Jin et al

Disruption of co-transcriptional splicing suggests RBM39 is a new therapeutic target in acute lymphoblastic leukemia

SUPPLEMENTAL INFORMATION

Supplemental Methods

Research with PDX cells and cell lines

DS-ALL PDX cells were derived in Dr. Sebastien Malinge's lab while the B-ALL PDX samples were obtained from the PROPEL repository at St. Jude Children's Hospital. Briefly, $1x10^6$ cells were transplanted into NSG mice and after mice showed signs of disease including hunching and reduced locomotion, leukemic cells were harvested from bone marrow using human CD45 selection EasySep kits (Stem Cell Technologies) and then cultured ex vivo. B-ALL PDX samples were cultured in StemSpan™ SFEM II (Stem cell technologies, Catalog # 09655) with human IL7, SCF, and FLT3-L at 10 ng/mL. T-ALL PDX cells was cultured in RPMI medium that was supplemented with 10% FCS, 10% heat inactivated human AB serum, human IL7 (10 ng/ml), human SCF (50 ng/ml), human FLT3 ligand (20 ng/ml), human IL2 (100 ng/ml). Healthy human $CD19^+$ cells and $CD3^+$ cells were purchased from STEMCELL Technologies or Cellero.com. $CD34^+$ cells from healthy controls were obtained from StemExpress.com. $CD34^+$ cells were cultured in StemSpan™ SFEM II with StemSpan™ $CD34^+$ Expansion Supplement (Stem cell technologies, Catalog # 02691). Healthy human $CD19^+$ cells and $CD3^+$ cells were cultured in RPMI supplemented with 10% FBS, Non-essential amino acid, Sodium Pyruvate, and 2-Mercaptoethanol. B-ALL cell lines (Kasumi-7, Kasumi-9, MUTZ-5, SEM, SUPB15) were gifts from Dr. Mullighan. HEK293T and NALM6 cells were purchased from ATCC. Cells were cultured in RPMI supplemented with 10% FBS (SEM, NALM6, Kasumi-7, Kasumi-9) or 20% FBS (MUTZ-5, SUPB15), 2mM L-glutamine, and penicillin/streptomycin (GIBCO). Cell identification was verified by STR determination, and cells were regularly tested to confirm the absence of mycoplasma. B-ALL PDX samples used for ex vivo and in vivo treatments are described in Table S8, and information of DS-ALL[1] and T-ALL PDX samples were previously described[2].

Flow cytometry

Cells were stained in FACS buffer containing PBS + BSA 0.5% with the antibodies listed under the list of reagents. Cells were washed with FACS buffer and data were acquired with a BD FACS Aria II or BD Fortessa. For xenografts, single-cell suspensions of PBMCs were prepared and stained in FACS buffer containing fluorochrome-conjugated antibodies for surface markers, including human CD19 (hCD19). Apoptosis was assayed with APC Annexin V (BD Bioscience) and performed according to manufacturer's specifications. Cell cycle analysis was determined using 5-ethynyl-2'-deoxyuridine (EdU) staining and performed according to manufacturer's specifications (ThermoFisher). Viability of $CD34^+$ cells after dinaciclib and EHT1610 treatment was assessed by flow cytometry with 4′,6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) staining. Flow cytometric analysis was performed with FlowJo software (Tree Star).

Western blot

Western blots were performed as previously described[3]. Briefly, cells were lysed with RIPA buffer (50 mM Tris HCl pH 8.0, 150 mM NaCl, 1% NP-40/IGEPAL, 0.5% sodium deoxycholate, 0.1% SDS, 1:100 protease inhibitor (Sigma-Aldrich, P8340), 1 mM NaV, and 1 mM NaF in H2O) on ice for 20 min, and spun down at 4°C, max speed, for 10 min to remove debris. Protein concentrations were determined via BCA assay (ThermoFisher) according to manufacturer's specifications. Up to 50 µg sample was boiled in 1X SDS loading dye (Bio-Rad) at 95°C for 10 min followed by loading into 4–15% Tris-glycine polyacrylamide gels (Bio-Rad) according to manufacturer's specifications. After transfer, Nitrocellulose membranes were probed with indicated antibodies at 1:1000 and visualized by Li-Cor Odyssey Clx Imaging System.

PCR reactions

A total of 250,000 cells under various conditions was isolated for RNA by using RNA extraction kit (Takara). cDNA synthesis was performed using the High-Capacity cDNA Reverse

Transcription Kit (Themofisher). PCR analysis was performed by using GoTaq® Green Master Mix, 2X (Promega) according to manufacturer's specifications. Primers are listed in Table S9.

Plasmids and viral transduction

RBM39 was PCR amplified from NALM6 cDNA and cloned into pRRL-EF1α-GFP**.** All the plasmids were amplified to 1mg/ml before transfection. HEK293T cells were grown in DMEM medium with 10% FBS and 1% penicillin streptomycin. All transfections were performed in HEK293T cells using TransIT-LT1 Transfection Reagent (Mirus bio) at 4:2:3 ratios of plasmid: pVSVG: pPax2 in OPTI-MEM solution. Viral supernatant was collected 48 hrs and 72 hrs post-transfection. Spin infections were performed at room temperature at 2,500 RPM for 1 hr and 30 mins with polybrene reagent (6 μg/ml, Santa Cruz). Puromycin (1 μg/ml, ThermoFisher) and blasticidin (5 μg/ml, ThermoFisher) were used for selection after 48 hrs of transduction. High GFP^+ cells were sorted after transduction. After selecting puromycin-resistant *shRBM39*-expressing cells, we used Ficoll-Paque to isolate viable cells. We achieved at least 80% viability after Ficoll-Paque selection for subsequent in vitro and in vivo studies.

Cell viability assays

100K cells were plated on 96 wells plate. All the drugs were added at 1:1000 dilution for indicated concentration. The cell viability was determined by CellTiter-Glo® Luminescent Cell Viability Assay (Promega) according to manufacturer's specifications.

CRISPR/Cas9 screen

A previously described sgRNA domain-focused approach was used in the screen[4]. The RNA binding factor CRISPR KO library[4] (a gift from Dr. Eric Wang and Iannis Aifantis) was expanded and validated by the St. Jude Center for Advanced Genome Engineering as described in the Broad GPP protocol (https://portals.broadinstitute.org/gpp/public/resources/protocols) with one exception being the use of Endura DUOs electrocompetent cells. All sequencing was performed in the St. Jude Hartwell Center. Single-end 100 cycle sequencing was performed on a NovaSeq 6000 (Illumina). Validation to check gRNA presence and representation was performed using calc_auc_v1.1.py (https://github.com/mhegde/) and count_spacers.py[5]. Viral particles were produced by the St. Jude Vector Development and Production laboratory. Cas9-expressing NALM6 cells were transduced with sgRNA library virus at a low MOI (~0.3). On Day 4 post-transduction, GFP percentage was assessed to determine infection efficiency (~30%) and sgRNA coverage (~300-500X). The remaining 300,000-500,000 cells were returned to culture after each passage until 20 days' post-transduction. Genomic DNA (gDNA) of cells containing 300-500X coverage were harvested on Day 4 and Day 20 using Qiagen DNA kit based on manufacturer's protocol. gDNA was amplified according to Broad GPP protocols and submitted for Illumina next generation sequencing as described above. CRISPR KO screens were analyzed using Mageck-Vispr/0.5.7[6]

RNA-seq

RNA was isolated using RNA extraction kit (Takara). Quality check, library preparation and sequencing were performed by the Hartwell Center at St Jude Children's Research Hospital and RNA quality was assessed by an Agilent 2100 Bioanalyzer. The Illumina TruSeq mRNA library preparation kit was used to prepare sequencing libraries and the NovaSeq 6000 system was used to generate paired-end, 150 bp reads at the depth of 100 M reads per sample. Read alignment to the Gencode human reference genome (GRCh38) and annotation (v31) was performed with STAR (version 2.7) software. Aligned reads were used as input to rMATS turbo (v4.1.0) using default parameters for alternative splicing analysis[7]. Junction and exon counts (JCEC) output from rMATS for every event type (A3SS, A5SS, MXE, RI, and SE) were filtered in R (v4.3.3). For comparison between $CD19^+$ and B-ALL patient samples, IncLevelDifference > 0.15 and FDR <

0.05 were applied to filter the alternative splicing events. For all the other AS analysis, first, we defined effect size as the absolute value of IncLevelDifference > 0.25 and required the FDR < 0.05. Every event has a skipping and an inclusion isoform for each condition (treated and untreated). Next, to remove events with low read coverage we required at least 1 of the 4 isoforms for every event to have an average >10 reads. These filtered events were the basis of all downstream analysis. We generated unique IDs for each event containing: gene symbol, chromosome, event start, event stop, upstream exon start, upstream exon stop, downstream exon start, downstream exon stop, event type. These unique IDs were used for all event-based comparative analyses including Venn diagrams (VennDiagram v1.7.3) and scatter plots (ggplot2 v3.5.). Gene Ontology were analyzed on the website: http://geneontology.org/.

