Supplemental information

High frequency of inactivating tetraspanin *CD37* mutations in diffuse large B-cell lymphoma at immune-privileged sites. Suraya Elfrink, Charlotte M de Winde, Michiel van den Brand *et al.*

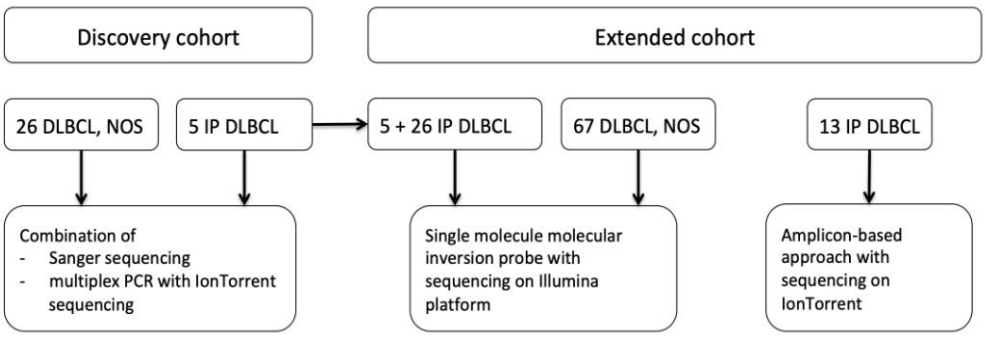
Supplemental materials and methods

Collection of DLBCL samples. The discovery cohort of 31 primary DLBCL diagnosis cases from 2001-2011 was collected from the archive of the Department of Pathology of the Radboudumc (Nijmegen, The Netherlands). For the validation cohort, primary DLBCL samples (n=106) diagnosed between 2001-2019 were collected from the archives of the Department of Pathology at the Radboudumc and CWZ hospital (n=62) (Nijmegen, The Netherlands), the Department of Pathology at the VU University Medical Center (n=8) (Amsterdam, The Netherlands), the Department of Pathology at the Leiden University Medical Center (n=13) (Leiden, The Netherlands), and the Dutch-Belgian Cooperative Trial Group for Hematology Oncology Group (HOVON)-46 and HOVON-84 (n=23) (Supplemental Table 1).

RNA isolation and cDNA generation. Frozen sections (10µm) of human DLBCL tumors were cut, resuspended in TRIzol reagent (Life Technologies, Bleiswijk, The Netherlands) and stored at -80°C until RNA isolation, according to manufacturer's instructions. 2µg RNA was treated with DNase I (amplification grade; Invitrogen, Bleiswijk, The Netherlands) and reverse transcribed to cDNA by using random hexamers and Moloney murine leukemia virus reverse transcriptase (Invitrogen).

Mutation analysis. Genomic DNA was extracted from formalin-fixed paraffin-embedded or frozen tissue. Mutation analysis for the *CD37* gene was performed using 1) Sanger sequencing of all exons (discovery cohort, cDNA and patient-matched control samples), 2) a combination of Sanger sequencing (exon 8) with a multiplex PCR and amplicon-based Ion Torrent sequencing for exons 1-7 (discovery cohort), 3) single molecule molecular inversion probe (smMIP) Next Generation Sequencing (NGS) analysis (validation cohort and PTL of discovery cohort) or (4) amplicon-based Ion

Torrent sequencing (samples obtained at Leiden University Medical Center). A summary of the different samples and analysis methods is indicated in the image below.



For Sanger sequencing of a part of the discovery cohort and the patient-matched control samples, the specific exons of *CD37* were amplified in a singleplex PCR with 50ng of DNA input for each reaction in a total volume of 50µl using the primers specified in Supplemental Table 3. For mutation analysis of cDNA, a PCR reaction was performed using the following primers: forward (located on *CD37* exon 1) 5'-AAGTACTTCCTCTTCGTTTTCAACCT, reverse (located on *CD37* exon 8) 5'-ATGAGGACTTGGAACCAGTC. Sequences were analyzed with SnapGene version 3.2.1 (GSL Biotech, Chicago, IL, USA).

For the multiplex PCR, three individual PCR reactions were performed with 100ng of DNA input for each reaction in a total volume of 25µL. Exons 1 through 7 were amplified in two multiplex PCRs (Pool A and B) and exon 8 was amplified in a singleplex PCR (Pool C) (see Supplemental Table 3 for primer sequences). The PCR product from exon 8 was analyzed by Sanger sequencing as sufficient number of reads could not be achieved by Ion Torrent sequencing using multiplex PCR. The PCR products from exons 1 through 7 were pooled and prepared for sequencing analysis using the NEXTflex DNA sequencing kit (Bioo Scientific, Austin, Texas, USA) according to the manufacturer's instructions. Sequencing was performed on an Ion Personal Genome machine (Life Technologies) according to the manufacturer's instructions. Sequences were analyzed with the SeqNext software package (JSI Medical Systems, Kippenheim, Germany). Variants were called if the absolute number of variant reads was 25 or higher and if the variant constituted at least 5% of the reads. Intronic

mutations, known single nucleotide polymorphisms (SNPs), mutations in untranslated regions, and silent mutations were filtered out.

For single molecule molecular inversion probe (smMIP) analysis, smMIPs were designed against *CARD11* exon 4-10, *CD37* exon 1-8, *CD79A* exon 5, *CD79B* exon 5-6 and *MYD88* exon 3-5 as part of a larger pool (see Supplementary Table 4 for smMIP sequences). smMIP pool and library preparation were performed as described for manual experiments in Eijkelenboom et al.[1] A total of 100ng genomic DNA was used as input divided over two capture reactions each containing 50ng genomic DNA and pooled before purification. Ligation was performed for 24 hours. Sequencing was performed on a NextSeq500 instrument (Illumina, San Diego, CA) according to manufacturer's protocol. Sequences were analyzed with the SeqNext software package (JSI Medical Systems). A minimal absolute coverage of 20 combined reads was required. This enabled detection of variants with at least 5% variant reads with 95% confidence. A sufficient number of reads for *CD37* exon 8 could not be obtained using smMIP. Variants were called at >5% variant reads. Intronic mutations (excluding splice sites), known SNPs, mutations in untranslated regions and silent mutations were filtered out. Variants were allocated a summary pathogenicity score as PA1 (validated polymorphism), PA2 (likely benign), PA3 (unknown significance), PA4 (likely pathogenic), and PA5 (pathogenic) (Supplemental Table 2).

For Ion Torrent amplicon based next generation sequencing of tumor samples obtained and analyzed at Leiden University Medical Center, at least 10ng per sample was used for library preparation with the LYMFv1 panel (1362 amplicons for 52 B-cell lymphomas associated genes). Samples were barcoded and pooled, before loading on a sequencing chip by the Ion Chef System (Thermo Fisher Scientific, Bleiswijk, The Netherlands) and sequencing was performed on the Ion S5 (Thermo Fisher Scientific). Generated sequencing data alignment was performed along the human reference genome (GRCh37/hg19) with TMAP 5.0.7 software. Quality control of sequencing runs were based on ratio and proportions of transition versus transversion variants, and read depth of the amplicons of interest. Variants were called by the Torrent Variant Caller (Thermo Fisher Scientific)

and analyzed by the Geneticist Assistant NGS Interpretive Workbench (GA; SoftGenetics, State College, PA, USA). Variants were called at >5% variant reads. Variants found in less than 100 reads, intronic mutations (excluding splice sites), known SNPs, mutations in untranslated regions and silent mutations were filtered out. Variants were categorized based on their pathogenicity as described above (Supplemental Table 2).

