Molecular characterization of the microenvironment in CLL-like monoclonal B cell lymphocytosis and early-stage chronic lymphocytic leukemia

Gonzalo Blanco Ares

TESI DOCTORAL UPF / 2017

DIRECTORA DE LA TESI

Dra. Blanca Espinet Solà

Laboratori de Citogenètica Molecular
Servei de Patologia
Hospital del Mar (Barcelona)

Grup de Recerca Translacional en Neoplàsies Hematològiques (GRETNHE)
Programa de Recerca en Càncer
Institut Hospital del Mar d'Investigacions Mèdiques (IMIM)

DEPARTAMENT DE CIÈNCIES
EXPERIMENTALS I DE LA SALUT

upf. Universitat Pompeu Fabra Barcelona

AGRADECIMIENTOS / ACKNOWLEDGEMENTS

Es bonito ver que ese proyecto que empezaste hace casi cuatro años llega a su fin, y que lo que en un principio era una página en blanco es ahora una interesante historia que contar. Siento una gran satisfacción al ver que toda la implicación y el esfuerzo durante estos años han dado sus frutos, pero soy consciente de que no sería posible sin el soporte profesional y humano de muchas personas. Me alegra tener esta oportunidad para mostrarles mi agradecimiento a todas ellas.

La historia comienza allá por enero de 2013, cuando tuve la oportunidad de conocer, por aquel entonces, a la Dra. Espinet. Trabajar con Blanca no sólo me permitió aprender de ella, sino que me puso delante a una gran mentora que me abrió la puerta a muchas oportunidades, profesionales y personales. Muchas gracias, Blanca, por tu generosidad, y por encontrar siempre un hueco para mí por muy apretada que tengas la agenda. Pero el "equipo LLC" no está completo sin Anna, que, como buena "becaria mayor", me enseñó todo lo necesario para desenvolverme en el laboratorio, desde la técnica del "churrito" hasta complicadísimos trucos de Excel, y siempre con una infinita paciencia, esperando a que yo volviese de "mi mundo". Muchas gracias, Blanca y Anna, por todos los buenos momentos de estos años (no sólo los de Sídney o Nueva York, sino los de todas esas reuniones que se acababan prolongando horas).

Gracias a mi "becaria pequeña", Andrea, por ser mi apoyo en el día a día, por todo el trabajo que me quitaste de encima (esta tesis no sería posible sin ti) y por los buenos ratos de oficina y cafés. A las demás compañeras del despacho 164: las antiguas (Alba y Sílvia), las de paso (Aida), y las nuevas generaciones (Silvia y Laura), por los buenos momentos. Gracias, Andrea y Silvia, por ocuparos de gran parte del trabajo común en los últimos meses para que yo me pudiese dedicar a la tesis.

Al resto de compañeros de Citogenética: Marta, Sergi, María y Carme, por ser la alegría de la huerta. Sergi y Marta, qué falta nos hace que nos vengáis a revolucionar el gallinero de vez en cuando. Gracias, Mary, por estar siempre de tan buen humor, y Carme, per les teves classes accelerades de català!

A Lara, Eulàlia, Marta y Magda, para mí ya inmortalizadas como "las chicas del SAM", súper heroínas que te salvan de cualquier apuro. Gracias por vuestra energía positiva y por todo vuestro trabajo y ayuda, y por el chocolate que os suelo robar a última hora del día con mi llave maestra. Siempre recordaré nuestros debates "científicos e intelectuales" de la hora de la comida.

Quiero agradecer también al resto de compañeros del hospital: Moltes gràcies, Lourdes y Dra. Woessner, por ser dos grandes ejemplos, y Ana,

por ofrecerme siempre tu disposición y ayuda, y por ser la encargada oficial de revisar los aspectos clínicos, también los de esta tesis. Gracias a Xavi, Leo y Lluís, por vuestro entusiasmo y por lo que he podido aprender de vosotros; Xavi, no me olvido de las cosas que tenemos pendientes. A Ramón, por tus comentarios críticos. A Bea, por dejarme siempre un huequecito en tu campana, y a Rosa, por dejarme otro hueco en tu centrífuga. Gracias a Asun, Montse, Yolanda, Raquel, Dulce y demás compañeros del hospital que amenizan la rutina diaria.

Gracias también a Eugenia y a Eva, por hacer posible este proyecto proporcionado las muestras de los pacientes, y a Alicia, Mary, Laia, Sara y demás "hematolocas", por los buenos momentos compartidos. A Conchi, por tu entusiasmo en las sesiones bibliográficas y tu "apoyo audiovisual". Al resto de miembros de los Servicios de Hematología y Anatomía Patológica del Hospital del Mar de los que he tenido la oportunidad de aprender, al Servicio de Hematología del Hospital Vall d'Hebron por ayudarnos en la recolección de las muestras de pacientes y a los miembros de los grupos cooperativos que trabajan en LLC/LBM con los que he colaborado. También quería agradecer al Servicio de Inmunología del Hospital CU Virgen de la Arrixaca de Murcia por su provechosa contribución a este estudio.

Gracias al personal del IMIM por toda su gestión y por facilitarme la vida y a la Fundació IMIM por la ayuda en la encuadernación de la tesis.

Besides, I am really grateful to Kostas, for giving me the great opportunity to stay at his lab in Thessaloniki and for all the help and support on this journey. Many thanks to Anna as well, for that unexpected meeting in which she taught me so many interesting things about the TCR, and for all her effort over the last years. I am also very grateful to Liz and Despoina, who showed me the Greek way of life (I miss having a good bougatsa in the morning), and to Aliki, Anastasia, Evangelia and all the other colleagues from Thessaloniki who were helpful with my needs.

Later, I was lucky enough to have another stay at the Feinstein Institute in New York. I really thank Nick for allowing me to be a part of his excellent team, and for giving free rein to my scientific "fantasies"; we both agree that imagination is essential for any advance in science. Many thanks to Barbara, Gerardo, Paulina, Flor, Joy, Yasmine and all the other colleagues from New York who made this project possible. I also want to add Jesús, Manolo and Santi, for making my free time at Feinstein pleasant, and Freya and Cristina, for taking good care of me.

Pero más allá de las amistades surgidas en el trabajo, no me puedo olvidar de todas las demás personas que también hicieron posible esta tesis, empezando por los amigos de nuestra "Asociación", mi familia de Barcelona. No sería posible acabar con éxito este proyecto sin todos esos fines de semana de desconexión, risas y cervezas. Gracias, en concreto,

Rita, por haber decidido compartir tu día a día conmigo (y escribiendo esto me acuerdo de Ibiza y me río) y a todos los buenos amigos del instituto y de la época universitaria que seguís ahí, con Estrella Galicia fresquita y nuevas historias que contar, cada vez que vuelvo a casa. Como diriades vós, “seica Gonzaliño xa rematou a tese!”. Tampoco me quiero olvidar de los genetistas de la UB ni de los amigos que hice en Londres y Nueva York.

Muchas gracias a la abuela, tíos, primos, madrina y demás familia, por empeñaros en que coma más y en que “tú estás más alto” cada vez que me veis; sois los responsables de que volviese siempre a Barcelona con las pilas cargadas para seguir dándole duro. Gracias especiales a Elenita, que con su felicidad innata me alegró desde la distancia las duras tardes escribiendo tesis. Y a Chema, por recibirme siempre con tu “¡va, venga, Gon, que te invito a una caña!” y ganarte a pulso la categoría de hermano mayor.

A mis hermanas, por todas las veces que os ocupasteis de mis cosas para que yo pudiese estudiar, trabajar, y en resumen, acabar terminando esta tesis. Pero, sobre todo, por ser tan geniales y haber influido tan positivamente en mí, y hacer que me acuerde de ello cada vez que me veo el tobillo.

Ao meu sobriño, que aterrará neste mundo case ao mesmo tempo que esta tese. Pensar na súa chegada deume moitas forzas para seguir nos momentos máis duros desta recta final. Ao seu pai, polos bos momentos compartidos nesta viaxe.

Y, sobre todo, gracias a mis padres, por respetarme y animarme siempre en todas las decisiones que tomo, y decirme cada vez que siento que se complican las cosas, “acuérdate de aquella vez que también pensabas que no podías y al final todo salió bien”.

Quería agradecer finalmente a todos los pacientes que desinteresadamente ofrecen sus muestras (sin las cuales nada de esto sería posible) y a La Fundación La Caixa y la Sociedad Española de Hematología y Hemoterapia por permitirme desarrollar esta tesis doctoral con todo el soporte necesario.

A todos, agradeceros el esfuerzo y el tiempo empleados para conseguir que esta historia tuviese un final feliz.

ABSTRACT

The analysis of the microenvironment in CLL-like monoclonal B cell lymphocytosis (MBL) and early-stage chronic lymphocytic leukemia (CLL) is relevant for understanding the natural history of CLL. To this end, a total of 58 MBL, 54 early-stage CLL and 31 healthy subjects were extensively characterized by various immunological and molecular methods. Purified $CD4^+$ and $CD8^+$ mononuclear cells from peripheral blood were subjected to gene expression studies and T cell receptor (TR) repertoire analysis, whereas cytokine immunoassays were performed in serum samples. Gene expression studies in $CD4^+$ cells revealed increased cytotoxic and inflammatory pathways, which were higher in MBL than in early-stage CLL. Gene dysregulation was not remarkable in $CD8^+$ cells. Increased serum levels of cytokines such as IL8, IFNγ and TNFα were also observed in MBL, while early-stage CLL generally displayed lower cytokine levels, especially amongst cases bearing somatically hypermutated IGHV genes. TR analysis demonstrated oligoclonality in both entities with persisting T cell clones over time and increasing clonality within $CD4^+$ T cells concurrently with the expansion of neoplastic B cells. Besides, identical T cell clonotypes were identified in different MBL/CLL cases. All these findings implicate inflammatory processes and antigenic elements in the immune background of CLL, whose effects are significantly altered during progression from MBL to CLL.

RESUMEN

El estudio del microambiente en la linfocitosis B monoclonal (LBM) de tipo LLC y en estadios iniciales de la leucemia linfática crónica (LLC) tiene gran relevancia para entender la historia natural de la enfermedad. Con este objetivo se caracterizaron 58 casos de LBM, 54 de LLC en fases iniciales y 31 sujetos sanos. Se analizó la expresión génica y el repertorio del receptor de células T (TR) en fracciones de células mononucleares $CD4^+$ y $CD8^+$ purificadas a partir de sangre periférica. Además, se realizaron inmunoensayos para medir niveles de citoquinas en suero. Los estudios de expresión génica revelaron patrones citotóxicos e inflamatorios aumentados en células $CD4^+$, superiores en LBM con respecto a LLC en fases iniciales. En las células $CD8^+$ no se observó ninguna disfunción remarcable en la expresión génica. Se detectaron niveles aumentados de citoquinas como IL8, IFNγ y TNFα en sueros de sujetos con LBM, mientras que en LLC en fases iniciales los niveles de citoquinas fueron generalmente inferiores, principalmente debido a los casos con hipermutaciones del gen IGHV. El análisis del TR mostró la existencia de oligoclonalidad en ambas entidades y de clones T persistentes en el tiempo, así como niveles de clonalidad en la fracción T $CD4^+$ que aumentan conjuntamente con la expansión de las células B malignas. Asimismo, se identificaron clonotipos comunes en diferentes casos con LBM/LLC. Todos estos hallazgos implican un papel clave de los procesos inflamatorios y de los elementos antigénicos desde las etapas más tempranas de la enfermedad, cuyos efectos varían notablemente durante la progresión desde LBM a LLC.

PREFACE

Chronic lymphocytic leukemia (CLL) is the most common leukemia of adults in Western countries. Over the last decades, extensive research has been performed in order to elucidate why some patients remain asymptomatic for long periods of time whereas others experience rapid progression and aggressive disease. Concurrently with advances in the knowledge about the molecular mechanisms of the disease, different therapy options have been developed, improving response rates and life expectancy. However, CLL still remains incurable.

In recent years, several investigations reported that interactions between CLL cells and tumor microenvironment are crucial for disease development. On the other hand, a pre-leukemic condition of CLL named CLL-like monoclonal B cell lymphocytosis (MBL) was formally defined in the 2008 Classification of Tumors of the Haematopoietic and Lymphoid Tissues of the World Health Organization as the presence of circulating CLL-like cells below the cut-off value ($<5.0\times10^9$/L) required for establishing the diagnosis of CLL, in the absence of other clinical features of the disease. Although this entity attracted great scientific interest, the role of microenvironmental interactions in MBL remains uncharacterized, although it obviously appears as highly relevant for understanding the molecular mechanisms underlying the etiology of CLL as well as potentially the indolent course of non-progressive CLL clones.

In the present study, an increased inflammatory response was demonstrated in MBL. Strikingly, this inflammation was decreased in early CLL, revealing a different pathophysiology between both entities. Besides, the identification of antigen-mediated T cell

restrictions highlights the role of antigens from the first stages of the disease.

All these findings offer novel insights into MBL and CLL, which will allow the development of future investigations addressed to completely understand the natural history of the disease and, ultimately, to develop pre-emptive strategies.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. Chronic lymphocytic leukemia

1.1.1. General aspects

Chronic lymphocytic leukemia (CLL) is a lymphoproliferative disorder characterized by the proliferation and accumulation of mature, typically $CD5^+$ B cells within the peripheral blood, bone marrow, lymph nodes and other lymphoid tissues (Rozman et al, 1995). It is the most common type of leukemia of adults in Western countries and 95% of cases are diagnosed in individuals aged 50 years or older, with a median age at diagnosis of 71 years (Miller et al, 2016). Clinical features are very variable. Although most patients are asymptomatic, some of them present with fatigue, autoimmune hemolytic anemia, infections, splenomegaly, hepatomegaly, lymphadenopathy or extra-nodal infiltrates (Swerdlow et al, 2008). The median survival time from diagnosis is around 10 years; however, it presents a highly variable clinical course. Whereas some cases are stable and asymptomatic, other patients undergo a rapidly progressive and aggressive disease. This heterogeneity has an impact on clinical approaches and determines survival times from diagnosis, ranging from some months to a normal life span (Cramer et al, 2011). Its etiology is still considered unknown, although the presence of restricted B cell receptors (BCR) argues in favor of a limited set of antigens which promote division of precursor cells and clonal evolution (Fais et al, 1998; Agathangelidis et al, 2012).

1.1.2. Diagnosis

The 2008 World Health Organization (WHO) classification includes CLL in the category of mature B cell lymphoid neoplasms (Swerdlow et al, 2008). As most patients are asymptomatic at presentation, CLL is often diagnosed incidentally after routine blood tests. However, it is essential to perform an accurate diagnosis and verify that the patient has CLL and not another mature B cell neoplasm. In order to achieve this, the blood count, blood smear and phenotype of the peripheral lymphoid cells must be evaluated (Hallek et al, 2008).

1.1.2.1. B cell counts and morphology

A requirement for CLL diagnosis is the detection of at least 5×10^9/L clonal B cells (5000/μL) in the peripheral blood for a minimum period of three months. In addition, the blood smear typically shows small and mature lymphocytes with scanty cytoplasm and densely compacted chromatin (Hallek et al, 2008) (Figure 1).

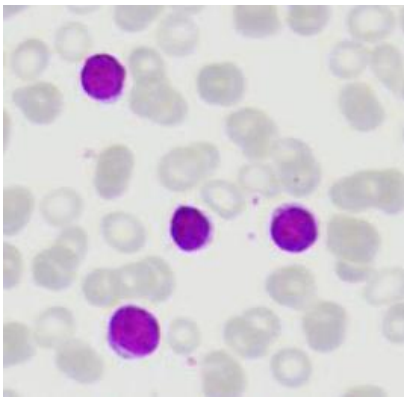

Figure 1. Peripheral blood smear showing typical B cells found in CLL.

1.1.2.2. Immunophenotype

The clonality of the B cells needs to be verified by flow cytometry, disclosing a specific phenotype. In particular, CLL cells display an

aberrant expression of the T cell antigen CD5, besides the B cell surface markers CD19, CD20, CD23 and CD79b. The levels of surface immunoglobulins (sIg), as well as those of CD20 and CD79b, are lower than in normal B cells. A restriction of either kappa or lambda immunoglobulin light chains must be also identified (Ginaldi et al, 1998; Hallek et al, 2008). Some cases may have an atypical immunophenotype (for instance, CD5-, CD23- or strong sIg) (Swerdlow et al, 2008).

A variant of CLL in which clonal B cells with the same immunophenotype are mainly found in solid lymphoid tissues ($<5\times10^9$/L in peripheral blood) and with no cytopenias due to bone marrow infiltration is referred as small lymphocytic lymphoma (SLL). It implies the presence of lymphadenopathy and/or splenomegaly; thus, histopathologic evaluation of the affected tissues is required (Hallek et al, 2008; Swerdlow et al, 2008).

1.1.3. Prognostic factors

1.1.3.1. Clinical staging

The clinical heterogeneity of CLL prompted the development of two independent clinical risk stratification approaches, the Rai and Binet staging systems (Rai et al, 1975; Binet et al, 1981). They are based on different clinical parameters, summarized in Table 1, which can be easily assessed in routine clinical practice and allow the stratification of patients in distinct risk categories: low (median survival: 10 years), intermediate (median survival: 5-7 years) and high risk (median survival: 2-3.5 years) (Table 1).

RAI	Clinical features	Risk
0	Lymphocytosis only	Low
I	Lymphocytosis with lymphadenopathy	Intermediate
II	Lymphocytosis with splenomegaly and/or hepatomegaly (with or without adenopathy)	
III	Anemia (hemoglobin <11 g/dL)	High
IV	Thrombocytopenia (platelets <100x10^9/L)	
BINET		
A	No anemia or thrombocytopenia, <3 enlarged lymphoid areas	Low
B	No anemia or thrombocytopenia, ≥3 enlarged lymphoid areas	Intermediate
C	Anemia (hemoglobin <10 g/dL) and/or thrombocytopenia (platelets <100x10^9/L)	High

Table 1. Rai and Binet staging systems and risk stratification (Rai et al, 1975; Binet et al, 1981).

1.1.3.2. Serum and phenotypic markers

Many studies in the past decades identified several serum and immunophenotypic markers with associated prognostic value, among them:

- Serum markers: Elevated serum levels of beta-2 microglobulin (B2M), lactate dehydrogenase (LDH), thymidine kinase and soluble CD23 were reported to predict an adverse prognosis in CLL (Han et al, 1985; Reinisch et al, 1994; Hallek et al, 1996; Hallek et al, 1999). More recently, increased serum levels of the cytokines CCL3 and CCL4 were also correlated to aggressive disease (Burger et al, 2009; Sivina et al, 2011).
- Immunophenotypic markers: Overexpression of phenotypic markers such as CD38, ZAP-70 and CD49d were also associated with a poor prognosis in patients with CLL (Damle et al, 1999; Crespo et al, 2003; Gattei et al, 2008).

1.1.3.3. IGHV gene somatic hypermutation status

Somatic hypermutations within the variable domain of the immunoglobulin genes constitute a physiological process during B cell development. Circulating mature B cells that are exposed to antigens migrate to germinal centers of peripheral lymphoid tissues, where somatic hypermutations occur (Jaffe et al, 2008). In CLL, mutations within the immunogloibulin heavy variable (IGHV) genes are closely linked to the biology of tumor B cells with a relevant impact on prognosis. In 1999, two independent groups first reported that CLL patients with mutated IGHV (M-CLL, >2% nucleotide differences from the most similar germline gene) usually show an indolent disease, as opposed to CLL cases with unmutated IGHV (U-CLL, ≤2%), associated with a poor prognosis (Damle et al, 1999; Hamblin et al, 1999). The molecular basis underlying these distinct clinical outcomes is related to the BCR signaling pathways: CLL B cells with mutated IGHV display attenuated BCR signaling and less proliferation. On the other hand, cases with unmutated IGHV activate key signal transduction pathways in response to BCR stimulation, which confers a higher proliferative potential (Zenz et al, 2010). In this regard, U-CLL has been correlated to increased CD38 and ZAP-70 expression levels (Damle et al, 1999; Crespo et al, 2003) (Figure 2).

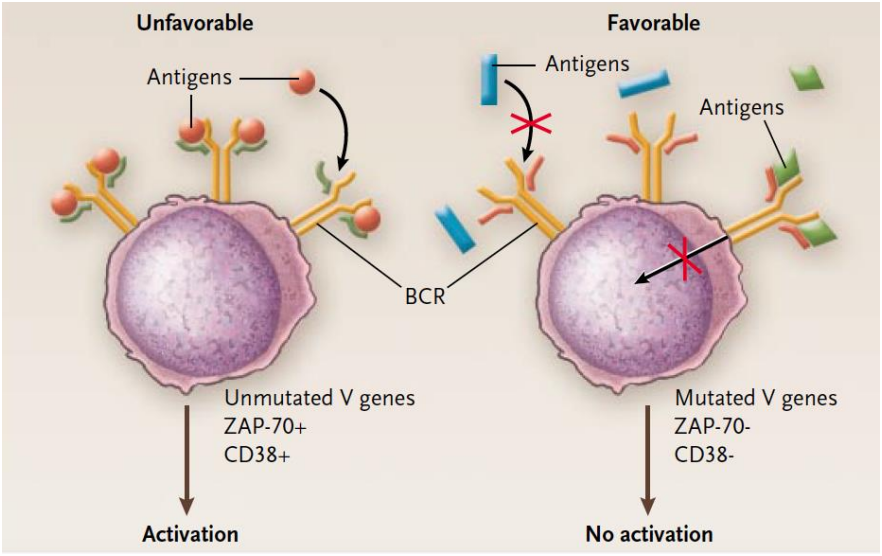

Figure 2. Association between the IGHV mutational status, the BCR signaling pathway and the prognosis (adapted from Chiorazzi et al, 2005).

The mutational status is not the only feature of the IGHV gene related to prognosis in CLL. An important restriction of the heavy chain complementary determining region 3 (CDR3) sequences (the BCR region mainly involved in antigen binding) is observed in CLL, with approximately one third of cases harboring stereotyped BCR. Remarkably, some of these subsets of patients with quasi-identical BCR represent distinct subgroups associated with specific clinical courses and outcomes (Agathangelidis et al, 2012; Agathangelidis et al, 2014). For instance, IGHV3-21 gene usage confers an aggressive disease and poor prognosis independently of the mutational status when belonging to subset #2 (Baliakas et al, 2015).

1.1.3.4. Genetic aberrations

The identification of cytogenetic subgroups (Döhner et al, 2000) defined by fluorescence *in situ* hybridization (FISH) that are associated with distinct prognosis constituted another turning point in the prognostic stratification of CLL patients. The Döhner hierarchical model established four chromosomal aberrations evaluated by FISH, namely deletions of 11q, 13q, 17p as well as trisomy 12, which have been the gold standard to determine prognosis in routine practice. Poor clinical outcome is associated with cases that show deletion of 17p and 11q, whereas a good prognosis is correlated to cases harboring deletion of 13q as a sole abnormality. Finally, trisomy 12 cases display an intermediate risk (Döhner et al, 2000; Van Dyke et al, 2016).

Besides cytogenetic abnormalities, mutations in several genes were also associated with prognosis in CLL. Thus, mutations within *TP53*, *NOTCH1*, *BIRC3*, *SF3B1* and *ATM* genes were reported to confer an inferior prognosis in CLL (Puente et al, 2011; Rossi et al, 2013; Nadeu et al, 2016). Combination of mutations and cytogenetic aberrations resulted in an integrated model to stratify CLL patients in different risk groups (Rossi et al, 2013) (Figure 3).

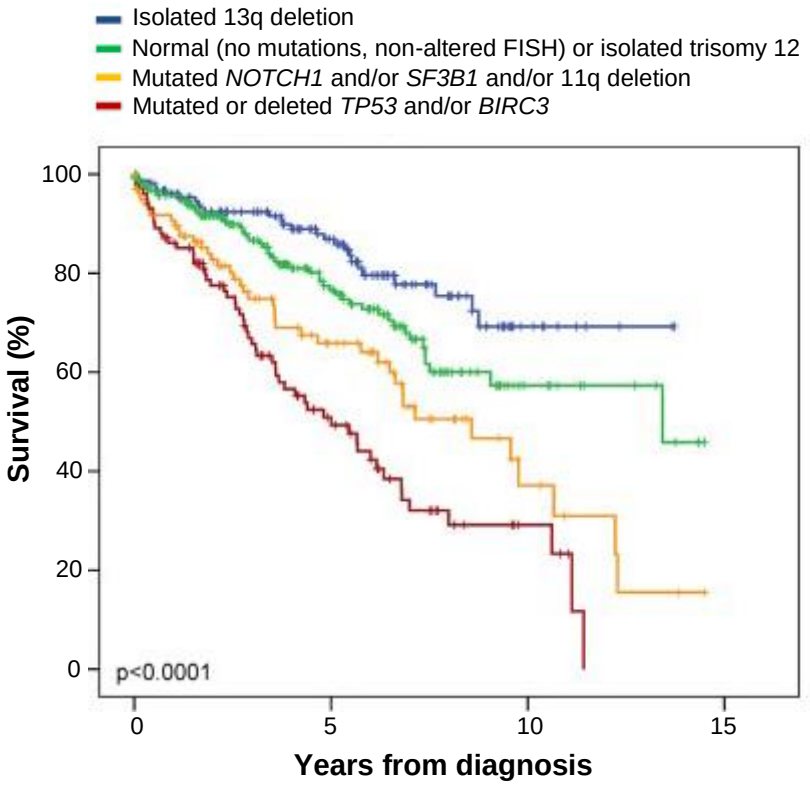

Figure 3. Overall survival in CLL according to the integrated mutational and cytogenetic model proposed by Rossi et al in 2013 (adapted from Rossi et al, 2013).

Finally, the identification of complex karyotypes, defined as the presence of three or more chromosomal abnormalities by conventional cytogenetics, has recently gained traction as a useful prognostic tool in CLL, since it confers poor clinical outcome even in *TP53* altered cases (Delgado et al, 2014; Blanco et al, 2016; Puiggros et al, 2017).

Some other parameters correlated to poor prognosis in CLL, as well as the most relevant prognostic factors previously detailed, are summarized in Table 2.

Older	Predict for poor prognosis
Clinical staging	Advanced (Rai III-IV, Binet C)
Lymphocyte doubling time (LDT)	<6 months (when lymphocyte counts >30×10^9/L)
Bone marrow infiltration	Diffuse pattern
Serum beta-2 microglobulin	Elevated (≥3.5 mg/L)
Serum lactate dehydrogenase	Elevated (≥2x normal)
Serum thymidine kinase	Elevated (>10 U/L)
ZAP-70 expression	High (≥20%)
CD38 expression	High (≥30%)
IGHV gene mutational status	U-CLL (≤2% of mutations)
FISH cytogenetics	17p or 11q deletion
Newer	
CD49d expression	High (≥30%)
Gene mutations	*TP53, NOTCH1, BIRC3, SF3B1* or *ATM*
Conventional cytogenetics	Complex karyotype
Stereotyped BCR	Specific subsets (#1, #2, #8)
Micro-RNA 155 expression	High
CCL3 and CCL4 in serum	Elevated levels

Table 2. Older and newer prognostic factors in CLL (adapted from Rai et al, 2016).

1.1.4. Treatment

1.1.4.1. Treatment criteria

The biological heterogeneity of CLL cases results in a broad range of clinical presentations and evolutions. Thus, approximately one third of all CLL patients will never require treatment. Those patients newly diagnosed and with early-stage disease (Rai 0 or Binet A) and even intermediate risk patients (Rai I-II or Binet B) should be monitored without treatment until they become symptomatic or show evidence of progressive disease, since there is no proof of benefit after an early therapy intervention. Criteria for treatment

administration in CLL include (Hallek et al, 2008; Wierda et al, 2017):

- Anemia and/or thrombocytopenia.
- Massive (>6 cm), progressive or symptomatic splenomegaly.
- Massive (>10 cm), progressive or symptomatic adenopathy.
- Lymphocyte doubling time (LDT) of less than 6 months or a lymphocytosis increase of more than 50% after a two-month period.
- Autoimmune anemia and/or thrombocytopenia unresponsive to corticosteroids.
- Presence of disease-related symptoms (B-symptoms), such as weight loss, fatigue, fevers or night sweats.

1.1.4.2. Current treatment strategies

Throughout the last decades, therapeutic strategies have evolved significantly due to improved understanding of the biology of CLL. Nonetheless, although treatment administration can prolong progression-free and overall survival, CLL is still considered an incurable disease.

In the past, chemotherapy (chlorambucil, cyclophosphamide, fludarabine or bendamustine) and monoclonal antibodies (anti-CD52 such as alemtuzumab and anti-CD20 such as rituximab or more recently, obinutuzumab or ofatumumab) were administrated as sole agents or in combination to treat CLL patients. However, the relevant percentage of refractory cases, which was even higher in patients with *TP53* deletion and/or mutation, prompted the development of new drugs. In this context, the *TP53* independent

kinase inhibitors ibrutinib and idelalisib, as well as the BCL-2 inhibitor venetoclax, are currently available to treat patients with *TP53* alterations, which has led to a new era of therapeutic approaches in CLL (Hallek, 2017; Wierda et al, 2017) (Figure 4).

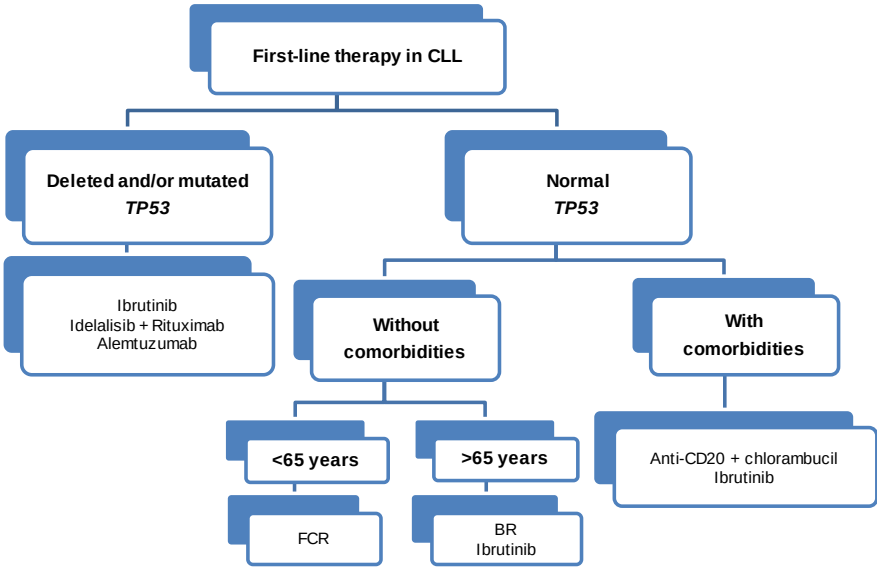

Figure 4. Current first-line therapeutic strategies in CLL (adapted from Sakr et al, 2017). FCR: fludarabine + cyclophosphamide + rituximab; BR: bendamustine + rituximab.

1.2. CLL-like monoclonal B cell lymphocytosis

1.2.1. Description and epidemiology

In the early 2000s, CLL-like monoclonal B cell lymphocytosis (MBL) was formally described as the presence of a small population of clonal B lymphocytes with a phenotype consistent with CLL in the peripheral blood of healthy subjects. The established diagnostic criteria included (Rawstron et al, 2002; Ghia et al, 2004; Marti et al, 2005; Strati et al, 2015):

- Identification of a stable clonal B cell population over a three-month period.
- Clonal B cell count $<5x10^9$ cells/L.
- Presence of a CLL-specific immunophenotype.
- No other features of lymphoproliferative disorder or autoimmune disease such as:
    - Lymphadenopathy or organomegaly.
    - B-symptoms (fatigue, weight loss or night sweats).
    - Autoimmune or infectious disease.

Although the first studies reported that this newly defined entity was present in approximately 3.5% of healthy adult population (Rawstron et al, 2002), its frequency increases with age and is highly dependent on the sensitivity of the flow cytometry method of detection (Nieto et al, 2009; Scarfò et al, 2010; Shanafelt et al, 2010). Thus, the reported CLL-like MBL frequency in healthy population studies ranges from 0.14% to 12% when all ages are included and up to 20% when the analysis is restricted to individuals older than 60 years (Nieto et al, 2009; Shanafelt et al, 2010) (Figures 5 and 6).

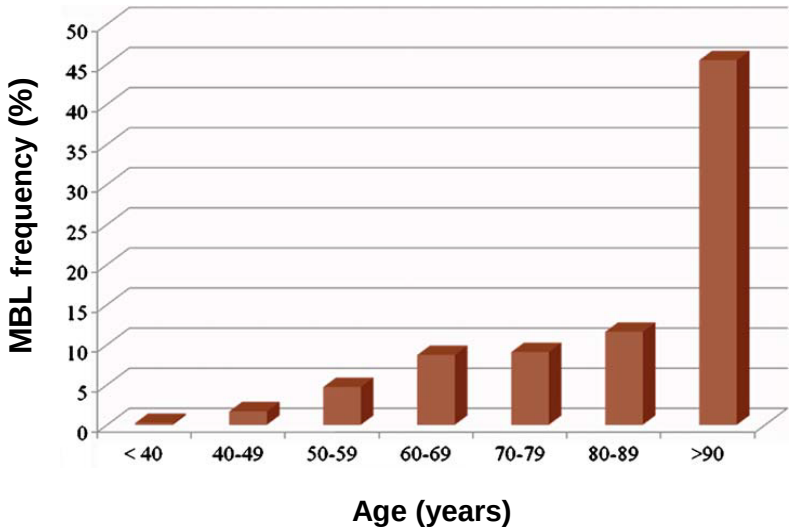

Figure 5. Frequency of CLL-like MBL based on age groups (Scarfò et al, 2010).

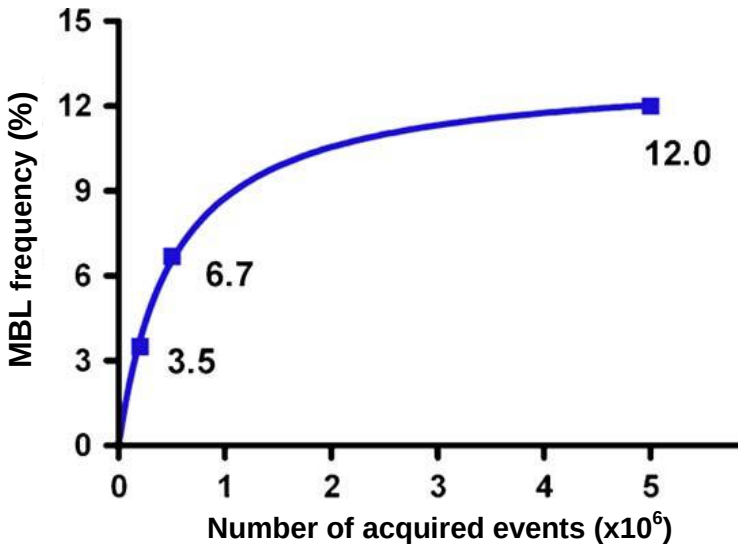

Figure 6. Relationship between the frequency of MBL detection and the number of acquired events (sensitivity) in the flow cytometric analysis (Scarfò et al, 2010).

Finally, it is worth noting that in addition to CLL-like MBL, there are two other groups of MBL depending on the immunophenotype of the B cell clone: MBL with atypical CLL phenotype and $CD5^-$ MBL (Lanasa et al, 2011) (Table 3).

Type of MBL	Immunophenotype					Relative proportion
	CD19	CD5	CD20	CD23	sIg	
CLL-like	+	+	Dim	+	Dim	69%
Atypical CLL	+	+	+	+/Dim	+	9%
$CD5^-$	+	-	+	-	+	22%

Table 3. Subtypes of MBL (Lanasa et al, 2011).

1.2.2. Risk factors

Healthy subjects who have first-degree relatives with CLL have a risk of developing the disease that is two to seven times greater than the general population (Capalbo et al, 2000; Goldin et al, 2009). Thus, family history of CLL must be considered as a risk factor and is associated with the presence of single-nucleotide polymorphisms, which were detected in over twenty genes after performing familial CLL studies (Sellick et al, 2007; Slager et al, 2011). Two additional risk factors are age and infections. As mentioned above, MBL frequency in the general population increases with age (Rawstron et al, 2002). On the other hand, prior investigations showed that MBL is significantly more frequent in subjects who experienced common infections and less frequent in those who were vaccinated against the related pathogens (Fazi et al, 2010; Casabonne et al, 2012). However, it should be noted that MBL is not perceived as an inherited condition nor directly associated with specific infections as causative factors and its etiology is still considered unknown.

1.2.3. Types, relation to CLL and risk of progression

Two different types of CLL-like MBL can be distinguished. Thus, “low-count” MBL (LC-MBL) usually shows <500 clonal cells/µL of peripheral blood, can be detected in the general population only through high-sensitivity methods and has relevant differences from CLL. In addition, it carries extremely reduced, if any, risk of progression to CLL. On the contrary, most cases with “clinical” or “high-count” MBL (HC-MBL) show >500 clonal cells/µL of peripheral blood. Besides, HC-MBL has very similar biological features with early-stage CLL (Rai 0/Binet A) and can potentially evolve to CLL requiring therapy with a progression rate of 1% to 2% per year (Table 4) (Rawstron et al, 2008; Rawstron et al, 2010; Vardi et al, 2013; Strati et al, 2015).

Some of the biological features that distinguish LC-MBL from HC-MBL are the different distribution of poor prognosis genetic abnormalities (deletion of 11q, 17p and *NOTCH1* mutations), which are infrequent in LC-MBL whereas they are detected in HC-MBL at almost similar frequencies to those observed in early-stage CLL (Rawstron et al, 2010; Fazi et al, 2011; Lionetti et al, 2014). Both subtypes of MBL have similar IGHV mutational status (more than 75% of cases harboring somatic hypermutations); however, the IGHV gene repertoire of LC-MBL shows significant differences compared to HC-MBL, such as reduced IGHV1-69, IGHV4-34 and IGHV3-23 and increased IGHV4-59/61 gene frequencies. By contrast, HC-MBL displays a very similar IGHV gene repertoire to that of early-stage CLL, with higher frequencies of IGHV1-69, IGHV3-23 and IGHV4-34. Besides, stereotyped BCR are not a common feature of LC-MBL, whereas they are detected at similar frequencies (20%) in both HC-MBL and early-stage CLL (Vardi et

al, 2013). All these findings strongly support that, as opposed to HC-CLL, LC-MBL is likely an age-related process associated with immune senescence instead of a truly preleukemic condition (Table 4).

Characteristics	CLL all stages	CLL Rai 0	HC-MBL	LC-MBL
		%		
Annual risk of progression	-	5.2	1.1	0
13q deletion	50	~40	~40	~30
Trisomy 12	16	~20	~20	~10
11q deletion	18	~5	~5	0
17p deletion	7	2–3	0–3	0
Unmutated IGHV genes	~45	~25	~25	~25
IGHV1-69	~13	~5	~8	~3
IGHV4-59/61	<5	~5	<5	~20
BCR stereotypy	~30	~20	~20	~5
NOTCH1 mutation	>15	13	11	0

Table 4. Main clinical, cytogenetic and molecular characteristics of the general CLL population, CLL Rai 0, HC-MBL and LC-MBL (adapted from Kalpadakis et al, 2014).

1.2.4. Clinical management of MBL

LC-MBL is only detected when employing high-sensitivity methods in population-screening studies. As previously detailed, subjects with this condition do not have a higher risk of progression to CLL than the general population and, consequently, they do not require further evaluation or follow-up. As for HC-MBL, most cases (around 85%) are diagnosed after the identification of an abnormal blood count and, since they can potentially progress to CLL, they require

further evaluation by a hematologist. However, as the risk of progression to CLL requiring treatment is very low (1-2% per year), only an annual follow-up is recommended unless symptoms or signs of progression appear (Shanafelt et al, 2010) (Figure 7).

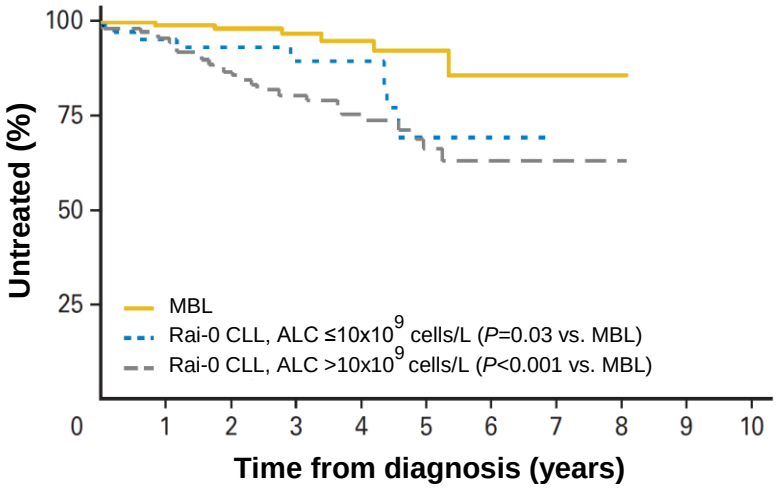

Figure 7. Time from diagnosis to treatment for subjects with clinical MBL and Rai-0 CLL patients with an absolute lymphocyte count (ALC) below or greater than $10x10^9$ cells/L (adapted from Shanafelt et al, 2009).

Of note, currently there are no tools to predict those MBL cases that will progress to CLL. Although in previous small series some of the prognostic parameters in CLL were found to be useful to predict the need for treatment in clinical MBL, such as 17p deletion or high CD38 expression (Rossi et al, 2009; Shanafelt et al, 2009), outside of research purposes there is no evidence to recommend routine prognostic testing in MBL.

1.3. T cell mediated immunity

As components of the adaptive immune system, T cells play a major role in the specific detection and neutralization of pathogens and abnormal cells, such as those from tumors. In order to perform their function, T cells harbor highly specialized receptors (T cell receptors, TR) that specifically bind antigens exposed by antigen-presenting cells (APC) on the surface of major histocompatibility complex (MHC) molecules. Once antigen recognition by TR occurs, T cells are activated, which triggers their proliferation and the initiation of several immune responses (Figure 8) (Janeway et al, 2001).

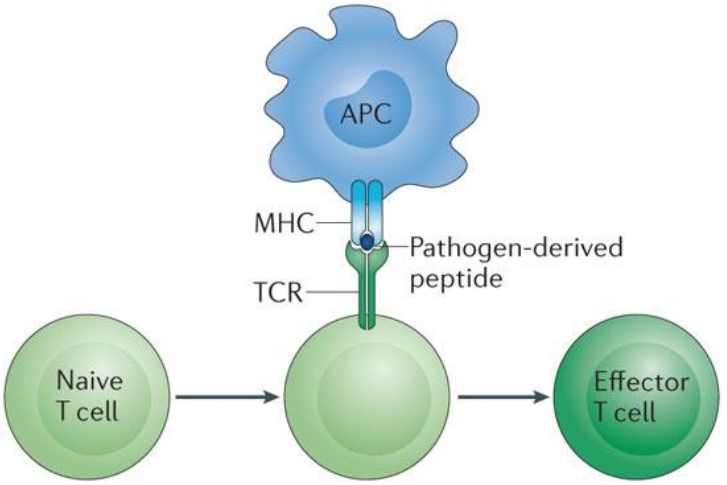

Figure 8. Antigen-dependent T cell activation. APC: antigen-presenting cell; MHC: major histocompatibility complex; TCR: T cell receptor (TR) (adapted from Sewell, 2012).

