Cell cycle corruption in a preleukemic ETV6-RUNX1 model exposes RUNX1 addiction as a therapeutic target in acute lymphoblastic leukemia

Jason P Wray[1,8], Elitza M Deltcheva[1,8], Charlotta Boiers[2], Simon E Richardson[3], Jyoti Bikram Chettri[1], Sladjana Gagrica[4], Yanping Guo[1], Anuradha Illendula[5], Joost HA Martens[6], Hendrik G Stunnenberg[6], John H Bushweller[5], Rachael Nimmo[7], Tariq Enver[1,2,9]*

[1] *Department of Cancer Biology UCL Cancer Institute, UCL, London, UK*

[2] *Molecular Medicine and Gene Therapy, Lund Stem Cell Center, Lund University, Lund, Sweden*

[3] *Wellcome-MRC Cambridge Stem Cell Institute, Cambridge, UK*

[4] *IMED Oncology, AstraZeneca, Cancer Research UK Cambridge Institute, Cambridge, UK*

[5] *Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22908, USA*

[6] *Department of Molecular Biology, Faculty of Science, Nijmegen Centre for Molecular Life Sciences, Radboud University, Nijmegen, the Netherlands*

[7] *Oxford Biomedica (UK) Ltd, Windrush Court, Transport Way, Oxford, OX4 6LT, UK.*

[8] *These authors contributed equally*

[9] *Lead contact*

** correspondence: t.enver@ucl.ac.uk*

Summary

ETV6-RUNX1 is the most common translocation in pediatric B-acute lymphoblastic leukemia (B-ALL), shown to arise *in utero* and to initiate a clinically silent pre-leukemic state. Here, we integrate chromatin immunoprecipitation sequencing (ChIP-seq) and RNA-sequencing (RNA-seq) to characterize the authentic ETV6-RUNX1 regulome and explore its interplay with the native RUNX1 as a substrate for preleukemia. We show that preleukemic initiation is primarily mediated through competition for RUNX1 binding sites and transcriptional repression of the RUNX1 program. Using mass cytometry we functionally demonstrate that as a "first-hit" ETV6-RUNX1 induces a cell cycle impairment in a human pluripotent stem cell (hPSC) model, antagonizing RUNX1-mediated cell cycle regulation. Strikingly, in overt leukemia, genetic perturbation of native RUNX1 protein or its co-factor CBFβ leads to cell death of pediatric and adult B-ALL cells. We demonstrate that RUNX1 addiction can be therapeutically exploited by pharmacological inhibition using an allosteric CBFβ inhibitor.

Keywords

B-cell acute lymphoblastic leukemia, transcription factors, regulome, ETV6-RUNX1, CBF complex, first-hit, mass cytometry, cell cycle, allosteric inhibition, p53, Myc, E2F

Introduction

Childhood B-acute lymphoblastic leukaemia (B-ALL) is the most common pediatric cancer and is clinically distinct from adult ALL, characterized by a distinct mutational spectrum, higher incidence and overall good prognosis. However, despite high cure rates in low risk cases, B-ALL remains a leading cause of cancer-related death in children. Current chemotherapy regimens are highly toxic, associated with severe long-term sequelae and relapse occurs in ~20% of patients. While immunotherapeutic approaches have contributed to improved outcomes, toxicity and relapse remain significant concerns.

Most pediatric leukemias are likely to originate during embryonic and fetal development and are often associated with chromosomal rearrangements that disrupt the normal function of hematopoietic transcription factors (Greaves, 2018). In B-ALL, a paradigmatic example of an oncogene arising *in utero* is the ETS translocation variant 6 (ETV6) - runt-related transcription factor 1 (RUNX1) fusion, affecting two major hematopoietic regulators. ETV6-RUNX1 accounts for ~25% of pediatric B-ALL cases and is detected in 1% of all newborns, as demonstrated by studies of neonatal blood spots and monochorionic twin (Greaves and Wiemels, 2003). However, the translocation is not sufficient to induce clinically overt disease and only 1% of children carrying the fusion will transition from the clinically silent pre-leukaemic state into full blown leukemia. This transition is dependent on the acquisition of secondary mutations that evolve in a branching evolutionary pattern and contribute to the genetic heterogeneity of B-ALL, a confounding factor for both conventional and targeted therapeutic approaches. While key second-hit mutations in ETV6-RUNX1^{+} leukemia involve loss of *CDKN2A* and the second copy of *ETV6* (Sulong et al., 2009, Tsuzuki et al., 2007, Mullighan et al., 2007), the fusion gene and the native RUNX1 allele are rarely lost, suggesting that the ETV6-RUNX1 - RUNX1 axis may be required for leukemic maintenance.

