blood®
Letter to *Blood*

TO THE EDITOR:

Genetic drivers in the natural history of chronic lymphocytic leukemia development as early as 16 years before diagnosis

P. Martijn Kolijn,[1,2,*] Florentin Späth,[3,*] Mouhamad Khouja,[4,*] Paul J. Hengeveld,[1] Lina van der Straten,[1] Nikos Darzentas,[5] Magnus Hultdin,[6] James D. McKay,[7] Christiane Pott,[4] Roel C. H. Vermeulen,[2,†] and Anton W. Langerak[1,†]

[1]Department of Immunology, Laboratory Medical Immunology, Erasmus MC, Rotterdam, The Netherlands; [2]Division of Environmental Epidemiology and Veterinary Public Health, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands; [3]Department of Radiation Sciences, Oncology, Cancer Center, Department of Hematology, Umeå University, Umeå, Sweden; [4]Second Medical Department and [5]Department of Hematology, University Hospital Schleswig-Holstein, Kiel, Germany; [6]Department of Medical Biosciences, Pathology, Umeå University, Umeå, Sweden; and [7]Genomic Epidemiology Branch, International Agency for Research on Cancer, Lyon, France

Chronic lymphocytic leukemia (CLL) is preceded by a prolonged premalignant stage referred to as monoclonal B-cell lymphocytosis (MBL).[1-4] MBL can be detected in up to 17% of the elderly population.[4-6] High-count MBL, defined as a persisting monoclonal B-cell count ≥0.5 × 10^9 cells/L, progresses to CLL requiring treatment at a rate of around 1% per year.[3] Previously, genetic driver mutations have been described in MBL up to 6 years prior to progression to CLL.[7-9] Whole-genome sequencing has been performed in small cohorts of low- and high-count MBL cases.[10] However, pathobiological drivers during the earliest stages of MBL development remain largely elusive.[11]

Recurrently mutated genes in CLL include *SF3B1*, *NOTCH1*, *ATM*, and *TP53*, whereas most other putative CLL driver mutations are present at low frequency (<5% of cases).[12-14] Genome-wide DNA methylation studies have identified profiles correlating with the cell of origin of CLL (pre- or post-germinal center) and the proliferative history of the cell.[15] Integration of genomic, transcriptomic, and epigenomic data has enhanced our understanding of pathobiological diversity in CLL.[14] The most important factors contributing to risk stratification of patients with CLL include the somatic hypermutation (SHM) status of the immunoglobulin heavy variable (IGHV) gene, *TP53* aberration, and stereotypy of the B-cell receptor immunoglobulins (BCR IGs).[16]

Recently, we described the BCR IG gene repertoire during the early stages of CLL in peripheral blood samples drawn up to 22 years before CLL diagnosis.[17] We observed significant BCR IG repertoire skewing and clonotypic evolution regardless of IGHV mutational status or stereotypy, representing the earliest detection of a clonotypic CLL precursor cell.[17] Here, we aim to deepen our insight into driver mutations during the natural history of early CLL development through in-depth study of prediagnostic longitudinal blood samples.

Study subjects were part of the Northern Sweden Health and Disease Study (NSHDS) and the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. For the current study, we selected all participants with longitudinal samples and a prediagnostic clonotype above 2% of the IGH gene repertoire (n = 16), the threshold for prediagnostic IGH gene repertoire skewing we previously identified. The patients had 2 to 6 (median 2) longitudinal samples available. CLL diagnosis ranged from 5 months to 16 years after first blood sampling. Genomic DNA was isolated from buffy coats obtained at blood sampling. Hybrid-capture targeted sequencing was conducted using the EuroClonality-NGS DNA Capture (EuroClonality-NDC) protocol.[18] Annotation was performed through the EuroClonality-NGS–developed ARResT/Interrogate software tool. We identified somatic variants present (variant allele frequency [VAF] > 3%) at CLL diagnosis and traced these variants in earlier samples to study their evolution. Variants with a VAF > 3% prior to CLL diagnosis that remained present at CLL diagnosis were also included. For individuals without a diagnostic sample available (n = 4), the prediagnostic sample drawn closest to diagnosis was used as a reference instead. We screened for variants in 24 recurrently mutated genes in CLL using COSMIC and gnomAD (VAF < 0.01). Low-frequency variants (<3%) in genes of interest were validated using digital droplet polymerase chain reaction (ddPCR). For details, see supplemental Methods, available on the *Blood* website.

