ETV6::ACSL6 translocation-driven super-enhancer activation leads to eosinophilia in acute lymphoblastic leukemia through IL-3 overexpression

Wenqian Xu,[1*] Feng Tian,[2*] Xiaolu Tai,[3] Gaoxian Song,[1] Yuanfang Liu,[1] Liquan Fan,[1] Xiangqin Weng,[1] Eunjeong Yang,[4] Meng Wang,[5] Martin Bornhäuser,[6] Chao Zhang,[3] Richard B. Lock,[7] Jason W.H. Wong,[4] Jin Wang,[1#] Duohui Jing[1#] and Jian-Qing Mi[1#]

[1]Shanghai Institute of Hematology, State Key Laboratory of Medical Genomics, National Research Center for Translational Medicine at Shanghai, Ruijin Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China; [2]Hebei Key Laboratory of Medical Data Science, Institute of Biomedical Informatics, School of Medicine, Hebei University of Engineering, Handan, Hebei Province, China; [3]Department of Plastic and Reconstructive Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [4]School of Biomedical Sciences, University of Hong Kong, Hong Kong, China; [5]Songjiang Research Institute, Songjiang District Central Hospital, Institute of Autism & MOE-Shanghai Key Laboratory for Children's Environmental Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China; [6]Medical Clinic I, University Hospital Carl Gustav Carus, TU Dresden, Dresden, Germany and [7]Children's Cancer Institute, Lowy Cancer Research Centre, School of Clinical Medicine, UNSW Medicine & Health, UNSW Centre for Childhood Cancer Research, UNSW Sydney, Sydney, New South Wales, Australia

WX and FT contributed equally as first authors.

#JW, DJ and J-QM contributed equally as senior authors.

Correspondence: D. Jing
jdh12262@rjh.com.cn

J. Wang
jinwang@shsmu.edu.cn

J-Q. Mi
jianqingmi@shsmu.edu.cn

Received: August 22, 2023.
Accepted: February 2, 2024.
Early view: February 15, 2024.

https://doi.org/10.3324/haematol.2023.284121

Supplementary Tables

The following tables were uploaded as individual Excel files.

Supplementary Table S1: List of SVs detected by ONT sequencing

Description of all SVs detected in the *ETV6::ACSL6* ALL cells, including the chromosome position, variation type, reference sequence, variant sequence, overall information and Variant Allele Frequency (VAF) for each SV. INS, insertion; DEL, deletion; INV, inversion; DUP, duplication; BND, breakend.

Supplementary Table S2: Genetic aberrations and functional annotations by RNA sequencing and ONT sequencing

Lists of recurrent leukemia-associated mutations detected by RNA-seq as well as the functional annotation of somatic SVs detected by ONT sequencing. In the left "RNA sequencing" section, we listed the results of some recurrent gene variations in leukemia. In the right "ONT sequencing" section, we showed the somatic SVs after filtering out germline variations, including variation types, chromosomal positions, VAF values, functional regions, and their adjacent genes.

Supplementary Table S3: Patient information

Description of ALL patients in the study. Bone marrow samples from five B-ALL patients were used in the study. The table showed their basic information, including gender, age at diagnosis, diagnosis date, initial white blood cell count, cytogenetics, and risk stratification.

Supplementary Methods

Nanopore long-read sequencing

High-molecular-weight (HMW) DNA was extracted from *ETV6::ACSL6* ALL cells with the MagAttract HMW DNA kit (Qiagen; 67563) according to the manufacturer's instructions. Briefly, 1×10^7 frozen cells were lysed with 200 μL of buffer ATL and 20 μL of proteinase K and incubated overnight at 56 °C and 900 r.p.m. 4 μL of RNase A was added to cleave RNA. 150 μL of buffer AL, 280 μL of buffer MB, and 40 μL of MagAttract Suspension G beads were then added to capture the HMW DNA. Next, the beads were cleaned up with 700 μL of buffer MW1, buffer PE, and NFW and eluted with 200 μL of buffer AE. The HMW was sequenced on a MinION Mk1b device (ONT).

Read sequences were extracted from base-called FAST5 files by Poretools (version 0.5.1) to generate a FASTQ file. Raw reads were aligned to hg38 with minimap2 (version 2.24). Structural variants were called using cuteSV (version 1.0.8) and Sniffles2 (version 2.0.7). Filtering steps of structural variants are described below: (1) Structural variants (chr1-22, X, and Y) with Allele Frequency > 0.3, Reads >= 10, and marked as PRECISE were included using snpSift (version 5.1d); (2) After depth information was added to individual structural variant by uphold (version 0.2.3), deletions with DHFFC > 0.7 and duplications with DHBFC <1.3 were excluded by bcftools (version 1.9); (3) Structural variants from different callers were merged by SURVIVOR (version 1.0.7); (4) Insertions and deletions in the merged call set overlapping with Simple repeat and Microsatellite regions from UCSC genome browser were removed by bedTools (version 2.30.0); (5) All structural variants in the merged call set overlapping with gnomAD structural variants were removed by bedTools (version 2.30.0). Genomic coordinates (hg19 to hg38) of gnomAD structural variants were lifted over using CrossMap (version 0.6.3).

RNA-seq

Total RNA was extracted and reverse transcribed using Yeasen. The cDNA was synthesized and added with adaptors. The amplified sample libraries were then paired-end sequenced (2×150 bp) on the Novaseq 6000 system and aligned against the human genome (hg19). Other ALL datasets were generated from previous publications(1-3) (accession no. GSE57795, GSE207057, and GSE164072).

