Whole exome sequencing reveals *NOTCH1* mutations in anaplastic large cell lymphoma and points to Notch both as a key pathway and a potential therapeutic target

Hugo Larose,[1,2] Nina Prokoph,[1,2] Jamie D. Matthews,[1] Michaela Schlederer,[3] Sandra Högler,[4] Ali F. Alsulami,[5] Stephen P. Ducray,[1,2] Edem Nuglozeh,[6] Mohammad Feroze Fazaludeen,[7] Ahmed Elmouna,[6] Monica Ceccon,[2,8] Luca Mologni,[2,8] Carlo Gambacorti-Passerini,[2,8] Gerald Hoefler,[9] Cosimo Lobello,[2,10] Sarka Pospisilova,[2,10,11] Andrea Janikova,[2,11] Wilhelm Woessmann,[2,12] Christine Damm-Welk,[2,12] Martin Zimmermann,[13] Alina Fedorova,[14] Andrea Malone,[15] Owen Smith,[15] Mariusz Wasik,[2,16] Giorgio Inghirami,[17] Laurence Lamant,[18] Tom L. Blundell,[5] Wolfram Klapper,[19] Olaf Merkel,[2,3] G. A. Amos Burke,[20] Shahid Mian,[6] Ibraheem Ashankyty,[21] Lukas Kenner[2,3,22] and Suzanne D. Turner[1,2,10]

[1]Department of Pathology, University of Cambridge, Cambridge, UK; [2]European Research Initiative for ALK Related Malignancies (ERIA; *www.ERIALCL.net*); [3]Department of Pathology, Medical University of Vienna, Vienna, Austria; [4]Unit of Laboratory Animal Pathology, University of Veterinary Medicine Vienna, Vienna, Austria; [5]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, UK; [6]Molecular Diagnostics and Personalised Therapeutics Unit, Colleges of Medicine and Applied Medical Sciences, University of Ha'il, Ha'il, Saudi Arabia; [7]Neuroinflammation Research Group, Department of Neurobiology, A.I Virtanen Institute for Molecular Sciences, University of Eastern Finland, Finland; [8]University of Milano-Bicocca, Monza, Italy ; [9]Diagnostic and Research Institute of Pathology, Medical University of Graz, Graz, Austria ; [10]Center of Molecular Medicine, CEITEC, Masaryk University, Brno, Czech Republic; [11]Department of Internal Medicine – Hematology and Oncology, University Hospital Brno, Czech Republic; [12]University Hospital Hamburg-Eppendorf, Pediatric Hematology and Oncology, Hamburg, Germany; [13]Department of Pediatric Hematology/Oncology and Blood Stem Cell Transplantation, Hannover Medical School, Hannover, Germany; [14]Belarusian Center for Pediatric Oncology, Hematology and Immunology, Minsk, Belarus; [15]Our Lady's Children's Hospital, Crumlin, Ireland; [16]Perelman School of Medicine, Philadelphia, PA, USA; [17]Department of Pathology and Laboratory Medicine, Cornell University, New York, NY USA; [18]Institut Universitaire du Cancer Toulouse, Oncopole et Université Paul-Sabatier, Toulouse, France; [19]Department of Pathology, Hematopathology Section, UKSH Campus Kiel, Kiel, Germany; [20]Department of Paediatric Oncology, Addenbrooke's Hospital, Cambridge, UK; [21]Department of Medical Technology Laboratory, College of Applied Medical Sciences, King Abdulaziz University, Jeddah, Saudi Arabia and [22]Ludwig-Boltzmann Institute for Cancer Research, Vienna, Austria

 doi:10.3324/haematol.2019.238766

Received: September 19, 2019.

Accepted: April 9, 2020.

Pre-published: April 23, 2020.

Correspondence: *SUZANNE D. TURNER* - sdt36@cam.ac.uk

Supplemental Material

Supplemental Methods

Supplemental Figures

Supplementary Figure 4. ChIP-seq demonstrating binding of STAT3 at the *notch1* promoter region.

Supplementary Figure 5. Changes in gene expression upon GSI treatment in T-ALL.

Supplementary Figure 6. Proliferation of cells lines when treating with GSI 1.

Supplemental Figure 7. Quality Controls for bioinformatic processing.

Supplemental Tables

Supplementary Table S1: Characteristics of the patient tumour samples obtained from the Children's Cancer and Leukaemia Group (CCLG) tissue bank for which WES was conducted or whose WES data was downloaded from the Sequence Read Archive (SRP044708).

Supplementary Table S2: Characteristics of the 78 patient samples that were employed to validate the presence of Notch1 variants in a larger patient population.

Supplementary Table S3: Characteristics of the 89 patient tissues that were stained for NOTCH1 expression on the tissue microarray, from which clinical data allowed computation of the 10-year EFS.

Supplementary Table S4: WES details for the three different patient sample cohorts, including sample names and coverage detail. The Childrens' Cancer and Leukaemia Group (CCLG), is a UK-based charity and tissue bank.

Supplementary Table S5: Detailed list of antibodies used in this study.

Supplementary Table S6: Detailed list of variants found in at least 10% of the WES cohort.

Supplementary Table S7 (Excel file): Details of copy number variations detected in more than 1 patient. Sheet 1: Copy number altered autosomal regions detected in patient samples. Sheet 2: Genes involved in recurrent copy number gains. Sheet 3: Genes involved in recurrent copy number losses.

Supplementary Table S8: Details of the mutations found in the NOTCH1 pathway by GSEA.

Supplementary Table S9: Clinical data pertaining to the validation cohort.

Supplementary Table S10: Detailed list of oligos used in this study.

Supplementary Table S11: Cell line description.

Supplementary Table S12: Detailed list of plasmids used in this study.

Supplemental Methods

RNA Extraction and RT-qPCR

RNA was extracted using a standard Phenol/Chloroform protocol using TRI reagent (Sigma-Aldrich) before DNA degradation with TURBO DNAse (Thermo Fisher Scientific), following which 2 µg of total RNA was reverse transcribed into cDNA using SuperScript III Reverse Transcriptase (ThermoFisher

Scientific). SYBR-Green qPCR analysis was then performed using the QuantStudio™ 6 Flex Real-Time PCR System in accordance with the manufacturer's instructions.

Sanger Sequencing

PCR was used to amplify the region of interest (for oligo sequences see Supplementary Table S10) after which primer dimers were removed using ExoSAP-IT (Thermo Scientific), fragments of interest were labelled using the BigDye Terminator v3.1 Cycle Sequencing Kit (Thermo Scientific) and excess removed using DynaBeads (Thermo Scientific), all according to the manufacturer's instructions. Finally, samples were sequenced using an ABI3730 Sequencer (48 capillaries) and analysed using SeqScanner 2 software. Samples were sequenced in duplicates when variant validation was not immediately evident.

Western Blot

Cells were lysed in Pierce RIPA buffer (Thermo Scientific) supplemented with Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific). Laemmli buffer with a 5% final concentration of 2-mercaptoethanol was added to the samples, which were then boiled and proteins separated by TGX 10% Acrylamide Gel (Bio-Rad, Hercules, California, US). Proteins were transferred to a 0.45 µm PVDF membrane (Immobilon, Burlington, Massachusetts, US) using the BioRad Trans-Blot Turbo system in 1X Transfer Buffer (Bio-Rad) for 7 minutes at 25 V and 1.3 A. Membranes were then incubated in 5% BSA (Acros Organics, Hampton, New Hampshire, US) in TBST before exposing to the indicated antibodies (Supplementary Table S5) diluted in 3% BSA in TBST and finally visualised with HRP substrate (Millipore) using an LAS4000 imager (Fujifilm, Minato, Japan).

Modelling Molecular Structure and Predicting the Impact of Mutations

Multiple templates from the Protein Data Bank were selected for the extracellular (PDB IDs: 4XL1, 4XBM, 5MWB, 4D90) and intracellular domains (PDB IDs: 1YMP, 1OT8) of NOTCH1 using FUGUE[1]. The model was derived using MODELLER, which was used to predict the structure of the mutant protein.

The same process was used to build the JAG1 model (PDB ID: 2VJ2). SDM[2], DynaMut[3] and mCSM[4] were used to predict the impact of mutations on NOTCH1 stability and NOTCH1 -ligand interaction.

Bioinformatics Analysis

Paired-end sequencing data from 7 previously analysed ALK+ ALCL samples (with matched peripheral blood) were retrieved from the Sequence Read Archive (SRP044708)[5]. Reads from these and the 18 *de novo* sequenced samples were first analysed using FastQC for quality control. Reads were trimmed to remove nucleotide calls with a Quality score inferior to 30. Reads with a length of less than 50 nt were removed. The Burrows-Wheeler Alignment algorithm ('bwa mem' version 0.7.12-5) was used to align reads to the reference sequence of the human genome (version hg38), which was also indexed using the Burrows-Wheeler alignment tool using default algorithm settings. Aligned files were then sorted and indexed (SAMtools v1.4), after which duplicate reads were removed and reads around known InDel loci were realigned (using Picard's 'MarkDuplicates', 'RealignerTargetCreator' and 'IndelRealigner' v2.5.0), again using default algorithm settings. The Coverage of the ensuing files were computed, and is displayed in Supplementary Table S4. SNVs (Single-Nucleotide Variants) were called using CaVEMan[6] (version 1.9.5) and InDels (Insertions and Deletions) using Pindel[7] (version 0.2.5b8). InDels on all chromosomes were called, and the same reference genome as above was used. A configuration file containing the required data (sample type and insert size) is required, as detailed in the software manual. All files were called in parallel to increase call accuracy. With respect to CaVEMan, samples were called (where possible) against their peripheral blood counterparts to screen out germline variants. CaVEMan requires the same reference genome, along with a reference file containing regions of the human genome to ignore – a file was compiled using all the intergenic regions as determined by the University of California Santa Cruz (USCS) Genomics Institute. Output files of CaVEMan and Pindel were concatenated into a single file.

Variants were annotated using Annovar's[8] 'table_annovar.pl' function (both SNVs and InDels), against the human genome hg38 build as a reference. The variants were annotated against the refGene (build

77), Cosmic (build 78), AVSNP (build 147) and dbSNP (build 148) databases. Variants were also annotated against a number of variant prediction databases: SIFT[9], Polyphen[10], LRT[11], MutationTaster[12], MutationAssessor[13], FATHMM[14], PROVEAN[15], MetaSVM[16], MetaLR[16], CADD[17], dbSNP, 1000Genome (phase 3 release), ExAc and Clinvar. Variants annotated as part of an intronic, intergenic, a UTR or immediately upstream or downstream of a gene were filtered out. Synonymous SNVs were also filtered out. Highly variable genes with low likelihood of pathogenicity were also screened out at this stage, using published literature[18]. Variants contained in any of the 11 matched peripheral blood samples were filtered out from all samples. Variants identified as present in the population at a frequency of >0.1% as determined by either dbSNP or the 1000 Genome project were also filtered out. Finally, variants were screened based on pathogenicity, using annotated variant effect prediction scores. Variants predicted to be damaging by both MetaSVM and MetaLR were retained, while those predicted to be tolerated were filtered out. For variants not annotated by MetaSVM/MetaLR, custom scores combining 8 Variant Effect Prediction software (SIFT, Polyphen, LRT, MutationTaster, MutationAssessor, FATHMM, PROVEAN, CADD) were derived: scores from each software were translated to a 0-1 scale and averaged out, following which variants with scores in the ten most damaging percentiles were retained.

To exclude batch effects, the number of post-filtering somatic variants per patient was analysed (looking at the sequencing cohort to which they pertain (i.e. ALCL99 samples, CCLG tissue bank or the published dataset[5]), so the differences between samples was independent of these variables (Figure S7A). Similarly, the proportion of variant type was independent of the sequencing cohort (Figure S7B).

Copy-number variation was studied using CNVkit 0.94[19].

Gene Set Enrichment Analysis

All genes which were mutated in at least two of the 25 patients were included in Gene Set Enrichment Analysis. Two different software programmes were used; DAVID[20] for protein domains, and

Reactome[21] for pathway analysis. Reactome analysis combined 5 different databases: IPA (Qiagen), PantherDB[22], KEGG[23] and Reactome's own database. Analysis of protein domains using DAVID employed four different databases, GO[24], Seq-Feat (National Center for Biotechnology Information, US), SMART[25] and InterPro[26]. The number of databases in which each pathway or domain was found to be enriched was then displayed along with *p*-values and the size of the network using ggplot2 in R.

