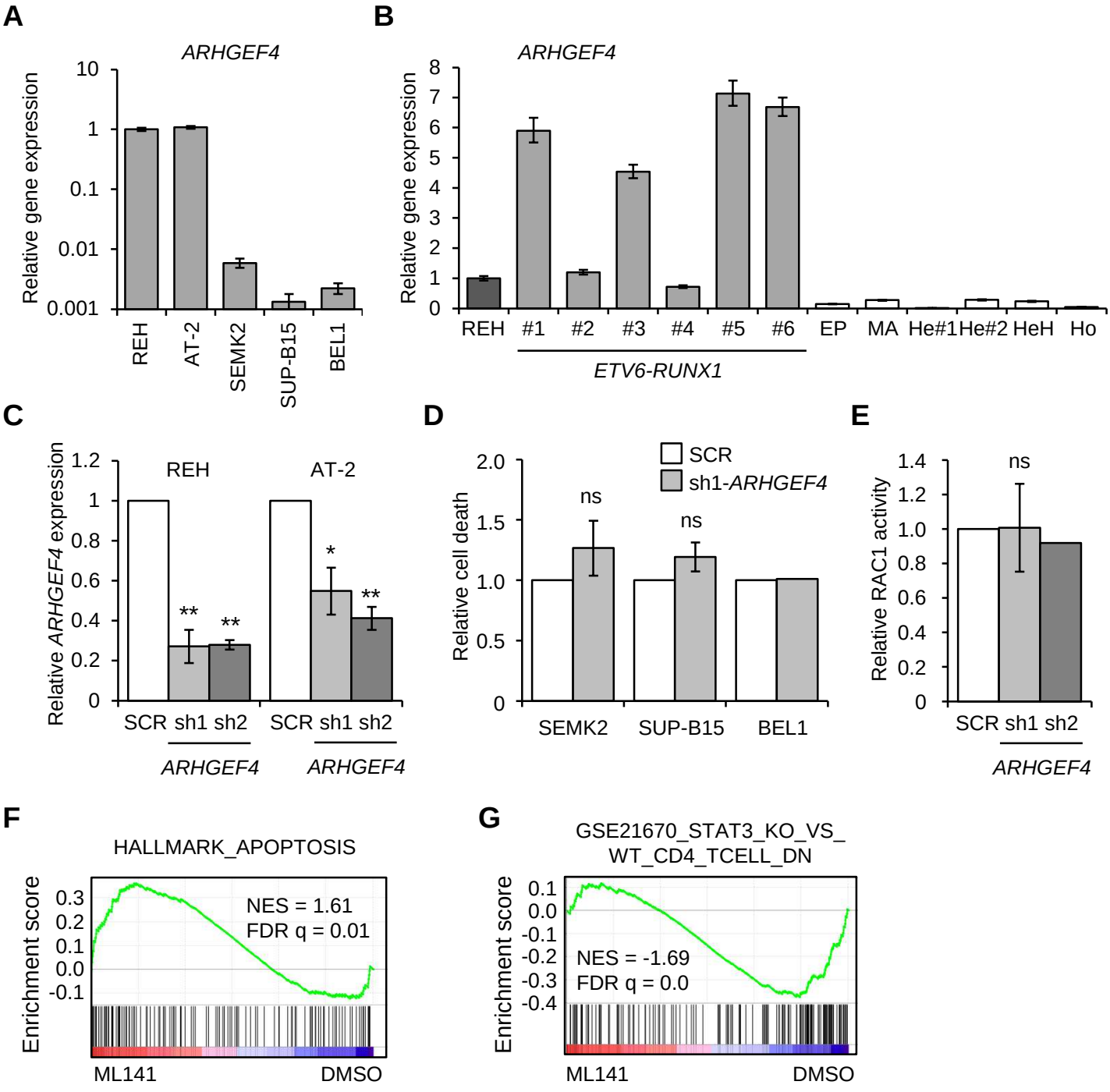
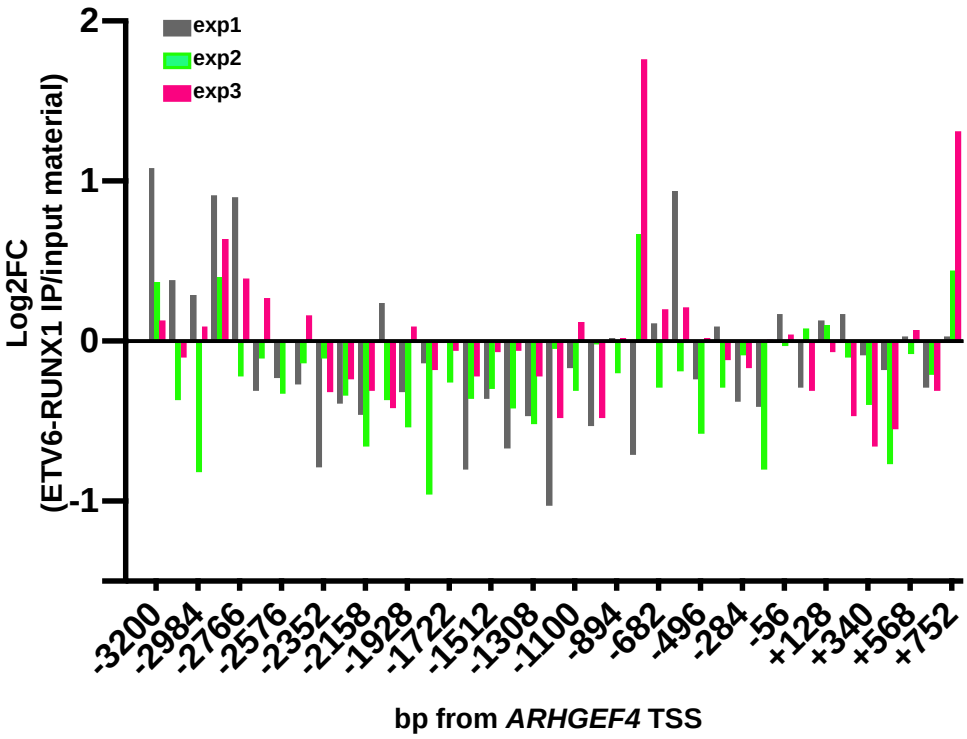


