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, Pathology Laboratory, 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 Zürich, Zürich, Switzerland; [6]European Molecular Biology Laboratory, Heidelberg, Germany; [7]Department of Hematology, Institut Català d'Oncologia, Hospital Germans Trias i Pujol, Josep Carreras Research Institute, Universitat Autònoma de Barcelona, Badalona, Spain; [8]Molecular Cytogenetics Laboratory, Department of Pathology, 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]Dipartimento Ematologia, Oncologia e Medicina Molecolare, Anatomia Istologia Patologica e Citogenetica, 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]Department of Pathology, Hospital Nostra Senyora de Meritxell, Escaldes-Engordany, Principat d'Andorra; [13]Hematology Laboratory, Hospital Universitari Bellvitge, Institut d'Investigació Biomèdica de Bellvitge (IDIBELL), L'Hospitalet De Llobregat, Spain; [14]Department of Hematology, Hospital Joan XXIII, Institut Català d'Oncologia, Tarragona, Spain; [15]Department of Hematology, 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, Toulouse, 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, USA; [26]Department of Pathology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA; [27]Department of Fonaments Clinics, Universitat de Barcelona, Barcelona, Spain and [28]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

[*]AC-M, FG, LW and MG contributed equally as first authors.
[#]FN and EC contributed equally as senior authors.

Correspondence: E. Campo
ecampo@clinic.cat

Received: March 23, 2022
Accepted: July 31, 2023.
Early view: August 10, 2023.

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

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

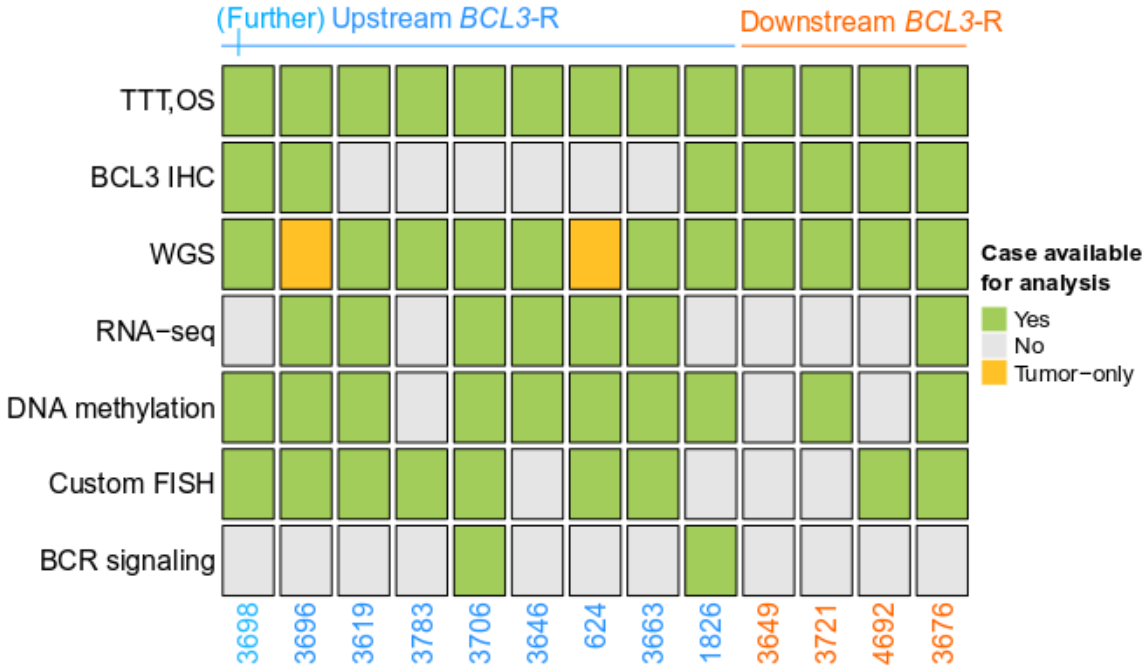
15	SMZL*	3’ downstream	Bright B-cell markers CD5- CD10- CD43- CD23- FMC7+ PAX5+ Cyclin D1- Bcl6-		Spleen:46,XX,add(14)(q32)[15]/46,XX[5] Lymph node: 45,XX,-14,der(19)(19q13.3→19p13.3::14q13→14q32::19p13:),der(19)(19p13.3→19q13::14q32)[11]/46,XX[5]
16	SMZL	3’ downstream	Bright B-cell markers CD5- CD23+ 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