ChIP-seq

ChIP-seq were performed in NAML6 cells with minor modifications (two biological replicates for each condition)[8]. Briefly, 30 million NALM6 cells were crosslinked with 1% formaldehyde (Thermo Fisher) in PBS for 10 min at room temperature. Chromatin was sonicated with the Covaris E220, and DNA-protein crosslinked complexes were immunoprecipitated overnight at 4°C with antibodies (see list of reagents) and Dynabeads Protein G (Invitrogen). Protease treatment and reverse crosslinking reaction were performed at 65°C overnight, followed by DNA purification using PCR purification kit (QIAGEN). Quality check, library preparation and sequencing were performed by the Hartwell Center at St Jude Children's Research Hospital. Quality was assessed with an Agilent 2100 Bioanalyzer. The NovaSeq 6000 system was used to generate paired-end, 50 bp reads at the depth of 100 M reads per sample. Raw ChIP-seq reads were mapped to GRCh38 using bwa-mem (Version 0.7.16a) sorting, indexing and duplicates were marked using samtools(Version 1.7). MACS2(Version 2.1.1) were used for peak calling on unique mapped reads and ENCODE blacklist regions were used to filter the peaks. bamCoverage command from deeptools (Version 3.5.4) were used to generate RPGC normalized bigwig tracks. mergePeaks command from homer (Version v4.11.1) were used to merge the replicates and only peaks existing in both replicates were used for downstream analysis. GetDifferentialPeaks command from homer(Version v4.11.1) were used for differential peak analysis from summits peak location by extending 200pb flanking regions. annotatePeaks command from homer were used to annotate the peaks. The metagene plot for ChIP-seq analysis is normalized to the total mapped reads.

Intravenous xenograft and mice studies

The animals used in this study were housed in the animal facilities of St. Jude Children's Research Hospital. For NALM6 B-ALL xenograft studies, 1 million (*shRBM39* or *control*) or 0.15 million (inducible *shRBM39* or *control*) cells were intravenously transplanted into female NSG (NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ) mice aged 8 to 12 weeks. Leukemia burden was determined by weekly bioluminescence imaging using a Xenogen IVIS-200 system and Living Image software (Caliper Life Sciences). Doxycycline (2 mg/ml in 5% sucrose) was provided on day 6 after transplant. In the context of silencing RBM39 in Ph-like patient samples, 0.5 million cells were used for the xenograft study. The cells were transduced with lentivirus for 48 hours, after which an additional 48 hours of puromycin selection was provided. Subsequently, these modified cells were intravenously transplanted in NSG mice for further analysis. For the MEF2D-HNRNPUL1 patient sample, 2 million cells were intravenously transplanted in NSG mice, and the mice were randomized. The E7820 were administrated daily with oral gavage (50 mg/kg) or intraperitoneal injection (25 mg/kg) after 2 weeks of transplantation for the time as indicated. In the context of dinaciclib and EHT1610 in B-ALL PDX sample studies, the Ph-like PDX sample (2 million cells) or the relapsed PDX sample (0.5 million cells) were intravenously transplanted in NSG mice, and the mice were randomized. Dinaciclib (5 mg/kg) and EHT1610 (20 mg/kg) were administrated daily with intraperitoneal injection after detection of 0.5% of $CD19^+$ cells in the peripheral blood

for 10 days. Leukemia burden was determined by the percent of $CD19^+$ cells in the peripheral blood.

Immunofluorescence
NALM6 cells were transferred to poly-l-lysine pre-treated coverslips for 5 minutes to allow them to adhere. Cells were then rapidly rinsed with 1X PBS-5mM EGTA, followed by incubation at -20 °C in 95% Methanol-5 mM EGTA for 30 minutes. The primary antibodies used were rabbit anti-SF3B1 (Sigma-Aldrich, HPA050275, used at 1:200), rat anti-RNA polymerase II subunit B1 phospho CTD Ser-2 (Sigma-Aldrich, clone 3E10, used at 1:400), and rat anti-RNA polymerase II subunit B1 phospho-CTD Ser-5 (Sigma-Aldrich, clone 3E8, used at 1:400). The secondary antibodies used were donkey and conjugated to Alexa Fluor Rhodamine Red™-X or Alexa Fluor 647 (Jackson ImmunoResearch) at 1:300 diluted in 5% normal donkey serum. Cells were counterstained with DAPI (4',6-Diamidino-2-Phenylindole, Dihydrochloride, Invitrogen) diluted in PBS (300 nM) for 2 minutes, and then mounted onto glass slides with antifade solution (90% glycerol, 0.5% N-propyl gallate).

Confocal microscopy Imaging
All microscopic images were acquired on a 3i Marianas system (Denver, CO) configured with a Yokogawa CSU-W spinning disk confocal microscope utilizing a 100x Zeiss objective and Slidebook 6.0 (3i). 3D images of cells were captured as z stacks with 0.2 mm spacing between planes, spanning 12.2 mm in total. All fixed cell imaging was performed at room temperature.

DNA FISH
Unfixed cells were applied to slides by cytocentrifugation to an appropriate cell density. Cells are fixed in 4% PFA followed by a 10-minute treatment in 0.2N HCl containing 0.5% Triton X 100 and then denatured in 70% formamide, 2X SSC at 80ºC for 10 minutes. Slides were then dehydrated in a graded ethanol series. The immunofluorescence staining of SF3B1 was performed as described in immunofluorescence section. Probes were prepared by nick translation of genomic DNA clones using fluorophore labeled dUTPs. Labeled DNA was suspended in a hybridization buffer composed of 50% formamide, 10% dextran, 2X SSC and human Cot DNA as competitor to suppress repetitive sequences. The probe mixture was denatured by incubation at 70ºC for 5 minutes and then applied to dehydrated slides and incubated at 37ºC overnight. Post hybridization washing was performed in 50% formamide, 2X SSC at 37ºC for 5 minutes followed by mounting slides in Vectashield mounting medium containing DAPI. Slides were then imaged to identify DNA and SF3B1 signals and slide coordinates of imaged fields are recorded.

Image analysis
Fluorescent microscopy images were segmented and quantified using the Punctatools pipeline[9]. In brief, cell nuclei were segmented in individual z-layers then combined into 3-D stacks using the Cellpose algorithm[10]. Puncta were then segmented by filtering respective channels with a scale adapted Laplacian of Gaussian (LoG) filter, thresholding the result, and applying watershed segmentation using the maxima of the LoG filtered image as seeds.

Manders' coefficients were calculated by with the following equation on a per-cell basis:

$$MCC_A = \frac{Sum\ Intensity^*\ of\ A\ in\ the\ intersect\ of Mask_A\ and\ Mask_B}{Sum\ Intensity^* of\ A\ in\ Mask_A}$$

Where fluorescent intensity is adjusted by subtracting the average background intensity in each cell.

Volume overlap was calculated with the following equation on a per-cell basis:

$$Volume\ Fraction\ Colocalized_A = \frac{Volume\ of\ A\ intersect\ B}{Volume\ of\ A}$$

NanoBRET target engagement and kinase assays

A screen of 240 kinases in 293T cells using cellular target engagement assays with EHT1610 was performed by Promega. 240 kinases in-frame with a nanoluciferase (NanoLuc) tag were transfected in 293T cells. A cell permeable NanoBRET fluorescent tracer was then added to the cells which reversibly binds the target-NanoLuc Fusion protein in live cells to result in a BRET signal. EHT1610 (5 μM) or vehicle were then added to each cell and the degree of drug-target protein binding was assessed via loss of NanoBRET signal. Percent occupancy = [1 – (X – Z)/(Y – Z)]*100 where X = BRET in the presence of the EHT1610 and tracer, Y = BRET in the presence of tracer only (zero occupancy control), and Z = BRET for the full occupancy control (ie., BRET from the untagged NLuc fusion). Negative % occupancy values may result from noise in the assay or changes in regulation of the kinase from the compound treatment. A full kinome screen (464 kinases) with 2 μM EHT1610 was performed by Eurofin discovery service.

Genetically Engineered NALM6 Cell lines

NALM6 genetically modified cell lines were created using CRISPR-Cas9 technology in the St. Jude Center for Advanced Genome Engineering. Briefly, 400,000 NALM6 cells were transiently co-transfected with precomplexed ribonuclear proteins (RNPs) consisting of 100 pmol of chemically modified sgRNA (Synthego) and 33 pmol of 3X NLS SpCas9 protein (St. Jude Protein Production Core), 3ug of ssODN donor (IDT) or 500ng of plasmid donor (BioBasic), and 200ng of pMaxGFP™(Lonza). The transfection was performed via nucleofection (Lonza, 4D-Nucleofector™ X-unit) using solution P3 and program CV-104 in a small (20ul) cuvette according to the manufacturer's recommended protocol. Single cells were sorted based on transfected cells (GFP+) five days post-nucleofection into 96-well plates containing prewarmed media and clonally expanded. Clones were screened and verified for the desired modification using PCR-based assays and/or targeted deep sequencing. NGS analysis of clones was performed using CRIS.py[11,12]. Final clones tested negative for mycoplasma by the MycoAlert™Plus Mycoplasma Detection Kit (Lonza) and were authenticated using the PowerPlex® Fusion System (Promega) performed at the St. Jude Hartwell Center.

