n=9). *p<0.05, **p<0.01, ***p<0.001. **(C)** A representative Western blot showing pLyn (Tyr396), pCD79a (Tyr182), pSyk (Tyr525/526) and pShip1 (Tyr1020) levels,as well as total amount of these proteins. Actin was used to verify equal loading. **(D)** Quantification of Lyn, pLyn, pCD79a, pSyk and pShip1 levels in C by normalization to actin using myImageAnalysis™ Software (n=8). *p<0.05, **p<0.01, ***p<0.001.

control were statistically significant after 48 hours of incubation (P<0.049). These results confirm that BTK inhibition in the therapeutic range induces only mild apoptosis in CLL.

BTK has a role in mediating signal transduction that modulates migration and adhesion of B-cells to the tissue microenvironment, promoting cell survival and proliferation (30, 31). Given that fact, we analyzed the effect of NC-1 PROTAC on CLL cells migration toward stromal cell-derived factor-1 (SDF-1). It is already known that CLL cells express high levels of the chemokine receptor CXCR4 and that stromal cells

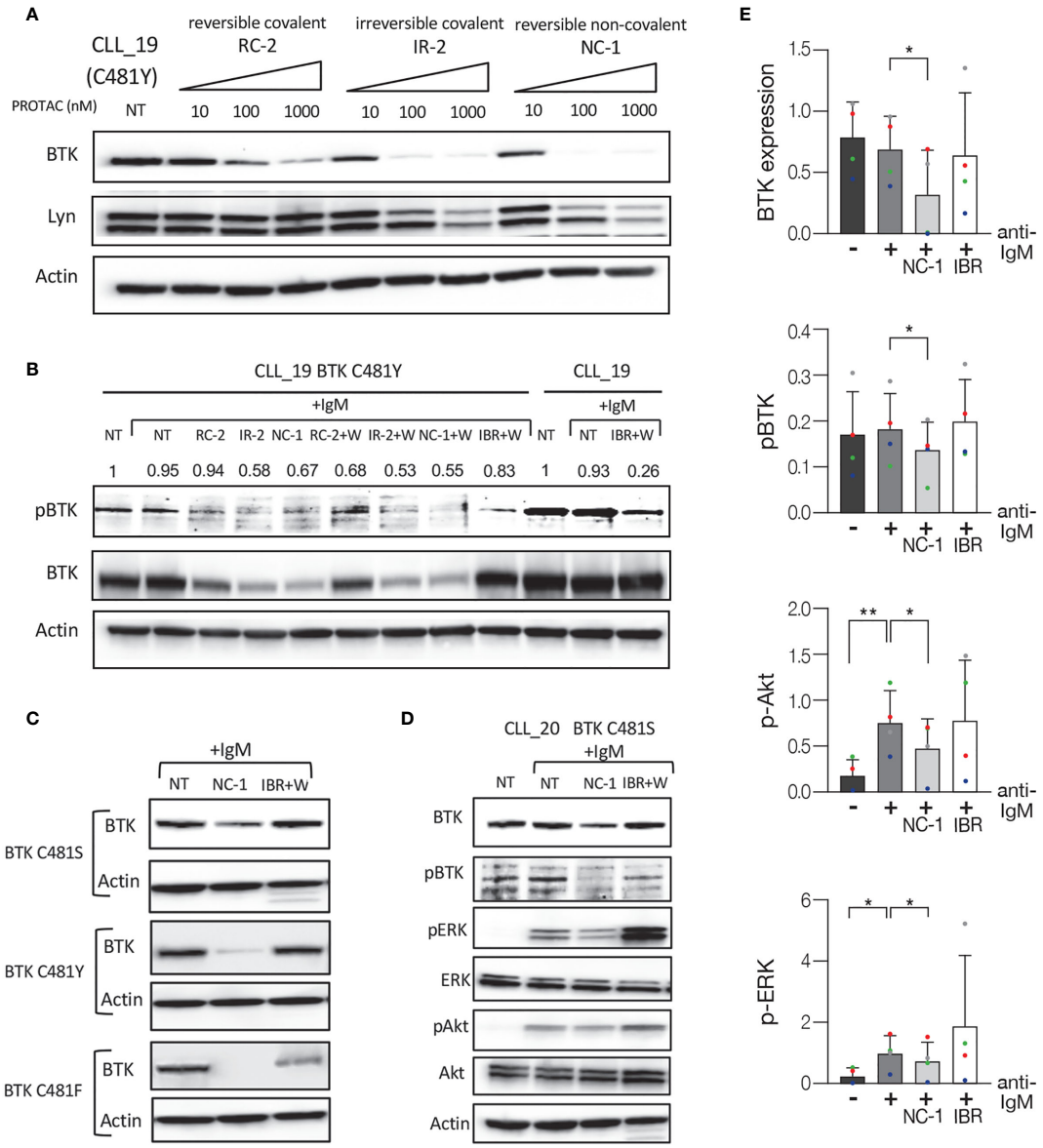


FIGURE 3 | Efficient BCR signaling pathway inhibition in ibrutinib resistant CLL cells. Peripheral blood CLL cells of ibrutinib resistant patients were cultured in 6-well dishes (20x10^6 cells/mL in RPMI 10% FCS) and incubated with BTK PROTACs for 18 hours or with ibrutinib (IBR) for 1 hour at 37° in a humidified 5% CO2 atmosphere. DMSO treated cells served as controls. Following treatment, cells were incubated with goat F(ab') 2 anti-human IgM (10 μg/mL) for 15 minutes or left unstimulated. Cells treated with ibrutinib were washed before activation. Then, proteins were extracted and analyzed by Western blot analysis. **(A)** A representative Western blot showing BTK levels in C481Y CLL cells in response to various concentrations (1000 nM, 100 nM and 10 nM) of RC-2, IR-2 and NC-1 PROTACs. Total Lyn levels were evaluated to confirm specificity of the PROTACs to BTK. Actin was used to verify equal loading. **(B)** A Western blot analysis showing pBTK and BTK levels in response to treatment with 100 nM RC-2, IR-2, NC-1 or ibrutinib and activation with goat F(ab')2 anti-human IgM in CLL cells before and after acquiring BTK C481Y mutation. PROTACs-treated cells were washed (W) or not washed prior to activation. **(C)** Total BTK levels in BTK mutated cells (C481S, C481Y and C481F) in response to 100nM NC-1 or ibrutinib. **(D)** A representative Western blot showing BTK (Tyr223), ERK (T202/Y204) and Akt (S473) phosphorylation, as well as total amount of these proteins, following treatment with 100 nM NC-1 or ibrutinib and activation with goat F(ab')2 anti-human IgM in BTK C481S CLL cells. **(E)** Quantification of BTK, pBTK, pERK and pAkt levels in ibrutinib resistant cells (representative analysis, D) by normalization to actin using myImageAnalysis™ Software (n=4). *p<0.05, **p<0.01.

secrete high amounts of SDF-1, thereby they can attract CLL cells *via* this receptor. This process can govern the homing and survival of CLL cells *in vivo* (31, 32). CLL cells were incubated with 100 nM NC-1 for 18 hours or left untreated. A total of 100 μL, containing 5 × 10^5 cells was added to the top chamber of transwell culture inserts. Filters then were transferred to wells containing medium with 100 ng/mL SDF-1α. Wells containing medium without SDF-1α served as a negative control. After 2

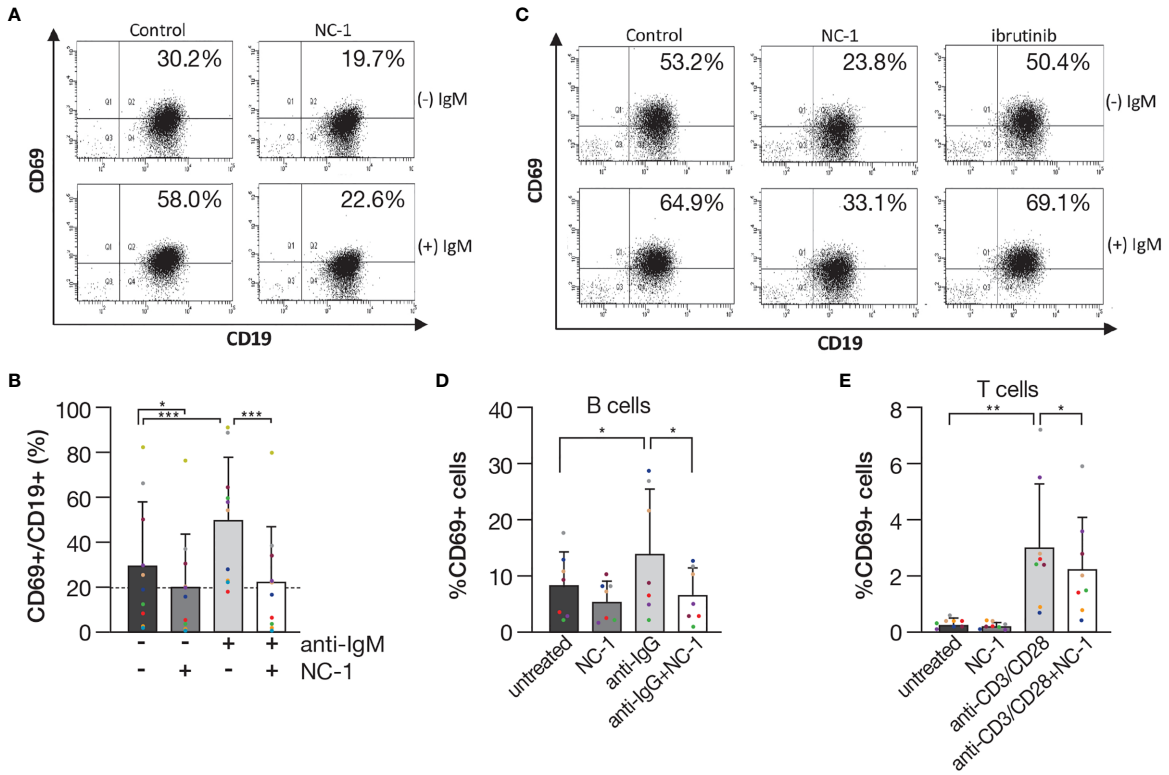


FIGURE 4 | The effect of PROTAC on activation marker expression. Peripheral blood CLL cells were cultured in 6-well dishes (5x10^6 cells/mL in RPMI 10% FCS) and incubated with 100 nM NC-1 BTK PROTAC for 18 hours or with 100nM ibrutinib for 1 hour at 37° in a humidified 5% CO2 atmosphere. DMSO treated cells served as controls. Following treatment, cells were incubated with goat F(ab')2 anti-human IgM (10 µg/mL) for 3 hours or left unstimulated. After incubation, the cells were stained with APC Mouse Anti-Human CD19 and PE Mouse Anti-Human CD69 and incubated for 30 minutes on ice. Isotype controls were APC Mouse IgG1, κ and PE Mouse IgG1, κ. Samples were acquired by a FACSCanto II (BD) and analyzed using BD FACSDiva software. **(A)** Flow cytometric dot-plots of CD19-APC versus CD69-PE expression on samples of one representative CLL case. **(B)** Quantification of CD69+/CD19+ cells in samples obtained from CLL patients (n=10). *p<0.05, ***p<0.001. **(C)** Flow cytometric dot-plots of CD19-APC versus CD69-PE expression on samples of ibrutinib resistant CLL patient (C481S BTK). **(D, E)** Peripheral blood mononuclear cells (PBMC) isolated from healthy donors were incubated with 100 nM NC-1 for 18 hours. DMSO treated cells served as controls. Following treatment, cells were incubated with goat F(ab')2 anti-human IgG (10 µg/mL) or with Dynabeads Human T-Activator CD3/CD28 for 3 hours or left unstimulated. After incubation, the cells were stained with APC Mouse Anti-Human CD19 or CD3, respectively, and PE Mouse Anti-Human CD69. Percentages of CD69+ in CD19+ (n=7) and in CD3+ (n=8) populations were calculated. *p<0.05, **p<0.01.

hours of incubation, the cells in the lower chamber were counted. The results are the average ± S.D. of seven individual patients. The chemotaxis of CLL cells to SDF-1α was significantly inhibited by preincubation of the input cells with 100 nM NC-1 PROTAC (**Figure 5C**).

To study the capability of BTK PROTAC to overcome the supporting effect of mesenchymal stromal cells on CLL cells, CLL cells were co-cultured with HS-5, human bone marrow-derived stromal cells, and treated with 100 nM NC-1 for 48 hours. Proteins were extracted from CLL cells treated with PROTAC and controls and analyzed by Western blotting. Cleaved PARP levels were lower in CLL cells co-cultured with stromal cells compared to those incubated in medium alone. The addition of NC-1 to the co-culture system abrogated the anti-apoptotic effect of the stromal cells (**Figures 5D, E**). Taken together, these results demonstrate that PROTAC can overcome supporting effects of the tissue microenvironment on CLL cells.

DISCUSSION

In this study we demonstrate that the BCR pathway in CLL cells can be efficiently inhibited with PROTACs directed towards BTK. We show that non-covalent (NC-1), irreversible covalent (IR-2) and also reversible covalent (RC-2) PROTACs are able to degrade BTK in CLL cells. The NC-1 compound has a higher degradation potency compared with the IR-2 and RC-2 PROTACs. The efficacy of NC-1 was also observed in patients with poor genomic abnormalities, including del17p and *TP53* mutation. The higher potency of non-covalent NC-1 may be attributed to its rapid binding and dissociation equilibrium (25), as well as potentially improved stability and permeability. Furthermore, the non-covalent binding of NC-1 may explain its more prominent reduction of Lyn levels. While ibrutinib is a 400-fold weaker inhibitor of Lyn than BTK (26), efficient degradation of kinases by PROTACs was shown also based on

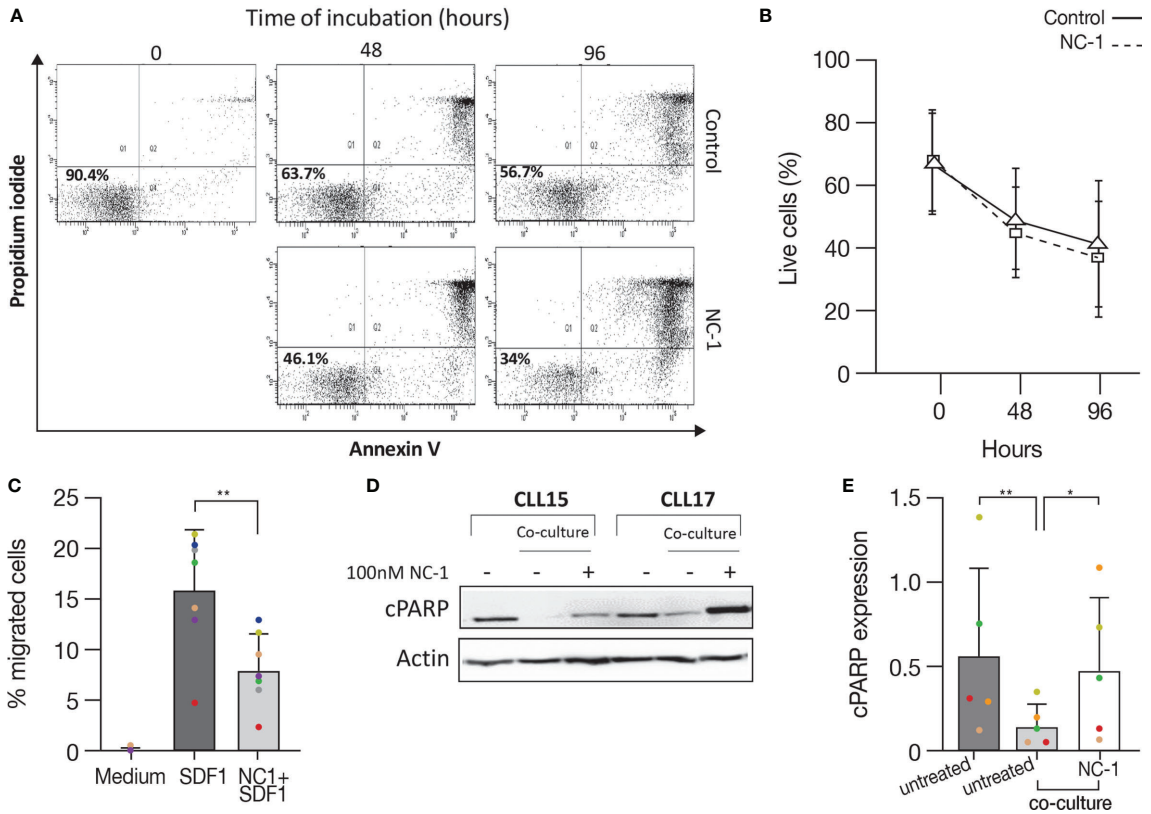


FIGURE 5 | BTK PROTAC induces apoptosis, inhibits migration and overcomes anti-apoptotic protection of stromal cell on CLL cells. Peripheral blood CLL cells were cultured in 24-well dishes (5x10^6 cells/mL in RPMI 10% FCS) and incubated with 100 nM NC-1 PROTAC at 37° in 5% humidified CO2 atmosphere. DMSO treated cells served as controls. Cells viability was evaluated by flow cytometry immediately after thawing, after 48 and 96 hours of incubation, using an Annexin V/PI apoptosis detection kit. **(A)** A representative flow cytometry analyses of annexin V-propidium iodide double staining CLL-treated cells and controls at 2 and 3 time points respectively. The left lower quadrant represents remaining live cells (percentages are indicated). The right lower quadrant represents the population of early apoptotic cells. The right upper quadrant represents the accumulation of late apoptotic cells. **(B)** Viability percentages of cells treated with PROTAC and controls (n=11) measured by flow cytometry after thawing, after 48 and 96 hours of incubation. **(C)** Peripheral blood CLL cells were cultured in 6-well dishes (5x10^6 cells/mL in RPMI 10% FCS and incubated with 100 nM NC-1 BTK PROTAC for 18 hours. DMSO treated cells served as controls. A total of 100 µL, containing 5 × 10^5 cells, was added to the top chamber of a 6.5-mm diameter Transwell culture inserts with a pore size of 5 µm. Filters then were transferred to wells containing medium with 100 ng/mL SDF-1α.Wells containing medium without SDF-1α served as a negative control. The chambers were incubated for 2 hours at 37°C in 5% CO2. After this incubation, the cells in the lower chamber were suspended and divided into aliquots for counting. The experiments were performed in triplicates (n=7). The averages + S.D. are shown. **p<0.01*. **(D, E)** Peripheral blood CLL cells (10x10^6 cells/mL) were co-cultured with HS-5 bone marrow stromal cells in 6-well dishes and treated with 100 nM NC-1. DMSO treated cells served as control. After 48 hours of incubation, proteins were extracted and analyzed by Western blot. A representative Western blot analysis showing cleaved PARP levels as a marker to measure apoptosis. Actin was used to verify equal loading (n=5). **p<0.05*, ***p<0.01*.

weak binding inhibitors (13). Nevertheless, the reduction in Lyn levels can in fact have a beneficial therapeutic effect in CLL. Lyn plays a crucial role in the onset and progression of CLL and its targeting by dasatinib has been shown to inhibit BCR signaling in CLL cells (33).

Because of the higher NC-1 potency, we focused on this PROTAC to investigate its further effects on BCR signaling, activation, apoptosis and migration. In accordance with BTK degradation by NC-1, phosphorylation of BTK was abolished and activation of its downstream elements (PLCγ2, Akt and ERK) was partially blocked in response to IgM engagement. Although treatment with ibrutinib resulted in a reduction in BTK phosphorylation and also inhibition of BCR downstream elements, the latter effect was less prominent, to some extent, than that of NC-1. Given that Lyn is an off target of NC-1 compound, we also analyzed the effect of this PROTAC on phosphorylation of proximal BCR elements. The treatment led to decrease in phosphorylation of CD79a and Lyn, which can contribute to further inhibition of BCR pathway.

Treatment of ibrutinib-resistant CLL cells with mutations at C481 using NC-1 led to a decrease in BTK and pBTK levels, as well as in phosphorylation of Akt and ERK. As expected in BTK mutated cells, ibrutinib did not inhibit the phosphorylation of these elements. Taken together, our results demonstrate that the PROTACs are effective in degradation of BTK as well as inhibiting the downstream elements of the BCR signaling, including inhibitory activity in ibrutinib resistant cells with mutations at C481. From a therapeutic point of view, although

PRTOACs have impressive activity in the targeting of BTK in CLL cells, over time compensatory molecules and pathways may arise that will oppose their activity, so in the future a combination of drugs should be studied.

CLL cells often present a phenotype of activated B cells with overexpression of surface activation markers including CD69 (28). Therefore, we analyzed the effect of the BTK PROTAC NC-1 on the expression of the early activation marker CD69. Treatment with NC-1 decreased the baseline CD69 surface levels, and completely abrogated the upregulation of CD69 following BCR activation both in BTK WT and mutated CLL cells. These results reinforce our findings that PROATCs directed to BTK are capable of blocking both "tonic" and "antigen-triggered" BCR signals in CLL cells. The effect of the PROTAC on this activation marker expression was also tested in normal B and T cells. As we expected, NC-1 compound decreased CD69 surface levels in activated normal B cells as they express BTK as part of their BCR complex. The PROTAC also decreased CD69 expression in activated normal T cells. This result can be explained by the possibility that NC-1 degraded the known ibrutinib off-target ITK in the T cells (26). In contrast to CLL cells, in normal cells there was no significant effect on CD69 levels in resting cells treated with 100 nM NC-1 compound. Our results indicate that in activated cells the effect of PROTAC is more pronounced and therefore it is an effective treatment strategy for CLL patients in whom the B cells have both "tonic" and "antigen –triggered" signals.

We also tested the effect of the PROTAC on apoptosis of CLL cells. Cleaved PARP analysis revealed a mild apoptosis after 18 hours of PROTAC treatment. Flow cytometry analysis using annexin V/PI staining revealed a decrease in cell viability after 48 and 96 hours of incubation, an effect that was approximately 10% higher in cells that were pre-treated with 100 nM NC-1 PROTAC, which is compatible with the mild apoptotic effect of ibrutinib using equivalent therapeutic plasma concentrations in CLL (50-100 nM) (29). It has been reported that ibrutinib can induce more marked apoptosis of CLL cells, but this has been achieved at supra-therapeutic concentrations up to 100 times higher than the concentration we used (6). These findings further support the concept that a central in-vivo mechanism of action of ibrutinib is displacement of CLL cells from their supportive microenvironment, leading to CLL cell redistribution from tissues into the blood, followed by cell death as a result of "death by neglect" (34).

It is already been previously published that CLL cells express high levels of CXCR4 and the stromal cells secrete SDF-1, thereby they can attract the CLL cells and subsequently govern their homing and survival (31, 32). BTK has a role in trafficking of CLL cells, and its inhibition results in impaired CXCR4 expression, signaling and function in CLL, as was seen while using ibrutinib (35, 36). In this study, the effect of NC-1 PROTAC on CLL cells migration toward SDF-1 was tested, using transwell migration assay. The chemotaxis of CLL cells to SDF-1α was significantly inhibited by preincubation of the cells with 100 nM NC-1 PROTAC. These results demonstrate that the PROTAC is effective in inhibiting an additional aspect of CLL cells physiology, which may cause *in vivo* blockade of CLL cell migration to secondary lymphoid tissues and bone marrow.

Given the importance of the stromal niche in CLL cells survival *in vivo*, the impact of PROTAC treatment on apoptosis in the presence of stromal support was investigated. Our results further show that NC-1 can overcome the anti-apoptotic protection of co-cultured stromal cells on CLL cells.

BTK PROTACs have already been developed (19–24) and the first designed BTK-targeted PROTACs were reported by Sun et al. (20). These studies showed a reduction in the level of BTK protein, inhibition of ERK phosphorylation in Ramos cell line (24), and degradation of BTK in CLL patient samples including C481S (22). The novelty of our work that we show the more comprehensive effects of PROTACs to BTK on proximal and downstream BCR signaling, activation, migration and apoptosis in patients derived CLL cells including those with different BTK mutations at C481. In addition, the use of PROTAC designed by us, has made it possible to further investigate the effect of compounds with different binding properties to BTK on CLL cells.

Taken together, PROTACs represent a very promising and powerful approach for the development of targeted therapy drugs, as was recently underscored by the first PROTAC, ARV-110, to enter clinical trials (12). For BTK, PROTACs such as NC-1 have advantages over irreversible inhibitors since they are not sensitive to mutations in C481 (**Figure 3**) as well as other reversible BTK inhibitors due to their longer duration of action. Furthermore, this work demonstrates the potential of BTK PROTACs to inhibit the BCR pathway, and paves the way for future development of this novel therapeutic modality in CLL. However, more efforts will be required to obtain deeper insight into the efficacy and safety of PROTACs in clinical settings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

YSA, YH, and B-ZK designed the study, analyzed data, and wrote the paper. YSA performed and analyzed Western blot, flow cytometry and cell migration experiments. YH conducted statistical analysis. NL, AS, and RG designed the PROTACs. RG synthesized the PROTACs. All authors contributed to the article and approved the submitted version.

FUNDING

YH is supported by grants from the Israeli Science Foundation (1707/19), Israeli Cancer Association and Sackler Faculty of Medicine, Tel-Aviv University. NL is the incumbent of the

Alan and Laraine Fischer Career Development Chair. NL would like to acknowledge funding from the Israel Science Foundation (grant no. 2462/19), The Rising Tide Foundation, The Israel Cancer Research Fund, the Israeli Ministry of Science Technology (grant no. 3-14763), and the Moross Integrated Cancer Center. NL is also supported by the Helen and Martin Kimmel Center for Molecular Design, Joel and Mady Dukler Fund for Cancer Research, the Estate of Emile Mimran and Virgin JustGiving, and the George Schwartzman Fund. RG was supported by the state of Israel, ministry of Aliyah, Center for Integration in Science.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.646971/full#supplementary-material

REFERENCES

1. Caligaris-Cappio F, Hamblin TJ. B-Cell Chronic Lymphocytic Leukemia: A Bird of a Different Feather. *J Clin Oncol* (1999) 17(1):399–408. doi: 10.1200/JCO.1999.17.1.399
2. Schwarzbich MA, McClanahan F, Gribben J. Allogeneic Transplantation for Chronic Lymphocytic Leukemia in the Age of Novel Treatment Strategies. *Oncology* (2016) 30(6):526–33.
3. Herishanu Y, Perez-Galan P, Liu D, Biancotto A, Pittaluga S, Vire B, et al. The Lymph Node Microenvironment Promotes B-cell Receptor Signaling, NF-kappaB Activation, and Tumor Proliferation in Chronic Lymphocytic Leukemia. *Blood* (2011) 117(2):563–74. doi: 10.1182/blood-2010-05-284984
4. Woyach JA, Johnson AJ, Byrd JC. The B-cell Receptor Signaling Pathway as a Therapeutic Target in CLL. *Blood* (2012) 120(6):1175–84. doi: 10.1182/blood-2012-02-362624
5. Wiestner A. Emerging Role of Kinase-Targeted Strategies in Chronic Lymphocytic Leukemia. *Hematol Am Soc Hematol Educ Program* (2012) 2012:88–96. doi: 10.1182/asheducation.V2012.1.88.3801172
6. Herman SE, Gordon AL, Hertlein E, Ramanunni A, Zhang X, Jaglowski S, et al. Bruton Tyrosine Kinase Represents a Promising Therapeutic Target for Treatment of Chronic Lymphocytic Leukemia and is Effectively Targeted by PCI-32765. *Blood* (2011) 117(23):6287–96. doi: 10.1182/blood-2011-01-328484
7. Woyach JA, Furman RR, Liu TM, Ozer HG, Zapatka M, Ruppert AS, et al. Resistance Mechanisms for the Bruton's Tyrosine Kinase Inhibitor Ibrutinib. *N Engl J Med* (2014) 370(24):2286–94. doi: 10.1056/NEJMoa1400029
8. An S, Fu L. Small-Molecule PROTACs: An Emerging and Promising Approach for the Development of Targeted Therapy Drugs. *EBioMedicine* (2018) 36:553–62. doi: 10.1016/j.ebiom.2018.09.005
9. Toure M, Crews CM. Small-Molecule PROTACS: New Approaches to Protein Degradation. *Angew Chem* (2016) 55(6):1966–73. doi: 10.1002/anie.201507978
10. Schneekloth AR, Pucheault M, Tae HS, Crews CM. Targeted Intracellular Protein Degradation Induced by a Small Molecule: En Route to Chemical Proteomics. *Bioorg Med Chem Lett* (2008) 18(22):5904–8. doi: 10.1016/j.bmcl.2008.07.114
11. Sun X, Gao H, Yang Y, He M, Wu Y, Song Y, et al. Protacs: Great Opportunities for Academia and Industry. *Signal Transduct Targeted Ther* (2019) 4:64. doi: 10.1038/s41392-019-0101-6
12. Mullard A. First Targeted Protein Degrader Hits the Clinic. *Nat Rev Drug Discov* (2019) 18:237–9. doi: 10.1038/d41573-019-00043-6
13. Bondeson DP, Smith BE, Burslem GM, Buhimschi AD, Hines J, Jaime-Figueroa S, et al. Lessons in PROTAC Design From Selective Degradation With a Promiscuous Warhead. *Cell Chem Biol* (2018) 25(1):78–87.e5. doi: 10.1016/j.chembiol.2017.09.010
14. Brand M, Jiang B, Bauer S, Donovan KA, Liang Y, Wang ES, et al. Homolog-Selective Degradation as a Strategy to Probe the Function of CDK6 in AML. *Cell Chem Biol* (2019) 26(2):300–6.e9. doi: 10.1016/j.chembiol.2018.11.006
15. Jiang B, Wang ES, Donovan KA, Liang Y, Fischer ES, Zhang T, et al. Development of Dual and Selective Degraders of Cyclin-Dependent Kinases 4 and 6. *Angew Chem* (2019) 58(19):6321–6. doi: 10.1002/anie.201901336
16. Huang HT, Dobrovolsky D, Paulk J, Yang G, Weisberg EL, Doctor ZM, et al. A Chemoproteomic Approach to Query the Degradable Kinome Using a Multi-kinase Degrader. *Cell Chem Biol* (2018) 25(1):88–99.e6. doi: 10.1016/j.chembiol.2017.10.005
17. You I, Erickson EC, Donovan KA, Eleuteri NA, Fischer ES, Gray NS, et al. Discovery of an AKT Degrader With Prolonged Inhibition of Downstream Signaling. *Cell Chem Biol* (2020) 27(1):66–73.se7. doi: 10.1016/j.chembiol.2019.11.014
18. Bondeson DP, Mares A, Smith IE, Ko E, Campos S, Miah AH, et al. Catalytic in Vivo Protein Knockdown by Small-Molecule Protacs. *Nat Chem Biol* (2015) 11(8):611–7. doi: 10.1038/nchembio.1858
19. Tinworth CP, Lithgow H, Dittus L, Bassi ZI, Hughes SE, Muelbaier M, et al. Protac-Mediated Degradation of Bruton's Tyrosine Kinase is Inhibited by Covalent Binding. *ACS Chem Biol* (2019) 14(3):342–7. doi: 10.1021/acschembio.8b01094
20. Sun Y, Zhao X, Ding N, Gao H, Wu Y, Yang Y, et al. PROTAC-Induced BTK Degradation as a Novel Therapy for Mutated BTK C481S Induced Ibrutinib-Resistant B-cell Malignancies. *Cell Res* (2018) 28(7):779–81. doi: 10.1038/s41422-018-0055-1
21. Zorba A, Nguyen C, Xu Y, Starr J, Borzilleri K, Smith J, et al. Delineating the Role of Cooperativity in the Design of Potent PROTACs for BTK. *Proc Natl Acad Sci U S A* (2018) 115(31):E7285–E92. doi: 10.1073/pnas.1803662115
22. Buhimschi AD, Armstrong HA, Toure M, Jaime-Figueroa S, Chen TL, Lehman AM, et al. Targeting the C481S Ibrutinib-Resistance Mutation in Bruton's Tyrosine Kinase Using PROTAC-Mediated Degradation. *Biochemistry* (2018) 57(26):3564–75. doi: 10.1021/acs.biochem.8b00391
23. Krajcovicova S, Jorda R, Hendrychova D, Krystof V, Soural M. Solid-Phase Synthesis for Thalidomide-Based Proteolysis-Targeting Chimeras (PROTAC). *Chem Commun* (2019) 55(7):929–32. doi: 10.1039/C8CC08716D
24. Dobrovolsky D, Wang ES, Morrow S, Leahy C, Faust T, Nowak RP, et al. Bruton Tyrosine Kinase Degradation as a Therapeutic Strategy for Cancer. *Blood* (2019) 133(9):952–61. doi: 10.1182/blood-2018-07-862953
25. Gabizon R, Shraga A, Gehrtz P, Livnah E, Shorer Y, Gurwicz N, et al. Efficient Targeted Degradation Via Reversible and Irreversible Covalent Protacs. *J Am Chem Soc* (2020) 142(27):11734–42. doi: 10.1021/jacs.9b13907
26. Honigberg LA, Smith AM, Sirisawad M, Verner E, Loury D, Chang B, et al. The Bruton Tyrosine Kinase Inhibitor PCI-32765 Blocks B-cell Activation and is Efficacious in Models of Autoimmune Disease and B-cell Malignancy. *Proc Natl Acad Sci U S A* (2010) 107(29):13075–80. doi: 10.1073/pnas.1004594107
27. Teicher BA, Ara G, Herbst R, Palombella VJ, Adams J. The Proteasome Inhibitor PS-341 in Cancer Therapy. *Clin Cancer Res* (1999) 5(9):2638–45.
28. Damle RN, Ghiotto F, Valetto A, Albesiano E, Fais F, Yan XJ, et al. B-Cell Chronic Lymphocytic Leukemia Cells Express a Surface Membrane Phenotype of Activated, Antigen-Experienced B Lymphocytes. *Blood* (2002) 99(11):4087–93. doi: 10.1182/blood.V99.11.4087
29. Chen LS, Bose P, Cruz ND, Jiang Y, Wu Q, Thompson PA, et al. A Pilot Study of Lower Doses of Ibrutinib in Patients With Chronic Lymphocytic Leukemia. *Blood* (2018) 132(21):2249–59. doi: 10.1182/blood-2018-06-860593
30. Burger JA, Burger M, Kipps TJ. Chronic Lymphocytic Leukemia B Cells Express Functional CXCR4 Chemokine Receptors That Mediate Spontaneous Migration Beneath Bone Marrow Stromal Cells. *Blood* (1999) 94(11):3658–67. doi: 10.1182/blood.V94.11.3658
31. Mohle R, Failenschmid C, Bautz F, Kanz L. Overexpression of the Chemokine Receptor CXCR4 in B Cell Chronic Lymphocytic Leukemia is Associated With Increased Functional Response to Stromal Cell-Derived Factor-1 (SDF-1). *Leukemia* (1999) 13(12):1954–9. doi: 10.1038/sj.leu.2401602

32. Burger JA, Kipps TJ. Chemokine Receptors and Stromal Cells in the Homing and Homeostasis of Chronic Lymphocytic Leukemia B Cells. *Leukemia Lymphoma* (2002) 43(3):461–6. doi: 10.1080/10428190290011921
33. ten Hacken E, Scielzo C, Bertilaccio MT, Scarfo L, Apollonio B, Barbaglio F, et al. Targeting the LYN/HS1 Signaling Axis in Chronic Lymphocytic Leukemia. *Blood* (2013) 121(12):2264–73. doi: 10.1182/blood-2012-09-457119
34. Burger JA, Li KW, Keating MJ, Sivina M, Amer AM, Garg N, et al. Leukemia Cell Proliferation and Death in Chronic Lymphocytic Leukemia Patients on Therapy With the BTK Inhibitor Ibrutinib. *JCI Insight* (2017) 2(2):e89904. doi: 10.1172/jci.insight.89904
35. Chen SS, Chang BY, Chang S, Tong T, Ham S, Sherry B, et al. BTK Inhibition Results in Impaired CXCR4 Chemokine Receptor Surface Expression, Signaling and Function in Chronic Lymphocytic Leukemia. *Leukemia* (2016) 30(4):833–43. doi: 10.1038/leu.2015.316
36. de Rooij MF, Kuil A, Geest CR, Eldering E, Chang BY, Buggy JJ, et al. The Clinically Active BTK Inhibitor PCI-32765 Targets B-cell Receptor- and Chemokine-Controlled Adhesion and Migration in Chronic Lymphocytic Leukemia. *Blood* (2012) 119(11):2590–4. doi: 10.1182/blood-2011-11-390989

Conflict of Interest: YH received honoraria from AstraZeneca and Janssen for work unrelated to the present study.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

BRIEF RESEARCH REPORT
published: 18 June 2021
doi: 10.3389/fonc.2021.703254

Predominant VH1-69 IgBCR Clones Show Higher Expression of CD5 in Heterogeneous Chronic Lymphocytic Leukemia Populations

Domenico Maisano[1†], Enrico Iaccino[1†], Alessandro D'Ambrosio[2], Federico Chiurazzi[2], Vincenzo Dattilo[3*], Mariangela Scalise[4], Massimo Gentile[5], Eleonora Vecchio[1], Nancy Nisticò[1], Annamaria Aloisio[1], Erika De Sensi[1], Giuseppe Fiume[1], Ileana Quinto[1‡] and Selena Mimmi[1*‡]*

[1] Laboratory of Immunology, Department of Experimental and Clinical Medicine, University Magna Graecia of Catanzaro, Catanzaro, Italy, [2] Hematological Clinic, Department of Clinical Medicine, University "Federico II" of Naples, Naples, Italy, [3] Genetics Unit, IRCCS Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy, [4] Laboratory of Molecular and Cellular Cardiology, Department of Experimental and Clinical Medicine, University Magna Graecia of Catanzaro, Catanzaro, Italy, [5] Hematology Unit, AO, Cosenza, Italy

The immunoglobulin B cell receptor (IgBCR) expressed by chronic lymphocytic leukemia (CLL) B cells plays a pivotal role in tumorigenesis, supporting neoplastic transformation, survival, and expansion of tumor clones. We demonstrated that in the same patient, two or more CLL clones could coexist, recognized by the expression of different variable regions of the heavy chain of IgBCR, composing the antigen-binding site. In this regard, phage display screening could be considered the easier and most advantageous methodology for the identification of small peptide molecules able to mimic the natural antigen of the tumor IgBCRs. These molecules, properly functionalized, could be used as a probe to specifically identify and isolate single CLL subpopulations, for a deeper analysis in terms of drug resistance, phenotype, and gene expression. Furthermore, CLL cells express another surface membrane receptor, the CD5, which is commonly expressed by normal T cells. Piece of evidence supports a possible contribution of CD5 to the selection and maintenance of autoreactivity in B cells and the constitutive expression of CD5 on CLL cells could induce pro-survival stimuli. In this brief research report, we describe a peptide-based single-cell sorting using as bait the IgBCR of tumor cells; in the next step, we performed a quantitative analysis of CD5 expression by qRT-PCR related to the expressed IgBCR. Our approach could open a new perspective for the identification, isolation, and investigation of all subsets of IgBCR-related CLL clones, with particular attention to the more aggressive clones.

Keywords: chronic lymphocytic leukemia, phage display, immunoglobulin B cell receptor, peptide-based sorting, gene expression

OPEN ACCESS

Edited by:
Andreas Agathangelidis, Institute of Applied Biosciences (INAB), Greece

Reviewed by:
Riccardo Bomben, Aviano Oncology Reference Center (IRCCS), Italy
Giovanna Cutrona, San Martino Hospital (IRCSS), Italy

***Correspondence:**
Domenico Maisano
maisano@unicz.it
Vincenzo Dattilo
vdattilo@fatebenefratelli.eu
Selena Mimmi
mimmi@unicz.it

†These authors share first authorship

‡These authors share senior authorship

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: *30 April 2021*
Accepted: *31 May 2021*
Published: *18 June 2021*

Citation:
Maisano D, Iaccino E, D'Ambrosio A, Chiurazzi F, Dattilo V, Scalise M, Gentile M, Vecchio E, Nisticò N, Aloisio A, De Sensi E, Fiume G, Quinto I and Mimmi S (2021) Predominant VH1-69 IgBCR Clones Show Higher Expression of CD5 in Heterogeneous Chronic Lymphocytic Leukemia Populations. Front. Oncol. 11:703254. doi: 10.3389/fonc.2021.703254

INTRODUCTION

CD5 is a membrane surface receptor expressed by thymocytes, mature T cells, B1a subset of B cells, and leukemic B cells of chronic lymphocytic leukemia (B-CLL) disease (1, 2). Called also Leu-1, it is a 67-kDa type I transmembrane glycoprotein monomer and a member of the scavenger receptor cysteine-rich (SRCR) family (3). The extracellular region is composed of three different domains (D1, D2, and D3) and represents the putative binding region, while the intracellular domain contains the Immunoreceptor Tyrosine-based Activation Motif (ITAM) sequence as the docking site for phosphorylated Src homology 2 (SH2) domain-containing proteins (4).

CD5 is not expressed in normal B cells, except the B1 subgroup, while it is mostly expressed in B-CLL cells (2); this suggests a possible critical role of CD5 in self-maintenance and progression of neoplastic B cells (5, 6). This hypothesis is supported by the evidence that CD5 activates multiple signaling pathways, including mitogen-activated protein kinase (MAPK) (Ras/Erk) pathway, the Ca2+–calmodulin–calcineurin–NFAT pathway, and the PI3-K/Akt/mTOR pathway (7). Further, in transgenic mice, the expression of CD5 correlates with the self-reactivity in B cell populations, supporting a possible contribution to the selection and maintenance of autoreactivity in B cells (8).

CLL is the most frequent adult leukemia in western countries, with a variable clinical course and the occurrence of a heterogeneous tumor population (9, 10). Increasing evidence supports the hypothesis that CLL pathogenesis is an antigen-driven process by a continuous triggering of the immunoglobulin B cell receptor (IgBCR) (11–15), resulting in the no random choice of heavy chain variable region (VH) family during B cell development and the consequential expression of the stereotyped IgBCRs, frequently found in different patients (16). The expression of peculiar IgBCRs is often related to the aggressiveness of the disease (17). Indeed, the unmutated CLL (U-CLL), with less than 2% of mutation in comparison with the germline sequence, seems to be more aggressive with respect to the mutated CLL (M-CLL), which shows a higher percentage of mutations in the variable region of the heavy chain (18). Furthermore, in the U-CLL subgroup, the rearrangement VH1-69 is the most representative (about 25%), and patients showed an aggressive disease with the expansion of CLL clones expressing the unmutated IgBCRs, drug resistance, and often a fatal outcome (16, 19).

Since CD5 seems to be located close to the surface IgBCR on the B cell surface (6) and seems to be a potential ligand of peculiar Ig heavy chain framework sequences in malignant B cells (5), these findings suggest that CD5 could be a self-antigen recognized by the CLL-IgBCRs, promoting survival and proliferation signaling.

In our last published work, we analyzed two CLL patients (named CLL1 and CLL5) for 2-years observation, demonstrating the coexisting of several leukemic subpopulations identified by different IgBCRs, but the most representative subpopulation identified by the rearrangement VH1-69 persisted during all the time (20). So, based on the evidence mentioned above, we asked whether the survival and progression of the VH1-69 subpopulation could be related to higher CD5 gene expression levels compared to the other coexisting clones. Taking advantage of the previously selected peptide (named p1), able to specifically target the leukemic cells expressing the rearrangement VH1-69 (20), we performed a peptide-based cell sorting, in order to isolate the VH1-69 clones from peripheral blood of both oligoclonal CLL1 and CLL5 patients. p1 positive sorted clones (corresponding to the VH1-69 clones) were analyzed by qRT-PCR for the expression of CD5, compared to the other CLL clones (p1 negative clones).

Our results demonstrate that this approach, extended to all IgBCR subsets, could open new strategies for a deeper comprehension of the most aggressive clones, analyzing a wide range of molecular mechanisms and drug resistance - related genes.

METHODS

Peptide-Based Cell Sorting

Frozen B-CLL cells previously isolated from CLL patients were gently thawed and 24h-cultured in RPMI medium supplemented with 10% fetal bovine serum. Then, cells (1×10^7 cells) were first labelled with anti CD19-APC (Miltenyi Biotec, Germany, cat.n. 170-078-090) and anti-CD5-PE. (Miltenyi Biotec – Germany, cat.n. 130-110-990) antibodies, for setting the gate of B-CLL cells. Bulk CLL cells were incubated with FITC-conjugated peptide p1 (1 ng/ml) at 4°C for 20 min, analyzed by flow cytometry, and p1-positive cells were sorted by BD FACSAria III TM (Becton Dickinson). Gating was done using the BD FACSDiva™ software (Becton, Dickinson Biosciences). Cell sorting was performed with 70-µm nozzle size and sorted directly into 5-ml tubes containing 3 ml of staining media in order to minimize cellular stress. Cells were gated in FSC-A vs SSC-A and single cells gated in FSC-H vs FSC-A. Stringent gating strategies to exclude debris or dead cells that exhibit autofluorescence and CD5 negative cells were applied.

qRT-PCR for CD5 Gene Expression

Total mRNA was extracted from CLL cells (bulk or sorted cells) by TRIzol RNA Isolation Reagents (Invitrogen) according to the manufacturer's instructions and quantified by spectrophotometer. 500ng of total mRNA was retrotranscribed into cDNA using the iScript™ cDNA Synthesis Kit (BioRad).

qRT-PCR was performed using the CD5 primers (forward 5′ CAGAAGAAGCAGCGCCAGT 3′; reverse 5′ TCCTGGGA GGTTGGCTGTATT 3′). The general reaction conditions were as follows: initial denaturation step at 95°C for 10 min; 40 cycles of denaturation at 95°C for 10 s; annealing at 57°C for 10 s; and elongation at 72°C. All reactions were performed in triplicate employing the CFX96 Touch Deep Well Real-Time PCR System (BioRad). The results normalized to the GAPDH housekeeping gene and determined by ΔΔCt method were represented as log10 fold expression ± SD of triplicate assessments. Statistical significance was evaluated using one-way analysis of variance (ANOVA), followed by Bonferroni's test for multiple

comparisons. Bars show mean values ± 95% confidence intervals based on three biological replications.

RESULTS

The two CLL patients (named CLL1 and CLL5) were previously analyzed on 2 years observation (20). In particular, the patient CLL1 was analyzed at Binet stage A at months 1 and 5 (CLL1A and CLL1B, respectively) and Binet stage C at month 8 (CLL1C); the patient CLL5, being Binet stage A all the time, was analyzed at months 1, 12, and 24 (CLL5A, CLL5B, and CLL5C, respectively) (**Table 1**). As reported by **Table 1**, the VH1-69 subpopulation persisted all the time, at different percentages of representativeness with respect to the other leukemic clones, related to the aggressiveness of the disease. We firstly analyzed by qRT-PCR the CD5 expression levels in total CD5 positive B-CLL cells of CLL1 and CLL5 patients at different times of disease. As shown in **Figure 1A**, relative gene expression of CD5 was slightly increased, but not statistically significant comparing CLL1A to CLL1B (both Binet stage A). However, a significant increase in CD5 expression was associated with the expansion of VH1-69 clone in the passage from Binet stage A to C (CLL1A fold 1.00± 0.19; CLL1C fold 1.37±0.06), passing from 50% to 60% up to 80% of representativeness. Differently, in CLL5 patient CD5 expression levels significantly decreased comparing CLL5A and CLL5B, associated to a representativeness decrease of VH1-69 clone, passing from 75% to 35%; further, relative gene expression of CD5 was slightly decreased, but not statistically significant comparing CLL5B to CLL5C (**Figure 1B**). These observations suggested that the expression level of CD5 correlated with the percentage of the existing VH1-69 CLL clones.

To deepen our analysis, we took advantage of the previously identified peptide p1 as a specific ligand of the VH1-69 unmutated subpopulation in CLL1 and CLL5 patients (20). Indeed, in this work, we used the peptide p1 as a probe to sort the VH1-69 subpopulation from the total B-CLL cells of CLL1 and CLL5 patients. **Figure 2** shows the coexisting of a p1 positive and a p1 negative population in the total CLL cells of CLL1 (**Figure 2A**) and CLL5 (**Figure 2B**) patients. After peptide-based sorting, IgBCRs sequence was analyzed both in p1-positive and p1-negative fraction, to validate the sorting procedures. No VH1-69 rearrangement was found in the p1-negative fraction.

TABLE 1 | Clinical and molecular data of CLL1 and CLL5 patients.

Patient	Sample (collection time)	WBC (% of CD19/CD5 positive)	Binet stage	VH1-69 subpopulation/total CLL cells (%)
CLL1 65-years old male	CLL1A (month 1)	40,410/mmc (90%)	A	60%
	CLL1B (month 5)	69,070/mmc (92%)	A	50%
	CLL1C (month 8)	92,670/mmc (99%)	C	80%
CLL5 80-years old female	CLL5A (month 1)	57,210/mmc (95%)	A	75%
	CLL5B (month 12)	119,999/mmc (98%)	A	45%
	CLL5C (month 24)	86,500/mmc (96%)	A	35%

Peripheral blood samples of CLL1 and CLL5 patients were collected at the indicated time. CLL stage was defined according to Binet classification (Cancer.Net Editorial Board, 10/2017). The percentage of VH1-69 subpopulation cells was determined by IgBCR sequencing as previously reported (20). The nucleotide sequences of CLL IgBCRs were deposited (GenBank accession numbers MT334403 to MT334414).

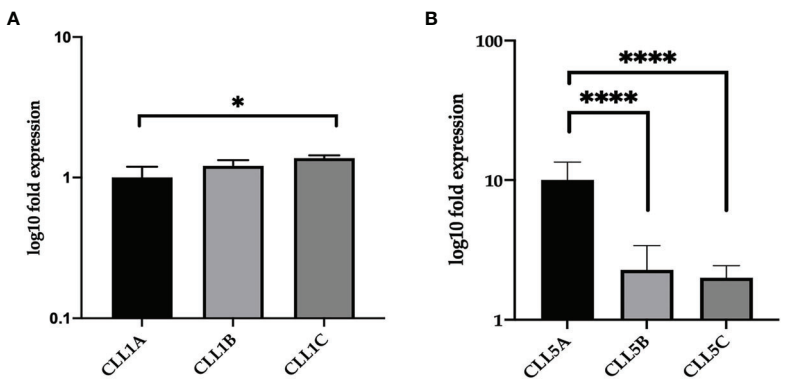


FIGURE 1 | Relative expression of CD5 in CLL clones of CLL1 and CLL5 patients. Relative CD5 gene expression in total B-CLL cells from CLL1 **(A)** and CLL5 **(B)** patient. The results were normalized to the GAPDH housekeeping gene, determined by ΔΔCt method, and represented as log10 fold expression ± SD of triplicate assessments. Statistical significance was evaluated using one-way analysis of variance (ANOVA), followed by Bonferroni's test for multiple comparisons. Bars show mean values ± 95% confidence intervals based on three biological replications. *P ≤ 0.01; ****P ≤ 0.0001.

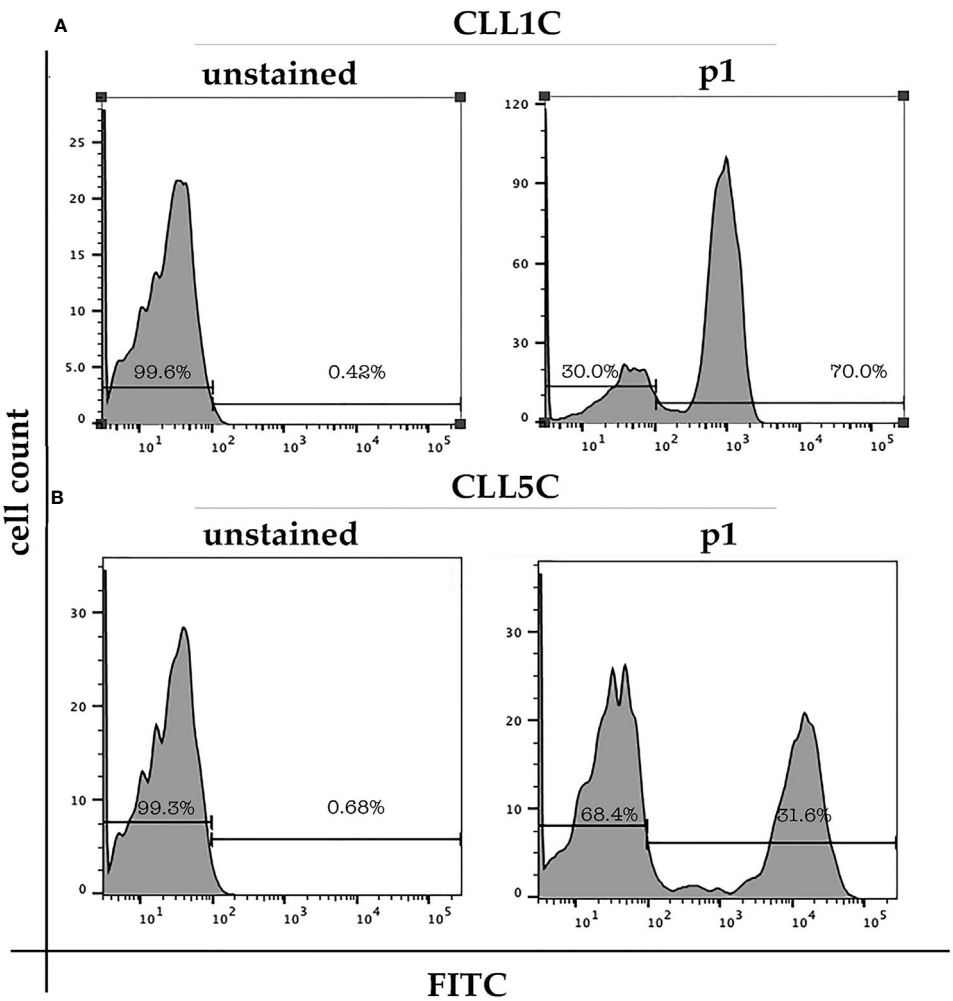


FIGURE 2 | CD19/CD5 positive B-CLL cells isolated from CLL1C (month 8, panel **A**) and CLL5C (month 24, panel **B**) patients were stained with FITC-conjugated peptide p1 or unstained. p1-positive cells corresponding to the VH 1-69 clones were 70% of total B-CLL population in CLL1C patient and 32% in CLL5C patient.

Then, we analyzed by qRT-PCR the CD5 expression levels in p1-positive and p1-negative cells, compared to the bulk. As shown in **Figure 3**, a higher expression of CD5 was observed in p1-positive cells (representing the VH 1-69 clone) as compared to p1-negative cells (the remaining CLL clones with IgBCR rearrangements differing from VH 1-69), both in CLL1 (**Figure 3A**) and CLL5 patient (**Figure 3B**), indicating that CD5 expression levels were related to the expression of VH1-69 rearrangement of IgBCR.

DISCUSSION

CLL clinical course could be characterized by the presence of different tumor B cell clones that could appear or disappear over time, recognized only by the different variable regions of the expressed IgBCRs. These clonal populations may influence the prognosis of the disease, establishing a balanced condition in which the patient remains stable for many years without therapy requirement, or one of them could escape from the apoptosis and proliferation checkpoints, resulting in tumor progression, the need for therapy, and in some cases fatal outcome.

In this scenario, it is interesting the investigation of molecular mechanisms which allow that particular B cell tumor clones to be more aggressive compared to the other tumor populations coexisting in the same patient.

Several studies were conducted in the field of predicting prognosis factors, including the presence or absence of zeta-chain-associated protein kinase (ZAP)-70 or CD38 (21), genomic alterations (10), TP53 status (22), and mutational status of the IgBCR (18).

In particular, the mutational status was one of the first prognostic factors evaluated in CLL patients, observing that

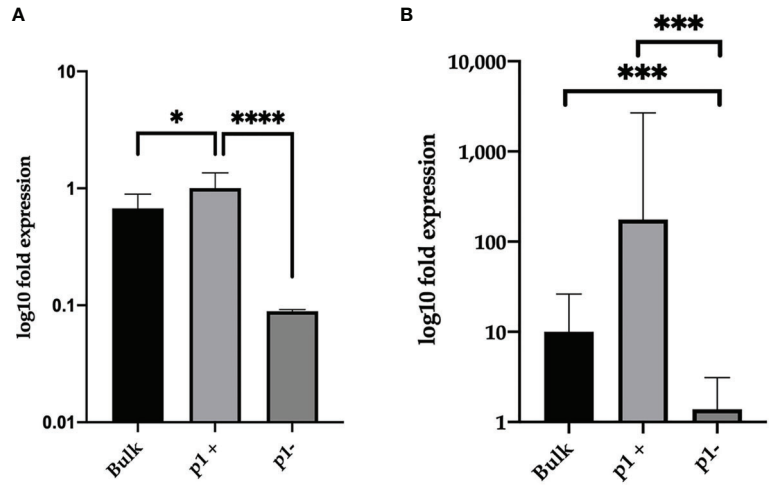


FIGURE 3 | Relative expression of CD5 in p1 positive and p1 negative CLL clones of CLL1 and CLL5 patients. Relative CD5 gene expression in B-CLL subpopulations of CLL1C **(A)** and CLL5C **(B)** sample. Bulk represents the total CLL population, p1-positive the sorted VH1-69 CLL cells, and p1-negative the remaining CLL cells. The results, normalized to the GAPDH housekeeping gene and determined by ΔΔCt method, are represented as log10 fold expression ± SD of triplicate assessments. Statistical significance was evaluated using one-way analysis of variance (ANOVA), followed by Bonferroni's test for multiple comparisons. Bars show mean values ± 95% confidence intervals based on three biological replicates. *P ≤ 0.01; ***P ≤ 0.001; ****P ≤ 0.0001.

patients with unmutated immunoglobulins showed a poor prognosis with drug resistance and rebound (18).

The keystone could be peculiar IgBCR rearrangements that result in self-stimulation through binding to intrinsic motifs of the IgBCR (14, 15) or, on the other side, the presence of a persistent stimulation guided and triggered by endogenous and exogenous epitopes (23–25).

Further, the sequence analysis of IgBCRs expressed by CLL cells, compared to normal B cells, revealed the expression of quasi-identical Ig receptors circa in 30% of diagnosed patients, defined stereotyped IgBCRs (26). These observations fit with the hypothesis of a common exogenous or endogenous antigen stimulating the tumor IgBCRs (14, 15, 23–25).

Reported data describe the capability of the IgBCR to induce the transcription of CD5 by B cells in a murine model (27), which, in turn, transduces pro-survival signaling, such as the IL-10 production (28). Furthermore, it was previously demonstrated that circulating CLL cells, that have been activated recently in proliferating centers, express high levels of surface CD5, which is progressively downregulated as the cells enter into an anergy state, suggesting that the CD5 expression levels could be correlated with the aggressive behavior of the CLL cells (29). Thus, CD5 seems to bind to the Ig heavy chain framework sequence of CLL cells, with a preference for the VH1-69 rearrangement (5), which could promote the selective expansion of B cell clones harboring specific VH genes in CLL.

Our study could represent the proof of concept of the potential use of specific peptide ligands of IgBCRs as probes for sorting and analyzing single tumor subpopulations in CLL patients. Thanks to this new methodology, and according to all well-reported data mentioned above, we observed that the CD5 expression level was increased with the expansion of a specific CLL subpopulation, in our case the unmutated VH1-69 clone, and this could be part of the mechanism of clonal expansion and persistence during the observation, representing additional information for CLL prognosis. More specifically, we demonstrated that single CLL clones could express variable CD5 expression levels, according to their tumorigenic behavior and IgBCR rearrangement.

In perspective, this research line could be extended to all tumor populations of CLL, allowing a wide gene expression analysis, for associating a peculiar IgBCR rearrangement to a specific panel of up or down- regulated genes. We are confident that this approach could get further insights into mechanisms of tumor progression and patient-specific molecular therapy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary materials. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

Experiments involving human subjects were approved by the Italian Regional "Calabria" Ethics Committee (Protocol N. 75, 23/03/17), in accordance with the ethical and safety rules and guidelines provided by the relevant Italian laws (art. 4–5 of D. lgs 116/92, DD.MM. of 29/09/1995 and 26/04/2000), and in accordance with the ethical guidelines of the European Community Council (directive n. 86/609/ECC). Blood samples from healthy donors or CLL patients were obtained upon written

and oral informed consent from the participants to the study. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

DM and SM designed and conducted the research, analyzed the data, and wrote the manuscript. EI assisted with the experiment and data analysis. VD designed qRT-PCR primers and protocols and analyzed data. AD'A, FC, and MG provided CLL samples and clinical analysis. MS performed cell sorting. EV, NN, AA, EDS, and GF focused on data analysis. IQ and SM supervised the research plan and data analysis, and reviewed the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the following grants: POR FES/FESR 2014-20-ATS ALCMEONE cup J18C17000610006 to IQ; MIUR-PRIN 2017MHJJ55_002 to IQ; GILEAD Fellowship 2018 to EI. SM was supported by funds from the EU project PON-AIM1897004-1; DM was supported by funds from the EU project FSE-FESR PON-RI2014-2020.

