Supplementary Information

Ph+B-ALL is defined by BCR::ABL1-induced enhancer reprogramming and hypersensitivity to enhancer-targeting drugs

Authors:

Han Leng Ng[1,*], Trudy Lee Glaser[1,*], Mark E. Robinson[2], Kadriye Nehir Cosgun[2], Valeriya Malysheva[3,4], Jintao Zhu[1], Ozgen Deniz[5], Nicholas T. Crump[1,6], Kaiyue Helian[1], Andrew J. Innes[1], Li Sun[1], George John[1], Haibin Zhou[7], Atsunori Kaneshige[7], Longchuan Bai[7], Shaomeng Wang[7], Mikhail Spivakov[3], Markus Mueschen[2] and Niklas Feldhahn[1,#]

Affiliations:

[1]Centre for Haematology, Department of Immunology and Inflammation, Faculty of Medicine, Imperial College London, London W12 0NN, UK

[2]Center of Molecular and Cellular Oncology, Yale University, New Haven, CT 06511, USA

[3]MRC London Institute of Medical Sciences, London W12 0NN, UK; Institute of Clinical Sciences, Faculty of Medicine, Imperial College, London W12 0NN, UK

[4]Present address: VIB Center for Molecular Neurology, University of Antwerp, Belgium; VIB Center for AI and Computational Biology, Leuven, Belgium

[5]Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK

[6]The Hugh and Josseline Langmuir Centre for Myeloma Research, Imperial College London, W12 0NN London, UK

[7]Department of Internal Medicine, Medical School, University of Michigan, Ann Arbor, Michigan 48109, USA

*Equal contribution

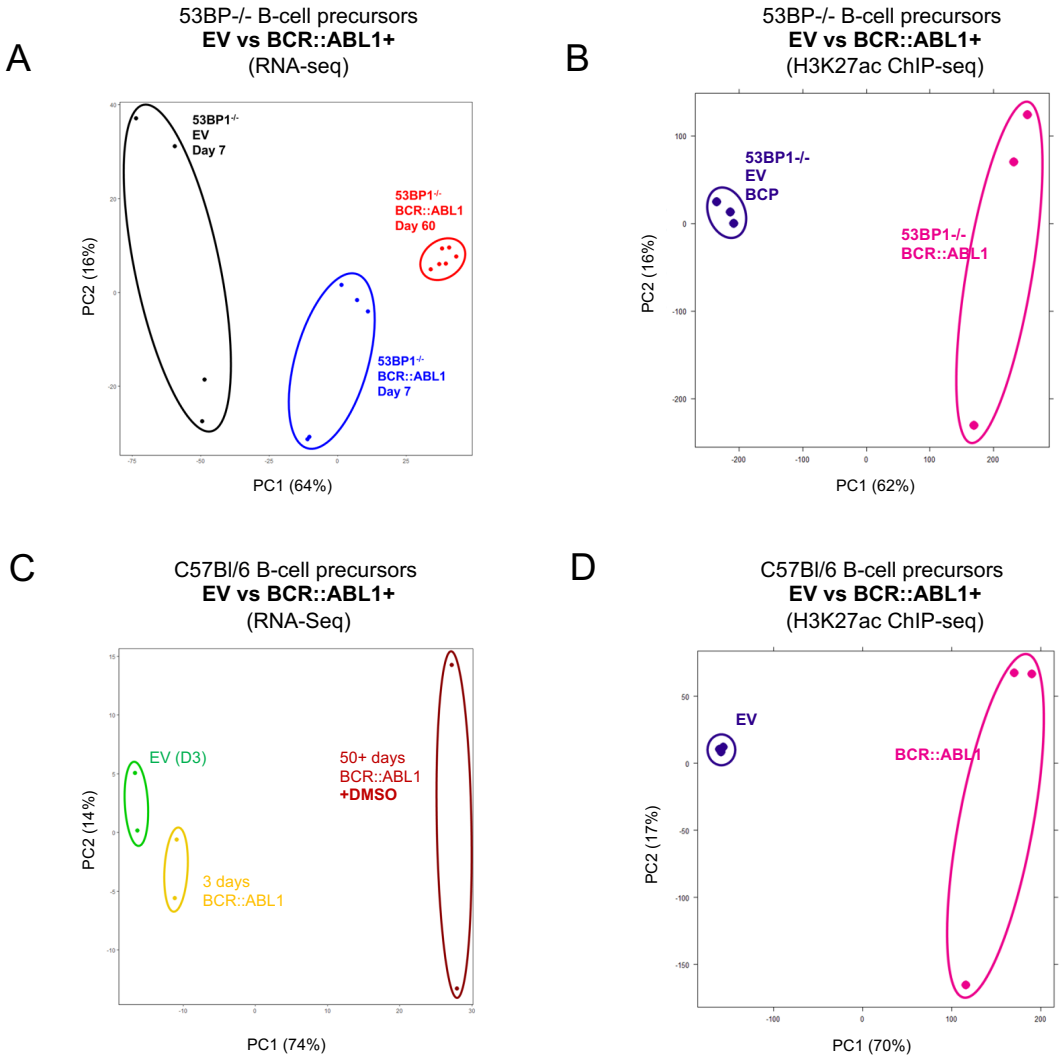
[#]Correspondence: n.feldhahn@imperial.ac.uk

This section includes:

Please note: A separate document contains Supplementary Tables S1-17

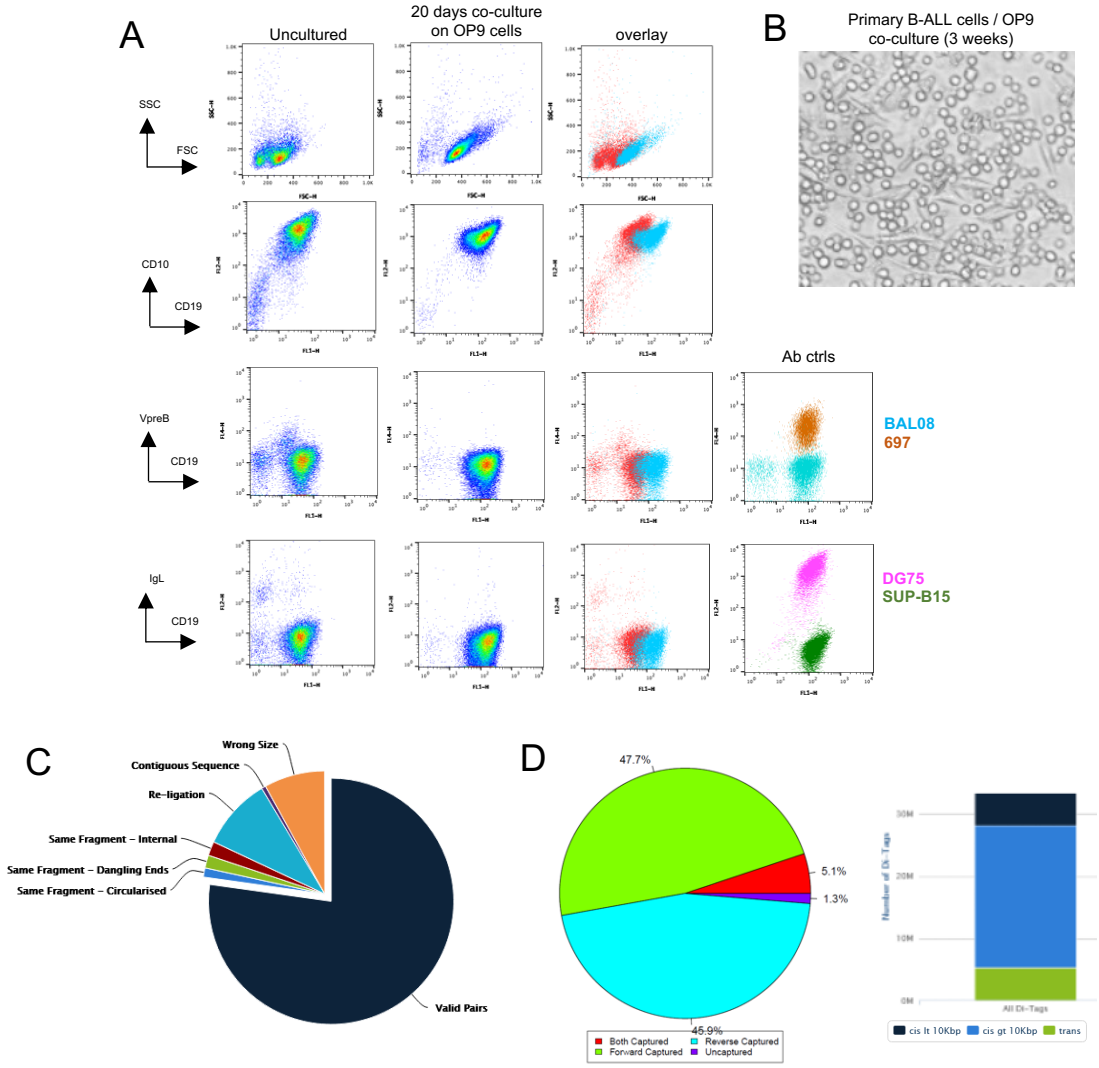
Supplementary Figures

Suppl. Figure 1 (related to Figure 1)