PRO-seq

NALM6 cells were treated with DMSO and EHT1610 for 4 h, and THAL-SNS-032 for 8 h (two biological replicates). Cell permeabilization were performed according to the protocol generated by the Nascent Transcriptomics Core at Harvard Medical School, Boston, MA. PRO-Seq library construction and data analysis[13] was performed by the Nascent Transcriptomics Core at Harvard Medical School, Boston, MA. The metagene plot for PRO-seq analysis is normalized to spike-in controls.

Annotated transcription start sites were obtained from human (GRCh38.99) GTF from Ensembl. After removing transcripts with (immunoglobulin, Mt, Mt tRNA, rRNA) biotypes, PRO-seq signal in each sample was calculated in the window from the annotated TSS to +150 nt downstream, using a custom script, make_heatmap.pl. This script counts each read one time, at the exact 3' end location of the nascent RNA. Given good agreement between replicates and similar return of spike-in reads, bedGraphs were merged within conditions, and depth-normalized, to generate bigWig files binned at 10bp. Reference genome = hg38;Spike genome = dm6

<u>Refinement of gene annotation (GGA) using PRO-seq and RNA-seq</u>

Paired-end RNA-seq reads were mapped to the hg38 reference genome via HISAT2 v2.2.1 (--known-splicesite-infile). To select gene-level features for differential expression analysis, and for pairing with PRO-seq data, we assigned a single, dominant TSS and transcription end site (TES) to each active gene. This was accomplished using a custom script, get_gene_annotations.sh (available at https://github.com/AdelmanLab/GetGeneAnnotation_GGA), which uses RNA-seq read abundance and PRO-seq R2 reads (RNA 5' ends) to identify dominant TSSs, and RNA-seq profiles to define most commonly used TESs. RNA-seq and PRO-seq data from all conditions were used for this analysis, to comprehensively capture gene activity in these samples.

Differential expression analysis

Reads were summed within the TSS to TES window for each active gene using the using the make heatmap script (https://github.com/AdelmanLab/NIH_scripts), which counts each read one time, at the exact 3' end location of the nascent RNA. DEseq2, using the Wald test, was used to determine statistically significant differentially expressed genes. Unless otherwise noted, the default size factors determined by DEseq2 were used.

Pausing Indices

Normalized data from replicate samples was merged. Promoter and gene body windows were defined as TSS/+150 and +250/+2250 (truncated at TES if shorter than 2250) respectively. RNA polymerase densities per kilobase were determined as (counts in window / window size in kb). A single pseudocount was added to both windows and pause index was calculated as PI = (promoter density +1) / (body density +1). Pause indexes were not calculated for features shorter than 400bp or with fewer than 8 reads in the promoter proximal region.

Statistical tests

For PRO-seq, statistical significance for comparisons was assessed by Wilcoxon (unpaired) or Mann-Whitney (pairwise) tests. The test used and error bars are defined in each figure legend.

Immunoprecipitation and LC-MS/MS

NAML6 cells with indicated treatment were lysed followed by immunoprecipitation of indicated antibody performed as previously described[3]. All the IP experiments used the IP lysis buffer (Thermo Fisher, cat. 87787) that is supplemented with Pierce™ Universal Nuclease for Cell Lysis (Thermo Fisher, cat. 88700). Western blot analysis was used for detection of IP lysates and input. For the mass spec analysis, the pull-down lysates were loaded in a 10% SDS gel. The proteins in the gel bands were reduced with dithiothreitol to break disulfide bond and the Cys residues are alkylated by iodoacetamide to allow the recovery of Cys-containing peptides. The gel bands were then washed, dried down in a speed vacuum, and rehydrated with a buffer containing a protease to allow the protease to enter the gel. For routine protein identification, trypsin was used for overnight proteolysis. The next day digested samples were acidified, and the peptides were extracted multiple times. Extracts were pooled, dried down and reconstituted in a small volume. The peptide samples were loaded on a nanoscale capillary reverse phase C18 column by a HPLC system (Thermo Ultimate 3000) and eluted by a gradient (75 min). The eluted peptides were ionized by electrospray ionization and detected by an inline mass spectrometer (Thermo Orbitrap Fusion). The MS spectra were collected first, and the 20 most abundant ions were sequentially isolated for MS/MS analysis. This process was cycled over the entire liquid chromatography gradient. Database searches were performed using JUMP search engine in our in-house SPIDERS software package. All matched MS/MS spectra were filtered by mass accuracy and matching scores to reduce protein false discovery rate to ~1%. Finally, all proteins identified in one gel lane were combined. The total number of spectra, namely spectral counts (SC), matching to individual proteins may reflect their relative abundance in one sample after the protein size is normalized. Moreover, the SC is useful for comparing the level of the same protein across a

number of samples (e.g. control and immunoprecipitated samples). A protein interaction network was generated using the STRING database[14]. The interaction network was generated with a required confidence score of 0.9 and ‘Active interaction sources’ based on experiments.

SUPPLEMENTAL REFERENCES

1. Laurent AP, Siret A, Ignacimouttou C, et al. Constitutive Activation of RAS/MAPK Pathway Cooperates with Trisomy 21 and Is Therapeutically Exploitable in Down Syndrome B-cell Leukemia. *Clin Cancer Res*. 2020;26(13):3307-3318.
2. Jin Q, Gutierrez Diaz B, Pieters T, et al. Oncogenic deubiquitination controls tyrosine kinase signaling and therapy response in acute lymphoblastic leukemia. *Sci Adv*. 2022;8(49):eabq8437.
3. Jin Q, Martinez CA, Arcipowski KM, et al. USP7 Cooperates with NOTCH1 to Drive the Oncogenic Transcriptional Program in T-Cell Leukemia. *Clin Cancer Res*. 2019;25(1):222-239.
4. Wang E, Lu SX, Pastore A, et al. Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. *Cancer Cell*. 2019;35(3):369-384 e367.
5. Joung J, Konermann S, Gootenberg JS, et al. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat Protoc*. 2017;12(4):828-863.
6. Li W, Koster J, Xu H, et al. Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol*. 2015;16:281.
7. Shen S, Park JW, Lu ZX, et al. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A*. 2014;111(51):E5593-5601.
8. Aoi Y, Takahashi YH, Shah AP, et al. SPT5 stabilization of promoter-proximal RNA polymerase II. *Mol Cell*. 2021;81(21):4413-4424 e4415.
9. Baggett DW, Medyukhina A, Tripathi S, et al. An Image Analysis Pipeline for Quantifying the Features of Fluorescently-Labeled Biomolecular Condensates in Cells. *Front Bioinform*. 2022;2:897238.
10. Stringer C, Wang T, Michaelos M, Pachitariu M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods*. 2021;18(1):100-106.
11. Connelly JP, Pruett-Miller SM. CRIS.py: A Versatile and High-throughput Analysis Program for CRISPR-based Genome Editing. *Sci Rep*. 2019;9(1):4194.
12. Narina S, Connelly JP, Pruett-Miller SM. High-Throughput Analysis of CRISPR-Cas9 Editing Outcomes in Cell and Animal Models Using CRIS.py. *Methods Mol Biol*. 2023;2631:155-182.
13. Mimoso CA, Goldman SR. PRO-seq: Precise Mapping of Engaged RNA Pol II at Single-Nucleotide Resolution. *Curr Protoc*. 2023;3(12):e961.
14. Jensen LJ, Kuhn M, Stark M, et al. STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res*. 2009;37(Database issue):D412-416.

Supplemental Figure Legends

Figure S1. A) Pie chart showing differential splicing events in MUTZ-5 cells treated with DMSO versus EHT1610 (5 μM) for 4 h. Events that passed FDR <0.05 and ΔPSI > 25% are included. **B)** Pie chart showing differential splicing events in NALM6 cells (left) treated with DMSO versus GNF2133 (5 μM) for 4 h. Pie chart showing differential splicing events in MUTZ-5 cells (right) treated with DMSO versus GNF2133 (5 μM) for 4 h. Events that passed FDR <0.05 and ΔPSI > 25% are included. **C, D)** scatter plot showing overlapping splicing events induced by EHT1610 and GNF2133 in NALM6 (**C**) and MUTZ-5 (**D**) cells. **E)** Sashimi plots representing splicing and exon–exon junctions for an alternative splicing event on *RBM39* transcript in MUTZ-5 cells treated with 5 μM EHT1610 for 4 h. **F)** MUTZ-5 cells were treated with different EHT1610 concentrations as indicated for 4 h. PCR reactions were performed for detection of *RBM39* splicing event. **G, H)** Sashimi plots representing splicing and exon–exon junctions for an alternative splicing event on *RBM39* transcript in NALM6 cells treated with 5 μM GNF2133 for 4 h (**G**), and in MUTZ-5 cells treated with 5 μM GNF2133 for 4 h (**H**). **I)** Transcripts per million (TPM) of *RBM39* in B-ALL (n=35) and $CD19^+$ B cells (n=4) from RNA-seq analysis. n.s. represents not significant. **J)** B-ALL PDX samples were treated with EHT1610 (5 μM) and GNF2133 (5 μM) for 4 h. PCR reactions were performed for detection of *RBM39* splicing event. **K)** DND41 and CUTLL1 cells were treated with EHT1610 (5 μM) and GNF2133 (5 μM) for 4 h. PCR reactions were performed for detection of *RBM39* splicing event. **L, M)** $CD19^+$ B cells and $CD3^+$ T cells were treated with DMSO or EHT1610 (5 μM). PCR reactions were performed for detection of RBM39 splicing event.

