University of Zurich[UZH]

Zurich Open Repository and Archive
University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2023

BCL3-rearrangements in B-cell lymphoid neoplasms occur in two breakpoint clusters associated with different diseases

Carbo-Meix, Anna ; Guijarro, Francesca ; Wang, Luojun ; Grau, Marta ; Royo, Romina ; Frigola, Gerard ; Playa-Albinyana, Heribert ; Bühler, Marco M ; Clot, Guillem ; Duran-Ferrer, Marti ; Lu, Junyan ; Granada, Isabel ; Baptista, Maria-Joao ; Navarro, Jose-Tomas ; Espinet, Blanca ; Puiggros, Anna ; Tapia, Gustavo ; Bandiera, Laura ; De Canal, Gabriella ; Bonoldi, Emanuela ; Climent, Fina ; Ribera-Cortada, Inmaculada ; Fernandez-Caballero, Mariana ; De la Banda, Esmeralda ; Do Nascimento, Janilson ; Pineda, Alberto ; Vela, Dolors ; Rozman, Maria ; Aymerich, Marta ; Syrykh, Charlotte ; Zenz, Thorsten ; et al

Abstract: The t(14;19)(q32;q13) often juxtaposes BCL3 with IGH resulting in overexpression of the gene. In contrast to other oncogenic translocations, BCL3-rearrangement (BCL3-R) has been associated with a broad spectrum of lymphoid neoplasms. Here we report an integrative whole-genome sequence, transcriptomic, and DNA methylation analysis of 13 lymphoid neoplasms with BCL3-R. The resolution of the breakpoints at single base-pair revealed that they occur in two clusters at 5' (n=9) and 3' (n=4) regions of BCL3 associated with two different biological and clinical entities. Both breakpoints were mediated by aberrant class switch recombination of the IGH locus. However, the 5' breakpoints (upstream) juxtaposed BCL3 next to an IGH enhancer leading to overexpression of the gene whereas the 3' breakpoints (downstream) positioned BCL3 outside the influence of the IGH and were not associated with its expression. Upstream BCL3-R tumors had unmutated IGHV, trisomy 12, and mutated genes frequently seen in CLL but had an atypical CLL morphology, immunophenotype, DNA methylome, and expression profile that differ from conventional CLL. In contrast, downstream BCL3-R neoplasms were atypical splenic or nodal marginal zone lymphomas (MZL) with mutated IGHV, complex karyotypes and mutated genes typical of MZL. Two of the latter 4 tumors transformed to a large B-cell lymphoma. We designed a novel FISH assay that recognizes the two different breakpoints and validated these findings in 17 independent tumors. Overall, upstream or downstream breakpoints of BCL3-R are mainly associated with two subtypes of lymphoid neoplasms with different (epi)genomic, expression, and clinicopathological features resembling atypical CLL and MZL, respectively.

DOI: https://doi.org/10.3324/haematol.2023.283209

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-239064
Journal Article
Published Version

The following work is licensed under a Creative Commons: Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) License.

Originally published at:

Carbo-Meix, Anna; Guijarro, Francesca; Wang, Luojun; Grau, Marta; Royo, Romina; Frigola, Gerard; Playa-Albinyana, Heribert; Bühler, Marco M; Clot, Guillem; Duran-Ferrer, Marti; Lu, Junyan; Granada, Isabel; Baptista, Maria-Joao; Navarro, Jose-Tomas; Espinet, Blanca; Puiggros, Anna; Tapia, Gustavo; Bandiera, Laura; De Canal, Gabriella; Bonoldi, Emanuela; Climent, Fina; Ribera-Cortada, Inmaculada; Fernandez-Caballero, Mariana; De la Banda, Esmeralda; Do Nascimento, Janilson; Pineda, Alberto; Vela, Dolors; Rozman, Maria; Aymerich, Marta; Syrykh, Charlotte; Zenz, Thorsten; et al (2023). BCL3-rearrangements in B-cell lymphoid neoplasms occur in two breakpoint clusters associated with different diseases. Haematologica, 109(2):493-508.
DOI: https://doi.org/10.3324/haematol.2023.283209

haematologica
Journal of The Ferrata Storti Foundation

BCL3-rearrangements in B-cell lymphoid neoplasms occur in two breakpoint clusters associated with different diseases

by Anna Carbó-Meix, Francesca Guijarro, Luojun Wang, Marta Grau, Romina Royo, Gerard Frigola, Heribert Playa-Albinyana, Marco M. Bühler, Guillem Clot, Martí Duran-Ferrer, Junyan Lu, Isabel Granada, Maria-Joao Baptista, José-Tomás Navarro, Blanca Espinet, Anna Puiggros, Gustavo Tapia, Laura Bandiera, Gabriella De Canal, Emanuela Bonoldi, Fina Climent, Inmaculada Ribera-Cortada, Mariana Fernández-Caballero, Esmeralda De la Banda, Janilson Do Nascimento, Alberto Pineda, Dolors Vela, Maria Rozman, Marta Aymerich, Charlotte Syrykh, Pierre Brousset, Miguel Perera, Lucrecia Yañez, Jesús Xavier Ortin, Esperanza Tuset, Thorsten Zenz, James R. Cook, Steven H. Swerdlow, José I. Martín-Subero, Dolors Colomer, stella Matutes, Sílvia Beá, Dolors Costa, Ferran Nadeu, and Elías Campo

Received: March 23, 2023.
Accepted: July 31, 2023.

Citation: Anna Carbó-Meix, Francesca Guijarro, Luojun Wang, Marta Grau, Romina Royo, Gerard Frigola, Heribert Playa-Albinyana, Marco M. Bühler, Guillem Clot, Martí Duran-Ferrer, Junyan Lu, Isabel Granada, Maria-Joao Baptista, José-Tomás Navarro, Blanca Espinet, Anna Puiggros, Gustavo Tapia, Laura Bandiera, Gabriella De Canal, Emanuela Bonoldi, Fina Climent, Inmaculada Ribera-Cortada, Mariana Fernández-Caballero, Esmeralda De la Banda, Janilson Do Nascimento, Alberto Pineda, Dolors Vela, Maria Rozman, Marta Aymeich, Charlotte Syrykh, Pierre Brousset, Miguel Perera, Lucrecia Yañez, Jesús Xavier Ortin, Esperanza Tuset, Thorsten Zenz, James R. Cook, Steven H. Swerdlow, José I. Martín-Subero, Dolors Colomer, Estella Matutes, Sílvia Beá, Dolors Costa, Ferran Nadeu, and Elías Campo. BCL3-rearrangements in B-cell lymphoid neoplasms occur in two breakpoint clusters associated with different diseases. Haematologica. 2023 Aug 10. doi: 10.3324/haematol.2023.283209 [Epub ahead of print]

Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication.
E-publishing of this PDF file has been approved by the authors. After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval; the final version of the manuscript will then appear in a regular issue of the journal.
All legal disclaimers that apply to the journal also pertain to this production process.

BCL3-rearrangements in B-cell lymphoid neoplasms occur in two breakpoint clusters associated with different diseases

Anna Carbó-Meix[1,*], Francesca Guijarro[1,2,*], Luojun Wang[2,*], Marta Grau[1,*], Romina Royo[3], Gerard Frigola[1,2], Heribert Playa-Albinyana[1,4], Marco M. Bühler[5], Guillem Clot[1], Martí Duran-Ferrer[1,4], Junyan Lu[6], Isabel Granada[7], Maria-Joao Baptista[7], José-Tomás Navarro[7], Blanca Espinet[8], Anna Puiggros[8], Gustavo Tapia[9], Laura Bandiera[10], Gabriella De Canal[10], Emanuela Bonoldi[10], Fina Climent[11], Inmaculada Ribera-Cortada[12], Mariana Fernández-Caballero[7], Esmeralda de la Banda[13], Janilson do Nascimento[14], Alberto Pineda[15], Dolors Vela[16], María Rozman[2], Marta Aymerich[1,2], Charlotte Syrykh[17], Pierre Brousset[17,18,19], Miguel Perera[20], Lucrecia Yáñez[21], Jesús Xavier Ortin[22], Esperanza Tuset[23], Thorsten Zenz[24], James R. Cook[25], Steven H. Swerdlow[26], José I. Martín-Subero[1,4,27,28], Dolors Colomer[1,2,4,27], Estella Matutes[2], Sílvia Beà[1,2,4,27], Dolors Costa[1,2,4], Ferran Nadeu[1,4,#] and Elías Campo[1,2,4,27,#]

[1]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain.
[2]Hematopathology Section, laboratory of Pathology, Hospital Clínic de Barcelona, Barcelona, Spain.
[3]Barcelona Supercomputing Center (BSC), Barcelona, Spain.
[4]Centro de Investigación Biomédica en Red de Cáncer (CIBERONC), Madrid, Spain.
[5]Department of Pathology and Molecular Pathology, University Hospital Zurich, Zurich, Switzerland.
[6]European Molecular Biology Laboratory, Heidelberg, Germany.
[7]Department of Hematology-Laboratory, Institut Català d'Oncologia, Hospital Germans Trias i Pujol, Josep Carreras Research Institute, Universitat Autònoma de Barcelona, Badalona, Spain.
[8]Molecular Cytogenetics Laboratory, Pathology Department, Hospital del Mar, Barcelona, Spain and Translational Research on Hematological Neoplasms Group (GRETNHE) - Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain.
[9]Department of Pathology, Hospital Germans Trias i Pujol, Badalona, Spain.
[10]Anatomia Istologia Patologica e Citogenetica, Dipartimento Ematologia, Oncologia e Medicina Molecolare, Niguarda Cancer Center, Milano, Italy.
[11]Department of Pathology, Hospital Universitari de Bellvitge, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), L'Hospitalet De Llobregat, Spain.
[12]Hospital Nostra Senyora de Meritxell, Escaldes-Engordany, Principat d'Andorra.
[13]Laboratory of Hematology, Hospital Universitari Bellvitge, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), L'Hospitalet De Llobregat, Spain.
[14]Hospital Joan XXIII, Institut Català d'Oncologia, Tarragona, Spain.
[15]Fundació Hospital de l'Esperit Sant, Badalona, Spain.
[16]Hematologia Clínica, Hospital General de Granollers, Granollers, Spain.
[17]Department of Pathology, Toulouse University Hospital Center, Cancer Institute University of Toulouse-Oncopole, 1 avenue Irène Joliot-Curie, 31059, Toulouse CEDEX 9, France.
[18]INSERM UMR1037 Cancer Research Center of Toulouse (CRCT), ERL 5294 National Center for Scientific Research (CNRS), University of Toulouse III Paul-Sabatier, Toulouse, France.
[19]Institut Carnot Lymphome CALYM, Laboratoire d'Excellence 'TOUCAN', Toulouse, France.
[20]Hematology Department, Hospital Dr Negrín, Las Palmas de Gran Canaria, Spain
[21]Hematology Department, Hospital Universitario Marqués de Valdecilla-Instituto de Investigación Valdecilla (IDIVAL), Santander, Spain.
[22]Hematology Department, Hospital Verge de la Cinta, Tortosa, Spain

[23]Hematology Department, Institut Català d'Oncologia, Hospital Dr. Josep Trueta, Institut d'Investigació Biomèdica de Girona (IDIBGI), Girona, Spain.
[24]Department of Medical Oncology and Hematology, University Hospital and University of Zürich, Zurich, Switzerland.
[25]Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, OH.
[26]Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA.
[27]Universitat de Barcelona, Barcelona, Spain.
[28]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.
*Contributed equally.
#These authors jointly supervised this work.

Running title: *BCL3* breakpoints identify two B-cell neoplasms

Corresponding author: Elías Campo, Hematopathology Unit, Laboratory of Pathology, Hospital Clinic of Barcelona, Villarroel 170, 08036 Barcelona, Spain; e-mail: ecampo@clinic.cat; phone: +34 93 2275450.

Data-sharing statement: Whole genome sequencing, RNA-seq, and DNA methylation data are available from the European Genome–phenome Archive (http://www.ebi.ac.uk/ega/) under accession no. EGAS00001007465.

Counts: Abstract: 250 words; Main text: 3984 words; Figures: 7; Tables: 1; References: 50. Supplementary files associated: 2.

Authors' contributions

A.C.-M. analyzed and interpreted the WGS, RNA-seq, and DNA methylation data and wrote the manuscript. F.G. collected the samples and clinical data, reviewed the histology, and contributed to manuscript preparation. L.W. reviewed the pathology and contributed to the manuscript preparation. M.G. performed custom FISH experiments and contributed to the manuscript preparation. R.R. designed and performed the bioinformatics pipelines for WGS and RNA-seq data analyses and contributed to the manuscript preparation. G.F. performed the immunohistochemistry experiments and contributed to manuscript preparation. H.P. performed the calcium flux analyses and contributed to manuscript preparation. M.B. contributed to the cases, reviewed the pathology, and prepared the manuscript. G.C., M.D.-F., J.L., I.G., M.-J.B., J.T.N., B.E., A.Pu., G.T., L.B., G.D.C., E.B., F.C., I.R.-C., M.F.-C., E.D.B., J.D.N., A.P., D.V., M.R., M.A., C.S., P.B., M.P., L.Y., J.X.O., E.S., T.Z., J.R.C., S.H.S., J.I.M.-S., D.C., E.M., S.B., and D.C. provided samples and/or data, performed experiments, and interpreted data. F.N. analyzed and interpreted the data, supervised the bioinformatic analyses,

wrote the manuscript, and contributed to the design of the study. E.C. reviewed and supervised the pathology, analyzed and interpreted the data, wrote the manuscript, and designed the study.

Authors' disclosures

MJB is currently an employee of Swedish Orphan Biovitrum. F.N. received honoraria from Janssen, AbbVie, and SOPHiA GENETICS for speaking in educational activities. E.C. has been a consultant for Takeda, NanoString, AbbVie and Illumina; has received honoraria from Janssen, EUSA Pharma and Roche for speaking at educational activities and research funding from AstraZeneca and is an inventor on 2 patents filed by the National Institutes of Health, National Cancer Institute: "Methods for selecting and treating lymphoma types," licensed to NanoString Technologies, and "Evaluation of mantle cell lymphoma and methods related thereof", not related to this project. F.N. and E.C. licensed the use of the protected IgCaller algorithm for Diagnóstica Longwood. The remaining authors declare no competing financial interests.

Acknowledgments

The authors thank the Hematopathology Collection registered at the Biobank of Hospital Clínic - Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), the Biobank HUB-ICO-IDIBELL (PT17/0015/0024), integrated in the Spanish Biobank Network and funded by Instituto de Salud Carlos III (PT17/0015/0024), and Xarxa de Bancs de Tumors de Catalunya sponsored by Pla Director d'Oncologia de Catalunya (XBTC), and the Molecular Cytogenetics Platform of IMIM, Hospital del Mar (Barcelona) for providing BAC clones. This work was partially developed at the Center Esther Koplowitz (CEK, Barcelona, Spain).

Financial support

This study was supported by the "la Caixa" Foundation (CLLEvolution - LCF/PR/HR17/52150017 [HR17-00221LCF] and CLLSYSTEMS - LCF/PR/HR22/52420015 [HR22-00172] Health Research 2017 and 2022 Programs, to E.C.), the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (810287, BCLLatlas, to E.C.), Ministry of Science and Innovation (MCIN) /AEI/10.13039/501100011033/ and European Regional Development Fund "Una manera de hacer Europa" (PID2021-123054OB-I00 to E.C) and the Generalitat de Catalunya Suport Grups de Recerca AGAUR (2021-SGR-01172 to E.C. and 2021-SGR-01293 to S.B.). H.P.-A. is a recipient of a pre-doctoral fellowship from the Spanish Ministry of Science, Innovation and Universities (FPU19/03110). M.D.-F acknowledges the research support from the AECC Scientific Foundation. F.N. acknowledges research support from the American Association for Cancer Research (2021 AACR-Amgen Fellowship in Clinical/Translational Cancer Research, 21-40-11-NADE), European

Hematology Association (EHA Junior Research Grant 2021, RG-202012-00245), and Lady Tata Memorial Trust (International Award for Research in Leukaemia 2021-2022, LADY_TATA_21_3223). E.C. is an Academia Researcher of the "Institució Catalana de Recerca i Estudis Avançats" (ICREA) of the Generalitat de Catalunya.

Abstract

The t(14;19)(q32;q13) often juxtaposes *BCL3* with IGH resulting in overexpression of the gene. In contrast to other oncogenic translocations, *BCL3*-rearrangement (*BCL3*-R) has been associated with a broad spectrum of lymphoid neoplasms. Here we report an integrative whole-genome sequence, transcriptomic, and DNA methylation analysis of 13 lymphoid neoplasms with *BCL3*-R. The resolution of the breakpoints at single base-pair revealed that they occur in two clusters at 5' (n=9) and 3' (n=4) regions of *BCL3* associated with two different biological and clinical entities. Both breakpoints were mediated by aberrant class switch recombination of the IGH locus. However, the 5' breakpoints (upstream) juxtaposed *BCL3* next to an IGH enhancer leading to overexpression of the gene whereas the 3' breakpoints (downstream) positioned *BCL3* outside the influence of the IGH and were not associated with its expression. Upstream *BCL3*-R tumors had unmutated IGHV, trisomy 12, and mutated genes frequently seen in CLL but had an atypical CLL morphology, immunophenotype, DNA methylome, and expression profile that differ from conventional CLL. In contrast, downstream *BCL3*-R neoplasms were atypical splenic or nodal marginal zone lymphomas (MZL) with mutated IGHV, complex karyotypes and mutated genes typical of MZL. Two of the latter 4 tumors transformed to a large B-cell lymphoma. We designed a novel FISH assay that recognizes the two different breakpoints and validated these findings in 17 independent tumors. Overall, upstream or downstream breakpoints of *BCL3*-R are mainly associated with two subtypes of lymphoid neoplasms with different (epi)genomic, expression, and clinicopathological features resembling atypical CLL and MZL, respectively.

Introduction

The t(14;19)(q32;q13) is a balanced translocation found in less than 1% of lymphoid neoplasms that often leads to the juxtaposition of *BCL3* (B-cell leukemia/lymphoma 3) with regulatory elements of the immunoglobulin heavy chain (IGH) gene, resulting in the overexpression of the gene.[1] *BCL3* encodes an IκB-like nuclear protein that regulates NF-κB activity apparently as a molecular adaptor between NF-κB transcription factors and nuclear coactivator and corepressor complexes.[2] Although the function of BCL3 in B cells is not fully understood, this gene seems to be involved in regulation of cell proliferation, differentiation, and survival.[3,4] In transgenic mice, Bcl3 overexpression promoted accumulation of mature B cells but it was not sufficient to drive malignant transformation.[5]

Chromosomal translocations activating oncogenes in lymphoid neoplasms are usually associated with relatively specific tumor subtypes. However, the t(14;19) and *BCL3*-rearrangement (*BCL3*-R) have been identified in a broad spectrum of different tumor subtypes.[6,7] Most patients have been diagnosed with chronic lymphocytic leukemia (CLL), atypical CLL, or transformed CLL. These tumors frequently have an unmutated IGHV (U-IGHV) and trisomy 12. However, they also have atypical features for CLL, including cytology and immunophenotype not characteristic of CLL, frequent IGHV stereotype #8, and aggressive behavior in some series.[6–9] Some authors have suggested that B-cell neoplasms carrying the t(14;19) could represent an entity different from CLL.[6] In addition to these tumors resembling CLL, the t(14;19) and *BCL3*-R have been also identified in diffuse large B-cell lymphomas (DLBCL), marginal zone lymphomas (MZL), splenic small B-cell lymphomas, and tumors diagnosed as B-cell non-Hodgkin lymphomas, some of them with evidence of transformation.[6,7] Whether this diversity of entities associated with *BCL3*-R corresponds to a real biological promiscuity is not clear. Some reports included tumors with the t(14;19) by cytogenetics without the specific analysis of *BCL3*-R. Since this translocation may rearrange genes other than *BCL3*, it is possible that some of the series reported may have included tumors that did not involve *BCL3*. Furthermore, some studies included tumors for which the pathological features were not thoroughly reviewed.[6,7]

The purposes of this study were to characterize the genomic configuration of *BCL3*-R in B-cell neoplasms and to understand the clinical and biological significance of this alteration using an integrative (epi)genomic and transcriptomic analysis in a cohort of patients with available clinical and pathological characteristics.

Methods

Patients and samples

We searched the cytogenetic files of lymphoid B-cell neoplasms with t(14;19) or *BCL3*-R in three institutions from 2008 to 2019. FISH with dual-color break-apart probes for IGH and *BCL3* genes (XL IGH BA and XL BCL3 BA, Metasystems) was performed in patients with available material. Patients with t(14;19), but lacking confirmation of *BCL3*-R, were excluded. Overall, 13 B-cell neoplasms carrying *BCL3*-R, with available material for genomic studies, were identified (Table 1, *Supplementary Figure S1, Supplementary Table S1*). These cases represent 0.28% of all small B-cell lymphomas studied genetically. The initial diagnoses were atypical CLL (aCLL) (n=5), SLL/CLL (n=3), splenic marginal zone lymphoma (SMZL) (n=3), lymphoplasmacytic lymphoma (n=1), and unclassifiable low-grade B-cell leukemic neoplasm (n=1). Tumor DNA was obtained from cryopreserved blood cells or frozen tumor tissue in all patients, germline DNA from non-neoplastic blood cells or saliva (n=11), and RNA from peripheral blood purified cells (n=5) or frozen tissue (n=2). Informed consent was obtained from all patients, and the study was approved by the Ethics Committee of the Hospital Clínic of Barcelona.

