Regulatory mechanisms and context-dependent roles of TAL1 in T-cell acute lymphoblastic leukemia

Jolynn Zu Lin Ong,[1] Tze King Tan,[1] Lu Wang,[1] Shi Hao Tan[1] and Takaomi Sanda[1,2,3]

[1]Cancer Science Institute of Singapore, National University of Singapore, Singapore; [2]Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore and [3]Department of Hematology and Oncology, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan

Correspondence: T. Sanda
takaomi_sanda@u.nus.edu

Received: April 30, 2023.
Accepted: October 10, 2023.
Early view: October 19, 2023.

https://doi.org/10.3324/haematol.2023.283450

Supplemental Information

Supplemental Methods

Cell samples

Human T-ALL cell lines were cultured in RPMI-1640 medium (BioWest) supplemented with 10% fetal bovine serum (FBS)(BioWest). 293T cells were cultured in DMEM medium supplemented with 10% FBS, penicillin and streptomycin. All cells were cultured in a humidified incubator containing 5% CO_2 at 37°C.

Reagents

Antibodies specific to the following proteins were used for immunoblotting: HA-Tag (Cell Signaling Technology), TAL1 (Millipore), LMO1 (Bethyl), MYC (Cell Signaling Technology), cleaved notch (Cell Signaling Technology), CDK4 (Cell Signaling Technology) and GAPDH-HRP (Santa Cruz). Dimethyl sulfoxide (DMSO) and doxycycline were purchased from Sigma–Aldrich. The dTAG-13 molecule was purchased from Tocris Bioscience.

dTAG cloning and cell line establishment

The degradation tag (dTAG) system was established using plasmids cloned following methods described previously in detail (1). Briefly, endogenous CRISPR-mediated knock-in of the $FKBP12^{F36V}$ domain depended on the microhomology-mediated PITCh (Precise Integration into Target Chromosome) vector system (2), where microhomology arms (20 base pairs) and single-guide RNA (sgRNAs) specific to the C-terminus of the TAL1 coding domain were cloned into pCRIS-PITChv2 (Addgene #91796) and px330A (Addgene #91794) plasmids, respectively. The sgRNA oligonucleotides and PCR primers used for dTAG cloning are listed in Supplementary

Table 5. To establish the dTAG system in the Jurkat cell line, 3 µg of pCRIS-PITChv2 and 6 µg of px330A plasmids were transiently transfected into 1 million cells via electroporation (Neon system, Invitrogen). The cells were selected with puromycin 48 hours post-transfection. Successful endogenous knock-in was validated by genomic DNA PCR followed by Sanger sequencing. The primers used for genomic PCR validation are listed in Supplemental Table 1.

Gene overexpression

Full-length *TAL1* and *LMO1* cDNA sequences were cloned into a doxycycline-inducible pCW57.1 lentiviral vector (Addgene) for overexpression in HPB-ALL. Full-length *TAL1* and *LMO1* cDNA were cloned into MSCV-hygromycin and MSCV-neomycin vectors (Addgene) respectively for overexpression in DND-41 and KOPT-K1. MYC and corresponding empty vector plasmids were purchased from VectorBuilder. Cells were transduced with lentivirus produced by cotransfecting 293T cells with pMDLg/pRRE, pRSV-Rev, and pMD2.G using FuGENE6 (Promega) and Opti-MEM (Thermo Fisher) reagents.

Cell cycle analysis

To induce doxycycline-inducible TAL1/LMO1 overexpression, a total of 500,000 cells were induced with either DMSO or 1 µM doxycycline and harvested, rinsed with cold 1x PBS and fixed with 70% ethanol at –20 °C overnight. The fixed cells were rinsed twice with cold 1x PBS and incubated with RNAse A (Qiagen, 10 mg/ml) at 37 °C for 30 mins. The cells were then stained with propidium iodide (Sigma, 50 mg/ml) and analyzed using an LSRII flow cytometer (BD Bioscience). Cell cycle phases were determined using FlowJo software.

Apoptosis Assay

Cells were rinsed with cold 1xPBS and staining buffer, incubated with an APC-conjugated anti-Annexin V antibody (BioLegend) and propidium iodide (Sigma), and analyzed using an LSRII flow cytometer (BD Biosciences). The percentage of Annexin V-positive cells was determined using FlowJo software.

Western blot analysis

Cell pellets were lysed in RIPA buffer (Cell Signaling Technology) supplemented with a protease inhibitor cocktail (Roche). Proteins were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and transferred onto a polyvinylidene difluoride membrane (Bio–Rad Laboratories). Bound antibodies were visualized using ECL Prime Western blotting reagents (GE Healthcare, Amersham) with an ImageQuant LAS500 gel imager (GE Healthcare).

Genomic DNA extraction and polymerase chain reaction

Genomic DNA was harvested from Jurkat cells after endogenous TAL1-FKBP12^{F36V} knock-in using a QIAamp DNA Blood Mini Kit (Qiagen). PCR was performed using primers flanking the C-terminus of the TAL1 CDS. The primer sequences used for PCR are listed in Supplemental Table 1.

RNA sequencing (RNA-seq)

Sample preparation — Total RNA was extracted from cell samples using QIAzol lysis reagent (Qiagen) and cleaned using an RNeasy Kit (Qiagen). RNA samples were treated with TuRBO DNase (Ambion). Strand-specific library construction and

sequencing of 100-bp paired-end reads on the DNBSEQ platform were performed by BGI Biotech Solutions Co., Ltd.

Data analysis — RNA-seq reads were aligned to the hg19 human reference genome using STAR 2.5.2a with outFilterMultimapNmax set to 1. The total number of mapped reads was quantified using featureCounts version 2.0.1, and count tables were generated based on Ensembl hg19 gene annotation gtf files. Differential expression analysis was conducted using the Bioconductor package DESeq2 version 1.12.4. Gene expression in each sample was estimated in transcripts per million (TPM) using Kallisto software version 0.43.1. All RNA-Seq data were normalized using the Sleuth R package from the Pachter laboratory (3). Detailed information on dataset availability can be found in Supplemental Table 3.

4C sequencing (4C-seq)

Sample preparation — 4C-seq was performed according to a protocol described previously in detail (4). Briefly, a total of 10 million cells were harvested, rinsed with warm FBS-free medium twice and crosslinked in a 1% formaldehyde solution for 10 minutes at room temperature. Glycine (2.5 M) was then added immediately to quench the formaldehyde reaction. The cells were washed twice in ice-cold 1x PBS before being snap frozen in liquid nitrogen and stored at –80 °C. Crosslinked cells were resuspended in cell lysis buffer and subjected to two rounds of enzymatic digestion, ligation and purification steps as follows: first, restriction enzyme digestion was performed with by N1aIII (NEB) followed by overnight ligation with 50 U of T4 ligase (NEB) and subsequent reverse crosslinking at 65 °C using 10 mg/ml of Proteinase K (Promega). DNA fragments were purified using Nucleomag P-beads (Macherey-Nagel) to yield the 3C template. The 3C template was then subjected to a second restriction

enzyme digestion by DpnII (NEB), overnight ligation and sample purification were repeated to yield the 4C template. Library preparation involved two PCR steps as follows: first, PCR with inverse primers was performed to amplify the fragments ligated to the sequence of interest, followed by a second PCR to incorporate sequencing adapters. The primer sequences used for 4C-seq are listed in Supplemental Table 4.

Data analysis — 4C-Seq reads were aligned to human indexed genome hg19 using bwa aligner with default parameter. The aligned reads were processed using r3CSeq R package (5). The forward and reverse primer sequence listed in Supplemental Table 4 were utilized to get the viewpoint-of-interest and with the getReadCountPerRestrictionFragment and getInteractions functions in r3Cseq package, the interactions from viewpoint were quantified and a heatmap were plot using plotDomainogramNearViewpoint function where the interactions intensity were visualized. Detailed information on dataset availability can be found in Supplemental Table 3.