REFERENCES

1. Boumsell L, Coppin H, Pham D, Raynal B, Lemerle J, Dausset J, et al. An Antigen Shared by a Human T Cell Subset and B Cell Chronic Lymphocytic Leukemic Cells. Distribution on Normal and Malignant Lymphoid Cells. *J Exp Med* (1980) 152:229–34. doi: 10.1084/jem.152.1.229
2. Agee JF, Garver FA, Faguet GB. An Antigen Common to Chronic Lymphocytic and Hairy Cell Leukemia Cells Not Shared by Normal Lymphocytes or by Other Leukemic Cells. *Blood* (1986) 68:62–8. doi: 10.1182/blood.V68.1.62.62
3. Padilla O, Calvo J, Vilà JM, Arman M, Gimferrer I, Places L, et al. Genomic Organization of the Human CD5 Gene. *Immunogenetics* (2000) 993–1001. doi: 10.1007/s002510000235
4. Brown MH, Lacey E. A Ligand for CD5 is CD5. *J Immunol* (2010) 185:6068–74. doi: 10.4049/jimmunol.0903823
5. Pospisil R, Silverman GJ, Marti GE, Aruffo A, Bowen MA, Mage RG. CD5 is A Potential Selecting Ligand for B-cell Surface Immunoglobulin: A Possible Role in Maintenance and Selective Expansion of Normal and Malignant B Cells. *Leuk Lymphoma* (2000) 36:353–65. doi: 10.3109/10428190009148857
6. Jaseb K, Purrahman D, Shahrabi S, Ghanavat M, Rezaeean H, Saki N. Prognostic Significance of Aberrant CD5 Expression in B-Cell Leukemia. *Oncol Rev* (2019) 13:400. doi: 10.4081/oncol.2019.400
7. Garaud S, Morva A, Lemoine S, Hillion S, Bordron A, Pers JO, et al. CD5 Promotes IL-10 Production in Chronic Lymphocytic Leukemia B Cells Through STAT3 and NFAT2 Activation. *J Immunol* (2011) 186:4835–44. doi: 10.4049/jimmunol.1003050
8. Raman C. CD5, an Important Regulator of Lymphocyte Selection and Immune Tolerance. *Immunol Res* (2002) 26:255–63. doi: 10.1385/IR:26:1-3:255
9. Mimmi S, Vecchio E, Iaccino E, Rossi M, Lupia A, Albano F, et al. Evidence of Shared Epitopic Reactivity Among Independent B-cell Clones in Chronic Lymphocytic Leukemia Patients. *Leukemia* (2016) 30:2419–22. doi: 10.1038/leu.2016.245
10. Hallek M. Chronic Lymphocytic Leukemia: 2020 Update on Diagnosis, Risk Stratification and Treatment. *Am J Hematol* (2019) 94:1266–87. doi: 10.1002/ajh.25595
11. Kil LP, Yuvaraj S, Langerak AW, Hendriks RW. The Role of B Cell Receptor Stimulation in CLL Pathogenesis. *Curr Pharm Des* (2012) 18:3335–55. doi: 10.2174/138161212801227041
12. Schleiss C, Ilias W, Tahar O, Güler Y, Miguet L, Mayeur-Rousse C, et al. BCR-Associated Factors Driving Chronic Lymphocytic Leukemia Cells Proliferation Ex Vivo. *Sci Rep* (2019) 9:701. doi: 10.1038/s41598-018-36853-8
13. Zhang S, Kipps TJ. The Pathogenesis of Chronic Lymphocytic Leukemia. *Annu Rev Pathol* (2014) 9:103–18. doi: 10.1146/annurev-pathol-020712-163955
14. Binder M, Müller F, Frick M, Wehr C, Simon F, Leistler B, et al. Cll B-cell Receptors can Recognize Themselves: Alternative Epitopes and Structural Clues for Autostimulatory Mechanisms in CLL. *Blood* (2013) 121:239–41. doi: 10.1182/blood-2012-09-454439
15. Dühren-von Minden M, Übelhart R, Schneider D, Wossning T, Bach MP, Buchner M, et al. Chronic Lymphocytic Leukaemia Is Driven by Antigen-Independent Cell-Autonomous Signalling. *Nature* (2012) 489:309–12. doi: 10.1038/nature11309
16. Agathangelidis A, Psomopoulos F, Stamatopoulos K. Stereotyped B Cell Receptor Immunoglobulins in B Cell Lymphomas. *Methods Mol Biol* (2019) 1956:139–55. doi: 10.1007/978-1-4939-9151-8_7
17. Packham G, Krysov S, Allen A, Savelyeva N, Steele AJ, Forconi F, et al. The Outcome of B-cell Receptor Signaling in Chronic Lymphocytic Leukemia: Proliferation or Anergy. *Haematologica* (2014) 99:1138–48. doi: 10.3324/haematol.2013.098384
18. Rotbain EC, Frederiksen H, Hjalgrim H, Rostgaard K, Egholm GJ, Zahedi B, et al. IGHV Mutational Status and Outcome for Patients With Chronic Lymphocytic Leukemia Upon Treatment: A Danish Nationwide Population-Based Study. *Haematologica* (2020) 105:1621–9. doi: 10.3324/haematol.2019.220194
19. 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) 94:1848–54. doi: 10.1182/blood.V94.6.1848
20. Mimmi S, Maisano D, Nisticò N, Vecchio E, Chiurazzi F, Ferrara K, et al. Detection of Chronic Lymphocytic Leukemia Subpopulations in Peripheral Blood by Phage Ligands of Tumor Immunoglobulin B Cell Receptors. *Leukemia* (2021) 35:610–4. doi: 10.1038/s41375-020-0885-y
21. Basabaeen AA, Abdelgader EA, BaHashwan OS, Babekir EA, Abdelateif NM, Bamusa SA, et al. Combined Analysis of ZAP-70 and CD38 Expression in Sudanese Patients With B-Cell Chronic Lymphocytic Leukemia. *BMC Res Notes* (2019) 12:282. doi: 10.1186/s13104-019-4319-8
22. Morabito F, Gentile M, Monti P, Recchia AG, Menichini P, Skafi M, et al. TP53 Dysfunction in Chronic Lymphocytic Leukemia: Clinical Relevance in the Era of B-cell Receptors and BCL-2 Inhibitors. *Expert Opin Investig Drugs* (2020) 29:869–80. doi: 10.1080/13543784.2020.1783239
23. Lupia A, Mimmi S, Iaccino E, Maisano D, Moraca F, Talarico C, et al. Molecular Modelling of Epitopes Recognized by Neoplastic B Lymphocytes in Chronic Lymphocytic Leukemia. *Eur J Med Chem* (2020) 185:111838. doi: 10.1016/j.ejmech.2019.111838
24. Os A, Bürgler S, Ribes AP, Funderud A, Wang D, Thompson KM, et al. Chronic Lymphocytic Leukemia Cells Are Activated and Proliferate in Response to Specific T Helper Cells. *Cell Rep* (2013) 4:566–77. doi: 10.1016/j.celrep.2013.07.011
25. Darwiche W, Gubler B, Marolleau JP, Ghamlouch H. Chronic Lymphocytic Leukemia B-Cell Normal Cellular Counterpart: Clues From a Functional Perspective. *Front Immunol* (2018) 9:683. doi: 10.3389/fimmu.2018.00683
26. Colombo M, Bagnara D, Reverberi D, Matis S, Cardillo M, Massara R, et al. Tracing CLL-Biased Stereotyped Immunoglobulin Gene Rearrangements in Normal B Cell Subsets Using a High-Throughput Immunogenetic Approach. *Mol Med* (2020) 26:25. doi: 10.1186/s10020-020-00151-9
27. Cong YZ, Rabin E, Wortis HH. Treatment of Murine CD5- B Cells With anti-Ig, But Not LPS, Induces Surface CD5: Two B-cell Activation Pathways. *Int Immunol* (1991) 3:467–76. doi: 10.1093/intimm/3.5.467
28. Gary-Gouy H, Harriague J, Bismuth G, Platzer C, Schmitt C, Dalloul AH. Human CD5 Promotes B-Cell Survival Through Stimulation of Autocrine IL-10 Production. *Blood* (2002) 100:4537–43. doi: 10.1182/blood-2002-05-1525

29. Calissano C, Damle RN, Hayes G, Murphy EJ, Hellerstein MK, Moreno C, et al. In Vivo Intraclonal and Interclonal Kinetic Heterogeneity in B-Cell Chronic Lymphocytic Leukemia. *Blood* (2009) 114:4832–42. doi: 10.1182/blood-2009-05-219634

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

BRIEF RESEARCH REPORT
published: 26 August 2021
doi: 10.3389/fonc.2021.740083

Consistent B Cell Receptor Immunoglobulin Features Between Siblings in Familial Chronic Lymphocytic Leukemia

*P. Martijn Kolijn[1†], Alice F. Muggen[1†], Viktor Ljungström[2,3,4], Andreas Agathangelidis[5,6], Ingrid L. M. Wolvers-Tettero[1], H. Berna Beverloo[7], Karol Pál[8], Paul J. Hengeveld[1], Nikos Darzentas[9], Rudi W. Hendriks[10], Jacques J. M. van Dongen[11], Richard Rosenquist[3,4] and Anton W. Langerak[1]**

[1] Laboratory Medical Immunology, Department of Immunology, Erasmus MC, University Medical Center, Rotterdam, Netherlands, [2] Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden, [3] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden, [4] Department of Clinical Genetics, Karolinska University Laboratory, Karolinska University Hospital, Solna, Sweden, [5] Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece, [6] Department of Biology, School of Science, National and Kapodistrian University of Athens, Athens, Greece, [7] Department of Clinical Genetics, Erasmus MC, University Medical Center, Rotterdam, Netherlands, [8] CEITEC – Central European Institute of Technology, Masaryk University, Brno, Czechia, [9] Department of Hematology, University Hospital Schleswig-Holstein, Kiel, Germany, [10] Department of Pulmonary Medicine, Erasmus MC, University Medical Center, Rotterdam, Netherlands, [11] Department of Immunology, LUMC, Leiden, Netherlands

OPEN ACCESS

Edited by:
Guru Prasad Maiti,
Oklahoma Medical Research Foundation, United States

Reviewed by:
Sergiu Pasca,
Iuliu Haţieganu University of Medicine and Pharmacy, Romania
Alain Chebly,
Saint Joseph University, Lebanon

***Correspondence:**
Anton W. Langerak
a.langerak@erasmusmc.nl

[†]These authors share first authorship

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: 12 July 2021
Accepted: 09 August 2021
Published: 26 August 2021

Citation:
Kolijn PM, Muggen AF, Ljungström V, Agathangelidis A, Wolvers-Tettero ILM, Beverloo HB, Pál K, Hengeveld PJ, Darzentas N, Hendriks RW, Dongen JJMv, Rosenquist R and Langerak AW (2021) Consistent B Cell Receptor Immunoglobulin Features Between Siblings in Familial Chronic Lymphocytic Leukemia. Front. Oncol. 11:740083. doi: 10.3389/fonc.2021.740083

Key processes in the onset and evolution of chronic lymphocytic leukemia (CLL) are thought to include chronic (antigenic) activation of mature B cells through the B cell receptor (BcR), signals from the microenvironment, and acquisition of genetic alterations. Here we describe three families in which two or more siblings were affected by CLL. We investigated whether there are immunogenetic similarities in the leukemia-specific immunoglobulin heavy (IGH) and light (IGL/IGK) chain gene rearrangements of the siblings in each family. Furthermore, we performed array analysis to study if similarities in CLL-associated chromosomal aberrations are present within each family and screened for somatic mutations using paired tumor/normal whole-genome sequencing (WGS). In two families a consistent IGHV gene mutational status (one IGHV-unmutated, one IGHV-mutated) was observed. Intriguingly, the third family with four affected siblings was characterized by usage of the lambda IGLV3-21 gene, with the hallmark R110 mutation of the recently described clinically aggressive IGLV3-21^{R110} subset. In this family, the CLL-specific rearrangements in two siblings could be assigned to either stereotyped subset #2 or the immunogenetically related subset #169, both of which belong to the broader IGLV3-21^{R110} subgroup. Consistent patterns of cytogenetic aberrations were encountered in all three families. Furthermore, the CLL clones carried somatic mutations previously associated with IGHV mutational status, cytogenetic aberrations and stereotyped subsets, respectively. From these findings, we conclude that similarities

in immunogenetic characteristics in familial CLL, in combination with genetic aberrations acquired, point towards shared underlying mechanisms behind CLL development within each family.

Keywords: CLL (Chronic Lymphocytic Leukemia), Familial CLL, BCR stereotypy, IGLV3-21 R110, CLL development

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is the most common leukemia in Western countries (1). Sex and age are important risk factors for CLL, with a two-fold increased risk of developing CLL for men compared to women and a median age at CLL diagnosis of around 70 years (2). Although no single genetic lesion drives CLL, a range of recurrent cytogenetic aberrations and somatic mutations have been identified in CLL (2–4).

Cytogenetic aberrations are common in CLL, with around 80% of CLL patients carrying at least one of the four common chromosomal alterations, i.e. del(13q), del(11q), del(17p) and trisomy 12 (2, 5). Of these four alterations, del(13q) is the most frequent and, as a sole aberration, is associated with indolent disease (6). Del(11q) and del(17p) are associated with an unfavorable prognosis, through loss of function of the *ATM* and *TP53* gene, respectively (3, 7–9). Lastly, trisomy 12 is associated with an intermediate prognosis (10, 11). Several key whole-exome sequencing (WES) and whole-genome sequencing (WGS) studies have revealed over 50 recurrently mutated genes (4, 12–15). However, the majority of these putative CLL driver mutations are present at low frequency (<5% of cases), with only a handful of more common mutations in genes such as *TP53*, *ATM*, *SF3B1, NOTCH1* and *BIRC3* (4, 12).

Another important facet of risk stratification of patients with CLL is the somatic hypermutation (SHM) status of the B cell receptor (BcR) immunoglobulin heavy variable (IGHV) gene (16). CLL patients with a mutated IGHV-gene (M-CLL), i.e. showing lower than 98% IGHV gene similarity to its closest germline counterpart, generally have a more indolent disease course than CLL patients with an unmutated IGHV gene with a germline identity equal to or above 98% (U-CLL) (2). Furthermore, stereotyped or (quasi)identical BcR IGs are observed in more than 40% of CLL patients (16). Patients with shared BcR IG motifs can be assigned to distinct stereotyped subsets associated with particular presentation and outcomes (17, 18). One of the stereotyped subsets with the worst clinical outcome is subset #2 (IGHV3-21/IGLV3-21), which displays a mixed IGHV mutation status and an enrichment of *SF3B1* mutations (17–19). An important new subset is the clinically aggressive IGLV3-21^{R110} subset, which also includes subset #2, that is characterized by shared usage of the lambda IGLV3-21*01 or *04 allele, along with a hallmark substitution of Gly to Arg at amino acid position 110 at the very end of the IGLJ gene (20, 21). The IGLV3-21*01 and *04 alleles encode a Lys at position 16 and two Asp residues at position 50 and 52 in the CDR2 region of the light chain variable region (VL), which interact with the R110 light chain residue, resulting in constitutive autostimulation of the BcR, putatively contributing to CLL pathogenesis (20).

Although the aforementioned genetic features mostly occur sporadically, evidence exists for germline predisposition for CLL (17, 22). The incidence of CLL varies geographically, with highest incidence among individuals with European ancestry (23). This hereditary element of CLL is also reflected in familial predisposition, as relatives of CLL patients have an increased risk of developing CLL as well as other B-cell malignancies (24). Furthermore, monoclonal B-cell lymphocytosis (MBL), the asymptomatic pre-stage to CLL, is more often seen in first-degree relatives of CLL patients and is particularly common among healthy relatives of patients with high-risk familial CLL (i.e. families with two or more relatives with CLL) with a prevalence of around 15% among individuals older than 40 years (23, 25). Genome-wide association studies (GWAS) have captured part of this familial predisposition by screening for single nucleotide polymorphisms (SNP) associated with familial CLL, yielding low-risk SNPs distributed over nearly 30 loci (22, 26–33).

In this context, through a combination of immunogenetic, SNP-array and WGS analysis, we here aimed to gain insight into the contribution of BcR composition, cytogenetic aberrations and CLL driver mutations to familial CLL occurrence by studying three families with multiple siblings diagnosed with CLL.

MATERIALS AND METHODS

Samples

Peripheral blood was obtained from ten CLL patients from three families (**Figure 1**). Informed consent was provided in accordance with the declaration of Helsinki and the study was approved by the hospital medical ethics committee (METC2015-741). Familial connection was confirmed through STR analysis. Peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll Paque (GE Healthcare, Little Chalfont, UK) gradient centrifugation. CLL cells and T lymphocytes were sorted from PBMCs using a FACSAria cell sorter (BD Biosciences, San Jose, CA, USA). Immediately after sorting, cells were lysed in RLT+ buffer (Qiagen, Valencia, CA, USA) complemented with β-mercapto-ethanol and stored at -80°C until further processing. DNA and RNA was isolated with the DNA/RNA/miRNA easy kit (Qiagen) according to the manufacturer's protocol. In the event that DNA was isolated from total PBMC, spin-column kits and the QIAcube platform (Qiagen) were used. cDNA was synthesized using the SuperScript™ III First-Strand Synthesis System (Thermo Fisher Scientific, Waltham, MA, USA), according to manufacturer's instructions.

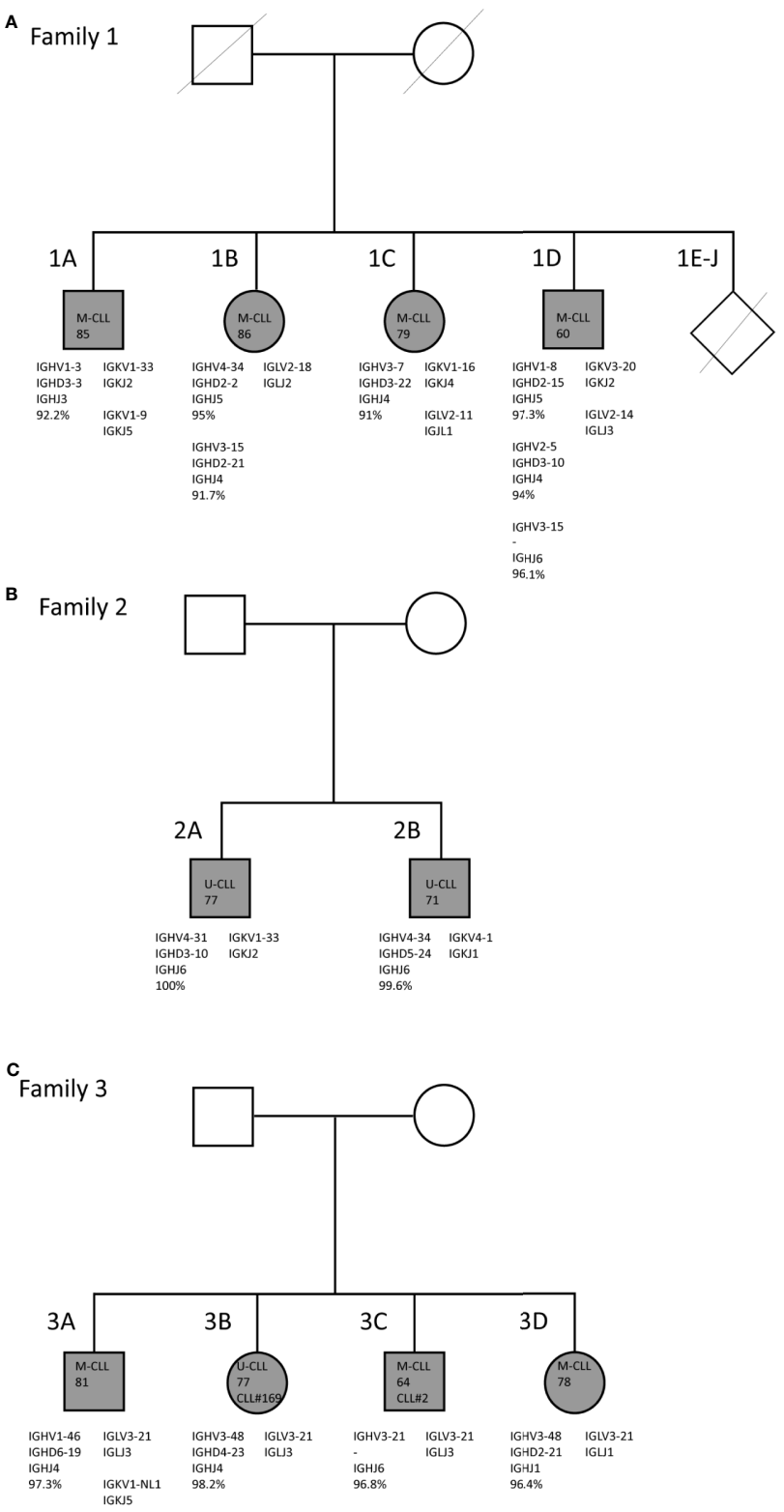


FIGURE 1 | Family trees and BcR IG characteristics of familial CLL cases. **(A)** Family 1 consists of two brothers and two sisters who carry mutated IGHV genes. **(B)** Family 2 consists of two brothers who both carry unmutated IGHV genes. **(C)** Family 3 consists of two brothers and two sisters. In all four siblings, the CLL clone utilizes the IGLV3-21*04 gene with the characteristic R110 mutation and the K16 and YDSD motifs. Additionally, siblings 3B and 3C express similar IGHV genes, i.e. IGHV3-21 and IGHV3-48, and belong to stereotyped subsets #2 and #169, respectively.

IG Gene Rearrangement Analysis

Immunoglobulin heavy (IGH) and IG kappa/lambda (IGK/IGL) gene rearrangements were amplified from 100 ng gDNA isolated from the total PBMC fraction with multiplex PCR utilizing the BIOMED-2 IGH primers and IG light chain consensus primers, following ERIC guidelines (34, 35). Clonal PCR products were separated by heteroduplex gel electrophoresis and were purified by gel extraction. Rearrangements were determined through Sanger sequencing on an ABI 3130xl instrument (ThermoFisher Scientific, Waltham, MA, USA). Sequencing results were analyzed using the IMGT/V-QUEST tool on the IMGT website (www.imgt.org, version 3.3.1). Stereotyped subsets were defined by the following parameters: (1) usage of IGHV genes from the same phylogenetic clan, (2) a minimum of 50% amino acid identity and 70% similarity within the heavy chain CDR3, (3) identical heavy chain CDR3 length and, (4) identical offset of the shared amino acid pattern (28). The IGLV3-21^{R110} mutation was confirmed using IGLV3-21 primers on cDNA for 3 out of 4 members of family 3. As no RNA was available for sibling 3C, the R110 mutation was instead confirmed based on the WGS results analyzed by an extension of the ARResT/Interrogate immunoprofiler for the analysis of IG/TR rearrangements in non-amplicon sequencing data such as from WGS, WES and RNA-seq (36, 37).

SNP Array Analysis

Two hundred fifty ng of genomic DNA was used for single nucleotide polymorphism (SNP) array analysis on the Illumina Human OmniExpress Beadchip (Illumina, San Diego, CA, USA). Data were analyzed with Beadstudio software (Illumina). The log R ratio and B allele frequency data were analyzed using Nexus Copy Number (Nexus BioDiscovery, El Segundo, CA, USA). The results were compared with a database of known copy-number variations (Department of Clinical Genetics, Erasmus MC, Rotterdam, The Netherlands) and a public copy-number variations dataset containing approximately 3500 healthy controls (dataset of genomic variants). The affected locations detected were analyzed in Ensembl Genome Browser 95 (www.ensembl.org) and screened for loci previously linked to CLL in GWAS studies. The used SNP array contained more than 700K probes, and the genome was analyzed with an average resolution of 150 kb, or smaller when it contained at least 10 consecutive probes.

Whole-Genome Sequencing

One hundred ng of genomic DNA was used for construction of WGS libraries using the TruSeq Nano Kit (Illumina Inc.) and sequenced in paired-end mode (2x150bp) on the Illumina HiSeqX Ten system (Illumina Inc.) with 30× target coverage. The bcl files were converted to FASTQ using bcl2fastq and subsequently processed using Piper, a pipeline built on top of GATK queue. Reads from each library were aligned to the Grch37 reference genome using BWA mem and merged and de-duplicated using Picard. Re-alignment around known and novel indel-sites was performed with GATK. All SAM/BAM-conversion steps were completed using SAMtools. Germline samples (T-lymphocytes or PBMC) were compared to reference genome GRCh37 using GATK. However, as the PBMC samples also included CLL cells, no distinction could be made between somatic mutations or novel germline variants for these patients; instead, the PBMC samples of patients 1B and 2B were used to confirm if germline variants identified in sibling(s) were shared. Somatic variation in CLL clone *vs* germline was annotated by the Strelka2 Small Variant Caller. The Variant Call Format (VCF) files were filtered for PASS variants, annotated with VEP and converted to Mutation Annotation Format (MAF) files using VCF2MAF. MAF-files were analyzed using the maftools R package (38). Somatic mutations in CLL-associated genes were annotated by the Ensembl Variant Effect Predictor (VEP, ensembl.org/info/docs/tools/vep/index.html). The panel of CLL driver genes was based on landmark WGS and WES studies (4, 12), for the full panel see **Supplementary Table 1**. Additional screening was then performed for genes in KEGG pathways related to DNA replication, DNA repair, BcR, p53 signaling, cell cycle and the spliceosome. Germline variants were filtered based on clinical significance in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), allele frequency, SIFT, PolyPhen and CADD score. All somatic mutations were screened for disease recurrence in CLL and cancer in COSMIC (39) (cancer.sanger.ac.uk) and Intogen (www.intogen.org). The WGS dataset and immunogenetic sequencing data are available upon request to the corresponding author through the SciLifeLab repository (DOI: 10.17044/scilifelab.14932062).

RESULTS

Families With Multiple CLL Patients

In family 1 (**Figure 1A**) four (out of a total ten) siblings, i.e. two brothers and two sisters, suffered from CLL. They were diagnosed at advanced age [85 (1A), 86 (1B), 79 (1C), and 60 (1D)] and were followed until late age (98, 91, 84 and 82 years, respectively) (40). All ten siblings grew up on a Dutch farm, where cattle breeding and agriculture were practiced. No record was kept of pesticide use at the farm. All of the other six siblings had passed away at time of inclusion, without showing clinical signs of hematological or immunological disease. Both male patients (1A and 1B) moved out during adolescence, while the female patients 1C and 1D remained at the farm until they were middle-aged. Only patient 1A, who also presented with lymphadenopathy, received treatment for CLL (chlorambucil), twelve years after diagnosis (40). The two brothers of family 2 (**Figure 1B**) were diagnosed with CLL at age 77 (2A) and 71 (2B) years. Family 3 also consisted of two brothers and two sisters with CLL, who were diagnosed in the age range from 64 until 81 years (**Figure 1C**). Sibling 3B was treated with fludarabine. Members of both family 2 and family 3 had the Dutch nationality and were Caucasian. Additional clinical data and descriptive information were unfortunately not available for families 2 and 3.

Familial CLL Shows Consistent BcR IG Characteristics

Through IG Sanger sequencing of genomic DNA from total PBMC fractions from CLL patients in each of the three CLL

families, we discovered strikingly similar immunogenetic features within each family (**Figure 1**). For family 1, a consistent somatic hypermutation (SHM) status of the IGHV gene was observed, with each of the four siblings harboring a M-CLL clone with an IGHV gene germline identity below 98% (**Table 1**). Moreover, two siblings, 1B and 1D had multiple CLL clones, each of which expressed a mutated IGHV gene. Although CLL is generally of monoclonal origin, multiple productive IGH rearrangements have been observed in around 2% of CLL cases (16). These can arise from a single CLL clone (biallelic rearrangement) or reflect biclonal CLL disease (41, 42). Interestingly, family member 1D appeared to have biclonal CLL consisting of a SmIgκ+ and a SmIgλ+ CLL clone as determined by flow cytometry (data not shown). Since we detected three productive IGH rearrangements, one of the two CLL clones likely expresses two IGH alleles. The multiple productive IGH rearrangements identified for family member 1B may also be biallelic but could not be discerned as only one rearranged Ig light chain gene was expressed. Previously, multiple additional IGH bands were detected for these family members in Southern blot analysis (40), but these were now all found to be unproductive. Altogether, family 1 is characterized by M-CLL, with multiple productive and unproductive rearrangements in two individuals.

In family 2, both siblings expressed unmutated IGHV genes (U-CLL). Notably, each sibling expressed IGHV4 (IGHV4-31 or IGHV4-34), IGHJ6 and IGK light chain genes, but no BcR IG stereotypy was observed (**Table 1**). Hence, the key defining feature of family 2 is the U-CLL type.

Finally, the CLL clone of all siblings of family 3 expressed an IGHV gene with (near) borderline IGHV mutational status (germline identity ranging from 96.4 - 98.2%; borderline IGHV mutational status is classically defined as 97-97.9% germline identity) (43). Notably, the CLL clone in all four siblings utilized the lambda IGLV3-21*04 gene, suggestive of membership of the recently discovered IGLV3-21^{R110} subset (20, 21), which usually has a borderline mutation status. As the BIOMED-2 IGLV/IGLJ light chain consensus primers did not capture the essential final nucleotide of the IGLJ gene to verify the R110 status, we repeated sequencing with adapted primers on cDNA in cases where RNA was available. We confirmed the somatic R110 mutation and germline configuration of the K16 and YDSD motifs in all four members of family 3 (**Supplementary Figure 1**). Regarding the heavy chain, two family members (3B and 3C) belonged to the closely related and clinically aggressive subsets #2 and #169, respectively (44). The respective IGHV genes of these heavy chain stereotypic CLL subsets, IGHV3-21 and IGHV3-48, were 97% identical. The CLL clone of sibling 3B also expressed the IGHV3-48 gene, though the variable heavy CDR3 (VH CDR3) of this patient did not match a stereotyped subset (**Table 1**). Thus, family 3 is paradigmatic for the IGLV3-21^{R110} subset with a borderline IGHV mutation status.

CLL Families Show Similar Genomic Profiles

To further explore the genomic profiles in these immunogenetically paradigmatic families, we utilized SNP array analysis. We detected genomic aberrations in all three families (**Table 2**). For family 1 and 3 we observed the most common deletion in CLL, del(13q), in the CLL clone(s) of all members, whereas the two brothers of family 2 carried trisomy 12. Additionally, sibling 3C carried del(11q), which is in line with previous reports of subset #2 patients having an increased incidence of 11q deletions (45). Lastly, sibling 1C carried a 2q34-2q35 deletion, a chromosomal aberration not previously associated with CLL, though deletions of 2q37 encompassing *SP140* and *SP110* have been reported (12, 46). Furthermore, the SNP array revealed a distinct loss of heterozygosity (LOH) profile for each family, composed of loci previously linked to CLL in GWAS (**Supplementary Table 3**) (26, 29–33). All three families shared LOH in the MHC locus (6p22.1) and the *CASP8* and *CASP10* locus (2q33.1). LOH of chromosome region 11q22.3, where the *ATM* gene is located, was detected in members of family 1. Additionally, we observed LOH of 14q32.2-q32.33 in family 2, which is interesting as 10% of CLL patients with trisomy 12 were previously observed to have an additional translocation in 14q32 (11). Furthermore, we observed LOH in the 2q22.1 locus in family 2 and family 3, which was recently identified as a novel CLL risk locus using shared genomic segment analysis and was found to include the full *CXCR4* gene. Although there were no cytogenetic data available, we have used SNP array data to define genomic

TABLE 1 | Overview of BcR IG sequencing results.

Family member	Stereotyped subset	IGHV gene	HCDR3	IGLV/IGKV gene	LCDR3
1A	–	V1-3	CARGVRFLEFLLYGDDAFDIW	IGKV1-33 IGK1-9	CQQYDNLPPALATVCQQVNSYPRITF
1B	–	V4-34 V3-15	CARSLVVPAAYGPNSWFDSW CATGGHCGGACYSPYFDYW	IGLV2-18	CSLYTGTKTIF
1C	–	V3-7	CAKHDNTGDFHLDNW	IGKV1-16 IGLV2-11	CQQYNSYPALTF CCSYAGSHTYVF
1D	–	V1-8 V2-5 V3-15	CARHPSRRCSGDFCSTGNWFDPW CLGHWVRGIMTPFDYW CNYYVMDVW	IGKV3-20 IGLV2-14	CQQYGSSPNTF CSSYTSSNTLVF
2A	–	V4-31	CARLLAGLHYYYYYAMDVW	IGKV1-33	CQQYDNLPPYTF
2B	–	V4-34	CARERRDSNYGSGIFYYYYGMDVW	IGKV4-1	CQQYYSTPRTF
3A	IGLV3-21^{R110}	V1-46	CARAWSSAWKYYFDY	IGLV3-21	CQVWDSGSDHPWVF
3B	IGLV3-21^{R110}/#169	V3-48	CARDGVGAPY	IGLV3-21	CQVWDSGTDHPWVF
3C	IGLV3-21^{R110}/#2	V3-21	CARDQNGMDV	IGLV3-21	CQVWDSSSDHPWVF
3D	IGLV3-21^{R110}	V3-48	CARDGGPCGDCYQ	IGLV3-21	CLVWDSGSDHPYVF

TABLE 2 | Cytogenetic aberrations encountered for each of the three families.

Family member	del(13q)	+12	del(11q)	del(17p)	del(2)(q34q35)	total abberations*
Sibling 1A	yes	no	no	no	no	2
Sibling 1B	yes	no	no	no	no	1
Sibling 1C	yes	no	no	no	yes	4
Sibling 1D	yes	no	no	no	no	1
Sibling 2A	no	yes	no	no	no	1
Sibling 2B	no	yes	no	no	no	1
Sibling 3A	yes	no	no	no	no	1
Sibling 3B	yes	no	no	no	no	1
Sibling 3C	yes	no	yes	no	no	3
Sibling 3D	yes	no	no	no	no	1

Total number of aberrations >5 Mb, including the recurrent FISH aberrations shown here.

complexity. In none of the patients a high genomic complexity was indicated (**Table 2**), defined as 5 or more unbalanced aberrations according to Leeksma et al. (47–49). Only two patients (1C and 3C) presented with three or four aberrations (**Supplementary Table 4**). In conclusion, SNP-array analysis revealed shared CLL-associated chromosomal aberrations within each family and LOH in several CLL risk loci, and no complex karyotype cases in any of the families.

Whole-Genome Sequencing Identifies Germline Variants in CLL-Related Pathways in All Three Families

To investigate somatic mutations in the CLL clones and review potential contributing germline variants, we performed WGS on both sorted CLL samples and normal T cells of all three families. Unfortunately, for patients 1B, 2B and 3B, sufficient CLL-derived genomic material for WGS was not available (**Supplementary Table 3**). However, for patient 1B and 2B, we were able to sequence leftover DNA from unsorted PBMCs, allowing us to screen for potential shared germline variants that were found in the families or their members. We performed an initial screen for germline variants and somatic mutations and in CLL driver genes previously identified in WGS and WES studies and then followed up with KEGG pathway analysis to screen for novel CLL-related genes (**Figure 2** and **Supplementary Table 5**).

First, we catalogued the germline variants in each of the families. We identified a germline frameshift deletion in *CHEK2* (p.T410Mfs*15) in siblings 1A, 1B and 1D, but not in sibling 1C (**Figure 2A** and **Supplementary Table 5**). Deleterious germline CHEK2 variants have been associated with an increased risk of developing primarily breast cancer and colorectal cancer (50). Moreover, somatic *CHEK2* alterations have been reported in CLL (4, 51). Additionally, we identified a rare germline missense variant (p.R325C) in *PIK3R3* in sibling 1A, 1C and 1D, but not in sibling 1B. *PIK3R3* is a regulatory subunit of phosphatidylinositol 3-kinase (PI3K) and thus an essential part of the *PI3K/AKT* signaling pathway involved in cell survival and proliferation (52–54).

Notably, we observed a rare germline missense variant in *NFKBIA* (p.T185M) in both siblings of family 2, predicted to be pathogenic by variant effect predictor tools. *NFKBIA* inhibits NF-κB/REL complexes during inflammatory response. *NFKBIA* is also a part of the BcR signaling pathway (55). In family 3, sibling 3A and 3C carried a germline missense variant in the *ERCC6* gene (p.R666C), which encodes a protein involved in the base excision repair pathway. Altogether, some interesting germline variants were observed, but many were of unknown significance and most variants were not shared by all siblings with CLL, making a strong causal relationship in familial CLL less straightforward than for previously described somatic mutations in CLL.

Known and Novel Somatic Mutations in CLL-Driver Genes and Related Pathways in All Three Families

Next, we characterized somatic mutations specific to the CLL clone (**Figure 2B**). We encountered a somatic frameshift deletion and missense mutation in *ATM* in sibling 1A, which in combination with the LOH of chromosome region 11q22.3 results in bi-allelic loss of *ATM*. This same clone had an additional p.G13D somatic missense *KRAS* mutation and a somatic frameshift deletion in the *BAX* gene. Lastly, we observed a somatic missense mutation (p.D470H) in sibling 1A in the *PRKCB* gene, involved in many different signaling pathways, including B-cell activation.

The CLL clone of sibling 1C carried two somatic missense mutations of interest, a p.D594N mutation in *BRAF* previously observed in CLL, and a novel *CD19* mutation (p.L495P). Sibling 1D presented with biclonal CLL, one SmIgκ$^+$ and one SmIgλ$^+$ CLL clone. In each CLL clone, a known CLL driver gene was affected; the IGK+ clone carried a truncating mutation in *MED12* (56), while the IGL$^+$ clone carried a missense mutation at the CLL hotspot (L273P) in *MYD88* (57). Furthermore, we observed a somatic missense mutation (p.G2R) in both the SmIgκ$^+$ and SmIgλ$^+$ clones in *LYN*, a gene directly downstream of the BcR.

In family member 2A we detected somatic frameshift deletions in *FBXW7* and *NOTCH1* and a missense mutation in *KRAS*, all of which have been previously associated with the occurrence of trisomy 12 in CLL (11, 58–60). Unfortunately, the lack of somatic data from patient 2B prevented us from confirming if the somatic mutational profile matched between siblings.

In family 3, we observed a somatic mutation in one of the CLL hotspots (p.G742D) of *SF3B1* for sibling 3A. *SF3B1* mutations are common in CLL and particularly associated with subset #2 and the IGLV3-21^{R110} subset (20, 21). Sibling 3C carried somatic mutations in several low-frequency mutated genes in CLL: *IGLL5*, *DYRK1A* and *BAZ2A* (4, 12). The somatic mutation in *IGLL5* is likely the result of aberrant SHM (61). Additionally,

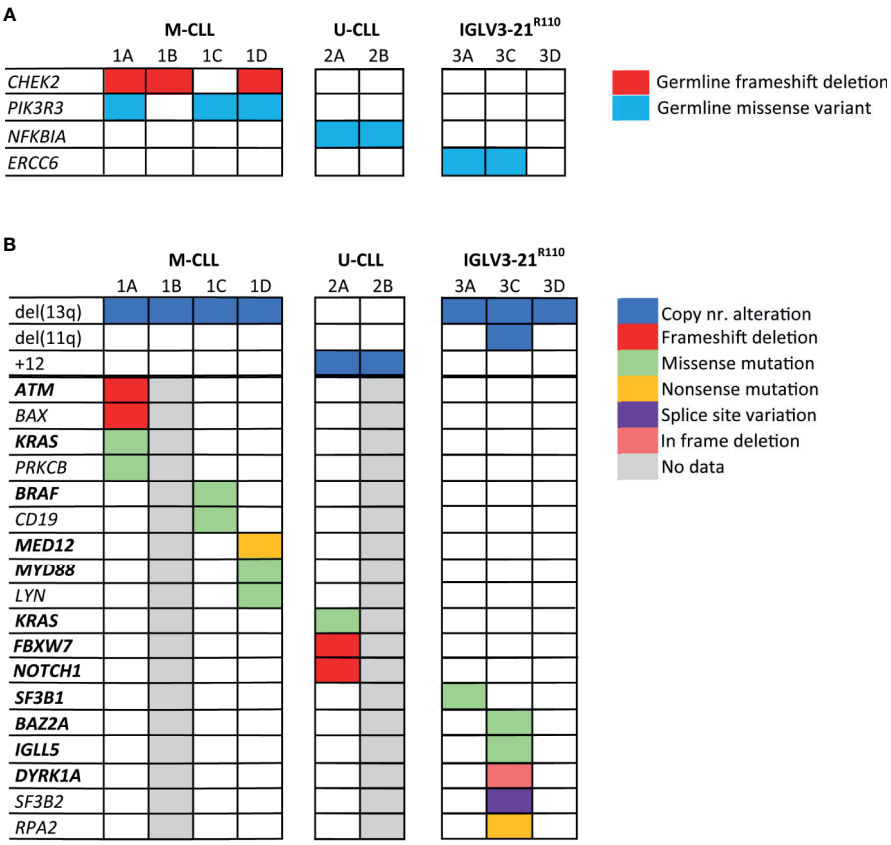


FIGURE 2 | Somatic mutations and germline variants detected by whole genome sequencing of the CLL families. In this figure, both germline variants **(A)** and somatic genetic alterations **(B)** detected in the CLL families are shown. The genes highlighted in bold text are genes, which have previously been identified as CLL driver genes. Genes that are not in bold text were identified during KEGG pathway analysis. Only mutations/variants with likely functional consequences related to CLL development are shown; mutations/variants that were previously reported to be benign or evaluated as benign by variant effect predictors were not shown.

sibling 3C carried a somatic truncating mutation in *RPA2*, a gene involved in DNA replication and repair. In contrast, no noteworthy somatic mutations were observed in sibling 3D.

In summary, the WGS results yielded several somatic mutations in recurrently mutated genes in CLL, as well as four germline variants in genes in CLL-associated pathways, though there was limited overlap in the genes affected by the somatic mutations in members within and across families.

DISCUSSION

In this study, we describe three families that represent distinct immunogenetic subgroups of CLL, presenting a unique opportunity to study the contribution of genetics and immunogenetics in CLL pathobiology. Each of the three families developed CLL with a consistent IGHV SHM status, encompassing one of three prototypes of the IGHV SHM spectrum: i.e. U-CLL, M-CLL and borderline mutated CLL. While families 1 and 2 reflect the M-CLL and U-CLL subgroups, respectively, family 3 presented with borderline mutated CLL and all family members carried the lambda IGLV3-21 light chain. Furthermore, family 3 expressed the IGLV3-21*04 allele and displayed the R110 mutation characteristic of the IGLV3-21^{R110} subset. This light chain was paired with a stereotyped VH CDR3 of the immunogenetically related subsets #2 and #169, both of which belong to the broader IGLV3-21^{R110} category. We observed distinct profiles of genetic alterations for each of these families, with further unique somatic mutations for each sibling. While our results are consistent with previous associations between IGHV SHM mutational status and specific genetic aberrations in CLL driver genes, the similarities in (immuno)genetic features within each family highlight their important contribution to the onset and evolution of familial CLL.

The dichotomy between U-CLL and M-CLL is thought to originate from the B-cell maturation process after antigen activation (62). For M-CLL, the antigen-activated B cell follows the traditional path of T cell-dependent germinal center B cell maturation. For U-CLL, the antigen-activated B cell is thought to mature largely independent of the T cell influence (62). Throughout these processes, chronic antigenic stimulation through (auto)antigens would keep the B cell in a constant state of activation. For the IGLV3-21^{R110} subset, this constant activation is most probably the result of autostimulation through BcR aggregates on the cell surface.

As would have been expected based on the association of IGLV3-21 with CLL with limited SHM activity, the IGLV3-21^{R110} subset is characterized by a (near) borderline mutational status (63). Correspondingly, no cases of IGLV3-21^{R110} with 100% IGHV germline identity have been encountered, thus supporting SHM as the mechanism for the introduction of the somatic R110 mutation (20). The IGHV germline identities of the IGLV3-21^{R110} CLL family 3 follow a similar pattern, ranging from a germline identity of 96.4% to 98.2%. Interestingly, usage of the IGLV3-21*01 or *04 alleles gives an inherent risk of IGLV3-21^{R110}-related CLL, due to the germline presence of the K16 and YDSD motifs (20). Our findings in the current study would support the theory that this inherent risk contributes to the increased incidence of CLL among relatives of CLL patients.

We additionally observed a somatic mutation in *SF3B1* for sibling 3A. *SF3B1* mutations are common in CLL and particularly associated within the IGLV3-21^{R110} subset (20, 21). As the SF3B1 protein is a component of the spliceosome, we screened for additional mutations in the spliceosome pathway. We discovered that sibling 3C carried a splice site alteration in *SF3B2*. Unlike *SF3B1*, *SF3B2* has never been independently linked to CLL. The finding of a splice site alteration in *SF3B2* in sibling 3C suggests that the alterations in other genes involved in the spliceosome may be relevant for the IGLV3-21^{R110} subset as well, although this awaits further confirmation in larger cohorts.

We identified several germline variants of unknown significance (VUS) in each of the families by KEGG pathway analysis. Family 1 presented with germline variants in *CHEK2* and *PI3KR3*, while family 2 carried a germline variant in *NFKBIA* and two siblings of family 3 carried a germline variant in *ERCC6*. *CHEK2* is a gene associated with DNA damage and repair as well as cell cycle regulation and apoptosis in response to DNA damage (51). Somatic *CHEK2* mutations have been identified as putative CLL drivers, while *CHEK2* germline variants have recently been indicated as a novel predisposition gene in CLL, implying that CLL may belong to the spectrum of malignancies associated with germline variant in CHEK2 (54). In addition, three out of four siblings with CLL carried a rare germline variant in *PIK3R3*, an essential component of the PI3K/AKT signaling pathway. Recently, altered activation of the PI3K/AKT signaling-pathway was identified as a critical component of sustained proliferation and survival in CLL (64). During this process, autonomous autoreactive BcR signaling typically converges with activation of the PI3K/AKT signaling-pathway (64). While germline variants in components of the PI3K/AKT pathway could theoretically contribute to this aspect of CLL development, no convincing supporting evidence for a role of any germline variant has this far been reported. NFKBIA is part of the NF-κB and BcR signaling pathways and its expression has been suggested as a biomarker for risk stratification in DLBCL (55). *ERCC6* has a role in base excision repair, particularly during transcription.

Our study was limited by its sample size as well as by the amount of material available for each patient. Additionally, clinical follow up data was not available for family 2 and 3 and no material was available from healthy family members. Lastly, the absence of conventional chromosomal analysis may have affected the identification of complex rearrangements (>3 abnormalities), a prognostic factor in CLL, although we feel that based on SNP array data we could exclude the occurrence of complex karyotype cases. Nevertheless, we feel that the three families are paradigmatic for the main CLL subgroups and as such provide a platform for further studies into the link between immunogenetics and genetic predisposition. That said, environmental factors like pesticides, herbicides and pathogens could be relevant risk factors in familial CLL as well. This would especially apply to the siblings of family 1, who all grew up on the same farm (65). Unfortunately, as no toxicological or biological measurements were done, the contribution of these factors to CLL development in family 1 remains unclear.

In summary, we evaluated immunogenetic, cytogenetic, germline and somatic lesions in familial CLL. In each family, a consistency of IGHV mutational status was observed, with the particularly intriguing finding that all individuals in one of the families belonged to the IGLV3-21^{R110} CLL subset. Furthermore, we highlight the co-occurrence of specific genetic aberrations and germline variants within each family, pointing towards shared underlying mechanisms in CLL development. Our data warrants a more comprehensive evaluation of this potential association between germline predisposition and immunogenetic features in the development of CLL.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://figshare.com/s/f8576aca6650fab99393, 10.17044/scilifelab.14932062.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Erasmus MC Medical Ethics Review Committee. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

AM and IW-T performed the experiments. PK, AM, VL, KP, and ND analyzed the data. AA validated CLL subsets. PK, AM, VL, RR, and AL interpreted results. PK and AL wrote the manuscript. AM, VL, AA, PH, KP, ND, HB, RH, JD, and RR critically reviewed and edited the manuscript. AL designed and supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation. RR is supported by SciLifeLab, the Swedish Cancer Society, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, Karolinska Institutet, Karolinska University Hospital, and Radiumhemmets Forskningsfonder, Stockholm. PK and AL are supported by a EU TRANSCAN-2/ Dutch Cancer Society grant (179;NOVEL consortium).

ACKNOWLEDGMENTS

The authors gratefully acknowledge Birna Thorvaldsdottir for academic discussion. Sequencing was performed by the SNP&SEQ Technology Platform in Uppsala. The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.740083/full#supplementary-material

Supplementary Figure 1 | Validation of the IGLV3-21*04^{R110} mutation in family 3. Identification of the IGL rearrangement in WGS data of sibling 3C by the ARResT/ Interrogate immunoprofiler (upper segment) and a complimentary view in Integrated Genome Viewer (IGV, lower segment) highlighting the g>c missense mutation at the last base in IGLJ3 resulting in the G110R substitution.

Supplementary Table 1 | Overview of CLL driver genes and KEGG pathways used to analyze the WGS data.

Supplementary Table 2 | Genomic profile for CLL loci for each family.

Supplementary Table 3 | Samples utilized for WGS.

Supplementary Table 4 | SNP array analysis results with relevant aberrations for complex karyotypes indicated.

Supplementary Table 5 | Overview of somatic mutations and germline variants that likely contribute to CLL pathogenesis.

REFERENCES

1. Hallek M. Chronic Lymphocytic Leukemia: 2020 Update on Diagnosis, Risk Stratification and Treatment. *Am J Hematol* (2019) 94(11):1266–87. doi: 10.1002/ajh.25595
2. Kipps TJ, Stevenson FK, Wu CJ, Croce CM, Packham G, Wierda WG, et al. Chronic Lymphocytic Leukaemia. *Nat Rev Dis Primers* (2017) 3:1–12. doi: 10.1038/nrdp.2016.96
3. Baliakas P, Jeromin S, Iskas M, Puiggros A, Plevova K, Nguyen-Khac F, et al. Cytogenetic Complexity in Chronic Lymphocytic Leukemia: Definitions, Associations, and Clinical Impact. *Blood* (2019) 133(11):1205–16. doi: 10.1182/blood-2018-09-873083
4. Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, et al. Mutations Driving CLL and Their Evolution in Progression and Relapse. *Nature* (2015) 526(7574):525–30. doi: 10.1038/nature15395
5. Dohner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, et al. Genomic Aberrations and Survival in Chronic Lymphocytic Leukemia. *New Engl J Med* (2000) 343(26):1910–6. doi: 10.1056/NEJM200012283432602
6. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, et al. The DLEU2/miR-15a/16-1 Cluster Controls B Cell Proliferation and Its Deletion Leads to Chronic Lymphocytic Leukemia. *Cancer Cell* (2010) 17(1):28–40. doi: 10.1016/j.ccr.2009.11.019
7. Dohner H, Stilgenbauer S, James MR, Benner A, Weilguni T, Bentz M, et al. 11q Deletions Identify a New Subset of B-Cell Chronic Lymphocytic Leukemia Characterized by Extensive Nodal Involvement and Inferior Prognosis. *Blood* (1997) 89(7):2516–22. doi: 10.1182/blood.V89.7.2516
8. Guarini A, Marinelli M, Tavolaro S, Bellacchio E, Magliozzi M, Chiaretti S, et al. ATM Gene Alterations in Chronic Lymphocytic Leukemia Patients Induce a Distinct Gene Expression Profile and Predict Disease Progression. *Haematologica* (2012) 97(1):47–55. doi: 10.3324/haematol.2011.049270
9. Marinelli M, Peragine N, Di Maio V, Chiaretti S, Stefania De Propris M, Raponi S, et al. Identification of Molecular and Functional Patterns of P53 Alterations in Chronic Lymphocytic Leukemia Patients in Different Phases of the Disease. *Haematologica* (2013) 98(3):371–5. doi: 10.3324/haematol.2012.069906
10. Abruzzo LV, Herling CD, Calin GA, Oakes C, Barron LL, Banks HE, et al. Trisomy 12 Chronic Lymphocytic Leukemia Expresses a Unique Set of Activated and Targetable Pathways. *Haematologica* (2018) 103(12):2069–78. doi: 10.3324/haematol.2018.190132
11. Roos-Weil D, Nguyen-Khac F, Chevret S, Touzeau C, Roux C, Lejeune J, et al. Mutational and Cytogenetic Analyses of 188 CLL Patients With Trisomy 12: A Retrospective Study From the French Innovative Leukemia Organization (FILO) Working Group. *Genes Chromosomes Cancer* (2018) 57(11):533–40. doi: 10.1002/gcc.22650
12. Puente XS, Bea S, Valdes-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, et al. Non-Coding Recurrent Mutations in Chronic Lymphocytic Leukaemia. *Nature* (2015) 526(7574):519–24. doi: 10.1038/nature14666
13. Ljungstrom V, Cortese D, Young E, Pandzic R, Mansouri L, Plevova K, et al. Whole-Exome Sequencing in Relapsing Chronic Lymphocytic Leukemia: Clinical Impact of Recurrent RPS15 Mutations. *Blood* (2016) 127(8):1007–16. doi: 10.1182/blood-2015-10-674572
14. Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, et al. SF3B1 and Other Novel Cancer Genes in Chronic Lymphocytic Leukemia. *N Engl J Med* (2011) 365(26):2497–506. doi: 10.1056/NEJMoa1109016
15. Quesada V, Conde L, Villamor N, Ordóñez GR, Jares P, Bassaganyas L, et al. Exome Sequencing Identifies Recurrent Mutations of the Splicing Factor SF3B1 Gene in Chronic Lymphocytic Leukemia. *Nat Genet* (2011) 44(1):47–52. doi: 10.1038/ng.1032
16. Rosenquist R, Ghia P, Hadzidimitriou A, Sutton L-A, Agathangelidis A, Baliakas P, et al. Immunoglobulin Gene Sequence Analysis in Chronic Lymphocytic Leukemia: Updated ERIC Recommendations. *Leukemia* (2017) 31(7):1477–81. doi: 10.1038/leu.2017.125
17. Agathangelidis A, Chatzidimitriou A, Gemenetzi K, Giudicelli V, Karypidou M, Plevova K, et al. Higher-Order Connections Between Stereotyped Subsets: Implications for Improved Patient Classification in CLL. *Blood* (2021) 137 (10):1365–76. doi: 10.1182/blood.2020007039
18. Jaramillo S, Agathangelidis A, Schneider C, Bahlo J, Robrecht S, Tausch E, et al. Prognostic Impact of Prevalent Chronic Lymphocytic Leukemia Stereotyped Subsets: Analysis Within Prospective Clinical Trials of the German CLL Study Group (GCLLSG). *Haematologica* (2020) 105(11):2598–607. doi: 10.3324/haematol.2019.231027
19. Baliakas P, Agathangelidis A, Hadzidimitriou A, Sutton L-A, Minga E, Tsanousa A, et al. Not All IGHV3-21 Chronic Lymphocytic Leukemias Are Equal: Prognostic Considerations. *Blood* (2015) 125(5):856–9. doi: 10.1182/blood-2014-09-600874
20. Maity PC, Bilal M, Koning MT, Young M, van Bergen CAM, Renna V, et al. IGLV3-21*01 is an Inherited Risk Factor for CLL Through the Acquisition of a Single-Point Mutation Enabling Autonomous BCR Signaling. *Proc Natl Acad Sci U S A* (2020) 117(8):4320–7. doi: 10.1073/pnas.1913810117
21. Nadeu F, Royo R, Clot G, Duran-Ferrer M, Navarro A, Martín S, et al. IGLV3-21R110 Identifies an Aggressive Biological Subtype of Chronic Lymphocytic Leukemia With Intermediate Epigenetics. *Blood* (2021) 137(21):2935–46. doi: 10.1182/blood.2020008311

22. Cerhan JR, Slager SL. Familial Predisposition and Genetic Risk Factors for Lymphoma. *Blood* (2015) 126(20):2265–73. doi: 10.1182/blood-2015-04-537498
23. Goldin LR, Landgren O, Marti GE, Caporaso NE. Familial Aspects of Chronic Lymphocytic Leukemia, Monoclonal B-Cell Lymphocytosis (MBL), and Related Lymphomas. *Eur J Clin Med Oncol* (2010) 2(1):119–26.
24. Goldin LR, Bjorkholm M, Kristinsson SY, Turesson I, Landgren O. Elevated Risk of Chronic Lymphocytic Leukemia and Other Indolent non-Hodgkin's Lymphomas Among Relatives of Patients With Chronic Lymphocytic Leukemia. *Haematologica* (2009) 94(5):647–53. doi: 10.3324/haematol.2008.003632
25. Goldin LR, Lanasa MC, Slager SL, Cerhan JR, Vachon CM, Strom SS, et al. Common Occurrence of Monoclonal B-Cell Lymphocytosis Among Members of High-Risk CLL Families. *Br J Haematol* (2010) 151(2):152–8. doi: 10.1111/j.1365-2141.2010.08339.x
26. Di Bernardo MC, Crowther-Swanepoel D, Broderick P, Webb E, Sellick G, Wild R, et al. A Genome-Wide Association Study Identifies Six Susceptibility Loci for Chronic Lymphocytic Leukemia. *Nat Genet* (2008) 40(10):1204–10. doi: 10.1038/ng.219
27. Sellick GS, Goldin LR, Wild RW, Slager SL, Ressenti L, Strom SS, et al. A High-Density SNP Genome-Wide Linkage Search of 206 Families Identifies Susceptibility Loci for Chronic Lymphocytic Leukemia. *Blood* (2007) 110 (9):3326–33. doi: 10.1182/blood-2007-05-091561
28. Slager SL, Goldin LR, Strom SS, Lanasa MC, Spector LG, Rassenti L, et al. Genetic Susceptibility Variants for Chronic Lymphocytic Leukemia. *Cancer Epidemiol Biomarkers Prev* (2010) 19(4):1098–102. doi: 10.1158/1055-9965.EPI-09-1217
29. Berndt SI, Skibola CF, Joseph V, Camp NJ, Nieters A, Wang Z, et al. Genome-Wide Association Study Identifies Multiple Risk Loci for Chronic Lymphocytic Leukemia. *Nat Genet* (2013) 45(8):868–76. doi: 10.1038/ng.2652
30. Crowther-Swanepoel D, Broderick P, Di Bernardo MC, Dobbins SE, Torres M, Mansouri M, et al. Common Variants at 2q37.3, 8q24.21, 15q21.3 and 16q24.1 Influence Chronic Lymphocytic Leukemia Risk. *Nat Genet* (2010) 42 (2):132–U159. doi: 10.1038/ng.510
31. Sava GP, Speedy HE, Di Bernardo MC, Dyer MJS, Holroyd A, Sunter NJ, et al. Common Variation at 12q24.13 (OAS3) Influences Chronic Lymphocytic Leukemia Risk. *Leukemia* (2015) 29(3):748–51. doi: 10.1038/leu.2014.311
32. Slager SL, Rabe KG, Achenbach SJ, Vachon CM, Goldin LR, Strom SS, et al. Genome-Wide Association Study Identifies a Novel Susceptibility Locus at 6p21.3 Among Familial CLL. *Blood* (2011) 117(6):1911–6. doi: 10.1182/blood-2010-09-308205
33. Speedy HE, Di Bernardo MC, Sava GP, Dyer MJS, Holroyd A, Wang Y, et al. A Genome-Wide Association Study Identifies Multiple Susceptibility Loci for Chronic Lymphocytic Leukemia. *Nat Genet* (2014) 46(1):56–+. doi: 10.1038/ng.2843
34. Langerak AW, Davi F, Ghia P, Hadzidimitriou A, Murray F, Potter KN, et al. Immunoglobulin Sequence Analysis and Prognostication in CLL: Guidelines From the ERIC Review Board for Reliable Interpretation of Problematic Cases. *Leukemia* (2011) 25(6):979–84. doi: 10.1038/leu.2011.49
35. van Dongen JJ, Langerak AW, Bruggemann M, Evans PAS, Hummel M, Lavender FL, et al. Design and Standardization of PCR Primers and Protocols for Detection of Clonal Immunoglobulin and T-Cell Receptor Gene Recombinations in Suspect Lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* (2003) 17(12):2257–317. doi: 10.1038/sj.leu.2403202
36. Bystry V, Reigl T, Krejci A, Demko M, Hanakova B, Grioni A, et al. ARResT/Interrogate: An Interactive Immunoprofiler for IG/TR NGS Data. *Bioinformatics* (2017) 33(3):435–7. doi: 10.1093/bioinformatics/btw634
37. van der Velden VHJ, Bruggemann M, Cazzaniga G, Scheijen B, Tops B, Trka J, et al. Potential and Pitfalls of Whole Transcriptome-Based Immunogenetic Marker Identification in Acute Lymphoblastic Leukemia; a EuroMRD and EuroClonality-NGS Working Group Study. *Leukemia* (2021) 35(3):924–8. doi: 10.1038/s41375-021-01154-z
38. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res* (2018) 28(11):1747–56. doi: 10.1101/gr.239244.118
39. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res* (2019) 47 (D1):D941–7. doi: 10.1093/nar/gky1015
40. Fernhout F, Dinkelaar RB, Hagemeijer A, Groeneveld K, van Kammen E, van Dongen JJ. Four Aged Siblings With B Cell Chronic Lymphocytic Leukemia. *Leukemia* (1997) 11(12):2060–5. doi: 10.1038/sj.leu.2400874
41. Kriangkum J, Motz SN, Mack T, Beiggi S, Baigorri E, Kuppusamy H, et al. Single-Cell Analysis and Next-Generation Immuno-Sequencing Show That Multiple Clones Persist in Patients With Chronic Lymphocytic Leukemia. *PloS One* (2015) 10(9):e0137232. doi: 10.1371/journal.pone.0137232
42. Hengeveld PJ, Levin MD, Kolijn PM, Langerak AW. Reading the B-Cell Receptor Immunome in Chronic Lymphocytic Leukemia: Revelations and Applications. *Exp Hematol* (2021) 93:14–24. doi: 10.1016/j.exphem.2020.09.194
43. Raponi S, Ilari C, Della Starza I, Cappelli LV, Cafforio L, Piciocchi A, et al. Redefining the Prognostic Likelihood of Chronic Lymphocytic Leukaemia Patients With Borderline Percentage of Immunoglobulin Variable Heavy Chain Region Mutations. *Br J Haematol* (2020) 189(5):853–9. doi: 10.1111/bjh.16434
44. Gemenetzi K, Psomopoulos F, Carriles AA, Gounari M, Minici C, Plevova K, et al. Higher-Order Immunoglobulin Repertoire Restrictions in CLL: The Illustrative Case of Stereotyped Subsets 2 and 169. *Blood* (2021) 137(14):1895–904. doi: 10.1182/blood.2020005216
45. Marincevic M, Cahill N, Gunnarsson R, Isaksson A, Mansouri M, Göransson H, et al. High-Density Screening Reveals a Different Spectrum of Genomic Aberrations in Chronic Lymphocytic Leukemia Patients With 'Stereotyped' IGHV3-21 and IGHV4-34 B-Cell Receptors. *Haematologica* (2010) 95 (9):1519–25. doi: 10.3324/haematol.2009.021014
46. Fraschilla I, Jeffrey KL. The Speckled Protein (SP) Family: Immunity's Chromatin Readers. *Trends Immunol* (2020) 41(7):572–85. doi: 10.1016/j.it.2020.04.007
47. Leeksma AC, Baliakas P, Moysiadis T, Puiggros A, Plevova K, Van der Kevie-Kersemaekers A-M, et al. Genomic Arrays Identify High-Risk Chronic Lymphocytic Leukemia With Genomic Complexity: A Multi-Center Study. *Haematologica* (2021) 106(1):87–97. doi: 10.3324/haematol.2019.239947
48. Rack KA, van den Berg E, Haferlach C, Beverloo HB, Costa D, Espinet B, et al. European Recommendations and Quality Assurance for Cytogenomic Analysis of Haematological Neoplasms: Reponse to the Comments From the Francophone Group of Hematological Cytogenetics (GFCH). *Leukemia* (2020) 34(8):2262–4. doi: 10.1038/s41375-020-0736-x
49. Schoumans J, Suela J, Hastings R, Muehlematter D, Rack K, van den Berg E, et al. Guidelines for Genomic Array Analysis in Acquired Haematological Neoplastic Disorders. *Genes Chromosomes Cancer* (2016) 55(5):480–91. doi: 10.1002/gcc.22350
50. Kleiblova P, Stolarova L, Krizova K, Lhota F, Hojny J, Zemankova P, et al. Identification of Deleterious Germline CHEK2 Mutations and Their Association With Breast and Ovarian Cancer. *Int J Cancer* (2019) 145 (7):1782–97. doi: 10.1002/ijc.32385
51. Rudd MF, Sellick GS, Webb EL, Catovsky D, Houlston RS. Variants in the ATM-BRCA2-CHEK2 Axis Predispose to Chronic Lymphocytic Leukemia. *Blood* (2006) 108(2):638–44. doi: 10.1182/blood-2005-12-5022
52. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, et al. Whole Genomes Redefine the Mutational Landscape of Pancreatic Cancer. *Nature* (2015) 518(7540):495–501. doi: 10.1038/nature14169
53. Yang X, Fu YJ, Hu FQ, Luo XL, Hu JB, Wang GH. PIK3R3 Regulates PPAR Alpha Expression to Stimulate Fatty Acid Beta-Oxidation and Decrease Hepatosteatosis. *Exp Mol Med* (2018) 50:e431–e431. doi: 10.1038/emm.2017.243
54. Moore NS, Aldubayan SH, Taylor-Weiner A, Stilgenbauer S, Getz G, Wu CJ, et al. Inherited DNA Repair and Cell Cycle Gene Defects in Chronic Lymphocytic Leukemia. *J Clin Oncol* (2019) 37(15):1508–1508. doi: 10.1200/JCO.2019.37.15_suppl.1508
55. Derenzini E, Mazzara S, Melle F, Motta G, Fabbri M, Bruna R, et al. A 3-Gene Signature Based on MYC, BCL-2 and NFKBIA Improves Risk Stratification in Diffuse Large B-Cell Lymphoma. *Haematologica* (2020). doi: 10.3324/haematol.2019.236455
56. Wu B, Slabicki M, Sellner L, Dietrich S, Liu X, Jethwa A, et al. MED12 Mutations and NOTCH Signalling in Chronic Lymphocytic Leukaemia. *Br J Haematol* (2017) 179(3):421–9. doi: 10.1111/bjh.14869
57. Improgo MR, Tesar B, Klitgaard JL, Magori-Cohen R, Yu L, Kasar S, et al. MYD88 L265P Mutations Identify a Prognostic Gene Expression Signature and a Pathway for Targeted Inhibition in CLL. *Br J Haematol* (2019) 184 (6):925–36. doi: 10.1111/bjh.15714

58. Herling CD, Klaumunzer M, Rocha CK, Altmüller J, Thiele H, Bahlo J, et al. Complex Karyotypes and KRAS and POT1 Mutations Impact Outcome in CLL After Chlorambucil-Based Chemotherapy or Chemoimmunotherapy. *Blood* (2016) 128(3):395–404. doi: 10.1182/blood-2016-01-691550
59. Vendramini E, Bomben R, Pozzo F, Benedetti D, Bittolo T, Rossi FM, et al. KRAS, NRAS, and BRAF Mutations Are Highly Enriched in Trisomy 12 Chronic Lymphocytic Leukemia and Are Associated With Shorter Treatment-Free Survival. *Leukemia* (2019) 33(8):2111–5. doi: 10.1038/s41375-019-0444-6
60. Hoofd C, Huang SJ, Gusscott S, Lam S, Wong R, Johnston A, et al. Ultrasensitive Detection of NOTCH1 C.7544_7545delct Mutations in Chronic Lymphocytic Leukemia by Droplet Digital PCR Reveals High Frequency of Subclonal Mutations and Predicts Clinical Outcome in Cases With Trisomy 12. *J Mol Diagn* (2020) 22(4):571–8. doi: 10.1016/j.jmoldx.2020.01.008
61. Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, et al. Whole-Genome Sequencing Reveals Activation-Induced Cytidine Deaminase Signatures During Indolent Chronic Lymphocytic Leukaemia Evolution. *Nat Commun* (2015) 6:1–12. doi: 10.1038/ncomms9866
62. Klein U, Dalla-Favera R. New Insights Into the Pathogenesis of Chronic Lymphocytic Leukemia. *Semin Cancer Biol* (2010) 20(6):377–83. doi: 10.1016/j.semcancer.2010.10.012
63. Ghia EM, Jain S, Widhopf GF, Rassenti LZ, Keating MJ, Wierda WG, et al. Use of IGHV3-21 in Chronic Lymphocytic Leukemia Is Associated With High-Risk Disease and Reflects Antigen-Driven, Post-Germinal Center Leukemogenic Selection. *Blood* (2008) 111(10):5101–8. doi: 10.1182/blood-2007-12-130229
64. Ecker V, Stumpf M, Brandmeier L, Neumayer T, Pfeuffer L, Engleitner T, et al. Targeted PI3K/AKT-Hyperactivation Induces Cell Death in Chronic Lymphocytic Leukemia. *Nat Commun* (2021) 12(1):3526. doi: 10.1038/s41467-021-23752-2
65. Goldin LR, Slager SL. Familial CLL: Genes and Environment. *Hematol Am Soc Hematol Educ Program* (2007) 1:339–45. doi: 10.1182/asheducation-2007.1.339

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Oncology

ORIGINAL RESEARCH
published: 26 October 2021
doi: 10.3389/fonc.2021.723722

The Hydropathy Index of the HCDR3 Region of the B-Cell Receptor Identifies Two Subgroups of IGHV-Mutated Chronic Lymphocytic Leukemia Patients With Distinct Outcome

Arancha Rodríguez-Caballero[1,2,3*], *Blanca Fuentes Herrero*[1,2,3], *Guillermo Oliva Ariza*[1,2,3], *Ignacio Criado*[1,2,3], *Miguel Alcoceba*[2,3,4], *Carlos Prieto*[5], *María Pérez Caro*[6], *Andrés C. García-Montero*[1,2,3], *Marcos González Díaz*[2,3,4], *Francesco Forconi*[7], *Ana Bela Sarmento-Ribeiro*[8,9], *Julia Almeida*[1,2,3†] *and Alberto Orfao*[1,2,3*†]

[1] Translational and Clinical Research Program, Cancer Research Center Institute of Cancer Molecular and Cellular Biology (IBMCC), University of Salamanca-The Spanish National Research Council (USAL-CSIC), Department of Medicine and Cytometry Service, Nucleus Research Support Platform from University of Salamanca (NUCLEUS), University of Salamanca, Salamanca, Spain, [2] CIBERONC Program of Liquid Biopsy, Hematologic Tumors, Centro de Investigación Biomédica en Red de Cáncer CB16/12/00400 and CB16/12/00233 (CIBERONC), Madrid, Spain, [3] Molecular and Cellular Biology of Hematologic Tumors, Biomedical Research Institute of Salamanca (IBSAL), Salamanca, Spain, [4] Department of Hematology, University Hospital of Salamanca/Biomedical Research Institute of Salamanca (HUS/IBSAL), Salamanca, Spain, [5] Bioinformatics Service, Nucleus Research Support Platform from University of Salamanca (NUCLEUS), University of Salamanca, Salamanca, Spain, [6] Spanish National DNA Bank Carlos III, University of Salamanca, Salamanca, Spain, [7] Haematology Oncology Group, School of Cancer Sciences, Cancer Research UK Centre and National Institute for Health Research Experimental Cancer Medicine, University of Southampton, Faculty of Medicine, Southampton, United Kingdom, [8] Faculty of Medicine, University of Coimbra (FMUC), Coimbra, Portugal, [9] Centro Hospitalar e Universitário de Coimbra (CHUC), Coimbra, Portugal

The HCDR3 sequences of the B-cell receptor (BCR) undergo constraints in length, amino acid use, and charge during maturation of B-cell precursors and after antigen encounter, leading to BCR and antibodies with high affinity to specific antigens. Chronic lymphocytic leukemia consists of an expansion of B-cells with a mixed immature and "antigen-experienced" phenotype, with either a mutated (M-CLL) or unmutated (U-CLL) tumor BCR, associated with distinct patient outcomes. Here, we investigated the hydropathy index of the BCR of 138 CLL patients and its association with the IGHV mutational status and patient outcome. Overall, two clearly distinct subgroups of M-CLL patients emerged, based on a neutral (mean hydropathy index of -0.1) vs. negatively charged BCR (mean hydropathy index of -1.1) with molecular features closer to those of B-cell precursors and peripheral/mature B-cells, respectively. Despite that M-CLL with neutral HCDR3 did not show traits associated with a mature B-cell repertoire, important differences in IGHV gene usage of tumor cells and patient outcome were observed in this subgroup of patients once compared to both U-CLL and M-CLL with negatively charged HCDR3 sequences. Compared to M-CLL with negatively charged HCDR3 sequences, M-CLL with neutral HCDR3 sequences showed predominance of men, more advanced stages of the disease, and a greater frequency of genetic alterations—e.g., del(17p)—together with a higher rate of disease progression and shorter time to therapy (TTT), independently of other

OPEN ACCESS

Edited by:
Franco Fais, University of Genoa, Italy

Reviewed by:
Nicholas Chiorazzi, Feinstein Institute for Medical Research, United States
Manlio Ferrarini, University of Genoa, Italy

***Correspondence:**
Alberto Orfao
orfao@usal.es
Arancha Rodríguez-Caballero
arocab@usal.es

[†]*These authors have contributed equally to this work and share last authorship*

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: *11 June 2021*
Accepted: *05 October 2021*
Published: *26 October 2021*

Citation:
Rodríguez-Caballero A, Fuentes Herrero B, Oliva Ariza G, Criado I, Alcoceba M, Prieto C, Pérez Caro M, García-Montero AC, González Díaz M, Forconi F, Sarmento-Ribeiro AB, Almeida J and Orfao A (2021) The Hydropathy Index of the HCDR3 Region of the B-Cell Receptor Identifies Two Subgroups of IGHV-Mutated Chronic Lymphocytic Leukemia Patients With Distinct Outcome. Front. Oncol. 11:723722. doi: 10.3389/fonc.2021.723722

prognostic factors. Our data suggest that the hydropathy index of the HCDR3 sequences of CLL cells allows the identification of a subgroup of M-CLL with intermediate prognostic features between U-CLL and the more favorable subgroup of M-CLL with a negatively charged BCR.