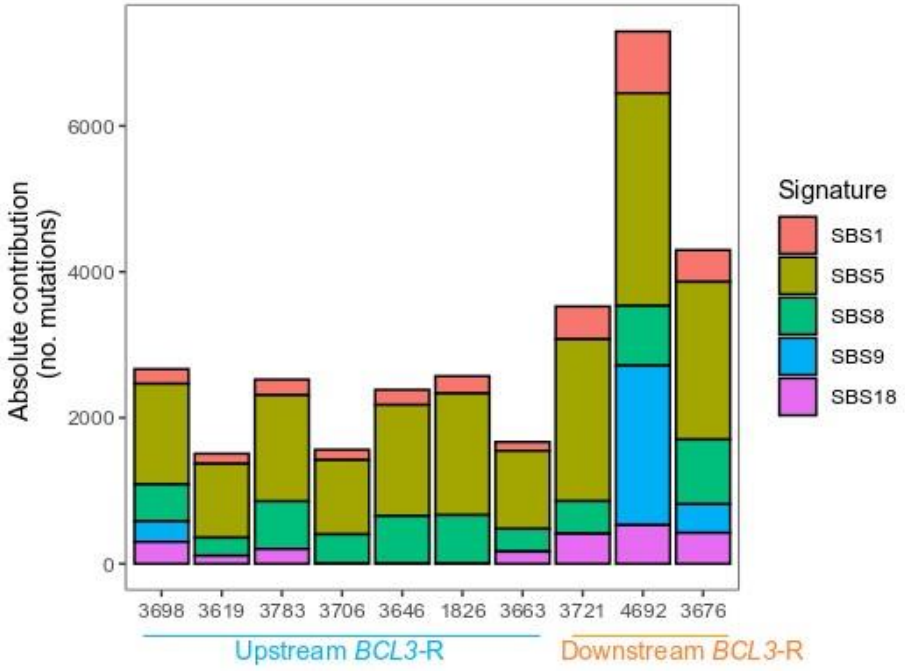
Supplementary Figure S1. Schema of the analyses performed in each *BCL3*-R tumor.

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.



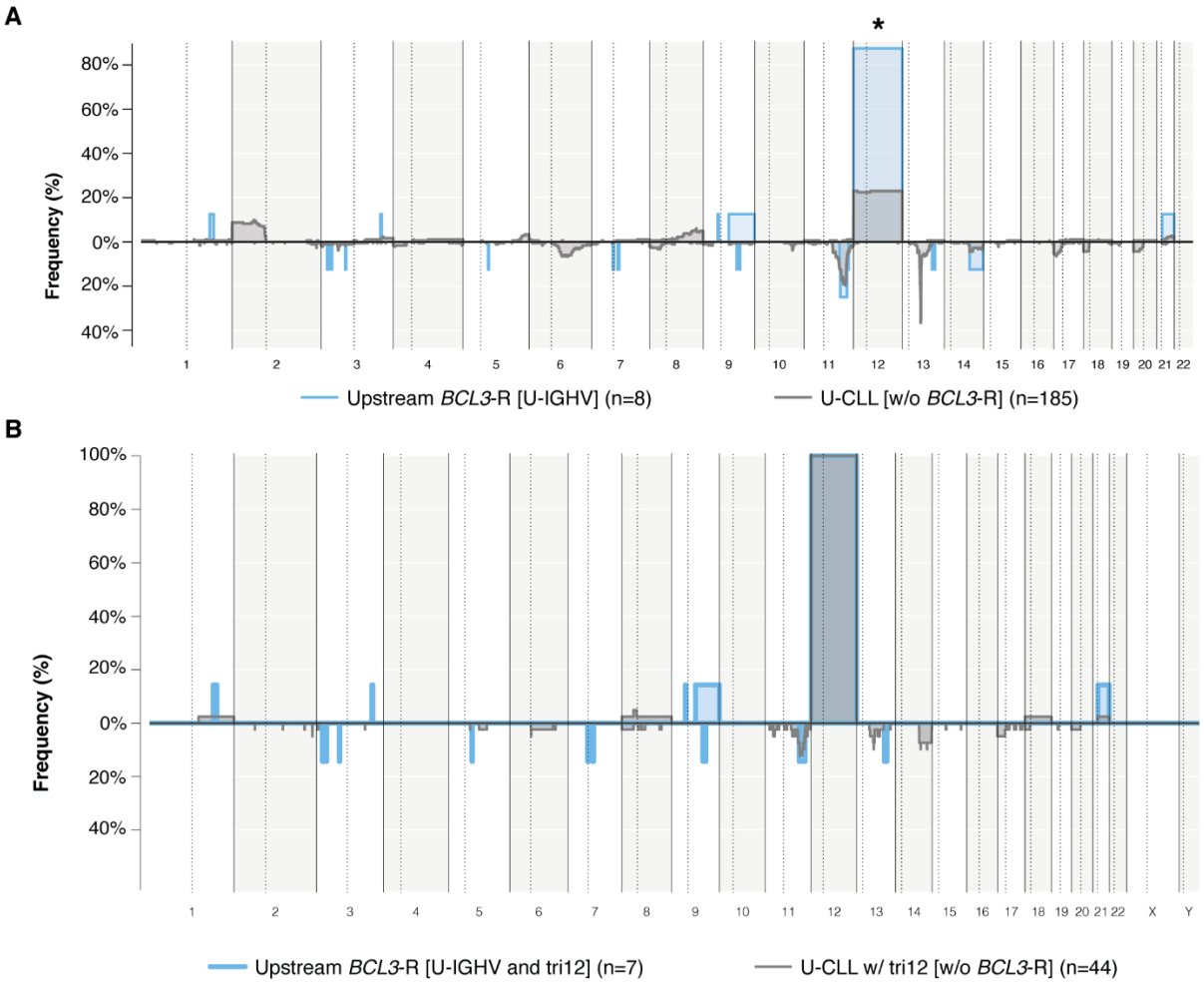
Supplementary Figure S2. Mutational signature analysis performed in *BCL3*-R tumors.