Cut-and-run assay

Sample preparation — Antibodies used for cut-and-run included an α-GATA3 antibody (Invitrogen) and an α-RUNX1 antibody (Abcam). Cut-and-Run-seq was performed utilizing a Cutana Cut-and-Run assay kit (EpiCypher). Briefly, a total of 500,000 cells were washed in buffer containing 1% digitonin and bound to 10 µl of activated Concanavalin A beads. Bead-bound cells were incubated with α-GATA3 (1 µg/µl) or α-RUNX1 (1 µg/µl) antibodies overnight and washed with cell permeabilization buffer the following day. After washing, the cells were incubated with CUTANA pA/G-MNase (EpiCypher), and targeted chromatin digestion was initiated by the addition of 100 mM CaCl2 and allowed to proceed for 2 hours at 4 °C. Released

chromatin fragments were purified utilizing a Cutana DNA purification kit (EpiCypher). Libraries were generated using a NEBNext Ultra II DNA Library Prep Kit and sequenced on an Illumina NovaSeq at GENEWIZ Azenta Life Sciences Co. Ltd. (Suzhou, China).

Data analysis — CUT&RUN-seq data were preprocessed and analyzed using the cutRunTools pipeline (6). This pipeline automated the alignment of CUT&RUN-seq data to the hg19 genome. CutRunTools used MACS2 and SEACR software to identify peaks in the data using default optimized cutoffs. The paired control data were then used by the cutRunTools step2 script to normalize the MACS2 and SEACR output peaks. Each bedGraph output was further normalized by subtracting the treatment file from the corresponding control file using MACS2 bdgcmp and floored to 0.1 before being converted to bigWig format using the UCSC bedGraphToBigWig tool. Detailed information on dataset availability can be found in Supplemental Table 3.

Chromatin immunoprecipitation (ChIP)-seq and ChIP–qPCR

Sample preparation — For ChIP-Seq, a total of 10 million TAL1-FKBP12^{F36V} Jurkat cells were harvested, washed twice in FBS-free RPMI-1640 medium, and crosslinked in a 1% formaldehyde solution for 10 minutes at room temperature. Glycine (2.5 M) was then added immediately to quench the formaldehyde reaction. The cells were washed twice in ice-cold 1x PBS before being snap frozen in liquid nitrogen and stored at –80 °C. The cells were resuspended in lysis buffer as previously described and sonicated with a Bioruptor sonicator (Diagenode) for 50 cycles (1 cycle = 30 sec on + 30 sec off). Antibodies were added to the sonicated lysate and incubated overnight at 4 °C on a rotating platform to promote binding. Magnetic Dynabeads (Sigma) preblocked with 0.5% BSA (w/v) in PBS were added the next day and incubated for at

least 8 hours to promote binding. The beads were washed with sonication buffer as previously described and eluted in ChIP elution buffer at 65 °C. Library construction from the ChIP samples and sequencing of single-end 50-bp-long reads by DNB-seq were performed at BGI Biotech Solutions Co., Ltd. (Hong Kong, China). Antibodies used for ChIP-seq included an α-HA (CST) antibody. For ChIP–qPCR, qPCR primers targeting specific genomic loci were designed to measure the enrichment of transcription factors occupying DNA sites. Antibodies used for ChIP–qPCR included an α-GATA3 (Invitrogen) antibody. The qPCR primers are listed in Supplemental Table 5.

Data analysis — All ChIP-seq reads were mapped to the hg19 human reference genome using Bowtie2 ver 2.4.2 with --no-unal --very-sensitive -q --phred33 --end-to-end parameters. Duplicates were removed from aligned ChIP-seq reads using the Picard MarkDuplicate package with --REMOVE_DUPLICATES true. ChIP-seq peaks were called with MACS2 software version 2.2.7.1 with --keep-dup=1 --SPMR -p 1e-9 -f BAM. The output bedGraph data were normalized by subtracting the corresponding background values using MACS2 bdgcmp -m subtracat. BedGraph signal were further floored to 0.1, then convert to bigWig format using UCSC bedGraphToBigWig tool. Detailed information on dataset availability can be found in Supplemental Table 3.

Hi-ChIP

Sample preparation — Antibodies used for Hi-ChIP included an α-H3K27ac antibody (Abcam). Hi-ChIP was performed using an Arima Hi-C+ assay kit (A160168 EAP, Arima Genomics). Briefly, a total of 3 million cells were harvested, rinsed with 1x PBS at room temperature and crosslinked with 2% formaldehyde for exactly 10 minutes. Stop buffer was added immediately to quench the formaldehyde reaction. Crosslinked

cells were digested with a restriction enzyme cocktail. Next, DNA 5'-overhangs were filled with a biotinylated nucleotide. Spatially proximal digested ends of DNA were ligated to capture the chromatin structure. The proximally ligated chromatin was then sonicated with a Bioruptor sonicator (Diagenode) for 20 cycles (1 cycle = 30 sec on + 30 sec off) and incubated with 2.5 µg of α-H3K27ac antibody (Abcam) for subsequent immunoprecipitation. DNA was purified to yield fragmented proximally ligated DNA bound to the H3K27ac histone mark. The proximally ligated DNA was then subjected to a custom library preparation protocol utilizing a Swift Biosciences Accel-NGS 2S Plus DNA Library Kit (A160169 EAP, Arima Genomics). The library was sequenced on the DNBSEQ performed at BGI Biotech Solutions Co. Ltd. (Hong Kong, China).

Data analysis – Hi-ChIP data were processed using the HiC-Pro pipeline version 2.11.1. The reads in the fastq files were aligned to the human reference genome (GRCh37/hg19) using Bowtie2 version 2.4.1. HiC-Pro used the restriction sites ^GATC and G^ANTC, which are present across the human genome, to cut and process the data using the HiC-Pro Utilities Python script "digest genome". The HiC-Pro configuration was set to optimize for ARIMA kit data using the ^GATC and G^ANTC ligation sites. The aligned data were further analyzed using hichipper to identify interactions. Hichipper version 2.7.9 was used to extract valid interactions from the HiC-Pro data and generate a bedpe interaction file by utilizing anchors from peaks called by MACS2 broadPeak with the --keep-dup all -p 1e-9 --extsize 147 --nomodel -f bed parameters. Hichipper was also used to process interactions with the -l 50 -mi 5000 -ma 2000000 parameters. The output bedpe files were converted to bigInteract tracks and visualized in the UCSC Genome Browser. Detailed information on dataset availability can be found in Supplemental Table 3.

ATAC sequencing

Sample preparation — ATAC-seq was performed according to protocols described previously in detail (7). Briefly, a total of 300,000 cells were harvested, rinsed once with cold 1x PBS and resuspended in cold lysis buffer. The cells were incubated with Tn5 transposase (NEB, Nextera kit) at 37 °C for 45 minutes. Tagmented DNA fragments were purified with a DNA purification kit (Zymo Research) and subjected to 10 cycles of PCR amplification with NEBNext High Fidelity 2x master mix and Illumina-compatible barcoded primers (Nextera). The amplified library was purified with AMPure XP beads (Beckman Coulter) and sequenced on an Illumina NovaSeq at GENEWIZ Azenta Life Sciences Co., Ltd. (Suzhou, China).

Data analysis – Paired-end ATAC-seq reads were aligned to the human reference genome (hg19) using Bowtie2 with -k 10 --very-sensitive parameters. Duplicate reads were removed using the Picard MarkDuplicates tool with the --REMOVE_DUPLICATES true option. Non-uniquely aligned reads with MAPQ<10 were then filtered out using Samtools view -q 10. MACS2 version 2.2.7.1 was used to identify peaks in the data with the –keep-dup=all -f BAMPE -p 1e-5 -B and --SPMR parameters. The signal tracks were further processed using MACS2 bdgcmp in subtract mode to subtract the background signal and floor the resulting values to 0.1. The signal tracks were then converted to BigWig format using the UCSC bedGraphToBigWig script and visualized in the UCSC genome browser. Detailed information on dataset availability can be found in Supplemental Table 3.

Gene Ontology (GO) and pathway analyses

All genes that were significantly downregulated (absolute log2-fold change ≤ -0.5) in the HPB-ALL cell line after TAL1/LMO1 overexpression were imported into Enrichr

(http://amp.pharm.mssm.edu/Enrichr/) to analyze molecular signature database enrichment (MSigDB hallmark 2020). The category terms were sorted based on the combination of the p value (Fisher's exact test) and z score, which were computed by assessing the deviation from the expected rank for each term in the gene set library.