1.3.1. Genetic and protein organization of the TR

The TR is a complex formed by one each alpha (α) and beta (β) or one each gamma (γ) and delta (δ) chain. The TR loci are always

rearranged and expressed in the same order; if delta and gamma rearrangements, which occur first, are not functional, the T cell will rearrange the beta and alpha loci. If these last rearrangements are functional, αβ T cells are generated, accounting for the great majority of all circulating T cells. Both α- and β-chains are associated with other invariant molecules that form the CD3 complex, which is essential for TR stability and signaling (Figure 9) (Janeway et al, 2001).

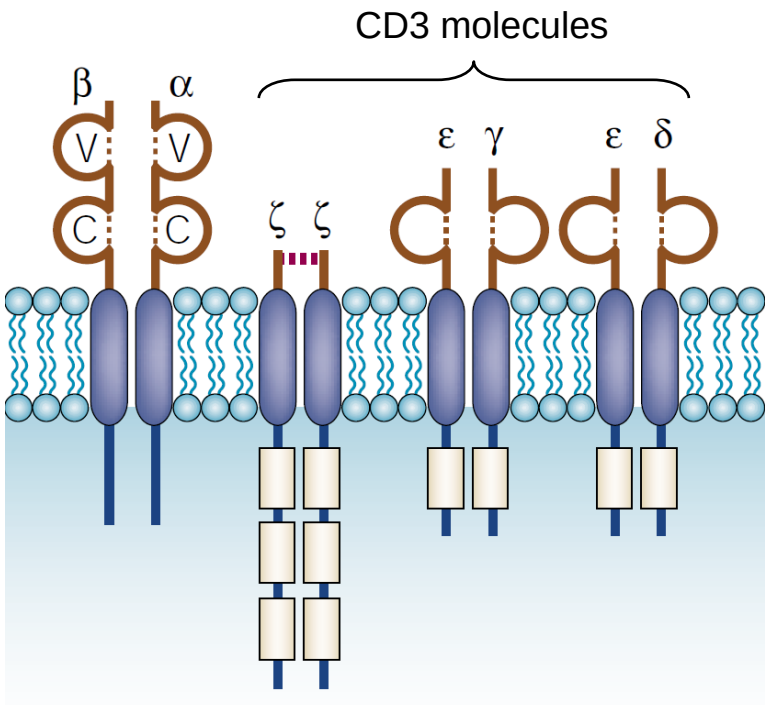

Figure 9. Protein structure of the αβ TR. V: variable region; C: constant region (adapted from Sadelain et al, 2003).

An extremely high diversity of TR is required to specifically recognize all the possible pathogens to which an individual can be

exposed. In order to achieve this, the immune system has developed mechanisms for the somatic diversification of the TR. Thus, in addition to the invariant molecules that form the TR, α- and β-chains harbor variable domains (Figure 9) (Janeway et al, 2001).

Concerning the genomic organization, the TR beta locus (TRB) is composed by four types of genes: V (variable), D (diversity), J (joining) and C (constant). V, D and J genes code for the variable segment of the β-chain whereas the C genes code for the invariant region. During T cell development, extensive recombination occurs between the V, D, and J genes; since there are several genes of each type and only one of them is present in the functional V-D-J rearrangement, a wide range of different rearrangements can occur. Besides, random addition of nucleotides in the junction of the V, D and J genes results in additional increased diversity (Figure 10). A similar rearrangement process occurs for the TR alpha locus (TRA). Collectively, this results in a vast TR repertoire, which is estimated to be around $2.5x10^7$ unique αβ TR per individual (Arstila et al, 1999; Janeway et al, 2001).

As a consequence of the different gene combinations and random nucleotide insertions, the region of the TR β-chain that spans the V-D-J junction is the most variable part and governs antigen recognition. Thus, the specificity of the TR mainly depends on this region, which is called “complementarity determining region 3” (CDR3) (Figure 10). In addition, it is unique not only to each TR β-chain but also to each T cell (as only one rearranged β-chain is expressed). Consequently, the CDR3 region is frequently employed as a measure of T cell diversity. On the other hand, TR α-chain is less appropriate for T cell repertoire characterization due to the possibility of dual-receptors: up to 30% of rearranged T cells express two distinct functional α-chains as opposed to only one β-

chain (Padovan et al, 1993; Janeway et al, 2001; Freeman et al, 2009; Li et al, 2013).

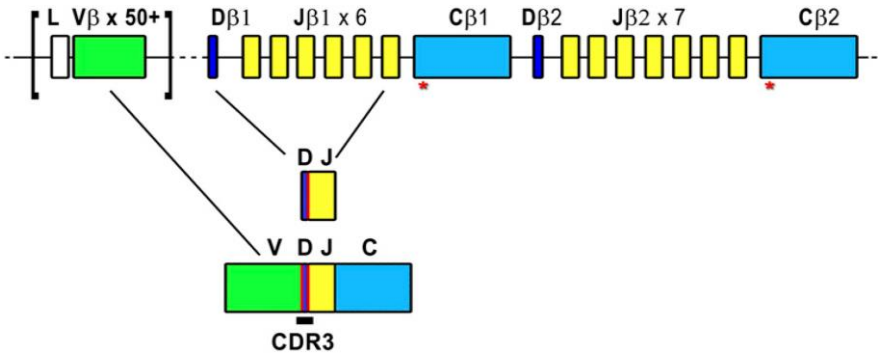

Figure 10. Genomic organization of the TR beta locus (TRB) at chromosome 7q34. It harbors around 50 TRBV (green), 2 TRBD (dark blue), 13 TRBJ (yellow) and 2 TRBC (light blue) genes. Addition of random nucleotides in the rearranged TRBV-TRBD-TRBJ junction is represented by red lines (adapted from Freeman et al, 2009).

1.3.2. Antigen-presenting cells and HLA molecules

Since T cells are not able to directly recognize pathogens, antigen processing and presentation by APC must occur. Thus, APC present peptide fragments derived from the pathogen on the surface of MHC molecules. The most important APC are highly specialized dendritic cells, whose only known function is antigen internalization and presentation. Macrophages and B cells can also function as APC (Janeway et al, 2001).

The MHC proteins are encoded by the human leukocyte antigen (HLA) system. The different HLA genes can be categorized into two main groups:

- HLA class I (A, B, and C), whose proteins present peptides from inside the cell.
- HLA class II (DP, DQ and DR), whose proteins present peptides from outside of the cell.

Each HLA protein can bind a different range of peptides. Besides, HLA genes are extremely polymorphic and each individual is equipped with several of them. Altogether, this variability allows the presentation of a broad range of antigens within each individual (Figure 11) (Janeway et al, 2001).

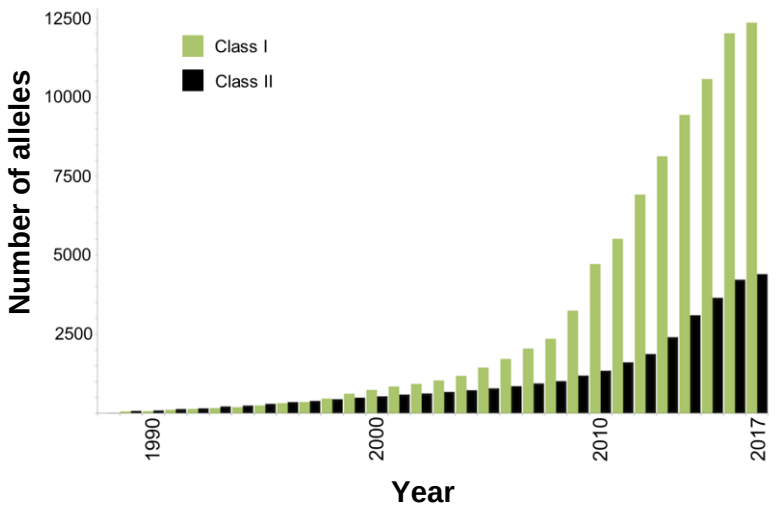

Figure 11. Number of estimated HLA Class I and Class II alleles (adapted from HLA Nomenclature; http://hla.alleles.org/nomenclature).

The different genes of the HLA system are closely linked and are inherited as haplotypes (groups of genes). Of note, distinct combinations of HLA genes within each individual and within populations make them more susceptible or resistant to the development of different diseases, including infections,

autoimmune processes and tumors (Sette et al, 2001; Matzaraki et al, 2017).

1.3.3. Diversity and function of T cells

Once naïve T cells are stimulated by antigen exposure, they become effector T cells. Peptides that are presented into MHC class I molecules are exposed to T cells that harbor the TR co-receptor CD8, which differentiate into cytotoxic T cells. These $CD8^+$ cytotoxic T lymphocytes (CTL) kill abnormal cells such as infected (particularly with viruses), damaged or cancer cells. In order to perform their function, they release a subset of cytotoxic molecules (perforin, granzymes and granulysin) that eventually trigger target cell death (Janeway et al, 2001; Golubovskaya et al, 2016).

On the other hand, those peptides presented into MHC class II molecules are exposed to T cells that express the TR co-receptor CD4. These cells differentiate into T helper cells, since they help the function of other immune cells by releasing different cytokines, modulating immune responses. Naïve $CD4^+$ T cells can differentiate into several subtypes of T helper cells after TR activation, co-stimulation signaling and different types of cytokine stimuli. Thus, T helper cells with different effector commitment can arise, including Th1, Th2, Th17, follicular helper (T_{FH}) and regulatory (Treg) cells. Besides, $CD4^+$ T cells can also differentiate into cells with cytotoxic properties ($CD4^+$ CTL, producing granzyme and perforin). Each of these subsets is characterized by specific functional properties and cytokine profiles and has been found to play distinct roles not only in infectious or immune disorders (Figure 12) but also in several types of tumors (Figure 13) (Janeway et al, 2001; Swain et al, 2012; Kim et al, 2014).

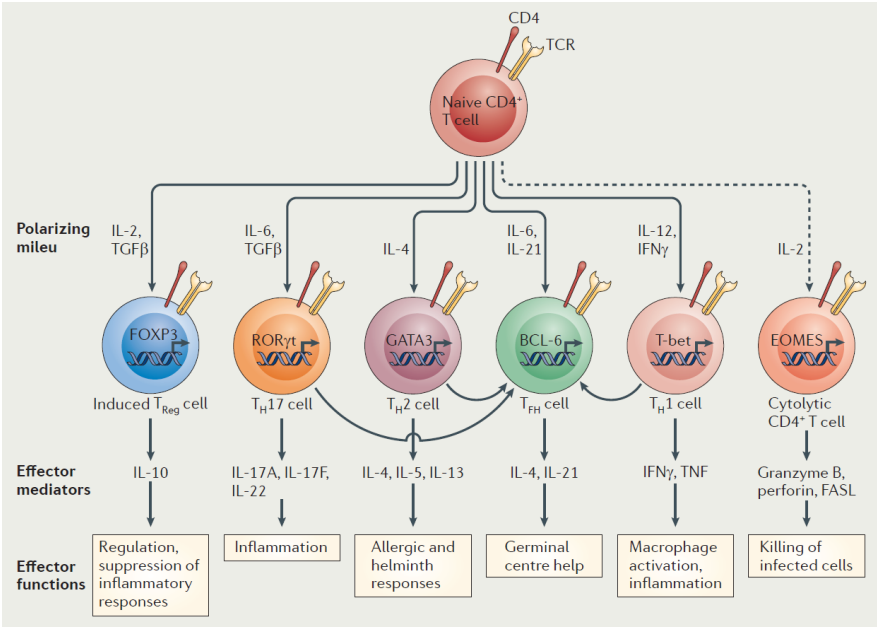

Figure 12. Polarizing cytokines, effector mediators and effector functions related to the different $CD4^+$ T cell subsets (Swain et al, 2012).

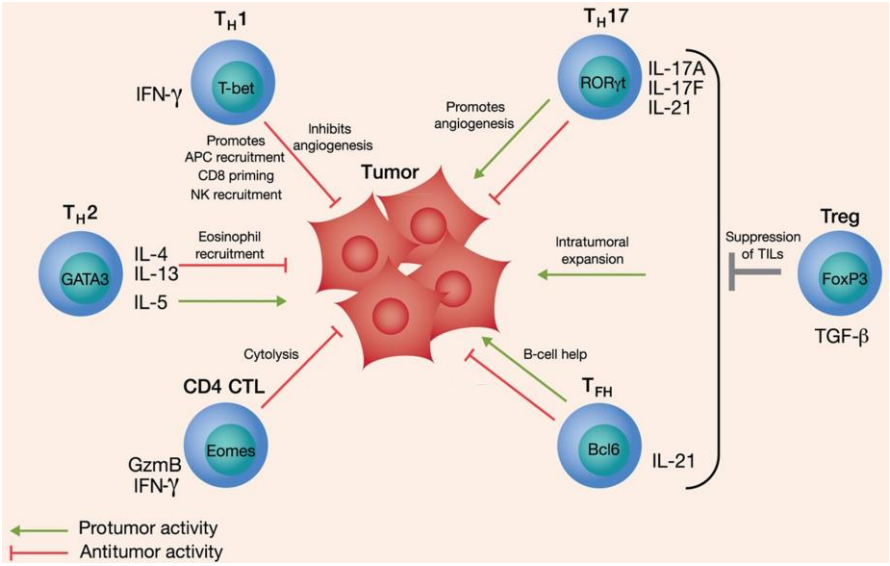

Figure 13. Contribution of the different $CD4^+$ T cell subsets to tumor immunity. APC: Antigen-presenting cell; CTL: Cytotoxic T lymphocyte; TIL: Tumor-infiltrating lymphocyte (adapted from Kim et al, 2014).

1.3.4. Interaction between T cells and B cells

B cells can act as APC by exposing peptides on the surface of HLA-II molecules, which present peptides coming from outside of the cell. Thus, after specific BCR antigen-recognition, the antigen is internalized, processed and linked to HLA-II molecules that are finally exposed on the cell surface. The peptide-HLA complex can be recognized by helper T cells that, in turn, trigger B cell proliferation and differentiation into antibody-secreting cells, which initiates the humoral immune response (Figure 14). Notably, IGHV gene somatic hypermutation depends on the interaction of antigen-stimulated B cells with helper T cells (Janeway et al, 2001).

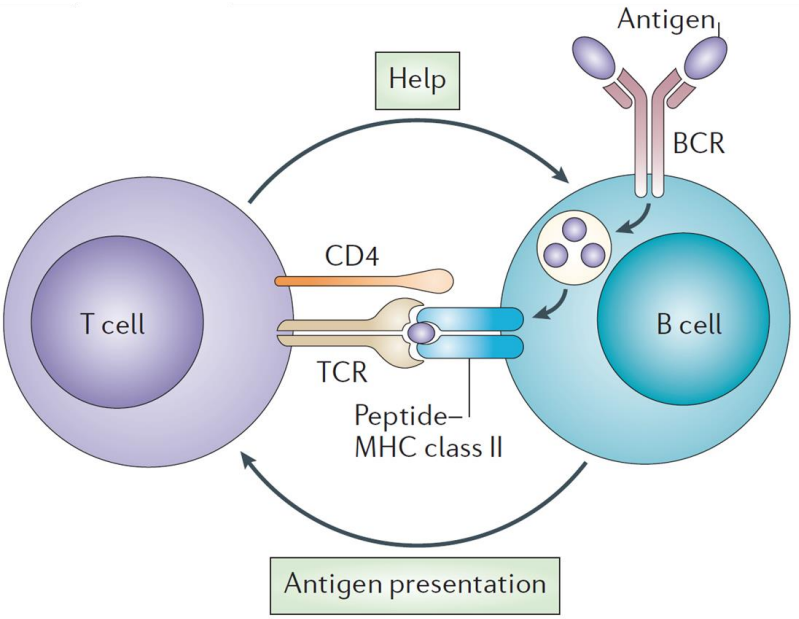

Figure 14. Autologous interaction between T and B cells (adapted from Edwards et al, 2006).

1.4. Tumor microenvironment (TME) in CLL

The tumor microenvironment (TME) is composed by all the accessory cells, soluble factors and extracellular matrix that interact and co-evolve together with the clonal B cells. Prior investigations supported that TME confers a protective niche for CLL cells, promoting cell survival and drug resistance. This, together with several evidences of an antigenic drive in CLL etiology, highlights the importance of the TME in the pathogenesis and evolution of CLL (Agathangelidis et al, 2012; Ten Hacken et al, 2014).

1.4.1. Supportive effects of the TME

Despite the fact that CLL cells circulate freely in peripheral blood, they are attracted to stromal microenvironments in the bone marrow and secondary lymphoid tissues. There, CLL cells interact with multiple cellular types, including bone marrow stromal cells, monocyte-derived nurse-like cells, follicular dendritic cells, endothelial cells and T cells. All these cellular components interact with the clonal B cells through direct cell-to-cell interactions, antigen-mediated interactions and cytokines. As a consequence, CLL cells obtain nourishment, protection from drugs, cytotoxic agents and apoptosis, chemotaxis towards germinal centers and survival signals. A summary of the main components of the TME that contribute to CLL cell survival is shown in Figure 15 (Fabbri et al, 2016).

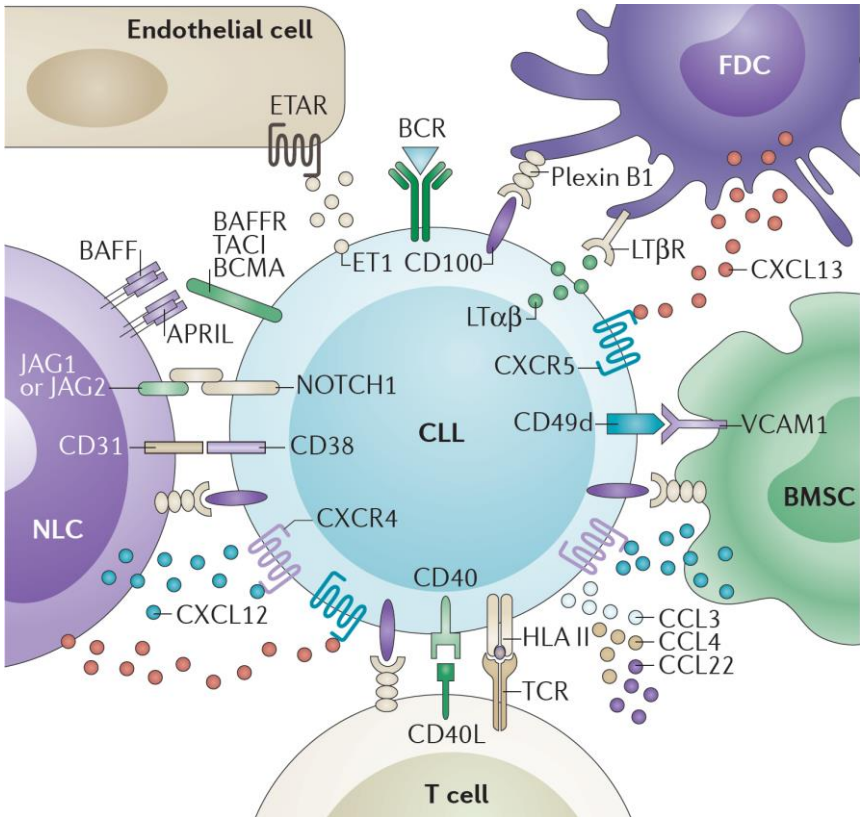

Figure 15. Cell types and molecules involved in the supportive interactions between CLL and TME (Fabbri et al, 2016). BMSC: bone marrow stromal cell; NLC: monocyte-derived nurse-like cell; FDC: follicular dendritic cell.

1.4.2. Alterations of the TME

As a consequence of the CLL-TME cross-talk, the clonal B cells are able to modify the components of the TME, which is essential for their survival. Nonetheless, reflecting the well known heterogeneity of CLL, alterations in the TME can differ among patients and some of them have been associated with a prognostic value.

First, CLL cells trigger modifications in the cellular components of the protective microanatomical sites. For instance, bone marrow

stromal cells co-cultured with CLL cells acquire a phenotype reminiscent of cancer-associated fibroblasts and activate NF-κB signaling, which promotes survival of malignant B cells (Lutzny et al, 2013). Nurse-like cells of CLL patients express higher levels of the hepatocyte growth factor receptor, which activates STAT3 signaling and ultimately favors tumor evasion (Giannoni et al, 2014). CLL cells are also able to stimulate follicular dendritic cells through lymphotoxin-β-receptor activation, resulting in increased CXCL13 secretion that contributes to homing of CLL cells (Heinig et al, 2014).

Second, as a consequence of chronic exposure to tumor burden, several alterations in the immune system are elicited, leading to immune suppression that allows tumor expansion. In this sense, T cells were extensively reported to have an abnormal function in CLL (Scrivener et al, 2003). Thus, despite the presence of elevated absolute T cell counts and oligoclonal expansions (Rezvany et al, 1999; te Raa et al, 2012; Palma et al, 2017), T cells from CLL patients show different gene expression profiles (Görgün et al, 2005), are not capable to form functional immune synapses (Ramsay et al, 2008; Ramsay et al, 2012), display a reduced motility (Ramsay et al, 2013; Ysebaert, 2013) and an increased expression of exhaustion markers such as programmed cell death-1 (PD1) (Riches et al, 2013). Notably, elevated numbers of $CD4^+$ $PD1^+$ T cells in the peripheral blood of CLL patients were correlated with a shortened time to the first treatment and advanced disease (Rusak et al, 2015). Moreover, T cell tolerance induced by CLL seems to be critical for clonal expansion and, interestingly, could potentially be reverted by immunomodulating drugs (Ramsay et al, 2012).

A bias in the $CD4^+$ T cell subpopulations was also described in CLL. Regarding this, prior studies identified increased levels of Th17 cells, mainly in early-stage and good prognosis patients (Jain et al, 2012; Hus et al, 2013). During disease progression, Treg cells are expanded, which promotes immune suppression and allows the proliferation of the malignant clone (Jadidi-Niaragh et al, 2013). Thereby, increased Treg levels were associated with advanced CLL (Lad et al, 2013). In addition, previous studies detected a switch from Th1 to Th2 predominance during disease evolution (Podhorecka et al, 2002). Cytotoxic $CD4^+CD28^-$ T cells, expressing granzyme and perforine, have been observed to be increased in CLL patients as well (Porakishvili et al, 2001; Porakishvili et al, 2004).

Defects in other immune cells were also reported in CLL. Thus, natural killer cell function is impaired in CLL, with a defective cytotoxicity that promotes evasion of malignant cells from antitumor responses (Reiners et al, 2013). Non-cytotoxic innate lymphoid cells, a heterogeneous group of cells with lymphoid morphology and without rearranged antigen receptors, were recently reported to be functionally altered in CLL as well (de Weerdt et al, 2016). Besides, an elevated number of monocytes was detected in the peripheral blood of CLL patients, displaying a deregulation of genes involved in phagocytosis and inflammation (Maffei et al, 2013). Myeloid-derived suppressor cells (MDSC), a subtype of myeloid cells with strong immunosuppressive activities, were shown to be increased in CLL, contributing to tumor expansion by inhibiting T cell responses and promoting Treg (Jitschin et al, 2014).

As expected, serum cytokine levels in patients with CLL are altered, too (Yan et al, 2011). Interestingly, increased levels of

some cytokines, such as CCL3 and CCL4, were associated with a worse outcome and with U-CLL cases (Sivina et al, 2011; Sivina et al, 2016).

Remarkably, while in CLL B cell expansions are accompanied by immune suppression that restrains antitumor immunity, prior investigations showed that the function of T cells in MBL is only slightly deviated. Thus, although immunosuppressive features are already detectable in MBL and evolve toward a more and more suppressive profile across the disease stages, the functional T cell compartment may play relevant roles in the first stages of the disease (Figure 16) (Rissiek et al, 2014; Liu et al, 2015).

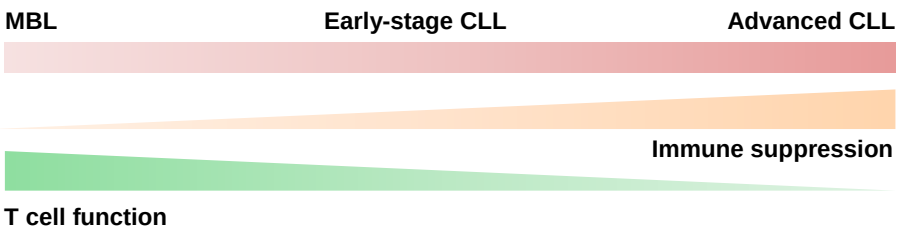

Figure 16. Schematic illustration of the gradual changes in the immune microenvironment with the progression of the disease.

1.4.3. Role of antigens

1.4.3.1. Common infections and self-antigens

Infections are the cause of death in 30-50% of cases with CLL and are mainly due to immune dysregulation. First, aberrant B cells that accumulate and replace normal B cells are not able to properly produce immunoglobulins. Thus, hypogammaglobulinemia is a common feature in CLL that consequently leads to higher risk of infections. Besides, the defects in the TME previously described,

which generally increase over time after diagnosis, contribute to immune dysfunction as well. Finally, immunosuppressive therapies are another source of opportunistic infections in CLL; for this reason, prophylactic strategies are essential in CLL patients who undergo immunosuppressive or innovative therapies. Some of the most common infections related to treatment are those caused by common bacterial pathogens (*Streptococci*, *Staphylococci*, enteric GRAM-negative bacteria), opportunistic infections (*Candida*, moulds, *Pneumocystis jiroveci*), herpesviruses (in particular cytomegalovirus) and other sporadic viral infections (Nosari et al, 2012; Barr et al, 2014; Whitaker et al, 2014). Concerning MBL, even though the immune system of these subjects is not so disturbed, some studies demonstrated that MBL individuals have also an increased risk of infections, although to a lesser extent (Rossi et al, 2009; Moreira et al, 2013).

Importantly, in addition to being a consequence of the tumor, infectious processes were proposed to have a causative role in MBL/CLL. On these lines, previous studies showed that MBL was significantly less frequent in individuals who had been subjected to vaccination against common infections such as influenza or pneumococcal disease (Casabonne et al, 2012) and more frequent in individuals infected with hepatitis C virus (Fazi et al, 2010). Furthermore, an increased risk of CLL was reported in subjects with a personal history of pneumonia (Landgren et al, 2007). All these findings suggest that B cell stimulation by particular pathogens could play a role in leukemogenesis (Figure 17).

Lastly, in addition to foreign antigens, antigens derived from the patient (self-antigens) were also proposed to have a role in CLL development (Catera et al, 2008; Chu et al, 2008; Lanemo Myhrinder et al, 2008) (Figure 17). Moreover, the binding of

intrinsic BCR motifs was proposed as an autostimulatory mechanism related to CLL pathobiology (Dühren-von Minden et al, 2012; Binder et al, 2013).

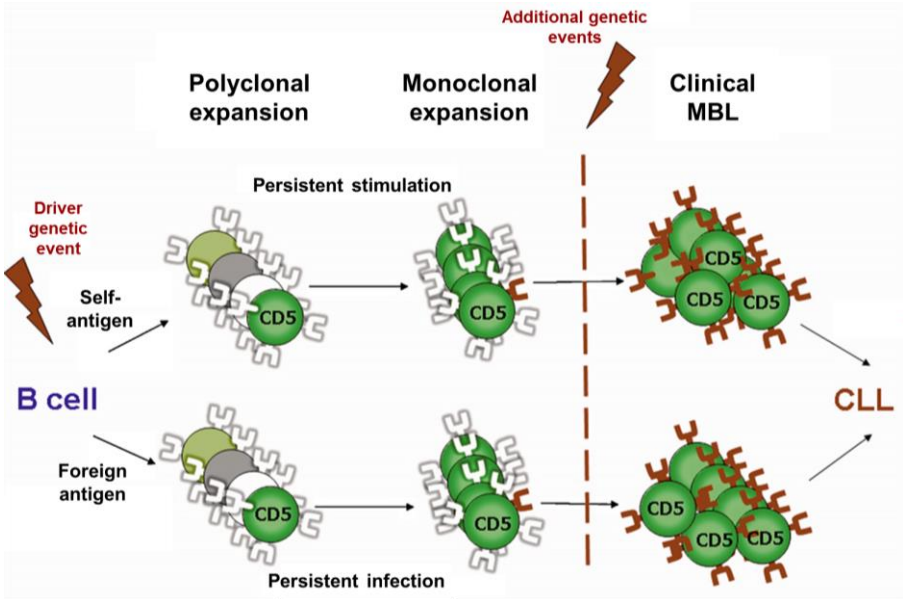

Figure 17. Hypothetical pathogenic model for MBL (adapted from Scarfò et al, 2010).

1.4.3.2. Molecular evidences of antigen restriction

Antigen selection has been strongly demonstrated in CLL, in which tumor B cells show a remarkable restriction of the immunoglobulins (IG) expressed by the BCR (Fais et al, 1998), including the existence of subsets with quasi-identical, stereotyped BCR IGs in one third of all CLL cases (Figure 18) (Agathangelidis et al, 2012; Agathangelidis et al, 2014). Recently, a role for antigen selection has also emerged for T cells in CLL, with oligoclonality and T cell clones that persist over time and are shared by different patients. These T cell clones were proposed to be disease-

specific, since they were not found in healthy individuals neither public databases (Vardi et al, 2016; Vardi et al, 2017). All these findings further extend the implication of antigen selection in the natural history of CLL.

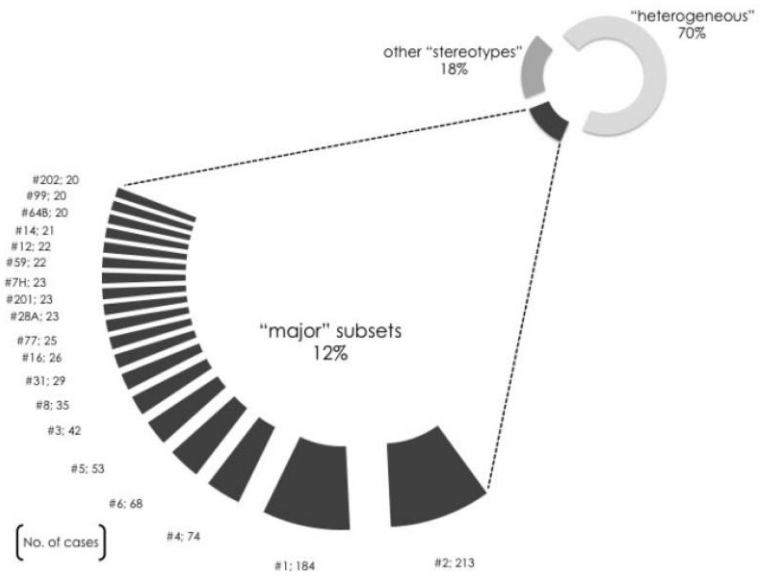

Figure 18. Stereotyped BCR in one third of all CLL cases. Major subsets account for a relevant proportion (12%) of the IG repertoire (Agathangelidis et al, 2012).

Besides, HC-MBL displays a very similar IGHV gene repertoire to that of early-stage CLL, as well as stereotyped BCR, both findings supporting antigenic restriction from the first stages of the disease (Vardi et al, 2013).

Finally, the importance of antigenic determinants in CLL is also supported by a bias in HLA usage, mainly concerning the development of severe disease (Shah et al, 2011; Di Bernardo et al, 2013; Gragert et al, 2014).

2. HYPOTHESIS AND AIMS

2.1. Hypothesis

CLL-like HC-MBL is generally perceived as a requisite precursor of CLL. Although in recent years many studies have been performed to characterize MBL, many aspects of its biology are still unknown, mainly concerning the TME. In CLL, prior investigations have reported dysfunctional $CD4^+$ and $CD8^+$ T cells (Görgün et al, 2005) and altered cytokine patterns (Yan et al, 2011). Gradual changes in the TME following clinical progression as well as numerical increases of clonal B cells have also been reported (Jadidi-Niaragh et al, 2013; Sivina et al, 2016). Besides, antigen-mediated interactions between the aberrant B cells and T cells are crucial for disease evolution. In this sense, antigenic elements are known to play a major role in CLL pathogenesis and shape not only the malignant clone (Agathangelidis et al, 2012) but also the T cell repertoire (Vardi et al, 2017). However, although MBL shares biological characteristics with CLL, it is an asymptomatic entity and displays a low progression rate to CLL (Rawstron et al, 2008). Therefore, it could be hypothesized that TME in MBL may harbor different biological features. In addition, as an early-stage of CLL ontogeny, MBL may hold important clues about leukemogenesis.

On these grounds, the hypotheses of the present study are:

1. As described in CLL, the TME plays an important and probably different role in MBL, which is reflected in the gene expression profile of T cells and in serum cytokine patterns.

2. There is a restricted set of antigens that shapes not only the malignant clone but also the T cell population in MBL, which can be detected by analyzing the TR repertoire.

2.2. Aims

The general aim is to characterize the TME, focusing on $CD4^+$ and $CD8^+$ T cells, in CLL-like HC-MBL.

The more specific aims are:

1. To define the gene expression profiles of $CD4^+$ and $CD8^+$ T cells from MBL subjects and compare them with those from early-stage CLL patients and healthy controls.

2. To describe the serum cytokine patterns of MBL subjects and compare them with those observed in early-stage CLL and healthy controls.

3. To characterize the TR repertoire of the $CD4^+$ and $CD8^+$ T cell subpopulations in MBL and early-stage CLL.

3. PATIENTS AND SAMPLES

3.1. Patients

A total of 58 subjects with CLL-like HC-MBL and 54 untreated early-stage CLL patients (Binet A/Rai 0-I and <23 $x10^9$ clonal B cells/L) were included in this study. All cases met the iwCLL diagnostic criteria (Hallek et al, 2008). In addition, 31 healthy subjects were studied as controls.

Taken together, samples from 143 different individuals were collected. No case had evidence of infection at sampling. In order to obtain a large number of study subjects (mainly MBL), different centers collaborated in the present study:

- Hospital del Mar (Barcelona): 25 MBL and 18 CLL.
- Hospital Vall d'Hebron (Barcelona): 6 MBL and 5 CLL.
- The Feinstein Institute for Medical Research (New York): 27 MBL and 31 CLL.

The distribution of the total study cohort into different types of experiments was as follows:

- Gene expression arrays: 15 MBL, 14 CLL and 9 controls.
- qPCR validation: The same cohort as that of the gene expression arrays and an additional independent cohort consisting of 10 MBL and 5 CLL cases.
- Cytokine immunoassays: 41 MBL (8 of these subjects were also included in the gene expression arrays, one of them

categorized as MBL in the cytokine immunoassays but as CLL in the gene expression arrays studies due to the employment of different longitudinal samples with an increase of lymphocytosis; 2 of them also included in the qPCR independent cohort), 45 CLL (7 of them also characterized by gene expression arrays and 2 of them belonged to the qPCR independent cohort as well) and 24 healthy controls (2 of them also studied by gene expression arrays).

- TR repertoire analysis by subcloning/Sanger sequencing: 16 MBL (the same 15 MBL cases included in the gene expression arrays and an additional case) and 13 CLL (all of them also characterized by gene expression arrays).

The demographic, clinical and immunogenetic characteristics of the subjects included in each part of the study are shown in the respective sections of the Results.

The study was performed in accordance with national and international guidelines (Professional Code of Conduct, Declaration of Helsinki) and approved by the Ethics Committee of Hospital del Mar (2011/4317/I). Cytokine studies were approved by the Northwell Health IRB.

3.2. Samples

Depending on the methodology, different samples were employed, as follows:

- Gene expression arrays: Purified $CD4^+$ and $CD8^+$ T cell fractions obtained from peripheral blood.
- qPCR validation: Purified $CD4^+$ T cell fractions from peripheral blood.
- Cytokine immunoassays: Serum samples obtained from peripheral blood. In 4 selected cases, sequential serum samples were also studied after 6 (stable M-CLL), 19 (aggressive U-CLL), 24 (stable MBL) and 60 (MBL that progressed to M-CLL) months.
- TR repertoire analysis by subcloning/Sanger sequencing: Purified $CD4^+$ and $CD8^+$ T cell fractions from peripheral blood. In 3 MBL cases, purified $CD4^+$ and $CD8^+$ T cells were analyzed over time (median sampling interval: 19 months; range: 3.5-32 months).

A schematic representation of the different subjects and samples employed in each part of the study is shown in Figure 19.

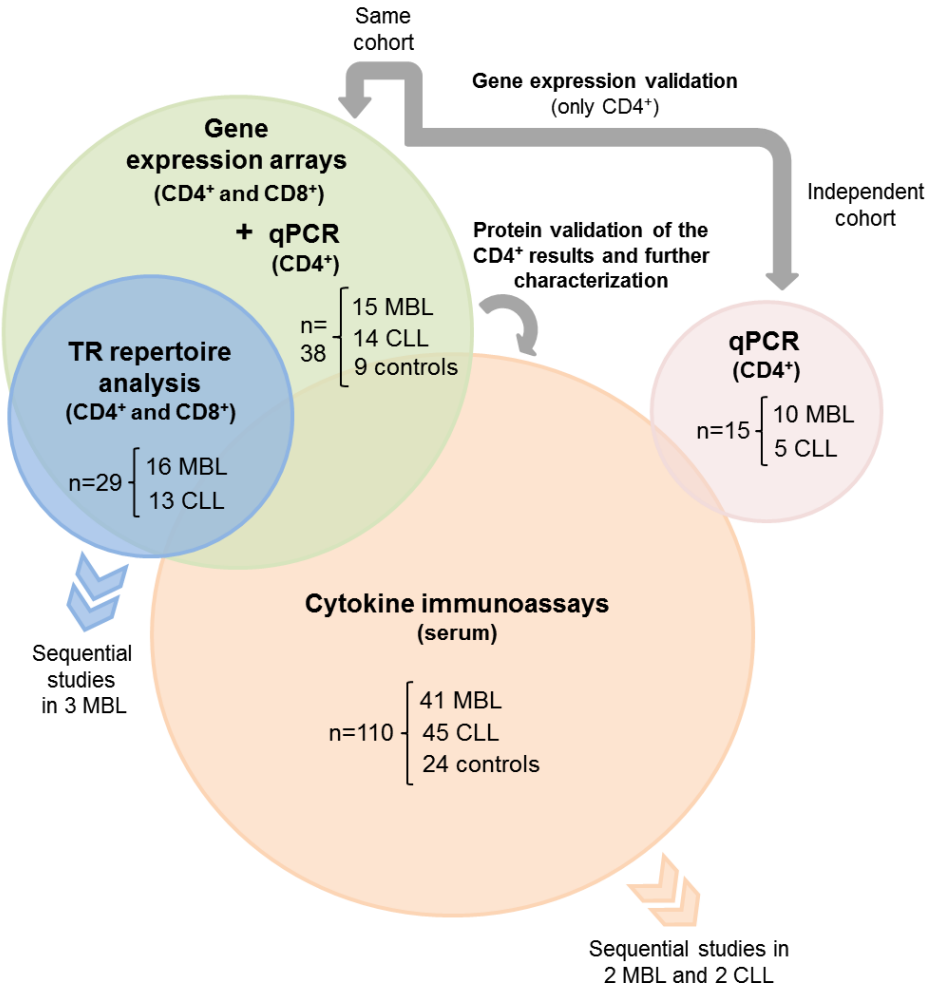

Figure 19. Schematic illustration showing the different subjects and samples employed in each part of the study. Overlapping areas represent individuals whose samples were analyzed by different methodologies.

4. METHODS

4.1. Clinical and biological data collection

For all subjects included in the different studies, demographic (sex and age) and clinical data (clonal B cell counts, absolute lymphocyte counts, Binet and Rai stages) were collected whenever possible, if applicable. Of note, age, cell counts and clinical staging were registered at the time point of the analysis.

Data regarding IGHV gene somatic hypermutation status was also collected when available. Concerning this, the 98% cut-off value for IGHV gene identity to germline was followed, which is used to discriminate between M-CLL and U-CLL (Damle et al, 1999).

For the group of MBL and CLL individuals included in the TR subcloning analysis, genetic data regarding the HLA-A, -B and -C loci and the HLA-DRB1 locus was also collected.

4.2. Gene expression studies

4.2.1. Isolation of purified $CD4^+$ and $CD8^+$ T cell fractions

For all studied subjects (MBL, CLL or healthy controls), a specific protocol was designed in order to simultaneously purify four different cell fractions: malignant B cells (MBL and CLL) or normal B cells (healthy controls), $CD4^+$ T cells, $CD8^+$ T cells and NK cells. Although only the $CD4^+$ and $CD8^+$ T cell fractions were included in this study, the methodological procedure will be explained together (Figure 20). In addition, as detailed in the section 4.2.5, the gene expression results of the B cell fraction were employed to filter the raw $CD4^+$ and $CD8^+$ T cell data. All the purification steps were

based on positive selection methods employing immunomagnetic beads and autoMACS technology, both from Miltenyi Biotec.

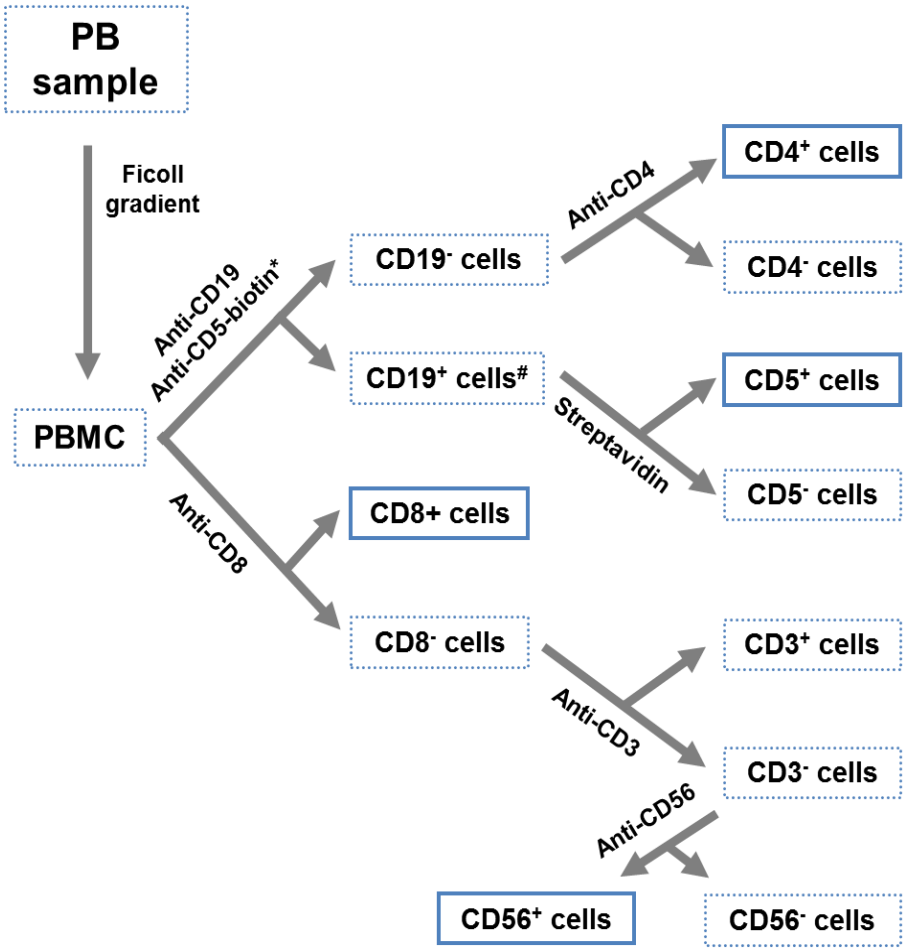

Figure 20. Designed procedure for the simultaneous purification of B cells (malignant or normal), $CD4^+$ T cells, $CD8^+$ T cells and NK cells. The final cell fractions employed for gene expression arrays are highlighted with blue solid lines. *Not employed for healthy controls. #Final cell fraction for healthy controls instead of $CD19^+CD5^+$. PBMC: Peripheral blood mononuclear cells.