Hi-C Sequencing

Hi-C was performed using a modified version of a previously described protocol(4). Cells were crosslinked with formaldehyde, and DNA was digested with MboI (NEB, R0147) that leaves a 5′overhang; the 5′overhang was filled, including a biotinylated residue; and the resulting blunt-end fragments were ligated under dilute conditions that favor ligation events between the cross-linked DNA fragments. The resulting DNA sample contained ligation products consisting of fragments that were originally in close spatial proximity in the nucleus, marked with biotin at the junction. A Hi-C library was created by shearing the DNA and selecting the biotin-containing fragments with streptavidin beads. The library was then analyzed by using massively parallel DNA sequencing.

CUT&Tag library preparation

CUT&Tag assays were performed as described previously(5). In brief, approximately

10,000 PDX cells in each biological replicate were harvested and centrifuged (400 × g) at room temperature for 5 min. The supernatant was washed and resuspended in Wash Buffer supplemented with protease inhibitors (Sigma). Concanavalin A-coated magnetic beads in each sample were washed and resuspended in binding buffer. Then beads were added to the cells, gently vortexed, and incubated in a shaker for 10 min at room temperature. The unbound supernatant was removed. The bead-bound cells were resuspended in precooling antibody buffer [2 mM EDTA, 0.1% BSA in DIG Wash Buffer (0.05% digitonin in wash buffer)] and incubated with (1:50 dilution) primary antibody against H3K27ac (ab177178, Abcam), H3K4me1 (ab176877, Abcam), H3K4me3 (ab213224, Abcam), p300 (ab275378, Abcam) and BRD4 (ab128874, Abcam) in a shaker overnight at 4°C. The primary antibody on the magnet stand was removed, and then a secondary antibody (Goat Anti-Rabbit IgG antibody, Vazyme) was diluted (1:100) in DIG Wash buffer and incubated with cells at room temperature for 1 h. Cells were washed in DIG Wash buffer using the magnet stand to remove unbound antibodies. A dilution of hyperactive pA-Tn5 transposon complex (0.04 μM) was prepared in DIG-300 buffer supplemented with 0.01% digitonin and protease inhibitors. Then cells were incubated with pA-Tn5 transposon complex in a shaker at room temperature for 1 h. Subsequently, cells were resuspended in tagmentation buffer (10 mM MgCl2 in DIG-300 buffer) and incubated at 37°C for 1 h. To terminate tagmentation, we added 10 μl of 0.5 M EDTA, 3 μl of 10% SDS, and 2.5 μl of 20 mg/ml proteinase K to each sample, followed by overnight incubation at 37°C. Purified DNA was amplified and indexed. The libraries were cleaned using Agencourt AMPure XP (Beckman) and sequenced on the Illumina Novaseq 6000 (Illumina). We applied 150-bp pair-end sequencing with a sequencing depth of 6G base pair raw data (generated approximately 20 million mapped paired reads).

Overexpression of *ETV6::ACSL6*

Human *ETV6::ACSL6* fusion cDNA was cloned into a lentiviral expression vector (GL121, pSLenti-EF1-EGFP-CMV-MCS-WPRE). Cells were transduced with lentivirus and then sorted to obtain GFP^+ cells.

Cytotoxicity Assays

PDX cells were seeded in 96-well U bottomed plates at 3,500,00 cells per 100 μl medium per well. JQ1 (Selleck) or vehicle control was added in triplicate wells. The cells in triplicate wells were incubated for a series of time points. Following the incubation, cck8 reagent (Yeasen) was added. After an additional 4 h incubation, fluorescence was measured, and cell viability was expressed relative to vehicle-treated

control cells. The additive effect of combination treatment (drugs 1 and 2) was calculated as: [relative survival fraction drug 1×relative survival fraction drug 2]. The *p* values were calculated by comparing the combination treatment to the presumed additive effect at each time point.

Cytometric Beads Array for Detecting Cytokines Levels

The level of IL-3, IL-5, and GM-CSF in the supernatant of cultured Nalm6 cells was determined using Cytometric Beads Array (BD Pharmingen) according to the manufacturer's instructions. Data were acquired on FACS Aria-II (BD Bioscience). Data were analyzed using FCAP Array software (BD Bioscience).

Luciferase Reporter Assay

The luciferase reporter assay was performed using pGL3P vectors from Promega (Madison, Wisconsin, USA), and luminescence was detected using the Dual-Glo system from Vazyme. Gene enhancer sequences were synthesized as dsDNA by IDT. The dsDNA was inserted into the pGL3P vector between Sal1 and BamH1 cutting sites, which is 1 kb downstream of the 3' end of the firefly luciferase gene. The cloned vector was co-transfected with pRL-TK renilla luciferase control reporter vectors (Promega) into 293T cells. The firefly and renilla luminescence were detected in the transfected cells after 24 hours of transfection using the Dual-Glo Luciferase Assay System (Vazyme). The firefly luminescence was normalized to renilla luminescence for each condition. Fold inductions were then calculated by normalizing to pGL3P control.

Western blotting

Radio immunoprecipitation (RIPA) lysis buffer (Yeasen) was used for the lysing of Nalm6 cells. Cell lysates were subjected to 12.5% SDS-polyacrylamide gel electrophoresis and transferred onto polyvinylidene difluoride membranes. Then, the membranes were incubated with the indicated antibodies. Primary antibodies against ETV6, flag, IL3, and GAPDH were purchased from Abcam. The secondary antibodies conjugated with HRP were used at 1:2,000 dilutions (CST). The protein bands were detected with the enhanced chemiluminescence reagent.