Supplemental Tables

Supplemental Table 1. List of PCR primers used in this study for TAL1-FKBP^{36v} endogenous knock-in and genomic validation; the data are presented in Excel file format.

Supplemental Table 2. List of qRT–PCR primers used in this study presented in Excel file format.

Supplemental Table 3. Details on dataset availability presented in Excel file format.

Supplemental Table 4. List of 4C-seq primers used in this study presented in Excel file format.

Supplemental Table 5. List of ChIP–qPCR primers used in this study presented in Excel file format.

Supplemental Table 6. Top 25 downregulated genes in Groups A and B, and Top 25 upregulated genes in Group C. Gene expression changes after dTAG treatment at the 1, 2, 4, 6, 8, 16, 24, 48, 72-hour time points. Group A genes expression gradually decreased; the thresholds were set as a log2-fold change ≤ -0.7 and p ≤ 0.05 at the 72-hour time point. Group B genes were downregulated and plateaued; the thresholds were set as a -0.7 ≤ log2- fold change ≤ -0.3 and p ≤ 0.05 at the 72-hour timepoint. Group C gene expression gradually increased; the thresholds were set as a log2-fold change ≥ 0.3 and p ≤ 0.05 at the 72-hour timepoint. An RNA-seq analysis was

performed for two control and two dTAG-treated samples at each timepoint. Gene expression changes were analyzed with DESeq2 version 1.12.4.

Supplemental Table 7. RNA-seq gene expression of Group A genes in Jurkat and HPB-ALL cell lines. These data were available in the CCLE database and presented in Excel file format.

Supplemental Table 8. RNA-seq expression of genes depicted in Venn diagram. Upregulated genes (a log2-fold change ≥ 0.5 and $p \leq 0.05$) in the HPB-ALL cell line after TAL1, LMO1 and TAL1/LMO1 were overexpressed, as measured at the 48-hour time point, and downregulated genes (a log2-fold change ≤ -0.3 and $p \leq 0.05$) in the TAL1-FKBP12^{F36V} Jurkat cell line after 1 μM dTAG treatment, as measured at the 48-hour timepoint; the data are presented in Excel file format.

Supplemental Table 9. RNA-seq expression of genes depicted in heatmap. Representative Group A and B genes that were significantly upregulated, representative Group C genes, MYC and CDK4 that were significantly downregulated after TAL1/LMO1 were overexpressed in the HPB-ALL cell line, measured at the 48-hour time point; the data are presented in Excel file format.

Supplemental References:

1. Nabet B, Roberts JM, Buckley DL, et al. The dTAG system for immediate and target-specific protein degradation. Nat Chem Biol. 2018 May;14(5):431-441.
2. Nakade S, Tsubota T, Sakane Y, et al. Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. Nat Commun. 2014 Nov 20;5(1):5560.
3. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods. 2017 Jul;14(7):687-690.
4. Krijger PHL, Geeven G, Bianchi V, Hilvering CRE, de Laat W. 4C-seq from beginning to end: A detailed protocol for sample preparation and data analysis. Methods. 2020 Jan 1;170:17-32.
5. Thongjuea S, Stadhouders R, Grosveld FG, Soler E, Lenhard B. r3Cseq: an R/Bioconductor package for the discovery of long-range genomic interactions from chromosome conformation capture and next-generation sequencing data. Nucleic Acids Res. 2013 Jul;41(13):e132.
6. Zhu Q, Liu N, Orkin SH, Yuan G-C. CUT&RUNTools: a flexible pipeline for CUT&RUN processing and footprint analysis. Genome biology. 2019;20(1):1-12.
7. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr Protoc Mol Biol. 2015 Jan 5;109(1):21 29 21-21 29 29.
8. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet. 2016 Oct;48(10):1193-1203.
9. Sanda T, Lawton LN, Barrasa MI, et al. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. Cancer Cell. 2012 Aug 14;22(2):209-221.
10. Seita J, Sahoo D, Rossi DJ, et al. Gene Expression Commons: an open platform for absolute gene expression profiling. PloS one. 2012;7(7):e40321.

Supplemental Figure Legends

Supplemental Fig. 1. TAL1 differentially regulates gene expression in human T-ALL cells. (A) Schematic diagram depicting the genomic knock-in of the FKBP12^{F36V} domain into the C-terminus of the TAL1 locus. The red box highlights the 3' end of the TAL1 CDS. PCR primers were designed to amplify this region to confirm that the endogenous FKBP12^{F36V} knock-in was in frame (Sanger sequencing results are shown). **(B)** Relative cell viability (%) of TAL1-FKBP12^{F36V} Jurkat cells treated with either 1 μM DMSO or dTAG-13 and measured 2 and 5 days later using a CellTiter Glo luminescence assay kit; n=3. Error bars are the mean ± SEM; ns: not significant as determined by Student's two-tailed t test compared to DMSO-treated cells. **(C)** TAL1-FKBP12^{F36V} Jurkat cells treated with 1 μM dTAG-13 for 8 or 24 hours. Whole-cell lysates were collected and subjected to immunoblot analysis using antibodies specific to TAL1, total RNA polymerase II, phosphorylated RNA polymerase II at Serine 5 (S5), phosphorylated RNA polymerase II at Serine 2 (S2) or β-actin (the loading control). **(D)** Relative mRNA expression levels of representative Group A genes (*ALDH1A2*, *TSPAN7*, and *SIX6*), Group B genes (*RUNX1*, *GIMAP2*, and *MYB*) and Group C genes (*RAG1*, *RORC*, and *SH3TC1*) in the TAL1-FKBP12^{F36V} Jurkat cell line after 1 μM dTAG treatment at different time points; n=2. These genes are depicted in the heatmap shown in Fig. 1c.

Supplemental Fig. 2. TAL1 induces the expression of a set of genes that are normally silenced in T-cells. (A) ATAC-seq tracks of hematopoietic stem cells (HSCs), multipotent progenitors (MPPs), common lymphoid progenitors (CLPs) and CD4+ T cells; the *SIX6* locus is magnified. The ATAC-seq dataset was previously reported(8). **(B)** ChIP-seq tracks of the TAL1, GATA3, and RUNX1 transcription

factors and the activating histone mark H3K27ac (top) in the Jurkat cell line; the *SIX6* loci are magnified. The ChIP-seq dataset was previously reported (9). H3K27ac Hi-ChIP was performed after DMSO treatment and dTAG treatment to evaluate changes in active chromatin loop formation (bottom). For **(A)** and **(B)**, yellow highlighting indicates the genomic region of interest: where the TAL1 transcription factor binds in the baseline Jurkat cell line. For the *SIX6* locus, the TSS and the –165-kb putative enhancer region are highlighted. **(C)** Relative gene expression of the *Aldh1a2*, *Pi16* and *Six6* genes in mouse hematopoietic cell populations measured by probe microarray, which was available from the Gene Expression Commons database (10). **(D)** ATAC-seq tracks of hematopoietic stem cells (HSCs), multipotent progenitors (MPPs), common lymphoid progenitors (CLPs) and CD4+ T cells; the *CD34, MYCN* and *ZNF521* loci are magnified. **(E)** Relative gene expression of *Cd34*, *Mycn* and *Zfp521* (the mouse homolog of human *ZNF521)* genes in mouse hematopoietic cell populations measured via probe microarray, available from the Gene Expression Commons database (10). Yellow highlighting indicates the genomic region of interest: where the TAL1 transcription factor binds in the baseline Jurkat cell line. For the CD34 locus, the TSS, +7-kb and +38-kb putative enhancer regions are highlighted. For the *MYCN* locus, the TSS and the +540-kb putative enhancer region are highlighted. For the *ZNF521* locus, the TSS is highlighted. **(F)** mRNA expression of representative Group A genes (*ALDH1A2*, TSPAN7, *PI16*, *SIX6* and *NKX3-1*) in a panel of TAL1-positive (JURKAT, RPMI-8402, CCRF-CEM, and MOLT4) and TAL1-negative (HPB-ALL, TALL-1, and DND-41) T-ALL cell lines measured by RNA-seq. **(G)** H3K27ac Hi-ChIP (top) and 4C-seq (bottom) performed after DMSO treatment and dTAG treatment; the *SIX6* locus is magnified. Green highlighting indicates the loss of chromatin interactions as determined by both H3K27ac Hi-ChIP and 4C-seq after dTAG

treatment. Yellow highlighting denotes the –165-kb putative enhancer site where the TAL1 transcription factor binds; this region was the basis for the 4C-seq primer design. A magnified view of the 4C-seq tracks near the SIX6 gene is also shown.