Supplementary Figure 1



Supplementary Figure 2

Supplementary Figure Legends

Supplementary Figure 1. *ARHGEF4* is downstream of ETV6-RUNX1 and is required for t(12;21) ALL survival and disease progression. (A) Quantitative RT-PCR (qPCR) analysis of *ARHGEF4* gene expression in a panel of human pre-B ALL cell lines: REH and AT2 (*ETV6-RUNX1*$^{+}$), SEMK2 and BEL1 (*MLL-AF4*$^{+}$), and SUP-B15 (*BCR-ABL*$^{+}$). Data is normalized to expression levels in REH cells. Bars and error bars are means and SD of quadruplicate measurements. (B) qPCR analysis of *ARHGEF4* gene expression in a panel of pre-B patient-derived xenograft (PDX) ALL samples, harvested from NSG mice transplanted with diagnostic primary ALL samples. The panel consisted of: 6 *ETV6-RUNX1*$^{+}$ (#1-#6), an *E2A-PBX1*$^{+}$ (EP), an *MLL-AF9*$^{+}$ (MA), 2 Hyperdiploid, a High-Hyperdiploid (HeH) and a Hypodiploid (Ho) ALL primograft samples. Data is normalized to expression levels in REH cells. Bars and error bars are means and SD of quadruplicate measurements. (C) qPCR analysis of *ARHGEF4* gene expression in REH and AT-2 cells 5 days after lentiviral transduction with control scramble (SCR) or *ARHGEF4*-specific (sh1 and sh2) shRNA. Bars and error bars are means and SD of 3 independent experiments. Data are normalized to expression levels in SCR cells. *$P < 0.05$; **$P < 0.01$ (relative to SCR), one sample *t*-test. (D) Cell death in *ETV6-RUNX1*$^{-}$ SEMK2, SUP-B15 and BEL1 ALL cells 5 days after lentiviral transduction with control scramble (SCR) or *ARHGEF4*-specific (sh1-*ARHGEF4*) shRNA. Bars and error bars are means and SD of 3 independent experiments for SEMK2 and SUP-B15, and one experiment for BEL1. Data are normalized to cell death in SCR cells. $^{ns}P > 0.05$, one sample *t*-test. (E) Analysis of RAC1 activity in REH cells 3 days after lentiviral transduction with control scramble (SCR) or *ARHGEF4*-specific (sh1 and sh2) shRNA. Bars and error bars are