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



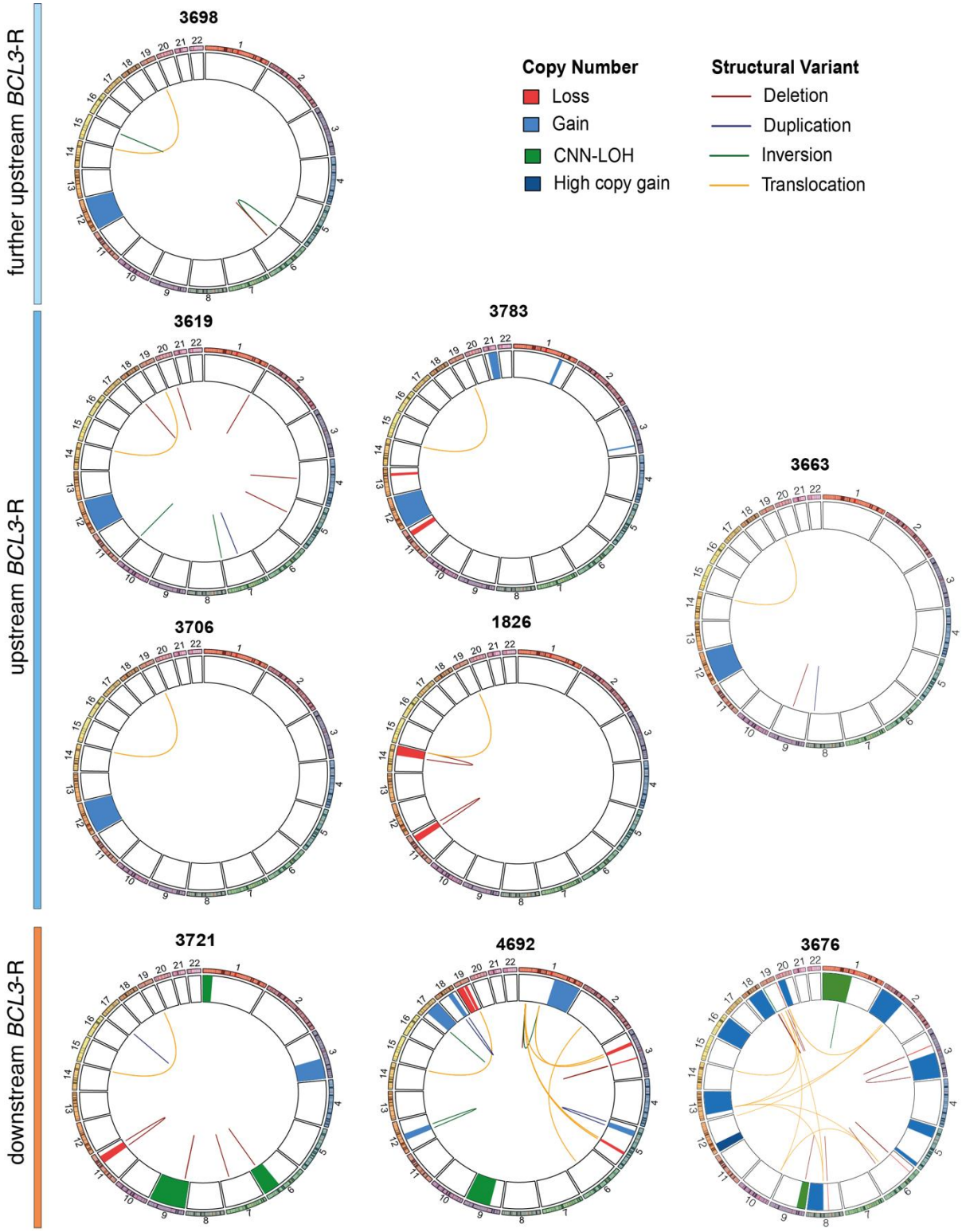
Supplementary Figure S3. Frequency of CNA in the upstream *BCL3*-R tumors vs CLL.