In order to obtain a sufficient amount of RNA from each of these cell fractions, a total of 25 mL of fresh peripheral blood in EDTA tubes were collected for each subject included in the analysis. The complete procedure is described below:

1. Peripheral blood was diluted (1:1) with phosphate-buffered saline (PBS) 1x.

2. In 15 mL tubes, 5 to 7 mL of the diluted blood were added very carefully over 3 mL of Ficoll (Lymphoprep, Axis-Shield) at room temperature. Importantly, two different and not mixed phases were obtained, as represented in Figure 21.

3. All tubes were centrifuged for 20 minutes at 2000 rpm. After this step, four different phases were obtained, as shown in Figure 21.

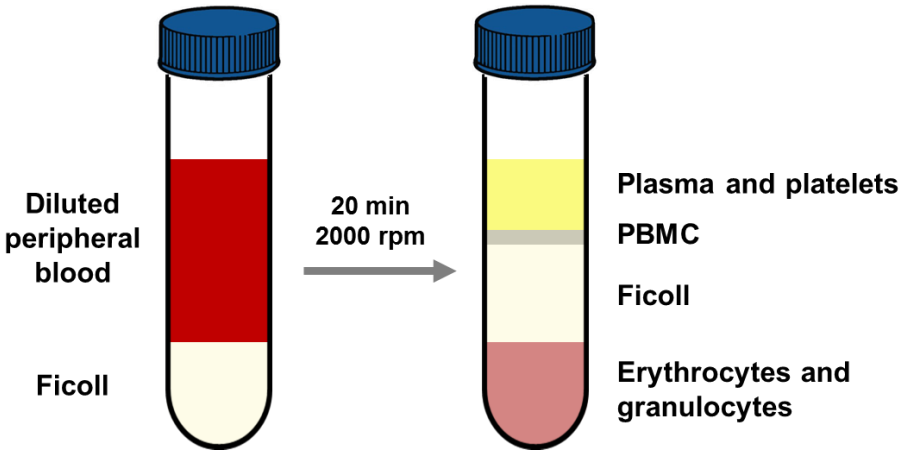

Figure 21. Representation of the Ficoll-based density-gradient centrifugation technique. PBMC: peripheral blood mononuclear cells.

4. The PBMC phases of all tubes from the same subject were isolated and placed in a new 15 mL tube, which was filled with PBS 1x.

5. The tube was centrifuged for 7 minutes at 1700 rpm.

6. The supernatant was discarded and the PBMC pellet was resuspended in 2-3 mL of PBS 1x.

7. PBMC were measured using a hematology analyzer coulter (Beckman Coulter) and divided in two halves that were placed in two different 15 mL tubes.

8. Both tubes were centrifuged for 7 minutes at 1700 rpm. Supernatants were discarded and PBMC pellets resuspended.

9. As represented in Figure 20, one of the tubes was incubated with CD8 MicroBeads, human (20 µL/10x10^6 cells) and BSA/EDTA buffer (80 µL/10x10^6 cells). The other tube was incubated with MultiSort CD19 MicroBeads (20 µL/10x10^6 cells), CD5-Biotin, human (10 µL/10x10^6 cells) and BSA/EDTA buffer (80 µL/10x10^6 cells). In the case of healthy subjects, only CD19 MicroBeads, human (20 µL/10x10^6 cells) and BSA/EDTA buffer (80 µL/10x10^6 cells) were added into the second tube. After mixing all the reagents, samples were incubated at 4ºC for 15 minutes.

10. After incubation, both tubes were filled with BSA/EDTA (3/4 of the 15 mL tube) and centrifuged for 10 minutes at 1800 rpm.

11. Supernatants were discarded and cell pellets were resuspended in 1 mL of BSA/EDTA buffer. Samples were then subjected to magnetic separation using autoMACS.

12. The $CD8^+$ cell fraction was measured using Hematology Analyzer Coulter and afterwards centrifuged (7 minutes at 1700 rpm). The obtained pellet was resuspended in 350 or 600 µL of 1% β-mercaptoethanol (BME) RLT-plus buffer (when obtaining <5x10^6 cells or 5-10x10^6 cells, respectively) and stored at -80ºC until RNA extraction. The $CD8^-$ cell fraction was employed to obtain $CD3^-$ $CD56^+$ NK cells, as shown in Figure 20.

13. In the case of healthy controls, the $CD19^+$ cell fraction was measured, centrifuged and stored at -80ºC in 1% BME RLT-plus buffer as described in the previous step. In the case of MBL and CLL subjects, magnetically selected $CD19^+$ cells were mixed with MultiSort Release Reagent (20 µL/mL) and incubated for 10 minutes at 4ºC.

14. After this incubation, the tube was centrifuged (7 minutes at 1700 rpm), the supernatant was completely removed and the cell pellet was resuspended in MultiSort Stop Reagent (30 µL/10x10^6 cells). Samples were mixed actively.

15. Streptavidin MicroBeads (10 µL/10x10^6 cells) and BSA/EDTA buffer (90 µL/10x10^6 cells) were added to the previous sample. The tube was incubated for 15 minutes at 4ºC.

16. After incubation, the tube was filled with BSA/EDTA (3/4 of the 15 mL tube) and centrifuged for 10 minutes at 1800 rpm.

17. The supernatant was discarded and the cell pellet was resuspended in 1 mL of BSA/EDTA buffer. The sample was then subjected to magnetic separation using autoMACS.

18. After this separation step, two cell fractions were obtained (Figure 20): $CD5^-$ and $CD5^+$. As previously described, both cell fractions were measured, centrifuged and stored at -80ºC in 1% BME RLT-plus buffer (instead of discarding the $CD5^-$ samples, they were stored as well for possible future applications).

19. In all cases (healthy controls, MBL and CLL), the $CD19^-$ cell fraction was measured using Hematology Analyzer Coulter and centrifuged (7 minutes at 1700 rpm).

20. After centrifugation, the supernatant was removed and the cell pellet was resuspended in a mix containing CD4 MicroBeads, human (20 µL/10x10^6 cells) and BSA/EDTA buffer (80 µL/10x10^6

cells). After mixing all the reagents, samples were incubated at 4ºC for 15 minutes.

21. After incubation, the tube was filled with BSA/EDTA (3/4 of the 15 mL tube) and centrifuged for 10 minutes at 1800 rpm.

22. The supernatant was discarded and the cell pellet was resuspended in 1 mL of BSA/EDTA buffer. The sample was then subjected to magnetic separation using autoMACS.

23. The $CD4^+$ cell fraction was measured employing Hematology Analyzer Coulter and centrifuged (7 minutes at 1700 rpm). Right after and as previously described, the pellet was resuspended in 1% BME RLT-plus and stored at -80ºC. The $CD4^-$ cell fraction was discarded.

For 11 of the 53 purified $CD4^+$ T cell fractions subjected to gene expression analysis (38 by gene expression arrays and all 53 samples by qPCR), the isolation process was performed by directly incubating PBMC, and not $CD19^-$ cells, with anti-CD4 antibodies.

Prior to resuspending the final purified cell fractions in 1% BME RLT-plus for storage, 100000 cells of each sample were isolated to check the efficacy of the purification procedure by flow cytometry. According to this, the panel detailed in Table 5 was applied.

Fraction	FITC (10µL)	PE (10µL)	PerCP-Cy5.5 (10µL)
PBMC	Control	Control	Control
$CD8^+$ T cells	CD8	-	CD3
NK cells	CD56	-	CD3
$CD4^+$ T cells	-	CD4	-
B cells	-	CD5	CD20

Table 5. Flow cytometry panel applied to check the purification procedure. All the antibodies were from Becton Dickinson. For the control tube, the tested sample was the initial unselected PBMC. At least 10000 events were acquired for each case.

The median purity of the $CD4^+$ and the $CD8^+$ cell fractions included in the gene expression studies was 95%, for both cases.

4.2.2. RNA extraction, quantification and quality analysis

Once the purified $CD4^+$ and $CD8^+$ T cell fractions were obtained and after storing at -80ºC in 1% BME RLT-plus buffer, RNA was extracted following the RNeasy Plus Mini Kit protocol (QIAGEN) as follows (Figure 22):

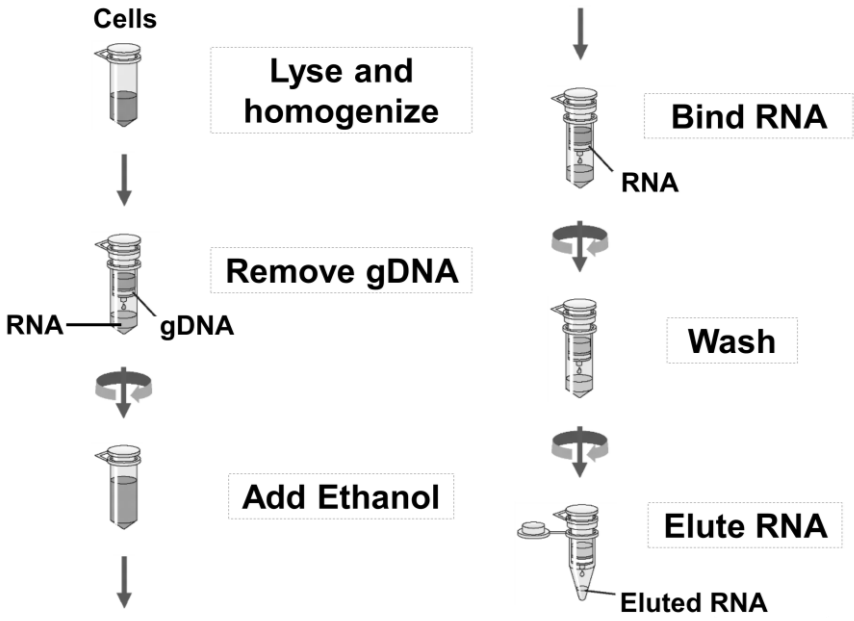

Figure 22. RNA extraction protocol (adapted from RNeasy Plus Mini Handbook, QIAGEN).

1. The homogenized lysed samples were transferred to the gDNA eliminator spin columns provided by the kit, which were placed into 2 mL collection tubes.

2. Samples were centrifuged for 30 seconds at ≥10000 rpm. Columns with the gDNA were then discarded and the flow-through with the total RNA corresponding to each sample was mixed with one volume (350 or 600 µL) of 70% ethanol.
3. Up to 700 µL of each sample were transferred to an RNeasy Mini spin column supplied by the kit placed into a 2 mL collection tube. Columns were centrifuged for 15 seconds at ≥10000 rpm.
4. After centrifugation, flow-through was discarded and 700 µL of RW1 buffer were added to the columns, which were centrifuged for 15 seconds at ≥10000 rpm.
5. Once again, flow-through was discarded and 500 µL of RPE buffer were added to the columns that were then centrifuged for 15 seconds at ≥10000 rpm.
6. After discarding the flow-through, 500 µL of RPE buffer were added again to the columns and centrifugation for 2 minutes at ≥10000 rpm was performed.
7. Columns were placed into new 2 mL collection tubes and centrifuged at full speed for 1 minute to further dry the membrane.
8. Finally, the columns were placed into 1.5 mL collection tubes and 30 to 50 µL of RNase-free H_2O were directly added to the column membrane. Centrifugation for 1 minute at ≥10000 rpm was performed to elute the RNA.

Once extracted, RNA quantity and purity was determined on the ND-2000 Spectrophotometer (NanoDrop Technologies) and RNA integrity was assessed by nanoelectrophoresis on the Nano and Pico lab-on-a-chip assays for total eukaryotic RNA using Agilent 2100 Bioanalyzer (Agilent Technologies). All samples had RNA integrity number (RIN) >7 and were subsequently employed for microarray experiments.

4.2.3. Gene expression arrays

To perform gene expression arrays, processing was designed to equilibrate batches including a representative number of samples of each condition. Samples were processed by the Microarray Analysis Service (SAM) from the Hospital del Mar Medical Research Institute (IMIM).

Array processing was performed with GeneChip Human Gene 2.0 ST arrays (Affymetrix), which provides a sensitive, accurate and comprehensive measurement of protein coding and long intergenic non-coding RNA transcripts. Briefly, total RNA from each sample was processed, labeled and hybridized to the array according to the NuGEN Ovation Pico WTA System V2 (M01224 v2) and Encore Biotin Module (M01111 v6) (NuGEN) and Affymetrix Expression Wash, Stain and Scan User Manual (P/N 702731 Rev. 3) (Affymetrix). The different steps are explained below:

Ribo-SPIA process for cDNA amplification

The Ribo-SPIA process, based on NuGEN's Single Primer Isothermal Amplification (SPIA) technology, is a three-step process that generates amplified cDNA from input RNA, as follows (Figure 23):

1. First strand cDNA synthesis: First strand cDNA was prepared from 30 ng of total RNA using reverse transcriptase (RT) and two types of DNA/RNA chimeric primers, random and oligo(dT). With this primer mix, priming occurred across the whole transcript. RT extended the 3' DNA end of each primer generating first strand cDNA. The resulting cDNA/RNA hybrid molecule contained a

unique RNA tag sequence (SPIA tag) at the 5' end of the cDNA strand which was used as a priming site for the SPIA process.

2. Generation of a DNA/RNA Heteroduplex Double Strand cDNA: Fragmentation of the RNA within the cDNA/RNA complex created priming sites for DNA polymerase to synthesize a second cDNA strand, which included DNA complementary to the 5' SPIA tag sequence from the first strand chimeric primers. The result was a double-stranded cDNA with a DNA/RNA heteroduplex corresponding to the SPIA tag at one end. Prior to cDNA amplification, the mix was purified using Agencourt RNAClean XP beads (Beckman Coulter).

3. SPIA Amplification: SPIA technology is a rapid amplification process developed by NuGEN that uses a DNA/RNA chimeric primer (SPIA primer), DNA polymerase and RNase H in an isothermal assay. First, RNase H removed the RNA portion of the heteroduplex SPIA tag sequence, revealing a site for binding the SPIA primer. Then, DNA polymerase synthesized cDNA starting at the 3' end of the primer, displacing the existing forward strand. Priming with the chimeric SPIA primer recapitulated the heteroduplex SPIA tag, creating a new substrate for RNase H and the initiation of the next round of cDNA synthesis. The process of SPIA DNA/RNA primer binding, DNA replication, strand displacement and RNA cleavage was repeated in a highly processive manner, resulting in rapid accumulation of micrograms of amplified cDNA from picograms of total RNA.

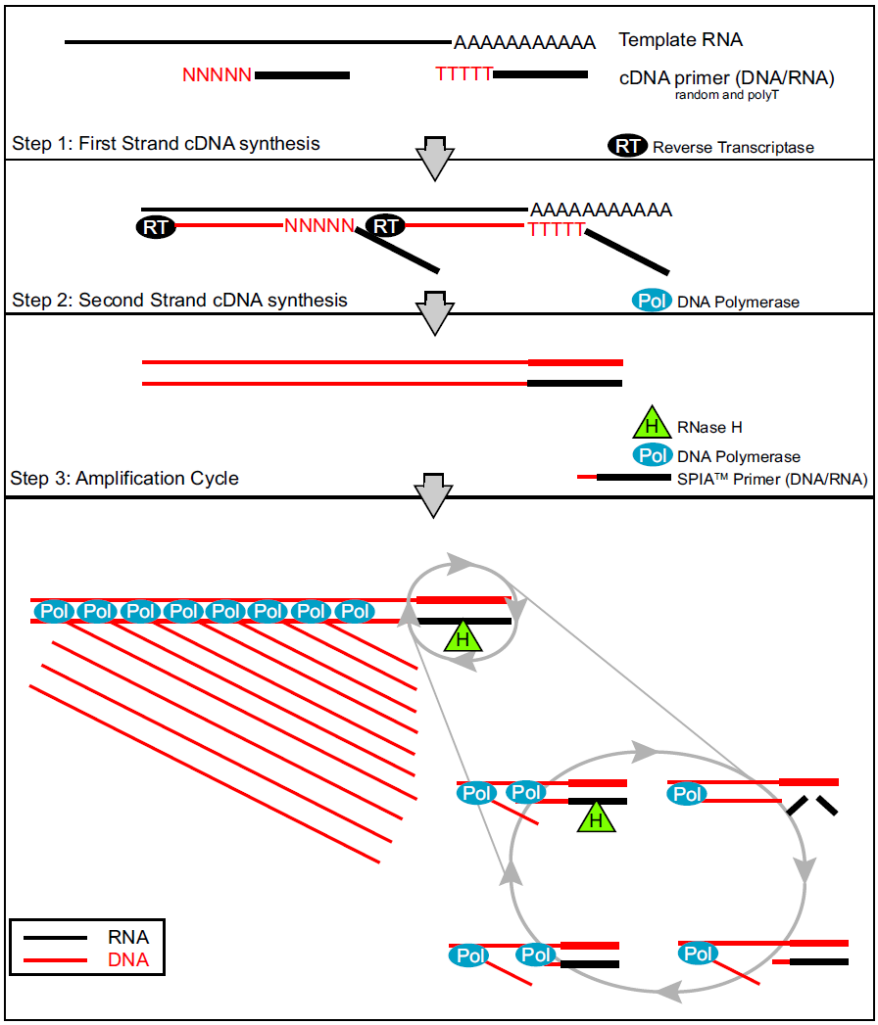

Figure 23. Representation of the Ribo-SPIA process for cDNA amplification (Watson et al, 2008).

cDNA fragmentation and labeling

The SPIA cDNA product was purified using MinElute Reaction Cleanup Kit (QIAGEN) and cDNA yield and purity were measured using the ND-2000 Spectrophotometer. Afterwards, 5 µg of cDNA were subsequently fragmented by employing a chemical and enzymatic process to obtain cDNA products in a 50-100 base

range. Finally, labeling was performed via enzymatic attachment of a biotin-labeled nucleotide to the 3-hydroxyl end of the fragmented cDNA.

Hybridization

The fragmented and biotinylated cDNA was added to the hybridization cocktail, loaded on a GeneChip Human Gene 2.0 ST array and hybridized for 16 hours at 45ºC and 60 rpm in an Affymetrix GeneChip Oven 645. Following hybridization, the array was washed and stained in the Affymetrix GeneChip Fluidics Station 450. The stained array was scanned using an Affymetrix GeneChip Scanner 3000 7G, generating CEL files for each array.

4.2.4. qPCR

Validation of 13 selected genes (*CCL2, CLU, CXCL5, DEFA1, FGFBP2, GZMH, ITGAM, NEXN, NKG7, PPBP, RAB31, TUBB1* and *VSTM1*) that were previously found to be differentially expressed in $CD4^+$ cells was performed by qPCR. In addition, *ACTB*, *GUSB* and *GAPDH* were initially selected as housekeeping genes. The Taqman Low Density Arrays (TLDA, Life Technologies) was the platform used for qPCR analysis in the 16-assay format (8 samples per card were interrogated for the expression of 16 genes; each gene was assayed in triplicate). The analysis was performed by the SAM service from the IMIM. A summary of the methodological procedure is represented in Figure 24 and detailed below:

1. cDNA was reversely transcribed from 400 ng of total RNA using the SuperScript VILO cDNA Synthesis Kit (Life Technologies), as described by the manufacturer.
2. Next, cDNA samples were subjected to TaqMan qPCR reactions using TaqMan Universal PCR Master Mix II on the TLDA platform, according to the manufacturer's instructions. Specifically, 100 ng of cDNA were loaded per reservoir and a total of 45 PCR cycles were performed. Fluorescence was detected employing an ABI PRISM 7900 HT system (Life Technologies).

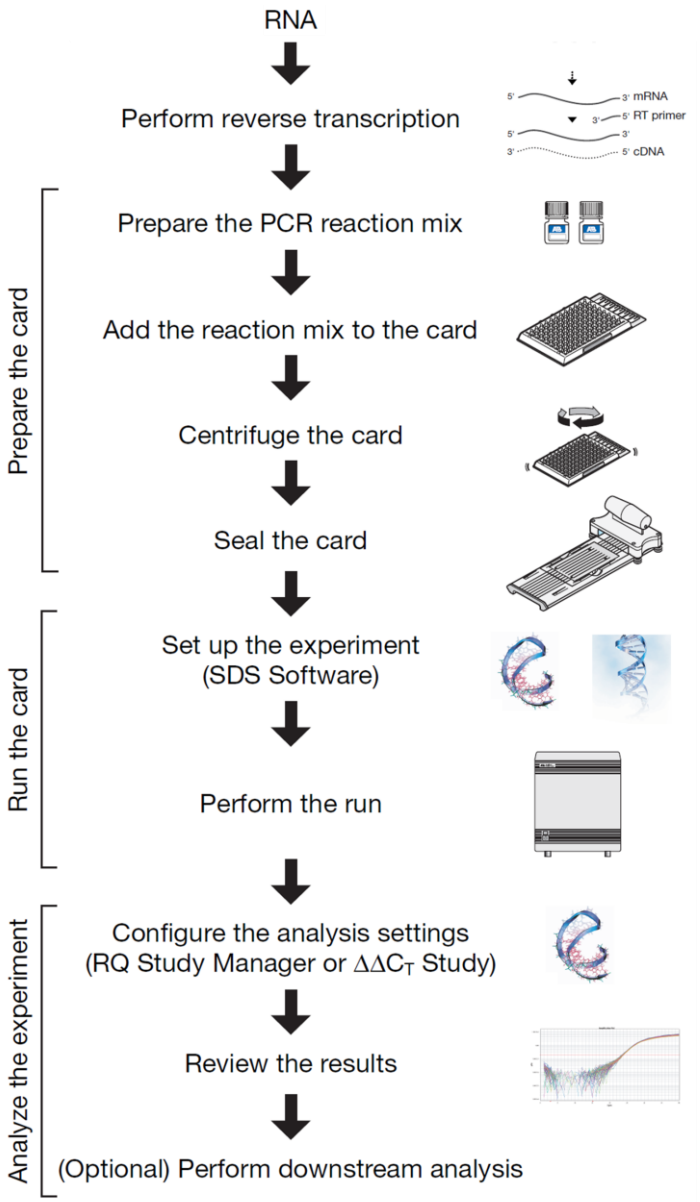

Figure 24. Workflow of the qPCR analysis (adapted from the TaqMan handbook).

4.2.5. Gene expression data analysis

Gene expression arrays

Quality Assessment was carried out in two steps. First, the Expression Console (Affymetrix, Santa Clara, CA) was used to check hybridization and processing controls. In the second step performed in R (v 3.1.1), scanned raw images and several quality plots were generated and checked. In particular, density plots, boxplots (raw and normalized intensities), Normalized Unscaled Standard Errors (NUSE) plots, Relative Log Expression (RLE) plots, dendrograms and Principal Component Analysis (PCA) were performed. After this first assessment, all $CD4^+$ and $CD8^+$ samples from the 38 subjects previously mentioned were considered for further analysis.

A Robust Multi-array Average (RMA) algorithm, included in the aroma.affymetrix package, was used for data normalization (Irizarry et al, 2003). To correct for batch effects, ComBat, implemented in the R package SVA, was employed (Johnson et al, 2007).

Differential expression analysis was performed using the Limma R package (Ritchie et al, 2015), which briefly applies linear models to detect differences between studied conditions. Pairwise t-test and Fisher's exact test showed none of the contrasts was influenced by age, but significant differences concerning sex were observed between CLL and healthy controls (P=0.036); therefore, sex was included in the linear model. In order to control for multiple testing, False Discovery Rate (FDR) adjustment was applied (Benjamini et al, 1995). However, non-adjusted P-values were finally considered in order to have enough and more robust results in all the comparisons. Thus, P-values <0.05, together with a Log Fold Change |logFC| >0.58 cut-off value (FC>1.5 or <-1.5), were used to

get the list of differentially expressed genes. Gene annotation was obtained from the Affymetrix annotation file Human Gene 2.0 ST array (NetAffix version na34, hg19). UCSC database was employed (Feb. 2009 hg19, GRCh37) to complement Affymetrix annotations.

Since the analysis of both $CD4^+$ and $CD8^+$ cell fractions revealed several genes typically expressed by B cells, the effect of a possible B cell contamination was corrected. With this purpose, the results of the gene expression arrays from B cells of the same cohort of subjects were used. Thus, those differentially expressed genes that were present in $CD4^+$ or $CD8^+$ cells and also showed a significant variation in the B cell fraction with the same logFC sign were excluded for further analysis. This filter was separately applied for the different comparisons (MBL vs. control, CLL vs. control and CLL vs. MBL).

In order to perform a functional analysis of the genetic signature of $CD4^+$ cells, Ingenuity Pathway Analysis (IPA) was employed, a software application that enables the integration of data from gene expression studies.

qPCR analysis

A first analysis of the three housekeeping genes (*ACTB, GUSB* and *GAPDH*) revealed that *GUSB* and *GAPDH* presented less variability and were subsequently selected for further analysis. Outlier replicates, defined by a standard deviation (SD) >0.5 and a qPCR threshold cycle (Ct) >40, were excluded. The ΔCt measures, defined as the mean of the expression of a studied gene minus the mean of the expression of the selected housekeeping genes, were compared using t-test. *P*-values <0.05 were considered significant.

4.3. Cytokine immunoassays

4.3.1. Sample processing

Peripheral blood was collected in tubes containing serum separation gel. After centrifuging for 15 minutes at 3700 rpm, two different phases separated by the gel were obtained. The upper phase, were the serum was located, was isolated, aliquoted and stored at -80ºC until used.

4.3.2. Immunoassays

The levels of the following 20 cytokines were measured: IL1β, IL2, IL4, IL5, IL6, IL8, IL10, IL12, IL15, IL17, IFNα, IFNγ, TNFα, GM-CSF, CCL3, CCL4, CCL19, CXCL10 and CXCL11 using the U-PLEX Platform (Meso Scale Discovery) and CXCL9 using the Human CXCL9/MIG Quantikine ELISA Kit (R&D Systems).

The assays performed with the U-PLEX Platform are based on biotinylated capture antibodies that are coupled to U-PLEX linkers, which self-assemble onto unique spots on the U-PLEX plate. These capture antibodies bind the analytes of the sample. Finally, detection antibodies conjugated with electrochemiluminescent labels bind the analytes to complete the sandwich immunoassay (Figure 25). Once completed, the plate is loaded into an instrument in which a voltage is applied, causing the captured labels to emit light. The intensity of the emitted light is proportional to the amount of analyte in the tested sample.

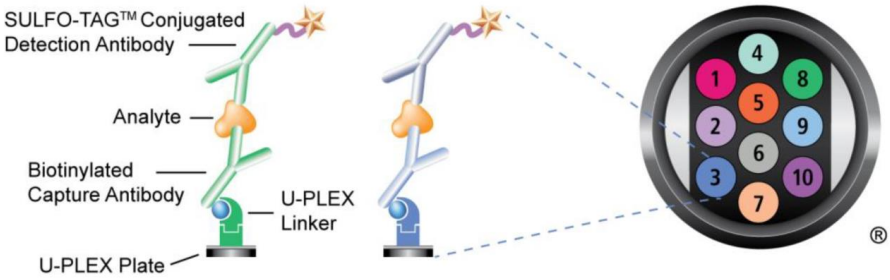

Figure 25. Representation of the sandwich immunoassay on a U-PLEX plate (obtained from the MSD U-PLEX Platform handbook).

The detailed procedure is described below:

1. All the reagents were brought to room temperature.

2. To create individual U-PLEX-coupled antibody solutions, 200 µL of each biotinylated antibody were added to 300 µL of the assigned linker. After mixing by vortexing, samples were incubated for 30 minutes at room temperature.

3. A total of 200 µL of Stop Solution were added to the previous samples. After vortexing, samples were incubated for 30 minutes at room temperature.

4. To prepare the multiplex coating solution, 600 µL of each U-PLEX-coupled antibody solution were combined into a single tube and vortexed. Up to 10 U-PLEX-coupled antibodies could be pooled. If combining less than 10 antibodies, Stop Solution was added to reach a final volume of 6 mL. The final solution was mixed by vortexing.

5. To coat the U-PLEX plate, 50 µL of the previous solution were added to each well of the U-PLEX plate, which was then sealed and incubated with shaking at room temperature for 1 hour.

6. Each well of the plate was washed 3 times with at least 150 µL of 1x PBS-T (PBS plus 0.05% Tween-20). The plate was then coated and ready for use.

7. To prepare the solutions required to obtain the calibration curve (Figure 26), the stock Calibrator solution was brought to room temperature and each vial was reconstituted by adding 250 µL of Diluent 43 (assay diluent). The reconstituted solution (5x) was equilibrated for 15-30 minutes at room temperature and then vortexed briefly.

8. To generate the Calibrator Standard 1 solution (1x, the high Standard), 50 µL of the reconstituted stock Calibrator solution were mixed with 200 µL of Diluent 43 and vortexed.

9. The subsequent 6 dilutions required to create the calibration curve were prepared by adding 75 µL of each previous solution to 225 µL of Diluent 43 (total volume: 300 µL). Between each serial dilution, tubes were vortexed.

10. The last solution for the calibration curve (zero Calibrator/blank) was prepared only with 300 µL of Diluent 43 (Figure 26).

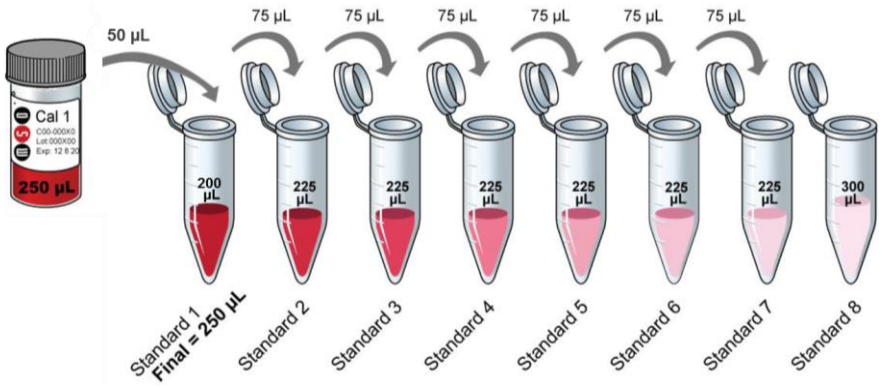

Figure 26. Preparation of the required dilutions to create the calibration curve. Standard 8 tube only had Diluent 43 (blank) (adapted from the MSD U-PLEX Platform handbook).

11. To prepare the detection antibody solution, 60 µL of the 100x stock solution were mixed with Diluent 3 to bring the final volume to 6 mL (1x working solution).
12. To prepare the read buffer T solution, 10 mL of the 4x stock were mixed with 10 mL of deionized H_2O (2x working solution).
13. To bind the analytes (cytokines) to the capture antibodies linked to the plates, firstly 25 µL of Diluent 43 were added to each well of the plate. The plate was gently tapped on all sides.
14. Then, 25 µL of each serum sample or the prepared Calibrator Standard were added to each well. Notably, two replicates of each sample were assayed. The plate was sealed and incubated at room temperature with shaking for 1 hour.
15. The plate was washed (x3) with ≥150 µL/well of 1x PBS-T.
16. A total of 50 µL of detection antibody solution were added to each well. The plate was sealed and incubated at room temperature with shaking for 1 hour.
17. The plate was washed (x3) with ≥150 µL/well of 1x PBS-T.
18. Finally, 150 µL of 2x read buffer T were added to each well. The plate was then ready to be analyzed on SECTOR Imager 2400 (Meso Scale Discovery).

A similar principle, based on a sandwich immunoassay technique, was applied for CXCL9 measurement as well. However, as opposed to the MSD U-PLEX technology, detection antibodies were not linked to electrochemiluminescent labels but to an enzyme. After adding a substrate that could be processed by the enzyme, color developed in proportion to the amount of CXCL9. The detailed procedure was as follows:
1. Reagents and samples were brought to room temperature.
2. Work solutions were prepared as follows:

- Wash Buffer: 20 mL of Wash Buffer Concentrate were added to deionized H_2O to prepare 500 mL of Wash Buffer work solution.
- Substrate Solution: Color Reagents A and B were mixed together in equal volumes within 15 minutes of use and protected from light.
- Calibrator Diluent RD5P: 10 mL of RD5P Concentrate were added to 10 mL of deionized H_2O to prepare 20 mL of RD5P (diluted 1:2).
- Human MIG (CXCL9) Standard: The Human MIG Standard was reconstituted with deionized H_2O, following the amount detailed in the vial label. As a result, a stock solution of 20000 pg/mL was produced. The Standard was mixed and sat for a minimum of 15 minutes with gentle agitation.
- Calibrator Standard dilutions: 900 µL of Calibrator Diluent RD5P were mixed with 100 µL of the stock Standard solution at 20000 pg/mL. As a result, the high Standard (2000 pg/mL tube) was generated. Then, 500 µL of the Calibrator Diluent RD5P were pipetted into the remaining tubes and a dilution series as shown in Figure 27 was produced. Calibrator Diluent RD5P served as the zero Standard (0 pg/mL).

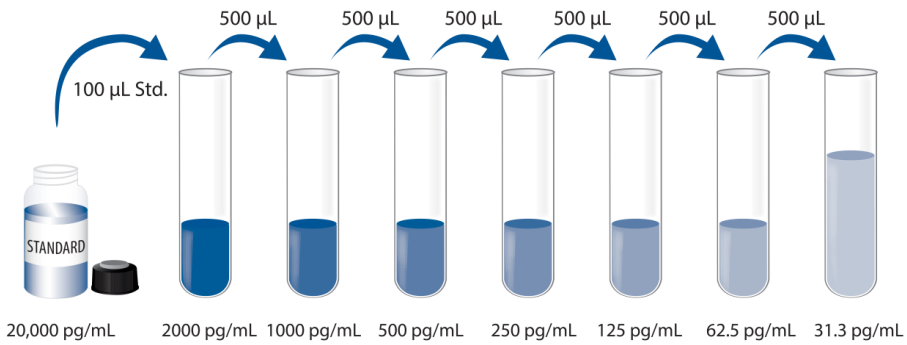

Figure 27. Dilution schema to create the calibration curve (obtained from the CXCL9/MIG Quantikine ELISA Kit handbook).

3. A total of 100 µL of Assay Diluent RD1W were added to each well of the plate.

4. Next, 100 µL of the sample or Standard were added to each well. The plate was covered with adhesive strip and incubated for 2 hours at room temperature. Samples were assayed in duplicate.

5. Each well was aspired and washed with 400 µL of Wash Buffer. The same process was repeated 3 times. After the last wash, the plate was inverted and blotted against clean paper towels.

6. Afterwards, 200 µL of Human MIG (CXCL9) Conjugate were added to each well. The plate was covered and incubated for 2 hours at room temperature.

7. The aspiration/wash step was repeated as in step 5.

8. Right after, 200 µL of Substrate Solution were added to each well. The plate was incubated for 30 minutes at room temperature, protected from light.

9. Next, 50 µL of Stop Solution were added to each well. The color in the wells changed from blue to yellow.

10. The optical density of each well was determined within the next 30 minutes using a microplate reader set to 450 nm (BioLinx 2.22, Dynex Technologies). Wavelength correction was set at 570 nm.

4.3.3. Data analysis

For each measured cytokine, three different plates had to be used due to the large number of studied cases. In order to check the reproducibility among plates, aliquots of the same sample from one U-CLL case were loaded onto the three plates. All values were given in pg/mL. Since samples were assayed in duplicate, mean values were considered.

Calibration curves were created using the values of the Standard solutions as shown in Figure 28. Those values below the lower limit of detection were replaced with the value of the lower limit of detection.

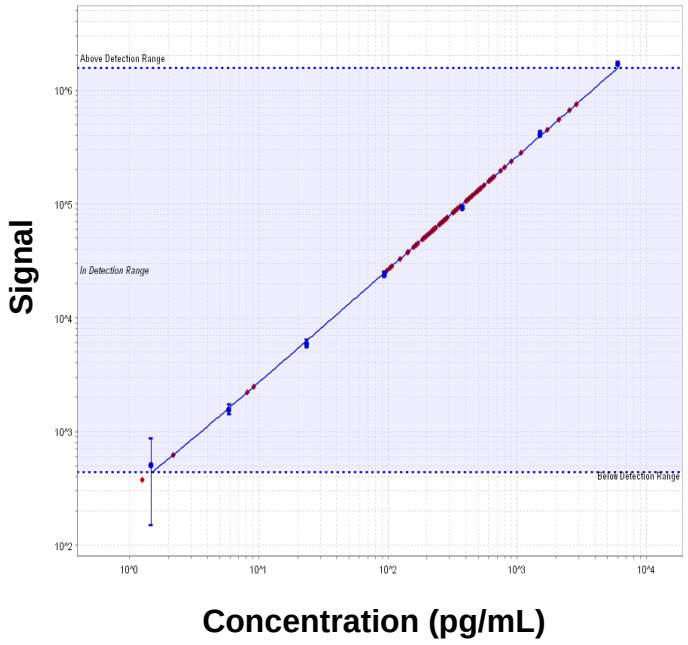

Figure 28. Calibration curve for CXCL10 (IP-10) measured in one plate. Values from Standard solutions are shown with blue dots. The created calibration curve is depicted as a blue line. Red dots correspond to the values of the tested samples. Blue dotted lines correspond to the limits of detection.

Concerning the statistical analysis, the Kruskal-Wallis test was used to compare multiple groups (healthy controls vs. MBL vs. CLL or MBL vs. M-CLL vs. U-CLL). In this case, $P<0.05$ was considered significant. Pairwise group comparisons were carried out using the Mann-Whitney test. In this case, in order to account for multiplicity in carrying out all the three pairwise comparisons on 20 cytokines, a Bonferroni-like adjustment was applied ($P<0.017$ was considered significant). Statistical analyses were performed using SPSS v.22 software (SPSS Inc., Chicago, IL, USA) and SAS, v9.3 (Cary, NC, USA).

4.4. TR repertoire analysis by subcloning/ Sanger sequencing

Subcloning is a useful strategy when characterizing a polyclonal or oligoclonal lymphoid cell population, since it allows the insertion of unique PCR products (in this case, unique TRBV-TRBD-TRBJ amplified rearrangements corresponding to unique T cells) into plasmid vectors. Then, vectors with the insert will transform competent *E.coli* bacteria in such a way that each bacterium will harbor a unique vector with a unique TRBV-TRBD-TRBJ rearrangement. After culturing the bacteria, isolated colonies will grow (each colony coming from a unique transformed bacteria), allowing the isolation of distinct TRBV-TRBD-TRBJ rearrangements which represent the clonal distribution of the T cell population of each analyzed individual. Once isolated and after re-amplification and purification steps, the insert of interest can be subjected to Sanger sequencing. The whole workflow of the TR subcloning analysis is shown in Figure 29.

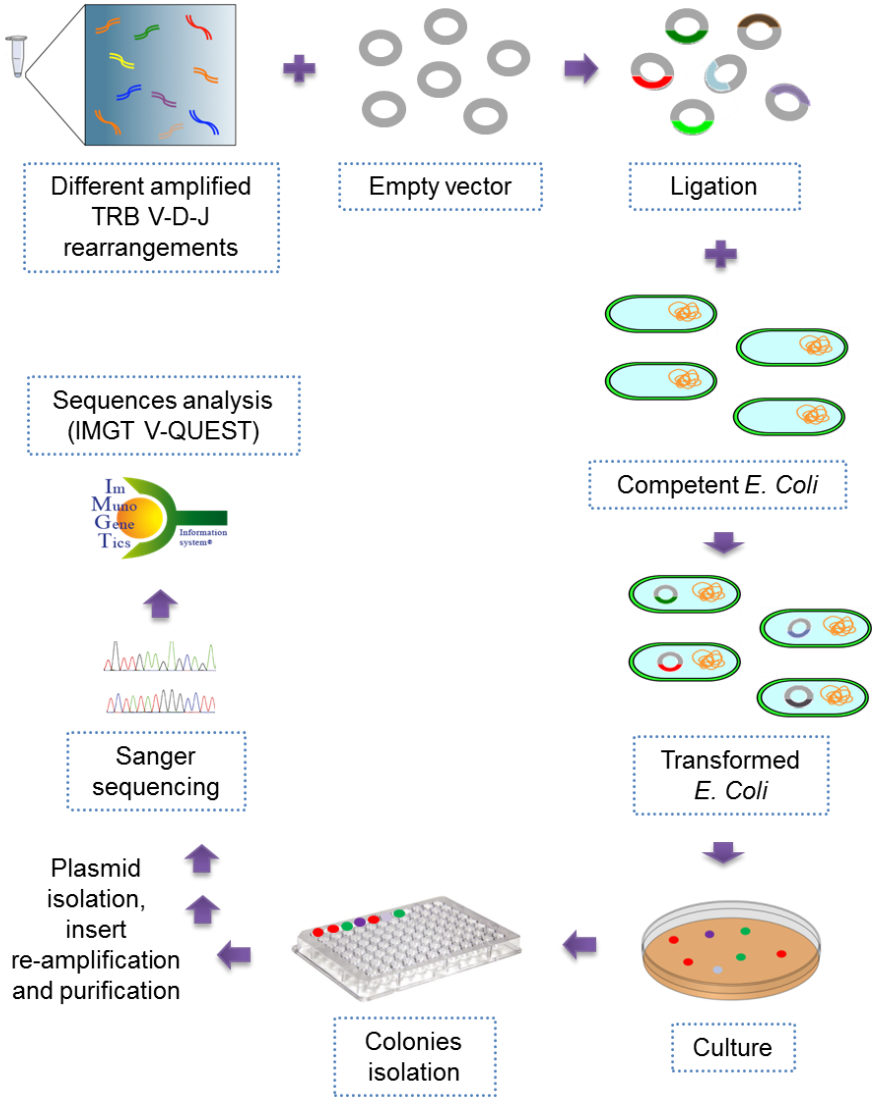

Figure 29. Workflow of the TR subcloning analysis.

4.4.1. Sample processing, RNA extraction and cDNA synthesis

Peripheral blood processing, purification of $CD4^+$ and $CD8^+$ cell fractions, RNA extraction and retrotanscription were performed as previously detailed (sections 4.2.1 to 4.2.3) even for the additional

MBL case not included in the gene expression arrays studies. However, for this additional MBL case, as well as for 3 of the 4 $CD4^+$ samples analyzed in sequential time points (not included in the gene expression arrays), the $CD4^+$ cell isolation process was performed by directly incubating PBMC, and not $CD19^-$ cells, with anti-CD4 antibodies. The median purity of all $CD4^+$ and the $CD8^+$ cell fractions included in the TR subcloning analysis was 95%.

For all sequential samples, the retrotranscription step was performed differently, as follows:

1. 1000 ng of RNA were diluted in 19 µL of DEPC H_2O.
2. RNA was denatured at 65ºC for 5 minutes.
3. 12 µL of M-MLV RT (200U/µL) and 6 µL of RNAsin (40U/µL) were added to 175 µL of the retrotranscription mix detailed in Table 6 (final volume: 193 µL).

Reagent	Volume (µL)
DEPC H_2O	420
Buffer 5x first strand	428
DTT 0.1M	21.5
dNTPs 25 mM	85.5
Hexanucleotide mix	65
Total volume: 1020 µL	

Table 6. Retrotranscription mix

4. A total of 21 µL of the obtained solution were mixed with the 19 µL of the previously denatured RNA (final volume: 40 µL).
5. Retrotranscription was achieved after subjecting the samples at 37ºC for 2 hours and 65ºC for 10 minutes. The obtained cDNA samples were held at 4ºC or at -20ºC for long storage.