ChIP-qPCR

ChIP-qPCR analysis for NOTCH1 and IRF4 was performed on 1 x 10^7 ALCL cells per sample using an anti-STAT3 or anti-GFP antibody (Supplementary Table S5). Following treatment with 1000 nM crizotinib or DMSO for 3 hrs in growth medium, 1 x 10^7 cells were fixed with 0.75% formaldehyde for 15 mins with orbital shaking at room temperature. Subsequently, glycine was added to a final concentration of 125 nM and incubated for 5 mins at room temperature. Next, cells were washed twice with cold PBS, collected by centrifugation and flash frozen in dry ice/isopropanol before storage at -80°C until use. Cell pellets were lysed in 650 µL ChIP lysis buffer (50 mM HEPES-KOH pH7.5, 140 mM NaCL, 1 mM EDTA pH8, 1% Triton X-100, 0.1% Sodium Deoxycholate, 0.1% SDS) supplemented with complete™ Mini EDTA-free Protease Inhibitor Cocktail (Roche, Basel, Switzerland) per 2 x 10^7 cells, followed by sonication for a total of 10 mins with 30 sec pulses on, followed by 30 sec off. Immunoprecipitation reactions were performed overnight with 3 µg STAT3 or GFP antibodies at 4 °C. Next, antibodies and chromatin were captured for 2 hrs at 4 °C using 50 µL of Protein G Dynabeads (Thermo Scientific) per sample. Beads were first washed three times with low salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCL pH 8.0, 150 mM NaCL), followed by three washes with high salt buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl pH 8.0, 500 mM NaCL), two washes with LiCl wash buffer (0.25 M LiCl, 1% NP-40, 1% Sodium Deoxycholate, 1 mM EDTA, 10 mM Tris-HCl pH 8.0) and two final washes with TE buffer (10 mM Tris pH 8.0, 1 mM EDTA). DNA was eluted with 200 µL elution buffer (1% SDS, 100mM $NaHCO_3$), RNA was digested using 2 µL RNase A (10 mg/mL, Roche) at 37 °C for 30 mins, before cross-links were reversed at 65 °C for 2 hrs with 2 µL proteinase K (20 mg/mL,

Thermo Scientific). De-crosslinked DNA was purified with a Zymo DNA Clean and Concentrator-5 kit according to the manufacturer's instructions (Zymo research, Cambridge, UK). ChIP and input DNA were analyzed with SYBR-Green qPCR analysis performed using a QuantStudio™ 6 Flex Real-Time PCR System in accordance with the manufacturer's protocol using qPCR primers as shown in Table S10.

Generation of Crizotinib Resistant ALCL Cell Lines

Crizotinib resistant Karpas-299, SUP-M2, SUDHL-1 and DEL cell lines were established as described previously[27]. Briefly, ALCL cells were seeded at approximately 0.5 x 10^6 cells/ml before Crizotinib was added at a concentration of 50 nM, which was replaced every 48-72 hrs. After every second passage, the concentration of Crizotinib was increased in half-log intervals. The maximum concentrations of Crizotinib reached for Karpas-299, SUP-M2, DEL and SU-DHL1 cell lines were: 0.6 µM, 0.3 µM, 0.2 µM and 0.1 µM respectively.

ChIP-Seq

BED files were downloaded from the GSE archive (accession GSE117164[28] for STAT3 ChIP-seq, GSE104261[29] and GSE29600[30] for NOTCH1 and NOTCH3 respectively). Files were sorted using BEDTools ("sort"), then converted into BEDGraph using BEDTools ('genomecov'), and then into BigWig track files using UCSC's "bedgraphToBigWig". The genome browser tracks were visualized in IGV v2.3.92.

Microarray data analysis

Microarray files were collected from the GSE archive (accession GSE5827[31], GSE104261[29] and GSE29600[30]) and analysed using NCBI's GEO2R online pipeline[32], by creating a group for samples treated with the vehicle control, and another for samples treated with GSI. The top 250 hits (according to adjusted *p*-values) were then extracted from GEO2R; hits present in at least two of the three datasets were retained for display on a heatmap.

Normalized, absolute, microarray expression data of ALK+ ALCL (n=64), ALK- ALCL (n=30) and reactive lymph nodes (n=12) for NOTCH1, MYC and DTX1 was downloaded from the GSE archive (GSE6338[33], GSE14879[34], GSE19069[5], GSE58445[35] and GSE78513[36]). Data were used to correlate mRNA expression of NOTCH1 with MYC, and NOTCH1 with DTX1, calculated by Pearson Correlation (using PRISM GraphPad 8).

Compounds, Cell Lines and Plasmids

The following compounds were used: GSI-1 (Abcam, Cambridge, UK); PF-03084014 (Sigma-Aldrich, St-Louis, Missouri, US), Crizotinib (Sigma-Aldrich), all dissolved in DMSO, and Ionomycin (Sigma-Aldrich) dissolved in water. HEK293FT were cultured in DMEM/10% FBS/1% Pen-strep. Karpas-299, SU-DHL1, SUP-M2 and DEL cell lines were obtained from the DSMZ, Braunschweig, Germany; FEPD were provided by Annarosa Del Mistro, University of Padua, Italy; Mac2A from Olaf Merkel, Medical University Vienna, Austria. These cell lines were cultured in RPMI 1640/ 10% FBS/ 1% Pen-Strep. OP9-DL1 cells (provided by Alison Michie, Glasgow) were cultured in α-MEM/20% FBS. All cell lines were incubated at 37 °C/5% CO_2, were certified mycoplasma free on a quarterly basis and are detailed in Supplementary Table S11. Details of all plasmid vectors used in the study are provided in Supplementary Table S12.

Cellular Proliferation and Apoptosis

Cell proliferation was measured by MTT (Sigma Aldrich) or RealTime-Glo (Promega, Madison, Wisconsin, US) assays according to the manufacturer's instructions. Using a SpectraMax i3 plate reader, absorbance at 570 nm for MTT assays and luminescence for RealTime-Glo were read. Apoptosis was assessed following incubation of 500,000 cells with 4 µL of an APC-conjugated Annexin V antibody (Biolegend, San Diego, California, US) for 45 mins at room temperature and/or Propidium Iodide (1 µg/mL) (Sigma-Aldrich) followed by flow cytometry on a FACSCalibur (BD Bioscience). All flow cytometry data were analysed with FlowJo (FlowJo, LLC). To assess apoptosis, cells were gated to filter out cell debris (FSC/SSC) and to analyse only single cells (SSC-Height/SSC-Area).

Site-Directed Mutagenesis

Plasmids were amplified in a reaction mix consisting of 1X Pfu buffer, 100ng plasmid, 0.5 µM of each primer (Supplementary Table S9), 500 µM dNTPs and 5U Pfu Taq polymerase, supplemented with 2.5% DMSO at 92°C for 30 secs before 16 cycles of 92°C for 30 secs, 55°C for 1 min and 68°C for 25 mins. The parental plasmids were digested by incubation for one hour at 37°C with 10U DpnI. The product (2 µL) was then transformed into XL-10 Gold competent bacteria (Agilent, Santa Clara, California, USA) before plasmid purification and Sanger sequencing to verify the presence of the desired variant.

Lentiviral Production, Transductions and shRNA Silencing

HEK293FT cells were seeded at 50% confluency 1 day before transfection in a T25 flask with 2.7 µg of the plasmid of interest, 1.5 µg pMD2.G (Addgene), 2.4 µg psPAX2 (Addgene) and 19.2 µL TransIT-293 (MirusBio, Madison, Wisconsin, USA) pre-mixed in Opti-MEM (Thermo Scientific, Waltham, Massachusetts, USA). Supernatant was collected 54 hrs later and overlaid onto the cells to be transduced, following which, after 24 hrs incubation, the appropriate antibiotic was added for 7 days (Supplementary Table S12). With respect to shRNA silencing, ALCL cell lines were transduced with shRNA constructs (Supplementary Table S12) and cells were then selected using the relevant antibiotic. Following 96 hrs of selection, RNA was extracted to verify gene silencing, further to which antibiotic selection was terminated and cells were cultured for downstream applications. The shRNA targeting STAT3 has been described and characterized previously[28]. A SU-DHL1 cell line and SUP-M2-derived cell line (also called 'TS') expressing a doxycycline-inducible NPM-ALK-targeting shRNA construct have been described and characterized previously[37].

Method References

1. Shi J, Blundell TL, Mizuguchi K. FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J. Mol. Biol. 2001;310(1):243–57.

2. Pandurangan AP, Ochoa-Montaño B, Ascher DB, Blundell TL. SDM: a server for predicting effects of mutations on protein stability. Nucleic Acids Res. 2017;45(W1):229–235.

3. Rodrigues CH, Pires DE, Ascher DB. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. Nucleic Acids Res. 2018;46(W1):350–355.

4. Pires DE V, Ascher DB, Blundell TL. mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics. 2014;30(3):335–42.

5. Crescenzo R, Abate F, Lasorsa E, et al. Convergent mutations and kinase fusions lead to oncogenic STAT3 activation in anaplastic large cell lymphoma. Cancer Cell. 2015;27(4):516–532.

6. Turinsky AL, Fanea E, Trinh Q, et al. CAVEman: Standardized anatomical context for biomedical data mapping. Anat. Sci. Educ. 2008;1(1):10–18.

7. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 2009;25(21):2865–2871.

8. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):164.

9. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. Nat. Protoc. 2016;11(1):1–9.

10. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res. 2002;30(17):3894–900.

11. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19(9):1553–1561.

12. Stenson PD, Mort M, Ball E V, et al. The Human Gene Mutation Database: 2008 update. Genome Med. 2009;1(1):13.

13. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 2011;39(17):118.

14. Shihab HA, Gough J, Cooper DN, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum. Mutat. 2013;34(1):57–65.

15. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLoS One. 2012;7(10):e46688.

16. Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. Hum. Mol. Genet. 2015;24(8):2125–37.

17. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. 2014;46(3):310–5.

18. Shyr C, Tarailo-Graovac M, Gottlieb M, et al. FLAGS, frequently mutated genes in public exomes. BMC Med. Genomics. 2014;7(1):64.

19. Talevich E, Shain AH, Botton T, Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLOS Comput. Biol. 2016;12(4):e1004873.

20. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 2009;4(1):44–57.

21. Fabregat A, Jupe S, Matthews L, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2018;46(D1):D649–D655.

22. Mi H, Huang X, Muruganujan A, et al. PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res. 2017;45(D1):183–189.

23. Goto S, Bono H, Ogata H, et al. Organizing and computing metabolic pathway data in terms of binary relations. Pacific Symp. Biocomput. 1997;175–86.

24. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. Nucleic Acids Res. 2019;47(D1):330–338.

25. Letunic I, Bork P. 20 years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46(1):493–496.

26. Mitchell AL, Attwood TK, Babbitt PC, et al. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic Acids Res. 2019;47(D1):351–360.

27. Ceccon M, Mologni L, Bisson W, Scapozza L, Gambacorti-Passerini C. Crizotinib-Resistant NPM-ALK Mutants Confer Differential Sensitivity to Unrelated Alk Inhibitors. Mol. Cancer Res. 2013;11(2):122–132.

28. Menotti M, Ambrogio C, Cheong T-C, et al. Wiskott-Aldrich syndrome protein (WASP) is a tumor suppressor in T cell lymphoma. Nat. Med. 2019;25(1):130–140.

29. Choi SH, Severson E, Pear WS, et al. The common oncogenomic program of NOTCH1 and NOTCH3 signaling in T-cell acute lymphoblastic leukemia. PLoS One. 2017;12(10):e0185762.

30. Wang H, Zou J, Zhao B, et al. Genome-wide analysis reveals conserved and divergent features of Notch1/RBPJ binding in human and murine T-lymphoblastic leukemia cells. Proc. Natl. Acad. Sci. 2011;108(36):14908–13.

31. Sanchez-Martin M, Ambesi-Impiombato A, Qin Y, et al. Synergistic antileukemic therapies in NOTCH1-induced T-ALL. Proc. Natl. Acad. Sci. 2017;114(8):2006–2011.

32. Wilhite SE, Barrett T. Strategies to Explore Functional Genomics Data Sets in NCBI's GEO Database. Methods Mol. Biol. 2012;802:41–53.

33. Piccaluga PP, Agostinelli C, Califano A, et al. Gene expression analysis of peripheral T cell lymphoma, unspecified, reveals distinct profiles and new potential therapeutic targets. J. Clin. Invest. 2007;117(3):823–34.

34. Eckerle S, Brune V, Döring C, et al. Gene expression profiling of isolated tumour cells from anaplastic large cell lymphomas: insights into its cellular origin, pathogenesis and relation to Hodgkin lymphoma. Leukaemia. 2009;23(11):2129–38.

35. Iqbal J, Wright G, Wang C, et al. Gene expression signatures delineate biological and prognostic subgroups in peripheral T-cell lymphoma. Blood. 2014;123(19):2915–23.

36. Pomari E, Basso G, Bresolin S, et al. NPM-ALK expression levels identify two distinct subtypes of paediatric anaplastic large cell lymphoma. Leukemia. 2017;31(2):498–501.

37. Piva R, Chiarle R, Manazza AD, et al. Ablation of oncogenic ALK is a viable therapeutic approach for anaplastic large-cell lymphomas. Blood. 2006;107(2):689–97.

38. DuBridge RB, Tang P, Hsia HC, et al. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell. Biol. 1987;7(1):379–87.

39. Fischer P, Nacheva E, Mason DY, et al. A Ki-1 (CD30)-positive human cell line (Karpas 299) established from a high-grade non-Hodgkin's lymphoma, showing a 2;5 translocation and rearrangement of the T-cell receptor beta-chain gene. Blood. 1988;72(1):234–40.

40. Epstein AL, Levy R, Kim H, et al. Biology of the human malignant lymphomas. IV. Functional characterization of ten diffuse histiocytic lymphoma cell lines. Cancer. 1978;42(5):2379–91.