Modeling of CD37 mutations. Point mutations having a direct effect on CD37 protein sequence and deletion of residues 122-133 were mapped on the full-length crystal structure of tetraspanin CD81.[2] Sequence alignment of CD81 (Uniprot: P60033) and CD37 (Uniprot: P11049) was performed using Clustal Omega[3] and edited in Jalview.[4] Figures displaying the mapped mutations on CD81 crystal structure were generated using Pymol (pymol.org/2/).

Culture and transfection of B-cell lines. The human Burkitt lymphoma B-cell line BJAB (ATCC) and human DLBCL cell line OCI-Ly8 (kind gift from M. Spaargaren, Academisch Medisch Centrum Amsterdam) were cultured in RPMI-1640 (Thermo Fischer Scientific) containing 10% fetal bovine serum (FBS, Greiner Bio-One, Alphen aan den Rijn, The Netherlands), 1% antibiotics/antimycotica, Thermo Fischer Scientific) and 1% Ultraglutamine1 (UG1, 200mM Glutamine in 0.85% sodium chloride solution, Lonza, Basel, Switzerland). CD37-WT-GFP and CD37-WT-mCherry constructs were generated by cloning human wild-type CD37 construct into psGFP2-C1 and pmCherry-C1. The CD37-Gly88Asp-GFP and CD37-Gly65Glu-GFP mutant constructs were generated by introducing a c.G263A or a c.G194A point mutation, respectively, in hCD37-WT-GFP using site-directed mutagenesis (Q5® Site-Directed Mutagenesis Kit, New England Biolabs) according to manufacturer's instructions. Transfection of these DNA constructs was performed in BJAB cells using the Neon™ transfection system according to manufacturer's protocol (Invitrogen), or with Mirus transfection buffer (Ingenio® Solution) using aluminum electrode cuvettes (Lonza). Transfection of OCI-Ly8 cells was performed

using SF Cell Line 4D-Nucleofector™ X Kit L (Lonza). Electroporation of the cells was done using the AMAXA Nucleofector™ biosystem (BJAB program M-013, OCI-Ly8 program DN-103, Lonza).

CRISPR/Cas9 mediated knock-out of CD37 in BJAB cells. Five guide RNA pairs targeting the first two coding exons of the human CD37 gene were designed using the MIT CRISPR design tool[5] and cloned into the px335 Cas9 vector as described before.[6] BJAB cells were transfected with 1.0 µg of each of the two gRNA plasmids forming a pair and 0.5 µg pSGFP2-C1 using a Nucleofector 2b system (Amaxa, Lonza) according to the manufacturers guidelines using program M013 and Mirus transfection buffer (Ingenio® Solution). One day after nucleofection, GFP-positive cells were flow sorted on a FACSAria (BD Biosciences, San Jose, CA, USA). After expanding the sorted culture for a week, cells were stained for CD37 (FITC-labeled, Clone M-B371, BioLegend) and CD37-negative cells were flow sorted on a FACSAria (BD Biosciences). Absence of CD37 in the resulting polyclonal CD37 knock-out culture was verified by Western Blot and flow cytometry.

Western Blot. Transfected cells were lysed 16 hours after transfection in 1x sample buffer supplemented with 2.5% β-mercaptoethanol, boiled at 95°C for 5 min and separated on a 10% SDS-PAGE gel. Separated proteins were blotted onto a Polyvinylidene fluoride (PVDF) membrane. After blocking, membranes were incubated with both 1µg/mL rabbit anti-GFP (polyclonal, Rockland, Limerick, PA, USA) and 0.2µg/mL rat anti-human-α-Tubulin (clone YOL1/34, Novus Biologicals, Centennial, CO, USA) antibodies, and subsequently stained with secondary goat anti-rat AF680 (polyclonal, Invitrogen) and goat anti-rabbit IRDye800 (polyclonal, LI-COR, Lincoln, NE, USA) antibodies. Western blots were imaged on Odyssey CLx (LI-COR) and analyzed with the Image Studio software (version 5.0, LI-COR).

Immunofluorescence microscopy. Cells were adhered on uncoated or PLL-coated coverslips and fixed in 4% paraformaldehyde (PFA). Fixed cells were blocked and subsequently incubated with 15µg/mL

mouse anti-CD20-Alexa647 antibody (clone 2H7, BioLegend, San Diego, CA, USA) or 10µg/mL mouse anti-human major histocompatibility complex-I (MHC-I) antibody (clone W6/32, Abcam, Cambridge, UK), followed by 2µg/mL secondary goat anti-mouse-IgG2a-Alexa568 antibody (Life Technologies). Afterwards cells were stained with 0.3µg/mL 4',6-diamidino-2-phenylindole (DAPI) and embedded in Mowiol. Cells were imaged using Olympus FV1000 or Leica SP8 confocal laser scanning microscope. Images were analyzed with Fiji[7] and Cell Profiler[8] software. The latter was used for quantification of the ratio of mean GFP-intensity in the membrane over the cytoplasm. The MHC-I staining was used to define the membrane and the cytoplasm was calculated by subtracting the nuclear (defined by DAPI signal) and membrane area from the cell area (defined by DAPI, GFP, and MHC-I signal).

Flow cytometry. Cells were incubated with 10µg/mL mouse anti-CD37 antibody (clone HH1, Santa Cruz Biotechnology, Dallas, TX, USA), followed by 2µg/mL secondary goat anti-mouse-IgG1-Alexa647 antibody (Life Technologies). Fluorescence intensities were measured on a Cyan ADP Flow cytometer (Beckman Coulter, Brea, CA, USA) and analyzed using FlowJo X software.

Immunohistochemistry. Immunohistochemistry staining on lymphoma tissues was performed on formalin-fixed, paraffin-embedded tissue sections or microarrays following manufacturer's instructions using a monoclonal antibody specific for CD37 (clone 2B8, ThermoFisher Scientific and Novus Biologicals), followed by hematoxylin counterstaining.

Statistics. Statistical differences were determined using a paired or unpaired t-test or two-tailed Fisher's exact test as indicated (GraphPad Prism version 5). Differences were considered to be statistically significant at $p \leq 0.05$.

Visual abstract. Visual abstract was created with BioRender (Biorender.com).

