High-throughput analysis of the T-cell receptor gene repertoire in low-count monoclonal B-cell lymphocytosis reveals a distinct profile from chronic lymphocytic leukemia

Monoclonal B-cell lymphocytosis (MBL) is an asymptomatic condition of monoclonal B-cell expansions in the blood of healthy, mostly elderly, individuals.[1,2] MBL is classified into three distinct subtypes: (i) "chronic lymphocytic leukemia (CLL)-like" MBL ($CD5^+CD23^+$), which accounts for the vast majority of cases; (ii) "atypical CLL-like" MBL ($CD5^+CD23\text{-}CD20^{brigh}$; and (iii) "non CLL-like" MBL ($CD5^-$).[3] "CLL-like" MBL is subdivided into two different categories based on clonal size; cases with 0.5-5x10^9 cells/L are categorized as "high-count MBL" (HC-MBL), whereas those with <0.5x10^9 cells/L as "low-count MBL" (LC-MBL).[4] HC-MBL progresses to CLL requiring treatment at a rate of 1-2% per year,[5] whereas the risk of progression for "CLL-like" LC-MBL is negligible despite persisting over time.[6,7]

Recently, we reported that the genomic profiles of LC-MBL, HC-MBL and 'ultra-stable' CLL (no disease progression for ≥10 years) are very similar.[8] On these grounds, we proposed that cross-talk between CLL progenitor cells and the microenvironment might represent a major driver in early stages of the disease. Relevant microenvironmental triggers might be provided by T cells, considering ample evidence for their implication in CLL pathogenesis[9-11] and the existence of T-cell expansions.[12] Such clonal expansions were also evident in HC-MBL and LC-MBL,[6,13] however, the relevant studies had important limitations, such as limited coverage of the expressed repertoire[6] and sequencing depth,[13] thus precluding firm conclusions from being drawn.

In order to overcome these limitations, we characterized the T-cell receptor beta (*TRB*) chain gene repertoire using a high-throughput sequencing approach. Blood samples were collected from individuals from Val Borbera, Italy, where a LC-MBL cohort is regularly followed up (*Online Supplementary Material*). Samples from aged-matched, healthy individuals without MBL from the same region were analyzed as controls; comparisons to CLL[12] were also performed. The research protocol was approved by the Ethics Committee of the San Raffaele Institute and all participants gave written informed consent in accordance with the Declaration of Helsinki.

Samples of blood (5 mL) were obtained from all individuals and processed within 24 h. LC-MBL was diagnosed following a standardized flow cytometry approach[6] (*Online Supplementary Methods*). We analyzed 48 samples from individuals with LC-MBL ("CLL-like" LC-MBL, n=41; "other" LC-MBL subtypes, n=7), and 17 samples from healthy controls (*Online Supplementary Table S1*). *TRBV-TRBD-TRBJ* gene rearrangements were amplified by polymerase chain reaction, sequenced on a MiSeq Sequencer and bioinformatically processed, as previously described.[12]

Overall, 2,357,648 distinct *TRB* clonotypes were identified. Of these, 1,006,126 (42.7%) were expanded (>1 read), whereas the remainder (1,351,522, 57.3%) concerned singletons (=1 read).

The relevance of age in shaping the *TRB* repertoire, reported in a recent next-generation sequencing study,[14] was also evident in our cohort; clonal expansions were present in all sample categories, including healthy donors. However, significantly different mean relative frequencies of expanded clonotypes were evident between sample categories [analysis of variance (ANOVA), $P<0.05$]. In more detail, expanded clonotypes were larger in "CLL-like" LC-MBL (0.018%; range, 0.003-0.14%) than in "other" LC-MBL (0.007%; range, 0.001-0.04%) or in the "healthy" category (0.01%; range, 0.002-0.012%) (Table 1, *Online Supplementary Figure S1*).