41. Dutil J, Chen Z, Monteiro AN, Teer JK, Eschrich SA. An Interactive Resource to Probe Genetic Diversity and Estimated Ancestry in Cancer Cell Lines. Cancer Res. 2019;79(7):1263–1273.

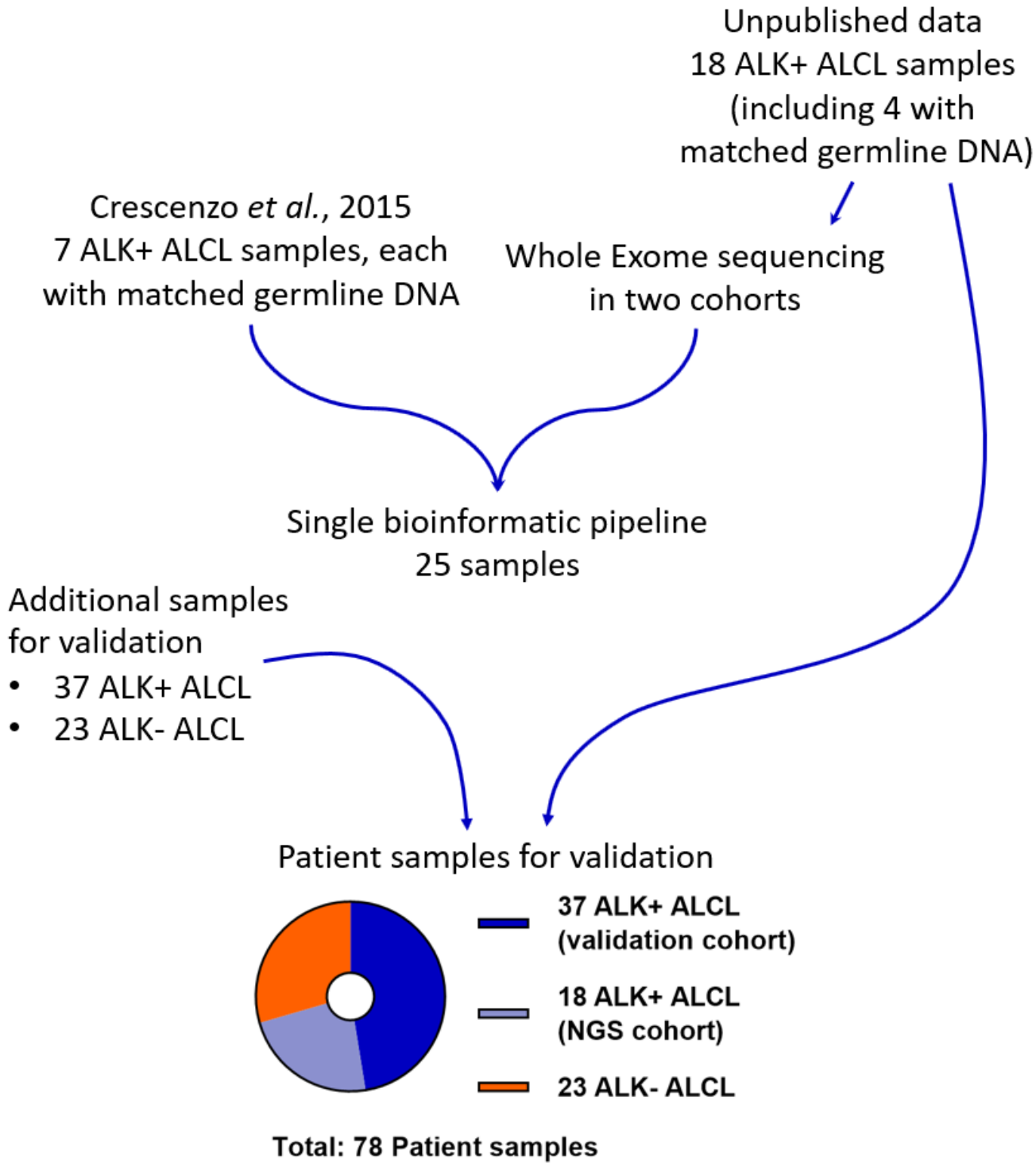
42. Barbey S, Gogusev J, Mouly H, et al. DEL cell line: a "malignant histiocytosis" CD30+ t(5;6)(q35;p21) cell line. Int. J. cancer. 1990;45(3):546–53.

43. Davis TH, Morton CC, Miller-Cassman R, Balk SP, Kadin ME. Hodgkin's disease, lymphomatoid papulosis, and cutaneous T-cell lymphoma derived from a common T-cell clone. N. Engl. J. Med. 1992;326(17):1115–22.

44. del Mistro A, Leszl A, Bertorelle R, et al. A CD30-positive T cell line established from an aggressive anaplastic large cell lymphoma, originally diagnosed as Hodgkin's disease. Leukemia. 1994;8(7):1214–9.

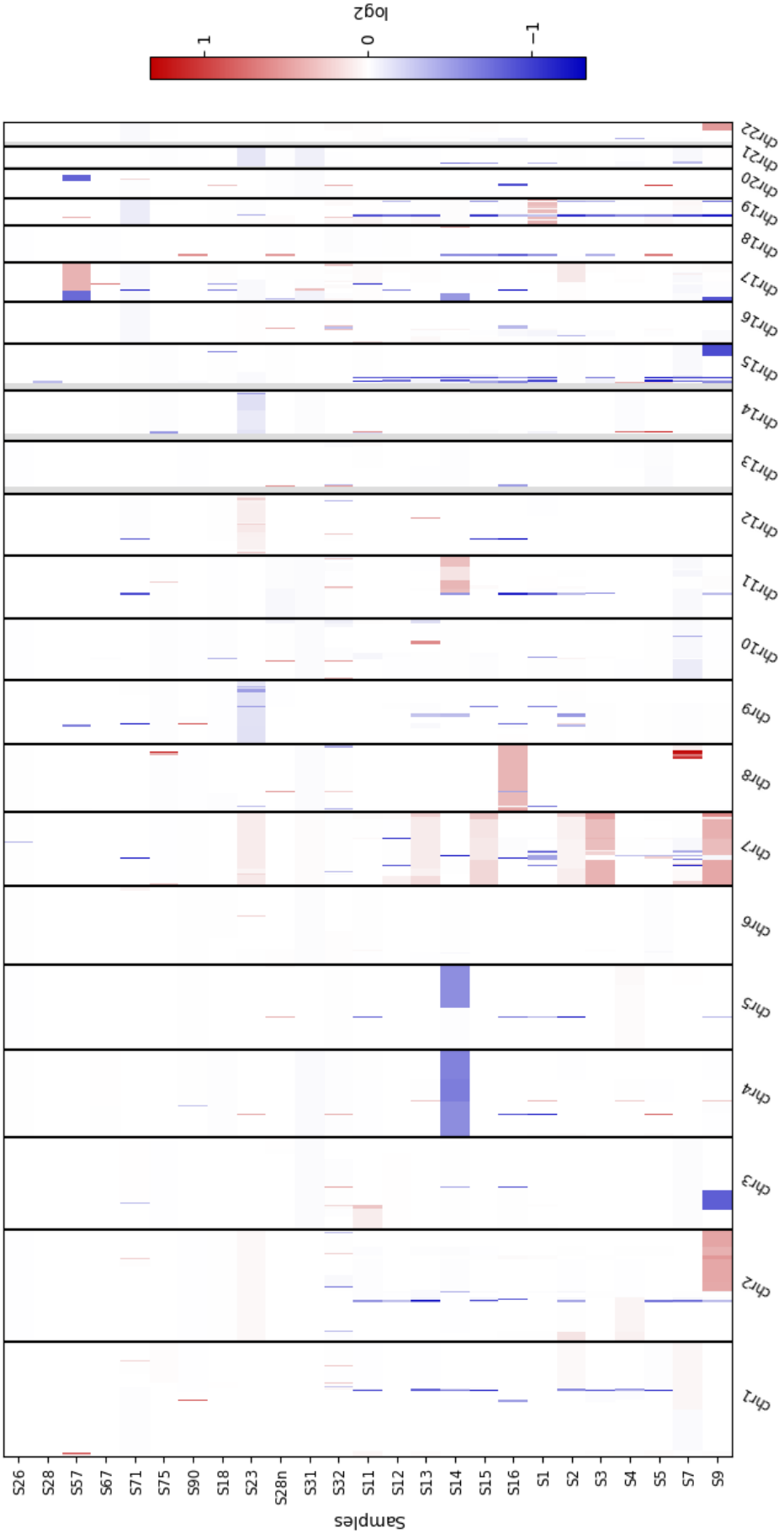
45. Nakano T, Kodama H, Honjo T. Generation of lymphohematopoietic cells from embryonic stem cells in culture. Science. 1994;265(5175):1098–101.