4.4.2. PCR of TRBV-TRBD-TRBJ gene rearrangements

PCR reaction

PCR amplification of TRBV-TRBD-TRBJ gene rearrangements was performed according to the BIOMED2 protocol (Van Dongen et al, 2003) as follows:

1. The tube with the mix of primers and other reagents to amplify the different TRBV-TRBD-TRBJ rearrangements (TCRB Tube A – Unlabeled, Invivoscribe) and the tube with the positive control for clonality (IVS-0009 Clonal Control DNA, Invivoscribe) were melted at room temperature.

2. On ice, 45 µL of Tube A solution were mixed with 5 µL of the previously obtained cDNA and 0.5 µL of AmpliTaqGold (Applied Biosystems). Additional tubes with 5 µL of clonal control DNA or 5 µL of H_2O were used as positive and negative controls, respectively.

3. The obtained solutions were subjected to thermal cycling, as follows:

94ºC 5 min
94ºC 45 sec
60ºC 45 sec
72ºC 90 sec
} x35
72ºC 10 min
Hold at 18ºC

Gel purification of the PCR products

Next, PCR products were gel-purified (QIAquick Gel Extraction Kit, QIAGEN) according to the following procedure:

1. A 1% agarose gel with BrEt staining was prepared. The total volumes of the samples with the PCR product (50.5 µL) as well as

the positive and negative control samples were loaded into the gel. Samples run for 90 minutes at 100V.

2. Once finished, the gel was visualized with UV illumination (Figure 30). Specific PCR products (240-285 bp) were then excised and placed individually into Eppendorf tubes.

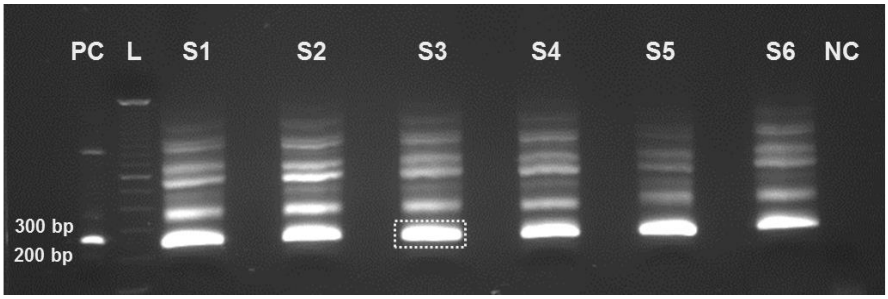

Figure 30. Visualization of the electrophoresis gel in which the amplified TRBV-TRBD-TRBJ rearrangements were loaded. The rectangle with dotted lines represents the slice of gel that was excised in a particular case. PC: positive control for clonality; L: ladder; S: test sample; NC: negative control; bp: base pairs.

3. A total of 600 µL of QG buffer were placed into each Eppendorf tube with the gel fragments. After vortexing, tubes were placed in a bath at 50ºC for 20 minutes or until the gel slices were completely dissolved.

4. From 200 to 600 µL of isopropanol were added to each tube and mixed until DNA precipitation.

5. Then, the samples were introduced into the columns provided by the kit, which were placed into 2 mL collection tubes and centrifuged for 2 minutes at full speed.

6. The flow-through was discarded and the columns were placed into the same collection tubes. Right after, 500 µL of QG buffer were added again to definitely dissolve all traces of agarose. Tubes were then centrifuged for 2 minutes at full speed.

7. After discarding the flow-through, samples were washed by adding 750 µL of PE buffer into each column. After 5 minutes of incubation at room temperature, tubes were centrifuged for 2 minutes at full speed.
8. In order to completely remove residual buffer, the flow-through was discarded and the same centrifugation step was performed.
9. The columns were placed into new Eppendorf tubes and 40 µL of elution buffer (EB) were added. After incubating for 5 minutes at room temperature, samples were centrifuged for 2 minutes at full speed.
10. DNA concentrations were measured and samples were held at 4ºC or at -20ºC for long storage.

4.4.3. Subcloning and Sanger sequencing

Ligation and transformation

Both steps were performed employing the TA Cloning Kit, with pCR2.1 Vector and One Shot TOP10F' Chemically Competent *E. coli* (Invitrogen). In order to insert separately each amplified TRBV-TRBD-TRBJ rearrangement into a plasmid vector, the following procedure was applied:
1. For each case, a ligation mix as detailed in Table 7 was prepared (reagents were supplied by the kit mentioned above).

Reagent	Volume (µL)
Sterile H_2O	3 to 4 µL
5x T4 DNA ligase buffer	2
pCR2.1 vector (25 ng/µL)	2
T4 DNA ligase	1
PCR product	1 to 2 µL
Total volume: 10 µL	

Table 7. Ligation mix.

2. The ligation reaction was then incubated at 14ºC overnight. Once finished, samples were stored at -20ºC or subjected to transformation.

To achieve the transformation of the competent *E.Coli*, the protocol below was followed:

1. The vials with the competent *E.Coli* provided by the kit were thawed on ice.
2. After a brief centrifugation, 2 µL of each ligation mix were added to 25 µL of the competent cells and stirred gently with the pipette.
3. The obtained samples were then incubated on ice for 30 minutes.
4. Right after, a heat shock step was applied: the samples were placed for 30 seconds at 42°C. Once finished, the vials were immediately transferred back to ice and incubated there for 10 minutes.
5. A total of 125 µL of room temperature S.O.C. medium were added to each vial.
6. The samples were then incubated horizontally at 37°C for 90 minutes in a shaking incubator at 200 rpm.
7. After incubation, the 152 µL of each sample were spread on LB agar (Invitrogen) plates containing 100 µg/mL of ampicillin. Before adding the sample, 20 µL of both X-Gal (40 mg/mL) and IPTG (100 mM) were spread over the agar plates. Ampicillin, X-Gal and IPTG were required for the selective isolation of the colonies, as explained in Figure 31.
8. Plates were incubated overnight at 37ºC. Before the isolation of the colonies, plates were placed at 4°C for 2-3 hours to allow for proper color development.

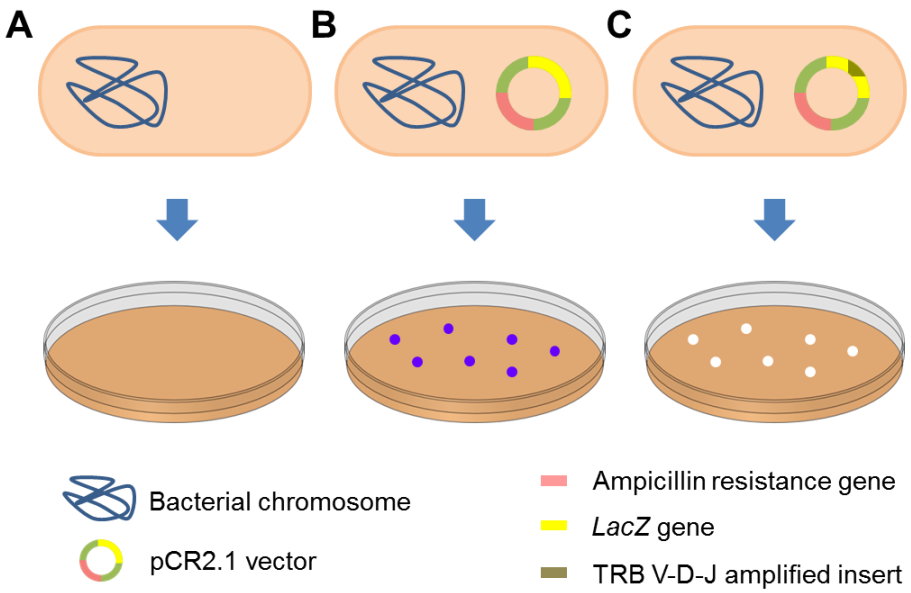

Figure 31. Schematic representation of the selective isolation method. In addition to LB agar, plates had IPTG (which induces *LacZ* transcription), X-gal (substrate of the enzyme codified by *LacZ*) and ampicillin. A) *E.coli* in which transformation did not occur do not grow since they do not have ampicillin resistance. B) *E.coli* transformed with the empty vector have ampicillin resistance and metabolize X-gal; as a result, colonies will grow with a blue color. C) *E.coli* transformed with the inserted vector have ampicillin resistance but do not metabolize X-gal, since the insert disrupts *LacZ* gene; as a consequence, colonies will grow with a white color.

Isolation of the colonies and the vector DNA

Once the colonies acquired the proper white or blue color, white colonies were picked up and the vector DNA was isolated as follows:

1. With a sterile pipette tip, each different colony was transferred to a unique well of a 96 multi-well culture plate with 200 µL of LB broth base (Invitrogen) containing 100 µg/mL of ampicillin. For

each sample ($CD4^+$ or $CD8^+$) of each analyzed subject, a total of 48 white colonies were picked up, if possible.
2. Once finished, the plate was cultured at 37ºC overnight.

The following day, in order to isolate the vector DNA, the protocol below was applied:
1. With a multi-pipette, the medium of each well of the culture plate was mixed and then transferred to a MW96 PCR plate.
2. The MW96 PCR plate was covered with a film and centrifuged for 10 minutes at 3700 rpm.
3. After centrifugation, precipitated DNA could be observed at the bottom of each well (vector + bacterial DNA). The supernatants were discarded and 40 µL of H_2O were added to each well to clean up. The plate was covered with film again and centrifuged for 10 minutes at 3700 rpm.
4. The film was removed and the supernatant discarded. 20 µL of H_2O were added to each well, but this time by mixing and resuspending the DNA pellet.
5. The plate was covered with sterilization tape and incubated for 10 minutes at 99°C followed by 10 minutes at 4°C.
6. After the incubation, the plate was centrifuged for 10 minutes at 3700 rpm.
7. As a result of steps 5 and 6, a supernatant with the vector DNA and a pellet with the bacterial chromosome were obtained. The sterilization tape was removed from the plate and the supernatants of each well, with the plasmid DNA, were isolated and transferred to a new MW96 PCR plate.

PCR re-amplification of the TRB V-D-J insert

Prior to sequence, the TRB V-D-J rearrangement must be re-amplified. Since at this point it is inserted into the *LacZ* region of the vector, specific M13 primers, located at both sides of the insert, were employed. The performed protocol was as follows:

1. On ice, a master mix as detailed in Table 8 was prepared:

Reagent	Volume (µL) x1 reaction
Buffer 10x	2.5
$MgCl_2$ (25 mM)	0.75
Primer M13 F (10 µM)	1.5
Primer M13 R (10 µM)	1.5
dNTPs (10 mM)	0.5
TaqPlatinum (Invitrogen)	0.25
H_2O	16
Total volume (x1): 23 µL	

Table 8. M13 PCR master mix.

2. On ice, 23 µL of master mix were added into each well of a MW96 PCR plate.

3. Then, 2 µL of the samples with isolated vector DNA were added separately to each well of the MW96 PCR plate.

4. The plate was subjected to thermal cycling, as follows:

95ºC 4 min
94ºC 40 sec
60ºC 30 sec } x35
72ºC 90 sec
72ºC 10 min
Hold at 4ºC

5. The PCR products were checked prior to sequencing in a 1% agarose gel with BrEt staining as shown in Figure 32.

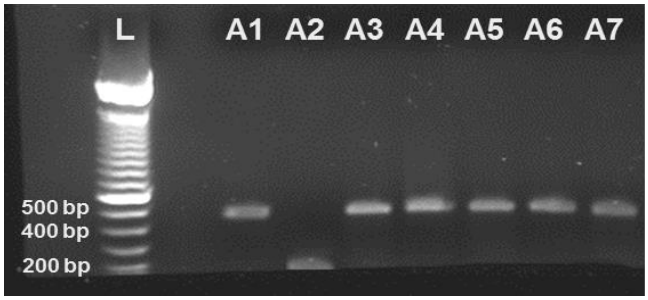

Figure 32. Visualization of the electrophoresis gel in which the re-amplified TRB V-D-J rearrangements were loaded. Notably, in addition to the 240-285 bp fragment corresponding to the TRB V-D-J rearrangement, these amplicons also have 199 bp from the pCR2.1 vector, resulting in a PCR product of approximately 450 bp. All the cases of this example (A1-A7) were correctly amplified except for A2, in which the vector did not have the insert of interest. L: ladder; bp: base pairs.

Purification of the PCR product and Sanger sequencing

Only the correctly amplified PCR products were subjected to purification and Sanger sequencing, as follows:

1. A total of 75 µL of 0.01M Tris-EDTA buffer were added to each well.
2. After mixing, the content of each well was transferred to a filter plate (MultiScreen-PCRµ96 Filter Plates, Millipore).
3. The filter plate was then connected to the vacuum system and filtered for 7 minutes.
4. After this, 25 µL of H_2O were added to each well, the plate was covered with parafilm and shaken for 10 minutes at 500 rpm.
5. The content of each well was then transferred to a new MW96 PCR plate and covered with film.

Once finished, the plate was stored at -20ºC until Sanger sequencing, which was performed as described below:

1. A master mix as detailed in Table 9 was prepared.

Reagent	Volume (µL) x1 reaction
BigDye mix (Thermo Fisher)	0.25
DNA Sequencing Buffer 5x	2
Primer M13 F (1 µM)	3.2
H_2O	4.40
Total volume (x1): 9.85 µL	

Table 9. Master mix for Sanger sequencing.

2. On ice, 9.85 µL of the master mix were placed into each well of a PCR MW96 plate.

3. Later, 0.15 µL of the purified PCR product (5-10 ng) were separately added to each well of the plate with the master mix.

4. The plate was covered with sterilization tape and subjected to thermal cycling, as follows:

94ºC 3 min
96ºC 10 sec
50ºC 5 sec
60ºC 4 min
} x40
Hold at 4ºC

5. The produced samples were submitted to the Genomics Core Facility of the Universitat Pompeu Fabra (Barcelona) for capillary electrophoresis.

4.4.4. Data analysis, definitions and interpretation

After removing the nucleotides corresponding to the pCR2.1 vector, subcloned sequences were introduced into the IMGT/V-QUEST tool (http://www.imgt.org) from the international ImMunoGeneTics information system for analysis. Only productive TRBV-TRBD-TRBJ rearrangements were considered (n= 2567).
Some definitions are required for the analysis and interpretation of the obtained data (Figure 33):

- Clonotype: Unique rearrangements carrying identical TRBV-TRBJ genes and CDR3 amino acid sequence.
- Expanded clonotype: When corresponding to ≥2 sequences within a sample (in clusters).
- Immunodominant (predominant) clonotype: The most expanded clonotype within each sample.
- Relative frequency of each clonotype within a sample: The number of rearrangements corresponding to a specific clonotype divided by the total number of rearrangements sequenced for that specific sample.
- Cumulative (relative) frequency of all expanded clonotypes within a sample: The number of rearrangements corresponding to all expanded clonotypes divided by the total number of rearrangements sequenced for that specific sample.

For TRBV gene repertoire studies, clonotypes instead of single rearrangements were considered in order to avoid skewing due to T cell expansions after antigenic stimulation (Figure 33).

In order to compare the results among independent groups (MBL vs. CLL), the Mann-Whitney test was applied, whereas $CD4^+$ and

$CD8^+$ T cell fractions, both belonging to the same subject, were compared using the Wilcoxon test. Simple linear regression analysis and Spearman correlation were calculated to analyze the relationship between T cell clonality and B cell clonal size. In all cases, $P<0.05$ was considered statistically significant. Statistical analyses were performed using SPSS v.22 software (SPSS Inc., Chicago, IL, USA).

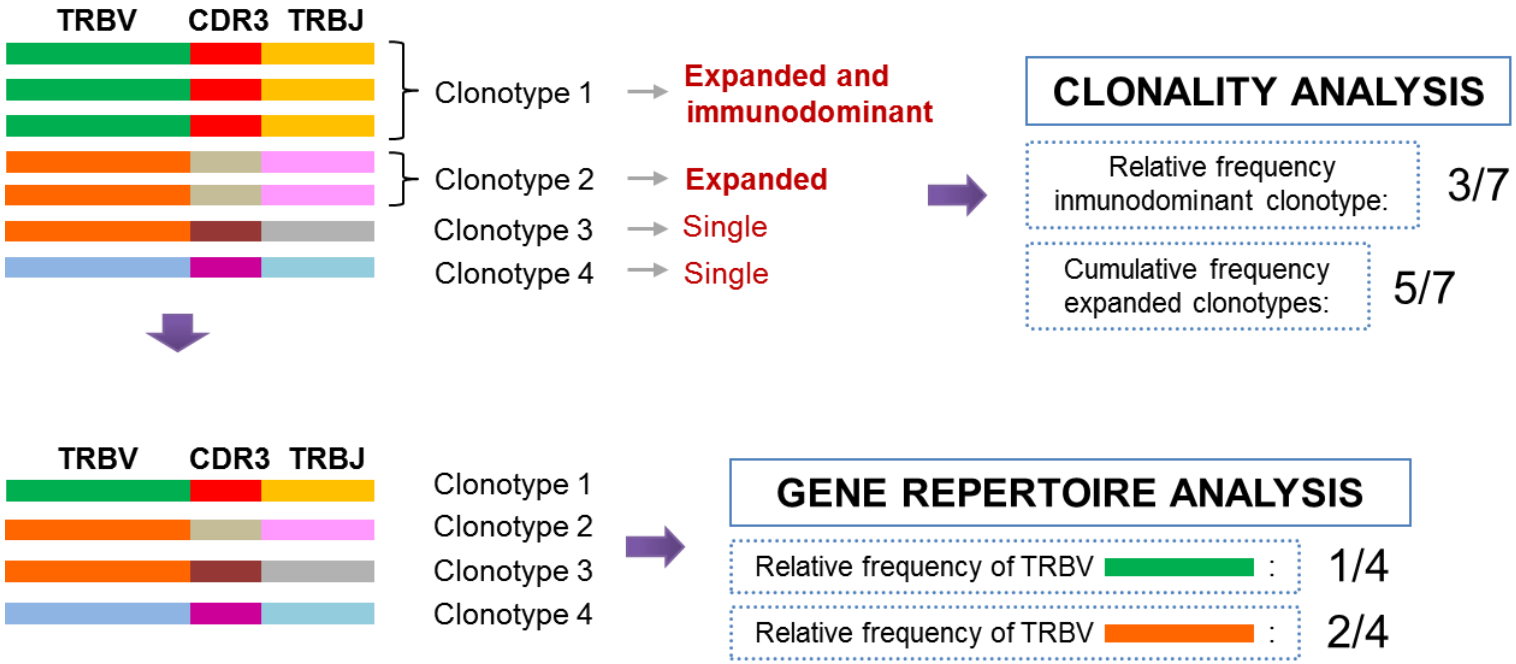

Figure 33. Simplified representation of the clonality and gene repertoire analysis.

4.4.5. Comparison to public data

Clonotype comparison across all MBL and CLL individuals was performed. Clonotypes extracted from $CD4^+$ and $CD8^+$ T cell samples or sequential time points of the same individual were concatenated in order to include all clonotypes belonging to the same subject in the comparison analysis. A panel of 5264 productive, unique and well-annotated TRBV-TRBD-TRBJ rearrangements from T cell clones of different entities (available to Dr. Stamatopoulos group, n=1262; or extracted from the IMGT/LIGM-DB sequence database, n=4002, http://www.imgt.org/IMGTindex/LGM.html) was also compared with all the clonotypes of this series.

5. RESULTS

5.1. Gene expression and cytokine studies

5.1.1. Gene expression arrays

5.1.1.1. Patient characteristics

The demographic, clinical and immunogenetic characteristics of the subjects included in the gene expression arrays studies are detailed in Table 10.

	Healthy controls (n=9)	**MBL (n=15)**	**CLL (n=14)**
Age (years)	64 (60-80)	76 (60-83)	71.5 (59-88)
Males*	2/9 (22.2%)	5/15 (33.3%)	10/14 (71.4%)
Absolute lymphocyte count ($x10^9$/L)	1.2 (0.5-1.8)	5 (1.2-7.2)	11.2 (8-18.5)
Clonal B cells ($x10^9$/L)	_	1.7 (0.2-4.3)	8.5 (5.8-15.9)
Rai 0	_	_	14/14 (100%)
IGHV mutated	_	12/14 (85.7%)	12/14 (85.7%)

Table 10. Characteristics of the cohort analyzed by gene expression arrays. Values are shown as median (range) or number of cases/total number of evaluated cases (percentage). *Significant differences were observed between CLL and healthy controls (*P*=0.036); therefore, the analysis was adjusted by sex.

5.1.1.2. Clustering analysis

As a first approach to the gene expression data obtained by the gene expression arrays, an unsupervised hierarchical cluster analysis was performed using the Ward's linkage method and correlation based

distance. This analysis showed the association among the $CD4^+$ and $CD8^+$ samples of all studied subjects. The produced dendrogram is represented in Figure 34.

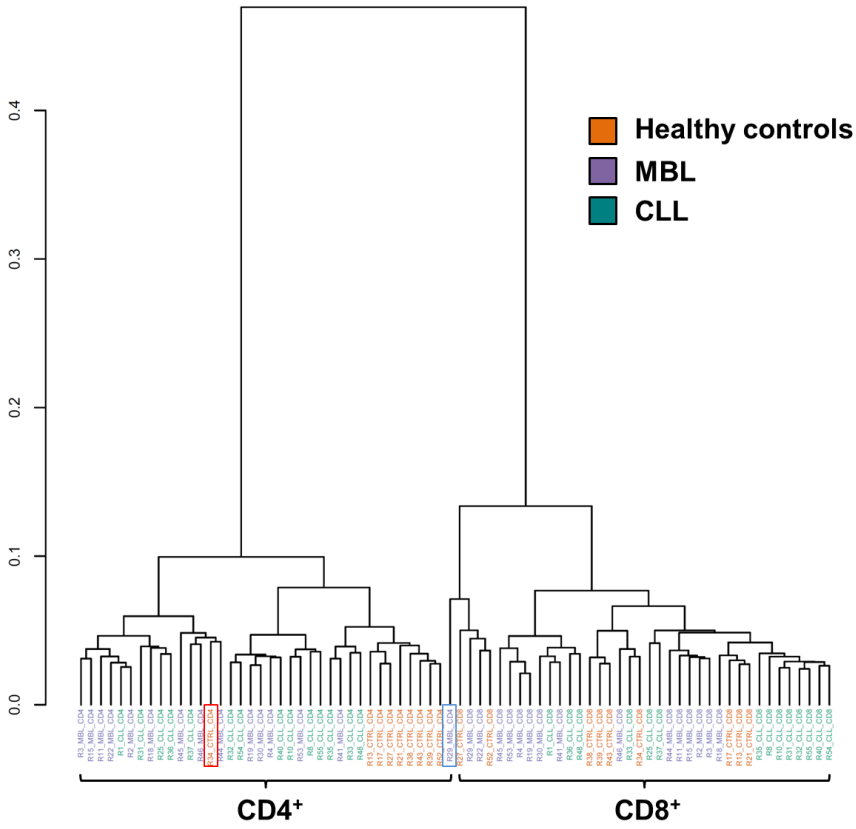

Figure 34. Dendrogram showing the association among all samples included in the gene expression arrays using the Ward's clustering method. $CD4^+$ and $CD8^+$ samples were clearly separated into two distinct groups. A $CD4^+$ healthy case that atypically grouped together with MBL and CLL is labelled in red, whereas a $CD4^+$ MBL case that atypically appeared in the same branch as $CD8^+$ samples is labelled in blue.

As observed, $CD4^+$ and $CD8^+$ samples could be clearly differentiated into two distinct groups. When focusing on the diagnosis (healthy subjects, MBL or CLL), differences between the two studied cell

fractions were found (Figure 34). Within the $CD4^+$ cell fraction, healthy controls grouped separately from MBL and CLL cases, except for one healthy case that appeared together with MBL and CLL. $CD4^+$ samples from MBL and CLL subjects were mixed together and separated from healthy controls, except for a MBL sample that grouped closer to healthy controls and actually belonged to the $CD8^+$ branch of the dendrogram. However, no clinical, biological or other reasons related to sample processing such as purity could be found to explain the distinct association of these $CD4^+$ samples and thus they could not be excluded from the analysis. As for the $CD8^+$ cell fraction, although groups of samples belonging to the same diagnostic category were detected, a mixed pattern was generally observed (Figure 34).

The $CD4^+$ MBL case that grouped closer to the $CD8^+$ samples could also be identified as an outlier by performing a principal component analysis (PCA). Notably, with this approach, it could be appreciated that this sample actually grouped separately from either $CD4^+$ or $CD8^+$ cases, as shown in Figure 35.

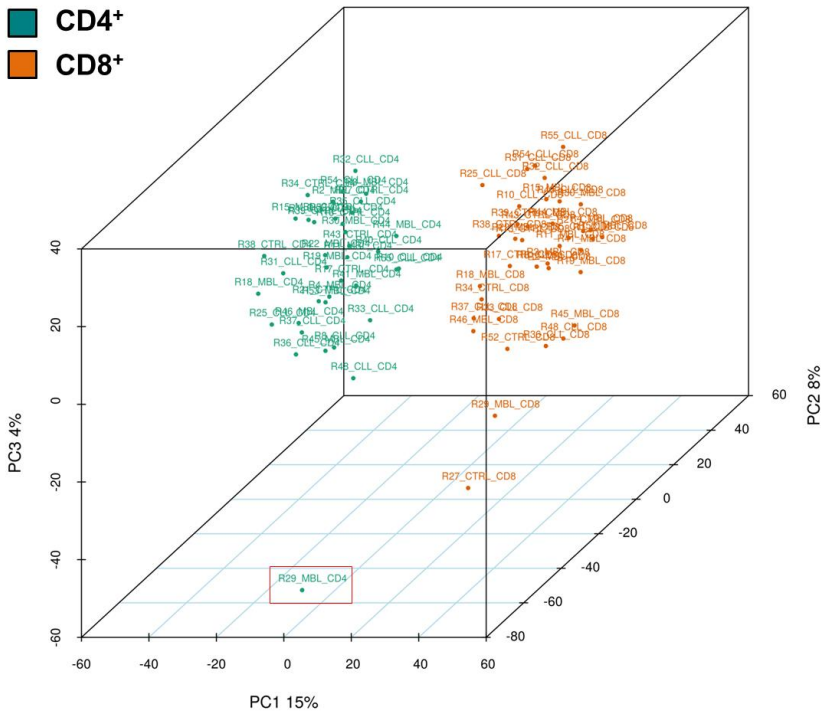

Figure 35. Principal component analysis (PCA) of the $CD4^+$ and $CD8^+$ samples included in the gene expression arrays. Although both cell types could be clearly differentiated, diagnostic groups appeared mixed. The $CD4^+$ MBL outlier that grouped closer to the $CD8^+$ samples in the dendrogram produced using the Ward's clustering approach is labeled in red.

5.1.1.3. Analysis of the differentially expressed genes among groups

Next, a gene expression comparison among the different diagnostic categories was carried out. Supervised heatmaps corresponding to pairwise comparisons are shown in Figures 36 ($CD4^+$) and 37 ($CD8^+$). Regarding the $CD4^+$ cell fraction, the best discrimination was observed for MBL vs. healthy controls. The separation between CLL vs. healthy subjects and CLL vs. MBL was not so clear (Figure 36). Concerning the $CD8^+$ cell fraction, the best discrimination was

detected for CLL vs. healthy controls, followed by MBL vs. healthy controls and finally CLL vs. MBL (Figure 37).

As for the $CD4^+$ cell fraction, a total of 247 genes were found to be differentially expressed (*P*<0.05 and FC>1.5 or <-1.5) between MBL and healthy controls, 90 between CLL and healthy controls and 59 between CLL and MBL. Concerning the $CD8^+$ cell fraction, 132 genes were differentially expressed between MBL and healthy controls, 138 between CLL and healthy controls and 126 between CLL and MBL (Figure 38). The detailed number of overexpressed and underexpressed genes for each comparison is summarized in Table 11.

	$CD4^+$			**$CD8^+$**		
	Over Exp.	**Under Exp.**	**Total**	**Over Exp.**	**Under Exp.**	**Total**
MBL vs. healthy controls	214	33	247	89	43	132
CLL vs. healthy controls	69	21	90	107	31	138
CLL vs. MBL	38	21	59	97	30	126*

Table 11. Number of differentially overexpressed and underexpressed genes for each cell type and comparison. **TRDV3* showed two different AffyIDs in this comparison, one of them overexpressed and the other underexpressed.

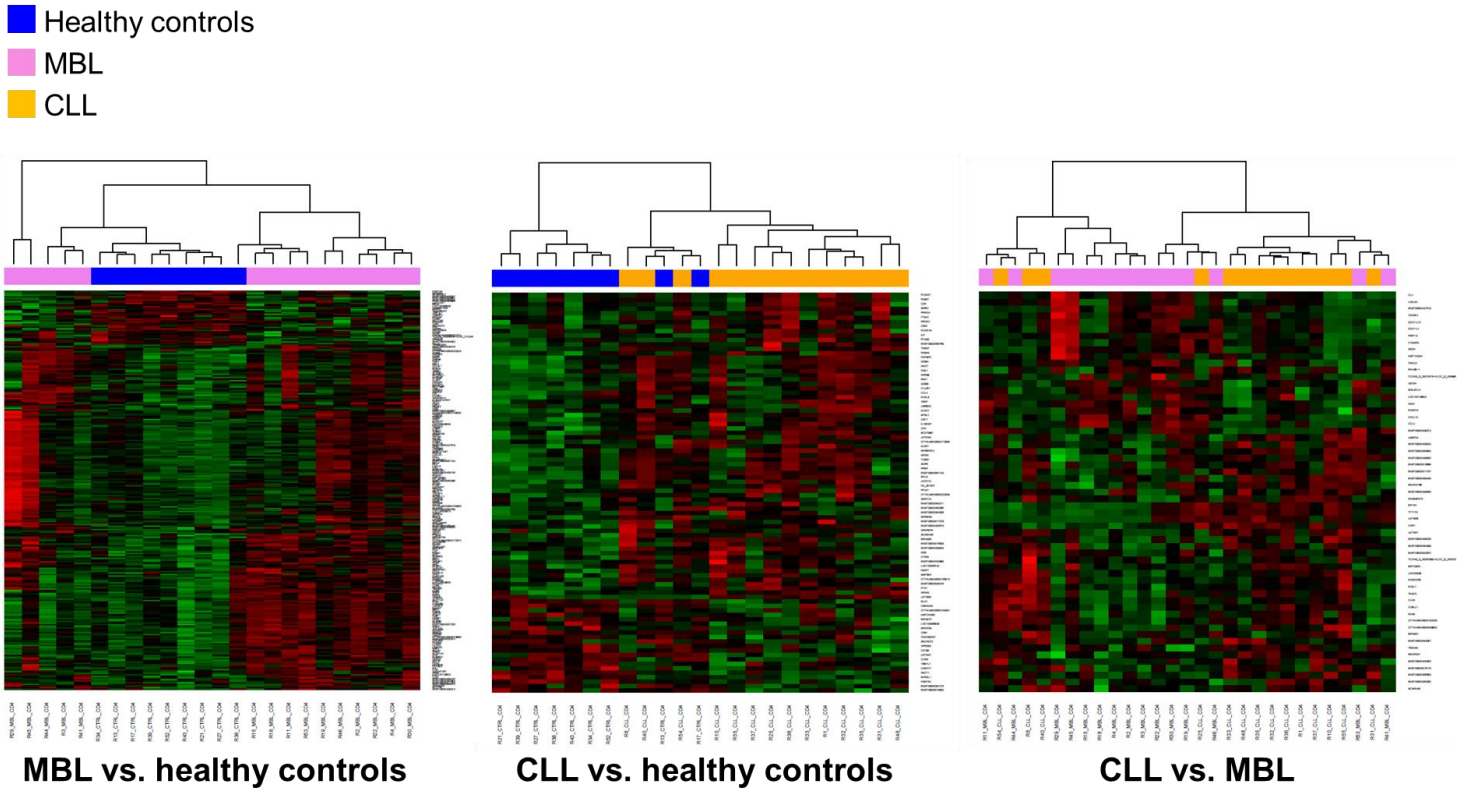

Figure 36. Supervised heatmaps for pairwise comparisons within the $CD4^+$ cell fraction. The differentially expressed genes of each comparison are represented in the different rows (P<0.05 and FC>1.5 or <-1.5) whereas each subject is shown as a different column. Red represents high expression whereas green represents low expression.

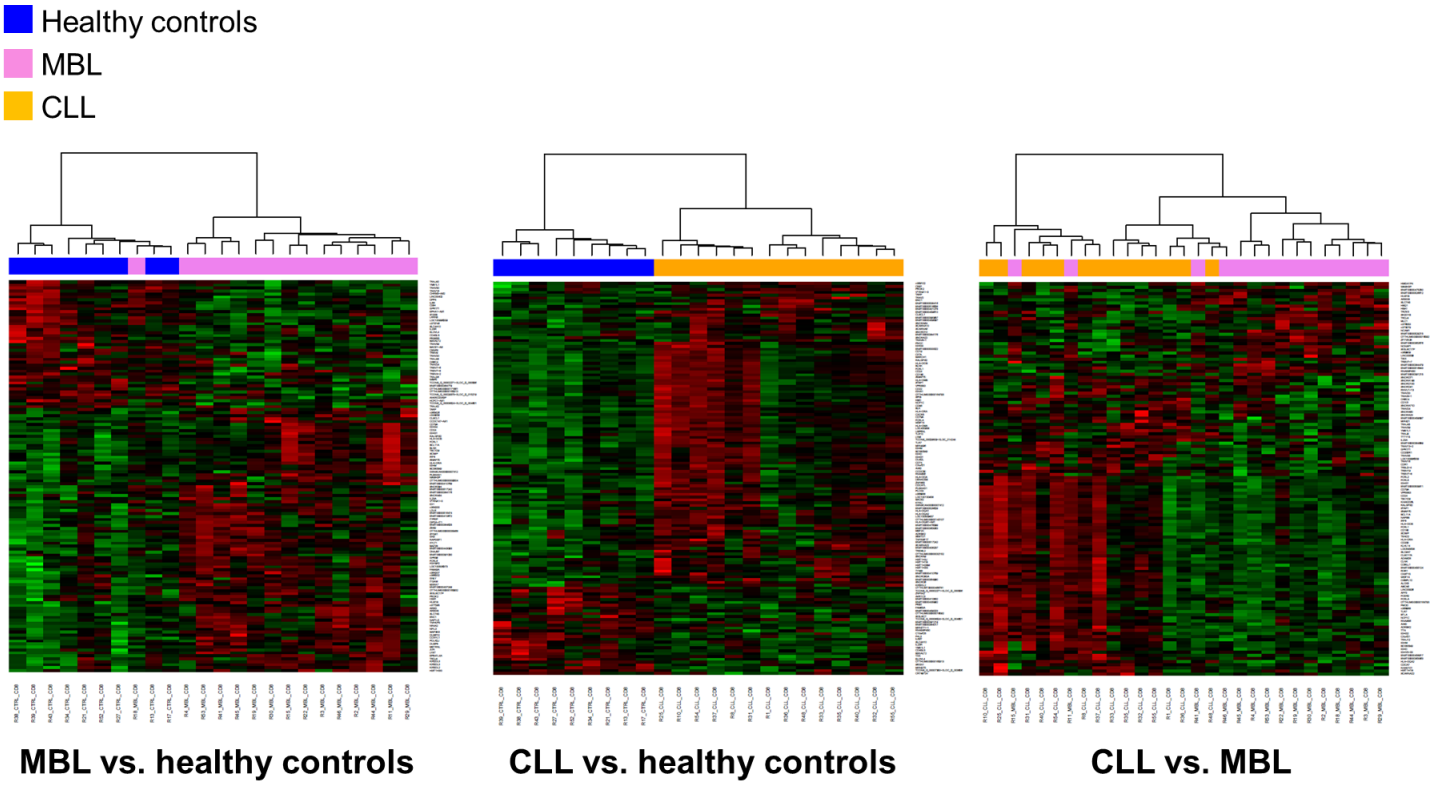

Figure 37. Supervised heatmaps for pairwise comparisons within the $CD8^+$ cell fraction. The differentially expressed genes of each comparison are represented in the different rows (*P*<0.05 and FC>1.5 or <-1.5) whereas each subject is shown as a different column. Red represents high expression whereas green represents low expression.

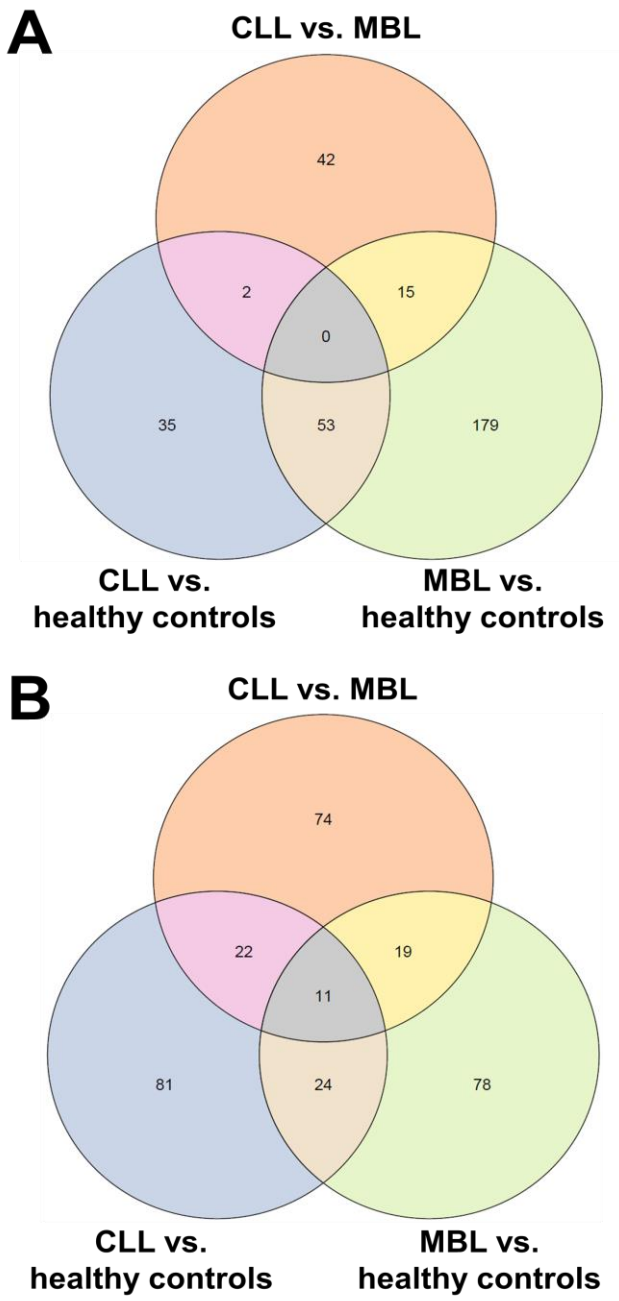

Figure 38. Venn diagrams showing the number of differentially expressed genes for the $CD4^+$ (A) and $CD8^+$ (B) cell fractions. Each pairwise comparison is represented by an independent circle. Overlapping areas represent the shared number of differentially expressed genes between independent comparisons.

Detailed analysis of the differentially expressed genes for the $CD4^+$ cell fraction revealed that many of them were expressed by myeloid cells rather than $CD4^+$ T cells. This is reasonable taking into account that in addition to helper T cells, myeloid cells such as monocytes and macrophages also express $CD4^+$ molecules. A detailed list of genes related to the inflammatory response and cell migration that were found to be differentially expressed between groups for the $CD4^+$ cell fraction is summarized in Table 12. As observed, most of the differentially expressed genes were detected in the MBL vs. healthy controls comparison, being overexpressed in MBL compared to controls. Some of these genes, such as *IL1B*, *CXCL8* (coding for IL8), *IL18* or *TREM1*, are key mediators of inflammatory pathways. A few of them were also upregulated in CLL compared to controls, although the number of differentially expressed genes and their FC values were lower than in the previous comparison. In this sense, *CCL2*, *CXCL10*, *VNN1* and *NEXN* were upregulated in MBL compared to CLL (Table 12).

Compared to healthy controls, $CD4^+$ cells of MBL and CLL also displayed upregulated genes related to cytotoxic pathways (*GZMB*, *PRF1*, *GNLY*, *FGFBP2, FCRL6* or *NKG7*). Again, FC values were higher in MBL vs. control than in CLL vs. control (Table 13).

Finally, several genes previously associated with a Th1 gene signature (Ono et al, 2014) were also upregulated in $CD4^+$ cells of both MBL and CLL compared to healthy subjects (Table 14).

	Gene	Description	Fold Change		
			MBL vs. control	CLL vs. control	CLL vs. MBL
Soluble factors	*IL1B*	Interleukin 1, beta	2.19		
	CXCL8	Interleukin 8	1.91	1.86	
	IL18	Interleukin 18	1.57		
	CCL2	Chemokine (C-C motif) ligand 2	2.07		-1.76
	CCL4	Chemokine (C-C motif) ligand 4	2.69	1.82	
	CCL5	Chemokine (C-C motif) ligand 5	1.68		
	CXCL5	Chemokine (C-X-C motif) ligand 5	2.46		
	PPBP	Chemokine (C-X-C motif) ligand 7	3.17	2.81	
	CXCL10	Chemokine (C-X-C motif) ligand 10			-1.65
Receptors	*TLR2*	Toll-like receptor 2	1.50		
	IL18RAP	Interleukin 18 receptor accessory protein	1.94	1.73	
	TREM1	Triggering receptor expressed on myeloid cells 1	1.56		
	CCR1	Chemokine (C-C motif) receptor 1	1.61		
	CX3CR1	Chemokine (C-X3-C motif) receptor 1	1.53		
	FCGR1A	Fc fragment of IgG, high affinity Ia, receptor (CD64)	1.51		
Cell dynamics	*RAB31*	RAB31, member RAS oncogene family	2.54	1.81	
	MYL9	Myosin, light chain 9, regulatory	2.73	2.12	
	VNN1	Vanin 1	1.97		-1.79
	NEXN	Nexilin (F actin binding protein)	1.68		-1.53
	TUBB1	Tubulin, beta 1 class VI	2.93	2.02	
	ITGAM	Integrin, alpha M	1.83		
	ITGB3	Integrin, beta 3	2.23		

Table 12. Differentially expressed genes related to the inflammatory response and cell migration for the $CD4^+$ cell fraction.

Gene	Description	Fold Change	
		MBL vs. control	CLL vs. control
GZMB	Granzyme B	2.27	1.80
GZMH	Granzyme H	4.06	3.44
PRF1	Perforin 1	1.64	1.55
GNLY	Granulysin	2.55	1.96
NKG7	Natural killer cell group 7 sequence	2.91	2.60
FCRL6	Fc receptor-like 6	1.86	1.71
KLRC4-KLRK1	*KLRC4-KLRK1* readthrough	1.85	
FGFBP2	Fibroblast growth factor binding protein 2	3.24	2.40

Table 13. Differentially expressed genes related to cytotoxic pathways identified in the $CD4^+$ cell fraction. None of the detailed genes was found to be differentially expressed in the CLL vs. MBL comparison.

