Long non-coding RNA mitophagy and ALK-negative anaplastic lymphoma-associated transcript: a novel regulator of mitophagy in T-cell lymphoma

Valentina Mularoni,[1] Benedetta Donati,[1] Annalisa Tameni,[1] Veronica Manicardi,[1] Francesca Reggiani,[1] Elisabetta Sauta,[2] Magda Zanelli,[3] Marco Tigano,[4] Emanuele Vitale,[1,5] Federica Torricelli,[1] Stefano Ascani,[6] Giovanni Martino,[6,7] Giorgio Inghirami,[8] Francesca Sanguedolce,[9] Alessia Ruffini,[10] Alberto Bavieri,[10] Stefano Luminari,[10] Marco Pizzi,[11] Angelo Paolo Dei Tos,[11] Cinzia Fesce,[12] Antonino Neri,[13] Alessia Ciarrocchi[1] and Valentina Fragliasso[1]

[1]Laboratory of Translational Research, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [2]IRCCS Humanitas Clinical and Research Center, Milan, Italy; [3]Pathology Unit, Department of Oncology, Oncology, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [4]Sidney Kimmel Medical College, Thomas Jefferson University, Philadelphia, PA, USA; [5]Clinical and Experimental Medicine Ph.D. Program, University of Modena and Reggio Emilia, Modena, Italy; [6]Pathology Unit, Azienda Ospedaliera Santa Maria di Terni, University of Perugia, Terni, Italy; [7]Institute of Hematology and CREO, University of Perugia, Perugia, Italy; [8]Department of Pathology and Laboratory Medicine, Weill Cornell Medicine, New York, NY, USA; [9]Pathology Unit, Policlinico Riuniti, University of Foggia, Foggia, Italy; [10]Hematology Unit, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy; [11]Surgical Pathology & Cytopathology Unit, Department of Medicine-DIMED, University of Padova, Padova, Italy; [12]Hematology Unit, University Hospital, Foggia, Italy and [13]Scientific Directorate, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy

Correspondence: V. Fragliasso
Valentina.Fragliasso@ausl.re.it

Received: December 14, 2022.
Accepted: June 20, 2023.
Early view: June 29, 2023.

https://doi.org/10.3324/haematol.2022.282552

Long non-coding RNA mitophagy and ALK$^-$ anaplastic lymphoma associated transcript: a novel regulator of mitophagy in T cell Lymphoma

Supplementary Methods

Cell culture and treatments

The human ALK$^-$ and ALK$^+$ALCL cell lines MAC2A, FePD, L82, SUPM2, SUDHL1, and KARPAS299 were a kind gift of Dr. Giorgio Inghirami. The human Breast Implanted Associated (BIA)-ALCL cell lines TLBR-2 and TLBR-3 were a kind gift of Dr. Alain Epstein. The human CUTLL1 were a kind gift from Dr. Iannis Aifantis while human K562, MJ, and KCL22 cell lines were a kind gift from Dr. Bruno Calabretta. The human cell lines TPC1, BCPAP, 8505, CAL62, H1299, H1975, H1650, MB231, OCI-LY10, OCI-LY13 and NUDUL1 were a kind gift of Dr. Ciarrocchi. Cell identity was determined yearly. All cell lines were genotyped and routinely tested for Mycoplasma contamination. Cell lines were cultured in RPMI-1640 medium (Gibco) supplemented with 10% FBS at 37 °C in an atmosphere of 5% CO_2. TLBR-2 cells were supplemented with IL2 (20U/ml). Doxycycline hyclate was purchased from Sigma and dissolved in H_2O.

Chloroquine was purchased from Sigma and dissolved in H_2O.

Cytoplasm/Nucleus fractionation

ALCL cells were resuspended in hypotonic buffer (20mM tris-HCl pH 7; 10mM NaCl; 3mN MgCl2; 0.3% NP40) supplemented with SUPERase•In (Ambion) for 15 min on ice. After a centrifugation at 800*g* for 10 min, the supernatant was collected as the cytoplasmic fraction. The pellet was resuspended in cell extraction buffer (10mM Tris pH 7.4; 2mM Na3VO4; 100mM NaCl; 1% Triton-x100; 1mM EDTA; 10% glycerol; 1mM EGTA; 0.1% SDS; 1mM NaF; 0.5% Na-deoxycholate; 20mM Na4P2O7) supplemented with RNase inhibitors for 30 min on ice. After a centrifugation at 14,000*g* for 30 min, the supernatant was collected as the nuclear

fraction. Chromatin was pelleted at maximum speed for 3 min. All fractions were resuspended in TRIzol (Invitrogen) and RNA was extracted following the standard protocol. All fractions were resuspended in TRIzol (Invitrogen) and RNA was extracted following the standard protocol.

Generation of TLBR-2 and MAC2A dCas9-KRAB MTAAT^{KD} cell lines

pHAGE TRE dCas9-KRAB was a gift from Rene Maehr & Scot Wolfe (Addgene plasmid # 50917; http://n2t.net/addgene:50917; RRID:Addgene_50917). Vector was packaged into lentiviral particles HEK 293T-cell line and used for infection of low passages MAC2A or TLBR-2. Cells were selected with 0.5 mg/ml of Geneticin™ Selective Antibiotic (G418 Sulfate, Gibco), for 3 days (both MAC2A and TLBR-2).

Annealed sgRNA oligomers were ligated into BsmB1 digested LRG2.1 plasmid (a gift from Christopher Vakoc, addgene plasmid #108098; http://n2t.net/addgene:108098; RRID:Addgene_108098) and lentiviral particles were created as previously described[16]. Viral particles were used to infect TLBR-2 dCas9-KRAB and MAC2A dCas9-KRAB previously derived cell lines. Infected cells were then purified by gating GFP+ cells using BD FACS Melody cell sorter. The list of sgRNA sequences is provided in **supplementary table 2**.

Gene Expression Profiling (GEP) by Nanostring

Total RNA was extracted by Maxwell® RSC RNA FFPE kit (Promega) starting from 5 slides of 5μm FFPE tissue. RNA quantity and quality were assessed by NanoDrop2000 (Thermo Fisher Scientific). For samples that reached the quality standards (A260/A280 ≥ 1.7 and A260/A230 ≥ 1.8), we evaluated the gene expression profile (GEP) by nCounter platform (NanoString Technologies) using a custom panel. This panel includes a total of 39 transcripts: 17/18 lncRNAs from the non coding-signature previously generated by RNA-seq platform[16], that showed a suitable sequence to generate unique and specific nCounter

probes, 4 ALCL-restricted coding transcripts (ALK, TMOD1, BATF, TNFRSF8), 1 ALCL-specific non-coding transcript (ERBB4)[20], BNIP3, BNIP3L and 15 housekeeping (**supplementary table 3**). Analysis of detected gene counts was performed by nSolver Analysis Software 4.0 (NanoString Technologies) as previously described[21]. Briefly, for samples that passed imaging quality controls, raw counts of coding genes were subjected to background subtraction as mean counts of negative controls plus two standard deviation and then normalized on synthetic positive controls. On the contrary, no background subtraction was applied to non coding transcript that were normalized on synthetic positive controls only. After that, counts normalized on technical controls were further normalized on the 3 housekeeping genes with the lowest coefficient of variation (COG7, DNAJC14, ERCC3) and log2 transformed. Applying the *3 gene model*[22] and considering the lack of ALK expression, 29 patients were classified as ALK- ALCL. Almost all these patients (n=26, 90%) showed the expression of at least one among TMOD, BATF and TNFRSF8 genes **(supplementary table 4).** On the contrary, 14 patients were classified as ALK^+ ALCL because of high ALK expression. This group included 3 ALK^- ALCL that were re-classified as ALK+ because of their high level of ALK mRNA. Moreover, one patient that showed a low level of both ALK and 3 genes model was classified as ALK^+ ALCL according to the current WHO diagnostic criteria classification [1].

Positive expression of the genes included in the 3 genes model (TMOD, BATF and TNFRSF8) was considered for level of expression over the 1^{st} quartile of normalized gene counts distribution.

To investigate lncRNAs differential expression a build ratio analysis was performed by comparing the transcriptomic profiles of ALK^- and ALK^+ samples. For each comparison, the p-value was calculated as Kruskal-Wallis test since data were not normal distributed. Correlation between normalized gene counts was evaluated by Pearson correlation coefficient. ERBB4, was used as positive technical control for lncRNA detection.

To evaluate the performance of MTAAT in predicting ALCL subtypes, we constructed the receiver operating characteristic (ROC) curve and calculated the area under the ROC curve (AUC) and the relative accuracy of prediction. Bioinformatic analyses on GEP were conducted by R Software v4.1.3 using the following R packages: ggplot2, ggbiplot (function prcomp), corrplot, pROC and ROCR.

Chromatin immunoprecipitation (ChIP) and ChIP-sequencing

TLBR-2 cells ($15*10^6$/IP) were crosslinked for 15 min with 1% formaldehyde, lysed and sonicated for 15 cycles (30 min ON, 30 min OFF) using Bioruptor Pico Sonicator (Diagenode, Denville, NJ, USA) to obtain 100-200 bp chromatin fragments. Chromatin was precipitated overnight using Dynabeads Protein G magnetic beads (Thermo Fisher Scientific) and 1,5ug of H3K4me3 (Rabbit Polyclonal, Abcam), 1,5ug of H3K27Ac (ab4729, Rabbit Polyclonal, Abcam), 2,5ug of RNAPII (Rabbit Monoclonal, #14958, Cell Signaling Technology), or IgG-isotype control (#66362, Cell Signaling Technology). A fraction equal to 0.25% of total chromatin was used as input. For Chip, each RT-qPCR value was normalized over the appropriate input control and reported in graphs as a % of input. The list of primers used is provided in **supplementary table 1**.

For ChIP-seq, samples were quantified with Qubit (Thermo Fisher Scientific), and the quality was evaluated by Bioanalyzer (Agilent Technologies). Library for sequencing was obtained following the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England BioLabs, Ipswich, MA, USA) using 3-10 ng ChIP DNA as starting material. Triplicates were sequenced on Illumina NextSeq500 high-output cartridge (single stranded, reads length 75 bp-1 × 75).

Library preparation and RNA-sequencing

RNA-seq libraries were obtained starting from 100 ng of total RNA following Illumina Stranded TotalRNA Prep Ligation with Ribo-zero Plus protocol. Sequencing was performed using Illumina NEXTSeq high-output cartridge (double-stranded, reads length 75bp-2 ×75).

Sequencing data processing

Sequencing data quality was assessed using the FastQC v0.11.8 software (www.bioinformatics.babraham.ac.uk/projects/fastqc/), low quality reads were discarded, and where residual adapters were present, they were removed using Trimmomatic (v.0.39) software.

For ChIP-seq, filtered reads were aligned to the human reference genome (GRCh38/hg38 assembly) by using Bowtie2 version 2.3.5.1. Picard tool (http://broadinstitute.github.io/picard) and samtools 1.9v (http://samtools.sourceforge.net/) were used to remove duplicates and unmapped reads and to retain uniquely aligned reads for downstream analyses. Peak calling was performed using MACS2 (v.2.1.3.3). Significant peaks ($q < 0.05$) were merged, and high confidence peaks enriched in at least 2 of 3 replicates were retained for the analysis.

ChIPseeker R package was used for assigning peaks to the nearest genes, according to GRCh38/hg38 annotation, using a transcription start site (TSS) window of ±3 kb.