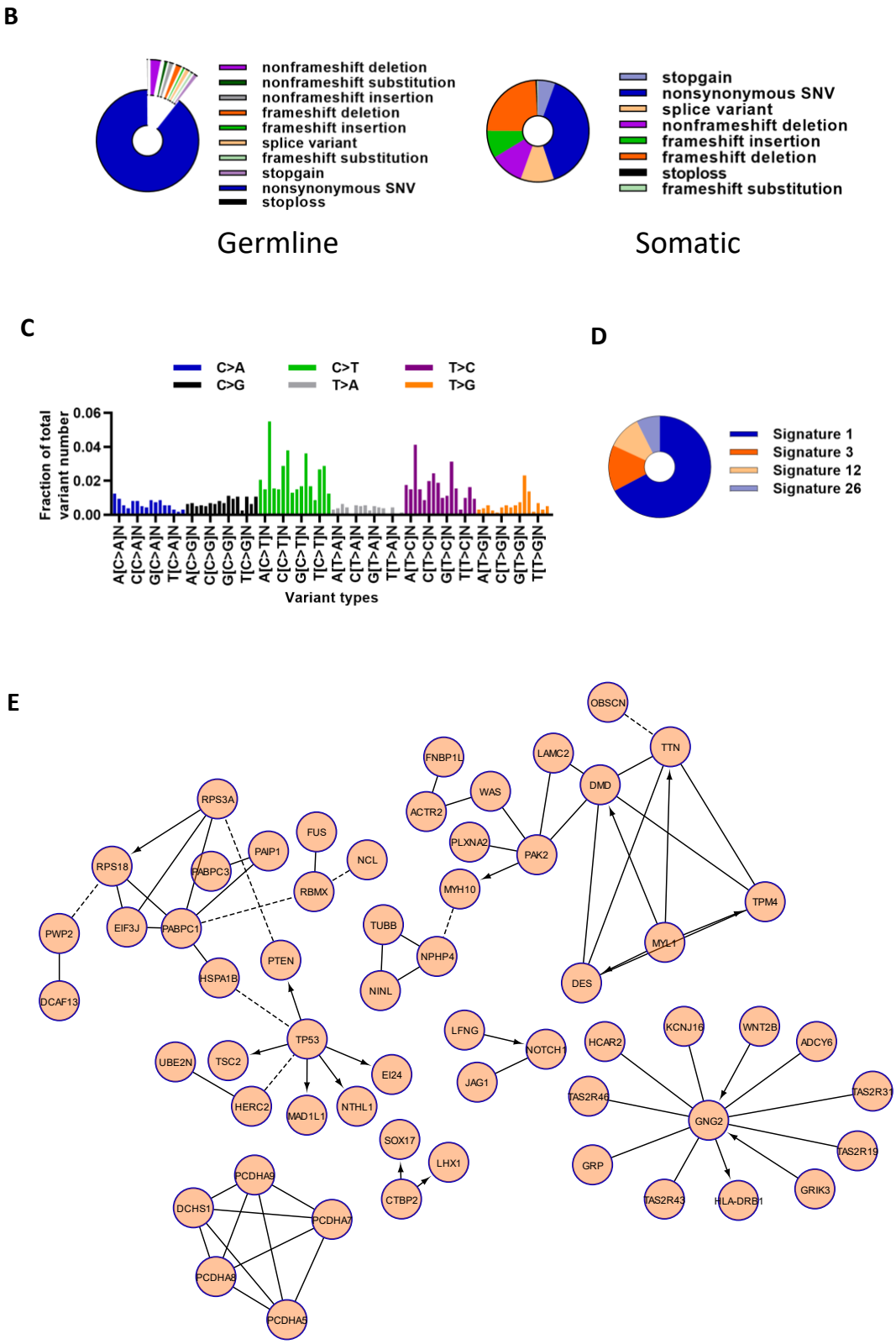
Supplemental Figures



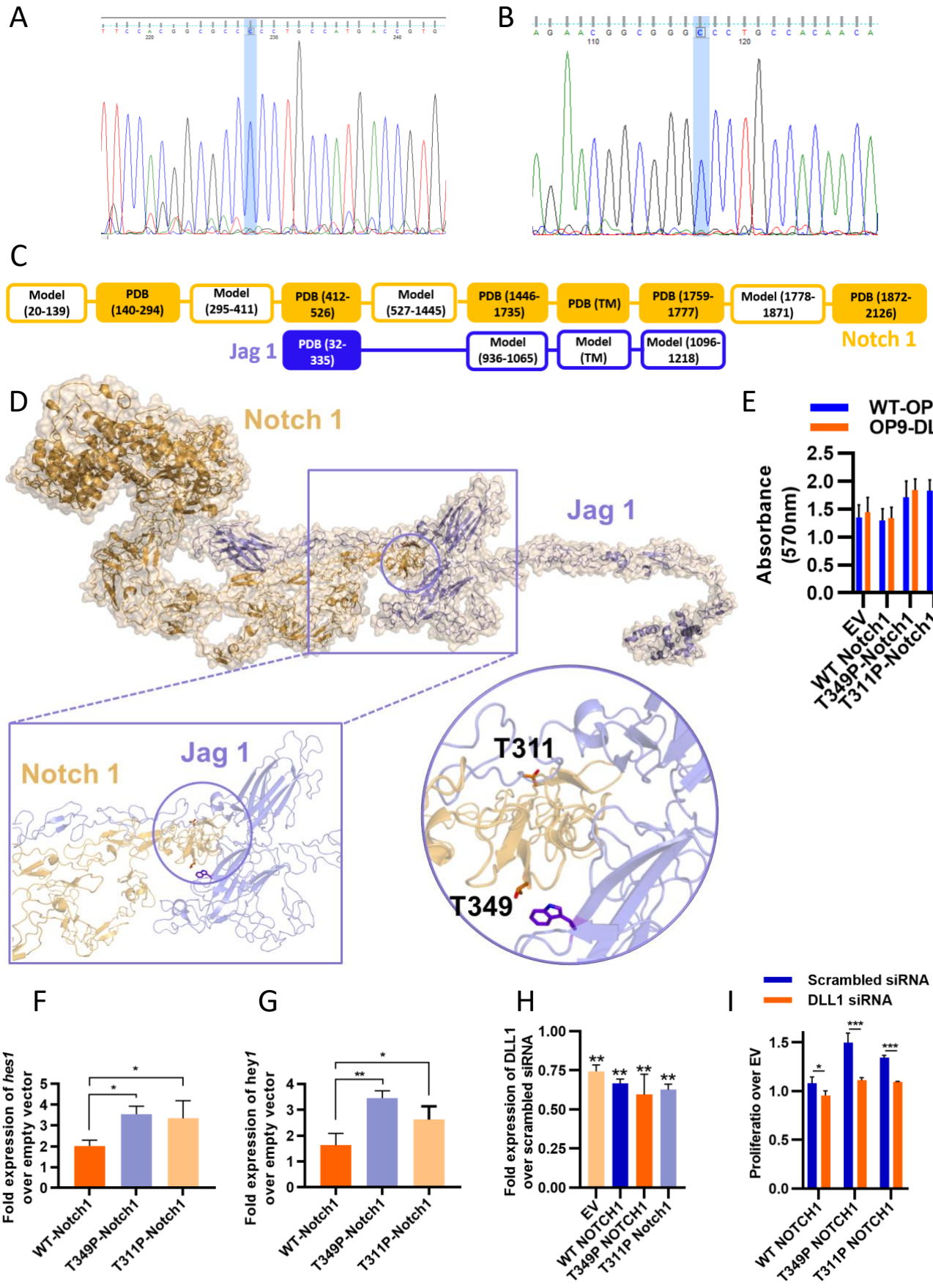
Supplementary Figure 1. Schematic representation of patient cohorts for which whole exome sequencing data were generated and analysed.

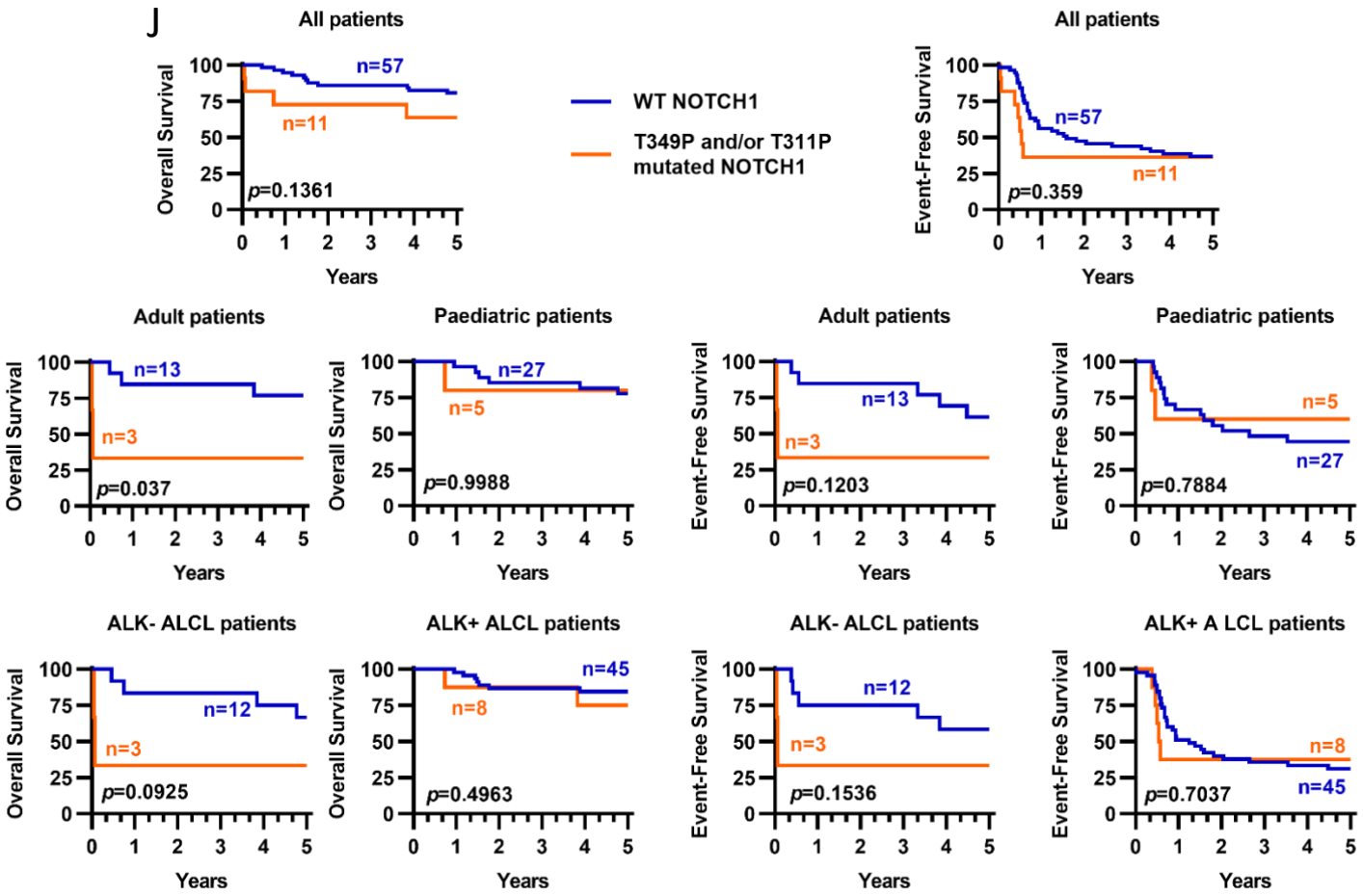
A


log2
1
0
−1
chr22
chr21
chr20
chr19
chr18
chr17
chr16
chr15
chr14
chr13
chr12
chr11
chr10
chr9
chr8
chr7
chr6
chr5
chr4
chr3
chr2
chr1
S26
S28
S57
S67
S71
S75
S90
S18
S23
S28n
S31
S32
S11
S12
S13
S14
S15
S16
S1
S2
S3
S4
S5
S7
S9
Samples

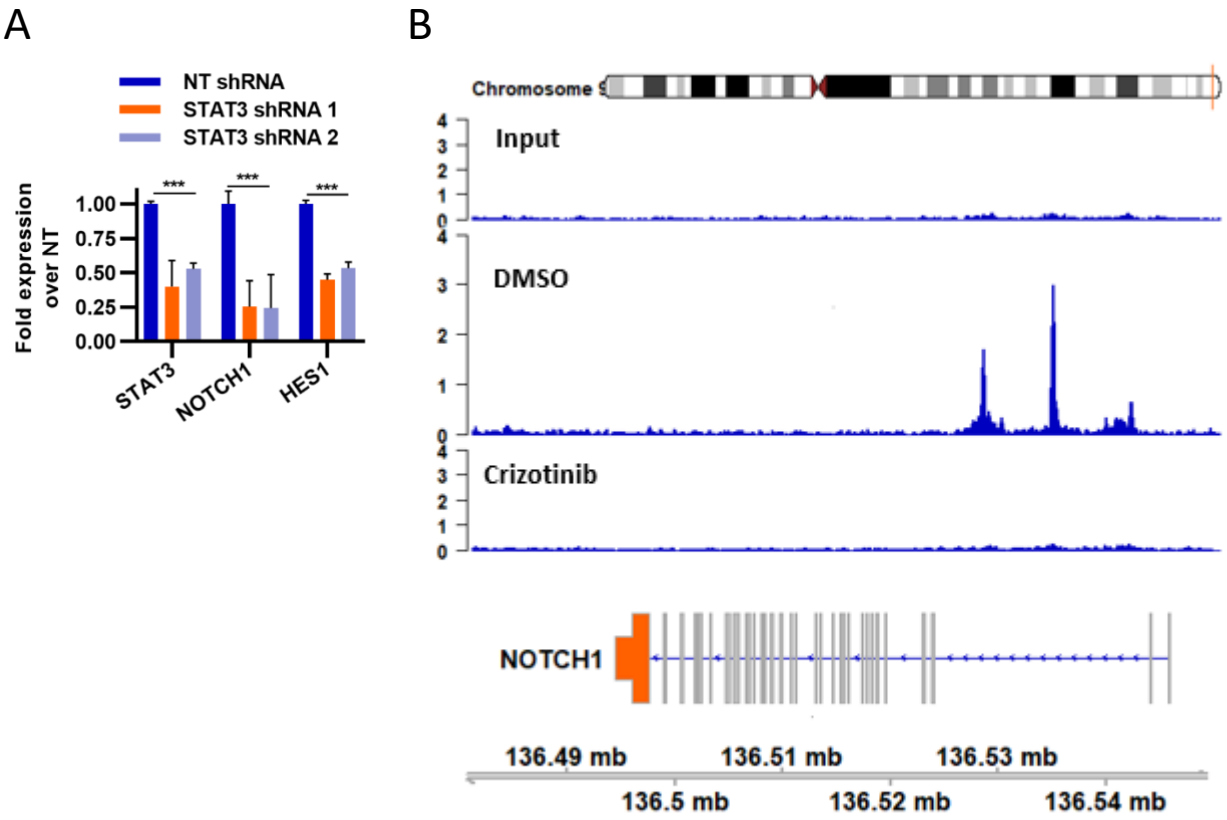


Supplementary Figure 2. Further details regarding WES bioinformatic analysis. (A) Autosome-wide heatmap of log2 copy number segments in patient samples. **(B)** Distribution of variant type, averaged for germline samples (n=11) and somatic samples (n=25). **(C)** Prevalence of each of the 96 variant types for representative patient sample S57; these were then used to derive the mutational signature, as displayed for patient sample S57 **(D). (E)** Nodular interaction plot showing connections between variants. The plot was designed using Reactome. Arrows indicate activating interactions, dotted lines indicate hypothesised interactions, lines ending with a perpendicular bar indicate inhibition.