Supplementary references

1. Eijkelenboom A, Kamping EJ, Kastner-van Raaij AW, et al. Reliable Next-Generation Sequencing of Formalin-Fixed, Paraffin-Embedded Tissue Using Single Molecule Tags. *J. Mol. Diagnostics*. 2016;18(6):851–863.
2. Zimmerman B, Kelly B, McMillan BJ, et al. Crystal Structure of a Full-Length Human Tetraspanin Reveals a Cholesterol-Binding Pocket. *Cell*. 2016;167(4):1041-1051.e11.
3. Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 2011;7(1):539.
4. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 2009;25(9):1189–91.
5. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 2013;31(9):827–832.
6. Pyzocha NK, Ran FA, Hsu PD, Zhang F. RNA-guided genome editing of mammalian cells. *Methods Mol. Biol.* 2014;1114:269–77.
7. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods*. 2012;9(7):676–682.
8. Carpenter AE, Jones TR, Lamprecht MR, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 2006;7(10):R100.
9. Livingstone CD, Barton GJ. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. *Comput. Appl. Biosci.* 1993;9(6):745–56.
10. Hans CP, Weisenburger DD, Greiner TC, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*. 2004;103(1):275–82.

Supplemental Figure Legends

Supplemental Figure 1. Sanger sequencing of cDNA of the DLBCL with a *CD37* splice-site mutation. Sanger sequencing results of cDNA of the DLBCL sample bearing *CD37* splice site mutation c.343-1G>C using a forward (left) and reverse (right) primer reaction. The 11-base-pair deletion is demonstrated by a double peak pattern in the cDNA sequence showing the exon 5 sequence in the forward reaction (left) and the exon 4 sequence in the reverse reaction (right) below the main peak pattern. White = exon 4, grey = exon 5. In the box the frame-shift as a result of the 11-base-pair deletion is visualized.

Supplemental Figure 2. *CD37* mutations mapped onto the crystal structure of tetraspanin CD81. (A) Sequence alignment of tetraspanins CD81 and CD37. Residues are colored coded in blue with shading intensity based on conservation of physical-chemical properties in the alignment.[9] Identical residues are highlighted in dark blue. Transmembrane domains are shown in a yellow frame, while mutated or deleted CD37 residues are shown in a red frame. (B) Ribbon representation of CD81 crystal structure[2] in two different views, where selected residues, corresponding to the *CD37* point mutations, have been mapped and shown as red spheres. Deletion of residues 122-133 in region A in the EC2 is highlighted in red. (C) Table displaying the *CD37* mutations affecting cDNA sequence and the respective protein sequence. Residues aligned in CD81 and mapped in (A) are shown including their location. In the right row, the predicted effects of the *CD37* mutations on the protein structure are listed.

Supplemental Figure 3. Mutations in *CD37* cause aberrant surface CD37 expression in OCI-Ly8 lymphoma cells. (A) Western blot analysis of CD37 protein expression in OCI-Ly8 lymphoma cells transfected with CD37-WT-GFP, CD37-Gly88Asp-GFP or CD37-Gly65Glu-GFP. The blots were probed with anti-GFP antibody to detect CD37-GFP expression (50-75kD; upper blot). α-Tubulin was used as a loading control (lower blot). The protein expression level of CD37-mutant-GFP was normalized to

CD37-WT-GFP for each experiment (n=4). *p=0.012, ns: p=0.11, paired t-test. Data represent mean±SEM. (B) Confocal microscopy images of OCI-Ly8 cells expressing CD37-WT-GFP and CD37-Gly88Asp-GFP or CD37-Gly65Glu-GFP (green) co-stained for MHC-I (red) to define the plasma membrane. Single cell images (bottom left) show representative cells for CD37-WT-GFP and CD37-Gly88Asp-GFP (n=2) or CD37-Gly65Glu-GFP (n=4). Scale bar = 5µm.

Supplemental Figure 4. Mutant CD37 has a dominant-negative effect on cell surface expression of wild-type CD37. (A) Sanger sequencing results of cDNA of the lymphoma with the CD37-Gly88Asp missense mutation with a forward (left) and reverse (right) primer reaction. Grey = exon 3, white = exon 4. The arrow indicates the location of the mutation. The additional presence of the wild-type sequence is indicative of a heterozygous mutation. (B) Representative histogram (left) and quantification (right) of flow cytometry analysis of CD37 knock-out BJAB cells co-transfected with CD37-WT-mCherry and CD37-Gly88Asp-GFP in an 1:1, 1:3 or 1:5 ratio or control vector psGFP2-C1 (ratio 1:5) showing cell surface expression of CD37. Viable cells were gated on successful transfection (high CD37-WT-mCherry fluorescence intensity and co-expression of GFP). gMFI=geometric mean fluorescence intensity. *p<0.05, unpaired t-test. Data represents mean±SEM of three independent experiments. (C) Confocal microscopy images of CD37 knock-out BJAB cells co-transfected in an 1:5 ratio with CD37-WT-mCherry (magenta) and CD37-Gly88Asp-GFP (not shown) or control vector psGFP2-C1 (not shown) co-stained for CD20 (cyan) to identify the plasma membrane. Representative overview images (left) and single cell images (right). Scale bar = 10µm.

Supplemental Figure 5. CD37 protein status determined by immunohistochemistry in IP-DLBCL. CD37 protein expression was determined in tissue sections of IP-DLCBL with and without CD37 mutation.

Supplemental Tables

Supplemental Table 1. IP-DLBCL patient characteristics.

	PCNSL (n=21)	**PTL (n=23)**
Male/female (n)	16/5	23/0
Median age and age range (years)	67 (48-81)	71 (51-84)
Ann Arbor stage (n)		
I-II	15	4
III-IV	3	7
Unknown	3	12
B-symptoms	2/18 (11%)	3/10 (30%)
ECOG		
0	2	7
1	4	1
2	4	2
3	3	0
4	2	0
Unknown	6	13
LDH elevated	3/14 (21%)	4/10 (40%)
CD37 protein expression (immunohistochemistry)		
Present	7	10
Absent	11	13
Unknown	3	0
Cell of origin classification*		
GCB	3	7
non-GCB	17	15
Unknown	1	1
University hospital		
Radboudumc Nijmegen	9	14
VU medical center Amsterdam	0	8
Leiden UMC	12	1

ECOG: Eastern Cooperative Oncology Group performance status; PCNSL: primary central nerve system lymphoma; PTL: primary testis lymphoma.
*Based on immunohistochemistry results from CD10, BCL6, and MUM1 according to the Hans algorithm.[10]

Supplemental Table 2. Mutation analysis of *CARD11, CD37, CD79A, CD79B* and *MYD88* in IP-DLBCL.