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.



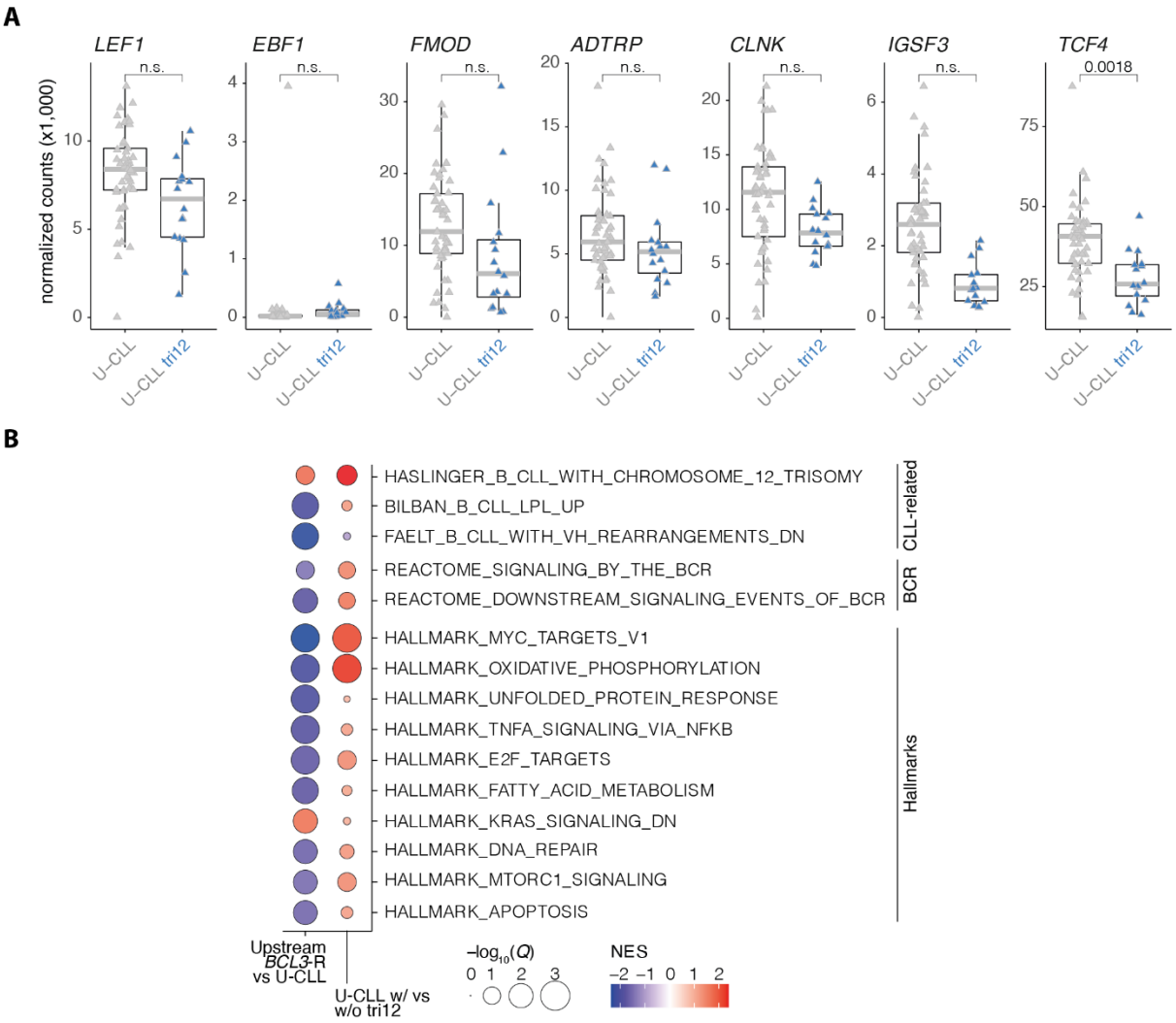
Supplementary Figure S4. Chromosomal landscape of *BCL3*-R tumors.