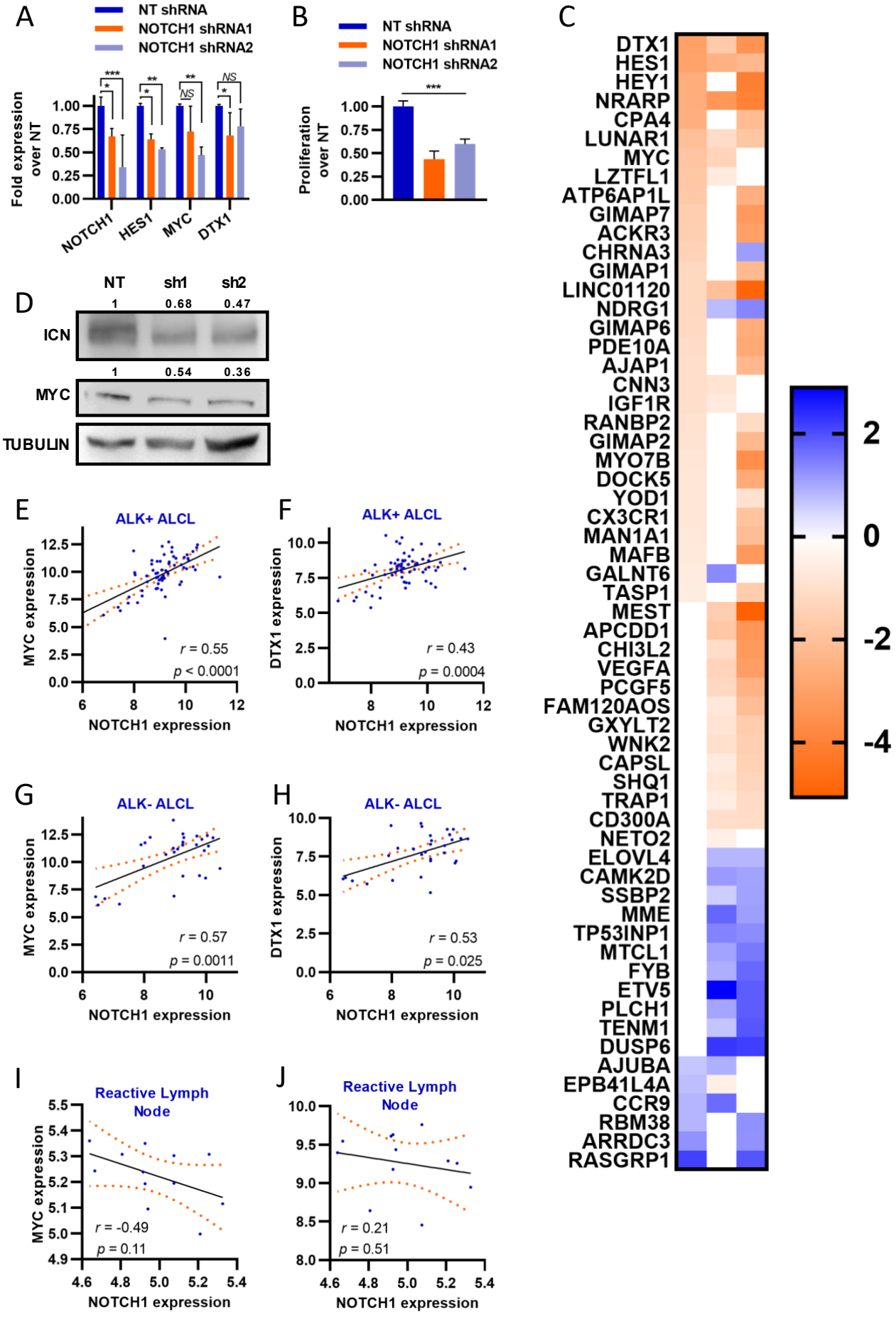

A
B
C
Model (20-139)
PDB (140-294)
Model (295-411)
PDB (412-526)
Model (527-1445)
PDB (1446-1735)
PDB (TM)
PDB (1759-1777)
Model (1778-1871)
PDB (1872-2126)
Notch 1
Jag 1
PDB (32-335)
Model (936-1065)
Model (TM)
Model (1096-1218)
D
Notch 1
Jag 1
T311
T349
E
WT-OP9
OP9-DLL1
Absorbance (570nm)
EV
WT Notch1
T349P-Notch1
T311P-Notch1
F
Fold expression of hes1 over empty vector
WT-Notch1
T349P-Notch1
T311P-Notch1
G
Fold expression of hey1 over empty vector
H
Fold expression of DLL1 over scrambled siRNA
EV
WT NOTCH1
T349P NOTCH1
T311P Notch1
I
Scrambled siRNA
DLL1 siRNA
Proliferatio over EV
WT NOTCH1
T349P NOTCH1
T311P NOTCH1



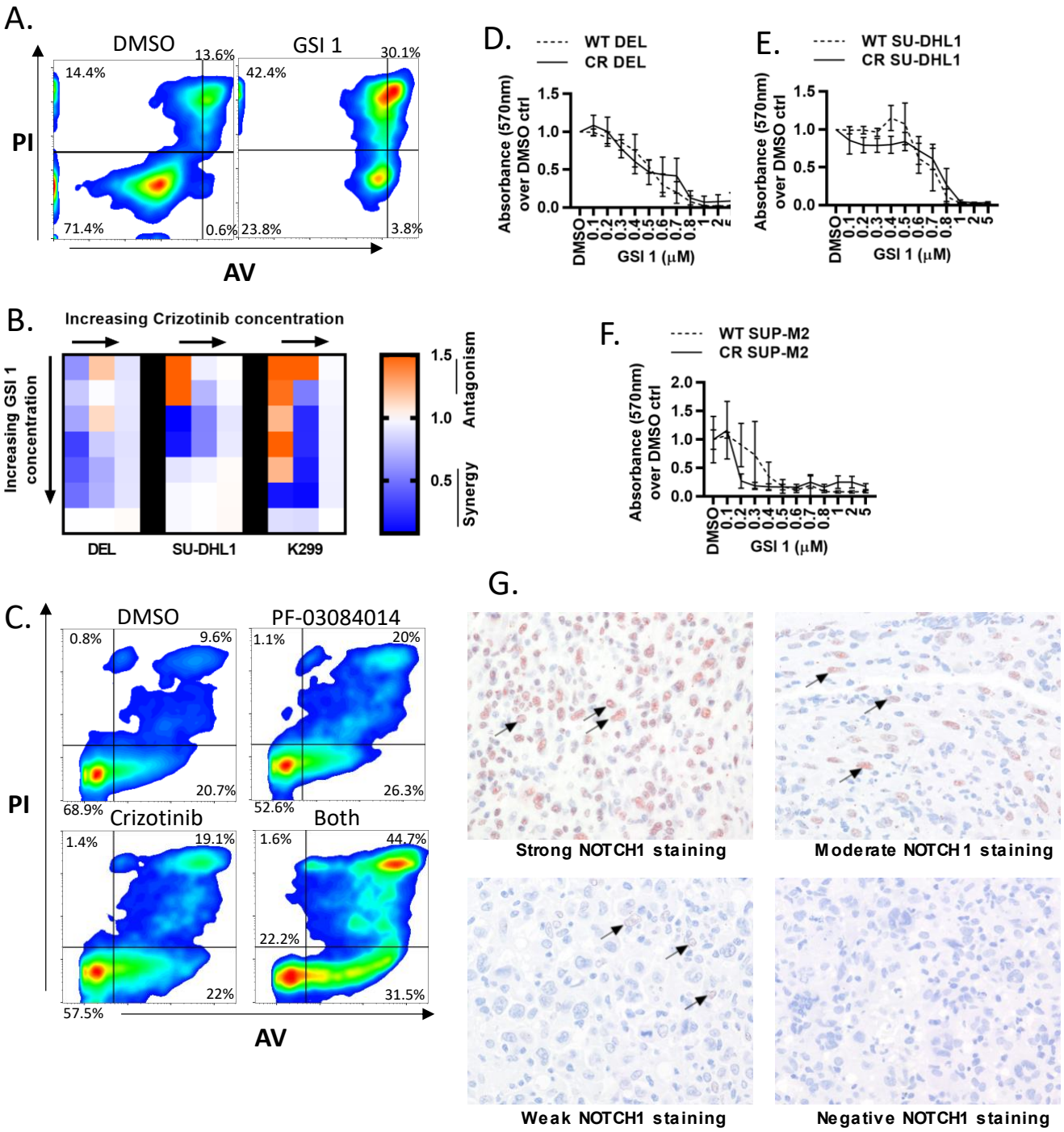
Supplementary Figure 3. Modelling of the NOTCH1-JAG1 interaction. Sanger chromatogram of a selected sequence of NOTCH1 showing mutations T311P **(A)** and T349P **(B)**. **(C)** Schematic representation of the domains of NOTCH1 and JAG1, detailing which domains were modelled from scratch ('Model'), and which were solved structures available on the PDB ('PDB') (PDB IDs: 4XL1, 4XBM, 5MWB, 4D90, 1YMP, 1OT8). **(D)** Interaction of full-length NOTCH1 (orange) bound to its receptor Jag1 (purple), also modelled in full-length. The interaction between JAG1 and NOTCH1 is enlarged for easier visualisation, with an additional magnification to clearly show NOTCH1 amino acids T311 and T349. **(E)** MTT assay of HEK293 cells transduced with an Empty Vector (EV), or the Wild-type (WT), T349P or T311P Notch1 mutants, co-cultured with wild-type or DLL1-expressing OP9 cells. Fold-change expression of HES1 **(F)** or HEY1 **(G)** over empty-vector control as assessed by qPCR in HEK293FT cells expressing the wild-type or mutant NOTCH1 proteins (normalized to GAPDH and PPIA; **p*<0.05; ***p*<0.01; n=3). HEK293 cells expressing either an EV, WT or mutated NOTCH1 were transfected with siRNA to DLL1 or a control, scrambled siRNA **(H)** fold change expression of DLL1 over scrambled siRNA was assessed by qPCR (normalized to GAPDH and PPIA (**p<0.01; n=3); and **(I)** fold change proliferation over EV was assessed by MTT assay (**p*<0.05; ***p*<0.01; ***p<0.001; n=3). **(J)** Kaplan-Meier Overall and Event-Free Survival plots of patients in our validation cohort for whom we hold at least 5-years of follow-up clinical data, comparing patients that are wild-type (WT), T349P and/or T311P-mutated for NOTCH1, comparing all patients; paediatric patients (18 years and under), adult patients (19 years and above), ALK+ ALCL and ALK- ALCL patients. *p*-value determined using the log-rank test.



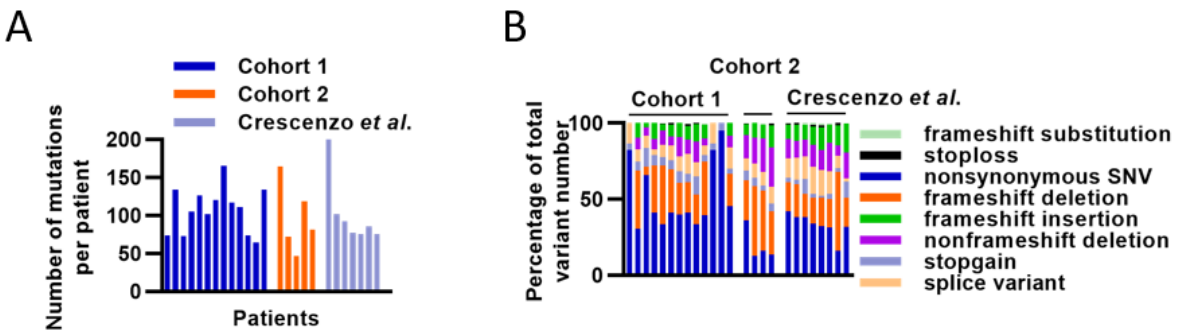
Supplementary Figure 4. ChIP-seq showing binding of STAT3 to the NOTCH1 promoter region. (A) Fold-change expression of the indicated genes over non-targeting control shRNA transduced cells (normalized to GAPDH) in the FEPD cell line 48 hours after transduction with control non-targeting (NT) shRNA, or one of two shRNAs targeting STAT3 as determined by qPCR (***p<0.001; n=3). All bar plots display the mean of biological replicates and error bars represent standard deviations; the bar plots are colour-coded as indicated in the Figure. **(B)** Binding of STAT3 to promoter regions of NOTCH1 in JB6 cells treated with a vehicle control (middle track) or Crizotinib (lower track); the upper track shows the input, data obtained by analysing previously published data[28].


A
NT shRNA
NOTCH1 shRNA1
NOTCH1 shRNA2
Fold expression over NT
NOTCH1
HES1
MYC
DTX1
B
Proliferation over NT
C
DTX1
HES1
HEY1
NRARP
CPA4
LUNAR1
MYC
LZTFL1
ATP6AP1L
GIMAP7
ACKR3
CHRNA3
GIMAP1
LINC01120
NDRG1
GIMAP6
PDE10A
AJAP1
CNN3
IGF1R
RANBP2
GIMAP2
MYO7B
DOCK5
YOD1
CX3CR1
MAN1A1
MAFB
GALNT6
TASP1
MEST
APCDD1
CHI3L2
VEGFA
PCGF5
FAM120AOS
GXYLT2
WNK2
CAPSL
SHQ1
TRAP1
CD300A
NETO2
ELOVL4
CAMK2D
SSBP2
MME
TP53INP1
MTCL1
FYB
ETV5
PLCH1
TENM1
DUSP6
AJUBA
EPB41L4A
CCR9
RBM38
ARRDC3
RASGRP1
2
0
-2
-4
D
NT
sh1
sh2
ICN
MYC
TUBULIN
E
ALK+ ALCL
MYC expression
NOTCH1 expression
r = 0.55
p < 0.0001
F
ALK+ ALCL
DTX1 expression
NOTCH1 expression
r = 0.43
p = 0.0004
G
ALK- ALCL
MYC expression
NOTCH1 expression
r = 0.57
p = 0.0011
H
ALK- ALCL
DTX1 expression
NOTCH1 expression
r = 0.53
p = 0.025
I
Reactive Lymph Node
MYC expression
NOTCH1 expression
r = -0.49
p = 0.11
J
Reactive Lymph Node
NOTCH1 expression
r = 0.21
p = 0.51

Supplementary Figure 5. Changes in gene expression upon GSI treatment in T-ALL. (A) Fold-change expression of the indicated genes over non-targeting control shRNA transduced cells (normalized to GAPDH) in the ALK- ALCL cell line FEPD, 48 hours after transduction with control non-targeting (NT) shRNA, or one of two shRNA targeting NOTCH1 as determined by qPCR (NS: Not Significant; **p*<0.05; ***p*<0.01; ****p*<0.001; n=3). All bar plots display the mean of biological replicates and error bars represent standard deviations; the bar plots are colour-coded as indicated in the Figure. **(B)** Proliferation of the ALK- ALCL cell line FEPD over the non-targeting shRNA control, determined using an MTT assay, 48 hours after transduction with control non-targeting (NT) shRNA, or one of two shRNAs targeting NOTCH1 (****p*<0.001; n=3). **(C)** Analysis of microarray data from three separate publications of the T-ALL cell line CUTLL-1, representing the fold-change in expression of the top 250 target genes (assessed by adjusted *p*-values) in the presence of GSI (over a vehicle control); analysing previously published data[29–31]. **(D)** Western Blot for MYC, intracellular NOTCH1 (ICN) and a loading control (TUBULIN), in DEL cell lines transfected either with non-Targeting (NT) or NOTCH1-targeting (sh1, sh2) shRNA-expression constructs. Only the relevant sections of the whole blot are displayed, the contrast was modified on the whole image to improve legibility. Pearson correlation of MYC or DTX1 and NOTCH1 mRNA expression in ALK+ ALCL (**E, F**; n=64), ALK- ALCL (**G, H**; n=30) and Reactive Lymph Nodes (**I, J**; n=12), including the *r*-score, *p*-value, linear regression line (in black) and 95% confidence interval (in orange), using published microarray data.