Genomic studies

Whole-genome sequencing (WGS) of the 13 tumors and 11 paired normal DNA samples was performed using the TruSeq DNA PCR Free or TruSeq DNA nano library preparation. Raw reads were mapped to the human reference genome (GRCh37) using the BWA-MEM algorithm.[10] Immunoglobulin gene rearrangements were extracted using IgCaller (version 1.1)[11] and annotated using IMGT/V-QUEST.[12] Genomic alterations were identified using a multi-caller bioinformatics approach (*Supplementary Methods*).[13] Driver mutations were studied considering a list of 247 recurrently mutated genes in B-cell neoplasms (*Supplementary Methods* and *Supplementary Table S2*). Total RNA-seq was performed in seven tumors with *BCL3*-R and nine CLL without *BCL3*-R. Raw data were analyzed as previously described (*Supplementary Methods*)[13] using the DESeq2 package.[14] mRNA-seq data from our International Cancer Genome Consortium (ICGC) CLL cohort were used for comparison.[15] DNA methylation profiles of 10 *BCL3*-R tumors were generated using EPIC methylation arrays. Similar data from 85 CLL were obtained for comparison from two previous publications: cohort 1 (C1) included 12 CLL from our institution,[13] and cohort 2 (C2), 73 CLL from University Hospital Heidelberg.[16] Data analyses were performed using minfi and limma packages.[17,18] The AME tool from the MEME suite[19] was used for enrichment analysis of known motifs (2022

JASPAR database, *Supplementary Methods*).[20] WGS, RNA-seq, and DNA methylation data are deposited in the European Genome-phenome Archive.

Immunohistochemistry

BCL3 protein expression was studied by immunohistochemistry in tumors with formalin-fixed paraffin-embedded tissue. Tissue sections (3µm) were stained using a Leica Bond-MAX stainer (Leica Biosystems) and the Anti-BCL3 primary antibody (23959-1-AP; Proteintech) (*Supplementary Methods*).

Custom *BCL3* FISH

Custom *BCL3* break-apart FISH probes to detect 5' and 3' *BCL3* breakpoints were designed using three differentially labeled BAC clones: RP11-927F16 (spectrum orange), CTD-2608C5 (spectrum aqua), and RP11-423N20 (spectrum green) from the Children´s Hospital Oakland Research Institute library obtained from the Molecular Cytogenetics Platform of IMIM (Barcelona, Spain) and Life Technologies. BAC extraction and labeling, slide preparation, and hybridization were performed according to standard procedures.[21]

Results

Genomic characterization of the *BCL3*-rearrangement

We first characterized the breakpoints of the *BCL3*-rearrangement at base-pair resolution using WGS data from 13 tumors (*Supplementary Table S3*). *BCL3* was rearranged with the IGH region as a clonal event in all but one tumor (3646), in which the number of reads suggested a subclonal distribution. All tumors had breakpoints on chromosome 14 within class switch recombination (CSR) regions of the IGH locus (Figure 1A). Breakpoints occurred in IGHA2 (n=1), IGHG2 (n=3), IGHA1 (n=4), IGHG1 (n=3), and IGHG3 (n=3). Breakpoints on chromosome 19 (chr19) were found upstream of the 5' untranslated region (UTR) of *BCL3* in 8/13 (61.5%) tumors (Figure 1A). These breakpoints occurred within a window of 13 kb, and the translocation juxtaposed *BCL3* downstream of the CSR (Figure 1B). Notably, all eight tumors had U-IGHV, six had 100%, and two had 99.6% IGHV identity with the germline (Figure 1A, *Supplementary Table S4*). One additional tumor (3698) with mutated IGHV (M-IGHV) (94.4% identity) had a breakpoint further upstream of *BCL3* truncating *CEACAM16,* although the result of the translocation also placed *BCL3* downstream of the CSR (Figures 1A, C). The four remaining tumors had breakpoints downstream of *BCL3*, two within *CBLC*, one in *BCAM*, and one after *NECTIN2* (Figure 1A). In these four translocations, *BCL3* was not located after the CSR of IGH;

therefore, it does not seem to be the target of the translocation (Figure 1C). All tumors carrying the breakpoint downstream of *BCL3* had M-IGHV with <98% germline identity.

To determine the influence of the chr19 breakpoint on *BCL3* expression, we studied 12 tumors, seven by RNA-seq (six with the breakpoint upstream and one downstream) and seven by immunohistochemistry (IHC; three upstream and four downstream). Two tumors were studied using both approaches (*Supplementary Table S5*). Eight tumors carrying the upstream *BCL3*-R, including one further upstream (3698), overexpressed *BCL3* in comparison to CLL without this rearrangement (Figures 2A-B). No protein expression was detected by IHC in 10 additional nodal CLL with unmutated IGHV (U-CLL) with trisomy 12 and without *BCL3*-R. In contrast, the four tumors downstream *BCL3*-R did not express *BCL3* (Figures 2A-B). The only downstream *BCL3*-R tumor with RNA available (3676) showed overexpression of *NECTIN2*, which was negative in all upstream *BCL3*-R tumors (Figures 2A).

Overall, the location of the chr19 breakpoint distinguishes two main subgroups: 1) tumors with upstream *BCL3*-R breakpoints, which overexpress *BCL3* and are enriched in U-IGHV, and 2) tumors with downstream *BCL3*-R breakpoints, which do not overexpress *BCL3* and carry M-IGHV.

Genomic landscape

Tumors with upstream *BCL3*-R, excluding the three patients lacking normal DNA, had significantly fewer somatic mutations (mean 2511, range 1825-3165; n=7) than tumors with downstream *BCL3*-R, excluding tumor 3646 with a subclonal *BCL3*-R (mean 6271.7, range 4535-9125; n=3) ($p<0.05$; Figure 3A-B, *Supplementary Table S6*). Mutational signature analysis identified the presence of SBS1, SBS5, and SBS8 in all cases, SBS18 in 7 cases, and SBS9 in 3 tumors with M-IGHV (*Supplementary Figure S2* and *Supplementary Table S7*). In addition, we have searched for AID motifs in the mutations occurring in IGH locus between the constant gene and class switch regions. We found that 17/25 (68%) mutations occurred in AID motifs (*Supplementary Table S8*).

The driver mutations in the upstream *BCL3*-R subgroup were very heterogeneous, with only *MED12* and *FAT4* recurrently mutated in 2 tumors each. Other mutated genes have also been frequently described in CLL (*ATM*, *NOTCH1*, *POT1*, *KHL6*). In the four downstream *BCL3*-R tumors, two carried mutations in *KMT2D*, *NOTCH2*, and *KLF2*, frequently seen in MZL, whereas the remaining two tumors had recurrent mutations in *TBL1XR1*, detected in aggressive lymphomas, but also in some MZL (Figure 3A).

In terms of CNA, upstream *BCL3*-R tumors had a significantly lower genomic complexity (mean 2.9, range 1-9; n=9) than downstream *BCL3*-R tumors (mean 11.7, range 5-19; n=3) ($p<0.05$; Figure 3B, *Supplementary Table S9*). All but one upstream *BCL3*-R tumor carried trisomy 12, but this aberration was not observed in any of the downstream *BCL3*-R tumors (Figures 3B-C, *Supplementary Figure S3*). In line with the CNA, the number of SV was lower in upstream *BCL3*-R tumors than in downstream *BCL3*-R tumors (mean 4.8 SV, range 2-10, n=6 vs 18 SV, range 8-28, n=3, respectively) (Figures 3B, *Supplementary Figure S4*, *Supplementary Table S10*).

Gene expression profiling

To determine the gene expression profile of the *BCL3*-R tumors, we compared the RNA-seq data of seven *BCL3*-R tumors (six with upstream and one downstream breakpoint) with nine CLL (four U-CLL and five mutated IGHV (M-CLL)). An unsupervised principal component analysis (PCA) suggested that upstream *BCL3*-R tumors displayed a distinct gene expression profile with some similarities with both M-CLL and U-CLL, whereas the downstream *BCL3*-R tumor did not cluster with any of the other tumors (Figure 4A). Then, we conducted a differential expression analysis (DEA) between upstream *BCL3*-R tumors, all U-IGHV with trisomy 12, four CLL with U-IGHV, and one with trisomy 12 (excluding tumor 3646 with subclonal *BCL3*-R). This analysis identified 1298 differentially expressed genes (DEG): 578 upregulated and 720 downregulated in the upstream *BCL3*-R subgroup ($q<0.05$; Figure 4B, *Supplementary Table S11*). These genes showed similar expression levels in U- and M-CLL (Figure 4B, *Supplementary Table S12*). Significant expression differences were found in genes previously described as characteristically down or upregulated in CLL compared with other B-cell neoplasms.[22–24] Among them, upstream *BCL3*-R tumors had significant overexpression of *EBF1*, usually not expressed in CLL, and, in contrast, downregulation of *LEF1*, *FMOD*, *ADTRP*, *CLNK*, *IGSF3*, and *TCF4*, frequently overexpressed in CLL (Figure 4C).

To rule out a potential confounding effect of trisomy 12, we performed a DEA between 16 U-CLLs with trisomy 12 and 49 U-CLLs without trisomy 12 using data from our ICGC CLL cohort.[15] These analyses identified 1527 DEG ($q<0.05$, absolute(log2FC)>0.1; *Supplementary Table S13*). Among them, only 129 (9.9%) were shared by the upstream *BCL3*-R tumors, suggesting that most DEG observed in *BCL3*-R tumors were not related to trisomy 12 (Figure 4D). Interestingly, most CLL-specific genes modulated in the upstream *BCL3*-R tumors appeared to be independent of trisomy 12 in U-CLL (Figure 4D, *Supplementary Figure S5A*).

Gene set enrichment analyses of upstream *BCL3*-R tumors and U-CLL with trisomy 12 showed that, while both subgroups of tumors shared some genes related to trisomy 12, most other pathways

identified were expressed at lower levels in *BCL3*-R tumors, such as BCR signaling or *TNFα* signaling via *NF-κB* (*Supplementary Figure S5B*, *Supplementary Tables S14-S15*). The lower BCR-signaling capacity of *BCL3*-R tumors was confirmed by measuring Ca^{2+} mobilization upon BCR stimulation with IgM (*Supplementary Figure S6*, *Supplementary Methods*). These findings suggest that, although upstream *BCL3*-R tumors share a subset of commonly expressed genes in CLL carrying trisomy 12, they also have a remarkably distinct profile.

DNA methylation

We analyzed the DNA methylation profile of eight upstream *BCL3*-R tumors, one of which was subclonal, and two downstream *BCL3*-R, and compared them with that of 85 CLL classified as naive-like CLL (n-CLL) (n=33), intermediate CLL (i-CLL) (n=7), or memory-like (m-CLL) (n=45),[25,26] and seven normal B-cell subsets (two naive, one germinal center, three memory, and one plasma cell). We first performed PCA using 764159 CpGs (Figure 5A). Principal component 1 (PC1) reflected the variability related to the proliferative history of the cells captured by the epiCMIT score,[26] whereas PC2 grouped samples based on the cell of origin, in which upstream *BCL3*-R clustered with n-CLL (Figure 5A). Upstream *BCL3*-R tumors had a higher proliferative history than n-CLLs. This observation was confirmed by comparing epiCMIT scores between *BCL3*-R and n-CLL in the C1 (p=0.0043) and C2 (p=0.00016) CLL cohorts (Figure 5B).

To gain further insight into the differences between upstream *BCL3*-R and CLL, we performed a differential methylation analysis between both subgroups of tumors adjusted for trisomy 12, IGHV status, epitope, and cohort (Figure 5C, *Supplementary Figure S7A*). This analysis showed 795 differentially methylated CpGs (DMCpGs), with 80 hyper- and 715 hypomethylated in *BCL3*-R tumors (q <0.05, LFC=0.25, *Supplementary Table S16*). A subset of 21 hypomethylated CpGs in *BCL3*-R tumors was modulated during B-cell differentiation, being hypomethylated in germinal center-experienced normal B cells and M-CLL. Unmethylated CpGs were enriched in heterochromatin and gene bodies, whereas hypermethylated CpGs were enriched in enhancer-promoter regions (*Supplementary Figure S7B*). Among the DMCpGs, 69 mapped to 37 DEG, with 45/64 (70%) hypomethylated CpGs located in the gene body (5'UTR/first exon/body/3'UTR, n=38) or promoter region (TSS1500/TSS200, n=7) of upregulated genes and 4/5 (80%) hypermethylated CpGs mapped to the gene body (n=2) or promoter (n=2) of downregulated genes (Figure 5C, *Supplementary Table S16*). These genes include *EBF1, CREBBP*, and genes associated with *NOTCH1* pathway (*EPS15L1*, *ZMIZ1*),[27,28] cell proliferation (*BHLHE40*, *TP63*),[29–31] cell motility and migration (*CORO1C*, *GAB1*, *GRAMD1B*, *ITGB2*),[32–36] and poor outcomes in CLL or other lymphoid neoplasms (*IMMP2L*, *OSBPL10*, *TP63*).[31,37,38] Notably, *TP63*, previously shown to be a pro-survival factor in CLL subset #8,[31] was

overexpressed in *BCL3*-R cases (Figure 5D). A subsequent transcription factor (TF) binding analysis in the hypomethylated CpGs revealed a significant enrichment in the binding sites of B-cell-related TF such as BCL11B, RUNX3, IRF, JUN/FOS, and FOX families (*Supplementary Table S17*).

Pathology and clinical characteristics

Given the marked genomic differences between upstream and downstream *BCL3*-R tumors, we reanalyzed their pathological and clinical features separately (Table 1, Figure 6, *Supplementary Tables S18-S19*).

Upstream BCL3-R tumors: Tumor cells in peripheral blood were small medium-sized with condensed non-clumped chromatin and broader pale cytoplasm than expected in typical CLL/SLL. Typical clumped chromatin was observed in only one tumor. Five tumors had cells with indented nuclei and seven tumors had prominent nucleoli (Figure 6A). Lymph node biopsies showed diffuse infiltration by small-to medium-sized cells in all tumors. In two tumors the cells had irregular nuclei and prominent nucleoli. Variable numbers of dispersed large cells were observed in all tumors, but clear proliferation centers were observed in only two. Flow cytometry showed expression of mature B-cell markers with CD5 and CD200 positivity in all tumors, but a typical CLL immunophenotype (CD19+, CD79b+, CD5+, CD23+, CD43+, CD200+ with dim expression of CD20, CD22, and FMC7-) was only found in three of the nine tumors (Table 1, *Supplementary Table S19*). The other tumors expressed bright B-cell antigens/surface IgG and/or were dim/negative for CD23 and CD43. In the tissue sections, the four tumors studied were LEF1 negative and no or very scant follicular dendritic networks were observed (Figure 6B).

Downstream BCL3-R tumors: Tumor cells in peripheral blood were larger than those in the previous group, and three of the four tumors had villi or clasmatosis (Figure 6A). The patient without villous lymphocytes had multiple chromosomal alterations that were not specific to any lymphoid neoplasm. The two lymph nodes examined in these patients had infiltration by atypical small cells that partially preserved the architecture, with open sinusoids and occasional residual germinal centers. Tumor cells expanded the perifollicular areas and colonized germinal centers. One patient showed marked monotypic plasmacytosis. The two spleens showed expansion of the white pulp and partial infiltration of the red pulp by small-to medium-sized lymphoid proliferation with occasional larger cells, consistent with SMZL (Figure 6C). The four tumors expressed strong B-cell markers and were CD5 and CD23 negative. CD200 and CD43 were positive in 1/3 of the tumors, and 2/3 expressed IgD. Follicular dendritic cells highlighted the presence of residual germinal centers in all tumors (Figure 6C).

Clinical characteristics

The main clinical difference between the two subgroups was the higher lymphocyte count in the upstream *BCL3*-R subgroup (p=0.01) and splenomegaly in 3 of the 4 patients with downstream *BCL3*-R (p=0.07) (Table 1). Two of the latter patients transformed to a large B-cell lymphoma 5 and 11 years after diagnosis. Transformations were not observed in the upstream *BCL3*-R subgroup with a similar median follow-up time as the downstream tumors (6.3 years versus 5.3 years, respectively, p=0.4).

All downstream *BCL3*-R patients required therapy, in contrast to only 5/9 from the upstream *BCL3*-R subgroup, although no significant differences were found in the median time to first treatment. There were six deaths in the whole cohort, four of which were disease-related, two in the upstream *BCL3*-R subgroup, and two in the downstream *BCL3*-R subgroup, without differences in median survival time. Patients with upstream *BCL3*-R tumors had a similar overall survival as patients with U-CLL and trisomy 12 in our ICGC CLL cohort (*Supplementary Figure S8*).

FISH validation of *BCL3*-rearrangement breakpoints and expanded cohort

As the current commercially available FISH probes do not distinguish between the 5' and 3' rearrangements of *BCL3*-R identified in this study, we designed a new three-color FISH assay that could identify the new *BCL3* 5' and 3' breakpoints (Figure 7A). We tested the new assay in 9 of our 13 tumors and confirmed the breakpoints concordantly with the WGS in all tumors, 7 upstream and 2 downstream (Figure 7B, *Supplementary Figure S9A, Supplementary Table S1*).

We used the new FISH assay in 17 additional B-cell neoplasms with t(14;19) or *BCL3*-R (*Supplementary Table S20-S21, Supplementary Figure S9B*). We identified an upstream breakpoint in 13 tumors and downstream in 4. In line with our previous observations, 11 tumors with upstream breakpoints were diagnosed as aCLL (n=8) or CLL (n=3) and two as leukemic non-nodal MCL. The aCLL had bright B-cell markers and LEF1 was negative in the 6 tumors studied. Trisomy 12 was present in 8/11 and 6/7 had U-IGHV. Lymph nodes examined in 4 cases were consistent with CLL, including prominent proliferation centers in two patients (Figure 7C, *Supplementary Figure S10A*). The two MCL were leukemic non-nodal, with *CCND1* rearrangement and overexpression, and SOX11 negative (*Supplementary Figure S10B*). Three of the four patients with downstream *BCL3*-R were SMZL, one of them with atypical features previously published,[6] splenomegaly and leukemic disease. Two cases carried del(7)(q32) and one case studied mutations frequent in SMZL (*TNFAIP3*, *NOTCH1*, *KMT2D*)(*Supplementary Table S21*).[39]

Discussion

In this study, we characterized the breakpoints of t(14;19) at base-pair resolution in 13 patients with B-cell neoplasms in whom the *BCL3*-rearrangement had been detected by FISH. These tumors showed marked molecular, pathological, and clinical differences according to the location of the breakpoint in the 5' or 3' *BCL3* region, suggesting that they correspond to different entities. Specifically, tumors upstream *BCL3*-R showed *BCL3* overexpression, unmutated IGHV, low genomic complexity, trisomy 12, gene mutations and mutational signatures typically observed in CLL. In contrast, tumors with downstream *BCL3*-R did not upregulate *BCL3* and carried M-IGHV, high genomic complexity, and mutations typically observed in MZLs. Intriguingly, all the breakpoints in the IGHV were mediated by aberrant CSR, but 8 of the 9 tumors with the 5' *BCL3* breakpoints had U-IGHV and 6 of them had 100% identity with the germline, consistent with the fact that CSR occurs before germinal cell commitment and initiation of somatic mutations in the immunoglobulin genes.[40,41]

The pathological features of both subgroups were atypical for CLL or MZL, raising difficulties in their precise taxonomic classification. Upstream *BCL3*-R tumors have characteristics supporting their relationship with CLL including the presence of nodal proliferation centers in some tumors, trisomy 12 in virtually all tumors, and mutations in genes seen in CLL and uncommon in other lymphoid neoplasms. However, the cytological and phenotypic features of most tumors are not completely typical of CLL with bright expression of B-cell antigens and surface Ig, weak or negative CD23 and the expression profile of a subset of genes different from that seen in U-CLL with trisomy 12 such as negative/low expression of *LEF1* and upregulation of *EBF1* among others. In addition, the BCR signaling response was lower than that in U-CLL with trisomy 12. These findings were confirmed in the validation cohort and suggest that lymphoid neoplasms with upstream *BCL3*-R may correspond to a distinct atypical subset of CLL.

Downstream *BCL3*-R tumors had features of MZL with the presence of villous lymphocytes and genetic alterations frequently seen in these tumors (*KLF2*, *NOTCH2, TBL1XR1*). However, they also had some atypical characteristics, such as the exclusive leukemic presentation for 5 and 11 years in two patients and large cell transformation in two of them, an event only seen in 10-15% of SMZL cases.[42] Three of the four tumors with downstream *BCL3*-R in the validation series were also SMZL, two of them with del(7q32).[6] The candidate gene of the downstream *BCL3*-R is unclear. We could only study one of these cases using RNA-seq, which overexpressed *NECTIN2*. This gene, also known as *PVRL2* or CD112, is a member of immunoglobulin-like cell adhesion molecules and a ligand for NK cells. Although its potential oncogenic role is unknown, translocations of this gene with IG and TCR

have been detected in occasional DLBCL and peripheral T-cell lymphomas, respectively.[43,44] Further studies are required to determine whether tumors with downstream *BCL3*-R are a homogeneous group within the marginal zone spectrum.

The biological and clinical differences between tumors with 5' and 3' *BCL3*-R observed in our study may explain the heterogeneity described in the literature. Most of the published tumors resemble our atypical CLL subgroup with an increased frequency of trisomy 12, U-IGHV, and atypical morphology and immunophenotype, although some tumors have also been described as having typical CLL features.[7–9,45] The other subgroup is more heterogeneous with frequent M-IGHV and also MZL characteristics, although with occasional atypical features. Some of the tumors had large B-cell morphology similar to our transformed 3' *BCL3*-R tumors.[6,7,46] The possible prognostic impact of *BCL3*-R in lymphoid neoplasms in the literature is also controversial. Some studies have indicated that CLL or aCLL with *BCL3*-R have an adverse prognosis[8,47–49] but this was not confirmed by others.[50] Our patients with upstream *BCL3*-R had a similar time to the first treatment and overall survival as U-CLL with trisomy 12.