Figure S2: A) Western blot analysis in NALM6 clone with a degradable C-terminally tagged *DYRK1A*-$FKBP12^{F36V}$-fusion after treatment with dTAG-13 (100 nM) for indicated time points. **B)** *DYRK1A*-$FKBP12^{F36V}$-fusion NALM6 clone was treated with dTAG13 (100 nM) for 6 hours. PCR reactions were performed for detection of *RBM39* splicing event. Three technical replicates (left) and quantification of RBM39 NMD isoform (right) were shown. **C)** NALM6 cells were treated with Laduviglusib or SB 216763 for 4 h. PCR reactions were performed for detection of *RBM39* splicing event. **D, E)** Western blot analysis (**D**) and editing efficacy (**E**) of NALM6 cells with multiple CRISPR knockout guides as indicated are shown. **F**) PCR reactions were performed for detection of *RBM39* splicing event upon CRISPR Knockout of *GSK3A* and *GSK3B* in NALM6 cells. **G-I)** NALM6 cells were treated with inhibitors as indicated. PCR reactions were performed for detection of *RBM39* splicing event.

Figure S3: A) Western blot to detect changes in phosphorylated SR proteins after treatment of NALM6 cells with increasing concentrations of SM09419 for 4 h. EHT1610 treatment was used for a negative control here. **B)** NALM6 cells were treated with OTS964 or EHT1610. PCR reactions were performed to detect the *RBM39* splicing event. **C)** Western blot analysis of extracts from NALM6 cells treated with escalated doses of EHT1610 for 4 h. **D)** Western blot analysis of Ph-like ALL patient sample (upper) or MUTZ-5 cells (lower) treated with increasing concentrations of EHT1610. **E)** Immunofluorescence analysis of p-Ser2 Pol II (upper) and p-Ser5 Pol II (lower) in NALM6 cells treated with DMSO or EHT1610 (4 μM). Scale bars depict 5 microns. **F)** Quantification of immunofluorescence signals in (**E**). *** $P < 0.001$. **G)** NALM6 cells were treated with DRB (30 μM) for 4 h. PCR reactions were used to detect the *RBM39* splicing event. **H)** NAML6 were treated with 10 μM of EHT1610 in a time-dependent manner. PCR reactions were performed for detection of *RBM39* splicing event (upper) and western blots were performed for detection of the indicated proteins (lower). **I)** Cell growth assay (n=3) for NALM6 cells treated with increasing doses of EHT1610. **J)** Additional biological replicates for the western blot analysis shown in Fig. 2J (left), along with the quantification of the analyses (right). **K)** NALM6 cells were

treated with THAL-SNS-032 (250 nM) for 4 h and 20 h. PCR reactions were performed to detect the *RBM39* splicing events (upper) and western blot analysis of extracts were shown (lower).

Figure S4. NALM6 cells were treated with DMSO and EHT1610 (5 μM) for 4 h, and THAL-SNS-032 (500 nM) for 8 h followed by PRO-seq. **A, B)** Scatter plot showing comparison of THAL-SNS-032 versus DMSO (**A**) or EHT1610 versus DMSO (**B**) by differential nascent mRNA expression analysis. FC represents fold change. **C)** Normalized counts of *RBM39* in NALM6 cells treated with DMSO, EHT1610, and THAL-SNS-032 from differential nascent mRNA expression analysis. n.s. represents not significant. * $P < 0.05$. **F)** The pausing indexes were calculated as the ratio of Pol II density at the promoter over gene body (See methods). **** $P < 0.0001$.

Figure S5. A) Western blot for RBM39 (left) and growth curves of control- and *shRBM39*-expressing NALM6 cells over a period of 5 days (n = 3, right; *$P < 0.05$). **B)** Annexin V staining (72 h) of NALM6 cells that expressed either control shRNA or *shRBM39*. A representative example (left) and quantification of Annexin V^+ cells (n = 3, right; *$P < 0.05$) are shown. **C)** Cell cycle analysis of NALM6 cells following introduction of control shRNA or *shRBM39* after 72 h (right). The experiment was repeated with three biological replicates; a representative example is shown (left). **D)** Western blot for RBM39 (left) and growth curves of control- and *shRBM39*-expressing SUPB15 cells over a period of 5 days (n = 3, right; *$P < 0.05$). **E)** Annexin V staining of SUPB15 cells that expressed either *control* shRNA or *shRBM39* after 72 hours. A representative example (left) and quantification of Annexin V^+ cells (n = 3, right; *$P < 0.05$) are shown. **F)** Cell cycle analysis following control shRNA or *shRBM39* treatment of SUPB15 cells for 72 hours (right). The experiment was repeated with three biological replicates; a representative example is shown (left).

Figure S6. A) Cell cycle analysis following inducible control shRNA or *shRBM39* treatment of NALM6 cells for 72 hours (right). The experiment was repeated with three biological replicates; a representative example is shown (left). **B)** Shown are the growth curves of inducible control- and *shRBM39*-expressing NALM6 cells without doxycycline treatment. **C)** Schematic representation of doxycycline administration following transplantation of NALM6 cells containing control hairpin RNA, *shRBM39.a*, or *shRBM39.c* in NSG mice.

Figure S7. A) Western blot analysis for RBM39 in Kasumi-7 cells treated with increasing concentrations of E7820 (upper) and for NALM6 cells treated with increasing concentrations of E7070 (lower). **B)** Annexin V staining of Kasumi-7 cells treated with E7070 or E7820 for 72 h. A representative example (left) and quantification of for three independent replicates are shown (*$P < 0.05$). **C)** Cell cycle analysis of Kasumi-7 cells following E7070 or E7820 treatment for 72 h of Kasumi-7 cells. Data depict three biological replicates; a representative example is shown (left). **D)** Cell cycle analysis of NALM6 cells following E7070 or E7820 treatment for 72 h (right). The experiment was repeated with three biological replicates; a representative example is shown (left). **E)** Annexin V staining of NALM6 cells treated with E7070 or E7820 for 72 h. A representative example (left) and quantification of three biological replicates are shown. (*$P < 0.05$) **F)** IC_{50} of E7070 across different B-ALL cell lines as indicated. n = 3; the mean value ± SD is shown.

Figure S8. A) Western blot analysis for RBM39 in MEF2D-HNRNPUL1 ALL and Ph-like ALL PDX cells treated with increasing concentrations of E7820 for 4 h (left). Quantification of the percent RBM39 that was calculated as RBM39/GAPDH normalized on vehicle (right). **B)** Schematic showing the administration of E7820 in MEF2D-HNRNPUL1 PDX model. **C)** Disease burden analysis for mice in (**B**) by assessment of human $CD19^+$ cells in PB. *$P < 0.05$. **D)** Western blot

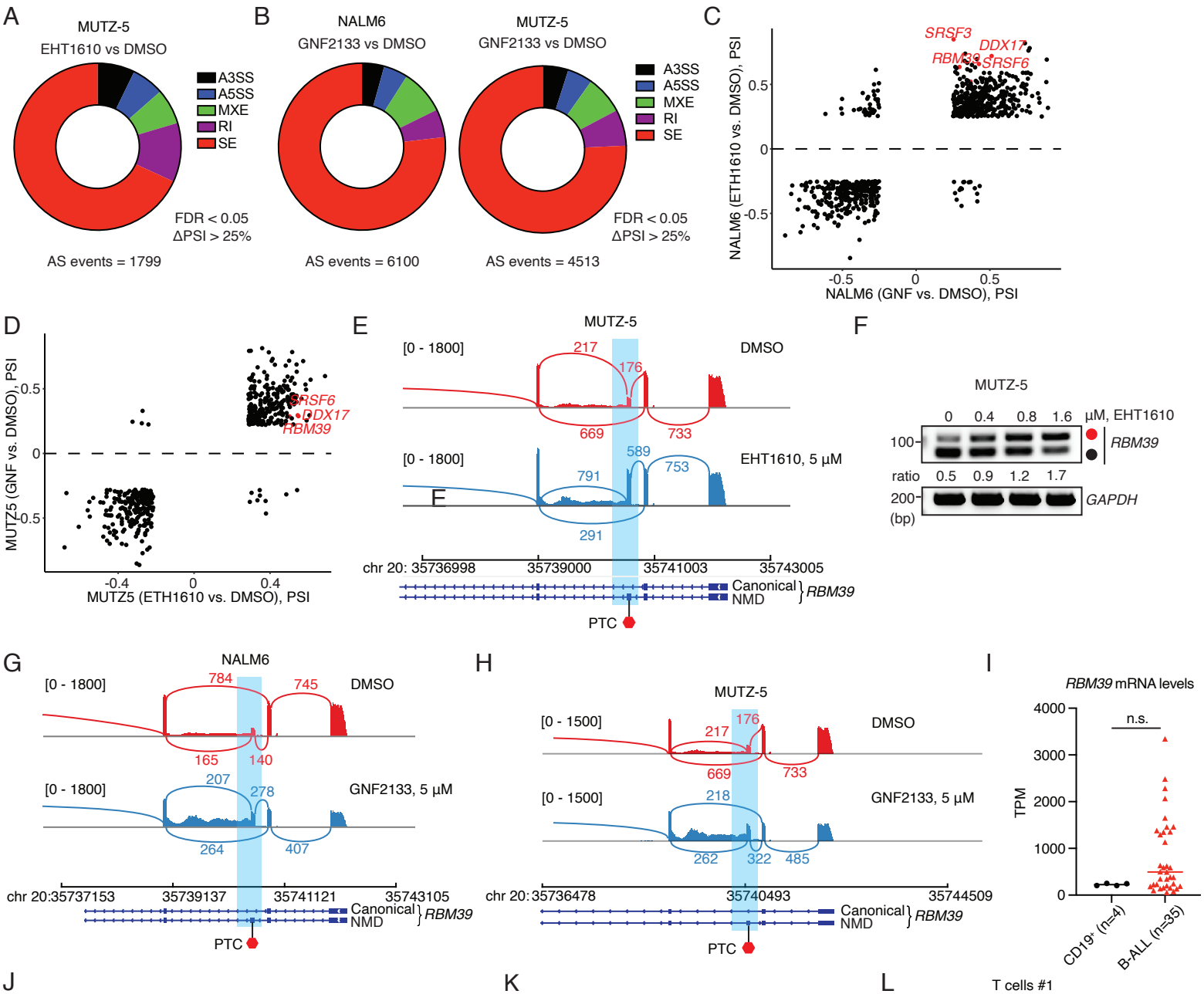
analysis of RBM39 in NALM6 cells treated with E7070 (500 nM), EHT1610 (5 μM) and the combination for 8 h. **E)** Synergy heatmaps for E7070 and EHT1610 treatment for 72 h in NALM6 cells.