Hi-C data processing

We first trimmed the adapters of the Hi-C raw FASTQ files and then mapped the trimmed files against the hg19 human reference genome using the runHi-C pipeline, which is based on the 4DN consortium. Specifically, Burrows-Wheeler Aligner was used for the FASTQ file alignment and aligned reads with low quality, and PCR

duplicates were filtered. Aligned reads were then paired on the basis of read pairs and filtered for fragments that contained ligations of at least two different restriction fragments. These reads were then binned at 5-kb resolution. To generate the contact matrix at multiple resolutions (5, 10, 25, 40, 50, 100, 250, 500 kb, 1, 2.5, 5, and 10 Mb), we used the run-cool2multirescool script from 4DN consortium, which performed the ICE normalization at the same time. We used Coolbox to visualize ICE-normalized genomic Hi-C data. The juicer tool was also used to generate multiresolution .hic files, which can be visualized using Juicebox.

RNA extraction

RNA was extracted using a Yeasen MolPure Cell RNA Kit. according to the manufacturer's instructions. The resulting RNA was stored at -80℃ until further use.

RT-qPCR

Extracted RNA was quantified using a Nanodrop Spectrophotometer. 1μg RNA was used to generate cDNA using the HiScript III RT SuperMix for qPCR (Vazyme) according to the manufacturer's instructions. Reactions were performed in technical duplicates using the ChamQ SYBR qPCR Master Mix (Vazyme) and indicated primer pairs on a machine (Applied Biosystems).

RT-qPCR primers

Name	**Sequence 5'-3'**
IL3-F	CGCGGATCCAAACATGAGCCGCCT
IL3-R	GCTCTAGATGCCGCAGGAAAAGGTGAA
IL5-F	ATGAGGATGCTTCTGCATTTG
IL5-R	TCAACTTTCTATTATCCACTCGGTGTTCATTAC
P4HA2-F	CAAACTGGTGAAGCGGCTAAA
P4HA2-R	GCACAGAGAGGTTGGCGATA
Actin-F	TATGAGCTGCCTGACGGC
Actin-R	CAGCAATGCCTGGGTACA

Statistics

The data presented in this study are expressed as the mean of at least three independent experiments ± SEM. Statistical analysis was carried out using GraphPad Prism 9 software, with unpaired two-sided Student's t-tests, one-way analysis of variance (ANOVA) with Kruskal-Wallis test with Dunn's multiple comparisons test, as

appropriate. Survival analysis was performed using Gehan-Breslow-Wilcoxon test. Statistical significance was considered to be achieved when $p < 0.05$. All experiments were repeated a minimum of three times unless otherwise indicated.

Data availability

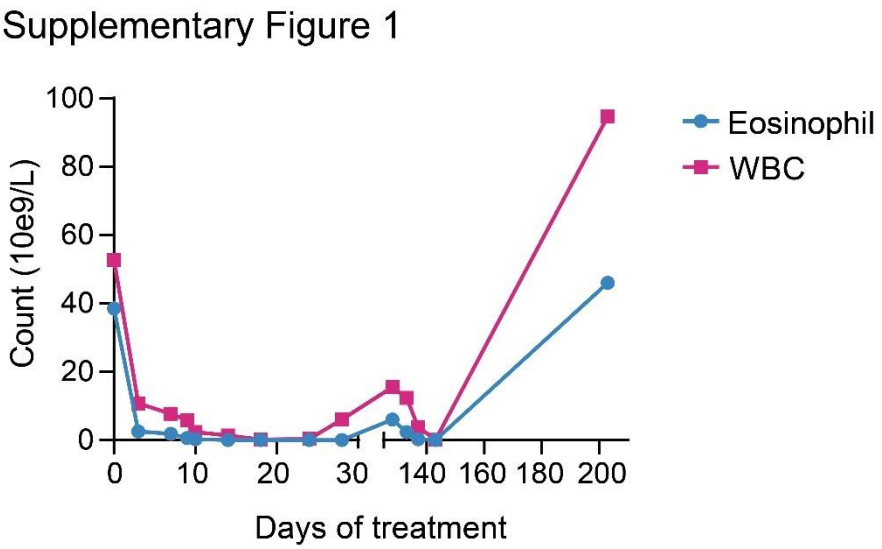
The raw sequencing data from this study have been deposited in the Genome Sequence Archive(6) in National Genomics Data Center(7), Beijing Institute of Genomics (BIG), Chinese Academy of Sciences, under the accession number: HRA004277, which is publicly accessible at https://ngdc.cncb.ac.cn/gsa.

We also host a UCSC browser session for easy access and viewing of genome-wide mapping at:

http://genome-asia.ucsc.edu/s/xuwenqian/RJ9_data

All other relevant data that support the conclusions of the study are available from the authors on request. Please contact jdh12262@rjh.com.cn.

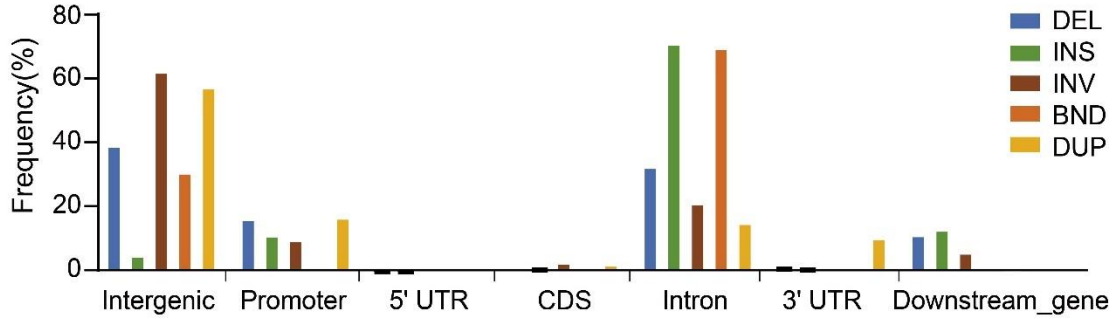
Supplementary Figures



Supplementary Fig. S1 Correlation between eosinophils and leukocytes in the patient sample.