Supplementary Figure 6. Proliferation of cell lines when treating with GSI 1. (A) Representative FACS plots for cells stained for Annexin V and PI when treated either with vehicle control or 1 μM GSI-I for 48 hrs. **(B)** BLISS matrix showing the combination index on treating the indicated ALK+ ALCL cell lines with Crizotinib and GSI-1 for 72 hrs (using a range of concentrations of 25-100 nM for Crizotinib, and 100 nM-1 μM GSI-1). A combination index of <1 indicates synergy between drugs, 1 indicates additive effects, >1 indicates antagonistic effects (n=3). **(C)** Representative FACS plots of cells stained for Annexin V and PI when treated either with vehicle control, 40 nM Crizotinib, 2 μM PF-03084014 or a combination of PF-03084014 and Crizotinib for 48 hours. Proliferation over vehicle control of wild-type and Crizotinib-resistant DEL **(D)**, SU-DHL1 **(E)** and SUP-M2 **(F)** cells when treated with increasing concentrations of GSI 1, as measured by MTT assay (*n*=3; CR= Crizotinib Resistant, WT=Wild-type). **(G)** ALCL patient tissue, taken at presentation, were stained for cleaved NOTCH1 and classified according to expression levels: strong, moderate, weak or negative cleaved NOTCH1 staining for which representative examples are shown.



Supplemental Figure 7. Quality Controls for bioinformatic processing. **(A)** Number of mutations per patient, colour-coded to reflect the three sequencing cohorts. **(B)** Distribution of variant types for each patient, separating out the three sequencing cohorts: the samples obtained from an online repository ('Crescenzo *et al.*'), and the two cohorts of patient samples sequenced for this publication (see Figure S1).

Characteristic	Subgroup	Number	Fraction of total
Gender	Male	8	32%
	Female	10	40%
	Unknown	7	28%
Age at diagnosis	0-15	16	64%
	16-25	3	12%
	26-40	1	4%
	>40	5	20%
ALK Status	+	25	100%
	-	0	0%
Matched blood	Yes	11	44%
	No	14	56%
Follow-up known?	5-year EFS known	18	72%
	5-year OS known	17	68%
	No	7	28%
Source	CCLG	18	72%
	Crescenzo et al; SRP044708	7	28%

Supplementary Table 1: Characteristics of the patient tumour samples obtained from the Children's Cancer and Leukaemia Group (CCLG) tissue bank for which WES was conducted or whose WES data was downloaded from the Sequence Read Archive (SRP044708).

Characteristic	Subgroup	Number	Fraction of total
Gender	Male	12	15.4%
	Female	15	19.2%
	Unknown	51	65.4%
Age at diagnosis	0-15	26	33.3%
	16-25	8	10.3%
	26-40	3	3.8%
	>40	24	30.8%
	Unknown	17	21.8%
ALK status	+	55	70.5%
	-	23	29.5%
Matched blood	Yes	12	15.4%
	No	66	84.6%
Follow-up known?	5-year EFS known	61	78.2%
	5-year OS known	46	59.0%
	No	17	21.8%
Source	UK	30	38.5%
	France	19	24.4%
	Czech Republic	10	12.8%
	Giessen	6	7.7%
	Ukraine	8	10.3%
	Graz	5	6.4%

Supplementary Table S2: Characteristics of the 78 patient samples that were employed to validate the presence of NOTCH1 variants in a larger patient population.

Characteristic	Subgroup	Number	Fraction of total
Gender	**Male**	65	73.0%
	Female	24	27.0%
Age at diagnosis	**<10**	33	37.1%
	11+	56	62.9%
ALK Status	**+**	89	100.0%
Central nervous system involvement	**Yes**	1	1.1%
	No	84	94.4%
	Unknown	4	4.5%
Bone marrow involvement	**Yes**	6	6.7%
	No	83	93.3%
Clinical Study	**NHL-BFM 90**	25	28.1%
	NHL-BFM 95	23	25.8%
	ALCL99	41	46.1%
Staging at diagnosis	**Stage I**	9	10.1%
	Stage II	17	19.1%
	Stage III	58	65.2%
	Stage IV	5	5.6%
Events	**Relapse/Progress**	29	32.6%
	Toxic Death	3	3.4%
	Complete Cytogenic Response	19	21.3%
	Lost to Follow-up in Complete Cytogenic Response	38	42.7%

Supplementary Table S3: Characteristics of the 89 patient tissues that were stained for NOTCH1 expression on the tissue microarray, from which clinical data allowed computation of the 10-year EFS.

Source	Library Prep Kit	Sequencing machine	Matched blood available	Sample ID	Coverage (x read depth)
SRP044708	SureSelect 50 Mb All Exon kit	Illumina HiSeq 2000	Yes	S26	187.02
				S28	156.00
				S57	228.53
				S67	223.40
				S71	241.77
				S75	210.48
				S90	92.04
CCLG tissue bank cohort 1	Nextera Rapid Capture Exome	Illumina HiSeq 2500	No	S18	284.79
			Yes	S23	163.38
			Yes	S28n	252.91
			Yes	S31	172.06
			Yes	S32	187.01
CCLG – tissue bank cohort 2 (all from ALCL99 clinical trial)	Nextera Rapid Capture Exome	Illumina MiSeq	No	S1	176.72
				S2	57.18
				S3	156.00
				S4	59.72
				S5	149.94
				S7	152.39
				S9	73.57
				S11	87.26
				S12	195.30
				S13	82.51
				S14	176.72
				S15	44.26
				S16	176.72

Supplementary Table S4: WES details for the three different patient sample cohorts (all fresh frozen tissues and all ALK+ ALCL), including sample names and coverage detail. CCLG = Childrens' Cancer and Leukaemia Group.

Name	Source	Application	Quantity used
PE-conjugated DLL-1	ThermoFisher Scientific, Cat# 12-5767-80	FACS	0.2 µg
(Mouse) NOTCH1 (intracellular)	ThermoFisher Scientific, Cat# 14-5785-81	Western blot	1:500
(Mouse) NOTCH1 (whole-length)	Sigma-Aldrich, Cat# N6786	Western blot	1:500
(Mouse) α-tubulin	Sigma-Aldrich, Cat# T9062	Western blot	1:2000
HRP anti-rabbit IgG	CiteAb, Cat# P0161	Western blot	1:10000
HRP anti-mouse IgG	CiteAb, Cat# P0448	Western blot	1:10000
(Rabbit) phospho-ALK (Tyr1278)	Cell Signaling Technology, Cat# 6941S	Western blot	1:1000
(Rabbit) ALK	Cell Signaling Technology, Cat# 3633S	Western blot	1:1000
(Rabbit) STAT3	Cell Signaling Technology, Cat# 4904SS	Western blot	1:1000
STAT3	Cell Signaling Technology, Cat# 9139	ChIP	3 µg
GFP	Abcam, Cat# ab290	ChIP	3 µg
Human NOTCH1 (intracellular)	ThermoFisher Scientific, Cat# 14-5785-81	IF	1:80

Supplementary Table S5: Detailed list of antibodies used in this study.

Gene	# Patients	Position	Exon	Ntde change	AA change	Type of mutation	COSMIC ID	dbSNP ID	Transcript ID
TYW1B	23	chr7:72728896	exon 9	G1118A	W373X	stopgain		rs3015858	NM_001145440
DEFB132	20	chr20:257795	exon 1	17-22 del	6-8 del	nonframeshift deletion	COSM1163662	rs37182593 8	NM_207469
KCNJ18	20	chr17:21703303	exon 3	G517A	D173N	nonsynonymous SNV			NM_001194958
MIR1-1HG	18	chr20:62565060	exon 3	T80C	V27A	nonsynonymous SNV	COSM3758712	rs6062251	NM_178463
ZNF283	17	chr19:43847015	whole gene	del GGAGAT		frameshift deletion	COSM1394324	rs71907168	NM_001297752
NRDC	17	chr1:51840392	exon 2	462-464 del	154-155 del	nonframeshift deletion	COSM1237693	rs35723519	NM_001101662
ZNF720	16	chr16:31759375	exon 5	381 ins A		frameshift insertion		rs34487972	NM_001130913
PYGL	16	chr14:50911873		del T		splicing variant		rs11356035	
MS4A14	15	chr11:60397880	exon 2	167-168 del		frameshift deletion	COSM1684267	rs3217518	NM_001079692
CLECL1	15	chr12:9733111	exon 1	153 ins ACTTA		frameshift insertion		rs11322262 1	NM_001253750
MYO15B	14	chr17:75601463	exon 14	A3437G	K1146R	nonsynonymous SNV	COSM4130628	rs11871553	NM_001309242
RIC8A	13	chr11:209895	exon 3	621-623 del	207-208 del	nonframeshift deletion	COSM1317342	rs3832797	NM_001286134
STK31	13	chr7:23717543	exon 4	G144C	Q48H	nonsynonymous SNV	COSM3762594	rs6945306	NM_001260504
TAAR9	13	chr6:132538470	exon 1	A181T	K61X	stopgain		rs2842899	NM_175057
CCDC129	12	chr7:31658299	exon 15	3097 ins T		frameshift insertion		rs35589779	NM_001257968
IRF5	12	chr7:128947298	exon 6	502-531 del	168-177 del	nonframeshift deletion	COSM5002496	rs19950896 4	NM_001098627
CYP3A5	12	chr7:99672916		T>C		splicing variant		rs776746	

ZNF3	11	chr7:100064889	exon 6	292-295 del		frameshift deletion	COSM5001700	rs56833874	NM_017715
ZNF219	11	chr14:21092594	exon 3	698-703 del	233-235del	nonframeshift deletion	COSM248523	rs11278664	NM_001101672
PSPH	11	chr7:56019672	exon 5	T203C	L68P	nonsynonymous SNV		rs78067484	NM_004577
PSPH	11	chr7:56019681	exon 5	G194A	R65H	nonsynonymous SNV		rs20044207 8	NM_004577
TYRO3	11	chr15:41570603		G>T		splicing variant	COSM1478102	rs20068435 0	
SCRN3	10	chr2:174427853	exon 8	1212-1224 del		frameshift deletion	COSM253915	rs14569907 7	NM_001193528
CDCP2	10	chr1:54139645	exon 4	1224 ins C		frameshift insertion		rs36013100	NM_201546
SYT15	10	chr10:46584599	exon 6	G927C	E309D	nonsynonymous SNV	COSM4144699	rs3127785	NM_031912
GXYLT1	10	chr12:42087868	exon 7	G1148A	C383Y	nonsynonymous SNV		rs20097303 0	NM_001099650
MROH5	10	chr8:141494938		C>T		splicing variant		rs6578193	
PTCHD3	9	chr10:27413327	exon 1	923 ins G		frameshift insertion		rs11206712 3	NM_001034842
HOMEZ	9	chr14:23275618	exon 2	1608-1610 del	536_537del	nonframeshift deletion	COSM954652	rs35076736	NM_020834
POTEE	9	chr2:131264056	exon 15	G2601T	E867D	nonsynonymous SNV	COSM4303570	rs7424029	NM_001083538
PSPH	9	chr7:56019607	exon 5	G268A	G90S	nonsynonymous SNV		rs75395437	NM_004577
PDE4DIP	9	chr1:149005097	exon 27	A4075G	K1359E	nonsynonymous SNV	COSM4590058	rs1747958	NM_001198834
TRPT1	9	chr11:64226062		G>C		splicing variant		rs2429457	
TYRO3	8	chr15:41570156	exon 10	1382 del		frameshift deletion			NM_006293
RNPC3	8	chr1:103533845	exon 3	347 del		frameshift deletion		rs772963253	NM_017619
CELA1	8	chr12:51329814	exon 7	628 ins C		frameshift insertion		rs17860363	NM_001971
UBE2N	8	chr12:93411143	exon 2	C187G	P63A	nonsynonymous SNV			NM_003348