means and SD of 5 (SCR and sh1) and 1 (sh2) independent experiments. $^{ns}P > 0.05$, one sample *t*-test. (F) GSEA of the HALLMARK_APOPTOSIS gene set (https://www.gsea-msigdb.org/gsea/msigdb/cards/HALLMARK_APOPTOSIS.html), and (G) of a STAT3 target gene set, derived from a list of genes whose expression was previously shown to decrease in *Stat3*-deficient CD4 T cells (Durant et al, 2010), in RNA-seq data of ML141 treated REH cells.

Supplementary Figure 2. Analysis of data from Linka et al (2013) demonstrating binding of ETV6-RUNX1 to the ARHGEF4 promoter. The figure illustrates the ETV6-RUNX1 binding peaks in three replicate experiments of the *ARHGEF4* promoter region in NALM6 cells, ectopically expressing the fusion. The data was reported in Linka et al (2013) using NimbleGen Human ChIP-chip 3x720K RefSeq Promoter Arrays. Data on the y-axis represent scaled log2 ratios (ETV6-RUNX1 IP/input material) and the x-axis depicts the distance from the *ARHGEF4* transcriptional start site (TSS) using the hg18 human genome assembly.

Supplementary Table 1. Patient sample characteristics.

Patient id	Sex	Age (yrs)	Cytogenetics
ETV6-RUNX1 #4	M	4.62	ETV6-RUNX1 rearrangement detected in 76% of interphase nuclei.
ETV6-RUNX1 #5	F	1.92	ETV6-RUNX1 rearrangement detected in 60% of interphase nuclei.
ETV6-RUNX1 #6	M	3.37	ETV6-RUNX1 fusion detected, showing clonal evolution involving gain of RUNX1 signal, by interphase FISH analysis.
EP	F	10.95	Abnormal karyotype with a der(19)t(1;19) unbalanced translocation by G-banded analysis. Unbalanced TCF3-PBX1 rearrangement in 52% of nuclei and mono-allelic loss of CDKN2A in 5.5% nuclei by interphase FISH analysis.
MA	F	2.52	KMT2A rearragment with concomitant deletion of the 3'KMT2A region. t(9;11)(p22;q23).
Hyperdyploid #1	M	5.1	Complex abnormal karyotype with multiple unbalanced structural rearrangements which included a PAX5 rearrangement, relative gain of 1q, loss of 1p, CDKN2A and MECOM signals.
Hyperdyploid #2	M	2.36	Gain of DXZ1, D9Z3, ABL1, RUNX1 and relative loss of one CDKN2A signal by FISH. Gain of chromosomes X, 9 and 21 by G-banded analysis.
High-Hyperdyploid	F	2.48	High-hyperdiploid karyotype with gain of chromosomes X, 3, 5, 8, 10, 11, 12, 14, 17, 18, 21 and 22.
Hypodyploid	M	11.8	Hypodiploid with a doubled-up clone in a sub-population showing additional structural abnormalities, indicative of clonal evolution. FISH analysis showed loss of ETV6 (73%), relative loss of centromere 17 (54%), and gain of signals for 11 and 17 centromeres (27%).

*ETV6-RUNX1 #1-3: patient samples characteristics previously reported in Mangolini et al (2013)

SUPPLEMENTARY MATERIALS AND METHODS

Mice

NOD-SCID-$\gamma^{-/-}$ (NSG) mice were maintained in the animal facilities of the UCL GOS ICH, London. All mouse experiments were performed according to and approved by the United Kingdom Home Office regulations and followed UCL GOSICH institutional guidelines. 6-12 week old NOD.Cg-PrkdcscidIl2rgtm1Wjl/SzJ (NSG) mice were used as recipients for transplantation of 1 x 10^5 viable REH cells, 3 days after transduction with SCR and *shARHGEF4* shRNA. Mice were sacrificed upon developing clinical signs of disease.