The study was approved by the local institutional medical ethical committee at the Erasmus MC (protocol number MEC 2019-0484) and the Ethical Review Board at Ume University (Dnr 2017/242-31). The EPIC steering committee approved the use of the material for the purpose of this study. All patients gave their written consent and the use of the material and data in this study were approved by the IARC Ethics Committee. The study was performed in compliance with the Declaration of Helsinki.

In total, 27 prediagnostic and 12 diagnostic PBMC samples from 16 individuals diagnosed with CLL were included in the study (supplemental Table 1). Of the 16 patients with CLL, 8 (50%) presented with variants of interest including *NOTCH1*,

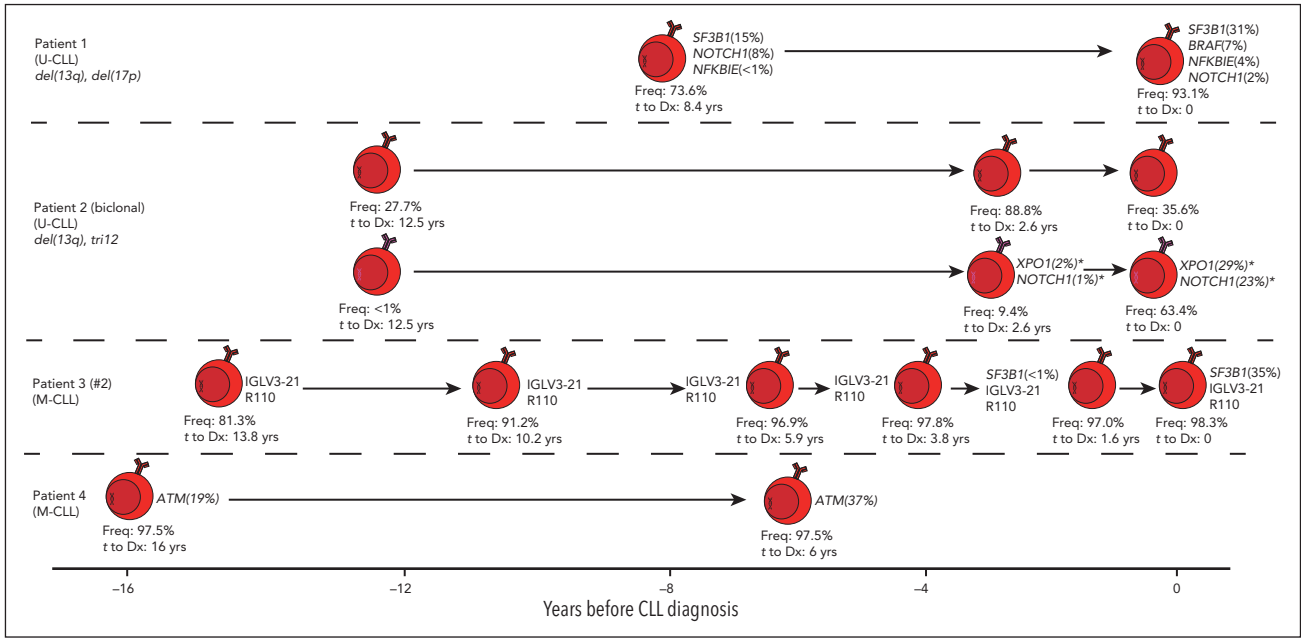

Figure 1. Longitudinal overview of somatic variants observed in CLL-associated genes in 4 patients during the early stages of MBL/CLL development. Only patients with samples likely to constitute a low-count MBL clone are shown (initial sample ranging from 16 to 8.4 years before CLL diagnosis). Each sample is depicted as an MBL/CLL clone for which variants present are indicated alongside the VAF, with variants with a VAF below 1% depicted as <1%. Low-frequency variants (<3%) are only shown if validated through ddPCR. Each sample is labeled with the dominant clonotype frequency (Freq.) as determined through NGS amplicon sequencing of the IGH gene repertoire. X-axis denotes time at which each blood sample was drawn in years before diagnosis. *Notably, *NOTCH1* and *XPO1* variants observed in biclonal CLL patient 2 (indicated with an asterisk) were annotated next to the expanding CLL clone corresponding with the expansion of these variants over time to diagnosis, although formally, we are not able to exclude the possibility that these variants could be present in the other CLL clone in this patient. Stereotypic subsets are indicated for the relevant patients. Cytogenetic aberrations observed at diagnosis are shown when available. The 8 other patients for whom no somatic variants were observed are not included in this graphic.