Supplemental Fig. 3. The TAL1 complex alters chromatin accessibility in T-ALL cells. (A) (Left) ATAC-seq tracks of the *SIX6* locus in basal-state HPB-ALL cells (black) and TAL1/LMO1 overexpression in the HPB-ALL cells, displayed as two independent clones (green). (Right) *SIX6* mRNA expression in HPB-ALL cells overexpressing TAL1/LMO1 under the Dox-inducible system as measured by RNA-seq. TPM: transcripts per million. **(B)** (Left) ATAC-seq tracks of the *CD34, MYCN* and *ZNF521* loci in basal-state HPB-ALL cells (black) and TAL1/LMO1 overexpression in the HPB-ALL, displayed as two independent clones (green). (Right) *CD34, MYCN* and *ZNF521* mRNA expression in HPB-ALL cells overexpressing TAL1/LMO1 under the Dox-inducible system as measured by RNA-seq. TPM: transcripts per million. In **(A)** and **(B)**, yellow highlighting indicates the genomic region of interest: where the TAL1 transcription factor binds in the baseline Jurkat cell line. For the *SIX6* locus, the –165-kb putative enhancer region is highlighted. For the *CD34* locus, the +7-kb and +37-kb regions are highlighted. For the *MYCN* locus, the +540-kb putative enhancer region is highlighted. For the *ZNF521* locus, the TSS, +113-kb and +130-kb putative enhancer sites are highlighted.

Supplemental Fig. 4. TAL1 promotes the expression of a set of genes that are highly expressed in different T-ALL subgroups. (A) ATAC-seq tracks of hematopoietic stem cells (HSCs), multipotent progenitors (MPPs), common lymphoid progenitors (CLPs) and CD4+ T cells; the *MYB* locus is magnified. The ATAC-seq

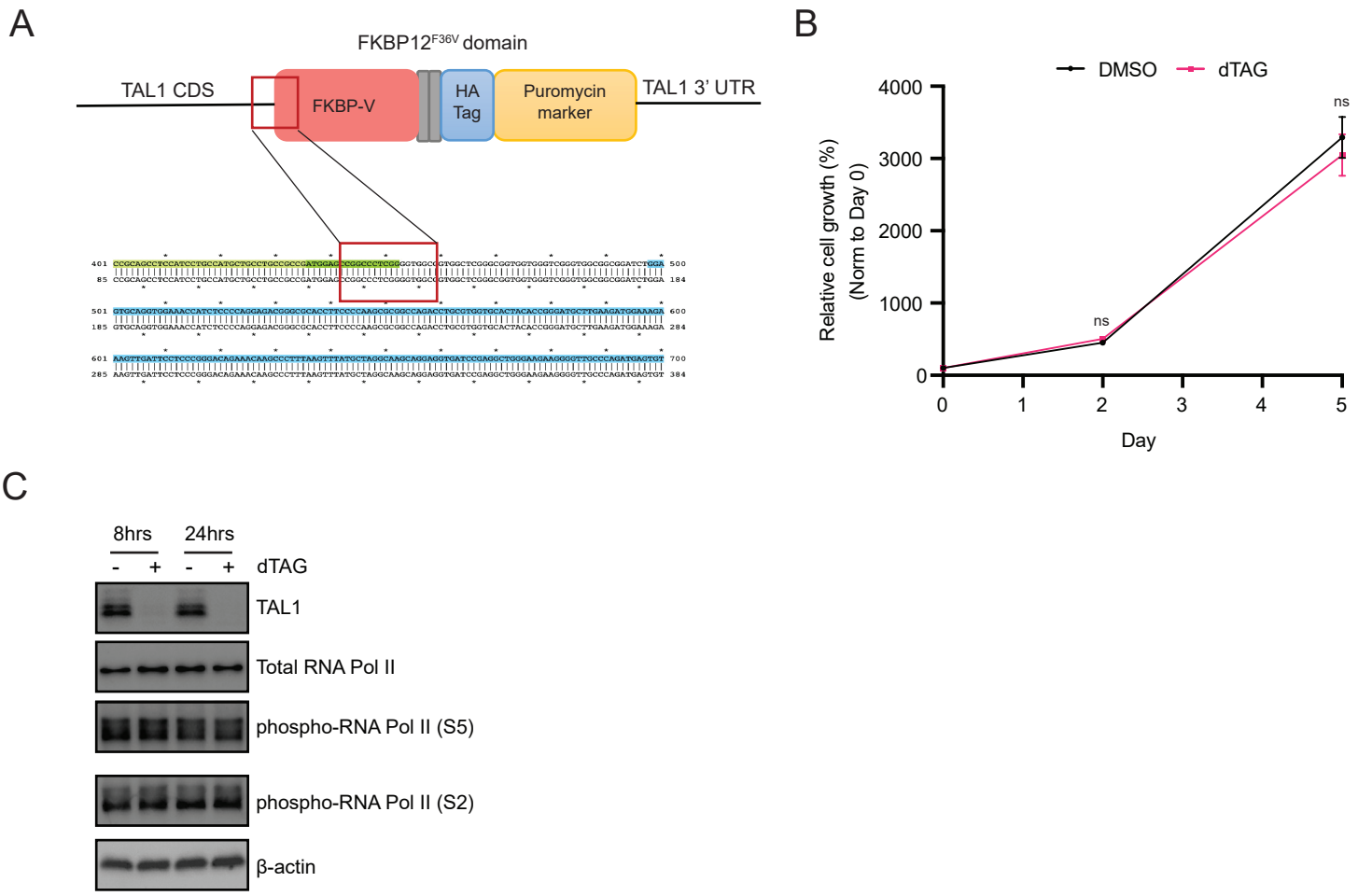
dataset was previously reported (8). **(B)** ChIP-seq tracks of the TAL1, GATA3, and RUNX1 transcription factors and activating histone mark H3K27ac (top) in the Jurkat cell line; the *MYB* family loci are magnified. The ChIP-seq dataset was previously reported(9). H3K27ac Hi-ChIP was performed after DMSO treatment and TAG treatment to evaluate changes in active chromatin loop formation (bottom). For **(A)** and **(B),** yellow highlighting indicates the genomic region of interest: where the TAL1 transcription factor binds in the baseline Jurkat cell line. For the *MYB* locus, the TSS and +143-kb putative enhancer region are highlighted. **(C)** Relative gene expression of *Gimap6*, *Runx1* and *Myb* genes in mouse hematopoietic cell populations measured by probe microarray, available from the Gene Expression Commons database (10). **(D)** mRNA expression of representative Group B genes (*GIMAP2, RUNX1, MYB, STAM,* and *STAT5B*) in a panel of TAL1-positive T-ALL cells (JURKAT, RPMI-8402, CCRF-CEM, and MOLT4) and TAL1-negative T-ALL cells (HPB-ALL, TALL-1, DND-41) as measured by RNA-seq. **(E)** ATAC-seq tracks of the basal-state HPB-ALL cells (black) and TAL1/LMO1 overexpression in the HPB-ALL cells, displayed as two independent clones (green). Yellow highlighting indicates the genomic region of interest: where the TAL1 transcription factor binds in the baseline Jurkat cell line. For the *MYB* locus, the TSS and +143-kb putative enhancer region are highlighted. **(F)** Relative mRNA expression levels of GIMAP family genes (*GIMAP2*, *GIMAP4*, GIMAP*6,* and *GIMAP8*) after TAL1-FKBP12^{F36V} Jurkat cells were treated with either DMSO, dTAG, and GSI or combined dTAG + GSI for 24 hours; n=2. **(G)** 4C-seq performed with TAL1-FKBP12^{F36V} Jurkat cells after DMSO, dTAG, GSI or combined dTAG + GSI treatment for 24 hours; the *GIMAP* family loci are magnified. Yellow highlighting denotes where the TAL1 transcription factor binds in the baseline Jurkat cell line; this region was the basis for the 4C-seq primer design.

