Mutagenic mechanisms in normal and neoplastic B cells: from AID-induced diversification to genome-wide patterns

Sepúlveda Yáñez, J.H.

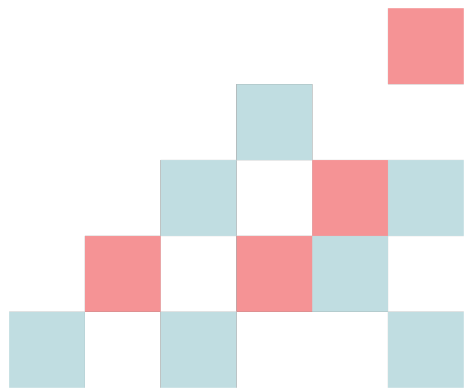
Isla Magdalena, Chile

CHAPTER 5

Autonomous B-cell receptor signaling and genetic aberrations in chronic lymphocytic leukemia-phenotype monoclonal B lymphocytosis in siblings of patients with chronic lymphocytic leukemia

Previously published as

Quinten, E., Sepúlveda-Yáñez, J. H., Koning, M. T., Eken, J. A., Pfeifer, D., Nteleah, V., de Groen, R. A., Saravia, D. A., Knijnenburg, J., Stuivenberg-Bleijswijk, H. E., Pantic, M., Agathangelidis, A., Keppler-Hafkemeyer, A., van Bergen, C. A. M., Uribe-Paredes, R., Stamatopoulos, K., Vermaat, J. S., Zirlik, K., Navarrete, M. A., Jumaa, H. & Veelken, H. Autonomous B-cell receptor signaling and genetic aberrations in chronic lymphocytic leukemia-phenotype monoclonal B lymphocytosis in siblings of patients with chronic lymphocytic leukemia. ***Haematologica*** **109.** Section: Articles, 824–834. doi:10.3324/haematol.2022.282542 (July 2023)



ABSTRACT

Clonal expansion of CD5-expressing B cells, commonly designated as monoclonal B lymphocytosis (MBL), is a precursor condition for chronic lymphocytic leukemia (CLL). The mechanisms driving subclinical MBL B-cell expansion and progression to CLL, occurring in approximately 1% of affected individuals, are unknown. An autonomously signaling B-cell receptor (BCR) is essential for the pathogenesis of CLL. The objectives of this study were functional characterization of the BCR of MBL in siblings of CLL patients and a comparison of genetic variants in MBL-CLL sibling pairs. Screening of peripheral blood by flow cytometry detected 0.2-480 clonal CLL-phenotype cells per microliter (median: 37/μL) in 34 of 191 (17.8%) siblings of CLL patients. Clonal BCR isolated from highly purified CLL-phenotype cells induced robust calcium mobilization in BCR-deficient murine pre-B cells in the absence of external antigen and without experimental crosslinking. This autonomous BCR signal was less intense than the signal originating from the CLL BCR of their CLL siblings. According to genotyping by single nucleotide polymorphism array, whole exome, and targeted panel sequencing, CLL risk alleles were found with high and similar prevalence in CLL patients and MBL siblings, respectively. Likewise, the prevalence of recurrent CLL-associated genetic variants was similar between CLL and matched MBL samples. However, copy number variations and small variants were frequently subclonal in MBL cells, suggesting their acquisition during subclinical clonal expansion. These findings support a stepwise model of CLL pathogenesis, in which autonomous BCR signaling leads to a non-malignant (oligo)clonal expansion of CD5+ B cells, followed by malignant progression to CLL after acquisition of pathogenic genetic variants.

5.1 | INTRODUCTION

Chronic lymphocytic leukemia (CLL) is a monoclonal expansion of more than 5000 mature CD5-expressing B cells per μL of peripheral blood. [259, 260] The clinical behavior of CLL is highly variable and depends strongly on structural and functional properties of the BCR as well as the presence of various recurrent genetic aberrations. [261] Autonomous signaling of the clonotypic BCR expressed by CLL cells, i.e. signaling without engagement of external antigen, is an indispensable oncogenic signal in CLL. [84] This peculiar property of CLL BCR is absent in other types of indolent B-cell lymphoma and in antigen-specific normal B cells. [84] Autonomous BCR signaling is caused by physical interaction between neighboring BCR complexes on the cell surface as demonstrated by seminal crystallographic studies in paradigmatic CLL BCR. [262] In the Eμ-TCL1 transgenic mouse model, autonomous BCR signaling is indispensable for CLL development even when all B cells primarily express an antigen-specific transgenic BCR and are stimulated by their cognate antigen. [85] CLL is characterized by remarkable skewing in the BCR immunoglobulin (BCR IG) gene repertoire, culminating in the existence of subsets of cases with restricted BCR IG features. [263] The largest immunogenetic CLL subset, stereotyped subset #2L, is defined by expression of a BCR with a light chain utilizing the IGLV3-21 gene that carries a G110R mutation at the junction between the variable and constant domains. [86] This seminal IGLV3-21 G110R mutation facilitates autonomous BCR signaling. [262] Numerous gene polymorphisms predispose to the development of CLL. [264–266] Formally, CLL develops through a precursor stadium, an subclinical expansion of mature B cells with the characteristic CD5+CD20lowCD79low CLL phenotype, commonly referred to as CLL-phenotype monoclonal B lymphocytosis (MBL). [76] The prevalence of MBL is approximately 100 times higher than CLL and is relatively increased in siblings of CLL patients. [77, 261] Similar to cases of indolent CLL, MBL cells carry relatively few genetic aberrations. [78, 267] However, the mechanisms leading to benign, longitudinally stable clonal expansion of CLL-phenotype B cells, and the hierarchy and sequence of events causing eventual progression to overt CLL are not fully elucidated yet. In order to investigate the early stages in the ontogeny of CLL, we screened siblings of CLL patients for the presence of MBL with CLL phenotype. Specifically, we sought to clarify whether these MBL cells express BCR with autonomous signaling capacity. In addition, we compared the prevalence of inherited risk loci and acquired genetic aberrations between CLL patients and their first-degree MBL relative.

5.2 | MATERIALS AND METHODS

PATIENTS, PROBANDS, AND SAMPLES

Siblings of CLL patients were asked to provide a single blood sample with informed consent. The study was conducted according to the Declaration of Helsinki and was approved by the Ethical Committees of both participating institutions (Leiden: P12.059; Freiburg: 44/08).

DETECTION, ISOLATION, AND PROCESSING OF CLL-PHENOTYPE CELLS

MBL was detected as discrete CD19+CD5+CD20lowCD79low populations (Figure 5.1) by six-color flow cytometry. DNA and RNA were isolated from MBL cells sorted to >95% purity (Allprep DNA/RNA mini kit; Qiagen, Hilden, Germany). Frequencies of MBL cells were calculated from gated cells and absolute lymphocyte counts.

IDENTIFICATION OF CLONAL BCR TRANSCRIPTS

VDJ and VJ sequences were determined by ARTISAN PCR. [268, 269] Ig sequences were analyzed by Geneious 10.2.6 (Geneious, Auckland, New Zealand) and IMGT/V-QUEST. [270, 271] Assignment to CLL stereotypes was performed by an established purpose-built algorithm. [263, 272]

ANALYSIS OF BCR SIGNALING ACTIVITY

Murine TKO pre-B cells were transduced with retroviral pMIZCC and pMIZYN vectors encoding paired Ig heavy and light chain sequences from individual cases. [84, 86] TKO cells are genetically deficient of ***Rag2, Lambda5,*** and ***Slp65***. ***Slp65*** function is reconstituted by a 4-hydroxytamoxifen-inducible ***Slp65-ERT2*** fusion gene. [74, 84, 273] Calcium mobilization was measured in Indo-1 AM-loaded (ThermoFisher, Waltham, MA), live-gated TKO cells by flow cytometry prior and after addition of 2 µM 4-hydroxytamoxifen (Sigma Aldrich, St Louis, MO). Maximum BCR signaling was measured after addition of crosslinking anti-IgM or anti-IgG antibodies (clones 2022-01 and 2042-01; Southern Biotech, Birmingham, AL). Autonomous BCR signaling was quantitated with correction for totally unresponsive cells during BCR crosslinking (supplementary material) at least twice at approximately 3 and 4 weeks after transduction.

GENETIC CHARACTERIZATION

Variants in 52 B-cell lymphoma-related genes were analysed by the LYMFv1 custom targeted sequencing panel on a Ion S5 system (ThermoFisher). [274] Sequencing reads were aligned to the GRCh37/hg19 human reference genome with TMAP 5.0.7 software under default parameters. CLL-associated copy number variations (CNV) were detected by SNP arrays. [275] Whole exome sequencing (WES) was performed with SureSelect Human All Exon V7 kit (Agilent, Santa Clara, CA) capture on the HiSeq2000 (Illumina, San Diego, CA) platform to an average coverage of 50x. If necessary, DNA was amplified by isothermal alkaline genome amplification (REPLI-g kit; Qiagen). Detected variants were annotated for predicted pathogenicity and occurrence in both CLL and MBL siblings (supplementary material). The global distribution of variant allele frequencies in CLL and MBL was estimated using the Wilcoxon rank sum test with continuity correction (Mann–Whitney U test) function 'wilcox.test()' in R. Monoallelic and subclonal variants present in CLL and MBL samples from siblings were tested using the Wilcoxon signed-rank test function 'wilcox.test(paired=TRUE)'. Both tests were only per-

formed after checking the corresponding test assumptions. Genotyping of inherited CLL risk loci was performed by targeted Sanger sequencing of 24 CLL susceptibility alleles genes and loci covered by WES and SNP array. [266, 276–280] Observed frequencies of risk alleles were compared to the most appropriate reference population from the gnomAD database by Fisher exact test. [281] CLL risk SNP were compared with WES-based data of Northwestern Europeans in gnomAD v2.1.1 and with WGS-based data of non-Finnish Europeans in gnomAD v3.1.2. Polygenic risk scores (PRS) were calculated for individual MBL probands and CLL patients for their genotyped risk loci based on the most recent reported odds ratios. [265, 266] The extent of genetic analyses was restricted in individual cases by insufficient amounts of DNA due to the paucity of clonal CLL-phenotype B cells.

5.3 | RESULTS

PREVALENCE AND FREQUENCY OF CLL-PHENOTYPE CELLS IN SIBLINGS OF CLL PATIENTS

Peripheral blood samples from 191 siblings of CLL patients were received and analyzed at two academic medical centers. Flow cytometry revealed a discrete population of CLL-phenotype CD19+CD5+CD20lowCD19low cells, in the majority with evident light chain restriction, in 34 siblings (17.8%) of 26 CLL patients (Figure 5.1, Supplementary Table 1). The median age of siblings with CLL-phenotype cells was 68 years (range: 43-80 years). The absolute number of CLL-phenotype MBL cells ranged from 0.2 to 1863 per µL blood, and 32 siblings (94%) were classified as low-count MBL with less than 500 clonal CLL-phenotype cells per µL (Figure 5.2A). [282]

SEQUENCE CHARACTERISTICS OF CLONALLY DOMINANT BCR IG OF CLL-PHENOTYPE CELLS IN CLL PATIENTS AND THEIR HEALTHY SIBLINGS

BCR IG transcripts from highly purified CLL-phenotype cells were sequenced in 17 siblings of CLL patients. In two siblings, two different productive VDJ gene rearrangements were identified within the CLL-phenotype MBL cell population (MBL01.1, MBL04.1); all other siblings carried an apparently monoclonal MBL population. Five MBL clones in healthy siblings could be assigned to CLL stereotyped subsets according to expanded immunogenetic criteria (Supplementary Table 1). [263, 272] In particular, the BCR of MBL07.1 and MBL27.2 belonged to CLL subset #2 and #2L, respectively, based on expression of a BCR light chain utilizing the IGLV3-21 gene with the G110R mutation. [86] An IgG-expressing MBL fulfilled the criteria of CLL subset #4. [283] These MBL BCR were predicted to exert autonomous signaling based on structural analyses. [262] Furthermore, all sequenced MBL BCR contained one or both of the FR2 VRQ and FR3 YYC motifs proposed as structural requirements for autonomous BCR signaling in the majority of CLL cases. [84, 284] The somatic mutation status of the clonotypic rearranged IGHV gene did not correlate to the degree of expansion of CLL-phenotype cells (Figure 5.2B). In addition, IGHV gene mutational status could differ between

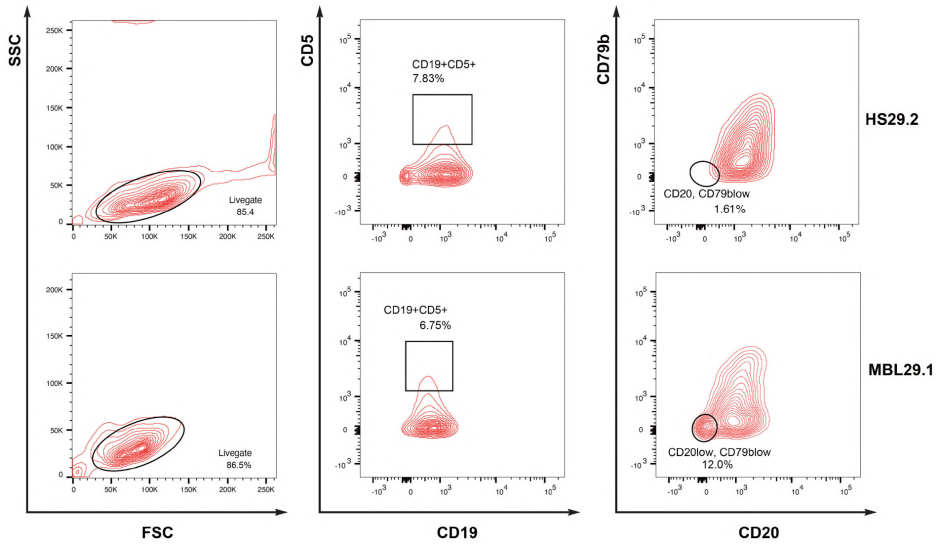


Figure 5.1. Detection of CLL-phenotype MBL cells. Flow cytometry screening of two healthy siblings of CLL patient 29.1. Upper row: Healthy sibling HS29.2 without detectable CLL-phenotype cells. Lower row:: Healthy sibling MBL29.1 with the lowest detected expansion of CLL-phenotype cells (CD19+CD5+CD20lowCD79low) in this study. SSC: Side scatter. FSC: Forward scatter.

CLL patients and their siblings; in particular, three non-subset 2L MBL (17.6%) expressed unmutated BCR IG whereas the corresponding CLL BCR was mutated (Figure 5.2C).

FUNCTIONAL CHARACTERISTICS OF MBL BCR FROM SIBLINGS OF CLL PATIENTS

BCR from clonally dominant MBL cells of 11 siblings of CLL patients, including all five MBL assigned to a CLL stereotyped subset, were expressed in TKO cells and tested for calcium mobilization prior and after tamoxifen-induced reconstitution of the BCR signaling cascade and after BCR crosslinking. As predicted, [86] both CLL subset #2L MBL BCR showed robust calcium flux upon addition of 4-hydroxytamoxifen without antigen exposure or BCR crosslinking (Figure 5.3). In addition, all other tested MBL BCR from siblings of CLL patients had autonomous BCR signaling activity. On average, autonomous BCR signaling was slightly stronger in CLL compared to MBL in each individual sibling pair ((Figure 5.4A). There was no quantitative correlation between BCR signaling activity of CLL and MBL sibling pairs ((Figure 5.4B). BCR signaling strength was positively correlated with MBL cell counts at the time of sampling (Figure 5.4C).

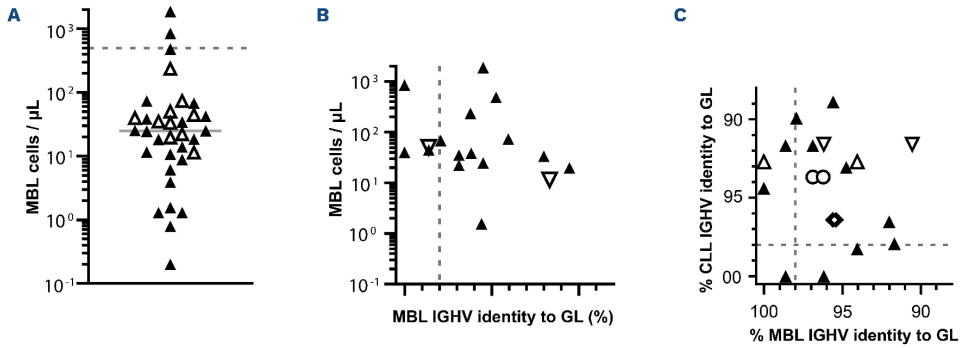


Figure 5.2. Characteristics of MBL cells in CLL siblings. A. MBL counts of all 34 siblings with a detectable CLL-phenotype B-cell population (CD19+CD5+CD20lowCD79low). Solid line: median. Dotted line: Commonly accepted threshold between low- and high-count MBL. Open triangles: MBL selected for functional BCR testing. B. IGHV gene SHM status (x axis) and counts of MBL cells (y axis). Dotted line: Commonly accepted threshold between mutated and unmutated BCR status in CLL. Red symbols: MBL cells from two CLL siblings expressing a CLL subset 2 BCR predicted to exert autonomous signaling.10 GL: Germ line. C. IGHV gene identity to germline in MBL cells (x axis) and CLL cells (y axis) in siblings. Dotted lines: Commonly accepted boundary between mutated and unmutated BCR status in CLL. Red and orange symbols: Bona fide biclonal MBL based on expression of two immunoglobulin heavy chains (MBL01.1, MBL04.1). Green and blue symbols: CLL patients (CLL24.1, CLL30.1) with two siblings with MBL, respectively (MBL24.1, MBL24.2; MBL30.1; MBL30.2). GL: Germ line.

COMPARATIVE GENETICS OF CLL PATIENTS AND MBL SIBLINGS

Analysis of known germ line variants from up to 26 different gene loci associated with increased incidence of CLL revealed a significantly higher prevalence of twelve risk alleles within our cohort (Supplementary Table 2). A polygenic CLL risk score composed of 24 risk loci was higher in both CLL and MBL sibling than the average reference score calculated from reference population data (Figure 5.5C). [281] There was no evidence for a higher PRS in CLL cases in a matched-pair comparison to their MBL siblings. Taken together, these findings suggest that CLL risk loci predispose to clonal expansion of CLL phenotype cells in both low-count MBL and CLL. Vice versa, we did not find evidence for a particular association of CLL risk alleles with malignant progression from MBL to clinical CLL. Fifteen MBL samples with sufficient DNA were genotyped by SNP arrays to detect recurrent CLL-associated CNV. Ten MBL carried recurrent CLL-associated CNV, including a del(17)(p13-q21) in one case and del(13)(q14) in nine cases. One additional MBL carried large areas of loss of heterozygosity, and another sample a del(14)(q22.2) in combination with a trisomy 18q12, resulting in a total of eleven MBL samples with informative CNV (Table 1). Eleven of 15 genotyped CLL cases within this study carried informative CNV, suggesting a similar overall prevalence of CNV between CLL cases and MBL in their siblings. Quantitative analysis of the SNP array data indicated that all malignant cells harbored the respective CNV in nine of eleven CLL samples. In contrast, individual CNV were present in only a minor fraction of the highly

purified MBL cells in ten of eleven samples (p=0.0019; Fisher exact test). We corroborated these quantitative CNV data with analysis of small variants, i.e. SNV and indels from WES in ten CLL-MBL sample pairs. Identical variants detected in a CLL case and its MBL sibling were presumed to be inherited and were considered noninformative with respect to differential pathogenesis between MBL and CLL. In contrast, variants that were not shared between siblings, i.e. variants observed exclusively in either a CLL sample or its MBL sibling, could either be inherited or acquired somatically and were considered as potentially informative on differential MBL-CLL pathogenesis. Initially, we focused on 120 genes recurrently mutated in CLL according to the COSMIC database. As determined by targeted panel sequencing and/or WES analysis in ten CLL-MBL sample pairs, CLL and MBL carried similar numbers of unique gene variants with predicted pathogenic relevance. Variants in these genes were infrequent, and their variant allele frequency (VAF) did not differ significantly between CLL and MBL (Figure 5.5B).

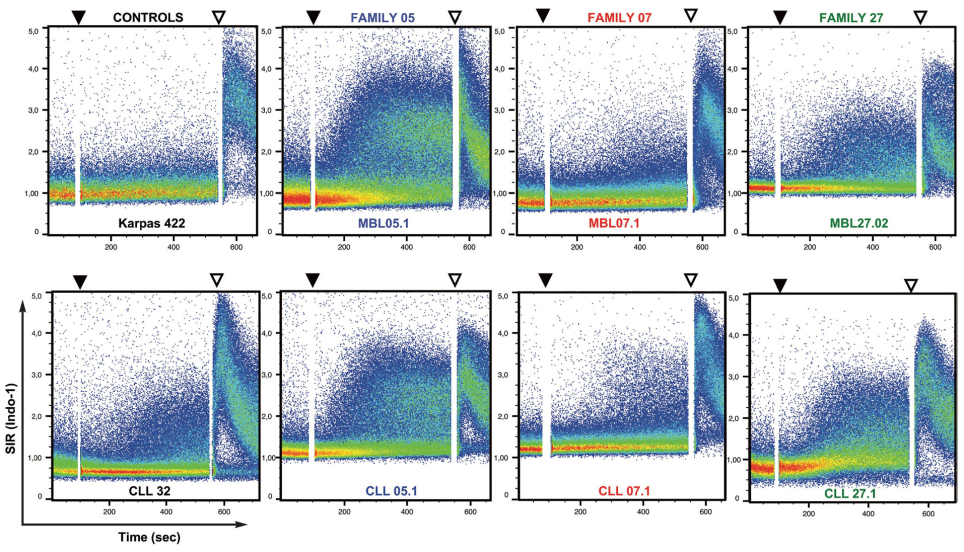


Figure 5.3. Autonomous BCR signaling of CLL and MBL cells in three representative sibling pairs. Calcium mobilization was measured by flow cytometry as Indo-1 AM signal intensity ratios of BCR-transduced TKO cells. Black triangle: Time point of addition of 4-hydroxytamoxifen to reconstitute a functional BCR signaling cascade. White triangle: Time point of addition of anti-IgM/IgG crosslinking antibody. K422: TKO cells transduced with BCR of lymphoma cell line Karpas 422 as negative control. CLL 32: TKO cells transduced with BCR of CLL 32 as positive control.7 MBL05.1: TKO cells transduced with BCR of MBL cells in a sibling expressing a CLL subset #60 BCR.5, 23 MBL07.1, MBL27.2: TKO cells transduced with BCR of MBL cells in siblings expressing a CLL subset #2 BCR. TKO cells transduced with CLL BCR are depicted below the MBL of their respective siblings.

Genome-wide, VAF of non-shared variants were significantly lower in MBL samples than in CLL cases ((Figure 5.7A). In order to test whether this difference could be attributed to a higher fraction of subclonal variants in MBL compared to CLL, we separately compared in

sibling pairs the numbers of non-shared monoallelic variants (VAF 0.43-0.55), presumably carried by all cells in a given sample, and numbers of monoallelic subclonal variants (VAF 0.1-0.33). In accordance with this hypothesis, monoallelic variants were detected at similar frequencies in CLL samples and their respective sibling MBL samples (Figure 5.7B). In contrast, the prevalence of subclonal variants was higher in MBL cells compared to their sibling CLL counterpart (Figure 5.7C).

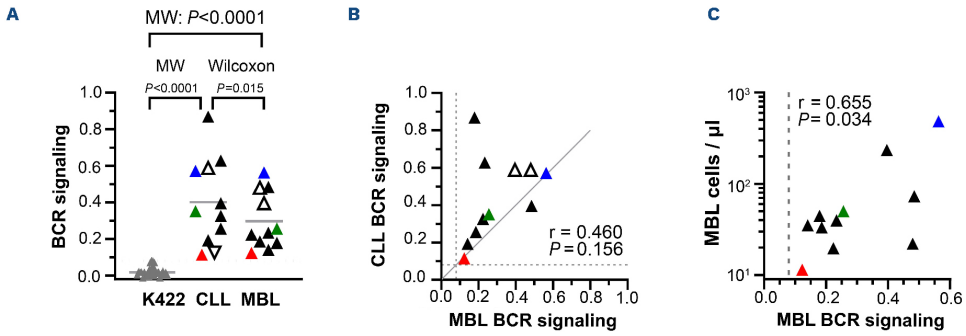


Figure 5.4. Quantitative analysis of autonomous BCR signaling of CLL and MBL cells. A. Autonomous BCR signaling strength in TKO cells transduced with clonal BCR. K422: All measurements of Karpas 422 BCR (negative control) performed during this study. Open triangle: CLL 32 (positive control). Orange triangles: A CLL with two MBL siblings. Other triangle colors correspond to Figure 3. Black triangles: All remaining tested CLL and MBL cases. M-W: Unpaired comparison by Mann-Whitney test. Wilcoxon: Paired comparison between CLL and MBL within siblings by Wilcoxon matched-pair signed rank test. B. Correlation of autonomous BCR signaling strength between CLL and MBL siblings. Colors correspond to Figure 3 and 4A. Dashed lines indicate maximum BCR signaling strength observed in Karpas 422 negative controls. C. MBL counts (y axis) according to autonomous BCR signaling strength (x axis). Colors correspond to Figure 3. Dashed line indicates maximum BCR signaling strength observed in Karpas 422 negative controls.

5.4 | DISCUSSION

This systematic study was undertaken to clarify whether the presence of an autonomous BCR signal could explain the different clinical behavior between non-malignant clonal expansions of cells with CLL phenotype (MBL) and clinically malignant CLL. In order to minimize potential confounding effects of inherited germ-line variants, we investigated this hypothesis in a comparative manner in siblings of CLL patients. Detection limit and range of clonal B cells with a CLL phenotype were very similar to a recently published study on members of CLL families. [261] However, the observed MBL prevalence of 17.8% was relatively high compared to previous reports on CLL siblings and in the range of individuals belonging to CLL families defined by more than one first-degree relative diagnosed with CLL. [77, 261]

Unbiased amplification of expressed heavy and light IG chain gene rearrangements revealed single dominant BCR in 15 and potential biclonality in two MBL subjects. Potential MBL oligoclonality has been recognized in MBL, especially in the low-count subtype. [285, 286]

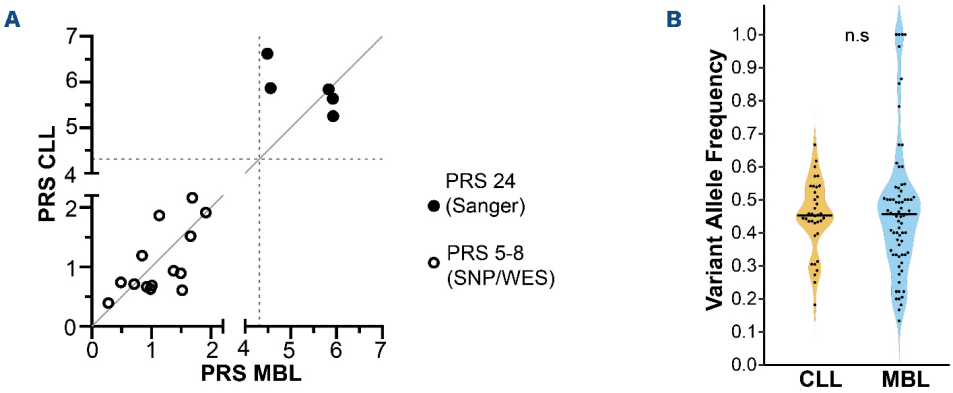


Figure 5.5. Comparison of chronic lymphocytic leukemia-associated variants in monoclonal B lymphocytosis-chronic lymphocytic leukemia siblings. (A) Polygenic risk score (PRS) for chronic lymphocytic leukemia (CLL) in probands with monoclonal B lymphocytosis (MBL) (x axis) and their CLL sibling (y axis). Depending on available data, the PRS was calculated for 24 CLL risk loci (black dots) or five to eight loci as assessed by single nucleotide polymorphism array and/or targeted Sanger sequencing (open circles). Dashed lines indicate the mean PRS of the reference European population for the 24 loci analyzed. (B) Distribution of variant allele frequencies of non-shared variants in 120 CLL driver genes as detected by whole exome sequencing in seven CLL cases and ten MBL siblings from seven families. CLL and MBL were compared by the Wilcoxon signed rank test. SNP: single nucleotide polymorphism; WES: whole exome sequencing; n.s.: not significant.

5

The identification of a fraction of MBL with unmutated BCR IG in our series is also in line with a published study. [286] Whereas low-count MBL was reported to belong only rarely to any immunogenetically defined CLL stereotype subset, [263, 286] we encountered five instances of BCR IG from low-count MBL (33.3%) that could be assigned to such CLL subsets. [263, 272] It remains hypothetical whether this apparent discrepancy is due to focusing on MBL carriers with CLL siblings in our study. In healthy persons who eventually developed CLL, app. 20% of dominant clonotypes identified during the prediagnostic period could be assigned to CLL stereotypes. [287]

With respect to the primary aim of our study, MBL BCR IG resembling certain CLL subsets were predicted to have autonomous BCR signaling activity akin to CLL. This prediction especially applied to the herein first reported low-count MBL expressing an isotype-switched BCR of CLL subset #4, and two subset #2L MBL cases expressing an IGLV3-21 light chain with the seminal G110R mutation. Both of these structural BCR IG characteristics mechanistically cause autonomous BCR signaling by facilitating steric interactions between neighboring BCR complexes on the cell surface. [86] Limited by the sample size, we found no evidence for correlations in structural features of the BCR, e.g. mutational status or IG gene usage, between CLL-MBL sibling pairs.

As already suggested by the MBL BCR IG sequences, our findings establish unequivocally by quantifiable functional testing that the BCR of MBL, i.e. a premalignant clonal expansion of

Family	MBL		CLL	
	Case	Aberration	Case	Aberration
1	MBL01.1	13q14.2q14.3x0[0.7]	CLL01.1	13q14.2q14.3x1
2	MBL02.1	13q14.2x1[0.6] 13q14.2q14.3x0[0.6] 13q14.3x1[0.6]	CLL02.1	13q14x1 8q21.3qterx3 15q26.1q26.3x1
3	MBL03.1	no CNV	CLL03.1	2p22.1p16.1x1 8q24.22x1
4	-	-	CLL04.1	no CNV
5	MBL05.1	no CNV	CLL05.1	13q14.13q14.2x1[0.5] 13q14.2q14.3x0[0.5] 13q14.3x1[0.5]
6	MBL06.1 MBL06.2	13q14.2q14.3x1 14q22.2q32.12x1[0.6] 18q12.3q23x3[0.6]	CLL06.1	14q23.2q32.13x1
7	MBL07.1 MBL07.2	13q14.2q14.3x1[0.15] 13q14.13q31.1x1[0.15]	CLL07.1	13q14.11q14.2x1
8	MBL08.1	13q14.2q14.3x1[0.85]	CLL08.1	5q33.1q35.3x3[0.3] 6q14.1q27x1 8q23.1q24.3x3 17p13.3p12x0[0,5]
9	MBL09.1	no CNV	CLL09.1	no CNV
10	MBL10.1	13q14.2q14.3x1[0.5]	CLL10.1	13q14x1
11	MBL11.1	13q14.2q14.3x1[0.7]	CLL11.1	no CNV
12	MBL12.1	13q14.2q14.3x0[0.5]	CLL12.1	13q14.2q14.3x1
13	-	-	CLL13.1	11q22.3q25x1 3q26.1q29x3 4p16.3p14x1 17q21.32q25.3x3
15			CLL15.1	no CNV
16	MBL16.1	2q36.3q37.3x1[0.7] 17p13.3q21.2x1[0.7]	CLL16.1	4p16.3p15.2x1[0.3] 4p14x1[0.3] 4p13x1[0.3] 4p13x1[0.3] 5p15.33p15.2x1[0.3] (11)cx[0.3~0.8] 13q14.11q14.2x1[0.5] 13q14.2q21.33x0[0.8] 13q21.33q22.3)x1[0.5]
Total analyzed	14		15	
With CNV	11		11	
With clonal CNV	1		9	
With subclonal CNV only	10		2	

MBL: monoclonal B lymphocytosis; CLL: chronic lymphocytic leukemia; CNV: copy number variation.

Figure 5.6. Copy number variations in monoclonal B lymphocytosis and chronic lymphocytic leukemia samples as detected by single nucleotide polymorphism array.

CD5-positive B cells with CLL phenotype, exerts antigen-independent, autonomous signaling. All known origins of autonomous BCR signaling in CLL were also encountered in our MBL cases, i.e. acquisition during primary V(D)J recombination in cases with unmutated BCR, acquisition by somatic hypermutation as exemplified best by the single G110R mutation of the IGLV3-21 gene, [86] and acquisition by class switch recombination. [262] Since our study specifically addressed CLL-MBL sibling pairs, we were able to compare the autonomous BCR signaling strength quantitatively in matched-pair fashion. This analysis revealed significantly lower intensity of BCR signaling in MBL compared to CLL. While autonomous BCR signaling is generally dependent on individual HCDR3 in conjunction with recurrent motifs in FR2 and interactions of defined amino acid restudies in some instances, [84, 86, 262, 284] quantitative

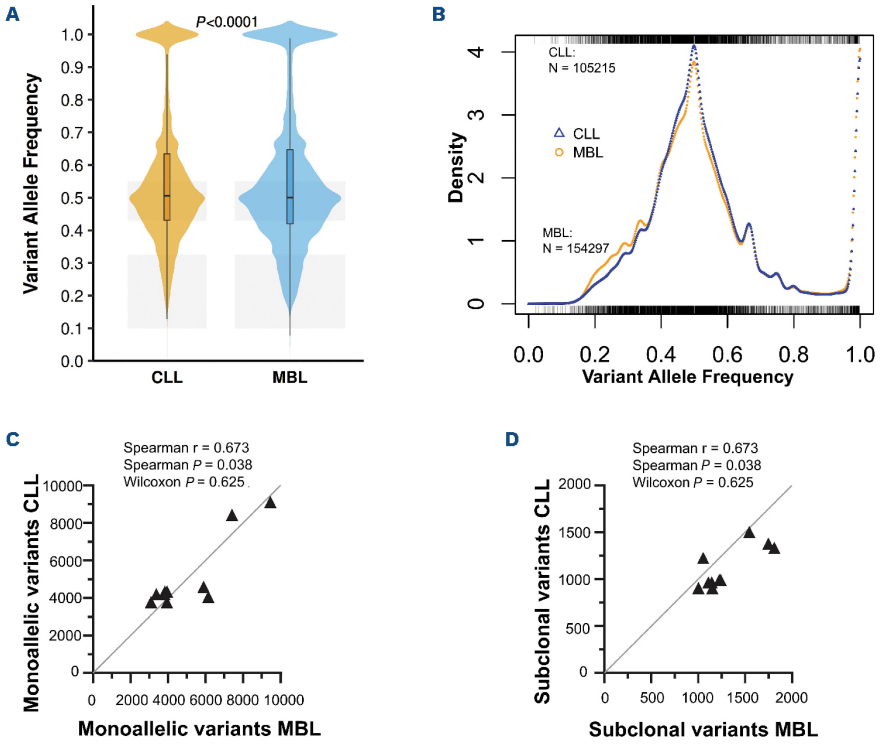


Figure 5.7. Comparison of genome-wide variants in siblings with monoclonal B lymphocytosis or chronic lymphocytic leukemia. (A) Global distribution of variant allele frequencies (VAF) in cases of chronic lymphocytic leukemia (CLL) and monoclonal B lymphocytosis (MBL). VAF were compared by the Wilcoxon rank sum test with continuity correction (Mann–Whitney U test). Shaded boxes indicate the ranges of VAF analyzed in (C) and (D). (B) Kernel density estimation distribution and rug plots of VAF in CLL and MBL. The rug plot marks along the x-axis (top and bottom) indicate positions of the VAF for CLL and MBL, respectively. A total of 105,215 CLL data points and 154,297 MBL data points were analyzed. (C) Numbers of monoallelic variants (VAF=0.43-0.55; present in all cells) of CLL and MBL samples from siblings. Correlations between CLL and MBL samples were calculated by the nonparametric Spearman correlation. CLL and MBL samples were compared by the Wilcoxon matched-pair signed rank test. (D) Numbers of subclonal variants (VAF=0.1-0.33) in CLL and MBL samples from siblings. Correlations between CLL and MBL samples were calculated by the nonparametric Spearman correlation. CLL and MBL samples were compared by the Wilcoxon matchedpair signed rank test. CLL: chronic lymphocytic leukemia; MBL: monoclonal B lymphocytosis.

differences in autonomous BCR signaling cannot be assigned yet to defined structural BCR characteristics. However, BCR signaling strength was directly correlated to the expansion of the MBL clone at the time of sampling. The novel observation requires validation in independent cohorts. Nevertheless, the quantitative differences of BCR signaling between CLL and MBL and in MBL expansion suggest a possible causal and differential role of BCR signaling in growth kinetics and subsequent clinical behavior of a neoplastic CLL-phenotype B-cell clone.

It would have been highly desirable to validate the correlation between BCR signaling strength and clonal B-cell expansion by longitudinal sampling. However, the approvals by the institutional ethics committees explicitly excluded longitudinal analyses and mandated to inform the healthy siblings with a categorical result of the screening. Since low-count MBL does not represent a definite pathological condition and rarely progresses to CLL,([261, 282] it was reported as absence of pathological findings. As a consequence, the ethical committees concluded that repeat sampling could potentially give rise to anxiety given the prior absence of pathology in the healthy siblings, and explicitly denied approval for longitudinal analysis.

In order to interpret BCR function in the context of genetic mechanisms in CLL pathogenesis, we addressed genetic susceptibility, genetic instability, and recurrent CLL-associated genetic changes in CLL cases and their MBL siblings. These studies were restricted by limited quantities of DNA extracted from the MBL samples. Nevertheless, analysis of CLL susceptibility loci indicated that both CLL and MBL probands have a similar PRS for CLL that is higher than in the general (Northern) European reference population from the gnomAD database. While these findings were predictable from the study design in CLL-MBL sibling pairs, on the other hand they do not provide any evidence that these CLL risk loci drive malignant progression rather than initial premalignant expansion of the CLL-phenotype clone. With respect to the possible inherited shared predisposition for MBL and CLL, it would have been highly desirable to strengthen this conclusion by showing lower PRS in the CLL siblings without detectable MBL cells. The study design, however, did not comprise informed consent for genetic analyses in non-MBL probands. Notwithstanding this limitation of our study, there appears no plausible causal relationship between the presence of any risk allele and the stochastic acquisition of a BCR with autonomous BCR signaling based on the known function of the genes located at the CLL risk loci.

The overall prevalence of CLL-associated CNV and gene variants was not evidently different between CLL and MBL in matched comparisons. In accordance with the literature, del(13)(q14) was the most abundant CNV in MBL, but the detected del(17)(p) has also been described occasionally. [288] Our observed low prevalence of potentially pathogenic variants in CLL driver genes is also in accordance with published data. [267] However, we noticed that CLL-associated CNV were frequently subclonal in MBL, suggesting that these aberrations might be gradually acquired at this stage. This observation is in accordance with findings from a longitudinal study in low-count MBL, wherein the categorical presence of CLL-associated CNV approximately doubled after a median interval of seven years compared to the primary sample. [289] The impression of gradual acquisition of mutations in MBL was independently supported by the observation that MBL cells carry more subclonal variants detected by next-generation sequencing than CLL, whereas the number of variants carried by all clonal cells did not differ between CLL and MBL. In CLL, the presence of subclonal drivers is associated with clinical outcome. [81]

In summary, our data demonstrate that autonomous BCR signaling operates in MBL in similarity to CLL. A hypothetical absence of this mechanism therefore does not explain the difference between benign MBL and malignant CLL. However, relatively low BCR signaling

strength in MBL may offer a plausible explanation for less aggressive expansion than in CLL. In addition, our data suggest that CLL risk loci likewise cannot account for the different clinical behavior between both conditions, and lend further support to the conclusion that prevalence and type of genetic pathogenic changes are indistinguishable between MBL and low-risk CLL. [267, 289] However, the observation of subclonal genetic CLL driver mutations in MBL supports a scenario of gradual clonal expansion driven by moderate autonomous BCR signaling, possibly leading to a certain degree of genetic instability that facilitates gradual acquisition of CLL driver mutations.

This conclusion supports the notion that MBL and CLL represent a spectrum of the same biological condition, wherein the cumulative stimuli from constitutive immune signaling through the BCR together with timing and type of eventually acquired genetic drivers govern the clonal dynamics from long-term balance between expansion and apoptosis over gradual expansion to clinically aggressive CLL.

5.5 | SUPPLEMENTARY MATERIAL

SUPPLEMENTAL METHODS

Patients, probands, and samples

CLL patients under management of the outpatient clinics at the participating centers were asked to inform their respective living siblings about this study and to provide their contact information. Approvals by the ethical committees explicitly excluded longitudinal analyses and mandated to provide siblings with a categorical screening result.

Detection, isolation, and processing of CLL-phenotype cells

Differential blood counts were obtained for each blood sample. DNA and RNA were directly isolated from sorted CLL-phenotype cells by spun-driven silica matrix columns (Allprep DNA/RNA mini kit; Qiagen, Hilden, Germany).

Identification of clonal BCR transcripts

VDJ and VJ sequences were determined by direct Sanger sequencing from gel-purified ARTISAN PCR products (Wizard SV gel and PCR clean-up kit; Promega, Madison, WI) or after molecular cloning (Topo TA cloning vector; ThermoFisher, Waltham, MA).

Analysis of BCR signaling activity

pMIZCC and pMIZYN expression vectors1 with inserted paired VDJ and VJ sequences from individual cases were synthesized (BaseClear BV, Leiden, The Netherlands) and packaged by liposomal cotransfection (FuGENE HD transfection reagent; Promega) of Phoenix-E cells with the pCL-eco helper vector. Murine TKO cells were transduced with virus-containing supernatant on fibronectin-coated plates at 48 h after transfection (RetroNectin; TaKaRa, Shiga, Japan). [86] TKO cells are arrested at the pro-B-cell stage by genetic deficiency of Rag2, Lambda5, and Slp65. Slp65 function is reconstituted by a ***Slp65-ERT2*** fusion gene that enables BCR signaling in the presence of 4- hydroxytamoxifen. [84, 273]

Calcium mobilization was measured in Indo-1 AM-loaded (ThermoFisher), live-gated TKO cells by flow cytometry as the ratio of signal intensities (SIR) at 405 and 485 nm for 90 sec. After addition of 2 μM 4-hydroxytamoxifen (Sigma Aldrich), calcium mobilization was measured for additional 7.5 min. Finally, maximum BCR signaling was induced by adding an unlabeled crosslinking anti-IgM or anti-IgG gamma antibody (clones 2022-01 and 2042-01; Southern Biotech, Birmingham, AL), and measurement was resumed for additional 3 min. Every BCR-transduced TKO population was measured at least twice at 3 approximately 3 and 4 weeks after transduction. The 2nd measurement was used for statistical analyses with two-tailed Mann-Whitney or Wilcoxon's rank sum tests for unpaired and paired comparisons, respectively, and Spearman's correlation coefficient.

Quantitative assessment of autonomous BCR signaling

The fraction of cells with a 405/485 SIR above the 95th percentile prior to addition of tamoxifen (Qaut) was determined with correction for totally unresponsive cells during BCR crosslinking. Subsequently, the calibrated median SIR observed in this Qaut fraction was determined (calSIRaut). Autonomous BCR signaling was quantitated as the arithmetic product of Qaut x

calSIRaut .

Details of recording (see Figure):

Measurements of BCR signaling activity is performed in 3 phases:

1. Baseline Phase (BP): Time interval from the start of measurements until the addition of 4-hydroxytamoxifen
2. Experimental Phase (EP): Time interval from addition of tamoxifen until the addition of cross-linking agent
3. Maximum Phase (MP): Time interval from addition of cross-linking agent to the end of measurement

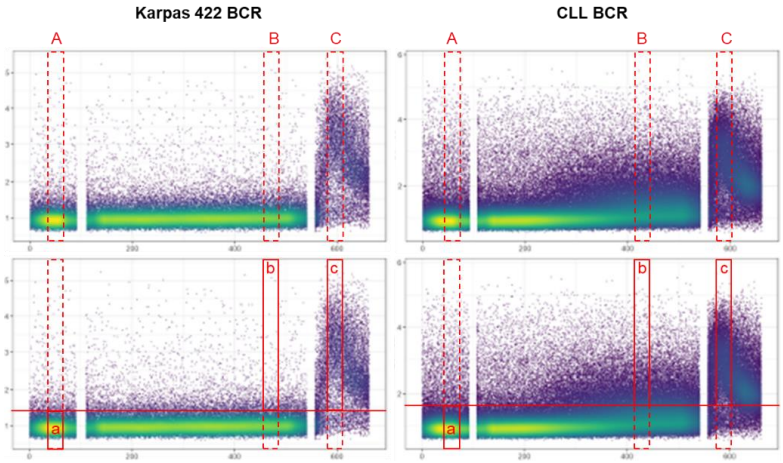
Three frames are defined and the number of events within each frame recorded:

1. A. Frame A: Stable interval in BP with number of events = NA
2. B. Frame B: Stable interval towards the end of EP with number of events = NB
3. C. Frame C: Interval starting at the time point of maximum increase signaling intensity ratio (SIR) and covering the interval of peak SIR in MP with number of events = NC.

Within each frame, a gate is defined and the number of events within each gate recorded:

a. Gate a: Part of frame A with upper limit (=SIR95) set to contain 95% of all events within frame A = an
b. Gate b: Part of frame B with lower limit set at SIR95 with number of events = nb
c. Gate c: Part of frame C with lower limit set at SIR95 with number of events = nc

Mean SIR are recorded for gates b (SIRb) and c (SIRc).



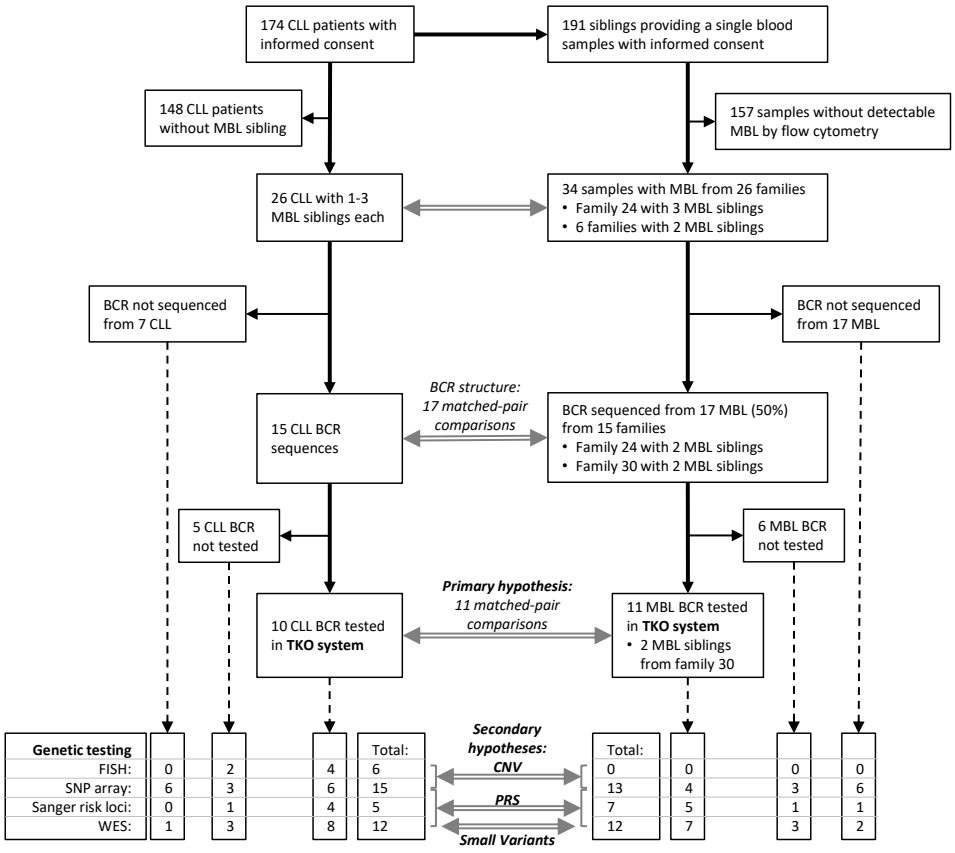
Interpretation:

- The fraction of false-positive events is set to 0.05.
- The fraction of non-responsive cells is calculated as (NC-nc) / NC.
- The fraction of cells with autonomous BCR signaling activity Qaut with correction for 5% false-positive cells and fraction of unresponsive cells is calculated as: Qaut = (nbNC – 0.05NBNC) / ncNC
- The calibrated autonomous BCR signal calSIRaut is: calSIRaut = SIRb/ SIRc
- BCR signaling strength is calculated as the product of the corrected fraction of cells with autonomous signalling and the calibrated autonomous BCR signal: QautcalSIRaut

Whole Exome Sequencing

FASTQ files from whole exome sequencing (WES) were processed using the Sarek workflow v2.7 and aligned to the human reference genome GRCh38 using Burrows Wheeler Algorithm (BWA) v0.7.17. [290, 291] Duplicated mapped reads were marked, local realignment of regions flanking indels and recalibration of base quality scores were performed to obtain more accurate bases according to the Genome Analysis ToolKit (GATK) best practices version v4.1.7.0. [292] Single nucleotide variants (SNV) and short insertions and deletions (INDELS) were called using Strelka2 v2.9.10. [293] Only high 5 confidence variants defined by quality scores (GQX) of at least 15 for SNV and 30 for INDELS were kept. The resulting variant call files were annotated by Ensembl-VEP (v103) with four filtering steps. [294] Variants were first filtered for the 120 most frequent mutated genes in CLL according the COSMIC database and an in-house cohort (Supplemental Table 3). [235, 295] Thereafter, variants were filtered based on their effect, i.e. frameshift, inframe deletion, missense, missense variant & splice region variant, splice region variant & synonymous variant, synonymous, in-frame insertion, stop gained, stop lost, frameshift variant & stop lost, missense variant & splice region variant, and coding sequence variant. Subsequently, variants were filtered for predicted deleterious effects

using a CADD phred score of >20, and benign variants annotated in Clinvar 202008 were discarded. Workflow quality control metrics were calculated and aggregated by MultiQC v1.8. 5 One sample (MBL24.3) with a high percentage (24.4%) of unmapped reads was discarded at this stage.



Supplementary Figure 1. Consort-like diagram of numbers of samples and experimental analyses. CNV: Copy number variations. PRS: Polygenic risk score.

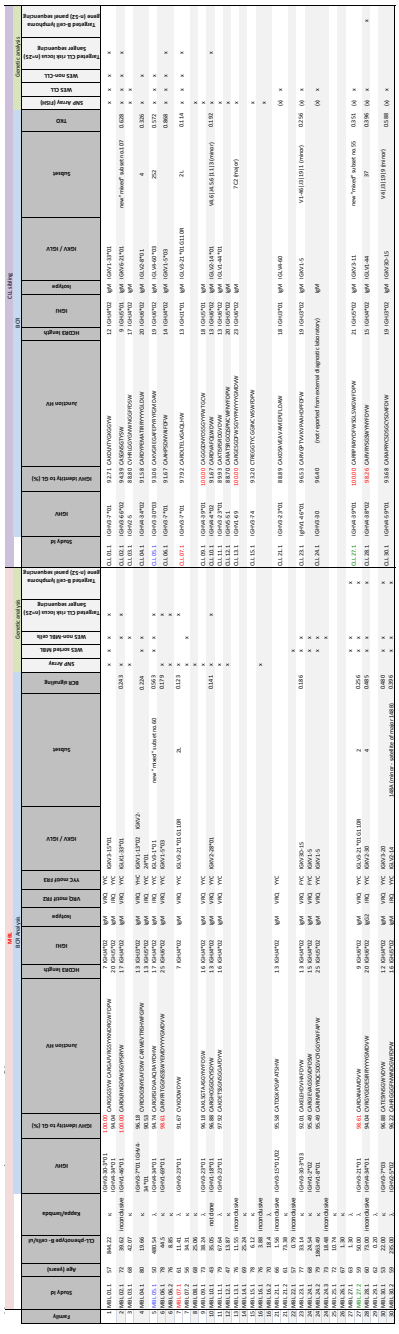
SUPPLEMENTARY TABLES

Supplementary Table 1. Characteristics and analyses of MBL and CLL in sibling pairs

Supplementary Table 2. Analysis of CLL susceptibility loci in CLL-MBL sibling pairs

Supplementary Table 3. CLL-related genes analyzed for small variants

Supplementary Table 4. Non-shared variants in 120 CLL-related genes



Supplementary Table 1. Characteristics and analyses of MBL and CLL in sibling pairs.

Supplementary Table 2. Analysis of CLL susceptibility loci in CLL-MBL sibling pairs.

ADAM19	DLG2	LINGO2	PCLO
ADGRL2	DLGAP2	LRMDA	PDE4D
ADGRL3	DMD	LRP1B	PDIA4
AGBL1	DPP10	LRRTM4	PDZD3
AGBL4	DPP6	LUZP2	PRKCH
AKAP13	DSCAM	MACROD2	PRKG1
ANKS1B	EPHA5	MAGI2	PRKN
ASTN2	EPHA6	MCL1	PTPRD
ATM	ERBB4	MDGA2	PTPRN2
ATRNL1	EYS	MYD88	PTPRT
BIRC3	FAM155A	MYH15	RBFOX1
BRINP3	FAT3	NAALADL2	RIN3
CACNA2D1	FHIT	NAGPA	ROBO1
CACNA2D3	FSTL5	NCAM2	ROBO2
CCSER1	GABRG3	NCKAP5	ROS1
CDH12	GALNTL6	NEGR1	RYR2
CDH13	GPC5	NELL1	SEMA6D
CDH18	GPC6	NKAIN2	SF3B1
CELSR1	GREB1L	NLGN1	SGCZ
CNTN4	GRID2	NOTCH1	SH3TC1
CNTN5	GRIK2	NPAS3	SNTG1
CNTNAP2	GRM7	NRG3	SPAG16
CNTNAP5	H1-2	NRXN1	SPOCK3
CSMD1	H1-3	NRXN3	THSD7B
CSMD3	H1-5	OPCML	TP53
CTNNA2	IGLL5	PARD3B	UNC5D
CTNNA3	IL1RAPL1	PARP4	USH2A
CTNND2	IL1RAPL2	PAX5	XPO1
DCC	KCNIP4	PCDH11X	ZFHX3
DGKB	KLHL1	PCDH15	ZFPM2

Supplementary Table 3. CLL-related genes analyzed for small variants.

Family	sample	CHROM	POS	SYMBOL	Gene	REF	ALT	QUAL	Allele	Consequence	IMPACT	cDNA_position	CDS_position	Protein_position	Amino_acids	Codons	Existing_variation
1	CLL01.1	chr1	49245770	AGBL4	ENSG00000186094	C	A	18	A	missense_variant&splice_region_variant	MODERATE	535	377	126	W/L	tGg/tTg	rs1370279151
1	CLL01.1	chr1	71407471	NEGR1	ENSG00000172260	T	G	46	G	missense_variant	MODERATE	1085	1040	347	Y/S	tAc/tCc	rs41289154&COSV60833851&COSV60866382
1	CLL01.1	chr2	137160352	THSD7B	ENSG00000144229	T	A	118	A	missense_variant	MODERATE	1687	1509	503	D/E	gaT/gaA	rs4954474&COSV55680598
1	CLL01.1	chr3	108470146	MYH15	ENSG00000144821	G	A	259	A	missense_variant	MODERATE	1567	1510	504	H/Y	Cac/Tac	rs9868484&COSV56308865
1	CLL01.1	chr4	61979743	ADGRL3	ENSG00000150471	C	G	73	G	missense_variant	MODERATE	1154	1156	386	L/V	Ctg/Gtg	rs125091100&COSV72232449&COSV99072504
1	CLL01.1	chr6	64307032	EYS	ENSG00000188107	T	C	47	C	missense_variant	MODERATE	6668	6129	2043	I/M	atA/atG	
1	CLL01.1	chr9	136510700	NOTCH1	ENSG00000148400	C	T	21	T	missense_variant	MODERATE	2955	2693	898	G/D	gGc/gAc	rs1166646402
1	CLL01.1	chr20	14085630	MACROD2	ENSG00000172264	C	T	122	T	missense_variant	MODERATE	568	173	58	T/I	aCt/aTt	rs29900505&COSV53960729
1	MBL01.1	chr1	216207290	USH2A	ENSG00000042781	T	A	18	A	missense_variant	MODERATE	3738	3299	1100	E/V	gAg/gTg	
1	MBL01.1	chr2	140274548	LRP1B	ENSG00000168702	C	G	15	G	missense_variant	MODERATE	13305	13018	4340	D/H	Gat/Cat	COSV101251919
1	MBL01.1	chr3	108441130	MYH15	ENSG00000144821	G	A	316	A	missense_variant	MODERATE	2903	2846	949	T/I	aCt/aTt	rs12638212&COSV56309895
1	MBL01.1	chr8	1565843	DLGAP2	ENSG00000198010	C	A	67	A	missense_variant	MODERATE	1571	1391	464	P/Q	cCa/cAa	rs2301963&COSV70102130
1	MBL01.1	chr14	92651884	RIN3	ENSG00000100599	C	T	337	T	missense_variant	MODERATE	987	835	279	R/C	Cgc/Tgc	rs117068593&COSV53649944
1	MBL01.1	chrX	31507453	DMD	ENSG00000198947	C	A	18	A	missense_variant&splice_region_variant	MODERATE	8455	8218	2740	D/Y	Gac/Tac	
2	CLL02.1	chr2	115753241	DPP10	ENSG00000175497	G	C	299	C	missense_variant	MODERATE	1156	1018	340	A/P	Gct/Cct	rs2053724&COSV59673831
2	CLL02.1	chr2	132783534	NCKAP5	ENSG00000176771	T	A	258	A	missense_variant	MODERATE	3654	3277	1093	N/Y	Aat/Tat	rs16841277&COSV58444336
2	CLL02.1	chr2	137160352	THSD7B	ENSG00000144229	T	A	99	A	missense_variant	MODERATE	1687	1509	503	D/E	gaT/gaA	rs4954474&COSV55680598
2	CLL02.1	chr2	140371210	LRP1B	ENSG00000168702	C	G	163	G	missense_variant	MODERATE	11131	10844	3615	G/A	gGa/gCa	rs76554185&COSV67206902&COSV67273739
2	CLL02.1	chr3	78667911	ROBO1	ENSG00000169855	G	T	154	T	missense_variant	MODERATE	2235	1938	646	S/R	agC/agA	
2	CLL02.1	chr3	108470146	MYH15	ENSG00000144821	G	A	1030	A	missense_variant	MODERATE	1567	1510	504	H/Y	Cac/Tac	rs9868484&COSV56308865
2	CLL02.1	chr3	108470780	MYH15	ENSG00000144821	C	T	139	T	missense_variant	MODERATE	1418	1361	454	R/Q	cGg/cAg	rs4299484&COSV56310353
2	CLL02.1	chr3	175627355	NAALADL2	ENSG00000177694	C	G	132	G	missense_variant	MODERATE	1939	1865	622	P/R	cCc/cGc	rs9866564
2	CLL02.1	chr4	161386160	FSTL5	ENSG00000168843	C	A	164	A	missense_variant	MODERATE	2533	2131	711	D/Y	Gat/Tat	rs3749598
2	CLL02.1	chr5	22078475	CDH12	ENSG00000154162	C	T	241	T	missense_variant	MODERATE	1011	202	68	V/M	Gtg/Atg	rs4371716&CM067358&COSV66394300
2	CLL02.1	chr5	157502808	ADAM19	ENSG00000135074	C	T	893	T	missense_variant	MODERATE	1382	1303	435	E/K	Gaa/Aaa	rs56384823
2	CLL02.1	chr5	157509356	ADAM19	ENSG00000135074	T	C	138	C	missense_variant	MODERATE	929	850	284	S/G	Agt/Ggt	rs14227958&CM093420&COSV57425399
2	CLL02.1	chr6	117301070	ROS1	ENSG00000047936	C	T	115	T	missense_variant	MODERATE	6905	6619	2207	D/N	Gac/Aac	rs529038&CM053396&COSV63850799
2	CLL02.1	chr8	50553159	SNTG1	ENSG00000147481	T	C	19	C	missense_variant	MODERATE	1717	790	264	S/P	Tca/Cca	
2	CLL02.1	chr8	112228859	CSMD3	ENSG00000164796	T	G	129	G	missense_variant	MODERATE	10946	10861	3621	N/H	Aat/Cat	rs1592624&COSV52174720
2	CLL02.1	chr8	112241725	CSMD3	ENSG00000164796	A	G	201	G	missense_variant	MODERATE	10548	10463	3488	L/P	cTa/cCa	rs61753736&COSV104398217
2	CLL02.1	chr10	82979088	NRG3	ENSG00000185737	G	T	81	T	missense_variant	MODERATE	1698	1551	517	R/S	agG/agT	rs761374063&COSV64583486
2	CLL02.1	chr11	92353347	FAT3	ENSG00000165323	C	T	404	T	missense_variant	MODERATE	1609	1235	412	S/F	tCt/tTt	rs10830902
2	CLL02.1	chr13	24455080	PARP4	ENSG00000102699	C	T	426	T	missense_variant	MODERATE	2769	2695	899	A/T	Gcc/Acc	rs2275660&COSV67960444
2	CLL02.1	chr15	85580133	AKAP13	ENSG00000170776	G	A	322	A	missense_variant	MODERATE	2272	2065	689	E/K	Gag/Aag	rs71771078&COSV63448301
2	CLL02.1	chr20	14085630	MACROD2	ENSG00000172264	C	T	82	T	missense_variant	MODERATE	568	173	58	T/I	aCt/aTt	rs29900505&COSV53960729
2	CLL02.1	chr21	21373867	NCAM2	ENSG00000154654	T	C	185	C	missense_variant	MODERATE	1204	1049	350	L/P	cTg/cCg	rs232518
2	CLL02.1	chr22	46535180	CELSR1	ENSG00000075275	G	C	398	C	missense_variant	MODERATE	2441	1991	664	S/W	tCg/tGg	rs4823850&COSV53084542
2	CLL02.1	chrX	32362843	DMD	ENSG00000198947	A	G	37	G	missense_variant	MODERATE	5507	5270	1757	I/T	aTc/aCc	rs201516290&COSV63773732
2	CLL02.1	chrX	105767516	IL1RAPL2	ENSG00000189108	A	G	186	G	missense_variant	MODERATE	2788	1916	639	H/R	cAt/cGt	rs368014818
4	MBL04.1	chr1	216070175	USH2A	ENSG00000042781	T	C	248	C	missense_variant	MODERATE	6414	5975	1992	Y/C	tAt/tGt	rs41303287&CM104140&COSV56354439
4	MBL04.1	chr2	115753241	DPP10	ENSG00000175497	G	C	204	C	missense_variant	MODERATE	1156	1018	340	A/P	Gct/Cct	rs2053724&COSV59673831
4	MBL04.1	chr2	132783032	NCKAP5	ENSG00000176771	G	T	281	T	missense_variant	MODERATE	4156	3779	1260	P/Q	cCa/cAa	rs13016342&COSV58445700
4	MBL04.1	chr2	132785012	NCKAP5	ENSG00000176771	C	G	291	G	missense_variant	MODERATE	2176	1799	600	S/T	aGt/aCt	rs17325719&COSV58430764&COSV58445709
4	MBL04.1	chr2	137160352	THSD7B	ENSG00000144229	T	A	351	A	missense_variant	MODERATE	1687	1509	503	D/E	gaT/gaA	rs4954474&COSV55680598
4	MBL04.1	chr3	108470146	MYH15	ENSG00000144821	G	A	155	A	missense_variant	MODERATE	1567	1510	504	H/Y	Cac/Tac	rs9868484&COSV56308865
4	MBL04.1	chr3	108470780	MYH15	ENSG00000144821	C	T	65	T	missense_variant	MODERATE	1418	1361	454	R/Q	cGg/cAg	rs4299484&COSV56310353
4	MBL04.1	chr4	61934862	ADGRL3	ENSG00000150471	A	T	15	T	missense_variant	MODERATE	303	305	102	N/I	aAc/aTc	
4	MBL04.1	chr5	22078475	CDH12	ENSG00000154162	C	T	182	T	missense_variant	MODERATE	1011	202	68	V/M	Gtg/Atg	rs4371716&CM067358&COSV66394300
4	MBL04.1	chr5	157490332	ADAM19	ENSG00000135074	A	C	523	C	missense_variant	MODERATE	2297	2218	740	S/A	Tcc/Gcc	rs10067096&COSV57429259
4	MBL04.1	chr5	157491842	ADAM19	ENSG00000135074	C	T	378	T	missense_variant	MODERATE	2058	1979	660	G/D	gGc/gAc	rs2287749
4	MBL04.1	chr7	83155814	PCLO	ENSG00000186472	C	T	142	T	missense_variant	MODERATE	1119	827	276	R/Q	cGa/cAa	rs7475932333&COSV61667530
4	MBL04.1	chr7	158138452	PTPRN2	ENSG00000155093	C	T	439	T	missense_variant	MODERATE	1128	974	325	S/N	aGt/aAt	rs11304098&COSV67024020
4	MBL04.1	chr8	1565843	DLGAP2	ENSG00000198010	C	A	69	A	missense_variant	MODERATE	1571	1391	464	P/Q	cCa/cAa	rs2301963&COSV70102130
4	MBL04.1	chr8	112228859	CSMD3	ENSG00000164796	T	G	117	G	missense_variant	MODERATE	10946	10861	3621	N/H	Aat/Cat	rs1592624&COSV52174720
4	MBL04.1	chr10	66766407	CTNNA3	ENSG00000183230	C	T	21	T	missense_variant	MODERATE	1339	1138	380	A/T	Gct/Act	rs772731501

Family	sample	CHROM	POS	SYMBOL	Gene	REF	ALT	QUAL	Allele	Consequence	IMPACT	cDNA_position	CDS_position	Protein_position	Amino_acids	Codons	Existing_variation
4	MBL04.1	chr11	20783740	NELL1	ENSG00000165973	G	A	202	A	missense_variant	MODERATE	383	245	82	R/Q	cGg/cAg	rs8176785&CM074381
4	MBL04.1	chr11	92831816	FAT3	ENSG00000165323	G	A	242	A	missense_variant	MODERATE	10050	9676	3226	V/I	Gtt/Att	rs766505847
4	MBL04.1	chr13	24455080	PARP4	ENSG00000102699	C	T	381	T	missense_variant	MODERATE	2769	2695	899	A/T	Gcc/Acc	rs2275660&COSV67960444
4	MBL04.1	chr22	46535180	CELSR1	ENSG00000075275	G	C	330	C	missense_variant	MODERATE	2441	1991	664	S/W	tCg/tGg	rs48238508&COSV53084542
5	CLL05.1	chr4	65365968	EPHA5	ENSG00000145242	C	T	234	T	missense_variant	MODERATE	2698	1951	651	A/T	Gca/Aca	rs36050417&COSV56636694
5	CLL05.1	chr4	65490601	EPHA5	ENSG00000145242	A	G	326	G	missense_variant	MODERATE	1925	1178	393	I/T	aTt/aCt	rs780551761
5	CLL05.1	chr7	14685305	DGKB	ENSG00000136267	A	C	16	C	missense_variant	MODERATE	1116	769	257	Y/D	Tat/Gat	COSV104385499
5	CLL05.1	chr11	20928414	NELL1	ENSG00000165973	C	T	120	T	missense_variant	MODERATE	1070	932	311	P/L	cCt/cTt	rs55926004
5	CLL05.1	chr20	14085630	MACROD2	ENSG00000172264	C	T	62	T	missense_variant	MODERATE	568	173	58	T/I	aCt/aTt	rs2990505&COSV53960729
5	CLL05.1	chr22	22893818	IGLL5	ENSG00000254709	G	A	44	A	missense_variant&splice_region_variant	MODERATE	563	325	109	G/S	Ggt/Agt	rs201857114&COSV101138586&COSV66537362
5	MBL05.1	chr1	71407471	NEGR1	ENSG00000172260	T	G	129	G	missense_variant	MODERATE	1085	1040	347	Y/S	tAc/tCc	rs41289154&COSV60833851&COSV60866382
5	MBL05.1	chr6	117301070	ROS1	ENSG00000047936	C	T	155	T	missense_variant	MODERATE	6905	6619	2207	D/N	Gac/Aac	rs529038&CM053396&COSV63850799
5	MBL05.1	chr7	158167251	PTPRN2	ENSG00000155093	G	C	166	C	missense_variant	MODERATE	744	590	197	A/G	gCc/gGc	rs140999965
6	CLL06.1	chr1	71407471	NEGR1	ENSG00000172260	T	G	138	G	missense_variant	MODERATE	1085	1040	347	Y/S	tAc/tCc	rs41289154&COSV60833851&COSV60866382
6	CLL06.1	chr2	115746095	DPP10	ENSG00000175497	G	A	47	A	missense_variant	MODERATE	1000	862	288	V/M	Gtg/Atg	rs36044503
6	CLL06.1	chr2	115753241	DPP10	ENSG00000175497	G	C	173	C	missense_variant	MODERATE	1156	1018	340	A/P	Gct/Cct	rs2053724&COSV59673831
6	CLL06.1	chr2	132783032	NCKAP5	ENSG00000176771	G	T	452	T	missense_variant	MODERATE	4156	3779	1260	P/Q	cCa/cAa	rs13016342&COSV58445700
6	CLL06.1	chr2	132785012	NCKAP5	ENSG00000176771	C	G	631	G	missense_variant	MODERATE	2176	1799	600	S/T	aGt/aCt	rs17325719&COSV58430764&COSV58445709
6	CLL06.1	chr2	141810341	LRP1B	ENSG00000168702	T	C	134	C	missense_variant	MODERATE	430	143	48	Q/R	cAg/cGg	rs12990449&COSV67182532
6	CLL06.1	chr3	78635874	ROBO1	ENSG00000169855	C	T	107	T	missense_variant	MODERATE	3569	3272	1091	S/N	aGc/aAc	rs35456279&COSV71397020
6	CLL06.1	chr3	108470146	MYH15	ENSG00000144821	G	A	1153	A	missense_variant	MODERATE	1567	1510	504	H/Y	Cac/Tac	rs9868484&COSV56308865
6	CLL06.1	chr3	175627355	NAALADL2	ENSG00000177694	C	G	273	G	missense_variant	MODERATE	1939	1865	622	P/R	cCc/cGc	rs9866564
6	CLL06.1	chr4	61979743	ADGRL3	ENSG00000150471	C	G	138	G	missense_variant	MODERATE	1154	1156	386	L/V	Ctg/Gtg	rs12509110&COSV72232449&COSV99072504
6	CLL06.1	chr5	22078475	CDH12	ENSG00000154162	C	T	416	T	missense_variant	MODERATE	1011	202	68	V/M	Gtg/Atg	rs4371716&CM067358&COSV66394300
6	CLL06.1	chr6	117301070	ROS1	ENSG00000047936	C	T	74	T	missense_variant	MODERATE	6905	6619	2207	D/N	Gac/Aac	rs529038&CM053396&COSV63850799
6	CLL06.1	chr7	149015000	PDIA4	ENSG00000155660	G	A	605	A	missense_variant	MODERATE	615	518	173	T/M	aCg/aTg	rs2290971&CM1511790
6	CLL06.1	chr7	158167203	PTPRN2	ENSG00000155093	C	T	30	T	missense_variant	MODERATE	792	638	213	R/H	cGc/cAc	rs1130496&COSV67046421
6	CLL06.1	chr8	1565843	DLGAP2	ENSG00000198010	C	A	16	A	missense_variant	MODERATE	1571	1391	464	P/Q	cCa/cAa	rs2301963&COSV70102130
6	CLL06.1	chr8	112228859	CSMD3	ENSG00000164796	T	G	258	G	missense_variant	MODERATE	10946	10861	3621	N/H	Aat/Cat	rs1592624&COSV52174720
6	CLL06.1	chr11	20783740	NELL1	ENSG00000165973	G	A	70	A	missense_variant	MODERATE	383	245	82	R/Q	cGg/cAg	rs8176785&CM074381
6	CLL06.1	chr13	24455080	PARP4	ENSG00000102699	C	T	639	T	missense_variant	MODERATE	2769	2695	899	A/T	Gcc/Acc	rs2275660&COSV67960444
6	CLL06.1	chr14	92651884	RIN3	ENSG00000100599	C	T	34	T	missense_variant	MODERATE	987	835	279	R/C	Cgc/Tgc	rs117068593&COSV53649944
6	CLL06.1	chr15	85580133	AKAP13	ENSG00000170776	G	A	546	A	missense_variant	MODERATE	2272	2065	689	E/K	Gag/Aag	rs71771078&COSV63448301
6	CLL06.1	chr15	85735078	AKAP13	ENSG00000170776	G	A	206	A	missense_variant	MODERATE	7576	7369	2457	G/S	Ggt/Agt	rs2241268&COSV63448267
6	CLL06.1	chr17	7670613	TP53	ENSG00000141510	A	C	178	C	missense_variant	MODERATE	1238	1096	366	S/A	Tcc/Gcc	rs17881470&CM078492&COSV52875421
6	CLL06.1	chr20	14085630	MACROD2	ENSG00000172264	C	T	80	T	missense_variant	MODERATE	568	173	58	T/I	aCt/aTt	rs2990505&COSV53960729
6	CLL06.1	chr21	21373867	NCAM2	ENSG00000154654	T	C	315	C	missense_variant	MODERATE	1204	1049	350	L/P	cTg/cCg	rs232518
6	CLL06.1	chr21	40353698	DSCAM	ENSG00000171587	C	T	41	T	missense_variant	MODERATE	1198	701	234	R/H	cGc/cAc	rs41395652
6	CLL06.1	chr22	46535180	CELSR1	ENSG00000075275	G	C	520	C	missense_variant	MODERATE	2441	1991	664	S/W	tCg/tGg	rs48238508&COSV53084542
21	CLL21.1	chr2	115753241	DPP10	ENSG00000175497	G	C	162	C	missense_variant	MODERATE	1156	1018	340	A/P	Gct/Cct	rs2053724&COSV59673831
21	CLL21.1	chr2	132783032	NCKAP5	ENSG00000176771	G	T	905	T	missense_variant	MODERATE	4156	3779	1260	P/Q	cCa/cAa	rs13016342&COSV58445700
21	CLL21.1	chr2	132785012	NCKAP5	ENSG00000176771	C	G	1173	G	missense_variant	MODERATE	2176	1799	600	S/T	aGt/aCt	rs17325719&COSV58430764&COSV58445709
21	CLL21.1	chr2	137160352	THSD7B	ENSG00000144229	T	A	843	A	missense_variant	MODERATE	1687	1509	503	D/E	gaT/gaA	rs4954474&COSV55680598
21	CLL21.1	chr2	141810341	LRP1B	ENSG00000168702	T	C	560	C	missense_variant	MODERATE	430	143	48	Q/R	cAg/cGg	rs12990449&COSV67182532
21	CLL21.1	chr2	213930019	SPAG16	ENSG00000144451	A	C	593	C	missense_variant	MODERATE	1294	1274	425	K/T	aAa/aCa	rs12623569&COSV59078990
21	CLL21.1	chr3	108470146	MYH15	ENSG00000144821	G	A	759	A	missense_variant	MODERATE	1567	1510	504	H/Y	Cac/Tac	rs9868484&COSV56308865
21	CLL21.1	chr3	108470780	MYH15	ENSG00000144821	C	T	557	T	missense_variant	MODERATE	1418	1361	454	R/Q	cGg/cAg	rs4299484&COSV56310353
21	CLL21.1	chr3	175627355	NAALADL2	ENSG00000177694	C	G	657	G	missense_variant	MODERATE	1939	1865	622	P/R	cCc/cGc	rs9866564
21	CLL21.1	chr5	157509356	ADAM19	ENSG00000135074	T	C	1981	C	missense_variant	MODERATE	929	850	284	S/G	Agt/Ggt	rs1422795&CM093420&COSV57425399
21	CLL21.1	chr6	26234599	H1-3	ENSG00000124575	G	C	49	C	missense_variant	MODERATE	389	335	112	A/G	gCg/gGg	rs140433263
21	CLL21.1	chr6	27866899	H1-5	ENSG00000184357	C	T	259	T	missense_variant	MODERATE	690	631	211	A/T	Gca/Aca	rs34144478
21	CLL21.1	chr6	117301070	ROS1	ENSG00000047936	C	T	240	T	missense_variant	MODERATE	6905	6619	2207	D/N	Gac/Aac	rs529038&CM053396&COSV63850799
21	CLL21.1	chr7	158138452	PTPRN2	ENSG00000155093	C	T	551	T	missense_variant	MODERATE	1128	974	325	S/N	aGt/aAt	rs1130498&COSV67024020
21	CLL21.1	chr8	1565843	DLGAP2	ENSG00000198010	C	A	547	A	missense_variant	MODERATE	1571	1391	464	P/Q	cCa/cAa	rs2301963&COSV70102130
21	CLL21.1	chr11	20783740	NELL1	ENSG00000165973	G	A	2263	A	missense_variant	MODERATE	383	245	82	R/Q	cGg/cAg	rs8176785&CM074381
21	CLL21.1	chr11	108310218	ATM	ENSG00000149311	G	C	283	C	missense_variant	MODERATE	5971	5821	1941	V/L	Gtt/Ctt	rs147187700&CM094667&COSV53739000&COSV53750671
21	CLL21.1	chr13	24455080	PARP4	ENSG00000102699	C	T	616	T	missense_variant	MODERATE	2769	2695	899	A/T	Gcc/Acc	rs2275660&COSV67960444

Family	sample	CHROM	POS	SYMBOL	Gene	REF	ALT	QUAL	Allele	Consequence	IMPACT	cDNA_position	CDS_position	Protein_position	Amino_acids	Codons	Existing_variation
21	CLL21.1	chr15	85735078	AKAP13	ENSG00000170776	G	A	417	A	missense_variant	MODERATE	7576	7369	2457	G/S	Ggt/Agt	rs2241268&COSV63448267
21	CLL21.1	chr15	86264612	AGBL1	ENSG00000273540	T	C	696	C	missense_variant	MODERATE	1543	1441	481	S/P	Tcc/Ccc	rs11857527&COSV66866512
21	CLL21.1	chr21	21373867	NCAM2	ENSG00000154654	T	C	1106	C	missense_variant	MODERATE	1204	1049	350	L/P	cTg/cCg	rs232518
21	CLL21.1	chr22	46535180	CELSR1	ENSG00000075275	G	C	1999	C	missense_variant	MODERATE	2441	1991	664	S/W	tCg/tGg	rs4823850&COSV53084542
22	MBL22.1	chr1	81907079	ADGRL2	ENSG00000117114	T	A	18	A	missense_variant	MODERATE	352	136	46	S/T	Tct/Act	
22	MBL22.1	chr2	213930019	SPAG16	ENSG00000144451	A	C	181	C	missense_variant	MODERATE	1294	1274	425	K/T	aAa/aCa	rs12623569&COSV59078990
22	MBL22.1	chr8	1565843	DLGAP2	ENSG00000198010	C	A	343	A	missense_variant	MODERATE	1571	1391	464	P/Q	cCa/cAa	rs2301963&COSV7010213 0
22	MBL22.1	chr8	112228859	CSMD3	ENSG00000164796	T	G	420	G	missense_variant	MODERATE	10946	10861	3621	N/H	Aat/Cat	rs1592624&COSV52174720
22	MBL22.1	chr13	94382448	GPC6	ENSG00000183098	A	G	26	G	missense_variant	MODERATE	1837	1187	396	K/R	aAa/aGa	
22	MBL22.1	chr21	40338287	DSCAM	ENSG00000171587	A	G	15	G	missense_variant	MODERATE	2094	1597	533	Y/H	Tac/Cac	
23	CLL23.1	chr5	157509356	ADAM19	ENSG00000135074	T	C	375	C	missense_variant	MODERATE	929	850	284	S/G	Agt/Ggt	rs1422795&CM093420&COSV57425399
23	CLL23.1	chr14	92651884	RIN3	ENSG00000100599	C	T	773	T	missense_variant	MODERATE	987	835	279	R/C	Cgc/Tgc	rs117068593&COSV53649944
23	CLL23.1	chr20	42104665	PTPRT	ENSG00000196090	G	T	825	T	missense_variant	MODERATE	3471	3471	1157	N/K	aaC/aaA	rs201162919&COSV61980909
23	CLL23.1	chr21	40353698	DSCAM	ENSG00000171587	C	T	210	T	missense_variant	MODERATE	1198	701	234	R/H	cGc/cAc	rs41395652
23	MBL23.1	chr1	237566732	RYR2	ENSG00000198626	A	G	624	G	missense_variant	MODERATE	3718	3380	1127	E/G	gAg/gGg	rs200525962
23	MBL23.1	chr4	8218302	SH3TC1	ENSG00000125089	G	A	238	A	missense_variant	MODERATE	1015	871	291	D/N	Gac/Aac	rs1281138
23	MBL23.1	chr4	8240900	SH3TC1	ENSG00000125089	C	A	239	A	missense_variant	MODERATE	4100	3956	1319	P/H	cCt/cAt	rs150702535
23	MBL23.1	chr6	27866899	H1-5	ENSG00000184357	C	T	278	T	missense_variant	MODERATE	690	631	211	A/T	Gca/Aca	rs34144478
23	MBL23.1	chr7	154052980	DPP6	ENSG00000130226	C	T	21	T	missense_variant	MODERATE	583	160	54	R/W	Cgg/Tgg	COSV66748904
24	MBL24.1	chr1	215728169	USH2A	ENSG00000042781	G	A	254	A	missense_variant	MODERATE	12366	11927	3976	T/M	aCg/aTg	rs142381713&CM073404&COSV56393885
24	MBL24.1	chr5	157509356	ADAM19	ENSG00000135074	T	C	353	C	missense_variant	MODERATE	929	850	284	S/G	Agt/Ggt	rs1422795&CM093420&COSV57425399
24	MBL24.1	chr6	117362658	ROS1	ENSG00000047936	G	A	555	A	missense_variant	MODERATE	3597	3311	1104	S/L	tCa/tTa	rs2229079&COSV63860865
24	MBL24.1	chr6	117404285	ROS1	ENSG00000047936	T	G	450	G	missense_variant	MODERATE	746	460	154	T/P	Act/Cct	rs1998206&COSV63854406
24	MBL24.1	chr11	92837690	FAT3	ENSG00000165323	G	T	422	T	missense_variant	MODERATE	10626	10252	3418	A/S	Gcc/Tcc	rs201449521&COSV53125665&COSV99973667
24	MBL24.1	chr13	24442624	PARP4	ENSG00000102699	G	A	74	A	missense_variant	MODERATE	3583	3509	1170	T/I	aCa/aTa	rs113301501&CM160622&COSV67960155
24	MBL24.1	chr13	24447185	PARP4	ENSG00000102699	A	G	87	G	missense_variant&splice_region_variant	MODERATE	3190	3116	1039	I/T	aTa/aCa	rs73172125&COSV67960144
24	MBL24.1	chr14	46957592	MDGA2	ENSG00000139915	C	T	632	T	missense_variant	MODERATE	1612	977	326	R/Q	cGa/cAa	rs553082374&COSV62105261
24	MBL24.1	chr17	7674917	TP53	ENSG00000141510	T	C	993	C	missense_variant	MODERATE	756	614	205	Y/C	tAt/tGt	rs1057520007&COSV52665440&COSV52677268&COSV52688506
24	MBL24.1	chr22	22888193	IGLL5	ENSG00000254709	G	A	243	A	missense_variant	MODERATE	378	140	47	S/N	aGc/aAc	rs540568540&COSV73249672
24	MBL24.2	chr2	197402110	SF3B1	ENSG00000115524	T	C	842	C	missense_variant	MODERATE	2127	2098	700	K/E	Aaa/Gaa	rs559063155&COSV59205318&COSV59216586
24	MBL24.2	chr3	108470780	MYH15	ENSG00000144821	C	T	665	T	missense_variant	MODERATE	1418	1361	454	R/Q	cGg/cAg	rs4299484&COSV63610353
24	MBL24.3	chr1	215647626	USH2A	ENSG00000042781	A	T	27	T	missense_variant	MODERATE	15126	14687	4896	L/H	cTt/cAt	
24	MBL24.3	chr1	215790251	USH2A	ENSG00000042781	A	C	15	C	missense_variant	MODERATE	10429	9990	3330	N/K	aaT/aaG	
24	MBL24.3	chr1	237569176	RYR2	ENSG00000198626	A	G	23	G	missense_variant	MODERATE	3793	3455	1152	Y/C	tAt/tGt	
24	MBL24.3	chr1	237674166	RYR2	ENSG00000198626	A	T	61	T	missense_variant	MODERATE	8999	8661	2887	E/D	gaA/gaT	
24	MBL24.3	chr2	132783534	NCKAP5	ENSG00000176771	T	A	1027	A	missense_variant	MODERATE	3654	3277	1093	N/Y	Aat/Tat	rs16841277&COSV58444336
24	MBL24.3	chr2	141062161	LRP1B	ENSG00000168702	C	A	18	A	missense_variant	MODERATE	1413	1126	376	A/S	Gca/Tca	
24	MBL24.3	chr2	197396311	SF3B1	ENSG00000115524	G	T	34	T	missense_variant	MODERATE	3313	3284	1095	A/D	gCt/gAt	
24	MBL24.3	chr3	108410799	MYH15	ENSG00000144821	C	A	16	A	missense_variant	MODERATE	4396	4339	1447	V/F	Gtc/Ttc	COSV56321473
24	MBL24.3	chr5	21783409	CDH12	ENSG00000154162	C	T	33	T	missense_variant	MODERATE	2151	1342	448	D/N	Gac/Aac	
24	MBL24.3	chr5	59893589	PDE4D	ENSG00000113448	C	G	250	G	missense_variant	MODERATE	138	34	12	A/P	Gcg/Ccg	
24	MBL24.3	chr7	81991211	CACNA2D1	ENSG00000153956	C	T	36	T	stop_gained	HIGH	2088	1770	590	W/*	tgG/tgA	rs1554341873
24	MBL24.3	chr7	147132332	CNTNAP2	ENSG00000174469	A	T	31	T	missense_variant	MODERATE	1247	1171	391	N/Y	Aac/Tac	COSV62189039
24	MBL24.3	chr7	148147676	CNTNAP2	ENSG00000174469	A	G	25	G	missense_variant	MODERATE	2816	2740	914	T/A	Acc/Gcc	
24	MBL24.3	chr8	2978629	CSMD1	ENSG00000183117	G	A	36	A	missense_variant	MODERATE	9047	8549	2850	S/F	tCc/tTc	
24	MBL24.3	chr8	14324194	SGCZ	ENSG00000185053	C	A	26	A	missense_variant	MODERATE	1053	245	82	G/V	gGa/gTa	COSV65998392&COSV66047381
24	MBL24.3	chr9	117214604	ASTN2	ENSG00000148219	A	T	3070	T	missense_variant	MODERATE	888	769	257	S/T	Tct/Act	rs139148246
24	MBL24.3	chr11	92837708	FAT3	ENSG00000165323	C	G	175	G	missense_variant	MODERATE	10644	10270	3424	P/A	Cct/Gct	
24	MBL24.3	chr13	93830302	GPC6	ENSG00000183098	T	G	63	G	missense_variant	MODERATE	1118	468	156	N/K	aaT/aaG	
24	MBL24.3	chr14	61485534	PRKCH	ENSG00000027075	T	G	30	G	missense_variant	MODERATE	1693	1311	437	N/K	aaT/aaG	
24	MBL24.3	chr15	85645868	AKAP13	ENSG00000170776	C	A	29	A	missense_variant	MODERATE	4495	4288	1430	Q/K	Caa/Aaa	
24	MBL24.3	chr18	21384368	GREB1L	ENSG00000141449	A	C	149	C	missense_variant	MODERATE	475	320	107	Q/P	cAa/cCa	rs1425668691
24	MBL24.3	chr21	40187251	DSCAM	ENSG00000171587	C	A	30	A	missense_variant	MODERATE	3156	2659	887	D/Y	Gac/Tac	
24	MBL24.3	chr22	46386404	CELSR1	ENSG00000075275	A	G	505	G	missense_variant&splice_region_variant	MODERATE	7187	6737	2246	M/T	aTg/aCg	rs749104236
24	MBL24.3	chr22	46537122	CELSR1	ENSG00000075275	C	G	27	G	missense_variant	MODERATE	499	49	17	A/P	Gcc/Ccc	
27	CLL27.1	chr1	81990901	ADGRL2	ENSG00000117114	G	A	551	A	missense_variant	MODERATE	4184	3968	1323	R/K	aGa/aAa	rs41292984&COSV54685119
27	CLL27.1	chr2	115746095	DPP10	ENSG00000175497	G	A	542	A	missense_variant	MODERATE	1000	862	288	V/M	Gtg/Atg	rs36044503

Family	sample	CHROM	POS	SYMBOL	Gene	REF	ALT	QUAL	Allele	Consequence	IMPACT	cDNA_position	CDS_position	Protein_position	Amino_acids	Codons	Existing_variation
	27 CLL27.1	chr2	197402635	SF3B1	ENSG00000115524	C	A	139	A	missense_variant	MODERATE	2027	1998	666	K/N	aaG/aaT	rs377023736&COSV59205777&COSV59205799
	27 CLL27.1	chr8	1565843	DLGAP2	ENSG00000198010	C	A	572	A	missense_variant	MODERATE	1571	1391	464	P/Q	cCa/cAa	rs230196338&COSV70102130
	27 CLL27.1	chr10	115334318	ATRNL1	ENSG00000107518	A	C	923	C	missense_variant	MODERATE	3460	3074	1025	N/T	aAt/aCt	rs1050159510
	27 CLL27.1	chr11	99819645	CNTN5	ENSG00000149972	C	T	371	T	stop_gained	HIGH	688	157	53	R/*	cGa/Tga	rs122926598&COSV54311010
	27 MBL27.1	chr4	161386160	FSTL5	ENSG00000168843	C	A	810	A	missense_variant	MODERATE	2533	2131	711	D/Y	Gat/Tat	rs3749598
	27 MBL27.1	chr9	136515524	NOTCH1	ENSG00000148400	C	T	117	T	missense_variant	MODERATE	2124	1862	621	R/H	cGc/cAc	rs138504021&COSV99487877
	27 MBL27.1	chr11	83962888	DLG2	ENSG00000150672	G	A	769	A	missense_variant	MODERATE	1723	1337	446	T/I	aCc/aTc	rs140599607
	27 MBL27.2	chr1	21579051	USH2A	ENSG00000042781	G	T	1925	T	missense_variant	MODERATE	10253	9814	3272	P/T	Ccg/Acg	COSV100239515
	27 MBL27.2	chr3	17573328	NAALADL2	ENSG00000177694	A	G	571	G	missense_variant	MODERATE	1993	1919	640	Q/R	cAa/cGa	rs62286105
	27 MBL27.2	chr9	136523093	NOTCH1	ENSG00000148400	G	A	15	A	missense_variant	MODERATE	761	499	167	P/S	Cca/Tca	rs143903832&COSV53077398
	27 MBL27.2	chr13	24447125	PARP4	ENSG00000102699	T	C	54	C	missense_variant	MODERATE	3250	3176	1059	Q/R	cAg/cGg	rs77269056&COSV67960139
	27 MBL27.2	chr13	24447185	PARP4	ENSG00000102699	A	G	194	G	missense_variant&splice_region_variant	MODERATE	3190	3116	1039	I/T	aTa/aCa	rs731721258&COSV67960144
	28 CLL28.1	chr2	61492051	XPO1	ENSG00000082898	T	C	209	C	missense_variant	MODERATE	2672	1871	624	D/G	gAt/gGt	COSV68944745
	28 CLL28.1	chr3	54569985	CACNA2D3	ENSG00000157445	G	A	672	A	missense_variant	MODERATE	931	769	257	V/M	Gtg/Atg	rs372465588&COSV55530096&COSV55536891
	28 CLL28.1	chr5	157490332	ADAM19	ENSG00000135074	A	C	836	C	missense_variant	MODERATE	2297	2218	740	S/A	Tcc/Gcc	rs10067096&COSV57429259
	28 CLL28.1	chr5	157509356	ADAM19	ENSG00000135074	T	C	462	C	missense_variant	MODERATE	929	850	284	S/G	Agt/Ggt	rs142279558&CM093420&COSV57425399
	28 CLL28.1	chr5	157575687	ADAM19	ENSG00000135074	C	T	116	T	missense_variant	MODERATE	89	10	4	G/S	Ggc/Agc	rs11465228
	28 CLL28.1	chr8	112636939	CSMD3	ENSG00000164796	A	G	695	G	missense_variant	MODERATE	3678	3593	1198	F/S	tTt/tCt	rs765163979
	28 CLL28.1	chr15	85735078	AKAP13	ENSG00000170776	G	A	413	A	missense_variant	MODERATE	7576	7369	2457	G/S	Ggt/Agt	rs224126&COSV63448267
	28 MBL28.1	chr2	211630569	ERBB4	ENSG00000178568	T	A	238	A	missense_variant	MODERATE	2244	1972	658	I/F	Att/Ttt	rs190654033
	28 MBL28.1	chr7	158138452	PTPRN2	ENSG00000155093	C	T	1271	T	missense_variant	MODERATE	1128	974	325	S/N	aGt/aAt	rs113049&COSV67024020
	28 MBL28.1	chr11	92797959	FAT3	ENSG00000165323	C	T	638	T	missense_variant	MODERATE	5320	4946	1649	P/L	cCg/cTg	rs375193261&COSV104590304&COSV53099591
	28 MBL28.1	chr13	24447185	PARP4	ENSG00000102699	A	G	21	G	missense_variant&splice_region_variant	MODERATE	3190	3116	1039	I/T	aTa/aCa	rs731721258&COSV67960144
	28 MBL28.1	chr13	91693325	GPC5	ENSG00000179399	C	T	367	T	missense_variant	MODERATE	890	464	155	A/V	gCg/gTg	rs553717&COSV65586039
	28 MBL28.1	chr15	85580133	AKAP13	ENSG00000170776	G	A	764	A	missense_variant	MODERATE	2272	2065	689	E/K	Gag/Aag	rs71771078&COSV63448301
	30 CLL30.1	chr2	124869717	CNTNAP5	ENSG00000155052	G	A	535	A	missense_variant	MODERATE	3752	3388	1130	V/I	Gta/Ata	rs1240405090
	30 CLL30.1	chr11	100061213	CNTN5	ENSG00000149972	C	G	145	G	missense_variant&splice_region_variant	MODERATE	1513	982	328	P/A	Ccc/Gcc	rs201910584&COSV54330963
	30 CLL30.1	chr13	24447185	PARP4	ENSG00000102699	A	G	63	G	missense_variant&splice_region_variant	MODERATE	3190	3116	1039	I/T	aTa/aCa	rs731721258&COSV67960144
	30 CLL30.1	chr20	14085630	MACROD2	ENSG00000172264	C	T	133	T	missense_variant	MODERATE	568	173	58	T/I	aCt/aTt	rs29905058&COSV53960729
	30 CLL30.1	chr22	22888259	IGLL5	ENSG00000254709	G	A	162	A	missense_variant&splice_region_variant	MODERATE	444	206	69	R/K	aGg/aAg	rs74548574&COSV73251123
	30 MBL30.2	chr2	211722434	ERBB4	ENSG00000178568	G	T	367	T	missense_variant	MODERATE	1114	842	281	A/E	gCa/gAa	
	30 MBL30.2	chr7	158138452	PTPRN2	ENSG00000155093	C	T	142	T	missense_variant	MODERATE	1128	974	325	S/N	aGt/aAt	rs113049&COSV67024020
	30 MBL30.2	chr15	85735078	AKAP13	ENSG00000170776	G	A	1813	A	missense_variant	MODERATE	7576	7369	2457	G/S	Ggt/Agt	rs224126&COSV63448267
	30 MBL30.1	chr1	81990901	ADGRL2	ENSG00000117114	G	A	326	A	missense_variant	MODERATE	4184	3968	1323	R/K	aGa/aAa	rs412929&COSV54685119
	30 MBL30.1	chr2	124524333	CNTNAP5	ENSG00000155052	C	T	416	T	missense_variant	MODERATE	1719	1355	452	S/L	tCg/tTg	rs177272618&COSV70497527
	30 MBL30.1	chr2	137160352	THSD7B	ENSG00000144229	T	A	1207	A	missense_variant	MODERATE	1687	1509	503	D/E	gaT/gaA	rs495447&COSV55680598
	30 MBL30.1	chr13	24442624	PARP4	ENSG00000102699	G	A	230	A	missense_variant	MODERATE	3583	3509	1170	T/I	aCa/aTa	rs113301501&CM160622&COSV67960155
	30 MBL30.1	chr18	21473100	GREBL1	ENSG00000141449	G	A	420	A	missense_variant	MODERATE	2080	1925	642	R/H	cGt/cAt	rs13435799561
	30 MBL30.1	chr22	46364248	CELSR1	ENSG00000075275	A	C	22	C	missense_variant	MODERATE	9233	8783	2928	I/S	aTc/aGc	

Supplementary Table 4. Non-shared variants in 120 CLL-related genes.