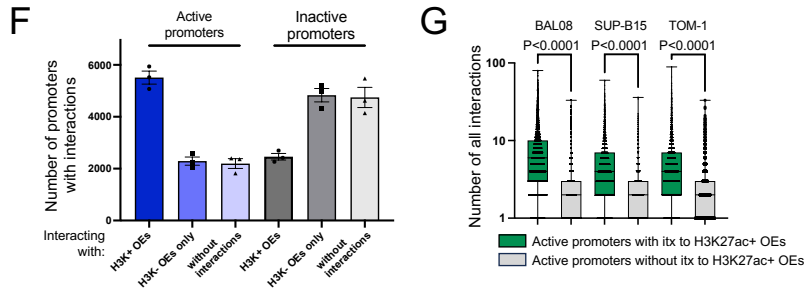
Suppl. Figure 1: Transcriptional and enhancer reprogramming during BCR::ABL1-induced malignant transformation of murine B-cell precursors. XY plots of Principal Component Analyses (PCA) for RNA-Seq (A/C) and H3K27ac ChIP-Seq (B/D) experiments is shown. Analyses were done on B-cell precursors (BCPs) from *53BP1*$^{-/-}$ mice (A/B) or C57BL/6 wild type mice (C/D) that were either transduced with empty MIGR1 vectors (EV) or *BCR::ABL1*-encoding MIGRI vectors. Cells represent different experimental time points as outlined in Figure 1, with EV and early transformed cells collected either at 7 days post *BCR::ABL1* transduction (for *53BP1*$^{-/-}$ BCPs) or 3 days post transduction (for C57BL/6 BCPs), and fully transformed cells collected at ~60 days post transduction.

Suppl. Figure 2 (related to Figure 2)



E

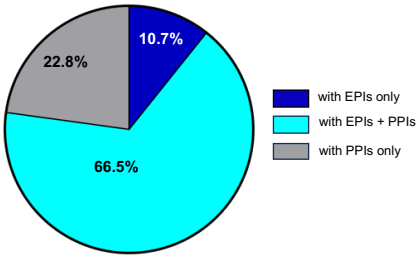
Total Reads	Pair reads	Valid pairs	Unique di-tag	% captured reads		
52,561,988	43,202,417	33,355,746	26,253,108	98.7%	+DMSO	SUP-B15
65,138,622	53,703,619	42,746,086	29,570,479	98.48%	+PON	
68,927,830	57,002,345	43,710,624	29,660,677	98.51%	+DMSO	TOM-1
76,008,211	61,144,156	40,564,123	25,226,269	97.67%	+PON	
65,271,817	54,989,654	40,788,584	27,393,198	98.36%	+DMSO	BAL08
68,593,925	57,777,489	44,245,792	30,292,832	98.83%	+PON	



Suppl. Figure 2 (continued [1])

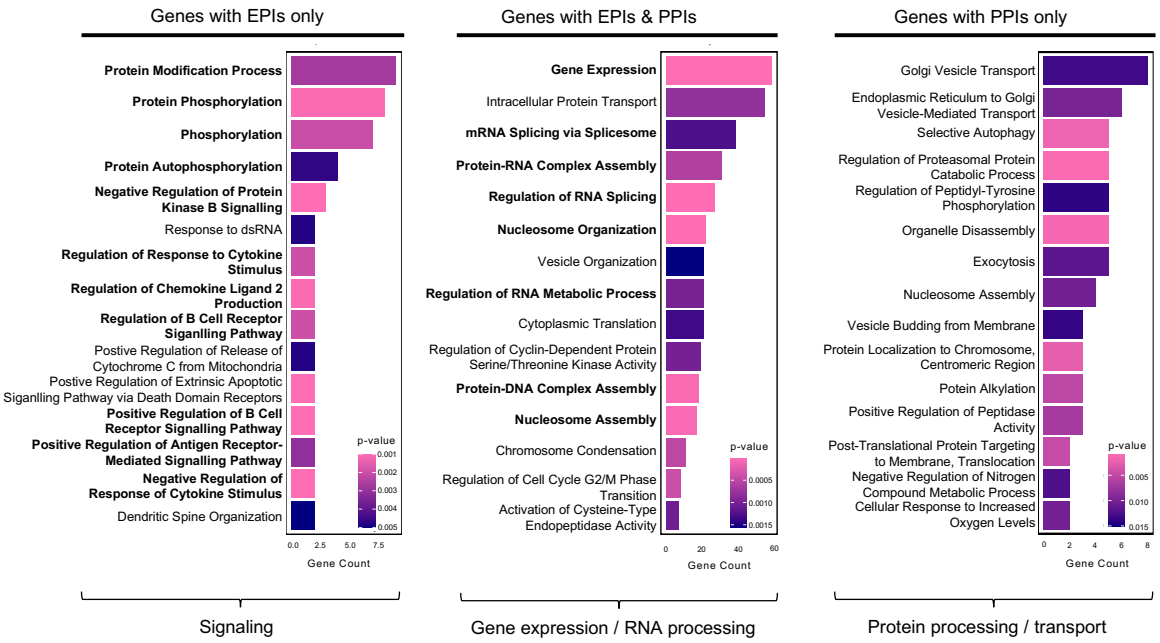
H

Active genes with H3K27ac+ OEs:

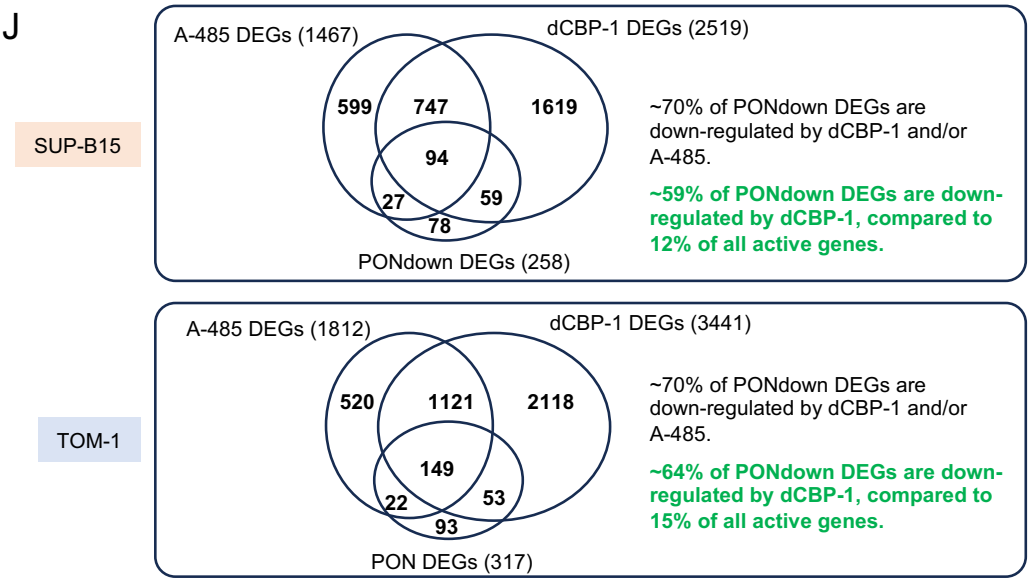


I

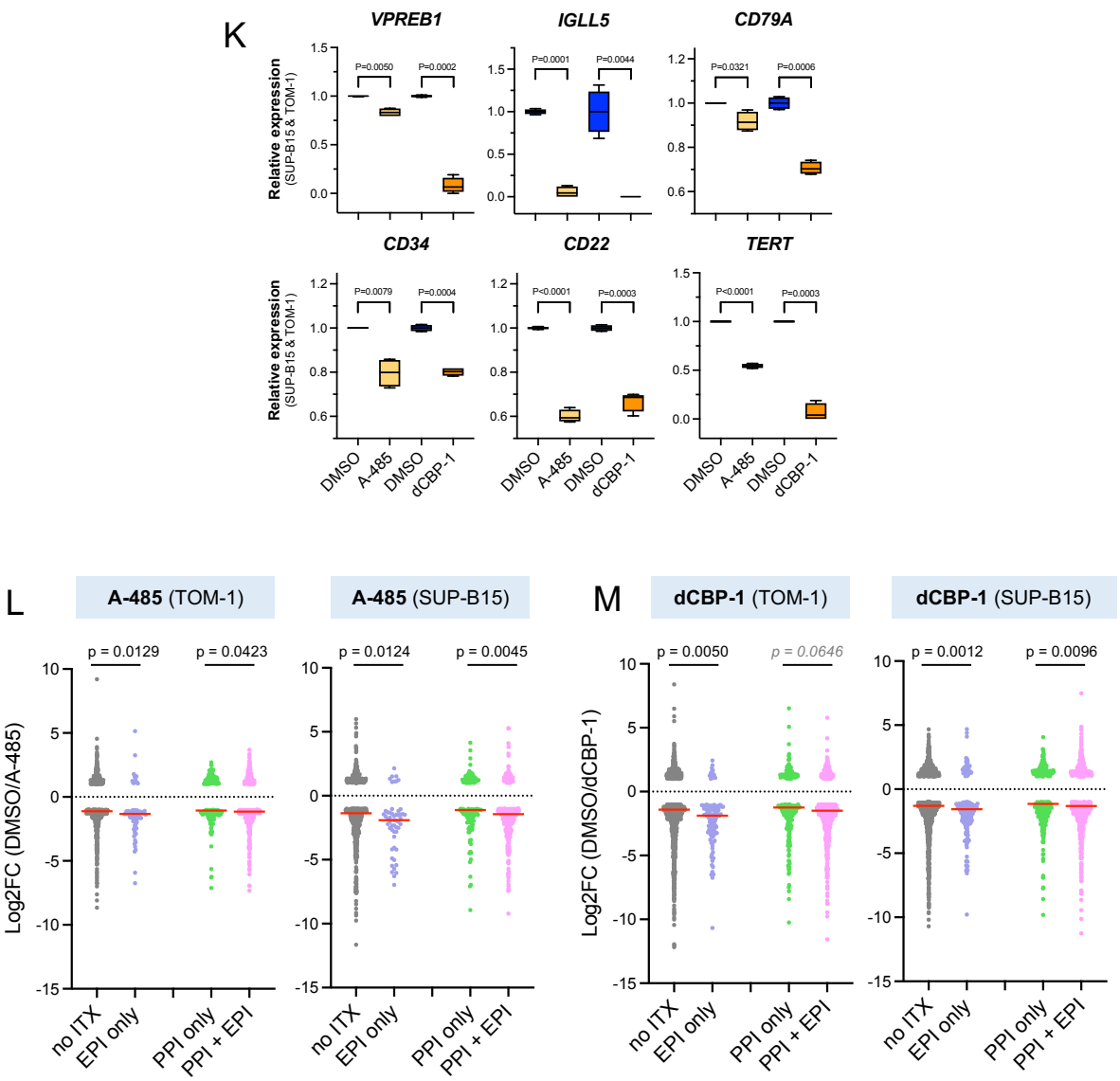
Gene ontology (GO) analysis of genes with PCHiC-defined long-range interactions:



J



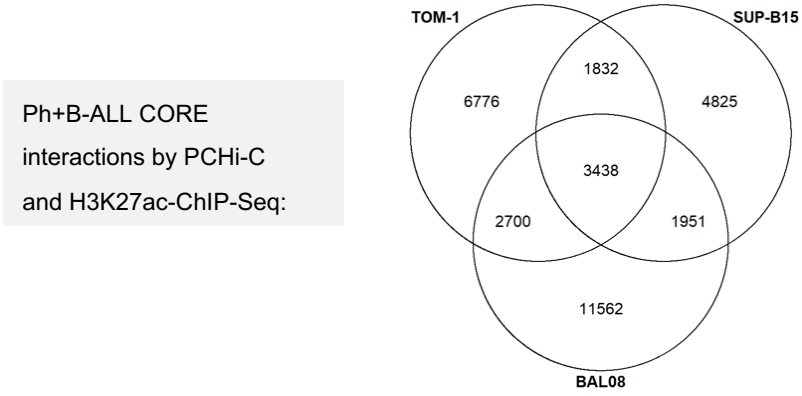
Suppl. Figure 2 (continued [2])



Suppl. Figure 2: Analysis of enhancer-promoter interactions by PCHi-C and associated A-485 treatments. (A) Flow cytometry plots showing the preservation of primary Ph+B-ALL cells (BAL08) during 20 days of *ex vivo* co-culture on sub-lethally irradiated OP9 cells. Cells were stained with antibodies as indicated. First column shows uncultured BAL08 cells, second column shows BAL08 after 20 days of co-culture, third column shows an overlay of first and second columns (red = uncultured, turquoise = 20 days cultured). Cultured BAL08 cells are slightly larger in size (by FSC) and express slightly higher levels of CD19 compared to uncultured cells. The fourth column shows antibody positive control staining for the VpreB and IgL antibodies. Pre-B cell receptor-positive 697 Ph-negative B-ALL cells (brown) were used as positive control for VpreB, DG75 Burkitt's lymphoma cells (pink) were used as positive control for IgL. Positive controls were shown as overlays with cells that are negative for the respective antibodies in the colors indicated. (B) A representative image of primary Ph+B-ALL cells (BAL08) *ex vivo* co-cultured for 20 days on sub-lethally irradiated OP9 cells is shown. (C-E) Different parts of the quality control of PCHi-C samples are shown, with (C) indicating the quality of the Hi-C procedure, (D) indicating the quality of the promoter-capture step, and (E) indicating the total numbers of reads sequenced and processed. Note that the proportion of captured di-tags was ~98.7% (E), of which an average of 84.3% were *in cis* interactions. (F) Bar chart showing the number of active versus inactive promoters with PCHi-C interactions that do or do not display interactions with at least one H3K27ac+ OE (i.e., H3K+ OE vs H3K- OE) and promoters without PCHi-C-defined interactions. Individual dots represent numbers obtained for the three samples analyzed by PCHi-C (i.e., BAL08, SUP-B15, TOM-1). (G) Box plot showing the total numbers of interactions for active promoters that do or do not display interactions (itx) with at least one H3K27ac+ OE (i.e., H3K+ OE vs H3K- OE). Number of interactions