The *x* axis represents the treatment days of the *ETV6::ACSL6* ALL patient, and the *y* axis represents the cell count in the peripheral blood of the patient.

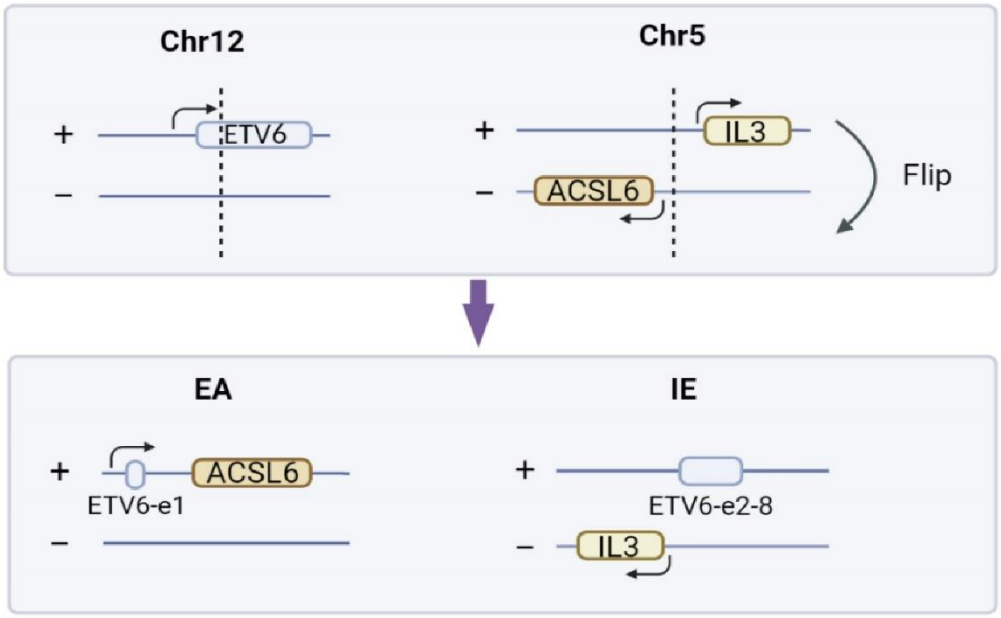
Supplementary Figure 2



Supplementary Fig. S2 ONT detecting structural variations in *ETV6::ACSL6* ALL

Breakpoint enrichment from each SV type distributed in different genomic regions. DEL, deletion; INS, insertion; INV, inversion; BND, breakpoint end of translocation; DUP, duplication.

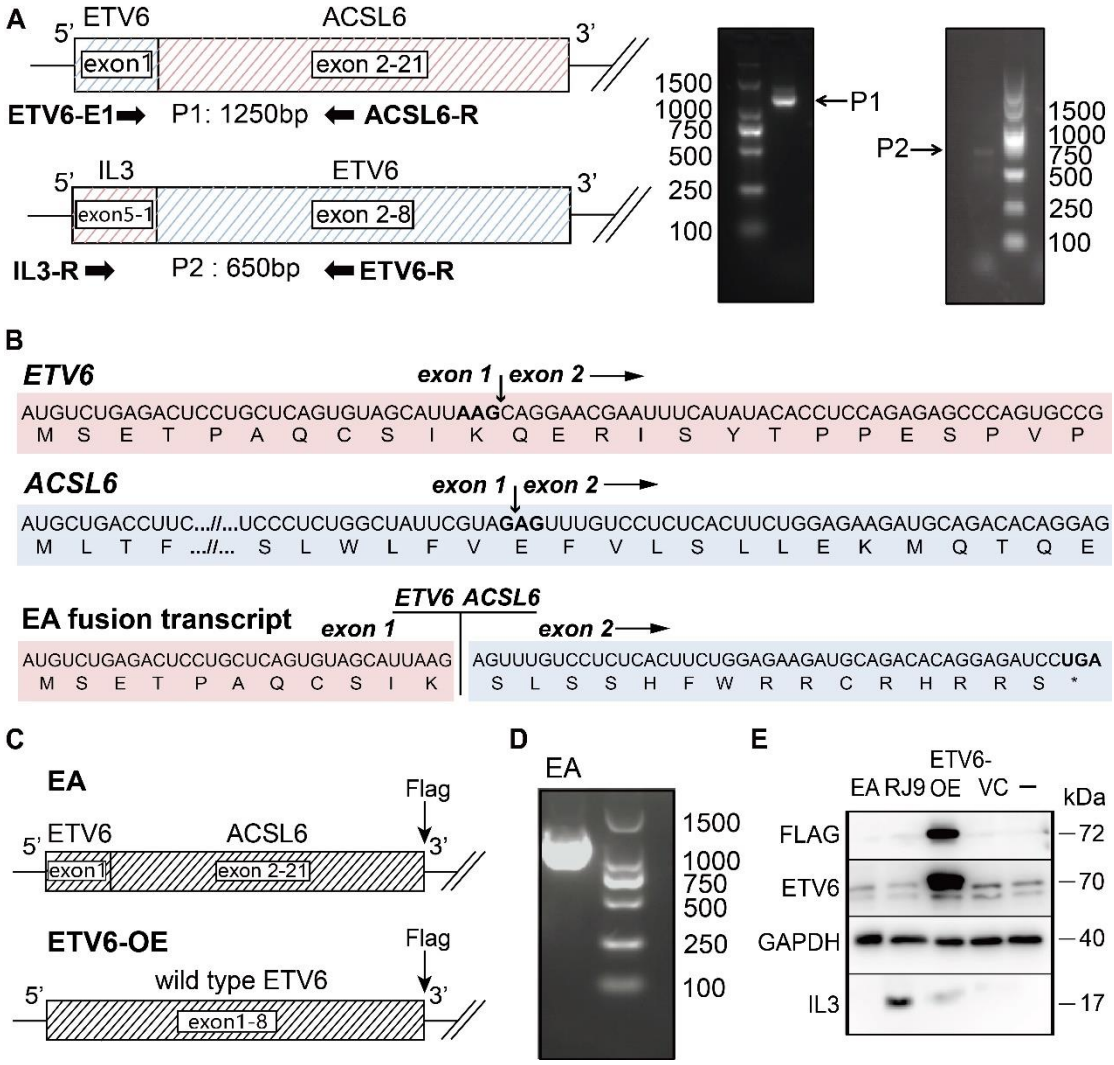
Supplementary Figure 3



Supplementary Fig. S3 Graphical illustration of newly formed chromosomes

Blue square, *ETV6*, yellow square, *IL-3*, orange square, *ACSL6*. e, exon.