Figure S9. A) Robust Rank Aggregation (RRA) score of genes from a pooled kinase domain-focused negative selection screen in NALM6 cells. Labeled genes are targets of dinaciclib. **B, C)** NALM6 (**B**) or MUTZ-5 (**C**) cells were treated with EHT1610 (4 μM) for 16h followed by additional treatment with dinaciclib (5 nM) for 4 h. PCR reactions were performed for detection of the alternative *RBM39* splicing event. **D)** Additional biological replicates for the western blot analysis shown in Fig. 7B (left), along with the quantification of the analyses (right). **E)** Synergy heatmaps for THAL-SNS-032 and EHT1610 treatment for 72 h in NALM6 cells. **F)** Western blot analysis of RBM39 in control or RBM39 overexpressing (O/E) NALM6 cells treated with dinaciclib (10 nM) and EHT1610 (4 μM) for 8 h. The experiments were performed with three biological replicates with quantification of RBM39/GAPDH (upper); a representative western blot is shown (lower). **G)** Cell viability (determined by %DAPI negative cells) analysis of two $CD34^+$ donor cells treated with the indicated dinaciclib and EHT1610 concentrations.

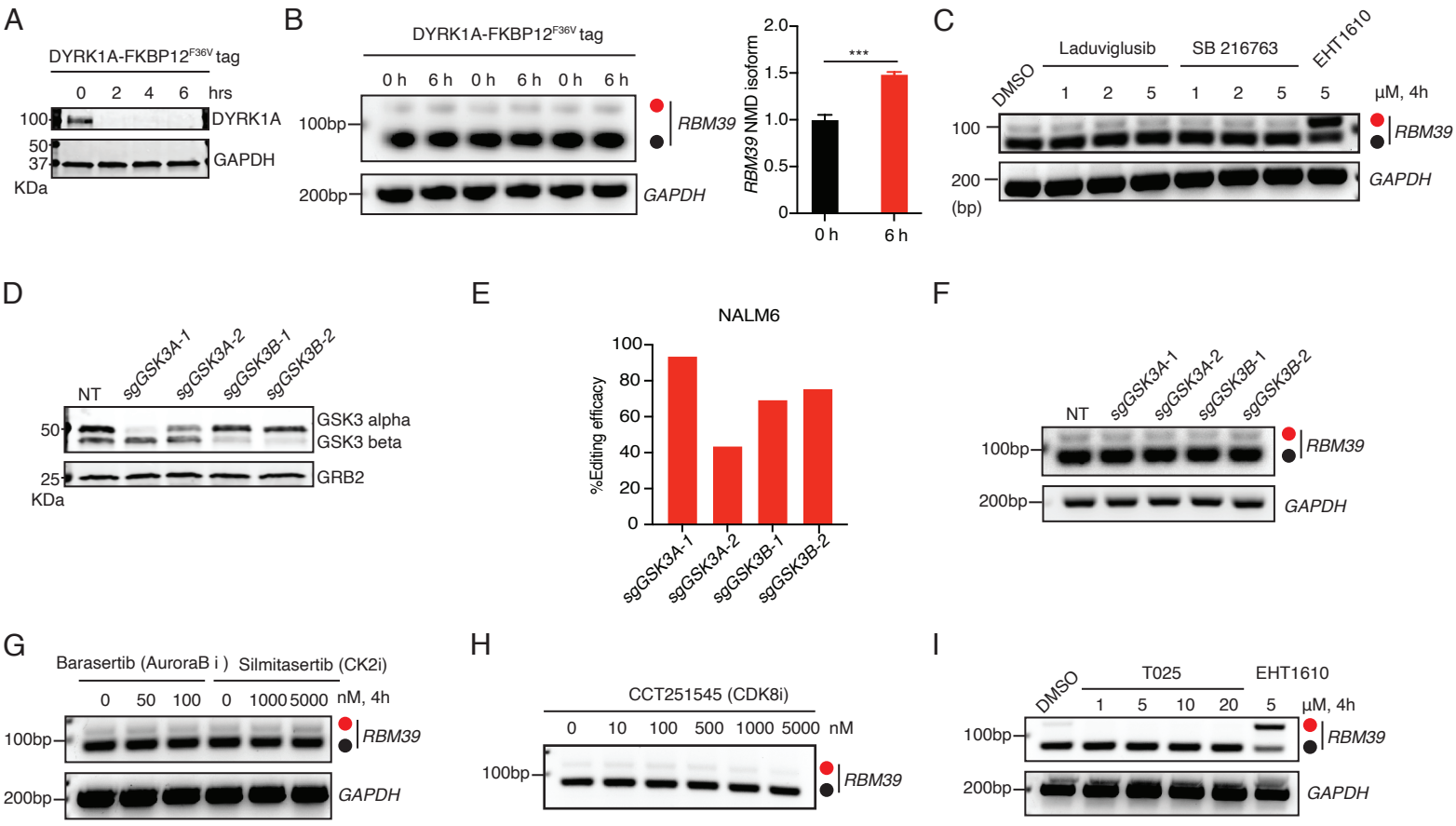
Figure S10. Synergy between EHT1610 and dinaciclib in B-ALL cells. A-G) Synergy heatmaps for EHT1610 and dinaciclib treatment for 3 days in NALM6 (**A**), MUTZ-5 (**B**) and 5 B-ALL PDX samples (**C-G**). ZIP analysis is shown.

Figure S11. Synergy between EHT1610 and dinaciclib in T-ALL patient samples. A-D) Synergy heatmaps for EHT1610 and dinaciclib treatment for 3 days in 4 T-ALL PDX samples. ZIP analysis is shown.