Two different approaches were followed to assess clonality in each sample: (i) the ten "major" (most expanded) clonotypes; and (ii) all expanded clonotypes with an individual frequency of >1%. The first approach led to the identification of slightly higher clonality levels in LC-MBL than in healthy individuals, yet lower than those reported in CLL[12] (ANOVA, $P=0.3$) (Figure 1, *Online Supplementary Table S2*). When considering clonotypes with a relative frequency >1%, the average number of expanded clonotypes for the "CLL-like" LC-MBL, "other" LC-MBL and healthy groups was five (range, 0-27), four (range, 2-6) and four (range, 1-11), respectively. The average sum of relative frequencies was 20% in "CLL-like" LC-MBL, 17% in "other" LC-MBL and 15% in the

Table 1. Detailed information regarding the size of *TRB* raw, processed and clonotype data for each sample category of the present cohort.

	"CLL-like" LC-MBL			"Other" LC-MBL			Healthy			All samples		
	total number	average	range	total number	average	range	total number	average	range	total number	average	range
Total raw reads	7,603,175	185,443	24,672 - 504,220	1,502,577	214,654	100,329- 460,168	4,493,126	264,302	94,256 - 700,301	13,598,878	209,214	24,672 - 700,301
Filtered-in reads	6,489,775	158,287	18,073 - 445,426	1,279,733	182,819	79,986 - 383,735	3,846,137	226,243	79,245 - 618,554	11,615,645	178,702	18,073 - 618,554
Filtered-out reads	1,113,400	27,156	3,540 - 72,911	222,844	31,835	14,712 - 76,433	646,989	38,058	9,493 - 82,690	1,983,233	30,511	3,540 - 82,690
Unique clonotypes	1,271,117	31,003	1,914 - 72,394	264,287	37,755	17,739 - 76,236	822,244	48,367	6,843 - 165,145	2,357,648	36,272	1,914 - 165,145
Expanded clonotypes (>1 read)	567,560	13,843	664 - 34,575	120,442	17,206	7,406 - 37,496	318,124	18,713	2,300 - 50,586	1,006,126	15,479	664 - 50,586
Singletons	703,557	17,160	1,250 - 37,934	143,845	20,549	10,333 - 38,740	504,120	29,654	4,543 - 114,559	1,351,522	20,793	1,250 - 114,559

CLL: chronic lymphocyte leukemia; MBL: monoclonal B-cell lymphocytosis.