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.



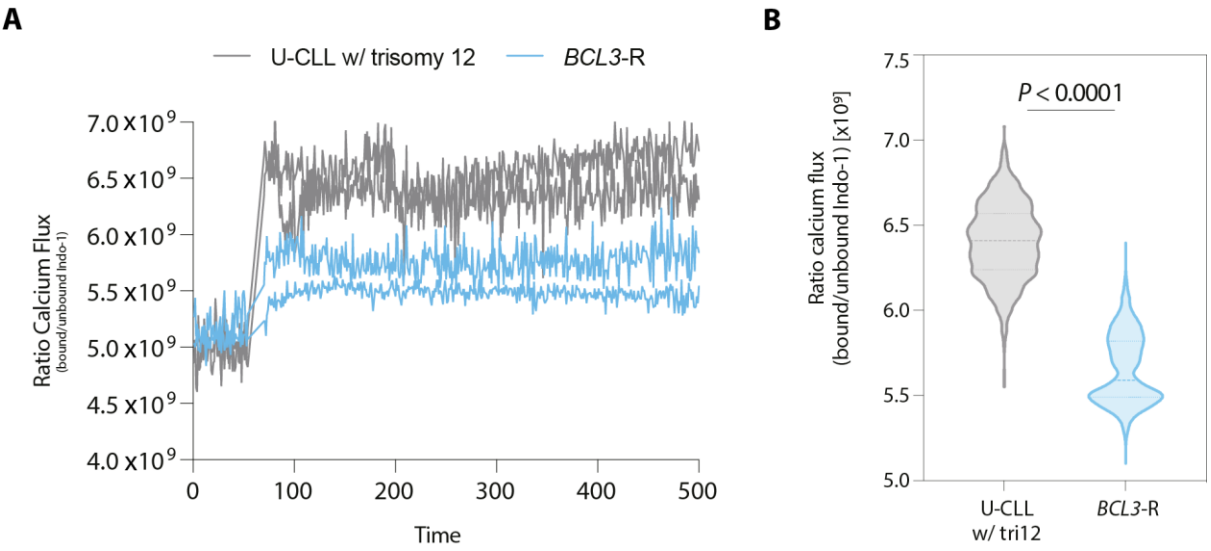
Supplementary Figure S5. Expression of CLL hallmark genes and GSEA.

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.