Keywords: hydropathy index, neutral HCDR3, negatively charged HCDR3, mutated CLL (M-CLL), disease progression

INTRODUCTION

Chronic lymphocytic leukemia (CLL) is the most prevalent leukemia in adults in the Western world, which is characterized by an expansion of mature-appearing $CD5^+CD20^{lo}$ B-cells showing an antigen-experienced $CD27^+$, IgM^+, and/or IgD^+ unswitched phenotype, in association with either an unmutated (U-CLL) or mutated (M-CLL) B-cell receptor (BCR) (1). At diagnosis, most CLL patients show stable disease with a variable number of tumor B-cells in blood (always >5,000 cells/μl) and bone marrow (BM), in the absence of organomegalies, and they do not require active therapy (2). Despite this, a significant fraction of patients shows more advanced disease already at diagnosis or they experience disease progression during follow-up, which translates into the need for active cytotoxic therapy (3).

In the last decades, the mutational status of the immunoglobulin (IG) heavy-chain variable (IGHV) genes that code for the BCR, together with disease stage and tumor cytogenetics, has emerged among other variables, as relevant prognostic factors in CLL (4). Thus, U-CLL patients show a significantly poorer outcome compared to M-CLL (4). Thereby, analysis of the IGHV status is currently part of the core variables investigated in the diagnostic workup of this disease (3, 5). Despite this, M-CLL patients have a heterogeneous outcome (6, 7).

From a pathogenic point of view, U-CLL cells resemble "pre-germinal center" (pre-GC) B-cells, whereas M-CLL cells mimic "post-GC" B-lymphocytes (8, 9). However, tumor cells from both CLL groups typically display a mixed immature ($CD5^+$ $CD23^+$) and "antigen-experienced" ($CD27^+$) B-cell phenotype (10), suggesting they might represent the leukemic counterpart of B-lymphocytes that might have undergone BCR stimulation in the GC (M-CLL) vs. peripheral tissues, following selection of B-cell precursors in BM (U-CLL). In line with this hypothesis, the IGHV1-69/IGHJ6 genes which show highly similar junctional regions to those of normal peripheral blood (PB) $CD5^+$ GC B-cells are more frequently represented among U-CLL, supporting a close relationship between U-CLL cells and the B-cells responsible for the natural antibody repertoire (11). This potential relationship is further supported by the fact that most normal $CD5^+$ B-cells isolated from blood correspond to immature and (early) naïve B-cells that express unmutated VH gene regions (12). In turn, B-cell activation *via* T-cell-dependent antigens leads to the expansion of hypermutated germinal center (GC)-derived B-cells (13), suggesting that M-CLL might be associated with the "classical unswitched memory B-cell" compartment, despite that some M-CLL also show BCR features that overlap with those of natural antibodies (14).

Another important biological feature of CLL is the usage of a biased IGHV-D-J repertoire (the so-called "stereotyped" BCR) (9) in around one-third of cases, particularly in U-CLL patients, with important pathogenic and prognostic implications (15, 16). In contrast to U-CLL, the higher load of somatic mutations in the BCR of M-CLL cases makes recognition of common amino acid (aa) patterns in the HCDR3 region more difficult (17). However, other HCDR3 characteristics, such as its overall charge and hydropathy index, might also contribute to better understand the ontogeny of tumor B-cells in CLL, the affinity and specificity profile of their BCR, and its relationship with antigen-driven B-cell responses, even at earlier stages of B-cell maturation (18). In fact, the HCDR3 sequence of the BCR undergoes constraints in length, amino acid use, and charge along the B-cell development and maturation (19). Consequently, the BCR repertoire of early B-cell progenitors is first focused into what appears to be a preferred range for functional antigen recognition by mature B-cells, and subsequently modified after antigen recognition, in order to generate high-affinity antigen-specific antibodies and memory B-cells (19, 20). Interestingly, receptor prototypes based on HCDR3 charge and its association with certain V gene characteristics have been defined in CLL cells with the possibility that such receptor restrictions could reflect selections of the BCR repertoire that have occurred among both antigen-experienced and naive B cells (21). Despite this, the hydropathy features of HCDR3 and its association with the IGHV mutational status and other clinical and biological features of the disease have not been systematically explored in large series of CLL and related with patient outcome.

Here we investigated the hydropathy index of the BCR of tumor cells from 138 CLL patients, and its potential association with other features of the disease, including the BCR mutational status and patient outcome.

MATERIALS AND METHODS

Patients and Samples

A total of 138 untreated CLL patients—81 males and 57 females; median age (range) at diagnosis of 63 years (y) (33–84 y)—diagnosed at the University Hospital of Salamanca (Salamanca, Spain) were studied. Most cases (95/138) had Binet stage A CLL, and 43 had more advanced CLL (Binet B, 22; and Binet C, 21 patients). Median follow-up at the time of study closure was 8 y; at that time, 65 patients (47%) had progressed and required therapy and 24 (17%) had died (**Table 1**). In every patient, genomic DNA (gDNA) from purified CLL B-cells was obtained for molecular investigations. The study was approved by the local institutional Ethics Committee (approval code: CEIC-PI4705/2017). All patients gave their written informed consent to participate to the study in agreement with the Declaration of Helsinki.

TABLE 1 | Clinical and biological features of CLL patients classified according to the HCDR3 hydropathy index.

Patient features	CLL with neutral HCDR3 (N = 65)	CLL with negatively charged HCDR3 (N = 73)	*p-value*
Age at diagnosis (years)	63 (33–84)	65 (35–84)	0.47
Men/women#	47/18 (72%/28%)	34/39 (47%/53%)	0.002
Binet stage#			
A	37 (57%)	58 (79%)	0.006
B	12 (18%)	10 (14%)	
C	16 (25%)	5 (7%)	
Rai stage#			
0	30/65 (46%)	54/73 (74%)	0.01
I	6/65 (9%)	7/73 (10%)	
II	9/65 (14%)	5/73 (7%)	
III	6/65 (9%)	2/73 (3%)	
IV	14/65 (21%)	5/73 (7%)	
Hemoglobin (g/L)	140 (70–180)	130 (90–170)	0.94
Anemia (<100 g of hemoglobin/L)#	7/65 (11%)	3/71 (4%)	0.13
N. of platelets ($\times 10^9$/L)	154 (15–429)	157 (48–448)	0.11
Thrombocytopenia (<100 × 10^9 platelets/L)#	14/65 (21%)	5/70 (7%)	0.01
N. of PB leukocytes ($\times 10^9$/L)	36 (6–352)	34 (8–576)	0.40
N. of PB total T-cells ($\times 10^9$/L)	3.2 (0.8–7.4)	2.8 (0.5–14)	0.68
N. of PB $CD4^+$ T-cells ($\times 10^9$/L)	1.6 (0.5–4.4)	1.7 (0.3–5.9)	0.64
N. of PB $CD8^+$ T-cells ($\times 10^9$/L)	0.8 (0.1–3.5)	0.9 (0.1–8)	0.60
N. of PB monocytes ($\times 10^9$/L)	0.6 (0.01–2)	0.4 (0.01–2.9)	0.24
N. of PB neutrophils ($\times 10^9$/L)	6.4 (0.9–20.7)	6.1 (0.2–19.4)	0.50
N. of PB basophils ($\times 10^9$/L)	0.07 (0.005–0.5)	0.06 (0.015–0.6)	1.0
N. of PB eosinophils ($\times 10^9$/L)	0.2 (0.02–1.7)	0.2 (0.01–1.2)	0.16
N. of PB NK cells ($\times 10^9$/L)	0.6 (0.04–2.1)	0.5 (0.04–3.2)	0.60
Tumor B-cell clone size in blood ($\times 10^9$/L)	41.4 (1.4–334.8)	20.9 (0.8–278.9)	0.11
U-CLL/M-CLL#	35 (54%)/30 (46%)	25 (34%)/48 (66%)	0.02
% IGHV homology with germ line counterpart	99% (86%–100%)	96% (85%–100%)	0.04
HCDR3 length of CLL clone (N. of aa)	19 (8–28)	16 (7–25)	0.004
Cytogenetically altered CLL patients#	55/65 (85%)	59/72 (82%)	0.43
Del(13q14)(*D13S25*)#	22/65 (34%)	32/72 (44%)	0.14
Trisomy12#	16/65 (25%)	11/72 (15%)	0.12
Del(11q)(*ATM*)#	6/65 (9%)	6/72 (10%)	0.54
Del(17p)#	3/62 (5%)	1/69 (1%)	0.29
N. of cytogenetically altered CLL cells ($\times 10^9$/L)	19 (0–233)	11 (0–185)	0.19
Disease progression#	39/65 (60%)	26/73 (36%)	0.003
TTT < 2 years#	20/65 (31%)	8/73 (11%)	0.004
Deaths#	11/65 (17%)	13/73 (18%)	0.54

Results expressed as median (range) or as # number of cases (percentage).

aa, amino acids; CLL, chronic lymphocytic leukemia; M, mutated IGHV; N, number; PB, peripheral blood; TTT, time to therapy; U, unmutated IGHV.

IGHV-D-J Gene Rearrangement Studies

Analysis of the tumor *IGHV-D-J* gene rearrangements was performed by polymerase chain reaction (PCR) of gDNA from fluorescence-activated cell sorting (FACS)-purified tumor CLL cells according to the ERIC protocols (22), as previously described in detail (23, 24). For *IGHV* sequencing, PCR amplicons were subjected to direct sequencing on both strands. Sequence data were analyzed using the IMGT databases and the IMGT/V-QUEST tool (http://www.imgt.org). Classification into the U-CLL vs. M-CLL categories was based on the well-established 98% cutoff identity to the germline sequence (U-CLL: 98%–100%; M-CLL: <98%) (21).

Calculation of the Hydropathy Index of HCDR3 Protein Sequences

To determine the hydropathy index—grand average of hydropathy (GRAVY) score—of the HCDR3 protein sequences, the ProtScale Tool from the ExPASy Bioinformatics Resource Portal (https://web.expasy.org/protscale/) and the amino acid (aa) scale values, as defined by Kyte and Doolittle (25), were used (**Supplementary Figures 1A, B**). The Gravy score (GS) was calculated for each HCDR3 sequence by summing up the hydropathy index value of each amino acid residue in the individual HCDR3 sequences and dividing the sum obtained by the number of amino acids in each specific sequence (23) (**Supplementary Table 1** and **Supplementary Figures 1B, C**). Since the HCDR3 hydropathy index in humans follows a Gaussian distribution centered in the neutral/hydrophilic range (average charge: -0.5) (19), each HCDR3 sequence was classified into the neutral HCDR3 (GS ≥ -0.5) or negatively charged HCDR3 (GS < -0.5) categories.

Cytogenetic Analyses

The most common cytogenetic alterations associated with CLL—i.e., del(13q14), trisomy 12, del(11q) (*ATM*), and del(17p) (*TP53*)—were investigated by iFISH on *FACS*-purified (single) tumor B cells (≥95% purity), as described elsewhere (23, 26).

Statistical Methods

For all continuous variables, median (range) values were calculated, while for categorical variables, frequencies were used. Either the Mann–Whitney U test or the chi-squared test was used to establish the statistical significance of differences observed between two groups, for continuous and categorical variables, respectively. To avoid associations occurring by chance due to multiple simultaneous comparisons, p-values were Bonferroni-adjusted for comparisons of continuous variables among three study groups. Time from diagnosis to (first) therapy (TTT) curves were built using the Kaplan–Meier method and compared by the (one-sided) log-rank test. Receiver operating characteristic (ROC) curve analysis was performed to determine the cutoff values of continuous variables that best distinguish disease progression. Multivariate analysis using the Cox regression model was performed to identify those variables independently associated with a greater/lower risk of disease progression (in terms of TTT) among M-CLL patients. All statistical analyses were performed with SPSS 25.0 (SPSS-IBM, Armonk, NY), and statistical significance was set at p-values ≤ 0.05, unless Bonferroni-adjusted p-values were applied (≤0.013).

RESULTS

HCDR3 Hydropathy Index in CLL and Its Relationship With Other Disease Features

Overall, a similar frequency of CLL patients with neutral (GS ≥ -0.5) and negatively charged (GS < -0.5) HCDR3 amino acid sequences was observed in our cohort: 65/138 (47%) *vs*. 73/138 (53%) CLL patients (p = 0.46), respectively (**Figure 1**). CLL patients with neutral or negatively charged HCDR3 showed no significant (p > 0.05) differences in age distribution, hemoglobin levels, leukocyte or platelet counts, CLL cell counts in blood, and the overall tumor cell cytogenetic profiles (**Table 1**).

In contrast, CLL patients with neutral HCDR3 sequences showed a significant predominance of men *vs*. women (72% *vs*. 47%, p = 0.002), together with a lower percentage of Rai stage 0 (46% *vs*. 74%, p = 0.01) and Binet stage A cases (57% *vs*. 79%, p = 0.006), a higher proportion of cases with thrombocytopenia (21% of cases *vs*. 7%, p = 0.01), a lower proportion of M-CLL cases (46% *vs*. 66%, p = 0.02), and a lower median percentage of IGHV mutations (1% *vs*. 4%, p = 0.04) with longer HCDR3 sequences (median: 19 *vs*. 16 amino acids, p = 0.004) compared to CLL patients with negatively charged HCDR3 (**Table 1**). This CLL profile with neutral HCDR3 sequences translated into a significantly (p = 0.003) higher risk of disease progression (60% *vs*. 36%) and thereby also a higher percentage of cases that had required therapy at 2 y from diagnosis (31% *vs*. 11%, p = 0.004) (**Table 1**), and a significantly shortened TTT—median (95% confidence interval): 6 y (4–8 y) *vs*. not reached, p = 0.003 (**Figure 2A**). Of note, the prognostic impact of the HCDR3 hydropathy index was specifically restricted to M-CLL patients, while it did not show an impact on the already poorer outcome of U-CLL cases (**Figure 2B**).

HCDR3 Hydropathy Index in M-CLL *vs*. U-CLL and Its Association With Other Disease Features

Based on the above findings, we subdivided M-CLL patients into cases with neutral HCDR3 (mean GS of -0.1) and patients with negatively charged HCDR3 sequences (mean GS of -1.1) and

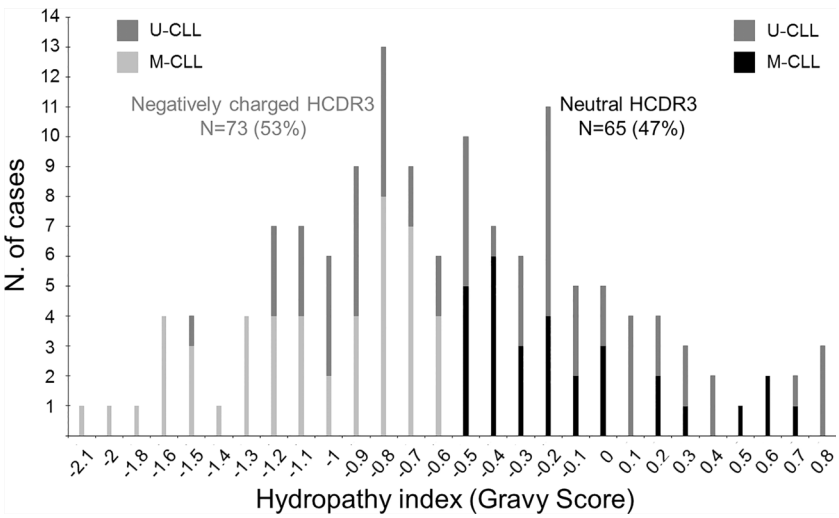


FIGURE 1 | Distribution of CLL patients according to the hydropathy index—Gravy score (GS)—of the HCDR3 aa sequence of their BCR. Bars represent the number of CLL patients (N = 138) with different HCDR3 GS. Black bars correspond to M-CLL patients with a GS ≥ -0.5 (neutral HCDR3; N = 65), light gray bars correspond to M-CLL patients with GS < -0.5 (negatively charged HCDR3; N = 73), and dark gray bars represent cases with a ≥98% identity to the V(H) germline (U-CLL) independently of their GS.

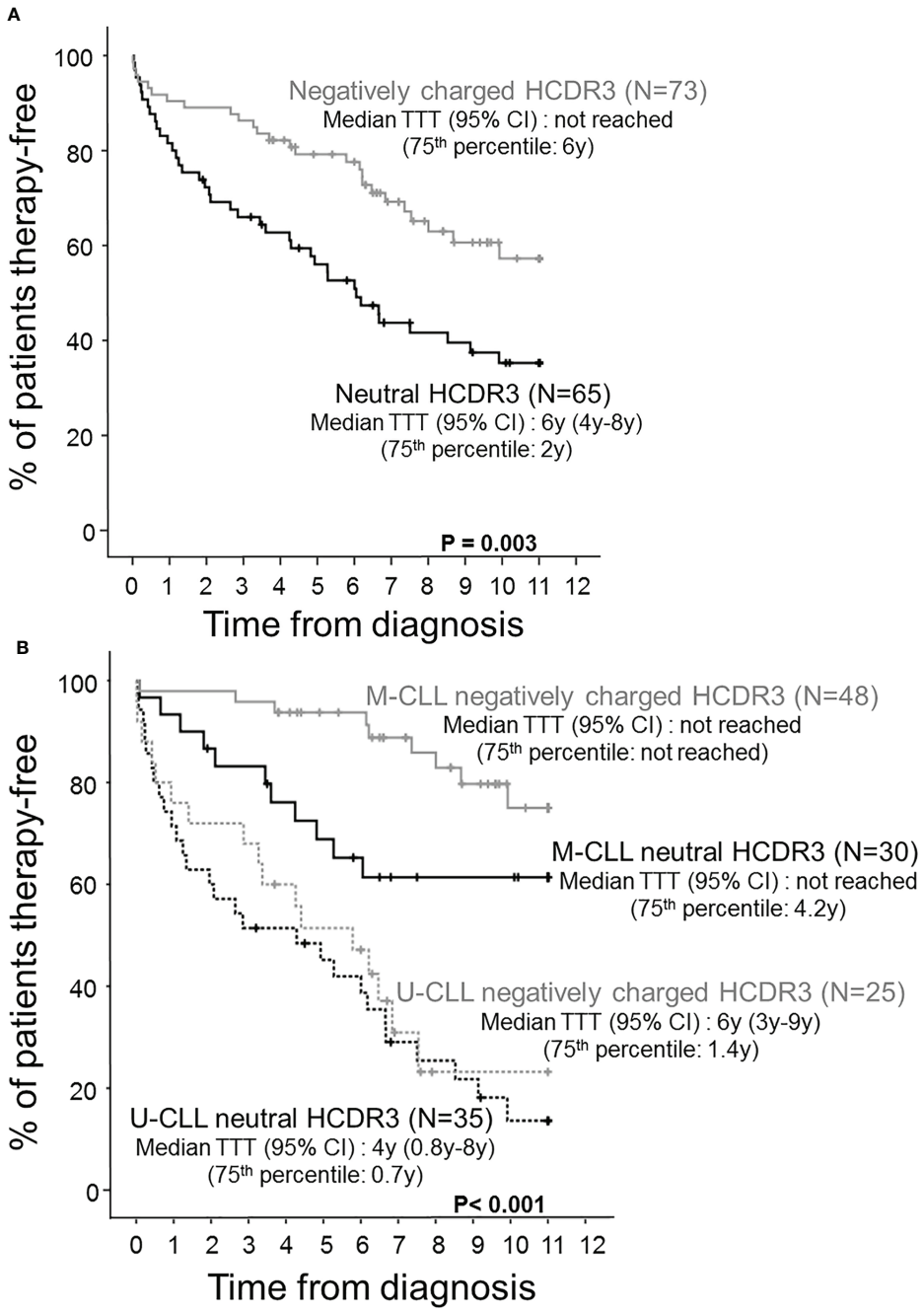


FIGURE 2 | Time to therapy (TTT) survival curves of CLL patients distributed according to the hydropathy index of the HCDR3 aa sequence and the mutational status of their tumor cell BCR. Prognostic impact of the HCDR3 hydropathy index **(A)** and both the HCDR3 hydropathy index and the IGHV mutational status **(B)** on the outcome of CLL patients assessed by their survival from diagnosis to first therapy (TTT).

compared the features of these two subgroups of M-CLL *vs.* U-CLL cases (**Table 2**). Thus, Rai stage 0 (p = 0.007) predominated in the two M-CLL patient subgroups *vs.* U-CLL (**Table 2**). In contrast, greater median hemoglobin levels (150 *vs.* 130 g/L, p = 0.004) were found in M-CLL patients with a neutral HCDR3 (but not within those with a negatively charged HCDR3) *vs.* U-CLL. Overall, the number of PB leukocytes, total T cells, $CD8^+$ T-cells, and tumor CLL cells, in blood, were all significantly increased in U-CLL compared with the two M-CLL patient groups, in the absence of significant differences between the later M-CLL groups (**Table 2**). Despite this, M-CLL with neutral HCDR3 sequences showed an intermediate frequency of cytogenetically

altered cases between U-CLL and M-CLL with negatively charged HCDR3 (80% *vs*. 92% and 74%, respectively) together with a significantly greater proportion of del(17p)$^+$ patients (11% *vs*. 0% and 2%, respectively; p = 0.02) (**Table 2**).

Regarding outcome, M-CLL with neutral HCDR3 sequences showed an intermediate rate of disease progression (37%) compared to both U-CLL patients (75%) (p < 0.001) and M-CLL with negatively charged HCDR3 sequences (19%), after a similar median follow-up (**Table 2**). This was associated with a significantly lower percentage of M-CLL cases with negatively charged HCDR3 sequences that required therapy during the first 2 years after diagnosis (2%) compared to U-CLL (37%, p < 0.001) and M-CLL with a neutral HCDR3 (17%, p = 0.03) (**Table 2**). This translated into significantly prolonged TTT among M-CLL with negatively charged HCDR3 sequences compared to both M-CLL patients with a neutral HCDR3 sequence and U-CLL patients—75th percentile TTT (95% confidence interval): not reached *vs*. 4.2 and 0.9 y, respectively; p < 0.001) (**Figure 2B**).

Based on the results above, we specifically investigated the prognostic impact of the hydropathy index of the HCDR3 sequence of the tumor cell BCR compared to other clinical and laboratory variables in patients with M-CLL. Among all variables analyzed, Binet stage (p < 0.001), the number of total T-cells (p < 0.001), CD4$^+$ T-cells (p = 0.002), CD8$^+$ T-cells (p = 0.001), basophils (p = 0.04), the size of the tumor B-cell clone in blood (p = 0.003), del(11q) and/or del(17p) (p = 0.04) and the number of cytogenetically altered CLL cells (p = 0.001) in addition to the hydropathy index of the HCDR3 sequences of the tumor B-cell clone (p < 0.001) all showed a prognostic impact in the univariate analysis (**Table 3**). Multivariate analysis confirmed the

TABLE 2 | Clinical and biological features of CLL patients classified according to their IGHV mutational status and the HCDR3 hydropathy index.

Patient features	U-CLL (N = 60)	M-CLL (N = 78)		p-value
		Neutral HCDR3 (N = 30)	Negatively charged HCDR3 (N = 48)	
Age at diagnosis (years)	64 (33–84)	61 (38–81)	63 (35–84)	0.21
Men/women[#]	38/22 (63%/37%)	20/10 (67%/33%)	23/25 (48%/52%)	0.16; 0.08[b,c]
Binet stage#				
A	37 (62%)	21 (70%)	37 (77%)	0.14
B	14 (23%)	2 (7%)	6 (13%)	
C	9 (15%)	7 (23%)	5 (10%)	
Rai stage[#]				
0	29/60 (48%)	19/30 (63%)	36/48 (75%)	0.007; 0.02[a]; 0.04[b]
I	6/60 (10%)	4/30 (13%)	3/48 (6%)	
II	12/60 (20%)	0/30 (0%)	2/48 (4%)	
III	6/60 (10%)	0/30 (0%)	2/48 (4%)	
IV	7/60 (12%)	7/30 (23%)	5/48 (10%)	
Hemoglobin (g/L)	130 (70–170)	150 (100–180)	130 (90–170)	0.010; 0.004[a]
Anemia (<100 g of hemoglobin/L)[#]	7/59 (12%)	0/30 (0%)	3/47 (6%)	0.12
N. of platelets (×10^9/L)	149 (15–448)	158 (62–429)	164 (48–344)	0.63
Thrombocytopenia (<100 × 10^9 platelets/L)[#]	7/58 (12%)	7/30 (23%)	5/47 (11%)	0.25
N. of PB leukocytes (×10^9/L)	55 (10–576)	23 (6–245)	26 (8–241)	0.001; 0.010[b]
N. of PB total T-cells (×10^9/L)	4 (0.5–14)	2.7 (0.9–7.4)	2.6 (0.6–9)	0.006; 0.003[b]
N. of PB CD4$^+$ T-cells (×10^9/L)	1.8 (0.6–6)	1.5 (0.5–4)	1.5 (0.3–4.3)	0.04
N. of PB CD8$^+$ T-cells (×10^9/L)	1.4 (0.1–8)	0.7 (0.2–5)	0.8 (0.1–5)	0.007; 0.005[a]
N. of PB monocytes (×10^9/L)	0.7 (0.01–3)	0.5 (0.02–1.4)	0.4 (0.01–2)	0.12
N. of PB neutrophils (×10^9/L)	6.3 (0.9–21)	6.3 (0.9–16)	6.0 (0.2–12)	0.47
N. of PB basophils (×10^9/L)	0.07 (0.005–0.6)	0.06 (0.02–0.5)	0.06 (0.01–0.3)	0.66
N. of PB eosinophils (×10^9/L)	0.2 (0.01–1.2)	0.2 (0.02–1.7)	0.2 (0.05–1)	0.66
N. of PB NK cells (×109/L)	0.7 (0.04–3)	0.6 (0.05–1.5)	0.4 (0.04–2.2)	0.014
Tumor B-cell clone size in blood (×10^9/L)	47 (0.8–335)	16 (1.4–238)	17 (4.4–219)	0.001; 0.011[b]
% IGHV homology with germ line counterpart	99.6 (98.8–100)	93.7 (86–98)	93.3 (85–98)	<0.001; <0.001[a,b]
HCDR3 length of CLL clone (N. of aa)	21 (12–28)	16 (8–24)	16 (7–24)	<0.001;<0.001[a,b]
Cytogenetically altered CLL patients[#]	55/60 (92%)	24/30 (80%)	35/47 (74%)	0.05; 0.02[b]
Del(13q14)(*D13S25*)[#]	25/60 (42%)	11/30 (37%)	18/47 (38%)	0.88
Trisomy12[#]	17/60 (28%)	3/30 (10%)	7/47 (15%)	0.07; 0.04[a]
Del(11q)(*ATM*)[#]	9/60 (15%)	1/30 (3%)	2/47 (4%)	0.07; 0.06[b]
Del(17p)[#]	0/57 (0%)	3/28 (11%)	1/46 (2%)	0.02; 0.03[a]
N. of cytogenetically altered CLL cells (×10^9/L)	28 (0–208)	8.6 (0–233)	6.3 (0–185)	0.12
Disease progression[#]	45/60 (75%)	11/30 (37%)	9/48 (19%)	<0.001; <0.001[a,b]
TTT < 2 years[#]	22/60 (37%)	5/30 (17%)	1/48 (2%)	<0.001; 0.04[a]; <0.001[b]; 0.03[c]
Deaths[#]	16/60 (27%)	2/30 (7%)	6/48 (12%)	0.03; 0.02[a]

Results expressed as median (range) or as # number of cases (percentage).

[a]*U-CLL vs. M-CLL with neutral HCDR3.*

[b]*U-CLL vs. M-CLL with negatively charged HCDR3.*

[c]*M-CLL with neutral HCDR3 vs. M-CLL with negatively charged HCDR3.*

CLL, chronic lymphocytic leukemia; M, mutated IGHV; N, number; PB, peripheral blood; TTT, time to therapy; U, unmutated IGHV.

independent adverse prognostic impact of neutral HCDR3 sequence of BCR (hazard ratio (HR), 12; 95% confidence interval (CI), 1.8 to 81; p = 0.01) together with an advanced Binet stage B/C (HR, 42.8; 95% CI, 1.7 to 1,073; p = 0.02) (**Table 3**).

Distinctive Molecular Features of the BCR of M-CLL With Neutral *vs*. Negatively Charged HCDR3 Sequences

Interestingly, no significant differences were found between the two groups of M-CLL patients defined by having a neutral *vs*. negatively charged HCDR3, as regards the frequency of V(H) gene families used (**Table 4**). Despite this, both groups of M-CLL patients (with neutral and negatively charged HCDR3 sequences) more frequently used the VH3 gene at the expense of a lower frequency of VH1 gene usage compared to U-CLL patients—53% and 52% *vs*. 32%, (p = 0.05) and 17% and 8% *vs*. 47%, (p < 0.001), respectively (**Table 4**). In more detail, usage of the VH1-69 gene family was significantly associated with U-CLL—27% *vs*. 0% and 4%, p < 0.001—while VH4-34 was more frequently used in the two groups of (neutral and negatively charged HCDR3) M-CLL patients *vs*. U-CLL—17% and 21% *vs*. 3%, p = 0.02, respectively. Interestingly, VH3-7 was significantly associated with M-CLL with neutral HCDR3 sequences (17%) while rarely found in U-CLL (2%) (p = 0.01) (**Table 3**). Likewise, usage of the D(H)2 genes was more frequently observed in M-CLL with neutral HCDR3 sequences (43%) than in M-CLL with negatively charged HCDR3 (19%, p = 0.02) and U-CLL (22%, p = 0.03) patients (**Table 4**). Within the D(H)2 gene family, D(H)2-15 and D(H)2-21 were those family members more frequently expressed in M-CLL with neutral HCDR3 sequences *vs*. U-CLL (13% and 20% *vs*. 2% and 5%, respectively) (**Table 4**). In turn, U-CLL showed a higher frequency of D(H)3 than M-CLL patients (p = 0.001), the D(H)3-3 gene family mostly accounting for these differences as it was found in 30% of U-CLL *vs*. 7% of M-CLL with neutral HCDR3 and 12% of M-CLL with negatively charged HCDR3 sequences (p = 0.01) (**Table 4**).

Regarding J(H) gene usage, M-CLL with negatively charged HCDR3 sequences showed a significantly higher frequency of J(H)4 genes (60%) than both M-CLL with a neutral HCDR3 (23%) (p = 0.001) and U-CLL (33%) (p = 0.004) patients (**Table 4**). Likewise, a significantly lower percentage of M-CLL with negatively charged HCDR3 sequences showed J(H)6 gene usage (21%) compared to U-CLL (48%, p = 0.003) (**Table 4**). Interestingly, the lower use of JH6 and DH2 genes in M-CLL with negatively charged HCDR3 sequences was associated with different charges of the specific HCDR3 amino acids comprised by these coding genes *vs*. U-CLL patients (p = 0.010 and p = 0.003, respectively) (**Table 4**). In contrast, the charge of the HCDR3 amino acids comprised by JH4 and DH3 genes did not show differences between U-CLL and both M-CLL groups (**Table 4**). It should be noted, however, that the HCDR3 fraction comprised by nucleotides distinct to those included in the above referred JH and DH gene sequences showed always a significantly lower charge in M-CLL patients with negatively charged sequences compared to M-CLL with neutral HCDR3 sequences (**Table 4**).

As expected, U-CLL had longer HCDR3 (median: 21 amino acids) than M-CLL (median: 16 amino acids) (p < 0.001) (**Table 2**). However, since the length of the BCR HCDR3 sequences differs between B-cells from younger and older subjects (27), we grouped our patients into younger adults (≤65 y) and older (>65 y) patients. Interestingly, older M-CLL patients with a negatively charged HCDR3 showed shorter HCDR3 sequences (median: 15 amino acids) than M-CLL with neutral HCDR3 (median: 18 amino acids) (p = 0.013) (**Table 4**). Furthermore, older M-CLL patients with neutral HCDR3 had similarly longer HCDR3 sequences (median: 18 amino acids) to those of U-CLL patients (median: 20 amino acids) (p = 0.47) (**Table 4**). Finally, M-CLL with negatively charged HCDR3 sequences showed a significantly lower frequency of stereotyped IGHV sequences (10%) compared to U-CLL (28%) (p = 0.02) (**Table 4**), with a similarly low incidence of stereotyped IGHV sequences corresponding to the more aggressive (#1, #2, #8) CLL subsets (**Table 4**).

DISCUSSION

B-cells are a key component of the adaptive immune system (28). Their function is typically triggered through BCR-mediated recognition of specific antigens (28). Specific binding of BCR to antigens is mostly mediated through unique HCDR3 (and also LCDR3) regions capable of identifying and attaching to complementary epitopes in the recognized antigen (28). For adequate binding to the antigen, electrostatic links with the BCR are required (29). Thereby, the HCDR3 charge plays a critical role in antigen binding to the BCR and recognition by B-cells (30). Importantly, during antigen-driven maturation, B-cells modify their HCDR3 sequences to enhance their affinity for specific antigen triggers (30). This includes acquisition of somatic mutations involving the HCDR3 region, which progressively confer more negatively charged amino acid sequences for higher affinity antigen binding by both the BCR and the future B-cell derived (higher-affinity) antibodies (31). In addition, due to its key role in antigen recognition, the interaction of the BCR with the BM microenvironment also plays a critical role at an earlier stage, during lymphopoiesis, in selecting B-cell precursors that carry a functional BCR (31).

For decades now, studies have accumulated which support an important role for BCR-mediated expansion of tumor cells in CLL (32, 33) in the absence of a common genetic driver (6). Thus, CLL cells show biased usage of specific *IGHV(D)J* gene families, with overrepresentation of some genes such as *IGHV1-69*, *IGHV4-34*, and *IGHV3-21* (34). Of note, these genes are differentially distributed among the two major prognostic subgroups of CLL defined according to the mutational status of the BCR (U-CLL and M-CLL) (35). Accordingly, U-CLL cells have polyreactive BCRs that may respond to a wide spectrum of epitopes (36, 37), as typically required during selection of recently produced immature B-lymphocytes in BM (38), whereas M-CLL cells are more mature B-cells that have undergone somatic hypermutation, whose BCRs are (potentially) less responsive to external signals, while more

TABLE 3 | Univariate and multivariate analyses of prognostic factors with an impact on disease progression in M-CLL (N = 78).

Variables	N	Univariate analysis on disease progression		Multivariate analysis on disease progression	
		Median (95% CI) (y)	*p-value*	HR (95% CI)	*p-value*
Age at diagnosis					
≤ 65 y	50	NR	0.12		
> 65 y	28	NR			
Sex					
Men	43	NR	0.88		
Women	35	NR			
Binet stage					
A	58	NR			
B/C	20	8 (1.6–14.4)	<0.001	42.8 (1.7–1,073)	0.02
Anemia					
<100 g hemoglobin/L	3	9.9 (5–15.9)	0.27		
≥100 g hemoglobin/L	74	NR			
Thrombocytopenia					
<100 platelets × 10^9/L	12	9.9	0.08		
≥100 platelets × 10^9/L	65	NR			
N. of PB leukocytes (×10^9/L)					
≤51.4	57	NR	0.06		
>51.4	20	NR			
N. of PB total T-cells (×10^9/L)					
≤3.2	56	NR	<0.001		
>3.2	21	6.2 (2.7–9.7)			
N. of PB $CD4^+$ T-cells (×10^9/L)					
≤2.1	60	NR	0.002		
>2.1	15	8.7 (4.1–13.2)			
N. of PB $CD8^+$ T-cells (×10^9/L)					
≤1.7	63	NR	0.001		
>1.7	12	6.1 (5.9–6.4)			
N. of PB monocytes (×10^9/L)					
≤1.04	67	NR	0.27		
>1.04	7	NR			
N. of PB neutrophils (×10^9/L)					
≤11.4	73	NR	0.07		
>11.4	3	6.1			
N. of PB basophils (×10^9/L)					
≤0.13	51	NR	0.04		
>0.13	5	8.7 (1.4–16)			
N. of PB eosinophils (×10^9/L)					
≤0.12	24	NR	0.08		
>0.12	50	NR			
N. of PB NK cells (×10^9/L)					
≤0.53	39	NR	0.31		
>0.53	32	NR			
Tumor B-cell clone size in blood (×10^9/L)					
≤37.7	53	NR	0.003		
>37.7	23	9.93			
HCDR3 hydropathy index					
Neutral	30	4.2 (75^{th} percentile: NR)	<0.001	12 (1.8–81)	0.01
Negatively charged	48	NR (75^{th} percentile: NR)			
% IGHV homology					
≤96	65	NR	0.56		
>96	13	NR			
HCDR3 length of CLL clone					
≤21 aa	73	NR			
>21 aa	5	NR	0.86		
DH2 gene usage	22	NR	0.70		
JH6 gene usage	21	NR	0.86		
Del (11q) and/or del(17p)	7	4.2 (2.8–5.6)	0.04		
N. of cytogenetically altered CLL cells (×10^9/L)					
≤33	57	NR			
>33	18	7.4 (2.2–12.5)	0.001		

aa, amino acids; CI, confidence interval; HR, hazard ratio; N, number of cases; NR, not reached; y, years.

TABLE 4 | *IGHV(D)J* gene usage in CLL patients classified according to their IGHV mutational status and the HCDR3 hydropathy index.

BCR features	U-CLL (N = 60)	M-CLL (N = 78)		p-value
		Neutral HCDR3 (N = 30)	Negatively charged HCDR3 (N = 48)	
V(H) gene family usage#				
V1	28/60 (47%)	5/30 (17%)	4/48 (8%)	<0.001; 0.004[a]; 0.000[b]
V1-69	16/60 (27%)	0/30 (0%)	2/48 (4%)	<0.001; 0.001[a,b]
V3	19/60 (32%)	16/30 (53%)	25/48 (52%)	0.05; 0.04[a]; 0.03[b]
V3-7	1/60 (2%)	5/30 (17%)	5/48 (10%)	0.03; 0.01[a]
V4	9/60 (15%)	8/30 (27%)	17/48 (35%)	0.05; 0.01[b]
V4-34	2/60 (3%)	5/30 (17%)	10/48 (21%)	0.02; 0.04[a]; 0.005[b]
D(H) gene family usage#				
D2	13/60 (22%)	13/30 (43%)	9/48 (19%)	0.04; 0.03[a]; 0.02[c]
D2-2	9/60 (15%)	3/30 (10%)	3/48 (6%)	0.34
D2-15	1/60 (2%)	4/30 (13%)	3/48 (6%)	0.08; 0.04[a]
D2-21	3/60 (5%)	6/30 (20%)	3/48 (6%)	0.04; 0.03[a]
D3	35/60 (58%)	6/30 (20%)	16/48 (33%)	0.001; 0.001[a]; 0.01[b]
D3-3	18/60 (30%)	2/30 (7%)	6/48 (12%)	0.01; 0.01[a]; 0.03[b]
J(H) gene family usage#				
J4	20/60 (33%)	7/30 (23%)	29/48 (60%)	0.002; 0.004[b]; 0.001[c]
J6	29/60 (48%)	11/30 (37%)	10/48 (21%)	0.01; 0.003[b]
Charge of HCDR3:				
J4 gene fraction	-0.167 (-0.296–0.029)	-0.076 (-0.325–0.115)	-0.209 (-0.946–0.115)	0.25
Non-J4 V(N1)D(N2) nucleotides	-0.392 (-1.150–0.447)	-0.107 (-0.346–0.670)	-0.936 (-1.702–-0.350)	<0.001; <0.001[b,c]
J6 gene fraction	-0.130 (-0.225–0.014)	-0.016 (-0.347–0.194)	-0.003 (-0.222–0.168)	0.010; 0.010[b]
Non-J6 V(N1)D(N2) nucleotides	-0.084 (-0.890–0.996)	-0.170 (-0.419–0.577)	-0.941 (-1.490–-0.572)	<0.001; <0.005[b,c]
D2 gene fraction	0.329 (-0.382–1.032)	0.010 (-0.286–0.698)	-0.258 (-0.596–0.029)	0.013; 0.003[b]
Non-D2 V(N1)(N2)J nucleotides	-0.502 (-1.380–-0.006)	-0.070 (-0.441–0.491)	-0.777 (-1.126–-0.471)	0.001; <0.010[a]; <0.001[c]
D3 gene fraction	-0.256 (-0.565–0.849)	-0.113 (-0.444–0.707)	-0.365 (-0.742–0.302)	0.10
Non-D3 V(N1)(N2)J nucleotides	-0.347 (-1.553–0.208)	0.007 (-0.570–0.351)	-0.712 (-1.057–-0.349)	0.001; <0.01[b]; 0.013[c]
HCDR3 length in patients aged ≤65y	21 (12–28)	14 (8–24)	16 (7–23)	0.001; 0.001[a]; 0.008[b]
HCDR3 length in patients aged >65y	20 (13–28)	18 (15–24)	15 (9–24)	<0.001; <0.001[b]; 0.013[c]
Stereotyped IGHV#	17/60 (28%)	5/30 (17%)	5/48 (10%)	0.06; 0.02[b]
Subsets (1,2,8)#	4/60 (7%)	2/30 (7%)	1/48 (2%)	0.50

Results expressed as median (range) or as # number of cases (percentage).

[a]*U-CLL vs. M-CLL with neutral HCDR3.*

[b]*U-CLL vs. M-CLL with negatively charged HCDR3.*

[c]*M-CLL with neutral HCDR3 vs. M-CLL with negatively charged HCDR3.*

CLL, chronic lymphocytic leukemia; M, mutated IGHV; N, number; U, unmutated IGHV.

specific for a given epitope (39–41). Among other factors, this might also contribute to explain the more aggressive clinical course (42, 43) and the shortened survival of U-CLL vs. M-CLL (44). As a consequence, the I*GHV* gene mutational status currently represents one of the most relevant prognostic determinants in CLL (44).

Similarly to normal B-cells (31), here we show that CLL cells also display a Gaussian distribution according to the hydropathy index of their BCR, slightly skewed toward negatively charged HCDR3 amino acid sequences. Interestingly, when we divided our patients into cases with neutral (mean GS of -0.1) vs. more negatively charged (mean GS of -1.1) HCDR3 sequences, two subgroups of CLL patients with clearly distinct clinical and biological features emerged. Thus, CLL patients with neutral HCDR3 sequences showed a clear predominance of men, U-CLL with longer HCDR3 sequences, a lower frequency of IGHV gene mutation, and higher frequency of more advanced stages of the disease, in association with a higher rate of disease progression and shorter TTT. Interestingly, shortening of HCDR3 sequences with a trend to negatively charged BCRs is a typical feature of selection of B-cell precursors in BM required for the survival of B-cells that will enter the mature B-cell repertoire (31). Based on these findings, our results suggest that expanded CLL cells in patients with neutral and longer HCDR3 sequences might reflect an earlier tumor cell origin in BM (45–47). Among other factors, this might also contribute to explain the greater frequency of more advanced stages of disease at diagnosis (48), together with an increased rate of disease progression vs. patients with negatively charged HCDR3 sequences. Nevertheless, these differences could be potentially due to the fact that CLL cases with neutral HCDR3 sequences included a higher fraction of U-CLL vs. M-CLL patients.

To investigate the potential independent value of both variables (the BCR mutational status and its hydropathy index), we separately studied the features of CLL patients with neutral vs. negatively charged HCDR3 sequences among U-CLL and M-CLL cases. Thus, U-CLL cases with a neutral and negatively charged HCDR3 showed similar clinical and biological features associated with a uniformly poorer outcome, in line with previous observations (49–51). In contrast, the HCDR3 hydropathy index identified two different prognostic subgroups of M-CLL. These included a subgroup of M-CLL with neutral HCDR3 who displayed intermediate clinical, genetic, and prognostic features between the classical U-CLL and M-CLL patients with a negatively charged BCR. Thus, M-CLL with

neutral HCDR3 showed predominance of men—similar to that found in U-CLL, but with significantly higher hemoglobin levels —in association with a higher frequency of thrombocytopenia and an intermediate frequency of cytogenetically altered cases between U-CLL and the other M-CLL patients, at the expense of a greater frequency of del(17p). At present, it is well established that progression of MBL toward CLL is associated with a more prominent male predominance and greater frequency of U-CLL (52). Male predominance among M-CLL cases with neutral HCDR3 might also contribute to explain the greater hemoglobin levels observed in these patients, which contrasts with the higher frequency of thrombocytopenia compared to M-CLL with negatively charged HCDR3 sequences. This together with the greater frequency of more advanced stage of the disease among M-CLL with a neutral vs. negatively charged HCDR3 would support a poorer outcome within M-CLL for the former patient group, as confirmed here *via* an adverse impact on the time elapsed from diagnosis to first therapy among M-CLL patients with neutral vs. negatively charged HCDR3 sequences.

From the molecular point of view, M-CLL with neutral HCDR3 showed DJ footprints compatible with a more immature BCR repertoire associated with preferential usage of D(H)2 *IGHV* gene segments (53), in the absence of a biased use of JH4 gene segments, as found in M-CLL cases with negatively charged HCDR3 sequences, being biased use of JH4 gene segments a typical feature of more mature PB B lymphocytes (54). In addition, we also observed longer HCDR3 sequences in older (>65 y) patients who had U-CLL and M-CLL with neutral HCDR3 vs. M-CLL with negatively charged HCDR3 sequences, in line with what might be expected among older subjects (27). Despite U-CLL and M-CLL with neutral HCDR3 shared HCDR3 sequences which typically had no traits associated with a mature B-cell repertoire, important differences were still observed in the IGHV repertoire of CLL cells of both patient groups as regards the usage of the VH1 and VH3 gene segments, further emphasizing also the biological differences between them.

Altogether, our findings show that based on the HCDR3 hydropathy index of HCDR3 sequences, two clearly distinct subgroups of M-CLL patients with different clinical, genetic, and prognostic features can be identified which are characterized by neutral vs. negatively charged BCRs, associated with molecular features of precursor vs. peripheral/mature B-cells, respectively. Further studies are needed to elucidate the precise mechanisms involved in determining the role of these different BCR profiles (compared to other prognostic factors such as ZAP70) in the distinct clinical behavior and outcome of both groups of M-CLL patients and facilitate implementation of assays for routine assessment of the HCDR3 hydropathy index in M-CLL in the clinical settings.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the local Institutional Ethics Committee, University Hospital of Salamanca (code of approval: CEIC-PI4705/2017). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AR-C, JA, and AO contributed to conception and design of the study. BF, GO, IC, and MA organized the database. CP contributed to bioinformatics calculations. MP and AG-M contributed to genomic collection, storage, and quality control. AG-M performed part of the statistical analysis and critical review of manuscript. MGD, FF and AS-R contributed in the clinical part of the manuscript and critical review of manuscript. AR-C wrote the first draft of the manuscript. All authors contributed to the manuscript revision and read and approved the submitted version.

FUNDING

This work was supported by the following grants: FS/37-2017, from the Fundación Memoria D. Samuel Solórzano, Universidad de Salamanca; FIS PI17/00399-FEDER, from the Fondo de Investigación Sanitaria of Instituto de Salud Carlos III, Madrid, Spain; 0639_IDIAL_NET_3_E, from cooperative network EP-INTERREG V A España Portugal (POCTEP); and ECRIN-M3, Accelerator Award Full, Cancer Research UK, Fundación Científica de la Asociación Española Contra el Cáncer (AECC), Fondazione AIRC per la Ricerca sul Cancro. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank all the staff of the Cytometry and Cell Purification Services of the University of Salamanca; the hematologists Almudena Navarro, Monica Baile, and Jose M. Bastida (University Hospital of Salamanca); the Spanish National DNA Bank Carlos III (Banco Nacional de DNA Carlos III, University of Salamanca: www.bancoadn.org); and the Genomic Unit of Cancer Research Centre (IBMCC, USAL-CSIC) for their support in the data collection for the preparation of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.723722/full#supplementary-material

Supplementary Figure 1 | Kyte-Doolittle numerical scale of aminoacids (aa) and the formula to calculate the Gravy Score (GS) of HCDR3 (hydropathy index). The biochemical nature of aa according to their Kyte-Doolittle value and the algorithm used to calculate the GS of the HCDR3 sequence in our patients are shown in **(A, B)**, respectively. In **(C)**, four different HCDR3 sequences with their corresponding GS (where the sum of the Kyte-Doolittle values for individual aa divided by HCDR3 length provides the specific GS of individual sequences). As shown, different patients can share the same GS of their HCDR3, despite having different aa sequences.