Our new *BCL3*-R FISH assay identified two breakpoints in 11 of the 12 (92%) initial tumors studied and in all 17 independent lymphoid neoplasms, 13 with a 5' breakpoint and four with a 3' breakpoint. Interestingly, 11 tumors with upstream *BCL3*-R had pathological and genetic features similar to those of aCLL/CLL with U-IGHV, trisomy 12, and negative LEF1 expression. The tumors with the 3' breakpoint were three SMZL, with some atypical features.[6] These results confirm the value of this new FISH assay in identifying different *BCL3* breakpoints and diseases. The finding of a 5' *BCL3*-R in two nnMCL suggests that, similar to other translocations in lymphoid neoplasms, *BCL3*-R is not specific to a single entity and needs to be interpreted in the appropriate context.

In conclusion, identification of breakpoints upstream or downstream of *BCL3* revealed two different subgroups of lymphoid neoplasms. Tumors with a 5' breakpoint may correspond to a distinct subset of aCLL/CLL with distinct (epi)genomic, transcriptomic, and clinicopathological features, whereas 3' rearranged tumors appear to be in the MZL spectrum. We developed a novel FISH assay that recognizes these two *BCL3* breakpoints and is therefore useful in clinical practice to identify the two subgroups of patients.

References

1. Michaux L, Mecucci C, Stul M, et al. BCL3 rearrangement and t(14;19)(q32;q13) in lymphoproliferative disorders. Genes Chromosomes Cancer. 1996;15(1):38-47.

2. Palmer S, Chen YH. Bcl-3, a multifaceted modulator of NF-kappaB-mediated gene transcription. Immunol Res 2008;42(1–3):210–218.

3. Liu H, Zeng L, Yang Y, Guo C, Wang H. Bcl-3: A Double-Edged Sword in Immune Cells and Inflammation. Front Immunol. 2022;13:847699.

4. Zhang X, Paun A, Claudio E, Wang H, Siebenlist U. The tumor promoter and NF-κB modulator Bcl-3 regulates splenic B cell development. J Immunol. 2013;191(12):5984-5992.

5. Ong ST, Hackbarth ML, Degenstein LC, Baunoch DA, Anastasi J, McKeithan TW. Lymphadenopathy, splenomegaly, and altered immunoglobulin production in BCL3 transgenic mice. Oncogene. 1998;16(18):2333-2343.

6. Soma LA, Gollin SM, Remstein ED, et al. Splenic small B-cell lymphoma with IGH/BCL3 translocation. Hum Pathol. 2006;37(2):218-230.

7. Martín-Subero JI, Ibbotson R, Klapper W, et al. A comprehensive genetic and histopathologic analysis identifies two subgroups of B-cell malignancies carrying a t(14;19)(q32;q13) or variant BCL3-translocation. Leukemia. 2007;21(7):1532-1544.

8. Kelly RJ, Wright D, Patil K, et al. t(14;19)(q32;q13) incidence and significance in B-cell lymphoproliferative disorders. Br J Haematol. 2008;141(4):561-563.

9. Huh YO, Schweighofer CD, Ketterling RP, et al. Chronic lymphocytic leukemia with t(14;19)(q32;q13) is characterized by atypical morphologic and immunophenotypic features and distinctive genetic features. Am J Clin Pathol. 2011;135(5):686-696.

10. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-1760.

11. Nadeu F, Mas-de-les-Valls R, Navarro A, et al. IgCaller for reconstructing immunoglobulin gene rearrangements and oncogenic translocations from whole-genome sequencing in lymphoid neoplasms. Nat Commun. 2020;11(1):3390.

12. Brochet X, Lefranc M-P, Giudicelli V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. Nucleic Acids Res. 2008;36(Web Server issue):W503-508.

13. Nadeu F, Royo R, Massoni-Badosa R, et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. Nat Med. 2022;28(8):1662-1671.

14. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.

15. Puente XS, Beà S, Valdés-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature. 2015;526(7574):519-524.

16. Lu J, Cannizzaro E, Meier-Abt F, et al. Multi-omics reveals clinically relevant proliferative drive associated with mTOR-MYC-OXPHOS activity in chronic lymphocytic leukemia. Nat Cancer. 2021;2(8):853-864.

17. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.

18. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363-1369.

19. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res. 2015;43(W1):W39-W49.

20. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 2022;50(D1):D165-D173.

21. Ventura RA, Martin-Subero JI, Jones M, et al. FISH analysis for the detection of lymphoma-associated chromosomal abnormalities in routine paraffin-embedded tissue. J Mol Diagn. 2006;8(2):141-151.

22. Navarro A, Clot G, Martínez-Trillos A, et al. Improved classification of leukemic B-cell lymphoproliferative disorders using a transcriptional and genetic classifier. Haematologica. 2017;102(9):e360-e363.

23. Seifert M, Sellmann L, Bloehdorn J, et al. Cellular origin and pathophysiology of chronic lymphocytic leukemia. J Exp Med. 2012;209(12):2183-2198.

24. Gutierrez A, Tschumper RC, Wu X, et al. LEF-1 is a prosurvival factor in chronic lymphocytic leukemia and is expressed in the preleukemic state of monoclonal B-cell lymphocytosis. Blood. 2010;116(16):2975-2983.

25. Kulis M, Heath S, Bibikova M, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet. 2012;44(11):1236-1242.

26. 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.

27. Griffen TL, Dammer EB, Dill CD, et al. Multivariate transcriptome analysis identifies networks and key drivers of chronic lymphocytic leukemia relapse risk and patient survival. BMC Med Genomics. 2021;14(1):171.

28. Pinnell N, Yan R, Cho HJ, et al. The PIAS-like Coactivator Zmiz1 Is a Direct and Selective Cofactor of Notch1 in T Cell Development and Leukemia. Immunity. 2015;43(5):870-883.

29. Cook ME, Jarjour NN, Lin C-C, Edelson BT. Transcription Factor Bhlhe40 in Immunity and Autoimmunity. Trends Immunol. 2020;41(11):1023-1036.

30. Rauschmeier R, Reinhardt A, Gustafsson C, et al. Bhlhe40 function in activated B and TFH cells restrains the GC reaction and prevents lymphomagenesis. J Exp Med. 2022;219(2):e20211406.

31. Papakonstantinou N, Ntoufa S, Tsagiopoulou M, et al. Integrated epigenomic and transcriptomic analysis reveals TP63 as a novel player in clinically aggressive chronic lymphocytic leukemia. Int J Cancer. 2019;144(11):2695-2706.

32. Roadcap DW, Clemen CS, Bear JE. The role of mammalian coronins in development and disease. Subcell Biochem. 2008;48:124-135.

33. Seda V, Vojackova E, Ondrisova L, et al. FoxO1-GAB1 axis regulates homing capacity and tonic AKT activity in chronic lymphocytic leukemia. Blood. 2021;138(9):758-772.

34. Khanna P, Lee JS, Sereemaspun A, Lee H, Baeg GH. GRAMD1B regulates cell migration in breast cancer cells through JAK/STAT and Akt signaling. Sci Rep. 2018;8(1):9511.

35. Hutterer E, Asslaber D, Caldana C, et al. CD18 (ITGB2) expression in chronic lymphocytic leukaemia is regulated by DNA methylation-dependent and -independent mechanisms. Br J Haematol. 2015;169(2):286-289.

36. Goldin LR, McMaster ML, Rotunno M, et al. Whole exome sequencing in families with CLL detects a variant in Integrin β 2 associated with disease susceptibility. Blood. 2016;128(18):2261-2263.

37. Dobashi A, Togashi Y, Tanaka N, et al. TP53 and OSBPL10 alterations in diffuse large B-cell lymphoma: prognostic markers identified via exome analysis of cases with extreme prognosis. Oncotarget. 2018;9(28):19555-19568.

38. Schweighofer CD, Coombes KR, Majewski T, et al. Genomic Variation by Whole-Genome SNP Mapping Arrays Predicts Time-to-Event Outcome in Patients with Chronic Lymphocytic Leukemia: A Comparison of CLL and HapMap Genotypes. J Mol Diagn. 2013;15(2):196-209.

39. Grau M, López C, Navarro A, et al. Unraveling the genetics of transformed splenic marginal zone lymphoma. Blood Adv. 2023;7(14):3695-3709.

40. Oppezzo P, Vuillier F, Vasconcelos Y, et al. Chronic lymphocytic leukemia B cells expressing AID display dissociation between class switch recombination and somatic hypermutation. Blood. 2003;101(10):4029-4032.

41. Roco JA, Mesin L, Binder SC, et al. Class-Switch Recombination Occurs Infrequently in Germinal Centers. Immunity. 2019;51(2):337-350.

42. Bastidas-Mora G, Beà S, Navarro A, et al. Clinico-biological features and outcome of patients with splenic marginal zone lymphoma with histological transformation. Br J Haematol. 2022;196(1):146-155.

43. Otto C, Scholtysik R, Schmitz R, et al. Novel IGH and MYC Translocation Partners in Diffuse Large B-Cell Lymphomas. Genes Chromosomes Cancer. 2016;55(12):932-943.

44. Almire C, Bertrand P, Ruminy P, et al. PVRL2 is translocated to the TRA@ locus in t(14;19)(q11;q13)-positive peripheral T-cell lymphomas. Genes Chromosomes Cancer. 2007;46(11):1011-1018.

45. Chapiro E, Radford-Weiss I, Bastard C, et al. The most frequent t(14;19)(q32;q13)-positive B-cell malignancy corresponds to an aggressive subgroup of atypical chronic lymphocytic leukemia. Leukemia. 2008;22(11):2123-2127.

46. Salido M, Baró C, Oscier D, et al. Cytogenetic aberrations and their prognostic value in a series of 330 splenic marginal zone B-cell lymphomas: a multicenter study of the Splenic B-Cell Lymphoma Group. Blood. 2010;116(9):1479-1488.

47. Busschots AM, Mecucci C, Stul M, et al. Translocation (14;19)(q32;q13.1) in a Young Patient who Developed a Large Cell Lymphoma after an Initial Diagnosis of CLL. Leuk Lymphoma. 1991;5(4):281-286.

48. Michaux L, Dierlamm J, Wlodarska I, Bours V, Van Den Berghe H, Hagemeijer A. t(14;19)/BCL3 rearrangements in lymphoproliferative disorders: A review of 23 cases. Cancer Genet Cytogenet. 1997;94(1):36-43.

49. Fang H, Reichard KK, Rabe KG, et al. IGH translocations in chronic lymphocytic leukemia: Clinicopathologic features and clinical outcomes. Am J Hematol. 2019;94(3):338-345.

50. Rossi D, Deambrogi C, Monti S, et al. BCL3 translocation in CLL with typical phenotype: assessment of frequency, association with cytogenetic subgroups, and prognostic significance. Br J Haematol. 2010;150(6):702-704.

Table 1: Clinical and pathological features of 13 patients with *BCL3*-R

	Total (n=13)	Upstream *BCL3*-R (n=9)	Downstream *BCL3*-R (n=4)	p-value
Clinical data at diagnosis				
Median age (years)	69 (50-81)	69 (50-78)	64 (53-81)	1
Male	7 (54%)	5 (56%)	2 (50%)	1
ALC ($x10^9$/L)	8.1 (0.9-161.8)	9 (2.9-161.8)	1 (0.9-2.9)	0.01
Lymphadenopathy	3 (25%)	3 (38%)	0	0.49
Splenomegaly	4 (33%)	1 (13%)	3 (75%)	0.07
B symptoms	3 (25%)	1 (13%)	2 (50%)	0.24
Clinical data at follow-up				
Need for treatment	9 (69%)	5 (56%)	4 (100%)	0.23
Median time to first treatment (years)	2.6	4.6	0.9	0.7
Large B-cell lymphoma transformation	2 (15%)	0	2 (50%)	0.08
Median survival time (years)	10.5	11.1	10.5	0.5
Genetics				
Trisomy 12	7 (54%)	7 (78%)	0	0.04
Complex karyotype	8 (61%)	6 (67%)	2 (67%)	1
Unmutated IGHV status	8 (61%)	8 (89%)	0	0.007
Deletion 11q	2 (15%)	1 (11%)	1 (25%)	1
TP53 mutation	0	0	0	
Deletion 17p	0	0	0	
Phenotype				
Flow cytometry				
Typical for CLL*	3 (23%)	3 (33%)	0	
Bright sIg	8 (61%)	4 (44%)	4 (100%)	
Bright B-cell markers	9 (69%)	5 (56%)	4 (100%)	
CD5 +	9 (69%)	9 (100%)	0	
CD43 +	7 (58%)	6 (67%)	1 (33%)	
CD23 +	6 (50%)	6 (67%)	0	
CD200 +	9 (75%)	8 (89%)	1 (33%)	
Immunohistochemistry				
CD5 +	3/7	3/4	0/3	
CD23 +	4/7	2/4 strong, 2/4 weak	0/3	
LEF1 +	0/7	0/4	0/3	
BCL3 +	3/6	3/3	0/3	

Quantitative parameters are expressed as median (range).

Abbreviations: ALC, Absolute lymphocyte count; sIg, surface immunoglobulins.

*Typical CLL immunophenotype CD19+ with dim expression of CD20, CD22, CD79b, CD5+, CD23+, CD43+, CD200+, and FMC7 −.

Figures

Figure 1. Characterization of the breakpoints derived from the t(14;19) at base-pair resolution. **A,** Representation of the breakpoints on chromosome 19 (chr19) and chromosome 14 (chr14) for each patient (red vertical line). Unmutated IGHV (U-IGHV) and mutated IGHV (M-IGHV) tumors are represented in gray and black labels, respectively. Tumors are classified based on the breakpoint on chr19: further upstream *BCL3* (pale blue), upstream *BCL3* (blue), and downstream *BCL3* (orange). **B,** Schema of the most recurrent translocation pattern observed in the upstream *BCL3* subgroup with IGH and its corresponding CSR located upstream *BCL3*, suggesting a constitutive upregulation of *BCL3*. **C,** Depiction of five patients, one with the translocation further upstream *BCL3* and four with the translocation downstream *BCL3*. In the further upstream tumor (3698), the t(14;19) truncates *CEACAM16* and, similar to upstream *BCL3* translocations, IGH is located 5' of *BCL3* suggesting a constitutive upregulation of this gene. In the downstream tumors, the t(14;19) affects three different genes (*CBLC*, *BCAM*, *NECTIN2*) located downstream of *BCL3*. The resulting derivatives of the t(14;19) suggest that *BCL3* is not placed under the regulation of the enhancers of the IGH and, therefore, its expression remains unchanged.

Figure 2. BCL3 expression in upstream and downstream *BCL3*-R tumors. **A,** RNA-seq data shows that BCL3 is upregulated in the upstream *BCL3* tumors, except tumor 3646 carrying a subclonal t(14;19), compared to U-CLL. Contrarily, the downstream *BCL3* tumor (3676) upregulated NECTIN2 while showed lower BCL3 expression than any of the upstream and CLL tumors. **B,** Immunohistochemistry images (400x) displaying a positive BCL3 expression in the further upstream tumor and in a representative upstream tumor, but not in a representative downstream tumor.

Figure 3. Mutations and structural alterations in B-cell neoplasms with t(14;19) and *BCL3*-R identified by WGS. **A,** Oncoprint representation of driver gene mutations frequently observed in CLL (red) or in other B-cell lymphomas (blue-green). Total number of mutations are not reported in sample 3696 and 624 due to the lack of germline DNA and in sample 3649 due to its low tumor cell content (20%) (Supplementary Methods, Supplementary Table S1). **B,** Comparison of the number of mutations, CNA and SV between upstream *BCL3*-R and downstream *BCL3*-R tumors. **C,** Copy number profile of *BCL3*-R tumors. Tumors are shown in rows and chromosomes in columns. The IGHV mutational status, breakpoint location on chromosome 19, and number of CNA are shown on the right. Sample 3649 had an estimated tumor cell content of 20% that allowed the detection of driver somatic mutations and the *BCL3*-R but was not sufficient for a proper analysis of CNA (Supplementary Methods, Supplementary Table S1).

Figure 4. Gene expression profile of upstream *BCL3*-R tumors. **A,** PCA of RNA-seq data of 6 upstream *BCL3*-R tumors, one downstream *BCL3*-R tumor, and 9 CLL (first component is shown against second, third, fourth, and fifth components). **B,** Heatmap of the differential gene expression analysis between 5 upstream *BCL3*-R tumors and 4 U-CLL, also compared to one tumor with downstream *BCL3*-R tumor and CLL without t(14;19). Tumor 3646 was excluded from the analysis due to its subclonal B*CL3*-R. Hallmark CLL genes differentially expressed between *BCL3*-R tumors and CLL are flagged. **C,** Expression of CLL hallmark genes in the upstream *BCL3*-R tumors compared to U-CLL. Q-values are from the DEA. **D,** Venn diagram showing the overlap of the differentially expressed genes among upstream *BCL3*-R vs U-CLL and U-CLL with vs without trisomy 12. Hallmark CLL genes are highlighted.

Figure 5. DNA methylation profile of upstream *BCL3*-R tumors. **A,** PCA of DNA methylation data of 10 B-cell neoplasms with *BCL3*-R, 85 CLL, and 7 normal B-cells subsets (first and second components are shown). The shape corresponds to the tumor types while the color represents the proliferative history (epiCMIT score). NBC: naive B cell, GC: germinal center B cell, MBC memory B cell, PC: plasma cell, n-CLL: naive-like CLL, i-CLL: intermediate CLL, m-CLL: memory-like CLL. **B,** Comparison of the epiCMIT score between upstream *BCL3*-R tumors and n-CLL from cohorts C1 and C2, respectively. The upstream *BCL3*-R subgroup of tumors does not include the tumor 3646 carrying a subclonal t(14;19). **C,** Heatmap of the differential methylated CpGs between 7 upstream *BCL3*-R tumors and 85 CLL. The chromatin state of each CpG is shown on the right. Differentially methylated CpGs mapping to differentially expressed CLL genes of interest are labeled. **D,** TP63 expression in the upstream *BCL3*-R subgroup compared to CLL.

Figure 6. Images of representative upstream and downstream *BCL3*-R tumors. **A,** Cells in peripheral blood smears from representative tumors. Both tumors show features such as nuclear irregularities and lobulation, non-clumped chromatin, central nucleoli, ample cytoplasm, or villi, which are atypical for conventional CLL. 1000x oil immersion, light microscope and camera, Leishman stain. **B-C,** Histology (H&E staining) and immunohistochemistry images were obtained from scanned slides (Ventana DP200 scanner, Roche Diagnostics). The upstream *BCL3*-R tumor had a diffuse growth pattern, resembling CLL, but without proliferation centers (100x). At high power (600x), the cells were small, with scarce cytoplasm, distinct irregular nuclei, and central nucleoli. Larger scattered cells were observed. The immunophenotype is atypical for a CLL tumor (CD5-, CD23+ weak, and CD43+ weak), and the cells are LEF1 negative. CD5 was negative in the lymph nodes by immunohistochemistry but positive in the peripheral blood according to flow cytometry. The downstream *BCL3*-R tumor has a perifollicular growth pattern (100x), leaving residual germinal

centers (400x), with a residual follicular dendritic network on CD21 and germinal center cells on BCL6, resembling marginal zone lymphoma. This tumor has a non-specific B-cell phenotype and plasma cell differentiation with kappa light-chain restriction.

Figure 7. Custom FISH assay to map the breakpoints of the *BCL3*-R and images of a representative tumor from the validation cohort. **A,** Schematic representation of the custom design of *BCL3* break-apart FISH probe. *BCL3* gene and *BCL3* FISH probe are annotated based on GRCh37/hg19 assembly. **B,** Interphase nucleus of tumor 3783 (left panel) and 3676 (right panel). Tumor 3783 shows a positive signal constellation indicating a break upstream of *BCL3* since the BAC-clone RP11-927F16 is split from CTD2608C5 and RP11-423N20. Tumor 3676 displays a positive signal constellation suggesting a break downstream of *BCL3* with the BAC-clone RP11-423N20 split from CTD2608C5 and RP11-927F16. **C,** Histology (H&E stain) and immunohistochemistry images of tumor 1 from the validation cohort. Low power magnification (50x) of lymph node shows clear proliferation centers. CD20 shows diffuse positivity (100x). CD23 is only partially and faintly expressed in proliferation centers (100x). CD3 highlights few admixed T cells (100x). CD5 shows few admixed T cells (strong staining intensity) and low expression in tumor cells in the lymph node (100x). LEF1 shows expression in T cells and few cells in proliferation centers but mainly negative in tumor cells (100x).