Gene	Description	Fold Change		
		Th1 vs. Th2 (Ono et al, 2014)	MBL vs. control	CLL vs. control
CCL5	Chemokine (C-C motif) ligand 5	23.7	1.68	
GZMH	Granzyme H	7.6	4.06	3.44
IL18RAP	Interleukin 18 receptor accessory protein	5.3	1.94	1.73
NKG7	Natural killer cell group 7 sequence	3.7	2.91	2.60
CTSW	Cathepsin W	2.9		1.51
CST7	Cystatin F (leukocystatin)	2.8	1.55	1.50
MAP3K8	Mitogen-activated protein kinase kinase kinase 8	2.6	2.12	1.84
TARP	TR gamma alternate reading frame protein	2.3	2.05	1.71
GNLY	Granulysin	2.3	2.55	1.96
ADRB2	Adrenoceptor beta 2, surface	2.1	1.75	
CFH	Complement factor H	2.1	1.74	1.66
PLEK	Pleckstrin	2.0	1.85	

Table 14. Matches with overexpressed genes in $CD4^+$ Th1 cells compared to $CD4^+$ Th2 cells of healthy subjects according to Ono et al. No matches were found for CLL vs. MBL.

Next, Ingenuity Pathway Analysis (IPA) was employed in order to obtain insight into the possible functional significance of the genetic signature of $CD4^+$ cells in MBL and early-stage CLL. All differentially expressed genes (*P*<0.05 and FC>1.5 or <-1.5) were included for IPA analysis. IPA recognized 209/247 genes for MBL vs. control, 69/90 for CLL vs. control and 36/59 for CLL vs. MBL. Since the number of included genes in the last comparison was very low for meaningful analysis, results are not shown. Relevant results concerning the other two comparisons are detailed below following the different IPA sections.

Diseases and disorders

The top five disorders summarized by IPA for both comparisons are detailed in Table 15.

Diseases and Disorders	**MBL vs. control**		**CLL vs. control**	
	***P*-value**	**Genes**	***P*-value**	**Genes**
Inflammatory response	<0.0003	77	<0.012	27
Hematological disease	<0.0004	72	–	–
Infectious diseases	<0.0002	50	–	–
Immunological disease	<0.0004	79	<0.012	23
Connective tissue disorders	<0.0004	53	<0.012	16
Inflammatory disease	–	–	<0.012	22
Organismal injury and abnormalities	–	–	<0.012	41

Table 15. Top five diseases and disorders summarized by IPA for the $CD4^+$ cell fraction of MBL or CLL vs. control comparisons.

Canonical pathways

Some canonical pathways highlighted by IPA that are relevant for the pathobiology of MBL and CLL are detailed in Table 16.

Canonical Pathway	MBL vs. control			CLL vs. control		
	***P*-value**	**Z-score**	**Genes**	***P*-value**	**Z-score**	**Genes**
Granulocyte adhesion and diapedesis	2.0×10^{-13}	–	17	1.2×10^{-3}	–	4
Agranulocyte adhesion and diapedesis	7.9×10^{-11}	–	15	1.5×10^{-3}	–	4
Communication between innate and adaptive immune cells	6.9×10^{-9}	–	10	1.9×10^{-3}	–	3
TREM1 signaling	2.4×10^{-7}	2.8	8	1.7×10^{-2}	–	2
Differential regulation of cytokine production in macrophages and T helper cells by IL17A and IL17F	1.9×10^{-5}	–	4	5.0×10^{-2}	–	1
Role of pattern recognition receptors in recognition of bacteria and viruses	2.8×10^{-5}	2.2	8	NS	–	1
Inflammasome pathway	3.6×10^{-5}	2.0	4	NS	–	1
IL8 signaling	8.2×10^{-5}	3.0	9	2.3×10^{-3}	2.0	4
Role of NFAT in regulation of the immune response	2.7×10^{-4}	1.9	8	NS	–	2
IL10 signaling	4.2×10^{-4}	–	5	NS	–	1
Toll-like receptor signaling	6.2×10^{-4}	2.0	5	–	–	–

Table 16. Relevant canonical pathways highlighted by IPA for the $CD4^+$ cell fraction of MBL or CLL vs. control comparisons. Z-score ≥2 predicts activation. NS: not significant.

Upstream regulators

Table 17 shows some of the most relevant upstream regulators detailed by IPA for MBL vs. control, obtained by integrating the gene expression results of $CD4^+$ cells. Of note, inflammatory regulators related to Th1 responses, such as *TNF, IFNG* or *IL2,* were predicted to be activated. Receptors and co-receptors involved in pathogen recognition, such as *TLR2, TLR4* or *CD14*, were also predicted as activated, as well as *CCL5*, which plays an important role in recruiting leukocytes into inflammatory sites, and *IL1A*, which initiates many immune responses that trigger inflammation. The transcription factor STAT3 and the NF-κB complex, both important triggers of survival and differentiation, were predicted to be activated as well. Interestingly, *PDCD1* (coding for PD1) and *CTLA4*, both negative regulators of T cell function, were predicted to be inhibited (Table 17).

Concerning the CLL vs. control comparison for $CD4^+$ cells, only *PDCD1* was listed as an upstream regulator with significant *P*-value (*P*= $6.6x10^{-10}$) and Z-score ≥2 or ≤-2. As well as in the previous case, *PDCD1* was predicted to be inhibited (Z-score= -2.4, 6 target genes).

Upstream regulator	Predicted state	Z-score	*P*-value	Target genes
PDGF BB	Activated	3.4	2.20×10^{-13}	13
TGM2	Activated	3.8	4.36×10^{-13}	18
TNF	Activated	3.8	2.47×10^{-10}	22
TLR2	Activated	2.6	3.64×10^{-10}	10
CCL5	Activated	2.8	3.06×10^{-9}	8
SELPLG	Activated	2.6	6.21×10^{-9}	7
CAMP	Activated	2.6	8.03×10^{-9}	8
CD14	Activated	2.4	9.04×10^{-9}	6
IL2	Activated	2.0	1.02×10^{-8}	14
TLR4	Activated	2.6	2.71×10^{-8}	9
PDCD1	Inhibited	-2.6	3.15×10^{-8}	7
IFNG	Activated	2.3	3.28×10^{-8}	16
JUN	Activated	2.4	3.90×10^{-8}	11
CTLA4	Inhibited	-2.4	8.18×10^{-8}	6
VIPAS39	Activated	2.0	9.94×10^{-8}	4
CEBPA	Activated	2.6	2.09×10^{-7}	10
STAT3	Activated	2.6	2.59×10^{-7}	11
NFkB (complex)	Activated	2.4	3.03×10^{-7}	13
IL1A	Activated	2.2	5.63×10^{-7}	8
PTGS2	Activated	2.4	5.81×10^{-7}	7

Table 17. Upstream regulators generated by IPA for $CD4^+$ cells of the MBL vs. control comparison. The 20 cases with lowest *P*-value and Z-score ≥2 or ≤-2 are shown. Last column corresponds to the target genes of a specific regulator that were found to be differentially expressed in the MBL vs. control comparison.

Gene networks

IPA also offers the possibility of visualizing gene interactions. An example of a gene network related to the inflammatory response obtained with the gene expression results of $CD4^+$ cells for the MBL vs. control comparison is represented in Figure 39.

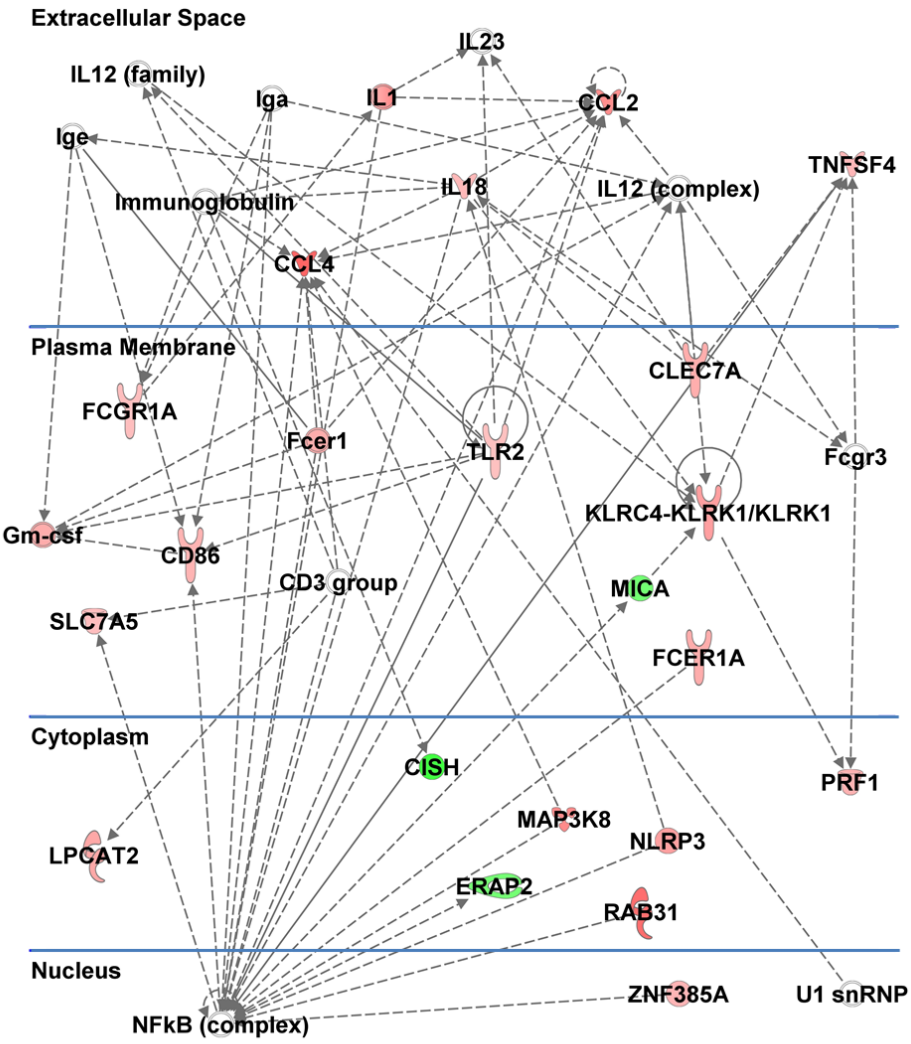

Figure 39. Inflammatory gene network related to NF-κB obtained with the results of $CD4^+$ cells. Those genes in red are overexpressed whereas those in green are underexpressed in MBL compared to controls. The intensity of the color is proportional to the expression value. The molecules coded by the different genes are indicated in their cellular or extracellular location.

Other type of networks focusing on the predicted global effect of the differentially expressed genes on specific functions can also be generated by IPA, as shown in Figure 40.

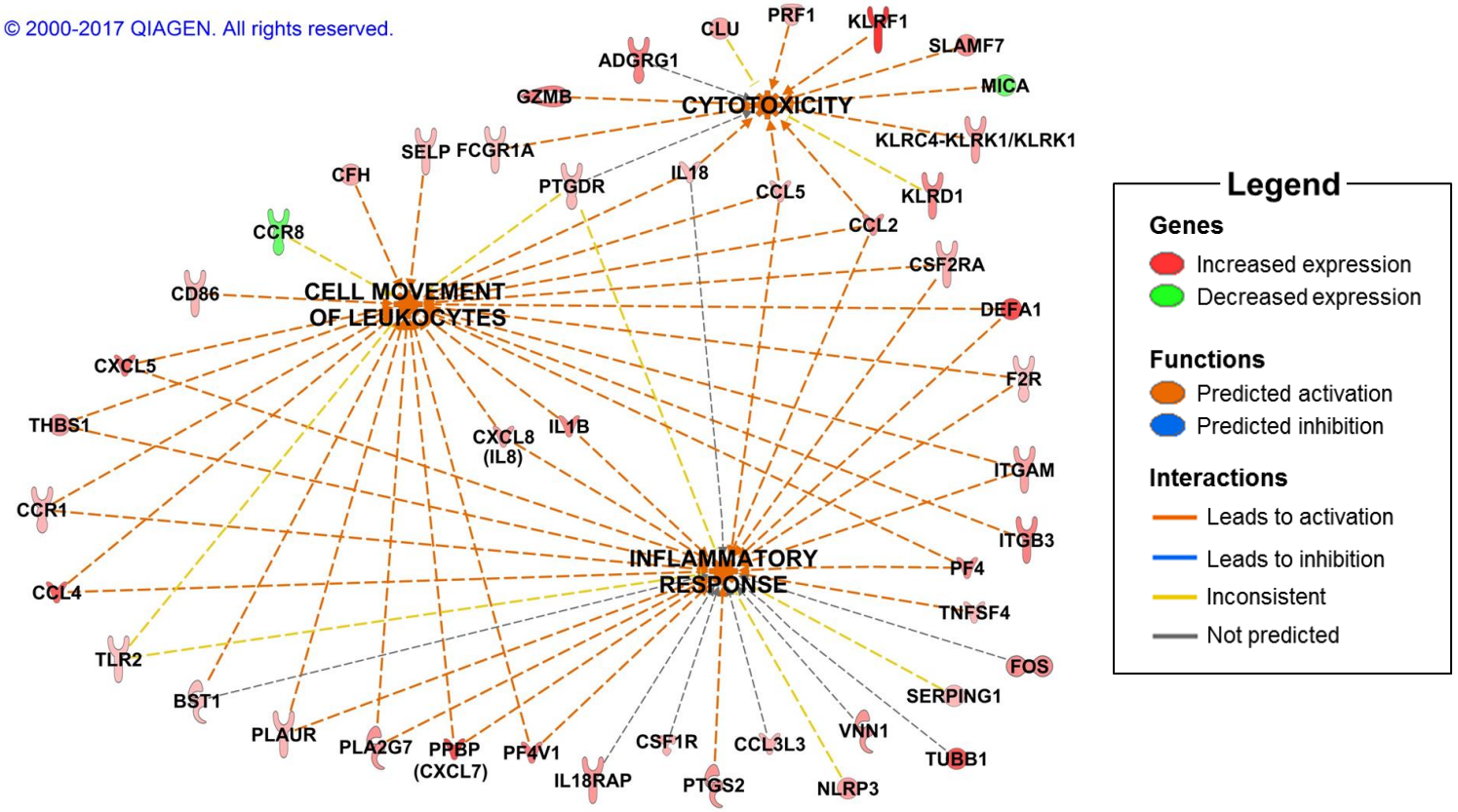

Figure 40. Combined network performed with the gene expression results of $CD4^+$ cells for the MBL vs. control comparison. The three biological functions are predicted to be activated.

With regard to $CD8^+$ cells, despite obtaining more than 125 differentially expressed genes for each comparison, neither the detailed analysis of those genes nor the functional analysis with IPA gave promising hints in the studied cohort. This could be attributed to the fact that, although the median purity for the $CD8^+$ cell fraction was 95% and despite filtering raw data to avoid possible B cell contamination as explained in section 4.2.5, many of the differentially expressed genes were found to be typically expressed by B cells. The possible reasons of these results are further discussed in section 6.1.3.

5.1.2. Gene expression validation by qPCR

Since concrete results were not obtained after performing gene expression arrays of $CD8^+$ cells, gene expression validation by qPCR was only carried out for the $CD4^+$ cell fraction.

5.1.2.1. Patient characteristics

The demographic, clinical and immunogenetic characteristics of the individuals included in the qPCR analyses to validate the gene expression arrays results of $CD4^+$ cells are detailed in Table 10 (same cohort as the one characterized by gene expression arrays) and Table 18, shown below (independent cohort).

	MBL (n=10)	**CLL (n=5)**
Age (years)	76 (51-90)	79 (60-89)
Males	8/10 (80%)	3/5 (60%)
Absolute lymphocyte count ($x10^9$/L)	3.9 (2.1-7.6)	11.2 (10.2-14)
Clonal B cells ($x10^9$/L)	1.3 (0.07-3.7)	9.6 (7.7-10.1)
Rai 0	_	5/5 (100%)
IGHV mutated	1/2 (50%)	3/3 (100%)

Table 18. Characteristics of the independent cohort analyzed by qPCR. Values are shown as median (range) or number of cases/total number of evaluated cases (percentage).

5.1.2.2. Gene expression validation results

A group of 13 genes (*CCL2, CLU, CXCL5, DEFA1, FGFBP2, GZMH, ITGAM, NEXN, NKG7, PPBP, RAB31, TUBB1* and *VSTM1*) which were differentially overexpressed (P<0.05 and FC>1.5) by gene expression arrays in $CD4^+$ cells of MBL subjects compared to healthy controls was selected for qPCR validation. With the exception of *CCL2* and *VSTM1*, all other genes (84.6%) were also significantly more expressed in MBL than in healthy individuals by qPCR.

Of the 13 selected genes, 6 (*FGFBP2, GZMH, NKG7, PPBP, RAB31* and *TUBB1*) were also differentially overexpressed by gene expression arrays in $CD4^+$ cells of CLL patients compared to healthy controls. In this case, only 2 of the 6 genes (*FGFBP2* and *NKG7,* 33.3%) were significantly more expressed in CLL than in healthy controls by qPCR.

Finally, 4 of the selected 13 genes (*CCL2, CLU, NEXN* and *VSTM1*) were found to be differentially overexpressed by gene expression arrays in $CD4^+$ cells of MBL subjects compared to CLL patients. In this case, 2/4 genes (*CLU* and *NEXN*, 50%) were also significantly more expressed in MBL than in CLL by qPCR. Besides, *PPBP* and *TUBB1* were also significantly more expressed in MBL than in CLL by this technique, although these differences were not observed by gene expression arrays.

The detailed list of validated genes and *P*-values is shown in Table 19, whereas boxplots representing qPCR expression levels for those cases with significant differential expression are depicted in Figures 41 (MBL vs. control) and 42 (CLL vs. control and MBL vs. CLL).

	P-value		
Gene	**MBL vs. control**	**CLL vs. control**	**CLL vs. MBL**
CCL2	0.301	0.481	0.095
CLU	**0.006**	0.880	**0.003**
CXCL5	**0.011**	0.081	0.181
DEFA1	**0.020**	0.075	0.302
FGFBP2	**0.010**	**0.036**	0.550
GZMH	**0.033**	0.052	0.735
ITGAM	**0.041**	0.999	0.151
NEXN	**0.012**	0.615	**0.023**
NKG7	**0.018**	**0.022**	0.903
PPBP	**0.010**	0.249	**0.024**
RAB31	**0.012**	0.737	0.190
TUBB1	**0.007**	0.383	**0.006**
VSTM1	0.258	0.875	0.339
% validated	84.6%	33.3%	50%
Total validated	65.2%		

Table 19. Validation results of the $CD4^+$ gene expression arrays by qPCR. *P*-values corresponding to the analysis of the qPCR results are detailed, and those that are significant (<0.05) are highlighted in bold. Boxes in grey correspond to those genes that were differentially overexpressed by gene expression arrays for each comparison.

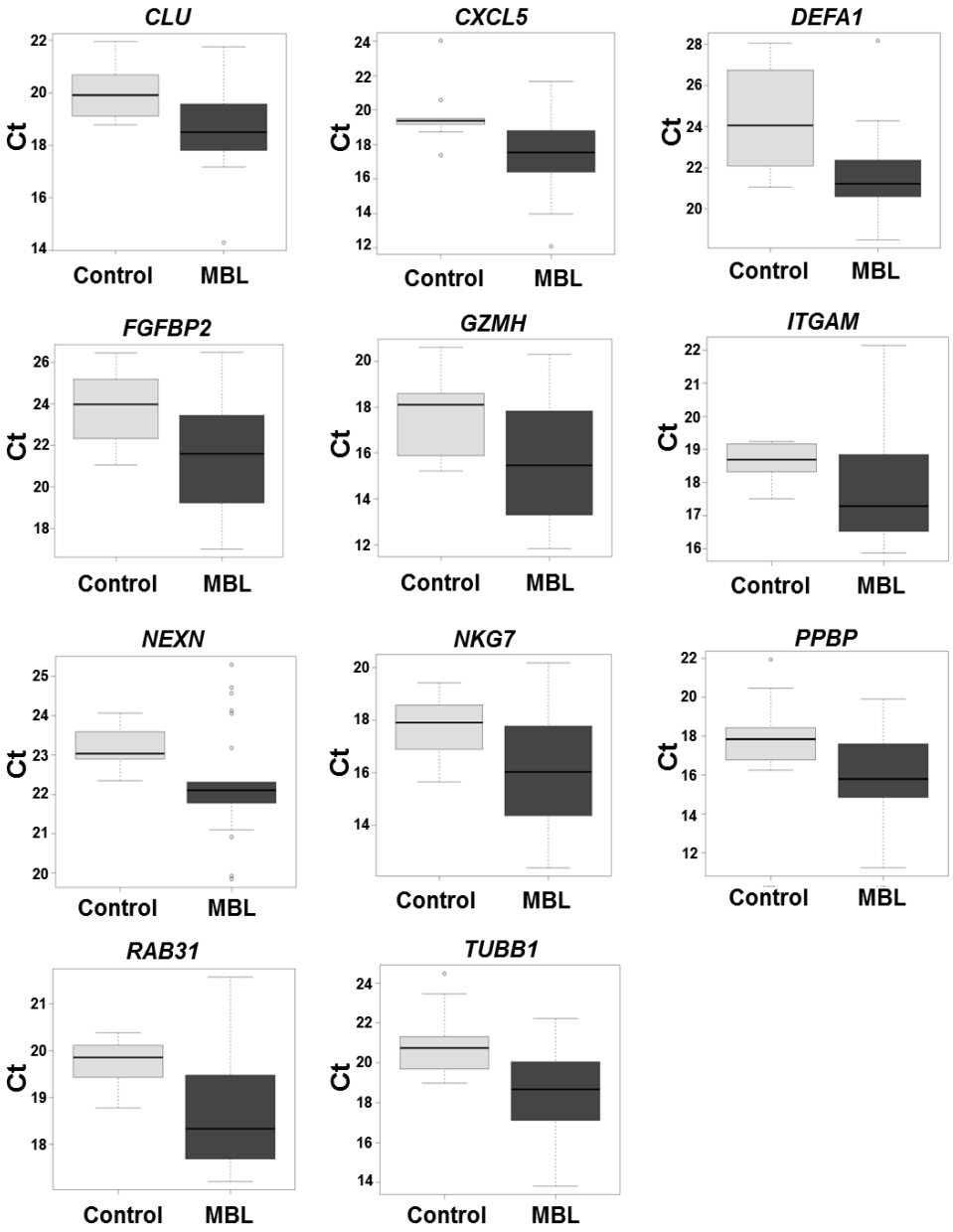

Figure 41. Boxplots corresponding to the qPCR expression levels of those genes with significant differential expression between $CD4^+$ cells of MBL and healthy controls. Ct: threshold cycle (lower Ct values indicate higher amounts of targeted nucleic acid and thus, higher gene expression).

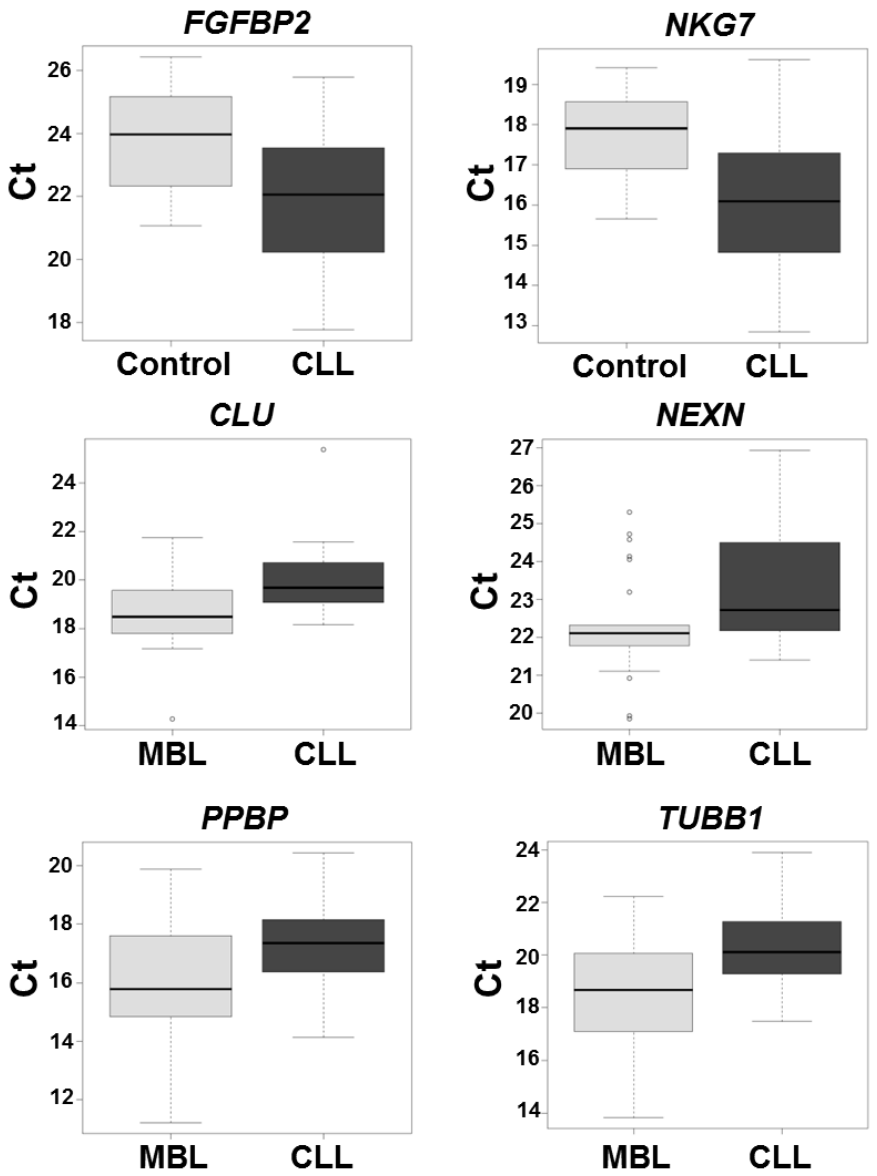

Figure 42. Boxplots corresponding to the qPCR expression levels of those genes with significant differential expression between $CD4^+$ cells of CLL and controls or MBL and CLL. Ct: threshold cycle.

5.1.3. Protein validation and further characterization by cytokine immunoassays

5.1.3.1. Patient characteristics

The demographic, clinical and immunogenetic characteristics of the subjects included in the cytokine immunoassays studies in serum are detailed in Table 20.

	Healthy controls (n=24)	MBL (n=41)	CLL (n=45)
Age* (years)	64 (43-88)	68 (45-89)	63 (37-92)
Males	10/24 (41.7%)	20/41 (48.8%)	26/45 (57.8%)
Absolute lymphocyte count ($x10^9$/L)	2.4 (1.7-3.4)***	4.3 (0.9-9.5)	14.6 (9.1-26)
Clonal B cells ($x10^9$/L)	_	2.1 (0.13-4.8)	11.8 (6.4-22.4)
Rai 0**	_	_	34/45 (75.5%)
IGHV mutated	_	9/9 (100%)	25/45 (55.5%)

Table 20. Characteristics of the cohort analyzed by cytokine immunoassays. Values are shown as median (range) or number of cases/total number of evaluated cases (percentage). *MBL was significantly (*P*<0.011) older than the CLL group. **The remaining cases were Rai I. ***Only available data from 9/24 cases.

Despite the fact that the MBL group was significantly older than the CLL group, adjusting the analysis for age was not required due to the following reasons:

- Since MBL is invariably discovered serendipitously when a complete blood count is analyzed for other reasons, MBL subjects will likely be older than CLL patients in the normal aging population.
- There is no indication in the literature suggesting a major difference in cytokine levels within individuals in the aging population as the subjects included in this analysis.

Besides, in 4 selected cases, sequential serum samples were studied after 6 (stable M-CLL), 19 (aggressive U-CLL), 24 (stable MBL) and 60 (MBL that progressed to M-CLL) months.

5.1.3.2. Comparison of cytokine levels among groups

The levels of 8/20 measured cytokines (IL5, IL8, IFNγ, TNFα, CCL3, CXCL9, CXCL10 and CXCL11) were found to be significantly different among healthy controls, MBL and CLL (Table 21). Of note, healthy subjects displayed lower median levels of these 8 cytokines than MBL and CLL. Besides, with the exception of CXCL11 that was increased in CLL compared to MBL, and CXCL9 that displayed similar levels, for the other 6 cases, median cytokine levels were higher in MBL than in CLL (Table 21, Figure 43).

Cytokine	Control	MBL	CLL	*P*-value
IL1β	0.26	0.30	0.47	0.208
IL2	1.49	1.49	1.66	0.613
IL4	0.21	0.21	0.08	0.430
IL5	1.44	1.99	1.53	**0.007**
IL6	2.04	3.60	2.86	0.071
IL8	22.05	179.49	24.83	**0.020**
IL10	0.74	0.93	0.95	0.081
IL12	1.07	1.07	1.07	0.466
IL15	3.21	2.95	2.90	0.971
IL17	2.51	2.52	2.52	0.414
IFNα	1.23	1.76	2.20	0.582
IFNγ	11.24	16.66	13.36	**0.028**
TNFα	2.25	5.00	3.30	**0.030**
GM-CSF	0.23	0.40	0.26	0.407
CCL3	22.67	46.80	39.20	**0.010**
CCL4	114.53	171.99	157.49	0.054
CCL19	109.36	131.83	108.49	0.157
CXCL9	70.98	163.00	158.80	**0.003**
CXCL10	188.10	309.17	231.02	**0.041**
CXCL11	32.37	45.60	55.69	**0.042**

Table 21. Serum cytokine levels of healthy controls, MBL subjects and CLL patients. Cytokine values (median) are reported in picograms per milliliter (pg/mL). Significant *P*-values (<0.05) are highlighted in bold.

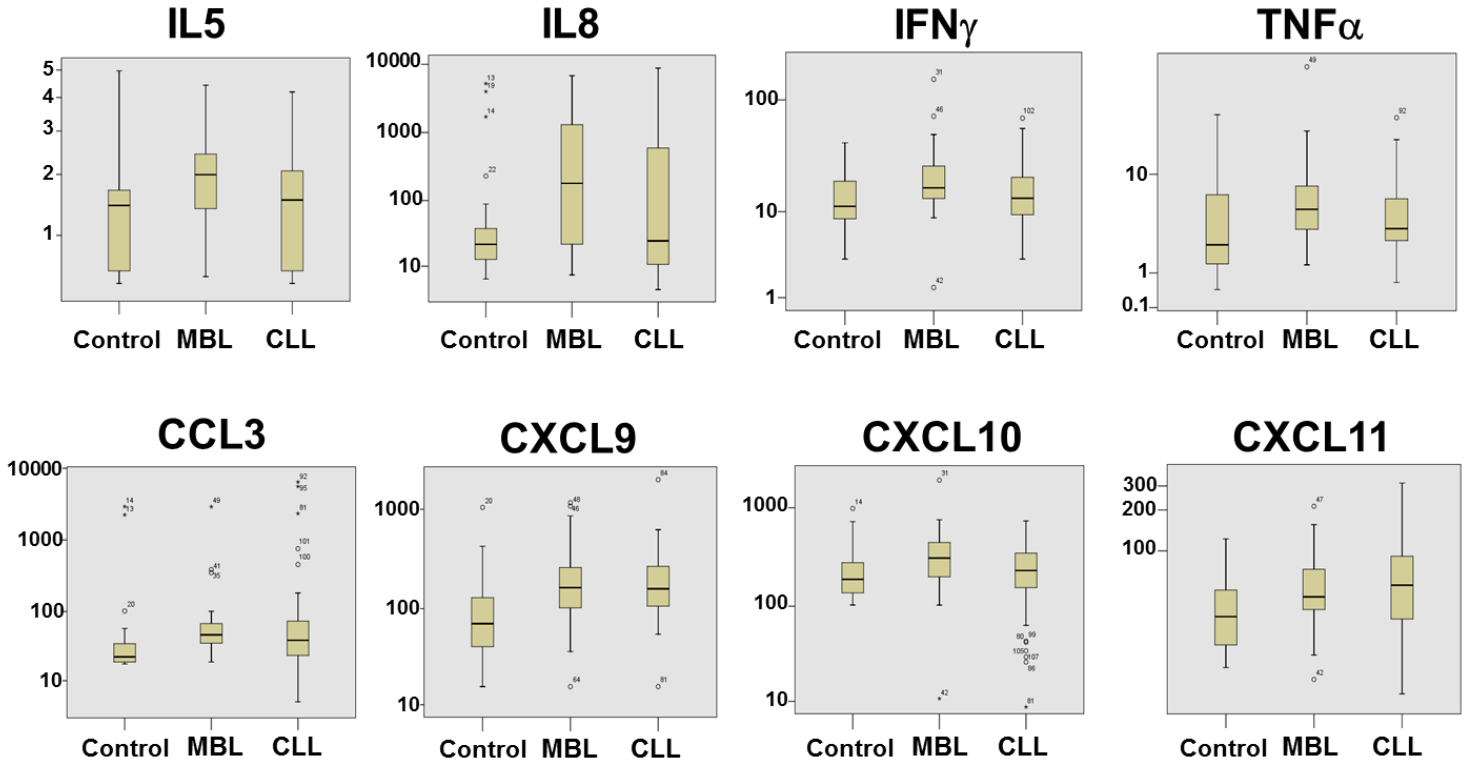

Figure 43. Boxplot representation of the cytokine levels in healthy controls, MBL and CLL. The levels of the 8 detailed cytokines showed significant (P<0.05) differences among groups. Cytokine values (pg/mL) represented in the *y*-axis were transformed into logarithmic scale.

Next, pairwise comparisons were carried out for those cytokines with significant different levels among controls, MBL and CLL. Interestingly, 7/8 cytokines were significantly increased in MBL compared to controls; however, in the CLL vs. control comparison only the levels of CXCL9 significantly differed. Concerning MBL vs. CLL, only IL5 achieved statistical significance (Table 22).

Cytokine	MBL vs. control	CLL vs. control	MBL vs. CLL
IL5	**0.005**	0.483	**0.012**
IL8	**0.009**	0.940	0.027
IFNγ	**0.013**	0.500	0.043
TNFα	**0.012**	0.131	0.122
CCL3	**0.001**	0.058	0.424
CXCL9	**0.004**	**0.001**	0.785
CXCL10	0.019	0.504	0.058
CXCL11	**0.016**	0.033	0.700

Table 22. Pairwise comparisons for those cytokines whose levels were significantly different among healthy controls, MBL and CLL. Significant *P*-values (<0.017) are highlighted in bold.

Finally, in order to address if the cytokine profile could vary depending on the mutational status of IGHV genes, CLL cases were separated into two groups (M-CLL and U-CLL) and compared with MBL. Differences in IL5, IL8 and CCL3 levels were identified (Table 23). Of note, for IL8 and CCL3, these differences were mainly due to decreased levels in M-CLL compared to MBL and U-CLL. In the case of IL5, similar levels were observed for both M-CLL and U-CLL, which were lower than those of MBL (Table 23, Figure 44).

Cytokine	MBL	M-CLL	U-CLL	*P*-value
IL5	1.99	1.53	1.49	**0.042**
IL8	179.49	15.40	138.35	**0.013**
CCL3	46.80	32.14	70.26	**0.007**

Table 23. Differences in cytokine levels among MBL, M-CLL and U-CLL. Only those cytokines whose levels were significantly different (*P*<0.05) among the three groups are detailed. Cytokine values (median) are reported in pg/mL.

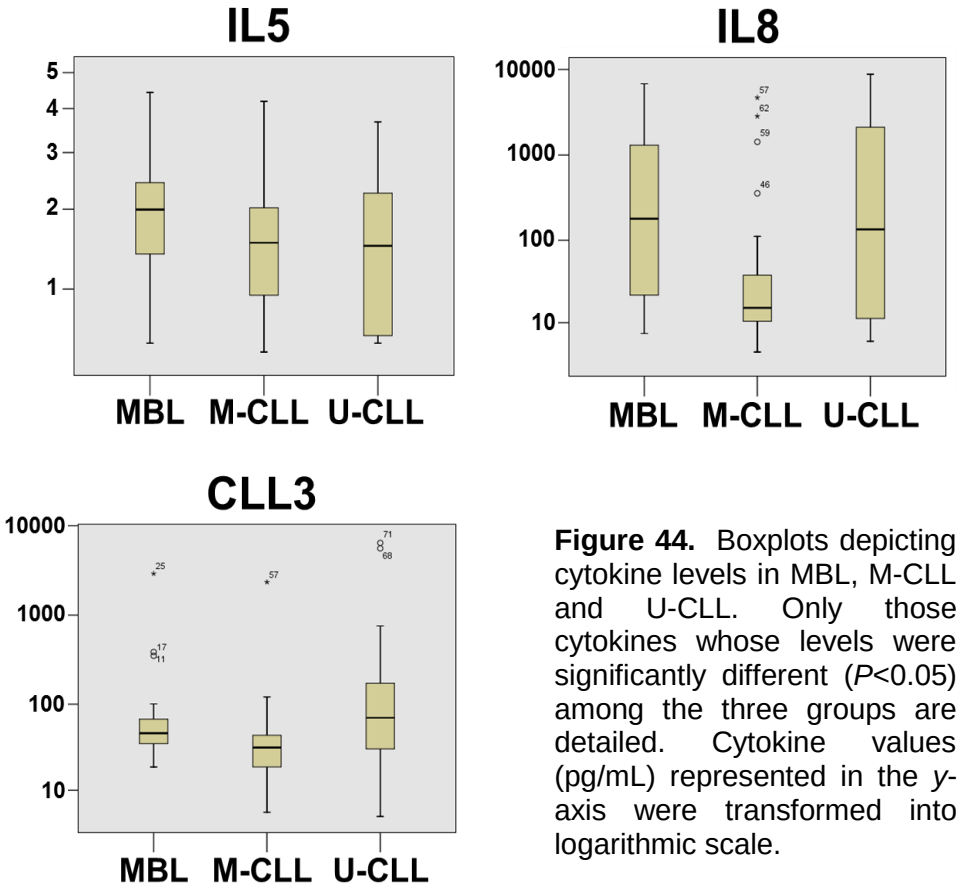

Figure 44. Boxplots depicting cytokine levels in MBL, M-CLL and U-CLL. Only those cytokines whose levels were significantly different (*P*<0.05) among the three groups are detailed. Cytokine values (pg/mL) represented in the *y*-axis were transformed into logarithmic scale.

When performing pairwise comparisons, no differences were observed between MBL and U-CLL. However, MBL displayed significantly higher levels of IL8 and CCL3 than M-CLL. Besides, CCL3 levels of U-CLL were significantly increased compared to M-CLL (Table 24, Figure 44).

Cytokine	MBL vs. M-CLL	MBL vs. U-CLL	M-CLL vs. U-CLL
IL5	0.029	0.056	0.873
IL8	**0.002**	0.730	0.098
CCL3	**0.010**	0.119	**0.011**

Table 24. Pairwise comparisons for those cytokines whose levels were significantly different among MBL, M-CLL and U-CLL. For pairwise comparison, significant *P*-values (<0.017) are highlighted in bold.

5.1.3.3. Longitudinal cytokine assessment

Results concerning sequential cytokine assessment are summarized in Table 25 and Figure 45. Fluctuations of the cytokine levels over time were identified, mainly involving IL8, CXCL9, CXCL10, CXCL11, CCL19 and IFNγ. Interestingly, whereas the aggressive U-CLL case experienced an increase of these cytokine levels between the two time-points, the stable M-CLL presented a decrease of the levels of these cytokines, with the exception of IL8. Regarding MBL studies, an increase of the levels of this subset of cytokines was identified for the stable case, whereas in the MBL subject that progressed to M-CLL this group of cytokines experienced a decrease of their levels (Table 25, Figure 45).

	Stable M-CLL		Aggressive U-CLL		Stable MBL		MBL → M-CLL	
	Time 1	Time 2 (+6 months)[1]	Time 1	Time 2 (+19 months)[2]	Time 1	Time 2 (+24 months)[3]	Time 1	Time 2 (+60 months)[4]
ALC/Rai	12.1/0	11.5/0	15.3/I	134/II	3.9/-	3.1/-	4.3/-	25.6/NA
IL1β	0.30	0.20	0.20	0.60	0.47	0.47	0.47	0.47
IL2	1.10	1.00	0.44	1.60	1.66	1.66	1.66	1.66
IL4	0.21	0.21	0.21	0.21	0.06	0.06	0.06	0.33
IL5	1.78	0.45	3.67	3.64	1.32	0.52	4.42	2.40
IL6	1.20	1.20	5.37	4.42	1.16	0.34	3.86	3.46
IL8	7.9	113.5	10.9	12.8	619.5	913.6	622.6	180.7
IL10	1.16	0.58	2.02	6.19	0.39	0.39	1.15	1.44
IL12	0.91	0.91	2.78	3.01	0.60	0.60	1.84	1.82
IL15	1.00	2.30	3.60	3.50	3.43	1.29	4.51	3.55
IL17	0.51	5.20	0.51	2.10	2.52	2.52	2.52	2.52
INFα	0.76	0.76	0.76	0.76	2.20	2.20	2.20	2.20
IFNγ	14.16	3.30	30.32	89.81	1.44	6.39	33.61	16.51
TNFα	0.70	0.25	6.00	7.60	1.34	1.34	4.34	5.43
GM-CSF	0.45	0.45	0.45	0.45	0.23	0.23	0.48	0.23
CCL3	15.6	16.7	153.6	112.3	19.3	16.4	46.8	138.6
CCL4	94.0	58.5	186.7	273.5	82.9	100.6	316.0	362.2
CCL19	143.4	3.7	689.6	1063.9	3.2	101.2	237.1	150.9
CXCL9	391.9	15.7	347.2	3594.4	36.6	111.4	321.7	123.0
CXCL10	247.0	1.7	564.3	2701.8	10.7	78.6	515.7	472.5
CXCL11	51.7	1.9	148.4	714.0	10.6	13.8	116.1	84.9

[1]Untreated after 7.5 years [2]Deceased after 9 months due to progression [3]Next 4 years without progression (still MBL) [4]Untreated

Table 25. Longitudinal cytokine studies. Cytokine values are reported in pg/mL. ALC: absolute lymphocyte count ($x10^9$ cells/L); NA: not available.