For RNA-seq, filtered paired-end reads were aligned to the human reference transcriptome (GRCh38, Gencode release 30 using STAR version 2.7), and gene abundances were estimated with RSEM algorithm (v1.3.1). Differential analysis was performed with R package DESeq2, considering a false discovery rate (FDR) of 10% and excluding genes with low read counts. Significantly deregulated genes underwent enrichment analysis, performed through enrichR package on Gene Ontology biological processes using a significance threshold of 0.05 on P value adjusted for multiple testing using the Benjamini–Hochberg correction.

Prediction of TF binding sites

TFs binding sites prediction was performed applying FIMO algorithm on a selected region of 500bp around MTAAT TSS. HOCOMOCO and JASPAR were used as reference motifs databases and q-value <0.1 was considered for significant motif enrichment.

Western blot

Western blot analysis was performed using standard techniques. The primary antibodies were: Caspase-3 (#9662, Rabbit Monoclonal, Cell Signaling Technology), PARP-1 (ab137653, Rabbit Polyclonal, Abcam), p62/SQSTM (ab91526, Rabbit Polyclonal, Abcam), LC3 A/B (#4108, Rabbit Monoclonal, Cell Signaling Technology), HA (C29FA, Rabbit, Cell Signaling Technology), B-Actin (AC-15, Mouse, Sigma-Aldrich). All secondary antibodies (rabbit and mouse) were HRPconjugated (GE Healthcare) and diluted 1:3000. Densitometric analysis was performed using ImageJ software.

Cell cycle analysis

For cell cycle analysis, the hypotonic propidium iodide (PI) method was used. All flow cytometry analyses were performed with FACSCanto™ II Cell Analyzer (BD Biosciences).

Generation of promoter and enhancer plasmids

pGL3 basic and pGL3 promoter plasmids (Promega) with the selected putative promoter/enhancer region of MTAAT were generated as follow: the selected region was amplified by PCR from genomic DNA of TLBR-2 using PhusionTM Plus DNA Polymerase kit (Thermo Fisher Scientific) with an upstream oligomer containing a 5′-KpnI site and a downstream oligomer containing a 3′-XhoI site (**supplementary table1**). PCR product of the right length was isolated from a 1% agarose gel, purified, and finally cloned into KpnI-XhoI-digested pGL3 basic/ pGL3 promoter vectors. The plasmids were sequenced and checked for right sequence.

Plasmid transfection

p3XFlagCMV10_BNIP3 was obtained cloning the gBlock Gene Fragment (IDT) for BNIP3 into the p3XflagCMV10 (Sigma Aldrich). For EGFP-LC3 expression, gBlock Gene Fragment (Integrated DNA Technologies, IDT) for LC3 was cloned into the pEGFP-C3 (Diatech Lab Line). 48 hours after induction with doxycycline, cells were nucleofected with a total of 400ng of plasmids and the co-localization of the two proteins was evaluated by immunofluorescence 24 hours after nucleofection.

Luciferase Assay

MAC2A and TLBR-2 cells (1 × 10^6) were transfected with 400 ng of reporter pGL3-luciferase plasmids using the Cell Line Nucleofector Kit SF and 4D Amaxa Nucleofector (program FI115 for MAC2A, DS-130 for TLBR-2). Twenty-four hours after transfection, cells were harvested, and luciferase activity was measured using the Dual-Luciferase Reporter Assay System (Promega) in a GloMax Discover Luminometer (Promega) according to the manufacturer's instructions. For each sample, firefly luciferase activity was normalized on *Renilla* luciferase activity and transactivation of the various reporter constructs was expressed as fold induction on empty vector (pGL3-basic or pGL3-promoter) activity.

Immunohistochemistry (IHC)

Formalin-fixed paraffin-embedded (FFPE) tissue sections were stored and classified by the Pathology Division of the Arcispedale S. Maria Nuova IRCCS, Reggio Emilia, Italy. IHC was performed on 4 µm-thick formalin-fixed, paraffin-embedded sections of 3 ALK^- ALCL using anti-SOD1 (#37385, Rabbit Monoclonal, Cell Signaling Technology, 1:800) monoclonal primary antibody. The immunohistochemical staining was developed on platform Ventana Bench-ULTRA using the OptiView DAB Detection Kit (Ventana-Roche,Tucson, Az) and

amplified with OptiView Amplification Kit (Ventana-Roche,Tucson Az). Slides were counterstained with Hematoxylin II counterstain and Bluing Reagent. The IHC have been evaluated independently by two Pathologists (M.Z and S.A.). ALK$^-$ ALCL FFPE sections were considered SOD1-positive when at least 50% of the cells expressed this antigen. Diagnoses were assigned according to the WHO classification[1].

Supplementary Figure Legends

Supplementary Figure 1. Coding and non-coding profiling of ALCLs

(**A**) Boxplots representing the expression of ALK, BATF3, TMOD1, TNFRSF8 among ALK$^+$ (red) and ALK$^-$ (grey) ALCL evaluated by nCounter platform in the FFPE cohort. Comparisons were considered significant for p≤0.05 (*). (**B**) Boxplot representing the expression of ERBB4 among ALK$^+$ (red) and ALK$^-$ (grey) ALCL evaluated by nCounter platform in the FFPE cohort. Comparison was considered significant for p≤0.05 (*). (**C**) Dot plots show gene expression counts distribution of the six not statistically significant lncRNAs among ALK$^+$ ALCL (red dots) and ALK$^-$ ALCL (black dots) of the FFPE training cohort. (**D**) Graphs show the expression of lncRNAs among ALK$^+$ (red) and ALK$^-$ (grey) ALCL evaluated by RT-qPCR in the validation cohort. (**E**) RT-qPCR analysis expression of MTAAT level in donor resting (n=3) and activated (n=3) CD4+ T-lymphocytes. (**F**) Graph shows the high discriminatory accuracy (70%) of the ROC curve.

Supplementary Figure 2. Regulatory features of MTAAT

(A) RT-qPCR analysis expression of MTAAT level in a panel of cell lines. (**B**) ChIP-PCR detection of IgG on MTAAT fragments in ALK$^-$ ALCL cell lines. GAPDH promoter was used as CTR+ whereas a non-coding intergenic region (CTR−) served as negative control. The values are representative of three independent experiments. (**C**) ChIP-PCR detection of RNAPII, H3K4me3, and H3K27Ac markers on MTAAT fragments in MTAAT-negative cell line

CUTLL1. GAPDH promoter was used as CTR+ whereas a noncoding intergenic region (CTR−) served as negative control. The values are representative of three independent experiments.

Supplementary Figure 3. Functional validation of MTAAT KD through CRISPR dCAS9-KRAB system

(**A**) Relative expression of OIPS5-AS1, DANCR, NEAT1, and KCNQ1OT1 used as controls for cytoplasmic, nucleic, and chromatinic fractions, respectively (representative three independent experiment). **(B)** Dot plots show physical parameters and setting strategy to sort ALCL dCas9-KRAB cells expressing MTAAT-sgRNA-GFP guides. **(C)** Western blot shows time course expression of dCAS9-KRAB-HA after doxycycline treatment. (**D**) RT-qPCR analysis of MTAAT expression 48 hours after doxycycline treatment in TLBR-2 dCAS9-KRAB ($MTAAT^{KD}$) and MAC2A dCAS9-KRAB ($MTAAT^{KD}$). Comparison was considered significant for $p \leq 0.01$ (**). (**E**) RT-qPCR expression analysis of a panel of MTAAT-target genes in ALCL dCas9-KRAB cells expressing sgRNA#3 (48 hours after doxycycline treatment). Comparison was considered significant for $p \leq 0.05$ (*) and $p \leq 0.01$ (**).

Supplementary Figure 4. Loss of MTAAT promotes RNAPII recruitment on its target genes.

(**A**) Graphs show the TSS and fragments around TSS of a set of genes deregulated after $MTAAT^{KD}$ ChIP-qPCR detection of H3K4me3 (**B**), H3K27Ac (**C**) and RNAPII (**D**) on MTAAT-genes target fragments in TLBR-2 $MTAAT^{KD}$ cell lines. The values are representative of three independent experiments. (**E**) RT-qPCR analysis of OPTN expression after MTAAT-KD (48 hours after doxycycline treatment) in TLBR-2 dCAS9-KRAB cells. (**F**) Graph shows interaction propensities of MTAAT with a set of chromatin modifier enzyme. In red, a representative group of H3K4-modifier proteins and in blue, a representative group of H3K9-

modifier proteins. In green, the most important catalytic subunit of RNAPII. The interaction propensity between EZH2 and HOTTIP was used as CTR+ whereases the interaction propensity between Wdr5 and HYMAI was set as CTR-. (**G**) RT-qPCR expression analysis of MTAAT and BNIP3 expression in a panel of non-TCL cell lines (thyroid cancer, purple; lung cancer, blue; breast, red; diffuse large B cell lymphoma, green).

Supplementary Figure 5. Loss of MTAAT does not affect cell cycle nor apoptosis

(**A**) Representative immunohistochemistry analysis of SOD1 in FFPE section of ALK^{-}ALCL patients expressing different level of MTAAT. Magnification 200X. (**B**) Western blots show the expression of autophagic markers after 48hr of MTAAT KD in TLBR-2 and MAC2A dCas9-KRAB sgRNA#3 cells. Densitometric analysis of LC3 refers to LC3II band. Cloroquine (CQ) was added to cells for the last for 2 hours of doxycycline -DOX- treatment (20μM for TLBR-2 and 40μM for MAC2A). (**C**) Histograms show the percentage of cell in each cell cycle phase (72 hours after doxycycline -DOX- treatment). (**D**) Western blot shows the expression of CASP3 and PARP-1 in TLBR-2 MTAAT^{KD} (72 hours after doxycycline treatment).