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
CARD11	83	c.1071C>A p.(Asp357Glu)	52	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 0.995)	COSM41655	PA4
	123	c.377G>A p.(Gly126Asp)	35	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 0.998)	COSM85863	PA4
CD37	26	c.343-1G>C	33	Splice site mutation	NA	PA5S
	58	c.194G>A p.(Gly65Glu)*	35	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	NA	PA4^{+}
	59	c.194G>A p.(Gly65Glu)*	49	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	NA	PA4^{+}
	65	c.263G>A p.(Gly88Asp)	25	ALIGN-GVGD: C0 SIFT: tolerated (score 0.28) POLYPHEN 2.0: probably damaging (score 1.000)	NA	PA3/PA4^{+}

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	75	c.109C>T p.(Leu37Phe)	5.7	ALIGN-GVGD: C0 SIFT: tolerated (score 0.17) POLYPHEN 2.0: probably damaging (score 1.000)	NA	PA3
	82	c.656T>G p.(Val219Gly)	22	ALIGN-GVGD: C0 SIFT: tolerated (score 0.06) POLYPHEN 2.0: benign (score 0.001)	NA	PA3
	85	c.266_267+10	22	Splice site mutation	NA	PA5S
	85	c.279G>T (p.Met93Ile)	41	ALIGN-GVGD: C0 SIFT: tolerated (score 0.72) POLYPHEN 2.0: benign (score 0.000)	NA	PA3
	85	c.364_399del p.(Val122_Asn133del)	28	36 bp in-frame deletion, causing loss of 12 amino acids, including strongly conserved amino acids.	NA	PA4
	127	c.827G>A p.(Arg276Gln)	49	ALIGN-GVGD: C0 SIFT: tolerated (score 0.12) POLYPHEN 2.0: possibly damaging (score 0.921)	NA	PA3
	130	c.268-2A>C	46	Splice site mutation	NA	PA4S

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	130	c.279G>A p.(Met93Ile)	44	ALIGN-GVGD: C0 SIFT: tolerated (score 0.72) POLYPHEN 2.0: benign (score 0.000)	NA	PA3
	130	c.337G>A p.(Ala113Thr)	42	ALIGN-GVGD: C0 SIFT: tolerated (score 0.56) POLYPHEN 2.0: benign (score 0.000)	NA	PA3
	132	c.460G>A p.(Gly154Ser)	8.3	ALIGN-GVGD: C55 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	NA	PA4
CD79A	70	c.568-2_610del	39	Splice site mutation	NA	PA4S
	82	c.568-2_610del	20	Splice site mutation	NA	PA4S
CD79B	26	c.589T>G p.(Tyr197Asp)	40	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM1737939	PA5
	64	c.590A>C p.(Tyr197Ser)	31	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220733	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	64	c.623A>G p.(Tyr208Cys)	32	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	NA	PA4
	65	c.555T>A p.(Asp185Glu)	84	ALIGN-GVGD: C0 SIFT: tolerated (score 0.22) POLYPHEN 2.0: probably damaging (score 0.999)	NA	PA3
	65	c.571A>G (p.Met191Val)	84	ALIGN-GVGD: C0 SIFT: tolerated (score 0.43) POLYPHEN 2.0: benign (score 0.244)	NA	PA3
	65	c.589T>A p.(Tyr197Asn)	84	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM144393	PA5
	66	c.590A>T p.(Tyr197Phe)	13	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 0.999)	COSM220735	PA5
	67	c.590A>C p.(Tyr197Ser)	21	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220733	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	68	c.553-1G>A	12	Splice site mutation.	NA	PA5S
	69	c.590A>C p.(Tyr197Ser)	94	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220733	PA5
	72	c.590A>T p.(Tyr197Phe)	26	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 0.999)	COSM220735	PA5
	73	c.590A>C p.(Tyr197Ser)	34	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220733	PA5
	76	c.589T>G p.(Tyr197Asp)	37	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM1737939	PA5
	76	c.676C>T p.(His22Tyr)	37	ALIGN-GVGD: C0 SIFT: deleterious (score 0.01) POLYPHEN 2.0: probably damaging (score 0.999)	NA	PA4

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	77	c.590A>G p.(Tyr197Cys)	37	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220736	PA5
	78	c.590A>C p.(Tyr197Ser)	28	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220733	PA5
	81	c.589T>A p.(Tyr197Asn)	96	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM144393	PA5
	85	c.589T>A p.(Tyr197Asn)	32	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM144393	PA5
	123	c.599T>A p.(Leu200Gln)	44	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM6904276	PA4
	126	c.589T>G p.(Tyr197Asp)	48	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM1737939	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	128	c.553-2A>C	34	Splice site mutation	NA	PA4S
	130	c.127G>A p.(Ala43Thr)	44	ALIGN-GVGD: C0 SIFT: tolerated (0.52) POLYPHEN 2.0: benign (score 0.008)	NA	PA3
	131	c.589T>G p.(Tyr197Asp)	14	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM1737939	PA5
	132	c.590A>G p.(Tyr197Cys)	88	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220736	PA5
	133	c.562A>T p.(Lys188*)	27	Nonsense mutation	COSM7338571	PA5
	134	c.589T>C p.(Tyr197His)	42	ALIGN-GVGD: C0 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM220734	PA5
MYD88	26	c.794T>C p.(Leu265Pro)	35	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	63	c.794T>C p.(Leu265Pro)	14	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	65	c.794T>C p.(Leu265Pro)	45	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	66	c.794T>C p.(Leu265Pro)	12	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	67	c.794T>C p.(Leu265Pro)	22	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	69	c.728G>A p.(Ser243Asn)	95	ALIGN-GVGD: C45 SIFT: deleterious (score 0) POLYPHEN 2.0: possibly damaging (score 0.770)	COSM85943	PA4
	70	c.794T>C p.(Leu265Pro)	44	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	71	c.794T>C p.(Leu265Pro)	17	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	72	c.794T>C p.(Leu265Pro)	61	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	73	c.794T>C p.(Leu265Pro)	22	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	76	c.794T>C p.(Leu265Pro)	35	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	77	c.794T>C p.(Leu265Pro)	27	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	78	c.794T>C p.(Leu265Pro)	15	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	80	c.794T>C p.(Leu265Pro)	27	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	81	c.794T>C p.(Leu265Pro)	53	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	82	c.794T>C p.(Leu265Pro)	46	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	83	c.794T>C p.(Leu265Pro)	51	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000).	COSM85940	PA5
	87	c.794T>C p.(Leu265Pro)	65	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	91	c.794T>C p.(Leu265Pro)	7.2	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	126	c.794T>C p.(Leu265Pro)	67	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	128	c.794T>C p.(Leu265Pro)	31	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	129	c.656C>G p.(Ser219Cys)	49	ALIGN-GVGD: C15 SIFT: deleterious (score 0.05) POLYPHEN 2.0: probably damaging (score 0.998)	COSM85944	PA4
	130	c.880A>C p.(Thr294Pro)	32	ALIGN-GVGD: C0 SIFT: deleterious (score 0.04) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85945	PA4
	131	c.794T>C p.(Leu265Pro)	18	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	132	c.794T>C p.(Leu265Pro)	53	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5

Gene	Sample no.	Nucleotide change[a]	Variant allele frequency (VAF, %)	Pathogenicity prediction[b]	COSMIC entry[c]	Pathogenicity summary[d]
	133	c.794T>C p.(Leu265Pro)	78	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	134	c.794T>C p.(Leu265Pro)	43	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5
	135	c.794T>C p.(Leu265Pro)	52	ALIGN-GVGD: C65 SIFT: deleterious (score 0) POLYPHEN 2.0: probably damaging (score 1.000)	COSM85940	PA5

[a]Numbering according to: *CARD11*: NM_032415 (hg19); *CD37*: NM_001774 (hg19); *CD79A*: NM_001783 (hg19); *CD79B*: NM_001039933 (hg19); *MYD88*: NM_002468 (hg19).