REFERENCES

1. Oppezzo P, Dighiero G. Role of the B-Cell Receptor and the Microenvironment in Chronic Lymphocytic Leukemia. *Blood Cancer J* (2013) 3:e149. doi: 10.1038/bcj.2013.45
2. Hallek M. Chronic Lymphocytic Leukemia: 2020 Update on Diagnosis, Risk Stratification and Treatment. *Am J Hematol* (2019) 94(11):1266–87. doi: 10.1002/ajh.25595
3. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner H, et al. iwCLL Guidelines for Diagnosis, Indications for Treatment, Response Assessment, and Supportive Management of CLL. *Blood* (2018) 131 (25):2745–60. doi: 10.1182/blood-2017-09-806398
4. Dighiero G, Hamblin TJ. Chronic Lymphocytic Leukemia. *Lancet* (2008) 371 (9617):1017–29. doi: 10.1016/S0140-6736(08)60456-0
5. International CLL-IPI working group. An International Prognostic Index for Patients With Chronic Lymphocytic Leukaemia (CLL-IPI): A Meta-Analysis of Individual Patient Data. *Lancet Oncol* (2016) 17(6):779–90. doi: 10.1016/S1470-2045(16)30029-8
6. Puente XS, Beá S, Valdés-Mas R, Villamor N, Gutiérrez-Abril J, Martín-Subero JI, et al. Non-Coding Recurrent Mutations in Chronic Lymphocytic Leukemia. *Nature* (2015) 526(7574):519–24. doi: 10.1038/nature14666
7. D´Avola A, Drennan S, Tracy I, Henderson I, Chiecchio L, Larrayoz M, et al. Surface IgM Expression and Function are Associated With Clinical Behavior, Genetic Abnormalities, and DNA Methylation in CLL. *Blood* (2016) 128 (6):816–26. doi: 10.1182/blood-2016-03-707786
8. Vardi A, Agathangelidis A, Sutton LA, Ghia P, Rosenquist R, Stamatopoulos K. Immunogenetic Studies of Chronic Lymphocytic Leukemia: Revelations and Speculations About Ontogeny and Clinical Evolution. *Cancer Res* (2014) 74(16):4211–6. doi: 10.1158/0008-5472.CAN-14-0630
9. Chiorazzi N, Stevenson FK. Celebrating 20 Years of IGHV Mutation Analysis in CLL. *Hemasphere* (2020) 4(1):e334. doi: 10.1097/HS9.0000000000000334
10. Damle RN, Ghiotto F, Valetto A, Albesiano E, Fais F, Yan XJ, et al. B-Cell Chronic Lymphocytic Leukemia Cells Express a Surface Membrane Phenotype of Activated, Antigen-Experienced B Lymphocytes. *Blood* (2002) 99(11):4087–93. doi: 10.1182/blood.v99.11.4087
11. Forconi F, Potter KN, Wheatley I, Darzentas N, Sozzi E, Stamatopoulos K, et al. The Normal IGHV1-69-Derived B-Cell Repertoire Contains Stereotypic Patterns Characteristic of Unmutated CLL. *Blood* (2010) 115(1):71–7. doi: 10.1182/blood-2009-06-225813
12. Dono M, Cerruti G, Zupo S. The $CD5^+$ B-Cell. *Int J Biochem Cell Biol* (2004) 36:2105–11. doi: 10.1016/j.biocel.2004.05.017
13. Küppers R, Zhao M, Hansmann ML, Rajewsky K. Tracing B Cell Development in Human Germinal Centres by Molecular Analysis of Single Cells Picked From Histological Sections. *EMBO J* (1993) 12(13):4955–67. doi: 10.1002/j.1460-2075.1993.tb06189.x
14. Chiorazzi N, Ferrarini M. Cellular Origin(s) of Chronic Lymphocytic Leukemia: Cautionary Notes and Additional Considerations and Possibilities. *Blood* (2011) 117(6):1781–91. doi: 10.1182/blood-2010-07-155663
15. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, et al. Stereotyped B-Cell Receptors in One-Third of Chronic Lymphocytic Leukemia: A Molecular Classification With Implications for Targeted Therapies. *Blood* (2012) 119(19):4467–75. doi: 10.1182/blood-2011-11-393694
16. Stamatopoulos K, Agathangelidis A, Rosenquist R, Ghia P. Antigen Receptor Stereotypy in Chronic Lymphocytic Leukemia. *Leukemia* (2017) 31(2):282–91. doi: 10.1038/leu.2016.322
17. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, Smilevska T, et al. Over 20% of Patients With Chronic Lymphocytic Leukemia Carry Stereotyped Receptors: Pathogenic Implications and Clinical Correlations. *Blood* (2007) 109:259–70. doi: 10.1182/blood-2006-03-012948
18. Dighiero G, Oppezzo P. What do Somatic Hypermutation and Class Switch Recombination Teach Us About Chronic Lymphocytic Leukemia Pathogenesis? *Curr Top Microbiol Immunol* (2005) 294:71–89. doi: 10.1007/3-540-29933-5_5
19. Ivanov I, Link J, Ippolito GC, Schroeder HWJr. "Constraints on Hydropathy and Sequence Composition of HCDR3 are Conserved Across Evolution". In: M Zanetti, editor. *The Antibodies*. London: Taylor and Francis (2002). p. 43–67.
20. Ivanov II, Schelonka RL, Zhuang Y, Gartland GL, Zemlin M, Schroeder HWJr. Development of the Expressed Ig CDR-H3 Repertoire is Marked by Focusing of Constraints in Length, Amino Acid Use, and Charge That Are First Established in Early B Cell Progenitors. *J Immunol* (2005) 174(12):7773–80. doi: 10.4049/jimmunol.174.12.7773
21. Fais F, Ghiotto F, Hashimoto S, Sellars B, Valetto A, Allen SL, et al. Chronic Lymphocytic Leukemia B Cells Express Restricted Sets of Mutated and Unmutated Antigen Receptors. *J Clin Invest* (1998) 102:1515–25. doi: 10.1172/JCI3009
22. Rosenquist R, Ghia P, Hadzidimitriou A, Sutton LA, Agathangelidis A, Baliakas P, et al. Immunoglobulin Gene Sequence Analysis in Chronic Lymphocytic Leukemia: Updated ERIC Recommendations. *Leukemia* (2017) 31(7):1477–81. doi: 10.1038/leu.2017.125
23. Henriques A, Rodríguez-Caballero A, Criado I, Langerak AW, Nieto WG, Lécrevisse Q, et al. Molecular and Cytogenetic Characterization of Expanded B-Cell Clones From Multiclonal Versus Monoclonal B-Cell Chronic Lymphoproliferative Disorders. *Haematologica* (2014) 99(5):897–907. doi: 10.3324/haematol.2013.098913
24. Criado I, Muñoz-Criado S, Rodríguez-Caballero A, Nieto WG, Romero A, Fernández-Navarro P, et al. Host Virus and Pneumococcus-Specific Immune Responses in High-Count Monoclonal B-Cell Lymphocytosis and Chronic Lymphocytic Leukemia: Implications for Disease Progression. *Haematologica* (2017) 102(7):1238–46. doi: 10.3324/haematol.2016.159012
25. Kyte J, Doolittle RF. A Simple Method for Displaying the Hydropathic Character of a Protein. *J Mol Biol* (1982) 157(1):105–32. doi: 10.1016/0022-2836(82)90515-0
26. Criado I, Rodríguez-Caballero A, Gutiérrez ML, Pedreira CE, Alcoceba M, Nieto W, et al. Low-Count Monoclonal B-Cell Lymphocytosis Persists After Seven Years of Follow Up and is Associated With a Poorer Outcome. *Haematologica* (2018) 103(7):1198–208. doi: 10.3324/haematol.2017.183954
27. Pickman Y, Dunn-Walters D, Mehr R. BCR CDR3 Length Distributions Differ Between Blood and Spleen and Between Old and Young Patients, and TCR Distributions can be Used to Detect Myelodysplastic Syndrome. *Phys Biol* (2013) 10(5):56001. doi: 10.1088/1478-3975/10/5/056001
28. Kwak K, Akkaya M, Pierce SK. B Cell Signaling in Context. *Nat Immunol* (2019) 20(8):963–9. doi: 10.1038/s41590-019-0427-9
29. DeKosky BJ, Lungu OI, Park D, Johnson EL, Charab W, Chrysostomou C, et al. Large-Scale Sequence and Structural Comparisons of Human Naïve and Antigen-Experienced Antibody Repertoires. *Proc Natl Acad Sci USA* (2016) 113(19):E2636–45. doi: 10.1073/pnas.1525510113
30. Khass M, Vale AM, Burrows PD, Schroeder HWJr. The Sequences Encoded by Immunoglobulin Diversity (DH) Gene Segments Play Key Roles in Controlling B-Cell Development, Antigen-Binding Site Diversity, and Antibody Production. *Immunol Rev* (2018) 284(1):106–19. doi: 10.1111/imr.12669
31. Kaplinsky J, Li A, Sun A, Coffre M, Koralov SB, Arnaout R. Antibody Repertoire Deep Sequencing Reveals Antigen-Independent Selection in Maturing B Cells. *Proc Natl Acad Sci USA* (2014) 111(25):E2622–9. doi: 10.1073/pnas.1403278111
32. Haselager MV, Kater AP, Eldering E. Proliferative Signals in Chronic Lymphocytic Leukemia; What Are We Missing? *Front Oncol* (2020) 10:592205. doi: 10.3389/fonc.2020.592205
33. Greaves M. Clonal Expansion in B-CLL: Fungal Drivers or Self-Service? *J Exp Med* (2013) 210(1):1–3. doi: 10.1084/jem.20122739

34. Hengeveld PJ, Levin MD, Martijn Kolijn P, Langerak AW. Reading the B-Cell Receptor Immunome in Chronic Lymphocytic Leukemia: Revelations and Applications. *Exp Hematol* (2021) 93:14–24. doi: 10.1016/j.exphem.2020.09.194
35. Mauerer K, Zahrieh D, Gorgun G, Li A, Zhou J, Ansén S, et al. Immunoglobulin Gene Segment Usage, Location and Immunogenicity in Mutated and Unmutated Chronic Lymphocytic Leukemia. *Br J Haematol* (2005) 129(4):499–510. doi: 10.1111/j.1365-2141.2005.05480.x
36. Hwang HK, Trama AM, Kozink DM, Chen X, Wiehe K, Cooper AJ, et al. IGHV1-69 B Cell Chronic Lymphocytic Leukemia Antibodies Cross-React With HIV-1 and Hepatitis C Virus Antigens as Well as Intestinal Commensal Bacteria. *PLoS One* (2014) 9(3):e90725. doi: 10.1371/journal.pone.0090725
37. Hatzi K, Catera R, Moreno Atanasio C, Fischetti VA, Allen SL, Kolitz JE, et al. Chronic Lymphocytic Leukemia Immunoglobulins Display Bacterial Reactivity That Converges and Diverges From Auto-/Poly-Reactivity and IGHV Mutation Status. *Clin Immunol* (2016) 172:44–51. doi: 10.1016/j.clim.2016.08.020
38. Cariappa A, Chase C, Liu H, Russell P, Pillai S. Naïve Recirculating B Cells Mature Simultaneously in the Spleen and Bone Marrow. *Blood* (2007) 109 (6):2339–45. doi: 10.1182/blood-2006-05-021089
39. Burger JA, Chiorazzi N. B Cell Receptor Signaling in Chronic Lymphocytic Leukemia. *Trends Immunol* (2013) 34(12):592–601. doi: 10.1016/j.it.2013.07.002
40. Herve M, Xu K, Ng YS, Wardemann H, Albesiano E, Messmer BT, et al. Unmutated and Mutated Chronic Lymphocytic Leukemias Derive From Self-Reactive B Cell Precursors Despite Expressing Different Antibody Reactivity. *J Clin Invest* (2005) 115(6):1636–43. doi: 10.1172/JCI24387
41. ten Hacken E, Burger JA. Molecular Pathways: Targeting the Microenvironment in Chronic Lymphocytic Leukemia—Focus on the B-Cell Receptor. *Clin Cancer Res* (2014) 20(3):548–56. doi: 10.1158/1078-0432.CCR-13-0226
42. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, et al. Ig V Gene Mutation Status and CD38 Expression as Novel Prognostic Indicators in Chronic Lymphocytic Leukemia. *Blood* (1999) 94(6):1840–7. doi: 10.1182/blood.V94.6.1840
43. Hamblin TJ, Davis Z, Gardiner A, Oscier DG, Stevenson FK. Unmutated Ig VH Genes are Associated With a More Aggressive Form of Chronic Lymphocytic Leukemia. *Blood* (1999) 94:1848–54. doi: 10.1182/blood.V94.6.1848
44. Chiorazzi N, Chen S-S, Rai KR. Chronic Lymphocytic Leukemia. *Cold Spring Harb Perspect Med* (2021) 11(2):a035220. doi: 10.1101/cshperspect.a035220
45. Alizadeh AA, Majeti R. Surprise! HSC Are Aberrant in Chronic Lymphocytic Leukemia. *Cancer Cell* (2011) 20(2):135–6. doi: 10.1016/j.ccr.2011.08.001
46. Kikushige Y, Ishikawa F, Miyamoto T, Shima T, Urata S, Yoshimoto G, et al. Self-Renewing Hematopoietic Stem Cells Is the Primary Target in Pathogenesis of Human Chronic Lymphocytic Leukemia. *Cancer Cell* (2011) 20(2):246–59. doi: 10.1016/j.ccr.2011.06.029
47. Rodríguez-Caballero A, Henriques A, Criado I, Langerak AW, Matarraz S, López A, et al. Subjects With Chronic Lymphocytic Leukemia-Like B Cell Clones With Stereotyped B-Cell Receptors Frequently Show MDS-Associated Phenotypes on Myeloid Cells. *Br J Haematol* (2015) 168(2):258–67. doi: 10.1111/bjh.13127
48. Pasikowska M, Walsby E, Apollonio B, Cuthill K, Philips E, Coulter E, et al. Phenotype and Immune Function of Lymph Node and Peripheral Blood CLL Cells are Linked to Transendothelial Migration. *Blood* (2016) 128(4):563–73. doi: 10.1182/blood-2016-01-683128
49. Görgün G, Holderried TAW, Zahrieh D, Neuberg D, Gribben JG. Chronic Lymphocytic Leukemia Cells Induce Changes in Gene Expression of CD4 and CD8 T Cells. *J Clin Invest* (2005) 115:1797–805. doi: 10.1172/JCI24176
50. van Attekum MH, Eldering E, Kater AP. Chronic Lymphocytic Leukemia Cells are Active Participants in Microenvironmental Cross-Talk. *Haematologica* (2017) 102(9):1469–76. doi: 10.3324/haematol.2016.142679
51. Huergo-Zapico L, Acebes-Huerta A, González-Rodríguez AP, Contesti J, Gonzalez-García E, Payer AR, et al. Expansion of NK Cells and Reduction of NKG2D Expression in Chronic Lymphocytic Leukemia. Correlation With Progressive Disease. *PLoS One* (2014) 9(10):e108326. doi: 10.1371/journal.pone.0108326
52. Catovsky D, Wade R, Else M. The Clinical Significance of Patients´ Sex in Chronic Lymphocytic Leukemia. *Haematologica* (2014) 99(6):1088–94. doi: 10.3324/haematol.2013.101378
53. Mroczek ES, Ippolito GC, Rogosch T, Hon Hoi K, Hwangpo TA, Brand MG, et al. Differences in the Composition of the Human Antibody Repertoire by B Cell Subsets in the Blood. *Front Immunol* (2014) 5:96. doi: 10.3389/fimmu.2014.00096
54. Huang C, Stewart AK, Schwartz RS, Stollar BD. Immunoglobulin Heavy Chain Gene Expression in Peripheral Blood B Lymphocytes. *J Clin Invest* (1992) 89(4):1331–43. doi: 10.1172/JCI115719

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Oncology

REVIEW
published: 11 November 2021
doi: 10.3389/fonc.2021.769612

Understanding Monoclonal B Cell Lymphocytosis: An Interplay of Genetic and Microenvironmental Factors

*Chrysi Galigalidou[1,2], Laura Zaragoza-Infante[1,3], Anastasia Iatrou[1], Anastasia Chatzidimitriou[1,4], Kostas Stamatopoulos[1,4] and Andreas Agathangelidis[1,5]**

[1] Institute of Applied Biosciences (INAB), Centre for Research and Technology Hellas (CERTH), Thessaloniki, Greece, [2] Department of Molecular Biology and Genetics, Democritus University of Thrace, Alexandroupolis, Greece, [3] Hematology Department, University General Hospital of Thessaloniki AHEPA, Thessaloniki, Greece, [4] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden, [5] Department of Biology, School of Science, National and Kapodistrian University of Athens, Athens, Greece

OPEN ACCESS

Edited by:
Dimitar G. Efremov, International Centre for Genetic Engineering and Biotechnology, Italy

Reviewed by:
Paolo Strati, University of Texas MD Anderson Cancer Center, United States
Stefano Molica, Azienda Ospedaliera Pugliese Ciaccio, Italy
Riccardo Bomben, Aviano Oncology Reference Center (IRCCS), Italy

***Correspondence:**
Andreas Agathangelidis
agathan@biol.uoa.gr

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: *02 September 2021*
Accepted: *26 October 2021*
Published: *11 November 2021*

Citation:
Galigalidou C, Zaragoza-Infante L, Iatrou A, Chatzidimitriou A, Stamatopoulos K and Agathangelidis A (2021) Understanding Monoclonal B Cell Lymphocytosis: An Interplay of Genetic and Microenvironmental Factors. Front. Oncol. 11:769612. doi: 10.3389/fonc.2021.769612

The term monoclonal B-cell lymphocytosis (MBL) describes the presence of a clonal B cell population with a count of less than 5×10^9/L and no symptoms or signs of disease. Based on the B cell count, MBL is further classified into 2 distinct subtypes: 'low-count' and 'high-count' MBL. High-count MBL shares a series of biological and clinical features with chronic lymphocytic leukemia (CLL), at least of the indolent type, and evolves to CLL requiring treatment at a rate of 1-2% per year, whereas 'low-count' MBL seems to be distinct, likely representing an immunological rather than a pre-malignant condition. That notwithstanding, both subtypes of MBL can carry 'CLL-specific' genomic aberrations such as cytogenetic abnormalities and gene mutations, yet to a much lesser extent compared to CLL. These findings suggest that such aberrations are mostly relevant for disease progression rather than disease onset, indirectly pointing to microenvironmental drive as a key contributor to the emergence of MBL. Understanding microenvironmental interactions is therefore anticipated to elucidate MBL ontogeny and, most importantly, the relationship between MBL and CLL.

Keywords: monoclonal B cell lymphocytosis (MBL), chronic Lymphocutic Leukemia (CLL), genetics, immunogenetics, tumor microenvironment, ontogenesis, B cell receptor, immunoglobulin

INTRODUCTION

Epidemiology

Monoclonal B cell lymphocytosis (MBL) is an asymptomatic hematological condition characterized by the presence of clonal B cell expansion(s) in otherwise healthy individuals. The immunophenotype of the B cell clones in approximately 75% of all MBL cases is similar to that observed in chronic lymphocytic leukemia (CLL): $CD5^+$, $CD19^+$, $CD20^{dim}$, $CD23^+$ and low levels (dim) of surface immunoglobulin (sIg^{dim}) that in most cases is IgM with or without IgD. Other immunophenotypic characteristics include the absence of FMC7 and the weak or no expression of CD79b and CD22. Based on this constellation of features, the most common subtype of MBL is

defined as 'CLL-like'. Another immunophenotype occasionally observed in MBL is characterized by strong expression of CD20 and sIg along with absence of CD23 ($CD5^+$, $CD19^+$, $CD20^{high}$, $CD23^+$, sIg^{high}); this MBL subtype is described as 'atypical' MBL. The last, most infrequent, MBL subtype is 'CD5-negative' MBL, in which the complete immunophenotypic characterization is $CD5^{neg}$, $CD19^+$, $CD20^+$ with moderate to high sIg expression, with the ratio of κ: λ sIg being strongly skewed (either more than 3:1 or less than 1:3) (1–4).

Immunophenotypic Subtypes of MBL

Significant differences exist between the three MBL immunophenotypic subtypes regarding the expression of several cell surface markers, such as CD23, CD79b, and FMC7. Moreover, several established CLL risk factors are also differentially expressed between the immunophenotypic subtypes of MBL; for instance, markers associated with aggressive CLL, such as CD38, ZAP70, and CD49d are highly expressed in the majority of 'atypical' MBL cases yet are rarely expressed in 'CLL-like' MBL (1, 5). Moreover, the B cell memory marker CD27 shows higher expression in the 'CLL-like' type of MBL compared to 'atypical' and 'CD5-negative' MBL, while, in contrast, CD22, CD79a, CD79b, and CD1c (BDCA-1) are mainly expressed in the 'non CLL-like' types. Furthermore, FMC7 and CD43 are also characterized by differential expression among the three immunophenotypic types of MBL; more specifically, FMC7 exhibits low expression levels in 'CLL-like' MBL, intermediate levels in atypical, and high levels in the 'CD5-negative' type of MBL, while the exact opposite pattern is evident for CD43 expression (6). Differences are also observed regarding the expression of sIg, which is significantly lower in 'CLL-like' MBL compared to either the 'atypical' or the 'CD5-negative' types (1). Finally, expression of the *ROR1* gene was statistically different among the three immunophenotypic subtypes of MBL being significantly higher in 'CLL-like' MBL versus the others (1).

MBL Risk Factors

Moving from the immunophenotype, the diagnostic criteria for distinguishing between MBL from CLL are primarily based on the number of circulating monoclonal B cells. Two distinct MBL subtypes are recognized: high-count MBL (HC-MBL), or 'clinical' MBL, with $0.5\text{-}4.99\times 10^9$ clonal B cells/L and low-count MBL (LC-MBL), or 'general' population MBL, characterized by $<0.5\times 10^9$ clonal B cells/L (4). It should be noted that most individuals diagnosed with MBL display a normal absolute B cell count, with the clonal B cell population accounting for less than 10% of the total number of B cells (3).

The incidence of MBL gradually increases with advancing age; in fact, MBL is very rarely detected in individuals under 40 years old, whereas it is present in more than 20% of individuals older than 70 years old and can reach even 75% among patients aged more than 90 years (3). Age is not the only factor strongly associated with MBL, as a familial history of hematological or solid malignancies has also been reported; in fact, 17% of first-degree relatives with no personal history of lymphoproliferative disorders from families with at least two cases of CLL were found to have MBL, in some cases at a young age (less than 40 years old) (7). Gender is another factor connected to the prevalence of MBL as, similar to CLL, males have a significantly higher risk for developing MBL compared to females (8).

Exposure to infections and immunodeficiency may also predispose to MBL development, at least in certain instances, especially when combined with advanced age and male gender. A characteristic example concerns the high frequency of MBL among patients with hepatitis C virus (HCV) infection (9). On the other hand, MBL was found to be significantly less common among individuals vaccinated against pneumococcal or influenza infections (3).

Risk of Progression to CLL

Unmutated somatic hypermutation (SHM) status of the rearranged immunoglobulin heavy variable (IGHV) gene and/or 'CLL-related' cytogenetic aberrations related to aggressive disease represent risk factors of progression from MBL to CLL and shorter overall survival (OS) (10). Furthermore, over 50% of HC-MBL cases carry at least one gene mutation in a CLL putative driver, evident up to 41 months before the progression to CLL. Thus, genomic characterization can be used to identify those MBL cases who will progress to CLL requiring treatment; this was more pronounced in cases with subclonal expansion of driver mutations (11, 12).

Along these lines, the genetic distance between LC-MBL and CLL is clearly greater than that of HC-MBL (13) and, thus, the question whether LC-MBL represents a very early stage in the natural history of CLL, remains elusive.

HC-MBL PROGRESSING TO CLL

Monoclonal B cell populations can be detected long before CLL diagnosis, even up to 6.4 years (13). However, the majority of 'CLL-like' HC-MBL cases remain stable overtime; progression to CLL has been estimated at a rate of 1-4% per year (1). In more detail, a longitudinal analysis in 185 individuals with 'CLL-like' HC-MBL revealed a progressive lymphocytosis in 28% of the cases with a median follow-up of 6.7 years. However, only 7% of the cases progressed to CLL requiring treatment with the overall rate of progression to CLL being 1.1% per year (14). In a retrospective study by the Mayo Clinic, 302 'CLL-like' HC-MBL cases were monitored for a median of 18 months with ~1.4% of cases per year requiring treatment due to disease progression. This study proposed a B cell count threshold of 10×10^9/L for discriminating between two categories of 'CLL-like' HC-MBL cases with higher and lower probability of treatment. This proposal was subsequently tested and validated in the GIMEMA database (15, 16). Finally, in an Italian study, 123 'CLL-like' HC-MBL cases were monitored for a median of 43 months with 4% per year requiring treatment in the first 7 years (17). Irrespective of progression to CLL, periodic hematological follow-up was recommended after the detection of HC-MBL, since this group has shorter overall survival compared to the general population (18).

Many of the biological characteristics of HC-MBL are similar to CLL Rai stage 0 (19). Nowadays, the strongest predictive

marker for MBL progressing to CLL is the B cell count. This has emerged after the establishment of a B cell count-based system transition for discriminating MBL from CLL. This has led to a sizeable fraction of patients, considered as Rai stage 0 CLL, ending up classified as HC-MBL (20, 21). Relevant findings suggested that the B cell count appears to be a better predictor of TFS and OS when a defined lymphocyte threshold is applied (10). In a relevant study, the authors proposed a distinction of 2 'CLL-like' cell count thresholds for the identification of a very low risk group (<1.2x10^9/L) and a high risk group (>3.7x10^9/L) within HC-MBL (10, 17).

Subsequent studies appraised the impact of the size of the MBL clone on several clinical metrics, such as progression-free-survival (PFS), TFS and OS. In regard to the former, it is rather difficult to draw robust conclusions, given that a slight increase in cell numbers in HC-MBL cases with clone sizes near the mathematical cutoff of 5.0x10^9/L could be classified as disease progression. On the other hand, the assessment of TFS through standard NCI/iwCLL criteria may be more informative (22).

In CLL, the study of prognostic markers for time-to-first-treatment (TTFT) and OS has been of great interest, especially towards the development of a multifactorial risk score. In this context, 28 individual prognostic variables were assessed in a series of 969 individuals with CLL Rai 0 or HC-MBL leading to the development of the CLL international Prognostic Index (CLL-IPI). Implementation of the CLL-IPI score led to the identification of 4 risk groups (low, intermediate, high, and very-high risk) with different 5-year OS (93%, 79%, 63%, and 23%, respectively). Multivariate analysis of absolute B-cell count with individual factors of the CLL-IPI showed that five parameters, namely age, Rai stage, serum beta-2 microglobulin, unmutated IGHV genes, and del(17p) or *TP53* mutation were associated with shorter TTFT and OS. Overall, the CLL-IPI risk score, despite some limitations, is able to predict TTFT and OS in previously untreated CLL patients whose only disease symptom is the presence of a circulating clonal B cell population (23). Despite these advances, the relevance of CLL-IPI for the evolution of HC-MBL to CLL remains to be fully elucidated.

Several markers have been studied at MBL diagnosis for their capacity to predict progression to CLL, including B cell receptor immunoglobulin (BcR IG) stereotypy, ZAP70 expression, Hb, platelet count and LDH, however none of them was clearly associated with progression to CLL ($p \geq 0.05$ in all instances). In contrast, the SHM status of the clonotypic IGHV gene, high expression of CD38 (≥30%), high expression of CD49d (≥30%) and chromosomal abnormalities were found to be significant on univariate analysis regarding prediction of TTFT. However, on multivariate analysis trisomy 12 and del (16) (p13) were the sole independent predictors of treatment-free survival (TFS) in MBL after adjusting for IGHV SHM status and CD38 expression (5, 14).

PERSISTENCE OF LC-MBL OVER TIME

HC-MBL represents the vast majority of MBL cases (approximately 85%) identified in the context of clinical practice following the investigation of lymphocytosis, with a median 'CLL-like' B cell count above 1.9x10^9/L. On the contrary, around 85% of MBL cases detected in population screening studies have a 'CLL-like' B cell count below 0.5x10^9/L, with 40% of them bearing fewer than 0.05x10^9 clonal B cells/L (24). Previous findings support a bimodal distribution with a lower peak below 0.05x10^9 clonal B cells/L, very few cases in the range of 0.05-0.5 x10^9 clonal B cells/L, and another peak accounting for cases with a clonal B cell count ranging between 0.5-5x10^9 clonal B cells/L. A possible explanation could be a bias towards the selection of cases with high lymphocyte counts against those with mild or borderline lymphocytosis in the hematology clinical routine (22). An alternative biological explanation is offered by the distinct immunogenetic profiles of HC-MBL and LC-MBL, whereby pronounced clonal expansions are seen only when cells with 'CLL-like' phenotype express BcR IG with certain distinctive features (19). Longitudinal studies support the latter explanation, since most of the LC-MBL cases persisted over large periods of time (up to 34 months), without any progression to clinically overt disease. In fact, LC-MBL seems to be stable over time, a characteristic that is more pronounced in 'CLL-like' LC-MBL (90%) compared to other immunophenotypic variants, such as atypical (44.4%) and CD5-negative MBL (66.7%) (2).

Even though progression to HC-MBL is very rare, LC-MBL cases may often display a small increase in the size of the B cell clone. Studies focusing on clonal dynamics revealed a correlation between clonal size and clonality; in specific, the majority of monoclonal cases increased in size over time, while a much smaller number of biclonal and multiclonal cases behaved accordingly (25). Mathematical modeling of these results led to the conclusion that, in most cases, the estimated time for the progression of LC-MBL to CLL far exceeds a normal life expectancy. However, assessment of the clinical impact of LC-MBL, if any, through the comparison against age- and sex-matched non-MBL subjects showed that the OS observed for individuals with LC-MBL was significantly shorter than that of the control group. Moreover, individuals with LC-MBL showed a significantly shortened survival compared to that of age-matched individuals of the general population from the same geographical region. Infections, cancer and cardiovascular diseases were the main causes of mortality among individuals with LC-MBL. On the other hand, non-infectious respiratory tract diseases or genitourinary diseases, diabetes, dementia or other nervous system disorders accounted for almost 30% of deaths in the age- and sex-matched general population cohort. In the entire cohort, advanced age, co-existing cardiovascular diseases, solid tumors and, to a lesser extent, the presence of LC-MBL clones, were independently associated with a shorter OS (25).

'CLL-SPECIFIC' CYTOGENETIC ABERRATIONS ARE DETECTED IN MBL

The cytogenetic profile of HC-MBL is highly similar to that of CLL, especially of the indolent type (Rai 0 stage), in that del (13q14) was found to be the most frequent aberration followed

by trisomy 12 (1). On the other hand, compared to Rai 0 CLL, HC-MBL displayed a lower prevalence of cytogenetic aberrations related to aggressive disease, such as del(11q) and del(17p), that were often acquired as secondary abnormalities (17).

In contrast, LC-MBL showed a significantly lower frequency of 'CLL-specific' cytogenetic aberrations compared to HC-MBL and CLL (26). At the individual aberration level, a tendency was observed towards a greater frequency of del(13q) and trisomy 12 from LC-MBL to HC-MBL and CLL, yet the individual frequencies were not significantly different between the three. On these grounds, one could speculate that these lesions occur rather early during the development of MBL and may be associated with the acquisition of the 'CLL-like' immunophenotypic profile rather than overt disease (26). A much lower prevalence of del(11q) and del(17p) was observed in LC-MBL; even so, such cases were detected, with the aberration being present in the majority of clonal B cells (2). This finding indicates that the presence of del(17p) per se does not axiomatically correlate with aggressive disease, as also reported for a small group of patients with CLL carrying this abnormality yet remaining stable for prolonged periods, especially when expressing mutated IGHV genes (27, 28). The distribution of the most common "CLL-related" cytogenetic aberrations among LC-MBL, HC-MBL and CLL is depicted in **Figure 1**. Of note, although LC-MBL appears to display a relatively simple cytogenetic profile, longitudinal analysis disclosed a significant increase in the overall frequency of cytogenetic aberrations after seven years of follow-up (29% at baseline versus 62% at follow-up): importantly, all cytogenetic aberrations observed at baseline also persisted at follow-up (25).

Overall, these findings support that the progression from LC-MBL to HC-MBL and CLL is accompanied by the progressive acquisition of recurrent cytogenetic aberrations, each one of them enriched at specific stages along the natural history of the disease. In more detail, del(13q) and trisomy 12 occur rather early, at the LC-MBL stage, whereas del(17p) and del(11q) seem to emerge later, as secondary events at either the HC-MBL or the CLL stages. In line with this, the study of telomere length (TL) in both subtypes of MBL, indolent CLL and healthy subjects revealed the presence of significantly shorter telomeres already at the level of HC-MBL, suggesting that it may be part of the initial events in the process of CLL pathogenesis (26, 29).

Very few data is available for the profile of cytogenetic aberrations in 'atypical' and 'CD5-negative' subtypes of MBL; having said that, cases with del(13q), trisomy 12 and even del (17p) have been identified (1).

RECURRENT GENE MUTATIONS IN MBL

TP53 aberrations, including mutations, have an established role in shaping the outcome of CLL and, for this reason, their assessment is considered mandatory before treatment initiation (30). In recent years, the advent of NGS technologies led to the identification of additional gene mutations with putative clinical relevance in CLL. In particular, recurrent mutations in the *NOTCH1*, *SF3B1*, and *BIRC3* genes were associated with distinct outcomes and were utilized in risk stratification schemes in the pivotal studies (31–33). Subsequent studies by several groups, including ours, have identified mutations in additional genes, e.g. *RPS15*, *NFKBIE* and *EGR2*, revealing a much more complicated genetic landscape for CLL (34–36). To complicate things even more, the detailed study of the clonal architecture in CLL in comparison to treatment revealed that the presence of mutations in putative CLL driver genes can adversely impact the clinical outcome of the disease, even when present in minor subclones, at least for certain mutations (37).

Some of the same recurrent gene mutations have been also found in HC-MBL. That said, at the individual gene level, *NOTCH1* mutations exhibited a significantly lower prevalence

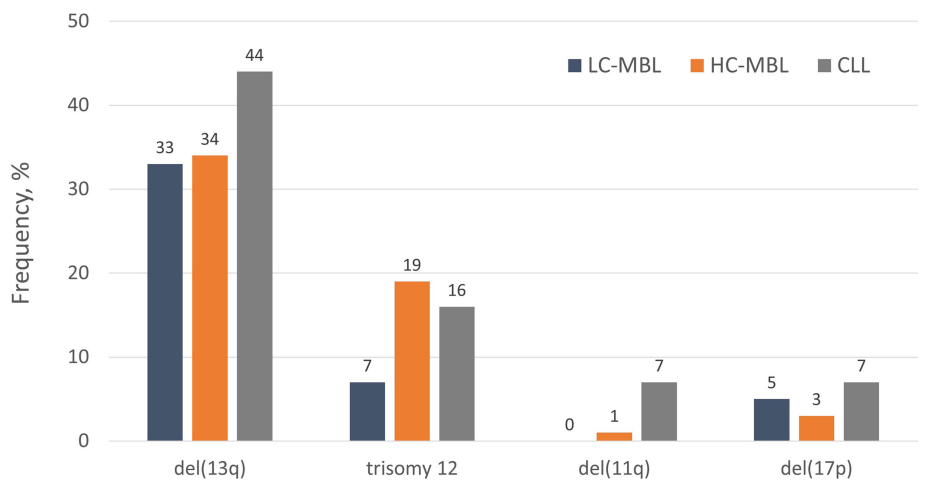


FIGURE 1 | The distribution of "CLL-specific" cytogenetic aberrations in LC-MBL, HC-MBL and indolent CLL. Data was extracted from the studies by Rossi et al (17), Fazi et al. (2) and Henriques et al. (26).

in HC-MBL compared to CLL (3.2% versus 11.6%, respectively) (P=0.050) (38); moreover, *SF3B1* mutations were even more scarce in HC-MBL (1.5%) versus CLL (~10%) (39). Interestingly, however, when larger cohorts were analyzed, the incidence of *NOTCH1* and *SF3B1* mutations was not significantly different between HC-MBL and Rai 0-CLL (*NOTCH1*: 8.2% *vs.* 13.1%; *SF3B1*: 4.7% *vs.* 3.8%) (**Figure 2**) (21). This finding supports the claim that MBL and early stage CLL are closely similar, differing only in clonal size, with MBL simply requiring more time to expand. This claim is further corroborated by the fact that low variant allele frequencies (VAF) were observed in almost all MBL cases but only in about half of CLL cases (41).

At the clinical level, both CLL and MBL cases bearing a *NOTCH1* mutation had a shorter progression free survival (PFS) compared to wildtype ones (41). Thus, arguably, gene mutations could be used as biomarkers for identifying those MBL cases that would eventually progress to CLL requiring treatment. Indeed, in a longitudinal analysis, 60% of MBL cases that progressed to require treatment exhibited a subclonal expansion bearing a driver gene mutation, compared to only 14% of untreated cases (11). On the other hand, a longitudinal analysis of eight MBL cases by whole exome sequencing (WES) at two time points of 65 months apart, revealed that the four cases who developed detectable lymphadenopathy by physical exam yet without need for treatment carried mutations in putative CLL driver genes, including *ATM*, *DDX3X*, *EGR2*, *FBXW7*, *SAMHD1* and *SF3B1* (42). Targeted re-sequencing not only validated the WES results, but also led to the identification of additional mutations in the *BIRC3, POT1* and *NOTCH1* genes. Interestingly, a damaging mutation affecting *PRDM1* was identified in a fifth case, which is frequently inactivated in diffuse large B-cell lymphomas. Of all, *FBXW7* was the only gene recurrently mutated and all mutations were found in highly evolutionarily conserved regions and considered damaging by PolyPhen and MutationTaster (42). This data supports the notion that 'CLL-associated' gene mutations are not necessarily clinically relevant when detected in MBL.

Along these lines, in our recent characterization of the genetic landscape in both subtypes of MBL as well as indolent CLL, we were able to detect gene mutations in CLL putative driver genes in all 3 entities, yet these were infrequent and did not have any obvious impact on disease progression after a prolonged follow up. On these grounds, we proposed that gene mutations may represent late events related mostly to disease progression, whereas interactions between the cell clone and its microenvironment, such as those mediated through the B cell receptor (BcR), could represent the major driver in the early stages of the natural history of CLL (40).

B CELL CLONAL DYNAMICS IN MBL: MONOCLONAL VERSUS MULTICLONAL CASES

Findings from low-throughput studies using mostly subcloning techniques studies reported that oligoclonality is a common feature of 'CLL-like' MBL (43), present in both LC-MBL and HC-MBL cases. In 2011, similar findings were obtained in the first relevant NGS study of 9 individuals with HC-MBL in whom multiple, immunogenetically independent MBL clones were identified, clearly suggesting distinct origins (44). The coexistence of multiple B cell expansions was also observed in a subsequent study as well, where the authors concluded that different subsets of normal B cells in the same individual may be targets for oncogenetic events involved in lymphomagenesis, arguing that bi- and multi-clonal MBL cases may be associated with distinct chronic antigen-driven immune responses (44),. This could be particularly frequent at earlier stages (LC-MBL), further evidence for the potential reactive nature of MBL among

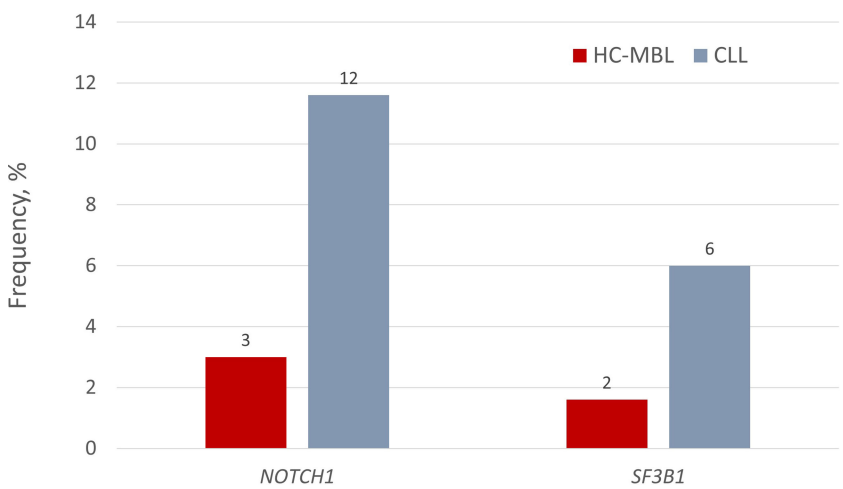


FIGURE 2 | Incidence of *NOTCH1* and *SF3B1* gene mutations in HC-MBL and indolent CLL. Respective values for LC-MBL were 0% in both cases. Data regarding the presence of gene mutations in LC-MBL was extracted from Agathangelidis et al. (40), data regarding *NOTCH1* mutations in HC-MBL and CLL from the study by Rasi et al. (38), and finally, the frequency of *SF3B1* mutations in HC-MBL and CLL from the studies by Greco et al. (39) and Rossi et al. (33), respectively.

individuals with normal lymphocyte counts, prior to the stepwise acquisition of genomic alterations and the (infrequent) progression to HC-MBL and CLL (45).

Of note, in the latter study, the median numbers of clonal B cells in both 'CLL-like' HC-MBL and CLL were significantly lower in multiclonal than in monoclonal cases (41). In contrast, the median number of clonal cells in 'CLL-like' LC-MBL was significantly higher in multiclonal than monoclonal cases. Yet, longitudinal analysis showed a significant increase in size in monoclonal compared to bi- and multi-clonal LC-MBL cases (86% versus 54%, respectively; p=0.004). In most studies, no statistically significant differences in IGHV gene usage were identified in monoclonal versus multiclonal cases. Interestingly, the number of multiclonal cases with stereotyped BcR IG was significantly higher than that of monoclonal cases. Finally, the frequency of cytogenetic alterations as well as genetic complexity in both 'CLL-like' MBL and CLL clones from multiclonal cases was significantly lower than that of monoclonal cases from these entities (45).

More recently, employing NGS, we systematically assessed the immunogenetic characteristics of LC-MBL. We found that 75% samples displayed a monoclonal profile characterized by the presence of a single clonotype dominating the repertoire (frequency range: 42.5-97.9%), whereas the remaining 25% of samples turned out to be oligoclonal (46). Clonality levels seemed to correlate with clone size, in line with previous studies reporting a lower frequency of oligoclonality along the transition from LC-MBL to HC-MBL and, eventually, CLL (47). **Figure 3** summarizes graphically the distribution of monoclonal and oligoclonal cases among the 2 subtypes of MBL and CLL.

Arguably, differences in clonality patterns among LC-MBL cases could indicate that these are captured at different stages along the ontogenetic trajectory. In this scenario, LC-MBL could initially involve a polyclonal B cell population that at some point acquired the CLL phenotype due to persistent antigenic stimulation. Microenvironmental interactions would induce further proliferation, eventually favoring the acquisition of genomic lesions, hence underlying the progressive transition to oligoclonal and, eventually, monoclonal LC-MBL cell populations. This evolutionary path could go in parallel with a continuous increase in clonal size, perhaps explaining the existence of multiclonal HC-MBL and CLL cases (25, 48).

LC-MBL HAS A DIFFERENT IMMUNOGENETIC PROFILE FROM EITHER HC-MBL OR CLL

The first studies reporting significant IGHV gene usage biases in the repertoire of CLL date from the 1990s (49). Subsequent studies in larger patient cohorts (50–52) cemented the notion that skewing of the IGHV repertoire in CLL results from strong selective pressures acting on the CLL progenitor cells by of a restricted group of antigens.

In order to obtain more insight into the reasons/ontogenetic timing repertoire skewing, several groups undertook comparative analyses of the IGHV gene repertoire in CLL and MBL. Early studies compared Rai 0 CLL with HC-MBL, yet the differences found were not statistically significant (17, 21). However, when comparisons extended to LC-MBL, that latter was found to display a distinct profile, including frequent expression of the IGHV4-59 and IGHV4-61 genes (rather infrequent in CLL), as well as pronounced scarcity of the IGHV4-34 and IGHV1-69 genes (26, 53). These findings suggested that HC-MBL and indolent CLL are closely related at the immunogenetic level, whereas LC-MBL is characterized by a rather unique repertoire, raising questions regarding its precise relationship to CLL (17, 43).

In the largest relevant study, our group applied Sanger sequencing in a cohort of 333 'CLL-like' MBL cases and

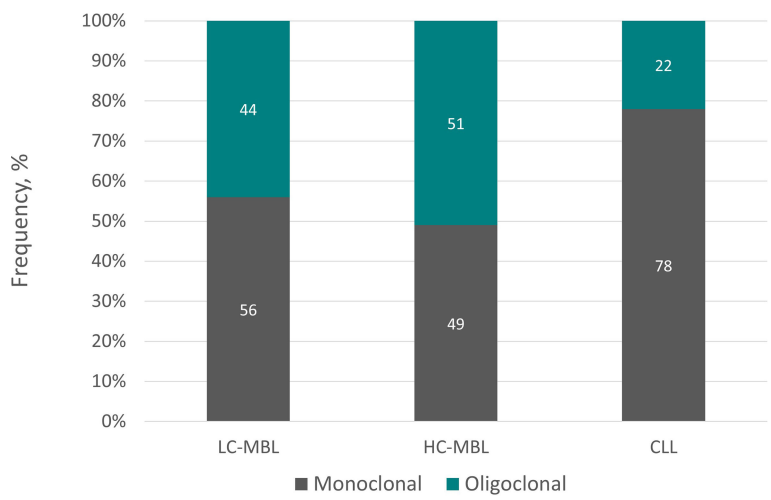


FIGURE 3 | Monoclonal versus oligoclonal case distribution in LC-MBL, HC-MBL and CLL. Data regarding LC-MBL were extracted from Agathangelidis et al. (46) and Henriques et al. (26), whereas the data concerning HC-MBL and CLL were obtained from the latter.

reported significant differences in the IGHV gene repertoire between HC-MBL versus LC-MBL; several genes were frequent only in the former (namely IGHV1-69, IGHV2-5, IGHV3-23, IGHV3-33, IGHV3-48 and IGHV4-34), whereas the IGHV4-59/61 genes was significantly more frequent in the latter. Comparison of both MBL subtypes against exclusively Rai-0 CLL cases or CLL cases from all stages disclosed several significant differences, again highlighting the distinct nature of the LC-MBL repertoire (**Figure 4**) (54). These findings strengthened the hypothesis that LC-MBL may not constitute a true pre-leukemic state but rather represent a manifestation of immune senescence.

More recently, we re-appraised the immunogenetic profile of LC-MBL employing a high-throughput approach (46). More particularly, employing NGS, we studied the IGHV gene repertoire from clonal and normal B cell samples from 23 individuals with LC-MBL, which we compared against the corresponding repertoires of naïve and memory B cell samples from 6 healthy individuals. The most intriguing finding of our study concerned the stronger repertoire skewing in normal B cells from individuals with LC-MBL compared to healthy individuals. This argues for different selection processes and functions of these B cell subpopulations in LC-MBL compared to healthy individuals (55).

DIFFERENT PATTERNS OF SOMATIC HYPERMUTATION IN MBL COMPARED TO CLL

The SHM status of the clonotypic IGHV gene is now considered one of the most accurate biomarkers for clinical decision making in CLL (30). Based on the SHM status, CLL patients can be robustly classified CLL into 2 distinct groups: (i) CLL with unmutated IGHV genes (U-CLL), with few or no SHM [i.e. cases with a germline identity equal to or higher than 98%]; and, (ii) CLL with mutated IGHV genes (M-CLL), with a considerable SHM load and a germline identity below 98%. This classification scheme holds strong prognostic relevance, with U-CLL cases generally experiencing a worse prognosis and a more aggressive disease accompanied by early need for treatment compared to M-CLL (56, 57). It has been postulated that these two groups might also have distinct ontogeny, with M-CLL possibly deriving from B cells that have been activated by antigen(s) and have participated in the germinal center (GC) reaction, and U-CLL cases possibly originating from B cells following GC-independent maturation pathways (48, 58).

Following the same, CLL-relevant 98% identity cut-off value, several studies analysed the distribution of mutated and unmutated cases in MBL. In our series, out of the 355 IGH gene rearrangements analysed from 333 'CLL-like' MBL cases, 267 (75.2%) were classified as mutated, while the rest (24.8%) were assigned to the unmutated subgroup. The distribution of unmutated and mutated IGH gene rearrangements in LC-MBL and HC-MBL was similar, and also comparable to Rai 0-CLL (26.4%, 24.4% and 25%, respectively). These 3 groups, however, differed significantly when compared against a CLL cohort containing 7,424 CLL patients from all stages, in which the frequency of unmutated cases was 45.1% (**Figure 5**) (54). In line with the above, both HC-MBL and Rai 0-CLL unmutated cases have been found to display distinct gene and miRNA expression profiles from the mutated MBL/CLL cases, independently of their classification (21).

STEREOTYPED BCR IG ARE INFREQUENT IN LC-MBL

A distinctive feature of the IG gene repertoire of CLL concerns the presence of (quasi)identical BcR IG shared by different patients, a phenomenon termed BcR IG stereotypy. At odds

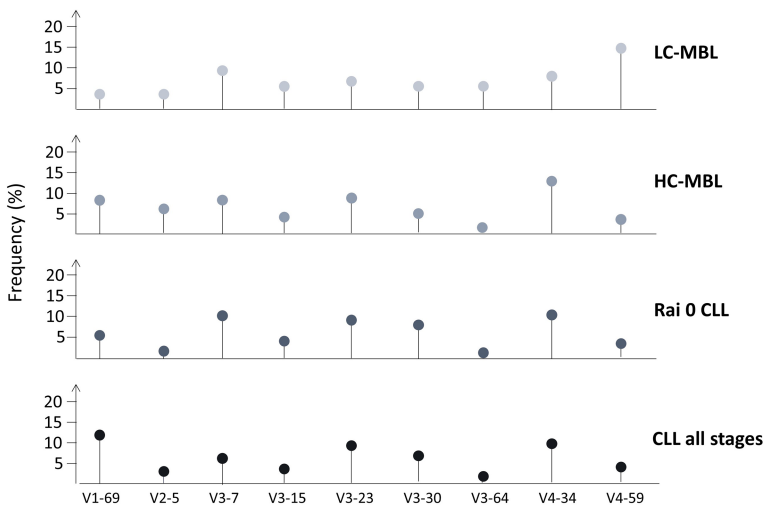


FIGURE 4 | IGHV gene repertoire differences between LC-MBL, HC-MBL, Rai 0 CLL as well as in CLL from all stages. Data was obtained from Vardi et al. (54).

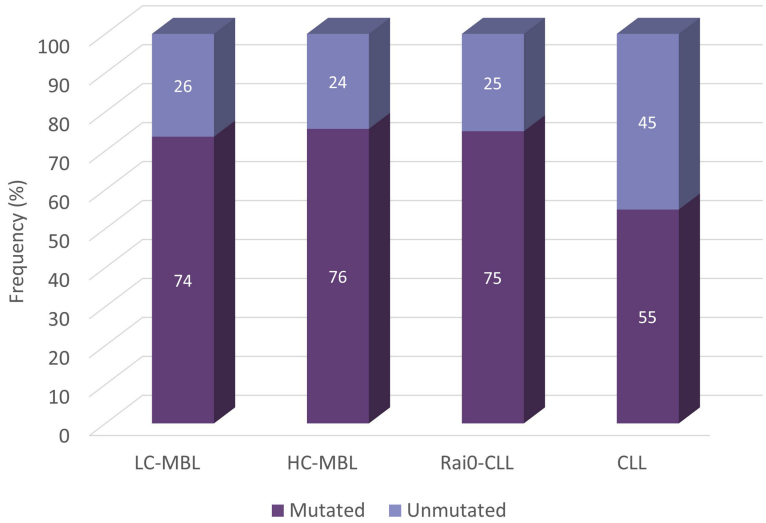


FIGURE 5 | Somatic hypermutation status in cohorts from LC-MBL, HC-MBL and CLL. Data was obtained from Vardi et al. (54).

with serendipity, whereby stereotyped BcR IG would be found with a negligible frequency (in the range of 10^{-12}-10^{-16}), stereotypy occurs in CLL at a remarkable frequency of 41% (59). Groups of patients expressing such restricted BCR IG are known as stereotyped subsets: they can be found in both U-CLL and M-CLL and vary in size from a handful to hundreds of cases, in which case they are deemed 'major' (52, 59). Notably, similarities between cases in stereotyped subsets extend from the primary IG sequences to biological and clinical profiles. Indeed, consistent clinical outcomes have been reported for patients assigned to the same subset (60–62), likely linked to consistent biological profiles: these pertain, amongst others, to distinctive gene expression, cytogenetic aberrations, BcR structure and signalling capacity, epigenetic modifications and antigenic reactivity profiles (63–71) that are also distinct between subsets but also from non-subset cases even of the same mutational category (i.e. U-CLL or M-CLL).

In this context, investigating whether 'CLL-like' stereotyped BcR IG could be detected also in MBL was a logical next step. When comparing HC-MBL with Rai 0-CLL, no significant differences were found, whereas, in sharp contrast, stereotyped BcR IG were exceedingly scarce in LC-MBL (17, 21). Interestingly, a correlation between the incidence of stereotypy and the absolute count of 'CLL-like' cells was identified: 5.5% in LC-MBL, 21.9% in HC-MBL, 20.2% in Rai 0-CLL, and 30.4% in CLL of all clinical stages (**Figure 6**) (54).

The scarcity of BcR IG stereotypes in LC-MBL was also evident in our recent high-throughput study. Of the few stereotyped clonotypes, virtually all were classified to minor subsets with a small minority corresponding to stereotypes defining major CLL stereotyped subsets, particularly those associated with aggressive disease. On these grounds, we proposed that LC-MBL is immunogenetically distinct from CLL, at least for those cases with aggressive clinical behavior (46).

T CELLS IN MBL AND CLL

Multifarious abnormalities in the T cell compartment constitute a well characterized feature of the tumor microenvironment in CLL (72–75). However, although our understanding of the interactions between T cells and CLL leukemic cells is continuously growing, several important questions remain, not lest regarding the precise role of T cells i.e. whether they exert pro- or anti-neoplastic actions.

CLL is characterized by a relative loss of naïve $CD4^+$ T cells accompanied by enrichment of antigen-experienced, memory and effector $CD4^+$ T cells (73, 76). $CD4^+$ T cells in CLL express higher levels of PD-1, human leukocyte antigen (HLA)-DR and Ki67, which are all activation markers (73, 77, 78). *In vitro* as well as *in vivo* studies in xenograft mouse models suggested that $CD4^+$ T cells could enhance the survival and proliferation of clonal cells in CLL, a finding that was supported by correlations between $CD4^+$ T cell counts and clinical outcome (79, 80).

$CD8^+$ T cells from patients with CLL display high expression of PD-1 and other inhibitory receptors, accompanied by defective immune synapse formation (also occurring in $CD4^+$ T cells as well) (78, 81). For these reasons, $CD8^+$ T cells in CLL have been described as "pseudoexhausted" (78). Relevant experiments in mice showed that the exhaustion of T cells was anatomically restricted, being observed mainly in spleen samples and to a much lower extent in the blood (81). Impaired immune synapse formation was observed *ex vivo* when T cells from healthy donors were co-cultured with clonal CLL cells, suggesting that this defect was largely induced by the CLL cells themselves. Of interest, blocking inhibitory ligands, such as CD200, PD-L1, or CD276 with neutralizing antibodies led to restored immune synapse formation capacity (82). Regarding the biological function of $CD8^+$ T cells in CLL, apparently they may recognize tumor specific antigens, however they ultimately fail to control the disease likely due to their functional exhaustion (76).

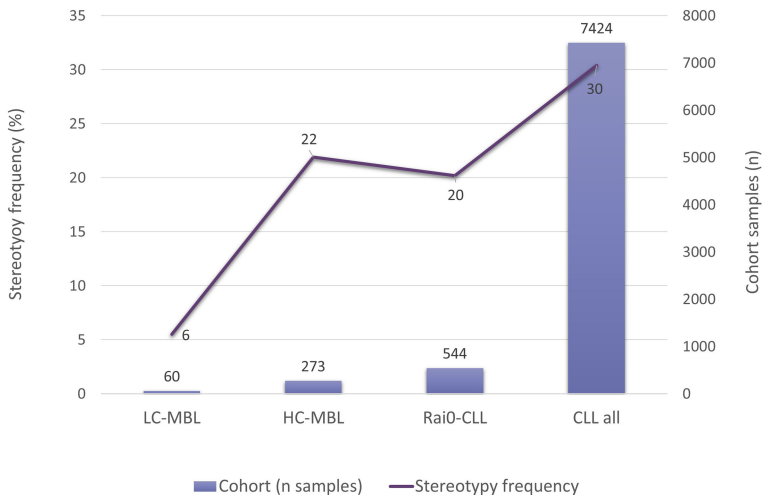


FIGURE 6 | Frequency of BcR IG stereotypy in LC-MBL, HC-MBL, Rai 0 CLL and CLL from all clinical stages. Data was extracted from Vardi et al. (54).

Characteristic T cell receptor (TR) gene repertoire restrictions have been established in CLL, mainly in the $CD8^+$ but also in the $CD4^+$ compartment (83). Even though clonal expansions occur frequently in the T cell repertoire of the elderly, restrictions were more pronounced in CLL, indicating the existence of stronger and perhaps different selective pressures. Whether these pressures are exerted by the same antigens interacting with the CLL clone or by CLL-derived antigens remains to be elucidated (83, 84).

Comparative analysis between HC-MBL, U-CLL and M-CLL cases revealed similar $CD4^+$ T cell counts in all three conditions. In contrast, the $CD8^+$ T cell count was higher in U-CLL compared to M-CLL patients and significantly higher than that of HC-MBL (85). Interestingly, both $CD4^+$ and $CD8^+$ clonal T cell expansions were identified in HC-MBL, similar to CLL. $CD4^+$ T cell clonal expansions appeared to follow the numerical increase of clonal B cells, thus being more pronounced in CLL compared to HC-MBL (85). This finding suggests that $CD4^+$ TR repertoire restrictions may be somehow influenced by the size of B cell clonal expansions, occurring early in clonal evolution and increasing concurrently with tumor progression. The fact that CLL displays higher T cell repertoire restrictions could also reflect the loss of effector T cell clones restraining CLL clonal expansions. In line with this, treatment with the Bruton's tyrosine kinase inhibitor ibrutinib, led to an increase in the T cell repertoire diversity; this suggests that the eradication of the CLL malignant cells sets the stage for T cell reconstitution (86).

Concerning the clonality of the $CD8^+$ T cell fraction, no significant differences were detected between HC-MBL and CLL (87). Furthermore, a longitudinal analysis showed that T cell clones, mainly from the $CD8^+$ compartment, persisted over time in all MBL samples analyzed, similar to what was previously reported in CLL (83, 88), suggesting that interactions between CLL and T cells are evident prior to/at the early stages of CLL development (87). Moreover, decreased numbers of $CD4^+CD8^+$ double positive T cells were identified in HC-MBL compared to healthy control samples, alluding to impaired immunosurveillance which may actually favor the emergence of MBL clones. Furthermore, Tregs appeared increased in CLL, but not in MBL, compared to healthy controls (89).

Regarding LC-MBL, early studies reported an increase in the size of clonal T cell populations compared to the general population, especially for double positive $CD4^+CD8^+$ T cells, suggesting a general deregulation of the immune system, possibly related to chronic antigenic stimulation (90). Recently, a follow-up study of LC-MBL revealed statistically significant increase of total T cells as well as $CD4^+$, $CD8^+$ and the double negative cell subpopulations over time. Interestingly, LC-MBL with larger clone sizes over time also showed significantly higher ($P<0.05$) numbers of the distinct normal residual T cell subsets at follow up compared to the baseline, especially for $CD4^+$ T cells (79, 91). This might imply that signals emanating from immune cells of the CLL microenvironment, potentially promoting activation, proliferation and/or survival of B cells, could contribute to the expansion of 'CLL-like' B cell clones at the earliest stages of CLL (25).

In our recent study of the TR beta chain gene repertoire in LC-MBL by NGS, we reported more pronounced T cell expansions in LC-MBL compared to aged-matched healthy individuals, yet lower than CLL. We attributed this finding to different antigenic pressures, leading to more pronounced T cell expansions in CLL versus LC-MBL versus healthy individuals (92). Moreover, a significant correlation between the level of T cell clonality and the size of the MBL clone was shown, in line with previous findings in CLL, further supporting the notion that T cell expansions could be driven by CLL-associated antigens (83, 88). Pairwise comparisons between all entities revealed distinct TRBV gene repertoire biases; this could perhaps imply that different antigen selection pressures operate in LC-MBL, CLL or healthy individuals. The latter assumption was further supported by the fact that shared clonotypes between different entities were scant. The nature of the implicated antigens and

whether they are related to the expanded 'CLL-like' cell clones remains to be clarified (92).

THE NUMBERS OF RESIDUAL NORMAL B CELLS ARE REDUCED PRIOR TO CLL ONSET

Several defects in both the innate and the adaptive immune responses have been identified in CLL, including hypogammaglobulinemia (93). The exact mechanisms underlying these defects are poorly characterized, whereas even less information is available regarding the disease stage during which they take place. In this context, Criado and colleagues (55) studied the residual normal B cell compartment in the blood of individuals with LC-MBL and HC-MBL compared to Rai 0-CLL in order to get insight into the mechanisms that may contribute to the emergence of hypogammaglobulinemia in CLL. Both HC-MBL and Rai 0-CLL showed significantly reduced normal B cell counts, mostly at the expense of pre-GC B cells, both immature and naïve, suggesting diminished production and/or release of these cells in the circulation at the early stages of CLL. Considering that no such decrease has been identified in circulating immature and naïve B cell counts in relation to ageing (94), the decreased numbers of circulating pre-GC B cells in individuals with HC-MBL alludes to impaired production of newly generated B cells in the BM, even prior to the development of CLL (55).

When focusing on the memory B cells (MBC) and plasma cells (PC), there were no significant differences in either MBL (both subtypes) or CLL compared to healthy controls with the exception of PC counts between HC-MBL and the healthy controls. Of note, both MBC and PC showed significantly different distribution among B cell subsets, expressing distinct IG subclasses in both MBL and CLL. Of importance, the extent of these changes was greater from LC-MBL to HC-MBL and finally Rai0-CLL. Thus, the possibility of a progressive impairment of B cell responses driven by newly encountered antigens along the transition from LC-MBL to HC-MBL and finally Rai 0-CLL was proposed; impaired pre-GC B cell production has been considered as largely responsible, since it could lead to a progressive reduction of the diversity of the BcR IG repertoire from LC-MBL to HC-MBL and Rai 0-CLL (55). This would be in line with other features of CLL but also MBL, such as the launch of active, but silent, responses against common pathogens, particularly viruses such as cytomegalovirus (CMV) and the Epstein Barr virus (EBV), as well as bacteria such as *S. pneumoniae* (95). Additional, longitudinal studies in larger MBL and CLL cohorts would be necessary to support this hypothesis.

CONCLUSIONS

MBL detection is based on the identification of B cell expansions in the blood circulation with the characteristic "CLL-specific" phenotype, yet of a smaller size than the one required for CLL diagnosis. This criterion, based solely on a mathematical cutoff is devoid of any biological context, perhaps hampering our understanding of the mechanisms driving disease onset. Thus, the in-depth characterization of MBL holds great potential for understanding the mechanisms that represent major drivers in the process of its transformation to CLL.

Nowadays, it is well established that CLL is characterized by a highly complicated genomic landscape, including a series of cytogenetic aberrations as well as gene mutations. That said, the overall simple genetic background of MBL pointed towards the BcR IG as one of the major "players" prior to disease onset. The study of BcR IG clonality, especially through NGS-based methodologies, revealed much a much higher frequency of oligoclonality in MBL, especially LC-MBL, compared to CLL, most likely reflecting different staged along the ontogenetic trajectory. Relevant to mention, the BcR IG repertoire in HC-MBL was similar to CLL regarding both the IGHV gene repertoire as well as BcR IG stereotypy, whereas LC-MBL showed a highly distinct repertoire characterized by the dominance of different IGHV genes and a very low prevalence of stereotypes, alluding to different antigenic pressures shaping the BcR IG repertoire in LC-MBL versus HC-MBL and CLL. Further support for an early role of the microenvironment in CLL ontogenesis came from the study of the TR repertoire in both LC- and HC-MBL. Similar to CLL, T cell expansions were evident in both subtypes of MBL yet at a lower extent. As above, the TR repertoire was characterized by different biases in the repertoire of TRBV genes and the absence of shared clonotyped among MBL and CLL pointing towards different selection mechanisms. Last, the study of residual B cells in MBL and CLL revealed an impaired production of newly generated B cells in the BM, even prior to the development of CLL, as well as a progressive impairment of B cell responses driven by newly encountered antigens along the transition from LC-MBL to HC-MBL and finally Rai 0-CLL that could account for the clinical impact of MBL detection.

Overall, even if major questions still remain, parallel analysis of the B cell clone(s) and the tumor microenvironment in MBL holds strong potential for characterizing the mechanisms that are required for the emergence and malignant transformation of a monoclonal B cell population.

AUTHOR CONTRIBUTIONS

All authors contributed to the article and approved the submitted version. CG, LZ-I, and AI wrote the manuscript. AC, KS, and AA edited the text and gave final approval.

FUNDING

This work was supported in part by the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT), under grant agreement No 336

(Project CLLon); COSMIC, a Marie Curie European Training Network funded from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 765158; the project ODYSSEAS (Intelligent and Automated Systems for enabling the Design, Simulation and Development of Integrated Processes and Products) implemented under the "Action for the Strategic Development on the Research and Technological Sector", funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union, with grant agreement no: MIS 5002462; NEoteRIC, funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No 871330; the Greek Precision Medicine Network (GPMN), a mission of the Research and Innovation sector of the Ministry of Education, Research and Religious Affairs of Greece.