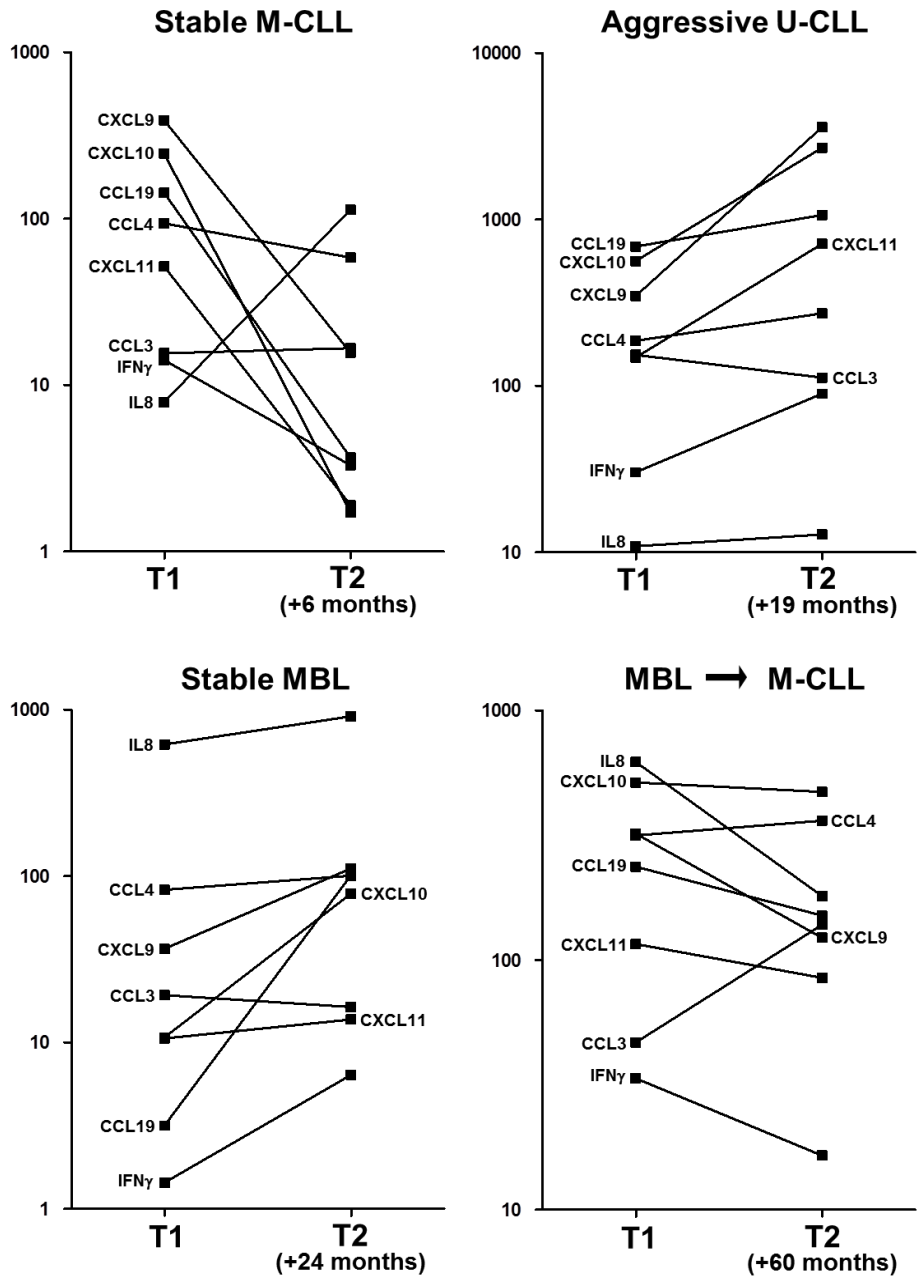

Figure 45. Longitudinal cytokine studies. Sequential serum samples were studied after 6 (stable M-CLL), 19 (aggressive U-CLL), 24 (stable MBL) and 60 (MBL that progressed to M-CLL) months. Only a subgroup of cytokines that underwent notable variations is shown. Cytokine values (pg/mL) represented in the *y*-axis are transformed into logarithmic scale. T1: first time-point; T2: second time-point.

5.2. TR repertoire analysis

5.2.1. Patient characteristics

The demographic, clinical and immunogenetic characteristics of the individuals included in the TR repertoire analysis are detailed in Table 26.

	MBL (n=16)	**CLL (n=13)**
Age (years)	76 (60-83)	73 (59-88)
Males	5/16 (31.3%)	9/13 (69.2%)
Absolute lymphocyte count ($x10^9$/L)	5 (1.2-7.2)	11.5 (8-18.5)
Clonal B cells ($x10^9$/L)	2.1 (0.2-4.3)	9 (5.8-15.9)
Rai 0	_	13/13 (100%)
IGHV mutated	13/15 (86.7%)	12/13 (92.3%)

Table 26. Characteristics of the MBL and CLL cases subjected to TR repertoire analysis. Values are shown as median (range) or number of cases/total number of evaluated cases (percentage).

Three MBL subjects were longitudinally studied. In detail, purified $CD4^+$ and $CD8^+$ T cell samples were analyzed over two sequential time points (median follow-up: 18 months) for two MBL cases and over three sequential time points for another MBL case (after 15 and 19 months).

5.2.2. Clonality analysis

A total of 2567 productive rearrangements were obtained (average: 39 rearrangements/sample, range: 21-62), corresponding to 1337 distinct clonotypes (887 from MBL, 449 from CLL and one shared by MBL/CLL). Expanded clonotypes were detected in all 65 samples analyzed. The number of expanded clonotypes/sample ranged from 1 to 13 (median: 6) whereas the cumulative frequency of all expanded clonotypes ranged from 9% to 97% (median: 64%). Further information regarding the numbers and frequencies of sequences and clonotypes for each group and cell type is shown in Table 27.

	Analyzed cases	Sequences/ sample	Expanded clonotypes	Cumulative frequency of all expanded clonotypes	
				%	*P*-value
T cell fraction	**$CD4^+$**				
MBL	16/16	41 (27-47)	6 (2-11)	41.4 (9.1-71.8)	0.033
CLL	12/13	39 (24-46)	7 (1-12)	60.0 (28.9-95.8)	
T cell fraction	**$CD8^+$**				
MBL	16/16	45 (29-62)	7 (1-10)	79.2 (49.0-91.5)	NS
CLL	13/13	43 (21-47)	6 (3-11)	79.6 (35.9-95.7)	

Table 27. Clonality analysis for the $CD4^+$ and the $CD8^+$ T cell fractions. Values are given as median (range). Data from longitudinal samples is not included. NS: not significant.

Concerning $CD4^+$ T cells, the MBL group exhibited a significantly lower cumulative frequency of all expanded clonotypes compared to CLL (median: 41.4% vs. 60%, *P*=0.033) (Table 27, Figure 46). In line with these observations, a significant positive correlation between the absolute count of clonal B cells and the cumulative frequency of all expanded $CD4^+$ T cell clonotypes was noted (*P*=0.015, ρ=0.46) (Figure 47). As for the $CD8^+$ T cell compartment, no differences in clonality between MBL and CLL or correlation with clonal B cell counts were identified. $CD8^+$ T cell samples showed a significantly higher cumulative frequency of all expanded clonotypes than $CD4^+$ T cell samples both in MBL (median: 79.2% vs. 41.4%, *P*=0.001) and CLL (median: 79.6% vs. 60%, *P*=0.010) (Table 27, Figure 46).

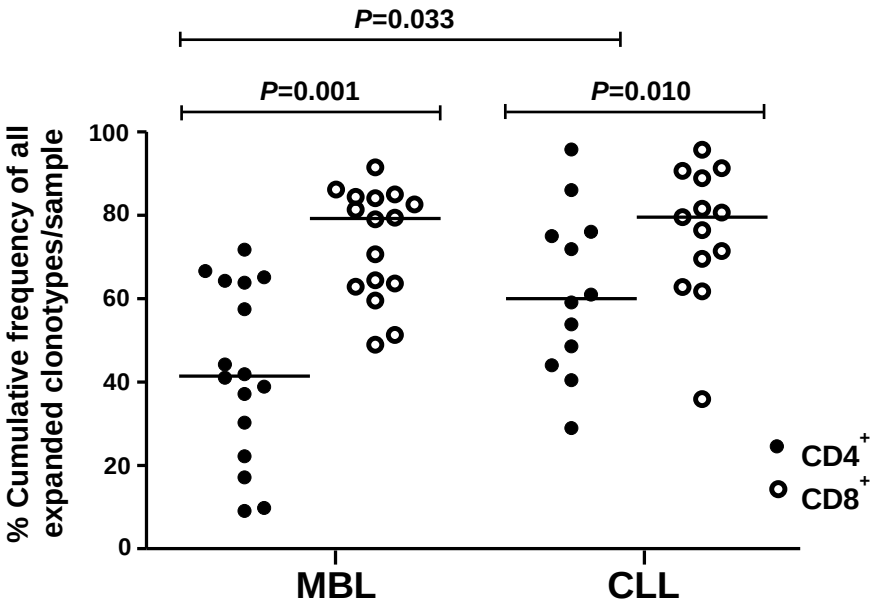

Figure 46. Percentage cumulative frequency of all expanded $CD4^+$ and $CD8^+$ T cell clonotypes in MBL subjects and CLL-A(0) patients. Horizontal lines correspond to the median value for each case.

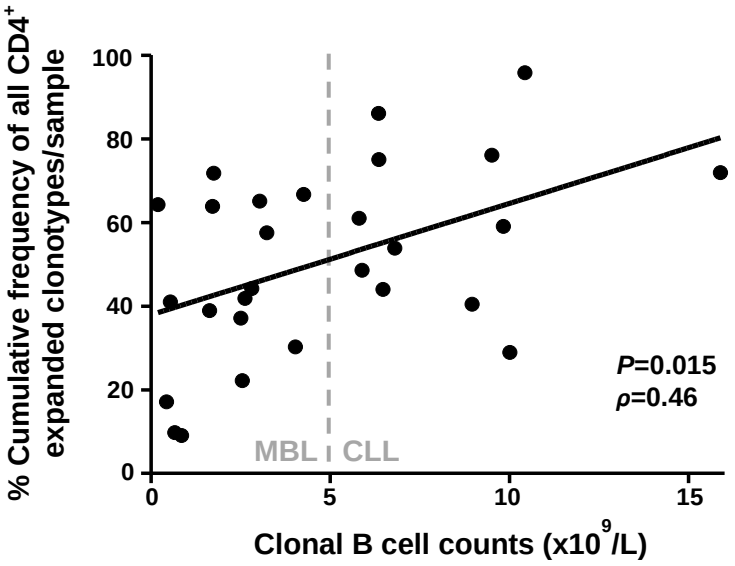

Figure 47. Correlation between the absolute count of malignant B cells and the percentage cumulative frequency of all expanded $CD4^+$ T cell clonotypes per sample. *ρ*: Spearman's rho correlation coefficient.

5.2.3. TRBV gene repertoire analysis

As for the TRBV gene repertoire of the $CD4^+$ T cell fraction, 34 functional genes were identified (Table 28). A remarkable bias in the TRBV gene usage was observed both for MBL and CLL, with only four genes (TRBV10-3, TRBV6-1, TRBV28 and TRBV19) accounting for the 40% of the entire repertoire in each group. However, when expanded clonotypes were considered, the frequencies of certain TRBV genes differed among groups (Figure 48A, Table 28). In detail, the TRBV6-2 or 6-3 gene was overrepresented in MBL compared to CLL (frequencies: 5.8% vs. 0% respectively, *P*=0.037) whereas TRBV20-1 was less frequent in

MBL than in CLL (frequencies: 2.3% vs. 10.3% respectively, *P*=0.031).

The TRBV gene repertoire of the $CD8^+$ T cell compartment was also skewed. A total of 32 functional genes were identified (Table 29). Similar to the $CD4^+$ T cell fraction, only few genes (TRBV6-5, TRBV10-3, TRBV28, TRBV6-2 or 6-3, TRBV27 and TRBV19) amounted for almost half of all clonotypic rearrangements in both MBL and CLL. When focusing on the expanded clonotypes, the frequencies of some TRBV genes were also different between the two groups. The main differences concerned higher frequencies of the TRBV10-3 and TRBV28 genes in MBL compared to CLL (frequencies: 14.3% vs. 6.6%, *P*=0.030 and 12.2% vs. 4.0%, *P*=0.024, respectively) (Figure 48B, Table 29).

Interestingly, the expanded clonotype repertoire also exhibited differences in TRBV gene usage between $CD4^+$ and $CD8^+$ T cells (Figure 48, Tables 28 and 29). In particular, the $CD4^+$ T cell fraction of MBL cases displayed higher TRBV6-1 and lower TRBV28 gene frequencies compared to the respective $CD8^+$ T cell fraction (frequency: 16.3% vs. 5.1%, *P*=0.036 and 4.7% vs. 12.2%, *P*=0.045, respectively). No significant differences were observed within the CLL group.

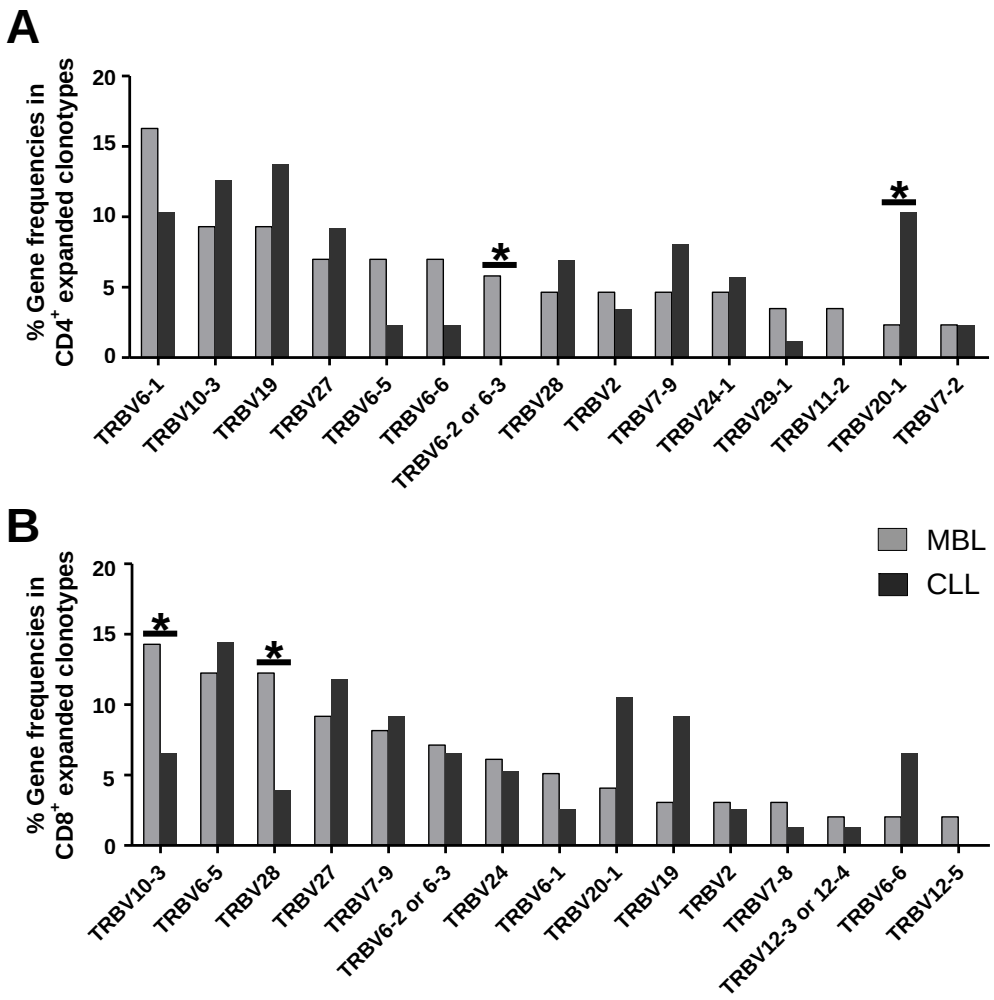

Figure 48. TRBV gene repertoire analysis of the $CD4^+$ (A) and $CD8^+$ (B) expanded T cell clonotypes in the MBL and CLL groups. The 15 most frequently detected genes within the expanded clonotypes of the MBL group are detailed in a decreasing order in the *x*-axis. Significant differences ($P<0.05$) are shown with *.

Gene	All clonotypes		Expanded clonotypes	
	MBL	CLL	MBL	CLL
TRBV10-3	13.57	10.51	9.30	12.64
TRBV6-1	11.09	12.06	16.28	10.34
TRBV28	8.37	8.95	4.65	6.90
TRBV19	7.01	10.12	9.30	13.79
TRBV6-2 or 6-3	6.11	3.11	5.81	ND
TRBV27	6.11	7.00	6.98	9.20
TRBV6-5	5.66	5.06	6.98	2.30
TRBV20-1	5.20	7.39	2.33	10.34
TRBV6-6	5.20	5.06	6.98	2.30
TRBV2	4.75	4.67	4.65	3.45
TRBV7-9	4.52	4.67	4.65	8.05
TRBV24-1	3.85	4.67	4.65	5.75
TRBV29-1	3.17	3.50	3.49	1.15
TRBV12-3 or 12-4	2.26	5.45	2.33	4.60
TRBV7-8	2.26	2.33	1.16	3.45
TRBV7-2	2.04	0.78	2.33	2.30
TRBV12-5	1.58	1.17	1.16	ND
TRBV11-2	1.36	0.78	3.49	ND
TRBV5-1	1.36	0.39	1.16	ND
TRBV14	0.90	0.39	1.16	ND
TRBV5-4	0.45	ND	ND	ND
TRBV30	0.45	0.78	ND	2.30
TRBV11-1	0.45	ND	ND	ND
TRBV15	0.45	ND	ND	ND
TRBV6-4	0.45	ND	ND	ND
TRBV13	0.23	ND	ND	ND
TRBV5-5	0.23	ND	ND	ND
TRBV6-8	0.23	ND	ND	ND
TRBV10-1	0.23	ND	1.16	ND
TRBV10-2	0.23	0.39	ND	ND
TRBV6-9	0.23	ND	ND	ND
TRBV18	ND	0.39	ND	1.15
TRBV7-6	ND	0.39	ND	ND

Table 28. TRBV gene repertoire of the $CD4^+$ T cell fraction when all clonotypes or only expanded clonotypes were considered. The frequencies (%) of all detected genes are shown for each case. In addition, TRBV11-3 was identified in a sequential MBL sample. ND: not detected.

Gene	All clonotypes		Expanded clonotypes	
	MBL	CLL	MBL	CLL
TRBV6-5	10.99	11.28	12.24	14.47
TRBV10-3	10.64	9.23	14.29	6.58
TRBV28	9.22	7.69	12.24	3.95
TRBV6-2 or 6-3	7.80	8.21	7.14	6.58
TRBV27	7.80	8.21	9.18	11.84
TRBV7-9	7.45	4.62	8.16	9.21
TRBV19	7.09	11.28	3.06	9.21
TRBV6-1	5.67	5.13	5.10	2.63
TRBV12-3 or 12-4	3.90	2.05	2.04	1.32
TRBV6-6	3.90	5.64	2.04	6.58
TRBV20-1	3.90	7.69	4.08	10.53
TRBV24	3.55	5.13	6.12	5.26
TRBV2	3.19	3.08	3.06	2.63
TRBV29-1	2.48	1.03	1.02	ND
TRBV7-8	2.13	1.54	3.06	1.32
TRBV14	1.42	1.03	1.02	1.32
TRBV12-5	1.42	ND	2.04	ND
TRBV5-1	1.42	ND	2.04	ND
TRBV7-2	1.42	1.03	ND	2.63
TRBV6-4	1.06	ND	ND	ND
TRBV18	0.71	ND	ND	ND
TRBV7-7	0.71	ND	ND	ND
TRBV11-2	0.71	1.54	1.02	1.32
TRBV30	0.71	0.51	1.02	ND
TRBV7-6	0.35	1.03	ND	1.32
TRBV15	0.35	1.03	ND	ND
TRBV4-3	ND	0.51	ND	ND
TRBV10-2	ND	0.51	ND	ND
TRBV11-3	ND	0.51	ND	1.32
TRBV5-4	ND	0.51	ND	ND

Table 29. TRBV gene repertoire of the $CD8^+$ T cell fraction when all clonotypes or only expanded clonotypes were considered. The frequencies (%) of all detected genes are shown for each case. In addition, TRBV10-1 and TRBV13 were identified in sequential MBL samples. ND: not detected.

5.2.4. Longitudinal analysis

Three MBL cases were longitudinally studied to investigate whether the small-sized MBL clones (<5000 cells/μL) would persistently affect T cell clonal dynamics.

With the exception of one case where the $CD4^+$ T cell fraction exhibited a relative stable TRBV gene repertoire, fluctuation in TRBV frequencies were detected over time in both $CD4^+$ and $CD8^+$ compartments, suggesting clonal drift (Figure 49). Regarding clonotype distribution, except for one case, a pronounced clonal drift was observed (Figure 50). Interestingly, in case #1, a highly expanded immunodominant clonotype of the $CD8^+$ T cell fraction persisted after 32 months (detected frequency at the first and second time point: 85% and 66% respectively). In case #11, the immunodominant clonotype of the $CD4^+$ T cell compartment at the first time point was detected in the other two sequential samples as well, being also the predominant clonotype after 19 months. In the remaining cases, the immunodominant clonotype at each time point differed, although at least one clonotype persisted over time (Figure 50).

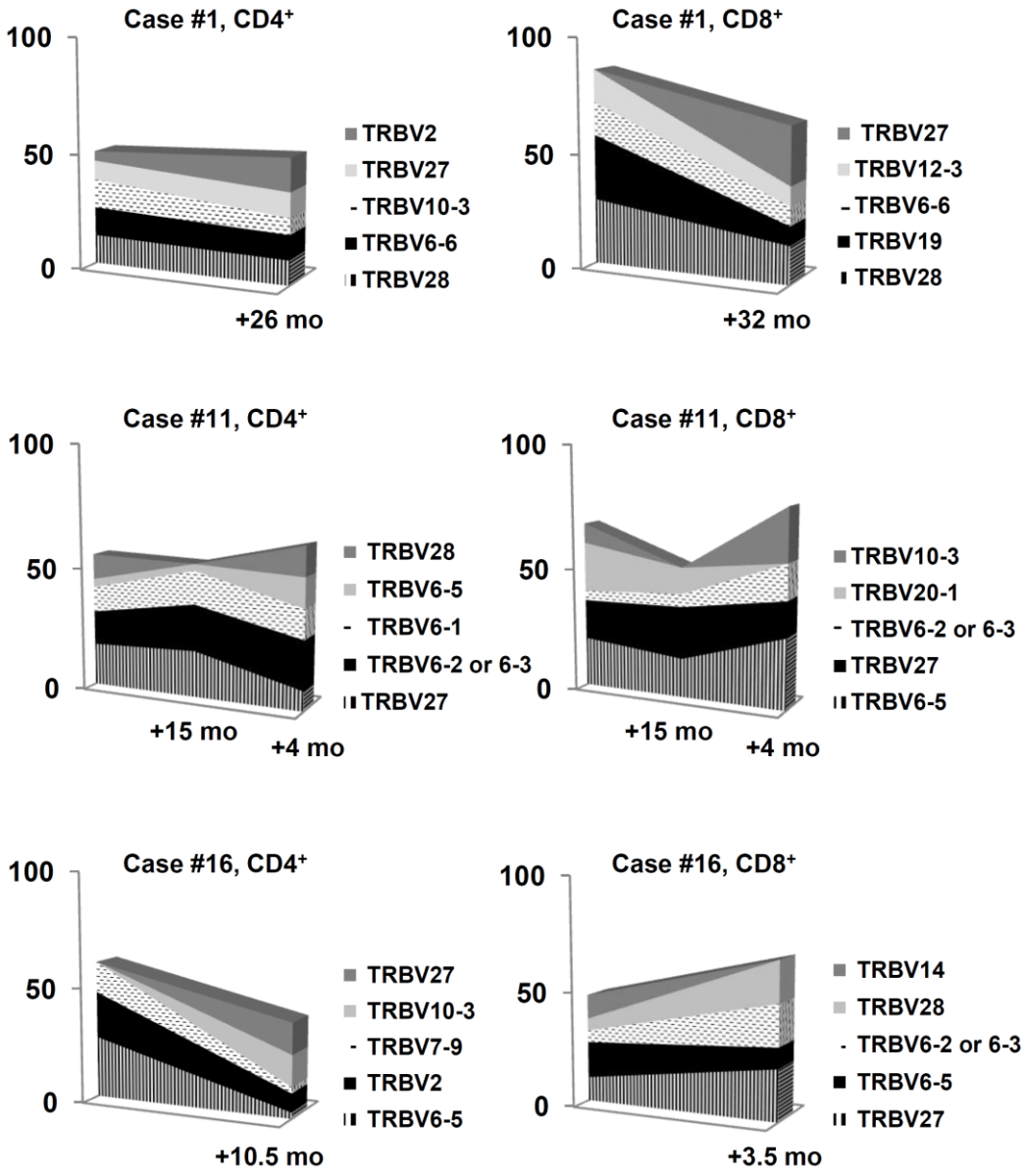

Figure 49. TRBV gene repertoire dynamics over time in three selected MBL cases for the $CD4^+$ and $CD8^+$ T cell fractions. Only the five most frequent genes are represented. Sequential time points are indicated in the *x*-axis whereas the frequency (%) of each gene is shown in the *y*-axis.

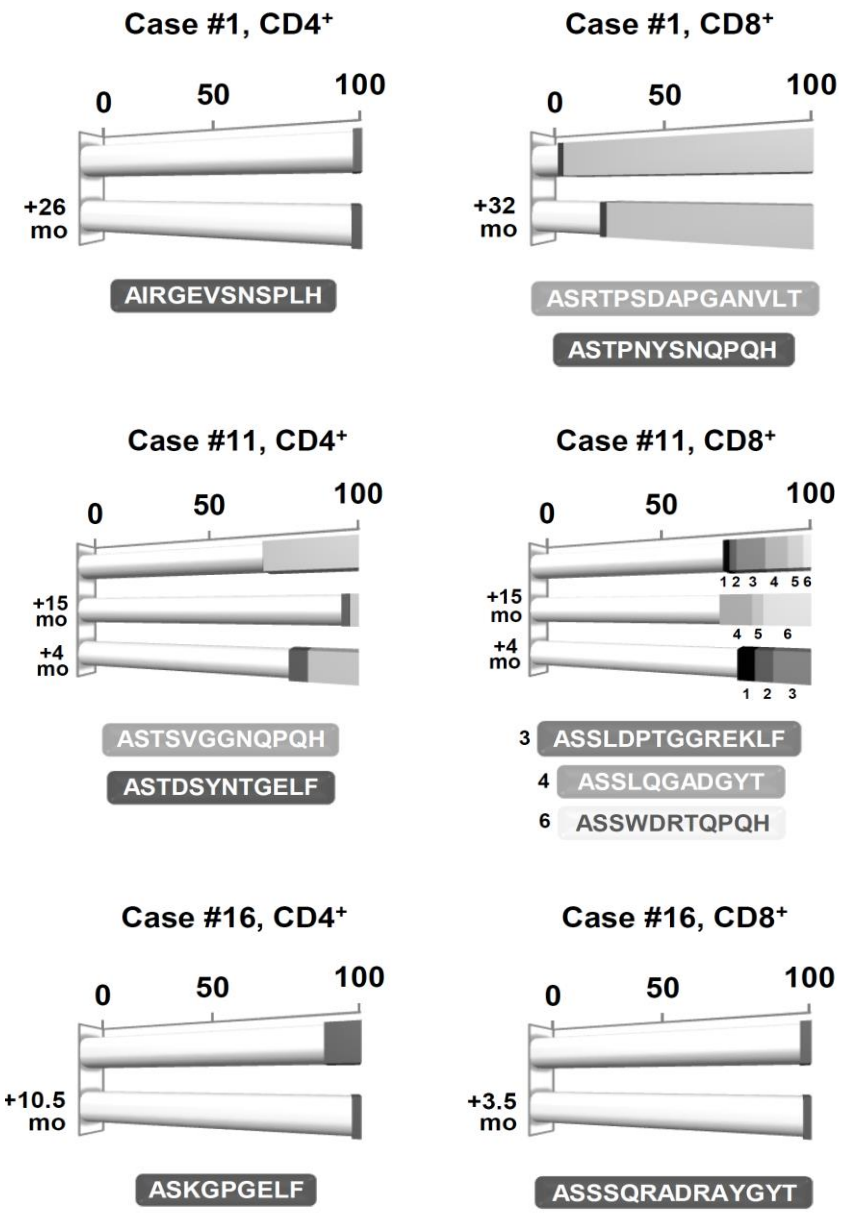

Figure 50. Clonal fluctuations over time in three selected MBL cases for the $CD4^+$ and $CD8^+$ T cell fractions. Each horizontal bar illustrates a different time point. White cylindrical parts of the bars account for the different clonotypes among the distinct time points whereas darker cubic parts represent persistent clonotypes. The frequency (%) of each clonotype is shown along the *x*-axis. Clonotypes shared by different time points, as well as their CDR3 amino acid sequence, are depicted in the same color.

5.2.5. Clonotype comparison among cases

All the obtained TRB CDR3 amino acid sequences were next compared across all the MBL and CLL subjects included in the study. Notably, two identical TRB CDR3 used by pairs of cases and another one shared by three MBL individuals were found (Table 30). In all but one case, shared CDR3 were also accompanied by identical TRBV genes, thereby consisting of “public” (“stereotyped”) clonotypes. Of note, the nucleotide sequences coding for two of the three shared CDR3 amino acid sequences were different, excluding the possibility of cross-contamination and highlighting the role of antigenic selection at the CDR3 amino acid level. Restricted HLA usage was confirmed between all cases sharing the same CDR3 regions (Table 30).

Case	Diagnosis	TRBV-D-J gene rearrangement			Shared CDR3 amino acid sequence	HLA haplotype			
		TRBV	TRBD	TRBJ		HLA-A	HLA-B	HLA-C	HLA-DRB1
#6	MBL	TRBV6-5	TRBD1	TRBJ1-5	ASSHGGSNQPQH	**02**, 11	**18**, 44	12, 16	11:01, 11:04
#16	MBL					**02**, 30	07, **18**	05, 07	03:01, 07:01
#12	MBL	TRBV6-5	TRBD1	TRBJ2-2	ASSYSGTGNTGELF	03, 11	35, **44**	12, 16	04:08, **07:01**
#22	CLL					24, 33	14, **44**	08, 08	01:02, **07:01**
#1	MBL	TRBV27	TRBD2			**02**, **30**	**18**, 49	**05**, 06	**03:01**, 11:04
#7	MBL	TRBV5-1	TRBD2	TRBJ1-5	ASSLEGDQPQH	24, 25	**07**, **07**	**07**, **07**	01:01, 01:01
#16	MBL	TRBV5-1	TRBD1			**02**, **30**	**07**, **18**	**05**, **07**	**03:01**, 07:01

Table 30. Groups of cases that showed common CDR3 amino acid sequences. Three distinct clonotypes (identical TRBV-TRBJ genes and CDR3 amino acid sequence) were shared between cases #6 and #16, #12 and #22 and #7 and #16. Case #1 displayed the same CDR3 but different TRBV gene than cases #7 and #16. A remarkable HLA restriction was observed between cases harboring identical CDR3; similarities are highlighted in bold.

5.2.6. Clonotype comparison with public databases

Finally, in order to obtain insight into the nature of the selecting antigens, all the clonotypes from the MBL and CLL cases (n=1337) were cross-compared with a panel of 5264 unique and productive TRBV-TRBD-TRBJ rearrangements from several entities obtained from the IMGT/LIGM-DB sequence database or from the database at CERTH supervised by Dr. Stamatopoulos. Nine hits sharing 100% CDR3 amino acid sequence identity were identified (Table 31): (i) a match between one MBL case and a CLL patient belonging to subset #2, both cases displaying the same TRBV gene, thus carrying a shared clonotype; (ii) a CLL case matched with an Epstein-Barr virus-specific T cell clone; (iii) a MBL case matched with a hepatitis C virus specific T cell clone; (iv) the aforementioned MBL case also showed a match with a T cell large granular lymphocyte leukemia patient, although with distinct TRBV genes. The remaining hits, mainly related to immunological disorders, as well as detailed information about CDR3 sequences and TRBV-TRBD-TRBJ genes, are summarized in Table 31.

When a less restrictive threshold (>85%) for CDR3 amino acid sequence identity was applied, additional matches were identified, including hits with CDR3 sequences of specific T cell clones in the context of human herpesvirus 2 infection, multiple sclerosis and cervical intraepithelial neoplasia (Table 31).

Match	ID	Entity/Condition	TRBV-D-J gene rearrangement			CDR3 amino acid sequence	CDR3 identity
			TRBV	TRBD	TRBJ		
1	P11840 #6	CLL patient subset #2 MBL	TRBV12-3	TRBD2 ND	TRBJ1-5	ASSPNYSNQPQH	100%
2	AM041151 #27	Epstein-Barr virus CLL	TRBV10-3	TRBD1	TRBJ1-5	AISTGDSNQPQH	100%
3	HM568209 #1	Hepatitis C virus MBL	TRBV10-3	TRBD1	TRBJ1-5	AISESTVGNQPQH	100%
4	P934 #1	T cell large granular lymphocyte leukemia MBL	TRBV19 TRBV6-6	TRBD1 TRBD2	TRBJ1-5	ASSPRGSNQPQH	100%
5	AF043185 #1	Early arthritis MBL	TRBV12-3	TRBD2	TRBJ1-5	ASTPNYSNQPQH	100%
6	S48146 #27	Immunodeficiency CLL	TRBV28	TRBD2 TRBD1	TRBJ2-7	ASSLGLHYEQY	100%
7	AY006257 #4	Organ post-transplantation MBL	TRBV5-1 TRBV27	TRBD1	TRBJ1-2	ASSLSGNYGYT	100%
8	AY006145 #27	Organ post-transplantation CLL	TRBV28	TRBD2	TRBJ2-2	ASSLTSAAGELF	100%
9	AM041177 #27	Structural limits CLL	TRBV10-3	TRBD1	TRBJ1-5	AISTGDSNQPQH	100%
10	EF592018 #4	Human herpesvirus 2 MBL	TRBV27 TRBV7-9	TRBD1	TRBJ1-4	ASRPQGANEKLF ASRPQGPNEKLF	91.6%
11	AJ405752 #27	Multiple sclerosis CLL	TRBV27 TRBV6-2 or 6-3	TRBD2 TRBD1	TRBJ1-5	ASSYEGSAQPQH ASSYEGSNQPQH	91.6%
12	CR1 #27	Cervical intraepithelial neoplasia CLL	TRBV10-3	TRBD1	TRBJ1-5	AISTGDVNQPQH AISTGDSNQPQH	91.6%

Table 31. Matches of the identified CDR3 regions with other entities. For each case, the first row represents the match identified in the IMGT/LIGM-DB or in Dr. Stamatopoulos group database and the second row corresponds to the MBL or CLL case from the studied cohort. ND: not detected.

6. DISCUSSION

6.1. Gene expression and cytokine studies

6.1.1. $CD4^+$ cells of MBL show increased inflammatory and cytotoxic pathways that are higher than in early-stage CLL

Several studies concerning $CD4^+$ T cells in CLL reported an altered function of this cell fraction, including dysfunctional gene expression patterns leading to decreased Th1 differentiation (Görgün et al, 2005), increased levels of $CD4^+$ $PD1^+$ T cells (Riches et al, 2013; Rusak et al, 2015) and impaired actin polymerization resulting in defective immunological synapses (Ramsay et al, 2008). In addition, in contrast to many malignancies where the numbers of circulating T cells are decreased, CLL shows an increased absolute number of T cells (te Raa et al, 2012; Palma et al, 2017). Unexpectedly, $CD4^+$ cells of the MBL and early-CLL subjects included in the present study did not display a similar dysfunctionality as previously described. In this sense, activation of cytotoxic and inflammatory pathways was identified in $CD4^+$ cells of the studied benign cases. Notably, prior investigations in MBL reported that T cell subset alterations are clone-size dependent and not present in MBL (te Raa et al, 2012). Rissiek et al detected an increasingly suppressive regulatory function already initiated at the MBL stage, although the effector function was impaired only after transition to CLL (Rissiek et al, 2014). In the aforementioned study, T cells from CLL patients poorly upregulated cytoplasmic IFNγ in response to specific stimulation, displaying a lower increment than healthy controls; on the contrary, T cells from MBL showed slightly increased IFNγ compared to healthy subjects (Rissiek et al, 2014). The gene

expression analysis of $CD4^+$ cells in MBL and early-stage CLL performed herein further supports these previous investigations, including the decreased T cell function in CLL compared to MBL. However, since $CD4^+$ T cells and monocytes/macrophages were studied together and many of the upregulated genes were typically expressed by myeloid cells, the activated cytotoxic and inflammatory pathways could reflect alterations of the T cell or monocytic subsets, or probably both. In this sense, the analysis of highly purified, preferentially sorted, $CD4^+$ T cells and monocytes/macrophages would be of interest, either by performing gene expression or targeted functional studies.

Although cytotoxicity is typically observed in $CD8^+$ T cells, some $CD4^+$ T cell subsets have cytotoxic properties as well. Along these lines, $CD4^+CD28^-$ cytotoxic T cells may accumulate to up to 45% of total $CD4^+$ T cells in healthy individuals older than 65 years (Vallejo et al, 1998). These cells are a unique type of pro-inflammatory and cytotoxic T cells that play important roles in many immune-mediated and infectious diseases (Maly et al, 2015). They typically express granzyme B, perforin, as well as NK and Toll-like receptors such as those coded by *TLR2* and *KLRC4-KLRK1*, which were upregulated in the $CD4^+$ cells of the present study. However, CD28 did not show significant differences in gene expression. Interestingly, previous investigations in CLL reported that up to 50% of blood $CD4^+$ T cells had a cytotoxicity-related $CD28^-$ phenotype with high content of both granzyme B and perforin, which were able to kill autologous CLL cells via a perforin-mediated mechanism (Porakishvili et al, 2001; Porakishvili et al, 2004). In this sense, the increased cytotoxic pattern in $CD4^+$ cells of the MBL and CLL cases studied here could reflect an antitumor function. On the other hand, this $CD4^+$ cytotoxic population was also associated with human cytomegalovirus

infection in CLL (Walton et al, 2010). Related to this, increased viral infections were reported both in MBL and CLL (Fazi et al, 2010; Casabonne et al, 2012; Criado et al, 2017).

Another possibility could be that the increased cytotoxic function was due to $CD4^+$ macrophages. Remarkably, prior investigations described cytotoxic macrophages with tumoricidal activity whose function could be enhanced by lymphocyte mediators (Cameron et al, 1979). Interestingly, mice injected with a myeloma cell line were able to develop tumoricidal macrophages that expressed granzyme B in addition to other cytotoxic molecules (Haabeth et al, 2011). Lysis of leukemic cells by monocyte-derived macrophages activated by IFNγ was also previously reported (Kakita et al, 1989). In the context of CLL, antitumor effects of macrophages were described in a mouse model of the disease (Wu et al, 2009) and activated macrophages were able to kill CLL cells in the presence of rituximab (Lefebvre et al, 2006). Considering all these findings, another possibility would be that the cytotoxic pattern observed in the $CD4^+$ cells analyzed in the present study was due to monocytic/macrophagic involvement, with potential antitumor effects.

On the other hand, many of the $CD4^+$ upregulated genes involved inflammatory pathways, mainly in MBL. Among them, key mediators of inflammation such as *IL1B*, *CXCL8* or *IL18* were detected. In addition, most of the pathways, functions and upstream regulators identified by IPA were associated with inflammatory disorders. Concerning these results, it was previously demonstrated that the immune system has a dual role in tumors, both promoting and suppressing cancer (Shalapour et al, 2015). Thus, although chronic inflammation and pro-inflammatory cytokines such as IL1 are considered to be tumor promoting, B cell lymphoma models in mice

reported successful immunosurveillance related to elevated levels of both pro-inflammatory and Th1-associated cytokines (Haabeth et al, 2011). Specifically, the aforementioned study by Haabeth et al reported that cancer eradication was achieved by collaboration between tumor-specific Th1 cells and antigen-presenting macrophages, concluding that inflammation, when triggered by tumor-specific Th1 cells, may prevent rather than promote cancer. On the contrary, chronic inflammation was related to progression and advanced disease in many types of tumors, including CLL, in which drugs such as ibrutinib can arrest the inflammatory drive that supports malignant B cells (Bachireddy et al, 2016). However, the benign and non-progressive clinical courses that generally characterize MBL and early-stage CLL suggest that the increased inflammatory response identified herein may reflect an effective immune surveillance rather than a supporting inflammatory TME. This lack of immune dysfunction, together with the reduced BCR signaling and proliferative capacity of the small B cell clones that are found in MBL, may account for the low progression rates to CLL.

Finally, it is worth considering the hypothesis that, since many common cancers develop after years of chronic inflammation due to persistent infections by microorganisms (Moss et al, 2005; Afrasiabi et al, 2015), the high inflammation detected in the present study could have started before MBL, facilitating the emergence of the neoplastic B cell clone. Functional validation is currently underway in order to confirm all the hypotheses previously discussed.

6.1.2. $CD4^+$ cells of both MBL and early-stage CLL exhibit overexpression of Th1-related genes

Several genes previously associated with a Th1 gene signature (Ono et al, 2014) were upregulated in $CD4^+$ cells of both MBL and CLL compared to healthy subjects. Thus, 11 and 9 gene matches were identified when comparing our results with the Th1 specific gene signature reported by Ono et al for MBL or CLL vs. control, respectively (Table 14). However, the Th2 specific gene signature described by Ono et al could not be confirmed in our study. Besides, IPA analysis revealed inflammatory regulators related to Th1 responses, such as TNFα, IFNγ or IL2. All these results suggest that $CD4^+$ cells of the MBL and early-stage CLL subjects analyzed here were Th1 deviated. This is in contrast with previous data showing dysfunctional gene expression patterns in $CD4^+$ T cells leading to a decreased Th1 differentiation (Görgün et al, 2005) as well as increased levels of IL10 in CLL (Fayad et al, 2001), which inhibits Th1 type cytokines such as IFNγ or IL2 (Fiorentino et al, 1989). Nonetheless, other investigations aimed to characterize the progressive alterations in the $CD4^+$ T cell subsets with the expansion of CLL clones revealed a Th1 dominance that is shifted towards Th2 during disease progression (Podhorecka et al, 2002). This data agrees with the results of the present study, where very early stages of the disease, including MBL, show an expression profile suggesting a Th1 polarization. As previously mentioned, Th1 cells were described to interact with macrophages to induce a specific antitumor response (Haabeth et al, 2011). Along these lines, Th1 infiltration was associated with a good prognosis in many types of cancers (Fridman et al, 2012). All these findings suggest that the Th1 and macrophagic features observed in the $CD4^+$ gene expression studies, together with the low proliferative activity of the

small B cell clones, could be related to the benign clinical courses typically observed in MBL and early-stage CLL. However, functional validation is required to confirm previous results and hypotheses.

6.1.3. Gene expression profiles of $CD8^+$ cells suggest no significant involvement of this cell fraction

Unexpectedly, gene expression analysis concerning $CD8^+$ cells was not very informative as to implicated genes and/or pathways. Thus, despite the fact that purity of the $CD8^+$ cell fraction was high and filters were applied in order to solve the effect of possible B cell contamination, many of the differentially expressed genes were found to be typically expressed by B cells. This was not observed for the $CD4^+$ cell fraction and could be likely due to the different method of purification, since $CD4^+$ cells were selected from previously CD19-depleted PBMC, whereas $CD8^+$ cells were directly purified from total PBMC (Figure 20). However, since in the analyzed cell fractions the great majority of cells were $CD8^+$ (median purity: 95%), these results suggest that $CD8^+$ cells of both studied entities were not really altered at a gene expression level. In line with the aforementioned, previous investigations reported no significant deviation of $CD8^+$ cells in MBL (te Raa et al, 2012) and showed that the effector T cell function was impaired only after transition to CLL (Rissiek et al, 2014). Besides, a B cell count threshold of $10x10^9$ cells/L was reported to best stratify aberrant $CD8^+$ T cell numbers (te Raa et al, 2012). All these findings agree with the failure to detect relevant alterations in the $CD8^+$ cell compartment of the MBL and early-stage CLL cases with low lymphocytosis included in the present study. Nonetheless, highly purified, preferentially sorted, $CD8^+$ T cells should be studied and

functional validation should be performed to definitely draw this conclusion.