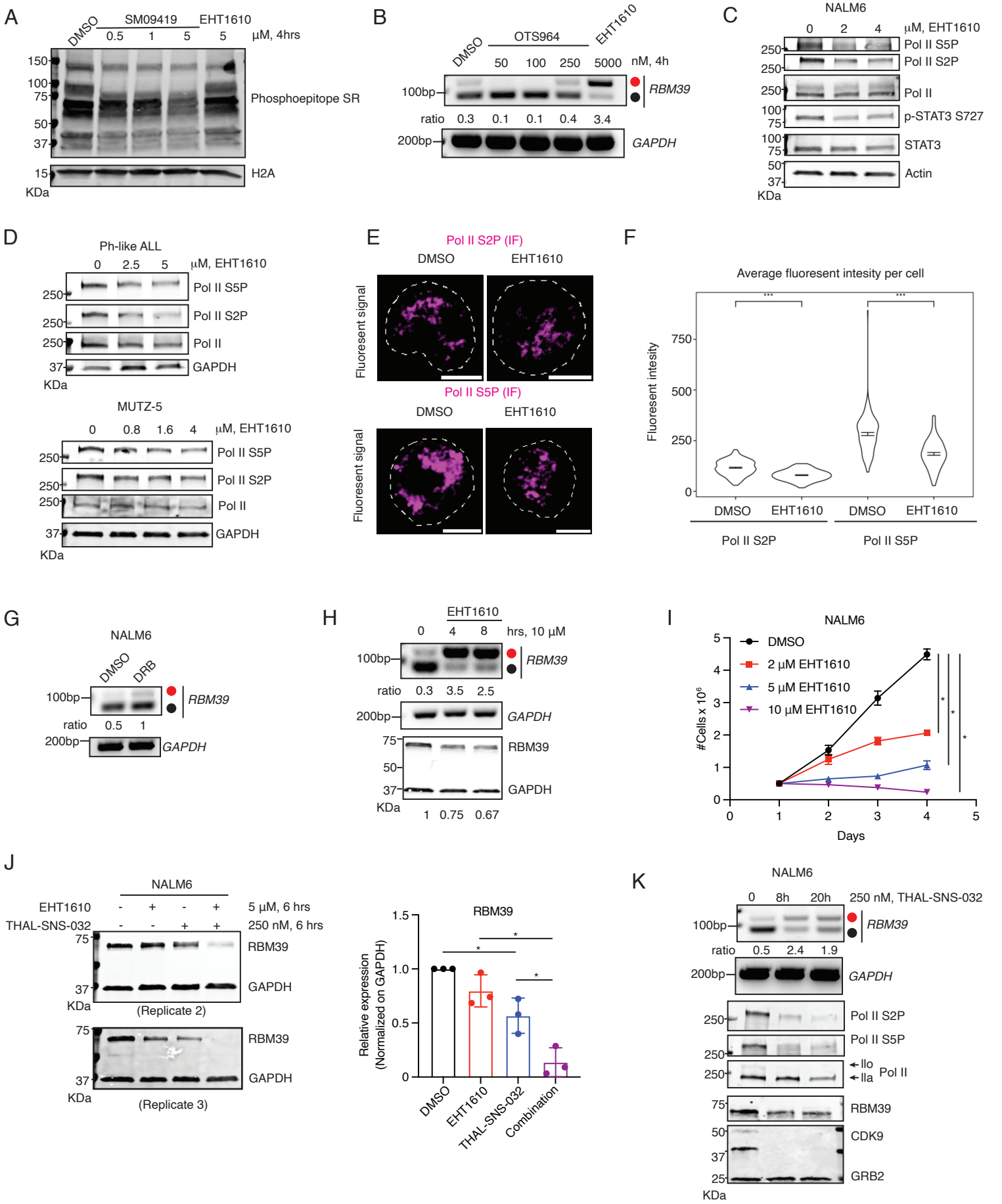
Supplementary Figure 1. Jin et al 2024



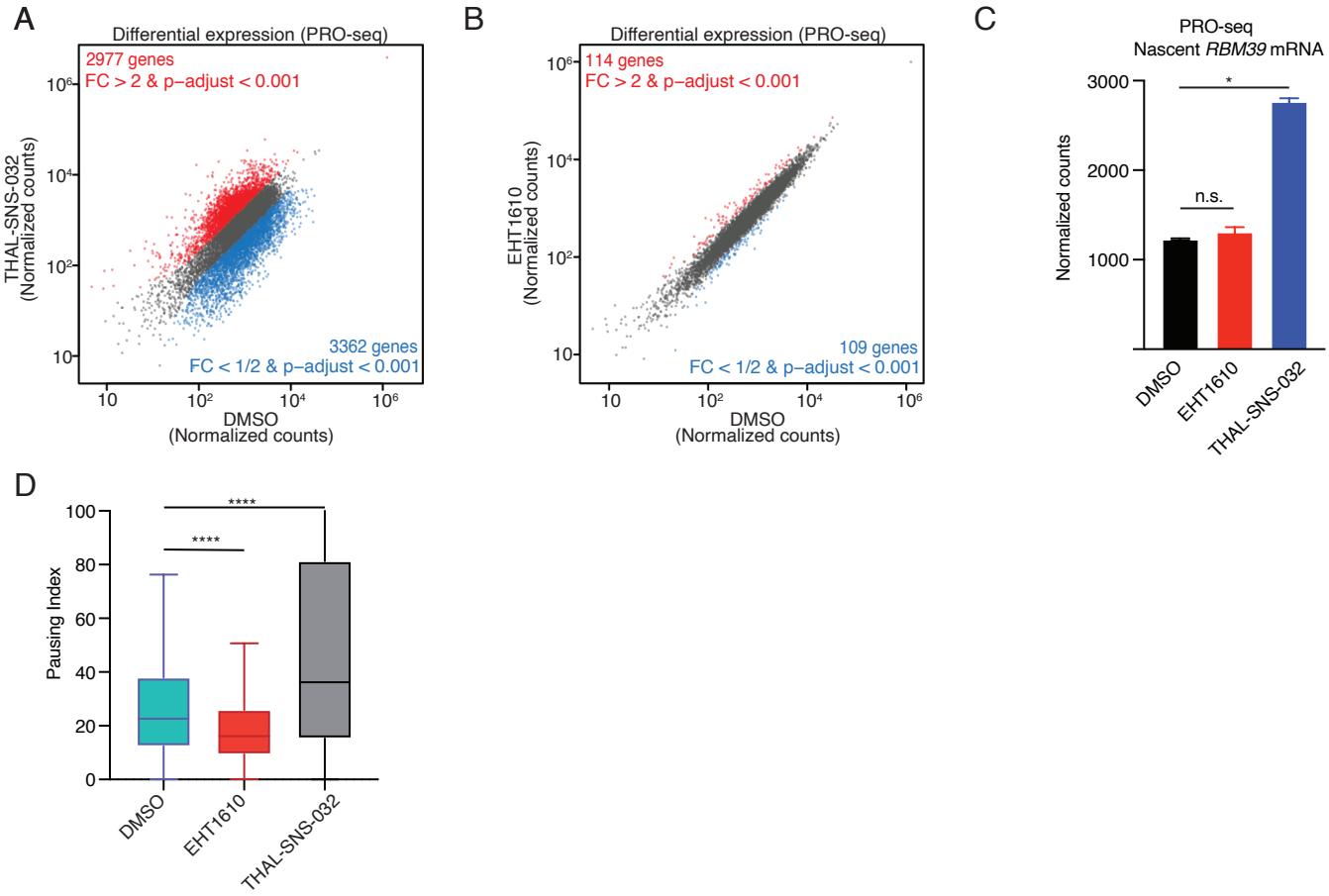
Supplementary Figure 2. Jin et al 2024



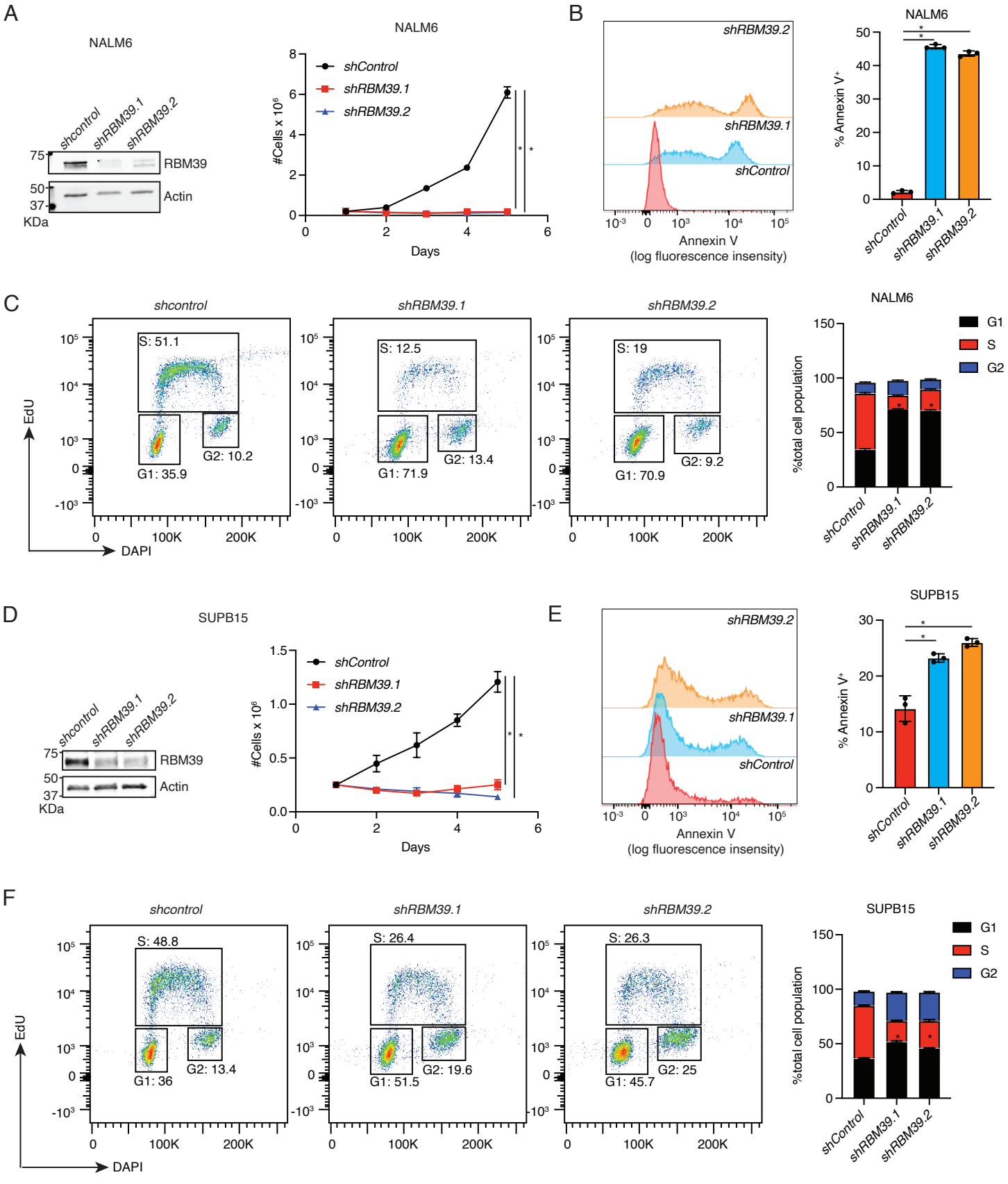
Supplementary Figure 3. Jin et al 2024



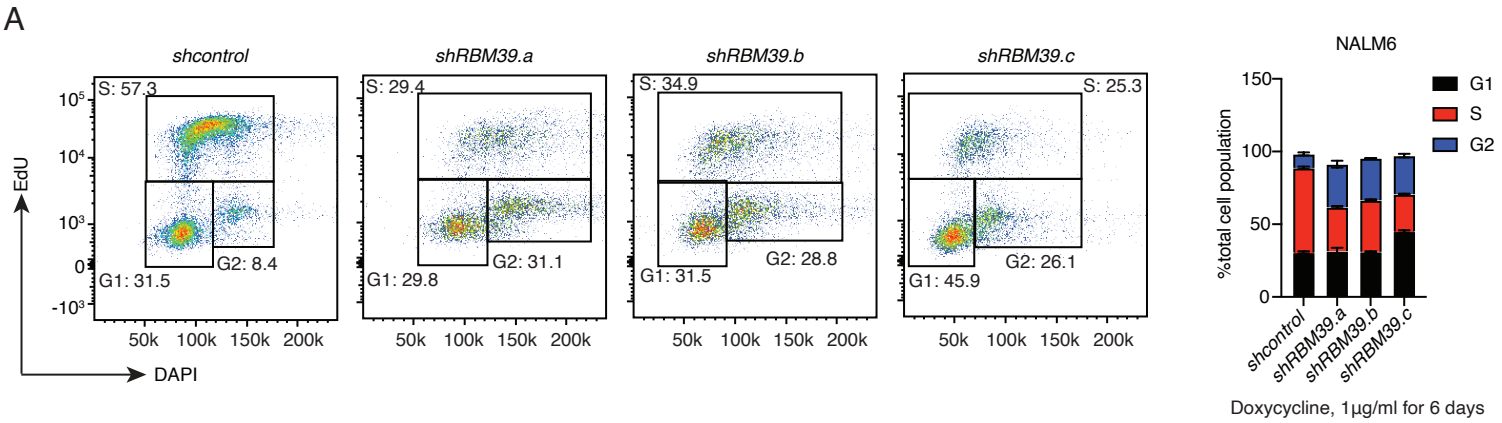
Supplementary Figure 4. Jin et al 2024

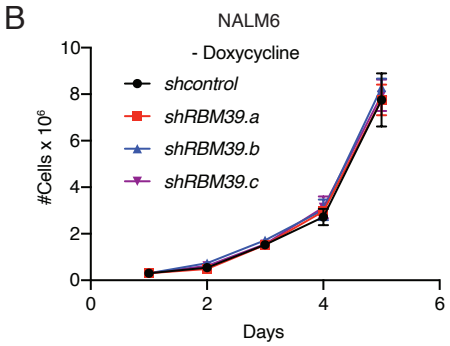