Structurally, ETV6-RUNX1 fuses almost the entire RUNX1 protein to the N-terminal portion of ETV6 and has been proposed to recruit repressors to RUNX1 target genes (Hiebert et al., 1996, Guidez et al., 2000, Hiebert et al., 2001, Morrow et al., 2007, Wiemels and Greaves, 1999). Several studies have examined the function of ETV6-RUNX1 by characterizing its targets in murine or human cell lines (Linka et al., 2013, Teppo et al., 2016, Fuka et al., 2011). Whilst providing useful insights, these do not address the role of ETV6-RUNX1 as a “first-hit” in a

preleukemic context and in the absence of second hit mutations. Furthermore, while multiple detailed studies have interrogated the relationship between native RUNX1 and the RUNX1-ETO fusion protein in acute myeloid leukemia (AML), such analyses are lacking in B-ALL (Ben-Ami et al., 2013, Ptasinska et al., 2014).

RUNX1 and its co-factor CBFβ, together comprising the Core Binding Factor (CBF) complex, are the most frequent targets of mutations in hematological malignancies. Genetic changes affecting RUNX1, such as translocations or point mutations, are generally linked to loss of function, broadly classifying it as a tumor suppressor (Osato, 2004). Notably, these mutations are usually heterozygous and the second, wild-type copy is rarely lost and often amplified, suggesting that retention of RUNX1 might offer a selective advantage for leukemic cells (Mikhail et al., 2002, Mullighan, 2012, Attarbaschi et al., 2004). Recent reports have indeed shown a supporting role for RUNX1 in leukemogenesis, primarily in AML, but also in T-ALL and mixed-lineage leukemia (MLL), offering a therapeutic rationale for targeting RUNX1 and its cofactor CBFβ (Wilkinson et al., 2013, Choi et al., 2017, Ben-Ami et al., 2013, Ptasinska et al., 2014). Accordingly, several structural derivates of an allosteric CBFβ inhibitor have been tested in these diseases, with promising results (Illendula et al., 2015, Oo et al., 2018). It remains to be assessed if B-ALL cells are dependent on RUNX1, but its requirement for the survival of early mouse B-cell progenitors suggests that its inhibition may represent a therapeutic opportunity (Niebuhr et al., 2013).

Here, we characterize the ETV6-RUNX1 chimeric- and native RUNX1-responsive regulomes and explore their relationship in the context of preleukemia and overt disease. Using mass cytometry we functionally demonstrate that as a “first-hit” ETV6-RUNX1 alters cell cycle of early B-cell progenitors by hijacking and repressing RUNX1 targets. The functional antagonism between the two transcription factors exposes an addiction to RUNX1-activity in overt leukaemia, spanning pediatric and adult B-ALL subtypes. We demonstrate that this vulnerability has therapeutic potential and can be exploited using an allosteric inhibitor of the CBF complex.

Results

Delineating the ETV6-RUNX1 regulome in childhood B-ALL

To identify direct targets of ETV6-RUNX1 we performed genome-wide ChIP-seq in the B-precursor cell line Reh, a widely used model system for t(12;21) ALL. As Reh expresses the fusion gene, but lacks the second ETV6 allele, ETV6-RUNX1 can be specifically immunoprecipitated using ETV6 antibodies. Two independent antibodies, HPA and PAS, were used to increase stringency of peak assignment. In addition, authenticity of peaks was further confirmed by integrating the ChIP-seq data with DNaseI hypersensitivity sequencing (DNaseI-HS) from the same cells. Only peaks bound by both antibodies and located within DNaseI sites were considered for further analysis (Figure 1A, upper panel, marked with red dashed line**,** Figure S1A). This approach identified 1171 high-confidence loci that were highly enriched for RUNX, ETS, Ascl2 (NHLH1) and SP1 binding motifs among others (Figure 1B, Table S1).