REFERENCES

1. Lanasa MC, Allgood SD, Slager SL, Dave SS, Love C, Marti GE, et al. Immunophenotypic and Gene Expression Analysis of Monoclonal B-Cell Lymphocytosis Shows Biologic Characteristics Associated With Good Prognosis CLL. *Leukemia* (2011) 25(9):1459–66. doi: 10.1038/leu.2011.117
2. Fazi C, Scarfò L, Pecciarini L, Cottini F, Dagklis A, Janus A, et al. General Population Low-Count CLL-Like MBL Persists Over Time Without Clinical Progression, Although Carrying the Same Cytogenetic Abnormalities of CLL. *Blood* (2011) 118(25):6618–25. doi: 10.1182/blood-2011-05-357251
3. Maitre E, Troussard X. Monoclonal B-Cell Lymphocytosis. *Best Pract Res Clin Haematol* (2019) 32(3):229–38. doi: 10.1016/j.beha.2019.06.002
4. Scarfò L, Ghia P. What Does It Mean I Have a Monoclonal B-Cell Lymphocytosis?: Recent Insights and New Challenges. *Semin Oncol* (2016) 43(2):201–8. doi: 10.1053/j.seminoncol.2016.02.013
5. Rawstron AC. Monoclonal B Cell Lymphocytosis - What Does It Really Mean? *Curr Hematol Malig Rep* (2013) 8(1):52–9. doi: 10.1007/s11899-012-0144-z
6. Kostopoulos IV, Paterakis G, Papadimitriou K, Pavlidis D, Tsitsilonis OE, Papadhimitriou SI. Immunophenotypic Analysis Reveals Heterogeneity and Common Biologic Aspects in Monoclonal B-Cell Lymphocytosis. *Genes Chromosomes Cancer* (2015) 54(4):210–21. doi: 10.1002/gcc.22234
7. de Tute R, Yuille M, Catovsky D, Houlston RS, Hillmen P, Rawstron AC. Monoclonal B-Cell Lymphocytosis (MBL) in CLL Families: Substantial Increase in Relative Risk for Young Adults [3]. *Leukemia* (2006) 20(4):728–9. doi: 10.1038/sj.leu.2404116
8. Goldin LR, Lanasa MC, Slager SL, Cerhan JR, Vachon CM, Strom SS, et al. Common Occurrence of Monoclonal B-Cell Lymphocytosis Among Members of High-Risk CLL Families. *Br J Haematol* (2010) 151(2):152–8. doi: 10.1111/j.1365-2141.2010.08339.x
9. Fazi C, Dagklis A, Cottini F, Scarfò L, Bertilaccio MTS, Finazzi R, et al. Monoclonal B Cell Lymphocytosis in Hepatitis C Virus Infected Individuals. *Cytom Part B - Clin Cytom* (2010) 78(SUPPL. 1):61–8. doi: 10.1002/cyto.b.20545
10. Shanafelt TD, Kay NE, Jenkins G, Call TG, Zent CS, Jelinek DF, et al. B-Cell Count and Survival: Differentiating Chronic Lymphocytic Leukemia From Monoclonal B-Cell Lymphocytosis Based on Clinical Outcome. *Blood* (2009) 113(18):4188–96. doi: 10.1182/blood-2008-09-176149
11. Barrio S, Shanafelt TD, Ojha J, Chaffee KG, Secreto C, Kortüm KM, et al. Genomic Characterization of High-Count MBL Cases Indicates That Early Detection of Driver Mutations and Subclonal Expansion Are Predictors of Adverse Clinical Outcome. *Leukemia* (2017) 31(1):170–6. doi: 10.1038/leu.2016.172
12. Angelillo P, Capasso A, Ghia P, Scarfò L. Monoclonal B-Cell Lymphocytosis: Does the Elderly Patient Need a Specialistic Approach? *Eur J Intern Med* (2018) 58:2–6. doi: 10.1016/j.ejim.2018.09.006
13. Landgren O, Albitar M, Ma W, Abbasi F, Hayes RB, Ghia P, et al. B-Cell Clones as Early Markers for Chronic Lymphocytic Leukemia. *N Engl J Med* (2009) 360(7):659–67. doi: 10.1056/NEJMoa0806122
14. Rawstron AC, Bennett FL, O'Connor SJM, Kwok M, Fenton JAL, Plummer M, et al. Monoclonal B-Cell Lymphocytosis and Chronic Lymphocytic Leukemia. *N Engl J Med* (2008) 359(6):575–83. doi: 10.1056/NEJMoa075290
15. Shanafelt TD, Kay NE, Rabe KG, Call TG, Zent CS, Maddocks K, et al. Brief Report: Natural History of Individuals With Clinically Recognized Monoclonal B-Cell Lymphocytosis Compared With Patients With Rai 0 Chronic Lymphocytic Leukemia. *J Clin Oncol* (2009) 27(24):3959–63. doi: 10.1200/JCO.2008.21.2704
16. Molica S, Mauro FR, Giannarelli D, Lauria F, Cortelezzi A, Brugiatelli M, et al. Differentiating Chronic Lymphocytic Leukemia From Monoclonal B-Lymphocytosis According to Clinical Outcome on Behalf of the GIMEMA Chronic Lymphoproliferative Diseases Working Group. *Haematologica* (2011) 96(2):277–83. doi: 10.3324/haematol.2010.030189
17. Rossi D, Sozzi E, Puma A, De Paoli L, Rasi S, Spina V, et al. The Prognosis of Clinical Monoclonal B Cell Lymphocytosis Differs From Prognosis of Rai 0 Chronic Lymphocytic Leukaemia and Is Recapitulated by Biological Risk Factors. *Br J Haematol* (2009) 146(1):64–75. doi: 10.1111/j.1365-2141.2009.07711.x
18. Parikh SA, Chaffee KG, Larson MC, Hampel PJ, Call TG, Ding W, et al. Outcomes of a Large Cohort of Individuals With Clinically Ascertained High-Count Monoclonal B-Cell Lymphocytosis. *Haematologica* (2018) 103(6):e237–40. doi: 10.3324/haematol.2017.183194
19. Strati P, Shanafelt TD. Monoclonal B-Cell Lymphocytosis and Early-Stage Chronic Lymphocytic Leukemia: Diagnosis, Natural History, and Risk Stratification. *Blood* (2015) 126(4):454–62. doi: 10.1182/blood-2015-02-585059
20. Mulligan CS, Thomas ME, Mulligan SP. Lymphocytes B Lymphocytes, and Clonal CLL Cells: Observations on the Impact of the New Diagnostic Criteria in the 2008 Guidelines for Chronic Lymphocytic Leukemia (CLL). *Blood* (2009) 113(25):6496–7. doi: 10.1182/blood-2008-07-166710
21. Morabito F, Mosca L, Cutrona G, Agnelli L, Tuana G, Ferracin M, et al. Clinical Monoclonal B Lymphocytosis Versus Rai 0 Chronic Lymphocytic Leukemia: A Comparison of Cellular, Cytogenetic, Molecular, and Clinical Features. *Clin Cancer Res* (2013) 19(21):5890–900. doi: 10.1158/1078-0432.CCR-13-0622
22. Rawstron AC, Shingles J, De Tute R, Bennett F, Jack AS, Hillmen P. Chronic Lymphocytic Leukaemia (CLL) and CLL-Type Monoclonal B-Cell Lymphocytosis (MBL) Show Differential Expression of Molecules Involved in Lymphoid Tissue Homing. *Cytom Part B - Clin Cytom* (2010) 78(SUPPL. 1):42–6. doi: 10.1002/cyto.b.20534
23. Parikh SA. Chronic Lymphocytic Leukemia Treatment Algorithm 2018. *Blood Cancer J* (2018) 8(10):93. doi: 10.1038/s41408-018-0131-2
24. Shanafelt TD, Ghia P, Lanasa MC, Landgren O, Rawstron AC. Monoclonal B-Cell Lymphocytosis (MBL): Biology, Natural History and Clinical Management. *Leukemia* (2010) 24(3):512–20. doi: 10.1038/leu.2009.287
25. Criado I, Rodríguez-Caballero A, Gutiérrez ML, Pedreira CE, Alcoceba M, Nieto W, et al. Low-Count Monoclonal B-Cell Lymphocytosis Persists After Seven Years of Follow Up and Is Associated With a Poorer Outcome. *Haematologica* (2018) 103(7):1198–208. doi: 10.3324/haematol.2017.183954
26. Henriques A, Rodríguez-Caballero A, Nieto WG, Langerak AW, Criado I, Lécrevisse Q, et al. Combined Patterns of IGHV Repertoire and Cytogenetic/Molecular Alterations in Monoclonal B Lymphocytosis Versus Chronic Lymphocytic Leukemia. *PloS One* (2013) 8(7):1–10. doi: 10.1371/journal.pone.0067751
27. Yu L, Kim HT, Kasar S, Benien P, Du W, Hoang K, et al. Survival of Del17p CLL Depends on Genomic Complexity and Somatic Mutation. *Clin Cancer Res* (2017) 23(3):735–45. doi: 10.1158/1078-0432.CCR-16-0594
28. Yi S, Li Z, Zou D, Xiong W, Li H, Cui R, et al. Del17p Does Not Always Significantly Influence the Survival of B-Cell Chronic Lymphoproliferative Disorders. *Oncotarget* (2018) 9(3):3353–64. doi: 10.18632/oncotarget.23261
29. Furtado FM, Scheucher PS, Santana BA, Scatena NF, Calado RT, Rego EM, et al. Telomere Length Analysis in Monoclonal B-Cell Lymphocytosis and

Chronic Lymphocytic Leukemia Binet a. *Braz J Med Biol Res* (2017) 50(5):1–4. doi: 10.1590/1414-431x20176019
30. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner H, et al. iwCLL Guidelines for Diagnosis, Indications for Treatment, Response Assessment, and Supportive Management of CLL. *Blood* (2018) 131 (25):2745–60. doi: 10.1182/blood-2017-09-806398
31. Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, et al. Clinical Impact of Clonal and Subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM Mutations in Chronic Lymphocytic Leukemia. *Blood* (2016) 127(17):2122–30. doi: 10.1182/blood-2015-07-659144
32. Cortese D, Sutton LA, Cahill N, Smedby KE, Geisler C, Gunnarsson R, et al. On the Way Towards a "CLL Prognostic Index": Focus on TP53, BIRC3, SF3B1, NOTCH1 and MYD88 in a Population-Based Cohort. *Leukemia* (2014) 28(3):710–3. doi: 10.1038/leu.2013.333
33. Rossi D, Bruscaggin A, Spina V, Rasi S, Khiabanian H, Messina M, et al. Mutations of the SF3B1 Splicing Factor in Chronic Lymphocytic Leukemia: Association With Progression and Fludarabine-Refractoriness. *Blood J Am Soc Hematol* (2011) 118(26):6904–8. doi: 10.1182/blood-2011-08-373159
34. Ljungström V, Cortese D, Young E, Pandzic T, Mansouri L, Plevova K, et al. Whole-Exome Sequencing in Relapsing Chronic Lymphocytic Leukemia: Clinical Impact of Recurrent RPS15 Mutations. *Blood* (2016) 127(8):1007–16. doi: 10.1182/blood-2015-10-674572
35. Mansouri L, Sutton LA, Ljungström V, Bondza S, Arngården L, Bhoi S, et al. Functional Loss of Iκbε Leads to NF-κb Deregulation in Aggressive Chronic Lymphocytic Leukemia. *J Exp Med* (2015) 212(6):833–43. doi: 10.1084/jem.20142009
36. Young E, Noerenberg D, Mansouri L, Ljungström V, Frick M, Sutton LA, et al. EGR2 Mutations Define a New Clinically Aggressive Subgroup of Chronic Lymphocytic Leukemia. *Leukemia* (2017) 31(7):1547–54. doi: 10.1038/leu.2016.359
37. Rossi D, Khiabanian H, Spina V, Ciardullo C, Bruscaggin A, Famà R, et al. Clinical Impact of Small TP53 Mutated Subclones in Chronic Lymphocytic Leukemia. *Blood* (2014) 123(14):2139–47. doi: 10.1182/blood-2013-11-539726
38. Rasi S, Monti S, Spina V, Foà R, Gaidano G, Rossi D. Analysis of NOTCH1 Mutations in Monoclonal B-Cell Lymphocytosis. *Haematologica* (2012) 97 (1):153–4. doi: 10.3324/haematol.2011.053090
39. Greco M, Capello D, Bruscaggin A, Spina V, Rasi S, Monti S, et al. Analysis of SF3B1 Mutations in Monoclonal B-Cell Lymphocytosis. *Hematol Oncol* (2013) 31(1):54–5. doi: 10.1002/hon.2013
40. Agathangelidis A, Ljungström V, Scarfò L, Fazi C, Gounari M, Pandzic T, et al. Highly Similar Genomic Landscapes in Monoclonal B-Cell Lymphocytosis and Ultra-stable Chronic Lymphocytic Leukemia With Low Frequency of Driver Mutations *Haematologica* (2018) 103(5):865–73. doi: 10.3324/haematol.2017.177212
41. Lionetti M, Fabris S, Cutrona G, Agnelli L, Ciardullo C, Matis S, et al. High-Throughput Sequencing for the Identification of NOTCH1 Mutations in Early Stage Chronic Lymphocytic Leukaemia: Biological and Clinical Implications. *Br J Haematol* (2014) 165(5):629–39. doi: 10.1111/bjh.12800
42. Ojha J, Secreto C, Rabe K, Ayres-Silva J, Tschumper R, Dyke DV, et al. Monoclonal B-Cell Lymphocytosis Is Characterized by Mutations in CLL Putative Driver Genes and Clonal Heterogeneity Many Years Before Disease Progression. *Leukemia* (2014) 28(12):2395–8. doi: 10.1038/leu.2014.226
43. Dagklis A, Fazi C, Sala C, Cantarelli V, Scielzo C, Massacane R, et al. The Immunoglobulin Gene Repertoire of Low-Count Chronic Lymphocytic Leukemia (CLL)-Like Monoclonal B Lymphocytosis Is Different From CLL: Diagnostic Implications for Clinical Monitoring. *Blood* (2009) 114(1):26–32. doi: 10.1182/blood-2008-09-176933
44. Lanasa MC, Allgood SD, Volkheimer AD, Gockerman JP, Whitesides JF, Goodman BK, et al. Single-Cell Analysis Reveals Oligoclonality Among Low-Count Monoclonal B-Cell Lymphocytosis. *Leukemia* (2010) 24(1):133–40. doi: 10.1038/leu.2009.192
45. Henriques A, Rodríguez-Caballero A, Criado I, Langerak AW, Nieto WG, Lécrevisse Q, et al. Molecular and Cytogenetic Characterization of Expanded B-Cell Clones From Multiclonal Versus Monoclonal B-Cell Chronic Lymphoproliferative Disorders. *Haematologica* (2014) 99(5):897–907. doi: 10.3324/haematol.2013.098913
46. Agathangelidis A, Galigalidou C, Scarfò L, Moysiadis T, Rovida A, Gounari M, et al. Infrequent "Chronic Lymphocytic Leukemia-Specific" Immunoglobulin Stereotypes in Aged Individuals With or Without Low-Count Monoclonal B-Cell Lymphocytosis. *Haematologica* (2021) 106(4):1178–81. doi: 10.3324/haematol.2020.247908
47. de Faria-Moss M, Yamamoto M, Arrais-Rodrigues C, Criado I, Gomes CP, de Lourdes Chauffaille M, et al. High Frequency of Chronic Lymphocytic Leukemia-Like Low-Count Monoclonal B-Cell Lymphocytosis in Japanese Descendants Living in Brazil. *Haematologica* (2020) 105(6):e298–301. doi: 10.3324/haematol.2019.230813
48. Gemenetzi K, Agathangelidis A, Zaragoza-Infante L, Sofou E, Papaioannou M, Chatzidimitriou A, et al. B Cell Receptor Immunogenetics in B Cell Lymphomas: Immunoglobulin Genes as Key to Ontogeny and Clinical Decision Making. *Front Oncol* (2020) 10:67. doi: 10.3389/fonc.2020.00067
49. Fais F, Ghiotto F, Hashimoto S, Sellars B, Valetto A, Allen SL, et al. Chronic Lymphocytic Leukemia B Cells Express Restricted Sets of Mutated and Unmutated Antigen Receptors. *J Clin Invest* (1998) 102(8):1515–25. doi: 10.1172/JCI3009
50. Murray F, Darzentas N, Hadzidimitriou A, Tobin G, Boudjogra M, Scielzo C, et al. Stereotyped Patterns of Somatic Hypermutation in Subsets of Patients With Chronic Lymphocytic Leukemia: Implications for the Role of Antigen Selection in Leukemogenesis. *Blood* (2008) 111(3):1524–33. doi: 10.1182/blood-2007-07-099564
51. Stamatopoulos K, Belessi C, Moreno C, Boudjograh M, Guida G, Smilevska T, et al. Over 20% of Patients With Chronic Lymphocytic Leukemia Carry Stereotyped Receptors: Pathogenetic Implications and Clinical Correlations. *Blood* (2007) 109(1):259–70. doi: 10.1182/blood-2006-03-012948
52. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, et al. Stereotyped B-Cell Receptors in One-Third of Chronic Lymphocytic Leukemia: A Molecular Classification With Implications for Targeted Therapies. *Blood* (2012) 119(19):4467–75. doi: 10.1182/blood-2011-11-393694
53. Scarfò L, Zibellini S, Tedeschi A, Maura F, Neri A, Bertazzoni P, et al. Impact of B-Cell Count and Imaging Screening in cMBL: Any Need to Revise the Current Guidelines. *Leukemia* (2012) 26(7):1703–7. doi: 10.1038/leu.2012.20
54. Vardi A, Dagklis A, Scarfo L, Jelinek D, Newton D, Bennett F, et al. Immunogenetics Shows That Not All MBL Are Equal: The Larger the Clone, the More Similar to CLL. *Blood* (2013) 121(22):4521–8. doi: 10.1182/blood-2012-12-471698
55. Criado I, Blanco E, Rodríguez-Caballero A, Alcoceba M, Contreras T, Gutiérrez ML, et al. Residual Normal B-Cell Profiles in Monoclonal B-Cell Lymphocytosis Versus Chronic Lymphocytic Leukemia. *Leukemia* (2018) 32 (12):2701–5. doi: 10.1038/s41375-018-0164-3
56. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, et al. Ig V Gene Mutation Status and CD38 Expression as Novel Prognostic Indicators in Chronic Lymphocytic Leukemia. *Blood* (1999) 94(6):1840–7. doi: 10.1182/blood.V94.6.1840
57. 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(6):1848–54. doi: 10.1182/blood.V94.6.1848
58. Chiorazzi N, Ferrarini M. Cellular Origin(s) of Chronic Lymphocytic Leukemia: Cautionary Notes and Additional Considerations and Possibilities. *Blood* (2011) 117(6):1781–91. doi: 10.1182/blood-2010-07-155663
59. Agathangelidis A, Chatzidimitriou A, Gemenetzi K, Giudicelli V, Karypidou M, Plevova K, et al. Higher-Order Connections Between Stereotyped Subsets: Implications for Improved Patient Classification in CLL. *Blood* (2021) 137 (10):1365–76. doi: 10.1182/blood.2020007039
60. Baliakas P, Hadzidimitriou A, Sutton LA, Minga E, Agathangelidis A, Nichelatti M, et al. Clinical Effect of Stereotyped B-Cell Receptor Immunoglobulins in Chronic Lymphocytic Leukaemia: A Retrospective Multicentre Study. *Lancet Haematol* (2014) 1(2):e74–84. doi: 10.1016/S2352-3026(14)00005-2
61. Baliakas P, Agathangelidis A, Hadzidimitriou A, Sutton LA, Minga E, Tsanousa A, et al. Not All IGHV3-21 Chronic Lymphocytic Leukemias Are Equal: Prognostic Considerations. *Blood* (2015) 125(5):856–9. doi: 10.1182/blood-2014-09-600874

62. Jaramillo S, Agathangelidis A, Schneider C, Bahlo J, Robrecht S, Tausch E, et al. Prognostic Impact of Prevalent Chronic Lymphocytic Leukemia Stereotyped Subsets: Analysis Within Prospective Clinical Trials of the German CLL Study Group. *Haematologica* (2020) 105(11):2598–607. doi: 10.3324/haematol.2019.231027
63. Marincevic M, Cahill N, Gunnarsson R, Isaksson A, Mansouri M, Göransson H, et al. High-Density Screening Reveals a Different Spectrum of Genomic Aberrations in Chronic Lymphocytic Leukemia Patients With "Stereotyped" IGHV3-21 and IGHV4-34 B-Cell Receptors. *Haematologica* (2010) 95 (9):1519–25. doi: 10.3324/haematol.2009.021014
64. Kanduri M, Marincevic M, Halldórsdóttir AM, Mansouri L, Junevik K, Ntoufa S, et al. Distinct Transcriptional Control in Major Immunogenetic Subsets of Chronic Lymphocytic Leukemia Exhibiting Subset-Biased Global DNA Methylation Profiles. *Epigenetics* (2012) 7(12):1435–42. doi: 10.4161/epi.22901
65. Marincevic M, Mansouri M, Kanduri M, Isaksson A, Göransson H, Smedby KE, et al. Distinct Gene Expression Profiles in Subsets of Chronic Lymphocytic Leukemia Expressing Stereotyped IGHV4-34 B Cell Receptors. *Haematologica* (2010) 95(12):2072–9. doi: 10.3324/haematol.2010.028639
66. Papakonstantinou N, Ntoufa S, Chartomatsidou E, Papadopoulos G, Hatzigeorgiou A, Anagnostopoulos A, et al. Differential microRNA Profiles and Their Functional Implications in Different Immunogenetic Subsets of Chronic Lymphocytic Leukemia. *Mol Med* (2013) 19(1):115–23. doi: 10.2119/molmed.2013.00005
67. Ntoufa S, Vardi A, Papakonstantinou N, Anagnostopoulos A, Aleporou-Marinou V, Belessi C, et al. Distinct Innate Immunity Pathways to Activation and Tolerance in Subgroups of Chronic Lymphocytic Leukemia With Distinct Immunoglobulin Receptors. *Mol Med* (2012) 18(9):1281–91. doi: 10.2119/molmed.2011.00480
68. Gounari M, Ntoufa S, Apollonio B, Papakonstantinou N, Ponzoni M, Chu CC, et al. Excessive Antigen Reactivity may Underlie the Clinical Aggressiveness of Chronic Lymphocytic Leukemia Stereotyped Subset 8. *Blood* (2015) 125 (23):3580–7. doi: 10.1182/blood-2014-09-603217
69. Ntoufa S, Papakonstantinou N, Apollonio B, Gounari M, Galigalidou C, Fonte E, et al. B Cell Anergy Modulated by TLR1/2 and the miR-17∼92 Cluster Underlies the Indolent Clinical Course of Chronic Lymphocytic Leukemia Stereotyped Subset 4. *J Immunol* (2016) 196(10):4410–7. doi: 10.4049/jimmunol.1502297
70. Minici C, Gounari M, Übelhart R, Scarfò L, Dühren-von Minden M, Schneider D, et al. Distinct Homotypic B-Cell Receptor Interactions Shape the Outcome of Chronic Lymphocytic Leukaemia. *Nat Commun* (2017) 8:15746. doi: 10.1038/ncomms15746
71. Gemenetzi K, Psomopoulos F, Carriles AA, Gounari M, Minici C, Plevova K, et al. Higher-Order Immunoglobulin Repertoire Restrictions in CLL: The Illustrative Case of Stereotyped Subsets 2 and 169. *Blood* (2021) 137(14):1895–904. doi: 10.1182/blood.2020005216
72. Vlachonikola E, Stamatopoulos K, Chatzidimitriou A. T Cells in Chronic Lymphocytic Leukemia: A Two-Edged Sword. *Front Immunol* (2021) 11:612244. doi: 10.3389/fimmu.2020.612244
73. Palma M, Gentilcore G, Heimersson K, Mozaffari F, Näsman-Glaser B, Young E, et al. T Cells in Chronic Lymphocytic Leukemia Display Dysregulated Expression of Immune Checkpoints and Activation Markers. *Haematologica* (2017) 102(3):562–72. doi: 10.3324/haematol.2016.151100
74. Scrivener S, Goddard RV, Kaminski ER, Prentice AG. Abnormal T-Cell Function in B-Cell Chronic Lymphocytic Leukaemia. *Leuk Lymphoma* (2003) 44(3):383–9. doi: 10.1080/1042819021000029993
75. Brusa D, Serra S, Coscia M, Rossi D, D'Arena G, Laurenti L, et al. The PD-1/PD-L1 Axis Contributes to T-Cell Dysfunction in Chronic Lymphocytic Leukemia. *Haematologica* (2013) 98(6):953. doi: 10.3324/haematol.2012.077537
76. Roessner PM, Seiffert M. T-Cells in Chronic Lymphocytic Leukemia: Guardians or Drivers of Disease? *Leukemia* (2020) 34(8):2012–24. doi: 10.1038/s41375-020-0873-2
77. Elston L, Fegan C, Hills R, Hashimdeen SS, Walsby E, Henley P, et al. Increased Frequency of CD4+ PD-1+ HLA-DR+ T Cells Is Associated With Disease Progression in CLL. *Br J Haematol* (2020) 188(6):872–80. doi: 10.1111/bjh.16260
78. Riches JC, Davies JK, McClanahan F, Fatah R, Iqbal S, Agrawal S, et al. T Cells From CLLpatients Exhibit Features of T-Cell Exhaustion But Retain Capacity for Cytokine Production. *Blood* (2013) 121(9):1612–21. doi: 10.1182/blood-2012-09-457531
79. Os A, Bürgler S, Ribes AP, Funderud A, Wang D, Thompson KM, et al. Chronic Lymphocytic Leukemia Cells Are Activated and Proliferate in Response to Specific T Helper Cells. *Cell Rep* (2013) 4(3):566–77. doi: 10.1016/j.celrep.2013.07.011
80. Bagnara D, Kaufman MS, Calissano C, Marsilio S, Patten PEM, Simone R, et al. A Novel Adoptive Transfer Model of Chronic Lymphocytic Leukemia Suggests a Key Role for T Lymphocytes in the Disease. *Blood J Am Soc Hematol* (2011) 117(20):5463–72. doi: 10.1182/blood-2010-12-324210
81. Hanna BS, Roessner PM, Scheffold A, Jebaraj BMC, Demerdash Y, Öztürk S, et al. Pi3kδ Inhibition Modulates Regulatory and Effector T-Cell Differentiation and Function in Chronic Lymphocytic Leukemia. *Leukemia* (2019) 33(6):1427–38. doi: 10.1038/s41375-018-0318-3
82. Ramsay AG, Clear AJ, Fatah R, Gribben JG. Multiple Inhibitory Ligands Induce Impaired T-Cell Immunologic Synapse Function in Chronic Lymphocytic Leukemia That Can Be Blocked With Lenalidomide: Establishing a Reversible Immune Evasion Mechanism in Human Cancer. *Blood* (2012) 120(7):1412–21. doi: 10.1182/blood-2012-02-411678
83. Vardi A, Vlachonikola E, Karypidou M, Stalika E, Bikos V, Gemenetzi K, et al. Restrictions in the T-Cell Repertoire of Chronic Lymphocytic Leukemia: High-Throughput Immunoprofiling Supports Selection by Shared Antigenic Elements. *Leukemia* (2017) 31(7):1555–61. doi: 10.1038/leu.2016.362
84. Rawstron AC. Monoclonal B-Cell Lymphocytosis. *Hematol Am Soc Hematol Educ Progr* (2009) 430–9. doi: 10.1182/asheducation-2009.1.430
85. Lanasa MC, Weinberg JB. Immunologic Aspects of Monoclonal B-Cell Lymphocytosis. *Immunol Res* (2011) 49(1):269–80. doi: 10.1007/s12026-010-8188-4
86. Yin Q, Sivina M, Robins H, Yusko E, Vignali M, O'Brien S, et al. Ibrutinib Therapy Increases T Cell Repertoire Diversity in Patients With Chronic Lymphocytic Leukemia. *J Immunol* (2017) 198(4):1740–7. doi: 10.4049/jimmunol.1601190
87. Blanco G, Vardi A, Puiggros A, Gómez-Llonín A, Muro M, Rodríguez-Rivera M, et al. Restricted T Cell Receptor Repertoire in CLL-Like Monoclonal B Cell Lymphocytosis and Early Stage CLL. *Oncoimmunology* (2018) 7(6):e1432328. doi: 10.1080/2162402X.2018.1432328
88. Vardi A, Agathangelidis A, Stalika E, Karypidou M, Siorenta A, Anagnostopoulos A, et al. Antigen Selection Shapes the T-Cell Repertoire in Chronic Lymphocytic Leukemia. *Clin Cancer Res* (2016) 22(1):167–74. doi: 10.1158/1078-0432.CCR-14-3017
89. Te Raa GD, Tonino SH, Remmerswaal EBM, Van Houte AJ, Koene HR, Van Oers MH, et al. Chronic Lymphocytic Leukemia Specific T-Cell Subset Alterations Are Clone-Size Dependent and Not Present in Monoclonal B Lymphocytosis. *Leuk Lymphoma* (2012) 53(11):2321–5. doi: 10.3109/10428194.2012.698277
90. Ghia P, Prato G, Stella S, Scielzo C, Geuna M, Caligaris-Cappio F. Age-Dependent Accumulation of Monoclonal CD4+ CD8+ Double Positive T Lymphocytes in the Peripheral Blood of the Elderly. *Br J Haematol* (2007) 139 (5):780–90. doi: 10.1111/j.1365-2141.2007.06867.x
91. Purroy N, Wu CJ. Coevolution of Leukemia and Host Immune Cells in Chronic Lymphocytic Leukemia. *Cold Spring Harb Perspect Med* (2017) 7(4): a026740. doi: 10.1101/cshperspect.a026740
92. Agathangelidis A, Galigalidou C, Scarfò L, Moysiadis T, Rovida A, Vlachonikola E, et al. High-Throughput Analysis of the T-Cell Receptor Gene Repertoire in Low-Count Monoclonal B-Cell Lymphocytosis Reveals a Distinct Profile From Chronic Lymphocytic Leukemia. *Haematologica* (2020) 105(10):515–8. doi: 10.3324/haematol.2019.221275
93. Parikh SA, Leis JF, Chaffee KG, Call TG, Hanson CA, Ding W, et al. Hypogammaglobulinemia in Newly Diagnosed Chronic Lymphocytic Leukemia: Natural History, Clinical Correlates, and Outcomes. *Cancer* (2015) 121(17):2883–91. doi: 10.1002/cncr.29438
94. Blanco E, Pérez-Andrés M, Arriba-Méndez S, Contreras-Sanfeliciano T, Criado I, Pelak O, et al. Age-Associated Distribution of Normal B-Cell and Plasma Cell Subsets in Peripheral Blood. *J Allergy Clin Immunol* (2018) 141 (6):2208–19. doi: 10.1016/j.jaci.2018.02.017
95. Criado I, Muñoz-Criado S, Rodríguez-Caballero A, Nieto WG, Romero A, Fernández-Navarro P, et al. Host Virus and Pneumococcus-Specific Immune Responses in High-Count Monoclonal B-Cell Lymphocytosis and Chronic

Lymphocytic Leukemia: Implications for Disease Progression. *Haematologica* (2017) 102(7):1238–46. doi: 10.3324/haematol.2016.159012

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 23 June 2022
doi: 10.3389/fonc.2022.894419

Characterizing Features of Human Circulating B Cells Carrying CLL-Like Stereotyped Immunoglobulin Rearrangements

Davide Bagnara[1†], Monica Colombo[2†], Daniele Reverberi[2], Serena Matis[2], Rosanna Massara[2], Niccolò Cardente[1], Gianluca Ubezio[3], Vanessa Agostini[3], Luca Agnelli[4,5], Antonino Neri[6], Martina Cardillo[7], Stefano Vergani[8], Fabio Ghiotto[1,2], Andrea Nicola Mazzarello[1], Fortunato Morabito[9,10], Giovanna Cutrona[2], Manlio Ferrarini[1] and Franco Fais[1,2]*

[1] Department of Experimental Medicine, University of Genoa, Genoa, Italy, [2] Molecular Pathology, IRCCS Ospedale Policlinico San Martino, Genoa, Italy, [3] Transfusion Centre, IRCCS Ospedale Policlinico San Martino, Genoa, Italy, [4] Department of Pathology, IRCCS Istituto Nazionale dei Tumori, Milan, Italy, [5] Department of Oncology and Hemato-oncology, University of Milan, Milan, Italy, [6] Scientific Directorate, Azienza Unità Sanitaria Locale (USL)-IRCCS di Reggio Emilia, Reggio Emilia, Italy, [7] Karches Center for Oncology Research, The Feinstein Institute for Medical Research, Manhasset, NY, United States, [8] Developmental Immunology Unit, Division of Molecular Hematology, Department of Laboratory Medicine, Lund Stem Cell Center, Lund University, Lund, Sweden, [9] Hematology and Bone Marrow Transplant Unit, Hemato-Oncology Department, Augusta Victoria Hospital, East Jerusalem, Israel, [10] Biothecnology Research Unit, AO of Cosenza, Cosenza, Italy

OPEN ACCESS

Edited by:
Dimitar G. Efremov,
International Centre for Genetic Engineering and Biotechnology, Italy

Reviewed by:
Anton W. Langerak,
Erasmus Medical Center, Netherlands
Emanuela Ghia,
University of California, San Diego, United States

***Correspondence:**
Davide Bagnara
davide.bagnara@edu.unige.it

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: 11 March 2022
Accepted: 18 May 2022
Published: 23 June 2022

Citation:
Bagnara D, Colombo M, Reverberi D, Matis S, Massara R, Cardente N, Ubezio G, Agostini V, Agnelli L, Neri A, Cardillo M, Vergani S, Ghiotto F, Mazzarello AN, Morabito F, Cutrona G, Ferrarini M and Fais F (2022) Characterizing Features of Human Circulating B Cells Carrying CLL-Like Stereotyped Immunoglobulin Rearrangements. Front. Oncol. 12:894419. doi: 10.3389/fonc.2022.894419

Chronic Lymphocytic Leukemia (CLL) is characterized by the accumulation of monoclonal $CD5^+$ B cells with low surface immunoglobulins (IG). About 40% of CLL clones utilize quasi-identical B cell receptors, defined as stereotyped BCR. CLL-like stereotyped-IG rearrangements are present in normal B cells as a part of the public IG repertoire. In this study, we collected details on the representation and features of CLL-like stereotyped-IG in the IGH repertoire of B-cell subpopulations purified from the peripheral blood of nine healthy donors. The B-cell subpopulations were also fractioned according to the expression of surface CD5 molecules and IG light chain, IGκ and IGλ. IG rearrangements, obtained by high throughput sequencing, were scanned for the presence of CLL-like stereotyped-IG. CLL-like stereotyped-IG did not accumulate preferentially in the $CD5^+$ B cells, nor in specific B-cell subpopulations or the $CD5^+$ cell fraction thereof, and their distribution was not restricted to a single IG light chain type. CLL-like stereotyped-IG shared with the corresponding CLL stereotype rearrangements the IGHV mutational status. Instead, for other features such as IGHV genes and frequency, CLL stereotyped-IGs presented a CLL-like subset specific behavior which could, or could not, be consistent with CLL stereotyped-IGs. Therefore, as opposed to the immuno-phenotype, the features of the CLL stereotyped-IG repertoire suggest a CLL stereotyped subset-specific ontogeny. Overall, these findings suggest that the immune-genotype can provide essential details in tracking and defining the CLL cell of origin.

Keywords: chronic lymphocytic leukemia (CLL), immunoglobulin repertoire, B-cells, CD5, IGHV somatic mutations, Ig light chain, CLL stereotyped BCR

INTRODUCTION

Chronic Lymphocytic Leukemia (CLL) is characterized by the accumulation of monoclonal B lymphocytes expressing CD5, CD23, and low surface immunoglobulin in blood, bone marrow, and lymphoid tissues (1, 2). Analyses of many CLL clones demonstrated that the IG gene rearrangements encoding the CLL B cell receptor (BCR) exhibit a striking skewed use of IGHV genes resulting in an IG repertoire different from that of normal B lymphocytes (3–5). Moreover, despite the enormous diversity potentially generated by the recombination of IGHV-IGHD-IGHJ genes, up to 40% of the CLL clones (6, 7) exhibit highly similar stereotyped BCR, which has led to the categorization of the CLL clones with stereotyped BCR into subsets based on their similarities. Although several hundred CLL subsets have been identified, those most frequently encountered, defined as "major subsets," are limited in numbers. Stereotyped BCRs are determined based on the VH CDR3 features of at least 50% of amino acids identity and 70% of amino acid motif similarities, identical VH CDR3 length and location of a shared pattern(s) among sequences of the same group (6, 8), and the use of IGHV genes belonging to the same phylogenetic clan. IGHV clans are IGHV family genes with structural similarities (9). Conceptually, the stereotyped gene rearrangements should be considered part of the public IG repertoire because different individuals share them (10). In addition, CLL clones of the same stereotyped subsets show IG light chain restrictions (e.g., #1, #2, #4, #6, #8, #64b, and #99) (6, 11–13) and IGKV-IGKJ/IGLV-IGLJ gene rearrangements presenting stereotypy features similar to those of IGHV rearrangements with limitations in IG light chain gene usage and VL CDR3 composition.

The above observations support the notion of a role for BCR stimulation in CLL ontogeny (14–16); moreover, the results of therapies with inhibitors of BCR-associated kinases suggest that stimulation *via* BCR may be critical for the survival/proliferation of CLL cells in full-blown leukemia (17).

Previous studies have identified IGHV-IGHD-IGHJ rearrangements sharing features with that characteristic of CLL subsets in splenic and circulating B cells from normal subjects (18–20). These rearrangements, from now onward defined as CLL-like stereotyped-IG or CLS-IG, can be observed in different B cell subpopulations, even though they accumulate in the $CD5^+$ B-cells (20, 21).

This study used high-throughput sequencing technology on peripheral blood B cell subsets to elucidate CLS-IG's features and cellular distribution. For this purpose, the cells, separated into defined subsets, were also fractionated according to CD5 or IG light chain expression. The data obtained provide a new perspective for interpreting the origin of CLL cell repertoire and possibly for disease ontogeny.

MATERIALS AND METHODS

Samples

Peripheral blood cells were obtained from the leukopak of anonymous blood donors (nine donors aged 55 to 64 years old) at the San Martino Hospital Blood Center presentation. Each leukopak is derived from ~500ml of blood. B cells were enriched with RosetteSep Human B Cell Enrichment Cocktail (Stemcell Technologies, Vancouver, Canada), obtaining on average $42x10^6$ B cells per donor ($24x10^6$ to $60x10^6$) (**Supplementary Table S1**).

Isolation of PB B Cell Subpopulations and Fractionation of $CD5^+$, $CD5^-$ B Cells, and IGκ and IGλ

B cells enriched cell fractions were stained with the following combination of mAbs: anti-IgD Alexa Vio770 (BioLegend, San Diego, CA, USA); anti-IgM PerCP_Cy5.5, anti-CD27 PE-CF594, anti-CD38 PE-Cy7, anti-CD24 Alexa Fluor 647 and anti-CD5 BV 421, anti-IGκ FITC, anti-IGλ PE, anti-IgA VioGreen (BD). B cell subsets were isolated by FACS sorting (FACSAria, Becton Dickinson, Franklin Lakes, NJ, USA) after depleting IgA^+ and dead cells with a two-step sorting approach: 1) a four-way pre-sort with yield setting was used to separate enriched B cells into $IGκ^+/CD5^+$, $IGκ^+/CD5^-$, $IGλ^+/CD5^+$, and $IGλ^+/CD5^-$ B cells; 2) each of the above cell fractions were then sorted into six main B cell subpopulations (after excluding $CD38^{high}CD24^-$ plasmablasts): $CD24^{high}CD38^{high}$ transitional (TR), $IgD^{high}IgM^+CD38^-CD27^-$ naive (N), IgD^{low} $IgM^+CD38^-CD27^+$ marginal zone-like (MZ), $IgM^+IgD^-CD38^-CD27^+$ IgM-only memory (MO), $IgM^-IgD^-CD38^-CD27^+$ switch-memory (SM), and $IgM^-IgD^-CD38^-CD27^-$ double negative (DN) B cells. See also **Figure 1** and **Supplementary Figure S1** for details.

Library Preparation and Sequencing

IGH sequencing and analysis were performed as previously reported in detail (22). Briefly, the library was prepared from mRNA with a multiplex approach with IGHV-specific forward primers on the leader sequence and reverse primers on the constant region. The primer set was kindly provided by TIB Molbiol srl (Genoa, Italy). The amplicons obtained included the entire IGHV-IGHD-IGHJ gene and enough constant region to

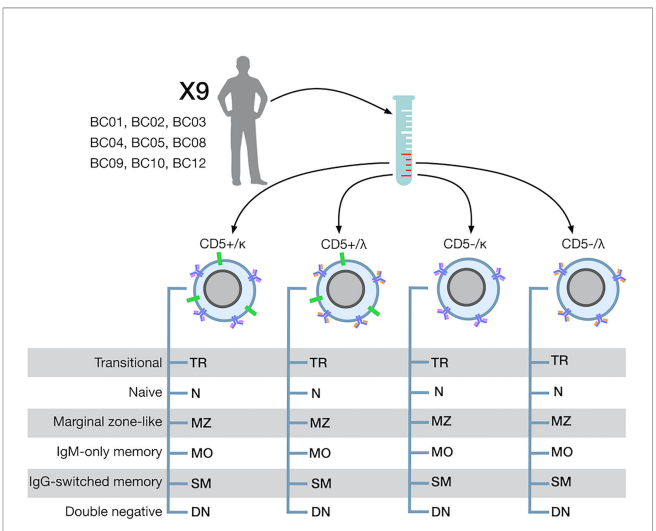


FIGURE 1 | B-cell separation strategy.

assign the isotype; UMIs (14 to 16 nucleotides) were inserted during ds-cDNA synthesis. The libraries were indexed with Illumina Nextera XT V2 kit (Illumina, San Diego, CA, USA) and sequenced on Illumina MiSeq (MiSeq V3, 2x300 kit, Illumina). DNA sequences were deposited on the NCBI Sequence Read Archive (SRA) portal with BioProject ID: PRJNA807871.

Bioinformatics Analysis

Raw reads were processed with a custom-built workflow using pRESTO (23) as previously described (20); processed sequences were then annotated by IMGT/HighV-QUEST (24). Only productive rearrangements derived from the consensus of two or more raw reads without N nucleotide passed the quality filter. ChangeO (25) was used to define and annotate clonotypes as sequences with identical amino acid VH CDR3 sequence using the same IGHV gene and IGHJ gene.

CLL Subsets Assignment

To identify CLL-like stereotyped sequences (CLS-IG), we first selected the sequences in our database consistent with the core features (IGHV clan, IGHV mutational status, and VH CDR3 length) (26) of each of the 19 major CLL stereotyped subsets and then submitted them to ARResT/AssignSubsets (26) for the assignment. Sequences assigned to CLL stereotyped subsets with a confidence "average" or higher were considered CLS-IG. Sequences not assigned to CLL stereotyped subsets were defined as non-CLS-IG. The entire analysis was then performed at the clonal level, i.e., for each clonotype, a single representative sequence was considered.

To identify VH CDR3 aa sequences consistent with a CLL stereotyped subset but with the reverse (r) mutational status compared to the core feature of CLL stereotypies (rCLS-IG), we re-submitted to aRResT the IGHV-IGHD-IGHJ rearrangements sequences in which the original IGHV gene region was replaced with one with the opposite mutational status, generating an *in silico* chimeric sequence.

Statistics

Statistical analyses were performed in R. Paired Wilcoxon tests or binomial test with Bonferroni correction applied to assess differences in CLS-IG or VH gene frequency (*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$). The paired Wilcoxon test was calculated only if three or more donors presented data for both points. CLS-IGs' frequency was calculated only when the number of CLS-IG sequences in the group was two or more. Frequencies obtained from only one CLS-IG sequence within a group were not considered informative and therefore not plotted unless specified.

RESULTS

Identification of CLL-Like Stereotyped-IGs in B-Cell Subpopulations

Peripheral blood B cell samples from each of nine donors were separated into 24 phenotypically distinct cell fractions. First, B cells were divided into four fractions according to the presence or absence of surface CD5 and surface IG light chain expression (IGκ or IGλ). From each of the four B cell fractions, six different B-cell subpopulations (B-subset) were isolated: transitional (TR), naive (N), marginal zone-like (MZ), IgM only memory (MO), IgG switched memory (SM), and double-negative (DN); see **Figure 1**, Methods and **Supplementary Figure S1** for phenotypes and details. After quality filtering, a total of 2,679,224 productive, unique IGHV-IGHD-IGHJ sequences were obtained from 8,043,500 sorted B cells. Curated sequences were subsequently clustered into 2,184,656 clonotypes (detailed in **Supplementary Table S2**), of which 1754 (0.08%) were assigned to one of the major CLL stereotyped subsets (CLL-subset) with ARResT/AssignSubsets (26) (as detailed in Methods) and defined as CLL-like stereotypes-IG (CLS-IG) (detailed in **Supplementary Table S3**). Clonal families were used as references for the entire analysis.

Correlation of the Higher CLL-Like Stereotyped-IG Representation in $CD5^+$ B Cells With an Asymmetrical Distribution of U and M IGHV Rearrangements in $CD5^+$ and $CD5^-$ B Cells

First, sequences from all $CD5^+$ or $CD5^-$ subpopulations respectively were pooled and analyzed for the presence of CSL-IG to see whether the CSL-IG were predominant in the $CD5^+$ cell fraction, and we found that the proportion of CLS-IG was significantly higher in $CD5^+$ than in $CD5^-$ B cells (**Figure 2A**). However, when the IGHV gene rearrangements were separated into mutated (≥2% IGHV gene mutations, M-IG) and unmutated (<2% mutations, U-IG), there was no difference in the frequency of CLS-IG between $CD5^+$ and $CD5^-$ B cells within a single mutational status group (**Figure 2B**). Furthermore, U-IG clonotypes had more CLS-IG than M-IG clonotypes in both $CD5^+$ and the $CD5^-$ populations. When looking at the average IGHV mutation frequency, the $CD5^+$ clonotypes appear to be enriched in U-IG, contrary to $CD5^-$ clonotypes enriched in M-IG (**Figure 2C**).

Presence of CLL-Like Stereotyped-IG in B-Cell Subpopulations

CLS-IG were found in all B-subsets, although in different proportions (**Figure 3A**); i.e., N and TR B cells had a significantly higher CLS-IG representation (0.09%) than MZ (0.05%) and SM (0.03%) B cells. N and TR B cells also had the highest U-IG sequences (**Figure 3C**). Further fractionation of each B-subset into $CD5^+$ and $CD5^-$ cells did not show a significant predominance of CLS-IG in any of the $CD5^+$ cell fractions (**Figure 3B**). In most cases, SM, DN, and MO B-subsets presented none or just one CLS-IG per donor. It must be noted that these B-subsets have the lower frequency of $CD5^+$ cells and, therefore, the least IGH sequences (**Supplementary Table S2**).

Different Frequencies of CLL-Like Stereotyped-IG Subsets in Normal B Cells and CLL Clones

Overall, the median frequency of CLS-IGs clonotypes of individual CLL subsets was within the range of 0.019% to

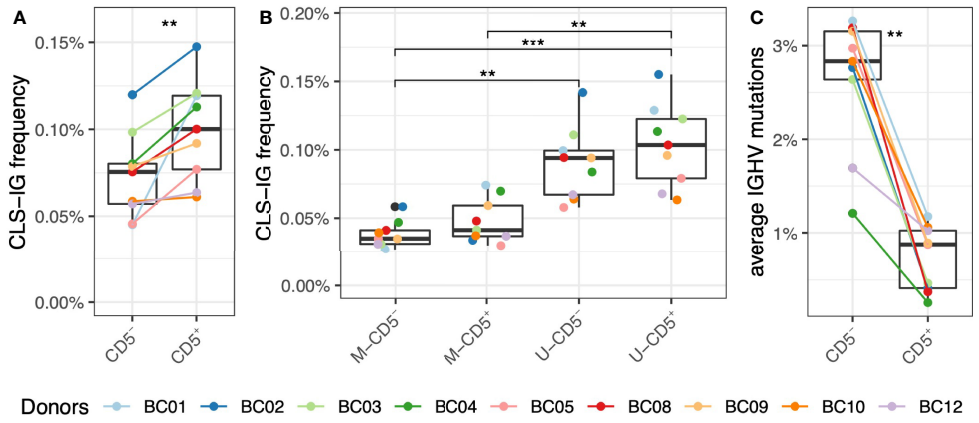


FIGURE 2 | Frequency of CLS-IG in $CD5^+$ and $CD5^-$ B cells. **(A)** CLS-IG frequency in $CD5^+$ and $CD5^-$ B cells analyzed in bulk. **(B)** Frequency of CLS-IG in the M and U IGHV rearrangements of $CD5^+$ and $CD5^-$ B cells analyzed separately. **(C)** IGHV genes mutation distribution in $CD5^+$ and $CD5^-$ B cells. Paired Wilcoxon test was performed, only significant statistics are shown (**p ≤ 0.01, ***p ≤ 0.001).

undetectable in circulating B cells (**Figure 4A**). Subset #5 was the most represented (403 clonotypes, 0.019%), followed by subset #2 (307 clonotypes, 0.017%), subset #64B (294 clonotypes, 0.014%), subset #3 (220 clonotypes, 0.008%), subset #14 (201 clonotypes, 0.006%), and subset #1 (103 clonotypes, 0.005%), whereas other CLL subsets were represented at a lower level. CLS-IG frequency was comparable between $CD5^+$ and $CD5^-$ B cells for every CLL subset except subset #5 where $CD5^+$ presented statistically more CLS-IGs than $CD5^-$, and subset #14, where $CD5^+$ presented statistically fewer CLS-IGs than $CD5^+$ (**Figure 4B**). For CLL subsets with U-IGs, we identified CLS-IG consistently only in N and TR B-subsets, whereas for CLL subsets with M-IGs (#2 and #14), CLS-IG were reproducibly detected also in MZ, MO, SM, and DN B-subsets (**Figure 4C**).

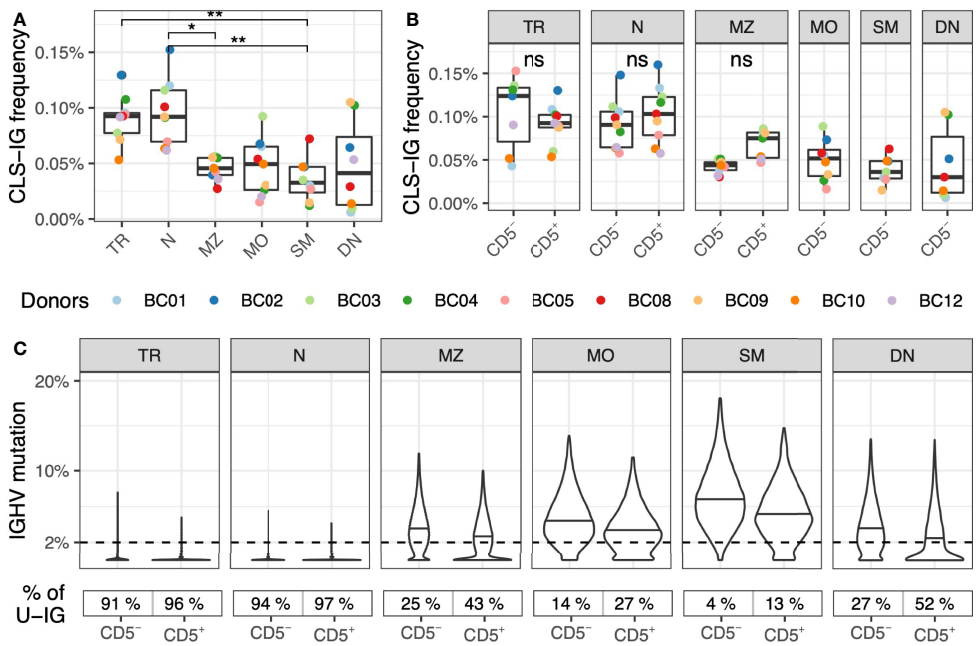


FIGURE 3 | Frequency of CLS-IG in circulating B-subset. **(A)** Frequency of CLS-IG in B-subset **(B)**. Frequency of CLS-IG in the same B-subset fractionated further for $CD5^+$ and $CD5^-$ B cells. Paired Wilcoxon test was performed, only significant statistics are shown (*p ≤ 0.05, **p ≤ 0.01; ns, not significant). CLS-IG identified in the $CD5^+$ fraction of MZ, MO, SM, and DN B cells were insufficient for statistical analysis. **(C)** IGHV genes mutation distribution in the B cell subsets-sub divided by CD5 expression. The percentage of unmutated IGs (< 2%) in each B cell population is shown at the bottom of the figure.

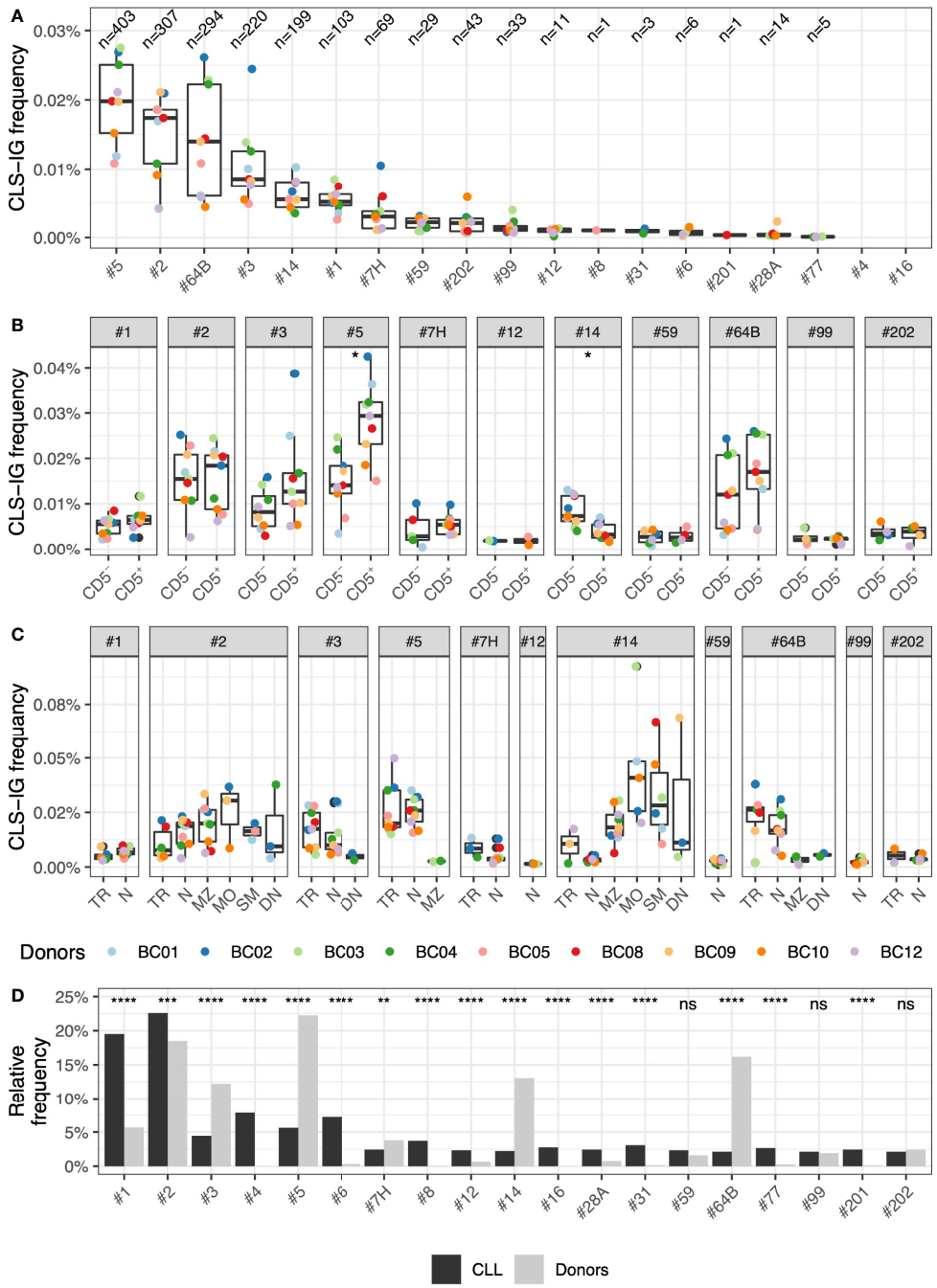


FIGURE 4 | Frequency of CLS-IG subsets. **(A)** Frequency of CLS-IGs for individual CLL subsets in healthy donors among all IGHV-IGHD-IGHJ clonotypes. **(B)** Frequency of CLS-IGs for each CLL subset in $CD5^+$ and $CD5^-$ B cells. Frequency data derived from only one donor was not considered informative and not plotted. Paired Wilcoxon test was performed, only significant statistics are shown (*$p \leq 0.05$). **(C)** Frequency of CLS-IGs for each CLL subset in B cell subsets. Frequency data derived from only one donor was not considered informative and not plotted. **(D)** Frequency of CLS-IGs for individual CLL subsets among all CLS-IGCLL subset frequency of CLS-IGs relative frequency compared with CLL stereotyped clones (6). Binomial test was performed, only significant statistics are shown (*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$; ns, not significant).

The CLS-IG frequency among the CLL subsets in our dataset was compared to CLL stereotyped-IGs reported in the study by Agathangelidis et al. (6). While the distribution of most CLL subsets did present statistically significant differences between CLL stereotyped IG and CLS-IG, only a few differed substantially (**Figure 4D**). Subset #1, which is relatively frequent in CLL, was much less represented in the CLS-IG of peripheral blood, whereas subsets #5, #14, and #64B, the most abundant

innormal B cells, were much less frequent in the CLL dataset. Other subsets were absent (#4 and #16) or rarely observed (#6, #8, #12, #31, #77, and #201).

Absence of IG Light Chain Restriction in CLL-Like Stereotyped-IGs

Many CLL clones with stereotyped receptors have restricted IG light chain usage (27, 28). This is particularly evident in leukemic clones utilizing subsets #1, #2, #4, #6, #8, #64b, and #99 (6, 12). We investigated whether CLS-IG from normal B cells showed light chain restriction. As each B-subset was fractionated starting from pre-sorted IGκ$^+$ or IGλ$^+$ B-cells and sequenced independently, we could obtain the CLS-IG distribution in both IGκ and IGλ-bearing cells. Among the CLL subsets for which clear evidence of IG light chain restriction is present in CLL clones, sufficient sequences for an informative analysis were obtained from subsets #1, #2, #64b, and #9 (**Figure 5**). There was no statistically significant difference in the IG light chain association for any CLS-IG analyzed. Only CLS-IG of subset #2 showed a trend indicating enrichment in IGκ$^+$ normal B cells; this contrasts with CLL clones of the same subset that consistently show a pairing with IGλ. The entire distribution of the IG light chain among CLS-IG stereotypes is reported in **Supplementary Figure S2**.

Similar Utilization of Mutated and Unmutated IGHV Genes by CLL Clones and CLL- Like Stereotyped-IGs

Most CLL stereotyped rearrangements are restricted to the utilization of U or M IGHV genes (6, 18). This characterizing feature is used by ARResT for CLL stereotyped subset assignment, meaning that a VH CDR3, potentially showing stereotyped features, will be excluded from a given subset if the IGHV gene, concomitantly utilized, fails to share the mutational status expected for that subset. Based on these considerations, we analyzed all the sequences from normal cells identified in this study to see whether there were restrictions to a specific mutational status for CLS-IG. To this end, we investigated whether VH CDR3 aa sequences, consistent with a CLL stereotyped subset, were present independently of the IGHV mutational status by extending the analysis to sequences with a reversed (r) mutational status compared to the core feature of the CLL subset (rCLS-IG) (see Methods). We observed that the "r" rearrangements were significantly less frequently classified as stereotyped than the native ones (**Figure 6A**). The same trend was observed for CLS-IG and rCLS-IG from individual CLL subsets when sufficient data were available for analysis (#1, #5, and #14, **Figure 6B**). The median mutation values of CLS-IG from the major CLL subsets followed a pattern like that of the corresponding CLL stereotyped-IGs (**Figure 6C**). For example, #1 was preferentially unmutated, #2 had around 2% mutations, and #14 was mutated. This pattern was similar for both the IGκ and IGλ B-cells.

Identification of "Typical" and "Non-Typical" CLL-Like Stereotyped-IGs

One of the core features defining each CLL stereotyped subset is represented by the VH clan utilized, and, in each CLL subset, the rearranged IGHV genes often present a restriction at this level (6). This restriction ranges from about half the genes within a clan to a

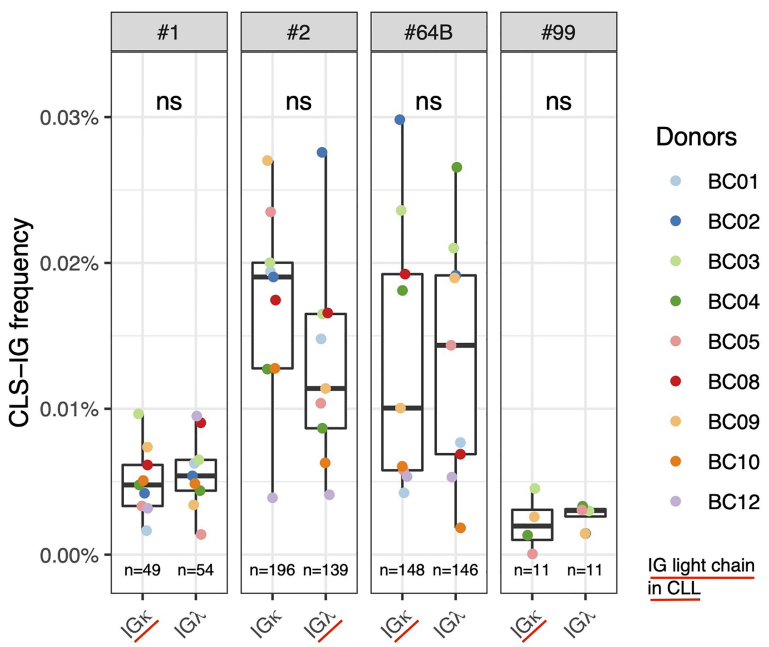


FIGURE 5 | Frequency of individual CLS-IG in sequences from normal B cells expressing either IGκ and IGλ chains in CLL stereotyped subsets with reported IG light chainIGL bias (6, 12). Paired Wilcoxon test was performed. ns, not significant

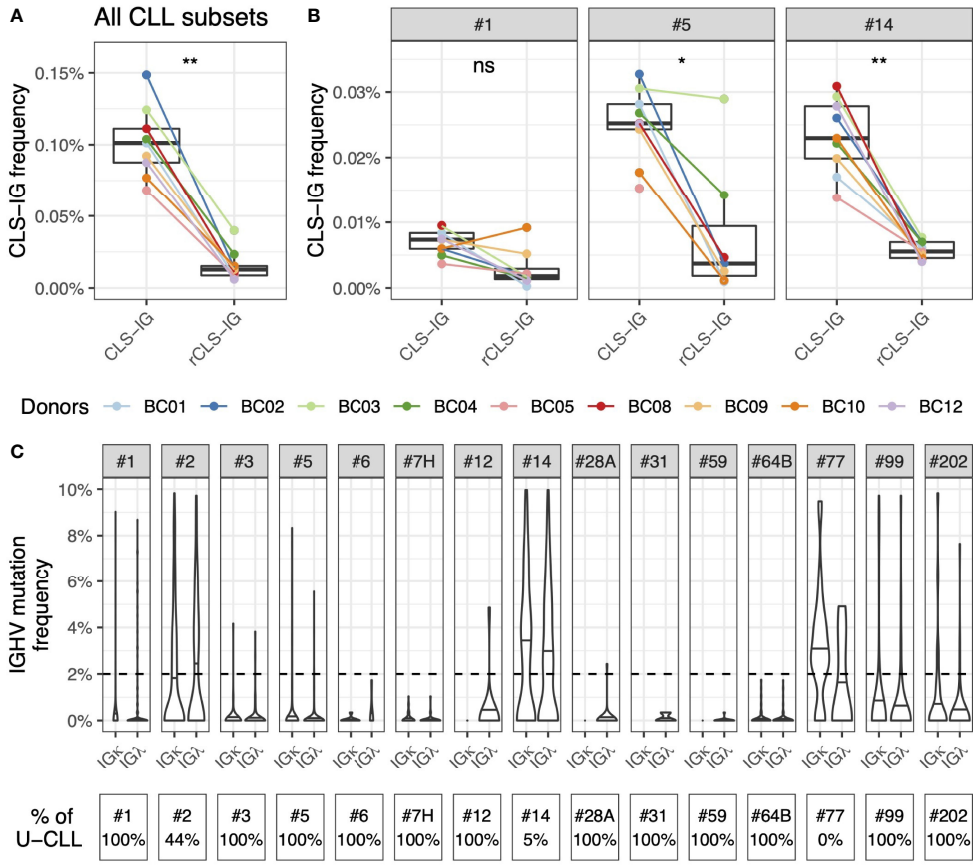


FIGURE 6 | Similar utilization of IGHV mutated and unmutated genes by CLS-IG and CLL stereotypes. **(A)** CLS-IG and rCLS-IG frequency in normal B cells. **(B)** Frequency of CLS-IG and rCLS-IG of the #1, #5 and #14 CLL subsets in normal B cells. Paired Wilcoxon test was performed, only significant statistics are shown (*p ≤ 0.05, **p ≤ 0.01; ns, not significant). **(C)** Mutation pattern of the IGHV genes utilized by CLS-IG from normal B cells subdivided for the indicated CLL subsets; the horizontal lines indicate the median mutation for each subset. The dotted line indicates the 2% mutation threshold separating U and M sequences. The percentage of CLL stereotypes with unmutated IGHV in each CLL subset is shown at the bottom of the figure.

single gene (i.e., #4 - IGHV4-34, #8 - IGHV4-39). Again, all the CLS-IG sequences of this study were analyzed and classified as "typical" if expressing an IGHV gene of those found at least once in their respective CLL counterparts (6) and "non-typical" if expressing a different IGHV gene within the VH clan. For each CLL subset, the CLS-IG VH genes usage distribution was compared to control sequences from the same dataset. Control sequences were defined as sharing the core features of the CLL subset in consideration (6) (e.g., #1 - the use of a clan I VH gene, a U-IGHV gene, and a VH CDR3 length of 13 aa). The patterns observed varied among subsets (see **Figure 7**) and can be summarized in three broad categories as follows: i) CLS-IG falling almost exclusively in the typical category as in subset #1, where even though all rearrangements but one (1/379) were typical, the IGHV gene usage followed a distribution that was not consistent with that observed in either CLL or control non-CLS-IG sequences. ii) CLS-IG distributed in typical and non-typical categories with an IGHV gene distribution like non-CLS-IG control sequences and not presenting the CLL IGHV bias. Two examples of this condition were the CLS-IG sequences from subsets #2 and #3, which utilized several IGHV genes (with a frequency similar to that of non-CLS-IG control sequences) rather than the IGHV3-21 and the IGHV1-69 genes characterizing the #2 and #3 CLL subset, respectively. iii) CLS-IG showing the same bias toward the utilization of a specific IGHV gene marking the CLL stereotyped subset. An example is subset #14, where most CLS-IG expressed the IGHV4-4 gene like the CLL stereotyped-IG and, in this respect, were different from non-CLS-IG control sequences. When separating the sequences by the IG light chain used, some fluctuation in the IGHV gene used could be observed; nevertheless, the general pattern described above remained valid for sequences from both the IGκ and IGλ group (see **Supplementary Figure S3**). **Figures 7B, C** summarize some statistics of what is shown in **Figure 7A**. In half of the informative CLL subsets (#1, #2, #5, #7H, #12, #14, and #99), we observed a significantly higher representation of typical IGHV genes in CLS-IG compared to controls (**Figure 7B**). As for the predominant IGHV gene often noticed in each single CLL subset, we observed that this bias generally was not present in CLS-IG. Indeed nine out of the 12 CLS-IG did not show evidence for this condition (**Figure 7C**).