[b]COSMIC: Catalogue Of Somatic Mutations In Cancer. Cancer.sanger.ac.uk/cosmic

[c]ALIGN-GVGD: Grantham Variation Grantham Deviation; risk classes are C0, C15, C25, C35, C45, C55, and C65, with higher classes indicating a higher risk of a damaging mutation. SIFT: Sorts Intolerant From Tolerant; scores range from 0 to 1, with a score of ≤0.05 predicted to be damaging. POLYPHEN 2.0: scores ranging from 0 to 1, with higher scores predicted to be more damaging.

[d]Pathogenicity summary score ranging from PA1 to PA5. PA1: known validated polymorphism with a frequency of >1%; PA2: probably not pathogenic, neutral mutation with no effect on amino acid sequence or splicing; PA3: variants of unknown significance, possibly pathogenic, amino acid change outside of a conserved domain or an amino acid change with ambiguous result from different prediction programs; PA4: likely pathogenic, amino acid substitutions with a predicted pathogenic effect, small in-frame deletions in conserved domains or mutations in consensus splice sites (PA4S); PA5: pathogenic, truncating mutations with loss of conserved domains, or known pathogenic mutations from literature.

* Germline *CD37* mutation analysis was performed for the patient who presented with both PTL (#58) and PCNSL (#59), and c.194G>A was detected in both germline DNA and the two lymphoma samples, while this was not the case for germline control analysis for patients #65 and #82 with c.263G>A and c.656T>G mutations, respectively.

[+] Functional analysis of mutations c.263G>A (p.Gly88Asp) and c.194G>A (p.Gly65Glu) showed aberrant CD37 cell surface expression (Figure 2).

Supplemental Table 3. Primers targeting *CD37* used for single- and multiplex PCR and Sanger sequencing of genomic DNA

Target exon	Pool	Forward primer sequence (5'--> 3')	Reverse primer sequence (5'--> 3')	Genomic coordinates (GRCh37)
Exon 1+2	A	CTGTCTCCCCCACTGTCAGC	GGGTAGCCATGAGGCAACT	49838733...49838888
Exon 1+2	B	TGAGTTGCCTCATGGCTACC	GAGACCCCAGGAAGTTGCTT	49838868...49839104
Exon 3	B	CTGTGCCACACAGCTCATCA	CACTCACCAGGCCCAGGA	49840124…49840297
Exon 3	A	CCTCTGCAGATCTGGTCCAA	CCATTTTTGAGGGTGGGGATA	49840180...49840349
Exon 4	B	TCCTATCCCCACCCTCAAAA	GGCCACTGCAGGAAGCTC	49840326...49840498
Exon 5	B	GTGGGGCGGGGAAGATAAG	AGTCCGGGAGGGAACTGGAG	49841137...49841319
Exon 6	A	CCCTCTGACTCGTATCCCTCT	CGTTGGTCGCCGACAAGTT	49841931…49842074
Exon 6	B	ACTGGTTCCAAGTCCTCATCC	ATTCGACAGGTTTATGCTCCGAA	49841988...49842227
Exon 7	A	GGGCTTGAGAGTCCCAGAAA	CAGGGCCATCTAGGGATTTAGG	49842535...49842724
Exon 8	C	CACCCCTTCGACTTCTCTGA	GCGAACTGGGGATTAAACAA	49843467...49843677

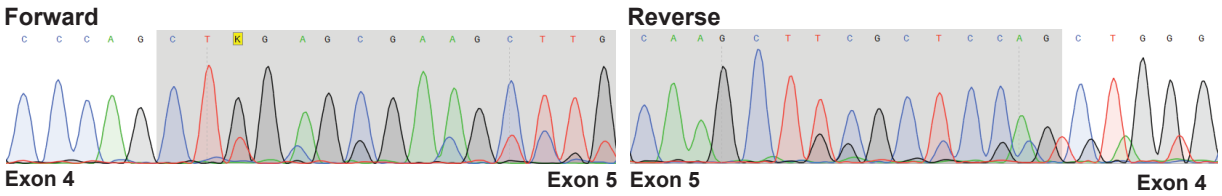
Supplemental Table 4. Overview of all MIPs targeting *CARD11*, *CD37, CD79A*, *CD79B* and *MYD88*

ROI	Amp	Chr	probe_ strand	mip_scan_ start_position	mip_scan_ stop_position	MIP Sequence
CARD11-E04	1	7	-	2985515	2985626	CCCAGAACTGTACAAACTGGTGANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCTTGCGTTCCATCAGAT
|	2	7	+	2985481	2985592	AGGGGAAAGACCAGATCACNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGTGGAGAATCTCCGAGTGGG
|	3	7	+	2985446	2985557	GCCCCTTGGTATGTAGAATGTCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGTTCCCAATGGCAAAAG
|	4	7	-	2985435	2985546	GGGAACATCTGGACTCTAANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTACCAAGGGGCAAAGGGGCTAT
CARD11-E05	1	7	+	2984150	2984261	CTCACCAACGTTGTGCCAAGCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGTTCATCAGGAAGTGCGTG
|	2	7	-	2984101	2984212	GCCAAGGACCTGCAACGCTGCGAGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAACCTCCTTTTTCCAC

ROI	Amp	Chr	probe_ strand	mip_scan_ start_position	mip_scan_ stop_position	MIP Sequence
|	3	7	+	2984049	2984160	GCCTTCCTCCACTGTGGAGAGGGGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTATCTGCTTCTTCTCAT
|	4	7	-	2984035	2984146	ATGACGCTGACGCGCGTGGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGAAGGCCACGAGGGCCTCAC
|	5	7	+	2983988	2984099	CCTTCATCTGCTGCTGCAGCTTGANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCTCTTCCTTCATCTTG
|	6	7	-	2983943	2984054	GGTGAAGGACGACAACTANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCTGCTGGCCAGGTTGCGGCAGC
|	7	7	+	2983883	2983994	GCTCCTGGAAGGTTAGCAGCTNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCTCCTCATGACCGCCATG
|	8	7	-	2983869	2983980	GGAGCCGAGACCTCCAACTCGANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGAGCGGTACTACAAGATG
|	9	7	-	2983838	2983949	TGCCCAAGGGCTACTCCGCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGGACAGCTACAATGACGAGC
|	10	7	+	2983787	2983898	GCTGTGCGTAGCGCATGGCTAAGTNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTATGGGAGAATTGAGCC
CARD11-E06	1	7	-	2979502	2979613	AGTCTCTAAAACTGAAGAATGACNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAAGCGTTGCCTTTTCTG
|	2	7	+	2979460	2979571	CCCAAGGGTGCCCAGGGGAGANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGTTCCAGAACCTGCTCCT
|	3	7	-	2979390	2979501	CATCCAGGTGAGGGCCTGGGGGCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTAAGCTGGAGAGAAATC
|	4	7	+	2979354	2979465	GCCGATTTTCAATGTCATTCTTCANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCTCACCCTCCCACTGC
CARD11-E07	1	7	+	2978368	2978479	GGAGGTTTCAGAGGGTCAGGGTNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGCAGGTTGTAGATCCTG
|	2	7	-	2978358	2978469	AACCTGCAGGAGGAGGCCCGCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGAAACCTCCTCCCTCCCAC
|	3a	7	+	2978312	2978423	CAGGATGGCCTTGTCTGAGTCTGGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCGATCCCCACTCCcA
|	3b					CAGGATGGCCTTGTCTGAGTCTGGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCGATCCCCACTCCtA
|	4	7	-	2978281	2978392	CTCGGCCGGCCCCCTTCACANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTACACGACCGCAAGGAGGCCC
CARD11-E08	1	7	-	2977622	2977733	GGAAAGGACTGTGAAATGTNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGGCGGGAAGGCACGAGCAGA
|	2	7	+	2977576	2977687	CAATATGCGCATCGAGCAATCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTCCACCTCCTCCAGCTGCA
|	3	7	-	2977519	2977630	GCAGGGCGGCGAGGCTGCANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGAGGACCTGGAGCTCAAGTGC
|	4	7	+	2977464	2977575	GCATGACCGTGTTCATGCGGTGCTNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTACTGGACAAAACACTC
CARD11-E09	1	7	+	2976848	2976959	GGGCCAGCCACATCGTCTGTGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCGAGTACTGTGTCTGAGC
|	2	7	-	2976752	2976863	ATGGTGCGGCGGGAGGCCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCCGCTTCTCCCTGGCAGGCCTT
|	3	7	+	2976743	2976854	CGGGAGTGGAAGGCCTGCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGTTGACGATGCAGGCCTCCCG
|	4	7	+	2976676	2976787	GCGGATCTGCTTCCTGTACTTGTCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTCCCGCCCCACCTGGT
|	5a	7	-	2976661	2976772	ACGGGCAGGACCCCACGCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGATCCGCGAgCTGGAGGAGAAG
|	5b					ACGGGCAGGACCCCACGCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGATCCGCGAaCTGGAGGAGAAG
|	6a	7	+	2976571	2976682	GTTGCTGTCCTTGGAGAGGCGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCAGGGGGCACCTGGCTTCa
|	6b					GTTGCTGTCCTTGGAGAGGCGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCAGGGGGCACCTGGCTTCt
CARD11-E10	1	7	+	2974185	2974296	GGGCTTGGCGATGCGCCACCAGTNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGTGGAAGAATCGTCAG
|	2a	7	-	2974173	2974284	CCACCTCgGAGGAGTCACCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCCAAGCCCCACCACTAACGC
|	2b					CCACCTCaGAGGAGTCACCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCCAAGCCCCACCACTAACGC
|	3	7	+	2974086	2974197	GGTCCTGGGGCTGGCATCCCCAAANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCCTCCACACCCATGCA
|	4	7	-	2974073	2974184	AGGGAGGGGGCGCAGGTGCANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCAGGACCAATGGTCAAGAAG
CD37-E01	1	19	+	49838759	49838870	GTTGCCTCATGGCTACCCANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTCCCCCACTGTCAGCACCTCT
|	2	19	-	49838785	49838896	GGTAAGCGGTCCACTCACNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGGGTTAGCGGCAGGGGCTGGG
CD37-E02	1	19	-	49838925	49839036	GAGGGCTGGGTTAGCGGCAGGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGCCCCCCTCACCCACAAA
|	2a	19	-	49838960	49839071	AGATGcGGATACGGGGGGTGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGTTGCTTGGGTTGGGGGAGG
|	2b					AGATGtGGATACGGGGGGTGNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGTTGCTTGGGTTGGGGGAGG
|	3	19	+	49838971	49839082	CCAAGCAACTTCCTGGGGTCTCCCNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCATCTCCTCTCCCCAG
|	4	19	+	49839027	49839138	GCAGGCAGGCTACAGGCTNNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGATCCTCATTGACAAGACCAG
CD37-E03E04	1	19	-	49840148	49840259	GTGATGATGAGCTGTGTGGCANNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCAGCGGAGCTCCTTGAGG

ROI	Amp	Chr	probe_ strand	mip_scan_ start_position	mip_scan_ stop_position	MIP Sequence
|	2	19	+	49840159	49840270	TCCGCTGCCTCCTGGGCCTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTCATCATCACTCCCCTCACCT
|	3	19	-	49840187	49840298	GCAGAGGCACGAAGGCCAAGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTAGGGACGGAGAGGGATGGT
|	4	19	+	49840257	49840368	GGTGGTCAGCCTGATCTCTCCACTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTCCTGGGTTGTGTGGG
|	5	19	+	49840364	49840475	CAGGTGAGCTTCCTGCAGTGGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAATGGCCTGGGCCGGGGTT
|	6	19	-	49840370	49840481	CCCTGCAACCCCGGCCCAGGCCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGGCCACTGCAGGAAGCT
|	7	19	+	49840425	49840536	GTGGGAGAGGGTGGAGAGAGACNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTTTGGGATGCTGCTGCTC
CD37-E05	1	19	+	49841141	49841252	GGCCGAGGAGAGCTGGGACNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCCTGAGAGGGGGAGGGGTGG
|	2a	19	-	49841195	49841306	AGCTTCGCTCCAGCTGAGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGGCGGGTCAGTCCGGgAGGGA
|	2b					AGCTTCGCTCCAGCTGAGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGGCGGGTCAGTCCGGcAGGGA
|	3	19	-	49841203	49841314	GTCCCGCAAGCTTCGCTCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCAGGGCTGAGGCGGGTCAGTCC
CD37-E06	1	19	-	49841936	49842047	GAGGGACCGGGACTGGAGANNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTCGCCGACAAGTTGTAGCAGG
|	2	19	+	49841954	49842065	CGACCAACGACTCCACAANNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCCTCTGACTCGTATCCCTCTCC
|	3	19	+	49842048	49842159	CGCTGTCCCTGCAGAGAGCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGGCGCACCGCGTGCCCTGCT
|	4	19	-	49842054	49842165	AGCAGGAGCAGGGCACGCGGTGCGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGTGGCTCTCTGCAGGG
|	5	19	+	49842095	49842206	CGGAGCATAAACCTGTCGAATGGGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCACAATCCTAGATAAG
|	6	19	-	49842111	49842222	GCTGGGGCAAGATCACCTTATCTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGGCCCCGCCCCATTCG
CD37-E07	1	19	+	49842548	49842659	GCCTACTCGAGGTGATCTGGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGGGCCTGGGCTTGAGAGTC
|	2	19	-	49842562	49842673	GGGATTCTTTCTGGGACTCTCAANNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGGGGGCGGGGCCAGAT
|	3a	19	+	49842582	49842693	GCGATCGGCCCTAAATCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCCCTTTAACTTTTCCCTACACcC
|	3b					GCGATCGGCCCTAAATCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCCCTTTAACTTTTCCCTACACtC
CD37-E08	1	19	-	49843483	49843594	AGAAGTCGAAGGGGTGCTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGGCGGGGCGGGACTTTGGGGA
CD79A-E05	1	19	-	42384927	42385038	GGAAATGAGGCAGCGAACATCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGTAGGGGTGTCACGGCTT
|	2	19	+	42384940	42385051	CCTACTCCTGCCAGGCTGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCTCATTTCCATCCCAGGGCCTG
|	3	19	-	42384976	42385087	GGAGATGTCCTCATACATNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGAGCTGAGACACTGGAGCTGG
CD79B-E05	1	17	+	62006777	62006888	CCTGCTGGGCCAGGCTGGGGANNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTAGGACTCAGAGCTGCTGGG
|	2	17	-	62006747	62006858	GCCACTATCTGCTGGTGTGGTTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGGGACACTAACACTCTGA
CD79B-E06	1	17	-	62006635	62006746	GACAGGGGAAGTGAAGTGGTCTGTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGTCCTCGGGGTCAGTG
|	2	17	+	62006599	62006710	CGGGACCACACCCCAACCACACCANNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCTCTCACTCCTGGCCT
MYD88-E03	1	3	-	38181858	38181975	GGTGGTCAGAGCTCCATGGGAAGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCACACAACTTCAGTCGA
|	2	3	+	38181867	38181984	GTTGTGTGTGTCTGACCGCGATGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTGACCACCACCCTTGTG
|	3	3	+	38181952	38182069	AGGCCACGGGGTGGGTGCGTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCAGTTTGTGCAGGAGATGAT
|	4	3	-	38181962	38182079	CAGTTGCCGGATCATCTCCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTTCATGCATCCACGCACCCAC
MYD88-E04	1	3	+	38182203	38182320	GCACTCAGCCTCTCTCCAGGTAAGNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTTCCCAGGGGATATGCT
|	2	3	-	38182226	38182343	CAGGTCCTGTGGCAACTTAGTTCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCCAGAGCAGGGTTGAGC
MYD88-E05	1	3	-	38182577	38182688	GTGGGACAAGCCCAGTCTTCNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTGCAGACAGTGATGAACCTCA
|	2	3	+	38182619	38182730	AATCTTGGTTCTGGACTCGCCTNNNNNNNNCTTCAGCTTCCCGATATCCGACGGTAGTGTCATGGCACCCCTTGGCTT