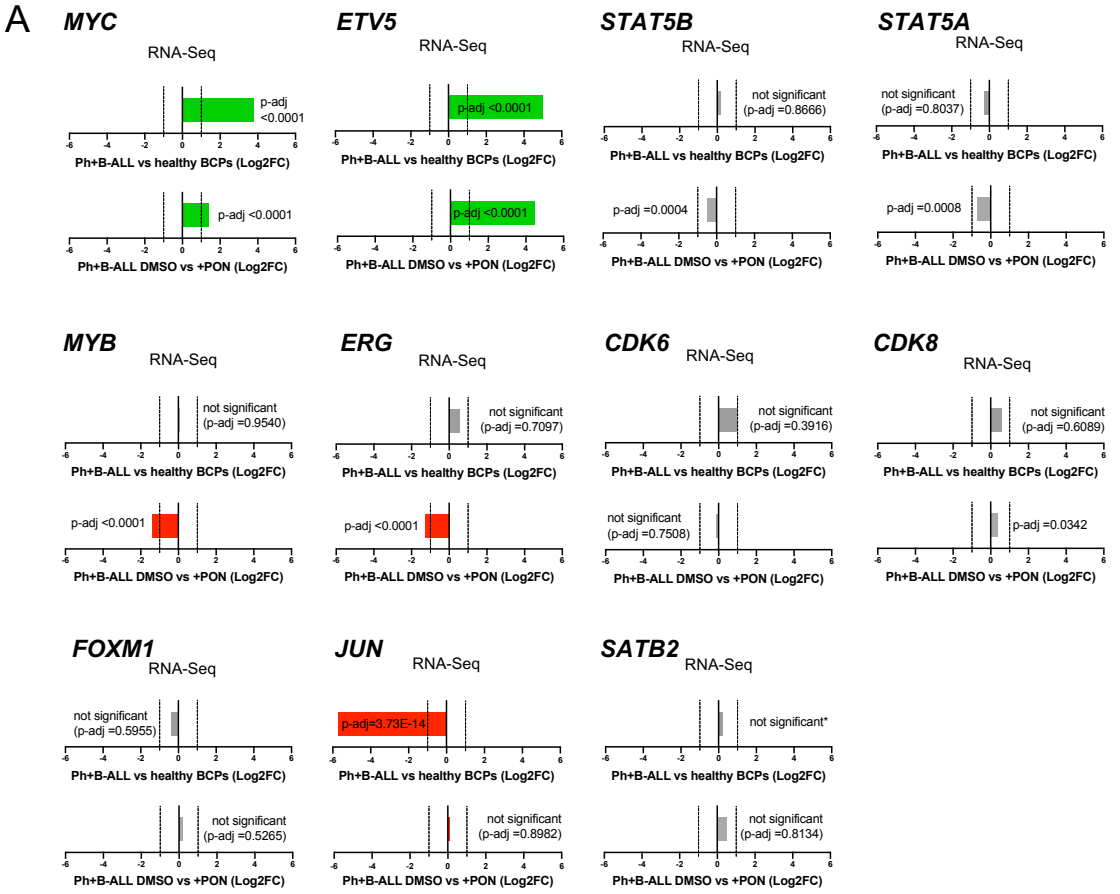
include promoter interactions without H3K27ac+ OEs. Statistical analysis was performed using GraphPad PRISM unpaired Student's t-test. (H) A pie chart is shown indicating the average % of active genes with H3K27ac+ OEs in Ph+B-ALL cells with EPIs only, EPIs+PPIs, or PPIs only. (I) Gene ontology (GO) analysis of active genes with PCHi-C defined long-range chromatin interactions to H3K27ac+ other ends (OE). Analyses were performed on genes that were common in three Ph+B-ALL samples (BAL08, SUP-B15 and TOM-1) and for the respective subgroups of interactions as indicated. (J) A diagram is shown indicating the overlap of dCBP-1 DEGs, A-485 DEGs and Ponatinib (PON) DEGs for the experiments performed in TOM-1 and SUP-B15 cells. (K) Bar diagrams showing relative expression values from A-485 or dCBP-1 treated compared to DMSO treated TOM-1 and SUP-B15 cells defined by RNA-Seq (DMSO and 0.5 μM A-485 for 48 h, or DMSO and 0.25 μM dCBP-1 for 24 h). To allow joint comparison, values were normalized to average DMSO. Indicated genes are lineage-, cell type- or differentiation stage-specific genes. (L/M) Log2 fold change (Log2FC) values of A-485/DMSO DEGs (0.5 μM for 48 h) (L) or dCBP-1/DMSO DEGs (0.25 μM for 24 h) (M) from TOM-1 and SUP-B15 cells (n=2) are plotted for genes with EPIs in comparison to genes with PPIs, genes with EPIs+PPIs, and genes with neither EPIs nor PPIs (no ITX). Statistical analysis was performed by paired Student's t-test and GraphPad PRISM.

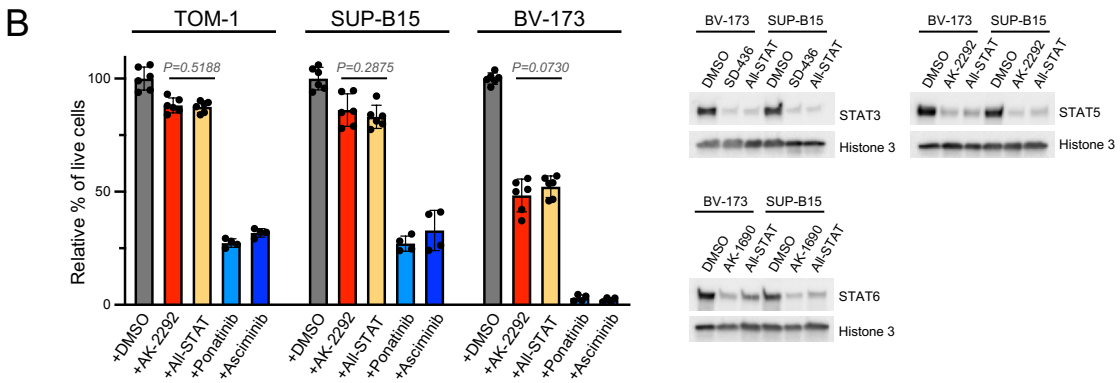
Suppl. Figure 3 (related to Figure 3)

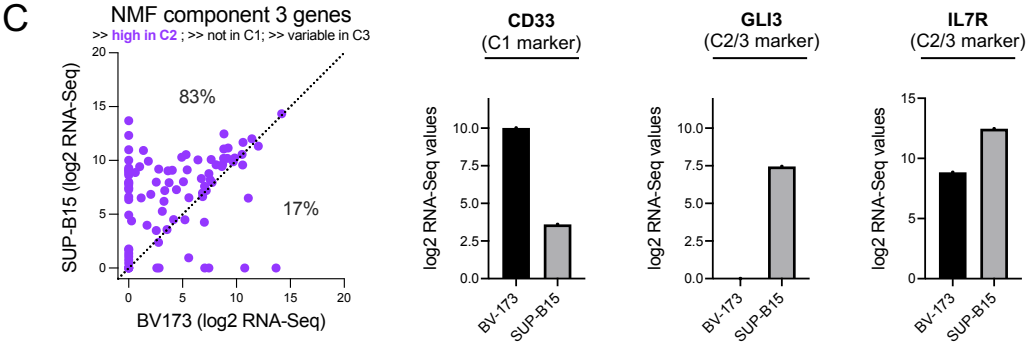


Suppl. Figure 3: Defining 'Ph+B-ALL core interactions' by PCHi-C and H3K27ac ChIP-seq. A Venn diagram is shown indicating the overlap of enhancer-promoter interactions defined by PCHi-C and H3K27ac ChIP-seq and detected in three Ph+B-ALL specimen as indicated. The center indicates interactions shared by all that were then considered as Ph+B-ALL core interactions.

Suppl. Figure 4 (related to Figure 5)



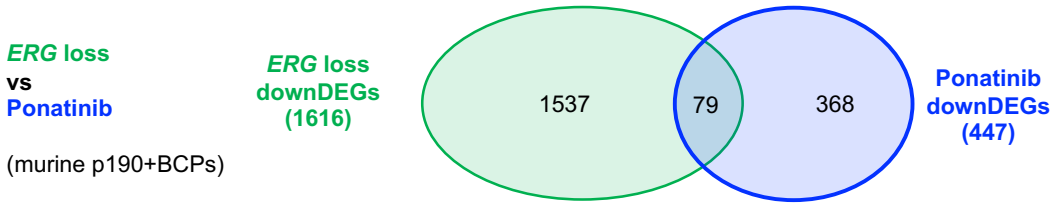




Suppl. Figure 4 (continued):