Supplementary Table

Supplementary Table 1. List of primers used in this study

Primer qPCR MTAAT Promoter	
KpnI_MTAAT Prom F	atcaaggtaccggtaccGTGTCTCCTCAACAGC TGTG
XhoI_MTAAT Prom R	gttcactcgagctcgagACAACGTGCCAGAATCT GTG

Primer lncRNAs	**Sequence (5'-3')**
XLOC_115902 F	TGGGTCACCAGTTCTGCTCT
XLOC_115902 R	AAGGGTGCTGTTTTGAGTGC
XLOC_066584 F	AGGAACTCTACTCAAGATTCTGGG
XLOC_066584 R	TCACTGTAAGTTAGATCAACGGGT
XLOC_211989 F	ATTTGACCCCTCTGCACCTG
XLOC_211989 R	TGCCAGGATGTATGGGTTCTG
XLOC_215396 F	GGATGCACTTGTCAGGGTAGG
XLOC_215396 R	GGTACTTGTCCCTACACCCCC
XLOC_261766 F	ACATACCTTTCCGTAGAGCAGT
XLOC_261766 R	CAAAGGTCTCAAAGCGGTCC
XLOC_136653 F	ATGCCAAACTGTACCCTGCC
XLOC_136653 R	TGCTTTTGTGTCCAAGGGGT
DANCR F	GAAGTGCAGCTGCCTCAGTTCTTA
DANCR R	AATGGCTTGTGCCTGTAGTTGTC
NEAT1 F	CTTCTTCCCTTTAACTTATCCATTCAC
NEAT1 R	CTCTTCCTCCACCATTACCAACAATAC
KCNQ1OT1 F	GGCTACGACCACAGGTGAAA
KCNQ1OT1 R	GTCTGCTGGCTTGTGTGTTG
OIP5-AS1 F	TTTCCTTGACCTTTAGGTGCTTT
OIP5-AS1 R	GAAGCAGGACTACCCACTCTAGG

Primers RNAseq validation	**Sequence (5'-3')**
BNIP3 F	CGGGATGCAGGAGGAGAG
BNIP3 R	TAGAAACCGAGGCTGGAACG

BNIP3L F	CAGCAGGGACCATAGCTCTC
BNIP3L R	TGATACCCAGTCCGCACTTT
MAP1LC3B F	TTCAGGTTCACAAAACCCGC
MAP1LC3B R	TCTCACACAGCCCGTTTACC
GABARAP F	CGAAAGAAATACCCGGACCG
GABARAP R	GAGATCAGAAGGCACCAGGT
XBP1 F	CTGGAGCTATGGTGGTGGTG
XBP1 R	CCCCGACAGAAGCAGAACTT
TRIB3 F	CAGCGGATGCAGAGGAGAGA
TRIB3 R	GCCGTCTGATGCCCTCG
ATF4 F	GAAGCGATTTAACGAGCGCC
ATF4 R	ATCTTGGTTCCTGCCACGTT
ALDOC F	GGGCGCTTACCTTCTCCTAT
ALDOC R	ACTGCCTTCATACTTGCCCT
PGM1 F	TGGGAAAGCAGCAGTTTGAC
PGM1 R	CCCACAACTCCATGCATAGC
PFKB3 F	CACTTGCATTACCGTCCCTG
PFKB3 R	ACTCTTCCGACCTTCCCAAG
TPI1 F	GGGGCTTTTACTGGGGAGAT
TPI1 R	CCAATGCAGGCGATTACTCC
GPI F	AAGACAATAGTGGGGTGGGG
GPI R	TTGTCCAGGAATTCACCCGA
MPI F	AGGAAAGGGAAAGGGTAGGC
MPI R	TGTAGGGAGGGCTCTTTTGG
OPTN F	TTGGAGTGACTTTTCCACAGGA

OPTN R	GGGGCTGTCCTCCTTTTCAG
CHMP2A F	ATGGACCTATTGTTCGGGCG
CHMP2A R	TCTCTAGTTTCTGTCGCTCGC

Primer ChIP	**Sequence (5'-3')**
P6 MTAAT F	TACTGGGGGAAACAGCAAAC
P6 MTAAT R	TGTGGGACACAGCTAAAGCA
P3 MTAAT F	GCCTCTGCCATGCCCTAATA
P3 MTAAT R	GATTCTTTATCTGCTGTCTGTGC
P4 MTAAT F	TCTCACCATTGTTCAAACTGCT
P4 MTAAT R	GGGTCTTGAGCTATGTGACG
P5 MTAAT F	TCCTCTGAATGATCTCTGTCTGT
P5 MTAAT R	TTGCTGCTCTGACTGTTCCA
P1 BNIP3 F	CTCTCCTCCTGCATCCCG
P1 BNIP3 R	CTGCCCTGTGAGTTCCTCC
P2 BNIP3 F	TGCTAGTGGGGAAACTGAGG
P2 BNIP3 R	TCCCGAGACGCTCAGCTC
P3 BNIP3 F	TCCCGAGACGCTCAGCTC
P3 BNIP3 R	GACCCCGGTTCAGCTTCT
P4 BNIP3 F	ACCGCCAGAGATACATAGCA
P4 BNIP3 R	GGAGTCTTTCTGTGTTGCCA
P1 BNIP3L F	TCCCCAAGACACCTATTGCA
P1 BNIP3L R	GCAGTGGGAGAGAGGATGAG
P2 BNIP3L F	GCGAGGAAAATGAGCAGTCT
P2 BNIP3L R	GCCAATGAGCTGCCTTCTC

P3 BNIP3L F	TCGTCTAGGGTTGGCTTCAG
P3 BNIP3L R	ATCTTGGGTGGTTCAGGAGG
P1 ATF4 F	GCACTTGAGCCGGATGAAAA
P1 ATF4 R	TTTCCAGAGGCCCCATTCAT
P2 ATF4 F	CAGGCCACAAATCACCACC
P2 ATF4 R	GGTCGCTGCTAGTCCTCAG
P3 ATF4 F	CGTCCTCGGCCTTCACAATA
P3 ATF4 R	TCACGAAAGGAGAGAGGTGT
P1 XBP1 F	CCAAACCGAGAGCTTTCCAG
P1 XBP1 R	GTCTTTTCGAACCCAAGGCC
P2 XBP1 F	GTTTCAGGACCGTGGCTATG
P2 XBP1 R	TCAGTCTGGAAAGCTCTCGG
P1 OPTN F	GCCTGGCATTCTCCTCTTTC
P1 OPTN R	GCTCTAAGGCGTCACTGTGA
P2 OPTN F	ATGGGCGGGGTATGGGAT
P2 OPTN R	GTCACTGTTTCCTCGGCATC
P3 OPTN F	GCGGCCTGAAAACGGTAC
P3 OPTN R	ACGTCACCTCCAAGTCTCTG
GAPDH Exon1 F	AAGACCTTGGGCTGGGACT
GAPDH Exon1 R	GCTGCGGGCTCAATTTATAG
RUNX2 Upstream F	TCTCAAGGTGCCTGTCTGC
RUNX2 Upstream R	TGAAGTTTGGCCTCTGGTCT
D17Z1 (α-Sat) F	CTTTGGATGGAGCAGGTTTGAGAC
D17Z1 (α-Sat) R	CCGTTTAGTTAGGTGCAGTTATCC

Supplementary Table 2. List of sgRNA and GapmeRs sequences

sg_RNA	Primer Forward	Primer Reverse
sgRNA_MTAAT #2	CACCGTAATCTTCTGAGGTTGCTGA	AAACTCAGCAACCTCAGAAGATTAC
sgRNA_MTAAT #3	CACCGGCACCTACTGTGTATTTCCT	AAACTCAGCAACCTCAGAAGATTAC

GapmeR	Sequence	Cat.no. * Qiagen
Negative control	AACACGTCTATACGC	339515
lncRNA_MTAAT #1	TCATTAGCTAGGAGTA	339511
lncRNA_MTAAT #2	TCAATAAAGCGGGATC	339511

Supplementary Table 3. List of genes analysed by nCounter platform

Gene	Type
XLOC_043524	lncRNA
XLOC_066584	lncRNA
XLOC_080485	lncRNA
XLOC_098415	lncRNA
XLOC_115902	lncRNA
XLOC_163319	lncRNA

XLOC_169868	lncRNA
XLOC_136653	lncRNA
XLOC_169876	lncRNA
XLOC_177839	lncRNA
XLOC_211989	lncRNA
XLOC_215396	lncRNA
XLOC_261766	lncRNA
XLOC_286804	lncRNA
XLOC_330767	lncRNA
XLOC_330840	lncRNA
XLOC_334219	lncRNA
ERBB4	lncRNA
ALK	Coding gene
TMOD	Coding gene
BATF3	Coding gene
TNFRSF8	Coding gene
BNIP3	Coding gene
BNIP3L	Coding gene
CHMP2A	Coding gene/housekeeping
REEP5	Coding gene/housekeeping
EMC7	Coding gene/housekeeping

COG7	Coding gene/housekeeping
DNAJC14	Coding gene/housekeeping
EIF2B4	Coding gene/housekeeping
ERCC3	Coding gene/housekeeping
G6PD	Coding gene/housekeeping
GUSB	Coding gene/housekeeping
MRPS5	Coding gene/housekeeping
PPIA	Coding gene/housekeeping
MTMR14	Coding gene/housekeeping
HPRT1	Coding gene/housekeeping
SF3A3	Coding gene/housekeeping
TLK2	Coding gene/housekeeping

Supplementary Table 4. Expression level of genes used to classify ALCL patients

Sample	ALK traslocation	BATF3	TMOD1	TNFRSF8	Other translocation	ALCL Classification
lt2	+	-	-	-	-	ALK+
lt3	+	+	-	+	-	ALK+
lt4	+	+	+	+	-	ALK-
lt5	+	+	+	+	-	ALK-
lt6	-	+	+	+	-	ALK-
lt7	+	+	+	+	-	ALK+
lt8	-	+	-	+	-	ALK-
lt9	+	+	+	+	-	ALK+
lt10	-	-	-	-	-	ALK-
lt11	+	+	+	+	-	ALK-
lt12	+	+	+	+	-	ALK-
lt13	+	-	-	-	-	ALK+
lt14	-	+	-	+	-	ALK-
lt15	+	+	+	+	-	ALK-
lt17	-	+	+	-	-	ALK-
lt19	+	+	+	+	-	ALK+
lt20	-	-	-	-	-	ALK-
lt21	+	+	+	+	-	ALK+
lt22	+	+	+	+	-	ALK-
lt23	+	+	+	+	-	ALK+
lt24	-	-	-	-	+	ALK+
lt25	+	+	-	+	-	ALK+
lt26	+	+	+	+	-	ALK+
lt27	+	+	+	+	-	ALK-
lt28	+	+	+	+	-	ALK+
lt29	-	-	-	-	-	ALK-
lt30	-	-	+	+	-	ALK-
lt31	+	+	+	+	-	ALK-
lt32	+	+	+	+	-	ALK-
lt34	+	+	+	+	-	ALK-
lt35	+	+	+	+	-	ALK-
lt36	+	+	+	+	-	ALK-
lt37	+	+	+	+	-	ALK-
lt38	+	+	+	+	-	ALK-
lt39	+	+	-	+	-	ALK+
lt40	+	-	-	+	-	ALK+
lt42	-	-	+	+	-	ALK-
lt44	+	+	+	+	-	ALK-
lt45	-	-	+	-	-	ALK-
lt47	+	+	+	+	-	ALK+
lt48	+	+	+	+	-	ALK-
lt50	+	+	+	+	-	ALK-
lt51	+	+	+	+	-	ALK-
lt52	-	-	+	-	-	ALK-

Supplementary Table 5. In silico analysis of MTAAT open reading frames (ORFs)

Label	Strand	Frame	Start	Stop	Lenght (nt|aa)
ORF1	+	1	4012	4182	171|56
ORF2	+	3	2127	2285	159|52
ORF3	-	1	5266	5072	195|64
ORF4	-	1	2620	2429	192|63

Supplementary Table 6. TF binding sites prediction by FIMO algorithm

motif	sequence	start	stop	score	p.value	q.value	matched_sequence	database
IRF3	chr3:180332773-180333239	347	366	18,5372	2,04E-07	0,000171	GAAGAGGAAAGGAAAAGGGT	HOCOMOCO
ZKSC1	chr3:180332773-180333239	270	288	17,3438	6,69E-07	0,000589	TAAGCACCTACTGTGTATT	HOCOMOCO
ZFP82	chr3:180332773-180333239	347	370	16,3833	0,00000119	0,000984	TTCCACCCTTTTCCTTTCCTCTTC	HOCOMOCO
IRF3	chr3:180332773-	346	365	16,0744	0,00000155	0,000648	AGAAGAGGAAAGGAAAAGGG	HOCOMOCO