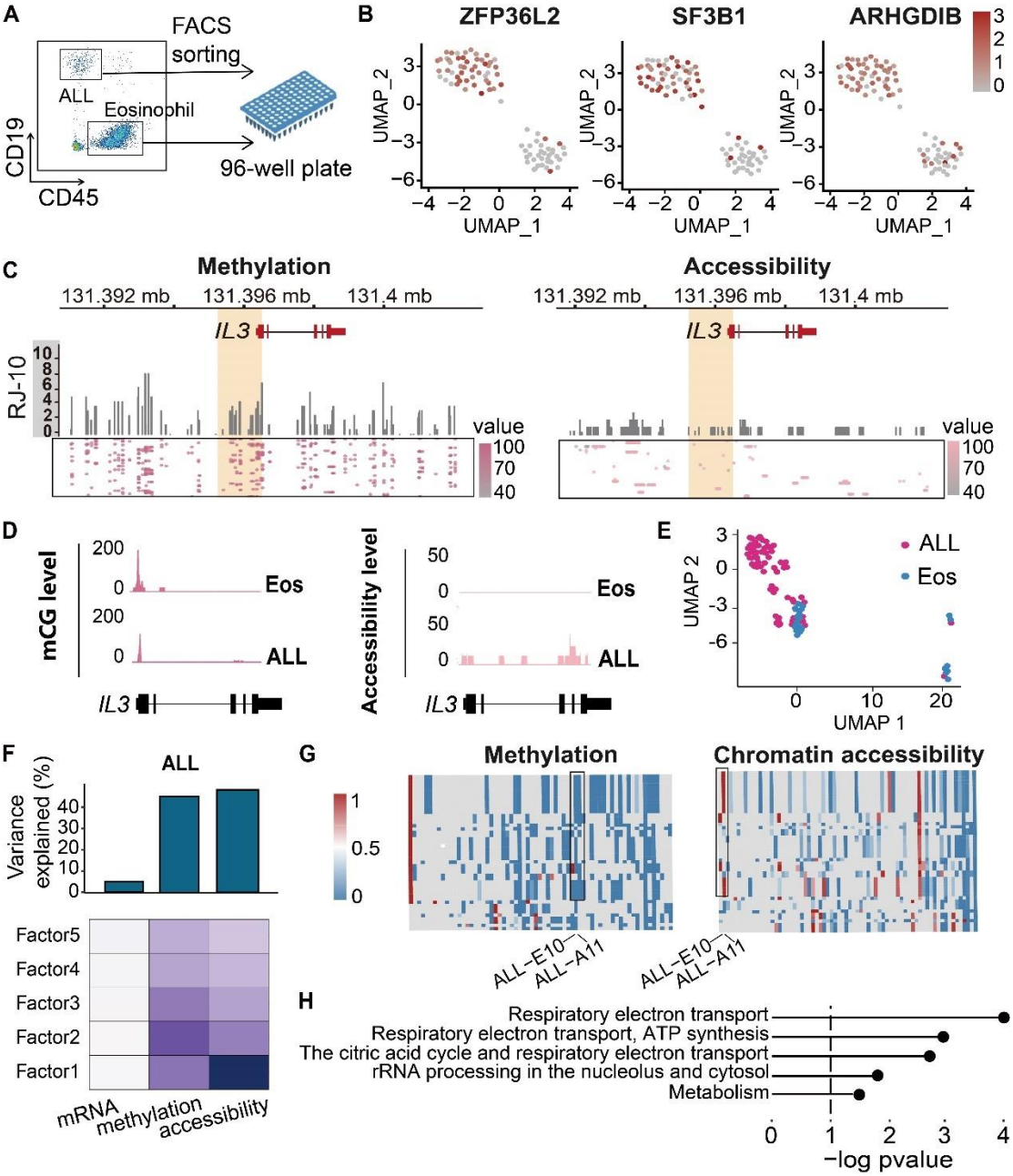
Supplementary Figure 4



Supplementary Fig. S4 Deciphering *ETV6::ACSL6* at the RNA level

(A) Schematic diagram of the fusion of two chromosomes at the cDNA level. Primers are shown in black arrows. Right, electropherogram of PCR products. (B) mRNA of *ETV6*, *ACSL6* and the fusion gene. In each box, the upper part represents the codon sequence, and the lower part represents the corresponding amino acids. Breakpoints at the mRNA level are indicated by vertical arrows. On *ETV6* and *ACSL6*, codons near the breakpoints are highlighted in bold. The premature stop codon is indicated in bold, and an asterisk represents transcription termination. (C) A schematic diagram of the overexpression vector in Nalm6. Upper part, 3×Flag is inserted upstream of the 3' UTR of the fusion gene; Lower part, overexpression of wild-type *ETV6*, with 3×Flag inserted at 3' end. OE, overexpression. (D) Electropherogram of PCR products validating expression of *ETV6::ACSL6* transcript. (E) Western blot analysis of IL3, Flag and ETV6 protein levels of RJ-9 PDX cells and Nalm6 after transfection. ETV6-OE, overexpression of wild-type *ETV6* with 3×Flag; VC, vehicle control; -, non-transfected Nalm6 cells.