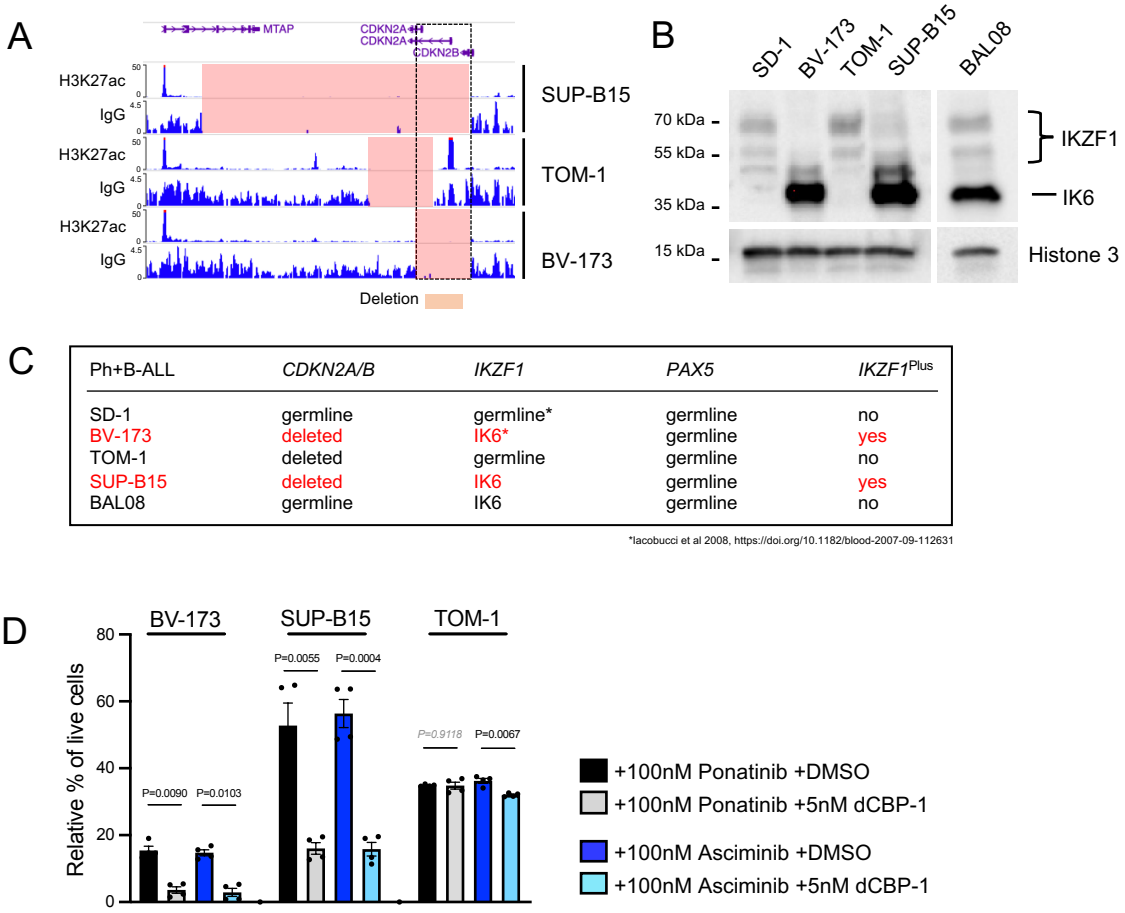
D

ERG loss vs Ponatinib-induced downregulation:



Suppl. Figure 4: Assessment of transcription factor (TF) expression and dependencies in Ph+B-ALL cells. (A) Bar charts showing the data from Figure 5A but separately for individual TFs previously linked to Ph+B-ALL as indicated. Top bars show Log2FC values for comparison of Ph+B-ALL cells to healthy bone marrow BCPs [HBM]. Bottom bars show Log2FC values for comparison of DMSO treated vs Ponatinib (PON) treated Ph+B-ALL cells. For DMSO/PON, three Ph+B-ALL cell lines (BV-173, TOM-1, SUP-B15) and one patient cells from one Ph+B-ALL patient (BAL08) were compared. For Ph+B-ALL/HBM, the same Ph+B-ALL cells were compared to CD19+CD10+ HBMs isolated from n=3 healthy donors. (B) (left) Bar diagram visualizing relative numbers of live cells obtained from 4-day Cell Titer Glo (CTG) experiments. The diagram summarizes the results of n=2 experiments with n=2 technical replicates each on TOM-1, SUP-B15, BV-173 and Ph+B-ALL cells treated with the STAT5 degrader AK-2292 (2μM), a mix of AK-2292 (2μM), the STAT3 degrader SD-436 (0.4μM) and the STAT6 degrader AK-1690 (1.5 μM) [All-STAT], the TKIs Ponatinib (100nM) and Asciminib (100nM), or DMSO as control. Drugs were replenished with a 1:1 media change on D2. Statistical analysis was performed using paired Student's t-test and GraphPad PRISM. (Right) Western blot validation of the degraders used by CTG, in combination or individually, performed on BV-173 and SUP-B15 cells. (C) (Left) XY plot showing log2-transformed normalized RNA-Seq expression values from BV-173 and SUP-B15 cells for NMF component 3 genes defined by Kim et al[1], which are highly expressed by the Ph+B-ALL C2 subgroup, not or little expressed by Ph+B-ALL C1 subgroup and variable expressed by Ph+B-ALL C3 subgroup. (Right) Bar charts of genes described by Kim et al[1] as expressed in the Ph+B-ALL C1-C3 subgroups as indicated. (D) Re-analysis of RNA-Seq data from Behrens et al[2]. Differentially expressed genes (DEGs) that were downregulated upon ERG loss in murine Ph+B-ALL cells are compared to genes that are downregulated in murine Ph+B-ALL cells upon 24h treatment with 100nM Ponatinib treatment (in-house data) and represented in a Venn diagram.

Suppl. Figure 5 (related to Figure 6)



C

Ph+B-ALL	*CDKN2A/B*	*IKZF1*	*PAX5*	*IKZF1*^{Plus}
SD-1	germline	germline*	germline	no
BV-173	deleted	IK6*	germline	yes
TOM-1	deleted	germline	germline	no
SUP-B15	deleted	IK6	germline	yes
BAL08	germline	IK6	germline	no

*Iacobucci et al 2008, https://doi.org/10.1182/blood-2007-09-112631

Suppl. Figure 5: *IKZF1*^{PLUS} classification on Ph+B-ALL lines and patients used in this study. (A) ChIP-Seq custom tracks for anti-H3K27ac and IgG control antibodies are shown for SUP-B15, TOM-1 and BV-173 cells depicting the CDKN2A/B locus. Pink areas show nearly complete absence of any background read alignment indicative of homozygous genomic deletion. (B) Western blot of IKZF1 and Histone 3 as loading control indicating cell lines/patients expressing full-length IKZF1 and/or the dominant-negative isoform IK6. (C) Summary of *CDKN2A/B* and *IKZF1* status defining cells as *IKZF1*^{PLUS} or not. (D) Bar diagram showing clinically relevant TKI concentrations from Figure 6G and comparing +DMSO vs +5nM dCBP-1 treatments. 100nM was defined for both TKIs as the molar concentration equivalent to the plasma concentrations of patients receiving standard dose treatment of Ponatinib (45mg/day) and Asciminib (80mg/day).

Extended Supplementary Methods

Cell lines, patient samples and mouse models

All cell lines were obtained from the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ, Germany). Human leukaemia and myeloma cell lines were cultured in RPMI 1640 media supplemented with 1% penicillin/streptomycin, 1X Glutamax and 20% FBS (all from Sigma) at 37°C in a humidified incubator with 5% CO_2, while HEK293t cells were cultured in DMEM media + 1% penicillin/streptomycin, 1X Glutamax and 20% FBS. OP9 cells were a kind gift of MM and cultured as the leukaemia cell lines above but using aMEM media (Invitrogen). Adherent HEK293T cells were split using Trypsin (Sigma Aldrich) when reaching a confluency of 80%. A549 and PC3 cells were a kind gift of Dr Richard Burt (Crick Institute/Imperial) and cultured as described for HEK293Y cells.

Human leukaemia samples used in this research project were deposited into, stored in and subsequently retrieved from the Imperial College Healthcare Tissue Bank (ICHTB). ICHTB is approved by NRES to release human material for research (12/WA/0196). A sub collection dedicated to this project was initiated (MEC_AR_16_030) and access to this sub collection for this project was granted. Healthy BM samples were provided as cryopreserved samples by The John Goldman Centre for Cellular Therapy of Hammersmith Hospital and collected under respective licenses. All patients gave specific informed consent for the use of surplus tissue. All samples were pseudo-anonymised on receipt and given a study ID.

For processing of primary B-ALL samples, whole blood, bone marrow or leukapheresis from these patients was separated by density gradient using centrifugation according to the manufacturer's protocol to retrieve mononuclear cells. Cells were then processed for ChIP-Seq, cryopreservation, RNA extraction etc. Purity of samples was checked by flow cytometry. Cytogenetic information to define if a sample was Ph+ or not was provided by the diagnostic lab of the Department of Haematology, Imperial College Healthcare NHS Trust. Ph+B-ALL cells from patient BAL08 were also expanded *ex vivo* on sub-lethally irradiated OP9 cells ($0.5x10^6$ OP9 cells/10 cm dish, mitotically inactivated by 13.6 Gy, 25×10^6 B-ALL cells/plate). After ~20 days of co-culture, cells were checked for viability and B-ALL phenotype by flow cytometry and then split to new OP9-coated plates for respective experiments.

For healthy controls, cryopreserved cells were thawed, sorted by flow cytometry using anti-human CD10 and CD19 antibodies (BioLegend, cl.HI10A and cl.HIB19, respectively) and subjected to RNA-Seq analysis. For one patient, where greater surplus material was available, mononuclear cells were purified as described above and B-cells were enriched by depletion of non-B-cells cells by magnetic separation using the human B Cell Isolation Kit II and LS columns followed by enrichment of CD10+ cells using anti-human

CD10 microbeads (all from Miltenyi Biotec). Enrichment was validated by flow cytometry as described above.