	180333239							
ZN467	chr3:18033 2773-180333239	347	368	13,734	0,000 00235	0,00 15	GAAGAGGAAAGG AAAAGGGTGG	HOCO MOCO
MAZ	chr3:18033 2773-180333239	361	382	14,55 08	0,000 00282	0,00 236	AAGGGTGGAACA GTCAGAGCAG	HOCO MOCO
NR1H3	chr3:18033 2773-180333239	165	183	14,84 56	0,000 00313	0,00 275	CTGAGGTTGCTGA AGGCCC	HOCO MOCO
FLI1	chr3:18033 2773-180333239	350	367	15,17 19	0,000 00316	0,00 267	GAGGAAAGGAAA AGGGTG	HOCO MOCO
ZN467	chr3:18033 2773-180333239	350	371	12,88 3	0,000 00358	0,00 15	GAGGAAAGGAAA AGGGTGGAAC	HOCO MOCO
ETS2	chr3:18033 2773-180333239	353	365	13,95 31	0,000 00376	0,00 324	GAAAGGAAAAGG G	HOCO MOCO
ETV5	chr3:18033 2773-180333239	352	365	14,04 76	0,000 00569	0,00 264	GGAAAGGAAAAG GG	HOCO MOCO
ZN394	chr3:18033 2773-180333239	352	371	13,36 62	0,000 00604	0,00 43	GGAAAGGAAAAG GGTGGAAC	HOCO MOCO

ETV5	chr3:18033 2773-180333239	347	360	13,95 24	0,000 0062	0,00 264	GAAGAGGAAAGG AA	HOCO MOCO
BC11A	chr3:18033 2773-180333239	346	362	13,89 17	0,000 00624	0,00 526	AGAAGAGGAAAG GAAAA	HOCO MOCO
CRX	chr3:18033 2773-180333239	226	238	13,87 5	0,000 00633	0,00 55	TCAGAGGATTAAG	HOCO MOCO
IRF1	chr3:18033 2773-180333239	70	89	14,09 38	0,000 00675	0,00 585	AAAAATGAAAATG AAAATGT	HOCO MOCO
OTX2	chr3:18033 2773-180333239	226	236	13,69 47	0,000 00678	0,00 593	AGAGGATTAAG	HOCO MOCO
KLF3	chr3:18033 2773-180333239	357	375	12,92 19	0,000 00718	0,00 615	GGAAAAGGGTGG AACAGTC	HOCO MOCO
ANDR	chr3:18033 2773-180333239	46	63	13,85 12	0,000 00723	0,00 631	AGTTCTGTCAAGT TTGCT	HOCO MOCO
ZFX	chr3:18033 2773-180333239	163	172	14,32 81	0,000 00851	0,00 771	GAAGGCCCCA	HOCO MOCO
ZN341	chr3:18033	353	374	13,07	0,000	0,00	GAAAGGAAAAGG	HOCO

	2773-180333239			69	00881	348	GTGGAACAGT	MOCO
ZN341	chr3:18033 2773-180333239	362	383	13,05 13	0,000 00894	0,00 348	AGGGTGGAACAG TCAGAGCAGC	HOCO MOCO
VEZF1	chr3:18033 2773-180333239	348	369	13,01 48	0,000 00902	0,00 482	AAGAGGAAAGGA AAAGGGTGGA	HOCO MOCO
CDX2	chr3:18033 2773-180333239	379	390	13,06 67	0,000 0091	0,00 828	CTTTATTGCTGC	HOCO MOCO
ZN394	chr3:18033 2773-180333239	347	366	12,81 69	0,000 0103	0,00 43	GAAGAGGAAAGG AAAAGGGT	HOCO MOCO
KLF15	chr3:18033 2773-180333239	350	368	11,25	0,000 0104	0,00 904	GAGGAAAGGAAA AGGGTGG	HOCO MOCO
SP4	chr3:18033 2773-180333239	356	375	12,75 21	0,000 0107	0,00 915	AGGAAAAGGGTG GAACAGTC	HOCO MOCO
MAZ	chr3:18033 2773-180333239	346	367	12,04 24	0,000 0115	0,00 479	AGAAGAGGAAAG GAAAAGGGTG	HOCO MOCO
VEZF1	chr3:18033 2773-	344	365	12,59 26	0,000 0116	0,00 482	GCAGAAGAGGAA AGGAAAAGGG	HOCO MOCO

	180333239							
ZN467	chr3:18033 2773-180333239	344	365	10,39 36	0,000 0117	0,00 324	GCAGAAGAGGAA AGGAAAAGGG	HOCO MOCO
ZN350	chr3:18033 2773-180333239	350	367	12,93 94	0,000 012	0,01 01	CACCCTTTTCCTT TCCTC	HOCO MOCO
IRF9	chr3:18033 2773-180333239	410	421	13,47 11	0,000 0121	0,01 1	AAAAGAGAAATT	HOCO MOCO
ZN341	chr3:18033 2773-180333239	343	364	12,46 15	0,000 0126	0,00 348	TGCAGAAGAGGA AAGGAAAAGG	HOCO MOCO
LYL1	chr3:18033 2773-180333239	158	171	12,46 15	0,000 0146	0,01 27	CCTTCTGGGGCC TT	HOCO MOCO
LEF1	chr3:18033 2773-180333239	346	359	12,15 15	0,000 0168	0,01 46	TCCTTTCCTCTTC T	HOCO MOCO
SPI1	chr3:18033 2773-180333239	345	361	12,42 19	0,000 0169	0,01 44	CAGAAGAGGAAA GGAAA	HOCO MOCO
VEZF1	chr3:18033 2773-180333239	343	364	11,71 11	0,000 0195	0,00 538	TGCAGAAGAGGA AAGGAAAAGG	HOCO MOCO

ZN257	chr3:180332773-180333239	349	360	12,5156	0,0000208	0,0182	TTCCTTTCCTCT	HOCOMOCO
IRF2	chr3:180332773-180333239	70	89	11,3906	0,0000209	0,0181	AAAAATGAAAATGAAAATGT	HOCOMOCO
FLI1	chr3:180332773-180333239	345	362	11,8906	0,0000214	0,00903	CAGAAGAGGAAAGGAAAA	HOCOMOCO
FOXJ3	chr3:180332773-180333239	103	115	13,0469	0,0000214	0,0188	ATGTTTTTGTTTG	HOCOMOCO
BC11A	chr3:180332773-180333239	423	439	12,2667	0,0000234	0,00988	GATACAGGAACTCAGAA	HOCOMOCO
STAT6	chr3:180332773-180333239	261	271	12,5619	0,0000235	0,0206	TGCCCAGGAAA	HOCOMOCO
PRDM1	chr3:180332773-180333239	353	366	12,4375	0,0000239	0,0207	GAAAGGAAAAGGGT	HOCOMOCO
ERG	chr3:180332773-180333239	347	359	12,5781	0,0000239	0,0169	GAAGAGGAAAGGA	HOCOMOCO
PATZ1	chr3:18033	350	371	10,84	0,000	0,02	GAGGAAAGGAAA	HOCO

	2773-180333239			03	0256	16	AGGGTGGAAC	MOCO
ZN140	chr3:18033 2773-180333239	360	383	9	0,000 0262	0,02 27	GCTGCTCTGACTG TTCCACCCTTT	HOCO MOCO
MAZ	chr3:18033 2773-180333239	349	370	10,41 53	0,000 0269	0,00 748	AGAGGAAAGGAA AAGGGTGGAA	HOCO MOCO
SPIB	chr3:18033 2773-180333239	345	361	10,62 5	0,000 0273	0,01 49	CAGAAGAGGAAA GGAAA	HOCO MOCO
ETV4	chr3:18033 2773-180333239	348	358	12,39 06	0,000 0274	0,02 39	AAGAGGAAAGG	HOCO MOCO
ETV2	chr3:18033 2773-180333239	350	365	11,54 41	0,000 028	0,02 44	GAGGAAAGGAAA AGGG	HOCO MOCO
IRF4	chr3:18033 2773-180333239	422	439	11,35 79	0,000 0294	0,01 36	GATACAGGAACTC AGAAT	HOCO MOCO
ZN341	chr3:18033 2773-180333239	344	365	10,82 91	0,000 0312	0,00 647	GCAGAAGAGGAA AGGAAAAGGG	HOCO MOCO
BATF3	chr3:18033 2773-	415	431	12,04 62	0,000 0312	0,02 77	CTCTTTTATTCTG AGTT	HOCO MOCO

	180333239							
ZIC3	chr3:18033 2773-180333239	273	287	11,02 74	0,000 0314	0,02 8	AAGCACCTACTGT GT	HOCO MOCO
IRF4	chr3:18033 2773-180333239	346	363	11,18 95	0,000 032	0,01 36	AGAAGAGGAAAG GAAAAG	HOCO MOCO
COE1	chr3:18033 2773-180333239	155	169	11,04 69	0,000 0336	0,02 94	AATCCTTCTGGGG CC	HOCO MOCO
KLF6	chr3:18033 2773-180333239	359	377	10,92 19	0,000 0337	0,02 95	AAAAGGGTGGAA CAGTCAG	HOCO MOCO
SPIB	chr3:18033 2773-180333239	424	440	10,03 12	0,000 0348	0,01 49	TGATACAGGAACT CAGA	HOCO MOCO
STAT2	chr3:18033 2773-180333239	68	86	11,17 19	0,000 0366	0,02 53	TTAAAAATGAAAA TGAAAA	HOCO MOCO
STF1	chr3:18033 2773-180333239	165	175	11,75 24	0,000 0372	0,03 33	GCTGAAGGCCC	HOCO MOCO
ZBT17	chr3:18033 2773-180333239	355	373	10,93 4	0,000 0384	0,01 14	AAGGAAAAGGGT GGAACAG	HOCO MOCO

ZBT17	chr3:180332773-180333239	350	368	10,8774	0,0000396	0,0114	GAGGAAAGGAAAAGGGTGG	HOCOMOCO
ERG	chr3:180332773-180333239	352	364	11,7031	0,000004	0,0169	GGAAAGGAAAAGG	HOCOMOCO
ETS1	chr3:180332773-180333239	352	364	11,5625	0,0000405	0,0351	GGAAAGGAAAAGG	HOCOMOCO
ZBT17	chr3:180332773-180333239	344	362	10,8302	0,0000407	0,0114	GCAGAAGAGGAAAGGAAAA	HOCOMOCO
MAZ	chr3:180332773-180333239	344	365	9,50847	0,0000423	0,00884	GCAGAAGAGGAAAGGAAAAGGG	HOCOMOCO
ZN502	chr3:180332773-180333239	353	372	2,90411	0,0000425	0,0373	GAAAGGAAAAGGGTGGAACA	HOCOMOCO
NR4A2	chr3:180332773-180333239	114	122	11,9195	0,0000432	0,0394	AAAGGTCAT	HOCOMOCO
SMCA5	chr3:180332773-180333239	38	52	11,339	0,0000433	0,0379	GAAGAGAGAGCAAAC	HOCOMOCO
NR4A1	chr3:18033	114	122	11,15	0,000	0,03	AAAGGTCAT	HOCO