Supplemental Figure 1



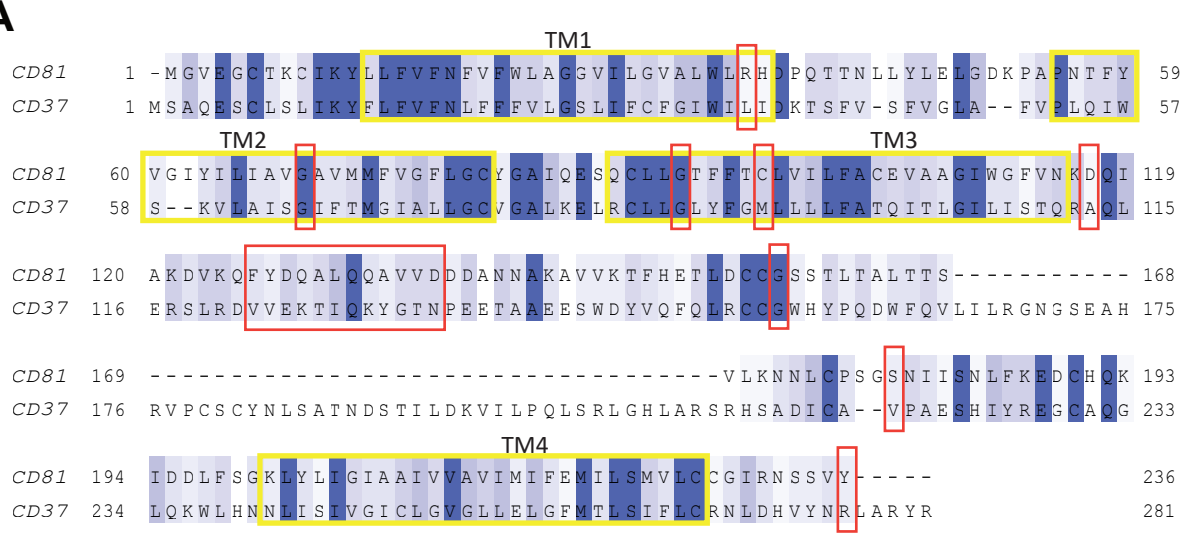
Wild-type RNA CC CAG CTG GAG CGA AGC TTG CGG GAC GTC