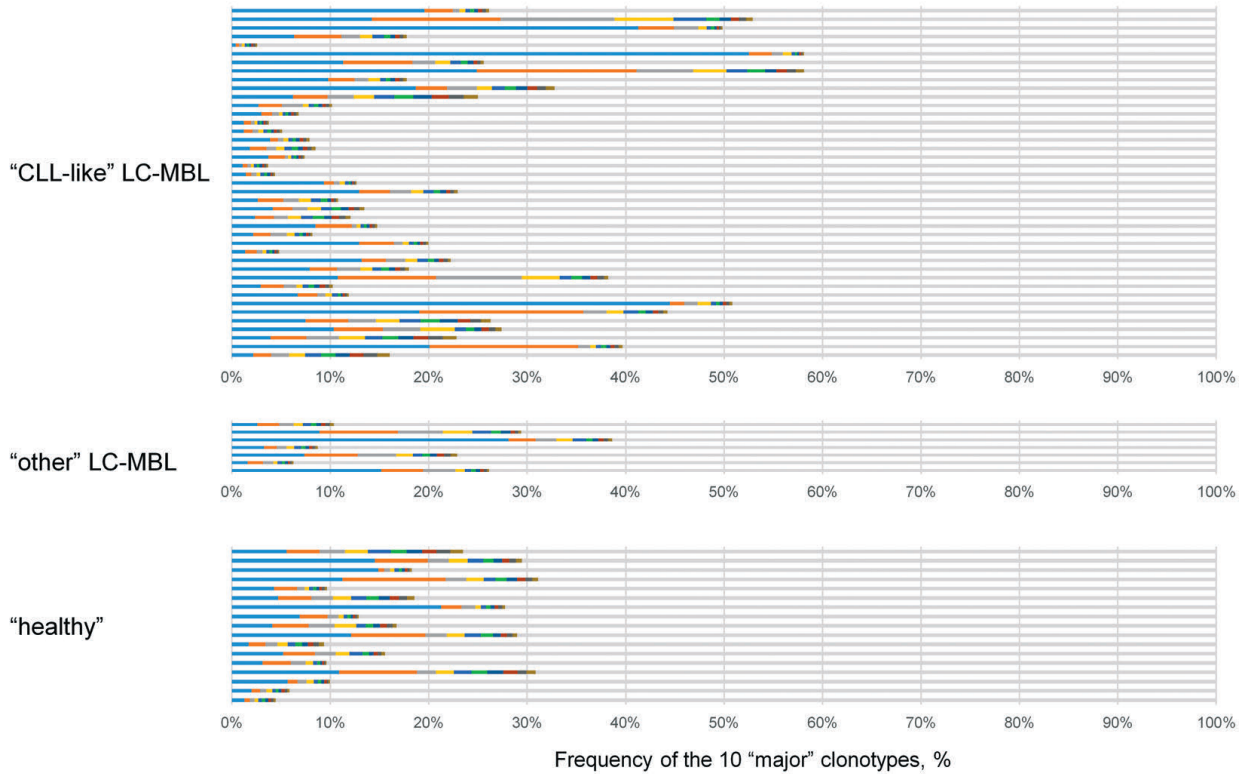

Figure 1. The *TRB* repertoire is oligoclonal in all entities of the present cohort, although to different extents. "Chronic lymphocytic leukemia (CLL)-like" low-count monoclonal B-cell lymphocytosis (LC-MBL) was mostly oligoclonal, followed by the "other" LC-MBL subtypes and the "healthy" compartment. Each lane corresponds to a single sample. Bright colors represent the ten most expanded clonotypes based on their frequency ranking, whereas the polyclonal background is represented by light gray (i.e., the dominant clonotype in all samples is colored blue).

healthy group (*Online Supplementary Table S3*). The average number of expanded clonotypes in CLL was eight (range, 2-27) and their average cumulative frequency was 28% (ANOVA, P=0.1). This trend raises the possibility of more persistent antigenic pressure, leading to more pronounced T-cell expansions in CLL than in LC-MBL or healthy individuals.

Moreover, we investigated whether the level of T-cell clonality could be related to the size of the MBL clone and analyzed the bivariate correlation between the clonal cell size of LC-MBL cases and the cumulative frequency percentage of: (i) the ten major clonotypes; and (ii) all clonotypes with an individual frequency of >1% by computing two correlation indices: the Pearson correlation coefficient and Spearman rho (ρ) correlation coefficient. A significant correlation was detected regarding the latter (Spearman ρ=0.286, P<0.05), in line with our previous claim in CLL[12] that T-cell expansions could be driven, at least in part, by B-cell clone-associated antigen(s); alternatively, the expanded T cells and the CLL/MBL cells could recognize the same antigens. However, the nature of the antigens selecting T cells in MBL remains largely unknown.

We also explored the relation between *TRB* clonality and age in LC-MBL and CLL in comparison to that in healthy donors. Following the approach of a recent next-generation sequencing study,[14] we divided the samples of each category into three subgroups on the basis of the age of the person from whom the sample was taken: "middle-aged" (average 39 years; range, 41-54 years), "aged" (average 62 years; range, 56-68 years), and "long-lived" (average 82 years; range, 69-93 years) (*Online Supplementary Table S4*). We observed a clear tendency for *TRB* clonality to increase with age in all categories (*Online Supplementary Figure S2*), with this effect being more pronounced in CLL (particularly) and LC-MBL than in healthy individuals. This would indicate that T-cell expansions are more relevant in these conditions.

TRBV gene repertoire analysis revealed strong biases in all sample categories (Figure 2), alluding to antigen selection. That said, in principle, this finding could also be attributed, at least in part, to the existence of a restricted gene pool in the study population, given that all samples were obtained from a restricted geographical area.