6.1.4. MBL subjects display increased levels of cytokines related to a Th1-macrophagic response

In order to validate previous gene expression results that pointed to an increased inflammatory response in MBL, the levels of 20 cytokines in serum were measured. Among the cytokines that were significantly increased in MBL compared to healthy subjects, IFNγ, IL8 and TNFα were detected. IFNγ is mainly produced by NK cells, CTL and $CD4^+$ Th1 cells and is an important activator of macrophages (Schoenborn et al, 2007). IL8 is chiefly produced by macrophages and induces chemotaxis in target cells, primarily neutrophils, in which also induces degranulation (Arango Duque et al, 2014). TNFα is mainly produced by activated macrophages as well, although other cell types such as $CD4^+$ T cells or neutrophils can also produce it. It is an important orchestrator of immune responses, inducing the recruitment of lymphocytes, neutrophils and monocytes to the inflammation sites by regulating chemokine release (Grivennikov et al, 2005; Arango Duque et al, 2014). CXCL9, CXCL10 and CXCL11, which were also higher in MBL than in healthy subjects, are produced by macrophages after stimulation with IFNγ and mediate T cell recruitment (Arango Duque et al, 2014). CCL3, whose levels were significantly higher in MBL than in controls, can be produced by a broad range of cell types, including monocytes/macrophages, T and B cells, NK cells and neutrophils. This cytokine is a very potent chemoattractant for monocytes, although it affects many cell types and induced chemotaxis was demonstrated *in vitro* for T cells, NK cells and granulocytes, among

others (Menten et al, 2002). Finally, IL5, which was also significantly increased in MBL, is mainly produced by T cells and granulocytes and leads to maintenance of survival and functions of B cells and eosinophils (Takatsu, 2011). A schematic representation of the possible interplay between all these immune components is represented in Figure 51.

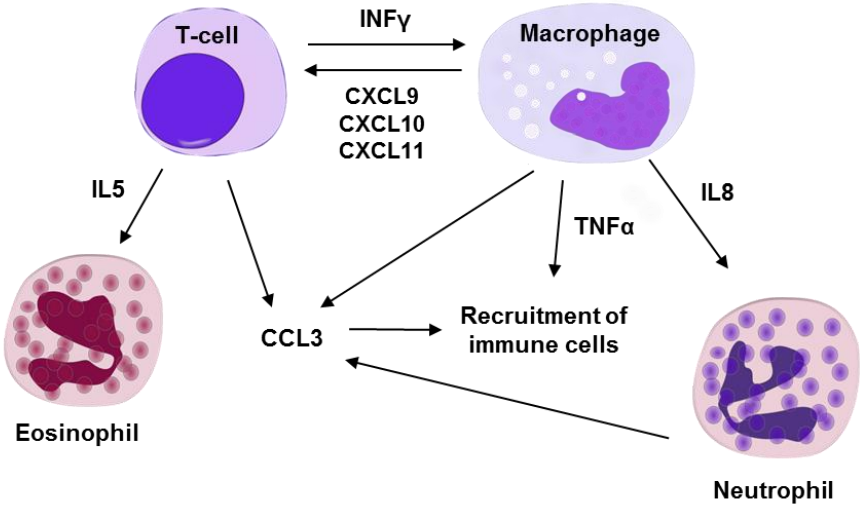

Figure 51. Schematic representation of the possible interactions mediated by the cytokines that were found to be increased in MBL.

Remarkably, most of the cytokines mentioned above are Th1 related. Th1 cells produce IFNγ which activates macrophages and evokes cell-mediated immunity and phagocyte-dependent protective responses (Romagnani, 1999). Phagocytes include several types of leukocytes, such as monocytes, macrophages, neutrophils and mastocytes. Taken together, this data suggests an increased Th1-macrophagic immunity in MBL, which could be potentially related to the asymptomatic and stable clinical course that characterizes this

entity. In this sense, several arguments support the potential antitumor effect of this characteristic TME detected in MBL. Among them, it is worth noting: (i) as previously mentioned, collaboration between tumor-specific Th1 cells and macrophages induced successful immunosurveillance in B cell lymphoma models in mice (Haabeth et al, 2011); (ii) Th1 infiltration was related to better outcome in many types of cancers (Fridman et al, 2012); (iii) T cell function in MBL was not reported to be really deviated (te Raa et al, 2012; Rissiek et al, 2014); (iv) TNFα was initially described for its ability to induce necrosis in tumors (Carswell et al, 1975) and (v) besides pathogens, phagocytic cells can also destroy dying and tumor cells, and CLL cells could be killed by activated macrophages in the presence of rituximab (Lefebvre et al, 2006). The fact that the increased cytokine levels are due to monocytes/macrophages displaying 'trained immunity' after antigen stimulation (Stevens et al, 2016), typically producing proinflammatory cytokines and showing effector functions such as phagocytosis, could also be hypothesized.

However, these interactions may not be as simple as depicted in Figure 51, and other immune cell types may also participate. Along these lines, IL5 can be produced by Th2 cells, which could also play a role in MBL. Dual functions with counteracting effects were also described for many cytokines; for instance, despite its name, TNFα (tumor necrosis factor) was associated with both cancer development and progression (Mocellin et al, 2008) and IFNγ was reported to inhibit apoptosis in CLL cells (Buschle et al, 1993). Supportive effects of T cells and macrophages in CLL were also largely described in the literature (Lad et al, 2013; McClanahan et al, 2015; Hanna et al, 2016; van Attekum et al, 2017). In this sense, although advanced CLL and MBL may have a very different

pathophysiology, supportive effects of T cells and myeloid cells may be already present in MBL to a lesser extent. Besides, Th1 cells were described to induce CLL cell proliferation and survival (Os et al, 2013) as well as CD38 expression in CLL cells (Bürgler et al, 2015); therefore, similar effects could be already present in MBL. In addition, both Treg and MDSC were detected to be increased in MBL, although to a lesser extent compared to CLL (Rissiek et al, 2014; Liu et al, 2015). Additionally, since many of the cytokines that were found to be increased in MBL subjects were previously described to play supportive roles in CLL and increased levels of some of them conferred poor prognosis and shorter times to treatment in CLL patients (Yan et al, 2011), the possibility that they also exert tumor promoting effects in MBL cannot be discarded. Nonetheless, the indolent clinical courses that characterize this entity suggest that other mechanisms may be acting to avoid tumor progression. As mentioned above, pleiotropic functions of cytokines, exerting both pro- and antitumor actions in different contexts (Rodriguez-Vita et al, 2010), may partially account for these observations. Finally, it is worth noting that a previous investigation also detected increased inflammatory cytokines in subjects with MBL (Aguilar-Santelises et al, 1992). Regarding the aforementioned study, Bürgler et al hypothesized that the inflammatory features associated with MBL could be a consequence of emerging supportive interactions between malignant and Th1 cells, which could allow progression from MBL to CLL (Bürgler et al, 2015).

Despite all these clues, whether Th1 cells display pro- or antitumor functions in MBL, as well as the exact role of monocytes/macrophages, is still unknown. Besides, serum levels can be affected by distinct circulating cell subsets in different

individuals; therefore, functional studies on specific cell subsets should be performed to address all these questions.

6.1.5. Serum cytokine levels may differ depending on the tumor burden and the IGHV mutational status

The analysis of 20 cytokines in serum revealed significant differences in the levels of IL5, IL8, IFNγ, TNFα, CCL3, CXCL9, CXCL10 and CXCL11 among healthy subjects, MBL and CLL. Remarkably, healthy individuals showed lower levels of these 8 cytokines than MBL and CLL. Pairwise comparisons revealed that with the exception of CXCL10, all the other cytokines were significantly increased in MBL compared to controls. However, for CLL vs. control comparison only the levels of CXCL9 were significantly higher in CLL.

These results validate the previous gene expression studies in $CD4^+$ cells, in which MBL cases displayed increased inflammatory pathways compared to healthy individuals, being higher than in CLL as well. Taken together, all this data alludes to an increased inflammatory response at the MBL stage that decreases with the progression to CLL.

When MBL and CLL cases were compared taking into account the IGHV mutational status, differences in the levels of IL5, IL8 and CCL3 arose. Although both M-CLL and U-CLL displayed similarly lower levels of IL5 than MBL, striking results were observed for IL8 and CCL3. Thus, despite of the fact that most of MBL cases are IGHV mutated, MBL and U-CLL displayed high levels of both cytokines, whereas M-CLL showed lower levels. Related to this, it was previously reported that U-CLL had higher cytokine levels than M-CLL. Sivina et al described that increased levels of CCL3 in CLL patients were associated with a worse outcome as well as high

CD38 expression and unmutated IGHV genes (Sivina et al, 2011). In this sense, BCR stimulation after interaction with the TME was related to CCL3 and CCL4 production by neoplastic B cells, which could be reduced by disrupting BCR signaling with ibrutinib (Burger et al, 2009; Ponader et al, 2012). Interestingly, in the present study CCL4 levels were also higher in U-CLL than in M-CLL, although differences were not significant. Thereby, BCR signaling, which is higher in U-CLL than in M-CLL, may induce the production of cytokines by the malignant clone to create a supportive TME. Concerning IL8, increased levels were associated with advanced disease and risk of death in CLL (Wierda et al, 2003). Interestingly, although IL8 is a cytokine typically produced by macrophages, CLL cells were described to express IL8 as well, likely reflecting the activation state of the neoplastic population (di Celle et al, 1994). Besides, expression of ZAP-70 in CLL cells increased BCR-induced IL8 production (Pede et al, 2013). Taken together, the abnormal production of CCL3 and IL8 by malignant B cell clones may be a pathological mechanism to promote their survival.

As previously mentioned, M-CLL typically shows a slower progression and longer survival, whereas U-CLL has a more unfavorable clinical course (Damle et al, 1999; Hamblin et al, 1999). Many data supports that these differences in clinical behavior are due to differences in BCR responsiveness to external signals. More specifically, M-CLL cells are thought to be stimulated by a restricted set of antigens, which may occur infrequently and induce anergy due to high-affinity binding. As a consequence, the M-CLL clone expands at a slower rate and remains overall stable. On the contrary, U-CLL cases harbor a polyreactive and low-affinity BCR that is able to recognize various environmental and auto-antigens in the TME. Therefore, BCR stimulation can occur more frequently in

U-CLL, leading to a more proliferative malignant clone (Stevenson et al, 2004; Burger et al, 2013). The results reported here concerning lower cytokine levels in M-CLL than in U-CLL are in line with the previous findings in the sense that cytokine production by malignant cells may depend on their capacity to interact with the TME. A schematic representation of this hypothesis is depicted in Figure 52.

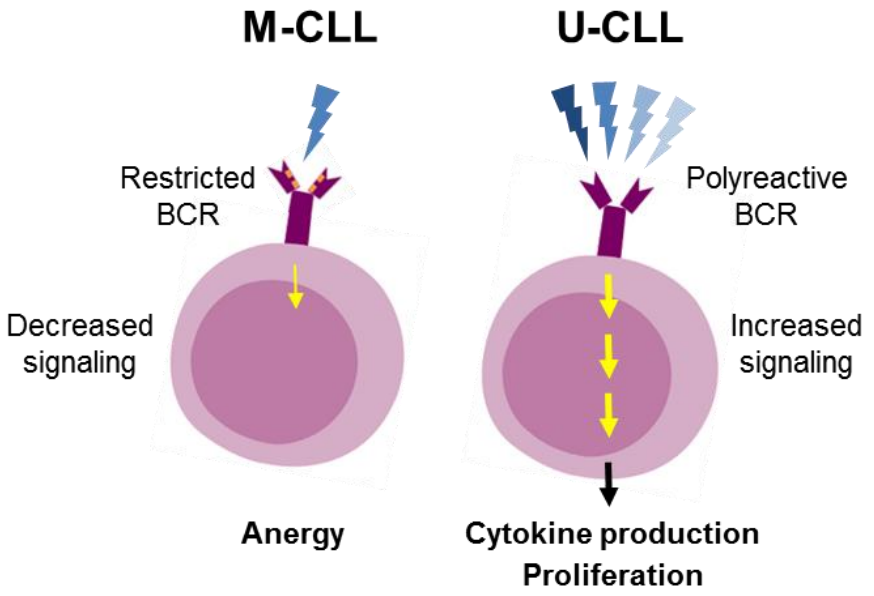

Figure 52. Different cytokine production by the malignant clone could depend on the BCR responsiveness to microenvironmental signals and therefore, on the IGHV mutational status.

Concerning MBL, increased cytokine levels were observed, even higher than in U-CLL for several of these cytokines. This could be surprising taking into account that most of MBL subjects show mutated IGHV genes as well as indolent and stable clinical courses, as opposed to U-CLL. Different arguments suggest that in MBL

cytokines may be actively produced by the TME rather than by the emerging B cell clone, among them: (i) the low tumor burden that defines MBL, (ii) previous investigations reporting that the TME in MBL is only slightly altered (te Raa et al, 2012; Rissiek et al, 2014), (iii) the fact that most MBL cases are IGHV mutated, and therefore, MBL clones may resemble anergic B cells, (iv) the gene expression results of the $CD4^+$ cell fraction of the present study, showing increased inflammatory pathways and the upregulation of genes coding for several cytokines, including IL8. Notably, IL8 was downregulated in malignant B cells of the same MBL cases analyzed by expression arrays (data not shown), although it may play a pivotal role orchestrated by the TME. In this sense, besides being overexpressed in $CD4^+$ cells of MBL and immunoassays revealed increased cytokine levels in serum, proteomic analysis in serum also revealed increased IL8 in MBL subjects (data not shown). Summarizing, the same cytokine may be produced by different cell compartments and may play different roles depending on the tumor burden (MBL vs. CLL) and the IGHV mutational status.
Taken together, two different models of progression from MBL to CLL could be hypothesized depending on the IGHV mutational status of the B cell clone. Regarding IGHV mutated clones, the TME may actively produce certain cytokines at the MBL stage, such as IL8 and IFNγ, which together with the low proliferative B cell clones may underlie the benign clinical courses that generally characterize this entity. In this sense, one could hypothesize that this increased inflammatory response could reflect an attempt of the TME to control neoplastic expansions. The progression to CLL implies suppression of the functional immunity together with a relative increase of the proliferative rate of the CLL-like MBL cells, although

the generally anergic state that usually characterizes mutated IGHV clones would produce a decrease in cytokine levels (Figure 53).

Concerning IGHV unmutated clones, information regarding MBL cases was not available. One could argue that the increased proliferation that unmutated B cell clones typically undergo may produce a rapid progression from MBL to CLL, which would account for the low frequency of unmutated MBL cases. Despite the immune suppression that occurs concurrently with clonal expansion, the IGHV unmutated clone interacts more actively with the TME in a BCR-dependent manner, which increases the levels of certain cytokines with supportive effects. Both hypothetical models are represented in Figure 53.

However, since the different cell populations may differ in composition between different subjects, also affected by concomitant diseases as well as other individual factors, cytokine levels may not depend exclusively on the IGHV status and the tumor burden. Consequently, functional studies focusing on specific cell subsets are required in order to confirm and further characterize this hypothetical model.

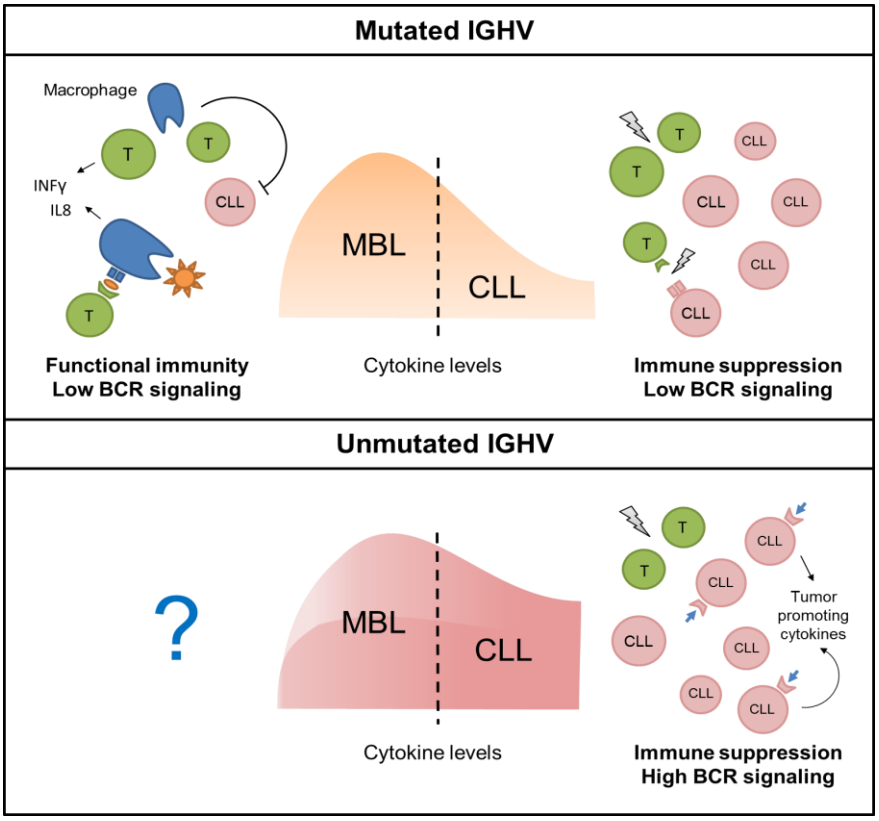

Figure 53. Hypothetical models of progression from MBL to CLL taking into account the IGHV mutational status.

6.1.6. Cytokine fluctuations over time may be associated with distinct clinical courses and outcomes

Fluctuations of the cytokine levels over time were identified for the 4 longitudinal studied cases, after 6 (stable M-CLL), 19 (aggressive U-CLL), 24 (stable MBL) and 60 (MBL that progressed to M-CLL) months. Interestingly, the stable M-CLL case showed a remarkable decrease of IFNγ, CXCL9, CXCL10, CXCL11 and CCL19 between the two analyzed time points. On the other hand, in the aggressive U-CLL case that deceased due to disease progression after 9

months of testing the second serum sample, levels of these 5 cytokines strongly increased. In line with these findings, Yan et al reported that high serum levels of a cluster of cytokines that included these 5 molecules correlated with shorter overall survival in CLL (Yan et al, 2011). As previously mentioned, IFNγ induces macrophage production of CXCL9, CXCL10 and CXCL11. CXCR3 is a common receptor for these three cytokines and was reported to be highly expressed on CLL cells. The consequences of CXCR3 stimulation in CLL cells could include regulation of tumor localization, chemotaxis or adhesion of tumor cells to protective microenvironments that ultimately induce tumor survival (Trentin et al, 1999; Jones et al, 2000). In CLL studies in mice, T cells were also reported to increase CXCR3 expression concurrently with increased levels of CXCL9 in monocytes (Hanna et al, 2016), contributing to the recruitment of supportive cells. In the aggressive U-CLL case, the concurrent remarkable increase of both absolute lymphocytosis (from 15.3 to 134 $x10^9$ cells/L) and cytokine levels suggests that malignant cells could also be the source of some of these cytokines, supporting the idea that clonal B cells may directly contribute to the increased inflammatory TME that promotes their survival. In line with this, previous studies in CLL suggested that the malignant cells were able to synthesize IFNγ, contributing to clonal accumulation (Buschle et al, 1993).

Regarding longitudinal analysis in the MBL subjects, the stable case displayed an increase of the levels of the aforementioned 5 cytokines. Thus, it could be argued that maintenance of increased cytokine levels may contribute to control malignant expansions at the MBL stage, where the BCR signaling and therefore the proliferation of the small B cell clone is still reduced. In this sense, Gustafsson et al described that monocyte-derived cells from CLL

patients that matured *in vitro* in the presence of IFNγ produced high amounts of CXCL9, CXCL10 and CXCL11 and showed an increased capacity to recruit cytotoxic lymphocytes, alluding to an antitumor response (Gustafsson et al, 2011). Concerning the MBL subject that progressed to M-CLL, a decrease of the levels of these 5 cytokines was noted, which could reflect a loss of this monocyte-derived immunity in the context of immune suppression that typically occurs concurrently with disease progression.

In summary, although sequential studies were performed in a limited number of cases, it seems that different cytokine fluctuations over time may be predictive of distinct clinical courses. It is worth noting that, as proposed by Yan et al, the pleiotropic nature of cytokines implies that cytokine clusters instead of individual cytokine levels must be considered to predict outcomes (Yan et al, 2011). However, as suggested by the present study, the possibility that similar changes in the same cluster of cytokines have distinct prognostic value depending on the stage of clonal development (MBL vs. CLL) must be considered and further evaluated in larger cohorts.

6.2. TR repertoire analysis

Molecular characterization of the TR confers a powerful tool for the detection of T cell clones potentially involved in immune surveillance and leukemogenesis. Although the T cell compartment was demonstrated to be dysfunctional in CLL (Görgün et al, 2005; Palma et al, 2017), allowing tumor expansion and disease progression, the indolent clinical courses observed in MBL and early-stage CLL may reflect a different T cell behavior, potentially underlying the molecular mechanisms of non-aggressive CLL cases. That said, MBL may represent an early-stage of CLL ontogeny, most likely concealing important clues about leukemogenesis. In the present study, the TR repertoire was analyzed, separately for both $CD4^+$ and $CD8^+$ T cell subpopulations, in the largest cohort of MBL subjects and early-stage CLL patients in order to gain insight into the role of T cells in CLL evolution.

6.2.1. The T cell repertoire of MBL and early-stage CLL is restricted, with persisting clones over time

The $CD4^+$ T cell fraction in CLL exhibited a significantly higher clonality than in MBL. In line with this, $CD4^+$ T cell clonal expansions followed the numerical increase of malignant B cells. These findings suggest that $CD4^+$ T cell repertoire restriction may be influenced by the extent of B cell clonal expansions, occurring early in clonal evolution and increasing concurrently with tumor progression, which possibly alludes to tumor-related selecting antigens. Prior investigations in CLL showed increasing frequencies of Treg cells accompanying tumor development (Dasgupta et al, 2013; Jadidi-Niaragh et al, 2013), contributing to the immunosuppressive TME that allows the neoplastic B cells to proliferate. Interestingly,

increased numbers of Treg have been detected in MBL as well, albeit to a lesser extent (D'Arena et al, 2011; Rissiek et al, 2014). However, in the present study the progressive increase of $CD4^+$ T cell clonality could not be demonstrated to be caused by Treg expansion. On the other hand, a recent investigation showed that the Bruton's tyrosine kinase inhibitor ibrutinib, a highly effective new therapy for CLL, increased T cell repertoire diversity in these patients (Yin et al, 2017), which was associated with immune reconstitution. Although cases reported in the present study are clinically very different from patients who receive ibrutinib therapy, the fact that MBL cases showed a lower restriction in their T cell repertoire compared to CLL could reflect a more effective T cell function in MBL. Concerning the $CD8^+$ T cell fraction, differences in clonality were not detected between MBL and CLL. However, this could be possibly attributed to the increased extent of clonality in the $CD8^+$ T cell compartment, which may not allow the discrimination of small variations when performing the methodology employed in this study. The $CD8^+$ T cell fraction exhibited more expanded clones compared to $CD4^+$ T cells, which was also recently confirmed in CLL patients by high-throughput immunoprofiling (Vardi et al, 2017) and consolidates subcloning as a useful tool for clonality analysis. Besides, a previous investigation that employed the same methodology as in the present study did not find T cell oligoclonality in age-matched healthy controls (median cumulative frequency of all expanded T cell clonotypes: 5%) (Vardi et al, 2016), which further supports that clonal $CD4^+$ and $CD8^+$ T cell expansions may occur in the presence of neoplastic B cells and are already detectable in MBL. Distinct TRBV gene usage between MBL and CLL groups was also identified, including significantly increased TRBV10-3 and TRBV28 frequencies in the expanded $CD8^+$ T cell repertoire of MBL

cases compared to CLL. Along this line, a previous subcloning study found that TRBV10-3 and TRBV28 genes were under-represented in CLL in comparison to healthy subjects (Vardi et al, 2016). Altogether, these findings provide evidence of progressive modifications in the architecture of the T cell compartment that may occur following clonal B cell expansions.

The persistence of certain T cell clones over time in all MBL samples longitudinally analyzed was also demonstrated. This is in agreement with previous studies in CLL (Vardi et al, 2016; Vardi et al, 2017) and extends the findings to MBL, suggesting that the selecting antigens may persist even from the early stages of CLL ontogeny. Of note, the $CD8^+$ T cell fraction of one MBL case displayed a significantly expanded immunodominant clonotype which persisted at similar frequency after 32 months. One might argue that it could correspond to a T cell expansion related to an infectious event. Indeed, infections are described to be more frequent in MBL with unclear role in its biology (Fazi et al, 2010; Casabonne et al, 2012; Criado et al, 2017). However, the significantly long follow-up time (almost three years) makes this a rather unlikely hypothesis. Potentially, it could represent a MBL-specific T cell clone restricting the expansion of the aberrant B cells.

6.2.2. Shared clonotypes and CDR3 regions between distinct cases, mostly MBL/CLL-specific

Another important finding concerned the identification of shared ("public") clonotypes between different MBL and/or CLL cases. Relevant to mention, the possibility of cross-contamination was essentially discarded due to several reasons, including subcloning experiments performed in different times or laboratories and distinct

nucleotide sequences of shared amino acid clonotypes. The possibility that unrelated individuals harbor shared clonotypes by chance is extremely low; hence, these public clonotypes may be associated with the pathophysiology of MBL/CLL, either by recognizing tumor-specific antigens or within the context of immune responses against as yet unidentified infectious agents with a potential relevance in the development of the disease (Fazi et al, 2010; Casabonne et al, 2012; Criado et al, 2017). These antigenic triggers likely occur very early in the natural history of CLL and may persist along with B cell clonal expansions, which was demonstrated herein by those shared clonotypes between MBL subjects but also between MBL and CLL cases. Of note, these public clonotypes were not detected in public databases, alluding to common antigenic stimulation that may be MBL/CLL-specific. Whether they correspond to the same antigens that are implicated in the selection of the malignant clone remains to be elucidated.

As expected, HLA restrictions were identified between cases with shared clonotypes. A bias in HLA usage, mainly concerning the development of severe disease, was reported in CLL (Shah et al, 2011; Di Bernardo et al, 2013; Gragert et al, 2014). In addition, a very recent study associated HLA specificities with prognosis in MBL (García-Álvarez et al, 2017), pointing to the existence of HLA-restricted T cell interactions involved in the control of tumor expansion. In line with this, two of the MBL cases with a shared TR CDR3 region also shared five of the eight analyzed HLA loci.

6.2.3. Matches with infectious and immune disorders

Several matches with infectious and immune disorders were also detected. Interestingly, a similar study performed with the same methodology in 58 CLL patients only found one match with a reactive $CD8^+$ Epstein-Barr virus-specific T cell clone (Vardi et al, 2016). The fact that considerably more matches were identified within the MBL and early-stage CLL cohort of the present study could reflect a potential role of infectious agents and immune alterations in the pathogenesis of the disease. Interestingly, the detection of a clonotype in a MBL case that matched a T cell clone found in an early arthritis patient and that persisted after 32 months might also be relevant considering several reports that CLL BCR IGs often exhibit rheumatoid factor reactivity (Sthoeger et al, 1989; Borche et al, 1990; Kostareli et al, 2012). All these findings suggest that chronic exposure to self or exogenous antigens could trigger immune reactions and processes leading to CLL-like clonal expansions.

6.3. Integration of the results

Taken together, increased inflammation, mostly Th1 and macrophagic related, as well as antigen restriction, were detected in MBL. These findings highlight a role of immune alterations and antigenic drivers from the very early stages of the disease. Although it was not directly assessed in this study, these antigenic elements, whose effects were reflected in the restricted TR repertoire, may potentially account for the increased inflammation observed in MBL.
In this sense, the role of antigens in the pathogenesis of MBL/CLL has been extensively explored. Prior investigations showed clear correlations between infections and MBL/CLL (Anderson et al, 2009; Fazi et al, 2010) and a key role of antigens in the development of CLL was highly supported (Fais et al, 1998; Catera et al, 2008; Chu et al, 2008; Lanemo Myhrinder et al, 2008; Agathangelidis et al, 2012; Hatzi et al, 2016). The present study corroborated these prior investigations and further characterized the antigenic drives in MBL and early-stage CLL. Related to this, several TR CDR3 matches with different antigen-related entities were identified in the MBL and early-stage CLL cohort studied here. These and/or other antigenic elements, including foreign antigens, self-antigens and/or potentially tumor-specific antigens, may cause a strong and initial inflammatory response, presumably with involvement of both adaptive (Th1) and innate immune cells (monocytes/macrophages). As observed both in the gene expression studies of $CD4^+$ cells and in the cytokine assessments in serum, this inflammatory response is decreased with progression to CLL, probably due to immune suppression that typically occurs concurrently with clonal expansions. At the MBL stage, the non-dysfunctional TME, as observed in the expression results of $CD4^+$ cells as well as in the less restricted $CD4^+$ TR

repertoire, may restrain B cell clonal expansions. Progressive T cell alterations, as those observed in $CD4^+$ T cells subjected to TR repertoire analysis, may contribute to impaired immune surveillance, allowing the progression from MBL to CLL.

Moreover, as suggested here by those T cell clonotypes shared by distinct MBL/CLL cases and not found in public databases, some of these antigenic elements may be MBL/CLL-specific. Whether these antigens are the same that interact with the neoplastic cells remains to be elucidated.

On the other hand, the proliferative capacity of the malignant clone may also be crucial for progression to CLL. In this sense, BCR responsiveness to external signals is influenced by the IGHV gene mutational status. Thereby, BCR binding in U-CLL clones can occur more frequently, leading to increased BCR signaling. As suggested in the present study, these differences in BCR signaling may also alter the inflammatory molecules of the TME.

Summarizing, a balance between external (microenvironmental) factors, with a remarkable role of antigens, and intrinsic properties of the emerging B cell clone, may be decisive from the first stages of the disease both in inducing and maintaining neoplastic expansions.

7. CONCLUSIONS

Concerning gene expression studies, the following conclusions can be drawn:

1. Gene expression arrays showed that the expression profile of $CD4^+$ cells from MBL subjects is characterized by an upregulation of genes related to the inflammatory response and cytotoxicity, many of them concerning myeloid cells. This overexpression pattern was higher in MBL than in early-stage CLL. A Th1 expression signature was also observed in $CD4^+$ cells of both entities by this technology. These results probably reflect a deviation in the architecture of $CD4^+$ mononuclear cells towards Th1 and monocytic/macrophagic cell subsets.

2. Expression results from $CD4^+$ cells were validated by qPCR in the same cohort studied by arrays and in an independent cohort. A subset of genes related to inflammation and cytotoxicity were selected for qPCR validation. A high number of genes were validated for MBL vs. control, whereas in the other two comparisons gene expression results were only partially validated. Globally, these results reaffirm that the increased inflammatory and cytotoxic pattern is higher in MBL than in early-stage CLL.

3. $CD8^+$ T cells did not show a clear alteration of their function when performing gene expression arrays as detailed in the present study and in the analyzed cohort of MBL and early-stage CLL subjects.

4. Functional validation is required to draw definitive conclusions.

Concerning cytokine immunoassays, the conclusions are:

1. The high cytokine levels detected in a larger MBL cohort confirmed that an increased inflammatory response characterizes this entity. In particular, the significantly higher levels of IL8 (typically produced by macrophages) and the Th1/macrophage related cytokines IFNγ and TNFα, together with the gene expression results of $CD4^+$ cells (including the upregulation of *CXCL8*, coding for IL8, besides other monocytic and Th1-related genes), reaffirm the increased inflammation in MBL and suggest a Th1 deviated response with involvement of monocytes/macrophages.

2. For those cytokines which showed significant differences in their serum levels among healthy controls, MBL and CLL, lower levels were mainly detected in early CLL compared to MBL, including IL8, IFNγ and TNFα. These results point to a decrease in the inflammatory response during disease evolution and confirm the decreased inflammation in early CLL compared to MBL as detected by expression studies.

3. The different cytokine profile of early M-CLL and U-CLL could reflect a different immune environment related to the progression to each condition.

4. Similar cytokine fluctuations over time are likely associated with distinct clinical courses and outcomes depending on the stage of clonal development (MBL vs. CLL), although sequential analyses must be performed in larger cohorts.

5. Since cytokine levels can be affected by different circulating cell subsets with variable distribution among individuals, functional validations focusing on specific cell populations are required to definitively confirm the previous results.

Concerning TR subcloning analysis, the conclusions are:

1. $CD4^+$ and $CD8^+$ T cells of MBL and early-stage CLL harbor oligoclonal TR repertoires, with persisting T cell clones over time and increasing clonality within the $CD4^+$ T cell subpopulation concurrently with the expansion of neoplastic B cells.

2. The identification of the same clonotypes in different MBL/early-stage CLL cases points to selection of T cell clones by common antigenic elements, very early in the clonal evolution process leading to CLL. The fact that these shared T cell clonotypes were not found in public databases alludes to common antigenic stimulation that may be potentially considered MBL/CLL-specific.

8. REFERENCES

Afrasiabi K, Zhou YH, Fleischman A. Chronic inflammation: is it the driver or is it paving the road for malignant transformation? Genes Cancer. 2015 May;6(5-6):214-9.

Agathangelidis A, Darzentas N, Hadzidimitriou A, Brochet X, Murray F, Yan XJ, Davis Z, van Gastel-Mol EJ, Tresoldi C, Chu CC, Cahill N, Giudicelli V, Tichy B, Pedersen LB, Foroni L, Bonello L, Janus A, Smedby K, Anagnostopoulos A, Merle-Beral H, Laoutaris N, Juliusson G, di Celle PF, Pospisilova S, Jurlander J, Geisler C, Tsaftaris A, Lefranc MP, Langerak AW, Oscier DG, Chiorazzi N, Belessi C, Davi F, Rosenquist R, Ghia P, Stamatopoulos K. Stereotyped B cell receptors in one-third of chronic lymphocytic leukemia: a molecular classification with implications for targeted therapies. Blood. 2012 May 10;119(19):4467-75.

Agathangelidis A, Vardi A, Baliakas P, Stamatopoulos K. Stereotyped B cell receptors in chronic lymphocytic leukemia. Leuk Lymphoma. 2014 Oct;55(10):2252-61.

Aguilar-Santelises M, Loftenius A, Ljungh C, Svenson SB, Andersson B, Mellstedt H, Jondal M. Serum levels of helper factors (IL-1 alpha, IL-1 beta and IL-6), T-cell products (sCD4 and sCD8), sIL-2R and beta 2-microglobulin in patients with B-CLL and benign B lymphocytosis. Leuk Res. 1992 Jun-Jul;16(6-7):607-13.

Anderson LA, Landgren O, Engels EA. Common community acquired infections and subsequent risk of chronic lymphocytic leukaemia. Br J Haematol. 2009 Nov;147(4):444-9.

Arango Duque G, Descoteaux A. Macrophage cytokines: involvement in immunity and infectious diseases. Front Immunol. 2014 Oct 7;5:491.

Arstila TP, Casrouge A, Baron V, Even J, Kanellopoulos J, Kourilsky P. A direct estimate of the human alphabeta T cell receptor diversity. Science. 1999 Oct 29;286(5441):958-61.

Bachireddy P, Wu CJ. Arresting the Inflammatory Drive of Chronic Lymphocytic Leukemia with Ibrutinib. Clin Cancer Res. 2016 Apr 1;22(7):1547-9.

Baliakas P, Agathangelidis A, Hadzidimitriou A, Sutton LA, Minga E, Tsanousa A, Scarfò L, Davis Z, Yan XJ, Shanafelt T, Plevova K, Sandberg Y, Vojdeman FJ, Boudjogra M, Tzenou T, Chatzouli M, Chu CC, Veronese S, Gardiner A, Mansouri L, Smedby KE, Pedersen LB, Moreno D, Van Lom K, Giudicelli V, Francova HS, Nguyen-Khac F, Panagiotidis P, Juliusson G, Angelis L, Anagnostopoulos A, Lefranc MP, Facco M, Trentin L, Catherwood M, Montillo M, Geisler CH, Langerak AW, Pospisilova S, Chiorazzi N, Oscier D, Jelinek DF, Darzentas N, Belessi C, Davi F, Ghia P, Rosenquist R, Stamatopoulos K. Not all IGHV3-21 chronic lymphocytic leukemias are equal: prognostic considerations. Blood. 2015 Jan 29;125(5):856-9.

Barr PM, Zent CS. Infection in chronic lymphocytic leukemia: parsimony has its limits. Leuk Lymphoma. 2014 Dec;55(12):2683-4.

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol. 1995; 57(1): 289-300.

Binder M, Müller F, Frick M, Wehr C, Simon F, Leistler B, Veelken H, Mertelsmann R, Trepel M. CLL B cell receptors can recognize themselves: alternative epitopes and structural clues for autostimulatory mechanisms in CLL. Blood. 2013 Jan 3;121(1):239-41.

Binet JL, Auquier A, Dighiero G, Chastang C, Piguet H, Goasguen J, Vaugier G, Potron G, Colona P, Oberling F, Thomas M, Tchernia G, Jacquillat C, Boivin P, Lesty C, Duault MT, Monconduit M, Belabbes S, Gremy F. A new prognostic classification of chronic lymphocytic leukemia derived from a multivariate survival analysis. Cancer. 1981 Jul 1;48(1):198-206.

Blanco G, Puiggros A, Baliakas P, Athanasiadou A, García-Malo M, Collado R, Xochelli A, Rodríguez-Rivera M, Ortega M, Calasanz MJ, Luño E, Vargas M, Grau J, Martínez-Laperche C, Valiente A, Cervera J, Anagnostopoulos A, Gimeno E, Abella E, Stalika E, Hernández-Rivas JM, Ortuño FJ, Robles D, Ferrer A, Ivars D, González M, Bosch F, Abrisqueta P, Stamatopoulos K, Espinet B. Karyotypic complexity rather than chromosome 8 abnormalities aggravates the outcome of chronic lymphocytic leukemia patients with *TP53* aberrations. Oncotarget. 2016 Dec 6;7(49):80916-80924.

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 Aug 1;76(3):562-9.

Burger JA, Chiorazzi N. B cell receptor signaling in chronic lymphocytic leukemia. Trends Immunol. 2013 Dec;34(12):592-601.

Burger JA, Quiroga MP, Hartmann E, Bürkle A, Wierda WG, Keating MJ, Rosenwald A. High-level expression of the T cell chemokines CCL3 and CCL4 by chronic lymphocytic leukemia B cells in nurselike cell cocultures and after BCR stimulation. Blood. 2009 Mar 26;113(13):3050-8.

Bürgler S, Gimeno A, Parente-Ribes A, Wang D, Os A, Devereux S, Jebsen P, Bogen B, Tjønnfjord GE, Munthe LA. Chronic lymphocytic leukemia cells express CD38 in response to Th1 cell-derived IFN-γ by a T-bet-dependent mechanism. J Immunol. 2015 Jan 15;194(2):827-35.

Buschle M, Campana D, Carding SR, Richard C, Hoffbrand AV, Brenner MK. Interferon gamma inhibits apoptotic cell death in B cell chronic lymphocytic leukemia. J Exp Med. 1993 Jan 1;177(1):213-8.

Cameron DJ, Churchill WH. Cytotoxicity of human macrophages for tumor cells. Enhancement by human lymphocyte mediators. J Clin Invest. 1979 May;63(5):977-84.

Capalbo S, Trerotoli P, Ciancio A, Battista C, Serio G, Liso V. Increased risk of lymphoproliferative disorders in relatives of patients with B cell chronic lymphocytic leukemia: relevance of the degree of familial linkage. Eur J Haematol. 2000 Aug;65(2):114-7.

Carswell EA, Old LJ, Kassel RL, Green S, Fiore N, Williamson B. An endotoxin-induced serum factor that causes necrosis of tumors. Proc Natl Acad Sci U S A. 1975 Sep;72(9):3666-70.

Casabonne D, Almeida J, Nieto WG, Romero A, Fernández-Navarro P, Rodriguez-Caballero A, Muñoz-Criado S, Díaz MG, Benavente Y, de Sanjosé S, Orfao A; Primary Health Care Group of Salamanca for the Study of MBL. Common infectious agents and monoclonal B cell lymphocytosis: a cross-sectional epidemiological study among healthy adults. PLoS One. 2012;7(12):e52808.

Catera R, Silverman GJ, Hatzi K, Seiler T, Didier S, Zhang L, Hervé M, Meffre E, Oscier DG, Vlassara H, Scofield RH, Chen Y, Allen SL, Kolitz J, Rai KR, Chu CC, Chiorazzi N. Chronic lymphocytic

leukemia cells recognize conserved epitopes associated with apoptosis and oxidation. Mol Med. 2008 Nov-Dec;14(11-12):665-74.

Chiorazzi N, Rai KR, Ferrarini M. Chronic lymphocytic leukemia. N Engl J Med. 2005 Feb 24;352(8):804-15.

Chu CC, Catera R, Hatzi K, Yan XJ, Zhang L, Wang XB, Fales HM, Allen SL, Kolitz JE, Rai KR, Chiorazzi N. Chronic lymphocytic leukemia antibodies with a common stereotypic rearrangement recognize nonmuscle myosin heavy chain IIA. Blood. 2008 Dec 15;112(13):5122-9.

Cramer P, Hallek M. Prognostic factors in chronic lymphocytic leukemia-what do we need to know? Nat Rev Clin Oncol. 2011 Jan;8(1):38-47.

Crespo M, Bosch F, Villamor N, Bellosillo B, Colomer D, Rozman M, Marcé S, López-Guillermo A, Campo E, Montserrat E. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. N Engl J Med. 2003 May 1;348(18):1764-75.

Criado I, Muñoz-Criado S, Rodríguez-Caballero A, Nieto WG, Romero A, Fernández-Navarro P, Alcoceba M, Contreras T, González M, Orfao A, Almeida J; Primary Health Care Group of Salamanca for the Study of MBL. Host virus and pneumococcus-specific immune responses in high-count monoclonal B cell lymphocytosis and chronic lymphocytic leukemia: implications for disease progression. Haematologica. 2017 Jul;102(7):1238-1246.

Damle RN, Wasil T, Fais F, Ghiotto F, Valetto A, Allen SL, Buchbinder A, Budman D, Dittmar K, Kolitz J, Lichtman SM, Schulman P, Vinciguerra VP, Rai KR, Ferrarini M, Chiorazzi N. Ig V gene mutation status and CD38 expression as novel prognostic indicators in chronic lymphocytic leukemia. Blood. 1999 Sep 15;94(6):1840-7.

D'Arena G, Rossi G, Minervini MM, Savino L, D'Auria F, Laurenti L, Del Principe MI, Deaglio S, Biagi A, De Martino L, De Feo V, Statuto T, Musto P, Del Poeta G. Circulating regulatory T cells in “clinical” monoclonal B cell lymphocytosis. Int J Immunopathol Pharmacol. 2011 Oct-Dec;24(4):915-23.

Dasgupta A, Mahapatra M, Saxena R. Flow cytometric immunophenotyping of regulatory T cells in chronic lymphocytic

leukemia: comparative assessment of various markers and use of novel antibody panel with CD127 as alternative to transcription factor FoxP3. Leuk Lymphoma. 2013 Apr;54(4):778-89.

Delgado J, Salaverria I, Baumann T, Martínez-Trillos A, Lee E, Jiménez L, Navarro A, Royo C, Santacruz R, López C, Payer AR, Colado E, González M, Armengol L, Colomer D, Pinyol M, Villamor N, Aymerich M, Carrió A, Costa D, Clot G, Giné E, López-Guillermo A, Campo E, Beà S. Genomic complexity and IGHV mutational status are key predictors of outcome of chronic lymphocytic leukemia patients with TP53 disruption. Haematologica. 2014 Nov;99(11):e231-4.

De Weerdt I, van Hoeven V, Munneke JM, Endstra S, Hofland T, Hazenberg MD, Kater AP. Innate lymphoid cells are expanded and functionally altered in chronic lymphocytic leukemia. Haematologica. 2016 Nov;101(11):e461-e464.

Di Bernardo MC, Broderick P, Harris S, Dyer MJ, Matutes E, Dearden C, Catovsky D, Houlston RS. Risk of developing chronic lymphocytic leukemia is influenced by HLA-A class I variation. Leukemia. 2013 Jan;27(1):255-8.