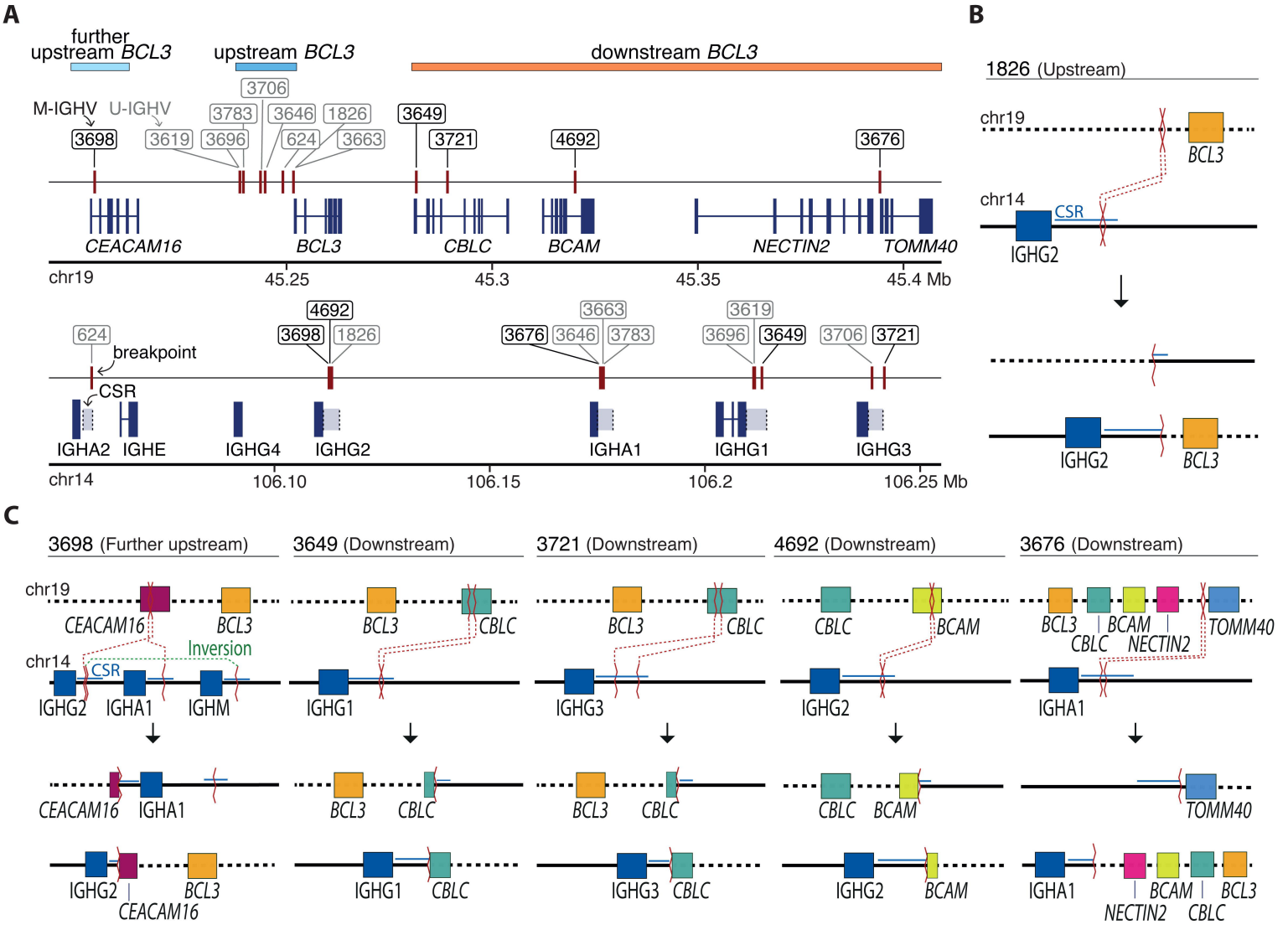

A
further upstream BCL3
upstream BCL3
downstream BCL3
M-IGHV
U-IGHV
3706
3783
3646
1826
3649
3698
3619
3696
624
3663
3721
4692
3676
CEACAM16
BCL3
CBLC
BCAM
NECTIN2
TOMM40
chr19
45.25
45.3
45.35
45.4 Mb
4692
3663
3619
624
3698
1826
3676
3646
3783
3696
3649
3706
3721
breakpoint
CSR
IGHA2
IGHE
IGHG4
IGHG2
IGHA1
IGHG1
IGHG3
chr14
106.10
106.15
106.2
106.25 Mb
B
1826 (Upstream)
chr19
BCL3
chr14
CSR
IGHG2
IGHG2
BCL3
C
3698 (Further upstream)
3649 (Downstream)
3721 (Downstream)
4692 (Downstream)
3676 (Downstream)
chr19
CEACAM16
BCL3
BCL3
CBLC
BCL3
CBLC
CBLC
BCAM
BCL3
BCAM
CBLC
NECTIN2
TOMM40
Inversion
chr14
CSR
IGHG2
IGHA1
IGHM
IGHG1
IGHG3
IGHG2
IGHA1
CEACAM16
IGHA1
BCL3
CBLC
BCL3
CBLC
CBLC
BCAM
TOMM40
IGHG2
CEACAM16
BCL3
IGHG1
CBLC
IGHG3
CBLC
IGHG2
BCAM
IGHA1
NECTIN2
BCAM
CBLC
BCL3

Figure 1

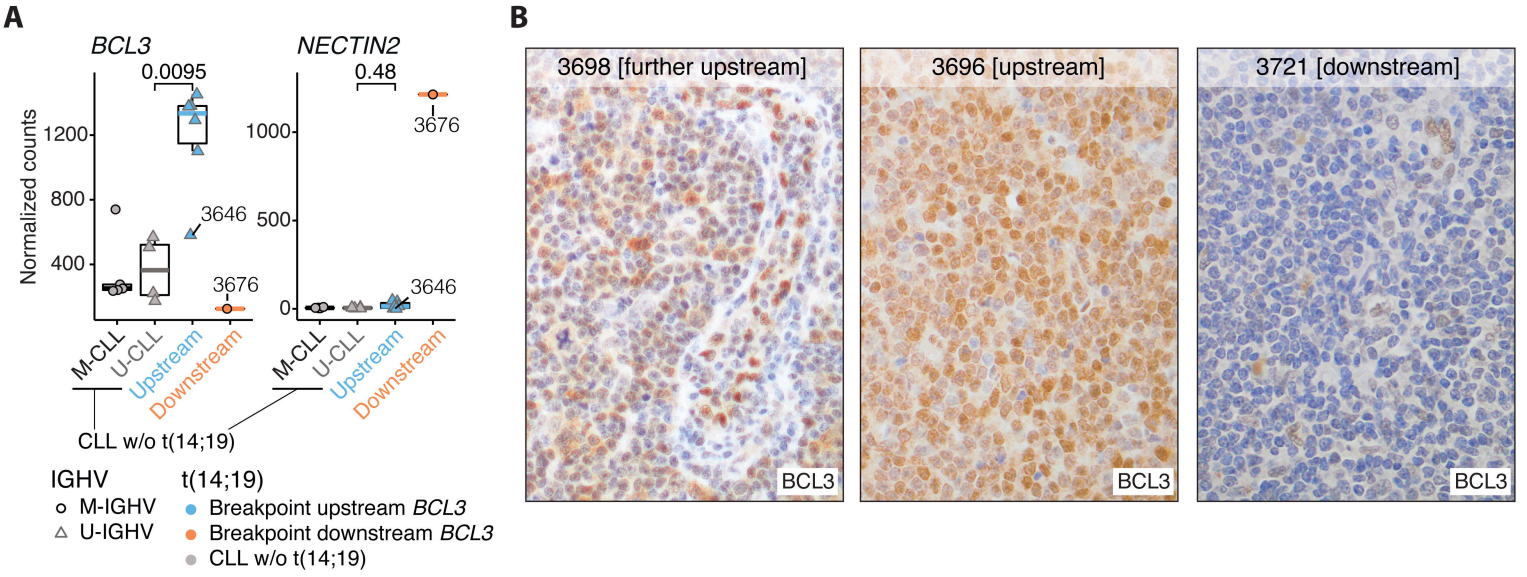

A
BCL3
NECTIN2
0.0095
0.48
Normalized counts
1200
800
400
1000
500
0
3646
3676
3646
3676
M-CLL
U-CLL
Upstream
Downstream
M-CLL
U-CLL
Upstream
Downstream
CLL w/o t(14;19)
IGHV
M-IGHV
U-IGHV
t(14;19)
Breakpoint upstream BCL3
Breakpoint downstream BCL3
CLL w/o t(14;19)
B
3698 [further upstream]
3696 [upstream]
3721 [downstream]
BCL3
BCL3
BCL3

Figure 2

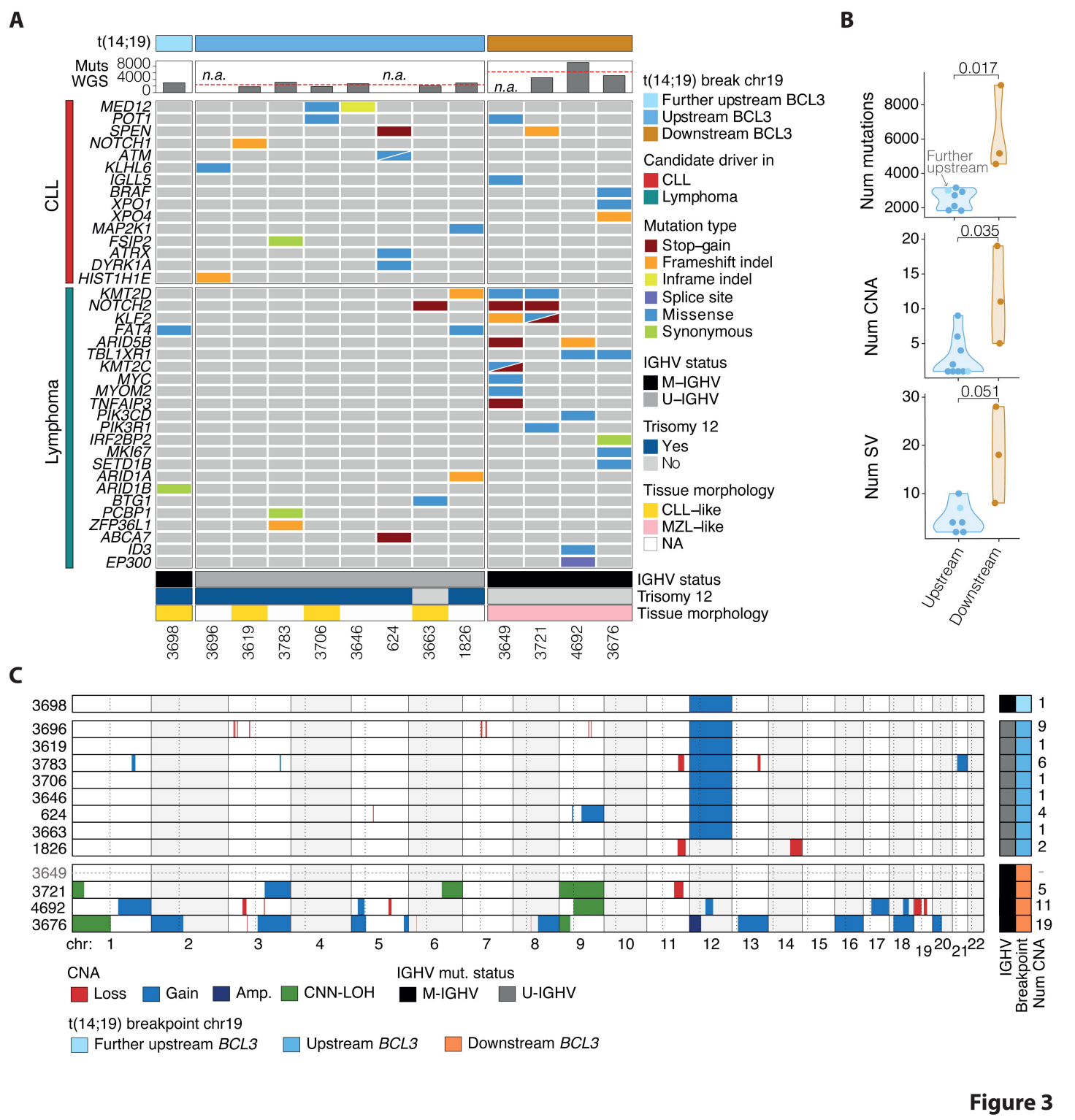

A
t(14;19)
Muts WGS
8000
4000
0
n.a.
n.a.
n.a.
CLL
Lymphoma
MED12
POT1
SPEN
NOTCH1
ATM
KLHL6
IGLL5
BRAF
XPO1
XPO4
MAP2K1
FSIP2
ATRX
DYRK1A
HIST1H1E
KMT2D
NOTCH2
KLF2
FAT4
ARID5B
TBL1XR1
KMT2C
MYC
MYOM2
TNFAIP3
PIK3CD
PIK3R1
IRF2BP2
MKI67
SETD1B
ARID1A
ARID1B
BTG1
PCBP1
ZFP36L1
ABCA7
ID3
EP300
3698
3696
3619
3783
3706
3646
624
3663
1826
3649
3721
4692
3676
t(14;19) break chr19
Further upstream BCL3
Upstream BCL3
Downstream BCL3
Candidate driver in
CLL
Lymphoma
Mutation type
Stop–gain
Frameshift indel
Inframe indel
Splice site
Missense
Synonymous
IGHV status
M–IGHV
U–IGHV
Trisomy 12
Yes
No
Tissue morphology
CLL–like
MZL–like
NA
IGHV status
Trisomy 12
Tissue morphology
B
Num mutations
0.017
8000
6000
4000
2000
Further upstream
Num CNA
0.035
20
15
10
5
Num SV
0.051
30
20
10
Upstream
Downstream
C
3698
3696
3619
3783
3706
3646
624
3663
1826
3649
3721
4692
3676
1
9
1
6
1
1
4
1
2
-
5
11
19
chr: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
IGHV
Breakpoint
Num CNA
CNA
Loss
Gain
Amp.
CNN-LOH
IGHV mut. status
M-IGHV
U-IGHV
t(14;19) breakpoint chr19
Further upstream BCL3
Upstream BCL3
Downstream BCL3

Figure 3

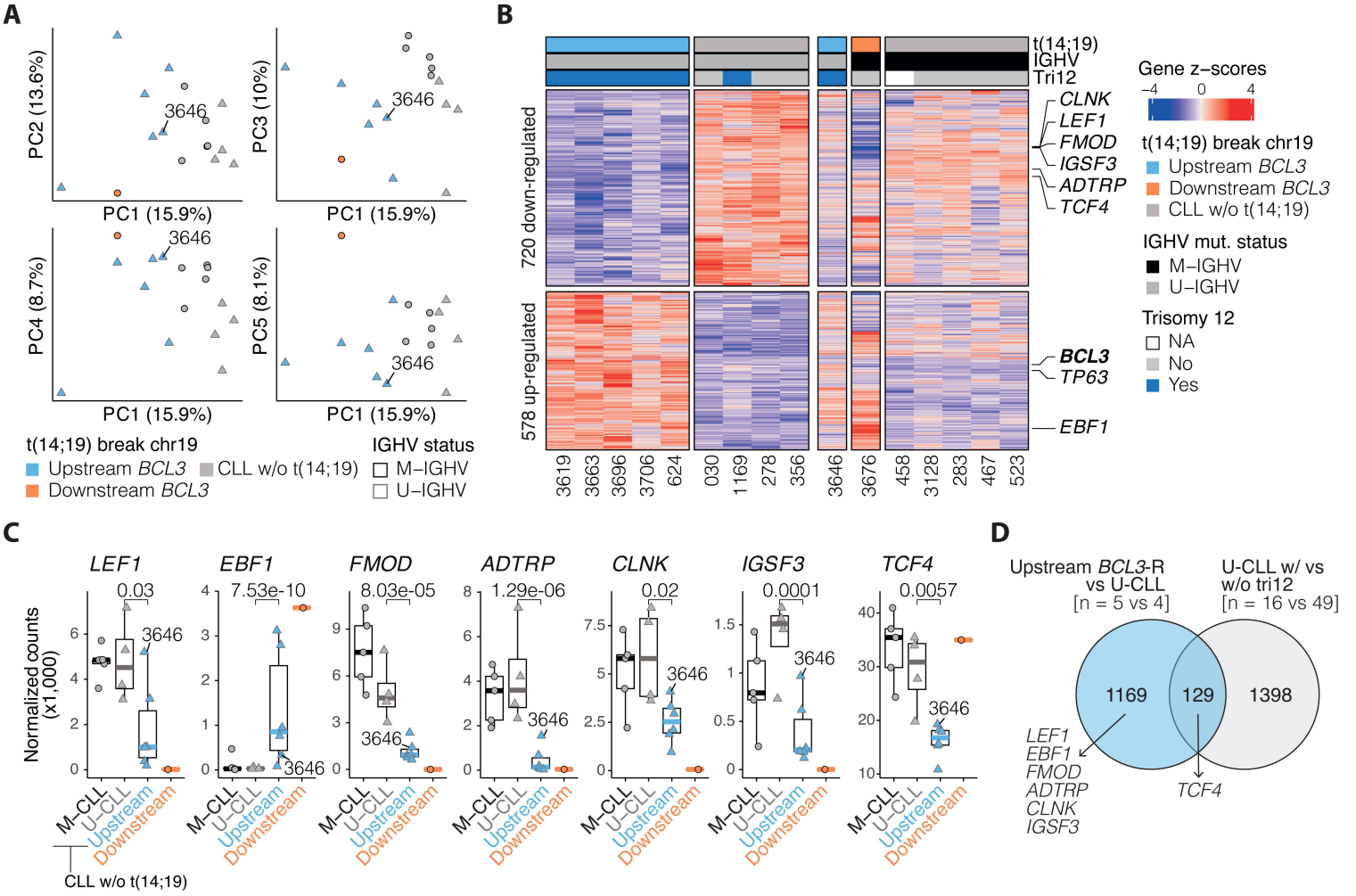

A
PC2 (13.6%)
PC1 (15.9%)
3646
PC3 (10%)
PC4 (8.7%)
PC5 (8.1%)
t(14;19) break chr19
Upstream BCL3
Downstream BCL3
CLL w/o t(14;19)
IGHV status
M–IGHV
U–IGHV
B
t(14;19)
IGHV
Tri12
Gene z–scores
−4
0
4
720 down-regulated
578 up-regulated
CLNK
LEF1
FMOD
IGSF3
ADTRP
TCF4
BCL3
TP63
EBF1
t(14;19) break chr19
IGHV mut. status
Trisomy 12
NA
No
Yes
3619
3663
3696
3706
624
030
1169
278
356
3646
3676
458
3128
283
467
523
C
LEF1
EBF1
FMOD
ADTRP
CLNK
IGSF3
TCF4
0.03
7.53e-10
8.03e-05
1.29e-06
0.02
0.0001
0.0057
Normalized counts (x1,000)
M–CLL
U–CLL
Upstream
Downstream
CLL w/o t(14;19)
D
Upstream BCL3-R vs U-CLL [n = 5 vs 4]
U-CLL w/ vs w/o tri12 [n = 16 vs 49]
1169
129
1398
LEF1
EBF1
FMOD
ADTRP
CLNK
IGSF3
TCF4

Figure 4

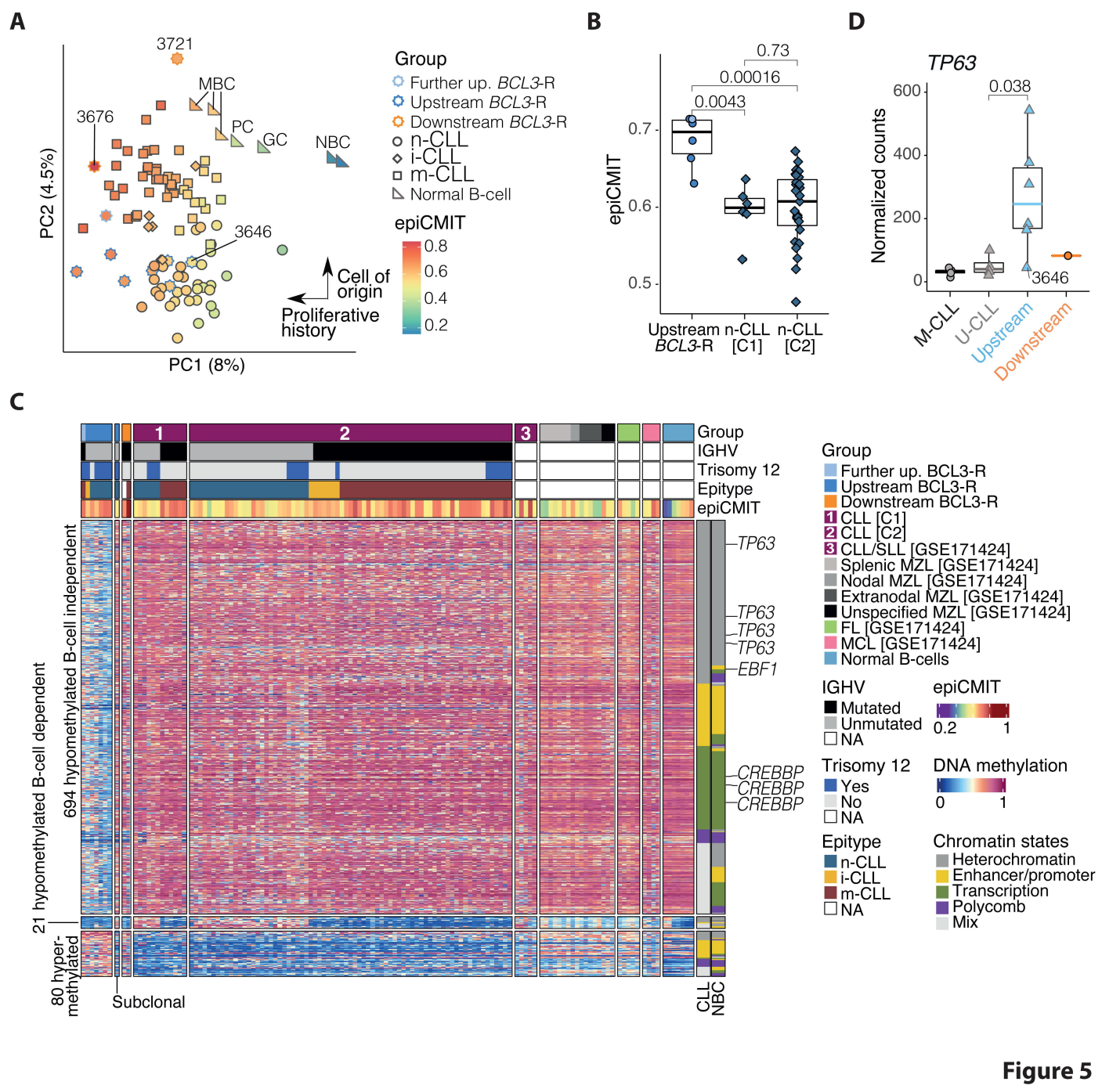

A
3721
MBC
3676
PC
GC
NBC
3646
PC2 (4.5%)
PC1 (8%)
Cell of origin
Proliferative history
Group
Further up. BCL3-R
Upstream BCL3-R
Downstream BCL3-R
n-CLL
i-CLL
m-CLL
Normal B-cell
epiCMIT
0.8
0.6
0.4
0.2
B
0.73
0.00016
0.0043
epiCMIT
0.7
0.6
0.5
Upstream BCL3-R
n-CLL [C1]
n-CLL [C2]
D
TP63
0.038
Normalized counts
600
400
200
0
3646
M-CLL
U-CLL
Upstream
Downstream
C
1
2
3
Group
IGHV
Trisomy 12
Epitype
epiCMIT
694 hypomethylated B-cell independent
21 hypomethylated B-cell dependent
80 hyper-methylated
Subclonal
CLL
NBC
TP63
TP63
TP63
TP63
EBF1
CREBBP
CREBBP
CREBBP
Group
Further up. BCL3-R
Upstream BCL3-R
Downstream BCL3-R
1 CLL [C1]
2 CLL [C2]
3 CLL/SLL [GSE171424]
Splenic MZL [GSE171424]
Nodal MZL [GSE171424]
Extranodal MZL [GSE171424]
Unspecified MZL [GSE171424]
FL [GSE171424]
MCL [GSE171424]
Normal B-cells
IGHV
Mutated
Unmutated
NA
epiCMIT
0.2
1
Trisomy 12
Yes
No
NA
DNA methylation
0
1
Epitype
n-CLL
i-CLL
m-CLL
NA
Chromatin states
Heterochromatin
Enhancer/promoter
Transcription
Polycomb
Mix