ATM, and *SF3B1*. No *TP53* variants were observed in our cohort, in line with previous MBL cohort studies.[7-9] Patients with CLL with variants of interest included 4 IGHV-unmutated (U-CLL) cases and 4 IGHV-mutated (M-CLL) cases and encompassed several stereotypic CLL subsets (IGLV3-21^{R110}, #2, #7 and #8; supplemental Table 1). SHM status was stable over time during CLL development in our cohort.

Patient 3 (IGLV3-21^{R110}), with stable MBL for over 10 years, acquired a subclonal *SF3B1* variant less than 2 years before CLL diagnosis, followed by a rapid expansion of this subclone (VAF 35%) at diagnosis (Figures 1 and 2A; supplemental Table 3), and the IGLV3-21^{R110} was present in all prediagnostic samples. This observation does not support speculations that *SF3B1* mutations may predispose a CLL clone to acquiring the IGLV3-21^{R110} mutation.[19] The VAF of the *SF3B1* subclone in patient 1, who does not carry the IGLV3-21^{R110} mutation, increased over time to CLL diagnosis from 15% to 31%. *SF3B1* variants were previously observed more frequently in progressive late-stage CLL (17%) vs CLL diagnosis (5%), suggesting *SF3B1* mutations are acquired during clonal evolution.[20]

Furthermore, we observed a rapid expansion of a *NOTCH1*-mutated subclone in biclonal CLL patient 2, within 3 years prior to diagnosis, and the *NOTCH1*-mutated subclones observed in patients 1 and 5 diminished or remained present at low VAF (Figures 1 and 2A; supplemental Table 3). Interestingly, patients with prediagnostic *NOTCH1*, *SF3B1*, or *XPO1* variants (n = 4) all progressed to CLL-requiring treatment (supplemental Table 4). The characteristic somatic *NOTCH1* variant (c.7541_7542delCT) was previously detected in 11% of MBL and 13.4% of patients with CLL.[21] *NOTCH1* variants previously reported in MBL cases were often subclonal.[21]

In our cohort, patients 4 and 6 presented with somatic *ATM* variants at high VAF (19%-37%) increasing in both patients over time to diagnosis (Figures 1 and 2A; supplemental Table 3). In patient 4, we observed the somatic ATM variant up to 16 years prior to CLL diagnosis, suggesting a role as a driver during the earliest stages of CLL development. In support of this hypothesis, an *ATM* mutation was previously reported in a low-count MBL case at a VAF of 20%, and the prevalence of *ATM* mutations in MBL was previously shown to be comparable to CLL up to 6 years before diagnosis.[7,10]

Patients 6 and 7 had *FAT1* or *PLCG2* variants detectable up to 10 years before CLL diagnosis, suggesting a potential role of these genes in early CLL development. However, *PLCG2* mutations in CLL are primarily described in the context of acquired resistance to *BTK* inhibitors (ibrutinib).[22] Similarly, 10% of fludarabine-refractory patients presented with *FAT1* variants compared with 1% at CLL diagnosis.[23] Hence, it remains unclear if *PLCG2* and *FAT1* variants truly contribute to CLL development or if they are merely passengers.