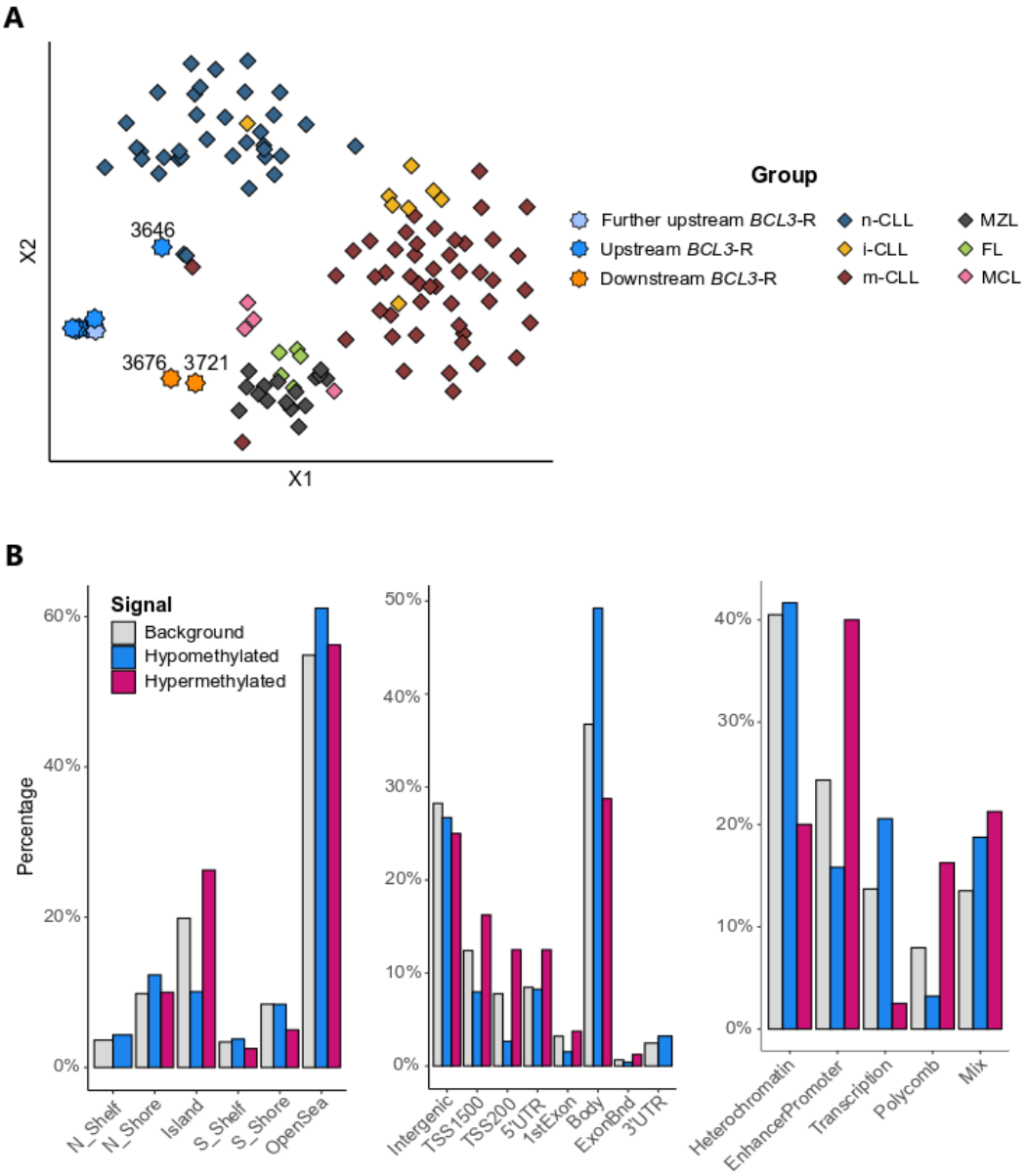
Supplementary Figure S6. Calcium flux of tumoral cells after BCR stimulation.

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.