Figure 5

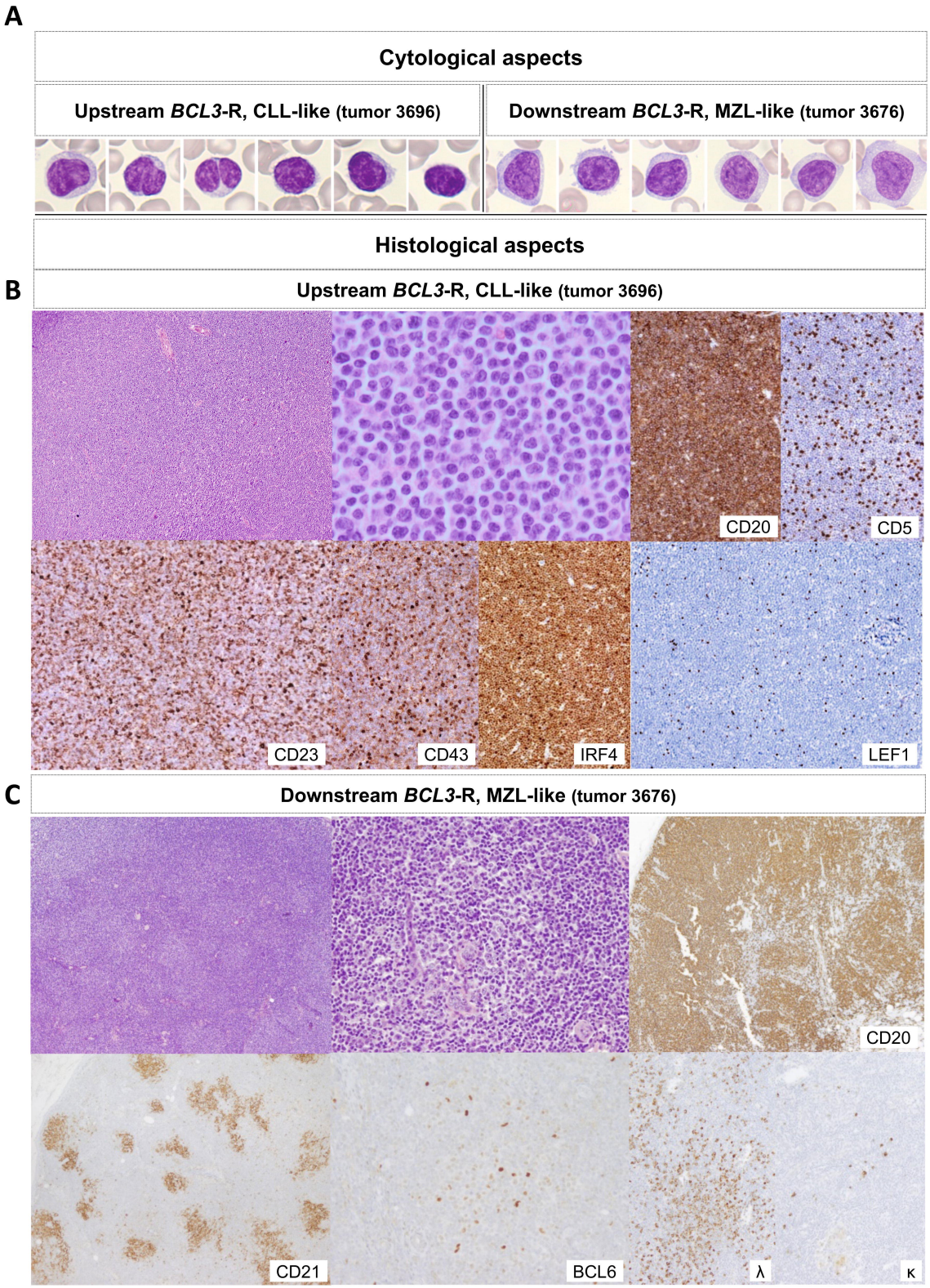

A
Cytological aspects
Upstream BCL3-R, CLL-like (tumor 3696)
Downstream BCL3-R, MZL-like (tumor 3676)
Histological aspects
B
Upstream BCL3-R, CLL-like (tumor 3696)
CD20
CD5
CD23
CD43
IRF4
LEF1
C
Downstream BCL3-R, MZL-like (tumor 3676)
CD20
CD21
BCL6
λ
κ

Figure 6

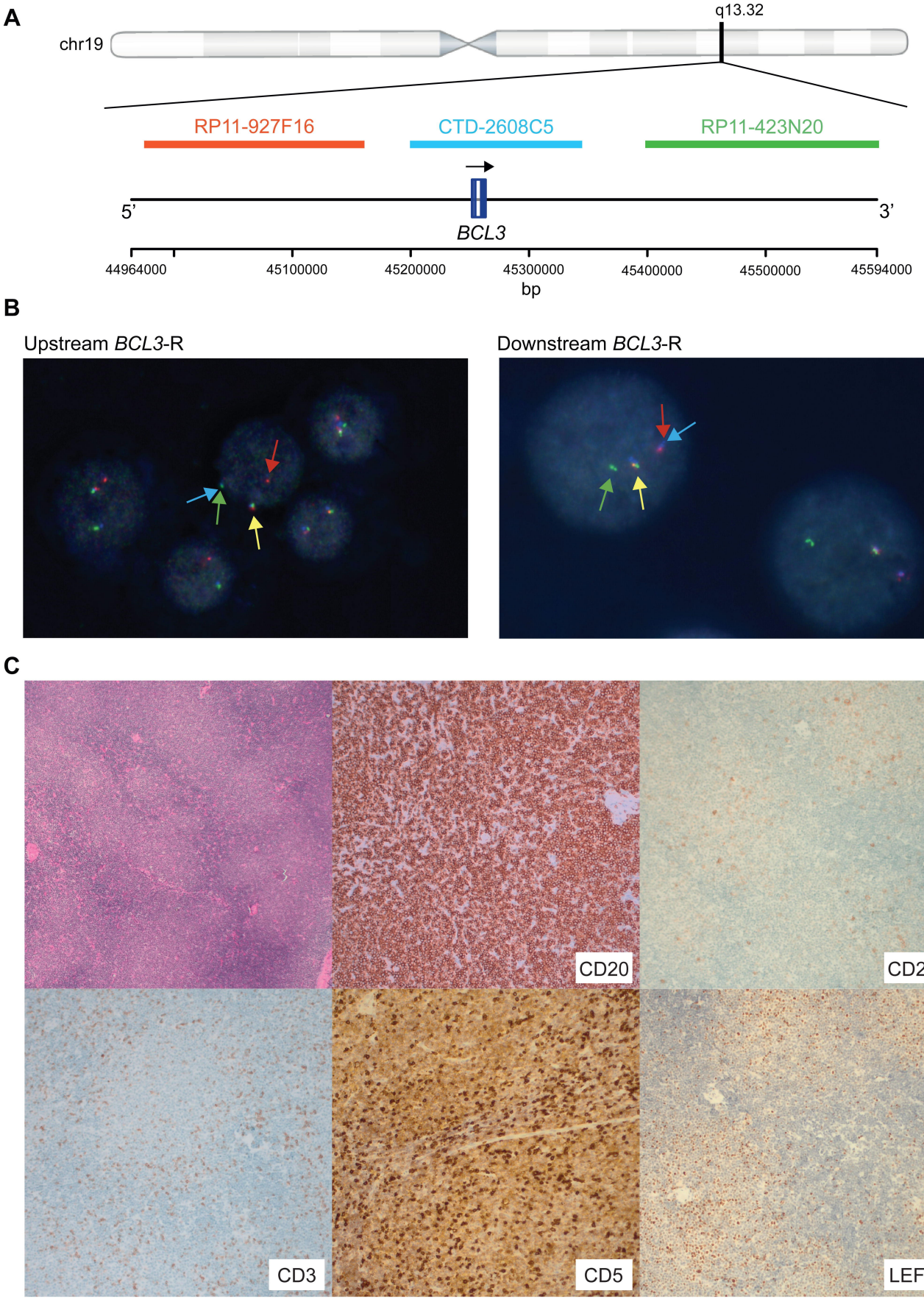

A
chr19
q13.32
RP11-927F16
CTD-2608C5
RP11-423N20
5'
3'
BCL3
44964000
45100000
45200000
45300000
45400000
45500000
45594000
bp
B
Upstream BCL3-R
Downstream BCL3-R
C
CD20
CD23
CD3
CD5
LEF1

Figure 7

Supplementary Materials

BCL3-rearrangements in B-cell lymphoid neoplasms occur in two breakpoint clusters associated with different diseases

Anna Carbó-Meix, Francesca Guijarro, Luojun Wang, Marta Grau, Romina Royo, Gerard Frigola, Heribert Playa-Albinyana, Marco M. Bühler, Guillem Clot, Martí Duran-Ferrer, Junyan Lu, Isabel Granada, Maria-Joao Baptista, José-Tomás Navarro, Blanca Espinet, Anna Puiggros, Gustavo Tapia, Laura Bandiera, Gabriella De Canal, Emanuela Bonoldi, Fina Climent, Inmaculada Ribera-Cortada, Mariana Fernández-Caballero, Esmeralda de la Banda, Janilson do Nascimento, Alberto Pineda, Dolors Vela, María Rozman, Marta Aymerich, Charlotte Syrykh, Pierre Brousset, Miguel Perera, Lucrecia Yánez, Jesús Xavier Ortin, Esperanza Tuset, Thorsten Zenz, James R. Cook, Steven H. Swerdlow, José I. Martín-Subero, Dolors Colomer, Estella Matutes, Sílvia Beà, Dolors Costa, Ferran Nadeu, and Elías Campo

SUPPLEMENTARY METHODS

Whole-genome sequencing analyses

Whole-genome sequencing (WGS) of paired tumor/normal samples (n=10) or tumor samples (n=3) from 13 B-cell neoplasms with the *BCL3*-rearrangement (*BCL3*-R) was performed using the TruSeq DNA PCR Free or the TruSeq DNA nano library preparation protocol based on material availability and sequenced in a NovaSeq6000 (2x150 bp). Samples were sequenced at a mean coverage of 30x, except three tumor samples [3646 (subclonal *BCL3*-R), 3649 (tumor cell content of 20%), and 4692 (tumor cell content of 40%)] that were sequenced at a mean coverage of 70-80x (Supplementary Table S1).

Raw reads were mapped to the human reference genome (GRCh37) using the BWA-mem algorithm (version 0.7.15).[1] BAM files were generated, sorted, indexed and optical or PCR duplicates flagged using biobambam2 (version 2.0.65).[2] Quality control and coverage metrics were extracted using FastQC (version 0.11.5, https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Picard (version 2.10.2, https://broadinstitute.github.io/picard/).

Tumor vs normal variant calling was performed as previously described.[3] Briefly, somatic single nucleotide variants (SNV) were analyzed using CaVEMan (cgpCaVEManWrapper, version 1.12.0),[4] Mutect2 (GATK, version 4.0.2.0),[5] and MuSE (version 1.0).[6] Caller-specific filters to remove low quality variants were identified by CaVEMan and Mutect2. Variants detected by CaVEMan with more than half of the mutant reads clipped (CLPM>0) and with supporting reads with a median alignment score (ASMD)<140 were excluded. Variants called by Mutect2 with MMQ<60 were eliminated. Finally, mutations detected by at least two algorithms were considered. Short insertions/deletions (indels) were called by Pindel (cgpPindel, version 2.2.3),[7] Platypus (version 0.8.1),[8] SvABA (version 7.0.2),[9] and Mutect2.[5] As performed for SNVs, caller-specific filters were applied: variants with mapping quality MMQ<60, MQ<60, and MAPQ<60 for Mutect2, Platypus, and SvABA, respectively, were removed. Only indels identified by at least two algorithms were retained for downstream analyses. Mutations identified were annotated using snpEff/snpSift (version 4.3t).[10,11] Copy number alterations (CNA) were called using Battenberg (cgpBattenberg, version 3.2.2)[12] and ASCAT (ascatNgs, version 4.1.0).[13] CNA within any of the immunoglobulin loci were not considered. Genome-wide structural variants (SV) were extracted using BRASS (version 6.0.5),[14] SvABA,[9] and DELLY2 (version 0.8.1).[15] Variants detected by BRASS with MAPQ<90 and those with MAPQ<60 for SvABA or DELLY2 were filtered out. Finally, SV

identified by at least two programs were kept. All SV were visually inspected using the Integrative Genomic Viewer (IGV).[16]

Tumor-only variant calling within the genomic regions of previously described candidate driver genes in chronic lymphocytic leukemia (CLL) and other lymphomas[3] was performed for 3 tumors without germline DNA available using an updated version of our tumor-only pipeline.[17] Briefly, mini-BAM files with the reads mapping to the driver genes were obtained using Picard tools and variant calling was performed using VarScan2 (version 2.4.3),[18] Mutect2,[5] VarDictJava (version 1.4),[19] LoFreq (version 2.1.3.1),[20] outLyzer (version 1.0),[21] and freebayes (version 1.1.0; https://github.com/freebayes/freebayes). Variants identified were annotated using snpEff/snpSift (version 4.3t).[10,11] Only variants identified as PASS by at least two algorithms were considered. Variants reported in the 1000 Genome Project, ExAC and/or gnomAD with a population frequency >1% were considered as potential polymorphisms and removed from downstream analyses. Similarly, variants reported as germline in our International Cancer Genome Consortium (ICGC)-CLL database were removed.[22] Tumor-only CNA were extracted using Control-FREEC (version 11.5) with default parameters.[23]

Note on sample 3649: Due to the low tumor cell content of sample 3649 (20%), we did not report CNA, SV and genome-wide mutations. Driver gene mutations were analyzed using the tumor-only variant calling pipeline.

Driver mutations and mutational signature analysis

Driver mutations were studied considering a list of 247 recurrently mutated genes in B-cell neoplasms, including CLL, SMZL, and DLBCL (Supplementary Table S2).[22,24–27]

Whole-genome analysis of single base substitutions (SBS) was performed in the 10 normal-tumor samples carrying *BCL3*-rearrangements using the MutationalPatterns package (version 3.8.1) in R.[28] For the analysis, signatures previously described in CLL and lymphomas were considered: SBS1, SBS5, SBS8, SBS9, and SBS18.[3,29] The cosine similarity between the original and reconstructed mutational profiles was measured to assess the robustness of the analysis.

Mutations found in the immunoglobulin heavy chains (IGH) constant genes and class switch regions (CSR) were visually examined using the Integrative Genome Viewer (version 2.16.1) to assess if they occurred in putative activation-induced cytidine deaminase (AID) target motifs, that is WRCY and RGYW.[30] IGH constant gene-CSR coordinates were defined as the region between the starting point of

the IGH constant gene and the end point of the CSR using the wgEncodeGencodeBasicV19 database for IGH constant genes and the previously defined consensus sequences for CSR.[31]

RNA-seq analyses

Total tumor RNA for RNA-seq could be obtained from 7 tumors with *BCL3*-R. We also extracted total RNA from 9 CLL lacking t(14;19) and *BCL3*-R, 4 with unmutated IGHV (U-IGHV) and 5 with mutated IGHV. Total RNA samples were quantified by Qubit RNA BR Assay kit (Thermo Fisher Scientific) and the RNA integrity was estimated by using RNA 6000 Nano Bioanalyzer 2100 Assay (Agilent) prior to library preparation. Stranded total RNA-seq libraries were prepared Illumina's recommendations. To further compare the expression data of tumors with *BCL3*-R and CLL we re-analyzed the RNA-seq data from 65 CLL [all U-IGHV] from a previous publication.[22] These tumors were negative for *BCL3*-R and carried less than 4 CNA to reduce the possible effects of CNA on gene expression.

The bioinformatic analysis was performed as previously described.[3] Briefly, sequencing reads were trimmed using trimmomatic (version 0.38)[32] and ribosomic RNA reads were filtered out using SortMeRNA (version 2.1b).[33] Gene-level counts (GRCh38.p13; Ensembl release 100) were calculated using kallisto (version 0.46.1)[34] and tximport (version 1.6.0).[35] A principal component analysis (PCA) was conducted to study the clustering of the samples. Differential expression analysis was conducted using DESeq2 (version 1.30.1).[36] Log fold change (FC) shrinkage was subsequently applied with the apeglm method.[37] Genes were considered as differentially expressed if $q<0.05$ and absolute($\log_2$FC)>0 in the upstream *BCL3*-R vs U-CLL comparison and if $q<0.05$ and absolute($\log_2$FC)>0.1 in the U-CLL with trisomy 12 vs U-CLL without trisomy 12 comparison. Finally, the variance stabilizing transformation (VST) was applied on the matrices of the normalized counts and used these transformed matrices for dimensionality reduction analyses. Gene set enrichment analysis (GSEA) was performed with the fgsea R package (version 1.20.0)[38] using the C2 and H collections from the MSigDB gene sets (version 7.4).

DNA methylation

The DNA methylation profile of 10 tumors with *BCL3*-R were examined using EPIC methylation arrays following manufacturer's recommendations. Similar data from 85 CLL negative for *BCL3*-R were obtained from two previous publications: cohort 1 (C1), which included 12 CLL from our institution,[3] and cohort 2 (C2), which consisted of 73 CLL from University Hospital Heidelberg.[39,40] Additionally, the DNA methylation profile of 5 CLL/small lymphocytic lymphoma (SLL), 21 marginal zone lymphoma (MZL) (7 splenic, 3 nodal, 6 extranodal, 5 unspecified), 5 follicular lymphomas (FL), and 4 mantle cell lymphomas (MCL), were obtained from the Gene Expression Omnibus

(https://www.ncbi.nlm.nih.gov/geo, accession number GSE171424).[41] In contrast to C1 and C2 DNA samples, which were obtained from frozen tissue, GSE171424 data were obtained from DNA from formalin fixed paraffin embedded tissue. Due to this difference, GSE171424 tumors were only used to visually compare their DNA methylation profile with that of the 10 tumors with the *BCL3*-R and the 85 CLL. The results were also visually compared to the methylation profile of 7 normal B-cells (NBC; 2 naive, 1 germinal center, 3 memory and 1 plasma cell, DNA from from fresh/frozen samples) from a previous publication.[42]

DNA methylation data were analyzed with particular use of the minfi package (version 1.36.0),[43] which was exclusively used for data preprocessing. EPIC arrays of the tumors with the *BCL3*-R, C1 and C2 cohorts and NBC were first combined into one object using the combineArrays function from the minfi package, while GSE171424 arrays were preprocessed separately. Only probes common to both datasets (i.e. *BCL3*-R, C1, C2 and NBC, on one side [A1], and GSE171424, on the other side [A2]) were used for the preprocessing. A total of 865859 probes present in each array were quantile normalized. Consecutively, 30435 CpGs representing SNPs, 2925 non-CpG probes, and 19205 CpGs present in sex chromosomes were excluded. From the remaining 813294 CpGs, 10555 CpGs in A1 and 48265 CpGs in A2 with a detection p-value of ≤1x10e-6 in more than 10% of the samples were removed. From a total of 802739 CpGs in A1 and 765029 CpGs in A2, 764159 CpGs common in both datasets A1 and A2 were kept. After implementing all filtering criteria, 10 B-cell neoplasms with the t(14;19) from our institution, 85 CLL from C1 and C2 cohorts, 7 NBC, and 5 CLL/SLL, 17 MZL (7 splenic, 2 nodal, 5 extranodal, 3 unspecified), 5 FL, and 4 MCL from GSE171424 were retained. In total, 102 samples in A1 and 31 samples in A2 were profiled with DNA methylation values for 764159 CpGs.