Supplementary Figure 5

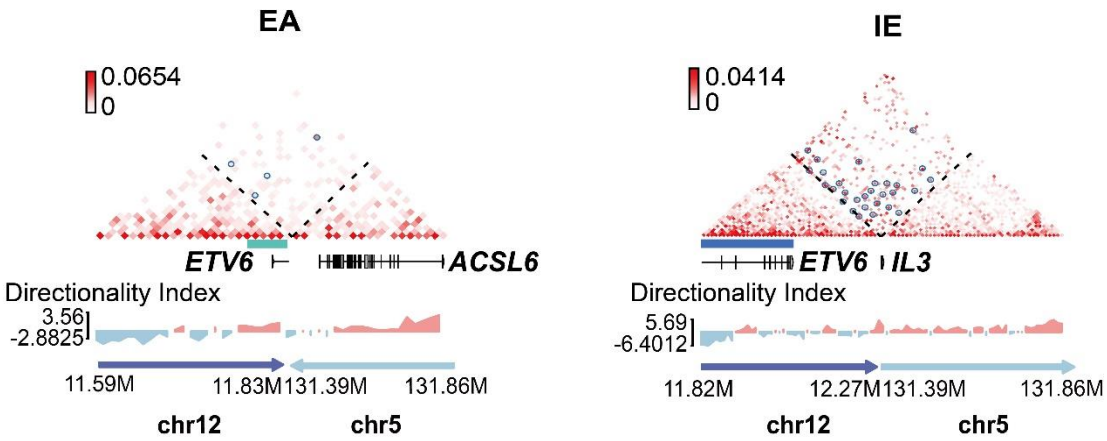


Supplementary Fig. S5 Single-cell sequencing dissecting heterogeneity within ALL cells.

(A) Single cells of blasts and eosinophils were sorted using FACS into 96-well plates. Cells shown were gated on live singlets. (B) UMAP dimensionality reduction displaying scRNA data of indicating genes. (C) Left, single-cell methylation data at IL-3 loci. Right, single-cell accessibility data at IL-3 loci. Bar plot showing the average level of all cells detected at indicated genomic loci. Each row in the box corresponds to a single cell. Each dot represents the methylation or accessibility level of the corresponding cell (horizontally) at the corresponding genomic position (vertically). RJ-10, a normal karyotype ALL patient. (D) Left, coverage tracks of single-cell methylation (mCpG) data at the IL-3 locus. Right, coverage tracks of single-cell accessibility data (GpCm) at the IL-3 locus. (E) UMAP projection based

on the MOFA factors that integrated RNA-seq, mCpG and GpCm datasets from ALLs and eosinophils (n=89). Cells are colored based on cell types. Pink dots, ALL cells; Blue dots, eosinophils. (F) Percentage of variance explained (R^2) by each MOFA factor (bottom rows) across all three layers. Top bar plot, the sum of all factors as a proportion. (G) Heatmap of methylation data in MOFA factor2 and chromatin accessibility data in MOFA factor1. Gray box, data not available. (H) GSEA analysis of characteristic genes of MOFA factor1 and MOFA factor2, overlapping pathway of two MOFA factors are shown.

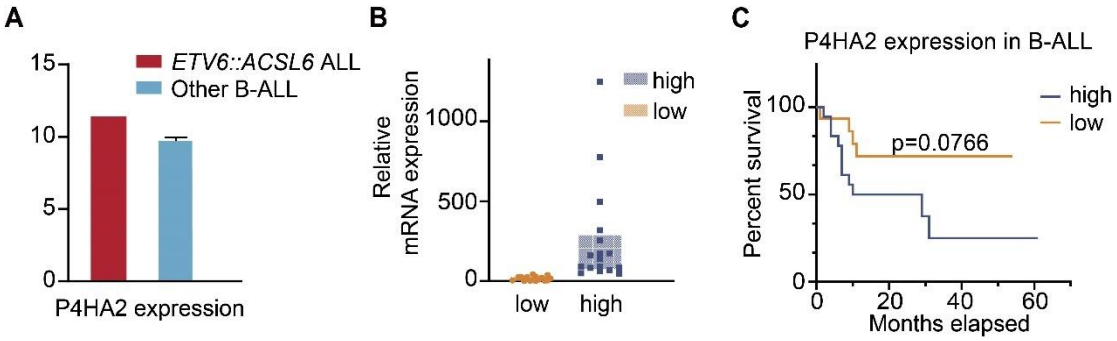
Supplementary Figure 6



Supplementary Fig. S6 Neo-TADs in *ETV6::ACSL6* ALL cells

Reconstructed Hi-C map indicating two neo-TADs following translocation. Neo-loops are showed in blue circles.

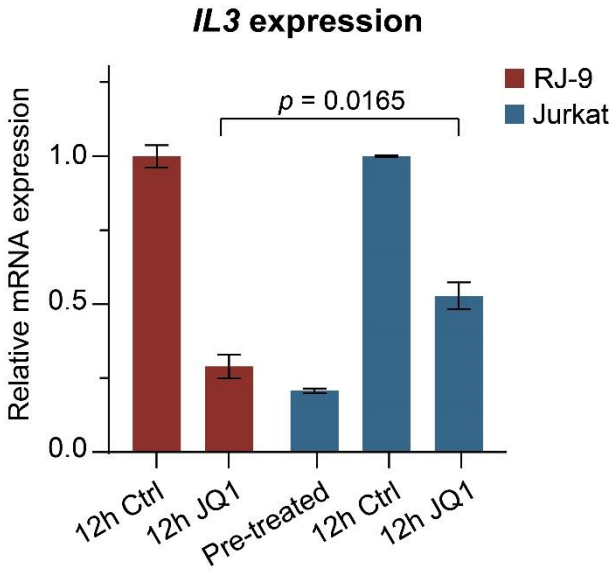
Supplementary Figure 7



Supplementary Fig. S7 P4HA2 expression level and outcome of ALL patients.