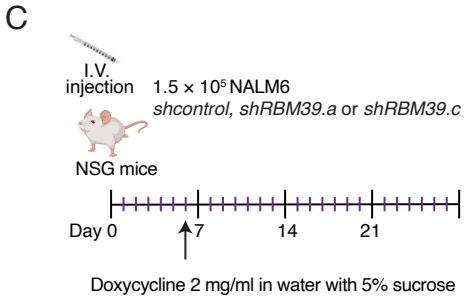
Supplementary Figure 5. Jin et al 2024



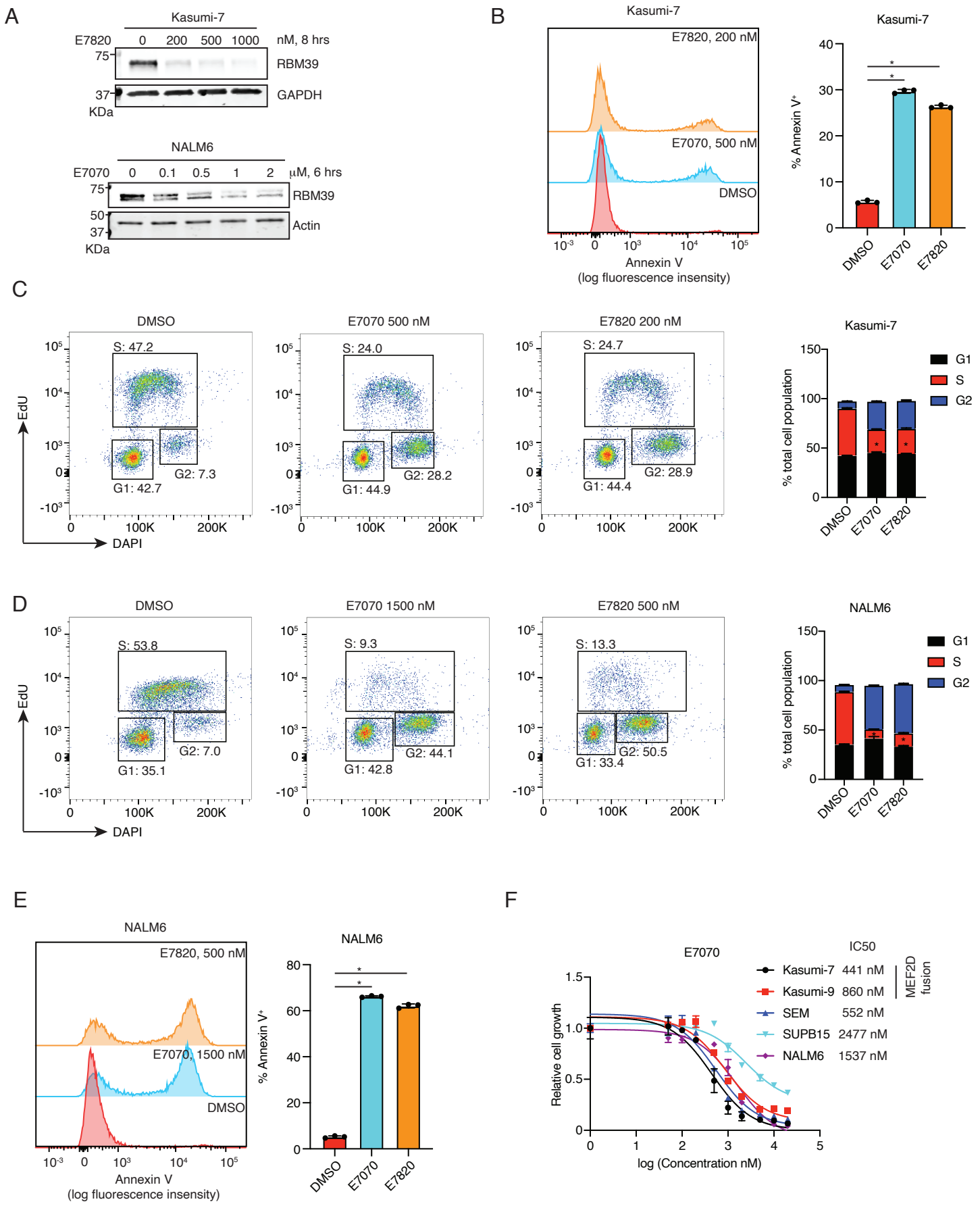
Supplementary Figure 6. Jin et al 2024





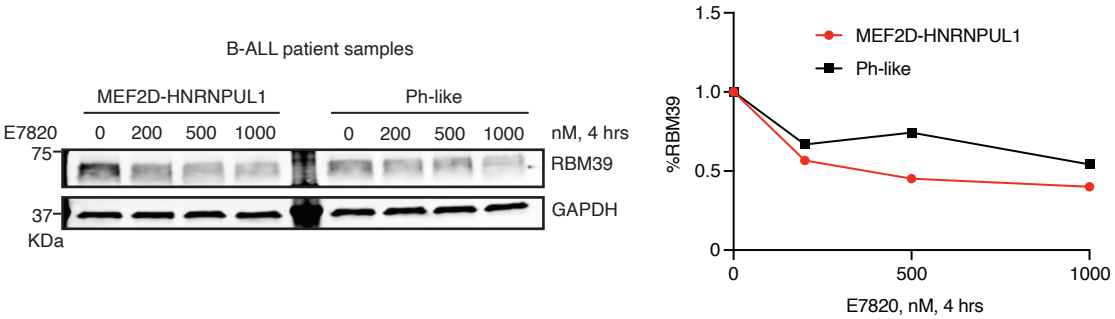


Supplementary Figure 7. Jin et al 2024

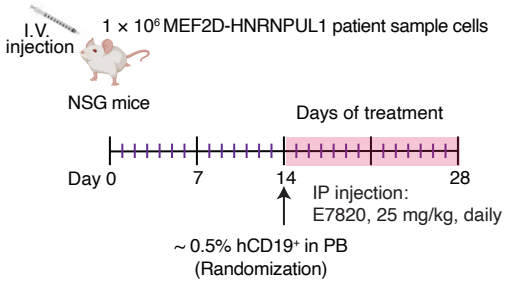


Supplementary Figure 8. Jin et al 2024