C1 and C2 CLL tumors and the 10 tumors with *BCL3*-R were classified using the CLL epitype classifier,[44] which categorizes CLL into 3 epigenetic subtypes (also known as epitypes), named naive-like CLL (n-CLL), intermediate CLL (i-CLL) and memory-like CLL (m-CLL),[45] based on 4 CpGs. A PCA was conducted on the beta values in A1. The epiCMIT score, a mitotic clock composed of 1348 CpGs that reflects the proliferative history of neoplastic B-cells,[44] was compared between upstream *BCL3*-R and n-CLL (1263/1348 (94%) and 1221/1348 (91%) CpGs were available in A1 and A2, respectively). Differential methylation analysis was performed between 7 upstream *BCL3*-R and 85 CLL, adjusting for IGHV, epitype, trisomy 12 and cohort, using the limma package (version 3.46.0).[46] CpGs were considered to be differentially methylated at $q<0.05$ (p-values were corrected with the Benjamini-Hochberg method) and an absolute difference in betas values of at least 0.25. T-distributed stochastic neighbor embedding analysis was carried out on the differentially methylated (DM)CpGs in all samples. DMCpGs were mapped to the genomic location (N_Shelf, N_Shore, Island, S_Shelf, S_Shore, OpenSea)

and to intergenic region or gene region (TSS1500, TSS200, 5'UTR, 1stExon, Body, ExonBnd, 3'UTR) using the getAnnotation function from the minfi package (version 1.36.0). Additionally, DMCpGs were mapped to chromatin states obtained from 7 CLL tumors and 5 normal naive B-cells from a previous publication.[47] Chromatin states were defined as: heterochromatin (H3K9me3_Repressed, Heterochromatin Low_Signal), polycomb (Posied_Promoter, H3K27me3_Repressed), enhancer/promoter (Active_Promoter, Strong_Enhancer1, Weak_Promoter, Weak_Enhancer, Strong_Enhancer), transcription (Transcription_Transition, Weak_Transcription, Transcription_ Elongation) and chromatin mix (when the tumors had several chromatin states. A transcription factor (TF) binding analysis was performed using 100 base-pair (bp) sequences around 715 CpGs showing hypomethylation (50 bp to each side of the CG). These sequences were obtained using the getAnnotation function from the minfi package. A total of 500000 randomly selected sequences were used as background. The frequency of A, T, C, and G in the background sequences was then calculated to account for the biases in the EPIC array. The AME tool from the MEME suite (version 5.5.0)[48] was used for the enrichment analysis of known motifs from the non-redundant vertebrate 2022 JASPAR database[49] using a one-tailed Wilcoxon rank-sum test with the maximum odds score as a sequence scoring method and a 0.05 false discovery rate (FDR) cutoff.

Calcium flux analysis

Calcium flux analysis were performed as previously described.[3] Cryopreserved cells were resuspended on RPMI-1640 medium with 10% FBS, 1% Glutamax and 5% penicillin (10,000 IU ml^{-1})/ streptomycin (10 mg ml^{-1}) (Thermo Fisher) at 10^6 cells ml^{-1}. After 6 h of incubation at 37 °C and 5% CO_2, cells were centrifuged and resuspended on RPMI-1640 with 4 µM Indo-1 AM (Thermo Fisher) and 0.08% Pluronic F-127 (Thermo Fisher) for 30 min at 37 °C and 5% CO_2. Cells were subsequently labeled for 20 min at room temperature with surface marker antibodies CD19 (Super Bright 600; Invitrogen) and CD5 (PE-Cy5; BD Biosciences) for the identification of tumoral cells ($CD19^+CD5^+$). Cells were next resuspended on RPMI-1640 before flow cytometry acquisition. Basal calcium was measured during 1 min before stimulation, then cells were incubated during 2 min at 37 °C with or without 10 µg ml^{-1} anti-human F(ab')2 IgM (Southern Biotech) and 3.3 mM H_2O_2 (Sigma-Aldrich). Finally, 2 µM 4-hydroxytamoxifen (4-OHT) (Sigma-Aldrich) was added to all conditions before continue recording for up to 8 min. Intracellular Ca^{2+} release was measured on LSRFortessa (BD Biosciences) using BD FACSDiva software (version 8) by exciting with ultraviolet laser (355 nm) and appropriate filters: Indo-1 violet (450/50 nm) and Indo-1 blue (530/30 nm). Bound (Indo-1 violet) and unbound (Indo-1 blue) ratiometric was calculated with FlowJo software (version 10.7.1). Gating analysis was as follows: cell identification in FSC-A versus SSC-A plot, singlet identification in FSC-A versus FCS-H plot, tumoral cells ($CD19^+CD5^+$)

in CD19 (Super Bright 600) versus CD5 (PE-Cy5) plot and Ca^{2+} release in time versus Indo-1 violet/Indo-1 blue plot using a kinetics tool.

Clinical analyses

Primary end points were overall survival and time to first treatment, calculated from the date of diagnosis. Median follow-up was calculated with the Kaplan-Meier estimate of potential follow-up.[50]

SUPPLEMENTARY TABLES

Tables not included in this PDF document can be found in the Supplementary Tables Excel file.

Supplementary Table S5. BCL3 expression according to the location of the translocation breakpoint on chromosome 19

BCL3 **Breakpoint**	**BCL3 overexpression**		**Total**
	RNA-seq	IHC	
Upstream	6/6	3/3	8/8
Downstream	0/1	0/4	0/4

Supplementary Table S18. Clinical characteristics of the *BCL3*-R tumors

Patient ID	Age, Sex	Clinical presentation	ALC ($x10^9$/L)	Treatment (in sequence)	Follow-up time (years)	Cause of death
Downstream *BCL3*-R tumors						
3721	81, Female	Splenomegaly, B symptoms	0.9	R-COPx2 + chlorambucil (PR); low dose cyclophosphamide (SD)	2.7, AR	
3649	70, Female	Splenomegaly, B symptoms	1	Splenectomy followed by FCRx5 (CR)	3.5, D	Unrelated to disease, COVID pneumonia, and therapy-related myelodysplastic syndrome
4692	53, Male	Splenomegaly	1	Splenectomy (CR); R-CHOPx6 (CR)	7.5, AR	
3676	59, Male	-	2.91	-	11, D	CNS hemorrhage before start of R-CHOP, with clinical transformation (splenomegaly, B-symptoms, lymphocytosis)
Upstream *BCL3*-R tumors						
3663	78, Female	Lymphocytosis	9.6	-	5.3, AD	
3646	71, Male	Lymphocytosis	19.5	-	1.8, AD	
3706	56, Female	Lymphadenopathy, splenomegaly, lymphocytosis and B symptoms	161.8	Ibrutinib (CR)	1.2, AR	
3619	78, Female	Lymphocytosis	8.5	-	0.5, AD	
3783	68, Female	Lymphadenopathy, lymphocytosis	8	-	0.9, AD	
624	50, Male	NA	NA	Chlorambucil (SD); DHAP + allogeneic HCT from matched unrelated donor (CR)	10.5, D	Graft versus host disease related complications
3696	69, Male	Lymphocytosis	8.3	Ibrutinib (PD)	5.1, D	Unrelated to disease, surgical complications
3698	56, Male	Lymphocytosis	12.4	R-CHOPx6 (CR); Ibrutinib (CR)	8.9, AR	
1826	77, Male	Lymphadenopathy	2.9	No treatment due to comorbidities	0.7, D	Cardiac arrest and mesentheric thrombosis, with progressive disease

Abbreviations: ID: identifier; ALC: absolute lymphocyte count; PR: partial response; SD: stable disease; CR: complete response, AR: alive in remission; AD: alive with disease; D: dead; HCT: hematopoietic cell transplantation; R-COP: Rituximab, cyclophosphamide, vincristine and prednisone; R-CHOP: Idem plus daunorubicin; FCR: Fludarabine, Rituximab and cyclophosphamide; DHAP: Dexamethasone, cysplatin and cytarabine; CNS: central nervous system.

Supplementary Table S19. Karyotype and immunophenotype of 13 tumors according to the 5' upstream or 3' downstream *BCL*3-R

Patient ID	Timing of study	Karyotype	*BCL3* FISH	PB FC Immunophenotype	Immunohistochemistry
		Downstream *BCL3*-R tumors			
3721	Diagnosis	46,XX,del(7)(q22q32),del(11)(q21q23), t(14;19)(q32;q13),add(16)(q24)[cp7]	Rearranged	CD19+, CD20+, CD22+, CD79b±, CD5-, CD23-, CD200-, CD43-, FMC7±, kappa+	CD20+, CD5-, CD23-, LEF1-, IgD weak, BCL3-
3649	Diagnosis	46,XX,t(14;19)(q32;q13)[8]/ 46,XX[12]	NA	CD19+^{dim}, CD20+, CD22+, CD79b-, CD5-, CD23-, CD200-, CD43+, FMC7-, kappa+	CD20+, CD5-, CD23-, LEF1-, IgD+; KI67 10%, BCL3-
4692	Progression (5 years after initial diagnosis)	NA	NA	B-antigens+, CD5-, CD23-, FMC7++, kappa+ (lymph node)	
3676	Progression (9 years after initial diagnosis)	48-52,XY, add(6)(q26), -8,add(10)(p11),+13, t(14;19)(q32;q13), +16,add(17)(q25),-18, der(?)t(?;2)(?;q12),+mar1,+mar2 ,+mar3,+mar4,+mar5,+mar6[cp17]	Rearranged	B-antigens+, CD5-, CD23-, CD200+, CD43-, FMC7+, lambda+	CD20+, CD5-, CD23-, LEF1-, IgD-; KI67 20%, BCL3-
		Upstream *BCL3*-R tumors			
3663	Diagnosis	47,XX,+12,del(12)(p13), t(14;19)(q32;q13)[6]/ 46,XX[8]	Rearranged	B-antigens+^{dim}, CD5+, CD23-, CD200+, CD43-, FMC7-, lambda+^{dim}	NA
3646	Diagnosis	47,XY,+12[14]/47,XY,+12, t(14;19)(q32;q13)[8]	Rearranged	B-antigens+^{dim}, CD5+, CD23+, CD200+, CD43+, FMC7-, kappa+^{dim}	NA
3706	Diagnosis	47,XX,+12[2]/ 47,XX,+12,t(14;19)(q32;q13)[11]/ 46,XX,+12,t(14;19)(q32;q13), der(21;22)(q10;q10)[2]	Rearranged	B-antigens+^{dim}, CD5+, CD23+, CD200+, CD43+, FMC7-, kappa+^{dim}	CD20+, CD5+, CD23+, LEF1-, IgD+; KI67 40%,
3619	Diagnosis	47,XX,+12,t(14;19)(q32;q13)[5]/ 46,XX[15]	Rearranged	B-antigens+, CD5+, CD23+, CD200+, CD43±, FMC7±, kappa+	NA
3783	Diagnosis	48,XX,+12,t(14;19)(q32;q13),+21[3]/ 46,XX[17]	NA	B-antigens+, CD5+, CD23+, CD200+, CD43+, FMC7-, lambda+	NA
624	Diagnosis made 6 years before in another center (no previous information available)	47,XY,+12,t(14;19)(q32;q13), add(21)(q22) [4]/ 46,XY[33]	Rearranged	B-antigens+, CD5+, CD23+, CD200+, CD43-, FMC7++, kappa+	NA
3696	Progression (2,5 years after initial diagnosis)	46,XY[20]*	Rearranged	B-antigens+, CD5+, CD23+^{dim}, CD200+, CD43+dim, FMC7-, kappa+	CD20+, CD5-, CD43+ weak, CD23 weak, LEF1-, Ki67 50%, BCL3+.
3698	Progression (2,5 years after initial diagnosis)	47,XY,+12,t(14;19)(q32;q13)[9]/ 47,XY,+12,t(14;19)(q32;q13), inv(6)(p25q13)[5]	Rearranged	B-antigens+, CD5+, CD23+^{dim}, CD200+, CD43-, FMC7+, kappa+^{dim}	CD20+, CD5+, CD23 weak, LEF1-, IgD weak; KI67 40%, BCL3+
1826	Progression (7 months after diagnosis)	46,XY,?del(10)(p?)[3], del(11)(q21q24)[11],del(14)(q?11q?31),t(14;19)(q32;q13)[9][cp11]/ 46,XY[10]	Rearranged	B-antigens+^{dim}, CD5+, CD23+, CD200+, CD43+, FMC7-, kappa+^{dim}	CD20+, CD5+, CD23+, LEF1-, IgD-; KI67 30%, BCL3+.

*Trisomy 12 detected by FISH in 5% of nuclei.

Abbreviations: FISH: Fluorescence in situ hybridization; PB: peripheral blood; FC: flow cytometry

Supplementary Table S20: Summary of the diagnosis, immunophenotypic and genetic characteristics of 17 B-cell lymphoid neoplasms with the *BCL3*-rearrangement in the validation cohort

	Total (n=17)	Upstream *BCL3*-R (n=13)	Downstream *BCL3*-R (n=4)
Diagnosis			
CLL	3 / 17	3 / 13	0 / 4
aCLL	8 / 17	8 / 13	0 / 4
nnMCL	2 / 17	2 / 13	0 / 4
SMZL	3 / 17	0 / 13	3 / 4
SCL, NOS	1 / 17	0 / 13	1 / 4
Phenotype			
Flow cytometry			
Typical for CLL**	3 / 17	3 / 13	0 / 4
Bright B-cell markers	12 / 17	8 / 13	4 / 4
CD5 +	13 / 17	13 / 13	0 / 4
CD43 +	7 / 10	6 / 7	1 / 3
CD23 +	16 / 17	9 / 13	1 / 3
Immunohistochemistry			
LEF1 +*	0 / 6	0 / 6	
BCL3 +[#]	2 / 4	2 / 2	0 / 2
Cyclin D1[+]	2 / 2	2 / 2	
SOX11[+]	0 / 2	0 / 2	
Genetics			
Unmutated IGHV status	6 / 7	6 / 7	
Trisomy 12	8 / 17	8 / 13	0 / 4
del(7)(q32)	2 / 17	0 / 13	2 / 4
Complex karyotype	12 / 17	9 / 13	3 / 4

Abbreviations: *BCL3*-R, *BCL3* rearrangement; CLL, chronic lymphocytic leukemia; aCLL, atypical chronic lymphocytic leukemia; SCL NOS, small B-cell lymphoma not otherwise specified; nnMCL, leukemic non-nodal mantle cell lymphoma; SMZL, splenic marginal zone lymphoma.

*The 6 cases examined were CLL or aCLL.

[#]The two *BCL3* positive tumors were one aCLL and one MCL. The two negative tumors were a SMZL and a nodal MZL.

[+]Cyclin D1 expression and SOX11 were studied in the 2 nnMCL.

Supplementary Table S21: Pathological and genetic features of 17 B-cell lymphoid neoplasms with *BCL3*-rearrangement included in the validation cohort

ID	Diagnosis	*BCL3* breakpoint	Phenotype	IGHV	Other FISH and molecular	Cytogenetics
1	CLL	5' upstream	Dim B-cell markers CD5+ CD23+ CD200+ CD43+ ROR1+	Unmutated	Mutations in *NOTCH1* and *XPO1*	46,XX,t(2;13)(q33;q14),t(14;19)(q32;q13)[9]/46,XX[2]
2	CLL	5' upstream	Dim B-cell markers CD5+ dim CD23+ dim CD200+ dim CD43+	Unmutated	*TP53 w*t	47,XY,+12,t(14;19)(q32;q13)[11]/48,idem,+2[6]/46,XY[3]
3	aCLL	5' upstream	Bright B-cell markers CD79b+ dim CD5+ CD23- LEF1 -	Unmutated	+12 13q,11q and17p wt *TP53, MYD88, SF3B*1 and *NOTCH1* wt	47,XX,t(3;8)(p21;p?21),+12,t(14;19)(q32;q13)[20]
4	aCLL	5' upstream	B-cell markers bright Lambda bright CD5+ CD23+ dim LEF1 - CCND1 -	Unmutated	+12 -13q 17p wt *IGH::CCND1* neg	47,XY,+12[4]/47,idem,t(14;19)(q32;q13)[6]/46,XY[10]
5	aCLL	5' upstream	Bright B-cell markers Kappa+ bright CD79b (partial dim) CD5+ dim CD23 - CD11c+ dim LEF1 - CCND1 -		12, 13q, 11q and 17p wt *IGH::CCND1* neg	46,XY,t(14;19)(q32;q13)[1]/46,idem,t(1;6)(p35;p25)[6]/46,XY[13]
6	aCLL / Richter's transformation	5' upstream	Dim B-cell markers CD5+ CD23+ dim CD200- CD43+ dim FMC7-			93,XXYY,+12,t(14;19)(q32;q13)[8] / 46,XY[12]
7	aCLL	5' upstream	B-cell markers dim kappa+ dim CD5+ CD23+ CD200+ CD43+ LEF1-	Unmutated		47,XX,+12,t(14;19)(q32;q13)[14]/47,idem,?t(12;13)(q24.1;q21)[5]/46,XX[1]

8	aCLL	5’ upstream	CD20+ CD79b+ dim Kappa+ CD5+ CD23+ dim CD200+ CD43- ROR1-			47,XY,+12,t(14;19)(q32;q13)[16]/47,XY,t(1;14;19)(p13;q32;q13.3),+12[4]
9	aCLL	5’ upstream	Bright B-cell markers CD5+ CD23+ dim CD43+ dim CCND1-	Unmutated	*TP53* wt *NOTCH1* mut	47,XX,+12,t(14;19)(q32;q13)[5]/47,idem,t(1;11)(p11;q24)[2]/46,XX[13]
10	aCLL	5’ upstream	Bright B-cell markers CD5+ CD23- CD200+ CD43+ dim CD11c+ dim LEF1- BCL3+			47,XX,+12,t(14;17)(q32;q24),t(14;19)(q32;q13),add(18)(p11)[16]
11	aCLL	5’ upstream	Dim B-cell markers Kappa+ CD5+ dim CD23- CD200+ FMC7+ dim CD11c+ LEF1-	Mutated	*TP53* wt, *MYD88* wt	46,XY,t(14;19)(q32;q13)[5]/46,XY,der(5)(5pter-->5q35::6q13-->6q22::14q31-->14q32::19q13-->19qter),del(6)(q13),del(13)(q14q34),der(14)(14pter-->14q31::6q23),der(19)t(14;19)(q32;q13)[10]
12	nnMCL	5’ upstream	Bright B-cell markers Lambda bright CCND1 + SOX11 - CD5 + CD23 + (subset) CD123 + (subset) CD10 -		-11q -13q -17p	45,XY,der(1)del(1)(p36)t(1;13)(q42;q22),del(2)(q33q35),add(6)(q13),der(11)t(11;14)(q13;q32),der(14)t(11;14)del(11)(q23q25), t(14;19)(q32;q13),-17[10]/46,XY[10]
13	nnMCL	5’ upstream	Bright B-cell markers Kappa+ CD5+ CD23+ dim CD200+ BCL3+		*TP53* wt	47,XX,t(11;22)(q13;q11),+12,t(14;19)(q32;?q13)[12]
14	SMZL	3’ downstream	Bright B-cell markers Kappa+ dim CD5-, CD23- CD11c-, BCL3-			46,XX,del(7)(q21q35),t(3;14)(q27;q32),t(14;19)(q32;q13)[cp18]/46,XX

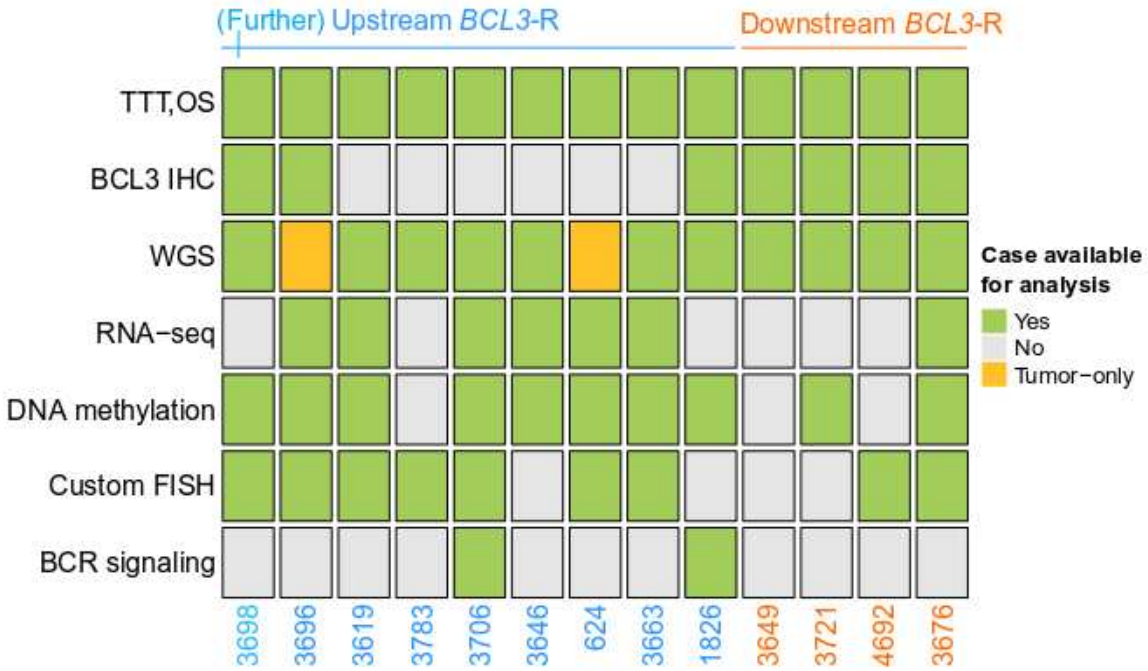
15	SMZL*	3’ downstream	Bright B-cell markers CD5- CD10- CD43- CD23- FMC7+ PAX5+ Cyclin D1- Bcl6-		Spleen:46,XX,add(14)(q32)[15]/46,XX[5] Lymph node: 45,XX,-14,der(19)(19q13.3→19p13.3::14q13→14q32::19p13:),der(19)(19p13.3→19q13::14q32)[11]/46,XX[5]
16	SMZL	3’ downstream	Bright B-cell markers CD5- CD23+ dim CD43+ FMC7+ BCL3-	Mutations in *TNFAIP3, NOTCH1, KMT2D, DNTM3A, CREBBP*	47,XX,add(1)(q23),+3,der(4)t(4;12)(q32;q14),del(6)(q23q26),del(7)(q21q32),add(12)(p11),der(14)t(1;14)(q22;q32),t(14;19)(q32;q13),dup(17)(q22q25),del(18)(p11)[12]/47,XX,add(1),+3,del(6),del(7),der(9)t(9;11)(q34;q14),add(12),der(14)t(14;19),dup(17),der(19)t(1;19)(q22;q13)[5]/46,XX[3]
17	SCL, NOS	3’ downstream	Bright B-cell markers CD5- CD200- CD43- CD11c+		46,XY,t(14;19)(q32;q13),der(22)t(1;22)(q11;p13)[6]/46,XY[14]

*Previous published in reference[51]. Abbreviations: *BCL3*-R, *BCL3* rearrangement; CLL, chronic lymphocytic leukemia; aCLL, atypical chronic lymphocytic leukemia; SCL, NOS, small B-cell lymphoma, not otherwise specified; nnMCL, leukemic non-nodal mantle cell lymphoma; SMZL, splenic marginal zone lymphoma.