VWF	8	chr12:6018901	exon 28	C4517T	S1506L	nonsynonymous SNV	COSM5313831	rs61750100	NM_000552
ACTR2	8	chr2:65261302	exon 7	C791T	A264V	nonsynonymous SNV			NM_005722
DNAJC28	7	chr21:33488443	exon 2	947-951 del		frameshift deletion		rs13985226 2	NM_001040192
SSPO	7	chr7:149806829	exon 58	8748 del		frameshift deletion		rs66470151	NM_198455
CYP4B1	7	chr1:46815075	exon 8	881-882 del		frameshift deletion		rs3215983	NM_000779
CHRNA7	7	chr15:32157674	exon 6	497-498 del		frameshift deletion	COSM5002499	rs37460373 4	NM_000746
TMEM254	7	chr10:80081665	exon 2	114 ins A		frameshift insertion		rs11317252 6	NM_001270371
FSIP2	7	chr2:185738878	exon 1	252 ins G		frameshift insertion		rs35617283	NM_173651
ANP32E	7	chr1:150226708	exon 4	453-458 del	151-153 del	nonframeshift deletion	COSM4770175		NM_001136478
THAP11	7	chr16:67842921	exon 1	367-369 del	123 del	nonframeshift deletion	COSM1479007	rs377516180	NM_020457
ACVR2A	7	chr2:147918575	exon 7	A621C	K207N	nonsynonymous SNV	COSM132653	rs37105918 4	NM_001278580
RPS3A	7	chr4:151102986	exon 4	A470C	Q157P	nonsynonymous SNV	COSM328158	rs13997982 8	NM_001006
NOTCH1	7	chr9:136518645	exon 6	A1045C	T349P	nonsynonymous SNV		rs20052008 8	NM_017617
PPFIA3	7	chr19:49133149		T>G		splicing variant	COSM135843	rs78135388 8	
PTEN	7	chr10:87864104		del T		splicing variant		rs71022512	
MPRIP	7	chr17:17154305		G>T		splicing variant			
OR51F1	6	chr11:4769644	exon 1	274 del		frameshift deletion		rs34672924	NM_001004752
OR10AD1	6	chr12:48203092	exon 1	200dupT		frameshift insertion		rs14424784 1	NM_001004134
ALOX5AP	6	chr13:30713841	exon 1	116 ins GTGT		frameshift insertion		rs36963648 3	NM_001204406

SETBP1	6	chr18:44876699	exon 4	675 ins TCTC		frameshift insertion			NM_0011301 10
PSORS1C1	6	chr6:31138723	exon 5	112 ins C		frameshift insertion		rs13847498 6	NM_014068
ZFYVE19	6	chr15:40807701	exon 1	112 ins GGGGC		frameshift insertion		rs14273057 4	NM_0010772 68
FAM205C	6	chr9:34893049	exon 4	354-355 CC>G-		frameshift substitutio n		rs71506187	NM_0013094 26
NUCB2	6	chr11:17330931	exon 13	1203-1205 del	401-402 del	nonframes hift deletion	COSM1237695	rs3842269	NM_005013
ATN1	6	chr12:6936729	exon 5	1462-1482 del	488-494 del	nonframes hift deletion	COSM1476884		NM_0010070 26
PLEKHA6	6	chr1:204249206	exon 11	T1652G	V551G	nonsynony mous SNV		rs20096198 0	NM_014935
PLXNA2	6	chr1:208098968	exon 6	T1609G	C537G	nonsynony mous SNV		rs20069876 5	NM_025179
L2HGDH	6	chr14:50283981	exon 5	T593G	V198G	nonsynony mous SNV		rs20169264 5	NM_024884
CAPN5	6	chr11:77084981	exon 2	T95C	F32S	nonsynony mous SNV	COSM3986477	rs20125654 7	NM_004055
RPS18	6	chr6:33276005	exon 4	A230G	Y77C	nonsynony mous SNV	COSM1131899	rs76983876 6	NM_022551
ERVV-1	6	chr19:53014925	exon 1	C835G	P279A	nonsynony mous SNV	COSM4132391	rs14087626 8	NM_152473
PDE4DIP	6	chr1:148931920	exon 3	G339T	Q113H	nonsynony mous SNV	COSM4593888	rs3961613	NM_0010028 10
PABPC1	6	chr8:100705591	exon 12	T1685C	L562S	nonsynony mous SNV		rs80006036	NM_002568
WAS	6	chrX:48685604	exon 3	A331C	T111P	nonsynony mous SNV			NM_000377
NPHP4	6	chr1:5875102		T>A		splicing variant		rs1287637	
SLC7A13	5	chr8:86214406	exon 4	1413-1511 del		frameshift deletion		rs56993779	NM_138817
PDE4DIP	5	chr1:149010509	exon 31	4994 del		frameshift deletion			NM_0011988 34

LFNG	5	chr7:2513247	exon 2	138- 139 ins GATG		frameshift insertion		rs34637446	NM_001166355
WDR73	5	chr15:84643646	exon 8	944-961 del	315-321 del	nonframeshift deletion	COSM1375060	rs11267906	NM_032856
FOXE1	5	chr9:97854419	exon 1	505-510 del	169-170 del	nonframeshift deletion	COSM1724903	rs71369530	NM_004473
PAK2	5	chr3:196782706	exon 2	C60G	S20R	nonsynonymous SNV	COSM1422033	rs76714248	NM_002577
TMTC2	5	chr12:82857341	exon 2	C415G	R139G	nonsynonymous SNV	COSM1188560	rs200268500	NM_152588
JAG1	5	chr20:10645368	exon 16	A2101C	T701P	nonsynonymous SNV		rs79176844	NM_000214
FRMD4A	5	chr10:13657338	exon 22	A2251C	T751P	nonsynonymous SNV		rs199968440	NM_018027
KCNJ18	5	chr17:21703568	exon 3	G782A	R261H	nonsynonymous SNV			NM_001194958
TUBB8	5	chr10:47467	exon 4	C925T	R309C	nonsynonymous SNV		rs78262855 6	NM_177987
PANK3	5	chr5:168561512	exon 5	G817A	G273R	nonsynonymous SNV		rs200317426	NM_024594
PABPC1	5	chr8:100704992	exon 13	G1752A	M584I	nonsynonymous SNV		rs11286810 1	NM_002568
EIF4EBP1	5	chr8:38057147	exon 2	C212T	P71L	nonsynonymous SNV			NM_004095
WAS	5	chrX:48685610	exon 3	T337C	F113L	nonsynonymous SNV			NM_000377
FLNB	5	chr3:58123679	exon 21	T3713A	I1238K	nonsynonymous SNV			NM_001164317
VPS50	5	chr7:93294638		T>G		splicing variant		rs75893203	
BCAP31	5	chrX:153724015		C>A		splicing variant		rs18470739 6	
TAS2R19	5	chr12:11021672	exon 1	A900G	X300W	stoploss		rs79475879	NM_176888
OR6C76	4	chr12:55427175	exon 1	922 del		frameshift deletion		rs39771996 5	NM_001005183
SPATA4	4	chr4:176184859	exon 6	836-839 del		frameshift deletion		rs28381989	NM_144644

MSH3	4	chr5:80654881	exon 1	154-171 del	52-57 del	nonframeshift deletion	COSM3718906	rs20187476 2	NM_002439
NCL	4	chr2:231460704	exon 4	774-776 del	258-259 del	nonframeshift deletion	COSM3736247	rs13977735 1	NM_005381
TSKS	4	chr19:49746676	exon 6	769-786 del	257-262 del	nonframeshift deletion	COSM5056834	rs55091696 0	NM_021733
NINL	4	chr20:25476414	exon 17	2872-2877 del	958-959 del	nonframeshift deletion	COSM1025361	rs34410422	NM_025176
CDSN	4	chr6:31117166	exon 2	447-449 del	149-150 del	nonframeshift deletion	COSM1077524	rs34182778	NM_001264
RNH1	4	chr11:502130	exon 2	19-33 del	7-11 del	nonframeshift deletion	COSM927774	rs71022920	NM_203384
POTEE	4	chr2:131263902	exon 15	G2447A	R816H	nonsynonymous SNV	COSM3836843	rs11546936	NM_001083538
POTEI	4	chr2:130462929	exon 15	G3115A	V1039M	nonsynonymous SNV		rs4850284	NM_001277406
MANEAL	4	chr1:37796745	exon 3	T662G	V221G	nonsynonymous SNV	COSM4143887	rs75705909	NM_001031740
KCNJ18	4	chr17:21703692	exon 3	G906T	M302I	nonsynonymous SNV			NM_001194958
CHD3	4	chr17:7905953	exon 28	T4499G	V1500G	nonsynonymous SNV	COSM4130771	rs201727011	NM_001005271
CDK11B	4	chr1:1636429	exon 16	C1202T	A401V	nonsynonymous SNV		rs1059811	NM_033487
FOXD4L1	4	chr2:113499719	exon 1	A463G	I155V	nonsynonymous SNV	COSM224838	rs19984579 2	NM_012184
PABPC3	4	chr13:25096638	exon 1	C440T	T147I	nonsynonymous SNV		rs78432860	NM_030979
PABPC1	4	chr8:100704954	exon 13	T1790C	L597P	nonsynonymous SNV		rs78146983	NM_002568
DCAF13	4	chr8:103440233	exon 9	C1504T	R502C	nonsynonymous SNV	COSM3412622		NM_015420
WNT2B	4	chr1:112516325	exon 3	C313T	R105C	nonsynonymous SNV		rs76236909 7	NM_001291880
TMPRSS13	4	chr11:117909838	exon 7	G972T	Q324H	nonsynonymous SNV			NM_001206789

COQ4	4	chr9:128325797	exon3	G221A	R74Q	nonsynonymous SNV	COSM5021646	rs2270203	NM_001305942
RABL6	4	chr9:136839820	exon13	G1888A	G630R	nonsynonymous SNV		rs14712472 5	NM_001173988
PCDHA5	4	chr5:140822899	exon1	A1124T	D375V	nonsynonymous SNV		rs13924549 6	NM_018908
PCDHGA5	4	chr5:141365274	exon1	A944G	Y315C	nonsynonymous SNV		rs19951270 8	NM_018918
RBSN	4	chr3:15090399	exon4	G289T	G97C	nonsynonymous SNV			NM_001302378
RNF175	4	chr4:153715537	exon7	T756G	C252W	nonsynonymous SNV		rs14222430 6	NM_173662
EZR	4	chr6:158787182	exon3	C118T	R40W	nonsynonymous SNV		rs77260842 8	NM_003379
SLC22A1	4	chr6:160122197	exon1	T262C	C88R	nonsynonymous SNV	COSM3928207	rs55918055	NM_003057
SCN7A	4	chr2:166405764	exon25	G4865A	R1622Q	nonsynonymous SNV		rs18878193 5	NM_002976
ALDH4A1	4	chr1:18883153	exon7	G469A	G157S	nonsynonymous SNV		rs78029802 7	NM_001161504
GATAD2A	4	chr19:19492680	exon3	A502G	S168G	nonsynonymous SNV		rs77997146 3	NM_001300946
NTHL1	4	chr16:2040195	exon5	G753C	W251C	nonsynonymous SNV			NM_002528
GJC2	4	chr1:228158846	exon2	C1088T	A363V	nonsynonymous SNV			NM_020435
CHRND	4	chr2:232531450	exon6	C337T	P113S	nonsynonymous SNV		rs14206332 8	NM_001311195
PYGB	4	chr20:25292528	exon17	G2092A	V698M	nonsynonymous SNV		rs15058250 2	NM_002862
DMD	4	chrX:31203978	exon6	C586T	R196W	nonsynonymous SNV		rs37344800 2	NM_004015
DMD	4	chrX:32699119	exon8	C800T	S267F	nonsynonymous SNV			NM_000109
SCUBE3	4	chr6:35239766	exon8	C841T	R281C	nonsynonymous SNV	COSM1265091	rs20195255 4	NM_001303136

COL9A2	4	chr1:40314260	exon4	A194G	K65R	nonsynonymous SNV		rs75663465 9	NM_001852
SPTBN5	4	chr15:41887281	exon6	G820A	V274I	nonsynonymous SNV		rs55830029	NM_016642
CYP4A11	4	chr1:46936786	exon4	G388A	G130S	nonsynonymous SNV		rs62621075	NM_000778
SKOR2	4	chr18:47248873	exon1	G311A	R104H	nonsynonymous SNV	COSM4595039	rs57468505 7	NM_001037802
GNG2	4	chr14:51966613	exon3	G142T	D48Y	nonsynonymous SNV	COSM1629335		NM_001243774
DNAH1	4	chr3:52353415	exon20	A3262T	I1088F	nonsynonymous SNV			NM_015512
ITIH3	4	chr3:52799805	exon9	T959C	L320P	nonsynonymous SNV			NM_002217
SOX17	4	chr8:54458151	exon1	G13C	D5H	nonsynonymous SNV			NM_022454
SOX17	4	chr8:54459282	exon2	G532T	G178C	nonsynonymous SNV		rs26760708 2	NM_022454
TRIM6	4	chr11:5611169	exon6	C769T	R257C	nonsynonymous SNV	COSM429249		NM_001198645
PHF3	4	chr6:63691742	exon4	T2195C	M732T	nonsynonymous SNV			NM_015153
MAP3K11	4	chr11:65607491	exon5	A1268G	E423G	nonsynonymous SNV			NM_002419
CLDN3	4	chr7:73769649	exon1	C401T	P134L	nonsynonymous SNV		rs13919132 8	NM_001306
GOLGA6L10	4	chr15:82344664	exon6	T1196C	L399P	nonsynonymous SNV		rs77331571 9	NM_001164465
FNBP1L	4	chr1:93551055	exon14	G1586C	G529A	nonsynonymous SNV			NM_001024948
SLC26A1	4	chr4:989077	exon3	T1862A	F621Y	nonsynonymous SNV			NM_022042
ZFP64	4	chr20:52164759		ins AA		splicing variant		rs20005997 8	
MTCH2	4	chr11:47622719	exon11	G780A	W260X	stopgain		rs79609634 7	NM_001317232
RETNLB	4	chr3:108757146	exon1	39 ins TAATCCCC	L14 insX	stopgain		rs36849766 0	NM_032579
KCNQ2	4	chr20:63407011	exon15	C2168A	S723X	stopgain			NM_004518