Supplemental Fig. 5. Transcriptional activity of GATA3 and RUNX1 depends on TAL1. (A) ChIP–qPCR was performed with TAL1-FKBP12^{F36V} Jurkat cells at the TAL1-binding site of several loci of interest, namely, – *ALDH1A2* (TSS), *MYCN* (+540-kb), *RUNX1* (+23.5-kb*)* and *ARID5B* (–137-kb), after DMSO treatment and dTAG treatment. The negative control was designed with primers and used for identifying the irrelevant *IGFBP3* locus. The data are presented as the percentages of input after GATA3 protein pulldown assay.

Supplemental Fig. 6. TAL1 induces lethality via the inhibition of the NOTCH1 pathway in TAL1-negative, TLX3-positive T-ALL cells. (A) Relative growth (%) of DND-41 overexpressing TAL1, LMO1 or TAL/LMO1 and of KOPT-K1 overexpressing TAL1, measured 0, 2 and 4 days after overexpression using a Cell TiterGlo assay kit; n=3. Error bars: the mean ± SEM; ns: not significant; ****p<0.001 by Student's two-tailed t test compared to empty vector (EV) control cells. **(B)** (Left) Representative flow cytometry plot after Annexin-V apoptosis assay was performed with Dox (-) or Dox (+) TAL1/LMO1-overexpressing HPB-ALL cells. (Right) The percentage (%) of apoptotic cells among the Dox (-) and Dox (+) TAL1/LMO1-overexpressing HPB-ALL cells displayed in a bar graph; n=3 for all conditions. Error bars are the mean ± SEM; ns: not significant by Student's two-tailed t test compared to Dox (-) cells. **(C)** (Top panel) DND-41 cells after TAL1/LMO1 overexpression and (bottom panel) KOPT-K1 cells after TAL1 overexpression. (Left) Whole-cell lysates were collected 48 hours after overexpression and then subjected to immunoblot analysis using antibodies specific for TAL1, LMO1/2, MYC, cleaved NOTCH, CDK4 and GAPDH (loading control). The blots are representative of n=3. (Right) The relative mRNA expression levels of gene

targets positively regulated (*PI16, ALDH1A2, ZNF521*) and negatively regulated (*MYC, NOTCH1, NOTCH3*) after TAL1/LMO1 and TAL1 overexpression in DND-41 and KOPT-K1 cell lines respectively; n=2. **(D)** HPB-ALL cells after Dox-induced TAL1/LMO1 overexpression were stably transduced with either an empty vector (EV) or the cDNA carrying the MYC gene. The relative mRNA expression levels of *NOTCH1* are shown. n=2. **(E)** Cell surface expression of CD1, CD3, CD4 and CD8a measured by flow cytometry in Dox (-) and Dox (+) TAL1-/LMO1-overexpressing HPB-ALL cells, analyzed by FlowJo software. n=3. ****p<0.0001 by Student's t test compared to Dox (-) cells.