In "CLL-like" LC-MBL, five *TRBV* genes were identified with a frequency greater than 5%, collectively accounting for a significant fraction of the entire cohort (35.5%) (*Online Supplementary Figure S3A*). A similar scenario was noted in "other" LC-MBL, in which seven genes were dominant (>5%) and collectively represented 44.9% of the repertoire (*Online Supplementary Figure S3B*). The healthy group exhibited somewhat weaker selection biases: four genes occurred with an individual frequency of >5% and a cumulative frequency of 27.9% (*Online Supplementary Figure S3C*).

Subsequently, we performed comparisons to CLL,[12] in which six genes displayed an individual frequency of >5% and collectively accounted for 38.7% of the repertoire. Statistically significant differences concerned the expression of 12/46 (26.1%) *TRBV* genes (ANOVA, P<0.001), of which three were among the most frequent (*TRBV28, TRBV5-1, TRBV19*). Pairwise comparisons

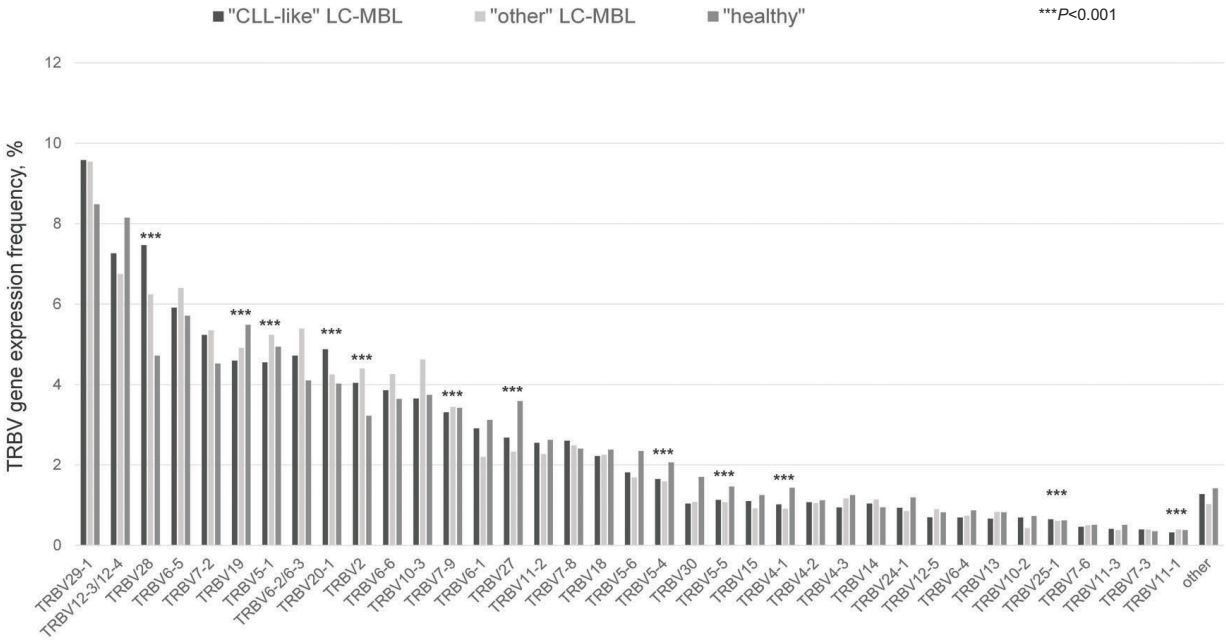

Figure 2. Total *TRBV* gene repertoire for all entities of the present cohort. Statistically significant differences in frequency of utilization were found for 12 of 46 identified *TRBV* genes. (analysis of variance, **P*<0.05, ***P*<0.01 and ****P*<0.001). With regards to the most frequently expressed *TRBV* genes, significant differences were identified for the *TRBV28, TRBV19* and *TRBV5-1* genes.

between all sample groups (*Online Supplementary Table S5*) revealed distinct *TRBV* gene repertoire biases, which might reflect different antigen selection pressures between LC-MBL, CLL and healthy individuals.