PDX B-ALL cells

Ethical approval was given (Research Ethics Committee reference 14/EM/0134) for use of appropriately consented material from patients with B-lineage ALL at Great Ormond Street Hospital for Children (London, UK). 1-2 x 10^6 mononuclear cells were intra-bone injected into non-irradiated 5- to 12-week-old NSG mice. Recipient mice were sacrificed upon developing clinical signs of disease. Human PDX B-ALL cells were harvested and purified from spleens using the mouse cell depletion kit (Miltenyi Biotec).

Cell Culture

The HEK 293FT (Thermofisher) packaging cell line was cultured in Dublecco's Modified Medium (DMEM), supplemented with 10% Foetal Bovine Serum (FCS), 100 U/ml Penicillin (Sigma-Aldrich), 100 µg/ml Streptomycin (Sigma-Aldrich) and 2 mM L-glutamine (Sigma-Aldrich). REH and SUP-B15 cells were from German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany), SEMK2 and BEL1 were kindly donated by Dr. R.W. Stam (Rotterdam, the Netherlands), and AT-2

cells were the kind gift of Prof R. Panzer-Grümayer (Vienna, Austria). Leukemic cell lines were cultured in Roswell Park Memorial Institute (RPMI), supplemented with 10% FCS, 100 U/ml Penicillin, 100 µg/ml Streptomycin and 2 mM L- glutamine. All cell lines were authenticated by short tandem repeat profiling using the PowerPlex 16 system (Promega) and mycoplasma negative status confirmed using the MycoAlert Mycoplasma Detection Kit (Lonza). For CDC42 inhibition, ML141 and CASIN were purchased from Sigma-Aldrich and Cayman Chemical, respectively.

Flow cytometry

Phospho-STAT3 (pY705) expression was measured 24 hours after treatment of REH cells and 16 hours after treatment of PDX samples with CDC42 inhibitors. For intracellular staining, human leukemic cells were lysed and fixed in a single step using BD™ Phosflow Lyse/Fix buffer (BD Phosflow), permeabilized with BD™ Phosflow Perm Buffer III (BD Phosflow) and stained with PE-conjugated anti-STAT3 (pY705) according to manufacturer's instructions (BD Biosciences). Data were normalized to expression levels in DMSO controls, after removal of isotype background. Apoptosis was measured in REH cells 5 days after transduction with *ARHGEF4*-specific shRNA using the Annexin V Apoptosis Detection Kit I (Thermo Fisher Scientific), 48 hours after treatment with the Dead cell removal kit (Miltenyi Biotec). Cell death was measured in ALL cell lines and PDX samples by flow cytometric analysis of cells stained with PI (Sigma-Aldrich) or TO-PRO-3 (Thermo Fisher). Cells were acquired on an LSRII (BD Bioscience) and the data was analyzed with Summit 4.3 software (Beckman Coulter Life Sciences).

Colony-forming assays

Colony-forming ability was assessed by plating REH cells in methylcellulose (HSC002, Bio-Techne) 3 days after lentiviral transduction with control scramble or *ARHGEF4*-

specific shRNA. Colonies were stained with 1 mg/ml p-iodonitrotetrazolium after 14 days of culture.

RNA sequencing (RNA-seq) and Gene set enrichment analysis (GSEA)

Total cellular RNA was purified from DMSO and 25μM ML141 treated REH samples from three independent experiments and submitted to UCL Genomics for RNA-sequencing. RNA was extracted using the RNeasy Plus Mini Kit (Qiagen) according to manufacturer's instructions. 100 ng of RNA per sample were analysed using Bioanalyser 2100 (Agilent Technologies, Santa Clara, CA) to verify RNA integrity prior to amplification. Samples were processed using the Illumina TruSeq RNA sample prep kit Version2 (p/n RS-122-2001) according to manufacturer's instructions (Illumina, Cambridge, UK). Briefly, mRNA was selected using paramagnetic dT beads and fragmented by metal hydrolysis to approximately 150 bp lengths. Random primed cDNA was then generated and adapters compatible with Illumina sequencing were ligated before being enriched by 12 cycles of PCR. Libraries were quantified, normalised and pooled before sequencing on an Illumina NextSeq 500, generating approximately 20 million 43 bp read pairs per sample. Fastq was then demultiplexed and generated using Illumina bcl2fastq v2.19 before pre-processing (trimmomatic) to remove adapter read-through and poor-quality sequences. Pre-processed data were then aligned to the genome (UCSC hg38) with Bowtie 2 and deduplicated using Picard. Reads-per-transcript were counted by FeatureCounts, before differential expression analysis by SARTools, a DESeq2 wrapper. All tools were invoked through the Galaxy Project for NGS analysis. GSEA (https://software.broadinstitute.org/gsea/) was used to examine enrichment of STAT3 and apoptosis gene sets.