	2773-180333239			33	0448	92		MOCO
HXC9	chr3:180332773-180333239	381	390	11,6355	0,0000458	0,0416	CTTTATTGCT	HOCOMOCO
ZN467	chr3:180332773-180333239	351	372	7,24468	0,0000461	0,00914	AGGAAAGGAAAAGGGTGGAACA	HOCOMOCO
BATF	chr3:180332773-180333239	97	114	11,5231	0,0000462	0,0411	TGTTTTTGTTTGACTAAT	HOCOMOCO
VEZF1	chr3:180332773-180333239	349	370	10,1704	0,0000462	0,00789	AGAGGAAAGGAAAAGGGTGGAA	HOCOMOCO
RXRG	chr3:180332773-180333239	114	126	8,79688	0,0000469	0,0419	CTTAAAAGGTCAT	HOCOMOCO
VEZF1	chr3:180332773-180333239	358	379	10,1185	0,0000475	0,00789	GAAAAGGGTGGAACAGTCAGAG	HOCOMOCO
ZN350	chr3:180332773-180333239	388	405	11,2626	0,0000481	0,0203	GAGTTCATTTCTTACCTT	HOCOMOCO
LEF1	chr3:180332773-	351	364	11,1212	0,0000498	0,0217	CCTTTTCCTTTCCT	HOCOMOCO

	180333239							
ZN143	chr3:18033 2773-180333239	139	160	7,151 16	0,000 05	0,04 45	AGGATTTGGGTAG ATGTATTTC	HOCO MOCO
PRDM6	chr3:18033 2773-180333239	353	365	11,54 69	0,000 0513	0,04 13	GAAAGGAAAAGG G	HOCO MOCO
ZN274	chr3:18033 2773-180333239	32	51	9,865 55	0,000 0515	0,02 89	GCCACTGAAGAG AGAGCAAA	HOCO MOCO
IRF1	chr3:18033 2773-180333239	408	427	10,21 88	0,000 0522	0,01 54	CAGAATAAAAGAG AAATTAT	HOCO MOCO
IRF1	chr3:18033 2773-180333239	347	366	10,17 19	0,000 0534	0,01 54	GAAGAGGAAAGG AAAAGGGT	HOCO MOCO
SPI1	chr3:18033 2773-180333239	424	440	9,953 12	0,000 0545	0,02 32	TGATACAGGAACT CAGA	HOCO MOCO
ZN467	chr3:18033 2773-180333239	362	383	6,829 79	0,000 0548	0,00 914	AGGGTGGAACAG TCAGAGCAGC	HOCO MOCO
BC11A	chr3:18033 2773-180333239	351	367	11,10 83	0,000 0553	0,01 55	AGGAAAGGAAAA GGGTG	HOCO MOCO

NR6A1	chr3:180332773-180333239	109	121	9,37209	0,0000562	0,0499	AAGGTCATGTTTT	HOCOMOCO
NR1I2	chr3:180332773-180333239	398	416	10,6977	0,0000588	0,052	AGAAATTATAAGAGTTCAT	HOCOMOCO
IRF3	chr3:180332773-180333239	70	89	10,562	0,0000589	0,0165	AAAAATGAAAATGAAAATGT	HOCOMOCO
RFX1	chr3:180332773-180333239	163	184	7,57812	0,0000592	0,0511	TCTGAGGTTGCTGAAGGCCCCA	HOCOMOCO
STAT2	chr3:180332773-180333239	350	368	10,2031	0,0000592	0,0253	GAGGAAAGGAAAAGGGTGG	HOCOMOCO
ELF5	chr3:180332773-180333239	345	359	10,8906	0,0000597	0,0515	CAGAAGAGGAAAGGA	HOCOMOCO
PATZ1	chr3:180332773-180333239	362	383	9,13445	0,0000623	0,0263	AGGGTGGAACAGTCAGAGCAGC	HOCOMOCO
ZN816	chr3:180332773-180333239	359	379	8,07812	0,0000628	0,0324	AAAAGGGTGGAACAGTCAGAG	HOCOMOCO
ZN121	chr3:18033	162	181	5,095	0,000	0,05	CTGGGGCCTTCA	HOCO

	2773-180333239			24	0632	53	GCAACCTC	MOCO
ZN263	chr3:180332773-180333239	366	385	8,4375	0,0000657	0,0566	TGGAACAGTCAGAGCAGCAA	HOCOMOCO
ZN335	chr3:180332773-180333239	354	375	8,67568	0,0000657	0,056	GACTGTTCCACCCTTTTCCTTT	HOCOMOCO
ZN490	chr3:180332773-180333239	332	355	-8,625	0,0000665	0,0571	TTCCTCTTCTGCAGGCAAAAGATA	HOCOMOCO
ZN274	chr3:180332773-180333239	341	360	9,36975	0,0000665	0,0289	CCTGCAGAAGAGGAAAGGAA	HOCOMOCO
NFAC1	chr3:180332773-180333239	350	364	11,0762	0,0000689	0,0591	GAGGAAAGGAAAAGG	HOCOMOCO
ZN341	chr3:180332773-180333239	359	380	9,26496	0,0000706	0,0117	AAAAGGGTGGAACAGTCAGAGC	HOCOMOCO
ETS2	chr3:180332773-180333239	348	360	10,9844	0,0000708	0,0305	AAGAGGAAAGGAA	HOCOMOCO
ETV4	chr3:180332773-	263	273	11,1562	0,0000715	0,0312	CCCAGGAAATA	HOCOMOCO

	180333239							
NR2C1	chr3:18033 2773-180333239	114	126	10,04 26	0,000 0721	0,06 33	CTTAAAAGGTCAT	HOCO MOCO
IRF2	chr3:18033 2773-180333239	408	427	8,515 62	0,000 0735	0,02 7	CAGAATAAAAGAG AAATTAT	HOCO MOCO
PAX6	chr3:18033 2773-180333239	1	12	10,85 95	0,000 0738	0,06 6	TCTCACTTGAGT	HOCO MOCO
COT2	chr3:18033 2773-180333239	111	123	10,84 38	0,000 0749	0,06 67	AAAAGGTCATGTT	HOCO MOCO
AP2A	chr3:18033 2773-180333239	164	178	9,656 25	0,000 0753	0,06 64	GGGGCCTTCAGC AAC	HOCO MOCO
ZN816	chr3:18033 2773-180333239	358	378	7,625	0,000 0758	0,03 24	GAAAAGGGTGGA ACAGTCAGA	HOCO MOCO
KLF15	chr3:18033 2773-180333239	362	380	6,609 38	0,000 0761	0,02 22	AGGGTGGAACAG TCAGAGC	HOCO MOCO
KLF15	chr3:18033 2773-180333239	351	369	6,593 75	0,000 0766	0,02 22	AGGAAAGGAAAA GGGTGGA	HOCO MOCO

GCR	chr3:18033 2773-180333239	367	381	10,4688	0,0000779	0,0679	TGCTCTGACTGTTCC	HOCOMOCO
ZN214	chr3:18033 2773-180333239	221	242	10,36	0,0000783	0,0687	CAGATCTTAATCCTCTGAATGA	HOCOMOCO
SMAD3	chr3:18033 2773-180333239	36	47	10,678	0,0000806	0,0718	CTCTCTCTTCAG	HOCOMOCO
PTF1A	chr3:18033 2773-180333239	158	175	9,70476	0,0000812	0,0701	CCTTCTGGGGCCTTCAGC	HOCOMOCO
PRDM1	chr3:18033 2773-180333239	347	360	10,2344	0,0000815	0,028	GAAGAGGAAAGGAA	HOCOMOCO
IRF7	chr3:18033 2773-180333239	353	362	10,9111	0,000082	0,0398	GAAAGGAAAA	HOCOMOCO
ETV2	chr3:18033 2773-180333239	260	275	8,98529	0,0000826	0,036	GTGCCCAGGAAATACA	HOCOMOCO
IRF3	chr3:18033 2773-180333239	414	433	9,90083	0,0000854	0,0179	GGAACTCAGAATAAAAGAGA	HOCOMOCO
ZNF76	chr3:18033	154	175	7,282	0,000	0,07	AAATCCTTCTGGG	HOCO

	2773-180333239			05	0859	58	GCCTTCAGC	MOCO
BCL6	chr3:18033 2773-180333239	197	209	10,4769	0,0000864	0,0772	TGTTGTCTAGGGA	HOCO MOCO
HSF1	chr3:18033 2773-180333239	160	174	9,75	0,0000873	0,0772	CTGAAGGCCCCA GAA	HOCO MOCO
NR5A2	chr3:18033 2773-180333239	165	175	10,1406	0,0000876	0,0787	GCTGAAGGCCC	HOCO MOCO
IRF7	chr3:18033 2773-180333239	76	85	10,8444	0,0000883	0,0398	GAAAATGAAA	HOCO MOCO
STAT2	chr3:18033 2773-180333239	74	92	9,29688	0,0000908	0,0259	ATGAAAATGAAAA TGTACT	HOCO MOCO
MAZ	chr3:18033 2773-180333239	350	371	7,92373	0,0000909	0,0152	GAGGAAAGGAAA AGGGTGGAAC	HOCO MOCO
KLF5	chr3:18033 2773-180333239	362	375	8,95312	0,000092	0,0822	AGGGTGGAACAG TC	HOCO MOCO
ZN467	chr3:18033 2773-	352	373	5,54255	0,0000923	0,0128	GGAAAGGAAAAG GGTGGAACAG	HOCO MOCO

	180333239							
PRDM6	chr3:18033 2773-180333239	348	360	10,875	0,0000929	0,0413	AAGAGGAAAGGAA	HOCOMOCO
IRF2	chr3:18033 2773-180333239	347	366	7,90625	0,000094	0,027	GAAGAGGAAAGGAAAAGGGT	HOCOMOCO
IRF8	chr3:18033 2773-180333239	420	439	8,60938	0,0000967	0,0421	GATACAGGAACTCAGAATAA	HOCOMOCO
PRDM1	chr3:18033 2773-180333239	352	365	9,89062	0,0000968	0,028	GGAAAGGAAAAGGG	HOCOMOCO
ELF3	chr3:18033 2773-180333239	346	359	10,0312	0,0000971	0,084	AGAAGAGGAAAGGA	HOCOMOCO
IRF8	chr3:18033 2773-180333239	70	89	8,5625	0,0000987	0,0421	AAAAATGAAAATGAAAATGT	HOCOMOCO
ZFP28	chr3:18033 2773-180333239	342	361	7,3125	0,0000994	0,0422	TTTCCTTTCCTCTTCTGCAG	HOCOMOCO
ZFP28	chr3:18033 2773-180333239	413	432	7,3125	0,0000994	0,0422	TTCTCTTTTATTCTGAGTTC	HOCOMOCO

ZNF263	chr3:18033 2773-180333239	349	369	13,5	0,000 0054	0,00 458	AGAGGAAAGGAA AAGGGTGGA	JASPA R
HIC2	chr3:18033 2773-180333239	260	268	13,67 92	0,000 00814	0,00 726	GTGCCCAGG	JASPA R
IRF1	chr3:18033 2773-180333239	71	91	13,7	0,000 00837	0,00 715	GTACATTTTCATTT TCATTTT	JASPA R
SP2	chr3:18033 2773-180333239	357	371	12,41 07	0,000 0108	0,00 916	GTTCCACCCTTTT CC	JASPA R
DMRT3	chr3:18033 2773-180333239	432	442	12,78 33	0,000 0116	0,01 03	CCTGTATCAAC	JASPA R
CRX	chr3:18033 2773-180333239	227	237	13,58 93	0,000 0136	0,01 2	CAGAGGATTAA	JASPA R
ZFX	chr3:18033 2773-180333239	158	171	12,23 44	0,000 0145	0,01 25	CCTTCTGGGGCC TT	JASPA R
IRF7	chr3:18033 2773-180333239	74	87	12,45	0,000 0148	0,01 31	ATGAAAATGAAAA T	JASPA R
ZNF263	chr3:18033	346	366	11,60	0,000	0,00	AGAAGAGGAAAG	JASPA