Next, we searched for common *TRB* clonotypes across all sample categories of the present cohort and CLL[12] considering the ten "major", hence, more relevant clonotypes of each sample. Most "major" *TRB* clonotypes (841/900, 93.4%) were present only in a single sample. Among the remaining, shared clonotypes, 28 were detected in two samples and one in three samples (*Online Supplementary Table S6*). Most shared clonotypes (15/28, 53.6%) were detected in samples of the same category; eight in "CLL-like" LC-MBL and seven in healthy individuals; the remaining 13 were found in samples from different categories. The average frequency of shared clonotypes was 4.4% with more than half (31/59, 52.5%) displaying a frequency of >1%.

In order to validate the relevance of shared clonotypes, we HLA-typed ten LC-MBL cases with available material that altogether expressed eight shared clonotypes. HLA gene restrictions were found in all sample pairs, indicating that T cells expressing shared clonotypes might recognize the same antigenic epitopes (*Online Supplementary Table S7*). These findings support two main conclusions: (i) shared clonotypes may recognize similar antigenic epitopes; and, (ii) LC-MBL and CLL display distinct T-cell repertoires, given the scarcity of shared clonotypes between the two entities.

In order to better understand the selection processes that drive T-cell expansions in MBL and CLL, we compared *TRB* clonotypes from the present study and CLL with: (i) 6,425 distinct clonotypes from different entities deposited in the IMGT/LIGM-DB and, (ii) expanded *TRB* clonotypes expressed by virus-specific T cells generated against Epstein-Barr virus, cytomegalovirus and BK virus overlapping peptides.[15] We identified 13 hits in the first case; six of these concerned viral infections, whereas the remainder were mostly reported in autoimmune conditions. All shared clonotypes were minor (mean relative frequency: 0.04%; range, 0.01-0.7%) (*Online Supplementary Table S8*). Only a single clonotype present in two samples of our cohort (from the "CLL-like" and "other" LC-MBL sample categories) was also expressed by virus-specific T cells. Hence, "major" *TRB* clonotypes appear to be associated with the intrinsic LC-MBL microenvironment rather than a microbial infection.

In conclusion, the existence of more pronounced T-cell expansions in LC-MBL than in aged-matched healthy individuals indicates that antigen drive could have an effect on the T-cell repertoire, although more limited than in CLL. The nature of the implicated antigen(s) and whether they are related to the LC-MBL clone remains to be clarified. Furthermore, despite overall similar *TRBV* gene usage and degree of oligoclonality, shared clonotypes were scant, thus indicating that the antigenic stimuli and/or immune processes shaping the T-cell receptor profiles in LC-MBL and CLL are probably different.

Andreas Agathangelidis,[1,2] Chrysi Galigalidou,[2,3] Lydia Scarfò,[1] Theodoros Moysiadis,[2,4] Alessandra Rovida,[1] Elisavet Vlachonikola,[2] Electra Sofou,[2] Fotis Psomopoulos,[2] Anna Vardi,[2] Pamela Ranghetti,[1] Alexandra Siorenta,[5] Alex Galanis,[3] Kostas Stamatopoulos,[2,4] Anastasia Chatzidimitriou[2,4] and Paolo Ghia[1*]*

[1]Strategic Research Program on CLL and B-Cell Neoplasia Unit, Division of Experimental Oncology, Università Vita-Salute San Raffaele and IRCCS Ospedale San Raffaele, Milan, Italy; [2]Institute of Applied Biosciences, Center for Research and Technology Hellas, Thessaloniki, Greece; [3]Department of Molecular Biology and Genetics, Democritus University of Thrace, Alexandroupolis, Greece; [4]Department of Molecular Medicine and Surgery, Karolinska Institutet,

Stockholm, Sweden and [3]Immunology and National Tissue Typing Center, [1]G. Gennimatas' General Hospital of Athens, Athens, Thessaloniki, Greece