PDX samples used for HiChIP were generated as described in[3]. Cells used to create them were collected from patients who gave informed consent and processed in compliance with the internal review boards of the Beckman Research Institute of City of Hope and all relevant ethical regulations. Experiments using murine cells were performed in agreement with ASPA guidelines and regulations and protocols approved by Home Office UK. Experiments using *53BP1*$^{-/-}$ mice[4] were performed as described in Boulianne et al[5]. All other mouse experiments were performed using cells from C57Bl/6 mice, with murine B-cell precursors isolated and cultured as described[5]. Respective time points of analysis are provided in the figure legends.

In vitro transformation

For all mouse experiments, cells were pre-cultured in Iscove's Modified Dulbecco's Medium (IMDM) containing 20% FCS, 1% Penicillin/Streptomycin, 1X Glutamax, 50 µM 2-mercaptoethanol (all from Sigma) and 10 ng/mL of recombinant mouse IL-7 (Peprotech) followed by transduction with BCR::ABL1 encoding retrovirus or empty vector (EV) as control. Virus production was performed as described in Boulianne et al[5] but using BCR::ABL1^{p190} or BCR::ABL1^{p210} encoding MIGR1 vectors, and respective EV. Viral transductions of *53BP1*$^{-/-}$ B-cell precursors (BCPs) were performed exactly as previously described[5]. For C57Bl/6 mice, cells were pre-cultured for 5 days with mIL7 to enrich for BCPs, followed by two transductions as in[5]. For RNA-Seq analysis, on day 1 after the second transduction, cells were sorted for GFP using flow cytometry (MA900 sorter, Sony), put back into culture, and harvested on day 3 or maintained at 0.5-2.5x10^6 cells/mL and harvested at ~50 days post transduction. For H3K27ac ChIP-Seq no sort of GFP+ cells was performed. IL7-induced enrichment of BCPs was monitored by flow cytometry and viability and density continuously monitored using trypan blue staining.

Flow cytometry

For validation of human healthy or leukemic B-cell precursors the following antibodies were used as indicated in the figures: Anti-CD10 (BioLegend, cl.HI10A), anti-CD19 (BioLegend, cl.HIB19), anti-CD179a (VpreB, BioLegend), anti-IgL (BioLegend). For staining of live/dead/apoptotic cells 7-aminoactinomycin D (7-AAD, Thermo Fisher Scientific) and Annexin-V-APC (BioLegend) were used. For cell cycle analysis, HOECHST 33342 dye (sigma) was used according to manufacturer’s instructions. Cells were then analysed by flow cytometry using a BD LSR-Fortessa cytometer.

Monitoring cell viability and proliferation using Cell Titer Glo (CTG)

Cells were seeded in 96-well plates at 50,000 cells/well and treated with the drugs indicated. On day 2, a partial media change adding new drugs and diluting cell counts 1:1 was performed. On day 4, cells were analysed using the Cell Titer Glo assay (Promega) and a TECAN plate reader.

Monitoring cell viability by 7-AAD and Annexin-V staining

Leukaemia and myeloma cells were seeded in 12-well plates at $0.5x10^6$ cells/mL and treated as indicated in the text and figure legends for 5 days. At 2 or 3 days of drug treatment, cells were split 1:3 to dilute cells, replenish media and add fresh drugs. On day 5 of treatment, cells were counted using Trypan blue and a haemocytometer, and remaining cells were stained for 7-AAD (Thermo Fisher Scientific) and Annexin-V-APC (BioLegend) according to manufacturer's instructions. Stained cells were analysed using a LSR-Fortessa flow cytometer (BD). For adherent cells, $0.7x10^6$ HEK293T or A549 cells were seeded per well of a 6-well plate while PC3 cells were seeded at $0.6x10^6$ cells/well into a 12-well plate. Cells were allowed to attach to the plate for 6 h before start of drug treatment. Drug treatment and analysis was done as described above for suspension cells, but cells were detached using trypsin and split 1:5 for A549 and HEK293t cells. PC3 cells were just replated.

Western blot

Western blot was performed as described in Boulianne et al[5] using 4-15% Mini-PROTEAN TGX Precast Gels and the Mini-PROTEAN blotting system (Bio-Rad), but using the following antibodies: Anti-phospho-CRKL (#3181), anti-H3K27ac (#4353), anti-Histone 3 (#9715), anti-Lamin B1 (#12586), anti-beta-Tubulin (#2128), anti-HA-tag (#3724), and anti-STAT5 (i.e., pan-STAT5A/B #94205), anti-STAT3 (#12640), anti-STAT6 (#5397), anti-P300 (#86377), anti-ETV5 (#16274) as primary antibodies (all from Cell Signaling), and goat anti-rabbit HRP (R&D systems) as secondary antibody. For visualisation, Pierce ECL Western Blotting Substrate (Thermo Fisher) and a Chemidoc Touch Imager (Bio-Rad) was used.

Drug treatments

Ponatinib, Asciminib and dCBP-1 were obtained from Cambridge Biosciences / MedchemExpress. AK-2292[6] was from Cambridge Biosciences / MedchemExpress and Sigma Aldrich. A-485 was from Sigma Aldrich (SML2192). SD-436/AK-1690[7,8] were kind gifts from Shaomeng Wang (University of Michigan). Cells were treated for the times indicated in the text and Figure legends. For all drug experiments, 1000X stocks (drug in DMSO) were used and added in a 1:1000 manner. DMSO (Sigma) treatment was performed as control.

RNA interference (RNAi)

ETV5 was depleted using lentivirus containing TRCN0000013938-encoding MISSION pLKO.1-puro-CMV-TurboGFP plasmids (MERCK) and custom-made 9-11 seed controls where nucleotides 9-11 of the shRNA seed sequence of TRCN0000013938 were mutated[9]. The 9-11 control plasmid was generated by first restoring the Age-I site of TRCN0000013938-pLKO.1-puro-CMV-TurboGFP plasmids using the Q5® Site-Directed Mutagenesis Kit (NEB) with GAAACACCGGtCGTGACACTT and GTCCTTTCCACAAGATATATAAAGC as primers, and then by replacing the original shRNA with the 9-11 mutated shRNA sequences CCGGGAGCGATACGAGAACAAATTTCTCGAGAAATTTGTTCTCGTATCGCTCTTTTTG and AATTCAAAAAGAGCGATACGAGAACAAATTTCTCGAGAAATTTGTTCTCGTATCGCTC. Lentivirus was generated as described in Pfeifer et al[10]. Transduced cells were sorted for GFP at 3 or 4 days after transduction using SONY MA900 and seeded as 5,000 cells/well into 96-well plates using the sorter. Three days later, cells were collected, counted and seeded at $1x10^6$ cells/mL into 12-wells and treated with DMSO or 2.5 μM AK-2292 for another 24 h before harvest for RNA-Seq or ChIP-Seq. For Cell Titer Glo (CTG) analysis, three days after sorting, cells were kept in 96-well plates but diluted 1:1 to add fresh media and respective DMSO or AK-2292. Cells were then treated as described for CTG analysis above (1:1 dilution to refresh media and add drugs on day 2 of drug treatment, analysis by CTG on day 4).

RNA-Seq, ChIP-Seq and ATAC-Seq library preparation

Samples for RNA-Seq were stored as snap-frozen cell pellets and RNA was extracted using the RNeasy mini kit (Qiagen). Total RNA was submitted to and processed for RNA-Seq library preparation by Novogene (UK). ChIP-Seq for H3K27ac was done as described in Boulianne et al[5] using anti-Histone H3 (acetyl K27) antibody – ChIP Grade (ab4729) (Abcam) and polyclonal rabbit IgG (isotype control) (sc2027) (Santa Cruz Biotechnologies), with the exception that initial cross-linking was performed at RT. ChIP-Seq was performed on 15 x 10^6 or 1 x 10^6 cells depending on cell availability but using the same protocol. ChIP-Seq libraries were generated from ChIP immunoprecipitated DNA using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos for Illumina (NEB) and sequenced either at Novogene or at the in-house sequencing facility of the MRC London Institute of Medical Sciences at Imperial College London. For ATAC-Seq, samples were processed according to the Omni-ATAC-seq protocol (https://doi.org/10.1038/protex.2017.096) and similarly sequenced Novogene. A total of 20 × 10^6 paired-end (PE) reads were sequenced per sample for all procedures.

PCHi-C and HiChIP library generation

PCHi-C was generated as previously described[11]. PCHi-C libraries were sequenced by paired-end (75 bp) sequencing at the BRC genomics facility at Imperial College London. HiChIP was performed using the Arima Genomics HiChIP kit with the recommended protocol. H3K27ac (Active Motif #91193) was used for immunoprecipitation of libraries. HiChIP libraries were prepared using the Arima Genomics HiChIP library preparation using Swift Biosciences® Accel-NGS® 2S Plus DNA library kit. HiChIP libraries were sequenced with PE (150 bp) sequencing at the in-house sequencing facility of the Yale School of Medicine.