	2773-180333239			42	015	634	GAAAAGGGT	R
IRF3	chr3:18033 2773-180333239	345	365	7,547 95	0,000 0155	0,01 09	CAGAAGAGGAAA GGAAAAGGG	JASPA R
NR4A1	chr3:18033 2773-180333239	114	123	13,88 24	0,000 0165	0,01 5	AAAAGGTCAT	JASPA R
ZNF24	chr3:18033 2773-180333239	453	465	12,24 76	0,000 0167	0,01 48	AATTCATTTATTC	JASPA R
NR2F6(VAR.2)	chr3:18033 2773-180333239	115	129	- 0,205 882	0,000 0175	0,01 56	CTGCTTAAAAGGT CA	JASPA R
NR2F2	chr3:18033 2773-180333239	113	123	13,35 9	0,000 0196	0,01 77	AAAAGGTCATG	JASPA R
GSC2	chr3:18033 2773-180333239	226	235	11,60 34	0,000 0211	0,01 86	CTTAATCCTC	JASPA R
CDX2	chr3:18033 2773-180333239	380	390	13,17 74	0,000 0215	0,01 95	CAGCAATAAAG	JASPA R
GATA1: :TAL1	chr3:18033 2773-	330	347	12,69 81	0,000 0216	0,01 91	CATATCTTTTGCC TGCAG	JASPA R

	180333239							
IRF3	chr3:18033 2773-180333239	68	88	6,191 78	0,000 0249	0,01 09	TTAAAAATGAAAA TGAAAATG	JASPA R
BCL6	chr3:18033 2773-180333239	196	209	12,33 33	0,000 026	0,02 33	GTCCCTAGACAAC A	JASPA R
KLF5	chr3:18033 2773-180333239	362	371	11,18 75	0,000 0301	0,02 65	GTTCCACCCT	JASPA R
SP1	chr3:18033 2773-180333239	361	371	10,33 33	0,000 0332	0,02 92	GTTCCACCCTT	JASPA R
IRF1	chr3:18033 2773-180333239	348	368	10,85	0,000 0348	0,01 49	CCACCCTTTTCCT TTCCTCTT	JASPA R
IRF9	chr3:18033 2773-180333239	73	87	0,660 377	0,000 0359	0,03 17	AATGAAAATGAAA AT	JASPA R
ZNF263	chr3:18033 2773-180333239	350	370	9,645 83	0,000 0399	0,01 13	GAGGAAAGGAAA AGGGTGGAA	JASPA R
IRF8	chr3:18033 2773-180333239	74	87	- 0,056 6038	0,000 0435	0,03 92	ATGAAAATGAAAA T	JASPA R

GSC	chr3:180332773-180333239	226	235	11,5	0,0000475	0,0431	CTTAATCCTC	JASPAR
NR1H2::RXRA	chr3:180332773-180333239	113	129	-4,82979	0,0000506	0,044	CTGCTTAAAAGGTCATG	JASPAR
CDX1	chr3:180332773-180333239	382	390	10,8302	0,0000507	0,0464	GCAATAAAG	JASPAR
ZIC1	chr3:180332773-180333239	274	287	3,73913	0,0000532	0,0464	AAGCACCTACTGTG	JASPAR
FOXH1	chr3:180332773-180333239	267	277	11,5882	0,0000534	0,0474	GGAAATACACA	JASPAR
ELF1	chr3:180332773-180333239	262	273	8,22642	0,0000574	0,0513	GCCCAGGAAATA	JASPAR
ZNF263	chr3:180332773-180333239	376	396	8,66667	0,0000634	0,0134	AGAGCAGCAATAAAGGTAAGA	JASPAR
RORA	chr3:180332773-180333239	115	124	12,0319	0,0000663	0,0599	TAAAAGGTCA	JASPAR
ESRRA	chr3:18033	114	124	6,647	0,000	0,06	TAAAAGGTCAT	JASPA

	2773-180333239			06	0742	75		R
TCF7	chr3:18033 2773-180333239	347	358	9,333 33	0,000 0751	0,06 62	GAAGAGGAAAGG	JASPA R
MEIS2	chr3:18033 2773-180333239	53	60	11,12 5	0,000 0769	0,06 93	TTGACAGA	JASPA R
KLF1	chr3:18033 2773-180333239	362	372	8,859 38	0,000 0774	0,06 93	TGTTCCACCCT	JASPA R
IRF4	chr3:18033 2773-180333239	74	88	0,156 25	0,000 0776	0,06 99	ATGAAAATGAAAA TG	JASPA R
POU2F 2	chr3:18033 2773-180333239	72	84	10,69 64	0,000 0799	0,07	TTCATTTTCATTT	JASPA R
NR2F1	chr3:18033 2773-180333239	111	123	10,14 58	0,000 0806	0,07 31	AAAAGGTCATGTT	JASPA R
VDR	chr3:18033 2773-180333239	399	406	10,66 42	0,000 0815	0,07 46	AGAGTTCA	JASPA R
SMAD4	chr3:18033 2773-	199	206	12,13 33	0,000 0817	0,07 41	TGTCTAGG	JASPA R

	180333239							
IRF2	chr3:18033 2773-180333239	75	92	4,701 92	0,000 0818	0,07 29	TGAAAATGAAAAT GTACT	JASPA R
IRF1	chr3:18033 2773-180333239	406	426	8,883 33	0,000 0818	0,02 33	TTATAATTTCTCTT TTATTCT	JASPA R
RXRB	chr3:18033 2773-180333239	115	128	- 0,924 528	0,000 0821	0,07 29	TGCTTAAAAGGTC A	JASPA R
GFI1B	chr3:18033 2773-180333239	175	185	10,67 86	0,000 0844	0,07 35	CAACCTCAGAA	JASPA R
HLTF	chr3:18033 2773-180333239	357	366	9,093 02	0,000 0859	0,07 76	ACCCTTTTCC	JASPA R
ELK4	chr3:18033 2773-180333239	264	274	9,490 57	0,000 0862	0,07 62	GTATTTCCTGG	JASPA R
PITX3	chr3:18033 2773-180333239	226	234	10,81 48	0,000 0862	0,07 9	CTTAATCCT	JASPA R
MEIS1	chr3:18033 2773-180333239	53	59	10,45 45	0,000 0863	0,07 93	TTGACAG	JASPA R

SPIB	chr3:18033 2773- 180333239	349	355	11,41 86	0,000 0863	0,07 77	AGAGGAA	JASPA R
ELF4	chr3:18033 2773- 180333239	262	273	7,2	0,000 0926	0,08 29	GCCCAGGAAATA	JASPA R
EHF	chr3:18033 2773- 180333239	262	273	8,816 67	0,000 0945	0,08 44	GCCCAGGAAATA	JASPA R

Figure S1

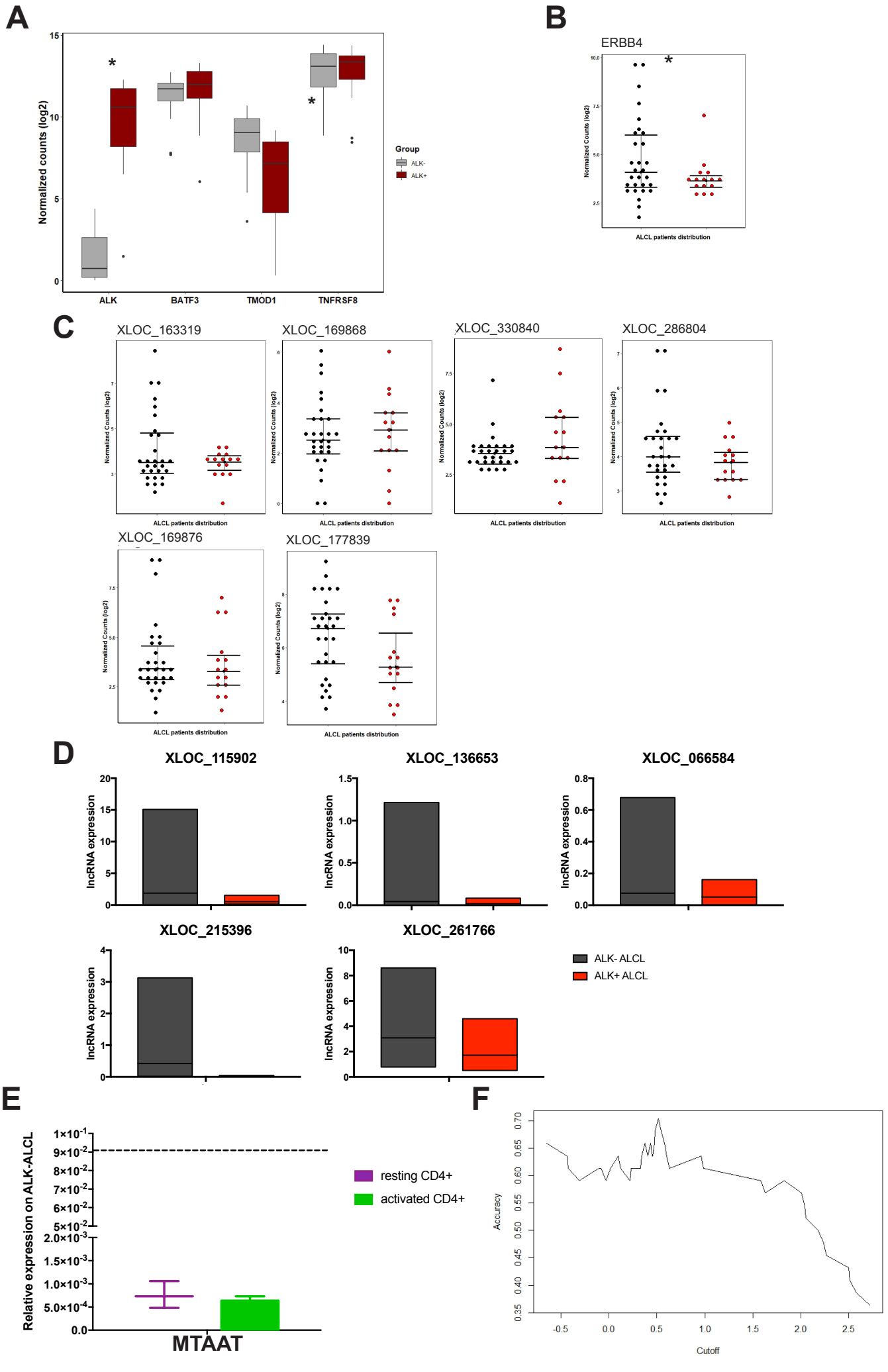

A
Normalized counts (log2)
Group
ALK-
ALK+
ALK
BATF3
TMOD1
TNFRSF8
B
ERBB4
Normalized Counts (log2)
ALCL patients distribution
C
XLOC_163319
XLOC_169868
XLOC_330840
XLOC_286804
XLOC_169876
XLOC_177839
D
XLOC_115902
XLOC_136653
XLOC_066584
XLOC_215396
XLOC_261766
lncRNA expression
ALK- ALCL
ALK+ ALCL
E
Relative expression on ALK-ALCL
resting CD4+
activated CD4+
MTAAT
F
Accuracy
Cutoff