(A) Expression analysis of *P4HA2* in *ETV6::ACSL6* ALL as compared with other 31 B-ALL patients. Gene expression was quantified by RNA-seq (log2 of counts normalized using DESeq2's median of ratios). (B) RT-qPCR validation of *P4HA2* expression in different ALL cases. Comparison analysis was performed by the Wilcoxon rank-sum test. (C) Kaplan-Meier curves for B-ALL patients grouped by *P4HA2* expression.

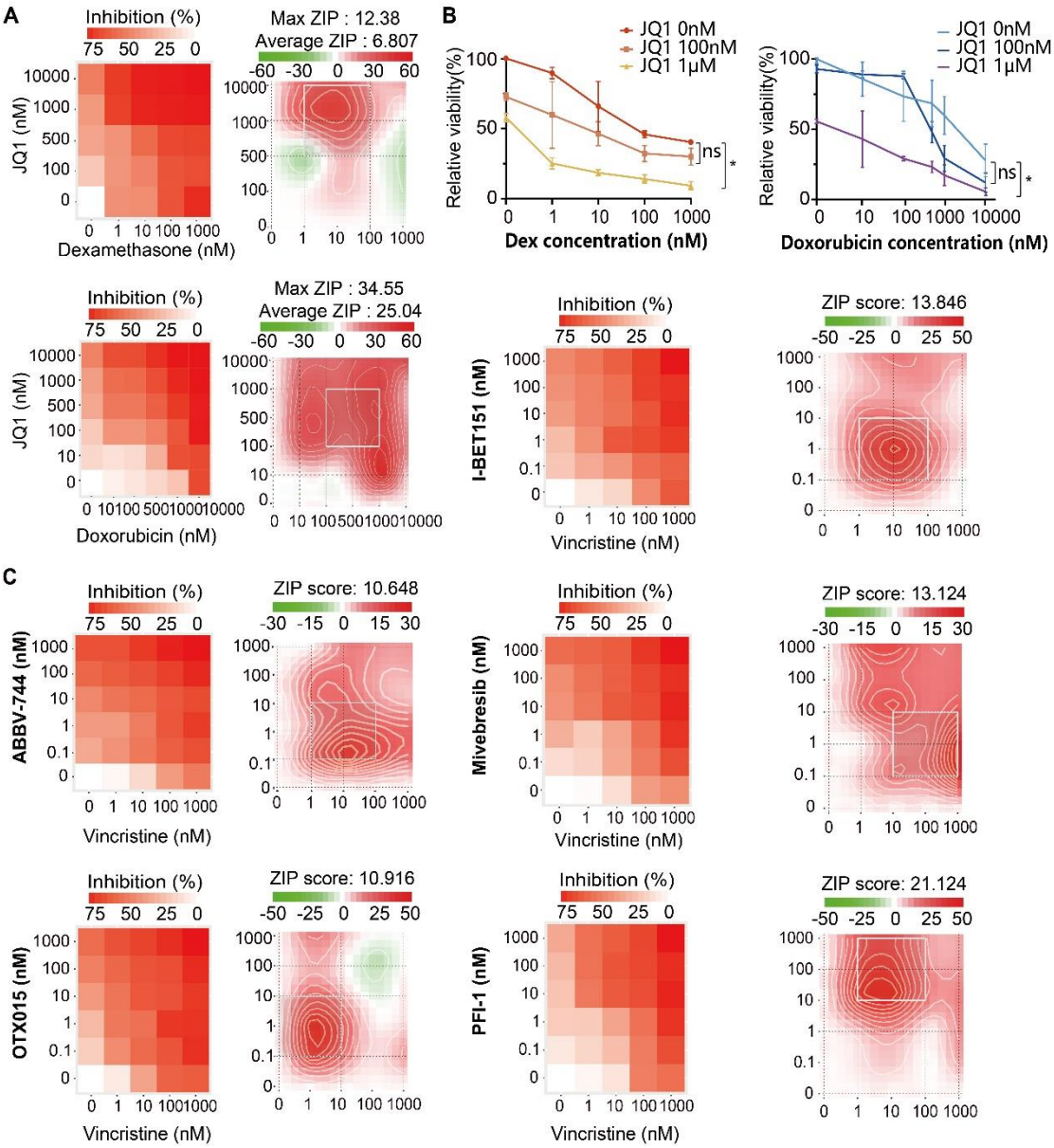
Supplementary Figure 8



Supplementary Fig. S8 BRD4 inhibitors demonstrated a mild impact on *IL3* expression in Jurkat cells

IL3 expression of RJ-9 PDX cells and Jurkat cell line treated with 1 μM JQ1 for 12h. RJ-9, *ETV6::ACSL6* ALL. Before JQ1, Jurkat was pre-stimulated with both PHA (2 μg/ml) and PMA (20 ng/ml) for 4 hours. 12h Ctrl and 12h JQ1 in blue, Jurkat cells treated by vehicle (0.02% DMSO) or JQ1 (1 μM) for 12 hours following the stimulation of *IL3* expression. A total of n = 3 independent experiments are plotted as mean ± SEM.