Patient RNA CC CAG CTT GCG GGA CGT CGT

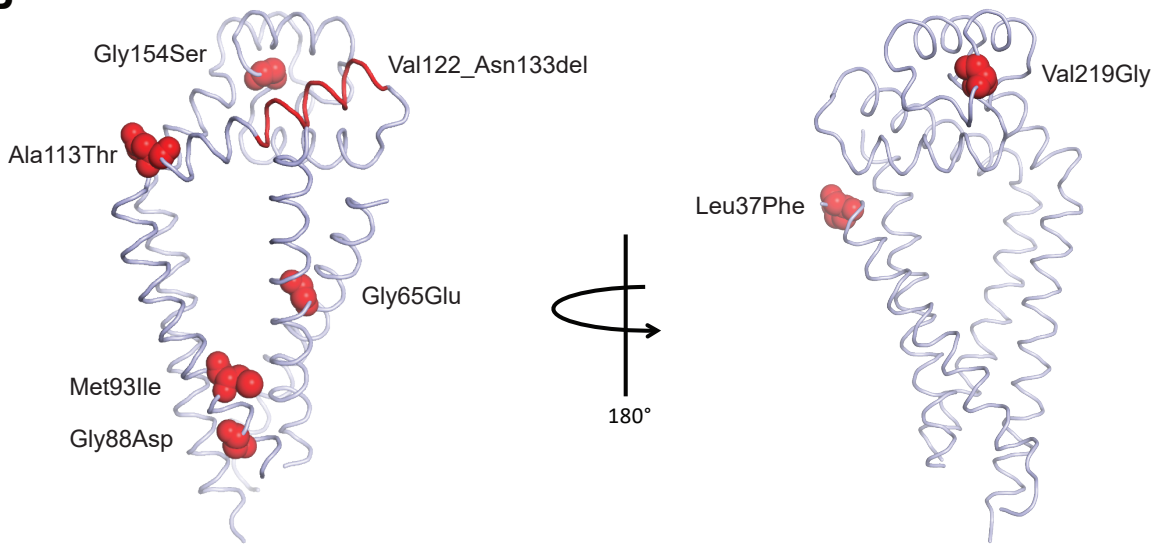
Exon 4 Exon 5

Supplemental Figure 2

A



B

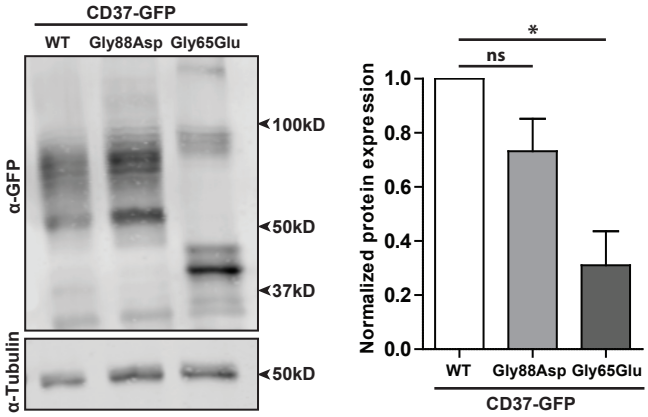


C

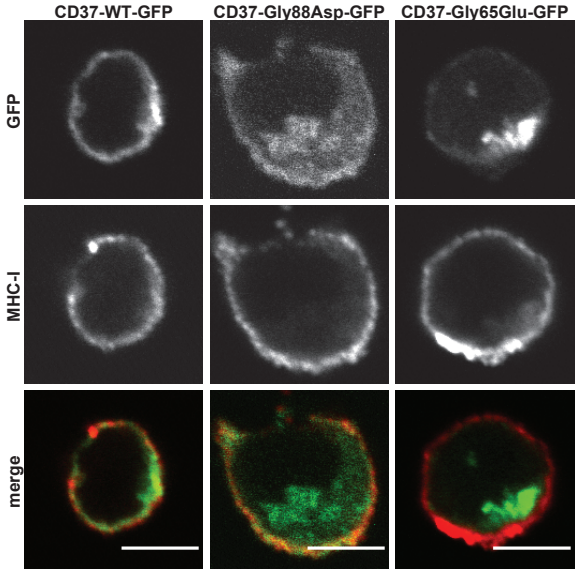
cDNA *CD37*	Protein residue CD37	Residue aligned in CD81	Location in CD81	Possible effect on protein structure
109C>T	Leu37Phe	Arg36	TM1	Cannot be correctly aligned; residue positioned near the water-lipid interface, possible minor effect
194G>A	Gly65Glu	Gly69; fully conserved	TM2	Major effect; introduces charge in transmembrane region
263G>A	Gly88Asp	Gly92; fully conserved	TM3	Major effect; introduces charge in transmembrane region
279G>A or T	Met93Ile	Cys97	TM3	Minor effect; high similarity between Met and Ile
337G>A	Ala113Thr	Asp117	EC2	Minor effect; solvent-exposed residue in EC2
364_399del	Val122-Asn133	Phe126-Asp137	EC2	Major effect; hampers correct folding of EC2
450G>A	Gly154Ser	Gly158; fully conserved	EC2	Major effect; introduces clash with C region
656T>G	Val219Gly	Ser179	EC2	Cannot be correctly aligned; harmful if pointing inside the protein core, but minor effect if pointing outwards
827G>A	Arg276Gln	Tyr236	C-terminus	No structural information available; this residue is found at the C-terminus of CD37 (rich in positively charged residues), therefore mutation to Gln is likely harmful

Supplemental Figure 3

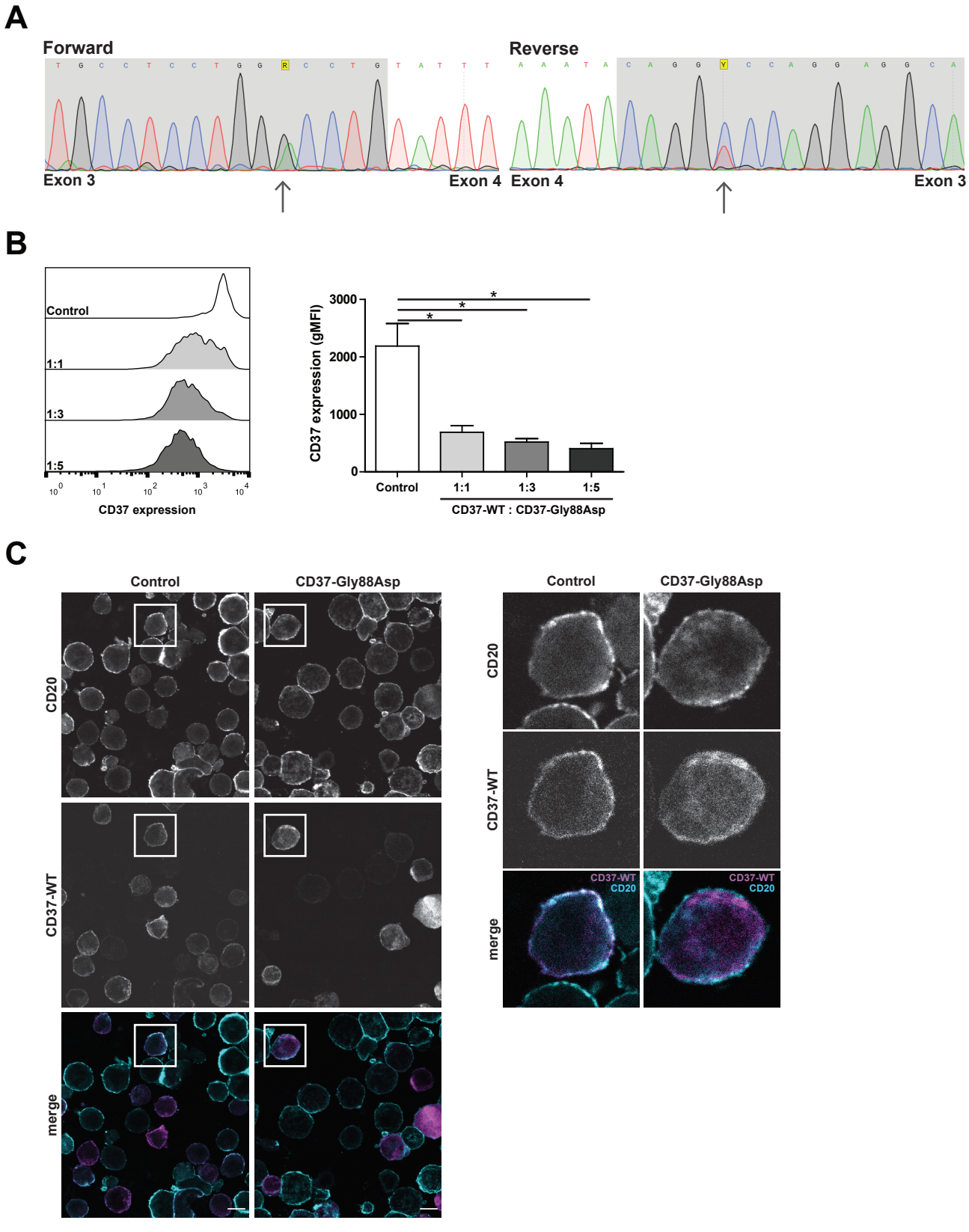
A



B



Supplemental Figure 4



Supplemental Figure 5