Figure S2

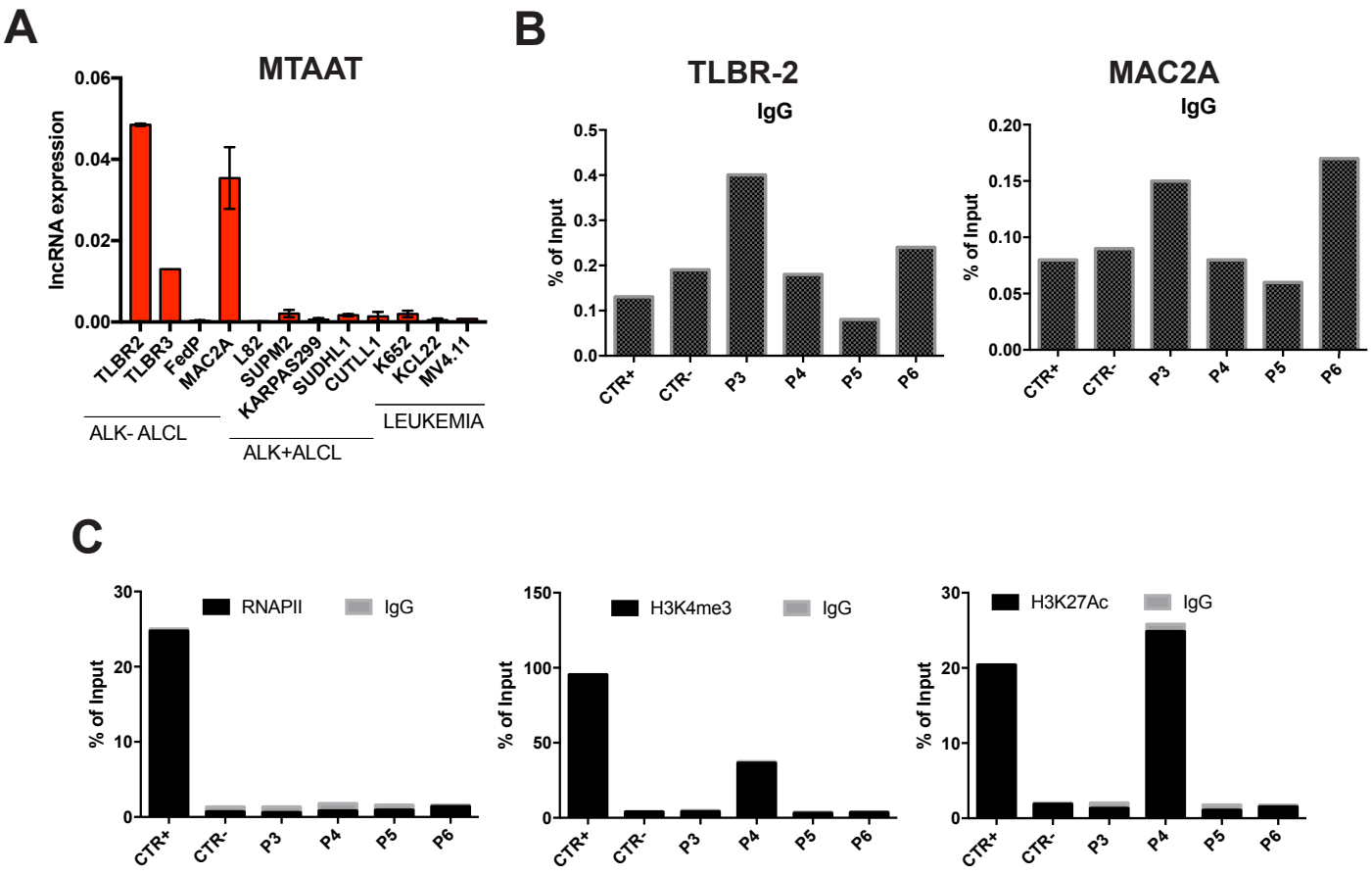
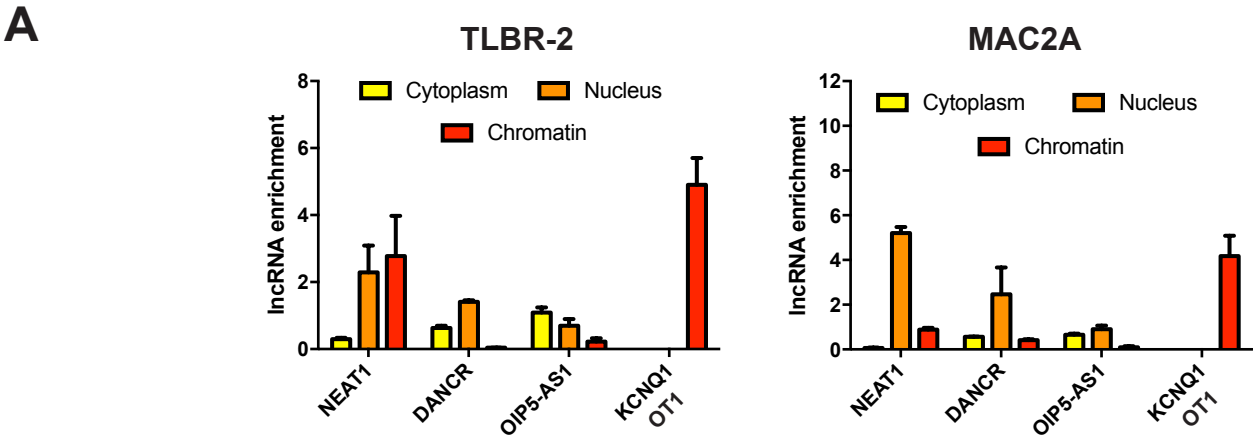
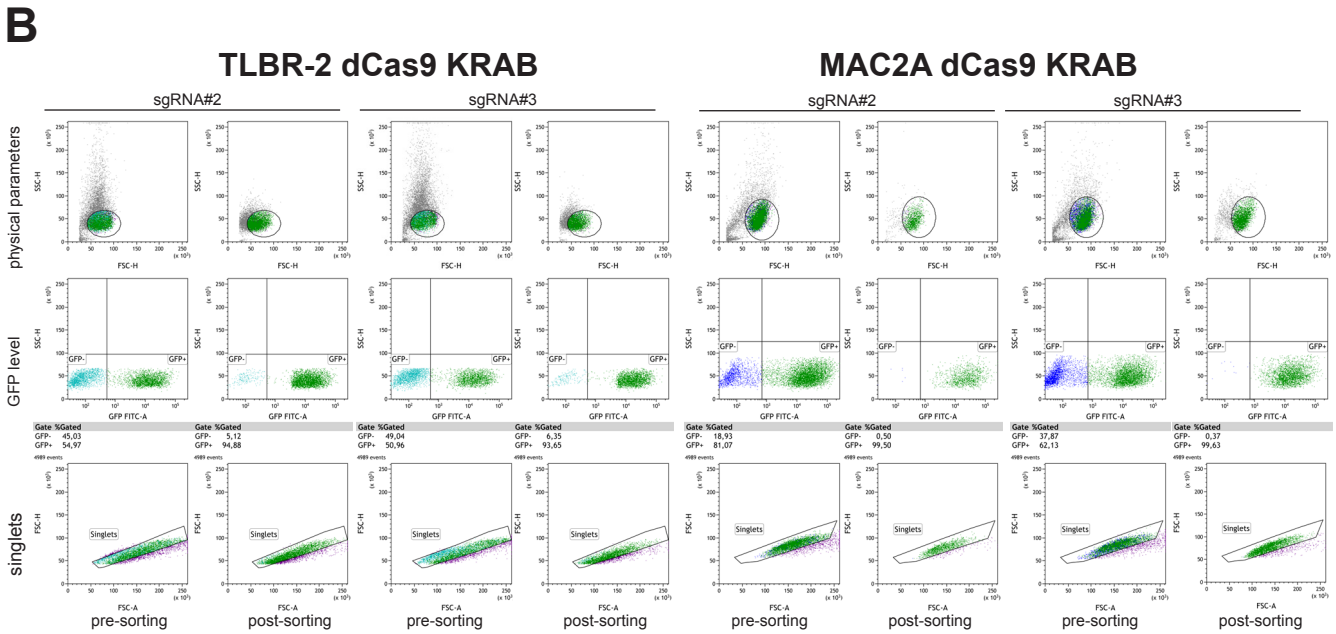
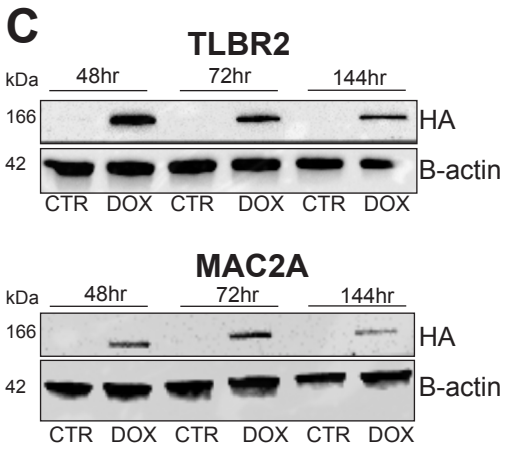
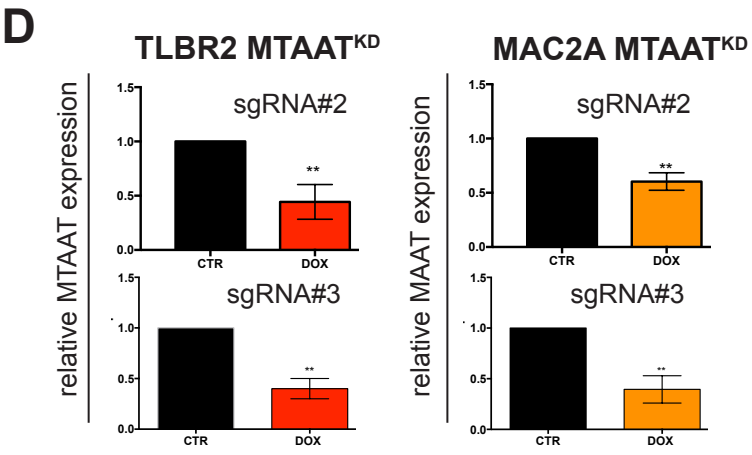

A
MTAAT
lncRNA expression
0.06
0.04
0.02
0.00
TLBR2
TLBR3
FedP
MAC2A
L82
SUPM2
KARPAS299
SUDHL1
CUTLL1
K652
KCL22
MV4.11
ALK- ALCL
ALK+ALCL
LEUKEMIA
B
TLBR-2
IgG
% of Input
0.5
0.4
0.3
0.2
0.1
0.0
CTR+
CTR-
P3
P4
P5
P6
MAC2A
IgG
% of Input
0.20
0.15
0.10
0.05
0.00
C
RNAPII
IgG
% of Input
30
20
10
0
H3K4me3
IgG
% of Input
150
100
50
0
H3K27Ac
IgG
% of Input
30
20
10
0

Figure S3


A
TLBR-2
MAC2A
Cytoplasm
Nucleus
Chromatin
lncRNA enrichment
NEAT1
DANCR
OIP5-AS1
KCNQ1 OT1