Although 50% (n = 8) of the patients with CLL in our cohort presented with a mutation in a recurrently mutated gene in CLL, no driver event was found for the remaining patients. One important driver may be the BCR itself, as recent evidence supports ubiquitous autonomous BCR signaling in CLL and MBL.[24] Additionally, epigenetic or small noncoding RNAs drivers have been described to contribute to molecular diversity

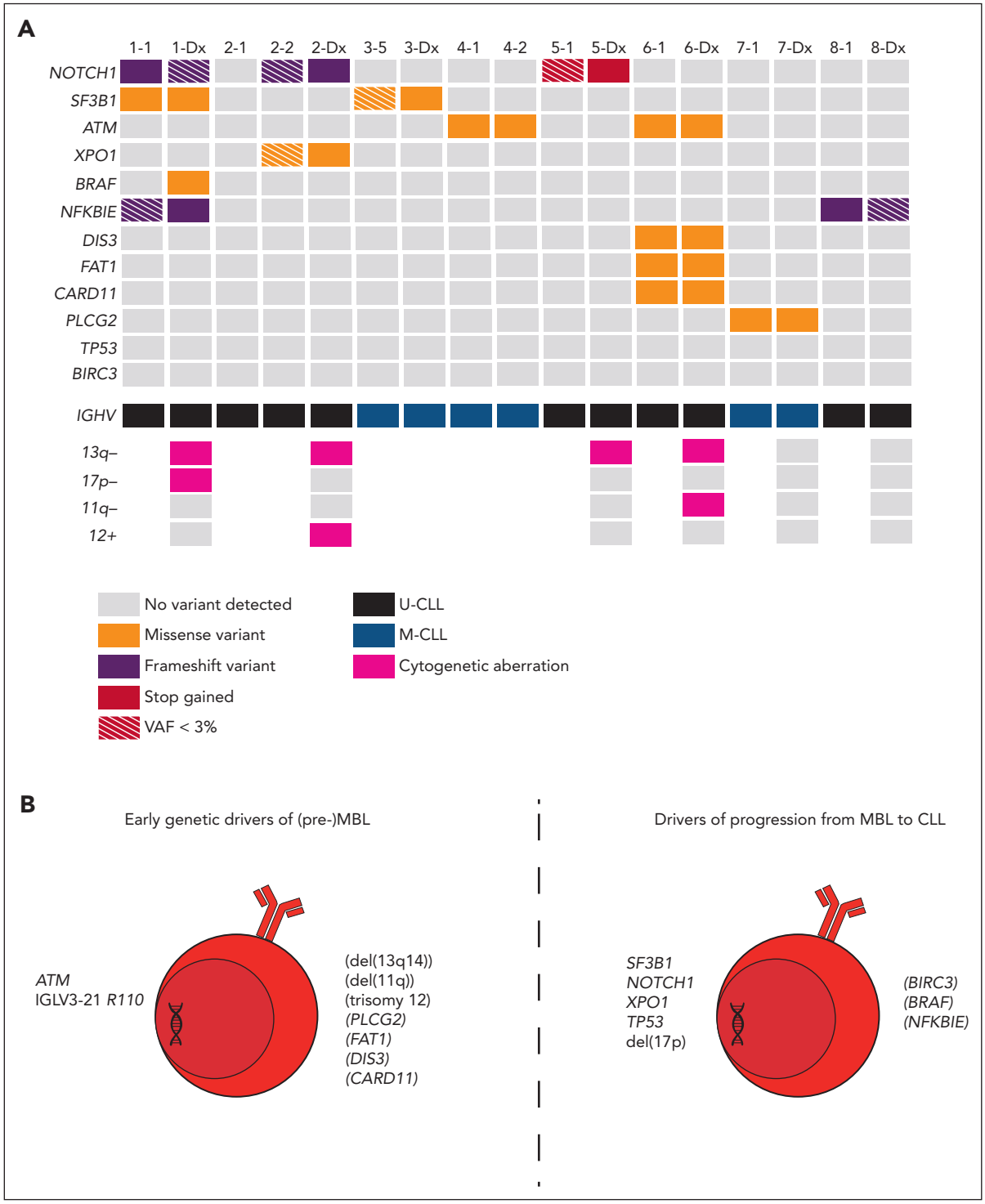

Figure 2. Landscape and presumed role of somatic variants in CLL-associated genes during early MBL development and progression from MBL to CLL. (A) Oncoplot indicating the impact of each of the variants identified. Variants below 3% VAF are indicated with a striped pattern and are only shown if validated through ddPCR. IGHV mutational status and cytogenetic aberrations are indicated for each patient. Cytogenetic aberrations are only available for diagnostic samples. *TP53* and *BIRC3* are shown to highlight the absence of somatic variants in these genes in the (pre)MBL stage. For further information on the identified variants, see supplemental Table 3. (B) Overview of the putative role of variants in the indicated genes. Variants with a putative role during the earliest stages of MBL development (left) are contrasted to variants with a putative role during progression from MBL to CLL (right), based on data from our cohort and literature. Variants with uncertainty (limited support from literature and/or our cohort) are indicated in parentheses.