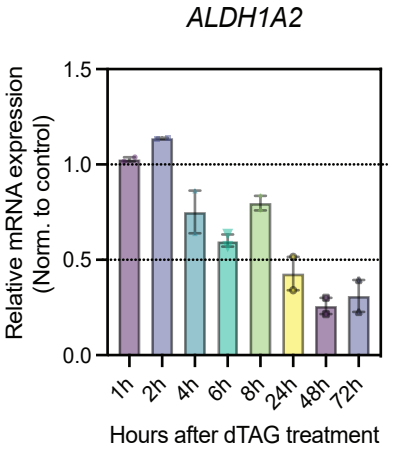
Supplemental Figure 1

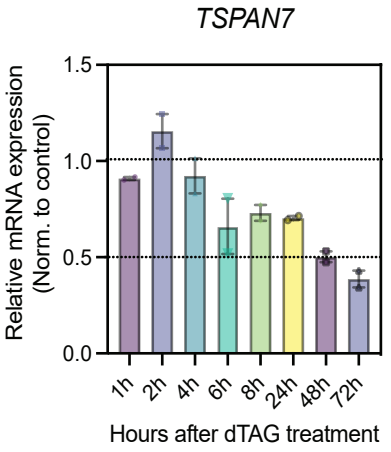


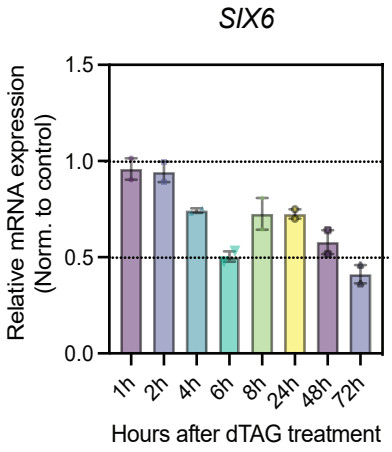
Supplemental Figure 1

D

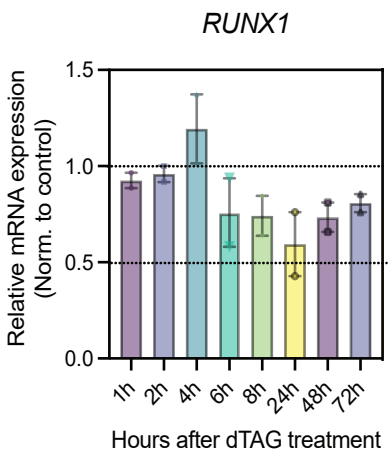
Group A

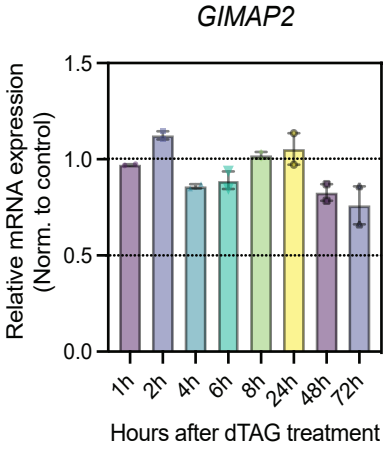


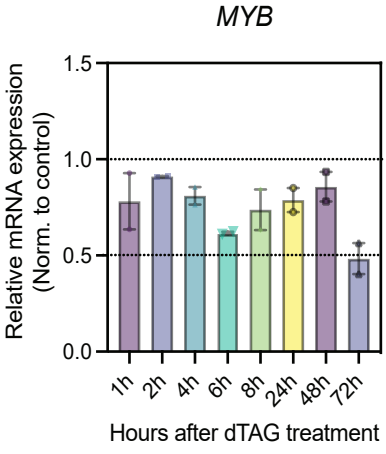




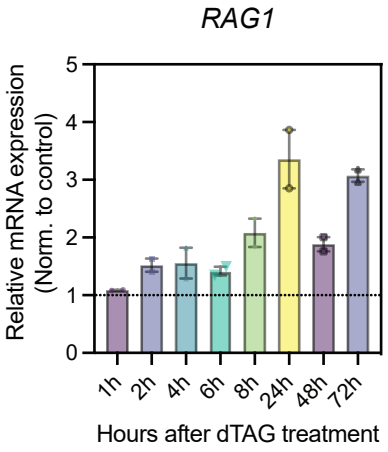