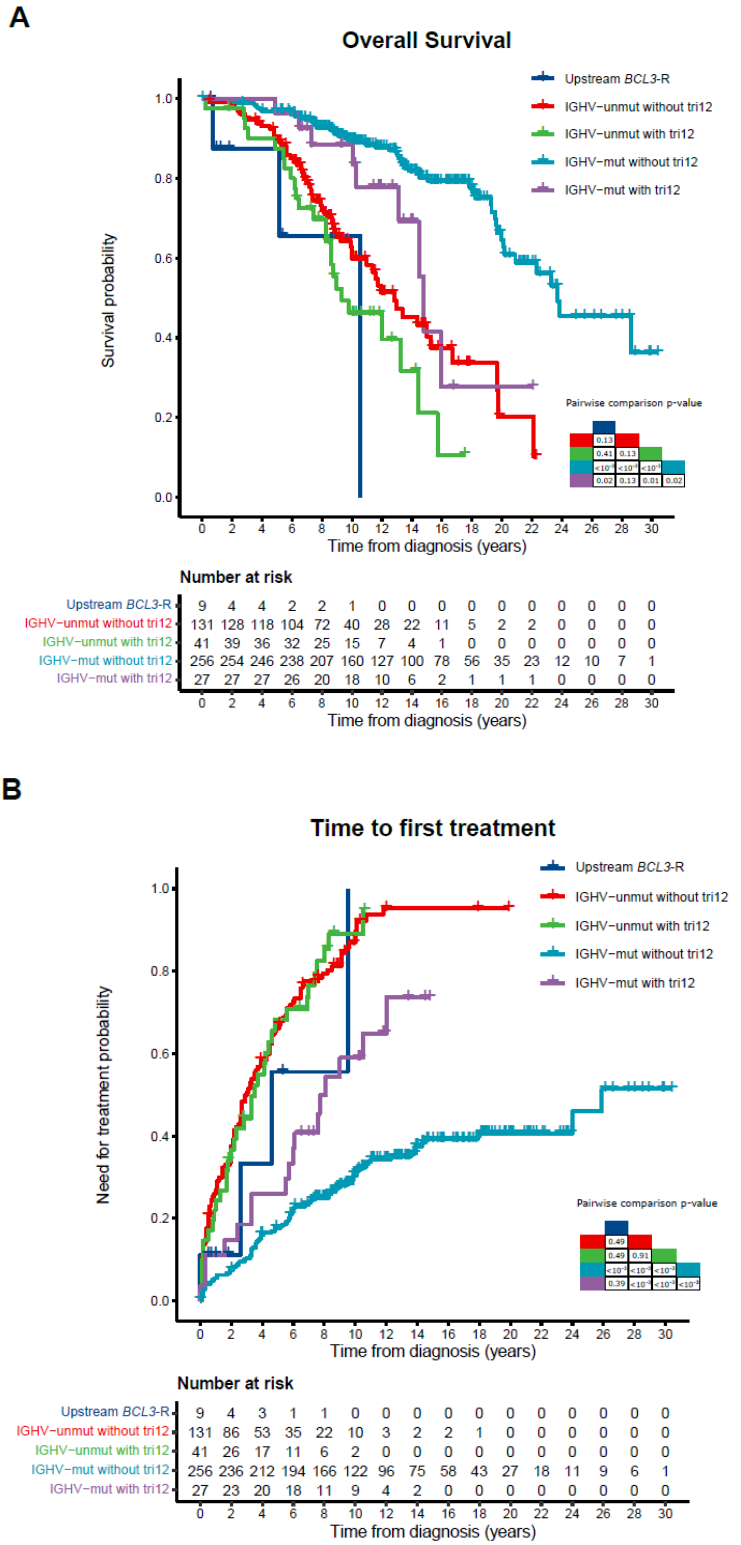


Supplementary Figure S7. T-distributed stochastic neighbor embedding analysis on the 795 DMCpGs and bar plots of their genomic location.

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).