B
TLBR-2 dCas9 KRAB
MAC2A dCas9 KRAB
sgRNA#2
sgRNA#3
physical parameters
GFP level
singlets
pre-sorting
post-sorting


C
TLBR2
MAC2A
kDa
48hr
72hr
144hr
166
42
HA
B-actin
CTR
DOX


D
TLBR2 MTAAT^KD
MAC2A MTAAT^KD
relative MTAAT expression
relative MAAT expression
sgRNA#2
sgRNA#3
CTR
DOX

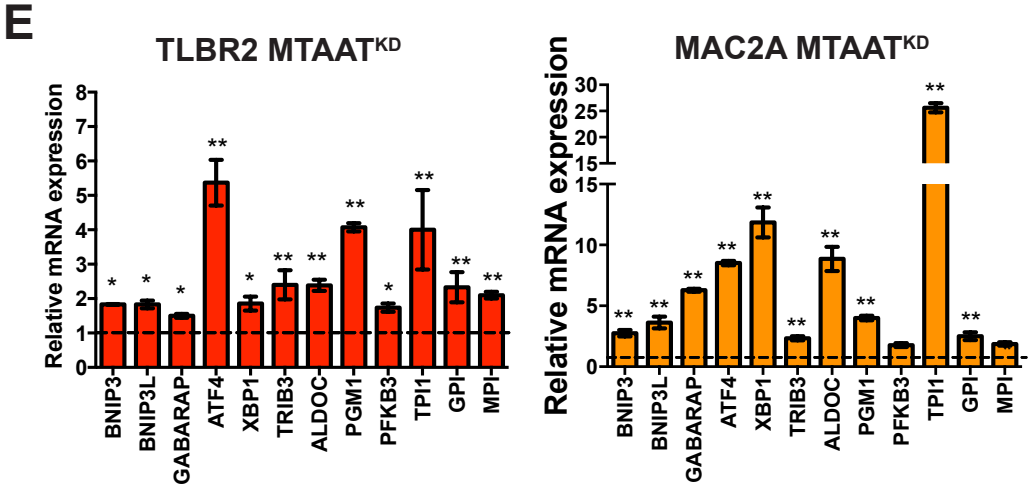

E
TLBR2 MTAAT^KD
MAC2A MTAAT^KD
Relative mRNA expression
BNIP3
BNIP3L
GABARAP
ATF4
XBP1
TRIB3
ALDOC
PGM1
PFKB3
TPI1
GPI
MPI

Figure S4

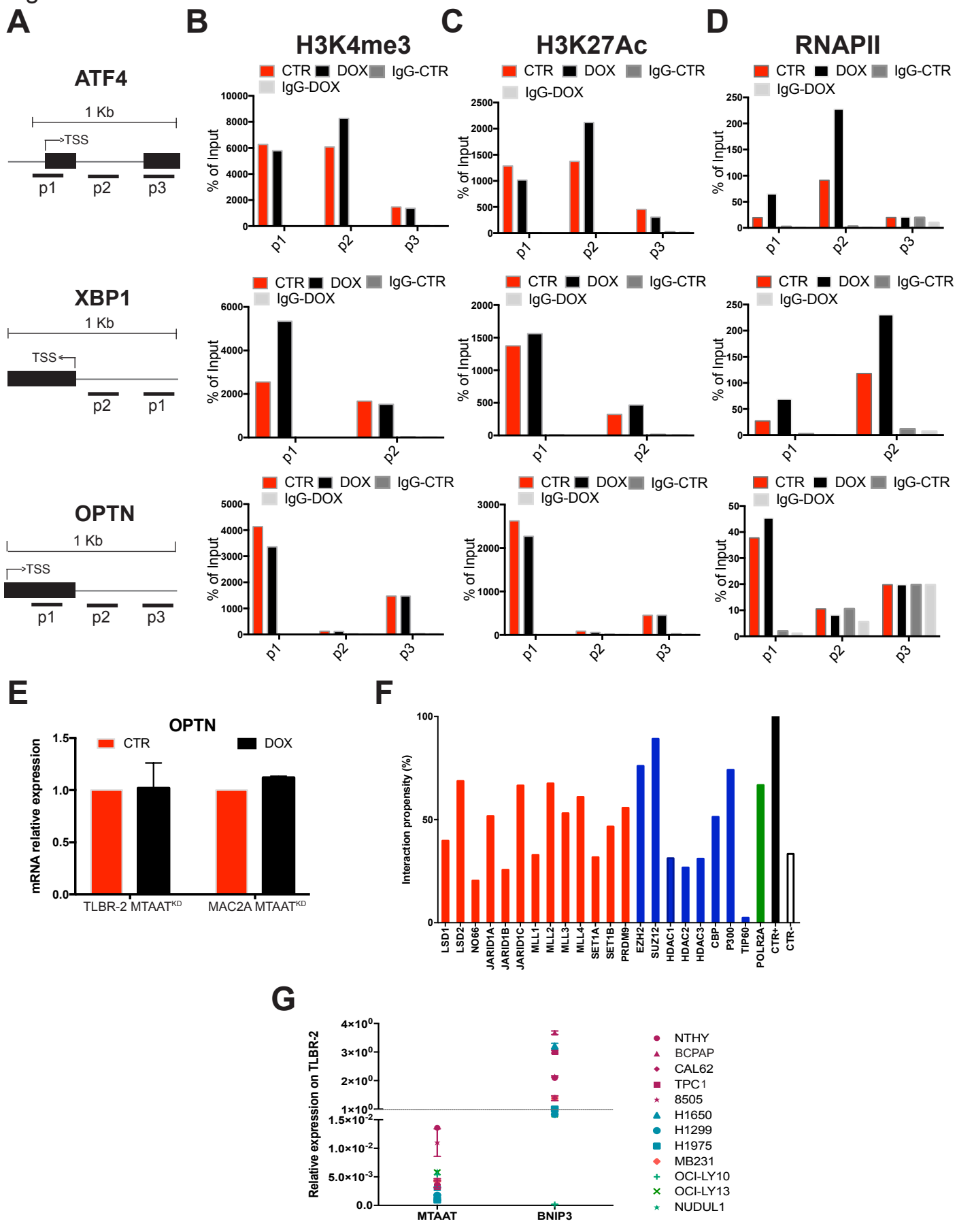
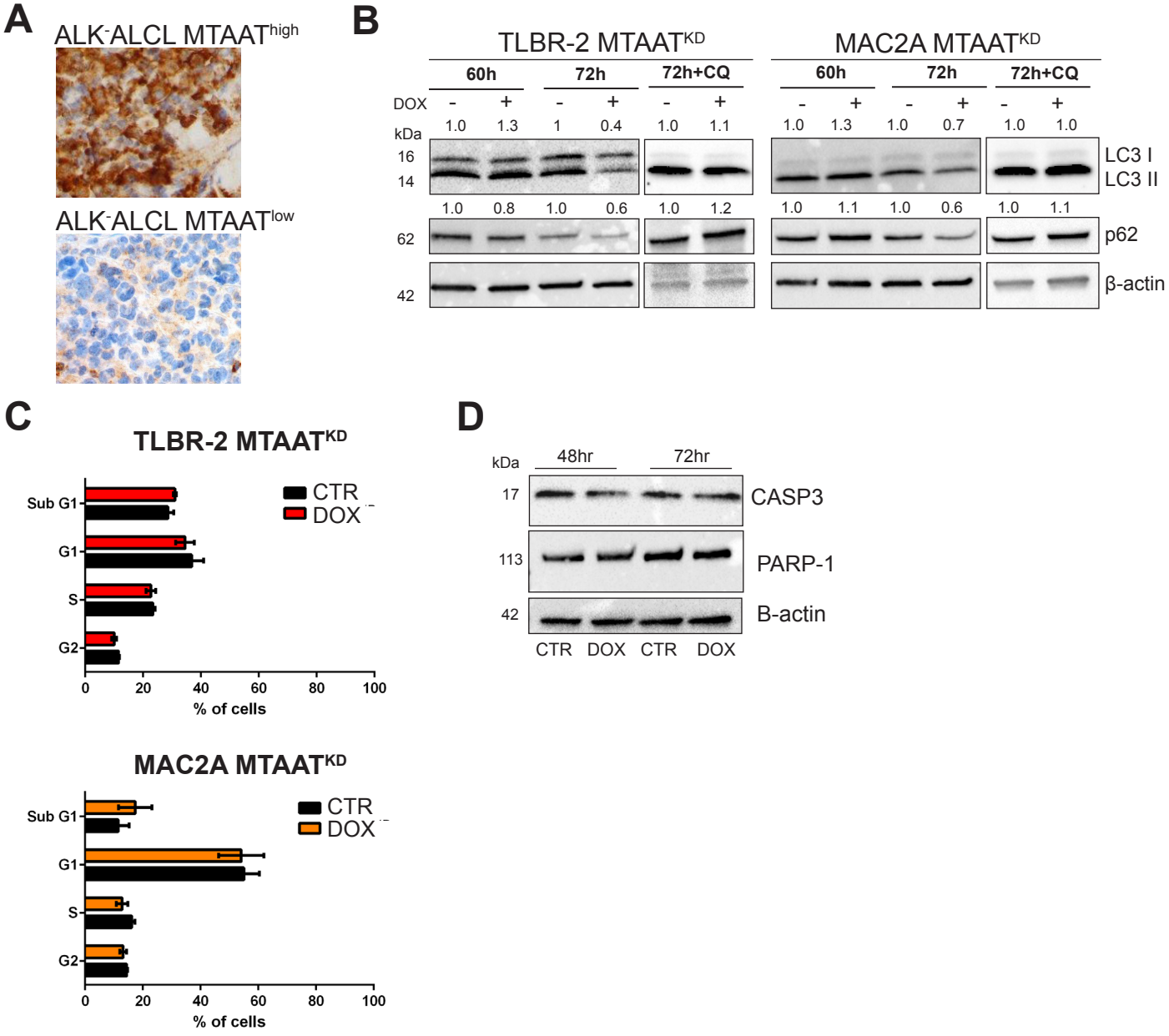

A
B
C
D
H3K4me3
H3K27Ac
RNAPII
ATF4
1 Kb
TSS
p1
p2
p3
XBP1
OPTN
CTR
DOX
IgG-CTR
IgG-DOX
% of Input
E
OPTN
mRNA relative expression
TLBR-2 MTAAT^KD
MAC2A MTAAT^KD
F
Interaction propensity (%)
LSD1
LSD2
NO66
JARID1A
JARID1B
JARID1C
MLL1
MLL2
MLL3
MLL4
SET1A
SET1B
PRDM9
EZH2
SUZ12
HDAC1
HDAC2
HDAC3
CBP
P300
TIP60
POLR2A
CTR+
CTR-
G
Relative expression on TLBR-2
MTAAT
BNIP3
NTHY
BCPAP
CAL62
TPC1
8505
H1650
H1299
H1975
MB231
OCI-LY10
OCI-LY13
NUDUL1

Figure S5


A
ALK-ALCL MTAAThigh
ALK-ALCL MTAATlow
B
TLBR-2 MTAATKD
MAC2A MTAATKD
60h
72h
72h+CQ
DOX
kDa
16
14
62
42
LC3 I
LC3 II
p62
β-actin
C
TLBR-2 MTAATKD
Sub G1
G1
S
G2
CTR
DOX
% of cells
MAC2A MTAATKD
D
48hr
72hr
17
113
CASP3
PARP-1
B-actin
CTR
DOX