Group B

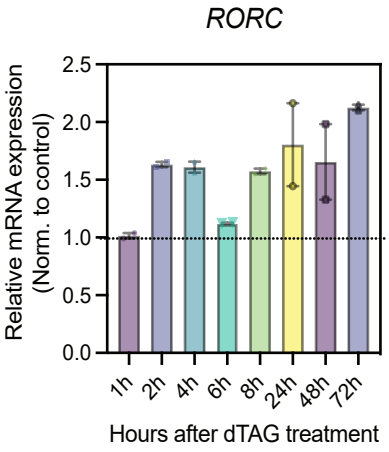


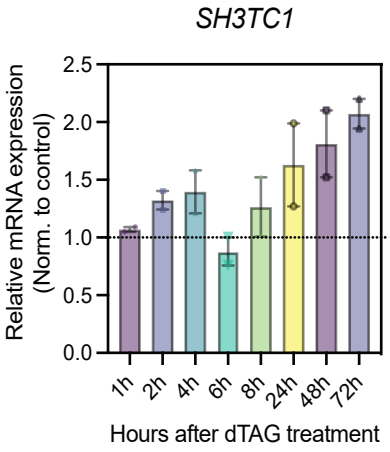




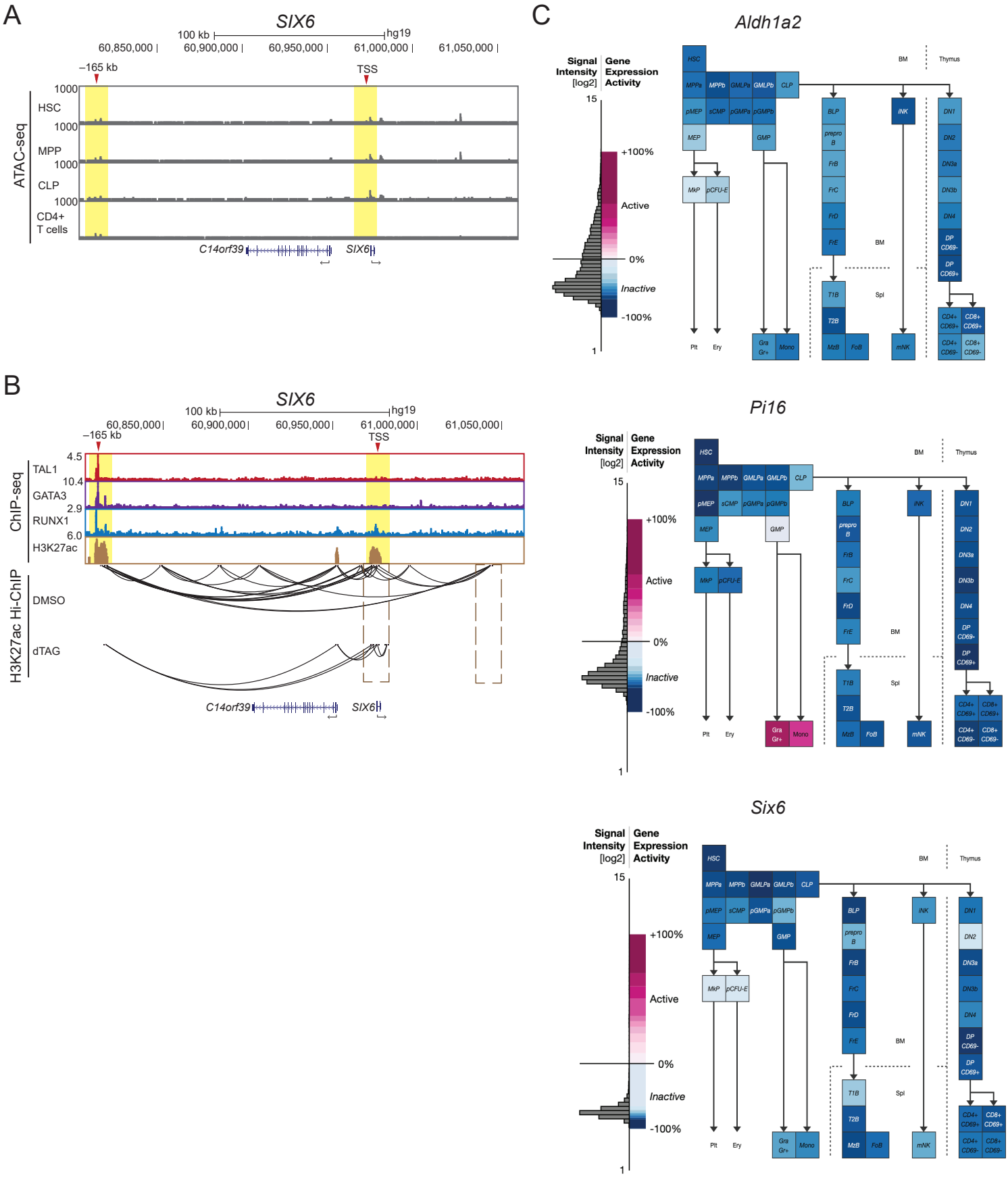
Group C



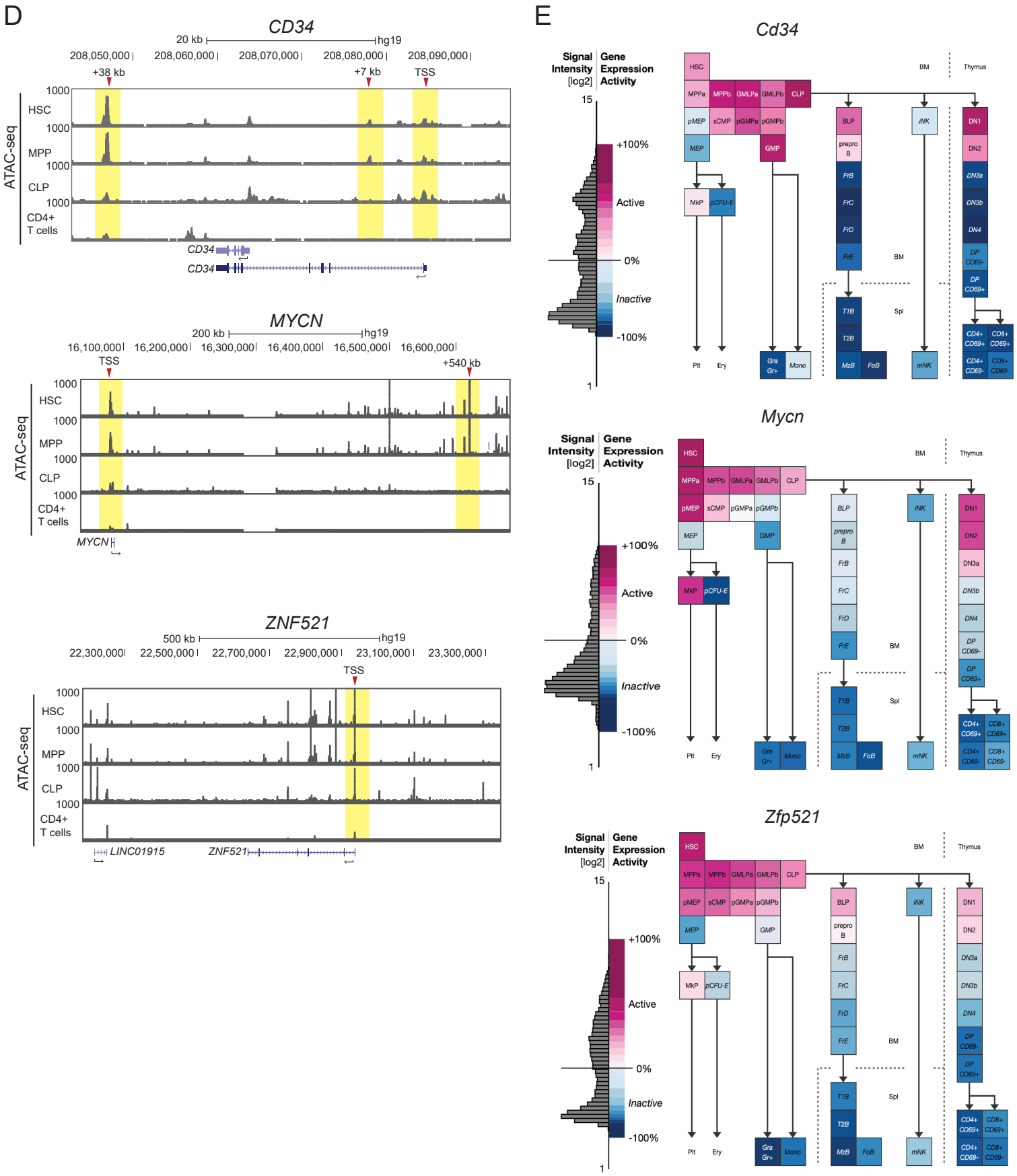




Supplemental Figure 2

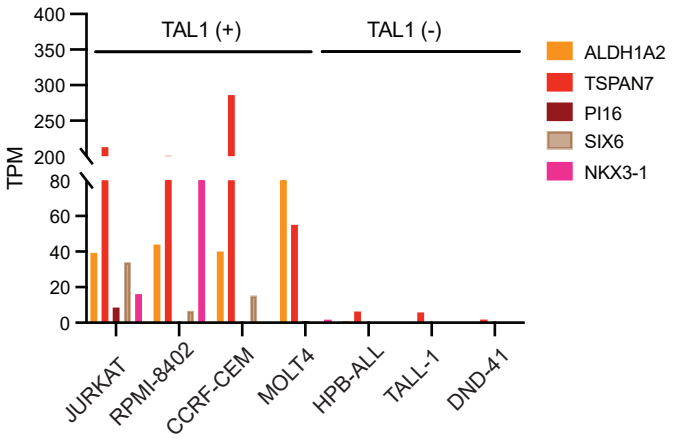


Supplemental Figure 2

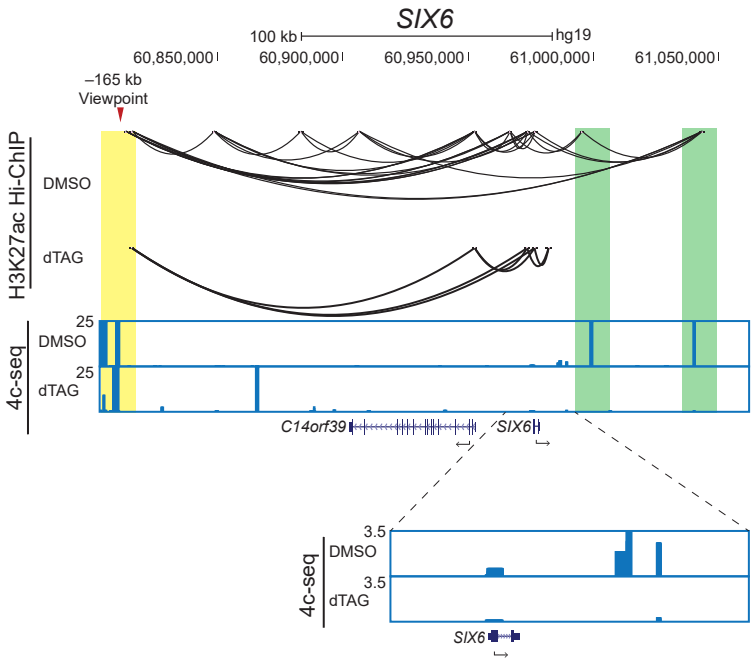