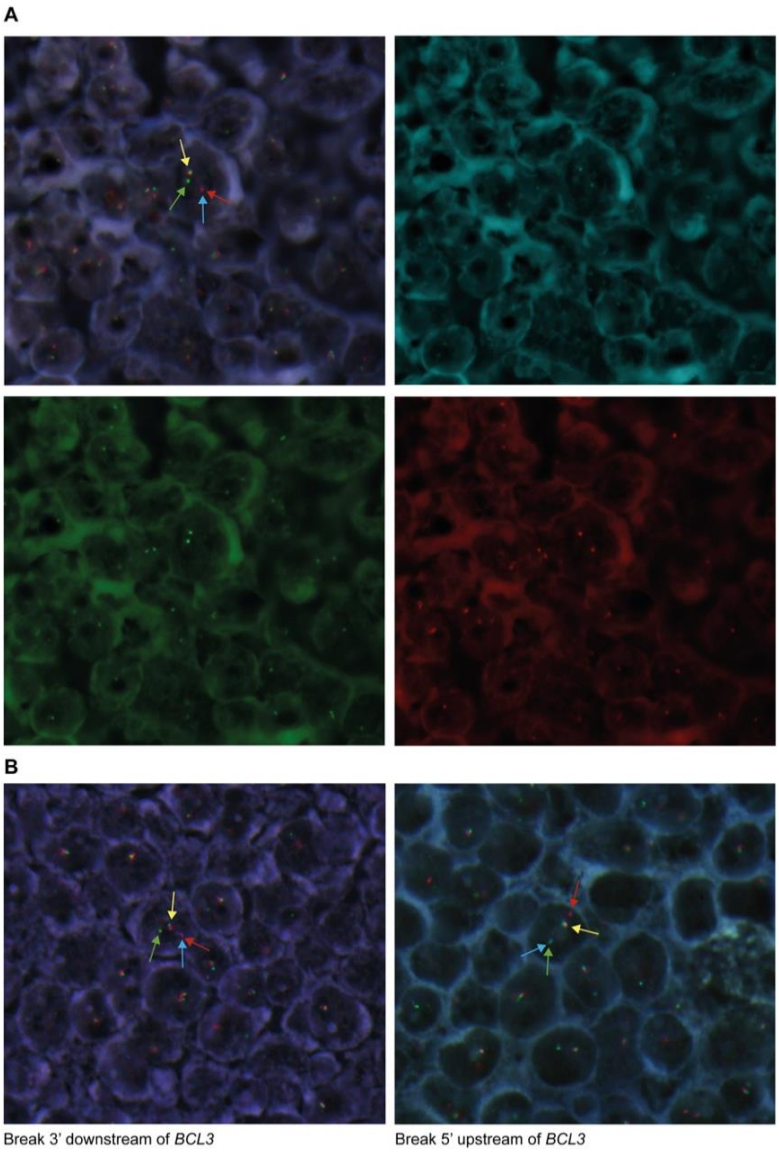


Supplementary Figure S8. Survival analysis between upstream *BCL3*-R and CLL. Comparison of A. overall survival and B. time to first treatment between upstream *BCL3*-R and CLL. CLL tumors have been stratified according to their IGHV mutational status and presence/absence of trisomy 12.

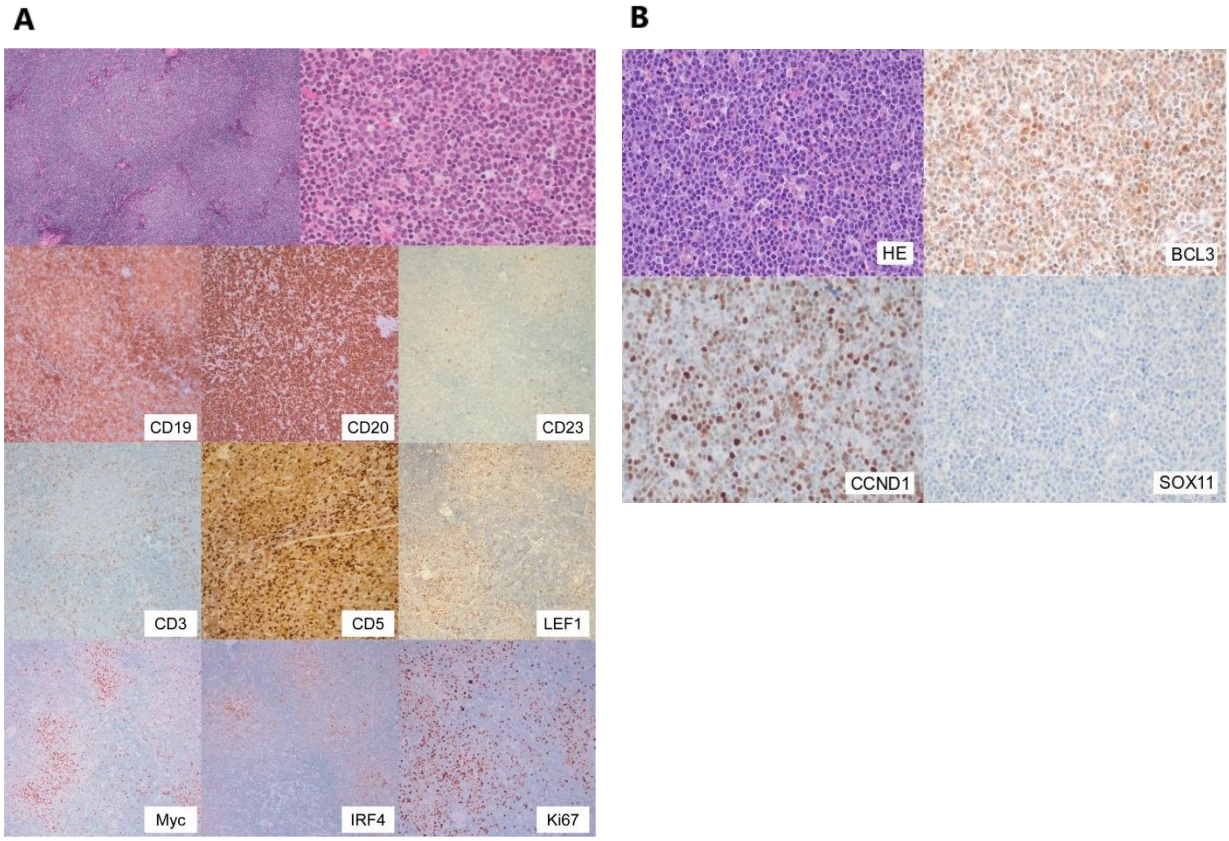


Supplementary Figure S9. FISH analysis of the 5' upstream and 3' downstream *BCL3* breakpoints using the custom FISH assay.

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.