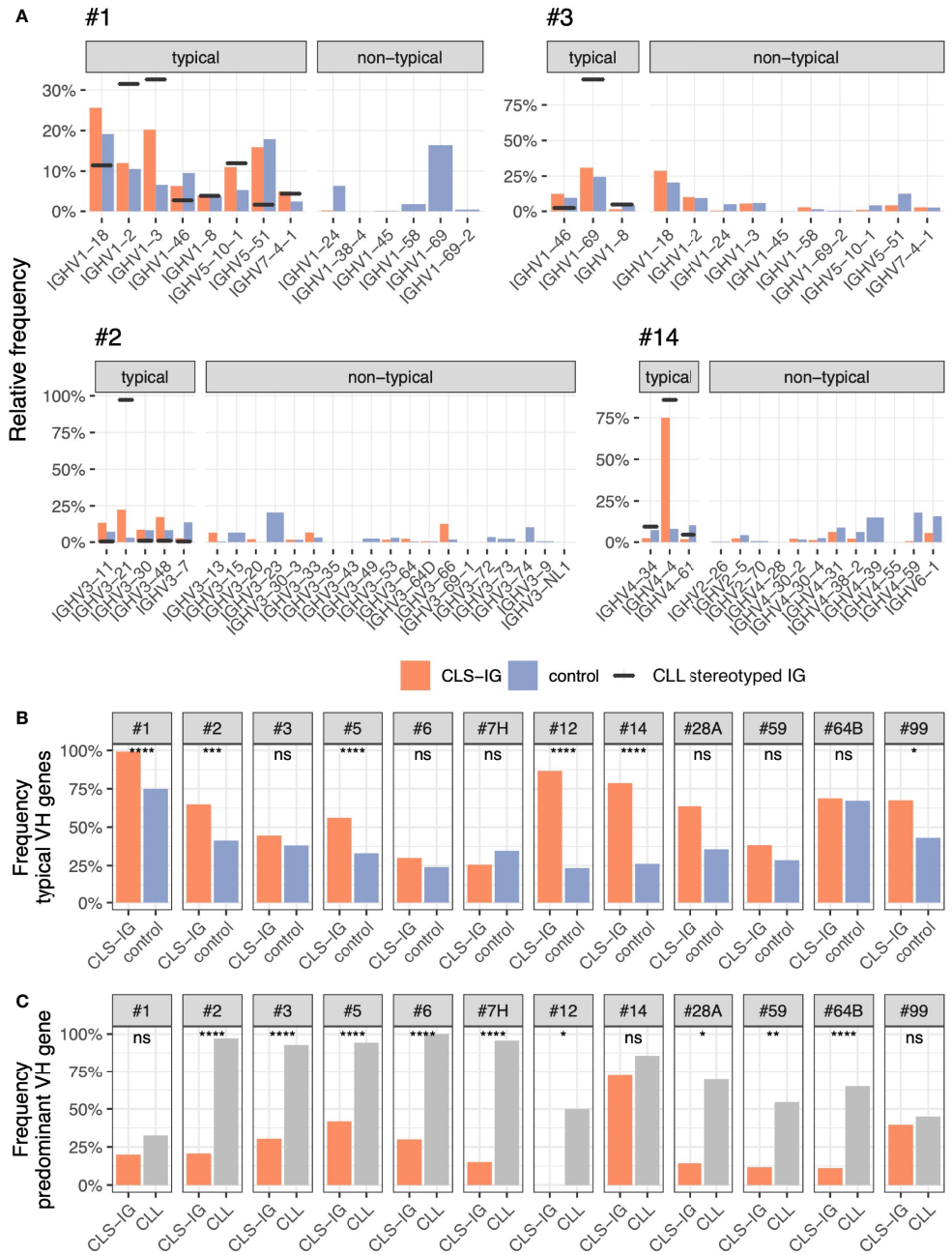


FIGURE 7 | **(A)** Frequency of IGHV gene used in typical and non-typical CLS-IG (red). Blue bars indicate non-CLS-IG control sequences (i.e., sharing the same core features as the CLL subset in consideration - IGHV clan, IGHV mutational status, and VH CDR3 length). These rearrangements were used as controls. Black horizontal lines indicate the level of IGHV representation in the reference CLL cohort. The predominant IGHV gene was identified as the most represented within a CLL subset. **(B)** Frequency of typical IGHV genes in CLS-IG clonotypes compared to control sequences. The control sequences are the same reported in panel **(A) (C)** Frequency of the predominant IGHV gene observed in CLL stereotyped IGs compared to CLS-IGs usage of the same IGHV gene. Binomial test was performed, only significant statistics are shown (*p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001; ns, not significant).

DISCUSSION

The identification of CLL clones' stereotypical IG receptors indicates that different individuals share immunoglobulin rearrangements with these specific features. This aspect reminds the existence of identical CDR3 aa sequences between individuals, which defines the so-called public repertoire (10, 29). More recently, the definition of public repertoire has been made less stringent by including rearrangements with a high degree of similarity (30). In this

respect, the public repertoire could enclose stereotypic rearrangements recognized in CLL clones.

The scope of this study was that of investigating whether, in healthy donors, CLL IG stereotyped sequences were present in the IG repertoire of any of the circulating B-subsets and to which extent. The donors analyzed did not have evidence of lymphocytosis, thus excluding the possibility of introducing biases due to some type of preleukemic stage, which may result in B-cell repertoire alterations (31). Blood samples were collected from donors with a median age of 58.8 (range 56-67), which is in the range of possible identification of clonotypes ascribable to CLL clones that became clinically evident later. Indeed, it has been reported that clonotypes attributable to the leukemic clone can be identified in the PB up to 16 years before CLL diagnosis (32). It is, however, unknown whether phenotypic changes were present in the PB B lymphocytes at this early stage. The age choice appears relevant for comparison with CLL, given that changes in B cell repertoire within single B-cell subsets may occur with advancing age and may influence the cell population in which leukemogenesis occurs. In addition, it has been reported that changes in the representation of certain CLS-IG can be observed in aging individuals (19).

Although CLL stereotyped receptors have been reported in healthy donors (18, 19, 33–35), our approach provides more details on their phenotypic features, considering, for instance, the expression of IG light chains and CD5 expression within B-subsets. The identification and characterization of CLL stereotyped IG in healthy donors could provide information on the cell(s) of origin of CLL. Attempts to identify a specific B-subset as a compartment of origin for CLL have been unsatisfying. Here we attempt to lay the ground for a better understanding of the CLL cell of origin.

Consistent with previous reports (20, 21), our data show a significantly higher presence of CLS-IG in bulk $CD5^+$ compared to bulk $CD5^-$ B cells. Nevertheless, analysis of U- and M-IGHV gene rearrangements within $CD5^+$ and $CD5^-$ B cells highlighted no enrichment of CLS-IG in $CD5^+$ compared to $CD5^-$ B cells within each mutational category (**Figure 2B**). Instead, CLS-IGs were enriched in U-IGHV genes regardless of being expressed by $CD5^+$ or $CD5^-$ B cells.

CLS-IG representation was also investigated in B-subset and their respective $CD5^+$ and $CD5^-$ fractions. The analysis showed a higher CLS-IG percentage in TR and N than in MZ and SM B-subset (**Figure 3A**), a finding in line with a previous study from our group on splenic CLS-IG carrying the IGHV-1 family genes (20). Since TR and N cells utilize almost exclusively U-IGHV rearrangements (**Figure 3C**), the higher representation of CLS-IG in these cell fractions is likely related to the predominance of U-IGHV rearrangements. The analysis of $CD5^+$ and $CD5^-$ cells from each B-subset separately showed no dominant presence of CLS-IG in any of the cell fractions studied (**Figure 3B**). However, MZ shows an evident trend where $CD5^+$ have more CLS-IG than $CD5^-$. It must be noticed that MZ presents a considerable amount of U-IGs concentrated in the CD5+ fraction (**Figure 3C**). It must be mentioned that because $CD5^+$ B cells are scarce in predominantly mutated B cells subsets, a limited number of IG rearrangements were obtained in some $CD5^+$ B-cell fractions (i.e., MO, SM, DN), impairing the possibility of statistical analysis (see **Supplementary Table S2** for details). Altogether, our data do not present evidence of a phenotypically distinct B-subset with a significant accumulation of CLS-IG. Their predominance in $CD5^+$ N and TR B cells depends upon the higher presence of U-IG rearrangements in these B-cell fractions.

The relative distribution of CLL subsets in normal B cells did not follow that typically reported for CLL (**Figure 4B**). For example, CLL subsets #4 and #8, relatively frequent among the CLL major subsets, were rarely (or never) identified. A preferential representation of individual CLS-IG in the IGκ or IGλ expressing B-subset was not observed. In contrast, specific CLL subsets (e.g., #1, #2, #4, #6, #8, #64b, and #99) show, in CLL, marked IG-light chain use restrictions (6, 12, 36). It might be worth mentioning that CLS-IG #2 showed a trend indicating a higher representation in $IG\kappa^+$ normal B-cells. This would be compatible with the observation that, in some instances, CLL clones belonging to subset #2 show functional rearrangements of the IGκ light chain inactivated with a Kappa Deleting Element (37). Thus IG-receptor editing may operate in this apparently large proportion of CLS-IG subset #2/$IG\kappa^+$ B cells. These observations indicate no major structural constraints contributing to certain IGH-IG light chain pairing in the CLS-IG of a normal B-cell repertoire. This is in keeping with what has been reported about the possibility of CLL subset #2 IGHV rearrangement to pair with "non-native" (e.g., other than IGLV3.21 encoded IGλ light chain) IG light chains (38).

The IGHV genes used by CLS-IG rearrangements within each CLL subset showed a heterogeneous pattern. Their overall utilization was not always as restricted as CLL stereotyped IG. For instance, the IGHV1-69 gene was virtually absent in CLS-IG from CLL subset #1, even though this gene is one of the IGHV1 genes highly represented in control IGs. CLS-IG in CLL subsets #1 and #3 exemplify the absence of IGHV selection, whereas CLS-IG in CLL subset #14 acts closely to what is observed in CLL with a highly prevalent representation of the IGHV4-4 gene. This indicates that specific VH CDR3 sequences may have a non-random association with IGHV genes within a VH clan, possibly related to restrictions during IGHV-IGHD-IGHJ rearrangements and/or positive and negative selection in ontogenesis or the course of early antigenic challenges. The absence of IGκ and IGλ restriction for each of these CLS-IG subsets points out, at this stage, negligible participation of the IG light chains. Altogether, the leukemogenic process likely involves a further selection of IGHV genes and the IG light chains in a CLL stereotype-specific manner.

The mutational status of CLS-IG deserves particular comment. When the CLS-IG analysis was extended by removing the IGHV mutational status as a prerequisite for classification (see Methods), the representation of CLS-IG mainly followed the mutational status characterizing the original CLL clones. For instance, subset #1 (always unmutated in CLL) was identified predominantly in the U-IG repertoire of normal donors and the majority of subset #14 CLS-IG (mutated in CLL) were recognized in the M-IG repertoire. Likewise, subset #2 CLS-IG were found in the U and M CLS-IG repertoire, as

observed in the CLL cohorts. The observation that the mutational status of CLS-IG parallels that of CLL IG stereotypes suggests that specific IG rearrangements could influence clonal function, e.g., by limiting the generation of a post-germinal center progeny (CLL subset #1) or by determining an accumulation of memory B cells (CLL subset #14).

Thus, single CLS-IG in recirculating B cell has only marginally superimposable features compared to those encountered in CLL clones and each one appears to have its characteristics. For instance, subsets #4 and #8 are substantially absent in the CLG-IG repertoire; subset #1 CLS-IG shows the utilization of IGHV genes closer to CLL stereotypes, whereas that of CLS-IG, subset #3 is more random. It can be presumed that the trajectory determining the emergence of CLL clones is very heterogeneous.

The above data demonstrate that CLS-IG detected in peripheral B cells from donors with no evidence of peripheral lymphocytosis have different features than those typically identified in leukemic clones, suggesting the shaping of CLL BCR repertoire and the emergence of the leukemic clones is dictated by numerous selecting factors. A recent study (32) showed that skewing of the B-cell repertoire is observable in some clusters before the clinical presentation of CLL. Thus, subjects in a pre-leukemic phase or predisposed to developing CLL are likely to have a different condition than the donors analyzed here. The observations that CLS-IGs are not enriched in PB derived B cells with a defined phenotype can be interpreted in different ways: 1) CLL cells originate from non-circulating B cells residing in solid lymphoid tissue, 2) CLL cells originate from B cells with a different phenotype than the ones explored in this study, 3) CLL cells originate from B cells without a defined immunophenotype. The above interpretations may not be mutually exclusive. In addition, it is possible that the leukemogenic process can be accompanied by immunophenotypic changes that encompass an elevated expression of CD5 typically observed in CLL and MBL cells and presumably the pre-monoclonal B-cell lymphocytosis described by Kolijn, P. et al. (32)

In this context, B-cell immune-genotype appears to be a relevant factor in the quest to identify the CLL clones' cell of origin, adding additional elements useful for understanding CLL emergence routes.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/genbank/, PRJNA807871.

AUTHOR CONTRIBUTIONS

DB designed the study, designed the library preparation protocol, analyzed data and wrote the manuscript. MoC performed FACS Sorting, prepared the library, sequenced and analyzed the data and wrote the manuscript. DR designed and performed FACS Sorting. SM and RM processed biological samples and prepared the library. GU and VA collected samples. LA, SV, and AM performed data and statistical analysis. AN, FM, and GC supervised research. MaC and FG prepared the figures. MF and FF designed the study and wrote the paper. All authors read and approved the final manuscript.

FUNDING

This work was supported by: Associazione Italiana Ricerca sul Cancro (AIRC) ID.15426 (to FF); AIRC and Fondazione CaRiCal co-financed Multi-Unit Regional Grant 2014 n.16695 (to FM); Italian Ministry of Health 5 × 1000 funds [2014 ([to GC), 2015 (to FF), 2016 (to FF and GC), Ricerca Corrente 2016 (to FF and GC); Gilead fellowship program 2016 (MC) and 2017 (GC); Associazione Italiana Ricerca sul Cancro (AIRC) Grant 5 x 1000 n. 9980 (to FM and AN); AIRC IG 24365 to AN; Associazione Italiana Leucemie, Cosenza (to FF). The primers for the determination of IGHV repertoire were kindly provided by TIB Molbiol srl (Genoa, Italy). AM received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement No. 101023721.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.894419/full#supplementary-material

Supplementary Figure 1 | FACS sorting gating strategy **(A)** Rosettesep enriched B cells were depleted of IgA^+ and aqua dead cells in all of the sorting procedures **(B)** four-way pre-sort with yield setting to separate enriched viable B cells into $IG\kappa^+/CD5^+$, $IG\kappa^+/CD5^-$, $IG\lambda^+/CD5^+$ and $IG\lambda^+/CD5^-$ B cells **(C)** $IG\kappa^+/CD5^-$ B cells subset (here shown as representative of all the other above subsets) was further sorted by gating $CD24^-CD38^{high}$ B cells to exclude plasmablasts and separated into $CD24^{high}CD38^{high}$ transitional (TR). Gating $CD38^{-/+}$ B cells were then separated based upon IgD and CD27 expression markers and sorted as $IgD^{++}CD27^-$ naive (N), $IgD^{low}CD27^+$ marginal zone-like (MZ), IgD^-CD27^- double negative (DN) B cells. IgD^-CD27^+ memory B cells were gated to isolate $IgM^+IgD^-CD27^+$ IgM-only memory (MO) and IgM^-IgD^-CD27+ switch-memory (SM) B cells.

Supplementary Figure 2 | Frequency of individual CLS-IG in sequences from normal B cells expressing either IGκ and IGλ chains in CLL stereotyped subsets.

Supplementary Figure 3 | Frequency of IGHV genes used in typical and non-typical CLS-IG (red). Blue bars indicate non-CLS-IG control sequences (i.e., sharing the same core features as the CLL subset in consideration - IGHV clan, IGHV mutational status, and VH CDR3 length). These rearrangements were used as controls. Black horizontal lines indicate the level of IGHV representation in the reference CLL cohort.

REFERENCES

1. Rozman C, Montserrat E. Chronic Lymphocytic Leukemia. *New Engl J Med* (1995) 333:1052–7. doi: 10.1056/nejm199510193331606
2. Chiorazzi N, Rai KR, Ferrarini M. Chronic Lymphocytic Leukemia. *New Engl J Med* (2005) 352:804–15. doi: 10.1056/nejmra041720
3. Schroeder HW, Dighiero G. The Pathogenesis of Chronic Lymphocytic Leukemia: Analysis of the Antibody Repertoire. *Immunol Today* (1994) 15:288–94. doi: 10.1016/0167-5699(94)90009-4
4. Johnson TA, Rassenti LZ, Kipps TJ. Ig VH1 Genes Expressed in B Cell Chronic Lymphocytic Leukemia Exhibit Distinctive Molecular Features. *J Immunol Baltim Md 1950* (1997) 158:235–46.
5. Fais F, Ghiotto F, Hashimoto S, Sellars B, Valetto A, Allen SL, et al. Chronic Lymphocytic Leukemia B Cells Express Restricted Sets of Mutated and Unmutated Antigen Receptors. *J Clin Invest* (1998) 102:1515–25. doi: 10.1172/jci3009
6. Agathangelidis A, Brochet X, Murray F, Giudicelli V, Bonello L, Janus A, et al. Stereotyped B-Cell Receptors in One-Third of Chronic Lymphocytic Leukemia: A Molecular Classification With Implications for Targeted Therapies. *Blood* (2012) 119:4467–75. doi: 10.1182/blood-2011-11-393694
7. Agathangelidis A, Chatzidimitriou A, Gemenetzi K, Giudicelli V, Karypidou M, Plevova K, et al. Higher-Order Connections Between Stereotyped Subsets: Implications for Improved Patient Classification in CLL. *Blood* (2020) 137:1365–76. doi: 10.1182/blood.2020007039
8. Darzentas N, Hadzidimitriou A, Murray F, Hatzi K, Josefsson P, Laoutaris N, et al. A Different Ontogenesis for Chronic Lymphocytic Leukemia Cases Carrying Stereotyped Antigen Receptors: Molecular and Computational Evidence. *Leukemia : Off J Leukemia Soc America Leukemia Res Fund UK* (2010) 24:125–32. doi: 10.1038/leu.2009.186
9. Elemento O, Lefranc M-P. IMGT/PhyloGene: An on-Line Tool for Comparative Analysis of Immunoglobulin and T Cell Receptor Genes. *Dev Comp Immunol* (2003) 27:763–79. doi: 10.1038/s41586-019-0934-8
10. Soto C, Bombardi RG, Branchizio A, Kose N, Matta P, Sevy AM, et al. High frequency of shared clonotypes in human B cell receptor repertoires. *Nature* (2019) 566:398–402. doi: 10.1038/s41586-019-0934-8
11. Ghiotto F, Fais F, Valetto A, Albesiano E, Hashimoto S, Dono M, et al. Remarkably Similar Antigen Receptors Among a Subset of Patients With Chronic Lymphocytic Leukemia. *J Clin Invest* (2004) 113:1008–16. doi: 10.1172/jci19399
12. Marcatili P, Ghiotto F, Tenca C, Chailyan A, Mazzarello AN, Yan X, et al. Igs Expressed by Chronic Lymphocytic Leukemia B Cells Show Limited Binding-Site Structure Variability. *J Immunol* (2013) 190:5771–8. doi: 10.4049/jimmunol.1300321
13. Stamatopoulos K, Agathangelidis A, Rosenquist R, Ghia P. Antigen Receptor Stereotypy in Chronic Lymphocytic Leukemia. *Leukemia : Off J Leukemia Soc America Leukemia Res Fund UK* (2017) 31:282–91. doi: 10.1038/leu.2016.322
14. Chiorazzi N, Ferrarini M. B Cell Chronic Lymphocytic Leukemia: Lessons Learned From Studies of the B Cell Antigen Receptor. *Annu Rev Immunol* (2003) 21:841–94. doi: 10.1146/annurev.immunol.21.120601.141018
15. Ghia P, Caligaris-Cappio F. The Origin of B-Cell Chronic Lymphocytic Leukemia. *Semin Oncol* (2006) 33:150–6. doi: 10.1053/j.seminoncol.2006.01.009
16. TenHacken E, Gounari M, Ghia P, Burger JA. The Importance of B Cell Receptor Isotypes and Stereotypes in Chronic Lymphocytic Leukemia. *Leukemia* (2019) 33:287–98. doi: 10.1038/s41375-018-0303-x
17. Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, Blum KA, et al. Targeting BTK With Ibrutinib in Relapsed Chronic Lymphocytic Leukemia. *New Engl J Med* (2013) 369:32–42. doi: 10.1056/nejmoa1215637
18. Forconi F, Potter KN, Wheatley I, Darzentas N, Sozzi E, Stamatopoulos K, et al. The Normal IGHV1-69-Derived B-Cell Repertoire Contains Stereotypic Patterns Characteristic of Unmutated CLL. *Blood* (2010) 115:71–7. doi: 10.1182/blood-2009-06-225813
19. Muggen AF, de Jong M, Wolvers-TetteroILM, Kallemeijn MJ, Teodósio C, Darzentas N, et al. The Presence of CLL-Associated Stereotypic B Cell Receptors in the Normal BCR Repertoire From Healthy Individuals Increases With Age. *Immun Ageing* (2019) 16:22. doi: 10.1186/s12979-019-0163-x
20. Colombo M, Bagnara D, Reverberi D, Matis S, Cardillo M, Massara R, et al. Tracing CLL-Biased Stereotyped Immunoglobulin Gene Rearrangements in Normal B Cell Subsets Using a High-Throughput Immunogenetic Approach. *Mol Med (Cambridge Mass)* (2020) 26:25. doi: 10.1186/s10020-020-00151-9
21. Seifert M, Sellmann L, Bloehdorn J, Wein F, Stilgenbauer S, Dürig J, et al. Cellular Origin and Pathophysiology of Chronic Lymphocytic Leukemia. *J Exp Med* (2012) 209:2183–98. doi: 10.1084/jem.20120833
22. Vergani S, Korsunsky I, Mazzarello AN, Ferrer G, Chiorazzi N, Bagnara D. Novel Method for High-Throughput Full-Length IGHV-D-J Sequencing of the Immune Repertoire From Bulk B-Cells With Single-Cell Resolution. *Front Immunol* (2017) 8:1157. doi: 10.3389/fimmu.2017.01157
23. Heiden JAV, Yaari G, Uduman M, Stern JNH, O'Connor KC, Hafler DA, et al. pRESTO: A Toolkit for Processing High-Throughput Sequencing Raw Reads of Lymphocyte Receptor Repertoires. *Bioinf (Oxford England)* (2014) 30:1930–2. doi: 10.1093/bioinformatics/btu138
24. Alamyar E, Giudicelli V, Li S, Duroux P. IMGT/HighV-QUEST: The IMGT® Web Portal for Immunoglobulin (IG) or Antibody and T Cell Receptor (TR) Analysis From NGS High Throughput and Deep Sequencing. *Immunome Res* (2012) 8(1):26. doi: 10.4172/1745-7580.1000056
25. Gupta NT, Heiden JAV, Yaari G, Kleinstein SH. Change-O: A Toolkit for Analyzing Large-Scale B Cell Immunoglobulin Repertoire Sequencing Data. *Bioinf (Oxford England)* (2015) 31:3356–8. doi: 10.1093/bioinformatics/btv359
26. Bystry V, Agathangelidis A, Bikos V, Sutton LA, Baliakas P, Hadzidimitriou A, et al. ARResT/AssignSubsets: A Novel Application for Robust Subclassification of Chronic Lymphocytic Leukemia Based on B Cell Receptor IG Stereotypy. *Bioinformatics* (2015) 31:3844–6. doi: 10.1093/bioinformatics/btv456
27. Widhopf GF, Goldberg CJ, Toy TL, Rassenti LZ, Wierda WG, Byrd JC, et al. Nonstochastic Pairing of Immunoglobulin Heavy and Light Chains Expressed by Chronic Lymphocytic Leukemia B Cells is Predicated on the Heavy Chain CDR3. *Blood* (2008) 111:3137–44. doi: 10.1182/blood-2007-02-073130
28. Matthews C, Catherwood MA, Morris TCM, Alexander HD. VH3–48 and VH3–53, as Well as VH3–21, Gene Rearrangements Define Unique Subgroups in CLL and are Associated With Biased Lambda Light Chain Restriction, Homologous LCDR3 Sequences and Poor Prognosis. *Leukemia Res* (2007) 31:231–4. doi: 10.1016/j.leukres.2006.03.028
29. Price DA, Asher TE, Wilson NA, Nason MC, Brenchley JM, Metzler IS, et al. Public Clonotype Usage Identifies Protective Gag-Specific CD8+ T Cell Responses in SIV Infection. *J Exp Med* (2009) 206:923–36. doi: 10.1084/jem.20081127
30. Tan TJC, Yuan M, Kuzelka K, Padron GC, Beal JR, Chen X, et al. Sequence Signatures of Two Public Antibody Clonotypes That Bind SARS-CoV-2 Receptor Binding Domain. *Nat Commun* (2021) 12:3815. doi: 10.1038/s41467-021-24123-7
31. Agathangelidis A, Galigalidou C, Scarfò L, Moysiadis T, Rovida A, Gounari M, et al. Infrequent "Chronic Lymphocytic Leukemia-Specific" Immunoglobulin Stereotypes in Aged Individuals With or Without Low Count Monoclonal B Cell Lymphocytosis. *Haematologica* (2020) 106:haematol.2020.247908. doi: 10.3324/haematol.2020.247908
32. Kolijn PMM, Hosnijeh FS, Späth F, Hengeveld PJ, Agathangelidis A, Saleh M, et al. HIGH-RISK SUBTYPES OF CHRONIC LYMPHOCYTIC LEUKEMIA ARE DETECTABLE AS EARLY AS 16 YEARS PRIOR TO DIAGNOSIS. *Blood* (2022) 139:1557–63. doi: 10.1182/blood.2021012890
33. Brezinschek H, Brezinschek RI, DÖrner T, Lipsky PE. Similar Characteristics of the CDR3 of VH1-69/DP-10 Rearrangements in Normal Human Peripheral Blood and Chronic Lymphocytic Leukaemia B Cells. *Brit J Haematol* (1998) 102:516–21. doi: 10.1046/j.1365-2141.1998.00787.x
34. Potter KN, Orchard J, Critchley E, Mockridge CI, Jose A, Stevenson FK. Features of the Overexpressed V1-69 Genes in the Unmutated Subset of Chronic Lymphocytic Leukemia are Distinct From Those in the Healthy Elderly Repertoire. *Blood* (2003) 101:3082–4. doi: 10.1182/blood-2002-08-2432
35. Colombo M, Cutrona G, Reverberi D, Bruno S, Ghiotto F, Tenca C, et al. Expression of Ig Receptors With Distinctive Features Indicating Antigen Selection by Marginal Zone B Cells From Human Spleen. *Mol Med (Cambridge Mass)* (2013) 19:1. doi: 10.2119/molmed.2013.00069
36. Ghiotto F, Fais F, Albesiano E, Sison C, Valetto A, Gaidano G, et al. Similarities and Differences Between the Light and Heavy Chain Ig Variable

Region Gene Repertoires in Chronic Lymphocytic Leukemia. *Mol Med* (2006) 12:300–8. doi: 10.2119/2006-00080.ghiotto

37. Ghia EM, Jain S, Widhopf GF, Rassenti LZ, Keating MJ, Wierda WG, et al. Use of IGHV3–21 in Chronic Lymphocytic Leukemia is Associated With High-Risk Disease and Reflects Antigen-Driven, Post–Germinal Center Leukemogenic Selection. *Blood* (2008) 111:5101–8. doi: 10.1182/blood-2007-12-130229
38. Ghia EM, Widhopf GF, Rassenti LZ, Kipps TJ. Analyses of Recombinant Stereotypic IGHV3-21–Encoded Antibodies Expressed in Chronic Lymphocytic Leukemia. *J Immunol* (2011) 186:6338–44. doi: 10.4049/jimmunol.0902875

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Oncology

ORIGINAL RESEARCH
published: 12 July 2022
doi: 10.3389/fonc.2022.897280

Impact of the Types and Relative Quantities of IGHV Gene Mutations in Predicting Prognosis of Patients With Chronic Lymphocytic Leukemia

Matthew Kaufman[1†‡], Xiao-Jie Yan[1‡], Wentian Li[2‡], Emanuela M. Ghia[3], Anton W. Langerak[4], Laura Z. Rassenti[3], Chrysoula Belessi[5], Neil E. Kay[6], Frederic Davi[7], John C. Byrd[8], Sarka Pospisilova[9], Jennifer R. Brown[10], Mark Catherwood[11], Zadie Davis[12], David Oscier[13], Marco Montillo[14], Livio Trentin[15], Richard Rosenquist[16], Paolo Ghia[17], Jacqueline C. Barrientos[1,18,19,20], Jonathan E. Kolitz[1,20], Steven L. Allen[1,20], Kanti R. Rai[1,18,19,20], Kostas Stamatopoulos[21], Thomas J. Kipps[3], Donna Neuberg[22] and Nicholas Chiorazzi[1,18,19,20]*

[1] Karches Center for Oncology Research, The Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, United States, [2] The Robert S. Boas Center for Genomics & Human Genetics, The Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, United States, [3] Center for Novel Therapeutics, Moores Cancer Center, University of California, San Diego, La Jolla, CA, United States, [4] Laboratory Medical Immunology, Department of Immunology, Erasmus MC, University Medical Center, Rotterdam, Netherlands, [5] Hematology Department, Nikea General Hospital, Pireaus, Greece, [6] Division of Hematology, Mayo Clinic, Rochester, MN, United States, [7] Department of Biological Hematology, Hôpital Pitié-Salpêtrière (AP-HP), Sorbonne Université, Paris, France, [8] Department of Internal Medicine, University of Cincinnati College of Medicine, Cincinnati, OH, United States, [9] Department of Internal Medicine - Hematology and Oncology and Department of Medical Genetics and Genomics, University Hospital Brno and Faculty of Medicine, Masaryk University, Brno, Czechia, [10] Chronic Lymphocytic Leukemia Center, Dana-Farber Cancer Institute, Boston, MA, United States, [11] Clinical Hematology, Belfast City Hospital, Belfast, Ireland, [12] Department of Molecular Pathology, Royal Bournemouth Hospital, Bournemouth, United Kingdom, [13] Department of Hematology, Royal Bournemouth Hospital, Bournemouth, United Kingdom, [14] Department of Hematology & Oncology, Niguarda Cancer Center, Niguarda Hospital, Milan, Italy, [15] Hematology Unit, Department of Medicine-(DIMED), University of Padua University Hospital, Padua, Italy, [16] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden, [17] Division of Experimental Oncology, IRCCS Ospedale San Raffaele, Milan, Italy, [18] Department of Molecular Medicine, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Uniondale, NY, United States, [19] Departments of Medicine, Donald and Barbara Zucker School of Medicine at Hofstra/Northwell, Uniondale, NY, United States, [20] Northwell Health Cancer Institute, Lake Success, NY, United States, [21] Institute of Applied Biosciences, Centre for Research and Technology Hellas, Thessaloniki, Greece, [22] Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, United States

Patients with CLL with mutated IGHV genes (M-CLL) have better outcomes than patients with unmutated IGHVs (U-CLL). Since U-CLL usually express immunoglobulins (IGs) that are more autoreactive and more effectively transduce signals to leukemic B cells, B-cell receptor (BCR) signaling is likely at the heart of the worse outcomes of CLL cases without/few IGHV mutations. A corollary of this conclusion is that M-CLL follow less aggressive clinical courses because somatic IGHV mutations have altered BCR structures and no longer bind stimulatory (auto)antigens and so cannot deliver trophic signals to leukemic B cells. However, the latter assumption has not been confirmed in a large patient cohort. We tried to address the latter by measuring the relative numbers of replacement (R) mutations that lead to non-conservative amino acid changes (Rnc) to the combined numbers of conservative (Rc) and silent (S) amino acid R mutations that likely do not or cannot change

OPEN ACCESS

Edited by:
Manlio Ferrarini, University of Genoa, Italy

Reviewed by:
Dimitar G Efremov, International Centre for Genetic Engineering and Biotechnology, Italy
Fortunato Morabito, Provincial Health Authority of Cosenza, Italy

***Correspondence:**
Nicholas Chiorazzi
NChizzi@Northwell.edu

†Present Address:
Matthew Kaufman, Department of Emergency Medicine, Jersey City Medical Center, Robert Wood Johnson-Barnabas Health, Jersey City, NJ, United States

‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: 15 March 2022
Accepted: 30 May 2022
Published: 12 July 2022

Citation:
Kaufman M, Yan X-J, Li W, Ghia EM, Langerak AW, Rassenti LZ, Belessi C, Kay NE, Davi F, Byrd JC, Pospisilova S, Brown JR, Catherwood M, Davis Z, Oscier D, Montillo M, Trentin L, Rosenquist R, Ghia P, Barrientos JC, Kolitz JE, Allen SL, Rai KR, Stamatopoulos K, Kipps TJ, Neuberg D and Chiorazzi N (2022) Impact of the Types and Relative Quantities of IGHV Gene Mutations in Predicting Prognosis of Patients With Chronic Lymphocytic Leukemia. Front. Oncol. 12:897280. doi: 10.3389/fonc.2022.897280

amino acids, "(S+Rc) to Rnc IGHV mutation ratio". When comparing time-to-first-treatment (TTFT) of patients with (S+Rc)/Rnc ≤ 1 and >1, TTFTs were similar, even after matching groups for equal numbers of samples and identical numbers of mutations per sample. Thus, BCR structural change might not be the main reason for better outcomes for M-CLL. Since the total number of IGHV mutations associated better with longer TTFT, better clinical courses appear due to the biologic state of a B cell having undergone many stimulatory events leading to IGHV mutations. Analyses of larger patient cohorts will be needed to definitively answer this question.

Keywords: chronic lymphocytic leukemia, CLL, somatic mutations, immunoglobulin variable domain, prognosis

INTRODUCTION

Patients with chronic lymphocytic leukemia (CLL) whose leukemic clone uses a mutated immunoglobulin heavy variable (IGHV) gene (M-CLL) typically have less aggressive disease than patients with CLL that use an unmutated IGHV gene (U-CLL) (1, 2). This observation has had a major direct impact on predicting the prognosis of CLL and a significant influence on its understanding and management (3, 4). Documenting this gene distinction is now considered a reliable prognostic factor, and the International Workshop on CLL (5) recommend checking this as a guideline for patient care and management. Moreover, IGHV-mutation status can help predict outcome for patients treated with chemoimmunotherapy (fludarabine, cyclophosphamide, and rituximab) (6–8), and for U-CLL patients treated with ibrutinib vs. chemoimmunotherapy (9). Moreover, IGHV-mutation status, along with other parameters, is incorporated into several prognostic algorithms (10–12).

There is speculation that the relationship between less aggressive disease and expression of mutated IGHVs is due to a loss or attenuation of autoreactivity of membrane immunoglobulin (IG), a major component of the B-cell antigen receptor (BCR), which limits the ability of the receptor to deliver trophic signals to the leukemic B cells. There is ample support for this concept. For example, U-CLL-derived IGs are extensively autoreactive, binding multiple self-molecules (13–15), especially those generated by apoptosis and protein catalysis (16–18). These are often referred to as natural autoantibodies. In contrast, M-CLL IGs are much less autoreactive. Notably, reverting M-CLL IGs to their germline sequence can lead to autoantigen binding, implying that those B cells that became leukemic *in vivo* might have been self-reactive prior to accumulating IGHV mutations (17, 19, 20). Thus, the process of losing autoantigen binding by somatic IGHV mutations can occur normally during B cell maturation, validating the speculation that this could explain the extended clinical course of patients with M-CLL. Additionally, CLL clones differ in their responsiveness to BCR engagement, with surrogate antigen binding, e.g., interaction with anti-IG antibodies, being more effective in stimulating U-CLL than M-CLL cases (21–23), and the ability to deliver a signal *via* the BCR correlating with worse clinical outcomes (24, 25). Finally, and possibly most convincingly, inhibition of BCR signaling by blocking the action of Bruton's tyrosine kinase (BTK) (26–29) or of phosphoinositide 3' kinase delta (PI3Kδ) (30–33) has a very significant effect on CLL cell survival, growth, and trafficking (9, 34–38). Such signaling inhibitors have had a major impact on the quality of patient lives, along with very high overall response rates and, in combination with other reagents, improving overall survival (39–42).

Nevertheless, the concept that the loss of polyreactive antigen binding and BCR signaling is at the root of better prognosis has not been directly confirmed in a large patient cohort, and that correlation is the intention of this investigation and report. For this process to be in play in most instances, only replacement (R) mutations and, in particular non-conservative R (Rnc) mutations, would be most relevant, since only R, and especially Rnc mutations can change the amino acid composition of an IGHV-IGHD-IGHJ rearrangement, thereby potentially altering (auto)antigen binding and eliminating or reducing BCR signaling. Conservative R (Rc) amino acid changes are less likely to alter protein structure and thereby (auto)antigen binding, and silent (S) mutations, by definition, cannot. Hence, since Rnc mutations more often alter amino acid structure, they are more prone to reduce BCR binding, and preempt cell signaling. In general terms, Rnc mutations yield an amino acid that has features opposite or distinct from those of the original one, e.g., hydrophilic vs. hydrophobic or non-polar vs. charged polar amino acids (43, 44). Additionally, in some instances, only a single R can lead to major alterations in protein structure and result in disease, e.g., cystic fibrosis and sickle cell anemia.

Here, we investigated the roles of S, Rc, and Rnc mutations on time-to-first-treatment (TTFT) of patients with CLL. This was addressed using a database of IGHV-IGHD-IGHJ DNA sequences with linked clinical information obtained from institutions in the United States of America and Europe. Our findings suggest that the relative frequencies of (S + Rc) IGHV mutations, which are less likely to create a major BCR structural change, are as important and correlate equally well with improved clinical course as Rnc mutations that are more likely to create a major BCR structural change. Moreover, our data suggest that the total number of mutations in the clonotypic rearranged IGHV gene of a CLL cell might be more central to better prognosis, suggesting that an overriding reason that IGHV mutations associate with better clinical course is the biologic state of a B cell that has undergone several rounds of stimulation

leading to germinal center reactions, possibly varying in the follicular and extra-follicular types.

METHODS

Patient Information and IGHV Sequence Data

Patient and corresponding IGHV-IGHD-IGHJ DNA sequence information (n = 3,598) were received from two large consortia studying CLL and BCR structure: the NIH-sponsored CLL Research Consortium (CRC) and the European Research Initiative on CLL/ImMunoGeneTics (ERIC/IMGT). CLL was defined as suggested by the latest guidelines from the International Workshops on CLL (5).

Analysis of Immunogenetic Characteristics

The same IGMT software and tools were used by both consortia to analyze IGHV-IGHD-IGHJ sequences from the leukemic clones of CLL patients and to characterize and define if IGHV mutations could change amino acid sequence (S vs. R) and if R mutations were conservative (Rc) or non-conservative (Rnc), as determined by charge, hydropathy, and size.

Analysis of Time to First Treatment

TTFT was defined as the number of years between the date of diagnosis and the date of initial therapy. The *survival* package of the R (statistical computing platform: https://www.r-project.org/) was used to estimate TTFT (Kaplan-Meier plots) through the *survfit* function. The impact of an independent variable on TTFT was determined by the Cox proportional hazard model (Cox regression) using the *coxph* function. To create graphic representations of TTFT in the figures in this document, Prism software and the log rank test were used. Nominal *P*-values are presented, without adjustment for multiplicity of testing.

Division of Individual IGHV-IGHD-IGHJ Sequences Into Groups Based on the Ratio of IGHV Mutations More or Less Prone to Lead to a Significant BCR Structural Change

IGHV-IGHD-IGHJ gene rearrangement sequences were segregated based on the ratio of the combined number of S + Rc IGHV mutations divided by the number of Rnc IGHV mutations: (S+Rc)/Rnc. To allow mathematical comparisons should any type of mutation be absent, a value of 0.05 was added to each ratio component, (S+Rc +0.05)/(Rnc+0.05). For simplicity, the latter is represented in the text as "(S+Rc/Rnc". An arbitrary (S+Rc)/Rnc percentage cutoff was chosen so that the calculated value would indicate the likelihood that amino acid change could appreciably alter BCR structure. Sequences with values of ≤ 1.0, based on the (S+Rc)/Rnc calculation, were considered more likely to lead to a major BCR structural change and are referred to as being in the "Low Ratio Group"; those sequences with values of > 1.0, were considered less likely to lead to a major structural change and are referred to as being in the "High Ratio Group". Also, when another threshold cutoff value was used, i.e., the median value of all (S+Rc)/Rnc percentages (1.91), the reported findings were essentially the unchanged (**Figure S1**). For sequences with no IGHV somatic mutation, (S+Rc)/Rnc is not defined.

Multiparameter Analysis of TTFT Comparing (S+Rc)/Rnc Percentage and Total Number of Mutations, Regardless of Type (S, Rc, Rnc)

A two-variable Cox regression was used to compare the total number of mutations (≥ 1) as the first variable, and logarithm-transformed mutation type ratio (S+Rc+0.5)/(Rnc+0.5) as the second variable, was used. A log-transformed ratio variable was employed as that more accurately followed a normal distribution than the ratio variable itself. The underlying assumption of the multivariable Cox regression was that the total number of mutations and the log-ratio jointly contribute to TTFT in an additive manner, and the contribution is averaged over the entire range of the variable values.

RESULTS

Comparing and Coalescing the IGHV-IGHD-IGHJ Gene Rearrangement Sequence Data From the Two Consortia

To assure that merging the IGHV-IGHD-IGHJ gene rearrangement data from the CRC and ERIC/IMGT was appropriate, the distribution of sequences bearing the various types of somatic mutations was compared (**Table 1**). The sequence data from the CRC were collected from 1,690 patients with CLL; 36% were 100% unmutated (0 Mut), 2% had only S mutations (S only), 7% had only R mutations (R only), and 55% had a combination of S and R mutations (S+R). The DNA sequences from ERIC/IMGT were from 1,908 patients; 34% were 0 Mut, 1.6% S only, 6.4% R only, and 58% S+R. Thus, the IGHV mutations were similar in types, patterns, and distributions between the two data sets.

Moreover, although the methodologies to obtain IGHV-IGHD-IGHJ sequence data were not uniform among all the institutions in the two consortia, the same IMGT tools were used to analyze the data.

Therefore, based on very similar patient mutation parameters and uniform analytic approaches, the data from the two sites were pooled and used in the findings described here. The breakdown for the combined IGHV sequence data for the combined 3,598 patients with CLL was 35% 0 Mut (n = 1,259), 1.8% S only (n = 65), 7% R only (n = 236), and 57% S+R (n = 2,038).

TTFT Defined by the Classical IGHV-Mutation Status Approach

As per convention, we first divided all 3,598 patients, using the standard 2% difference from germline cutoff, into IGHV-unmutated (U-CLL; n = 1,713) and IGHV-mutated (M-CLL; n = 1,885) subgroups, and then used the Kaplan-Meier approach

TABLE 1 | Distributions of mutation types between the CLL Research Consortium (CRC) and the European Research Initiative on CLL/ImMunoGeneTics (ERIC/IMGT).

	CRC	ERIC/IMGT	Total
0 mutations (100% homology with germline)	36.0% (609)	34.1% (650)	35.0% (1,259)
Silent (S) mutations only	2.0% (34)	1.6% (31)	1.8% (65)
Replacement (R) mutations only	6.8% (114)	6.4% (122)	6.6% (236)
Silent + Replacement mutations (S+R)	55.2% (933)	57.9% (1,105)	56.6% (2,038)
Total	100% (1,690)	100% (1,908)	100% (3,598)

to estimate TTFT. In this way, as expected, M-CLL patients had a significantly longer TTFT than U-CLL patients (**Figure 1A**; median TTFT: 9.00 vs. 2.22yrs; $P < 0.0001$), consistent with established, published clinical observations and assuring that the combined cohort was a fair representation of the real-world patient base.

Categorizing IGHV-IGHD-IGHJ Gene Rearrangement Sequences Containing Somatic IGHV Mutations Into Those With Mutation Types More or Less Likely to Change BCR Amino Acid Structure

The most direct way to address BCR structural change as being responsible for differences in TTFT would be to compare the clinical courses of CLL patients bearing somatically mutated IGHVs comprised of only S mutations to those patients whose IGHVs have only R mutations, or preferably, solely Rnc mutations. This would require having an extremely large database containing sequences with at least 5-6 somatic mutations of solely one type, since, depending on the individual IGHV gene expressed by a CLL clone, 5 or 6 mutations are needed to exceed the 2% mutation difference from germline and hence be tested in the standard IGHV-mutation analysis. This was not possible for our patient cohort, since none of the IGHV sequences exhibiting only S mutations reached the required ≥ 5 level, and only 44 sequences contained ≥ 5 R only mutations.

Therefore, we devised an approach that incorporated all patient sequences with IGHVs containing ≥ 1 mutation of any type (n = 2,339), and then segregated these based on the ratios of S + Rc mutations divided by Rnc mutations, (S+Rc)/Rnc. Since BCR ratios ≤ 1.0 ("Low Ratio Group") would have a greater number of Rnc changes, such mutations would more likely lead to significant alterations in (auto)antigen binding. Likewise, BCRs with ratios > 1.0 ("High Ratio Group") would be skewed to having a greater number of S and Rc mutations that would be less likely to change (auto)antigen binding (see **Methods** for details).

After dividing the CLL patients with IGHV mutations in the original cohort into these two ratio groups, we went on to analyze TTFT. Thus, in the following analyses, the TTFT for patients in the Low and High Ratio Groups (**Figure 2**), as well as a subgroup that was created based on equal numbers of samples and mutations per sample (**Figure 5**), were compared in two ways: independently; after dividing each by the 2% IGHV-mutation cutoff; and based on the data obtained disregarding the < 2% or the ≥ 2% cutoff categories.

TTFT for Patients in the Low and High Ratio Groups

When dividing the Low Ratio Group (n = 405) into categories below (worse outcome) and above (better outcome) the 2% cutoff, a clear and significant difference in TTFT was found (**Figure 2A**; median TTFT: 1.91 vs. 6.78 yrs; $P < 0.0001$). To illustrate how this result related to the standard IGHV-mutation

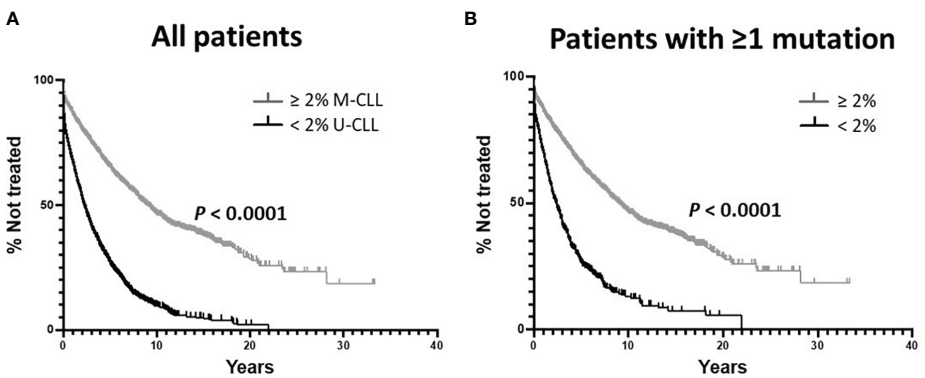


FIGURE 1 | Kaplan-Meier estimates of time to first treatment (TTFT) using the classical IGHV-mutation status parameters. **(A)** Comparison of TTFT based on the < 2% vs. ≥ 2% difference from the germline IGHV sequence. All 3,598 sequences were used without regard for the types of somatic IGHV mutations. M-CLL: 1,885 patients, 856 treated; U-CLL: 1,713 patients, 1,369 treated. Data analyzed using the Log-rank (Mantel-Cox) test. **(B)** Comparison of TTFT based on the < 2% vs. ≥ 2% difference from the germline IGHV sequence analyzing only those patients with ≥ 1 IGHV mutation. < 2%: 454 patients, 351 treated; ≥ 2% 1,885 patients, 856 treated.

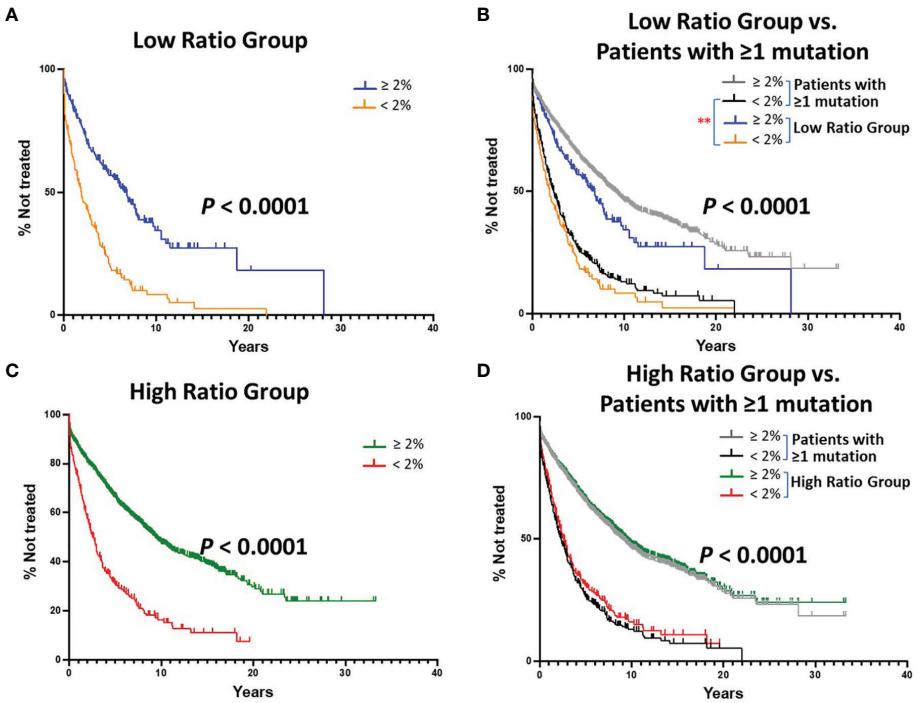


FIGURE 2 | Kaplan-Meier estimates of TTFT of patients with IGHV sequences bearing at least 1 somatic mutation divided into Low or High (S+Rc)/Rnc Ratio Groups. **(A)** TTFT of patients with IGHV genes falling into the "Low Ratio Group", ≤ 1.0 (S+Rc)/Rnc (n = 405) were compared based on the < 2% vs. ≥ 2% difference from the germline IGHV sequence. Number of cases in the < 2% difference group: 183, 152 treated (median TTFT = 1.91 years); number of cases in the ≥ 2% difference group: 222, 119 treated (median TTFT = 6.78 years) (P < 0.0001). **(B)** TTFT of all patients (**Figure 1B**) and those sequences in Low Ratio Group based on the < 2% vs. ≥ 2% difference from the germline IGHV sequence. ** = $P < 0.01$ Pair-wise Log-rank (Mantel-Cox) test. **(C)** TTFT of patients in the High Ratio Group, > 1.0 (S+Rc)/Rnc) (n = 1,934) compared based on the 2% cutoff. Number of cases in the < 2% difference group: 271, 199 treated (median TTFT = 2.59 years); number of cases in the ≥ 2% difference group: 1,663, 737 treated (median TTFT = 9.38 years) (P < 0.0001). **(D)** TTFT based on the < 2% vs. ≥ 2% difference from the germline IGHV sequence using all patient sequences (**Figure 1B**) and those sequences in High Ratio group. ** = $P < 0.01$.

analysis using all patient sequences (**Figure 1B**), the two sets of tracings were overlaid. This indicated that the > 2% category in the Low Ratio Group was different from the same category for all patients (**Figure 2B**; $P < 0.01$). Thus, the > 2% category for the Low Ratio Group has less patients in the better outcome category than the standard IGHV-mutation analysis using all patients; the statistical significance of this finding was not adjusted for multiple comparisons.

When segregating the High Ratio Group (n = 1,934) into the 2% cutoff categories and then comparing TTFT, again a highly significant difference was seen (**Figure 2C**; median TTFT: 2.59 vs. 9.38 years; $P < 0.0001$). Additionally, overlaying the curves from the High Ratio Group with those found using all IGHV-mutated patients (**Figure 1B**), TTFT for both the ≥ 2% and < 2% categories overlapped (**Figure 2C**). Thus, for the High Ratio Group, the relative numbers of patients in each of the two 2% cutoff categories are similar to those from the standard analysis using all patients (**Figure 2D**); this suggests that the High Ratio Group better reflects the IGHV mutation status distribution of the unseparated IGHV-mutated cohort.

Next, we calculated the degree of difference in TTFT of the ≥ 2% and < 2% categories between the Low and High Ratio Groups. This indicated that the High Ratio Group had significantly longer TTFTs for the ≥ 2% (**Figure 3A**; median TTFT: 6.78 vs. 9.38 yrs; $P = 0.0009$) and the < 2% (**Figure 3B**; median TTFT: 1.91 vs. 2.59 yrs; $P = 0.0053$) categories than the Low Ratio Group.

Finally, we compared TTFT for patients in the Low vs. High Ratio Groups without dividing them into < or ≥ 2% categories. This also indicated that the patients in the High Ratio Group had significantly better clinical courses than those in the Low Ratio Group (**Figure 4A**; median TTFT: 3.56 vs. 8.03 yrs; $P < 0.0001$).

Collectively, these calculations indicate that the both the High Ratio and the Low Ratio Groups contain patients with mutated IGHVs with better or worse clinical courses. Also, the inter-group comparisons suggest that the High Ratio Group might have a better TTFT than the Low Ratio Group in more accurately discerning important patient clinical outcomes.

Correcting Analyses for Differences in the Numbers of Sequences of Various Types and in the Number of Mutations Per Type

The above estimates of TTFT using the various comparisons were unexpected in that they suggested that IGHV mutations which would likely change or likely not change BCR structure

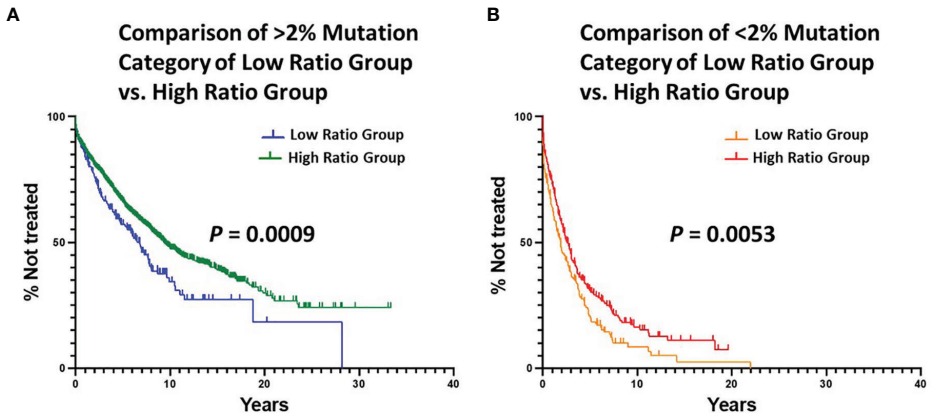


FIGURE 3 | Comparison of estimated TTFT in the < 2% and ≥ 2% mutation categories of the Low Ratio Group and the High Ratio Group. **(A)** Comparison of TTFT in ≥ 2% mutation category of Low vs. High Ratio Groups (P = 0.0009). Number of cases from Low Ratio group: 222, 119 treated (median TTFT = 6.78 years). Number of cases from High Ratio group: 1663, 737 treated (median TTFT = 9.38 years). **(B)** Comparison of TTFT in < 2% mutation category of Low vs. High Ratio Groups (P = 0.0053). Number of cases from Low Ratio group: 183, 152 treated (median TTFT = 1.91 years). Number of cases from High Ratio group: 271, 199 treated (median TTFT = 2.59 years).

impact clinical course in a similar manner in general, and that the TTFT of High Ratio Group is even better. One confounding factor in the comparisons, however, is that the total number of patients in the Low Ratio Group is significantly less than in the High Ratio Group (Low: n = 405; High: n = 1,934). In addition, the DNA sequences of the individual patients in the High Ratio Group had a higher number of IGHV mutations (of any type) than the Low Ratio Group (**Figure S2**). Thus, a significantly smaller number of cases fell into the ≥ 2% category in the Low Ratio than the High Ratio Group (n = 222 vs 1,663).

Therefore, to rule out that the TTFT findings above were artefactual and not reflecting a true biologic effect, we modified the Groups to assure that the numbers of cases in the Low and High Ratio Groups were equal and that the numbers of mutations per sequence in each Group were similar. This was achieved using an exact match approach (45). Specifically, for each sample in the Low Ratio Group, a sample from the High Ratio Group was randomly picked that had the same number of total mutations. The end result was a one-to-one matching in sample size (n = 405) and mutation number between the Low and High Ratio Groups (range: 1 – 36).

Since the number of cases in the Low Ratio Group was constant, the TTFT comparisons for this Group, based on the 2% cutoff, were those already shown in **Figure 2A** (P < 0.0001). The TTFT findings for the High Ratio Group were still very significant (**Figure 5A**; median TTFT: 2.42 vs. 8.50 yrs; P < 0.0001), and the curves based on the 2% cutoff overlapped that of the original High Ratio Group (**Figure 5B**). Moreover, this double matching approach reduced the differences between the Low and High Ratio Groups seen in **Figure 3** to insignificant

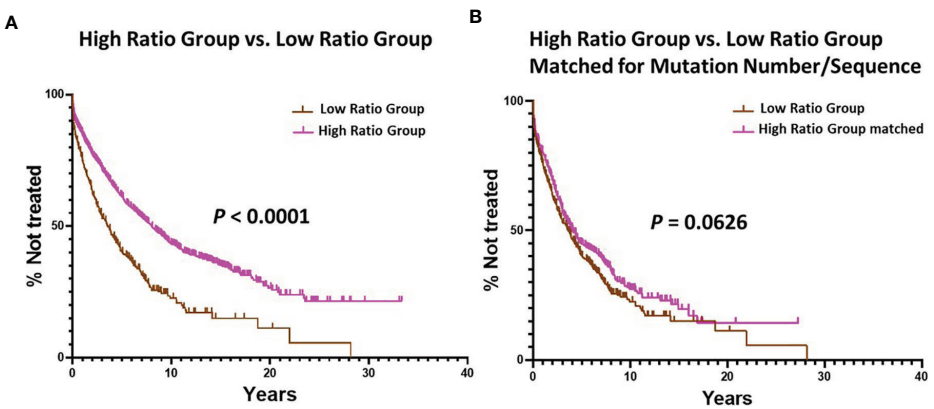


FIGURE 4 | Comparison of estimated TTFT between the Low Ratio Group and the High Ratio Group. **(A)** Comparison of TTFT between the Low and High Ratio Groups using all patients in the respective groups (P < 0.0001). Number in Low Group = 405, median = 3.56 years; number in High Group = 1934, median = 8.03 years. **(B)** Comparison of TTFT after matching the Low and High Ratio Groups for equal numbers of patients (n = 405) with equal numbers of mutations per sequence (range: 1 – 36). Low Group median = 3.56 years, and High Group median = 4.08 years; P = 0.0626.