in CLL during development and progression.[15] Altogether, our findings support the proposed stepwise model for CLL pathogenesis, in which autonomous BCR signaling in genetically predisposed individuals results in a monoclonal expansion of B cells, followed by accumulation of pathogenic somatic variants and progression to CLL.[24]

In conclusion, we provide insights in the occurrence of somatic variants in *NOTCH1*, *ATM*, and *SF3B1* during the pathogenesis of M-CLL, U-CLL, and stereotyped subsets #2 (IGLV3-21^{R110}), #7, and #8. We observed a lack of mutations in *TP53* and a low frequency of mutations in *NOTCH1* and *XPO1* during the earliest stages of CLL development (Figure 2A-B) in keeping with cohort studies in MBL.[7-9,25] Notably, *ATM* variants and the IGLV3-21^{R110} mutation were detected at a high frequency as early as 16 years before CLL diagnosis, indicating a role as early drivers of (pre)MBL.

Acknowledgments

The authors gratefully acknowledge Kostas Stamatopoulos for critical reading of the manuscript. In particular, the authors thank all participants in the initial EPIC and NSHDS studies for providing the samples that enabled our research.

This study was supported by TRANSCAN/Dutch Cancer Society grant 179, the NOVEL Consortium, research grants from the Swedish Cancer Society, the Cancer Research Foundation in Northern Sweden, and the regional agreement between Umeå University and Region Västerbotten.

Where authors are identified as personnel of the International Agency for Research on Cancer/World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer/World Health Organization.

Authorship

Contribution: P.M.K. and M.K. performed the experiments; P.M.K. and M.K. analyzed the data; P.M.K., F.S., M.K., P.J.H., L.v.d.S., M.H., J.D.M., C.P., R.C.H.V., and A.W.L. interpreted results; P.M.K., F.S., M.K., C.P., R.C.H.V., and A.W.L. wrote the manuscript; P.J.H., L.v.d.S., N.D., M.H., and J.D.M. critically reviewed and edited the manuscript; N.D. designed and built the bioinformatics pipeline; F.S. and M.H. facilitated acquisition of patient material and data; and C.P., R.C.H.V., and A.W.L. designed and supervised the study.

Conflict-of-interest disclosure: N.D., C.P., and A.W.L. are members of EuroClonality; EuroClonality receives royalties for the sale of the EuroClonality-NDC assay by Univ8 Genomics (Belfast, Northern Ireland). The remaining authors declare no competing financial interests.

ORCID profiles: F.S., 0000-0002-0711-0830; P.J.H., 0000-0002-0070-138X; A.W.L., 0000-0002-2078-3220.

Correspondence: Anton W. Langerak, Department of Immunology, Laboratory Medical Immunology, Erasmus MC University Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands; email: a.langerak@erasmusmc.nl.

Footnotes

Submitted 3 January 2023; accepted 8 July 2023; prepublished online on *Blood* First Edition 31 July 2023.

*P.M.K., F.S., and M.K. are joint first authors.

†R.C.H.V. and A.W.L are joint last authors.

The data presented in this paper has been uploaded to the Sequence Read Archive (SRA) database under accession ID PRJNA915327.

The online version of this article contains a data supplement.