TYRO3	3	chr15:41571117	exon 13	1659-1660 del		frameshift deletion			NM_006293
OR5B3	3	chr11:58403320	exon 1	90 del		frameshift deletion		rs200799158	NM_001005469
GRP	3	chr18:59230436	exon 3	394-397 del		frameshift deletion		rs14996206 8	NM_001012512
SETBP1	3	chr18:44876700	exon 4	676-677 del		frameshift deletion			NM_001130110
AP3S1	3	chr5:115866721	exon 2	121-124 del		frameshift deletion	COSM1319295	rs80118146	NM_001284
PEBP4	3	chr8:22713395	exon 7	659 del		frameshift deletion		rs35121552	NM_144962
CNTNAP3	3	chr9:39149858	exon 10	1597 del		frameshift deletion			NM_033655
DSPP	3	chr4:87615750	exon 5	3088-3105 del	1030-1035 del	nonframeshift deletion	COSM5547737	rs77355733 0	NM_014208
CTBS	3	chr1:84574316	exon 1	92-100 del	31-34 del	nonframeshift deletion	COSM1290235	rs14253476 2	NM_004388
FAM109A	3	chr12:111363023	exon 3	397-405 del	133-135 del	nonframeshift deletion	COSM430328	rs13903286 7	NM_001177997
DSPP	3	chr4:87615391	exon 5	2729-2737 del	910-913 del	nonframeshift deletion	COSM4603772	rs11145663 7	NM_014208
NPIPB6	3	chr16:28343070	exon 7	813-815 del	271-272 del	nonframeshift deletion	COSM5215813	rs37469258 8	NM_001282524
TRAK1	3	chr3:42210086	exon 13	1842-1844 del	614-615 del	nonframeshift deletion	COSM308433	rs75344077 4	NM_001265609
HRCT1	3	chr9:35906351	exon 1	64-66 del	22 del	nonframeshift deletion	COSM1490012	rs37060624 6	NM_001039792
PAK2	3	chr3:196803111	exon 4	A383G	K128R	nonsynonymous SNV	COSM1717568	rs78043821	NM_002577
ZBTB4	3	chr17:7465792	exon 3	A1010C	Y337S	nonsynonymous SNV		rs20126423 8	NM_001128833
AK2	3	chr1:33013299	exon 6	A578T	Y193F	nonsynonymous SNV		rs11371146 7	NM_001199199
KCNJ18	3	chr17:21703651	exon 3	G865C	E289Q	nonsynonymous SNV			NM_001194958
PRSS3	3	chr9:33795599	exon 1	T26C	F9S	nonsynonymous SNV			NM_002771

CACNA1G	3	chr17:50615426	exon25	A4723C	T1575P	nonsynonymous SNV		rs20082577 5	NM_198376
CHRNB2	3	chr1:154569495	exon2	T98G	V33G	nonsynonymous SNV	COSM3726999	rs20072932 8	NM_000748
LDLRAD3	3	chr11:36227304	exon4	A527C	H176P	nonsynonymous SNV	COSM1285901	rs75089692 5	NM_001304263
SLIT3	3	chr5:168685818	exon31	A3445C	T1149P	nonsynonymous SNV		rs20138639 6	NM_001271946
TFAM	3	chr10:58388704	exon4	T326G	V109G	nonsynonymous SNV	COSM5034010	rs77418790	NM_001270782
CNTNAP3	3	chr9:39078875	exon22	G3488A	G1163D	nonsynonymous SNV		rs75157819 6	NM_033655
RBMX	3	chrX:136875541	exon6	A586G	R196G	nonsynonymous SNV		rs13995433 3	NM_002139
PA2G4	3	chr12:56106718		ins A		splicing variant		rs34728522	
SLC3A1	3	chr2:44301129		del T		splicing variant		rs61179824	
LRRC37A3	3	chr17:64858885		ins A		splicing variant		rs54020713 8	
IQCK	3	chr16:19717694		G>C		splicing variant		rs4782272	
PIBF1	3	chr13:72835370		ins A		splicing variant		rs20068394 0	
CSF1	3	chr1:109924191		G>T		splicing variant			
PKD2L2	3	chr5:137936318		A>G		splicing variant			
MAD1L1	3	chr7:2014501		C>A		splicing variant			
PAIP1	3	chr5:43538923		C>T		splicing variant			
DGKZ	3	chr11:46378989		A>C		splicing variant			
PIGQ	3	chr16:582881		A>T		splicing variant			
DOCK8	3	chr9:463655	exon45	C5907A	Y1969X	stopgain	COSM3982891	rs79568455	NM_001190458
NPIPB15	3	chr16:74391889	exon7	G1141T	E381X	stopgain	COSM4592878	rs37577669 3	NM_001306094

BAGE 4	3	chr21:10 414915	exon 2	A120C	X40C	stoploss			NM_18 1704

Supplementary Table S6: Detailed list of variants found in at least 10% of the WES patient cohort. Ntde = nucleotide, aa = amino acid

Gene	Mutation (Amino Acid change)	Number of patients presenting with mutation
NOTCH1	T311P	6
	T349P	3
	H1190P	1
	G1503S	1
CTBP2	T731R	2
	G813S	1
TP53	R141C	2
	C44F	1
ACTR2	A264V	2
LFNG	Insertion of 4 nucleotides leading to a frameshift after amino acid 46	5
JAG1	T701P	5
NOTCH3	C864F	2
MAML3	Deletion of 11 nucleotides leading to a frameshift after amino acid 505	2
NCSTN	T476M	2
DLL3	C4553R	2
WNT2B	R105C	4

Supplementary Table S8: Details of the mutations detected in the Notch1 pathway by GSEA

	All patients		Adult patients		Paediatric patients		ALK+ ALCL patients		ALK- ALCL patients	
	WT NOTCH1	NOTCH1 T349P and/or T311P	WT NOTCH1	NOTCH1 T349P and/or T311P	WT NOTCH1	NOTCH1 T349P and/or T311P	WT NOTCH1	NOTCH1 T349P and/or T311P	WT NOTCH1	NOTCH1 T349P and/or T311P
# Patients	n = 57	n = 11	n = 13	n = 3	n = 27	n = 5	n = 45	n = 8	n = 12	n = 3
OS (years)	4.4	3.6	4.2	1.7	4.4	2.9	4.5	4.3	4.5	1.7
EFS (years)	2.6	2.1	4	1.7	4.2	3.2	2.3	2.2	3.6	1.7
Death (%)	25%	36%	38%	67%	26%	20%	18%	25%	50%	67%
Relapse (%)	61%	55%	38%	0%	52%	60%	67%	75%	42%	0%

Supplemental Table S9: Clinical data pertaining to the validation cohort. Data is based on 5-years of follow-up. Similarly, the average OS (Overall Survival) and EFS (Event-Free Survival) is based on a maximum of 5 years follow-up. Adult patients are defined as being older than 18 years of age.

Name	Direction	Application	Sequence
NOTCH1	Forward	Sanger validation	CTCTGCCTGGCGCTGCTG
NOTCH1	Reverse	Sanger validation	GGAAACAACTGCAAGAACGGG
U6	Forward	Sanger sequencing	AATGACTATCATATGCTTACCG
H1_TET	Forward	Sanger sequencing	TCGCTATGTGTTCTGGGAAA
CMV_F	Forward	Sanger sequencing	CGCAAATGGGCGGTAGGCGTG
SP6	Forward	Sanger sequencing	CGATTTAGGTGACACTATAG
NOTCH1	Forward	qPCR	TACAAGTGCGACTGTGACCC
NOTCH1	Reverse	qPCR	ATACACGTGCCCTGGTTCAG
HEY1	Forward	qPCR	GTTCGGCTCTAGGTTCCATGT
HEY1	Reverse	qPCR	CGTCGGCGCTTCTAATTATTC
HES1	Forward	qPCR	TCAACACGACACCGGATAAAC
HES1	Reverse	qPCR	GCCGCGAGCTATCTTTCTTCA
GAPDH	Forward	qPCR	CTGGGCTACACTGAGCACC
GAPDH	Reverse	qPCR	AAGTGGTCGTTGAGGGCAATG
PPIA	Forward	qPCR	GCTTTGGGTCCAGGAATG
PPIA	Reverse	qPCR	AGAAGGAATGATCTGGTGGTTAAG
DLL1	Forward	qPCR	GATTCTCCTGATGACCTCGCA
DLL1	Reverse	qPCR	TCCGTAGTAGTGTTCGTCACA
MYC	Forward	qPCR	GGCTCCTGGCAAAAGGTCA
MYC	Reverse	qPCR	CTGCGTAGTTGTGCTGATGT
DTX1	Forward	qPCR	GACGGCCTACGATATGGACAT
DTX1	Reverse	qPCR	CCTAGCGATGAGAGGTCGAG
STAT3	Forward	qPCR	CAGCAGCTTGACACACGGTA
STAT3	Reverse	qPCR	AAACACCAAAGTGGCATGTGA
NOTCH1_349	Forward	Site-Directed Mutagenesis	GGTCATGGCAGGGGCGCCGTGGAA
NOTCH1_349	Reverse	Site-Directed Mutagenesis	TTCCACGGCGCCCCTGCCATGACC
NOTCH1_311	Forward	Site-Directed Mutagenesis	GTGTTGTGGCAGGGCCCGCCGTTCTGG
NOTCH1_311	Reverse	Site-Directed Mutagenesis	CCAGAACGGCGGGCCCTGCCACAACAC
NOTCH1	Forward	ChIP	ATCAACCTGTTCCTCCCCTG
NOTCH1	Reverse	ChIP	TTCCCGACTACAAGCGGACT
IRF4	Forward	ChIP	CTCTAAACACCGCGGAGAGG
IRF4	Reverse	ChIP	CTTTGCAGAGCGTGTAACGG
Control	Forward	ChIP	ATTCCACCTTGTCCAGCCCT
Control	Reverse	ChIP	GGTTTTATCCCTCTCCCCGAC

Supplementary Table S10: Detailed list of oligos used in this study.

Cell line	Tissue of origin	Cell of origin	Sex	Karyotype	Species	Growth Medium	Doubling time	Growth Mode	Ref
HEK293FT	Kidney (foetal)	Epithelial	Female	Hypotriploid	Human	DMEM + 10% FBS	20 hrs	Adherent	38
Karpas 299	Lymph node	ALK+ ALCL	Male	Hypodiploid	Human	RPMI 1640 + 10% FBS	30 hrs	Suspension	39
SU-DHL1	Lymph node	ALK+ ALCL	Male	Octoploid	Human	RPMI 1640 + 10% FBS	45 hrs	Suspension	40
SUP-M2	Lymph node	ALK+ ALCL	Female	Near-Diploid	Human	RPMI 1640 + 10% FBS	45 hrs	Suspension	41
DEL	Lymph node	ALK+ ALCL	Male	Hyper-diploid	Human	RPMI 1640 + 10% FBS	35 hrs	Suspension	42
MAC2A	Metastatic lymph node	ALK- ALCL	Male	Near-Diploid	Human	RPMI 1640 + 10% FBS	50 hrs	Suspension	43
FEPD	Peripheral Blood	ALK- ALCL	Female	Unknown	Human	RPMI 1640 + 10% FBS	50 hrs	Suspension	44
OP9	Bone Marrow	Embryonic stem cell	?	Unknown	Mouse	α-MEM + 20% FBS	26 hrs	Adherent	45

Supplementary Table S11: Cell line description

Plasmid	Reference	Selection antibiotic
psPAX2	Addgene; Cat# 12260	-
PMD2.G	Addgene; Cat# 12259	-
pLJM1-EGFP-NOTCH1	-	Puromycin
MISSION® shRNA for NOTCH1	Sigma-Aldrich, Cat# SHCLNG-NM_017617 (TRCN0000003362, TRCN0000350253, TRCN0000350254)	Puromycin
MISSION® shRNA for STAT3	Sigma-Aldrich, Cat# SHCLNG-NM_003150 (TRCN0000020840, TRCN0000020842)	Puromycin
pLVTHM vector containing the H1 promoter ALK-shRNA (A5) cassette	Piva et al., 2006	Puromycin

Supplementary Table S12: Detailed list of plasmids used in this study