Quantitative RT-PCR analysis

Total RNA was isolated from cells using RNeasy Mini Kit (Qiagen) according to the

manufacturer's instructions. RNA was converted into cDNA using a cDNA synthesis kit (Thermo Fisher Scientific) according to the manufacturer's instructions. Samples were treated with DNase I (Thermo Fisher Scientific) prior to reverse transcription using the Moloney murine leukemia virus reverse transcriptase (Thermo Fisher Scientific). qPCR was performed on isolated mRNA using TaqMan probe based chemistry and an ABI Prism 7900HT fast Sequence Detection System (Thermo Fisher Scientific). All primer/probe sets were from Thermo Fisher Scientific.

GTPase activity assays

GTP-bound CDC42 and RAC1 were measured using G-LISA Activation Assay Kits (Cytoskeleton, CDC42 Cat#BK127, RAC1 Cat#BK126). REH cells were washed with cold PBS and lysed. Lysates were centrifuged at 10,000 × g for 1 min (CDC42) or at 17000 × g for 2 min (RAC1) at 4 °C. Supernatants were aliquoted, snap-frozen in liquid nitrogen, and stored at −80 °C. Protein concentrations were determined, and CDC42/RAC1 GTPase activity was assessed according to the manufacturer's instructions. The luminescence signal or the absorbance signal were quantified using a microplate reader (Infinite M200 Pro, TECAN).

Lentiviral transduction of human cell lines

293FT packaging cells (Thermo Fisher Scientific) were transiently co-transfected with the lentiviral expression vectors, the pCMV-PAX2 construct and the pVSV-G envelope construct (kind gifts of Prof D. Trono, Lausanne, Switzerland). Human leukemia cells were transduced with lentiviral supernatant by spinoculation at 700g, 25°C for 45 minutes in the presence of 5μg/mL polybrene. After 2 days, transduced cells were selected in puromycin for 1-3 days. The *ETV6-RUNX1* cDNA was cloned into the pCSGW-PIG vector,[1] made by replacing the GFP cDNA from pCSGW[2] with a puro-IRES-GFP cassette. Lentiviral MISSION shRNA constructs targeting *ETV6* (Clone ID:

NM_001987.x-309s1c1), *ARHGEF4* (sh1, Clone ID:NM_015320.2-2212s1c1; sh2, Clone ID:NM_015320.2-820s1c1; sh3, Clone ID:NM_015320.2-2098s1c1) and the scramble (SCR) non-silencing control (SHC002) were purchased from Sigma-Aldrich. The *ETV6*-specific shRNA was used to specifically silence *ETV6-RUNX1* fusion gene expression in REH cells, since these cells have deleted the wild-type *ETV6* allele.

Statistics

All data are represented as mean ± SD. Statistical significance was determined using Prism (GraphPad) software. Statistical analysis of survival curves was performed using the Mantel-Haenszel log-rank test. Statistical analysis of means was performed using the one sample t test or unpaired Student's t test, two-tailed P values < 0.05 being considered statistically significant.

References

1. Walf-Vorderwulbecke V, Pearce K, Brooks T, et al. Targeting acute myeloid leukemia by drug-induced c-MYB degradation. *Leukemia*. Apr 2018;32(4):882-889.
2. Demaison C, Parsley K, Brouns G, et al. High-level transduction and gene expression in hematopoietic repopulating cells using a human immunodeficiency [correction of imunodeficiency] virus type 1-based lentiviral vector containing an internal spleen focus forming virus promoter. Research Support, Non-U.S. Gov't. *Human gene therapy*. May 01 2002;13(7):803-813.