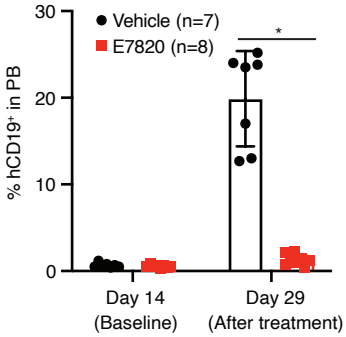
A



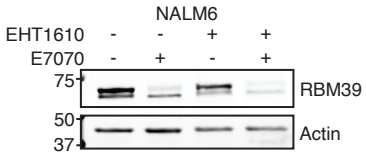
B



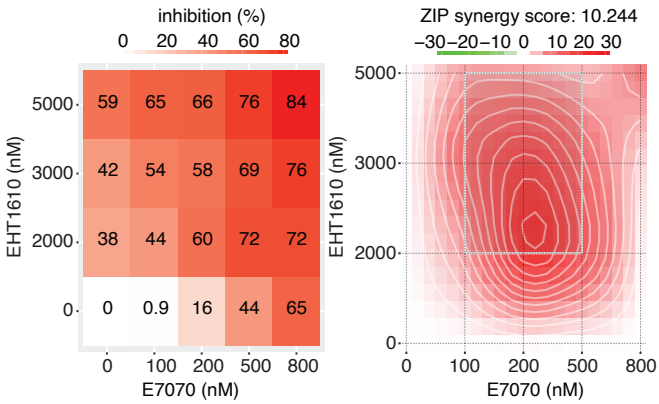
C



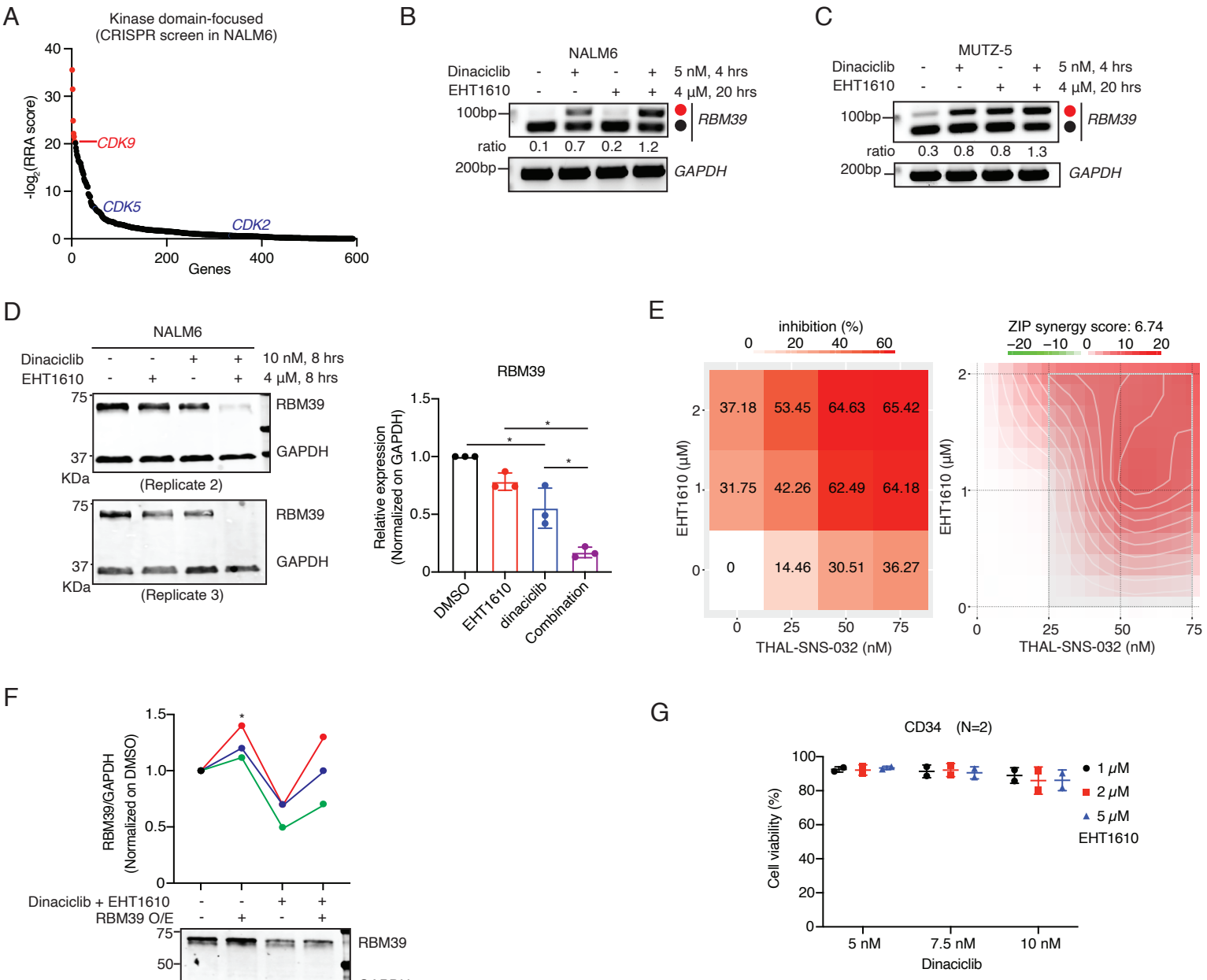
D



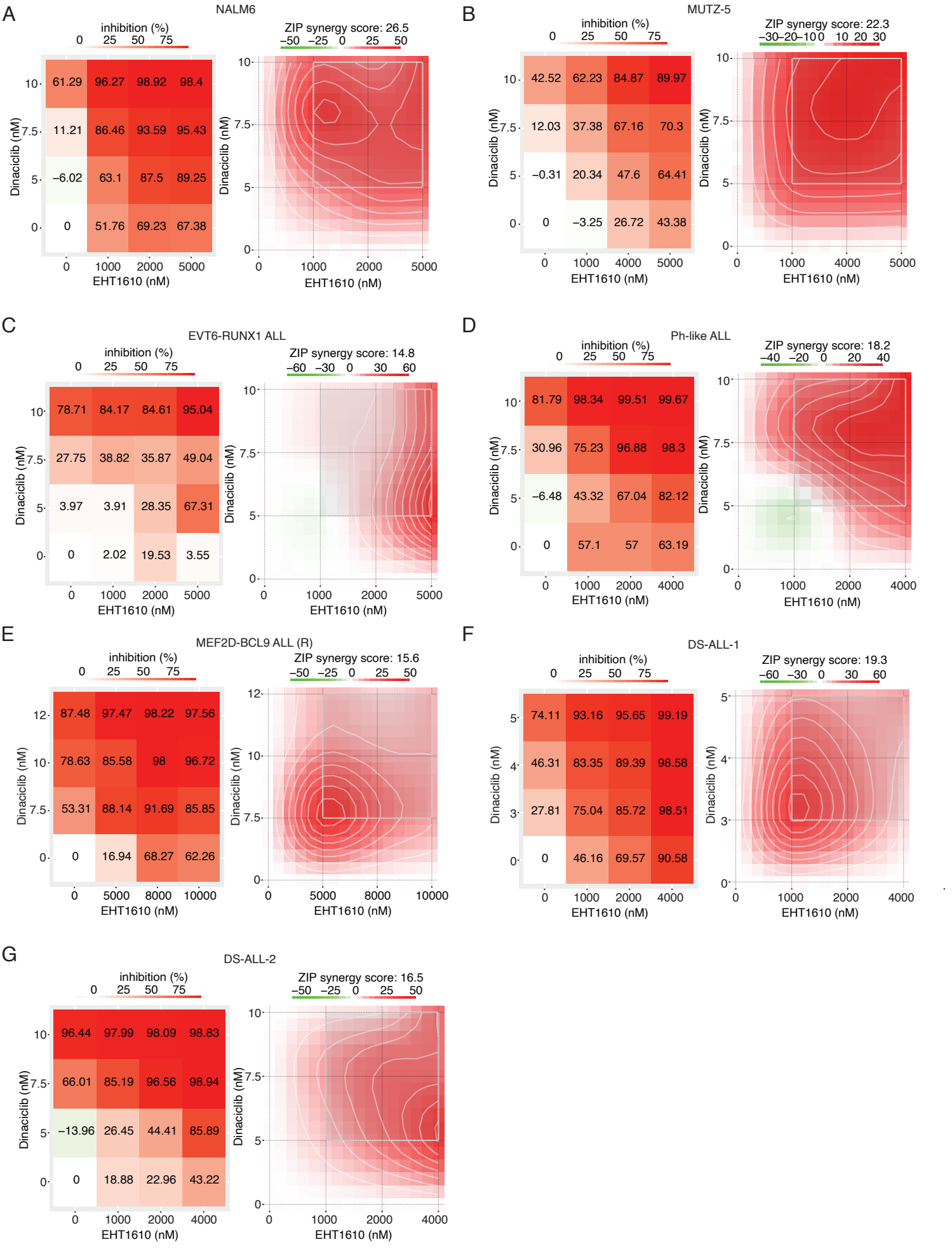
E



Supplementary Figure 9. Jin et al 2024



Supplementary Figure 10. Jin et al 2024



Supplementary Figure 11. Jin et al 2024