Supplemental Figure 2

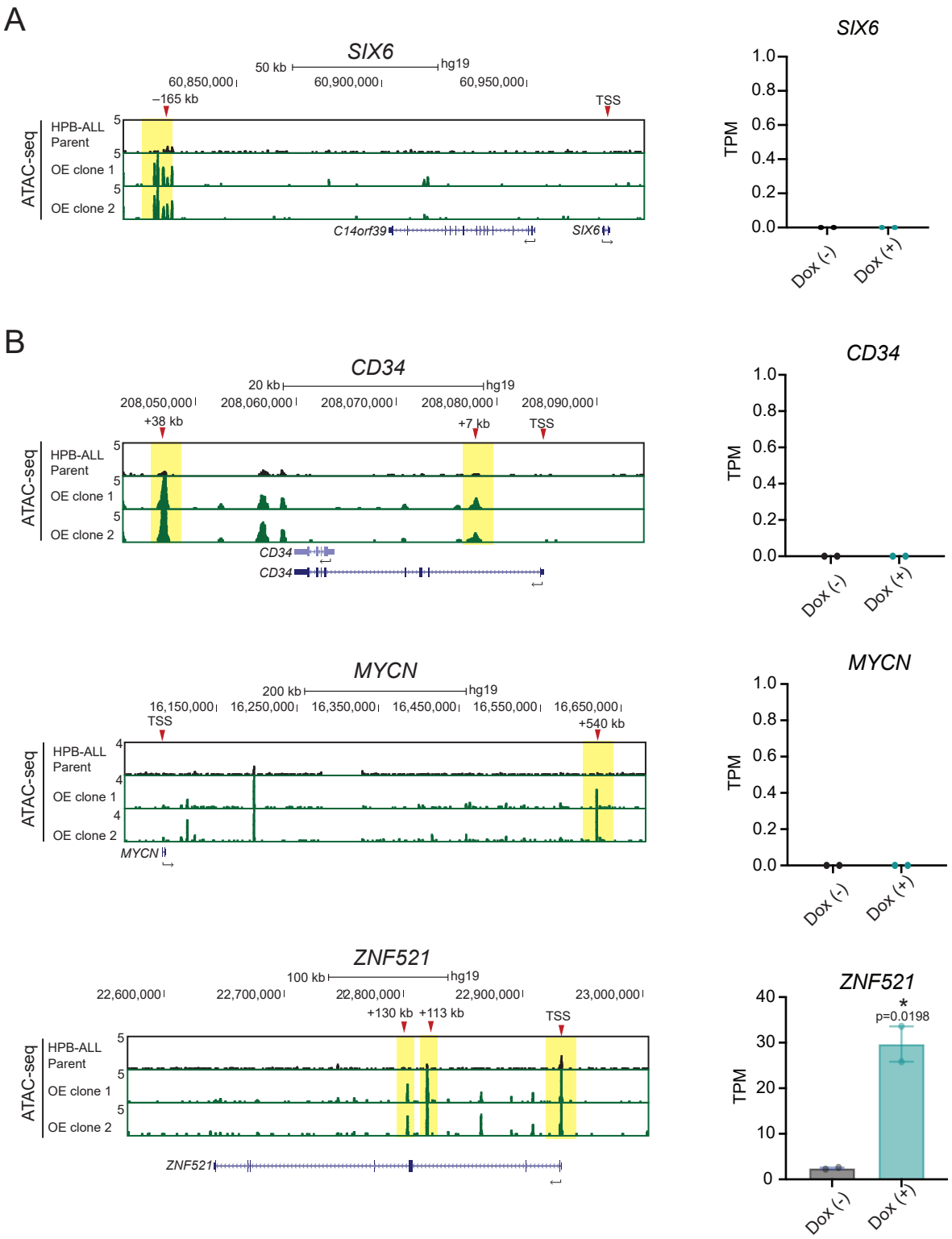
F



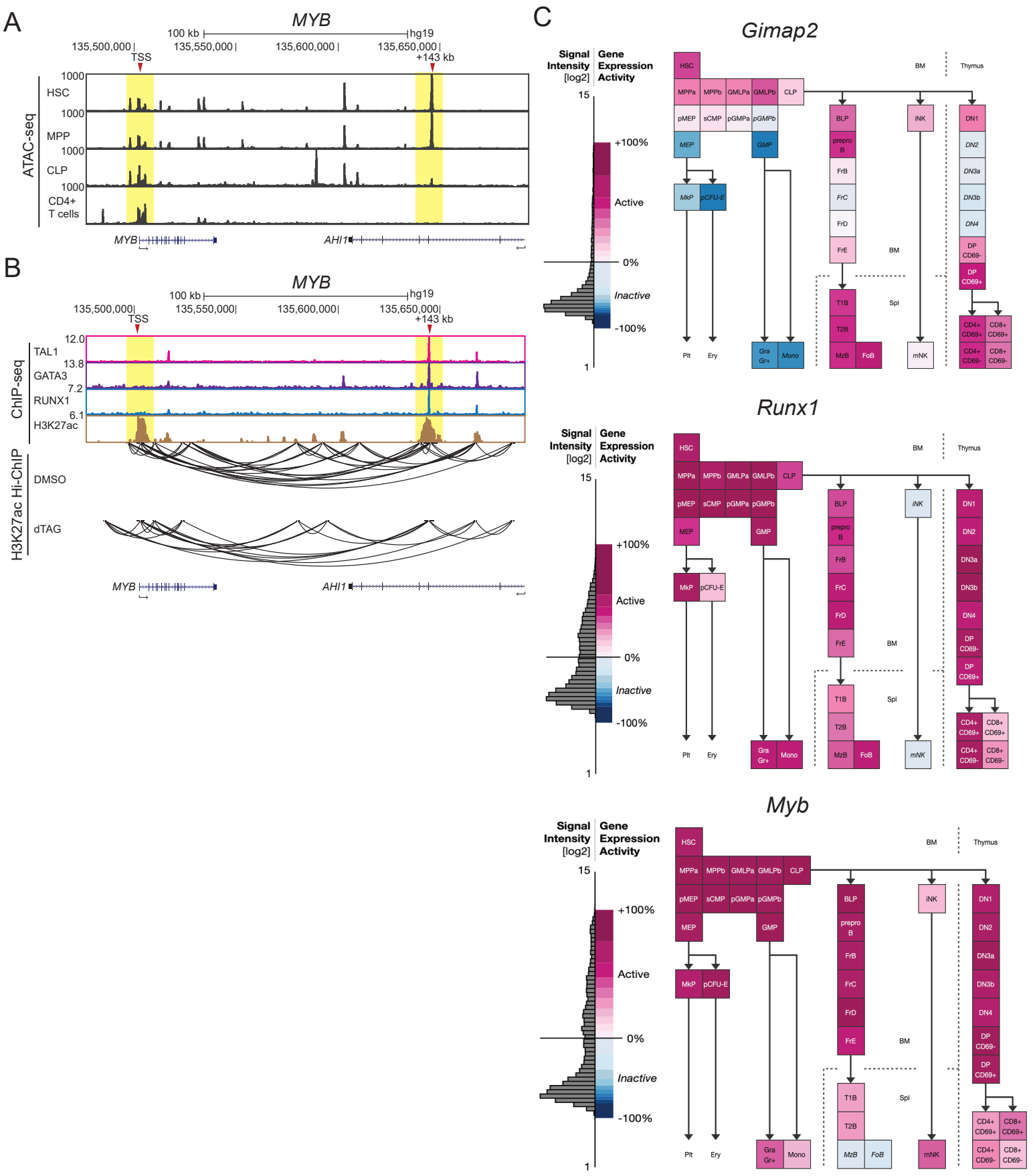
G



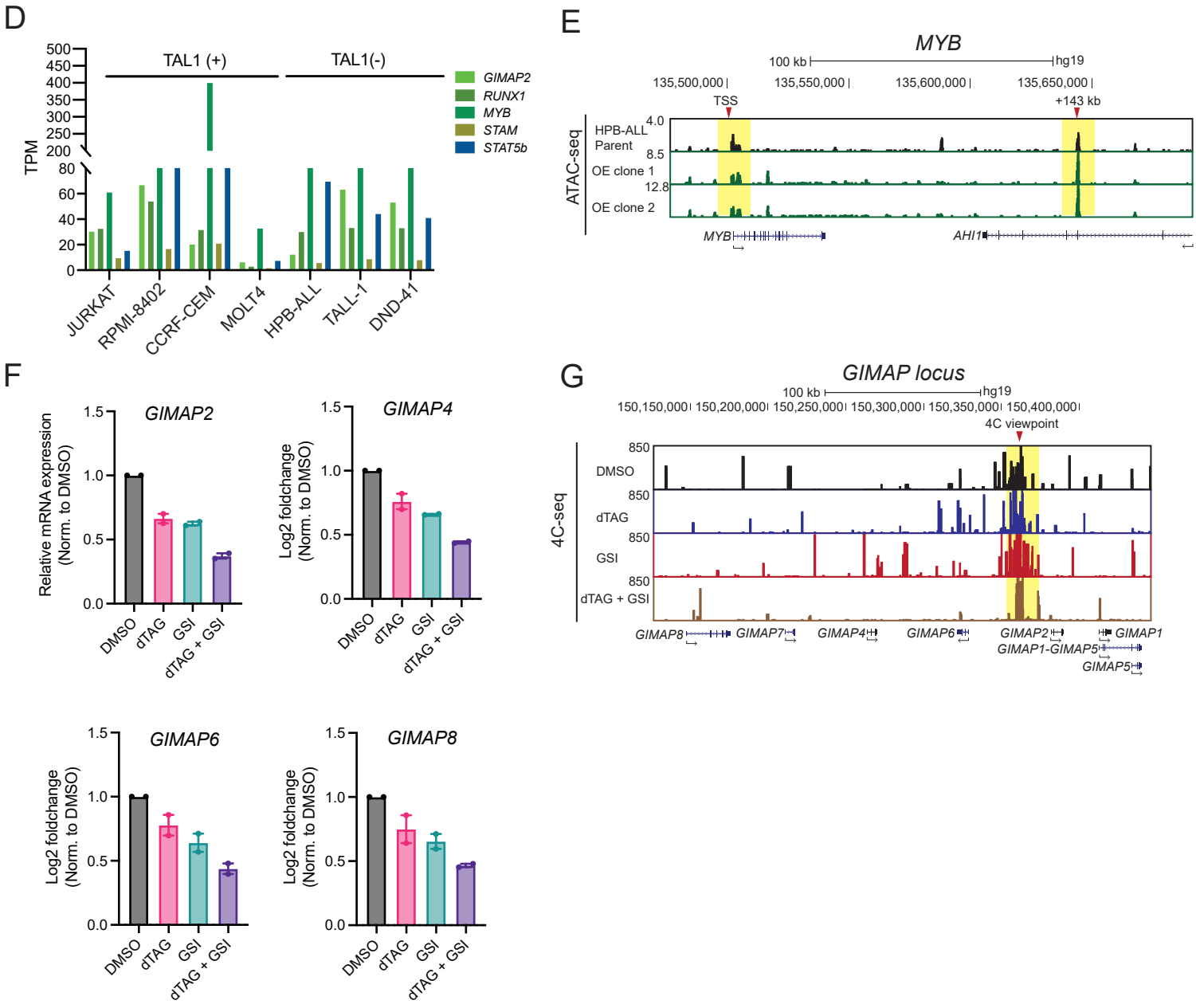
Supplemental Figure 3



Supplemental Figure 4

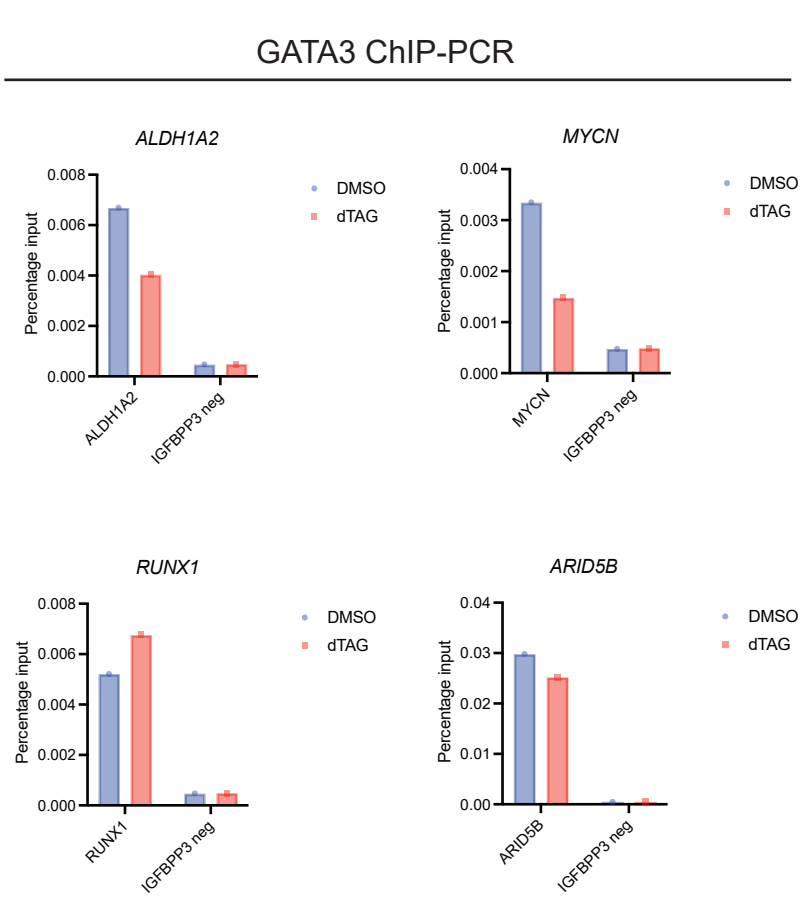


Supplemental Figure 4



Supplemental Figure 5

A



Supplementary Figure 6