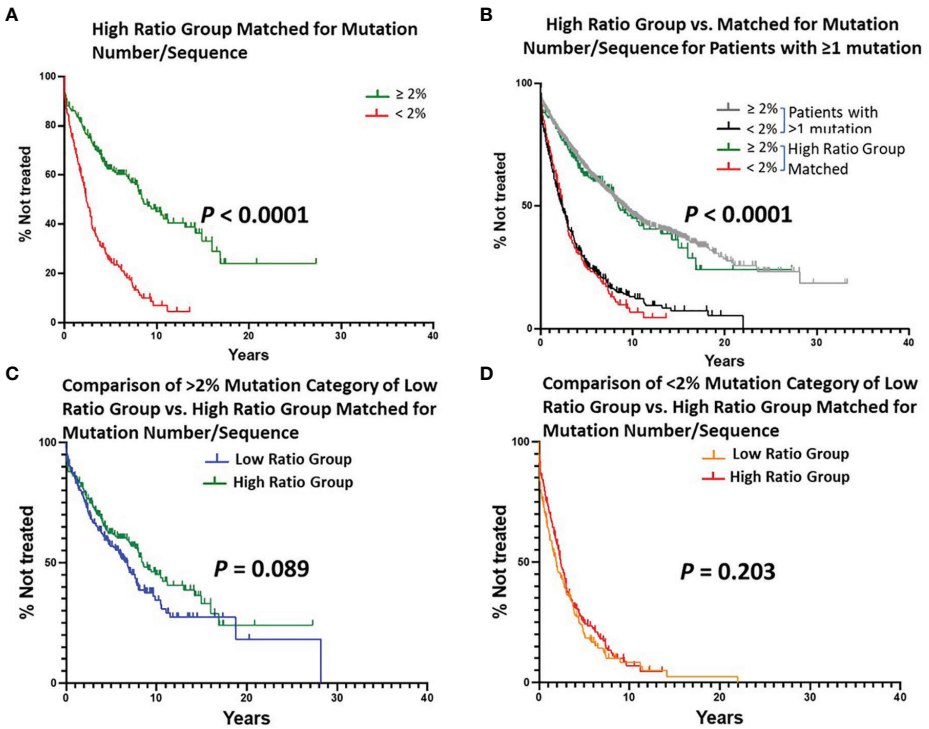


FIGURE 5 | Estimated TTFT of patients in the Low or High Ratio Groups matched for numbers of patients and mutations per sequence. A, B, C, **(D)** Layout of graphs as per **Figure 2**. An exact matching approach with random sampling was used to achieve equal numbers of patients (n = 405) with the same number of IGHV mutations per patient (1-36) in the Low Ratio and High Ratio Groups. **(A)** Number of cases in the < 2% difference group: 187, 148 treated (median TTFT = 2.42 years); number of cases in the ≥ 2% difference group: 218, 101 treated (median TTFT = 8.50 years) (P < 0.0001). **(B)** TTFT based on the < 2% vs. ≥ 2% difference from the germline IGHV sequence using all patient sequences (**Figure 1B**) and those sequences in Matched High Ratio group. **(C)** Number of cases from Low Ratio group: 222, 119 treated (median TTFT = 6.78 years). Number of cases from Matched High Ratio group: 218, 101 treated (median TTFT = 8.50 years). **(D)** Number of cases from Low Ratio group: 183, 152 treated (median TTFT = 1.91 years). Number of cases from High Ratio group: 187, 148 treated (median TTFT = 2.42 years).

levels (difference between > 2% categories: $P = 0.089$; difference between < 2% categories: $P = 0.203$).

Finally, we found that TTFT was equal for patients matched as above in the Low vs. High Ratio Groups without dividing them using the 2% cutoff (**Figure 4B**; TTFT: 3.56 vs. 4.08 yrs; $P = 0.0626$).

Collectively, these findings, especially when using cohorts of equal size and bearing the same number of mutations per sequence, suggest that the relative types of IGHV mutations with a higher (Low Ratio Group) or lower (High Ratio Group) probability of altering BCR structure have the same effect on TTFT. Thus, although somatic mutation altered BCR structure is likely a factor involved in lengthening TTFT in certain instances, it does not appear to be the most influential variable for this patient cohort.

Total Numbers of Mutations, Regardless of Type, Correlate With TTFT After Exceeding a Minimum and Reaching an Apparent Maximum

Since both types of mutations appear to affect TTFT similarly, this implies that the total number of IGHV mutations might correlate better with clinical course in CLL. To test this, we re-ran the IGHV-mutation analysis on the original total patient cohort using all samples with ≥ 1 IGHV mutation, choosing a series of arbitrary mutation number cutoffs, ranging from 1 to > 21 (**Figure 6**). Notably, this approach indicated that the extent of TTFT increased significantly after the number of IGHV mutations reached and exceeded 5-6 mutations (the classical 2% cutoff) and reached a plateau for TTFT at ≥ 10 mutations. This again highlights that the ≥ 2% cutoff is effective in segregating cases into a better outcome group. Additionally, outcomes continue to improve as the number of total mutations increases at least to the level of ~10 mutations/sequence. Another study of patients treated with chemoimmunotherapy found that clinical outcome improved progressively as IGHV mutations increased (46). Notably, incorporating the (S+Rc)/Rnc variable into these arbitrary mutation intervals did not improve clinical course prediction (**Figure S3**).

Multiparameter Analysis Comparing (S+Rc)/Rc Percentage and Total Number of Mutations, Regardless of Type, in Defining TTFT

Finally, we used a two-variable Cox regression to compare the total number of mutations (≥ 1) as the first variable, and logarithm-

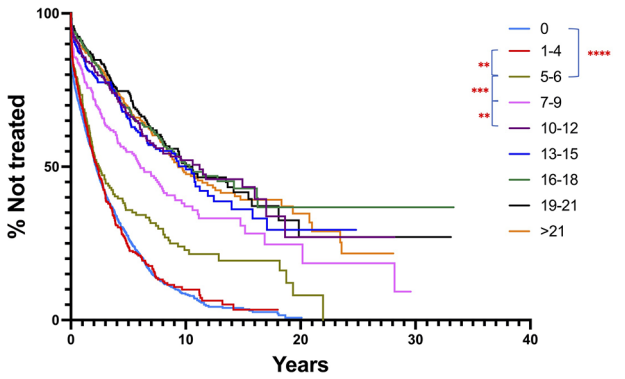


FIGURE 6 | Comparison of TTFT for patients based on IGHV nucleotide mutation number intervals, regardless of mutation type. All patients with ≥ 1 mutation per IGHV sequence were segregated into nucleotide mutation number ranges, and then TTFT compared without using a 2% cutoff. Patients bearing clones without any IGHV mutations are provided for comparison. TTFT and numbers of patients in the various intervals: 1-4: median TTFT - 2.19 yrs, n = 389; 5-6: median TTFT - 2.43 yrs, n = 173; 7-9: median TTFT - 6.06 yrs, n = 232; 10-12: median TTFT - 11.21 yrs, n = 221; 13-15: median TTFT - 10 yrs, n = 257; 16-18: median TTFT - 10.33 yrs, n = 297; 19-21: median TTFT - 10.58 yrs, n = 240; >21: median TTFT - 9.36 yrs, n = 539. ** < 0.01; *** < 0.001.

transformed mutation type ratio (S+Rc+0.5)/(Rnc+0.5) as the second variable (see **Methods**). The *P*-value for testing the null hypothesis that the total number of mutation variable does not contribute to the hazard ratio, is < 0.0001. The hazard ratio (increment of one mutation) is 0.955 (95% confidence interval: 0.948-0.962).

DISCUSSION

Defining the IGHV gene mutation status of a patient's leukemic B cell clone is a cornerstone of prognosis in CLL (3, 47). Although several studies have addressed the best cutoff to be used in this analysis (46, 48–52), there has not been a detailed investigation aimed at determining if a loss of (auto) antigen – BCR interaction, which could obviate or reduce transmission of ongoing survival signals to the leukemic B cell, is the feature that is most responsible for defining M-CLL patients with better clinical outcomes. Here, we have tried to address this issue.

Although the most robust way to address the issue would be to compare the clinical courses and outcomes of CLL patients with IGHV genes expressing solely S vs. patients with solely R, especially Rnc, mutations, our database did not contain sufficient numbers of such cases with only S or only R or Rnc mutations to allow this. Therefore, we compared IGHV mutated sequences segregated based on skewing of the relative frequencies towards (S+Rc) or Rnc mutations, as the former are less likely to make a significant change in BCR structure than the latter. This was done for the entire population of samples with one or more somatic mutations and after normalizing the number of cases within the two (S+Rc)/Rnc High and Low Ratio Groups for numbers of samples and mutations per sequence.

Both relative mutation ratio options, < 1 and ≥ 1 (S+Rc)/Rnc, indicated similar TTFT, suggesting that the ability or inability to carry out BCR-mediated signaling is not the only reason that somatic mutations that lead to a ≥ 2% difference in nucleotide sequence influence clinical course.

However, several caveats need to be considered before excluding that a change in BCR structure is a major driver in determining clinical course. First, it is clear that for some antigens changing a single or only a very few amino acids in the IG variable domain can significantly alter auto- and foreign-antigen binding (53–55). This is the case for CLL B cells as well (56–58). Moreover, changing a single amino acid residue can affect the transmission of "autonomous BCR signaling" (59) based on BCR-BCR autoreactivity (60, 61), which appears to be a key factor in CLL B cell survival (59). Also, the apparent autoreactivity of IGs encoded by unmutated IGHV genes depends upon the proper pairing of IG heavy and light chains and the somatically generated VH CRD3s. As such, autoreactivity should not be considered an intrinsic property of unmutated rearranged IGHV genes commonly used in CLL (e.g., IGHV1-69), suggesting that autoreactivity can be a selected binding activity (19). Finally, it is possible that Rnc mutations lead to enhanced (auto)antigen binding and continuous binding site occupancy, thereby resulting in an anergic state (22, 62) that reduces cell division, clonal burden, and disease progression.

Second, our analyses only relate to potential structural changes involving the IGHV portion of the BCR. Changes in other parts, including the IGK/LV-IGK/LJ gene rearrangement and the VH CDR3 are not considered in this examination, and it is well documented that changes in each of these BCR regions are critical for (auto)antigen binding. Indeed, a characteristic mutation in the IGLV3-21 gene is directly associated with clinical course for CLL patients (63, 64), albeit in the opposite, more adverse, direction than would be predicted by IGHV-mutation analysis.

Nevertheless, the finding that TTFT increases as the number of total mutations increases is compatible with clinical courses being affected by the number of times a CLL cell precursor was signaled to undergo cell division and potentially a germinal center reaction. In this regard, the number of mutations accrued by a normal B lymphocyte is a function of the number of times a cell is stimulated to experience a germinal center reaction, with each episode leading to one mutation (65, 66). In this scenario, a given clinical course would not necessarily be directly affected by the types of IGHV mutations; rather, mutation amounts would imply the number of BCR signals received and the cell divisions undergone. This is consistent with the documentation that clinical outcomes of patients treated with chemoimmunotherapy are reflected by a continuum in the number of mutations a CLL B cells exhibits [**Figure 6** and (46)]. Thus, the physiologic state and the biologic properties of a cell that has undergone many rounds of stimulation and

accumulated multiple mutations, regardless of a BCR structural change, could have a major prognostic impact. This situation is consistent with the phenotype (67), telomere (68), and methylation (69–71) features of CLL B cells, suggesting that all CLL cases, U-CLL as well as M-CLL, are derivatives of chronically stimulated, memory-like B cells. In this regard, it is important to recognize that there are several ways that antigen-experience and germinal center reactions can be initiated and where they can occur (72), e.g., within a lymphoid follicle in a classical germinal center with a well-defined cellular microenvironment, or outside of a follicle (extra-follicular), where the tissue architecture is not as rigidly set as in a germinal center within a follicle. Thus, the biologic characteristics of B cells that mutate in response to different types of (auto)antigenic challenges and within different tissue microenvironments might differ, as might the types and relative quantities of S, Rc, and Rnc mutations that are selected for or against. The similarities and differences in the follicular and extra-follicular B cell differentiation pathways are reviewed in (73, 74).

Finally, the two possibilities, i.e., BCR structural change and the number of CLL precursor B cell divisions, are not mutually exclusive. Certainly, the more often a cell undergoes a germinal center reaction the more likely a key R mutation could occur, especially since only a single or a few key amino acid changes can have dramatic influences on (auto)antigen binding. Moreover, the number of mutations in the IGHV gene, regardless of type, could be a surrogate for what has happened in the VH CDR3 or in the IGHK/LV-IGK/LJ genes. In this regard, one could postulate that the clinical courses of patients in the (S+Rc)/Rnc Low Ratio Group are a direct consequence of BCR structural changes in the IGHV domain of the BCR, and the clinical courses of those patients in the (S+Rc)/Rnc High Ratio Group could be the consequence of structural change that occurred outside of the IGHV domain. Future studies with larger numbers of CLL patient sequences might enable a more definitive understanding of the relative influences of these two parameters on the prognosis of CLL patients with the IGHV-mutated subtype.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.imgt.org/CLLDBInterface/query.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board, The Feinstein Institutes for Medical Research. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MK and NC conceived the project. MK, X-JY, WL, DN, and NC designed the experimental and analytic approaches. MK, X-JY, WL, LR, and DN performed data analyses. X-JY, EG, AL, LR, CB, NK, FD, JoCB, SP, JRB, MC, ZD, DO, MM, LT, RR, PG, JaCB, JK, SA, KR, KS, TK, and NC provided clinical information and IGHV-IGHD-IGHJ DNA sequences. MK, X-JY, WL, EG, CB, NK, FD, ZD, DO, LT, RR, JoCB, JK, SA, TK, and NC edited the manuscript. All authors read the manuscript and agreed with the content.

FUNDING

This study was supported in part by grants from the US National Institutes of Health, National Cancer Institute, PO1-CA81534 and RO1-CA236361 to TK and R01-CA238523 to NC. SP is supported by Ministry of Health of the Czech Republic, grant No. (NV19-03-00091). RR is supported by the Swedish Cancer Society, the Swedish Research Council, the Knut and Alice Wallenberg Foundation, Karolinska Institutet, Karolinska University Hospital, and Radiumhemmets Forskningsfonder, Stockholm. KS has received support from the Hellenic Precision Medicine Network in Oncology; and the project ODYSSEAS, funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014–2020) and co-financed by Greece and the European Union, with grant agreement no: MIS 5002462. NC and KR thank the Karches Family, The Nash Family Foundation, and the Jean Walton Fund for Leukemia, Lymphoma, and Myeloma Research for their support of the Feinstein Institutes' CLL Research & Treatment Program.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2022.897280/full#supplementary-material

REFERENCES

1. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, et al. Ig V Gene Mutation Status and CD38 Expression as Novel Prognostic Indicators in Chronic Lymphocytic Leukemia. *Blood* (1999) 94(6):1840–7. doi: 10.1182/blood.V94.6.1840
2. 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) 94:1848-54. doi: 10.1182/blood.V94.6.1848
3. Kipps TJ, Stevenson FK, Wu CJ, Croce CM, Packham G, Wierda WG, et al. Chronic Lymphocytic Leukaemia. *Nat Rev Dis Primers* (2017) 3:16096. doi: 10.1038/nrdp.2016.96
4. Chiorazzi N, Chen SS, Rai KR. Chronic Lymphocytic Leukemia. *Cold Spring Harb Perspect Med* (2020). doi: 10.1101/cshperspect.a035220
5. Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Dohner H, et al. iwCLL Guidelines for Diagnosis, Indications for Treatment, Response Assessment, and Supportive Management of CLL. *Blood* (2018) 131 (25):2745–60. doi: 10.1182/blood-2017-09-806398

6. Rossi D, Terzi-di-Bergamo L, De Paoli L, Cerri M, Ghilardi G, Chiarenza A, et al. Molecular Prediction of Durable Remission After First-Line Fludarabine-Cyclophosphamide-Rituximab in Chronic Lymphocytic Leukemia. *Blood* (2015) 126(16):1921–4. doi: 10.1182/blood-2015-05-647925
7. Thompson PA, Tam CS, O'Brien SM, Wierda WG, Stingo F, Plunkett W, et al. Fludarabine, Cyclophosphamide, and Rituximab Treatment Achieves Long-Term Disease-Free Survival in IGHV-Mutated Chronic Lymphocytic Leukemia. *Blood* (2016) 127(3):303–9. doi: 10.1182/blood-2015-09-667675
8. Fischer K, Bahlo J, Fink AM, Goede V, Herling CD, Cramer P, et al. Long-Term Remissions After FCR Chemoimmunotherapy in Previously Untreated Patients With CLL: Updated Results of the CLL8 Trial. *Blood* (2016) 127 (2):208–15. doi: 10.1182/blood-2015-06-651125
9. Herman SEM, Mustafa RZ, Gyamfi JA, Pittaluga S, Chang S, Chang B, et al. Ibrutinib Inhibits B-Cell Receptor and NF-κb Signaling and Reduces Tumor Proliferation in Tissue-Resident Cells of Patients With Chronic Lymphocytic Leukemia. *Blood* (2014). doi: 10.1182/blood-2014-02-548610
10. Bahlo J, Kutsch N, Bauer K, Bergmann MA, Byrd J, Chaffee G. An International Prognostic Index for Patients With Chronic Lymphocytic Leukaemia (CLL-IPI): A Meta-Analysis of Individual Patient Data. *Lancet Oncol* (2016) 17(6):779–90. doi: 10.1016/S1470-2045(16)30029-8
11. Delgado J, Doubek M, Baumann T, Kotaskova J, Molica S, Mozas P, et al. Chronic Lymphocytic Leukemia: A Prognostic Model Comprising Only Two Biomarkers (IGHV Mutational Status and FISH Cytogenetics) Separates Patients With Different Outcome and Simplifies the CLL-IPI. *Am J Hematol* (2017) 92(4):375–80. doi: 10.1002/ajh.24660
12. Cohen JA, Rossi FM, Zucchetto A, Bomben R, Terzi-di-Bergamo L, Rabe KG, et al. A Laboratory-Based Scoring System Predicts Early Treatment in Rai 0 Chronic Lymphocytic Leukemia. *Haematologica* (2020) 105(6):1613–20. doi: 10.3324/haematol.2019.228171
13. Broker BM, Klajman A, Youinou P, Jouquan J, Worman CP, Murphy J, et al. Chronic Lymphocytic Leukemic Cells Secrete Multispecific Autoantibodies. *J Autoimmun* (1988) 1(5):469–81. doi: 10.1016/0896-8411(88)90068-6
14. Sthoeger ZM, Wakai M, Tse DB, Vinciguerra VP, Allen SL, Budman DR, et al. Production of Autoantibodies by CD5-Expressing B Lymphocytes From Patients With Chronic Lymphocytic Leukemia. *J Exp Med* (1989) 169 (1):255–68. doi: 10.1084/jem.169.1.255
15. Borche L, Lim A, Binet JL, Dighiero G. Evidence That Chronic Lymphocytic Leukemia B Lymphocytes are Frequently Committed to Production of Natural Autoantibodies. *Blood* (1990) 76(3):562–9. doi: 10.1182/blood.V76.3.562.562
16. Lanemo Myhrinder A, Hellqvist E, Sidorova E, Soderberg A, Baxendale H, Dahle C, et al. A New Perspective: Molecular Motifs on Oxidized LDL, Apoptotic Cells, and Bacteria are Targets for Chronic Lymphocytic Leukemia Antibodies. *Blood* (2008) 111(7):3838–48. doi: 10.1182/blood-2007-11-125450
17. Catera R, Silverman GJ, Hatzi K, Seiler T, Didier S, Zhang L, et al. Chronic Lymphocytic Leukemia Cells Recognize Conserved Epitopes Associated With Apoptosis and Oxidation. *Mol Med* (2008) 14(11-12):665–74. doi: 10.2119/2008-00102.Catera
18. Chu CC, Catera R, Hatzi K, Yan XJ, Zhang L, Wang XB, et al. Chronic Lymphocytic Leukemia Antibodies With a Common Stereotypic Rearrangement Recognize Nonmuscle Myosin Heavy Chain IIA. *Blood* (2008) 112(13):5122–9. doi: 10.1182/blood-2008-06-162024
19. Martin T, Duffy SF, Carson DA, Kipps TJ. Evidence for Somatic Selection of Natural Autoantibodies. *J Exp Med* (1992) 175(4):983–91. doi: 10.1084/jem.175.4.983
20. Herve M, Xu K, Ng YS, Wardemann H, Albesiano E, Messmer BT, et al. Unmutated and Mutated Chronic Lymphocytic Leukemias Derive From Self-Reactive B Cell Precursors Despite Expressing Different Antibody Reactivity. *J Clin Invest* (2005) 115(6):1636–43. doi: 10.1172/JCI24387
21. Lanham S, Hamblin T, Oscier D, Ibbotson R, Stevenson F, Packham G. Differential Signaling *via* Surface IgM is Associated With VH Gene Mutational Status and CD38 Expression in Chronic Lymphocytic Leukemia. *Blood* (2003) 101(3):1087–93. doi: 10.1182/blood-2002-06-1822
22. Mockridge IC, Potter KN, Wheatley I, Neville LA, Packham G, Stevenson FK. Reversible Anergy of Sigm-Mediated Signaling in the Two Subsets of CLL Defined by VH-Gene Mutational Status. *Blood* (2007) 109(10):4424–31. doi: 10.1182/blood-2006-11-056648
23. Guarini A, Chiaretti S, Tavolaro S, Maggio R, Peragine N, Citarella F, et al. BCR-Ligation Induced by IgM Stimulation Results in Gene Expression and Functional Changes Only in IgVH Unmutated Chronic Lymphocytic Leukemia (CLL) Cells. *Blood* (2008) 12:127688. doi: 10.1182/blood-2007-12-127688
24. D'Avola A, Drennan S, Tracy I, Henderson I, Chiecchio L, Larrayoz M, et al. Surface IgM Expression and Function Associate With Clinical Behavior, Genetic Abnormalities and DNA Methylation in CLL. *Blood* (2016) 128:816-27. doi: 10.1182/blood-2016-03-707786
25. Mazzarello AN, Gentner-Göbel E, Dühren-von Minden M, Tarasenko TN, Nicolò A, Ferrer G, et al. B Cell Receptor Isotypes Differentially Associate With Cell Signaling, Kinetics, and Outcome in Chronic Lymphocytic Leukemia. *J Clin Invest* (2022) 132(2):e149308. doi: 10.1172/jci149308
26. O'Brien S BJ, Blum KA, Furman RR, Coutre SE, Sharman J, Flinn IW, et al. The Bruton Tyrosine Kinase Inhibitor PCI-32765 Induces Durable Responses in Relapsed or Refractory (R/R) Chronic Lymphocytic Leukemia/Small Lymphocytic Lymphoma (CLL/SLL): Follow-Up of a Phase Ib/II Study. *Blood* (2011) 118:983. doi: 10.1182/blood.V118.21.983.983
27. Farooqui MZ, Valdez J, Martyr S, Aue G, Saba N, Niemann CU, et al. Ibrutinib for Previously Untreated and Relapsed or Refractory Chronic Lymphocytic Leukaemia With TP53 Aberrations: A Phase 2, Single-Arm Trial. *Lancet Oncol* (2015) 16(2):169–76. doi: 10.1016/S1470-2045(14)71182-9
28. Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, Blum KA, et al. Targeting BTK With Ibrutinib in Relapsed Chronic Lymphocytic Leukemia. *New Engl J Med* (2013) 369(1):32–42. doi: 10.1056/NEJMoa1215637
29. Byrd JC, Furman RR, Coutre SE, Flinn IW, Burger JA, Blum K, et al. Ibrutinib Treatment for First-Line and Relapsed/Refractory Chronic Lymphocytic Leukemia: Final Analysis of the Pivotal Phase Ib/II PCYC-1102 Study. *Clin Cancer Res* (2020) 26:3918-27. doi: 10.1158/1078-0432.ccr-19-2856
30. Burger JA, Hoellenriegel J. Phosphoinositide 3'-Kinase Delta: Turning Off BCR Signaling in Chronic Lymphocytic Leukemia. *Oncotarget* (2011). doi: 10.18632/oncotarget.341
31. Brown JR, Byrd JC, Coutre SE, Benson DM, Flinn IW, Wagner-Johnston ND, et al. Idelalisib, an Inhibitor of Phosphatidylinositol 3-Kinase P110δ, for Relapsed/Refractory Chronic Lymphocytic Leukemia. *Blood* (2014) 123:3390–7. doi: 10.1182/blood-2013-11-535047
32. Flinn IW, Kahl BS, Leonard JP, Furman RR, Brown JR, Byrd JC, et al. Idelalisib, a Selective Inhibitor of Phosphatidylinositol 3-Kinase-δ, as Therapy for Previously Treated Indolent Non-Hodgkin Lymphoma. *Blood* (2014) 123:3406–13. doi: 10.1182/blood-2013-11-538546
33. Furman RR, Sharman JP, Coutre SE, Cheson BD, Pagel JM, Hillmen P, et al. Idelalisib and Rituximab in Relapsed Chronic Lymphocytic Leukemia. *New Engl J Med* (2014) 370(11):997–1007. doi: 10.1056/NEJMoa1315226
34. Herman SE, Gordon AL, Hertlein E, Ramanunni A, Zhang X, Jaglowski S, et al. Bruton Tyrosine Kinase Represents a Promising Therapeutic Target for Treatment of Chronic Lymphocytic Leukemia and is Effectively Targeted by PCI-32765. *Blood* (2011) 117(23):6287–96. doi: 10.1182/blood-2011-01-328484
35. Ponader S, Chen S-S, Buggy JJ, Balakrishnan K, Gandhi V, Wierda WG, et al. The Bruton Tyrosine Kinase Inhibitor PCI-32765 Thwarts Chronic Lymphocytic Leukemia Cell Survival and Tissue Homing *In Vitro* and *In Vivo*. *Blood* (2012) 119(5):1182–9. doi: 10.1182/blood-2011-10-386417
36. Niedermeier M, Hennessy BT, Knight ZA, Henneberg M, Hu J, Kurtova AV, et al. Isoform-Selective Phosphoinositide 3'-Kinase Inhibitors Inhibit CXCR4 Signaling and Overcome Stromal Cell-Mediated Drug Resistance in Chronic Lymphocytic Leukemia: A Novel Therapeutic Approach. *Blood* (2009) 165068:5549-57. doi: 10.1182/blood-2008-06-165068
37. Fiorcari S, Brown WS, McIntyre BW, Estrov Z, Maffei R, O'Brien S, et al. The PI3-Kinase Delta Inhibitor Idelalisib (GS-1101) Targets Integrin-Mediated Adhesion of Chronic Lymphocytic Leukemia (CLL) Cell to Endothelial and Marrow Stromal Cells. *PLoS One* (2013) 8(12):e83830. doi: 10.1371/journal.pone.0083830
38. Göckeritz E, Kerwien S, Baumann M, Wigger M, Vondey V, Neumann L, et al. Efficacy of Phosphatidylinositol-3 Kinase Inhibitors With Diverse Isoform Selectivity Profiles for Inhibiting the Survival of Chronic Lymphocytic Leukemia Cells. *Int J Cancer* (2015) 137:234–42. doi: 10.1002/ijc.29579
39. Shanafelt TD, Wang V, Kay NE, Hanson CA, O'Brien SM, Barrientos JC, et al. Ibrutinib and Rituximab Provides Superior Clinical Outcome Compared to

FCR in Younger Patients With Chronic Lymphocytic Leukemia (CLL): Extended Follow-Up From the E1912 Trial. *Blood* (2019) 134 (Supplement_1):33–. doi: 10.1182/blood-2019-126824
40. Munir T, Brown JR, O'Brien S, Barrientos JC, Barr PM, Reddy NM, et al. Final Analysis From RESONATE: Up to Six Years of Follow-Up on Ibrutinib in Patients With Previously Treated Chronic Lymphocytic Leukemia or Small Lymphocytic Lymphoma. *Am J Hematol* (2019) 94(12):1353–63. doi: 10.1002/ajh.25638
41. Burger JA, Barr PM, Robak T, Owen C, Ghia P, Tedeschi A, et al. Long-Term Efficacy and Safety of First-Line Ibrutinib Treatment for Patients With CLL/SLL: 5 Years of Follow-Up From the Phase 3 RESONATE-2 Study. *Leukemia* (2020) 34(3):787–98. doi: 10.1038/s41375-019-0602-x
42. Langerbeins P, Zhang C, Robrecht S, Cramer P, Fürstenau M, Al-Sawaf O, et al. The CLL12 Trial: Ibrutinib vs Placebo in Treatment-Naïve, Early-Stage Chronic Lymphocytic Leukemia. *Blood* (2022) 139(2):177–87. doi: 10.1182/blood.2021010845
43. Chan KF, Koukouravas S, Yeo JY, Koh DW, Gan SK. Probability of Change in Life: Amino Acid Changes in Single Nucleotide Substitutions. *Biosystems* (2020) 193-194:104135. doi: 10.1016/j.biosystems.2020.104135
44. Dayhoff MO, Schwartz RM, Chen HR, Barker WC, Hunt LT, Orcutt BC. Nucleic Acid Sequence Database. *DNA* (1981) 1(1):51–8. doi: 10.1089/dna.1.1981.1.51
45. Breslow NE, Day NE. Statistical Methods in Cancer Research. Volume I - The Analysis of Case-Control Studies. *IARC Sci Publ* (1980) 32):5–338.
46. Jain P, Nogueras Gonzalez GM, Kanagal-Shamanna R, Rozovski U, Sarwari N, Tam C, et al. The Absolute Percent Deviation of IGHV Mutation Rather Than a 98% Cut-Off Predicts Survival of Chronic Lymphocytic Leukaemia Patients Treated With Fludarabine, Cyclophosphamide and Rituximab. *Br J Haematol* (2018) 180(1):33–40. doi: 10.1111/bjh.15018
47. Chiorazzi N, Stevenson FK. Celebrating 20 Years of IGHV Mutation Analysis in CLL. *Hemasphere* (2020) 4(1):e334. doi: 10.1097/HS9.0000000000000334
48. Bomben R, Dal Bo M, Zucchetto A, Zaina E, Nanni P, Sonego P, et al. Mutational Status of IgV(H) Genes in B-Cell Chronic Lymphocytic Leukemia and Prognosis: Percent Mutations or Antigen-Driven Selection? *Leukemia* (2005) 19(8):1490–2. doi: 10.1038/sj.leu.2403830
49. Hamblin TJ, Davis ZA, Oscier DG. Determination of How Many Immunoglobulin Variable Region Heavy Chain Mutations are Allowable in Unmutated Chronic Lymphocytic Leukaemia - Long-Term Follow Up of Patients With Different Percentages of Mutations. *Br J Haematol* (2008) 140 (3):320–3. doi: 10.1111/j.1365-2141.2007.06928.x__
50. Davis Z, Forconi F, Parker A, Gardiner A, Thomas P, Catovsky D, et al. The Outcome of Chronic Lymphocytic Leukaemia Patients With 97% IGHV Gene Identity to Germline is Distinct From Cases With <97% Identity and Similar to Those With 98% Identity. *Br J Haematology* (2016) 173(1):127–36. doi: 10.1111/bjh.13940
51. Morabito F, Shanafelt TD, Gentile M, Reda G, Mauro FR, Rossi D, et al. Immunoglobulin Heavy Chain Variable Region Gene and Prediction of Time to First Treatment in Patients With Chronic Lymphocytic Leukemia: Mutational Load or Mutational Status? Analysis of 1003 Cases. *Am J Hematol* (2018) 93(9):E216–E9. doi: 10.1002/ajh.25206
52. Shi K, Sun Q, Qiao C, Zhu H, Wang L, Wu J, et al. 98% IGHV Gene Identity is the Optimal Cutoff to Dichotomize the Prognosis of Chinese Patients With Chronic Lymphocytic Leukemia. *Cancer Med* (2020) 9(3):999–1007. doi: 10.1002/cam4.2788
53. Diamond B, Scharff MD. Somatic Mutation of the T15 Heavy Chain Gives Rise to an Antibody With Autoantibody Specificity. *Proc Natl Acad Sci USA* (1984) 81(18):5841–4.
54. Katz JB, Limpanasithikul W, Diamond B. Mutational Analysis of an Autoantibody: Differential Binding and Pathogenicity. *J Exp Med* (1994) 180(3):925–32.
55. Panka DJ, Mudgett-Hunter M, Parks DR, Peterson LL, Herzenberg LA, Haber E, et al. Variable Region Framework Differences Result in Decreased or Increased Affinity of Variant Anti-Digoxin Antibodies. *Proc Natl Acad Sci USA* (1988) 85(9):3080–4. doi: 10.1073/pnas.85.9.3080
56. Catera R, Liu Y, Gao C, Yan XJ, Magli A, Allen SL, et al. Binding of CLL Subset 4 B-Cell Receptor Immunoglobulins to Viable Human Memory B Lymphocytes Requires a Distinctive IGKV Somatic Mutation. *Mol Med* (2017) 23:1-12. doi: 10.2119/molmed.2017.00003
57. Liu Y, Catera R, Gao C, Yan X-J, Magli A, Allen SL, et al. CLL Stereotyped Subset 4 Igs Acquire Binding to Viable B Lymphocyte Surfaces by Somatic Mutations, Isotype Class Switching, and With the Prerequisite of IG Self-Association. *Blood* (2017) 130(Supplement 1):58. doi: 10.1182/blood.V130.Suppl_1.58.58
58. Gounari M, Iatrou A, Kotta K, Sarrigeorgiou I, Lymberi P, Chatzidimitriou A, et al. Changes in N-Glycosylation Induced by Somatic Hypermutation Modulate the Antigen Reactivity of the Immunoglobulin Receptors in CLL Stereotyped Subset 201. *Blood* (2019) 134(Supplement 1):1733. doi: 10.1182/blood-2019-127852
59. Duhren-von Minden M, Ubelhart R, Schneider D, Wossning T, Bach MP, Buchner M, et al. Chronic Lymphocytic Leukaemia is Driven by Antigen-Independent Cell-Autonomous Signalling. *Nature* (2012) 489(7415):309–12. doi: 10.1038/nature11309
60. Binder M, Müller F, Frick M, Wehr C, Simon F, Leistler B, et al. CLL B-Cell Receptors can Recognize Themselves: Alternative Epitopes and Structural Clues for Autostimulatory Mechanisms in CLL. *Blood* (2013) 121(1):239–41. doi: 10.1182/blood-2012-09-454439
61. Minici C, Gounari M, Ubelhart R, Scarfo L, Duhren-von Minden M, Schneider D, et al. Distinct Homotypic B-Cell Receptor Interactions Shape the Outcome of Chronic Lymphocytic Leukaemia. *Nat Commun* (2017) 8:15746. doi: 10.1038/ncomms15746
62. Muzio M, Apollonio B, Scielzo C, Frenquelli M, Vandoni I, Boussiotis V, et al. Constitutive Activation of Distinct BCR-Signaling Pathways in a Subset of CLL Patients: A Molecular Signature of Anergy. *Blood* (2008) 09:111344. doi: 10.1182/blood-2007-09-111344
63. Thorselius M, Krober A, Murray F, Thunberg U, Tobin G, Buhler A, et al. Strikingly Homologous Immunoglobulin Gene Rearrangements and Poor Outcome in VH3-21-Utilizing Chronic Lymphocytic Leukemia Independent of Geographical Origin and Mutational Status. *Blood* (2006) 107:2889–4. doi: 10.1182/blood-2005-06-2227
64. Maity PC, Bilal M, Koning MT, Young M, van Bergen CAM, Renna V, et al. IGLV3-21*01 is an Inherited Risk Factor for CLL Through the Acquisition of a Single-Point Mutation Enabling Autonomous BCR Signaling. *Proc Natl Acad Sci USA* (2020) 117(8):4320–7. doi: 10.1073/pnas.1913810117
65. Kocks C, Rajewsky K. Stable Expression and Somatic Hypermutation of Antibody V Regions in B-Cell Developmental Pathways. *Annu Rev Immunol* (1989) 7:537–59. doi: 10.1146/annurev.iy.07.040189.002541
66. Kleinstein SH, Louzoun Y, Shlomchik MJ. Estimating Hypermutation Rates From Clonal Tree Data. *J Immunol* (2003) 171(9):4639–49. doi: 10.4049/jimmunol.171.9.4639
67. Damle RN, Ghiotto F, Valetto A, Albesiano E, Fais F, Yan XJ, et al. B-Cell Chronic Lymphocytic Leukemia Cells Express a Surface Membrane Phenotype of Activated, Antigen-Experienced B Lymphocytes. *Blood* (2002) 99(11):4087–93. doi: 10.1182/blood.V99.11.4087
68. Damle RN, Batliwalla FM, Ghiotto F, Valetto A, Albesiano E, Sison C, et al. Telomere Length and Telomerase Activity Delineate Distinctive Replicative Features of the B-CLL Subgroups Defined by Immunoglobulin V Gene Mutations. *Blood* (2004) 103(2):375–82. doi: 10.1182/blood-2003-04-1345
69. Oakes CC, Seifert M, Assenov Y, Gu L, Przekopowitz M, Ruppert AS. DNA Methylation Dynamics During B Cell Maturation Underlie a Continuum of Disease Phenotypes in Chronic Lymphocytic Leukemia. *Nat Genet* (2016) 48:253–64. doi: 10.1038/ng.3488
70. Queiros AC, Villamor N, Clot G, Martinez-Trillos A, Kulis M, Navarro A, et al. A B-Cell Epigenetic Signature Defines Three Biological Subgroups of Chronic Lymphocytic Leukemia With Clinical Impact. *Leukemia* (2015) 29:598-605. doi: 10.1038/leu.2014.252
71. Gaiti F, Chaligne R, Gu H, Brand RM, Kothen-Hill S, Schulman RC, et al. Epigenetic Evolution and Lineage Histories of Chronic Lymphocytic Leukaemia. *Nature* (2019) 569(7757):576–80. doi: 10.1038/s41586-019-1198-z
72. Ng A, Chiorazzi N. Potential Relevance of B-Cell Maturation Pathways in Defining the Cell(s) of Origin for Chronic Lymphocytic Leukemia. *Hematol Oncol Clin North Am* (2021) 35(4):665–85. doi: 10.1016/j.hoc.2021.03.002
73. Elsner RA, Shlomchik MJ. Germinal Center and Extrafollicular B Cell Responses in Vaccination, Immunity, and Autoimmunity. *Immunity* (2020) 53(6):1136–50. doi: 10.1016/j.immuni.2020.11.006

74. Jenks SA, Cashman KS, Woodruff MC, Lee FE-H, Sanz I. Extrafollicular Responses in Humans and SLE. *Immunol Rev* (2019) 288(1):136–48. doi: 10.1111/imr.12741

Conflict of Interest: The following authors received funding from other sources as listed. However, these funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

MK is Chief Executive Officer, StationMD; AL has received unrestricted research grants and/or speaker fees from Roche-Genentech, AbbVie, and Janssen. NK serves on Advisory Boards for AbbVie, Astra Zeneca, Beigene, Behring, Cytomx Therapy, Dava Oncology, Janssen, Juno Therapeutics, Onco tracker, Pharmacyclics and Targeted Oncology, is a member of Data Safety Monitoring Committees for Agios Pharm, AstraZeneca, BMS–Celgene, Cytomx Therapeutics, Janssen, Morpho-sys, Rigel, and received research funding from AbbVie, Acerta Pharma, Bristol Meyer Squib, Celgene, Genentech, MEI Pharma, Pharmacyclics, Sunesis, TG Therapeutics, Tolero Pharmaceuticals. JRB has served as a consultant for Abbvie, Acerta/Astra-Zeneca, BeiGene, Bristol-Myers Squibb/ Juno/Celgene, Catapult, Eli Lilly, Genentech/Roche, Hutchmed, Janssen, MEI Pharma, Morphosys AG, Novartis, Pfizer, Pharmacyclics, Rigel; received research funding from Beigene, Gilead, Loxo/Lilly, SecuraBio, Sun, TG Therapeutics. MM has received research funding from Roche-Genentech, has had advisory roles for Abbvie, Astra Zeneca, Gilead, Janssen Pharmaceuticals, and Verastem Oncology. RR has received honoraria from AbbVie, AstraZeneca, Janssen, Illumina, and Roche. PG has served on advisory boards or as a consultant for AbbVie, AstraZenenca, BeIgene, BMS, Janssen, Lilly/Loxo Merck, Roche Sanofy, and received research support from AbbVie, AstraZenenca, and Janssen. JoCB has received research support from Pharmacyclics/Abbvie and AstraZeneca, Oncternal, and Velosbio, and has served on advisory boards for Pharmacyclics/ Abbvie, Beigene, AstraZeneca, TG Therapeutics, and MEI. SA owns stock in Bristol Myers Squibb and C4 Therapeutics and has served on the Advisory Boards of Stemline Therapeutics and Sanofi Genzyme. KS has received support from the Hellenic Precision Medicine Network in Oncology and the project ODYSSEAS funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014–2020) and co financed by Greece and the European Union, with grant agreement no: MIS 5002462. TK has had consultancy/advisory roles for AbbVie, Genentech-Roche, Gilead, Pharmacyclics LLC, an AbbVie Company, and Celgene; and research funding from AbbVie, Genentech-Roche, Pharmacyclics LLC, an AbbVie Company, and Oncternal. DN has stock in Madrigal Pharmaceuticals and receives research funding from Pharmacyclics LLC, an AbbVie Company. NC has received research funding from Verastem Oncology, argenx, Janssen Pharmaceuticals, and AVA Lifesciences GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 17 March 2023
DOI 10.3389/fonc.2023.1112879

OPEN ACCESS

EDITED BY
Dimitar G. Efremov,
International Centre for Genetic Engineering and Biotechnology, Italy

REVIEWED BY
Valter Gattei,
Aviano Oncology Reference Center (IRCCS), Italy
Harry W. Schroeder,
University of Alabama at Birmingham, United States

*CORRESPONDENCE
Nicholas Chiorazzi
NChizzi@Northwell.edu

SPECIALTY SECTION
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

RECEIVED 30 November 2022
ACCEPTED 23 February 2023
PUBLISHED 17 March 2023

CITATION
Vergani S, Bagnara D, Agathangelidis A, Ng AKY, Ferrer G, Mazzarello AN, Palacios F, Yancopoulos S, Yan X-J, Barrientos JC, Rai KR, Stamatopoulos K and Chiorazzi N (2023) CLL stereotyped B-cell receptor immunoglobulin sequences are recurrent in the B-cell repertoire of healthy individuals: Apparent lack of central and early peripheral tolerance censoring. *Front. Oncol.* 13:1112879. doi: 10.3389/fonc.2023.1112879

CLL stereotyped B-cell receptor immunoglobulin sequences are recurrent in the B-cell repertoire of healthy individuals: Apparent lack of central and early peripheral tolerance censoring

Stefano Vergani[1], Davide Bagnara[1], Andreas Agathangelidis[2,3], Anita Kar Yun Ng[1], Gerardo Ferrer[1], Andrea N. Mazzarello[1], Florencia Palacios[1], Sophia Yancopoulos[4], Xiao-Jie Yan[1], Jaqueline C. Barrientos[1], Kanti R. Rai[1], Kostas Stamatopoulos[2] and Nicholas Chiorazzi[1]*

[1]Karches Center for Oncology Research, The Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, United States, [2]Centre for Research and Technology Hellas, Institute of Applied Biosciences, Thessaloniki, Greece, [3]Department of Biology, School of Science, National and Kapodistrian University of Athens, Athens, Greece, [4]New York Genome Center, New York, NY, United States

Introduction: The leukemic cells of patients with chronic lymphocytic leukemia (CLL) are often unique, expressing remarkably similar IGHV-IGHD-IGHJ gene rearrangements, "stereotyped BCRs". The B-cell receptors (BCRs) on CLL cells are also distinctive in often deriving from autoreactive B lymphocytes, leading to the assumption of a defect in immune tolerance.

Results: Using bulk and single-cell immunoglobulin heavy and light chain variable domain sequencing, we enumerated CLL stereotype-like IGHV-IGHD-IGHJ sequences (CLL-SLS) in B cells from cord blood (CB) and adult peripheral blood (PBMC) and bone marrow (BM of healthy donors. CLL-SLS were found at similar frequencies among CB, BM, and PBMC, suggesting that age does not influence CLL-SLS levels. Moreover, the frequencies of CLL-SLS did not differ among B lymphocytes in the BM at early stages of development, and only re-circulating marginal zone B cells contained significantly higher CLL-SLS frequencies than other mature B-cell subpopulations. Although we identified CLL-SLS corresponding to most of the CLL major stereotyped subsets, CLL-SLS frequencies did not correlate with those found in patients. Interestingly, in CB samples, half of the CLL-SLS identified were attributed to two IGHV-mutated subsets. We also found satellite CLL-SLS among the same normal samples, and they were also enriched in naïve B cells but unexpectedly, these were ~10-fold higher than standard CLL-SLS. In general, IGHV-mutated CLL-SLS subsets were enriched among antigen-experienced B-cell subpopulations, and IGHV-unmutated CLL-SLS were found mostly in antigen-inexperienced B cells. Nevertheless, CLL-SLS with an IGHV-mutation status matching that of CLL clones varied among the normal B-cell subpopulations, suggesting that specific CLL-SLS could originate from distinct subpopulations of normal B

cells. Lastly, using single-cell DNA sequencing, we identified paired IGH and IGL rearrangements in normal B lymphocytes resembling those of stereotyped BCRs in CLL, although some differed from those in patients based on IG isotype or somatic mutation.

Discussion: CLL-SLS are present in normal B-lymphocyte populations at all stages of development. Thus, despite their autoreactive profile they are not deleted by central tolerance mechanisms, possibly because the level of autoreactivity is not registered as dangerous by deletion mechanisms or because editing of L-chain variable genes occurred which our experimental approach could not identify.

KEYWORDS

CLL (chronic lymphocytic leukemia), B cell development and differentiation, B cell repertoire, stereotyped antigen receptors, VDJ sequencing

Introduction

Chronic lymphocytic leukemia (CLL) is a disease characterized by the expansion of a $CD5^+$ B cell clone in the peripheral blood, bone marrow (BM), and secondary lymphoid tissues (1). The development of the disease strongly correlates with age, with a median age at diagnosis of ~70 years. The antigen receptor on the surface membrane of a B cell (BCR) plays a key role in the development and evolution of CLL as indicated by multiple studies (2–7). Most extraordinary among these studies is the remarkable similarity in the amino acid sequences of the antigen binding domains of the BCRs from certain CLL patients (8, 9). Analyses of large patient cohorts indicate that this is a recurrent feature in at least 40% of CLL clones (10). Indeed, patients can be divided into specific stereotyped subsets based on similarity in the VH CDR3 of the IGHV-IGHD-IGHJ (IGHV-D-J) rearrangement, and patients bearing discrete stereotyped BCRs can have unique clinical features and outcomes and have leukemic clones with distinct specific genomic aberrations (11, 12).

Based on the structural distinctiveness and clinical importance of BCRs in the disease, recombinant CLL IGs have been studied for antigen reactivity, revealing binding to a variety of exo- and auto-antigens (13–17). Documentation that reversion to the germline IGHV sequence converted certain exo-reactive to auto-reactive IGs (13, 14) led to the notion that CLL derives from an autoreactive B lymphocyte.

Because of the potentially harmful capacities of autoreactive clones for healthy people, evolutionarily a series of immunologic censoring mechanisms have evolved to eliminate or to reduce the avidity of autoreactive B cells during the early phase of development (18). In this regard, the existence of apparently "CLL-specific IGHV-D-J rearrangements" in the healthy B-cell repertoire is an important but relatively unstudied issue as current information is only available for mature circulating and splenic B cells (19–22). Whether such B cells from normal individuals, which would be expected to be self-reactive, are subjected to immunologic censoring mechanisms during development is not known.

Using a sensitive IGHV-D-J deep-sequencing approach (23), we sought to identify stereotyped IGHV-D-J rearrangements in B lymphocytes from healthy people at various stages of B-cell maturation In particular, considering the auto-reactive nature of CLL IGs, we set out to determine at which checkpoints such stereotyped rearrangements were triaged from the B-cell repertoires of normal individuals to maintain immune tolerance.

Our studies indicate that IGHV-D-J gene sequences resembling stereotyped CLL BCRs and belonging to one of the 29 major CLL stereotyped subsets are present in the normal B-cell repertoire. They are found at different sites, such as cord blood (CB), BM, and peripheral blood. Despite their autoreactive features, they do not appear to be purged during early B-cell development, the first checkpoints to sustain immunologic tolerance in the healthy setting.

Material and methods

Samples

The study was approved by the Institutional Review Board of Northwell Health. Bone marrow (BM) samples were collected as discarded bone segments from anonymized patients who had undergone joint replacement surgery. Persons with a history of any autoimmune disease or condition and of any cancer were excluded from the study. Peripheral blood and umbilical cord blood samples were similarly collected from anonymous healthy donors. Mononuclear cell (MC) fractions were separated by density gradient centrifugation (Ficoll, GE Healthcare), frozen (10% DMSO 45% FBS and 45% RPMI1640) and stored in liquid nitrogen until used.

Processing of BM samples

BM samples were placed in a large Petri dish containing cold PBS with 2.5% BSA, and the tissue was gently dissociated using the

plunger of a 60 ml sterile plastic syringe. Bone fragments were broken into small pieces using scissors and rinsed with the same buffer to extract cells from tissue niches. Cell suspensions were then passed through a 70uM cell strainer into a 50 ml tube. To optimize the yield, after processing Petri dishes were rinsed with buffer used for the dissociation, and the contents were added to the previously filtered suspension. BMMCs were separated by Ficoll density gradient centrifugation, frozen (10% DMSO 90%FBS), and stored in liquid nitrogen until used.

Isolation of various B-cell subpopulations by cell sorting

BM cell suspensions were incubated with V500 anti-CD19 and with PE-cy7 anti-CD10 mAbs (both BD Biosciences) for 20 minutes at 4°C, and after washing were sorted into $CD10^+$ and $CD10^-$ fractions. Non-B cells were excluded by using efluor-450 anti-CD3 and anti-CD16 mAbs, and dead cells were triaged by Sytox Blue (ThermoFisher) staining. The $CD10^+$ fractions were then additionally stained with FITC anti-CD34 (BD Biosciences), PE anti-IgM (Ebioscience) and eflour-450 anti-CD27 (Ebioscience) to further prohibit contamination with mature B cells. Pro B cells (PRO, $CD34^+IgM^-$), Pre B cells (PRE, $CD34^-IgM^-$) and immature B cells (IMM, $CD34^-IgM^+$) were collected.

$CD10^-$ fractions were also stained with PerCPcy5.5 anti-CD38 (BioLegend), FITC anti-IgD (ThermoFisher), APC anti-CD27 (BD Bioscience), and PE anti-CD24 (Bioscience) to discriminate naïve (NAÏVE, $CD24^+IgD^+CD27^-$) and memory (MEM, $CD24^+IgD^-CD27^+$) B cells and plasmablasts/plasma cells (PB/PC, $CD24^-CD38^{++}$).

PBMCs from normal blood donors were incubated with the following anti-human Abs for 20 minutes at 4°C: V500 anti-CD19 (BD Biosciences), PerCPcy5.5 anti-CD38 (BioLegend), PE-cy7 anti-CD24 (BioLegend), FITC anti-IgD (ThermoFisher), and allophycocyanin anti-CD27 (BD Bioscience), and then sorted to isolate Transitional (TRANS; $IgD^+CD27^-CD10^+CD38^+$), NAIVE ($CD27^-IgD^+$), recirculating Marginal Zone (rcMZ; IgD^+CD27^+ (24), MEM (IgD^-CD27^+) and double negative (DN; IgD^-CD27^-) B cells.

Total $CD19^+IgD^+CD27^-$ B cells were sorted from umbilical cord blood samples. In both cases, non-B cells were excluded with efluor-450 anti-CD3 and anti-CD16 mAbs, and dead cells barred with Sytox Blue (ThermoFisher) staining.

For all samples, B cells were sorted directly into 200µl PCR tubes containing 100µl Dynabeads Oligo(dT) (ThermoFisher) lysis buffer and stored at -80°C.

Library preparation and sequencing

mRNA isolation from B-cell lysates was performed in 96-well plates using Dynabeads Oligo(dT) (ThermoFisher) according to the manufacture's protocol. mRNA was used in its entirely for reverse transcription in 10 µl (50°C 1h, 72°C 10min) using SuperScript III Enzyme (ThermoFisher) in solid phase with Dynabeads Oligo(dT) as primer. After RNase H treatment, second-strand synthesis was performed (37°C 20 min, 98°C 30s, 62°C 2min and 72°C 10min) in solid phase in 10µl using Q5 Polymerase (NEB) and a mix of 13 primers covering all *IGHV* leader sequence segments reported in the IMGT database (25); primers contained a maximum of one mismatch, along with 13 to 16 random nt and partial Illumina adaptor sequences. Double-stranded cDNA was washed 3 times in 10mM tris-HCl to remove the remaining primers, and the entire sample was utilized as template for PCR amplification in 10 µl using Q5 Polymerase with universal FW primer and mix of reverse isotype specific primers (98°C 30s; 10 cycles at 98°C for 10s, 58°C for 15s, and 72°C for 1min; 72°C 10min). Two µl of the PCR product were used for a semi-nested PCR with inner RV primers for the constant region which also introduce partial Ilumina adaptors. This reaction was carried out in 20µl (98°C 30s; 15 cycles at 98°C for 10s, 58°C for 15s, and 72°C for 1min; 72°C 10min). The PCR product was purified with Ampure XP beads at a ratio of 1:1, and 1 - 10ng were used to add Illumina Indices with Nextera XT kit (Illumina). The MiSeq Illumina (v3 2 x 300 kit, Illumina MS-102-3003) was used to sequence the library. The library was loaded at 12pM with 10% PhiX [14].

9G4 Antibody labeling

The 9G4 rat anti-IGHV4-34 mAb (26) was labeled with Alexa Fluor ™ 488 according to the manufacturer's recommendations (Alexa Fluor™ 488 Antibody Labeling Kit, ThermoFisher). Briefly, antibody solution was mix with1/10^{th} of 1M sodium bicarbonate and then incubated with Alexa Fluor™ 488 dye for 1h at room temperature. After the recommended period, the solution was placed in the provided purification column and labeled antibody was collected from the flow through.

Analysis of IGHV4-34 IG heavy and light chain rearrangements in single B cells

To analyze the paired IG heavy and light chain IGHV-D-J and IGLVκJκ sequences of B cells expressing IGHV4-34, we used 10x methodology. After exposing normal Naïve, MZ, MEM, and DN B cell populations to the Alexa Fluor ™ 488-labeled 9G4 mAb, labeled cells were enriched by FACS (Figure S6A). For each sample, cells were washed and resuspended in 31.7 ul of PBS (0.04% BSA) immediately after sorting. Single-cell libraries were then generated using the Chromium Controller, Chromium Single Cell 5' Library & Gel Bead Kit v2 and i7 Multiplex Kit (10x Genomics, Pleasanton, CA, USA), according to the manufacturer's protocols. Target enrichment from cDNA was performed using the Chromium Single Cell V(D)J Enrichment Kit, Human B Cell (10x Genomics), followed by adaptor ligation. Enriched libraries were quantified on an Agilent Bioanalyzer High Sensitivity chip, and then sequenced on an Illumina Nextseq500 instrument (Illumina, San Diego, CA, USA) with the paired-end (2x150bp) mid output kit (300 cycles) according to manufacturer's protocols.

Bioinformatic analysis of immunoglobulin repertoire

For bulk VDJ-seq, processing of raw reads was performed using a custom workflow built with pRESTO (REpertoire Sequencing TOolkit) (27). The IGHV-D-J sequences obtained were submitted to IMGT/HighV-QUEST and analyzed using ChangeO and custom R scripts (23, 27).

Cellranger vdj pipeline was used to analyze sequencing data obtained from 10x Chromium V(D)J libraries.

Attribution of IGHV-IGHD-IGHJ rearrangements to stereotyped CLL subsets

IGHV-IGHD-IGHJ gene rearrangements were analyzed for similarity to stereotyped CLL BCRs using our established bioinformatics method (10). In more specific, the subsequent clustering criteria were applied: (i) utilization of IGHV genes belonging to the same phylogenetic clan, (ii) ≥ 50% amino acid identity and ≥ 70% similarity within the VH CDR3, (iii) equal VH CDR3 length and, (iv) identical offset of the common amino acid motif.

Satellite CLL-SLS, i.e. sequences with strong immunogenetic similarities with major CLL subsets, were identified using a purpose-built bioinformatics algorithm, which is based on a set of previously described parameters (10) (1): utilization of phylogenetically associated IGHV genes (2), maximum VH CDR3 length difference of 2 amino acids, and (3), presence of the "subset-specific" VH CDR3 sequence motifs with an offset of ± 2 amino acids. This analysis was performed individually for each major subset.

Statistical analyses

Statistical analyses and identification of outliers were performed in Graphpad Prism 9. Tests for statistical significance are described in figure legends for the relevant graphs.

Results

Identification of CLL stereotyped IGHV-D-J sequences in the B-cell repertoires of normal individuals

First, we asked if and to what extent B lymphocytes expressing BCRs closely resembling CLL-stereotyped BCRs exist in healthy people. To do so, we collected samples from CB of neonates (n=5), BM (n=11) of elderly people (≥ 70 years of age) who had undergone hip replacement surgery, and peripheral blood mononuclear cells (PBMC) of adult volunteers (35-60 years of age; n=16) (Figures S1A, B).

Total CD19$^+$ B cells were FACS isolated from CB. Whereas from BM we isolated several B-cell subsets representative of the distinct B-cell developmental stages, using a combination of surface membrane markers (Figure S1A).

From each of the 3 different cell sources (CB, BM, PBMC), we were able to identify B cells bearing an IGHV-D-J rearrangement that corresponded to that of a known CLL stereotyped subset. For convenience, we refer to this type of rearrangement found in B cells from normal individuals as a "CLL stereotype-like sequence" (CLL-SLS). Table S1 summarizes the number of unique CLL-SLS obtained for each cell population sorted from the various sites. We identified a total of 123, 513, and 999 CLL-SLS in CB, BM, and PBMC, respectively. The average frequencies were relatively comparable among the three (CB: 0.044%; BM: 0.037%; PBMC: 0.051%) (Figures 1A, S1C).

We then compared the distribution in BM of CLL-SLS at the various stages of B-cell development and among mature B-cell subsets. The highest frequency values were found among B-cell populations representative of the first stages of maturation: PRO-B, PRE-B, IMM, and NAIVE B cells (average frequencies: 0.046%, 0.048%, 0.031%, and 0.037%, respectively). The average frequency of CLL-SLS in MEM (0.023%) was lower, although these values did not reach statistical significance despite the number of total sequences queried being similar (Table S1). Interestingly, the PB/PC population harbored cells carrying CLL-SLS rearrangements at higher frequencies than immature/naïve stages (0.059%) (Figures 1B, S1D).

CLL-SLS were identified in each of the circulating B-cell subsets in PBMC: TRANS, NAIVE, rcMZ, MEM, and DN B cells (average frequencies: 0.047%, 0.047%, 0.076%, 0.032%, and 0.03%, respectively) (Figures 1C, S1B). CLL-SLSs were significantly higher in rcMZ compared to MEM and DN B-cell subsets (rcMZ vs MEM and DN, $P \leq 0.05$) (Figures 1C, S1E). Lastly, we did not find a difference of CLL-SLS frequencies between NAÏVE from BM and NAÏVE from PBMC (Figure 1D).

In summary, normal B cells expressing standard CLL-SLS exist in the repertoires of healthy subjects and are present at similar frequencies from tissues that dramatically differs in age. Moreover, CLL-SLS are found at all stages of B-cell development, but, interestingly, at significantly higher frequencies in the PB/PC compartment in the BM and significantly higher frequencies in the rcMZ B-cell subset isolated from PBMCs.

Together, these findings suggest that CLL-SLS are not triaged during first stages of B cell development by clonal deletion. Since our data are generated from whole B-cell subpopulations, we cannot determine if receptor editing could have occurred in these dells, which also would allow them to transit through B-cell maturation. In addition, their presence in antigen-experienced B cell subsets, e.g., MZ and MEM, and among antibody-secreting B cells in the BM, could represent class switch recombination (CSR) and SHM along with positive selection by a particular antigen.

10.3389/fonc.2023.1112879

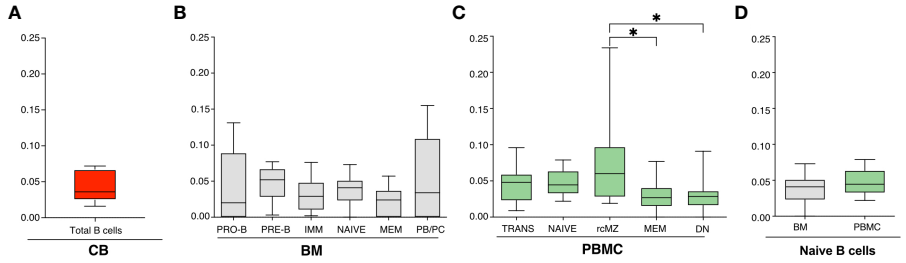


FIGURE 1
(A-C) Frequencies of CLL-SLS resembling standard stereotyped CLL BCRs identified in: **(A)** CB B cells (n=5); **(B)** BM B cell subsets (n=11); and **(C)** PBMC B cell populations (n=16). Statistical analyses were calculated with Kruskal-Wallis test. **(D)** Comparison of CLL-SLS in naïve B cells from BM and PBMC. Bars display minimum and maximum values. * $p \leq 0.05$.

Identification of satellite CLL-SLS in the B-cell repertoire of normal individuals

Next, we extended our analysis to IGHV-D-J sequences that resemble those referred to as "satellites" of known CLL stereotyped subsets (10). Satellite stereotyped sequences resemble the VH CDR3 motif of CLL stereotyped subsets but differ in certain amino acid residues at specific positions in the IGHV-D-J rearrangement or vary in VH CDR3 length. Although satellites are only in a minor component of the total number of stereotyped subsets found in CLL patients (10), we identified a ≥ 10-fold enrichment of "satellite" CLL-SLS compared to standard CLL-SLS in every tissue (CB: 3,545 vs 123; BM: 6,145 vs 513; PBMC: 11,707 vs 999; Tables S1, S2). Notably, the highest frequency of satellite CLL-SLS was measured in CB, where they reached an average frequency of 1.31% of the total IG sequence (Figures 2A, S2A).

When examining satellite CLL-SLS at the different stages of B-cell development, we were able to detect satellite sequences in every B-cell subset isolated from the BM (PRO-B: 0.27%; PRE-B: 0.42%; IMM: 0.40%; NAÏVE: 0.70%; MEM: 0.20%; PB/PC: 0.39%). Differently from the standard we observed a statistically significant enrichment in satellite CLL-SLS in BM NAÏVE compared to BM PRO-B (P = 0.0089) and NAÏVE to PBMC MEM (P = 0.0016) (Figures 2B, S2B).

Similarly, in PBMC samples, satellite CLL-SLS were present at discrete frequencies inn all the BB cell subsets analyzed but we only observed a significant difference in satellite CLL-SLS frequencies when comparing the average of NAÏVE to MEM and DN cells (NAÏVE: 0.798 vs. MEM: 0.37%, P = 0.0007; NAÏVE vs DN 0.61% vs MEM: 0.37%, P = 0.0011; Figures 2C, S2C).

Like what we observed in the case of standard CLL-SLS, we did not detect any differences in the frequencies of satellite CLL-SLS when comparing NAÏVE from BM and PBMC samples (Figure 2D).

Thus, we identified satellite CLL-SLS in the B-cell repertoires of healthy donors. However, unlike CLL patients, where satellite subsets comprised only 3% of the total cohort versus 13.5% assigned to major subsets (10), satellite CLL-SLS in normal individuals were found at higher frequencies and numbers than standard CLL-SLS. This difference might be the result of their immunogenetic properties together with the selection forces shaping the normal B cell repertoire.