Bioinformatic processing of NGS libraries

For RNA-Seq, ChIP-Seq and ATAC-Seq libraries, raw fastq files were trimmed using Trim-galore (v0.6.7) before further analysis. GRCh38 (hg38), or GRCm39 (mm39) reference genomes were used for alignment of reads for all human and mouse samples, respectively. For all analysis on R (v4.1.0), tidyverse (v1.3.1) and magrittr (v2.0.1) were loaded.

RNA-Seq

Salmon (v1.10.1) was used to quantify transcript count from RNA-Seq. A full decoy index was generated using whole genome and transcript sequences obtained from GENCODE genes database, with the hg38 (release 43) and mm39 (release 32) genomes, respectively. Trimmed fastq files were processed using Salmon (v1.10.1) with the following parameters: --gcbias, --writeUnmappedNames.

GenomicFeatures(v1.46.5) was used to generate transcript-to-gene objects in R, using GTF files obtained from the GENECODE gene database. Salmon quantification file was subsequently analysed using DESeq2 (v1.34.0). Briefly, tximport (v1.22.0) was used to import Salmon quantification files into R. DESeqDataSetFromTximport was used to process transcript count. As the human cell lines were inherently different in genome, this difference was factored in when using DESeq2 in our comparison between DMSO and ponatinib treatment. In DESeqDataSetFromTximport(), the following parameter was used in the "design" option: ~condition+cell_line. For all other analyses with DESeq2, ~condition was used in the "design" option. Subsequently, genes with low transcript counts below 5 were filtered out from the analysis using rowSums(counts()) >= 5. DESeq() and results() were used to determine differentially expressed genes (DEGs). Data generated by results() were filtered for DEGs with a p-adjusted value < 0.05 and a log2 fold change > 1 or < -1.

To generate DEG expression heatmaps, rlog() transformation, followed by assay()[DEG,] (where DEG is a vector of DEGs) generated a matrix. The matrix was used to plot a heatmap with pheatmap (v1.0.12), where the scale="row" option was added.
For principal component analysis (PCA) plots, plotPCA() was used with intgroup = c("Treatment") option.

ChIP-Seq

Trimmed fastq files were aligned to the human genome, hg38, using bowtie2 (v2.4.4), using the additional parameter: --no-unal. SAM files generated by bowtie2 were converted to BAM format using samtools (v1.14) view -S -b. Bam files were processed using samtools fixmate -m, samtools sort -o, samtools markdup -r, samtools collate -o, samtools sort -o again.

To generate a file for visualisation on a genome browser, the final bam files were processed by deeptools (v3.5.1) suite. To generate bigwig files, bamCoverage was used with the following parameters: --normalizeUsing RPKM, --extendReads, --effectiveGenomeSize 2913022398 (for hg38), or 2654621783 (for mm39).

To determine peak enrichment across the genome, macs2 (v2.2.7.1) callpeak was used, with an IgG ChIP used as the input control. The following parameters were used: -q 0.05, -B, --SPMR, --broad (only for H3K27ac ChIP), and -g hs (for hg38) or mm (for mm39).

To determine reads within specific region, the Homer (v4.11) package suite was used. The Homer suit was used to create tag directories using makeTagDirectory, with the added parameter: --keepAll.

ATAC-Seq

Trimmed fastq files were aligned to the human genome, hg38, using bowtie2 (v2.4.4) with the same parameter as for ChIP-Seq. Similarly, alignment SAM files were processed with samtools in the identical method as for ChIP-Seq. To generate bigwig files for viewing on a genome browser, bamCoverage (from the deeptools package) was used with the same parameters as for ChIP-Seq, except for the additional parameter of --extendReads.

ATAC-Seq were called using macs2 as for ChIP but without input control and not using --broad.

PCHi-C

Libraries were processed in the same method as previously described[11]. Raw fastq data was processed by HiCUP (v0.7.3) with the following parameters set in the configuration file: bowtie2, longest di-tag length: 850, shortest di-tag length: 100. The intermediate file generated by HiCUP was further processed in R using CHiCAGO (v1.18.0) and processed

as described in Freire-Pritchett et al[12] to generate interaction files and scores. Interaction files generated by CHiCAGO was visualised on the WashU epigenome browser https://epigenomegateway.wustl.edu/browser/. The makePeakMatrix.R pipeline was used to generate a matrix file of all interactions in all PCHi-C samples.

Computational analyses

Gene promoters 2 kb flanking region

Gene promoters were obtained from the GENCODE gene database for hg38 and mm39. Start coordinates of gene transcriptional start sites were obtained from GTF files and the 2 kb region flanking (both up- and downstream) each transcriptional start site were obtained using bedtools slop -b 1000.

H3K27ac ChIP-Seq heatmap

Non-promoter peaks were obtained using bedtools intersect between H3K27ac peaks called by macs2 with gene promoter coordinates.

Homer package was used to determine differential peaks between DMSO- and Ponatinib-treated samples using the getDifferentialPeaks command and the H3K27ac peaks called by macs2, with the parameters: -P 0.05, -F 2, -size 500, with -rev either added or not, to determine peaks gained or lost with Ponatinib treatment, respectively. H3K27ac peak genomic coordinates that were within 1 kb of each other were merged together using bedtools (v2.30.0) merge, with the added parameter: -d 1000. To determine consensus differential peak region between samples, bedtools multiinter was used, where the region overlapped by all three samples were filtered. To get the normalised peak score within the consensus region, Homer getPeakTags was used with the consensus differential peaks. Scores generated by getPeakTags were used to plot heatmap in R, using dplyr to stack peaks score and converted to a matrix for plotting with pheatmap, with the parameters: scale = “row”, cluster_cols = FALSE, cluster_rows = FALSE.

H3K27ac PCA plots

H3K27ac peaks called by macs2 were split into either promoter or non-promoter using bedtools intersect. DiffBind (v3.4.11) was used to generate PCA plots in R. In sum, a table is required containing macs2 peaks in bed format, together with bam files generated from ChIP-Seq analysis. The table was used in R with dba() and processed by dba.count(). The data was further normalised with dba.normalize() with the added options: normalize = DBA_NORM_LIB, and library = DBA_LIBSIZE_PEAKREADS. Finally, PCA plots were generated using dba.plotPCA() with the option: attributes = DBA_TREATMENT.

H3K27ac DiffBind analysis

Differential H3K27ac peaks were identified for the experimental groups defined in the text/figures using R and r-tidyverse and Bioconductor-diffbind packages, and H3K27ac ChIP-Seq bam files MACS2-defined peak bed files from n=2 sets of experiments as input.

PCHi-C analysis

Interaction data generated by CHiCAGO were used to determine overlap with H3K27ac peaks for both the bait (promoter) and other end (OE). For this, bedtools intersect was used to determine promoters and OE that overlap with H3K27ac peaks. In doing so, we were able to determine active promoters and inactive promoters by the presence, or absence of H3K27ac, respectively. Bedtools groupby was used to group PCHi-C interactions by either the presence of at least of one H3K27ac-overlapping OE, or the complete absence of H3K27ac-overlapping OE. As such, using bedtools groupby, we were able to classify four different types of interactions: 1) active promoter to active OE interaction(s); 2) active promoter to inactive OE interaction(s); 3) inactive promoter to active OE interaction(s); 4) inactive promoter to inactive OE interaction(s). Further, bedtools intersect was used to determine if any of the OE interactions overlap with any gene promoter. Together with bedtools groupby from the bait end, we were able to determine each bait interacting with the OE to be either one of three possible situations: enhancer-promoter interactions (EPIs), promoter-promoter interactions (PPIs), or a combination of EPIs and PPIs.

To determine PCHiC interactions and changes with Ponatinib treatment, RNA-Seq data was utilised. Prior to analysis, active genes were filtered for normalised transcript count of 10 or more using dplyr::filter(). By comparing between healthy human B-cell precursors and Ph+ samples, Ph-driven DEGs were identified from DESeq2. Additionally, TKI-sensitive DEGs were determined by DESeq2 by comparing between DMSO- and Ponatinib-treated Ph+ samples. Using both sets of DEGs, we were able to identify one of four groups of genes: 1) Ph-upregulated genes that are not TKI-sensitive; 2) Ph-upregulated and TKI-sensitive (down-regulated) genes; 3) TKI-sensitive (down-regulated) and not deregulated by Ph; 4) cell line specific genes that are not in one of the three previous groups.