**AC and PG contributed equally as co-serior authors.*

Correspondence:
ANASTASIA CHATZIDIMITRIOU - achatzidimitriou@certh.gr

doi:10.3324/haematol.2019.221275

Funding: this project received funding from the Hellenic Foundation for Research and Innovation (HFRI) and the General Secretariat for Research and Technology (GSRT), under grant agreement n. 336 (Project CLLon); the TRANSCAN-2 Joint Transnational Call for Proposals 2014 (JTC 2014), GCH-CLL #143, by the European Commission/DG Research and Innovation; the project "KRIPIS II ODYSSEUS" funded by the Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund); the Italian Association for Cancer Research (AIRC, Special Program on Metastasis, 5 per mille # 21198 to PG); PRIN 2015ZMRFEA, Italian Ministry of University and Research – MIUR, Roma, Italy ERA-NET TRANSCAN-2 JTC 2016 #179 NOVEL; open access project, ID number LM2011020, funded by the Ministry of Education, Youth and Sports of the Czech Republic under the activity "Projects of major infrastructures for research, development and innovations".

References

1. Landgren O, Albitar M, Ma W, et al. B-cell clones as early markers for chronic lymphocytic leukemia. N Engl J Med. 2009;360(7):659-667.
2. Rawstron AC, Bennett FL, O'Connor SJ, et al. Monoclonal B-cell lymphocytosis and chronic lymphocytic leukemia. N Engl J Med. 2008;359(6):575-583.
3. Ghia P, Prato G, Scielzo C, et al. Monoclonal CD5+ and CD5- B-lymphocyte expansions are frequent in the peripheral blood of the elderly. Blood. 2004;103(6):2337-2342.
4. Scarfo L, Ghia P. What does it mean I have a monoclonal B-cell lymphocytosis? Recent insights and new challenges. Semin Oncol. 2016;43(2):201-208.
5. Parikh SA, Kay NE, Shanafelt TD. Monoclonal B-cell lymphocytosis: update on diagnosis, clinical outcome, and counseling. Clin Adv Hematol Oncol. 2013;11(11):720-729.
6. Fazi C, Scarfo L, Pecciarini L, et al. General population low-count CLL-like MBL persists over time without clinical progression, although carrying the same cytogenetic abnormalities of CLL. Blood. 2011;118(25):6618-6625.
7. Criado I, Rodriguez-Caballero A, Gutierrez ML, et al. Low-count monoclonal B-cell lymphocytosis persists after seven years of follow up and is associated with a poorer outcome. Haematologica. 2018; 103(7):1198-1208.
8. Agathangelidis A, Ljungstrom V, Scarfo L, et al. Highly similar genomic landscapes in monoclonal B-cell lymphocytosis and ultra-stable chronic lymphocytic leukemia with low frequency of driver mutations. Haematologica. 2018;103(5):865-873.
9. Bagnara D, Kaufman MS, Calissano C, et al. A novel adoptive transfer model of chronic lymphocytic leukemia suggests a key role for T lymphocytes in the disease. Blood. 2011;117(20):5463-5472.
10. Os A, Burgler S, Ribes AP, et al. Chronic lymphocytic leukemia cells are activated and proliferate in response to specific T helper cells. Cell Rep. 2013;4(3):566-577.
11. Kokhaei P, Hojjat-Farsangi M, Mozaffari F, et al. Autologous T cells expressing the oncogenic transcription factor KLF6-SV1 prevent apoptosis of chronic lymphocytic leukemia cells. PLoS One. 2018; 13(2):e0192839.
12. Vardi A, Vlachonikola E, Karypidou M, et al. Restrictions in the T-cell repertoire of chronic lymphocytic leukemia: high-throughput immunoprofiling supports selection by shared antigenic elements. Leukemia. 2017;31(7):1555-1561.
13. Blanco G, Vardi A, Puiggros A, et al. Restricted T cell receptor repertoire in CLL-like monoclonal B cell lymphocytosis and early stage CLL. OncoImmunology. 2018;20(7):e1432328.
14. Britanova OV, Putintseva EV, Shugay M, et al. Age-related decrease in TCR repertoire diversity measured with deep and normalized sequence profiling. J Immunol. 2014;192(6):2689-2698.
15. Galigalidou C, Papadopoulou A, Stalika E, et al. High-throughput T cell receptor (TR) repertoire analysis of virus-specific T cells: implications for T cell immunotherapy and viral infection risk stratification. Blood. 2018;132(Suppl 1):2057.