SUPPLEMENTARY FIGURES

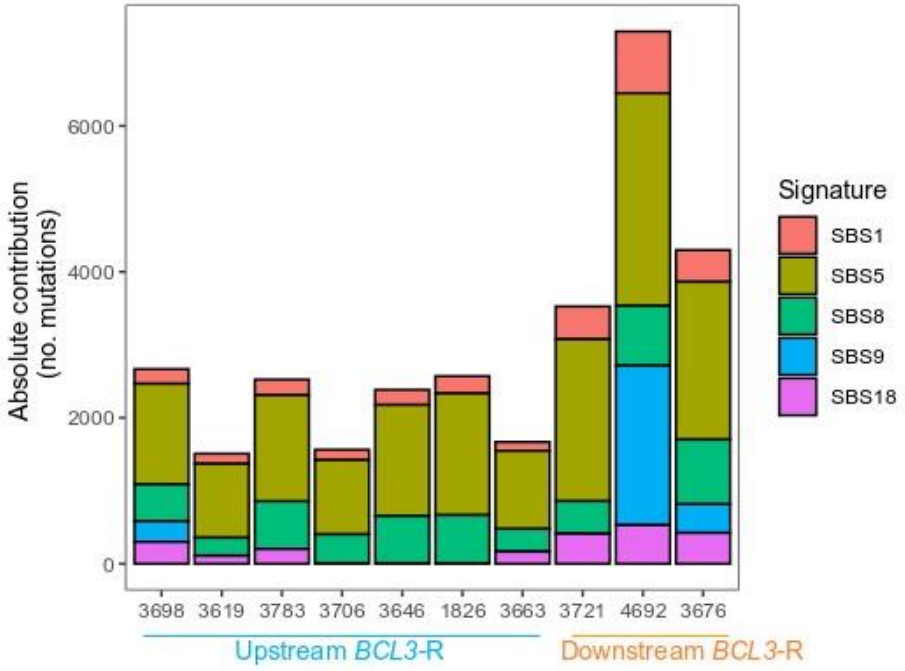
Supplementary Figure S1. Schema of the analyses performed in each *BCL3*-R tumor.

Diagram of the availability of the tumor samples in each analysis. Green squares represent availability of a tumor sample in that analysis; gray squares represent no availability, and golden squares represent tumor-only availability in whole-genome sequencing (WGS) analyses. TTT: time-to-first treatment; OS: overall survival; IHC: immunohistochemistry.



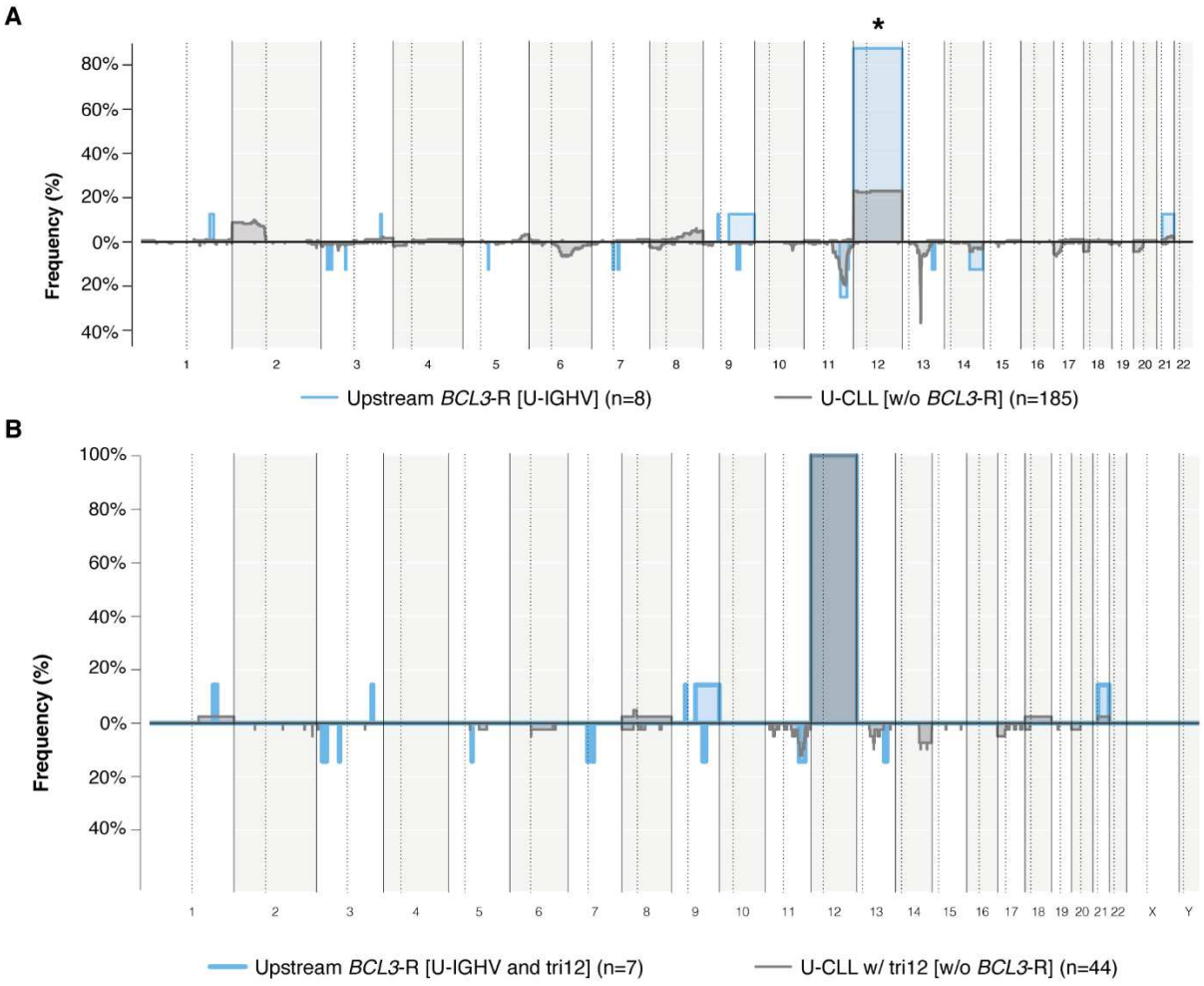
Supplementary Figure S2. Mutational signature analysis performed in *BCL3*-R tumors.

Mutational signature analysis in the upstream and the downstream *BCL3*-R tumors.



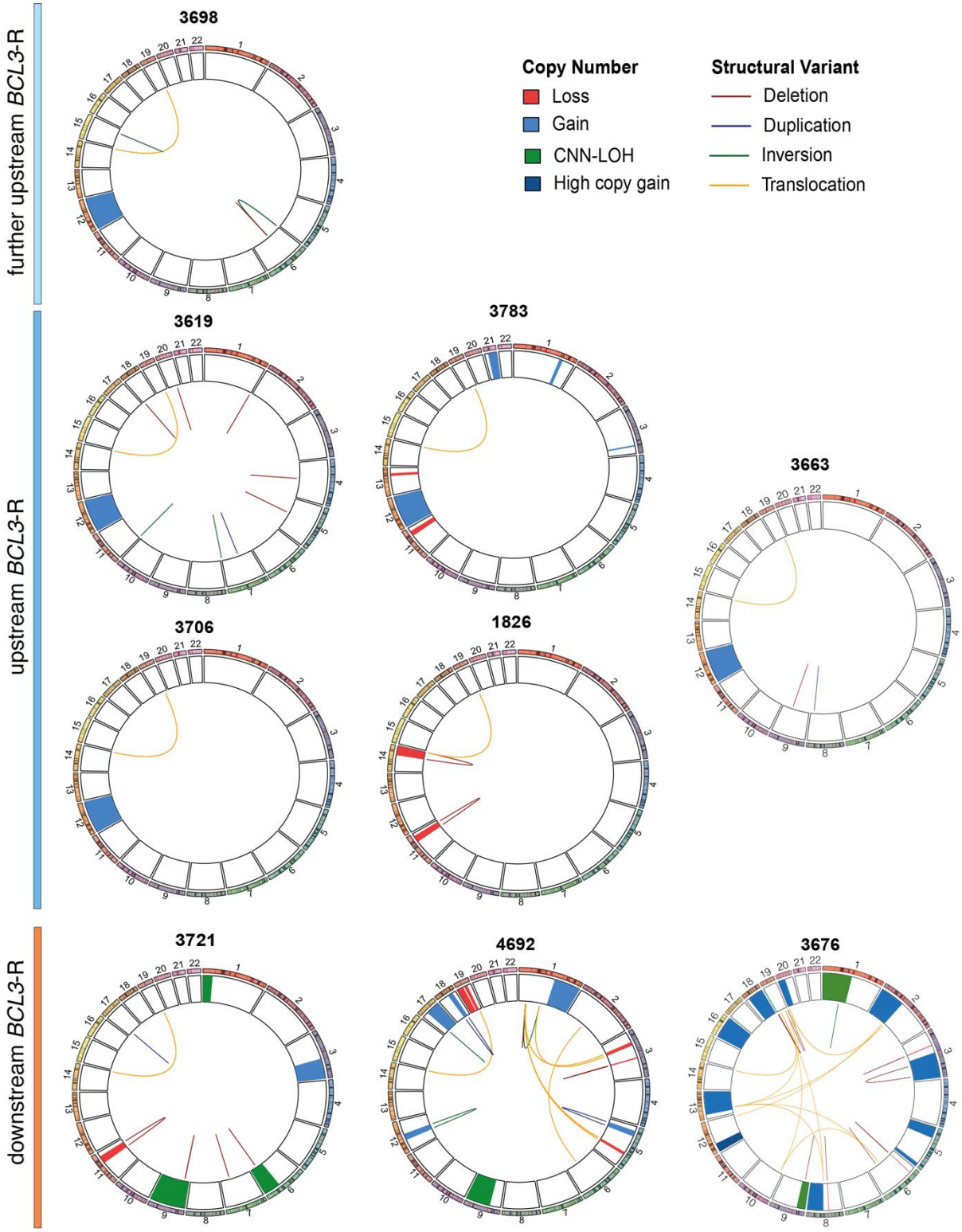
Supplementary Figure S3. Frequency of CNA in the upstream *BCL3*-R tumors vs CLL.

A. Comparison of CNA frequency between upstream *BCL3*-R tumors vs CLL [all unmutated IGHV]. **B.** Comparison of CNA frequency between upstream *BCL3*-R vs CLL [all unmutated IGHV and trisomy 12]. The x axis shows the 12 chromosomes, while the y axis displays the frequency of CNA.



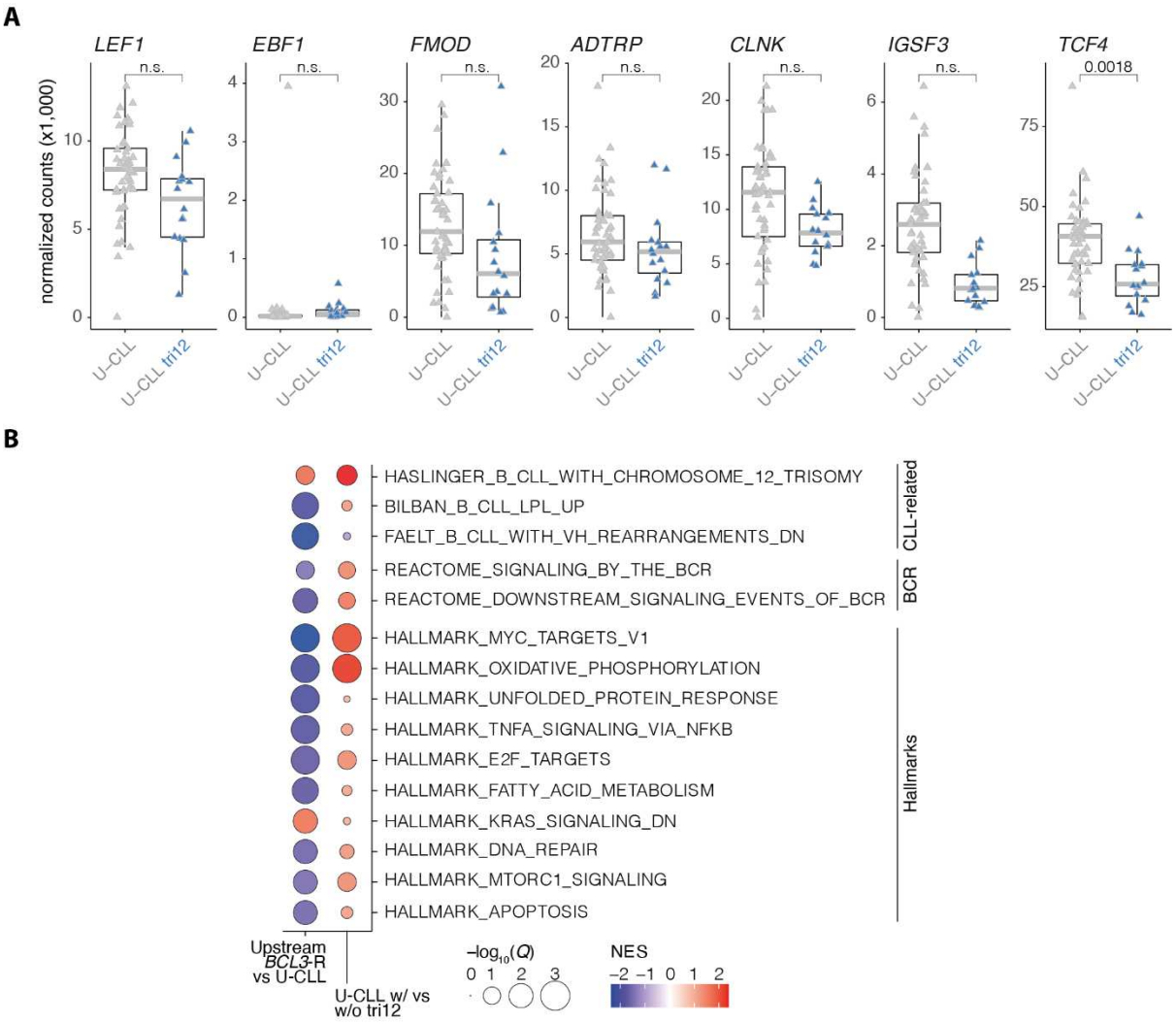
Supplementary Figure S4. Chromosomal landscape of *BCL3*-R tumors.

Illustration of SV and CNA in six upstream *BCL3*-rearranged (*BCL3*-R) tumors and three downstream *BCL3*-R tumors. The innermost layer depicts SV with lines, the middle layer shows CNA with boxes, and the outermost layer indicates the chromosome.



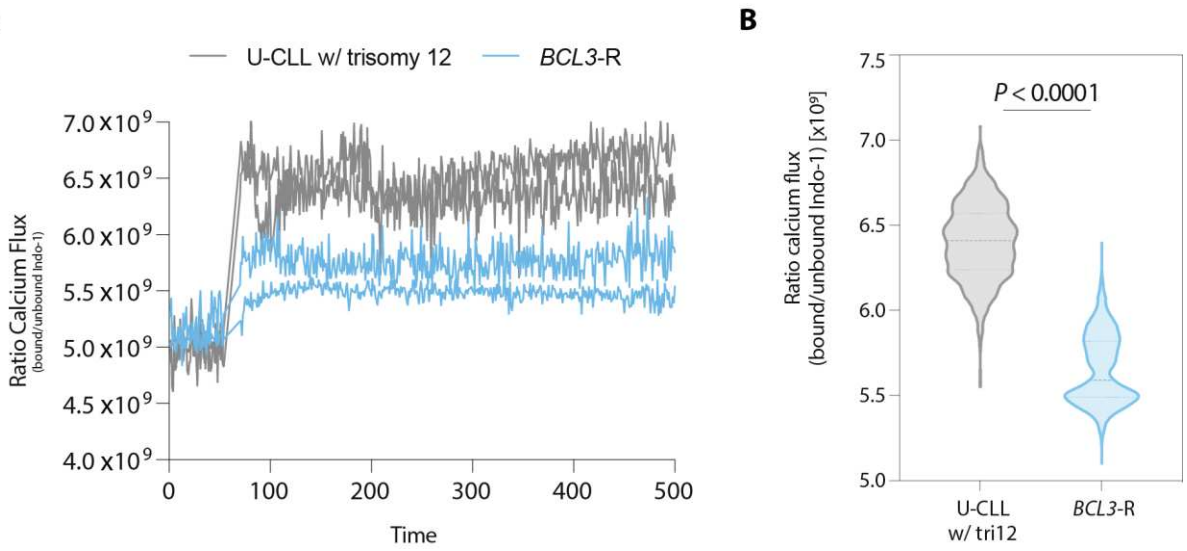
Supplementary Figure S5. Expression of CLL hallmark genes and GSEA.

A. Expression of CLL hallmark genes in U-CLL without trisomy 12 compared to U-CLL with trisomy 12. Q-values are from the DEA. n.s., not significant (Q-value < 0.05, absolute(log2FC) > 0.1). **B.** Representation of the most relevant significantly enriched pathways in the upstream *BCL3*-R tumors vs U-CLL and U-CLL with vs without trisomy 12.



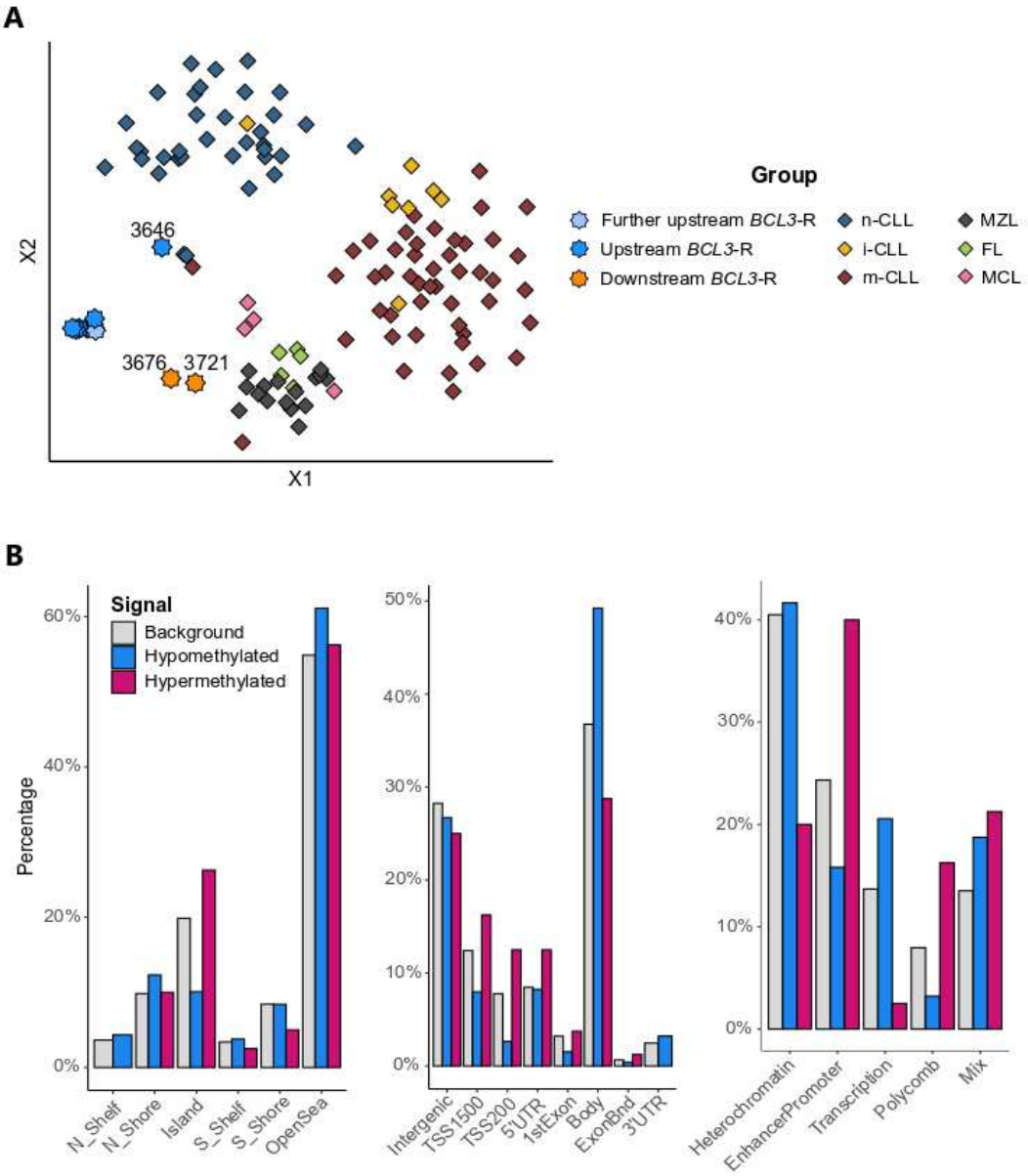
Supplementary Figure S6. Calcium flux of tumoral cells after BCR stimulation.