Using the list of all TKI-sensitive genes, we were able to determine TKI-sensitive interactions using dplyr::filter(). Additionally, we were able to separate the interactions into either EPIs, PPIs, or a mixture of EPIs and PPIs. For our analysis, all EPIs were grouped together and CHiCAGO scores determined based on their promoter baits and plotted each promoter in samples treated with DMSO or TKI using GraphPad Prism (v9). Additionally, using the H3K27ac signal at both ends of each interaction, changes in H3K27ac enrichment was determined for interactions in DMSO-treated samples and the corresponding interaction (or loss of interaction) after Ponatinib treatment.

From each sample, we were able to obtain a list of active promoter to active enhancer interactions (active EPIs) with dplyr::filter(). The lists of active EPIs were used to define a set of core Ph+ interactions. Using the number of interactions for each sample, a Venn diagram was generated in R using draw.triple.venn(). Additionally, the matrix generated by makePeakMatrix.R was filtered for only the core Ph+ interaction set and used to plot a hierarchical dendrogram in R. The matrix was transformed, t(), followed by using the dist() function with the option: method = “euclidean" and used with the hclust() function with the option: method = “average”. The same matrix was also used to generate a PCA plot. First, the matrix was transformed t(), followed by prcomp(), and finally plotted using autoplot().

A-485/dCBP-1 and PCHi-C analysis

DESeq2 was utilised to determine DEGs from cells treated with A-485 or dCBP-1. Using the PCHi-C interaction data, A-485 and dCBP-1 DEGs were separated into either having EPI only, PPI only, a mixture of both EPIs and PPIs and other genes with no interaction and results were analysed and visualised using GraphPad PRISM.

HiChIP analysis

The MAPS pipeline generated a BEDPE file of both ends of interactions, where a raw read count is calculated from the HiChIP fastq files, as well as different expected contact frequency calculated for interaction clusters. For each sample, the read count for each interaction was normalised to the library size of per 10 million reads to get an interaction score. Additionally, each interaction had different expected contact frequency for the cluster and as such, unique interactions were filtered for downstream analysis using the unique() function in R. Using the filtered list of unique interactions for all samples, we were able to define three types of interactions: 1) promoter-promoter interactions; 2) promoter-promoter interactions; 3) enhancer-enhancer interactions. These interactions were determined using bedtools intersect with known gene promoter coordinates for both ends of the interactions.

For our analysis, we examined how the samples clustered from each of the three types of interactions: promoter-promoter, promoter-enhancer and enhancer-enhancer. Hierarchical clustering was performed on R using t(), dist() and hclust() for downstream PCA analysis. Similar to PCHi-C PCA analysis, the data was processed in R using t(), prcomp() and autoplot() in the respective order.

TARGET RNA-Seq analysis

Publicly available RNA-seq data from the TARGET study (Therapeutically Applicable Research to Generate Effective Treatments initiative, phs000463 and phs000464, https://www.cancer.gov/ccg/research/genome-sequencing/target) was downloaded for B-ALL

with diagnosis known for either Ph+ (n=6), or KMT2A-r (n=4). The data used for this analysis are available at the Genomic Data Commons (https://portal.gdc.cancer.gov). DESeq2 was used to identify DEGs between Ph+ and KMT2A-r B-ALL.

HiChIP and RNA-Seq combined analysis

From the RNA-Seq analysis of Ph+ and KMT2A-r B-ALL (TARGET) samples, DEGs were used for HiChIP samples. Using dplyr::filter(), DEGs were used to filter for HiChIP interactions that overlap with the promoter at either end of the HiChIP interactions. From the filtered interactions, an average number of interactions was counted for each gene across all samples analysed. The average number of interactions was further processed to generate a log2 fold change value to define if a gene had more interactions in the Ph+ B-ALL or KMT2A-r B-ALL samples. Similarly, RNA-Seq log2 fold changes for these DEGs were filtered and combined with the average HiChIP log2 fold change interaction in a table in R. To generate a trend line for the graph, in ggplot, the function geom_smooth was added with the following option: method = “gam”, formula = y ~ s(x).

DepMap dependency analysis

DepMap compound screen values were obtained via the DepMap portal / Data Explorer at https://depmap.org/portal/interactive/ for the cancer types indicated and plotted as indicated via GraphPad PRISM.

Gene Ontology analysis

Ensemble gene IDs were converted to symbols in R and were further analysed by Enrichr to obtain GO Biological Process enrichment figures. Ensembl gene ID lists for enhancer – promoter interaction genes only (EPIonly), promoter – promoter interaction genes only (PPIonly) and enhancer – promoter with promoter – promoter interactions genes (EPIPPI) were obtained and loaded in R using read.table() with readr package (version 2.1.5). Because Ensemble gene IDs with decimal points were not valid in R, decimal points were removed using gsub(). The Ensembl gene IDs were converted into Gene symbols with org.Hs.eg.db package (version 3.19.1) and clusterProfiler package (version 4.12.0) with the added options: keytype = “ENSEMBL”, column = “SYMBOL”. In order to enable Enrichr to process the gene symbol lists, NAs were removed from the list by na.omit(). The gene symbol lists were exported as a txt file using writeLines(). The lists were then uploaded to Enrichr website (available at: https://maayanlab.cloud/Enrichr/) to obtain GO Biological Process 2023 enrichment table. Finally, the bar charts were generated based on the table by plotEnrich() using enrichR package (version 3.2) with the added options: showTerms = 15, numChar = 80, y = “Count”, orderBy = “P.value”.

Supplementary References:

1 Kim, J. C. *et al.* Transcriptomic classes of BCR-ABL1 lymphoblastic leukemia. *Nat Genet* **55**, 1186-1197 (2023). https://doi.org:10.1038/s41588-023-01429-4

2 Behrens, K. *et al.* ERG and c-MYC regulate a critical gene network in BCR::ABL1-driven B cell acute lymphoblastic leukemia. *Sci Adv* **10**, eadj8803 (2024). https://doi.org:10.1126/sciadv.adj8803

3 Chen, Z. *et al.* Signalling thresholds and negative B-cell selection in acute lymphoblastic leukaemia. *Nature* **521**, 357-361 (2015). https://doi.org:10.1038/nature14231

4 Ward, I. M., Minn, K., van Deursen, J. & Chen, J. p53 Binding protein 53BP1 is required for DNA damage responses and tumor suppression in mice. *Mol Cell Biol* **23**, 2556-2563 (2003).

5 Boulianne, B. *et al.* Lineage-Specific Genes Are Prominent DNA Damage Hotspots during Leukemic Transformation of B Cell Precursors. *Cell Rep* **18**, 1687-1698 (2017). https://doi.org:10.1016/j.celrep.2017.01.057

6 Kaneshige, A. *et al.* A selective small-molecule STAT5 PROTAC degrader capable of achieving tumor regression in vivo. *Nat Chem Biol* **19**, 703-711 (2023). https://doi.org:10.1038/s41589-022-01248-4

7 Kaneshige, A. *et al.* Discovery of AK-1690: A Potent and Highly Selective STAT6 PROTAC Degrader. *J Med Chem* **68**, 5125-5151 (2025). https://doi.org:10.1021/acs.jmedchem.4c01009

8 Xu, R. *et al.* Discovery of SD-436: A Potent, Highly Selective and Efficacious STAT3 PROTAC Degrader Capable of Achieving Complete and Long-Lasting Tumor Regression. *J Med Chem* **67**, 20495-20513 (2024). https://doi.org:10.1021/acs.jmedchem.4c01946

9 Buehler, E., Chen, Y. C. & Martin, S. C911: A bench-level control for sequence specific siRNA off-target effects. *PLoS One* **7**, e51942 (2012). https://doi.org:10.1371/journal.pone.0051942

10 Pfeifer, M. *et al.* SSB1/SSB2 Proteins Safeguard B Cell Development by Protecting the Genomes of B Cell Precursors. *J Immunol* **202**, 3423-3433 (2019). https://doi.org:10.4049/jimmunol.1801618

11 Ng, H. L. *et al.* Promoter-centred chromatin interactions associated with EVI1 expression in EVI1+3q- myeloid leukaemia cells. *Br J Haematol* **204**, 945-958 (2024). https://doi.org:10.1111/bjh.19322

12 Freire-Pritchett, P. *et al.* Detecting chromosomal interactions in Capture Hi-C data with CHiCAGO and companion tools. *Nat Protoc* **16**, 4144-4176 (2021). https://doi.org:10.1038/s41596-021-00567-5