Finally, the highest frequency of satellite CLL-SLS was found in NAÏVE B cells coming from BM and PBMC, in contrast to what we observed in the context of standard CLL-SLS where the highest frequencies were found in PB/PC and rcMZ. Thus, similarly to what observed for standard CLL-SLS, the presence of satellite CLL-SLS distributed along all B cell differentiation axis further strengthens the idea that CLL-like BCRs are not subjected to elimination by central tolerance mechanisms and are present in the normal B-cell repertoire.

Assignment of the CLL-SLS in normal, healthy people to specific, standard CLL stereotypes

The second tier of analysis was directed at understanding to which specific standard stereotyped CLL subset the CLL-SLS belong, and at determining if the distribution of the CLL-SLS differs from the standard stereotyped sequences observed in cohorts of CLL patient (Figures 3A–D). When comparing the distribution of standard CLL-SLS resembling the 29 most prominent subsets identified in patients with CLL across the 3 different cell sources (Figures 3B–D), we found at least one sequence belonging to each of the major CLL stereotyped subset from each site apart from subsets #16 and #7D3 (Table S3). When examining CB B cells, 39.8% of the CLL-SLSs were members of subset #14 and 13.01% of subset #73 (Figure 3B). Together these two subsets made up ~50% of all the CLL-SLS in the CB. Interestingly, these subsets were not the most prevalent ones among the major CLL subsets (Figure 3A). Finding these at increased frequencies in the CB, where most B cells have not encountered foreign antigen and have not undergone somatic hypermutation (SHM), was also surprising because, among CLL patients, these two subsets display mutated IGHVs. However, for subsets # 73 and 14, most of the CLL-SLS sequences were IGHV-unmutated (75% and 84%, respectively).

In the BM, the most recurrent subsets were # 148B, #14, #28A, #12 and #1 (15.8%, 12.3%, 12.1%, 7.4% and 5.1 respectively; Figure 3C). Like the CB, these abundant subsets were not the most frequently found in patients (Figure 3A). CLL B cells belonging to subsets #1 and

10.3389/fonc.2023.1112879

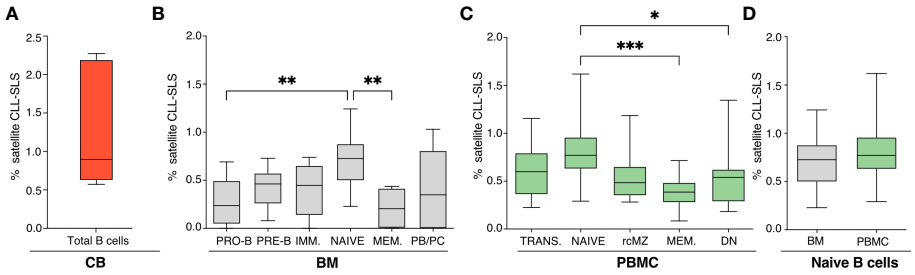
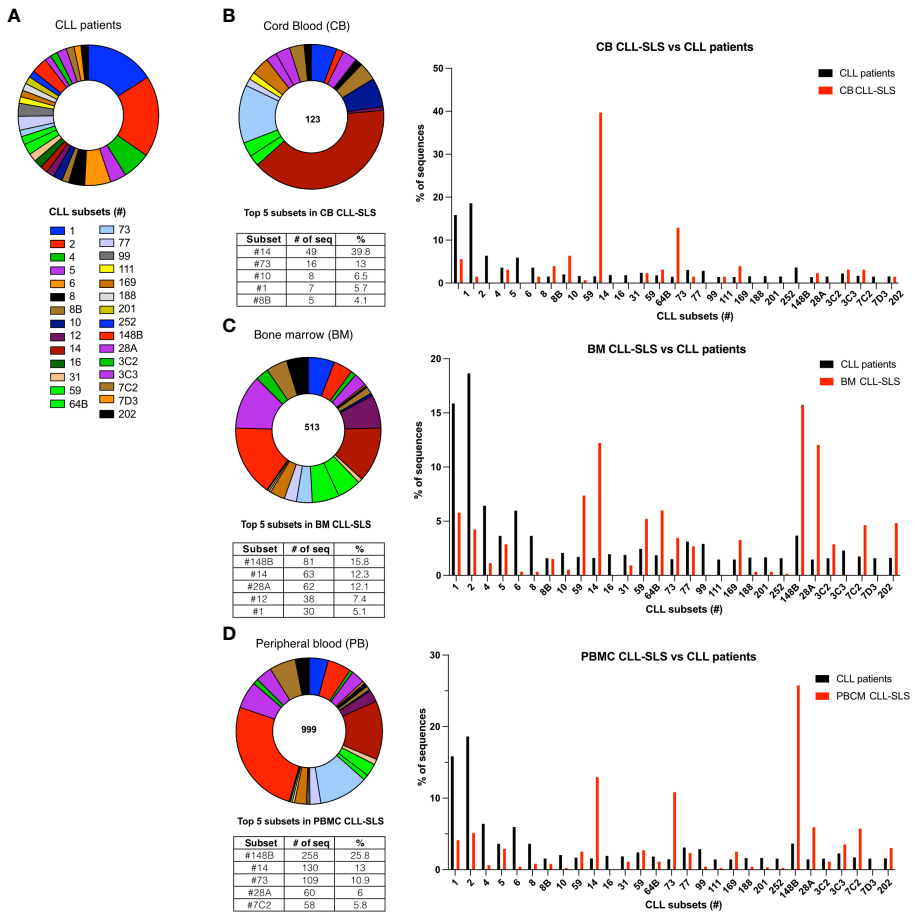


FIGURE 2
(A-C) Frequencies of CLL-SLS resembling satellite stereotyped CLL BCRs identified in: **(A)** CB B cells (n=5); **(B)** BM B cell subsets (n=11); and **(C)** PBMC B cell populations (n=16). Statistical analyses were calculated with Kruskal-Wallis test. **(D)** Comparison of CLL-SLS frequencies in naive B cells from BM and PBMC. Bars display minimum and maximum values. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

#28A are part of the IGHV-unmutated (U-CLL) group and are encoded by IGHV1-69 (subset #1) and IGHV1-2 (subset #28A), respectively. Conversely, CLLs falling into subsets #148B and #14 express IGHV2-5 and IGHV4-4, respectively, and belong to the IGHV-mutated (M-CLL) group. For those CLL-SLS identified among PBMCs, again subsets # 148b and #14 were highly represented (25.8% and 13%), followed by subsets # 73, #28A and #7C2 (10.9%, 6.0, and 5.8% of the total CLL-SLS) (Figure 3D).

Finally, when comparing among the three tissues, we found that BM and PBMC display a very similar distribution of CLL-SLS with



Subset	# of seq	%
#14	49	39.8
#73	16	13
#10	8	6.5
#1	7	5.7
#8B	5	4.1

Subset	# of seq	%
#148B	81	15.8
#14	63	12.3
#28A	62	12.1
#12	38	7.4
#1	30	5.1

Subset	# of seq	%
#148B	258	25.8
#14	130	13
#73	109	10.9
#28A	60	6
#7C2	58	5.8

FIGURE 3
(A) Frequency distribution of the 29 different major standard stereotyped subsets among CLL patients. Each pie slice identifies a specific subset in a different color. **(B)** Left panel: Frequency distribution of the standard CLL-SLS in the CB. Each pie slice identifies a specific subset following color code in **(A)** Table below the chart indicates the top 5 most frequently found standard CLL-SLS. Right panel: Comparison of the frequencies of major CLL stereotyped subsets in patients and in CLL-SLS from CB. **(C)** Left panel: Frequency distribution of the standard CLL-SLS in the BM. Right panel: Comparison between the frequencies of major CLL stereotyped subsets in patients and in CLL-SLS from BM. **(D)** Left panel: Frequency distribution of the standard CLL-SLS in the PBMC. Right panel: Comparison of the frequencies of major CLL stereotyped subsets in patients and in CLL-SLS from PBMC.

a statistically significant difference in the frequencies only for subset #148b (Figure S3A).

In conclusion, the distribution of certain standard CLL-SLS found in normal individuals can be dramatically different from the distribution found in CLL patients. However, since little is known about the immunogenetic properties of most of the subsets found frequently in the normal repertoire, it is difficult to speculate about the underlying biological process or force that preferentially selected for a B cell expressing a particular stereotyped sequence. In addition, we did not observe major differences in the distribution of CLL-SLS subsets when comparing BM and PMBC samples. Nevertheless, we noticed a discrete difference when comparing these to CB sample, where 50% of CLL-SLS are attributed to two subsets. Finding such high-level restriction in the CB is not necessarily paradoxical, given the absence of terminal deoxynucleotidyl transferase expression with the consequent lack of non-templated additions during the neonatal period, which often significantly limits diversity in the VH CDR3.

Assignment of CLL-SLS found in normal individuals to specific satellite CLL stereotypes

We next analyzed the distribution of satellite CLL-SLS as above. This identified several differences from that found for the standard CLL-SLS. For example, in CB, 50% of CLL-SLS were satellites of subsets # 77, #73, #31, #12, #2, and #169 (Figure 4A and Table S3). Within the BM, we also found that the most recurrent subsets were # 2, #169, #12, and #73 (17.5%, 12.9%, 8.8%, 9.1%, respectively; Figure 4B). Similarly, for the peripheral blood samples, subsets # 2 and #169 were among the most frequent (19.1% and 21.1%, respectively), making up to 40% of the total sequences (Figures 4C, S4A). It is noteworthy that the satellite sequence for subset #169 was found among the top-ranking group for each of the three sites of collection.

Furthermore, when we compared the frequencies of standard and satellites subsets, we observed two different scenarios. For some specific subsets we found a significant increase in satellite CLL-SLS compared to the standard ones. For example, #169 and its companion #2 comprise 30 to 40% of all the sequences in PBCM and BM whereas only a minor fraction of standard CLL-SLS were attributed to subset #2 and #169 in both tissues examined (BM: #2, 2.9% and #169, 2.5%; PBMC: #2, 4.8% and #169, 2.3%) (Figures 4D, E). However, we also found an opposite behavior for some of most frequent standard subsets that instead were underrepresented in the satellites, e.g., # 148b, 28A, and 14 (Figures 4D, E).

Overall, the satellite CLL-SLS subsets enriched in specific B-cell subpopulations can differ significantly from the standard CLL-SLS, especially in PBMC and BM. Among individual satellite subsets, # 2 and #169 seem to be over-represented in both PBCM and BM, and

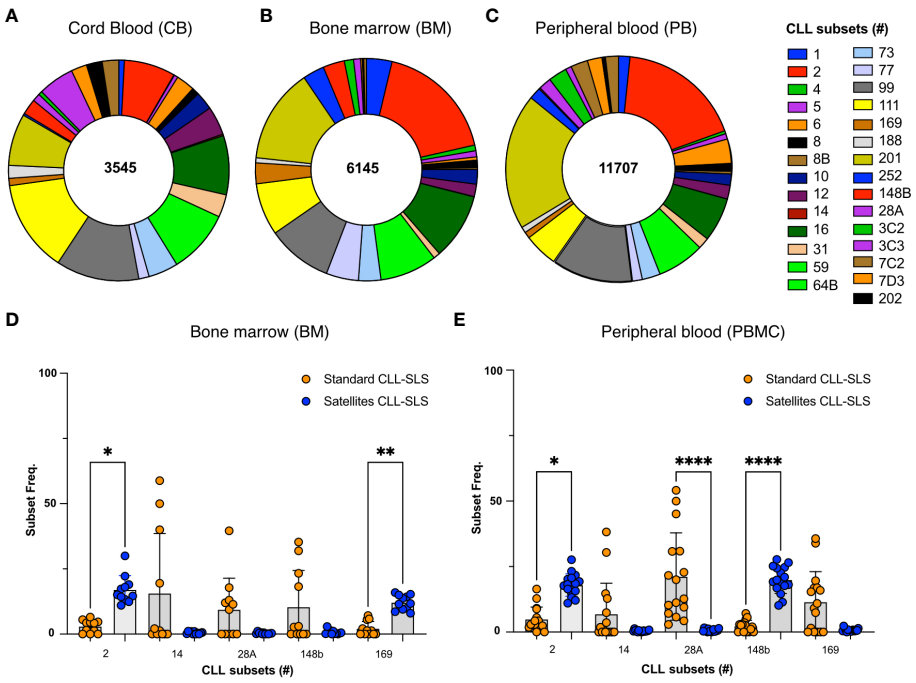


FIGURE 4
(A-C) Pie chart displays the frequency distributions of satellite CLL-SLS in **(A)** CB; **(B)** BM. Each pie slice identifies a specific satellite stereotyped subset following color code in the legend. **(D, E)** Plot summarizes the changes in frequencies of the depicted subsets between standard and satellite CLL-SLS in **(D)** BM and **(E)** PBMC. Statistical analyses were performed with Kruskal-Wallis test. * $p \leq 0.05$, ** $p \leq 0.01$, **** $p \leq 0.0001$, # = subset number.

they both significantly increase in frequencies when compared to their standard subset counterparts. Alternatively, some of the most frequent subsets in the standard CLL-SLS analysis were underrepresented in the satellite analysis. This last observation is in accord to what is observed in CLL patients, where satellite represent only a minor fraction of the total subset sequences. Instead, the significant increase in satellite CLL-SLS attributed to #2 and #169 suggest a selection pressure of these type of sequences compared to the standard counterpart that only occurs in the normal repertoire.

Distribution of CLL-SLS corresponding to IGHV-unmutated and IGHV-mutated CLL subsets among B-cell subpopulations at different stages of B-cell development

Like CLL clones (4), stereotyped subsets can be segregated based on IGHV-mutation status (10, 28). Therefore, we next determined if the U-CLL CLL-SLS were enriched in the earlier stages of B-cell development and if the M-CLL CLL-SLS were enriched in the later stages of B-cell maturation (Figures 5B, C). Since the incidence of standard CLL-SLS differed significantly from satellite CLL-SLS, for the following analysis, we only analyzed the former (Figure 5A).

Since we did not find proof for negative selection of any individual CLL-SLS at the early maturation stages in the BM and at the transitional stage in the blood (Figure 5B), we checked for such evidence at later stages of maturation, by examining specific patterns of distribution of individual CLL-SLS. This analysis revealed that CLL-SLS belonging to the IGHV-mutated subsets were found at different frequencies than the IGHV-unmutated among the various antigen-experienced B cell subsets (Figures 5B, C).

Using subsets #2 and #169 as examples of IGHV-mutated standard subsets, for subset #2, there was a statistically significant enrichment in NAIVE and rcMZ compared to B cells at the earlier stages of development; a significant level of different was not found with B cells at the later stages of maturation, although there was a trend in this regard. The principal was the same for subset #169, except there were also statistically significant differences for the MEM and DN. There was a similar trend for enrichment in PC/PB, but this did not reach statistical significance. Since subsets # 2 and #169 are members of the IGHV-mutated subtype, this pattern of distribution in the mature stages of B-cell development seems consistent. Similar trends were found for M-CLL subsets # 73, #77, and #188 (Figure S5A).

In contrast, those CLL-SLS expressing unmutated IGHVs were detected in B-cell subpopulations at the early stages of development and in TRANS and NAIVE; they were virtually absent from the IGHV-mutated MEM and rcMZ populations. Examples are CLL-SLS belonging to U-CLL subsets # 12 and 31 (among the most frequent ones, Figure 5D); a similar trend was observed for subsets # 8, 8B, and 59 (Figure S5B). Surprisingly, however, CLL-SLS attributed to U-CLL subsets were significantly enriched in DN B cells, an Ag-experienced, usually IGHV-mutated B cell subpopulation that plays a role in auto-reactive conditions and infectious disease (29). The reason for the abundance in this IGHV-mutated B-cell subpopulation is not obvious.

In summary, when examining individual CLL-SLS subsets, CLL-SLS bearing IGHV-mutated IGs are most frequent in naïve and Ag-experienced B cells, whereas CLL-SLS attributed to IGHV-unmutated subsets in general are restricted to the early stages of B-cell maturation and to naive B cells, with DN B cells being the exception.

Pattern of SHM found among CLL-SLS from normal, healthy people

Next, we examined if the CLL-SLS enriched at different B-cell maturation stages in PBMCs exhibited the same IGHV-mutation status as that found in standard CLL stereotyped subsets (Figures 6A, B).

First, we analyzed CLL-SLS attributed to U-CLL subsets, observing that ~94% of CLL-SLS found in TRANS and NAIVE B cells matched the IGHV-mutation status of the stereotyped subsets in CLL patients (Figure 6A). This situation changed when examining the more mature B-cell subsets in PBMC. In the case of MZ and DN, there was a fall in the frequencies of IGHV-unmutated CLL-SLS matching the IGHV-mutation status of the patient-defined CLL subsets (61% and 72%, Figure 6A). This drop was most evident for CLL-SLS in MEM B cells, where 72% displayed SHM and only 28% matched the original SHM status (Figure 6A).

When examining CLL-SLS belonging to M-CLL subsets (Figure 6B), we found the opposite: only a minor fraction present in the NAIVE and TRANS B-cell compartments displays SHMs and hence matches the IGHV-mutation status in CLL (8% and 5.7%, respectively). The majority in MEM (88.4%), MZ (68%), and DN (54.5%) were mutated and thus in agreement with CLL SHM status.

Thus, CLL-SLS that are IGHV-unmutated in patients are found more often in the normal repertoire among B cells at the earlier stages of B-cell maturation, which have usually not interacted with foreign antigens and therefore have not undergone SHM and developed IGV mutations. In contrast, CLL-SLS that are IGHV-mutated in patients are found more often in the normal repertoire among B-cell subpopulations at the later stages of B-cell maturation when SHM is common. So, in general, these results are consistent with U-CLL clones originating from and TRANS, NAÏVE, and MZ B-cell populations, and M-CLL clones coming from more antigen-experienced subsets such as MZ, MEM, and DN. However, the fact that some U-CLL-associated CLL-SLS can bear somatic IGHV mutations and can be enriched in antigen-experienced B cells suggests a positive selective driving away from the corresponding CLL-associated IGHV-mutation phenotype for these CLL-SLS.

We next expanded this type of analysis to individual CLL-SLS from the major CLL stereotyped subsets, examining the distribution patterns of U-CLL-like and M-CLL-like SLS sequences among the different B cell subpopulations (Figures 6C–F). This revealed 7 distinct patterns (Figures 6C–F). In the case of U-CLL subsets exhibiting pattern 1, the vast majority of IGHV-unmutated CLL-SLS belonging to subsets # 6, #7C2, #8, and #8B match the IGHV-mutation status of the CLL patients, independent of the B-cell subset in which they were found (Figures 6C, D). Regarding the other patterns, most were still characterized by high frequencies of

10.3389/fonc.2023.1112879

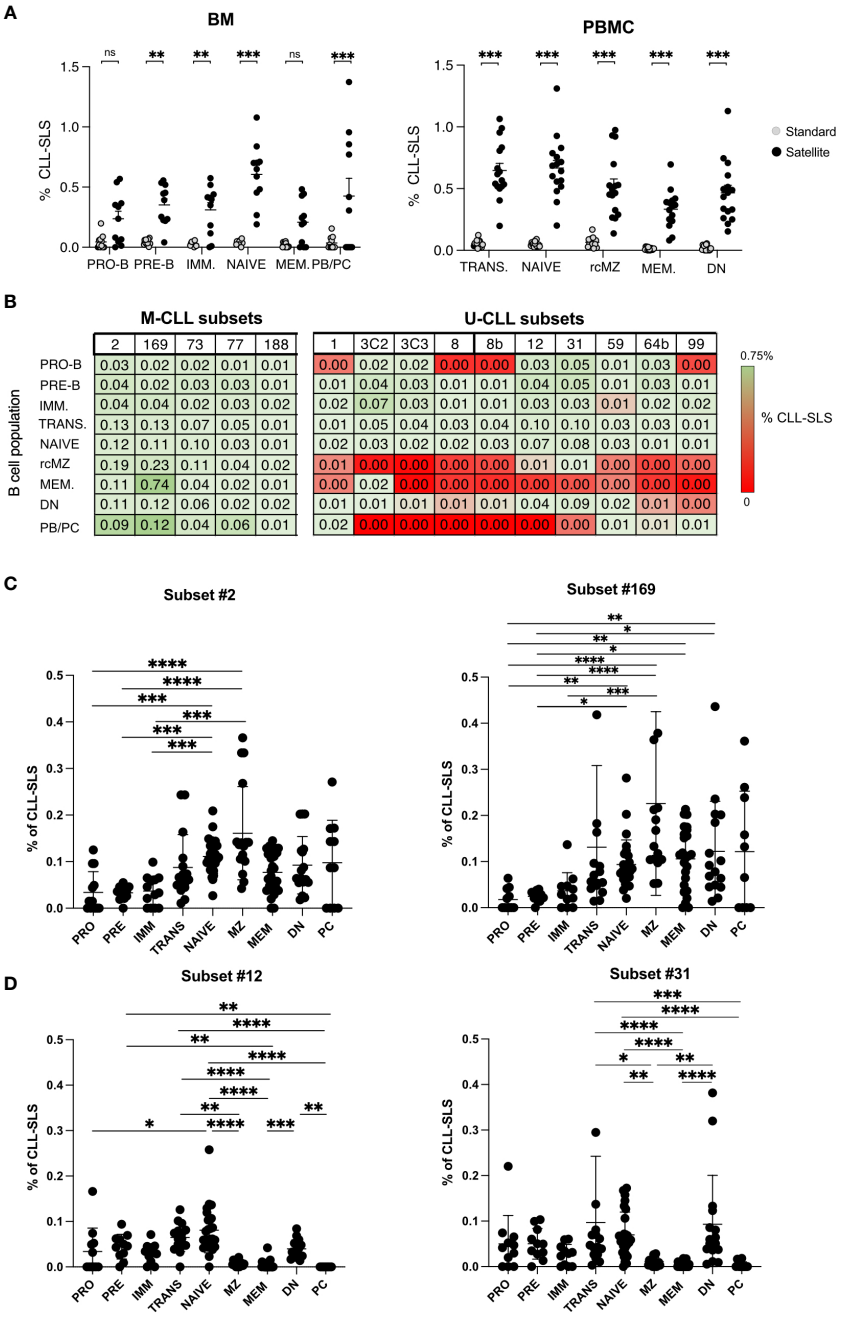
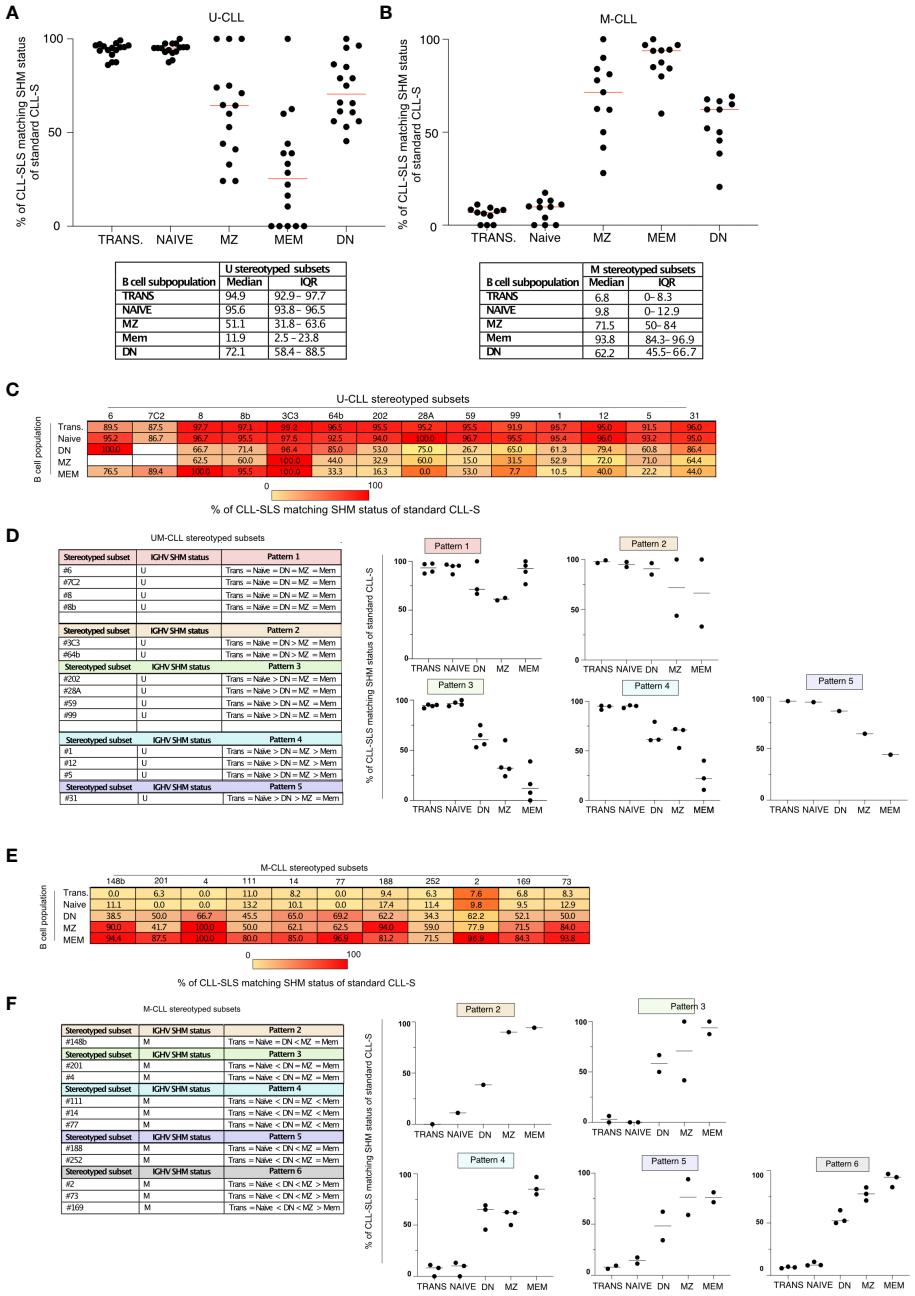


FIGURE 5
(A) Left Panel: Summary plot comparing frequencies of standard and satellites CLL-SLS in different B cell subsets in BM. Right Panel: Summary plot comparing frequencies of standard and satellites CLL-SLS in different B cell subsets in PBMC. **(B)** Upper panel: Table summarizes average frequency of the most frequent satellite CLL-SLS divided by subsets through different stages of B cell development. Lower panel: Table summarizes average frequency of the less frequent satellite CLL-SLS divided by subsets through different stages of B cell development. **(C)** Summary plot of satellite CLL-SLS frequencies attributed to subset #2 and #169 through different stage of B cell development (Mean with SEM). Statistical analysis and multiple comparisons were performed with Kruskal-Wallis test. **(D)** Summary plot of satellite CLL-SLS frequencies attributed to subset #12 and #31 through different stage of B cell development (Mean with SEM). Statistical analysis and multiple comparisons were performed with Kruskal-Wallis test.
* $p \leq 0.05$, **$p \leq 0.01$,***$p \leq 0.001$, ****$p \leq 0.0001$. ns, not significant.

CLL-SLS matching the SHM pattern of CLL patients; these were mainly restricted to the NAIVE and TRANS B cell stages. However, there were obvious differences in the other B-cell subsets. For example, in the case of subsets belonging to pattern 3 (subsets #202, #28A, #59, and #99), only a median of 65% DN, 32.9% MZ, and 7.7% MEM CLL-SLS matched the original CLL IGHV mutation status. Similarly, by looking at CLL-SLS attributed to M-CLL subsets, we observed that the most frequent ones found in our

10.3389/fonc.2023.1112879



B cell subpopulation	U stereotyped subsets Median	IQR
TRANS	94.9	92.9 – 97.7
NAIVE	95.6	93.8 – 96.5
MZ	51.1	31.8 – 63.6
Mem	11.9	2.5 – 23.8
DN	72.1	58.4 – 88.5

B cell subpopulation	M stereotyped subsets Median	IQR
TRANS	6.8	0 – 8.3
NAIVE	9.8	0 – 12.9
MZ	71.5	50 – 84
Mem	93.8	84.3 – 96.9
DN	62.2	45.5 – 66.7

FIGURE 6
(A) Summary plot of the frequencies of IGHV-unmutated CLL-SLS with matching U-CLL stereotyped subsets (median). **(B)** Summary plot of the frequencies of IGHV-mutated CLL-SLS with matching M-CLL stereotyped subsets (median). **(C)** Table with median frequencies of IGHV-unmutated CLL-SLS with matching U-CLL stereotyped subsets divided by subsets. **(D)** Left panel: CLL-SLS attributed to U-CLL subsets were grouped in pattern based on the matching frequencies. Right panel: Plots showing the median frequency of CLL-SLS matching U-CLL status. Each plot represents one of the patterns indicated in the left panel. **(E)** Table with median frequencies of IGHV-mutated CLL-SLS with matching M-CLL stereotyped subsets divided by subsets. **(F)** Left panel: CLL-SLS attributed to M-CLL subsets were grouped in pattern based on the matching frequencies. Right panel: Plots showing the median frequency of CLL-SLS matching M-CLL status. Each plot represents one of the patterns indicated in the left panel.

analysis, such as # 2, #73, and #169, group together in pattern 6 (Figures 6E, F). In this case, MEM display the highest median frequencies of CLL-SLS matching the original CLL IGHV mutation, whereas there was a progressive decline in both MZ and DN populations. In other cases, such as those subsets belonging to pattern 5, both MZ and MEM were similarly enriched in CLL-SLS matching the IGHV mutation status of the original CLL subsets (Figures 6E, F).

Thus, when examining the SHM status of individual CLL-SLS, different patterns can be identified. These patterns suggest that specific CLL stereotyped subsets might originate from particular subpopulations in the normal B-cell repertoire.

More precise documentation and assignment of CLL-SLS to normal B-cell subsets based on expression of IGKV and IGLV genes

The preceding data indicate that CLL-SLS are present in normal, healthy people, and that the frequencies at which these exist in the normal B-cell repertoire do not appear to decrease when progressing from developing to mature B cells. Since CLL B cells appear to derive from autoreactive precursors (13) and CLL IGs are often autoreactive (13–17), finding CLL-SLS in healthy individuals at all stages of maturation is not consistent with the elimination of autoreactive BCRs/IGs from the normal B-lymphocyte repertoire (18). Alternatively, the IGK/LV-J rearrangements paired with the CLL-SLS IGHV-D-J could differ from those in CLL cells, thereby neutralizing or preventing autoreactivity. Since several of the major stereotyped subsets display IGK/LV gene restrictions, we isolated cells based on membrane L chain expression (κ or λ) and defined the frequencies at which certain specific CLL-SLS IGHV-D-J rearrangements were found in the κ and λ chain populations.

As representatives of those subsets that display IGK/LV L chain restriction, we examined subset #2 and its companion subset #169, both of which always express lambda light chains encoded by IGLV3-21. Notably, when sorting NAIVE, MZ, and MEM cells from a normal, healthy person based on L chain isotype, we observed a similar distribution of CLL-SLS subsets # 2 and 169 within the κ and λ chain expressing B-cell populations. Thus, for these subsets there was a lack of skewing toward λ light chain use (Figure 7A).

We then looked at CLL-SLS subsets characterized by the expression of IGHV4-34, such as subsets #201, #77, and #4 (Figure 7B). For subset #201, which always uses IGLVλ1-44, we found CLL-SLS only in the λ-expressing fraction of MEM B cells. Whereas for subsets # 77 and 4, whose IGHVs always pair with IGLV10-54 or IGKV2-30, respectively, we did not observe a particular bias toward the usage of a specific light chain type (Figure 7A). Thus, we found the CLL-SLS in the appropriate light chain population depending on the specific subset. However, by taking this approach, we could not determine the specific IG light chain expressed by CLL-SLS.

To overcome this challenge, we performed single cell IGHV-D-J sequencing of B cells in PBMCs that express IGHV4-34 by sorting using the 9G4 mAb which reacts specifically with IGs bearing this gene (30). This strategy allowed us to enrich for subsets that use IGHV4-34 and, at same time, to identify the matching L chain and its DNA sequence. After sorting 20,000 NAIVE, 4,274 MZ, 9,547 MEM, and 340 DN B cells (Figure S6A), we identified IGHV-D-J sequences from 5,406 NAIVE, 1,205 MZ, 2,386 MEM, and 122 DN cells. A large majority of clonotypes identified expressed IGHV4-34, as depicted by V to J heatmap (Figure S6B). In this way, we identified CLL-SLS representative of several stereotyped subsets associated with IGHV4-34 (Figure 7C) in 110 Naïve, 2 rcMZ, 7 MEM, and 5 DN cells.

When looking specifically at CLL-SLS sequences resembling subset # 201, we identified 6 such sequences, all in naïve B cells. However, none were paired with a λ light chain (Figure 7D). However, although subset #201 CLL clones show a strong restriction for λ light chain use, there are a few identified instances where subset #201 stereotyped sequences were paired with the κ light chain gene IGKV4-1. Notably, one of our CLL-SLS sequences attributed to subset #201 had an IGKV4-1 gene partner. Moreover, the VK CDR3 sequence of that cell was remarkably like the CLL stereotype and the CLL-SLS (Figure 7D). Thus, this apparently normal B cell could be a precursor to a subset #201 CLL clone.

Similarly, when examining CLL-SLS resembling subset #4, we found an IGHV-D-J gene rearrangement paired with IGKV 2-30, the gene most often co-expressed in this leukemic subset (Figures 7E, F). However, this CLL-SLS did not bear IGHV mutations, which all subset #4 rearrangements have. Nor did it carry a characteristic amino acid at a specific position in the IGKV-J rearrangement corresponding to the standard stereotyped CLL BCRs, i.e., an aspartic acid at position 66 in the VK FR3 that is introduced by SHM in CLL cells. Thus, this apparent precursor of standard CLL stereotyped subsets # 4 does exhibit the complete subset #4 CLL sequence.

Discussion

Using our efficient IGHV-D-J sequencing approach that provides considerable depth of analysis (23), we demonstrated that CLL-SLS are present in B lymphocytes from normal individuals isolated from three sources that differ in B-cell composition and age. In line with a recent report finding such rearrangements in fetal liver-derived B cells (31), our studies indicate that CLL-SLS are present at the first stages of developmental time, in our instance, human cord blood. Since the median age of diagnosis of CLL is ~70, it might be expected that samples from aged individuals would contain higher frequencies of CLL-SLS. Notably, however, despite the age differences in the sites we sampled, the frequencies of CLL-SLS in the CB, PBMC and BM were similar, suggesting that CLL-SLS accumulation does not change with aging.

In addition, when focusing on the adult PBMC and BM repertoires, we did not find a fall in the frequency of normal B cells bearing CLL-SLS suggesting that censoring by central tolerance mechanisms had not occurred in the BM This was surprising since CLL IGHV-D-J rearrangements, including stereotyped rearrangements, generally derive from autoreactive B cells that normally would be eliminated (32, 33). This lack of censoring suggests that CLL-SLS do not recognize self-antigen with sufficient affinity to activate clonal deletion mechanisms. This conclusion, however, might be premature since, in the main, we analyzed solely IGHV-D-J rearrangements and not their accompanying IGKV-J and IGLV-J rearrangements, both of which are often needed for autoantigen binding. Thus, receptor editing of IGK/LV genes (18), another mechanism to maintain tolerance, could have taken place in B cell carrying CLL-like BCRs and could explain why CLL-SLS are relatively overabundant in normal people. Consistent with this possibility, certain major CLL stereotyped subsets use specific IGK/LV genes (34). So, this L chain

10.3389/fonc.2023.1112879

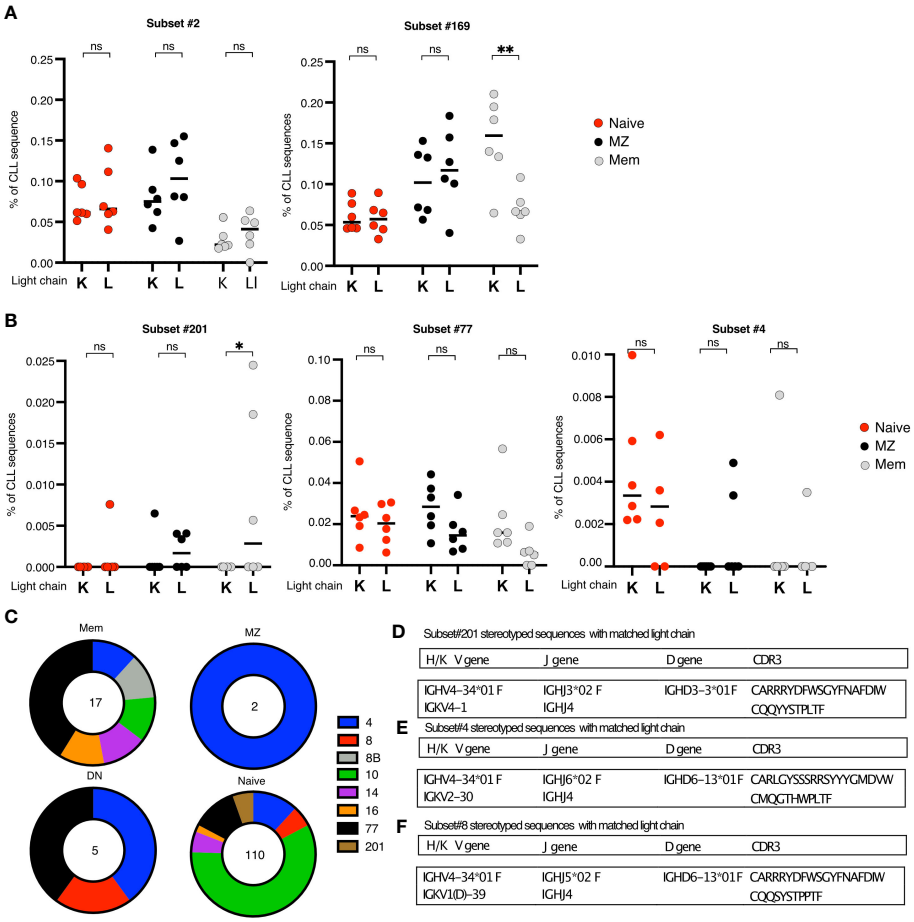


FIGURE 7
(A, B) Summary plots comparing the mean frequencies of CLL-SLS in the IG κ and IG λ fractions of the indicated B cell populations. Each plot refers to a particular CLL subsets. Statistical analysis and multiple comparisons were performed with Anova, Holm-Sidak test. **(C)** Pie charts of the frequency distribution of the CLL-SLS found after single cell RNA sequencing of the indicated B cell population. Each pie slice identifies a specific satellite subset following the color code in the legend. **(D-F)** Table of features of IGHV and matching IGK/LV attributed to **(D)**. subset #201; subset #4; Subset #8. *p ≤ 0.05, **p ≤ 0.01, ns, not significant.

feature might reflect the actions of an operative, normal tolerance mechanism and represent a level of negative selection for B cells bearing potentially harmful autoreactive BCRs in people without CLL. Finally, CLL-SLS might evade negative selection in the BM if the B cell carrying those rearrangement go through receptor revision of the H chain (35, 36).

We did find a fall in the frequencies of CLL-SLS among B-cell subsets in PBMC. Specifically, the highest level of such sequences was identified in NAIVE, TRANS, and rcMZ B cells with a decrease in the MEM and DN compartments. These observations suggest that CLL-SLS were not purged at the TRANS level, another point in early development where autoreactive B cells can be triaged from the repertoire (37, 38). Consequently, normal B cells bearing BCRs with acceptable, not intolerably high reactivity with autoantigens, could expand, possibly by tonic BCR signaling, and move into the mature B-cell pool (37–39). Hence, the lack of apparent negative selection at this stage would again be consistent with the CLL-SLS bearing lower affinity BCRs to autoantigens than registered by the immune system as dangerous.

However, the decrease in MEM and DN B cells suggests that the next set of immune tolerance mechanisms that prevent entry of unwanted, high affinity specificities B cells with into the more mature stages is effective in normal people. Hence, at least some CLL-SLS are prevented from engaging in germinal center-like responses. This could especially be the case for those CLL subsets using IGHV4-34, e.g., #4 and #201, since IGHV4-34-bearing normal B lymphocytes are usually excluded from GC reactions and prohibited to differentiate to antibody-secreting cells due their inherent autoreactive profile (40, 41). However, this does not totally exclude the possibility that processes such as SHM and CSR, occurring during GC reactions could redeem those potentially self-reactive CLL-SLS and allow them to mature to MEM and DN B cells. Lastly, CLL-SLS B cells that make it into the MEM and DN pool could become anergic, and therefore not increase numerically.

Finding cells with CLL-SLS BCRs in the MZ B subpopulation at frequencies higher than naïve could appear contrary to this principle. However, this might be explained by the features of the MZ B-cell subset. Indeed, even though IgM-expressing MZ B cells can be antigen-experienced, they can be generated through GC-independent process and they are poorly recruited to GC reaction and can be generated through a GC-independent process (42). Additionally, TRANS B cells can bypass the NAÏVE stage and differentiate in the MZ (43). Thus, high frequencies of CLL-SLS in the MZ B-cell compartment might be the combined result of a distinctive B-cell developmental route and of reduced involvement in GC reactions. This result differs from that reported recently, where CLL-SLS were present at significantly lower percentages in rcMZ B cells (21). This inconsistency might be due to different criteria and bioinformatics tools used to identify CLL-SLS.

In addition to those CLL-SLS attributed to standard stereotyped CLL subsets, we identified an unexpectedly high number of satellite CLL-SLS. Indeed, there was a ≥ 10-fold enrichment compared to standard CLL-SLS in every tissue. Interestingly, this is the opposite of what is seen in CLL patients, where satellites are only a minor component of the stereotyped subsets identified. A plausible explanation for the higher frequency of satellite CLL-SLS compared to the standard CLL-SLS is the more relaxed criteria used to identify satellite sequences. However, if that was the only factor involved, then the same observations would be also made in CLL. Hence, the frequency of standard and satellite CLL-SLS in our cohort could be the result of their immunogenetic properties together with selection forces shaping the normal B-cell repertoire.

Like standard CLL-SLS, satellite CLL-SLS were found at different frequencies throughout all stages of B-cell development with a significant enrichment in Naïve B cells in the BM and PBMC. This observation also suggests that satellite CLL-SLS are not subjected to negative selection during first steps of B-cell maturation. On the other end, a decrease in both MEM and DN B cells, is consistent with antigenic selection representing a barrier for B cells carrying CLL-like BCRs. Finally, this decrease might represent dilution of B cells bearing CLL-SLS in favor of positively selected non-CLL-SLS normal B cells by foreign antigen.

Regarding standard CLL-SLS, it was notable that those subsets found most recurrent in the CB, BM, and PB (#14, #73, #148b, #28A) were not those that are the most prominent in patients with CLL (#1, #2, #4, #6, and #8). Since there is not sufficient information available about the standard subsets found enriched in the normal repertoire, we can only propose that this distinction reflects a lesser necessity to remove or edit the former rearrangements and/or a greater need to remove the latter.

In this regard, it is noteworthy that the vast majority of CLL-SLS found in the CB are attributed to only two subsets (#14 and #73), both of which are in IGHV-mutated in CLL. However, most of the CLL-SLS sequences attributed to these two subsets are IGHV-unmutated in the CB.

Moreover, an uneven distribution of stereotyped subsets in the CB could reflect restrictions which, in many cases, are defined by unique combinations of IGHV, IGHD and IGHJ genes ('germline motifs') with less significant contribution by the IGHV-IGHD and IGHD-IGHJ gene junctions. Finding such high-level restriction in the CB is not necessarily paradoxical, given the absence of terminal deoxynucleotidyl transferase expression with the consequent lack of non-templated additions during the neonatal period, which can leads to severe limitations of diversity in the VH CDR3 (44, 45).

Of interest, when looking at satellite CLL-SLS, the findings were different. Strikingly, in both BM and PBMC, the most recurrent subsets identified were satellites of subsets #2 and #169, which fit into the set selected against in standard stereotyped instance, suggesting that negative selection for such satellite sequences had not occurred.

When assigning CLL-SLS to distinct normal B-cell subpopulations differing in foreign antigen experience based on IGHV mutations, we found some CLL-SLS predominate in subpopulations matching or not their IGHV-mutation status. For example, as expected, CLL-SLS attributed to the U-CLL type were mainly found in TRANS and NAIVE B-cell subpopulations (# 12, #31, #8, #8B and #59), whereas, unexpectedly, some CLL-SLS that do not carry mutations in CLL patients (#59 and #99) were highest in IGHV-mutated MEM cells. Likewise, CLL-SLS of the M-CLL type predominated in rcMZ, MEM, and DN (e.g., #201 and #4). Nevertheless, the majority of subset #2 and #169 CLL-SLS were IGHV-mutated and found in the MEM, and ~50% of the subset #2 and #169 CLL-SLS in the DN compartment exhibited somatically mutated BCRs. Thus, those B cells bearing BCRs that are discordant in IGHV-mutation status between the CLL setting and the normal setting would not be identified as the normal counterpart of leukemic clones. We can only speculate whether B cells carrying CLL-SLS that were found in M-CLL subsets but had discordant SHM status and were found in the NAÏVE compartment might represent a candidate precursor of CLL. Indeed, those cells have the potential to differentiate and accumulate SHMs and become identical to the original CLL counterpart.

As mentioned, several standard CLL-SLS can exhibit striking light chain gene sequence restriction. We took advantage of this issue by sorting B-cell subpopulations based on surface expression of κ or λ L chains and then asking if certain CLL-SLS, defined by the presence of either IGKV or IGLV genes in CLL, were enriched in normal B cells expressing that L chain isotype. Notably, for subsets # 2 and 169, we did not find such a restriction in L chain use as CLL-SLS attributed to these subsets were present in the λ-expressing and the κ-expressing fractions of the different B-cell populations sorted. Different, however, was the case for BCRs belonging to CLL subset #201, for which we found subset #201-like CLL-SLS only in the λ-expressing fraction of normal IgM memory B cells, consistent with the findings in patients with CLL. Thus, some CLL-SLS in normal B-cell populations express both the particular H and L chain subtypes reminiscent of a CLL cell and others do not. Subsets # 2 and 169, are examples of the latter and again suggest that these would not lead to CLL.

To formally address the possibility that a single normal B cell could express a CLL-SLS BCR IG carrying a single IGHV-D-J and IGKLV-J as found in patients with CLL, we performed single cell V

(D)J sequencing analysis of B-cell populations sorted for surface membrane expression of IGHV4-34 using the 9G4 mAb. When checking IGHV4-34$^+$ sequences bearing the subset #201 IGHV-D-J rearrangement, we did not find any of these paired with the expected λ gene. However, we did identify a subset #201 CLL-SLS along with a companion IGKV2-30 gene that is used in some CLL subset # 201 clones. However, this pair was only found in the naïve B-cell population, not in the more mature subsets that bear IGHV mutations as subset #201 usually does. Thus, finding a H-L BCR pair resembling the standard stereotype CLL BCR in the NAIVE but not in an antigen-experienced B-cell subset is consistent with effective peripheral tolerance censoring in normal individuals (40, 41).

Likewise, we identified other cells bearing the subset #4 CLL-SLS that were paired with the specific κ L chain gene rearrangement corresponding to that found in CLL. Interestingly, however, the expressed IGHV of this cell was not somatically mutated and did not carry a characteristic amino acid present at a specific position in the light chain variable region that is found in that standard stereotyped CLL BCR. Thus, in this instance, either there was a negative selection triaging against such specificities entering the mature B-cell repertoire or the antigenic drive needed to initiate these mutations did not occur in the normal setting.

These single cells analyses, thus, provided two examples of B lymphocytes within the B-cell repertoire of apparently normal people that differed in the potential to be a CLL precursor. The first example (CLL-SLS #201) was not consistent with this, suggesting either that the potential precursor was blocked from attaining or was negatively selected after attaining the canonical subset #4 CLL sequence. The second (CLL-SLS #4) is consistent with this finding and suggests that censoring of a CLL precursor does not necessarily occur. Single cell sequencing at depths greater than those we achieved will be necessary to determine which of these possibilities is correct.

Finally, finding certain, specific CLL-SLS in discrete normal B-cell populations raises the possibility that the final transformation event for that stereotyped subset occurred in that population or at that anatomic site. Thus, one could speculate that certain normal B-cell populations represent reservoirs in which specific stereotyped CLL clones are transformed. In this regard, SHM could act to control or promote CLL-SLS expression and transformation in the various mature B-cell repertoires. However, the possibility that transformation happens earlier but the transformed cells retain the ability to respond to specific types of antigens and to follow distinct maturation pathways, which lead to over or under abundance in distinct B-cell populations, cannot be excluded.

Collectively, our findings are consistent with CLL stereotypes not being sufficiently autoreactive to be censored by central and early (TRANS level) tolerance mechanisms, and therefore being permitted to enter the NAIVE subpopulation. Nevertheless, after arriving in the NAIVE repertoire, peripheral tolerance mechanisms for some CLL-SLS appear to restrict the number of cells entering the more mature B-cell repertoires in normal individuals, except possibly for the MZ. In patients with CLL, however, the later tolerance checkpoints might be faulty, allowing these CLL-SLS to be enriched in antigen experienced and memory B cells. In addition, some of those B cells with BCRs resembling those in CLL that do differentiate to MEM, DN, and PB/PC do not necessarily exhibit a H-L pairing, consistent with CLL, or differ in IGHV-mutation status or IG isotype from the CLL counterpart, thereby retaining tolerance constraints. The precision of our analysis at this level, however, is not sufficient to assert this with complete confidence. Thus, in patients with CLL, the effectiveness of receptor editing and GC reaction checkpoints might be reduced, allowing putatively dangerous H-L CLL pairing to occur and to be recruited into a GC response, where they can differentiate into antigen-experienced cells with or without accumulation of specific SHMs and isotypes.

Data availability statement

The data presented in the study are deposited in the SRA repository, accession number PRJNA931941 and PRJNA381394.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board of Northwell Health. The patients/participants provided their written informed consent to participate in this study.

Author contributions

SV, DB, and NC conceived the study. SV performed experiments and analyzed data. AA and KS carried out bioinformatic analyses. AN and GF performed the single cell IGHV-IGHD-IGHJ sequencing. DB, FP and AM helped with the experiments. X-JY and SY participated in manuscript preparation. KR and JB provided patient samples and clinical correlations NC directed the experiments. SV and NC wrote the initial and final versions of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

NC and KR received philanthropic support from The Nash Family Foundation, the Karches Foundation, The Marks Foundation, The Muriel Fusfeld Foundation, and the Jean Walton Fund for Leukemia, Lymphoma, & Myeloma Research. AM received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No 101023721.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2023.1112879/full#supplementary-material

References

1. Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. *N Engl J Med* (2005) 352(8):804–15. doi: 10.1056/NEJMra041720

2. Schroeder HWJr., Dighiero G. The pathogenesis of chronic lymphocytic leukemia: Analysis of the antibody repertoire. *Immunol Today* (1994) 15(6):288–94. doi: 10.1016/0167-5699(94)90009-4

3. Hashimoto S, Dono M, Wakai M, Allen SL, Lichtman SM, Schulman P, et al. Somatic diversification and selection of immunoglobulin heavy and light chain variable region genes in IgG+ CD5+ chronic lymphocytic leukemia b cells. *J Exp Med* (1995) 181(4):1507–17. doi: 10.1084/jem.181.4.1507

4. Fais F, Ghiotto F, Hashimoto S, Sellars B, Valetto A, Allen SL, et al. Chronic lymphocytic leukemia b cells express restricted sets of mutated and unmutated antigen receptors. *J Clin Invest.* (1998) 102(8):1515–25. doi: 10.1172/JCI3009

5. 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(6):1848–54. doi: 10.1182/blood.V94.6.1848

6. Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, et al. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. *Blood* (1999) 94(6):1840–7. doi: 10.1182/blood.V94.6.1840

7. Mazzarello AN, Gentner-Gobel E, Duhren-von Minden M, Tarasenko TN, Nicolo A, Ferrer G, et al. B cell receptor isotypes differentially associate with cell signaling, kinetics, and outcome in chronic lymphocytic leukemia. *J Clin Invest* (2022) 132(2). doi: 10.1172/JCI149308

8. Tobin G, Thunberg U, Karlsson K, Murray F, Laurell A, Willander K, et al. Subsets with restricted immunoglobulin gene rearrangement features indicate a role for antigen selection in the development of chronic lymphocytic leukemia. *Blood* (2004) 104(9):2879–85. doi: 10.1182/blood-2004-01-0132

9. Ghiotto F, Fais F, Valetto A, Albesiano E, Hashimoto S, Dono M, et al. Remarkably similar antigen receptors among a subset of patients with chronic lymphocytic leukemia. *J Clin Invest.* (2004) 113(7):1008–16. doi: 10.1172/JCI19399

10. Agathangelidis A, Chatzidimitriou A, Gemenetzi K, Giudicelli V, Karypidou M, Plevova K, et al. Higher-order connections between stereotyped subsets: implications for improved patient classification in CLL. *Blood* (2021) 137(10):1365–76. doi: 10.1182/blood.2020007039

11. Stamatopoulos K, Agathangelidis A, Rosenquist R, Ghia P. Antigen receptor stereotypy in chronic lymphocytic leukemia. *Leukemia* (2017) 31(2):282–91. doi: 10.1038/leu.2016.322

12. Sutton LA, Young E, Baliakas P, Hadzidimitriou A, Moysiadis T, Plevova K, et al. Different spectra of recurrent gene mutations in subsets of chronic lymphocytic leukemia harboring stereotyped b-cell receptors. *Haematologica* (2016) 101(8):959–67. doi: 10.3324/haematol.2016.141812

13. Herve M, Xu K, Ng YS, Wardemann H, Albesiano E, Messmer BT, et al. Unmutated and mutated chronic lymphocytic leukemias derive from self-reactive b cell precursors despite expressing different antibody reactivity. *J Clin Invest.* (2005) 115 (6):1636–43. doi: 10.1172/JCI24387

14. Catera R, Liu Y, Gao C, Yan XJ, Magli A, Allen SL, et al. Binding of CLL subset 4 b-cell receptor immunoglobulins to viable human memory b lymphocytes requires a distinctive IGKV somatic mutation. *Mol Med* (2017) 23:1–12. doi: 10.2119/molmed.2017.00003

15. Catera R, Silverman GJ, Hatzi K, Seiler T, Didier S, Zhang L, et al. Chronic lymphocytic leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. *Mol Med* (2008) 14(11-12):665–74. doi: 10.2119/2008-00102.Catera

16. Chu CC, Catera R, Zhang L, Didier S, Agagnina BM, Damle RN, et al. Many chronic lymphocytic leukemia antibodies recognize apoptotic cells with exposed nonmuscle myosin heavy chain IIA: Implications for patient outcome and cell of origin. *Blood* (2010) 115(19):3907–15. doi: 10.1182/blood-2009-09-244251

17. Chu CC, Catera R, Hatzi K, Yan XJ, Zhang L, Wang XB, et al. Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. *Blood* (2008) 112(13):5122–9. doi: 10.1182/blood-2008-06-162024

18. Nemazee D. Mechanisms of central tolerance for b cells. *Nat Rev Immunol* (2017) 17(5):281–94. doi: 10.1038/nri.2017.19

19. Colombo M, Bagnara D, Reverberi D, Matis S, Cardillo M, Massara R, et al. Tracing CLL-biased stereotyped immunoglobulin gene rearrangements in normal b cell subsets using a high-throughput immunogenetic approach. *Mol Med* (2020) 26(1):25. doi: 10.1186/s10020-020-00151-9

20. Forconi F, Potter KN, Wheatley I, Darzentas N, Sozzi E, Stamatopoulos K, et al. The normal IGHV1-69-derived b-cell repertoire contains stereotypic patterns characteristic of unmutated CLL. *Blood* (2010) 115(1):71–7. doi: 10.1182/blood-2009-06-225813

21. Bagnara D, Colombo M, Reverberi D, Matis S, Massara R, Cardente N, et al. Characterizing features of human circulating b cells carrying CLL-like stereotyped immunoglobulin rearrangements. *Front Oncol* (2022) 12:894419. doi: 10.3389/fonc.2022.894419

22. Muggen AF, de Jong M, Wolvers-Tettero ILM, Kallemeijn MJ, Teodosio C, Darzentas N, et al. The presence of CLL-associated stereotypic b cell receptors in the normal BCR repertoire from healthy individuals increases with age. *Immun Ageing.* (2019) 16:22. doi: 10.1186/s12979-019-0163-x

23. Vergani S, Korsunsky I, Mazzarello AN, Ferrer G, Chiorazzi N, Bagnara D. Novel method for high-throughput full-length IGHV-D-J sequencing of the immune repertoire from bulk b-cells with single-cell resolution. *Front Immunol* (2017) 8:1157. doi: 10.3389/fimmu.2017.01157

24. Weller S, Braun MC, Tan BK, Rosenwald A, Cordier C, Conley ME, et al. Human blood IgM "memory" b cells are circulating splenic marginal zone b cells harboring a prediversified immunoglobulin repertoire. *Blood* (2004) 104(12):3647–54. doi: 10.1182/blood-2004-01-0346

25. Giudicelli V, Chaume D, Lefranc MP. IMGT/GENE-DB: A comprehensive database for human and mouse immunoglobulin and T cell receptor genes. *Nucleic Acids Res* (2005) 33(Database issue):D256–61. doi: 10.1093/nar/gki010

26. Stevenson FK, Smith GJ, North J, Glennie MG, Hamblin TJ. Use of a cross reacting anti-idiotype to identify normal counterparts of neoplastic cells. *Nouv Rev Fr Hematol* (1978) 30(5-6):299–300.

27. Gupta NT, Vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, Kleinstein SH. Change-O: A toolkit for analyzing large-scale b cell immunoglobulin repertoire sequencing data. *Bioinformatics* (2015) 31(20):3356–8. doi: 10.1093/bioinformatics/btv359

28. Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, et al. Stereotyped b-cell receptors in one-third of chronic lymphocytic leukemia: A molecular classification with implications for targeted therapies. *Blood* (2012) 119 (19):4467–75. doi: 10.1182/blood-2011-11-393694

29. Li Y, Li Z, Hu F. Double-negative (DN) b cells: An under-recognized effector memory b cell subset in autoimmunity. *Clin Exp Immunol* (2021) 205(2):119–27. doi: 10.1111/cei.13615

30. Potter KN, Li Y, Pascual V, Williams RCJr., Byres LC, Spellerberg M, et al. Molecular characterization of a cross-reactive idiotope on human immunoglobulins utilizing the VH4-21 gene segment. *J Exp Med* (1993) 178(4):1419–28. doi: 10.1084/jem.178.4.1419

31. Roy A, Bystry V, Bohn G, Goudevenou K, Reigl T, Papaioannou M, et al. High resolution IgH repertoire analysis reveals fetal liver as the likely origin of life-long, innate b lymphopoiesis in humans. *Clin Immunol* (2017) 183:8–16. doi: 10.1016/j.clim.2017.06.005

32. Cashman KS, Jenks SA, Woodruff MC, Tomar D, Tipton CM, Scharer CD, et al. Understanding and measuring human b-cell tolerance and its breakdown in autoimmune disease. *Immunol Rev* (2019) 292(1):76–89. doi: 10.1111/imr.12820

33. Reijm S, Kissel T, Toes REM. Checkpoints controlling the induction of b cell mediated autoimmunity in human autoimmune diseases. *Eur J Immunol* (2020) 50 (12):1885–94. doi: 10.1002/eji.202048820

34. Hadzidimitriou A, Darzentas N, Murray F, Smilevska T, Arvaniti E, Tresoldi C, et al. Evidence for the significant role of immunoglobulin light chains in antigen recognition and selection in chronic lymphocytic leukemia. *Blood* (2009) 113(2):403–11. doi: 10.1182/blood-2008-07-166868

35. Wilson PC, Wilson K, Liu YJ, Banchereau J, Pascual V, Capra JD. Receptor revision of immunoglobulin heavy chain variable region genes in normal human b lymphocytes. *J Exp Med* (2000) 191(11):1881–94. doi: 10.1084/jem.191.11.1881

36. Chen C, Nagy Z, Prak EL, Weigert M. Immunoglobulin heavy chain gene replacement: A mechanism of receptor editing. *Immunity* (1995) 3(6):747–55. doi: 10.1016/1074-7613(95)90064-0

37. Loder F, Mutschler B, Ray RJ, Paige CJ, Sideras P, Torres R, et al. B cell development in the spleen takes place in discrete steps and is determined by the quality of b cell receptor-derived signals. *J Exp Med* (1999) 190(1):75–89. doi: 10.1084/jem.190.1.75

38. Carsetti R, Kohler G, Lamers MC. Transitional b cells are the target of negative selection in the b cell compartment. *J Exp Med* (1995) 181(6):2129–40. doi: 10.1084/jem.181.6.2129

39. Petro JB, Gerstein RM, Lowe J, Carter RS, Shinners N, Khan WN. Transitional type 1 and 2 b lymphocyte subsets are differentially responsive to antigen receptor signaling. *J Biol Chem* (2002) 277(50):48009–19. doi: 10.1074/jbc.M200305200

40. Pugh-Bernard AE, Silverman GJ, Cappione AJ, Villano ME, Ryan DH, Insel RA, et al. Regulation of inherently autoreactive VH4-34 b cells in the maintenance of human b cell tolerance. *J Clin Invest.* (2001) 108(7):1061–70. doi: 10.1172/JCI200112462

41. Cappione A 3rd, Anolik JH, Pugh-Bernard A, Barnard J, Dutcher P, Silverman G, et al. Germinal center exclusion of autoreactive b cells is defective in human systemic lupus erythematosus. *J Clin Invest.* (2005) 115(11):3205–16. doi: 10.1172/JCI24179

42. Martin F, Kearney JF. Marginal-zone b cells. *Nat Rev Immunol* (2002) 2(5):323–35. doi: 10.1038/nri799

43. Tull TJ, Pitcher MJ, Guesdon W, Siu JHY, Lebrero-Fernandez C, Zhao Y, et al. Human marginal zone b cell development from early T2 progenitors. *J Exp Med* (2021) 218(4). doi: 10.1084/jem.20202001

44. Zemlin M, Bauer K, Hummel M, Pfeiffer S, Devers S, Zemlin C, et al. The diversity of rearranged immunoglobulin heavy chain variable region genes in peripheral blood b cells of preterm infants is restricted by short third complementarity-determining regions but not by limited gene segment usage. *Blood* (2001) 97(5):1511–3. doi: 10.1182/blood.V97.5.1511

45. Marshall AJ, Doyen N, Bentolila LA, Paige CJ, Wu GE. Terminal deoxynucleotidyl transferase expression during neonatal life alters D(H) reading frame usage and ig-receptor-dependent selection of V regions. *J Immunol* (1998) 161 (12):6657–63. doi: 10.4049/jimmunol.161.12.6657

Frontiers in Oncology

Advances knowledge of carcinogenesis and tumor progression for better treatment and management

The third most-cited oncology journal, which highlights research in carcinogenesis and tumor progression, bridging the gap between basic research and applications to imrpove diagnosis, therapeutics and management strategies.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact