doi:10.1038/nature10113

Patients

The clinical and biological characteristics of the four patients selected for whole-genome sequencing (WGS) are shown in Supplementary Table 1. The patients were two males and two females with age at diagnosis ranging from 42 to 59 years. In three cases, the expression of ZAP-70 and CD38 on leukaemic cells was low and in one case was high. Two patients had *IGHV*-unmutated genes, and two *IGHV*-mutated genes (<98% homology) (Supplementary Table 1). The four CLL patients lacked common 11q (*ATM*) and 17p deletions, as well as TP53 mutations because we considered that in this initial discovery phase of the project, we should try to open the study to cases with unknown mutations. The tumour samples used for WGS were obtained in an advanced stage of the disease at the time of progression but before administration of any treatment. The interval between diagnosis and preparation of tumour samples for sequencing ranged from 0.4 to 6.1 years (Supplementary Table 1). All patients gave informed consent for their participation in the study following the International Cancer Genome Consortium (ICGC) guidelines[1].

Patients and samples for mutational screening and clinical validation

To evaluate the prevalence of mutations detected in the 4 CLL cases in which we performed WGS, pairs of tumour (≥70% of CLL cells) and normal DNA (<5% of CLL cells) (assessed by flow cytometry) were obtained from 169 CLL patients. For those mutations present in ≥3% of the patients of the screening series, additional cases were collected to perform clinical correlations. The tumour sample of all these cases had at least 30% of tumour cell content in the sample examined. Overall, 363 patients were included in the genomic studies and the clinical and biological data from these cases were retrieved from clinical files. The main features of these patients are summarized in Supplementary Table 10. All patients submitted to mutational screening and clinical validation gave their informed consent in agreement with an Institutional Review Board-approved informed consent for genetic studies.

Collection and preparation of samples

The tumour samples used for WGS were obtained from cryopreserved mononuclear cells. To purify the CLL fraction, samples were incubated with a cocktail

of magnetically labelled antibodies directed against T cells, NK cells, monocytes and granulocytes (CD2, CD3, CD11b, CD14, CD15, CD56), adjusted to the percentage of each contaminating population (AutoMACS, Miltenyi Biotec). The degree of contamination by non-CLL cells in the CLL fraction was assessed by immunophenotype and flow cytometry. When the contamination by normal cells was higher than 5%, the tumour samples were subjected to an immunomagnetic purification process (AutoMACS, Miltenyi Biotec). The final tumour cell purity of the CLL1-4 samples was 98.8%, 99.5%, 99.4% and 98%, respectively, as assessed by flow cytometry. Normal blood cells of the same patients were obtained at six months (CLL1), 3 years (CLL2), 3.7 years (CLL3) and 2 years (CLL4) after the treatment was completed. No (CLL1, CLL3) or insignificant (<0.05%) (CLL2, CLL4) evidence of tumour cells was found in these normal samples by using a sensitive minimal residual disease flow cytometry assay[2]. Whole blood was sedimented by 2% dextran and the leukocyte fraction was obtained. DNA was extracted from purified samples by using a Qiagen kit, and the quality of purified DNA was assessed by SYBR-green staining on agarose gels and quantified using Nanodrop ND-100 spectrophotometer.

Library construction and sequencing

Protocols for long and short-insert library construction and massively parallel paired-end sequencing on the Illumina/Solexa second-generation sequencing platform have been described elsewhere[3]. For whole-genome sequencing, at least two independent libraries per sample were constructed and sequenced with an average insert size of 400 bp. Two cases (CLL2 and CLL4) were sequenced using the Illumina GAIIx standard protocol with minor modifications, including fragmentation with a Covaris E210 instrument and pre-sizing, which results in a reduced number of chimeric reads and a tighter size distribution[4]; while the other two cases (CLL1 and CLL3) were sequenced using the amplification-free library protocol[5], resulting in reduced bias in GC-rich regions. For mate-pair libraries, 10 μg of genomic DNA were fragmented using a hydroshear instrument (Digilab) to achieve an insert size of 2.5 Kb. Size-fractionated DNA was processed using the Illumina kit PE-112-1002, according to the manufacturer's recommendations.

Exome-enrichment

Three μg of genomic DNA from each sample were sheared and used for the construction of a paired-end sequencing library as previously described in the Paired-End sequencing sample preparation protocol provided by Illumina[3]. Enrichment of exonic sequences was then performed for each library using the Sure Select Human All Exon Kit (Agilent Technologies) following the manufacturer´s instructions. Exon-enriched DNA was pulled down by magnetic beads coated with streptavidin (Invitrogen), and was followed by washing, elution and 18 additional cycles of amplification of the captured library. Exon enrichment was validated by real-time PCR in a 7300 Real-Time PCR System (Applied Biosystems) using a set of two pairs of primers to amplify exons and one pair to amplify an intron. Enriched libraries were sequenced using two lanes of an Illumina GAIIx.

Structural variations and DNA copy number analysis

For the detection of structural variants (SVs) we used the sequencing data from short- and long-insert libraries as previously described[6], as well as the information from three different genotyping platforms.

Computational analysis of SVs. To identify SVs we used a pipeline named PeSVFisher (Escaramis *et al.,* manuscript in preparation). A workflow describing all the process is shown in Supplementary Fig. 9. Briefly, this pipeline consists of five steps that lead to the identification of four different categories of SVs (deletions, insertions, inversions and translocations):

- First, from each BAM file, read-pairs corresponding to a single paired-end or mate–pair library are extracted using SAMtools. Unpaired reads and unmapped read-pairs are pulled out taking advantage of the flag field.

- Second, reads are assigned to three different files based on the following criteria: (1) read-pairs mapped in right order and right orientation, (2) read-pairs mapped in right order but wrong orientation (i.e., both reads mapped in either forward or reverse strands), (3) read-pairs mapped in different chromosomes. In this step, we discard read-pairs matching regions in the reference genome with a phred-based quality score lower than 30 at either end. This will ensure a mapping accuracy of at least 99.9%.

- Third, file (1) is used to construct the empirical distribution of the insert size (L). This will help to define cut-offs to discriminate concordant read-pairs (those falling into

the expected range) from discordant read-pairs. An upper cut-off (UC) and a lower cut-off (LC) are calculated as the 0.995 and 0.005 percentiles of the L distribution.

- Fourth, a clustering procedure is carried out to group indicators pointing to the same SV category. This clustering algorithm is based on Mario Cáceres approach to detect genomic inversions (Martinez *et al.*, in preparation) which we have adapted to call further types of SVs. The algorithm is based on geometrical rules in addition to the following conventional criteria: (i) Putative deletions are derived from overlapping read-pairs with insert sizes larger than UC. (ii) Putative insertions are derived from overlapping read-pairs with insert sizes smaller than LC. (iii) Putative inversions are derived from overlapping read-pairs mapped into the same strand. (iv) Putative translocations are derived from overlapping read-pairs mapped onto the same strands and chromosomes. In all four cases, any read-pair that has been included in one cluster is not considered for any other cluster.

- Fifth, predictions of presumed SVs are computed from each cluster that contains at least two read-pairs. Breakpoints are predicted as an interval where the actual breakpoint can occur. Two breakpoints are considered for the case of deletions, inversions and translocations, whereas a single point is predicted for the case of insertions.

Somatic SVs identification. To find out somatic SVs, we extracted those variants that involved genes present in the cluster files of tumour samples with at least 4 read-pairs and found in the normal samples with at most 2 read-pairs. The stringency of the analysis eliminated those structural variants that contained repeat sequences and segmental duplications. It is likely that this stringent approach may lead to the loss of several structural variants potentially associated with the evolution of CLL. The full characterisation of the multiple additional changes and the variability of SV associated with CLL will be further investigated in a follow-up study.

Agilent 1M arrays. Tumour and normal DNA (1 μg) from the same patient were fluorescently labelled and hybridized according to the manufacturer's protocol. After hybridization to an Agilent 1M array, slides were washed and fluorescence was assessed using an Agilent microarray scanner G2565CA (Agilent Technologies). Raw data were generated from scanned images using Agilent Feature Extraction Software (v10.7). Log2ratios of background corrected values for tumour over normal DNA were calculated. Normalization was carried out on Agilent's CGH Analytics, integrated on the Genomic Workbench suite (v5.0), using data from the internal set of control probes

included in the microarray. Detection of Copy Number Alterations (CNAs) was performed using the ADM-2 algorithm, also implemented within the Agilent's genomics suite Genomic Workbench v5.0, with a threshold of 6.5 and a minimum of 5 consecutive probes. Post-hybridization quality control reports included DLRspread values, signal intensity, background intensity in each channel, signal-to-noise ratio, and reproducibility. Any array with DLRspread over 0.3 was considered as low quality and consequently discarded. Array CGH analysis was outsourced to qGenomics (www.qgenomics.com/).

Affymetrix SNP6.0 arrays. Tumour and normal DNA (500 ng) were digested with *NspI* and *StyI* restriction enzymes and processed according to standard protocols for Affymetrix Genome-Wide Human SNP 6.0 microarrays. Arrays were washed using Affymetrix fluidics station and scanned with the Gene Chip Scanner 3000. Image data were analysed with Genotyping Console 4.0 to obtain the CEL data files, which were also exported to the Partek Genomic Suite (Partek Inc.) for data analysis and visualization. Quality controls assessed by Genotyping Console included Contrast Quality Control (CQC>0.4) produced by Birdseed and Median of the Absolute values of all Pairwise Differences (MAPD<0.35). Tumour and normal DNA samples were analysed individually with the Genotyping Console 4.0 using regional GC correction and segment reporting tool filters: minimum of 5 markers per segment and 100 Kb minimum genomic size of a segment, as well as by visual inspection. Paired analysis using constitutional DNA was obtained by the Partek software package in order to exclude inherited copy number variants (CNV). CNAs were scored with a Hidden Markov Model (HMM) and the segmentation method included in Partek's Suite. *Affymetrix SNP6.0* genotyping was outsourced to CeGen (www.cegen.org).

Illumina HumanOmni1-Quad arrays. Tumour and normal DNA samples (200 ng) were genotyped using Illumina HumanOmni1-Quad Beadchips following the manufacturer's standard recommendations. Illumina's BeadArray Reader was used to analyse fluorescence signals from each BeadChip. The fluorescence intensities analysis and decoding of SNP position were performed using Illumina's BeadScan software. Genotype calling was performed with GenomeStudio software (v2010.1) using the standard cluster file provided by Illumina (HumanOmni1-Quad_v1-0_B.egt). The GenCall score cut-off was 0.15. All samples passed Illumina internal controls. Normalized total signal intensity ratios (LRR) as well as normalized allelic intensity ratios (BAF) were also obtained from GenomeStudio software to further call CNVs.

CNV calls of tumour and normal DNA samples individually was performed with the PennCnv program using a wave adjustment procedure for genomic waves via the gcmodel argument and a filter of 5 markers per segment. Paired analysis was performed based on the ratio of tumour LRR values to normal LRR values. CNAs were then obtained using GADA package within R environment using the recommended settings. *Illumina HumanOmni1-Quad* genotyping was outsourced to CeGen (www.cegen.org).

Mapping and initial base calls

Reads from each library were mapped to the human reference genome (GRCh37) using BWA[7] with the sampe option, and a BAM file was generated using SAMtools[8]. Reads from the same paired-end libraries were merged, and optical or PCR duplicates were removed using Picard (http://picard.sourceforge.net/index.shtml). Finally, all BAM files from the same sample were merged using Picard. Bases were initially called using the Maq consensus model implemented in SAMtools. Statistics about the number of mapped reads, genome physical coverage and depth of coverage for each sample are shown in Supplementary Tables 2, 3 and 4.

Identification of somatic substitutions

To determine the base calling accuracy, we used the genotyping information from two independent genotyping platforms performed on the same samples used for WGS, Affymetrix 6.0 and Illumina OmniQuad, using a similar approach as that described for AML[9]. These platforms have 275,976 SNPs in common, and 275,012 of them were also present in the whole-genome data, resulting in an estimated coverage of 99.65%. Both platforms have the same call in 269,240 out of 270,367 positions common to them and whole-genome data (99.58% agreement between both platforms). Those positions with the same call in both platforms were called hqSNPs and used to determine the calling accuracy from whole-genome data. We found that 99.78% of the hqSNPs were called correctly by SAMtools (268,653 out of 269,240), and 99.37% of heterozygous hqSNPs were identified (78,594 out of 79,093), suggesting that both the coverage and calling accuracy was sufficient to identify somatic mutations present in heterozygosis. We estimated the false discovery rate to be about 0.02%. To identify somatic mutations we developed a mutation caller named *Sidrón* (Quesada *et al.,* to be published elsewhere),

which uses a binomial probabilistic model similar to other mutation callers recently described[10]. Basically, for every single variant base initially identified in the tumour genome the program calculates a probability for that base not being a reference base for both tumour and normal DNA, taking into account the number of reads supporting the variant and the base quality. Then, genotyping information for each sample is used to define thresholds for calling non-reference bases in the tumour and in the normal. Those bases considered non-reference in the tumour and reference in the normal DNA are considered somatic mutations. The procedure is detailed below. The program consists of five steps that lead to the identification of somatic mutations:

- First, the information from the hqSNPs is used to calibrate the sequencing error. Briefly, in a homozygous hqSNP, discordant bases identified in the pileup are likely due to sequencing artifacts. Thus, homozygous hqSNPs are used to compute the probability of getting a nucleotide read N_{read} if the read base quality is Q and the real genomic base is N_{ref} ($p(N_{read} \mid N_{ref}, Q)$).

- Second, the information from the heterozygous hqSNPs is used to determine the minimum and maximum coverage, as well as minimum mapping quality and SNP quality to identify >99% of all heterozygous hqSNPs, and these settings are applied to the whole tumour pileup.

- Third, for every variant base in the tumour genome, *Sidrón* computes the probability of getting the corresponding pileup line (C) given two possible genotypes: homozygous for the most frequent base (Hz) and heterozygous with the most frequent and second most frequent bases (Het). The output is the score S:

$$S = \log\left[\frac{p(C \mid Het)}{p(C \mid Hz)}\right]$$

This score is calculated in each variant position for the corresponding tumour and normal pileup lines using the values of $p(N_{read} \mid N_{ref}, Q)$ obtained for each one in the first step.

- Fourth, genotypes are called based on S, with cut-off values determined empirically from hqSNPs and validation data. We found optimal sensitivity and specificity by calling Het with S>14 and Hz with S<-4 in high coverage positions (>20) and Het with S>8, Hz with S<-2 otherwise. Borderline variants were manually inspected.

- Fifth, candidate somatic mutations are subjected to quality controls to filter out those owing to common sequencing and alignment artefacts by inspecting the reads supporting them. Thus, reads that supported a mutation by showing a novel stretch of more than ten consecutive identical nucleotides were not considered. Reads that could be better aligned to different sites in the genome were also eliminated. Reads present in repetitive regions of the genome were discarded. Finally, mutations unevenly distributed in the supporting reads were mostly due to the presence of small indels, and were filtered out.

We classified the somatic mutations into three different classes according to their potential functional effect: class 1, non-synonymous changes, indels causing frameshifts in coding regions and mutations affecting critical splicing sites; class 2, synonymous mutations and those located in untranslated regions (UTRs); and class 3, comprising all remaining mutations.

We were able to validate more than 96% (83 out of 86) of the identified class 1 and 2 variants for which PCR and capillary sequencing was obtained. Capillary sequencing of 384 random mutations (96 per case) revealed that more than 96% of them were true somatic mutations. In addition, the use of a combined strategy of WGS and exome sequencing was helpful to try to determine the sensitivity of this analysis, as the high depth of coverage in exome sequencing simplifies the identification of variants which might be difficult to detect at lower coverage. Using the WGS data we were able to identify 37 of the 42 somatic mutations detected and validated by exome sequencing, suggesting that the sensitivity of our method is around 88%, although the reduced number of mutations per tumor makes difficult to precisely estimate this parameter.

Somatic mutation identification in pooled samples (SMIPS)

To discover recurrent somatic mutations in CLL, we performed a screening of somatic mutations in a set of 169 additional CLL cases using a combination of pooled samples, PCR amplification and high-throughput sequencing. Expressed genes with class 1 somatic mutations detected and validated in one CLL sample using either whole-genome or exome sequencing were considered as candidates for driver mutations. We used a modified method for the analysis of pooled samples[11]. First, we pooled a normalized amount of genomic DNA from normal and tumour samples of CLL patients.

Then, we amplified exons of interest and pooled them again to be sequenced using Illumina Genome Analyzer. Genomic DNAs were quantified by Quant-iT dsDNA Broad Range assay (Invitrogen) and Nanodrop ND-1000 spectrophotometer. Two pools were made with either 88 cases (1100 ng per individual – Pool 1) or 84 cases (630 ng per individual – Pool 2), including three of the cases subjected to WGS. Normal and tumour samples were pooled separately in 4 different pools (1N, 1T, 2N and 2T). Final concentrations for these pools were 70.3 ng/μL (Pool 1N), 70.6 ng/μL (Pool 1T), 49.9 ng/μL (Pool 2N) and 48.8 ng/μL (Pool 2T).

PCR primers (Supplementary Table 11) were designed using Primer 3 (http://frodo.wi.mit.edu/primer3), and purchased from Sigma-Aldrich. PCR reactions were performed to amplify 200 amplicons, including 236 exons and spanning a total of 103,940 bp in each of four genomic DNA pools. These PCRs and subsequent pooling and sequencing were made in 3 different rounds of validation corresponding roughly to 80 amplicons and 40 Kb per round. Each target locus was amplified using 50 ng from each pooled DNA, corresponding to approximately 100 diploid genomes per case. Amplification was performed using Platinum Pfx DNA Polymerase (Invitrogen), and the reaction mix contained 1X final concentration of 10X Pfx Amplification Buffer (Invitrogen), 300 μM dNTPs (Invitrogen), 300 nM forward primer, 300 nM reverse primer, 1 mM $MgSO_4$ and 1 unit Platinum Pfx DNA polymerase in a final reaction volume of 50 μL. PCR conditions were 1 cycle of 2 min at 94 °C, followed by 35 cycles of 15 sec at 94 °C, 30 sec at 60 °C, and 30-60 sec (depending on amplicon length, 60 sec if larger than 500 bp) at 68 °C. If amplification failed, PCR was repeated with 2X final concentration of 10X Pfx Amplification Buffer or using Expand Long Polymerase (Roche). PCR products were then purified using QIAquick PCR Purification Kit (QIAGEN) and quantified using Nanodrop ND-1000 spectrophotometer.

From each pool, equimolecular amounts (40 x 10^{10} molecules) of each amplicon were pooled. Then, random ligation of amplicons was performed in order to obtain concatemers using 1X T4 DNA Ligase Buffer (New England Biolabs), 2400 units of T4 DNA ligase (New England Biolabs), 120 units T4 polynucleotide kinase (New England Biolabs), 15% (wt/vol) polyethylene glycol 8000 MW (Sigma-Aldrich-Fluka). The final volume was divided into four parts and incubated at 22 ºC for 17 h, followed by incubation at 65 ºC for 20 min. Agarose gel electrophoresis was performed with 1 μL of ligated and non-ligated pools to confirm the concatenation of amplicons. To remove

polyethylene glycol, we sedimented the ligated pools in a microcentrifuge at 15,000 x g and room temperature for 30 min in the presence of 10 mM $MgCl_2$ as previously described[12]. Pellets were washed twice with 70% ethanol followed by centrifugation to remove the ethanol. Once dried, DNA was dissolved in 50 µL of TE and quantified. Three µg of each pool were individually sonicated in 6 x 16 mm round bottom glass microtubes (Covaris) using the Covaris S2 sonicator (200 Cycles, Duty Cycle 10%, intensity 5 and 360 sec) (Covaris). Fragmentations were confirmed by Bioanalyzer 2100 (Agilent Technologies), and DNA libraries were prepared as previously described following the Paired-End sample preparation protocol from Illumina.

Analysis of SMIPS data

Reads were mapped to the reference genome with BWA using the same procedure as outlined above, but omitting the removal of PCR duplicates. The coverage per base was between 50,000-80,000X. A pileup file was generated with SAMtools, and only those bases with base quality ≥30 were used for the analysis. Due to the high depth of coverage, errors in the sequencing or introduced by the DNA polymerase during PCR amplification or library construction can be detected using this approach. In the PCR reaction for each amplicon, we used about 570 pg of DNA per case, which represents more than 100 diploid genomes for that particular patient. In the case that a mutation is introduced by the DNA polymerase during PCR, that mutation would be present in 1 out of 200 allele copies for that patient, or 1 out of >17,000 copies for the whole pool. However, a real mutation in heterozygosity would be present in about 100 of the 200 copies for that patient, or in 1 of the ~174 alleles present in the pool. Therefore, it would be possible to discriminate between real variants and PCR or sequencing errors as real variants should be supported by approximately one allele in the whole pool. To confirm that this method is able to estimate the correct number of alleles for a variant, we calculated the allele frequency for those SNPs present in our sample, with corresponding entries in dbSNP131, and for which the frequency in the Central European (CEU) population was known. We then compared the estimated allele frequency obtained from our analysis of 169 cases with those reported for the CEU population for 72 SNPs (Supplementary Fig. 10). We obtained a correlation coefficient r^2=0.974 between both datasets, suggesting that using this method we are able to

correctly estimate the number of alleles with a certain variant present in the pooled sample.

To identify somatic substitutions, we extracted those variants present in the pool of tumour DNAs with an allele frequency estimation of more than 0.5, and for which there were less than 0.2 alleles in the pool of normal DNAs. We selected these cut-offs because the tumour DNAs included in the pool had some degree of normal cell contamination (up to 30%), and normal DNA was obtained from peripheral blood leukocytes and it was required to have less than 5% contamination with tumour cells. In most cases some tumour cell contamination below this limit was present in the normal sample. Using these cut-offs, we were able to identify all somatic substitutions analysed and present in cases CLL1, CLL3 and CLL4, which had been included in the pool as positive controls, demonstrating the sensitivity of the analysis. We detected germline or somatic variants with high sensitivity (up to one mutant allele in a pool of 182 alleles and >93% of the somatic mutations from the three cases included as positive controls) and high specificity (>82% of mutations detected by SMIPS were validated as somatic by Sanger sequencing in tumour and normal DNA).

The identification of small indels in pooled samples with very high depth of coverage but present in a small fraction of alleles is challenging. We took advantage of the extended CIGAR field present in the BAM format[8], as this annotation reveals the presence of an indel in an individual read. We used custom scripts to process the CIGAR fields and compare the frequency of insertions and deletions between tumour and normal pools at each analysed position. We were able to correctly identify the somatic deletions present in *CD8A* and *NOTCH1*, although no additional indels were detected in any of the analysed genes.

Gene expression analysis

Total RNA was extracted with the TRIzol reagent following the recommendations of the manufacturer (Invitrogen Life Technologies). RNA integrity was examined with the Agilent 2100 Bioanalyzer (Agilent Technologies) and only high quality RNA samples were hybridized to Affymetrix GeneChip Human Genome U133 plus 2.0 arrays, according to Affymetrix standard protocols. The analysis of the scanned images and the determination of the detection call for each probe set of the array were obtained with the GeneChip Operating Software (GCOS, Affymetrix). Summarized expression values were computed using the robust multichip average (RMA) approach

implemented in the Expression Console software (Affymetrix). The supervised analysis was performed with the BRB-Array Tools, v.3.6.0 software. Previous to the supervised analysis of CLL samples, we selected the probe sets that were present in at least 10% of the cases (n=31,989). The differential gene expression analysis was performed using the Significance Analysis of Microarrays Data (SAM) method implemented in BRB-Array Tools. The level of significance to detect differentially expressed genes was a 90th percentile False Discovery Rate (FDR) <0.05. To identify potential KEGG pathways (http://www.genome.jp/kegg/pathway.html) that could show a differentially expression between *NOTCH1*-mutated and -unmutated CLL cases, we performed a gene set analysis applying three different tests, LS permutation test, Efron-Tibshirani's GSA maxmean test, and Goeman's global test, as implemented in BRB-Array Tools. To reduce multiple probe sets to one per gene symbol, the different probe sets for a gene were collapsed into a single vector of values using the most variable probe set measured by the interquartile range across arrays reducing the list to 13,239 genes. The pathways that were significant at level P<0.02 in the three enrichment tests performed were considered to be differentially expressed between classes. A t-test was performed to identify the genes included in the NOTCH1 pathway from KEGG that had a differential expression between *NOTCH1*-mutated and -unmutated tumours**.**

Immunoprecipitation and Western blotting

Tumour cells were lysed for 30 min in Triton buffer (1% Triton X-100, 50 mM Tris–HCl pH 7.6, 150 mM NaCl, 1 mM EDTA) supplemented with protease and phosphatase inhibitors (1 mM PMSF, 2 mM sodium pyrophosphate, 2 mM sodium beta-glycerophosphate, 1 mM sodium fluoride, 1 mM sodium orthovanadate, 10 μg/mL leupeptin and 10 μg/mL aprotinin). For the detection of phospho-STAT3, cells were lysed in CHAPS buffer (1% CHAPS, 100 mM NaCl, 5 mM Na_2HPO_4 and 2.5 mM EDTA). Lysates were cleared by centrifugation at 15,000 x g at 4 °C for 15 min, and protein concentrations were determined using the Bradford method. Fifty μg of protein were separated by SDS-polyacrylamide gel electrophoresis and transferred onto Immobilon-P membranes. Proteins were detected by using the following antibodies: anti-NOTCH1 (D1E11), cleaved NOTCH1 (Val1744) (D3B8), anti-IRAK1 (D51G7), anti-MyD88 (D80F5), anti-phospho-IκBα, anti-phospho-STAT3 (Tyr705) (D3A7), anti-STAT3 (79D7) (Cell Signaling Technology), anti-phospho-p65 (Ser536), anti-

cyclin D2 (M-20) (Santa Cruz Biotechnology) and anti-IκBα (Merck, Calbiochem). Antibody binding was detected using horseradish peroxidase-labelled anti-mouse (Sigma) or anti-rabbit (Cell Signalling Technology) antibodies and chemiluminescence was detected using a LAS4000 device (Fuji). Equal protein loading was confirmed with antibodies against β-actin or α-tubulin (Sigma). For MyD88 immunoprecipitation, CHAPS buffer was used (1% CHAPS, 100 mM NaCl, 5 mM Na_2HPO_4 and 2.5 mM EDTA) supplemented with protease and phosphatase inhibitors. Protein extracts were incubated overnight at 4 °C with an anti–MyD88 antibody (D80F5) and then, protein-A beads were added for one additional hour. Supernatant (nonimmunoprecipitated fraction) was recovered by centrifugation, and protein-A beads (immunoprecipitated fraction) were washed three times with CHAPS buffer.

NF-κB p65 DNA-binding analysis

Nuclear extracts were obtained from primary CLL cells and assayed for NF-κB p65 activity using the TransAM NF-κB Chemiluminescence kit (Active Motif). Two micrograms of nuclear extracts were incubated according to the manufacturer's protocol into 96-well plate coated with an oligonucleotide containing the NF-κB consensus DNA-binding site (ELISA-based method). DNA binding was detected by incubating with an antibody against p65 followed by a horseradish peroxidase conjugated secondary antibody. The acquisition and the quantification of the signal were done on a LAS4000 device as above.

TLR stimulation assay

A TLR stimulation assay was performed using tumour cells from either two patients with the L265P *MYD88* mutation, or from four patients with wild type *MYD88*, and purified B-lymphocytes from two patients with the previously reported inactivating *MYD88* mutation E52DEL[13]. All the samples contained more than 95% B-lymphocytes and they were cultured in Ex-vivo 15 medium (Lonza) for 48 h with TLR agonists or IL-1beta (IL-1β 10 ng/ml, R&D Systems Europe), as previously described[14]. TLR agonists were Pam3CSK4 1 μg/mL for TLR1/2, HKLM 10^8 cells/mL for TLR2, poly(I:C) LMW 10 μg/mL for TLR3, *Escherichia coli* K12 LPS 1 μg/mL for TLR4, *Salmonella typhimurium* flagellin 1 μg/mL for TLR5, FSL1 1 μg/mL for TLR6/2,

Imiquimod 1 μg/mL for TLR7, ssRNA40 1μg/mL for TLR8, and ODN2006 5 μM for TLR9. These TLR agonists were from Human TLR1-9 Agonist KitTM (InvivoGen). For all incubations except that with LPS, cells were first incubated with 10 μg/mL of polymyxin B (Sigma, P-4932) for 20 min. *In vivo* cytokine production was assessed in tissue culture supernatants using fluorescence-activated cell sorter analysis (Luminex 100 System) Cytokine Human 25-plex panel (Invitrogen), a multiplexed sandwich immunoassays based on flow-cytometry Luminex technology, was used to measure 25 cytokines (IL-1β, IL-1RA, IL-2, IL-2R, IL-4, IL-5, IL-6, IL-7, IL-10, IL-12 (p40), IL-13, IL-15, IL-17, TNF-α, IFN-α, IFN-γ, GM-CSF, MIG, CXCL8/IL-8, CXCL10/IP-10, CCL2/MCP-1 CCL3/MIP-1α, CCL4/MIP-1β, CCL5/RANTES and CCL11/Eotaxin). Data were analysed with the Luminex software.

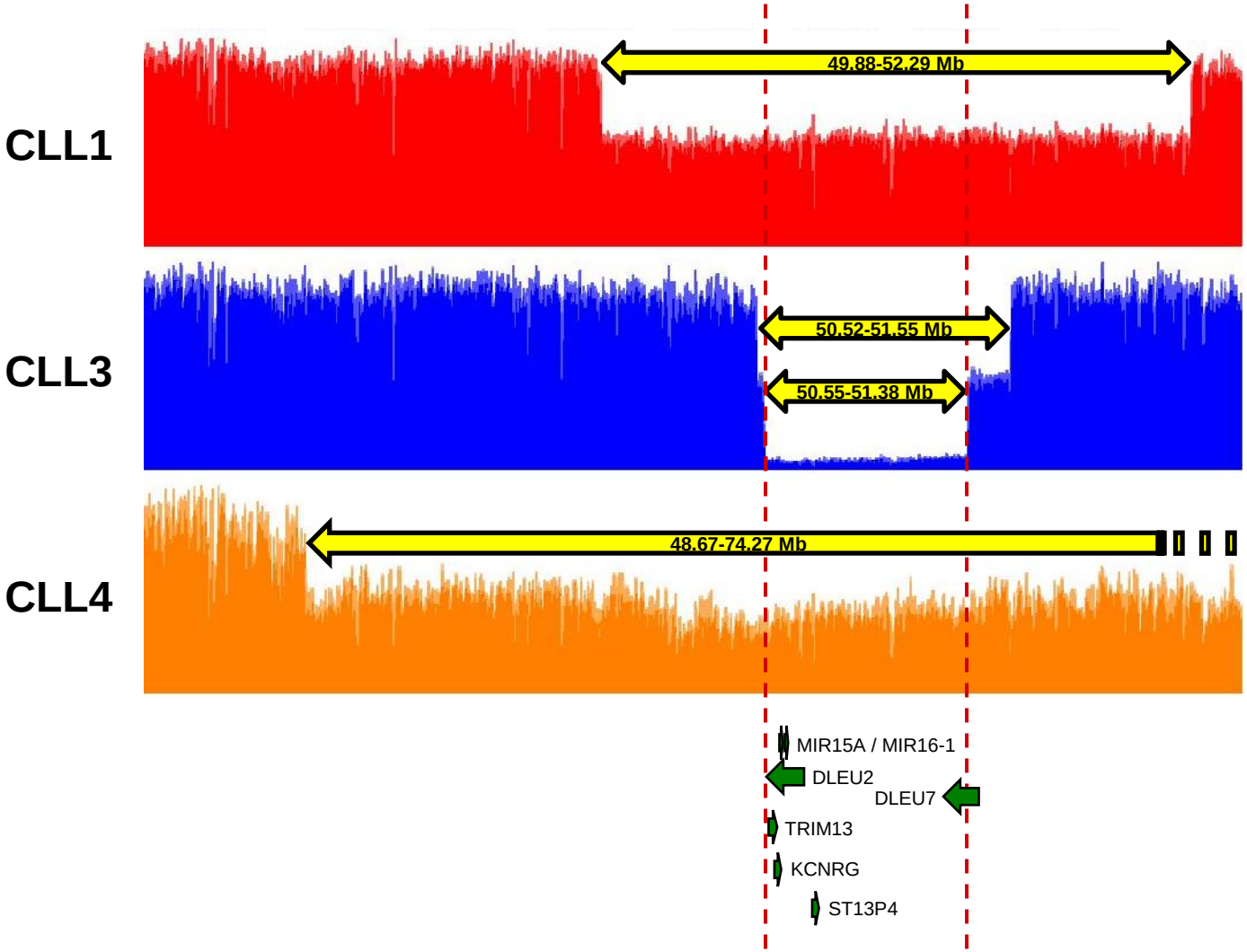
Statistical analysis

The SPSS Statistics 17.0 (SPSS Inc) package was employed to correlate clinical and biological variables by means of Fisher's test or non-parametric test when necessary. Survival curves were analysed according to the Kaplan and Meier method and compared using the log-rank test[15]. In a Cox regression analysis comparing *NOTCH1* mutations and transformation to DLBCL only the later ($P<0.001$) maintained its prognostic impact on overall survival, suggesting that the adverse survival impact of *NOTCH1* mutation in CLL may be due in part to the higher risk of transformation into DLBCL. All statistical tests were two-sided and the level of statistical significance was 0.05.

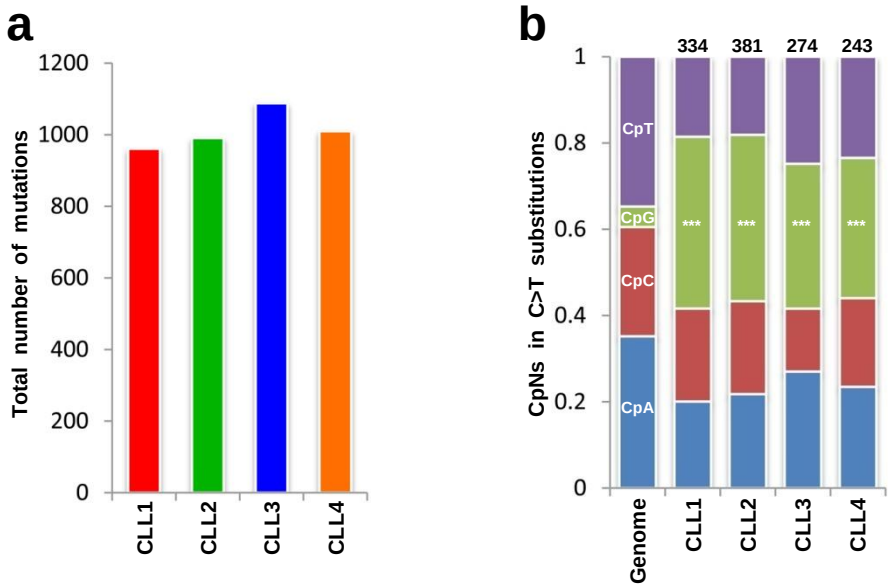
References

1. Hudson, T. J. *et al.* International network of cancer genome projects. *Nature* **464**, 993-998 (2010).
2. Rawstron, A. C. *et al.* International standardized approach for flow cytometric residual disease monitoring in chronic lymphocytic leukaemia. *Leukemia* **21**, 956-964 (2007).
3. Bentley, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53-59 (2008).
4. Quail, M. A. *et al.* A large genome center's improvements to the Illumina sequencing system. *Nat Methods* **5**, 1005-1010 (2008).
5. Kozarewa, I. *et al.* Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. *Nat Methods* **6**, 291-295 (2009).
6. Campbell, P. J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109-1113 (2010).
7. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).

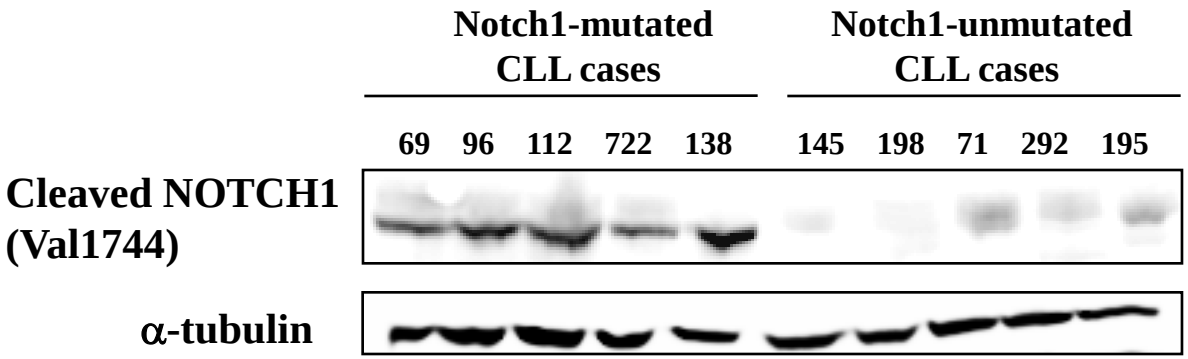
8. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-2079 (2009).
9. Ley, T. J. *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* **456**, 66-72 (2008).
10. Goya, R. *et al.* SNVMix: predicting single nucleotide variants from next-generation sequencing of tumors. *Bioinformatics* **26**, 730-736 (2010).
11. Druley, T. E. *et al.* Quantification of rare allelic variants from pooled genomic DNA. *Nat Methods* **6**, 263-265 (2009).
12. Paithankar, K. R. & Prasad, K. S. Precipitation of DNA by polyethylene glycol and ethanol. *Nucleic Acids Res* **19**, 1346 (1991).
13. von Bernuth, H. *et al.* Pyogenic bacterial infections in humans with MyD88 deficiency. *Science* **321**, 691-696 (2008).
14. von Bernuth, H. *et al.* A fast procedure for the detection of defects in Toll-like receptor signaling. *Pediatrics* **118**, 2498-2503 (2006).
15. Peto, R. & Pike, M. C. Conservatism of the approximation sigma (O-E)2-E in the logrank test for survival data or tumor incidence data. *Biometrics* **29**, 579-584 (1973).



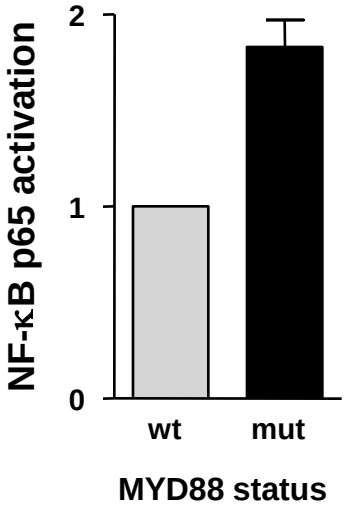
Supplementary Figure 1. Identification of recurrent deletions of chromosome 13q in cases CLL1, CLL3 and CLL4. WGS coverage with local regression smoothing of region 48 Mb to 52.5 Mb. Yellow arrows indicate the deleted region and the chromosomal breakpoints. For case CLL3, two independent deletion events are shown, and the corresponding breakpoints are also indicated. Dashed lines represent the minimum common region deleted in the three cases. Genes present in this region are shown at the bottom.



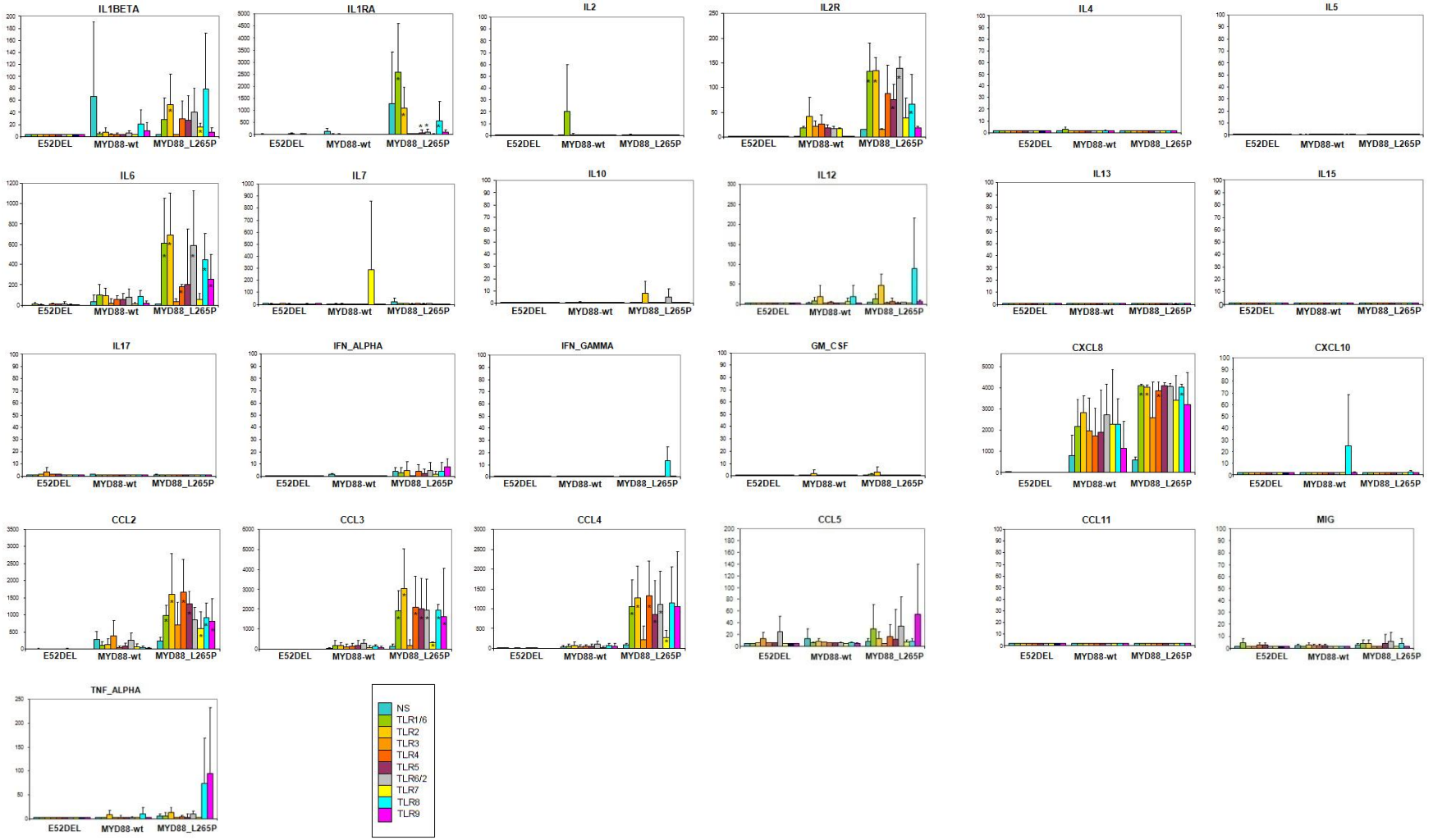
Supplementary Figure 2. a, Total number of somatic substitutions per genome. **b**, Distribution of the four possible CpN dinucleotides for the C>T transition in each tumour genome compared to the expected distribution across the genome. The total number of C>T substitutions per case is indicated at the top (***, $P<10^{-42}$).



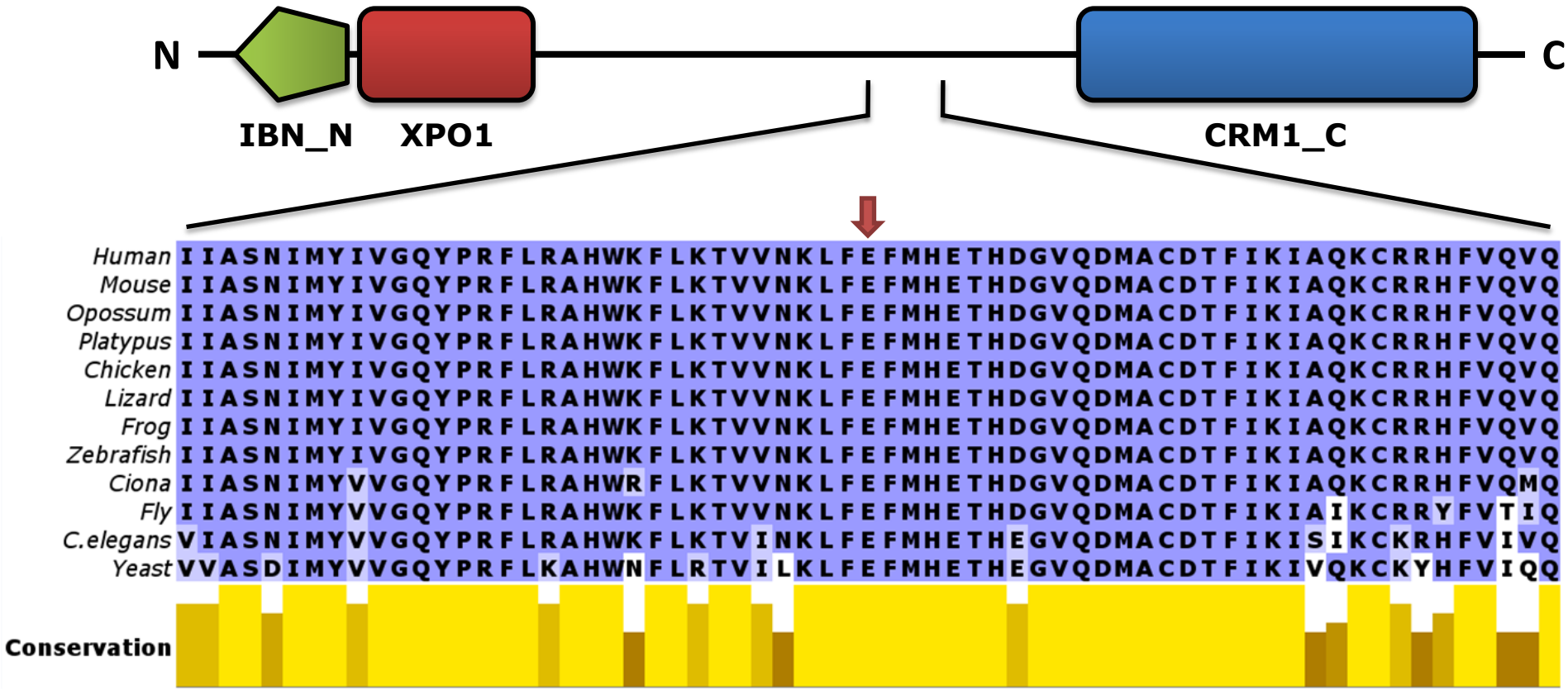
Supplementary Figure 3. Western blot showing NOTCH1 protein levels in CLL cases with or without the NOTCH1 p.P2515Rfs*4 mutation using an antibody specific for the cleaved form of NOTCH1.



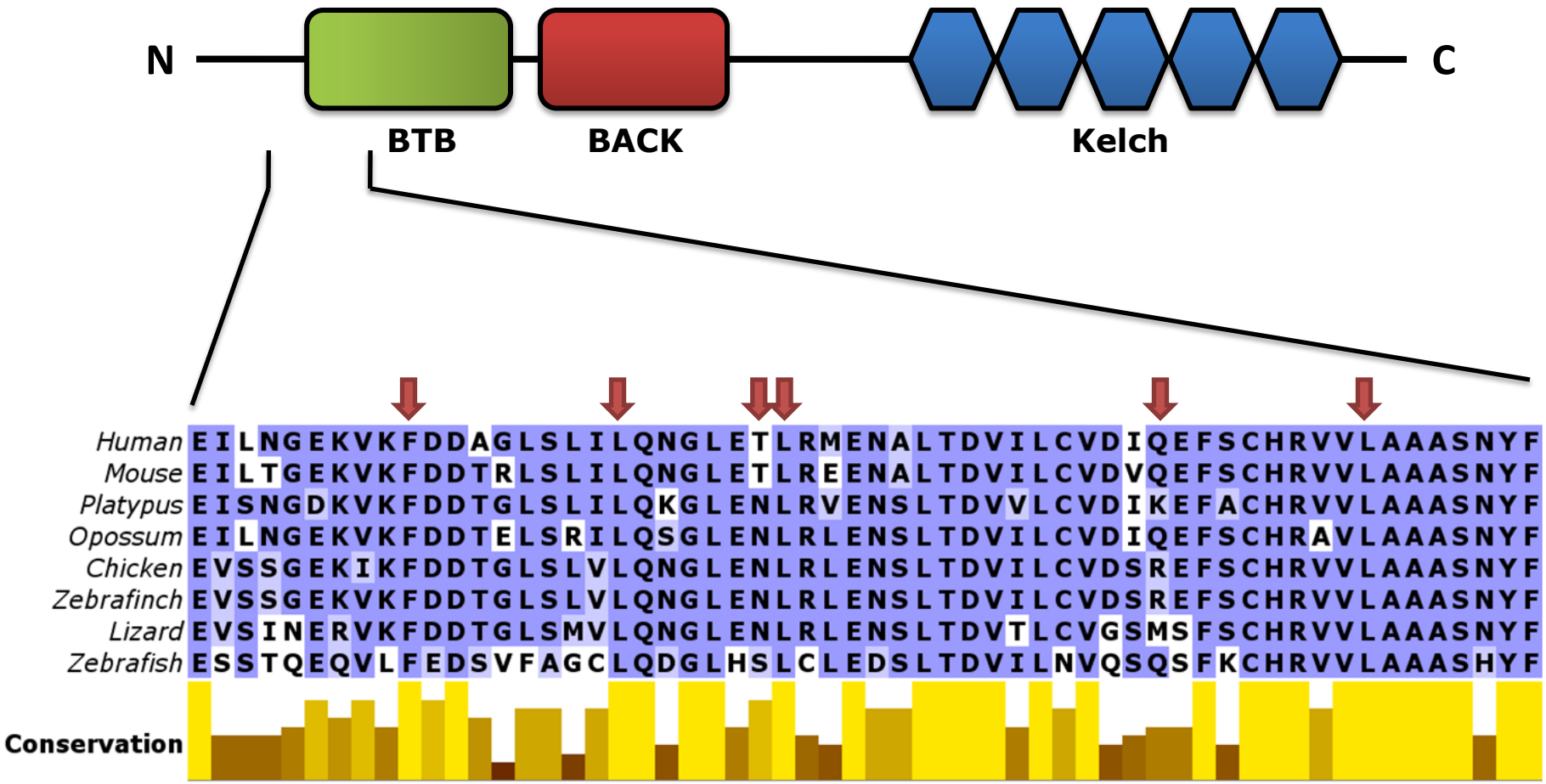
Supplementary Figure 4. p65 DNA-binding activity was analyzed by ELISA-based chemiluminescence in nuclear extracts obtained from 3 *MYD88*-mutated CLL cases and 3 *MYD88*-unmutated cases. DNA binding of mutated cases is expressed relatived to the mean of unmutated cases.



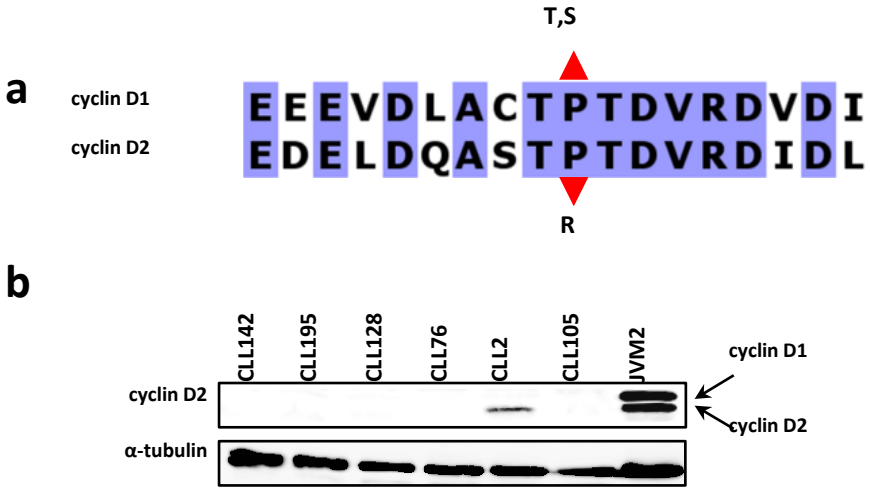
Supplementary Figure 5. Bar graphs displaying the median values of the cytokine levels detected in B-cell lymphocytes from patients with an inactivating *MYD88* mutation (E52DEL), CLL patients without *MYD88* mutation (*MYD88*-wt), and CLL patients carrying a mutated *MYD88* (L265P). In different colours are represented the stimulation experiment for each of the TLRs. Statistically differences between CLL patients with the *MYD88* (L265P) mutation *vs*. CLL patients without *MYD88* mutation are indicated with an asterisk (Mann-Whitney test, $P<0.05$).



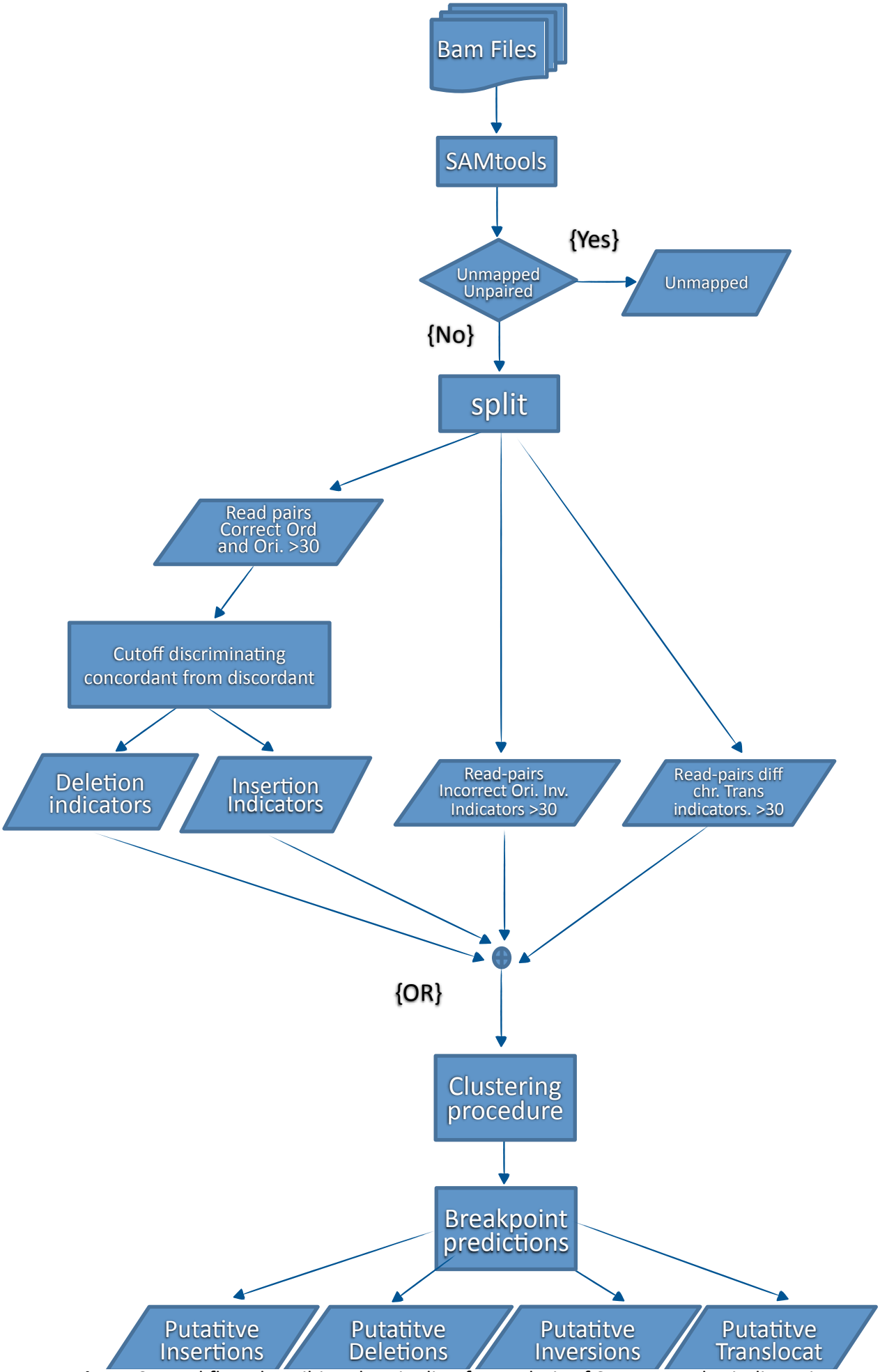
Supplementary Figure 6. Schematic representation of exportin 1 and multiple sequence alignment of the region containing the recurrent mutation identified in CLL tumours affecting residue E571 (indicated with a red arrow). IBN_N: importin-beta N-terminal domain; XPO1: XPO1 N-terminal domain; CRM1_C: CRM1 C-terminal domain.



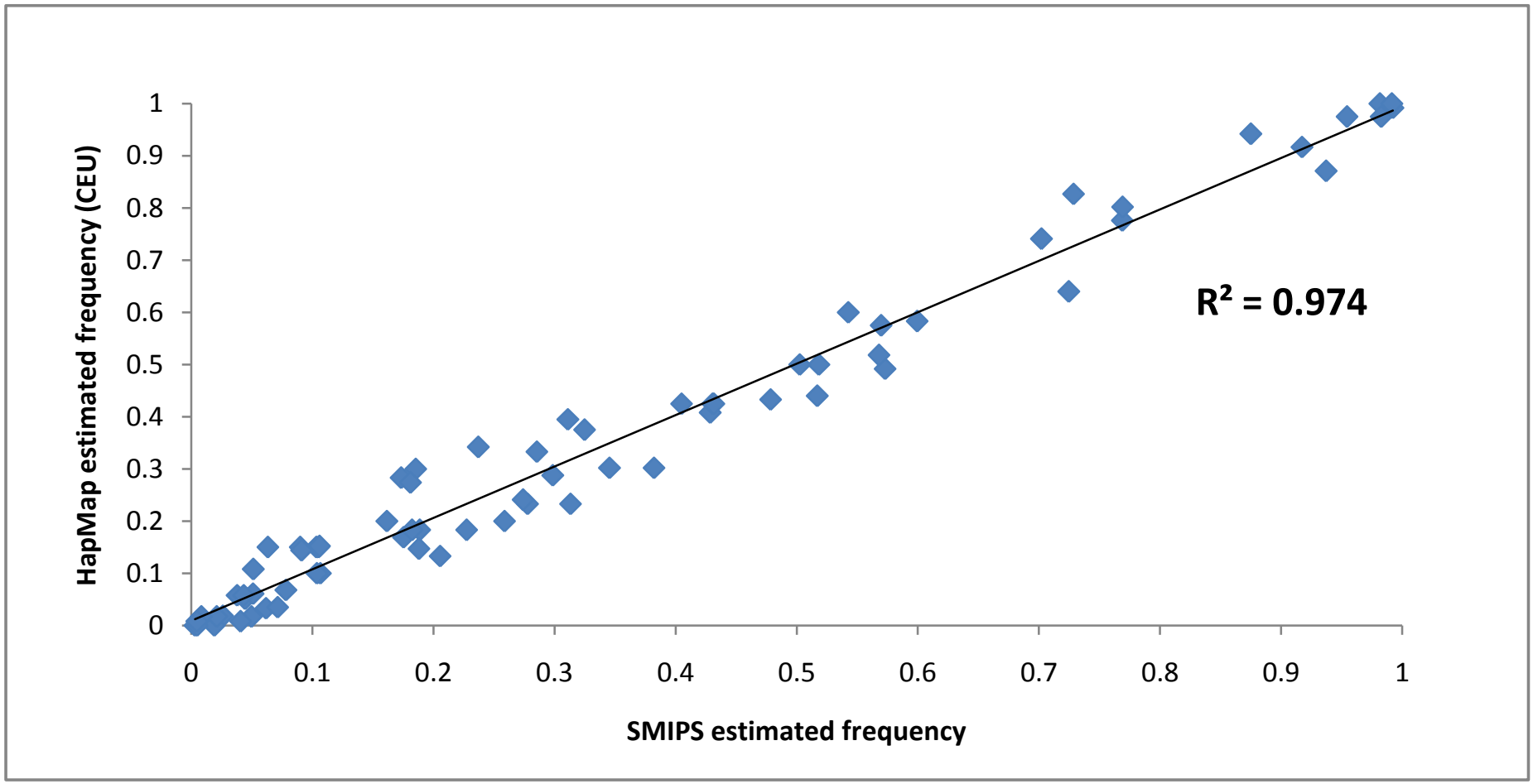
Supplementary Figure 7. Schematic representation of the domain structure of KLHL6 and multiple sequence alignment showing the location of six somatic mutations (red arrows) identified in CLL tumours. BTB: Broad-Complex, Tramtrack and Bric a brac domain; BACK: BTB and C-terminal Kelch domain.



Supplementary Figure 8. a, Sequence alignment of the C-terminus of human cyclin D1 and cyclin D2. Arrows indicate the somatic mutation identified in CLL in cyclin D2, and the two somatic mutations identified in endometrial cancer with the corresponding substitutions. **b**, Western blot analysis of cyclin D2 in tumour cells from six CLL patients, including patient CLL2 with the p.P281R mutation. Extracts from the cell line JVM2 were used as control.



Supplementary Figure 9. Workflow describing the pipeline for analysis of SV. Rectangles indicate intermediate processes; diamonds show condition steps and rhomboids are files generated within the process.



Supplementary Figure 10. Correlation between HapMap- and SMIPS-estimated allele frequencies for 72 SNPs included in the analysed regions.

Supplementary Table 1. Clinical and biological features at diagnosis, evolution and sample characteristics of the four CLL patients studied by whole-genome sequencing

	CLL1	CLL2	CLL3	CLL4
Age	44	59	45	42
Gender	Male	Female	Male	Female
Stage at diagnosis	A(0)	A(0)	C(IV)	A(0)
Lymphocytes x10^9/L	8.2	22	201.9	10
Haemoglobin g/L	139	132	91	151
Platelets x10^9/L	200	360	75	214
LDT	>1 year	>1 year	NA	>1 year
LDH (U/L)	338	299	328	360
β_2-microglobulin mg/L	ND	1.2	5.3	1.4
PB morphology	Atypical **	Typical	Typical	Typical
Immunophenotype	Typical	Typical	Typical	Typical
Bone marrow biopsy	Interstitial	Interstitial	Diffuse	Nodular+Interstitial
ZAP70 (%)	8	39	12	2
CD38 (%)	11	98	15	12
Genetic alterations	del13q/del6q	Complex alterations	del13q in homozygosis	del13q
IGHV homology	100%	100%	95.2%	93.1%
Progression	6 years C(IV)	7 years B(II)	NA	3 years A(I)
Treatment	FCM	R-FCM	FCM	R-FCM
Response	CR MRD+	CR MRD+	CR MRD-	CR MRD-
Time from diagnosis to tumour DNA sampling	5.5 years	6.1 years	0.4 years	3.9 years
Tumour DNA purity	98.8%	99.5%	99.4%	98%
Time from treatment to normal DNA sampling	0.5 years	3 years	3.7 years	2 years
Normal DNA purity	100%	100%	100%	100%

LDT, Lymphocyte Doubling Time; NA, not applicable; ND, not done; FCM, fludarabine, cyclophosphamide, mitoxantrone; R-FCM, rituximab and FCM; CR, complete remission; MRD, minimal residual disease; ** Mixed cell

Supplementary Table 2. Statistics for Whole-Genome Sequencing

Case	IGHV status	Sample	Number of reads mapped	Depth	Coverage	Cov. ≥10 reads	Number of SNPs	Not in dbSNP (%)
CLL1	unmutated	Normal	1,160,452,910	42X	99.0%	97.8%	3,614,143	261,995 (7.25%)
		Tumour	1,455,831,568	53X	99.0%	98.2%	3,629,099	265,329 (7.31%)
CLL2	unmutated	Normal	941,896,591	31X	98.5%	95.8%	3,469,363	213,168 (6.14%)
		Tumour	996,700,520	33X	98.5%	96.1%	3,456,394	210,318 (6.08%)
CLL3	mutated	Normal	1,551,483,389	49X	99.1%	96.8%	3,832,713	331,559 (8.65%)
		Tumour	861,785,635	31X	98.8%	98.0%	3,777,357	344,083 (9.10%)
CLL4	mutated	Normal	1,361,098,408	40X	99.0%	96.8%	3,454,075	202,905 (5.87%)
		Tumour	1,369,552,707	41X	99.1%	97.4%	3,394,946	199,026 (5.86%)

Supplementary Table 3. Statistics for Exome Sequencing

Case	IGHV status	Sample	Number of reads mapped	Depth	Exome Cov.	Cov. ≥10 reads
CLL1	unmutated	Normal	107,937,237	99X	99.3%	93%
		Tumour	100,606,508	96X	99.3%	93%
CLL2	unmutated	Normal	120,385,251	187X	99.3%	95%
		Tumour	114,365,205	179X	99.6%	96%
CLL3	mutated	Normal	141,399,386	119X	99.2%	93%
		Tumour	111,236,146	70X	99.1%	91%
CLL4	mutated	Normal	103,472,724	97X	99.3%	93%
		Tumour	113,362,146	107X	99.2%	93%

Supplementary Table 4. Statistics for Long Insert Paired-End Sequencing

Case	IGHV status	Sample	Number of reads mapped	Insert Size
CLL1	unmutated	Normal	58,163,212	2540
		Tumour	51,083,672	2476
CLL2	unmutated	Normal	80,703,724	2461
		Tumour	110,167,808	2531
CLL3	mutated	Normal	67,835,175	2559
		Tumour	60,419,198	2446
CLL4	mutated	Normal	69,367,176	2642
		Tumour	111,110,742	2542

Supplementary Table 5. Structural Variants in CLL tumours detected by paired-end analysis.

Sample	Chr	Cytoband	Start (bp)	End (bp)	Size (Mb)	CN Estimate	Structural Variant	CLL Alteration
CLL1	6	q14.1-q22.2	78688092	118428895	39.74	1	Deletion	Known
CLL1	13	q14.2-q14.3	49883705	52299307	2.41	1	Deletion	Known
CLL2	1	q21.1	144597148	147730898	3.13	3	Fold-back inversion	Novel
CLL2	1	q25.2	177162885	180234801	3.07	3	Tandem duplication	Novel
CLL2	2	p16.1-p15	59155728	61892257	2.73	3	Tandem duplication	Known
CLL2	3	q26.2-q26.31	168533864	171596966	3.07	3	Complex insertion	Novel
CLL2	6	q15-q16.1	89756638	92803277	3.05	3	Complex insertion	Novel
CLL2	12	q13.13-q13.2	52611909	55200910	2.59	3	Complex insertion	Known
CLL3	13	q14.3	50551421 50521570	51386741 51558441	0.83 1.03	0	Homozygous deletion	Known
CLL4	13	q14.2-q22.1	48679885	74276206	25.59	1	Deletion	Known

CN: copy number estimate by WGS and Affymetrix 6.0 and Agilent 1M arrays; Genomic positions are based on genome assembly GRCh37/hg19. All changes in antigen receptor loci were excluded (IGK@ 2p11.2, TRG@ 7p14.1, TRB@ 7q34, TRA@ 14q11.2, IGH@ 14q32.33, and IGL@ 22q11.2)

Supplementary Table 7. Class 1 and 2 mutations identified in the 4 CLL cases.

Case	Symbol	Ensembl Accession	Mutation Type	Effect	Chr	Position	Ref.	Obs	Protein Name	Expressed
CLL1	FUT9	ENSG00000172461	frameshift	K51*fs	6	96651182	AA	T	fucosyltransferase 9	Yes
CLL1	B3GAT2	ENSG00000112309	non_synonymous	Y83F	6	71665885	T	W	galactosylgalactosylxylosylprotein 3-beta-glucuronosyl	No
CLL1	CNOT3	ENSG00000088038	non_synonymous	E20K	19	54646887	G	R	CCR4-NOT transcription complex, subunit 3	Yes
CLL1	CYP2A7	ENSG00000198077	non_synonymous	L298F	19	41383838	G	R	cytochrome P450, family 2, subfamily A, polypeptide 7	No
CLL1	FSIP2	ENSG00000188738	non_synonymous	T2526R	2	186659440	C	S	fibrous sheath interacting protein 2	No
CLL1	OR1L8	ENSG00000171496	non_synonymous	N43K	9	125330628	G	K	olfactory receptor, family 1, subfamily L, member 8	No
CLL1	ROBO1	ENSG00000169855	non_synonymous	A897V	3	78700896	G	R	roundabout, axon guidance receptor, homolog 1	Yes
CLL1	SETD5	ENSG00000168137	non_synonymous	Q787*	3	9495435	C	Y	SET domain containing 5	Yes
CLL1	SYNE1	ENSG00000131018	non_synonymous	V1167I	6	152770673	C	Y	spectrin repeat containing, nuclear envelope 1	No
CLL1	TRPV4	ENSG00000111199	non_synonymous	V620I	12	110230201	C	Y	ransient receptor potential cation channel, subfamily V	No
CLL1	ABCB5	ENSG00000004846	synonymous	I956I	7	20778606	A	W	ATP-binding cassette sub-family B member 5 isoform 1	Yes
CLL1	DENND4A	ENSG00000174485	synonymous	C1291C	15	65983059	G	R	DENN/MADD domain containing 4A	Yes
CLL1	TRPM6	ENSG00000119121	synonymous	P933P	9	77400910	A	M	transient receptor potential cation channel, subfamily	No
CLL1	AK2	ENSG00000004455	3p_UTR	-	1	33474552	C	Y	adenylate kinase 2	Yes
CLL1	ARHGEF38	ENSG00000236699	3p_UTR	-	4	106599672	G	R	Rho guanine nucleotide exchange factor (GEF) 38	No
CLL1	CEP170	ENSG00000143702	5p_UTR	-	1	243319878	C	Y	centrosomal protein 170kDa	Yes
CLL1	CLN8	ENSG00000182372	3p_UTR	-	8	1733881	T	K	ceroid-lipofuscinosis, neuronal 8	Yes
CLL1	CYP4A22	ENSG00000162365	3p_UTR	-	1	47611843	G	R	cytochrome P450, family 4, subfamily A, polypeptide 2	No
CLL1	IL20RB	ENSG00000174564	5p_UTR	-	3	136676936	G	R	interleukin 20 receptor beta	Yes
CLL1	KDM3A	ENSG00000115548	5p_UTR	-	2	86667987	C	M	lysine (K)-specific demethylase 3A	Yes
CLL1	RASSF6	ENSG00000169435	3p_UTR	-	4	74440620	C	M	Ras association (RalGDS/AF-6) domain family member	No
CLL1	SCLT1	ENSG00000151466	5p_UTR	-	4	130014445	A	W	sodium channel and clathrin linker 1	Yes
CLL2	GPCPD1	ENSG00000125772	frameshift	I558*fs	20	5538748	CCC	A	glycerophosphocholine phosphodiesterase GDE1 hom	Yes
CLL2	NOTCH1	ENSG00000148400	frameshift	P2515R*fs4	9	139390648	AG	*	Notch1	Yes
CLL2	ACER2	ENSG00000177076	non_synonymous	R173L	9	19446293	G	K	alkaline ceramidase 2	Yes
CLL2	C14orf39	ENSG00000179008	non_synonymous	A548D	14	60903684	G	K	Protein SIX6OS1	No
CLL2	CCND2	ENSG00000118971	non_synonymous	P281R	12	4409147	C	S	cyclin D2	Yes
CLL2	CDH9	ENSG00000113100	non_synonymous	P363L	5	26902750	G	R	cadherin 9, type 2	No
CLL2	DMD	ENSG00000198947	non_synonymous	I2550L	X	31747760	T	K	dystrophin	No
CLL2	ELOVL6	ENSG00000170522	non_synonymous	A185T	4	110972739	C	Y	elongation of long chain fatty acids family member 6	No
CLL2	F13A1	ENSG00000124491	non_synonymous	R159H	6	6266886	C	Y	coagulation factor XIII, A1 polypeptide	No
CLL2	KIAA1324	ENSG00000116299	non_synonymous	S852R	1	109742530	C	S	KIAA1324	No
CLL2	LLGL1	ENSG00000131899	non_synonymous	R955C	17	18145294	C	Y	lethal giant larvae homolog 1	Yes
CLL2	MERIT40	ENSG00000105393	non_synonymous	Y217C	19	17387384	A	R	BRCA1-A complex subunit MERIT40	Yes
CLL2	MGA	ENSG00000174197	non_synonymous	R1242P	15	42021429	G	S	MAX gene associated	No
CLL2	PIK3R2	ENSG00000105647	non_synonymous	R539H	19	18277996	G	R	phosphoinositide-3-kinase, regulatory subunit 2	Yes
CLL2	RBBP9	ENSG00000089050	non_synonymous	D46N	20	18476488	C	Y	retinoblastoma binding protein 9	Yes
CLL2	SLC45A2	ENSG00000164175	non_synonymous	T412M	5	33947401	G	R	solute carrier family 45, member 2	No
CLL2	UGT2B11	ENSG00000213759	non_synonymous	P453A	4	70066391	G	S	UDP glucuronosyltransferase 2 family, polypeptide B1	No
CLL2	UNC119B	ENSG00000175970	non_synonymous	Q251R	12	121157831	AG	GT	unc-119 homolog B	Yes
CLL2	VWF	ENSG00000110799	non_synonymous	R2464H	12	6085323	C	Y	von Willebrand factor	No
CLL2	XPO1	ENSG00000082898	non_synonymous	E571K	2	61719472	C	Y	exportin 1	Yes
CLL2	HK3	ENSG00000160883	synonymous	S833S	5	176308431	G	R	hexokinase 3	No
CLL2	MAGEE1	ENSG00000198934	synonymous	T875T	X	75650948	C	S	melanoma antigen family E, 1	Yes
CLL2	PDE4DIP	ENSG00000178104	synonymous	A2374A	1	144856852	C	Y	phosphodiesterase 4D interacting protein	No
CLL2	PXDNL	ENSG00000147485	synonymous	P650P	8	52323922	G	R	peroxidasin homolog (Drosophila)-like	No
CLL2	TSHZ2	ENSG00000182463	synonymous	K1006K	20	51873015	A	R	teashirt zinc finger homeobox 2	Yes
CLL2	EPHA4	ENSG00000116106	3p_UTR	-	2	222283824	G	R	EPH receptor A4	No
CLL2	GDA	ENSG00000119125	3p_UTR	-	9	74863584	T	Y	guanine deaminase	No
CLL2	SPATA18	ENSG00000163071	3p_UTR	-	4	52962403	C	M	spermatogenesis associated 18 homolog	No
CLL2	SIKE1	ENSG00000052723	5p_UTR	-	1	115323009	T	K	suppressor of IKBKE 1	Yes
CLL2	ZNF518A	ENSG00000177853	5p_UTR	-	10	97911272	A	W	zinc finger protein 518A	No
CLL3	CD8A	ENSG00000153563	frameshift	E37*fs	2	87017743	C	+T	CD8 antigen alpha polypeptide isoform 1	Yes
CLL3	ZNF37A	ENSG00000075407	frameshift	T294*fs	10	38406960	C	-AG	zinc finger protein 37A	Yes
CLL3	ABI3BP	ENSG00000154175	non_synonymous	V678F	3	100499051	C	M	ABI gene family, member 3 (NESH) binding protein	No
CLL3	KIR3DL1	ENSG00000136856	non_synonymous	V272I	19	55333178	G	R	killer cell immunoglobulin-like receptor, three domain	No
CLL3	KLHL6	ENSG00000172578	non_synonymous	L65P	3	183273248	A	R	kelch-like 6	Yes
CLL3	KLHL6	ENSG00000172578	non_synonymous	F49L	3	183273297	A	R	kelch-like 6	Yes
CLL3	MYD88	ENSG00000172936	non_synonymous	L265P	3	38182641	T	Y	myeloid differentiation primary response gene (88)	Yes
CLL3	PPP1R9A	ENSG00000158528	non_synonymous	S1283Y	7	94917966	C	M	protein phosphatase 1, regulatory (inhibitor)	Yes
CLL3	SCN11A	ENSG00000168356	non_synonymous	R222P	3	38966953	C	S	sodium channel, voltage-gated, type XI, alpha	No
CLL3	SLC2A8	ENSG00000136856	non_synonymous	G337S	9	130167129	G	R	solute carrier family 2 (facilitated glucose transporter),	Yes
CLL3	EFR3A	ENSG00000132294	splicing-site	-	8	132989340	G	S	EFR3 homolog A	Yes
CLL3	CELA3A	ENSG00000142789	synonymous	L145L	1	22333395	C	S	elastase 3A	No
CLL3	SMC2	ENSG00000136824	synonymous	A858A	9	106889044	A	M	structural maintenance of chromosomes 2	Yes
CLL3	AGXT2L2	ENSG00000175309	3p_UTR	-	5	177638688	T	K	alanine-glyoxylate aminotransferase 2-like 2	Yes
CLL3	BMPR1B	ENSG00000138696	3p_UTR	-	4	96078173	A	R	bone morphogenetic protein receptor, type IB	No
CLL3	C10orf108	ENSG00000180525	3p_UTR	-	10	709729	C	M	chromosome 10 open reading frame 108	No
CLL3	DKK2	ENSG00000155011	3p_UTR	-	4	107844936	G	K	dickkopf homolog 2 precursor	No
CLL3	KRT31	ENSG00000094796	3p_UTR	-	17	39549987	T	Y	keratin 31	No
CLL3	MDGA2	ENSG00000139915	3p_UTR	-	14	47310653	T	W	MAM domain containing glycosylphosphatidylinositol	No
CLL3	POLK	ENSG00000122008	3p_UTR	-	5	74895632	C	S	DNA-directed DNA polymerase kappa	Yes
CLL3	RERGL	ENSG00000111404	3p_UTR	-	12	18234101	T	Y	RERG/RAS-like	No
CLL4	SNX7	ENSG00000162627	non_synonymous	L383I	1	99203814	C	M	sorting nexin 7 isoform a	Yes

CLL4	GRM3	ENSG00000198822	non_synonymous	R352Q	7	86416163	G	R	glutamate receptor, metabotropic 3 precursor	Yes
CLL4	SLC25A40	ENSG00000075303	non_synonymous	I149V	7	87477180	T	Y	mitochondrial carrier family protein	Yes
CLL4	C13orf23	ENSG00000120685	non_synonymous	I229T	13	39596507	A	R	Uncharacterized protein C13orf23	Yes
CLL4	MBP	ENSG00000197971	non_synonymous	I287F	18	74696742	T	W	myelin basic protein	Yes
CLL4	DOCK4	ENSG00000128512	synonymous	G913G	7	111474740	C	M	dedicator of cytokinesis 4	No
CLL4	FLRT2	ENSG00000185070	synonymous	I634I	14	86089760	T	Y	fibronectin leucine rich transmembrane protein 2	No
CLL4	NIPA2	ENSG00000140157	synonymous	S259S	15	23006527	T	K	on imprinted in Prader-Willi/Angelman syndrome	Yes
CLL4	MCFD2	ENSG00000180398	3p_UTR	-	2	47129227	A	R	multiple coagulation factor deficiency 2	Yes
CLL4	SYNPO2	ENSG00000172403	3p_UTR	-	4	119981524	C	Y	synaptopodin 2	No
CLL4	SEPHS1	ENSG00000086475	3p_UTR	-	10	13359781	C	Y	selenophosphate synthetase 1	Yes
CLL4	FAM171A1	ENSG00000148468	3p_UTR	-	10	15254137	C	M	family with sequence similarity 171, member A1	No
CLL4	SYT4	ENSG00000132872	3p_UTR	-	18	40850140	T	Y	synaptotagmin IV	No
CLL4	CXCR6	ENSG00000172215	5p_UTR	-	3	45987195	A	R	chemokine (C-X-C motif) receptor 6	No

Supplementary Table 8. Mutations and characteristics of the clinical validation series

Patient Code	*IGHV* Mutational Status	ZAP-70	CD38	*NOTCH1*	*MYD88*	*XPO1*	*KLHL6*	Validation Sets
1	Unmutated	8	11	NV	WT	WT	NV	SMIPS
2	Unmutated	39	98	P2515Rfs*4	NA	E571K	NA	EV
3	Mutated	12	15	WT	L265P	WT	F49L,L65P	SMIPS
4	Mutated	2	12	WT	WT	WT	WT	SMIPS
5	Unmutated	64	20	NV	WT	WT	WT	SMIPS
6	Mutated	16	2	WT	WT	WT	WT	SMIPS
7	Unmutated	42	85	WT	WT	WT	WT	SMIPS
8	Mutated	5	10	WT	WT	WT	WT	SMIPS
9	Unmutated	26	17	WT	WT	WT	WT	SMIPS
10	NA	36	10	NA	WT	NA	NA	EV
11	Unmutated	0	0	WT	WT	WT	WT	SMIPS
12	Unmutated	57	18	WT	WT	WT	NV	SMIPS
13	Mutated	3	10	WT	WT	WT	WT	SMIPS
14	Mutated	12	2	P2515Rfs*4	WT	WT	WT	SMIPS
15	Unmutated	67	0	WT	WT	WT	WT	SMIPS
16	NA	13	0	NV	WT	WT	WT	SMIPS
17	Mutated	10	88	P2515Rfs*4	WT	WT	WT	SMIPS
18	Unmutated	44	44	WT	WT	WT	WT	SMIPS
19	Mutated	0	6	WT	L265P	WT	NV	SMIPS
20	Unmutated	42	32	NV	WT	WT	WT	SMIPS
21	Unmutated	18	90	WT	WT	WT	WT	SMIPS
22	Unmutated	24	45	P2515Rfs*4	NA	NA	NA	EV
23	Unmutated	1	4	WT	WT	WT	WT	SMIPS
24	Unmutated	51	44	WT	WT	WT	WT	SMIPS
25	Unmutated	1	0	NV	WT	WT	WT	SMIPS
26	Unmutated	13	18	NV	WT	WT	WT	SMIPS
27	Unmutated	18	79	WT	WT	WT	WT	SMIPS
28	Mutated	1	97	WT	WT	WT	WT	SMIPS
29	Mutated	NA	25	WT	WT	WT	NV	SMIPS
30	Mutated	0	0	WT	L265P	WT	WT	SMIPS
31	Mutated	0	0	WT	WT	WT	WT	SMIPS
32	Unmutated	0	50	WT	WT	NA	NA	SMIPS
33	Unmutated	30	80	WT	WT	WT	WT	SMIPS
34	Mutated	0.4	2.5	WT	WT	WT	WT	SMIPS
35	Mutated	2	1	WT	WT	WT	WT	SMIPS
36	Mutated	36	11	WT	WT	WT	WT	SMIPS
37	Unmutated	14	13	NV	WT	WT	WT	SMIPS
38	Mutated	12	10	WT	WT	WT	WT	SMIPS
39	NA	26	26	WT	WT	WT	WT	SMIPS
40	Mutated	9	91	WT	WT	WT	NV	SMIPS
41	Mutated	2	4	WT	WT	WT	WT	SMIPS
42	Mutated	44	75	NV	WT	WT	WT	SMIPS
43	Unmutated	65	98	P2515Rfs*4	WT	NV	WT	SMIPS
44	Mutated	19	0.4	NV	WT	WT	L90F	SMIPS
45	Mutated	6	8	NV	WT	NV	WT	SMIPS
46	Unmutated	3	2	NV	WT	WT	WT	SMIPS
47	Unmutated	25	98	WT	WT	WT	WT	SMIPS
48	Unmutated	14	45	WT	WT	WT	WT	SMIPS
49	Unmutated	0	0	WT	WT	WT	WT	SMIPS
50	Unmutated	0	0	WT	WT	WT	WT	SMIPS
51	Mutated	0	22	WT	WT	WT	WT	SMIPS
52	Mutated	0	0	WT	WT	WT	WT	SMIPS
53	Unmutated	7	87	WT	WT	WT	WT	SMIPS
54	Mutated	0.9	16	NV	WT	NV	WT	SMIPS
55	Mutated	NA	1	WT	WT	WT	NV	SMIPS
56	Unmutated	2	0	WT	WT	WT	WT	SMIPS
57	Unmutated	NA	NA	WT	NV	WT	WT	SMIPS
58	Unmutated	73	NA	P2515Rfs*4	WT	WT	NV	SMIPS

59	Mutated	0	NA	WT	WT	WT	WT	SMIPS
60	Unmutated	23	20	WT	WT	WT	WT	SMIPS
61	Unmutated	54	15	WT	WT	WT	WT	SMIPS
62	Unmutated	31	50	WT	WT	WT	WT	SMIPS
63	Unmutated	73	NA	NV	WT	NV	WT	EV
64	NA	86	NA	NV	WT	WT	WT	SMIPS
65	NA	100	NA	WT	WT	WT	WT	SMIPS
66	NA	21	86	WT	WT	WT	WT	EV
67	Unmutated	2	70	WT	WT	WT	WT	SMIPS
68	Unmutated	60	97	NV	NV	WT	WT	SMIPS
69	Unmutated	60	NA	P2515Rfs*4	WT	NV	WT	SMIPS
70	Unmutated	51	NA	WT	WT	WT	WT	SMIPS
71	Mutated	10	NA	WT	WT	WT	WT	SMIPS
72	Mutated	14	NA	NV	WT	WT	WT	SMIPS
73	Unmutated	30	NA	P2515Rfs*4	WT	NV	NV	SMIPS
74	Mutated	3	0	WT	WT	WT	WT	SMIPS
75	Mutated	4	14	WT	WT	WT	WT	SMIPS
76	NA	2	NA	P2515Rfs*4	WT	NV	WT	SMIPS
77	Mutated	NA	NA	WT	WT	NV	WT	SMIPS
78	Unmutated	56	95	NV	WT	NV	NV	EV
79	Unmutated	78	NA	P2515Rfs*4	WT	WT	WT	SMIPS
80	NA	43	90	WT	WT	WT	WT	SMIPS
81	Unmutated	45	NA	WT	WT	WT	WT	SMIPS
82	Unmutated	39	NA	WT	WT	WT	WT	SMIPS
83	Unmutated	64	74	WT	WT	WT	WT	SMIPS
84	Unmutated	30	NA	P2515Rfs*4	WT	WT	WT	SMIPS
85	Unmutated	64	NA	WT	WT	WT	WT	SMIPS
86	Unmutated	16	67	WT	WT	WT	WT	SMIPS
87	Unmutated	15	2	WT	WT	WT	WT	SMIPS
88	Unmutated	4	NA	WT	WT	WT	WT	SMIPS
89	Unmutated	49	60	NV	WT	NV	NV	SMIPS
90	Mutated	0.5	0	NV	WT	WT	WT	SMIPS
91	Mutated	6	13	NV	WT	WT	WT	SMIPS
92	Mutated	0	0	NV	WT	WT	WT	SMIPS
93	Unmutated	71	97	P2515Rfs*4	WT	E571G	WT	SMIPS
94	Mutated	10	3	NV	WT	WT	WT	SMIPS
95	Mutated	0.3	46	WT	NV	WT	WT	SMIPS
96	Mutated	4	0	WT	WT	WT	WT	SMIPS
97	Unmutated	82	83	WT	WT	WT	WT	SMIPS
98	Mutated	7	8	NV	WT	WT	WT	SMIPS
99	NA	0.5	5	WT	L265P	WT	WT	SMIPS
100	Unmutated	1	80	WT	WT	WT	WT	SMIPS
101	Unmutated	5	NA	NV	WT	WT	WT	SMIPS
102	Mutated	1	2	WT	WT	WT	WT	SMIPS
103	Mutated	14	2	WT	WT	WT	WT	SMIPS
104	NA	17	63	NV	WT	NA	NA	SMIPS
105	Mutated	2	1	NV	WT	WT	WT	SMIPS
106	Unmutated	42	19	WT	WT	WT	WT	SMIPS
107	Unmutated	86	28	NV	WT	WT	WT	SMIPS
108	Unmutated	6	37	NV	WT	WT	WT	EV
109	Unmutated	73	81	P2515Rfs*4	WT	WT	WT	SMIPS
110	Unmutated	4	74	WT	WT	WT	WT	SMIPS
111	Mutated	4	0	WT	WT	WT	WT	SMIPS
112	Unmutated	24	87	WT	WT	WT	WT	SMIPS
113	Mutated	51.5	NA	WT	WT	WT	WT	SMIPS
114	Unmutated	8	99	WT	WT	WT	WT	SMIPS
115	Mutated	7	14	NV	WT	NA	NA	SMIPS
116	Mutated	18	10	NV	WT	WT	WT	SMIPS
117	Unmutated	48	90	WT	WT	WT	WT	SMIPS
118	Mutated	2	4	WT	WT	WT	WT	SMIPS
119	Mutated	14	7	WT	WT	WT	WT	SMIPS
120	Mutated	0	0	WT	WT	WT	WT	SMIPS

121	Unmutated	0	0	WT	WT	WT	WT	SMIPS
122	Unmutated	0	0	WT	WT	WT	WT	SMIPS
123	Mutated	0	54	NV	WT	WT	WT	SMIPS
124	Unmutated	0	13	WT	WT	WT	WT	SMIPS
125	Unmutated	0	0	WT	WT	WT	WT	SMIPS
126	NA	1	1.5	NV	WT	WT	WT	SMIPS
127	Mutated	1	14	WT	WT	WT	L58P, T64A, Q81P	SMIPS
128	Mutated	0	1	NV	WT	WT	WT	SMIPS
129	Mutated	60	78	P2515Rfs*4	WT	WT	WT	SMIPS
130	Mutated	6	5	WT	WT	WT	WT	SMIPS
131	Unmutated	84	90	NV	WT	WT	WT	SMIPS
132	Mutated	74	12	NV	WT	WT	WT	SMIPS
133	Mutated	13	2	NV	WT	WT	WT	SMIPS
134	Unmutated	33	38	WT	WT	WT	WT	SMIPS
135	Mutated	1.5	2	WT	WT	NA	NA	EV
136	Mutated	4	5	WT	WT	NA	NA	EV
137	NA	3	NA	NA	WT	NA	NA	EV
138	Mutated	3	0	WT	WT	NA	NA	EV
139	Mutated	4	2	WT	WT	NA	NA	EV
140	Mutated	1.5	3	WT	WT	NA	NA	EV
141	Unmutated	41	14	WT	WT	NA	NA	EV
142	Mutated	NA	94	WT	WT	NA	NA	EV
143	Mutated	4	2	WT	L265P	NA	NA	EV
144	Unmutated	18	21	NV	WT	NA	NA	EV
145	Unmutated	70	30	P2515Rfs*4	WT	NA	NA	EV
146	Mutated	1	0	WT	NV	NA	NA	EV
147	Unmutated	26	10	NV	WT	NA	NA	EV
148	Mutated	14	97	WT	NV	NA	NA	EV
149	Unmutated	14	1	WT	WT	NA	NA	EV
150	Mutated	17	5	NV	WT	NA	NA	EV
151	Mutated	0.3	0	WT	WT	NA	NA	EV
152	Mutated	4	0	WT	NV	NA	NA	EV
153	Mutated	14	4	WT	WT	NA	NA	EV
154	Mutated	1	15	WT	WT	NA	NA	EV
155	Unmutated	77	62	WT	WT	WT	NV	SMIPS
156	Mutated	78	98	WT	WT	WT	WT	SMIPS
157	Mutated	0	0	NV	WT	WT	NV	SMIPS
158	Unmutated	NA	100	WT	WT	E571K	WT	SMIPS
159	Unmutated	NA	5	WT	WT	WT	WT	SMIPS
160	Unmutated	0	0	P2515Rfs*4	WT	WT	WT	SMIPS
161	Unmutated	80	NA	WT	WT	WT	NV	SMIPS
162	Mutated	0	0	WT	WT	WT	NV	SMIPS
163	Unmutated	NA	NA	WT	WT	WT	NV	SMIPS
164	Unmutated	0	0	WT	WT	WT	WT	SMIPS
165	Unmutated	NA	NA	NV	WT	WT	WT	SMIPS
166	Unmutated	20	NA	WT	WT	WT	WT	SMIPS
167	Mutated	NA	NA	WT	WT	WT	WT	EV
168	Unmutated	0	0	NV	WT	WT	WT	SMIPS
169	NA	NA	NA	WT	WT	WT	WT	SMIPS
170	Unmutated	NA	NA	NV	WT	WT	WT	SMIPS
171	Unmutated	0	48	WT	WT	WT	WT	SMIPS
172	Mutated	NA	NA	NV	WT	WT	WT	SMIPS
173	Unmutated	NA	NA	WT	WT	E571K	WT	SMIPS
174	Unmutated	NA	0	NV	WT	WT	WT	SMIPS
175	Unmutated	NA	NA	WT	WT	WT	WT	EV
176	Unmutated	NA	0	NV	WT	WT	WT	SMIPS
177	Unmutated	NA	NA	NV	WT	WT	WT	SMIPS
178	Mutated	NA	9	WT	NV	WT	WT	EV
179	Mutated	3	5	WT	WT	NA	NA	EV
180	Unmutated	9	8	WT	WT	NA	NA	EV
181	NA	8	0	NA	WT	NA	NA	EV
182	Unmutated	11	1	NA	NV	NA	NA	EV

183	Unmutated	25	98	NV	L265P	WT	WT	SMIPS
184	Mutated	14	5	NV	NV	NA	NA	EV
185	NA	2	26	NV	NV	NA	NA	EV
186	Mutated	NA	15	NV	NV	NA	NA	EV
187	NA	37	0	NV	NV	NA	NA	EV
188	NA	1	0	NV	NV	NA	NA	EV
189	Unmutated	37	8	WT	WT	WT	WT	SMIPS
190	Mutated	0.9	0.35	NV	WT	WT	WT	SMIPS
191	Mutated	NA	NA	WT	WT	NA	NA	EV
192	Mutated	NA	NA	WT	NV	NA	NA	EV
193	NA	5.5	NA	WT	NV	NA	NA	EV
194	Mutated	NA	NA	WT	WT	NA	NA	EV
195	Mutated	NA	NA	WT	WT	NA	NA	EV
196	Mutated	4	0	NA	WT	NA	NA	EV
197	Mutated	NA	NA	WT	WT	NA	NA	EV
198	Mutated	NA	NA	WT	WT	NA	NA	EV
199	NA	NA	NA	WT	NV	NA	NA	EV
200	Unmutated	70	40	WT	WT	NA	NA	EV
201	Unmutated	NA	NA	P2515Rfs*4	NV	NA	NA	EV
202	NA	NA	NA	WT	NV	NA	NA	EV
203	NA	NA	NA	NV	NV	NA	NA	EV
204	Mutated	NA	NA	NA	NV	NA	NA	EV
205	Unmutated	12	NA	WT	WT	NA	NA	EV
206	Unmutated	NA	NA	F2482Ffs*2	WT	NA	NA	EV
207	Mutated	7	NA	WT	WT	NA	NA	EV
208	NA	NA	NA	WT	WT	NA	NA	EV
209	Unmutated	NA	NA	WT	NV	NA	NA	EV
210	NA	NA	NA	WT	WT	NA	NA	EV
211	NA	NA	NA	WT	NV	NA	NA	EV
212	NA	NA	NA	WT	NV	NA	NA	EV
213	Mutated	1.7	6	NA	NV	NA	NA	EV
214	NA	24	100	NV	WT	NA	NA	EV
215	NA	NA	NA	NV	WT	NA	NA	EV
216	NA	27	57	WT	WT	NA	NA	EV
217	NA	NA	NA	WT	NV	NA	NA	EV
218	Unmutated	56	NA	WT	WT	NA	NA	EV
219	NA	NA	NA	NA	WT	NA	NA	EV
220	NA	1.7	0	WT	WT	NA	NA	EV
221	NA	NA	NA	WT	WT	NA	NA	EV
222	Unmutated	NA	NA	P2515Rfs*4	WT	NA	NA	EV
223	NA	NA	NA	WT	WT	NA	NA	EV
224	NA	NA	NA	NV	WT	NA	NA	EV
225	Unmutated	6	NA	NV	NV	NA	NA	EV
226	Mutated	25	2	WT	WT	WT	WT	SMIPS
227	Unmutated	NA	NA	P2515Rfs*4	WT	NA	NA	EV
228	NA	NA	NA	WT	WT	NA	NA	EV
229	Mutated	41.5	NA	NV	WT	NA	NA	EV
230	NA	0	NA	NA	NV	NA	NA	EV
231	NA	1	3	WT	WT	NA	NA	EV
232	NA	17	NA	WT	WT	NA	NA	EV
233	Unmutated	9	6	WT	WT	NA	NA	EV
234	Unmutated	46	NA	WT	WT	NA	NA	EV
235	Mutated	1	1	NA	WT	NA	NA	EV
236	Mutated	NA	NA	WT	WT	NA	NA	EV
237	Mutated	3	70	NA	WT	NA	NA	EV
238	Mutated	0	52	P2515Rfs*4	WT	NA	NA	EV
239	NA	NA	NA	P2515Rfs*4	WT	NA	NA	EV
240	Mutated	1	100	NA	WT	NA	NA	EV
241	Mutated	NA	NA	WT	NV	NA	NA	EV
242	Mutated	2	98	NA	WT	NA	NA	EV
243	Mutated	10.5	NA	WT	WT	NA	NA	EV
244	NA	0.5	3	NV	WT	NA	NA	EV

245	NA	NA	NA	WT	WT	NA	NA	EV
246	Unmutated	NA	NA	WT	WT	NA	NA	EV
247	Mutated	4	1	NA	WT	NA	NA	EV
248	Unmutated	NA	NA	Q2503*	WT	NA	NA	EV
249	Mutated	3	9	WT	WT	WT	WT	SMIPS
250	Mutated	NA	NA	WT	WT	NA	NA	EV
251	NA	36	8	WT	WT	NA	NA	EV
252	NA	45	NA	NA	NV	NA	NA	EV
253	Unmutated	10	NA	NA	WT	NA	NA	EV
254	NA	NA	NA	WT	WT	NA	NA	EV
255	Mutated	2	3	WT	WT	NA	NA	EV
256	Unmutated	34	34	WT	WT	NA	NA	EV
257	Mutated	NA	NA	WT	WT	NA	NA	EV
258	Mutated	4	25	P2515Rfs*4	WT	NA	NA	EV
259	NA	14	4	WT	WT	NA	NA	EV
260	Unmutated	29	NA	WT	WT	NA	NA	EV
261	Unmutated	NA	NA	WT	WT	NA	NA	EV
262	Mutated	2	1	WT	WT	NA	NA	EV
263	NA	2	57	NV	WT	NA	NA	EV
264	Mutated	6	3	NA	WT	NA	NA	EV
265	Mutated	19	NA	WT	WT	NA	NA	EV
266	NA	2	0	WT	WT	NA	NA	EV
267	Unmutated	94	NA	WT	WT	NA	NA	EV
268	Mutated	1	NA	WT	WT	NA	NA	EV
269	NA	NA	89	WT	WT	NA	NA	EV
270	NA	4	NA	NA	WT	NA	NA	EV
271	Unmutated	29	NA	NA	WT	NA	NA	EV
272	Mutated	13	NA	WT	WT	NA	NA	EV
273	NA	26	NA	WT	WT	NA	NA	EV
274	NA	69	NA	WT	WT	NA	NA	EV
275	NA	0.5	1	NA	NV	NA	NA	EV
276	Mutated	2	3	WT	L265P	NA	NA	EV
277	Mutated	0	NA	NA	NV	NA	NA	EV
278	NA	87	NA	WT	WT	NA	NA	EV
279	NA	NA	NA	WT	WT	NA	NA	EV
280	Unmutated	80	NA	NV	WT	NA	NA	EV
281	NA	NA	NA	WT	WT	NA	NA	EV
282	NA	NA	NA	WT	WT	NA	NA	EV
283	NA	NA	NA	NV	NV	NA	NA	EV
284	NA	NA	NA	NA	WT	NA	NA	EV
285	Mutated	NA	NA	NA	WT	NA	NA	EV
286	Unmutated	83	67	NA	WT	NA	NA	EV
287	Unmutated	NA	NA	WT	WT	NA	NA	EV
288	Mutated	NA	NA	WT	WT	NA	NA	EV
289	Unmutated	54	NA	WT	WT	NA	NA	EV
290	NA	NA	NA	WT	WT	NA	NA	EV
291	NA	NA	NA	WT	WT	NA	NA	EV
292	NA	NA	NA	NV	WT	NA	NA	EV
293	NA	NA	NA	WT	WT	NA	NA	EV
294	Unmutated	16	17	WT	WT	NA	NA	EV
295	Mutated	15	2	WT	WT	NA	NA	EV
296	Mutated	13	0	WT	NV	NA	NA	EV
297	NA	1	NA	NA	NV	NA	NA	EV
298	Unmutated	43	NA	P2515Rfs*4	WT	NA	NA	EV
299	NA	0	NA	WT	NV	NA	NA	EV
300	Mutated	7	1	WT	NV	NA	NA	EV
301	Mutated	53	10	WT	WT	NA	NA	EV
302	Unmutated	3	NA	P2515Rfs*4	NV	NA	NA	EV
303	Mutated	16	NA	NA	WT	NA	NA	EV
304	Mutated	0	26	NA	WT	NA	NA	EV
305	NA	91	NA	WT	WT	NA	NA	EV
306	Mutated	1	NA	WT	NV	NA	NA	EV

307	Unmutated	88	NA	WT	WT	NA	NA	EV
308	Mutated	0.4	NA	NV	WT	NA	NA	EV
309	Unmutated	14	NA	WT	WT	NA	NA	EV
310	Mutated	6	5	WT	NV	NA	NA	EV
311	Unmutated	11	NA	P2515Rfs*4	NV	NA	NA	EV
312	Unmutated	7	NA	WT	NV	NA	NA	EV
313	NA	3	NA	NA	NV	NA	NA	EV
314	Mutated	3	1	WT	WT	NA	NA	EV
315	Unmutated	10	80	WT	NV	NA	NA	EV
316	NA	0.2	0	WT	WT	NA	NA	EV
317	Unmutated	34	NA	WT	WT	NA	NA	EV
318	Mutated	1	6	P2515Rfs*4	NV	NA	NA	EV
319	NA	59	22	WT	NV	NA	NA	EV
320	Mutated	6.5	7	WT	NV	NA	NA	EV
321	NA	56	NA	WT	WT	NA	NA	EV
322	Unmutated	30	36	P2515Rfs*4	WT	NA	NA	EV
323	NA	59	NA	WT	WT	NA	NA	EV
324	Mutated	12	NA	WT	WT	NA	NA	EV
325	Unmutated	31.5	NA	WT	WT	NA	NA	EV
326	Mutated	NA	NA	NV	NV	NA	NA	EV
327	Mutated	6	0	WT	WT	NA	NA	EV
328	Unmutated	70	NA	P2515Rfs*4	NV	NA	NA	EV
329	NA	NA	NA	NV	WT	NA	NA	EV
330	Unmutated	70	NA	WT	WT	NA	NA	EV
331	NA	NA	NA	NA	NV	NA	NA	EV
332	Mutated	5	NA	WT	WT	NA	NA	EV
333	NA	NA	NA	NV	WT	NA	NA	EV
334	Mutated	21.5	7	WT	WT	NA	NA	EV
335	Unmutated	1.5	2	WT	WT	NA	NA	EV
336	Mutated	25	NA	WT	NV	NA	NA	EV
337	Unmutated	39	NA	WT	WT	NA	NA	EV
338	Unmutated	35	NA	WT	WT	NA	NA	EV
339	Unmutated	1	2	WT	NV	NA	NA	EV
340	Mutated	6.5	1	WT	WT	NA	NA	EV
341	Mutated	3	1	NV	L265P	NA	NA	EV
342	Mutated	6	NA	WT	WT	NA	NA	EV
343	Unmutated	73	NA	WT	WT	NA	NA	EV
344	NA	61	NA	WT	WT	WT	WT	SMIPS
345	Unmutated	73	78	WT	WT	WT	WT	SMIPS
346	Mutated	17	0	NV	WT	WT	WT	SMIPS
347	Mutated	1	2	NV	WT	NV	WT	SMIPS
348	Mutated	29	26	WT	WT	WT	WT	SMIPS
349	Mutated	22	83	P2515Rfs*4	WT	WT	WT	SMIPS
350	NA	NA	NA	WT	WT	WT	WT	SMIPS
351	Unmutated	55	97	NV	WT	WT	WT	SMIPS
352	Mutated	14	28	NV	WT	WT	NV	SMIPS
353	Unmutated	73	21	NV	WT	WT	WT	SMIPS
354	Mutated	2	2	WT	WT	NA	NA	SMIPS
355	Unmutated	8	47	WT	WT	WT	WT	SMIPS
356	Mutated	14	3	NV	L265P	WT	WT	SMIPS
357	Mutated	10	6	NV	WT	NV	WT	EV
358	Unmutated	43	1	NV	WT	WT	WT	SMIPS
359	Unmutated	5	2	WT	WT	WT	WT	SMIPS
360	Unmutated	39	42	NV	WT	WT	WT	SMIPS
361	Mutated	1	5	WT	WT	WT	WT	SMIPS
362	Mutated	0	0	WT	WT	WT	WT	SMIPS
363	Mutated	3	0	WT	WT	WT	WT	SMIPS

Supplementary Table 9. Differentially expressed genes between CLL cases with NOTCH1-mutations or not.

Gene symbol	Probe set	Signal NOTCH1-unmutated	Signal NOTCH1-mutated	Fold-change
TUBB6	209191_at	44.34	207.12	0.21
LAG3	206486_at	38.71	164.28	0.24
LTK	217184_s_at	43.91	179.84	0.24
MGC24125	1554733_at	13.85	51.73	0.27
	1554732_at	15.3	32.17	0.48
MSI2	1552364_s_at	37.9	139.68	0.27
	225238_at	112.27	343.28	0.33
	243010_at	93.88	245.51	0.38
	225237_s_at	447.29	949.61	0.47
	225240_s_at	884.52	1619.29	0.55
SFTPB	209810_at	103.45	336.95	0.31
	37004_at	44.47	123.97	0.36
	214354_x_at	49.6	69.32	0.72
DNASE1L3	205554_s_at	13.59	42.96	0.32
LGALS1	201105_at	543.76	1637.77	0.33
YES1	202932_at	12.94	37.75	0.34
PDZD2	209493_at	7.95	21.87	0.36
ZAP70	1555613_a_at	73.41	193.03	0.38
ATF5	204999_s_at	52.73	132.42	0.4
	204998_s_at	99.24	246.62	0.4
CRIP1	205081_at	938.51	2357.79	0.4
EPPK1	232165_at	13.44	32.24	0.42
	232164_s_at	10.19	20.75	0.49
PAK1	209615_s_at	12.75	30.05	0.42
TTC39B	232000_at	8.17	19.3	0.42
	242477_at	12.07	18.73	0.64
	236826_at	10.01	12.88	0.78
UCP2	208997_s_at	235.15	545.39	0.43
ENTPD1	209474_s_at	50.53	113.77	0.44
	207691_x_at	64.51	122.03	0.53
FARP1	201911_s_at	16.73	37.75	0.44
ITGA4	205884_at	10.31	23.49	0.44
TMEM231	64900_at	19.36	44.19	0.44
TNFRSF1B	203508_at	271.04	618.25	0.44
NFKBID	1553042_a_at	111.43	245.08	0.45
	230052_s_at	283.88	591.72	0.48
RAB40B	204547_at	41.99	90.93	0.46
RIMKLB	225978_at	20.17	44.11	0.46
EFCAB4A	227429_at	54.11	115.3	0.47
LARP1	212193_s_at	34.04	72.88	0.47
	210966_x_at	85.1	133.37	0.64
SNRNP25	218493_at	276.73	594.32	0.47
DTX1	227336_at	82.74	171.72	0.48
PFN1	200634_at	504.91	1041.98	0.48
PRKACB	235780_at	15.15	31.6	0.48
CD79A	1555779_a_at	428.86	874.11	0.49
	205049_s_at	682.6	1283.38	0.53
FLNB	208613_s_at	36.01	73.31	0.49
	208614_s_at	156.53	300.65	0.52
MYO6	203216_s_at	7.9	16.25	0.49
PEA15	200787_s_at	53.35	109.07	0.49
	200788_s_at	364.29	592.14	0.62
GRN	216041_x_at	332.42	660.66	0.499
	200678_x_at	353.81	705.61	0.51
	211284_s_at	232.02	448.24	0.51
MLF2	200948_at	84.33	167.01	0.5
PVT1	1558290_a_at	156.62	310.96	0.5
AICDA	224499_s_at	7.72	15.21	0.51
AKIRIN2	223143_s_at	32.73	63.61	0.51
AP2M1	200613_at	161.41	315.46	0.51
CFL1	1555730_a_at	332.58	652.95	0.51
	200021_at	3958.73	4963.32	0.8
FKBP1A	214119_s_at	71.9	141.25	0.51
	210186_s_at	53.11	97.58	0.54
	200709_at	109.15	175.38	0.62
PKM2	201251_at	52.76	103.81	0.51
TRAPPC1	225294_s_at	31.9	62.05	0.51
BANF1	210125_s_at	79.07	152.98	0.52
DAZAP2	212595_s_at	813.93	1562.21	0.52
DOK3	223553_s_at	133.23	258.44	0.52

MX1	202086_at	395.32	754.44	0.52
NAPA	208751_at	19.01	36.84	0.52
	206491_s_at	66.76	95.7	0.7
CTSB	213274_s_at	23.71	44.32	0.53
	200838_at	116.45	200.44	0.58
	200839_s_at	225.24	349	0.65
DTX4	212611_at	127.99	239.52	0.53
NMB	205204_at	106.18	201.89	0.53
WSB2	201760_s_at	706.11	1328.98	0.53
ASNA1	202024_at	69.83	129.57	0.54
CLN8	222874_s_at	29.36	54.54	0.54
	219340_s_at	59.82	98.06	0.61
COPE	201264_at	165.28	306.93	0.54
DRAP1	203258_at	83.38	154.16	0.54
NINJ2	219594_at	27.33	50.8	0.54
PLEKHG2	225979_at	20.23	37.37	0.54
PPP1R9B	225124_at	73.26	136.37	0.54
PPP2R1A	200695_at	87.55	161.77	0.54
SH3GLB2	218813_s_at	47.99	88.89	0.54
	224907_s_at	46.72	70.1	0.67
UBA1	200964_at	244.82	455.28	0.54
ATP2A3	213042_s_at	35.54	65.17	0.55
	207521_s_at	36.76	59.55	0.62
CLPTM1	211136_s_at	65.45	119.22	0.55
	201640_x_at	104.77	142.47	0.74
COPZ1	222386_s_at	154.59	278.83	0.55
CYB5R3	1554574_a_at	36.21	66.26	0.55
GALNT2	217787_s_at	75.6	137.95	0.55
HOMER2	217080_s_at	18.43	33.65	0.55
RNF41	201962_s_at	354.59	645.37	0.55
	201961_s_at	121.25	219.7	0.55
TMED2	204426_at	121.11	219.61	0.55
	204427_s_at	286.87	488.41	0.59
ZMAT3	219628_at	104.92	192.12	0.55
	1555609_a_at	12.97	17.74	0.73
ARID1A	210649_s_at	97.37	175.18	0.56
COMMD4	209132_s_at	77.91	139.78	0.56
	206441_s_at	114.91	171.78	0.67
DCTN1	201082_s_at	24.49	44.03	0.56
DDA1	218260_at	59.6	106.16	0.56
DGKA	211272_s_at	94.9	170.74	0.56
IL2RG	204116_at	123.58	221.71	0.56
OGDH	201282_at	97.45	172.69	0.56
PSME3	209852_x_at	97.95	176.22	0.56
VAMP5	204929_s_at	83.47	149.9	0.56
ABHD6	45288_at	15.05	26.19	0.57
	221679_s_at	8.62	13.32	0.65
AP1M1	223025_s_at	71.09	125.21	0.57
CAPNS1	200001_at	308.5	537.54	0.57
CAPZB	201950_x_at	204.18	355.93	0.57
CS	208660_at	577.69	1013.88	0.57
EIF4E2	213570_at	79.35	139.2	0.57
	213571_s_at	242.79	369.56	0.66
NUCB1	200646_s_at	19.55	34.25	0.57
	200649_at	57.75	91.1	0.63
SEC61A2	228747_at	26.42	46.69	0.57
	219499_at	60.25	85.71	0.7
SGMS2	227038_at	7.46	13.13	0.57
TP53I3	210609_s_at	131.05	229.71	0.57
UNC93B1	225869_s_at	62.17	109.53	0.57
ABCC1	202805_s_at	64.4	110.53	0.58
	202804_at	492.04	691.14	0.71
AKIRIN1	222458_s_at	79.87	137.22	0.58
CMTM3	1555705_a_at	102.64	176.77	0.58
EIF5A	201123_s_at	792.77	1356.71	0.58
	201122_x_at	391.91	566.24	0.69
	213753_x_at	276.05	385.79	0.72
MCRS1	202556_s_at	101.99	174.69	0.58
MGAT4B	220189_s_at	123.86	212.74	0.58
MTA1	202247_s_at	20.16	34.57	0.58
PCGF1	210023_s_at	18.28	31.35	0.58
PPP1CA	200846_s_at	554.29	949.68	0.58
RAB5C	201156_s_at	103.27	179.39	0.58
	201140_s_at	134.12	205.43	0.65
ST6GALNAC4	221551_x_at	87.06	151.03	0.58
	220937_s_at	193.16	309.44	0.62
TLR7	222952_s_at	44.83	77.22	0.58

MAP2K2	213490_s_at	33.22	56.73	0.59
	202424_at	158.17	259.64	0.61
MYH10	212372_at	9.35	15.87	0.59
MYL6B	204173_at	101.22	170.28	0.59
NCKAP1L	209734_at	128.79	216.7	0.59
NCRNA00173	1563369_at	23.8	40.19	0.59
ODF3B	238327_at	40.79	69.64	0.59
PRKCSH	200707_at	62.04	105.04	0.59
PRR14	1559397_s_at	30.02	51.29	0.59
RNF130	217865_at	333.89	562.2	0.59
S100A13	202598_at	155.09	262.26	0.59
SPON2	218638_s_at	70.4	118.58	0.59
ALOX5	204445_s_at	76.53	126.96	0.6
AP2S1	208074_s_at	256.34	423.83	0.6
	211047_x_at	225.03	371.4	0.61
	202120_x_at	150.23	242.39	0.62
BLOC1S1	202592_at	148.16	248.65	0.6
CORO6	1552301_a_at	25.97	43.01	0.6
FBXO6	231769_at	66.48	110.97	0.6
MAPK1IP1L	212497_at	42.38	70.23	0.6
UBAP2L	201378_s_at	212.15	350.98	0.6
ARF3	211622_s_at	43.92	72.18	0.61
CALCOCO1	209002_s_at	164.22	271.18	0.61
DCAF15	221849_s_at	27.8	45.45	0.61
SMARCB1	212167_s_at	292.04	477.95	0.61
ATP6V0B	200078_s_at	705.64	1137.06	0.62
CCDC24	228994_at	70.54	113.59	0.62
DEDD	211255_x_at	39.59	64.23	0.62
	202480_s_at	48.62	65.35	0.74
FAM89B	32209_at	167.9	269.66	0.62
	212484_at	93.11	126.87	0.73
	218074_at	442.36	573.63	0.77
FRY	214318_s_at	31.43	50.82	0.62
GAPDH	213453_x_at	2216.14	3581.91	0.62
	217398_x_at	2210	3584.89	0.62
	212581_x_at	2391.59	3799.65	0.63
IRF9	203882_at	726.04	1166.29	0.62
OS9	215399_s_at	160.31	257.39	0.62
	200714_x_at	291.85	423.78	0.69
P4HB	1564494_s_at	44.41	71.57	0.62
	200656_s_at	156.89	245.37	0.64
	200654_at	460.27	641.09	0.72
PSMF1	201052_s_at	109.37	177.23	0.62
RFC5	203210_s_at	38.1	61.78	0.62
ROGDI	218394_at	30.94	50.29	0.62
TLN1	203254_s_at	43.66	70.55	0.62
TRAF4	211899_s_at	23.08	37.11	0.62
ATP11A	216488_s_at	49.23	77.99	0.63
CBX4	206724_at	61.04	96.12	0.63
CD99	201028_s_at	242.68	383.77	0.63
DDX41	217840_at	243.85	387.11	0.63
ERGIC1	223847_s_at	90.36	142.99	0.63
	224577_at	187.25	281.24	0.67
MAF1	222998_at	178.66	282.24	0.63
NDST1	1554010_at	36.53	58.36	0.63
PATL1	225468_at	64.35	101.78	0.63
	235235_s_at	18.88	25.84	0.73
PIGR	204213_at	52.32	82.56	0.63
PMF1	202337_at	145.74	231.95	0.63
RHBDD2	232053_x_at	63.53	101.03	0.63
	222995_s_at	161.52	241.91	0.67
SH3PXD2A	213252_at	47.1	74.86	0.63
SRPR	200918_s_at	183.9	290.9	0.63
STRN4	217903_at	50.51	79.98	0.63
TFDP1	242939_at	43.11	67.94	0.63
	204147_s_at	69.12	100.35	0.69
THRAP3	217847_s_at	90.92	144.78	0.63
WDR1	200611_s_at	451.75	712.16	0.63
ARHGDIA	213606_s_at	19.42	30.49	0.64
	201168_x_at	157.56	223.54	0.7
	211716_x_at	195.02	273.63	0.71
	201167_x_at	55.73	76.76	0.73
CIB1	201953_at	839.72	1318.07	0.64
CREBBP	211808_s_at	48.56	75.66	0.64
GNG2	224965_at	29.49	46.21	0.64
IPO4	218305_at	39.96	62.51	0.64
LSP1	203523_at	205.12	322.42	0.64

MGAT1	201126_s_at	150.74	235.04	0.64
MIER1	240227_at	15.89	25.03	0.64
MR1	207565_s_at	25.51	39.75	0.64
	210223_s_at	40.05	58.74	0.68
SF3B4	209044_x_at	106.54	165.24	0.64
SNX1	201716_at	183.56	285.77	0.64
	214531_s_at	105.55	156.2	0.68
TAPBP	1555565_s_at	111.44	175.1	0.64
	208829_at	768.69	1133.72	0.68
TMEM214	217899_at	79.8	125.56	0.64
ADPGK	224455_s_at	288.23	440.24	0.65
ARAF	201895_at	216.67	331.03	0.65
ASPSCR1	218908_at	57.1	87.5	0.65
ATAD2B	232908_at	40.81	62.87	0.65
ATF7	228830_s_at	56.65	87.53	0.65
ATP6V0C	36994_at	888.28	1358.69	0.65
BSG	208677_s_at	63.6	97.15	0.65
CIZ1	213977_s_at	53.97	83.64	0.65
	211358_s_at	36.54	48.82	0.75
ELMO2	220363_s_at	12.92	19.85	0.65
	221528_s_at	120.87	171.46	0.7
ERGIC3	216032_s_at	305.79	472.94	0.65
FMNL1	204789_at	78.27	120.22	0.65
GMIP	222782_s_at	31.88	48.72	0.65
	218913_s_at	317.62	449.91	0.71
HECTD3	218632_at	50.85	78.26	0.65
INO80B	65133_i_at	12.2	18.69	0.65
MEN1	202645_s_at	110.38	169.74	0.65
NOTCH2	210756_s_at	88.07	134.79	0.65
PKN1	202161_at	83.7	129.24	0.65
PNPLA6	203718_at	66.89	103.63	0.65
PRKAB1	201835_s_at	49.42	76.42	0.65
PSEN1	207782_s_at	45.2	69.17	0.65
REC8	218599_at	132.4	203.81	0.65
RHOG	203175_at	169.88	262.35	0.65
SH3BGRL3	221269_s_at	526.25	812.93	0.65
SUOX	204067_at	25.09	38.48	0.65
TCF3	209151_x_at	33.23	51.13	0.65
TOLLIP	217930_s_at	68.41	104.75	0.65
TRIM28	200990_at	164.43	253.7	0.65
TRRAP	214908_s_at	84.25	129.33	0.65
APH1A	1554417_s_at	80.46	122.68	0.66
CDK16	207239_s_at	19.42	29.27	0.66
NCOR2	207760_s_at	449	679.21	0.66
NUP62	207740_s_at	90.62	136.54	0.66
PAK2	208876_s_at	24.41	37.25	0.66
PCGF5	227935_s_at	283.67	426.82	0.66
RNH1	206050_s_at	240.06	362.15	0.66
SLC25A1	210010_s_at	75.58	114.76	0.66
SLC25A39	223649_s_at	111.51	168.78	0.66
YPEL3	223179_at	607.91	914.46	0.66
ATG9A	202492_at	129.12	193.03	0.67
BAT2	212081_x_at	131.9	196.21	0.67
	208132_x_at	155.13	208.29	0.74
CMIP	224991_at	40.45	59.95	0.67
COPG	217749_at	217.65	325.19	0.67
CSNK2A1	206075_s_at	73.69	109.9	0.67
	212072_s_at	310.85	415.75	0.75
GNB2	200852_x_at	237.4	356.96	0.67
LGALS9	203236_s_at	83.73	125.54	0.67
MAP2K3	207667_s_at	161.29	239.47	0.67
MAP2K7	226053_at	67.38	100.22	0.67
MRPL2	229884_s_at	41.97	62.29	0.67
MRPL4	223743_s_at	42.62	63.96	0.67
	218105_s_at	109.05	138.46	0.79
OBFC2B	218903_s_at	18.02	26.95	0.67
PDXP	223290_at	63.49	95.35	0.67
PPP4C	208932_at	237.66	353.87	0.67
RBM42	205740_s_at	91.01	136.11	0.67
RELA	209878_s_at	127.48	190.7	0.67
	201783_s_at	413.2	553.57	0.75
SLC35A4	224626_at	99.29	148.48	0.67
SRRM2	216629_at	20.49	30.56	0.67
TMEM208	221597_s_at	226.99	338.96	0.67
YIF1A	202418_at	104.58	155.05	0.67
ZNF384	212369_at	141.45	210.06	0.67
BCL7B	202518_at	160.88	237.33	0.68

CTBP1	203392_s_at	978.56	1438.52	0.68
EFHD2	217992_s_at	265.63	392.6	0.68
HSBP1	200942_s_at	220.57	325.3	0.68
MLST8	220587_s_at	60.44	88.37	0.68
MRPL38	225103_at	67.16	98.97	0.68
P2RX4	204088_at	95.91	140.85	0.68
SMARCD1	209518_at	139.65	204.26	0.68
SPEN	1556058_s_at	44.03	64.54	0.68
SRM	201516_at	117.76	172.91	0.68
ST6GAL1	214971_s_at	24.06	35.54	0.68
TPP1	200743_s_at	710.55	1043.42	0.68
UBE2M	203109_at	159.67	234.16	0.68
UBE2N	201523_x_at	435.55	653.49	0.67
	201524_x_at	1240.22	1648.75	0.75
YKT6	217785_s_at	17.42	25.53	0.68
ADI1	222400_s_at	452.29	653.49	0.69
ARF1	208750_s_at	329.07	477.36	0.69
ATP5J2	202961_s_at	592.03	860.14	0.69
AXIN1	212849_at	103.95	151.36	0.69
CD74	209619_at	5451.49	7917.39	0.69
CHCHD5	223479_s_at	47.69	69.15	0.69
FAM83H	226129_at	12.88	18.66	0.69
N4BP2	228242_at	265.77	386.18	0.69
NRBP1	217765_at	113.22	163.75	0.69
PLD3	201050_at	40.28	58.31	0.69
PRPF38A	1553709_a_at	167.77	243.74	0.69
PTK2B	203111_s_at	36.66	53.51	0.69
PTPN9	202958_at	40.14	58.27	0.69
SENP2	218122_s_at	97.48	140.52	0.69
TMEM115	216267_s_at	51.47	74.75	0.69
TOM1	202807_s_at	59.42	86.71	0.69
UCKL1	232675_s_at	74.84	109.06	0.69
WDR48	222157_s_at	89.97	130.68	0.69
XPO5	223057_s_at	48.09	69.68	0.69
ACO2	200793_s_at	197.14	282.46	0.7
AKT1	207163_s_at	124.65	177.84	0.7
AP3M2	203410_at	102.53	146.09	0.7
CACNB2	1555098_a_at	15.75	22.64	0.7
CARKD	217940_s_at	390.29	559.56	0.7
CCDC124	225454_at	80.5	114.2	0.7
CTSD	200766_at	25.24	36.26	0.7
DEAF1	209407_s_at	37.18	52.94	0.7
FBXW11	209456_s_at	15.64	22.35	0.7
HLA-J	217436_x_at	812.12	1159.07	0.7
MED12	203506_s_at	56.15	80.59	0.7
MXD4	212347_x_at	73.24	103.89	0.7
	212346_s_at	76.1	98.56	0.77
NUDT3	221579_s_at	153.87	220.64	0.7
PACS1	224658_x_at	105.22	149.82	0.7
PI4KB	210417_s_at	93.03	133.84	0.7
SCAMP4	213244_at	79.58	113.22	0.7
SRRT	222046_at	57.06	81.25	0.7
SS18	209954_x_at	50.82	73.11	0.7
STAT2	205170_at	54.62	77.72	0.7
TMEM127	222887_s_at	83.22	118.16	0.7
TUSC2	203272_s_at	121.37	174.19	0.7
	203273_s_at	117.15	166.85	0.7
TYK2	205546_s_at	342.03	486.5	0.7
UBE4B	202316_x_at	38.63	55.37	0.7
	215533_s_at	45.3	63.7	0.71
YIPF3	216338_s_at	103.36	147.39	0.7
ARFGAP2	211975_at	280.5	394.21	0.71
	227057_at	17.72	26.62	0.67
ARPC4	217817_at	332.6	468.4	0.71
ATP5B	201322_at	1265.46	1785.4	0.71
ATP6V1F	201527_at	573.51	802.36	0.71
CBX5	212126_at	187.78	265.44	0.71
CLTB	206284_x_at	78.87	111.65	0.71
	205172_x_at	74.7	94.77	0.79
	211043_s_at	132.07	167.6	0.79
CNPY3	217931_at	329.17	465.83	0.71
COPS7A	209029_at	76.68	108.22	0.71
CORO7	219040_at	103.71	147.1	0.71
DAXX	201763_s_at	136.31	191.4	0.71
DPP3	218567_x_at	61.76	87.47	0.71
	232510_s_at	171.09	237.06	0.72
EHBP1L1	221755_at	178.61	252.26	0.71

	91703_at	134.5	178.53	0.75
	1557228_at	121.68	157.56	0.77
ELMO1	204513_s_at	113.54	159.16	0.71
FGD2	1559091_s_at	59.86	83.93	0.71
GABARAPL2	209046_s_at	1381.64	1939.66	0.71
HLA-C	211799_x_at	1723.51	2429.1	0.71
	214459_x_at	7637.63	9321.63	0.82
HLA-G	210514_x_at	1100.72	1548.41	0.71
	211529_x_at	4128.35	5806.05	0.71
	211528_x_at	4988.82	6509.67	0.77
LETM1	222006_at	132.77	186.85	0.71
LRRC6	206483_at	15.62	21.98	0.71
NADK	213607_x_at	62.62	88.78	0.71
NONO	210470_x_at	348.76	492.91	0.71
	208698_s_at	229	320.2	0.72
OTUB1	201245_s_at	276.26	387.29	0.71
PARP9	227807_at	96.37	136.44	0.71
PDCL	204448_s_at	49.67	69.86	0.71
PIGG	1563842_at	57.1	80.98	0.71
PRELID1	223032_x_at	472.42	662.34	0.71
RNF10	207801_s_at	446.08	626.47	0.71
SAP30L	239802_at	29.12	40.94	0.71
SFRS17A	210269_s_at	145.85	206	0.71
SNRPB	213175_s_at	1028.1	1456.83	0.71
	208821_at	749.76	966.47	0.78
SRA1	224130_s_at	161.66	227.02	0.71
SUDS3	233841_s_at	251.8	352.81	0.71
THBS3	209561_at	62.87	88.5	0.71
TM9SF4	212198_s_at	97.07	137.18	0.71
	212194_s_at	59.02	76.9	0.77
TSC2	215735_s_at	77.72	109.72	0.71
UBXN6	220757_s_at	83.85	118.37	0.71
WDR45	209217_s_at	111.47	156.47	0.71
	209216_at	336.14	439.24	0.77
ACIN1	201715_s_at	239.23	332.59	0.72
ACTN4	200601_at	67.58	93.9	0.72
ALDOA	214687_x_at	1213.45	1679.09	0.72
	200966_x_at	1160.39	1614.23	0.72
	238996_x_at	101.94	140.73	0.72
AMDHD2	219082_at	64.57	89.08	0.72
ARL8A	225347_at	113.78	157.22	0.72
ASB8	218841_at	90.31	125.51	0.72
CDC37	209953_s_at	415.53	573.66	0.72
COX6B1	201441_at	925.22	1280.22	0.72
CSNK2B	201390_s_at	519.02	725.32	0.72
DCTN3	204246_s_at	273.71	381.64	0.72
DVL3	201908_at	217.93	304.09	0.72
EIF6	210213_s_at	236.86	330.34	0.72
ERAP1	210385_s_at	19.67	27.36	0.72
GOLGB1	201056_at	94.46	131.86	0.72
ITGB5	201124_at	9.36	13.08	0.72
NDUFA13	220864_s_at	681.87	946.66	0.72
NKIRAS2	218240_at	106.24	146.72	0.72
	222105_s_at	70.3	91.41	0.77
PRKCA	215195_at	19.22	26.74	0.72
PSPC1	226574_at	261.37	364.41	0.72
RAB4B	233385_x_at	157.45	219.02	0.72
RABEP2	77508_r_at	102.92	142.07	0.72
SF3A1	201357_s_at	110.09	153.17	0.72
SLC43A3	210692_s_at	11.1	15.43	0.72
STAG3L4	222801_s_at	158.58	221.03	0.72
TAF10	200055_at	865.97	1204.4	0.72
TBCB	216194_s_at	563.1	782.59	0.72
	201804_x_at	580.3	794.65	0.73
TUFM	201113_at	585.35	808.91	0.72
ZDHHC5	224868_at	31.57	43.64	0.72
COMT	208817_at	60.89	83.69	0.73
CORO1B	221754_s_at	56.91	77.9	0.73
CTAGE9	215549_x_at	59.05	80.97	0.73
DNM2	202253_s_at	148.21	202.41	0.73
ERAL1	212087_s_at	47.52	64.77	0.73
FAM134A	218037_at	217.61	299.64	0.73
FLOT1	208749_x_at	267.87	366.64	0.73
GLE1	206920_s_at	56.12	77.27	0.73
HPS1	217354_s_at	57.87	78.96	0.73
L3MBTL2	1555815_a_at	46.09	63.32	0.73
MAN2B1	209166_s_at	305.91	418.91	0.73

MICALL1	221779_at	168.16	229.55	0.73
	55081_at	149.59	203.97	0.73
MLX	217909_s_at	93.33	127.99	0.73
MYO9B	217297_s_at	188.26	259.12	0.73
NCOA5	234471_s_at	49.96	68.01	0.73
PGLS	218388_at	326.05	446.18	0.73
POTEKP	210926_at	20.05	27.59	0.73
PSMC4	201252_at	273.46	374.48	0.73
SUPT6H	208830_s_at	62.74	85.9	0.73
TBC1D5	201815_s_at	79.7	109.75	0.73
TMEM87A	212204_at	599.48	816.99	0.73
AAAS	218075_at	70.86	95.85	0.74
ADRM1	201281_at	228.54	308.96	0.74
API5	201686_x_at	47.15	63.53	0.74
ASCC2	215684_s_at	112.28	152.69	0.74
ATP5G3	207507_s_at	829.58	1122.39	0.74
BAT3	213318_s_at	517.11	699.08	0.74
	201255_x_at	582.86	746.18	0.78
	210208_x_at	639.42	805.77	0.79
CHCHD1	226896_at	296.63	400.86	0.74
CIAPIN1	208968_s_at	256.25	348.03	0.74
COPA	208684_at	445.16	601.46	0.74
CYC1	201066_at	281.24	379.35	0.74
DMWD	33768_at	104.63	140.54	0.74
	1554429_a_at	44.17	58.91	0.75
GPI	208308_s_at	290.63	390.73	0.74
HEXA	201765_s_at	297.52	402.91	0.74
HLA-F	204806_x_at	2439.5	3277.05	0.74
	221875_x_at	3680.15	4820.98	0.76
MPHOSPH9	237158_s_at	20.8	27.95	0.74
NAT14	223284_at	35.45	47.66	0.74
NCSTN	208759_at	112.28	151.78	0.74
PDK2	202590_s_at	53.81	72.95	0.74
PHB2	201600_at	1659.23	2232.17	0.74
PITPNM1	203826_s_at	114.37	153.55	0.74
POLDIP2	217806_s_at	156.74	211.33	0.74
	222425_s_at	83.98	101.04	0.83
POLE3	208828_at	386.66	524.25	0.74
POLR2D	214144_at	26.2	35.29	0.74
PSMC3	201267_s_at	202.93	273.61	0.74
RAN	200750_s_at	1767.89	2402.76	0.74
VAMP3	211749_s_at	177.18	238.17	0.74
ALKBH5	1553101_a_at	201.05	267.61	0.75
AP3M1	222516_at	224.02	299.93	0.75
CHFR	223931_s_at	249.57	333.17	0.75
GART	210005_at	30.06	39.95	0.75
GFM1	225161_at	95.58	127.89	0.75
GTF2F1	202356_s_at	105.34	140.73	0.75
HLA-B	208729_x_at	6829.14	9080.46	0.75
	211911_x_at	9645.62	12697.18	0.76
	209140_x_at	11756.25	13756.08	0.85
INO80D	227924_at	72.35	96.56	0.75
LINS1	231976_at	49.56	65.66	0.75
LRSAM1	227675_at	42.99	57.58	0.75
MBD4	214048_at	556.53	746.26	0.75
MYO1C	32811_at	118.23	156.76	0.75
QKI	214543_x_at	135.82	180.54	0.75
RAB1B	220964_s_at	232.54	309.5	0.75
RAB2B	225074_at	411.71	549.94	0.75
STK24	208854_s_at	463.11	620.53	0.75
TLR9	223903_at	44.76	59.49	0.75
TMEM138	223113_at	364.83	487	0.75
TMEM199	225375_at	90.88	120.81	0.75
TTYH3	224674_at	67.5	89.98	0.75
TYMP	204858_s_at	55.92	74.1	0.75
ZNF207	200828_s_at	957.2	1276.6	0.75
AP2A1	234068_s_at	11.04	14.43	0.76
ARPC2	208679_s_at	2978.61	3917.51	0.76
CPSF7	217866_at	389.49	510.59	0.76
CRIPAK	241408_at	113.25	149.69	0.76
DGKZ	207556_s_at	70.7	93.38	0.76
	239342_at	18.53	23.31	0.79
EFTUD2	222398_s_at	523.61	691.75	0.76
GHITM	1554510_s_at	308.28	407.38	0.76
HIGD2A	209329_x_at	863.83	1138.19	0.76

KLHDC3	208784_s_at	50.4	66.71	0.76
MAPK9	210570_x_at	100.26	131.83	0.76
MAZ	212064_x_at	84.19	110.39	0.76
NFX1	202585_s_at	80.04	105.98	0.76
PIGT	217770_at	176.16	231.31	0.76
PPP2R4	208874_x_at	49.65	65.07	0.76
PSMB7	200786_at	618.69	815.34	0.76
RCAN1	215253_s_at	16.51	21.58	0.76
RNF40	206845_s_at	91.02	119.21	0.76
STK11	41657_at	108.7	143.5	0.76
STXBP2	209367_at	91.95	121.06	0.76
TET3	214754_at	47.14	61.97	0.76
TP53	211300_s_at	17.51	22.9	0.76
ZFAND3	218020_s_at	152.04	199.76	0.76
ZMAT2	224782_at	480.56	632.22	0.76
AP2A2	211779_x_at	200.23	260.06	0.77
	212159_x_at	197.41	250.46	0.79
ASB6	221657_s_at	81.36	105.14	0.77
CHRNB1	206703_at	15.91	20.58	0.77
FKBP9	212169_at	20.8	27.1	0.77
GSS	211630_s_at	70.87	92.51	0.77
HCFC1R1	45714_at	57.03	74.26	0.77
MED24	213043_s_at	110.27	142.72	0.77
PACSIN2	1554691_a_at	37.64	48.65	0.77
PITPNM2	1552923_a_at	16.27	21.12	0.77
PPM1G	200913_at	209.09	270.51	0.77
RNF181	223064_at	419.52	541.71	0.77
SEPHS1	208941_s_at	106.16	137.97	0.77
SH3GLB1	209091_s_at	467.6	605.71	0.77
TBC1D9B	212054_x_at	201.07	261.79	0.77
	215994_x_at	252.77	313.03	0.81
VPS39	212156_at	247.7	322.88	0.77
ANKFY1	219868_s_at	32.56	41.64	0.78
CINP	218267_at	42.18	54.12	0.78
CLK4	1568836_at	10.21	13.12	0.78
GTDC1	238585_at	28.98	37.38	0.78
MINK1	215909_x_at	45.28	58.32	0.78
MRPL10	224671_at	242.34	310.13	0.78
MRPL37	222993_at	137.38	176.55	0.78
MT1X	208581_x_at	107.4	137.44	0.78
NOL6	218199_s_at	53.3	68.28	0.78
PDE7B	220343_at	13.52	17.36	0.78
PFDN6	222019_at	15.77	20.33	0.78
POLR2C	208996_s_at	214.06	274.54	0.78
RBM5	201394_s_at	627.69	802.62	0.78
RNF220	219988_s_at	213.33	271.96	0.78
SFRS9	201698_s_at	1050.28	1349.47	0.78
SHC1	201469_s_at	32.48	41.81	0.78
	214853_s_at	542.8	690.11	0.79
TBC1D10B	220947_s_at	95.85	122.5	0.78
TCEB2	200085_s_at	1005.06	1289.11	0.78
AGBL5	218480_at	38.74	49.32	0.79
BCL7C	219072_at	174.02	219.67	0.79
BECN1	208946_s_at	472.34	597.93	0.79
CHMP1A	201933_at	198.38	251.16	0.79
DAP	201095_at	173.1	220.08	0.79
INPP5K	202781_s_at	91.76	116.3	0.79
MAP4	200836_s_at	74.17	93.96	0.79
PES1	202212_at	108.04	137.19	0.79
PRR13	217794_at	966.55	1216.61	0.79
SART1	200051_at	240.29	302.84	0.79
SSBP4	229270_x_at	78.84	99.24	0.79
URM1	208101_s_at	78.7	99.91	0.79
USF2	202152_x_at	290.56	369.05	0.79
WIZ	52005_at	60.9	77.4	0.79
YWHAZ	200640_at	3434.94	4364.69	0.79
	200638_s_at	3561.13	4430.95	0.8
DAPK3	203890_s_at	8.93	11.12	0.8
DNAJC4	223371_s_at	36.28	45.49	0.8
EWSR1	210011_s_at	431.66	537.19	0.8
GUK1	200075_s_at	362.7	452.46	0.8
HDGF	200896_x_at	339.31	425.41	0.8
HLA-A	215313_x_at	8696.31	10850.22	0.8
	213932_x_at	8315.98	10246.72	0.81
HMGXB3	212431_at	155.42	195.46	0.8
LSM12	212532_s_at	190.07	237.95	0.8
MBD1	208595_s_at	97.66	121.51	0.8

PIAS2	1555513_at	6.77	8.48	0.8
RAB11B	217793_at	42.16	53	0.8
	34478_at	9.75	12.04	0.81
SAPS2	202791_s_at	115.44	144.58	0.8
SH3BP1	213633_at	82.43	102.99	0.8
CASP2	34449_at	23.12	28.64	0.81
HDAC3	216326_s_at	293.08	361.55	0.81
NFAT5	224984_at	1080.49	1338.21	0.81
BRMS1	215631_s_at	188.22	229.62	0.82
DDX49	31807_at	189.04	231.57	0.82
DNAJB12	202865_at	10.15	12.35	0.82
GTF3C5	217876_at	32.9	40.36	0.82
HSF1	213756_s_at	23.7	29.07	0.82
LYPLA2	202292_x_at	217.71	264.85	0.82
SLC25A46	212833_at	441.71	540.71	0.82
SMC1A	239688_at	16.17	19.83	0.82
TNPO2	226428_at	119.3	145.6	0.82
MT1P2	211456_x_at	155.51	187.28	0.83
UCK1	223142_s_at	37.74	45.71	0.83
VPS4A	217913_at	280.77	338.33	0.83
VPS33A	204590_x_at	36.98	44.25	0.84
PHF17	218517_at	345.51	250.37	1.38
	225816_at	337.11	216.98	1.55
PTPDC1	238841_at	21.9	14.81	1.48
EPM2AIP1	202909_at	1076.63	710.97	1.51
PLEKHF2	218640_s_at	1225.11	745.75	1.64
ZNF83	236429_at	36.03	21	1.72
MUDENG	232156_at	127.26	73.62	1.73
FARSB	232063_x_at	29.79	16.65	1.79
PMAIP1	204285_s_at	4922.05	2739.24	1.8
FAM115A	212979_s_at	38.42	21.05	1.83
NFATC3	210556_at	110.43	59.75	1.85
KLF2	219371_s_at	2013.43	1046.35	1.92
ZNF844	228346_at	241.51	125.51	1.92
ZNF652	205594_at	359.47	182.32	1.97
LRRC37A2	221740_x_at	150.13	73.85	2.03
RUFY2	241996_at	170.73	82.94	2.06
MALAT1	227510_x_at	51.72	24.86	2.08
	223940_x_at	1557.89	403.1	3.86
	224568_x_at	1050.4	271.31	3.87
HELLS	220085_at	84.46	36.08	2.34
FBXL3	242829_x_at	411.56	173.34	2.37
HNRPLL	225386_s_at	101.36	39.59	2.56
ARRDC3	224797_at	634.01	219.77	2.88

Genes from the NOTCH1 pathway are shown in bold.

Supplementary Table 10. Clinical and biological features of the 363 CLL patients included in the validation study

Parameter	Category	Cases (%)
Gender	Male	223 (61%)
Age (years), median (range)		63 (27-94)
Binet stage	A	286 (79%)
	B	54 (15%)
	C	23 (6%)
Rai stage	0	190 (52%)
	I-II	142 (39%)
	III-IV	31 (9%)
Lymphocytes (x10^9/L), median (range)		45 (1.2-410)
Haemoglobin (g/L), median (range)		141 (46-168)
Platelets (x10^9/L), median (range)		189 (21-470)
LDH	>UNL	46/335 (14%)
β_2-microglobulin (mg/L)	>UNL	81/294 (28%)
Lymphocyte doubling time	< 1 year	65/177 (37%)
CD38	High	68/225 (30%)
ZAP-70	High	103/289 (36%)
IGHV	Unmutated	134/255 (53%)
Genetic abnormality	del(13)(q14.3)	99/236 (42%)
	del(11)(q22.3)	35/228 (15%)
	+12	39/246 (16%)
	del(17)(p13.1)	13/236 (6%)
10-year time TTP (95% CI)	Binet stage A	50% (43-57)
10-year time to Richter (95% CI)	All	5 % (2-8)
10-year OS (95% CI)	All	55 % (48-62)
Follow-up (years), median (range)		7.4 (0.1-23)

CD38 high: >30% of positive CLL cells; ZAP-70 high: ≥ 20% of positive CLL cells; IGHV unmutated: ≥ 98% homology with germline; TTP: time to progression; OS: overall survival

Supplementary Table 11. Primers used for the clinical validation series

Round of validation	Amplicon	Genomic Position (GRCh37)	Amplicon Size (bp)	Forward Primer Sequence	Reverse Primer Sequence
CV_A	CNOT3_Ex1/2	chr19:54646572+54647006	435	CCCACCTACCTCACTATGCTG	CTGGTCAACACCCAGAGGTC
CV_A	CNOT3_Ex3/4	chr19:54647048+54647529	482	TCAGCTCTAAGATGGATTGGG	GGTGACCTTCCCACCTCTCT
CV_A	CNOT3_Ex5/6	chr19:54647645+54648152	508	AAGTAGGGTCACGAGGCTCAG	GGAGCATAAATGACTGGCCTC
CV_A	CNOT3_Ex7/8	chr19:54649248+54649865	618	AGGGACTGAGGACAGGTTCTG	ATAGCCACTTTGGAGTGACCC
CV_A	CNOT3_Ex9	chr19:54650198+54650521	324	TAAGGACAGCCATTTGACCAG	CAGCACTGATTTCTGAGCCAC
CV_A	CNOT3_Ex10/11	chr19:54651799+54652631	833	ACAAAGATGGAGCCTGAGGTG	GTGTAACACCCGAGGGAGATG
CV_A	CNOT3_Ex12	chr19:54653150+54653611	462	TCTTTCTCCCATCTGTCTGCC	GAAAGCCAGAGAAAGAGATCCAG
CV_A	CNOT3_Ex13/14	chr19:54655818+54656437	620	CCTGTCATGGGTAGATTGTGG	CCTCACAGGTTCAGCCACTG
CV_A	CNOT3_Ex15	chr19:54656492+54656930	439	GCCCTGGGTCTTTCTGTACC	TAACATCTCTGGGCTGGAAGG
CV_A	CNOT3_Ex16	chr19:54657311+54657724	414	CCCATCTCCAAGAGGATTGTC	AGCCAGAGGGAGGAGTACTTG
CV_A	CNOT3_Ex17	chr19:54658898+54659234	337	AGGATGGATGAGAGTGTGTGC	CCTCCCTCCAGTCTTCCAG
CV_A	ROBO1_Ex1	chr3:79638882+79639192	311	AACTTGATTTGCAACACACACAA	TTTGAACTGCAATGTTTCACAA
CV_A	ROBO1_Ex2	chr3:79174491+79174818	328	GTGGTTGTTAGCTGGAAGCAG	GCTGTGTGTTGTCCTTGTGTG
CV_A	ROBO1_Ex3	chr3:78987677+78988194	518	TCATTGCCATGATCAAACAAA	AAGGCTAGGCCAATGTTATGC
CV_A	ROBO1_Ex4	chr3:78795780+78796276	497	GCTGACGCAAGCCTTTATTG	TGCTGGCCTCTGAAGAATATC
CV_A	ROBO1_Ex5/6	chr3:78766248+78767127	880	AGCCTCAGAGAAAGCAGGATG	TGCTTAGAAAGCAGTCTTCACTTATG
CV_A	ROBO1_Ex7	chr3:78763263+78763759	497	GCACATTCTTCGCCTCTACTTG	TGTGCTTTGAAGCATTGTGTG
CV_A	ROBO1_Ex8	chr3:78737667+78738180	514	GGCACTCAGAGATTTGCCAG	TTTAGAATTCCTTTCATTATGTGGG
CV_A	ROBO1_Ex9	chr3:78734815+78735152	338	GAAATAAATATATCTACTGGCACCCTCTC	ACATGTAAATGCATAAGTGCCAA
CV_A	ROBO1_Ex10	chr3:78719175+78719527	353	TCAATGCCAGTTGTTGGAGG	ACAACTGCTGTTGCTTGGATG
CV_A	ROBO1_Ex11/12/13	chr3:78716908+78717773	866	CTTGGGCAACTTGAACAACA	TTCCAGCCCAGTGAGTTCTT
CV_A	ROBO1_Ex14	chr3:78710930+78711335	406	AAGTTCACTACTGCAGCACAATAGG	GTTTCACCATGCTTTCATTCC
CV_A	ROBO1_Ex15	chr3:78709970+78710521	552	GGTAGGCCACTATTGAAGCAA	GGATGAAGCCAGACAGAATTT
CV_A	ROBO1_Ex16	chr3:78708713+78709017	305	TGTTGCTATTTGGACCAGCAG	ACCTGCTTTCAGTGTTCTCTAGC
CV_A	ROBO1_Ex17	chr3:78706147+78706491	345	TTAAATTAGCAATGGTGGGTGG	GCAACCCTGTGTTGATCTTTG
CV_A	ROBO1_Ex18	chr3:78700729+78701149	421	GGATACCGCAGCACAATAAAC	AGCCACAAATGAGTGCGTG
CV_A	ROBO1_Ex19	chr3:78696622+78697073	452	TTAAACACATGCCTGGTGTGAG	AATGCAGGGTCATTCACAGG
CV_A	ROBO1_Ex20	chr3:78695158+78695500	343	TGAATGGAGTCTTCGTCAAGG	TCACTAGTTGAATCAATCTGTTACGTC
CV_A	ROBO1_Ex21	chr3:78688809+78689244	436	CCCATGTAGGGAGAGGGAAAG	ATGTAAATGGAGGCACATCCC
CV_A	ROBO1_Ex22	chr3:78684847+78685346	500	GACAAGTTTCAACATCTAGTCGAGG	CTGATGACTCGCAAAGCTACG
CV_A	ROBO1_Ex23	chr3:78682936+78683265	330	TTGCAGAGACAAGACACATGC	GAGAAAGCAGAGATGTGAAATCG
CV_A	ROBO1_Ex24	chr3:78680188+78680636	449	AAAGGCAACTGTTTGTGGACC	ATGGCAGTACCTTGCATGTTC
CV_A	ROBO1_Ex25	chr3:78676328+78676800	473	ATTGCCTGGGTTTACTATGGG	AAAGGTGACATTTCATTACTATCCAA
CV_A	ROBO1_Ex26	chr3:78666632+78667403	772	TGGACTGTCATTTCCTTAGGC	CAGAAAGTAACCCGGTGGAG
CV_A	ROBO1_Ex27	chr3:78663715+78664110	396	CAGTGAATGCATTTGCTAGTCC	CAAGCCTGTTATTTGCTGGAAG
CV_A	ROBO1_Ex28	chr3:78655804+78656272	469	CATGGCCATATAGAGCCTAACTGT	TGGCTCCAATGATAATCCAAA
CV_A	ROBO1_Ex29	chr3:78649122+78649578	457	TCAGTTTCTGCCATCATAGACATTAG	TTTATACCTTAAATCACTGATGCTTCC
CV_A	ROBO1_Ex30	chr3:78647980+78648292	313	TTATCTGGCGTCATGTGTCATC	CACTGTGCAGAACCCAGTTTG
CV_A	ROBO1_Ex1short	chr3:79067485+79067822	338	GGTTTGATCGTGCAAAGTGG	AACCTTGGAATTCTTCCTCTGC
CV_A	SETD5_Ex1	chr3:9470353+9470848	496	CATGGGACACTGTGGCTGTA	ACCTTCAATGGGGGTAAAGG
CV_A	SETD5_Ex2	chr3:9475306+9475676	371	TGCCTGTACTGGCTTCATGT	CAGGAAGATCACTGTTCTGGAG
CV_A	SETD5_Ex3/4	chr3:9475837+9476698	862	CTTTCCTGGTTGAAGCCAAA	CTTTGGCCTCCCTAAGTGGT
CV_A	SETD5_Ex5	chr3:9477307+9477706	400	GGAAGTAATGGCTTACTGCATGTG	AAAGTGGCCCTTCAAACACC
CV_A	SETD5_Ex6	chr3:9482020+9482484	465	TATGCCACATCAATTGCCAAG	GCCTTGTACATTTGGAAGGAAC
CV_A	SETD5_Ex7/8	chr3:9483108+9483967	860	AGAAAAATTGAGCTTTCATCACA	TGGTTCCAGTCAAAGGGAAG
CV_A	SETD5_Ex9	chr3:9484811+9485309	499	CAAGCCAGTGCAGTATAATAGTTTCAG	TGAGCAGTATCATACAAGGTATCCG
CV_A	SETD5_Ex10/11	chr3:9486655+9487535	881	TTTAAGAGGGTGAATGGATAATGTAAC	AACCTTCAGGTCAAATCAATGAG
CV_A	SETD5_Ex12	chr3:9488538+9489100	563	TTGCCTTTATTTACAACTATTCCC	TGTGATATGGAGCAAAGCTGA
CV_A	SETD5_Ex13/14	chr3:9489286+9490440	1155	TCTGAATGCTTAAAGTGAATAGGAAA	AACTCACATGGGCAGAGTGG
CV_A	SETD5_Ex15	chr3:9495261+9495745	485	CAATGTTCCAATATGCTGAGACAC	AGTCCTTCTTTGTCGTCCTGG
CV_A	SETD5_Ex16	chr3:9505967+9506457	491	TGGTTGGATTGGAGGAGATAAC	AAGATAGTAAAGAAACATTTGGAATGG
CV_A	SETD5_Ex17	chr3:9512112+9512705	594	CCCATGTCTCCGTTGATTTT	GTGCGGAACACATAGCAGAA
CV_A	SETD5_Ex18	chr3:9514847+9515330	484	CCTGTTGCTGGTGATCCACT	ATGTTCCCATCAAGTGTTCCC
CV_A	SETD5_Ex19/20	chr3:9516062+9516925	864	TGCACTGGTTAGGGCTTCAT	GCATCTCAATGGCAAAGGAT
CV_A	SETD5_Ex21	chr3:9517036+9517893	858	TGCCTTTGGTTTCCTAAAAAGA	ATCCTCTGGCACCTTGTACC
CV_A	KLHL6_Ex1	chr3:183273063+183273534	472	CAACTGTTCAAGATTGGGCTC	GACAGCTGGTGATAAGTGGAGG
CV_A	KLHL6_Ex2	chr3:183245539+183245884	346	GTACAGGCCTGGGAGCTAGAC	AGGGCCTGATGAGTCTGAAAG
CV_A	KLHL6_Ex3	chr3:183225769+183226425	657	TTTGCTTGCATCTCTGTGAGC	TTAGCTCATCCTGTACTGTTGGAC
CV_A	KLHL6_Ex4	chr3:183217292+183217697	406	TCTTCTAGGATGGTTGCCTGG	GACCCGACTCCTAATGGCTC
CV_A	KLHL6_Ex5	chr3:183211654+183212142	489	TCACCACTTGGAAGGACATTG	TGTGGGATATTATGGGAGAAGC
CV_A	KLHL6_Ex6/7	chr3:183209643+183210579	937	GACTGGAGGAGGGTGAGAGG	AGCTCTGTCCTGTATCCTGGG
CV_A	MYD88_Ex1	chr3:38180123+38180599	477	CGCAGGAGAAAGAGGAAGC	ATGGGAGACAGGATGCTGAG
CV_A	MYD88_Ex2	chr3:38181249+38181670	422	CTAGAACAACCCAGCCAGAGG	GCATCTCACAGCAATCCAGAG
CV_A	MYD88_Ex3/4/5	chr3:38181798+38182796	999	CCTAGGCAGGGGACTCTTG	CAGGGCCTCAGAACAGTCTT
CV_A	CD8A_Ex1/2	chr2:87017310+87018091	782	GCTGTGTCTGTGAAATGGGAA	GGTGACATCACATCCTCCAAA
CV_A	CD8A_Ex3/4	chr2:87016338+87016897	560	CAGAAAACTCAACCCCAAGC	AACCACGCTCTGTTGCTCTT
CV_A	CD8A_Ex5	chr2:87015447+87015835	389	CTAACCCGGGCCTCCTTT	CATCCAATTTCAGCCCATGA
CV_A	CD8A_Ex6	chr2:87012969+87013273	305	AATGAAAGGGAAGGACTTGCTC	AAGTAAGGCATTTAGCACGGG
CV_B	MIR15/16	chr13:50622812+50623375	564	TGTTTTATCCCAAGTAAAAATCTGAA	ATTCTTTAGGCGCGAATGTG
CV_B	NOTCH1 Ex34.1	chr9:139391644-139392228	585	TCTCTGGGTGGGTTTCAGAAG	ACTTCTTCCTCCGTGCCTTG
CV_B	NOTCH1 Ex34.2	chr9:139391097-139391857	761	CTGCTGGACGAGTACAACCTG	TAGCTCATCATCTGGGACAGG
CV_B	NOTCH1 Ex34.3	chr9:139390431-139391284	854	AGTTTGAATGGTCAATGCGAG	AAGGCTCCTCTGGTCGGC
CV_B	NOTCH1 Ex26	chr9:139398980-139399548	569	ACGACCAGTACTGCAAGGACC	AGGTCCTCTCGGAACCTCC
CV_B	NOTCH1 Ex27	chr9:139397540-139397936	397	CTGCTGTCAGACCTGGCTTC	GTAGCAACTGGCACAAACAGC
CV_B	NOTCH1 Ex28	chr9:139396558-139397061	504	GGAGGAGAGTGGGTGAGGAG	GAGAAGTGAGGCTGAGCGAG
CV_B	LLGL1 Ex1	chr17:18128767+18129266	500	CCCAGTTCACCATTGTCTGG	ACCCACCCAGAGTCCGAG
CV_B	LLGL1 Ex2	chr17:18133109+18133447	339	GCTGAGGCACAGAGCAAGTAG	GAGCTGGACACTGCCAAGAC
CV_B	LLGL1 Ex3/4	chr17:18135635+18136280	646	GTGCAGAGTGGGAGCTTTAGG	AGTAAAGCAGGGTCCAACAGC
CV_B	LLGL1 Ex5/6/7	chr17:18137003+18137879	877	GAGTTATCAGCAGTGGCCCAG	TACCTTACAGCTGCCTCCTGC
CV_B	LLGL1 Ex8/9/10	chr17:18137787+18138699	913	AACTGGCTGAGGAGGGACTTC	CCGGGCTTAGAACTGAAACTC
CV_B	LLGL1 Ex11	chr17:18138681+18138983	303	GTTTCAGTTCTAAGCCCGGAG	GCAAGGAGGAACTACCTGGTC
CV_B	LLGL Ex12/13	chr17:18139817+18140326	510	GCACGTGCAGTAGGTGCTTAG	ACCAGACCCTCCAGCTCATC
CV_B	LLGL1 Ex14	chr17:18140711+18141176	466	CTGCTTCAGAGGCTCCAGG	TGCCACACCAAACACTAGAGG
CV_B	LLGL1 Ex15/16	chr17:18141229+18142038	810	GGACATTCTCAGTACCACCTGC	ACCTCCTGGAAGGTGCAGTC

CV_B	LLGL1 Ex17	chr17:18143813+18144278	466	GACCTGCAGTCTGTGGGAAG	AGGAAGGGTTTCTTGAGGGTC
CV_B	LLGL1 Ex18/19	chr17:18144655+18145450	796	ATGCAGCTGGGCTTTGTG	TAGGTGGATGACTTGAGGCTG
CV_B	LLGL1 Ex20/21/22	chr17:18145404+18146225	822	AGTGGGAGGGCAAGGATCTA	GTAGAAGACCAGGGCCAAATG
CV_B	FUT9 Ex1.1	chr6:96650974+96651850	877	TCTCTTCATTCCCACCGCTAC	AGATGGTCCCAGAACAACAGG
CV_B	FUT9Ex1.2	chr6:96651830+96652296	467	CCTGTTGTTCTGGGACCATCT	AACCGAATGGAATGATTGCTC
CV_B	RBBP9 Ex1	chr20:18477561-18477914	354	GCAACTACAACTCCCAGGATG	GACTTTAAGAGAAGGCAGCGG
CV_B	RBBP9 Ex2	chr20:18476374-18476773	400	GCCTGTAGAGGCAAATATGGG	ATAGGCATCTTTGGTTCCTCC
CV_B	RBBP9 Ex3	chr20:18474493-18474792	300	AGTTGCCTGTTTGAAAGGTGG	GGTCAGCACTGTTTCCTCTGC
CV_B	RBBP9 Ex4	chr20:18470836-18471252	417	TCAACCTCAACAGCAGGTAGC	CAGGCCTCTAGGGTTCAGTTC
CV_B	RBBP9 Ex5	chr20:18470302-18470734	433	GCTCCTTTGCGTCTCTATGTG	GGCACTTACGGAACTTAACTGG
CV_B	UNC119B Ex1	chr17:26879181-26879657	477	CTGCGGATGCTCCCACTT	ACACAGAGATCTCAGACCTGGC
CV_B	UNC119B Ex2	chr17:26875527-26875858	332	GACTGAGGTCCCTCTTTCCAG	CTCTGTGGCCCAAGGATTC
CV_B	UNC119B Ex3/4	chr17:26874551-26875272	722	CACTATTTGGCCAATGGTCTG	ACGGATGATATGGGAATGGTG
CV_B	UNC119B Ex5	chr17:26874173-26874523	351	CTGTTTGGCTGTCCCTGTACC	CCTTCCTCCCAACATTGACTC
CV_B	CASP8AP2 Ex1	chr6:90556207+90556506	300	TTGAAGCTAGGTTTATTCTCTGAATTG	AGAGGATCCCTTGAGCTCCC
CV_B	CASP8AP2 Ex2	chr6:90562699+90563117	419	CTTAAGCGATCCACCTGCC	TGGGTTCTGGAAGTTACTGTGG
CV_B	CASP8AP2 Ex3	chr6:90564416+90564859	444	AAGATTTGCACCACTGCCAC	AGATGCTTAAATTGGGTCAAGC
CV_B	CASP8AP2 Ex4	chr6:90565082+90565398	317	GCCTTCTTAGTGAATTGATTTGC	GCTACCTTGCTAAGGCTGGG
CV_B	CASP8AP2 Ex5	chr6:90566683+90567116	434	GGCTTATAAATGTATGATGCCTGC	GGTGGAGGCTGCATTGAG
CV_B	CASP8AP2 Ex6.1	chr6:90571736+90572608	873	CTACCATGCCTGGCCAAAG	TGGAATGAGGAAGTGATGCTC
CV_B	CASP8AP2 Ex6.2	chr6:90572588+90573377	790	GAGCATCACTTCCTCATTCCA	TCTGAGATACCGATCCTGGG
CV_B	CASP8AP2 Ex6.3	chr6:90573358+90574227	870	CCCAGGATCGGTATCTCAGA	CATGTGCACTTCTAGATAACAGGG
CV_B	CASP8AP2 Ex7.1	chr6:90575539+90576441	903	AAACAAACAGTGAATTGCCTCAG	TGCAAACAAATAATCAAGTTGGTC
CV_B	CASP8AP2 Ex7.2	chr6:90576418+90577368	951	GACCAACTTGATTATTTGTTTGCA	ATCTGACTTCAGAGGCGATGG
CV_B	CASP8AP2 Ex7.3	chr6:90577348+90578161	814	CCATCGCCTCTGAAGTCAGA	AATTGCCTCCAGAACATCCC
CV_B	CASP8AP2 Ex7.4	chr6:90578142+90578869	728	GGGATGTTCTGGAGGCAATT	CAGCCAGACAGATTAGATTTCTGAG
CV_B	CASP8AP2 Ex8	chr6:90580912+90581213	302	TTAAATTTCCCTGGTTGGTGG	GCTGTGCTGGGATTATAGCTG
CV_B	CASP8AP2 Ex9	chr6:90583385+90583703	319	TTTGAGACCAATTGGAAGATGAG	CAAGCTGAACATGCTGATGTG
CV_B	CCND2 Ex1	chr12:4253352+4253743	392	TAGCCAAAGGAAGGAGGTCAG	CTCCCAGGTTTAGGGCTCC
CV_B	CCND2 Ex2	chr12:4255283+4255732	450	CTTATCACGCATTCTGGTCCC	TCTGGCAGTGGTGATAGAAGG
CV_B	CCND2 Ex3	chr12:4258065+4258459	395	TCTAGCAGCCACGTCCTAATG	CCCTCCTCCAAACTCCTGAAC
CV_B	CCND2_Ex4	chr12:4397856+4398229	374	TGAAAGAACTCGAGGGAGGA	TCAAATCCGTCGTCTCCTTC
CV_B	CCND2_Ex5	chr12:4408951+4409299	349	ACGCATGTTTTCTCCGTAGG	GATCTAGGTGGGGGCAGAAT
CV_C	SNX7_E1	chr1:99127207+99127514	308	GAGACGTGGGAGCCAATG	CGGAGTTGCCTTGGAAGTT
CV_C	SNX7_E2	chr1:99150208+99150727	520	GCAATCCTTGAGGGTTTATACTTT	GCTAAGATCTTGGAATGTTCAGAAT
CV_C	SNX7_E3/4	chr1:99156530+99157353	824	TTTCTTAGAGACCCATTTATTTCACAA	CCACTGCATGATTTGTAACGG
CV_C	SNX7_E5	chr1:99160892+99161351	460	GCCACTTAGAGGAAACACTGTACC	TGCTATATTACACCAACAAATGTGAAG
CV_C	SNX7_E6	chr1:99164162+99164645	484	CACTGTGAGAGAAATAAGGCCC	AACAACGGATTCTTTCCTTCATC
CV_C	SNX7_E7	chr1:99167254+99167648	395	CAAATGACTGATGGCTTGGG	AAGCTTGTAGACTCACATGGTATCC
CV_C	SNX_E8	chr1:99203605+99204022	418	CCTGACTCAACCTGACTGAACC	TCTTTCAACATTGGACGGACC
CV_C	SNX_E9	chr1:99225421+99225770	350	CCCTTACTTAGCAGTAGCATCCC	ATAAGGACACGGCAGATCCAG
CV_C	GRM3_E1	chr7:86394404+86395035	632	CTAGCATGACACATTGGCTCC	TTGATATAATCCGTTGATGACGTT
CV_C	GRM3_E2	chr7:86415556+86416443	888	TCCTTCTTCTCCCTCCCCTA	TCTTGGCTTACCCGTGAAGT
CV_C	GRM3_E3	chr7:86468152+86469278	1127	CAGCTCCATTCAACCCAAAT	CACAAGAAACACTGGAGGAGAA
CV_C	GRM3_E4	chr7:86479469+86479939	471	ACACAAATGGTGTTTGCTTGC	ATGTGTTGCCTGTTCTTGGC
CV_C	GRM3_E5	chr7:86493521+86493943	423	CCATTGTATCCTTCATGCTATTACC	GACTGACCATGTCAGACCCTG
CV_C	SLC25A40_E1	chr7:87487745+87488161	417	GGCAATCACGGTCCTACCA	CTAATGTGATGCTGCTCACCC
CV_C	SLC25A40_E2	chr7:87485441+87485876	436	GCCTACACCTCATCACGTACAAC	GAGAAGTAATTCGTCTGTCTTTCACTG
CV_C	SLC25A40_E3	chr7:87483420+87483837	418	GTTTAATATGATTCTTAACCCTCTTGC	CTGAGTTTAGAATGCATTGCAGTAG
CV_C	SLC25A40_E4	chr7:87479038+87479357	320	ATCATCAAACATGTGGGTGG	CCTGAAAGAATGAAATTACCAATCA
CV_C	SLC25A40_E5	chr7:87477102+87477516	415	GAATGTCACAACTTAAACAGAAGCTC	TGGATGGCAAATGCCTACTT
CV_C	SLC25A40_E6	chr7:87476107+87476684	578	GGAAAGATCTGAGAGCCCAAC	GCCCTTGAGAGAGTTGAAACC
CV_C	SLC25A40_E7	chr7:87472964+87473324	361	AGTGCAGCCAAAGTGGATCTC	CCCACCCTAGTGGGTGTCTAAG
CV_C	SLC25A40_E8	chr7:87470896+87471289	394	AAGGAAAAACAAAAATGGCATC	GGAATTGTGGCTCTGCAACT
CV_C	SLC25A40_E9/10	chr7:87465489+87466209	721	TGGTGAAAAACATTCTTGCCTA	AGGGCCATGGACACTATGTT
CV_C	C13orf23_E1	chr13:39600283+39600643	361	TTTCAATCAGTGGTGAGCCTG	GAGCAGATTGAATGTAGTGCTGTTC
CV_C	C13orf23_E2	chr13:39598464+39598830	367	TCCCTCATCCATTCACTCTCTC	AACTGAAATGGGAATGCAAGTC
CV_C	C13orf23_E3	chr13:39597013+39597338	326	ATCTGCTGCAGTGTTGTAGGC	TGAAAGACAAGTTGCCAGGAG
CV_C	C13orf23_E4	chr13:39596381+39596743	363	ACGTCATGCCAAACAACTGTC	TGGTTGTTTCATTTGGTGTCAG
CV_C	C13orf23_E5	chr13:39591434+39591911	478	AAATTCAGTTAACACGTGCATTTC	TCAGTTTCTGGTAATCCTCTTGC
CV_C	C13orf23_E6.1	chr13:39586752+39587475	724	TGATGACTTGAAATGAAATGTGAAC	ATTGGGCCGTGCATATACTTC
CV_C	C13orf23_E6.2	chr13:39587205+39588085	881	GGTGGCTATTAATGACCCTGG	CCTTACCACCAACATCCCAAG
CV_C	C13orf23_E6.3	chr13:39587662+39588810	1149	CCACAGCTAACTGGCACTGAC	AAACTGGATCTGACTGGCTCC
CV_C	C13orf23_E7	chr13:39586083+39586536	454	TCATCAAATCTCTATGCATTCTGC	GATTCATATCAGGGAATGGCTG
CV_C	C13orf23_E8	chr13:39585379+39585705	327	TCTCCAGGATTACCCTCATTCTC	TCTGTGTGGCATGGGTAGAAG
CV_C	MBP_E1	chr18:74728689+74729043	355	GCTACGTGCCAGTTCTTCCTC	CAAAGACAGGCCCTCTGAGTC
CV_C	MBP_E2	chr18:74721604+74721960	357	TCTTGGGCATTTGTCTCTTCC	AGATTTGCTGGAGGACTTTGG
CV_C	MBP_E3	chr18:74701710+74702096	387	CAATCCGTGGCAGATACAGTC	GAGAGATCCAGCTTCCCTGAG
CV_C	MBP_E4	chr18:74700735+74701076	342	GGTAGCTCGGAGCCTAACTCTC	GTGCTCTGGGAGATCTGCATC
CV_C	MBP_E5	chr18:74696648+74697078	431	CAGCCACCCTTGTACTCAGC	AAACCAGAGCCCTCCTCTCTC
CV_C	MBP_E6	chr18:74692236+74692537	302	GTGGCCTGACCCTACTACGTG	GAGAGAGACACCCAATGGCTC
CV_C	UNC119B_E1	chr12:121148232+121148709	478	ATAGTTACCGCGCTGTGGAG	CCTTCCCCACCAGAGCTAC
CV_C	UNC119B_E2	chr12:121150896+121151373	478	TTGGATCCCCTAATTCTGTCC	CAAACATCAGCCTCTGAACG
CV_C	UNC119B_E3/4	chr12:121154353+121154870	518	TGCTGTTGGGACTGGTGATA	GCTGTAAGGGTCAGGGGATA
CV_C	UNC119B_E5	chr12:121157564+121157908	345	CTGCCTCCCTCTGTTCTGAC	AGCTTCGATCATCTCCGTGT
CV_C	SLC2A8_E1/2/3	chr9:130159507+130160414	908	GACATGACGCCCGAGGAC	ACACTCGAGAGACGGGACAC
CV_C	SLC2A8_E4	chr9:130162064+130162374	311	TGGTCAGGCCAGTCTCAAAC	GTCCCTTGACCTCACTGAACC
CV_C	SLC2A8_E5	chr9:130164755+130165105	351	AAGCTGAGTCTTTGAGGCTGG	CTTGGGAACAGGCTGTGCTAC
CV_C	SLC2A8_E6/7	chr9:130165853+130166444	592	CAAGCCTTAGGACAGTAGACCC	TCATCACAAATGCTCCCTGTC
CV_C	SLC2A8_E8/9	chr9:130167094+130167899	806	CAGGTGTGGTCATGGTGTTC	GAAAGTGCTTCTCGGCTGAC
CV_C	SLC2A8_E10	chr9:130169311+130169753	443	GTGGCAGAGCTGTCTCAGTG	CTCGAAGGTGTTCCTGAGTCC
CV_C	PIK3R2_E1	chr19:18266611+18267086	476	AGCCCTGCTTCAACCAATG	CCTCACTGGACAGAAGAGGAGAC
CV_C	PIK3R2_E2/3/4	chr19:18271199+18272083	885	CTGAGGTTCAGCCAGCAAAG	GACAGATGGCCAGGGCAG
CV_C	PIK3R2_E5	chr19:18271875+18272437	563	AGCGACGTGGATCAGTGG	AGAACGCAGCTTCTCACTCTG
CV_C	PIK3R2_E6/7/8	chr19:18272639+18273400	762	CTGTATCATCTCCTCCTCCGC	AAACCGCACATGTTCTATCCC
CV_C	PIK3R2_E9/10	chr19:18273673+18274316	644	CTGGACAACAGAGCAGCAAG	CTCTGTGACCTGCAACTGGA
CV_C	PIK3R2_E11	chr19:18276891+18277218	328	GAGTTGAGATGTGCCTTTACCC	TACTCTTTGCCTCCAAGTGGC

CV_C	PIK3R2_E12	chr19:18277845+18278216	372	CTGAGGTAGGAGGGTCACCTG	GCCACTTGGTGGAGAGAGAG
CV_C	PIK3R2_E13/14/15	chr19:18279235+18280213	979	GCTCAGCACCCACACAACT	AGAAATGAGGACCCCTGGAT
CV_C	GPCPD1_E1	chr20:5584861+5585290	430	AAGCACATTTCTGCTGTCAGG	TTTCACTTGTTTCTTTCCCGC
CV_C	GPCPD1_E2	chr20:5579165+5579614	450	GGTTGAGAGGCAGGGATTAAC	ACCTGGGCCCAAACCTATATC
CV_C	GPCPD1_E3	chr20:5573854+5574283	430	GAGTTCTATTCCTTATTTGACCAAAGC	TGTCTACCCACACAGCAAAGG
CV_C	GPCPD1_E4	chr20:5566764+5567078	315	AATCCCACTAAGCCATTGTGC	GTGATCTGGTACATTCCAGGC
CV_C	GPCPD1_E5	chr20:5564782+5565109	328	CACGTGTGGCCTGTCTTTG	AAATATGCTTATTCACCTTGGGC
CV_C	GPCPD1_E6	chr20:5560481+5560966	486	CAACAGTAGCAATCTGAGCGG	ATGAACAACGGTTTAGGCAGG
CV_C	GPCPD1_E7	chr20:5558804+5559370	567	CTGGTCTTGATTGGAAACATCA	TTCGAAGATTTAGCATATTTGGG
CV_C	GPCPD1_E8/9	chr20:5555980+5556808	829	CCTTTGCTGAGAAAACCAGTG	AGACGTGAGCCACTGACCC
CV_C	GPCPD1_E10	chr20:5554421+5554844	424	GTGTTGGCCATGTCGGTATTC	TGTCTTCATTCCGATCTTCACC
CV_C	GPCPD1_E11	chr20:5550658+5550960	303	GGACTGGCAGTAATTCACCAAG	TTTCTTACGTTCTTGCAGTTTGAG
CV_C	GPCPD1_E12	chr20:5547970+5548410	441	CAGAACCCTTCTCCCAAACAG	AATATAAGGGCTGGGAGGTGG
CV_C	GPCPD1_E13	chr20:5547222+5547432	211	AATGCAAATTCATTTCCGAAC	TGTGTGCTTTCTAGCTCACTCA
CV_C	GPCPD1_E14	chr20:5545635+5545860	226	CTTGCCTTGGTTTTCCTCAA	TACTTATGCTCTCGCGCTTG
CV_C	GPCPD1_E15	chr20:5541965+5542354	390	AGGGTCCAATAAGCATTGTGC	CATCTACCAGGTTGTCACTTTGAG
CV_C	GPCPD1_E16	chr20:5540514+5540936	423	CACCTCCAGACTGCTGCC	CTCTATGTCTGAATTATGACCTGTGTC
CV_C	GPCPD1_E17	chr20:5539202+5539648	447	GAGCCACTTGTTTGCTACAGTT	TTGAACTCATGACCGCAAGT
CV_C	GPCPD1_E18	chr20:5538453+5538849	397	ACCCACAGGCAAATGTCAAG	CGCCTGCTGTACTTTGGAGAG
CV_C	GPCPD1_E19	chr20:5528221+5528573	353	GGCATAGATCAACAATGCTCAG	AAAGGTTGAGTCAGGCAGGTC
CV_C	ACER2_E1	chr9:19408885+19409280	396	TCTGCGAACGAGTAACCTCC	AGAGACCCACGTCCAGCTTC
CV_C	ACER2_E2	chr9:19423765+19424152	388	TGCCAACTTTGTTATCAAGCC	TCCCAGTTTGCAAGGACATAC
CV_C	ACER2_E3	chr9:19424542+19424938	397	TGGGTGGTGTGAGTTTCTTTC	AGGCTGAGTTCAATCATGTGC
CV_C	ACER2_E4	chr9:19434858+19435163	306	GCAATGCCTGTACCTTGCTAAC	TACAGGACAGCAAAGAAGCCC
CV_C	ACER2_E5	chr9:19446132+19446497	366	TGTGGAGTGACTGTGTGTTGG	ACCACTTGCAGGGTGAACAG
CV_C	ACER2_E6	chr9:19450375+19450707	333	AAGGAGCAGGCTAAACTCAGG	TGGCTGTCTTCCTAGCAAAGG
CV_C	XPO1_E12/13/14/15	chr2:61719087+61720305	1219	TTGCCCTCCTATTTCCATTG	CCTACACATGGCTGGCTTG
CV_C	GJA10	chr6:90605515+90606001	487	TGCACAGTGACTCAGGAAGC	TCCCTCCTTTGCTCACTCAT
CV_C	EFR3A	chr8:132989282+132989540	259	TTTTAGGGATTTGGGAACCA	TGAAGATCAAAGACAGTGACAGAA