A. Calcium flux was measured on tumoral cells (CD19+CD5+). Basal calcium was adjusted at $5x10^9$ ($Indo^{-1}$ ratio) for 60 seconds, then cells were stimulated with IgM + H_2O_2 at 37ºC and 4-hydroxytamoxifen (4-OHT), calcium flux was recorded up to 500 seconds. Two samples of upstream *BCL3*-rearranged tumors were compared to two CLL with unmutated IGHV (U-CLL) and trisomy 12. **B.** Violin plots of the calcium release after BCR stimulation in the two upstream *BCL3*-R tumors vs the two U-CLL with trisomy 12.

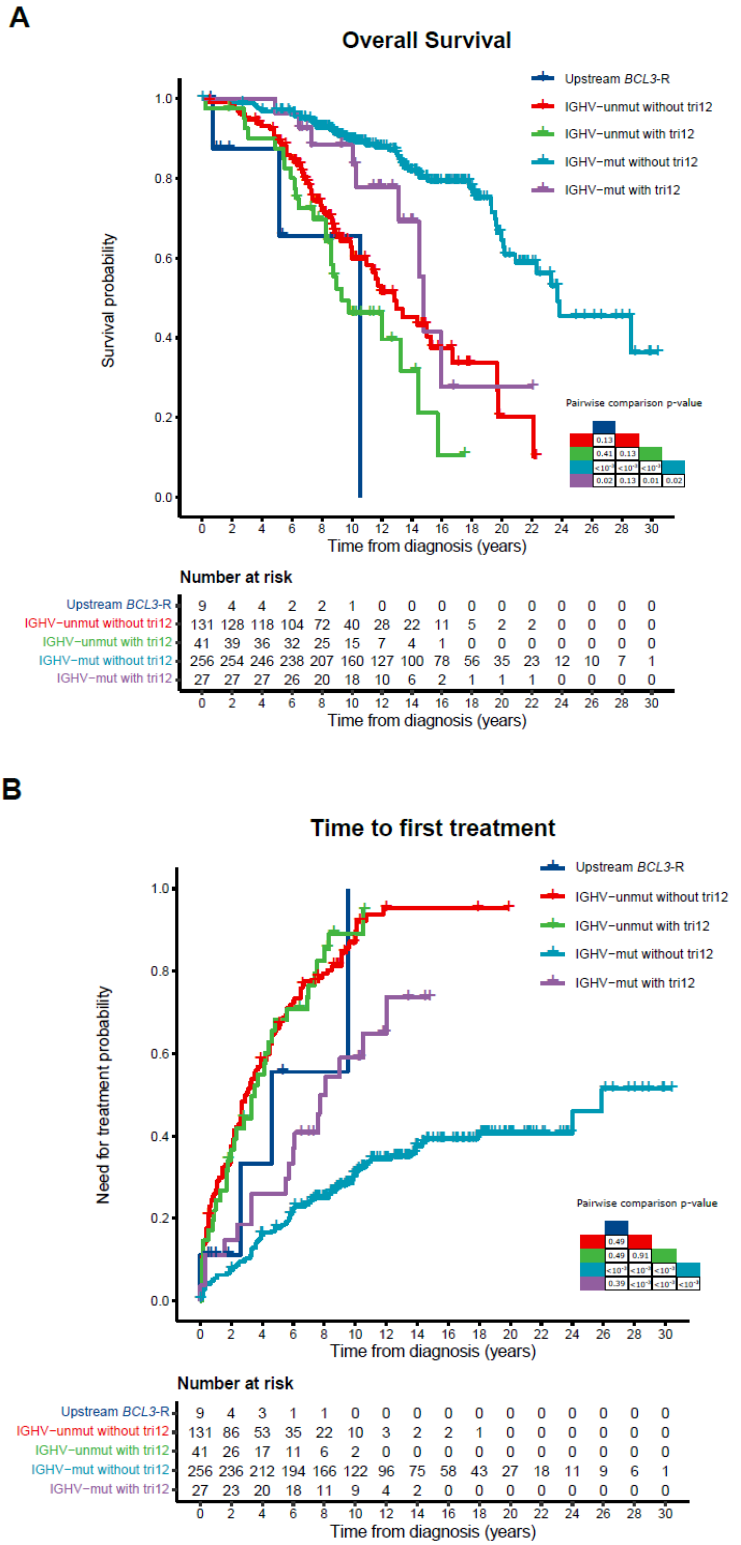


Supplementary Figure S7. T-distributed stochastic neighbor embedding analysis on the 795 DMCpGs and bar plots of their genomic location.

A. T-distributed stochastic neighbor embedding performed on the 795 DMCpGs for 10 B-cell neoplasms with *BCL3*-rearrangement (*BCL3*-R), 90 CLL [12 C1, 73 C2, 5 GSE171424[41]], 17 marginal zone lymphomas (7 splenic, 2 nodal, 5 extranodal, 3 not specified) from GSE171424,[41] 5 follicular lymphoma from GSE171424,[41] 4 mantle cell lymphoma from GSE171424,[41] and 7 normal B-cells (first and second components are shown). The tumor type is represented by stars (B-cell neoplasms with *BCL3*-R) and diamonds (B-cell neoplasms without *BCL3*-R), while color is representing the different groups. The tumor carrying a subclonal *BCL3*-R (3646) and the two with downstream *BCL3*-R (3676 and 3721) are labeled. **B.** Bar plots representing the distribution of the DMCpGs based on its genomic location (left and middle graphs) and CLL chromatin states (right graph).

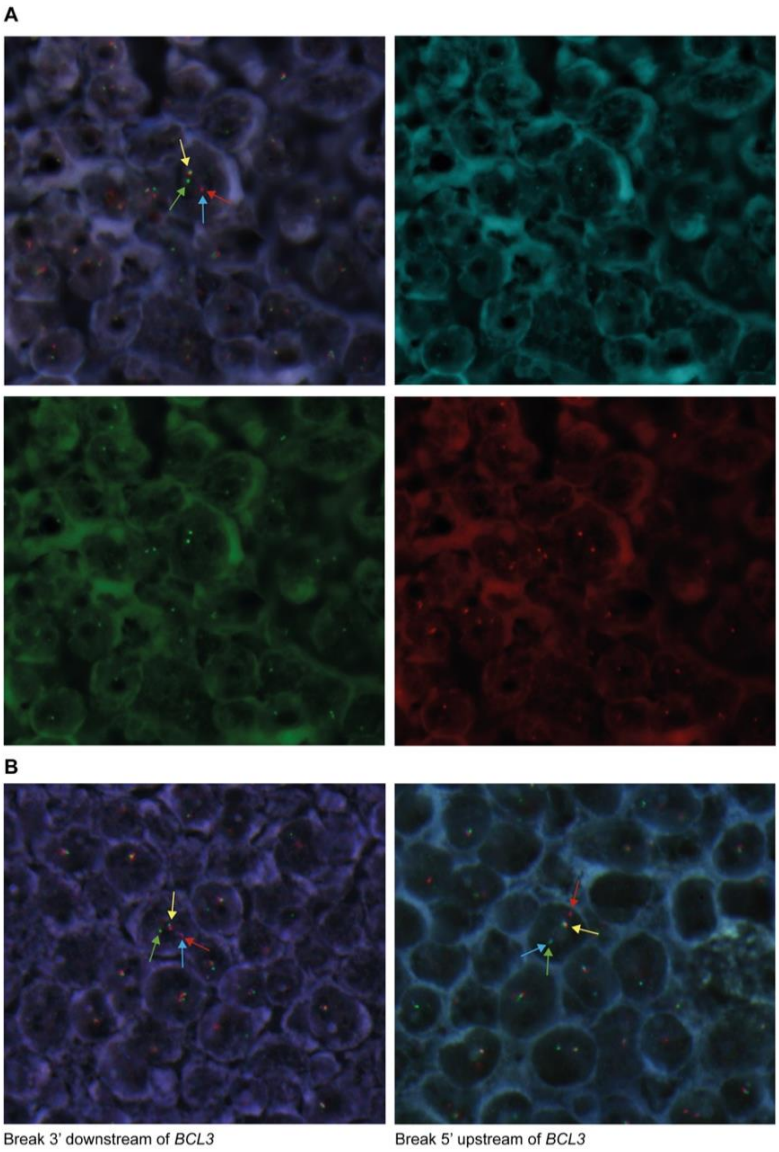


Supplementary Figure S8. Survival analysis between upstream *BCL3*-R and CLL. Comparison of A. overall survival and B. time to first treatment between upstream *BCL3*-R and CLL. CLL tumors have been stratified according to their IGHV mutational status and presence/absence of trisomy 12.

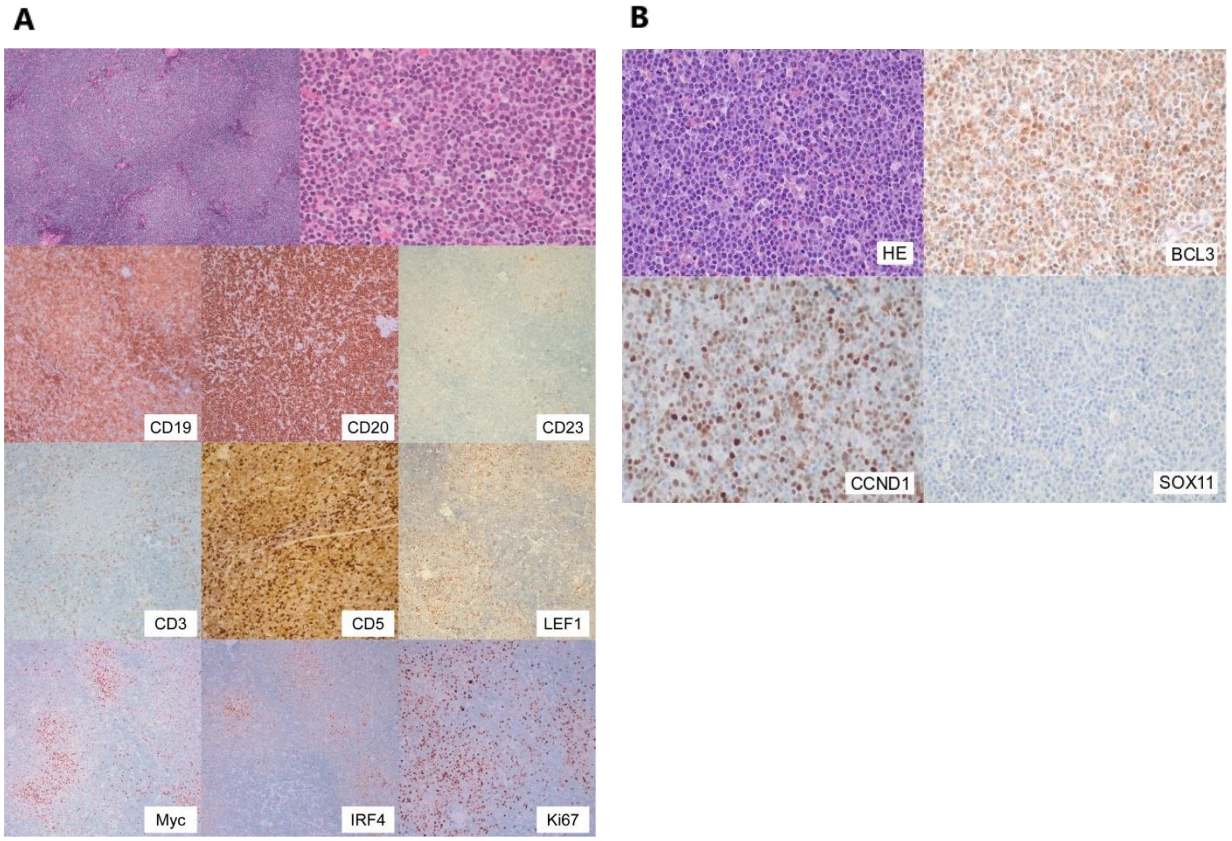


Supplementary Figure S9. FISH analysis of the 5' upstream and 3' downstream *BCL3* breakpoints using the custom FISH assay.

A. Tumor 4692 from the initial cohort. The first panel shows merging signal patterns of all BACs of the probe, indicating a positive signal constellation for a break downstream of *BCL3* gene, one colocalized signal (yellow arrow) and one red signal (red arrow) split from blue and green signals (blue and green arrows). The second, third and four panels display the signal pattern of each specific BAC clone of the *BCL3* probe. **B.** Tumors 4 (left) and 1 (right) from the validation cohort. Tumor 4 shows a positive signal constellation for a break downstream of *BCL3*, one colocalized signal (yellow arrow) and one green signal (green arrow) split from blue and red signals (blue and red arrows). Tumor 1 indicates a positive signal constellation for a break upstream of *BCL3*, one colocalized signal (yellow arrow) and one red signal (red arrow) split from blue and green signals (blue and green arrows).



Supplementary Figure S10. Histological sections of tumors 1 and 13 from the validation cohort. A. Tumor 1: chronic lymphocytic leukemia. Low power magnification (100x) shows proliferation centers. High power magnification (400x) of a proliferation center with increased number of paraimmunoblasts and prolymphocytes. CD19 immunohistochemistry (100x) shows slight decreased intensity in proliferation centers. CD20 immunohistochemistry (100x) shows high expression. CD23 immunohistochemistry (100x) shows few positive cells in proliferation centers. CD3 immunohistochemistry (100x) shows few admixed T cells. CD5 immunohistochemistry (100x) shows admixed T cells (high intensity) and faint positivity in tumor cells. LEF1 immunohistochemistry (100x) shows T cells and few cells in proliferation centers. MYC immunohistochemistry (100x) shows expression in proliferation centers. IRF4 immunohistochemistry (100x) shows expression in proliferation centers. Ki67 immunohistochemistry (100x) shows elevated proliferation rate in proliferation centers. **B.** Tumor 13: leukemic non-nodal mantle cell lymphoma with *CCND1* rearrangement due to a t(11;22)(q13.3;q11.21) and 5' *BCL3*-R. Lymph node diffusely infiltrated by a lymphoid proliferation composed of medium-sized cells with scant cytoplasm, and irregular nuclei. Occasional plasmacytic differentiation is seen, as described in some cases of SOX11-negative MCL variant[52] (HE 400x). The mitotic index is high (3 mitoses per high power field, 400x). Immunohistochemical staining show that the neoplastic cells are positive for BCL3 (400x) and cyclin D1 (400x). SOX11 is completely negative (400x).



SUPPLEMENTARY REFERENCES

1. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinforma Oxf Engl 2009;25(14):1754–1760.
2. Tischler G, Leonard S. biobambam: tools for read pair collation based algorithms on BAM files. Source Code Biol Med 2014;913.
3. Nadeu F, Royo R, Massoni-Badosa R, et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. Nat Med 2022;28(8):1662–1671.
4. Jones D, Raine KM, Davies H, et al. cgpCaVEManWrapper: Simple Execution of CaVEMan in Order to Detect Somatic Single Nucleotide Variants in NGS Data. Curr Protoc Bioinforma 2016;5615.10.1-15.10.18.
5. Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C, Lichtenstein L. Calling Somatic SNVs and Indels with Mutect2. 2019;861054.
6. Fan Y, Xi L, Hughes DST, et al. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. Genome Biol 2016;17(1):178.
7. Raine KM, Hinton J, Butler AP, et al. cgpPindel: Identifying Somatically Acquired Insertion and Deletion Events from Paired End Sequencing. Curr Protoc Bioinforma 2015;5215.7.1-15.712.
8. Rimmer A, Phan H, Mathieson I, et al. Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. Nat Genet 2014;46(8):912–918.
9. Wala JA, Bandopadhayay P, Greenwald NF, et al. SvABA: genome-wide detection of structural variants and indels by local assembly. Genome Res 2018;28(4):581–591.
10. Cingolani P, Platts A, Wang LL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin) 2012;6(2):80–92.
11. Cingolani P, Patel VM, Coon M, et al. Using Drosophila melanogaster as a Model for Genotoxic Chemical Mutational Studies with a New Program, SnpSift. Front Genet 2012;335.
12. Nik-Zainal S, Van Loo P, Wedge DC, et al. The Life History of 21 Breast Cancers. Cell 2012;149(5):994–1007.
13. Raine KM, Van Loo P, Wedge DC, et al. ascatNgs: Identifying Somatically Acquired Copy-Number Alterations from Whole-Genome Sequencing Data. Curr Protoc Bioinforma 2016;5615.9.1-15.9.17.
14. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534(7605):47–54.
15. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics 2012;28(18):i333–i339.
16. Robinson JT, Thorvaldsdóttir H, Winckler W, et al. Integrative genomics viewer. Nat Biotechnol 2011;29(1):24–26.
17. Nadeu F, Delgado J, Royo C, et al. Clinical impact of clonal and subclonal TP53, SF3B1, BIRC3, NOTCH1, and ATM mutations in chronic lymphocytic leukemia. Blood 2016;127(17):2122–2130.

18. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012;22(3):568–576.

19. Lai Z, Markovets A, Ahdesmaki M, et al. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. Nucleic Acids Res 2016;44(11):e108.

20. Wilm A, Aw PPK, Bertrand D, et al. LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. Nucleic Acids Res 2012;40(22):11189–11201.

21. Muller E, Goardon N, Brault B, et al. OutLyzer: software for extracting low-allele-frequency tumor mutations from sequencing background noise in clinical practice. Oncotarget 2016;7(48):79485–79493.

22. Puente XS, Beà S, Valdés-Mas R, et al. Non-coding recurrent mutations in chronic lymphocytic leukaemia. Nature 2015;526(7574):519–524.

23. Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics 2012;28(3):423–425.

24. 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.

25. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. N Engl J Med 2018;378(15):1396–1407.

26. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat Med 2018;24(5):679–690.

27. Grau M, López C, Navarro A, et al. Unraveling the genetics of transformed splenic marginal zone lymphoma. Blood Adv 2023;bloodadvances.2022009415.

28. Blokzijl F, Janssen R, van Boxtel R, Cuppen E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. Genome Med 2018;10(1):33.

29. Alexandrov LB, Kim J, Haradhvala NJ, et al. The repertoire of mutational signatures in human cancer. Nature 2020;578(7793):94–101.

30. Dörner T, Foster SJ, Farner NL, Lipsky PE. Somatic hypermutation of human immunoglobulin heavy chain genes: targeting of RGYW motifs on both DNA strands. Eur J Immunol 1998;28(10):3384–3396.

31. Hübschmann D, Kleinheinz K, Wagener R, et al. Mutational mechanisms shaping the coding and noncoding genome of germinal center derived B-cell lymphomas. Leukemia 2021;35(7):2002–2016.

32. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30(15):2114–2120.

33. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 2012;28(24):3211–3217.

34. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol 2016;34(5):525–527.

35. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research 2015;41521.

36. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550.

37. Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. Bioinforma Oxf Engl 2019;35(12):2084–2092.

38. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. Fast gene set enrichment analysis. 2021;060012.

39. Dietrich S, Oleś M, Lu J, et al. Drug-perturbation-based stratification of blood cancer. J Clin Invest;128(1):427–445.

40. Lu J, Cannizzaro E, Meier-Abt F, et al. Multi-omics reveals clinically relevant proliferative drive associated with mTOR-MYC-OXPHOS activity in chronic lymphocytic leukemia. Nat Cancer 2021;2(8):853–864.

41. Xia D, Leon AJ, Yan J, et al. DNA Methylation-Based Classification of Small B-Cell Lymphomas: A Proof-of-Principle Study. J Mol Diagn 2021;23(12):1774–1786.

42. Kulis M, Merkel A, Heath S, et al. Whole-genome fingerprint of the DNA methylome during human B-cell differentiation. Nat Genet 2015;47(7):746–756.

43. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30(10):1363.

44. 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.

45. Kulis M, Heath S, Bibikova M, et al. Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia. Nat Genet 2012;44(11):1236–1242.

46. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43(7):e47.

47. Beekman R, Chapaprieta V, Russiñol N, et al. The reference epigenome and regulatory chromatin landscape of chronic lymphocytic leukemia. Nat Med 2018;24(6):868–880.

48. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. Nucleic Acids Res 2015;43(W1):W39–W49.

49. Castro-Mondragon JA, Riudavets-Puig R, Rauluseviciute I, et al. JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles. Nucleic Acids Res 2022;50(D1):D165–D173.

50. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials 1996;17(4):343–346.

51. Soma LA, Gollin SM, Remstein ED, et al. Splenic small B-cell lymphoma with IGH/BCL3 translocation. Hum Pathol 2006;37(2):218–230.

52. Ribera-Cortada I, Martinez D, Amador V, et al. Plasma cell and terminal B-cell differentiation in mantle cell lymphoma mainly occur in the SOX11-negative subtype. Mod Pathol Off J U S Can Acad Pathol Inc 2015;28(11):1435–1447.