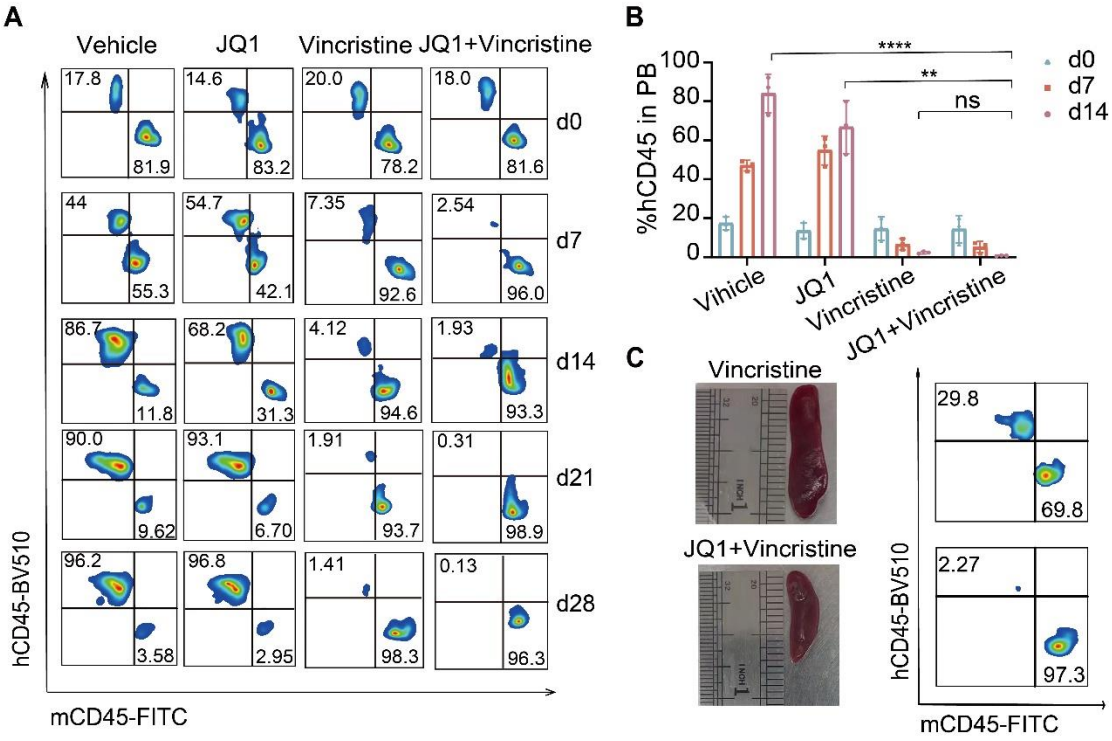
Supplementary Figure 9



Supplementary Fig. S9 ALL cells exhibited vulnerability to BRD4 inhibitors.

(A) Cytotoxicity assay determined the synergy of JQ1 and dexamethasone/doxorubicin in inducing cell death of RJ-9 cells. ZIP, ZIP synergy score. (B) Cytotoxic assay of JQ1 and dexamethasone/doxorubicin at indicating concentrations. *$p < 0.05$. (C) ZIP score of vincristine with indicating BRD4 inhibitors through *in vitro* cytotoxicity assay.

Supplementary Figure 10



Supplementary Fig. S10 BRD4 inhibitors demonstrated anti-tumor effects *in vivo* on ALL cells.

(A) Flowcytometry of leukemia cells in the peripheral blood on indicated days of treatment. (B) Proportion of leukemia cells in peripheral blood on indicated days of treatment. n.s., not significant, $^{**}p < 0.01$, $^{****}p < 0.0001$. (C) Tumor burden in the spleens of VCR alone (top) and VCR+JQ1 (bottom) at day 28 of treatment. Flow cytometry data of the corresponding leukemia cells in the spleen are on the right.

Supplementary References

1. Jing D, Bhadri VA, Beck D, Thoms JA, Yakob NA, Wong JW, Knezevic K, Pimanda JE, Lock RB. Opposing regulation of BIM and BCL2 controls glucocorticoid-induced apoptosis of pediatric acute lymphoblastic leukemia cells. Blood. 2015; 125(2):273-83.

2. Yu CH, Wu G, Chang CC, Jou ST, Lu MY, Lin KH, Chen SH, Wu KH, Huang FL, Cheng CN, Chang HH, Hedges D, Wang JL, Yen HJ, Li MJ, Chou SW, Hung CT, Lin ZS, Lin CY, Chen HY, Ni YL, Hsu YC, Lin DT, Lin SW, Yang JJ, Pui CH, Yu SL, Yang YL. Sequential Approach to Improve the Molecular Classification of Childhood Acute Lymphoblastic Leukemia. J Mol Diagn. 2022; 24(11):1195-1206.

3. Gao M, Wang J, Rousseaux S, Tan M, Pan L, Peng L, Wang S, Xu W, Ren J, Liu Y, Spinck M, Barral S, Wang T, Chuffart F, Bourova-Flin E, Puthier D, Curtet S, Bargier L, Cheng Z, Neumann H, Li J, Zhao Y, Mi JQ, Khochbin S. Metabolically controlled histone H4K5 acylation/acetylation ratio drives BRD4 genomic distribution. Cell Rep. 2021; 36(4):109460.

4. Belaghzal H, Dekker J, Gibcus JH. Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. Methods. 2017; 123:56-65.

5. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, Ahmad K, Henikoff S. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019; 10(1):1930.

6. Chen T, Chen X, Zhang S, Zhu J, Tang B, Wang A, Dong L, Zhang Z, Yu C, Sun Y, Chi L, Chen H, Zhai S, Sun Y, Lan L, Zhang X, Xiao J, Bao Y, Wang Y, Zhang Z, Zhao W. The Genome Sequence Archive Family: Toward Explosive Data Growth and Diverse Data Types. Genomics Proteomics Bioinformatics. 2021.

7. Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2022. Nucleic Acids Res. 2022; 50(D1):D27-d38.