Di Celle PF, Carbone A, Marchis D, Zhou D, Sozzani S, Zupo S, Pini M, Mantovani A, Foa R. Cytokine gene expression in B cell chronic lymphocytic leukemia: evidence of constitutive interleukin-8 (IL-8) mRNA expression and secretion of biologically active IL-8 protein. Blood. 1994 Jul 1;84(1):220-8.

Döhner H, Stilgenbauer S, Benner A, Leupolt E, Kröber A, Bullinger L, Döhner K, Bentz M, Lichter P. Genomic aberrations and survival in chronic lymphocytic leukemia. N Engl J Med. 2000 Dec 28;343(26):1910-6.

Dühren-von Minden M, Übelhart R, Schneider D, Wossning T, Bach MP, Buchner M, Hofmann D, Surova E, Follo M, Köhler F, Wardemann H, Zirlik K, Veelken H, Jumaa H. Chronic lymphocytic leukaemia is driven by antigen-independent cell-autonomous signalling. Nature. 2012 Sep 13;489(7415):309-12.

Edwards JC, Cambridge G. B cell targeting in rheumatoid arthritis and other autoimmune diseases. Nat Rev Immunol. 2006 May;6(5):394-403.

Fabbri G, Dalla-Favera R. The molecular pathogenesis of chronic lymphocytic leukaemia. Nat Rev Cancer. 2016 Mar;16(3):145-62.

Fais F, Ghiotto F, Hashimoto S, Sellars B, Valetto A, Allen SL, Schulman P, Vinciguerra VP, Rai K, Rassenti LZ, Kipps TJ, Dighiero G, Schroeder HW Jr, Ferrarini M, Chiorazzi N. Chronic lymphocytic leukemia B cells express restricted sets of mutated and unmutated antigen receptors. J Clin Invest. 1998 Oct 15;102(8):1515-25.

Fayad L, Keating MJ, Reuben JM, O'Brien S, Lee BN, Lerner S, Kurzrock R. Interleukin-6 and interleukin-10 levels in chronic lymphocytic leukemia: correlation with phenotypic characteristics and outcome. Blood. 2001 Jan 1;97(1):256-63.

Fazi C, Dagklis A, Cottini F, Scarfò L, Bertilaccio MT, Finazzi R, Memoli M, Ghia P. Monoclonal B cell lymphocytosis in hepatitis C virus infected individuals. Cytometry B Clin Cytom. 2010;78 Suppl 1:S61-8.

Fazi C, Scarfò L, Pecciarini L, Cottini F, Dagklis A, Janus A, Talarico A, Scielzo C, Sala C, Toniolo D, Caligaris-Cappio F, Ghia P. General population low-count CLL-like MBL persists over time without clinical progression, although carrying the same cytogenetic abnormalities of CLL. Blood. 2011 Dec 15;118(25):6618-25.

Fiorentino DF, Bond MW, Mosmann TR. Two types of mouse T helper cell. IV. Th2 clones secrete a factor that inhibits cytokine production by Th1 clones. J Exp Med. 1989 Dec 1;170(6):2081-95.

Freeman JD, Warren RL, Webb JR, Nelson BH, Holt RA. Profiling the T cell receptor beta-chain repertoire by massively parallel sequencing. Genome Res. 2009 Oct;19(10):1817-24.

Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. Nat Rev Cancer. 2012 Mar 15;12(4):298-306.

García-Álvarez M, Alcoceba M, López-Parra M, Puig N, Antón A, Balanzategui A, Prieto-Conde I, Jiménez C, Sarasquete ME, Chillón MC, Gutiérrez ML, Corral R, Alonso JM, Queizán JA, Vidán J, Pardal E, Peñarrubia MJ, Bastida JM, García-Sanz R, Marín L, González M. HLA specificities are associated with prognosis in IGHV-mutated CLL-like high-count monoclonal B cell lymphocytosis. PLoS One. 2017 Mar 1;12(3):e0172978.

Gattei V, Bulian P, Del Principe MI, Zucchetto A, Maurillo L, Buccisano F, Bomben R, Dal-Bo M, Luciano F, Rossi FM, Degan M, Amadori S, Del Poeta G. Relevance of CD49d protein expression as overall survival and progressive disease prognosticator in chronic lymphocytic leukemia. Blood. 2008 Jan 15;111(2):865-73.

Ghia P, Prato G, Scielzo C, Stella S, Geuna M, Guida G, Caligaris-Cappio F. Monoclonal CD5+ and CD5- B-lymphocyte expansions are frequent in the peripheral blood of the elderly. Blood. 2004 Mar 15;103(6):2337-42.

Giannoni P, Pietra G, Travaini G, Quarto R, Shyti G, Benelli R, Ottaggio L, Mingari MC, Zupo S, Cutrona G, Pierri I, Balleari E, Pattarozzi A, Calvaruso M, Tripodo C, Ferrarini M, de Totero D. Chronic lymphocytic leukemia nurse-like cells express hepatocyte growth factor receptor (c-MET) and indoleamine 2,3-dioxygenase and display features of immunosuppressive type 2 skewed macrophages. Haematologica. 2014 Jun;99(6):1078-87.

Ginaldi L, De Martinis M, Matutes E, Farahat N, Morilla R, Catovsky D. Levels of expression of CD19 and CD20 in chronic B cell leukaemias. J Clin Pathol. 1998 May;51(5):364-9.

Goldin LR, Björkholm 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 May;94(5):647-53.

Golubovskaya V, Wu L. Different Subsets of T Cells, Memory, Effector Functions, and CAR-T Immunotherapy. Cancers (Basel). 2016 Mar 15;8(3).

Görgün G, Holderried TA, 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 Jul;115(7):1797-805.

Gragert L, Fingerson S, Albrecht M, Maiers M, Kalaycio M, Hill BT. Fine-mapping of HLA associations with chronic lymphocytic leukemia in US populations. Blood. 2014 Oct 23;124(17):2657-65.

Grivennikov SI, Tumanov AV, Liepinsh DJ, Kruglov AA, Marakusha BI, Shakhov AN, Murakami T, Drutskaya LN, Förster I, Clausen BE, Tessarollo L, Ryffel B, Kuprash DV, Nedospasov SA. Distinct and nonredundant in vivo functions of TNF produced by T cells and

macrophages/neutrophils: protective and deleterious effects. Immunity. 2005 Jan;22(1):93-104.

Gustafsson K, Junevik K, Werlenius O, Holmgren S, Karlsson-Parra A, Andersson PO. Tumour-loaded α-type 1-polarized dendritic cells from patients with chronic lymphocytic leukaemia produce a superior NK-, NKT- and CD8+ T cell-attracting chemokine profile. Scand J Immunol. 2011 Sep;74(3):318-26.

Haabeth OA, Lorvik KB, Hammarström C, Donaldson IM, Haraldsen G, Bogen B, Corthay A. Inflammation driven by tumour-specific Th1 cells protects against B cell cancer. Nat Commun. 2011;2:240.

Hallek M. Chronic lymphocytic leukemia: 2017 update on diagnosis, risk stratification, and treatment. Am J Hematol. 2017 Sep;92(9):946-965.

Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner H, Hillmen P, Keating MJ, Montserrat E, Rai KR, Kipps TJ; International Workshop on Chronic Lymphocytic Leukemia. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Lymphocytic Leukemia updating the National Cancer Institute-Working Group 1996 guidelines. Blood. 2008 Jun 15;111(12):5446-56.

Hallek M, Langenmayer I, Nerl C, Knauf W, Dietzfelbinger H, Adorf D, Ostwald M, Busch R, Kuhn-Hallek I, Thiel E, Emmerich B. Elevated serum thymidine kinase levels identify a subgroup at high risk of disease progression in early, nonsmoldering chronic lymphocytic leukemia. Blood. 1999 Mar 1;93(5):1732-7.

Hallek M, Wanders L, Ostwald M, Busch R, Senekowitsch R, Stern S, Schick HD, Kuhn-Hallek I, Emmerich B. Serum beta(2)-microglobulin and serum thymidine kinase are independent predictors of progression-free survival in chronic lymphocytic leukemia and immunocytoma. Leuk Lymphoma. 1996 Aug;22(5-6):439-47.

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 Sep 15;94(6):1848-54.

Hanna BS, McClanahan F, Yazdanparast H, Zaborsky N, Kalter V, Rößner PM, Benner A, Dürr C, Egle A, Gribben JG, Lichter P, Seiffert M. Depletion of CLL-associated patrolling monocytes and macrophages controls disease development and repairs immune dysfunction in vivo. Leukemia. 2016 Mar;30(3):570-9.

Han T, Emrich LJ, Ozer H, Reese PA, Gajera R, Gomez GA, Henderson ES, Bloom ML, Bhargava A, Fitzpatrick J. Clinical significance of serum lactate dehydrogenase in chronic lymphocytic leukemia. N Y State J Med. 1985 Dec;85(12):685-90.

Hatzi K, Catera R, Moreno Atanasio C, Fischetti VA, Allen SL, Kolitz JE, Rai KR, Chu CC, Chiorazzi N. Chronic lymphocytic leukemia immunoglobulins display bacterial reactivity that converges and diverges from auto-/poly-reactivity and IGHV mutation status. Clin Immunol. 2016 Nov;172:44-51.

Heinig K, Gätjen M, Grau M, Stache V, Anagnostopoulos I, Gerlach K, Niesner RA, Cseresnyes Z, Hauser AE, Lenz P, Hehlgans T, Brink R, Westermann J, Dörken B, Lipp M, Lenz G, Rehm A, Höpken UE. Access to follicular dendritic cells is a pivotal step in murine chronic lymphocytic leukemia B cell activation and proliferation. Cancer Discov. 2014 Dec;4(12):1448-65.

HLA Nomenclature; accessed on 8 May 2017; http://hla.alleles.org/nomenclature.

Hus I, Bojarska-Junak A, Chocholska S, Tomczak W, Woś J, Dmoszyńska A, Roliński J. Th17/IL-17A might play a protective role in chronic lymphocytic leukemia immunity. PLoS One. 2013 Nov 1;8(11):e78091.

Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003 Apr;4(2):249-64.

Jadidi-Niaragh F, Yousefi M, Memarian A, Hojjat-Farsangi M, Khoshnoodi J, Razavi SM, Jeddi-Tehrani M, Shokri F. Increased frequency of CD8+ and CD4+ regulatory T cells in chronic lymphocytic leukemia: association with disease progression. Cancer Invest. 2013 Feb;31(2):121-31.

Jaffe ES, Harris NL, Stein H, Isaacson PG. Classification of lymphoid neoplasms: the microscope as a tool for disease discovery. Blood. 2008 Dec 1;112(12):4384-99.

Jain P, Javdan M, Feger FK, Chiu PY, Sison C, Damle RN, Bhuiya TA, Sen F, Abruzzo LV, Burger JA, Rosenwald A, Allen SL, Kolitz JE, Rai KR, Chiorazzi N, Sherry B. Th17 and non-Th17 interleukin-17-expressing cells in chronic lymphocytic leukemia: delineation, distribution, and clinical relevance. Haematologica. 2012 Apr;97(4):599-607.

Janeway CA, Travers P, Walport M, Shlomchik M. Immunobiology: The Immune System in Health and Disease. 5th edition. New York: Garland Science; 2001.

Jitschin R, Braun M, Büttner M, Dettmer-Wilde K, Bricks J, Berger J, Eckart MJ, Krause SW, Oefner PJ, Le Blanc K, Mackensen A, Mougiakakos D. CLL-cells induce IDOhi CD14+HLA-DRlo myeloid-derived suppressor cells that inhibit T cell responses and promote TRegs. Blood. 2014 Jul 31;124(5):750-60.

Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007 Jan;8(1):118-27.

Jones D, Benjamin RJ, Shahsafaei A, Dorfman DM. The chemokine receptor CXCR3 is expressed in a subset of B cell lymphomas and is a marker of B cell chronic lymphocytic leukemia. Blood. 2000 Jan 15;95(2):627-32.

Kakita T, Sasada M, Moriguchi T, Nishimura T, Yamamoto K, Uchino H. Lysis of human leukemic cells by monocyte-derived macrophages activated with interferon-gamma and interleukin-2. Jpn J Cancer Res. 1989 Jan;80(1):59-64.

Kalpadakis C, Pangalis GA, Sachanas S, Vassilakopoulos TP, Kyriakaki S, Korkolopoulou P, Koulieris E, Moschogiannis M, Yiakoumis X, Tsirkinidis P, Kyrtsonis MC, Levidou G, Papadaki HA, Panayiotidis P, Angelopoulou MK. New insights into monoclonal B cell lymphocytosis. Biomed Res Int. 2014;2014:258917.

Kim HJ, Cantor H. CD4 T cell subsets and tumor immunity: the helpful and the not-so-helpful. Cancer Immunol Res. 2014 Feb;2(2):91-8.

Kostareli E, Gounari M, Janus A, Murray F, Brochet X, Giudicelli V, Pospisilova S, Oscier D, Foroni L, di Celle PF, Tichy B, Pedersen LB, Jurlander J, Ponzoni M, Kouvatsi A, Anagnostopoulos A, Thompson K, Darzentas N, Lefranc MP, Belessi C, Rosenquist R, Davi F, Ghia P, Stamatopoulos K. Antigen receptor stereotypy across B cell lymphoproliferations: the case of IGHV4-59/IGKV3-20 receptors with rheumatoid factor activity. Leukemia. 2012 May;26(5):1127-31.

Lad DP, Varma S, Varma N, Sachdeva MU, Bose P, Malhotra P. Regulatory T cells in B cell chronic lymphocytic leukemia: their role in disease progression and autoimmune cytopenias. Leuk Lymphoma. 2013 May;54(5):1012-9.

Lanasa MC, Weinberg JB. Immunologic aspects of monoclonal B cell lymphocytosis. Immunol Res. 2011 Apr;49(1-3):269-80.

Landgren O, Rapkin JS, Caporaso NE, Mellemkjaer L, Gridley G, Goldin LR, Engels EA. Respiratory tract infections and subsequent risk of chronic lymphocytic leukemia. Blood. 2007 Mar 1;109(5):2198-201.

Lanemo Myhrinder A, Hellqvist E, Sidorova E, Söderberg A, Baxendale H, Dahle C, Willander K, Tobin G, Bäckman E, Söderberg O, Rosenquist R, Hörkkö S, Rosén A. A new perspective: molecular motifs on oxidized LDL, apoptotic cells, and bacteria are targets for chronic lymphocytic leukemia antibodies. Blood. 2008 Apr 1;111(7):3838-48.

Lefebvre ML, Krause SW, Salcedo M, Nardin A. Ex vivo-activated human macrophages kill chronic lymphocytic leukemia cells in the presence of rituximab: mechanism of antibody-dependent cellular cytotoxicity and impact of human serum. J Immunother. 2006 Jul-Aug;29(4):388-97.

Lionetti M, Fabris S, Cutrona G, Agnelli L, Ciardullo C, Matis S, Ciceri G, Colombo M, Maura F, Mosca L, Gentile M, Recchia AG, Ilariucci F, Musolino C, Molica S, Di Raimondo F, Cortelezzi A, Rossi D, Gaidano G, Morabito F, Ferrarini M, Neri A. High-throughput sequencing for the identification of *NOTCH1* mutations in early-stage chronic lymphocytic leukaemia: biological and clinical implications. Br J Haematol. 2014 Jun;165(5):629-39.

Li S, Lefranc MP, Miles JJ, Alamyar E, Giudicelli V, Duroux P, Freeman JD, Corbin VD, Scheerlinck JP, Frohman MA, Cameron

PU, Plebanski M, Loveland B, Burrows SR, Papenfuss AT, Gowans EJ. IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. Nat Commun. 2013;4:2333.

Liu J, Zhou Y, Huang Q, Qiu L. CD14(+)HLA-DR(low/-) expression: A novel prognostic factor in chronic lymphocytic leukemia. Oncol Lett. 2015 Mar;9(3):1167-1172.

Lutzny G, Kocher T, Schmidt-Supprian M, Rudelius M, Klein-Hitpass L, Finch AJ, Dürig J, Wagner M, Haferlach C, Kohlmann A, Schnittger S, Seifert M, Wanninger S, Zaborsky N, Oostendorp R, Ruland J, Leitges M, Kuhnt T, Schäfer Y, Lampl B, Peschel C, Egle A, Ringshausen I. Protein kinase c-β-dependent activation of NF-κB in stromal cells is indispensable for the survival of chronic lymphocytic leukemia B cells in vivo. Cancer Cell. 2013 Jan 14;23(1):77-92.

Maffei R, Bulgarelli J, Fiorcari S, Bertoncelli L, Martinelli S, Guarnotta C, Castelli I, Deaglio S, Debbia G, De Biasi S, Bonacorsi G, Zucchini P, Narni F, Tripodo C, Luppi M, Cossarizza A, Marasca R. The monocytic population in chronic lymphocytic leukemia shows altered composition and deregulation of genes involved in phagocytosis and inflammation. Haematologica. 2013 Jul;98(7):1115-23.

Maly K, Schirmer M. The story of $CD4^+$ $CD28^-$ T cells revisited: solved or still ongoing? J Immunol Res. 2015;2015:348746. Erratum in: J Immunol Res. 2015;2015:251657.

Marti GE, Rawstron AC, Ghia P, Hillmen P, Houlston RS, Kay N, Schleinitz TA, Caporaso N; International Familial CLL Consortium. Diagnostic criteria for monoclonal B cell lymphocytosis. Br J Haematol. 2005 Aug;130(3):325-32.

Matzaraki V, Kumar V, Wijmenga C, Zhernakova A. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. Genome Biol. 2017 Apr 27;18(1):76.

McClanahan F, Hanna B, Miller S, Clear AJ, Lichter P, Gribben JG, Seiffert M. PD-L1 checkpoint blockade prevents immune dysfunction and leukemia development in a mouse model of chronic lymphocytic leukemia. Blood. 2015 Jul 9;126(2):203-11.

Menten P, Wuyts A, Van Damme J. Macrophage inflammatory protein-1. Cytokine Growth Factor Rev. 2002 Dec;13(6):455-81.

Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, Rowland JH, Stein KD, Alteri R, Jemal A. Cancer treatment and survivorship statistics, 2016. CA Cancer J Clin. 2016 Jul;66(4):271-89.

Mocellin S, Nitti D. TNF and cancer: the two sides of the coin. Front Biosci. 2008 Jan 1;13:2774-83.

Moreira J, Rabe KG, Cerhan JR, Kay NE, Wilson JW, Call TG, Leis JF, Jelinek DF, Schwager SM, Bowen DA, Hanson CA, Slager SL, Shanafelt TD. Infectious complications among individuals with clinical monoclonal B cell lymphocytosis (MBL): a cohort study of newly diagnosed cases compared to controls. Leukemia. 2013 Jan;27(1):136-41.

Moss SF, Blaser MJ. Mechanisms of disease: Inflammation and the origins of cancer. Nat Clin Pract Oncol. 2005 Feb;2(2):90-7; quiz 1 p following 113.

Nadeu F, Delgado J, Royo C, Baumann T, Stankovic T, Pinyol M, Jares P, Navarro A, Martín-García D, Beà S, Salaverria I, Oldreive C, Aymerich M, Suárez-Cisneros H, Rozman M, Villamor N, Colomer D, López-Guillermo A, González M, Alcoceba M, Terol MJ, Colado E, Puente XS, López-Otín C, Enjuanes A, Campo E. Clinical impact of clonal and subclonal *TP53*, *SF3B1*, *BIRC3*, *NOTCH1*, and *ATM* mutations in chronic lymphocytic leukemia. Blood. 2016 Apr 28;127(17):2122-30.

Nieto WG, Almeida J, Romero A, Teodosio C, López A, Henriques AF, Sánchez ML, Jara-Acevedo M, Rasillo A, González M, Fernández-Navarro P, Vega T, Orfao A; Primary Health Care Group of Salamanca for the Study of MBL. Increased frequency (12%) of circulating chronic lymphocytic leukemia-like B cell clones in healthy subjects using a highly sensitive multicolor flow cytometry approach. Blood. 2009 Jul 2;114(1):33-7.

Nosari A. Infectious complications in chronic lymphocytic leukemia. Mediterr J Hematol Infect Dis. 2012;4(1):e2012070.

Ono C, Yu Z, Kasahara Y, Kikuchi Y, Ishii N, Tomita H. Fluorescently activated cell sorting followed by microarray profiling of helper T cell subtypes from human peripheral blood. PLoS One. 2014 Nov 7;9(11):e111405.

Os A, Bürgler S, Ribes AP, Funderud A, Wang D, Thompson KM, Tjønnfjord GE, Bogen B, Munthe LA. Chronic lymphocytic leukemia cells are activated and proliferate in response to specific T helper cells. Cell Rep. 2013 Aug 15;4(3):566-77.

Padovan E, Casorati G, Dellabona P, Meyer S, Brockhaus M, Lanzavecchia A. Expression of two T cell receptor alpha chains: dual receptor T cells. Science. 1993 Oct 15;262(5132):422-4.

Palma M, Gentilcore G, Heimersson K, Mozaffari F, Näsman-Glaser B, Young E, Rosenquist R, Hansson L, Österborg A, Mellstedt H. T cells in chronic lymphocytic leukemia display dysregulated expression of immune checkpoints and activation markers. Haematologica. 2017 Mar;102(3):562-572.

Pede V, Rombout A, Vermeire J, Naessens E, Vanderstraeten H, Philippé J, Verhasselt B. Expression of ZAP70 in chronic lymphocytic leukaemia activates NF-κB signalling. Br J Haematol. 2013 Dec;163(5):621-30.

Podhorecka M, Dmoszynska A, Rolinski J, Wasik E. T type 1/type 2 subsets balance in B cell chronic lymphocytic leukemia--the three-color flow cytometry analysis. Leuk Res. 2002 Jul;26(7):657-60.

Ponader S, Chen SS, Buggy JJ, Balakrishnan K, Gandhi V, Wierda WG, Keating MJ, O'Brien S, Chiorazzi N, Burger JA. The Bruton tyrosine kinase inhibitor PCI-32765 thwarts chronic lymphocytic leukemia cell survival and tissue homing in vitro and in vivo. Blood. 2012 Feb 2;119(5):1182-9.

Porakishvili N, Kardava L, Jewell AP, Yong K, Glennie MJ, Akbar A, Lydyard PM. Cytotoxic CD4+ T cells in patients with B cell chronic lymphocytic leukemia kill via a perforin-mediated pathway. Haematologica. 2004 Apr;89(4):435-43.

Porakishvili N, Roschupkina T, Kalber T, Jewell AP, Patterson K, Yong K, Lydyard PM. Expansion of CD4+ T cells with a cytotoxic phenotype in patients with B-chronic lymphocytic leukaemia (B-CLL). Clin Exp Immunol. 2001 Oct;126(1):29-36.

Puente XS, Pinyol M, Quesada V, Conde L, Ordóñez GR, Villamor N, Escaramis G, Jares P, Beà S, González-Díaz M, Bassaganyas L, Baumann T, Juan M, López-Guerra M, Colomer D, Tubío JM, López C, Navarro A, Tornador C, Aymerich M, Rozman M, Hernández JM, Puente DA, Freije JM, Velasco G, Gutiérrez-Fernández A, Costa D,

Carrió A, Guijarro S, Enjuanes A, Hernández L, Yagüe J, Nicolás P, Romeo-Casabona CM, Himmelbauer H, Castillo E, Dohm JC, de Sanjosé S, Piris MA, de Alava E, San Miguel J, Royo R, Gelpí JL, Torrents D, Orozco M, Pisano DG, Valencia A, Guigó R, Bayés M, Heath S, Gut M, Klatt P, Marshall J, Raine K, Stebbings LA, Futreal PA, Stratton MR, Campbell PJ, Gut I, López-Guillermo A, Estivill X, Montserrat E, López-Otín C, Campo E. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. Nature. 2011 Jun 5;475(7354):101-5.

Puiggros A, Collado R, Calasanz MJ, Ortega M, Ruiz-Xivillé N, Rivas-Delgado A, Luño E, González T, Navarro B, García-Malo M, Valiente A, Hernández JÁ, Ardanaz MT, Piñan MÁ, Blanco ML, Hernández-Sánchez M, Batlle-López A, Salgado R, Salido M, Ferrer A, Abrisqueta P, Gimeno E, Abella E, Ferrá C, Terol MJ, Ortuño F, Costa D, Moreno C, Carbonell F, Bosch F, Delgado J, Espinet B. Patients with chronic lymphocytic leukemia and complex karyotype show an adverse outcome even in absence of TP53/ ATM FISH deletions. Oncotarget. 2017 Apr 21.

Rai KR, Jain P. Chronic lymphocytic leukemia (CLL)-Then and now. Am J Hematol. 2016 Mar;91(3):330-40.

Rai KR, Sawitsky A, Cronkite EP, Chanana AD, Levy RN, Pasternack BS. Clinical staging of chronic lymphocytic leukemia. Blood. 1975 Aug;46(2):219-34.

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 Aug 16;120(7):1412-21.

Ramsay AG, Evans R, Kiaii S, Svensson L, Hogg N, Gribben JG. Chronic lymphocytic leukemia cells induce defective LFA-1-directed T cell motility by altering Rho GTPase signaling that is reversible with lenalidomide. Blood. 2013 Apr 4;121(14):2704-14.

Ramsay AG, Johnson AJ, Lee AM, Gorgün G, Le Dieu R, Blum W, Byrd JC, Gribben JG. Chronic lymphocytic leukemia T cells show impaired immunological synapse formation that can be reversed with an immunomodulating drug. J Clin Invest. 2008 Jul;118(7):2427-37.

Rawstron AC, Bennett FL, O'Connor SJ, Kwok M, Fenton JA, Plummer M, de Tute R, Owen RG, Richards SJ, Jack AS, Hillmen P. Monoclonal B cell lymphocytosis and chronic lymphocytic leukemia. N Engl J Med. 2008 Aug 7;359(6):575-83.

Rawstron AC, Green MJ, Kuzmicki A, Kennedy B, Fenton JA, Evans PA, O'Connor SJ, Richards SJ, Morgan GJ, Jack AS, Hillmen P. Monoclonal B lymphocytes with the characteristics of "indolent" chronic lymphocytic leukemia are present in 3.5% of adults with normal blood counts. Blood. 2002 Jul 15;100(2):635-9.

Rawstron AC, Shanafelt T, Lanasa MC, Landgren O, Hanson C, Orfao A, Hillmen P, Ghia P. Different biology and clinical outcome according to the absolute numbers of clonal B cells in monoclonal B cell lymphocytosis (MBL). Cytometry B Clin Cytom. 2010;78 Suppl 1:S19-23.

Reiners KS, Topolar D, Henke A, Simhadri VR, Kessler J, Sauer M, Bessler M, Hansen HP, Tawadros S, Herling M, Krönke M, Hallek M, Pogge von Strandmann E. Soluble ligands for NK cell receptors promote evasion of chronic lymphocytic leukemia cells from NK cell anti-tumor activity. Blood. 2013 May 2;121(18):3658-65.

Reinisch W, Willheim M, Hilgarth M, Gasché C, Mader R, Szepfalusi S, Steger G, Berger R, Lechner K, Boltz-Nitulescu G. Soluble CD23 reliably reflects disease activity in B cell chronic lymphocytic leukemia. J Clin Oncol. 1994 Oct;12(10):2146-52.

Rezvany MR, Jeddi-Tehrani M, Osterborg A, Kimby E, Wigzell H, Mellstedt H. Oligoclonal TRBV gene usage in B cell chronic lymphocytic leukemia: major perturbations are preferentially seen within the CD4 T cell subset. Blood. 1999 Aug 1;94(3):1063-9.

Riches JC, Davies JK, McClanahan F, Fatah R, Iqbal S, Agrawal S, Ramsay AG, Gribben JG. T cells from CLL patients exhibit features of T cell exhaustion but retain capacity for cytokine production. Blood. 2013 Feb 28;121(9):1612-21.

Rissiek A, Schulze C, Bacher U, Schieferdecker A, Thiele B, Jacholkowski A, Flammiger A, Horn C, Haag F, Tiegs G, Zirlik K, Trepel M, Tolosa E, Binder M. Multidimensional scaling analysis identifies pathological and prognostically relevant profiles of circulating T cells in chronic lymphocytic leukemia. Int J Cancer. 2014 Nov 15;135(10):2370-9.

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015 Apr 20;43(7):e47.

Rodriguez-Vita J, Lawrence T. The resolution of inflammation and cancer. Cytokine Growth Factor Rev. 2010 Feb;21(1):61-5.

Romagnani S. Th1/Th2 cells. Inflamm Bowel Dis. 1999 Nov;5(4):285-94.

Rossi D, Rasi S, Spina V, Bruscaggin A, Monti S, Ciardullo C, Deambrogi C, Khiabanian H, Serra R, Bertoni F, Forconi F, Laurenti L, Marasca R, Dal-Bo M, Rossi FM, Bulian P, Nomdedeu J, Del Poeta G, Gattei V, Pasqualucci L, Rabadan R, Foà R, Dalla-Favera R, Gaidano G. Integrated mutational and cytogenetic analysis identifies new prognostic subgroups in chronic lymphocytic leukemia. Blood. 2013 Feb 21;121(8):1403-12.

Rossi D, Sozzi E, Puma A, De Paoli L, Rasi S, Spina V, Gozzetti A, Tassi M, Cencini E, Raspadori D, Pinto V, Bertoni F, Gattei V, Lauria F, Gaidano G, Forconi F. 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 Jun;146(1):64-75.

Rozman C, Montserrat E. Chronic lymphocytic leukemia. N Engl J Med. 1995 Oct 19;333(16):1052-7. Erratum in: N Engl J Med 1995 Nov 30;333(22):1515.

Rusak M, Eljaszewicz A, Bołkun Ł, Łuksza E, Łapuć I, Piszcz J, Singh P, Dąbrowska M, Bodzenta-Łukaszyk A, Kłoczko J, Moniuszko M. Prognostic significance of PD-1 expression on peripheral blood CD4+ T cells in patients with newly diagnosed chronic lymphocytic leukemia. Pol Arch Med Wewn. 2015;125(7-8):553-9.

Sadelain M, Rivière I, Brentjens R. Targeting tumours with genetically enhanced T lymphocytes. Nat Rev Cancer. 2003 Jan;3(1):35-45.

Sakr R, Massoud M. Chronic Lymphocytic Leukemia (CLL): Upfront Treatment Options in 2017. Hematol Transfus Int J. 2017 Feb;4(2):00076

Scarfò L, Dagklis A, Scielzo C, Fazi C, Ghia P. CLL-like monoclonal B cell lymphocytosis: are we all bound to have it? Semin Cancer Biol. 2010 Dec;20(6):384-90.

Schoenborn JR, Wilson CB. Regulation of interferon-gamma during innate and adaptive immune responses. Adv Immunol. 2007;96:41-101.

Scrivener S, Goddard RV, Kaminski ER, Prentice AG. Abnormal T cell function in B cell chronic lymphocytic leukaemia. Leuk Lymphoma. 2003 Mar;44(3):383-9.

Sellick GS, Goldin LR, Wild RW, Slager SL, Ressenti L, Strom SS, Dyer MJ, Mauro FR, Marti GE, Fuller S, Lyttelton M, Kipps TJ, Keating MJ, Call TG, Catovsky D, Caporaso N, Houlston RS. A high-density SNP genome-wide linkage search of 206 families identifies susceptibility loci for chronic lymphocytic leukemia. Blood. 2007 Nov 1;110(9):3326-33.

Sette A, Chesnut R, Fikes J. HLA expression in cancer: implications for T cell-based immunotherapy. Immunogenetics. 2001 May-Jun;53(4):255-63.

Sewell AK. Why must T cells be cross-reactive? Nat Rev Immunol. 2012 Sep;12(9):669-77.

Shah N, Decker WK, Lapushin R, Xing D, Robinson SN, Yang H, Parmar S, Tung SS, O'Brien S, Fernandez-Viña M, Shpall EJ, Wierda WG. HLA homozygosity and haplotype bias among patients with chronic lymphocytic leukemia: implications for disease control by physiological immune surveillance. Leukemia. 2011 Jun;25(6):1036-9.

Shalapour S, Karin M. Immunity, inflammation, and cancer: an eternal fight between good and evil. J Clin Invest. 2015 Sep;125(9):3347-55.

Shanafelt TD, Ghia P, Lanasa MC, Landgren O, Rawstron AC. Monoclonal B cell lymphocytosis (MBL): biology, natural history and clinical management. Leukemia. 2010 Mar;24(3):512-20.

Shanafelt TD, Kay NE, Rabe KG, Call TG, Zent CS, Maddocks K, Jenkins G, Jelinek DF, Morice WG, Boysen J, Schwager S, Bowen D, Slager SL, Hanson CA. 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 Aug 20;27(24):3959-63.

Sivina M, Hartmann E, Kipps TJ, Rassenti L, Krupnik D, Lerner S, LaPushin R, Xiao L, Huang X, Werner L, Neuberg D, Kantarjian H, O'Brien S, Wierda WG, Keating MJ, Rosenwald A, Burger JA. CCL3 (MIP-1α) plasma levels and the risk for disease progression in chronic lymphocytic leukemia. Blood. 2011 Feb 3;117(5):1662-9.

Sivina M, Werner L, Rassenti L, Ferrajoli A, Wierda WG, Keating MJ, O'Brien S, Neuberg D, Kipps T, Burger JA. Dynamic changes in CCL3 and CCL4 plasma concentrations in patients with chronic lymphocytic leukaemia managed with observation. Br J Haematol. 2016 Oct 21.

Slager SL, Rabe KG, Achenbach SJ, Vachon CM, Goldin LR, Strom SS, Lanasa MC, Spector LG, Rassenti LZ, Leis JF, Camp NJ, Glenn M, Kay NE, Cunningham JM, Hanson CA, Marti GE, Weinberg JB, Morrison VA, Link BK, Call TG, Caporaso NE, Cerhan JR. Genome-wide association study identifies a novel susceptibility locus at 6p21.3 among familial CLL. Blood. 2011 Feb 10;117(6):1911-6.

Stevenson FK, Caligaris-Cappio F. Chronic lymphocytic leukemia: revelations from the B cell receptor. Blood. 2004 Jun 15;103(12):4389-95.

Stevens WB, Netea MG, Kater AP, van der Velden WJ. 'Trained immunity': consequences for lymphoid malignancies. Haematologica. 2016 Dec;101(12):1460-1468.

Sthoeger ZM, Wakai M, Tse DB, Vinciguerra VP, Allen SL, Budman DR, Lichtman SM, Schulman P, Weiselberg LR, Chiorazzi N. Production of autoantibodies by CD5-expressing B lymphocytes from patients with chronic lymphocytic leukemia. J Exp Med. 1989 Jan 1;169(1):255-68.

Strati P, Shanafelt TD. Monoclonal B cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification. Blood. 2015 Jul 23;126(4):454-62.

Swain SL, McKinstry KK, Strutt TM. Expanding roles for $CD4^+$ T cells in immunity to viruses. Nat Rev Immunol. 2012 Jan 20;12(2):136-48.

Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H, Thiele J, Vardiman JW. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. Lyon, France: IARC Press; 2008.

Takatsu K. Interleukin-5 and IL-5 receptor in health and diseases. Proc Jpn Acad Ser B Phys Biol Sci. 2011;87(8):463-85.

Ten Hacken E, Burger JA. Microenvironment dependency in Chronic Lymphocytic Leukemia: The basis for new targeted therapies. Pharmacol Ther. 2014 Dec;144(3):338-48.

Te Raa GD, Tonino SH, Remmerswaal EB, van Houte AJ, Koene HR, van Oers MH, Kater AP. Chronic lymphocytic leukemia specific T cell subset alterations are clone-size dependent and not present in monoclonal B lymphocytosis. Leuk Lymphoma. 2012 Nov;53(11):2321-5.

Trentin L, Agostini C, Facco M, Piazza F, Perin A, Siviero M, Gurrieri C, Galvan S, Adami F, Zambello R, Semenzato G. The chemokine receptor CXCR3 is expressed on malignant B cells and mediates chemotaxis. J Clin Invest. 1999 Jul;104(1):115-21.

Vallejo AN, Nestel AR, Schirmer M, Weyand CM, Goronzy JJ. Aging-related deficiency of CD28 expression in CD4+ T cells is associated with the loss of gene-specific nuclear factor binding activity. J Biol Chem. 1998 Apr 3;273(14):8119-29.

Van Attekum MHA, Terpstra S, Slinger E, von Lindern M, Moerland PD, Jongejan A, Kater AP, Eldering E. Macrophages confer survival signals via CCR1-dependent translational MCL-1 induction in chronic lymphocytic leukemia. Oncogene. 2017 Jun 29;36(26):3651-3660.

Van Dongen JJ, Langerak AW, Brüggemann M, Evans PA, Hummel M, Lavender FL, Delabesse E, Davi F, Schuuring E, García-Sanz R, van Krieken JH, Droese J, González D, Bastard C, White HE, Spaargaren M, González M, Parreira A, Smith JL, Morgan GJ, Kneba M, Macintyre EA. 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 Dec;17(12):2257-317.

Van Dyke DL, Werner L, Rassenti LZ, Neuberg D, Ghia E, Heerema NA, Dal Cin P, Dell Aquila M, Sreekantaiah C, Greaves AW, Kipps TJ, Kay NE. The Dohner fluorescence in situ hybridization prognostic classification of chronic lymphocytic leukaemia (CLL): the CLL Research Consortium experience. Br J Haematol. 2016 Apr;173(1):105-13.
Vardi A, Agathangelidis A, Stalika E, Karypidou M, Siorenta A, Anagnostopoulos A, Rosenquist R, Hadzidimitriou A, Ghia P, Sutton LA, Stamatopoulos K. Antigen Selection Shapes the T cell Repertoire in Chronic Lymphocytic Leukemia. Clin Cancer Res. 2016 Jan 1;22(1):167-74.

Vardi A, Dagklis A, Scarfò L, Jelinek D, Newton D, Bennett F, Almeida J, Rodriguez-Caballero A, Allgood S, Lanasa M, Cortelezzi A, Orlandi E, Veronese S, Montillo M, Rawstron A, Shanafelt T, Orfao A, Stamatopoulos K, Ghia P. Immunogenetics shows that not all MBL are equal: the larger the clone, the more similar to CLL. Blood. 2013 May 30;121(22):4521-8.

Vardi A, Vlachonikola E, Karypidou M, Stalika E, Bikos V, Gemenetzi K, Maramis C, Siorenta A, Anagnostopoulos A, Pospisilova S, Maglaveras N, Chouvarda I, Stamatopoulos K, Hadzidimitriou A. Restrictions in the T cell repertoire of chronic lymphocytic leukemia: high-throughput immunoprofiling supports selection by shared antigenic elements. Leukemia. 2017 Jan 3.

Walton JA, Lydyard PM, Nathwani A, Emery V, Akbar A, Glennie MJ, Porakishvili N. Patients with B cell chronic lymphocytic leukaemia have an expanded population of CD4 perforin expressing T cells enriched for human cytomegalovirus specificity and an effector-memory phenotype. Br J Haematol. 2010 Jan;148(2):274-84.

Watson JD, Wang S, Von Stetina SE, Spencer WC, Levy S, Dexheimer PJ, Kurn N, Heath JD, Miller DM 3rd. Complementary RNA amplification methods enhance microarray identification of transcripts expressed in the C. elegans nervous system. BMC Genomics. 2008 Feb 19;9:84.

Whitaker JA, Shanafelt TD, Poland GA, Kay NE. Room for improvement: immunizations for patients with monoclonal B cell lymphocytosis or chronic lymphocytic leukemia. Clin Adv Hematol Oncol. 2014 Jul;12(7):440-50.

Wierda WG, Johnson MM, Do KA, Manshouri T, Dey A, O'Brien S, Giles FJ, Kantarjian H, Thomas D, Faderl S, Lerner S, Keating M, Albitar M. Plasma interleukin 8 level predicts for survival in chronic lymphocytic leukaemia. Br J Haematol. 2003 Feb;120(3):452-6.

Wierda WG, Zelenetz AD, Gordon LI, Abramson JS, Advani RH, Andreadis CB, Bartlett N, Byrd JC, Caimi P, Fayad LE, Fisher RI, Glenn MJ, Habermann TM, Harris NL, Hernandez-Ilizaliturri F, Hoppe RT, Horwitz SM, Kaminski MS, Kelsey CR, Kim YH, Krivacic S, LaCasce AS, Martin MG, Nademanee A, Porcu P, Press O, Rabinovitch R, Reddy N, Reid E, Roberts K, Saad AA, Snyder ED, Sokol L, Swinnen LJ, Vose JM, Yahalom J, Dwyer MA, Sundar H. NCCN Guidelines Insights: Chronic Lymphocytic Leukemia/Small Lymphocytic Leukemia, Version 1.2017. J Natl Compr Canc Netw. 2017 Mar;15(3):293-311.

Wu QL, Buhtoiarov IN, Sondel PM, Rakhmilevich AL, Ranheim EA. Tumoricidal effects of activated macrophages in a mouse model of chronic lymphocytic leukemia. J Immunol. 2009 Jun 1;182(11):6771-8.

Yan XJ, Dozmorov I, Li W, Yancopoulos S, Sison C, Centola M, Jain P, Allen SL, Kolitz JE, Rai KR, Chiorazzi N, Sherry B. Identification of outcome-correlated cytokine clusters in chronic lymphocytic leukemia. Blood. 2011 Nov 10;118(19):5201-10.

Yin Q, Sivina M, Robins H, Yusko E, Vignali M, O'Brien S, Keating MJ, Ferrajoli A, Estrov Z, Jain N, Wierda WG, Burger JA. Ibrutinib Therapy Increases T Cell Repertoire Diversity in Patients with Chronic Lymphocytic Leukemia. J Immunol. 2017 Feb 15;198(4):1740-1747.

Ysebaert L. T cells in CLL: lost in migration. Blood. 2013 Apr 4;121(14):2580-2.

Zenz T, Mertens D, Küppers R, Döhner H, Stilgenbauer S. From pathogenesis to treatment of chronic lymphocytic leukaemia. Nat Rev Cancer. 2010 Jan;10(1):37-50.

9. ABBREVIATIONS

APC: Antigen-presenting cell

BCR: B cell receptor

CDR3: Complementarity-determining region 3

CLL: Chronic lymphocytic leukemia

CTL: Cytotoxic T lymphocyte

FC: Fold change

FISH: Fluorescence *in situ* hybridization

HC-MBL: High-count CLL-like monoclonal B cell lymphocytosis

HLA: Human leukocyte antigen complex

IFNγ: Interferon gamma

IG: Immunoglobulins

IGHV: V gene region of the immunoglobulin heavy chain

IL8: Interleukin 8

IPA: Ingenuity Pathway Analysis

LC-MBL: Low-count CLL-like monoclonal B cell lymphocytosis

MBL: CLL-like monoclonal B cell lymphocytosis

M-CLL: Chronic lymphocytic leukemia with mutated IGHV

MDSC: Myeloid-derived suppressor cells

MHC: Major histocompatibility complex

PBMC: Peripheral blood mononuclear cells

PD1: Programmed cell death-1

sIg: Surface immunoglobulins

Th1: Type 1 helper T cell

Th2: Type 2 helper T cell

TME: Tumor microenvironment

TNFα: Tumor necrosis factor alpha

TR: T cell receptor

TRB: T cell receptor beta locus

TRBD: T cell receptor beta diversity gene

TRBJ: T cell receptor beta joining gene

TRBV: T cell receptor beta variable gene

Treg: Regulatory T cell

U-CLL: Chronic lymphocytic leukemia with unmutated IGHV