REFERENCES

1. Hallek M. Chronic lymphocytic leukemia: 2020 update on diagnosis, risk stratification and treatment. *Am J Hematol*. 2019;94(11):1266-1287.
2. 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.
3. Strati P, Shanafelt TD. Monoclonal B-cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification. *Blood*. 2015;126(4):454-462.
4. Henriques A, Rodriguez-Caballero A, Nieto WG, et al. Combined patterns of IGHV repertoire and cytogenetic/molecular alterations in monoclonal B lymphocytosis versus chronic lymphocytic leukemia. *PLoS One*. 2013;8(7):e67751.
5. Slager SL, Parikh SA, Achenbach SJ, et al. Progression and survival of MBL: a screening study of 10 139 individuals. *Blood*. 2022;140(15):1702-1709.
6. Shanafelt TD, Ghia P, Lanasa MC, Landgren O, Rawstron AC. Monoclonal B-cell lymphocytosis (MBL): biology, natural history and clinical management. *Leukemia*. 2010;24(3):512-520.
7. Barrio S, Shanafelt TD, Ojha J, et al. Genomic characterization of high-count MBL cases indicates that early detection of driver mutations and subclonal expansion are predictors of adverse clinical outcome. *Leukemia*. 2017;31(1):170-176.
8. Ojha J, Secreto C, Rabe K, et al. Monoclonal B-cell lymphocytosis is characterized by mutations in CLL putative driver genes and clonal heterogeneity many years before disease progression. *Leukemia*. 2014;28(12):2395-2398.
9. Winkelmann N, Rose-Zerilli M, Forster J, et al. Low frequency mutations independently predict poor treatment-free survival in early stage chronic lymphocytic leukemia and monoclonal B-cell lymphocytosis. *Haematologica*. 2015;100(6):e237-e239.
10. 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.
11. Marti G. The clonal evolution and natural history of MBL. *Blood*. 2022;140(15):1660-1661.
12. Puente XS, Bea S, Valdes-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature*. 2015;526(7574):519-524.
13. Landau DA, Tausch E, Taylor-Weiner AN, et al. Mutations driving CLL and their evolution in progression and relapse. *Nature*. 2015;526(7574):525-530.
14. Knisbacher BA, Lin Z, Hahn CK, et al. Molecular map of chronic lymphocytic leukemia and its impact on outcome. *Nat Genet*. 2022;54(11):1664-1674.
15. Duran-Ferrer M, Clot G, Nadeu F, et al. The proliferative history shapes the DNA methylome of B-cell tumors and predicts clinical outcome. *Nat Cancer*. 2020;1(11):1066-1081.
16. Agathangelidis A, Chatzidimitriou A, Gemenetzi K, et al. Higher-order connections between stereotyped subsets: implications for improved patient classification in CLL. *Blood*. 2021;137(10):1365-1376.
17. Kolijn PM, Hosnijeh FS, Spath F, et al. High-risk subtypes of chronic lymphocytic leukemia are detectable as early as 16 years prior to diagnosis. *Blood*. 2022;139(10):1557-1563.
18. Stewart JP, Gazdova J, Darzentas N, et al. Validation of the EuroClonality-NGS DNA capture panel as an integrated genomic tool for lymphoproliferative disorders. *Blood Adv*. 2021;5(16):3188-3198.
19. Paschold L, Simnica D, Brito RB, et al. Subclonal heterogeneity sheds light on the transformation trajectory in IGLV3-21(R110) chronic lymphocytic leukemia. *Blood Cancer J*. 2022;12(3):49.
20. Rossi D, Bruscaggin A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. *Blood*. 2011;118(26):6904-6908.

21. Lionetti M, Fabris S, Cutrona G, et al. High-throughput sequencing for the identification of NOTCH1 mutations in early stage chronic lymphocytic leukaemia: biological and clinical implications. *Br J Haematol*. 2014; 165(5):629-639.

22. Quinquenel A, Fornecker LM, Letestu R, et al. Prevalence of BTK and PLCG2 mutations in a real-life CLL cohort still on ibrutinib after 3 years: a FILO group study. *Blood*. 2019;134(7):641-644.

23. Messina M, Del Giudice I, Khiabanian H, et al. Genetic lesions associated with chronic lymphocytic leukemia chemo-refractoriness. *Blood*. 2014; 123(15):2378-2388.

24. Sepulveda Yanez JH, Quinten E, Koning MT, et al. Autonomous BCR signaling and genetic aberrations in CLL-phenotype monoclonal B lymphocytosis in siblings of CLL patients [abstract]. *Blood*. 2022; 140(suppl 1):1526-1527.

25. Kleinstern G, O'Brien DR, Li X, et al. Tumor mutational load predicts time to first treatment in chronic lymphocytic leukemia (CLL) and monoclonal B-cell lymphocytosis beyond the CLL international prognostic index. *Am J Hematol*. 2020;95(8):906-917.

https://doi.org/10.1182/blood.2023019609