To validate the clinical relevance of the peaks identified in our cell line model, we performed ChIP-seq in three t(12;21) patient samples using the same two *ETV6*-specific antibodies and compared their binding patterns to Reh cells (Figure 1A, bottom panel, marked with blue dashed line, Figure S1B). Notably, our analysis revealed a strong correlation across datasets and the majority of high-confidence sites identified in Reh cells were also found in patients, thus verifying their authenticity in a patient context (Figure 1C). Peaks detected in all eight samples (711 in total) had the highest binding affinity (Figure S1C). Motif analysis confirmed a preference for canonical RUNX and ETS motifs with a prevalence of RUNX-motif in the higher affinity peaks, consistent with ETV6-RUNX1 interacting with chromatin primarily through the runt domain (Figure S1D and S1E). Further, to gain insight into the genomic distribution of ETV6-RUNX1, we assigned peaks to the nearest gene up to a maximum of 100kb from the transcriptional t site (TSS). The majority of high-confidence peaks were found to be intragenic or upstream of the TSS with a small proportion overlapping with the TSS (Figure 1D).

We next examined the relationship between ETV6-RUNX1 target binding and transcriptional regulation, taking advantage of publicly available transcriptome datasets from diagnostic childhood B-ALLs (Gu et al., 2019). We compared ETV6-RUNX1^{+} samples to all other subtypes, ranking genes according to their relative expression. A trend towards increased ETV6-RUNX1 binding was observed regardless of expression directionality with a statistically significant enrichment for downregulated genes (Figure 1E).

Taken together, our data suggest, in a patient-derived cell line and primary patient samples, that ETV6-RUNX1 binds to DNA primarily through the Runt domain and that binding is associated with both positive and negative regulation of transcription, with a tendency towards repression.

ETV6-RUNX1 **induces transcriptional changes indicative of cell cycle repression in pre-leukemia**

In order to link ETV6-RUNX1 genome occupancy with transcriptional regulation, we next performed RNA-seq following knockdown of the fusion gene in Reh cells using two independent shRNAs (Figure S2A). Differentially expressed genes (DEGs) were found to be significantly enriched for ETV6-RUNX1 binding, with a predominance of upregulated over downregulated genes (Figure 2A, left panel). These data suggest that, in a cell line model of overt leukaemia with acquired second hits, the fusion behaves primarily as a repressor, in agreement with our analysis of the patient cohort (Figure 1E).

As ETV6-RUNX1 is an initiating event arising *in utero*, its initial impact is on a fetal cell in the absence of additional mutations acquired during leukaemic progression. To explore this, we made use of a hIPSC ETV6-RUNX1 knock-in system developed in our lab, which displays a partial block in B-cell differentiation (Figures S2B and S2C) (Boiers et al., 2018). Comparison of our ChIP-seq data with the gene expression profile of iPS derived-ETV6-RUNX1 proB cells showed significant enrichment for ETV6-RUNX1 binding among the DEGs, with a prevalence of downregulated over upregulated genes (Figures 2A, right panel). In addition, differentially expressed genes from ETV6-RUNX1 knock-in and Reh knockdown were strongly anticorrelated demonstrating the preservation of the ETV6-RUNX1 regulome across these model systems (Figure 2B). Together, these data show that in a preleukemic context ETV6-RUNX1 acts predominantly, but not exclusively, as a repressor and that this function is preserved in overt leukemia.

We next used gene set enrichment analysis (GSEA) to identify pathways regulated by ETV6-RUNX1. Comparison to the Hallmark database identified “MYC targets V1”, “E2F targets” and “G2M checkpoint” as significantly enriched in the downregulated genes in ETV6-RUNX1-expressing proB cells (Figure 2C, upper panel, and S2D) (Subramanian et al., 2005, Liberzon et al., 2015). Conversely, these gene sets were upregulated following ETV6-RUNX1 knockdown, consistent with repression by ETV6-RUNX1 (Figure 2C, lower panel). Since E2F transcription factors are regulators of the G1 to S phase transition and MYC is broadly

implicated in cell cycle regulation, these results suggest that ETV6-RUNX1 has a negative impact on the cell cycle of proB cells in a preleukemic setting.

While ETV6-RUNX1$^+$ ALL invariably has an immature B-cell phenotype, various experimental systems have reported an impact of the fusion on other compartments of the haematopoietic hierarchy (BaF3, OwenW). We therefore examined the transcriptional profiles of ETV6-RUNX1-expressing HSCs and IL7R$^+$ progenitors, found upstream of proB cells, and of pre-B cells. Intriguingly, MYC - and cell cycle - related gene signatures were similarly perturbed in the HSC and IL7R$^+$ progenitors. However, in the preB population MYC, but not cell cycle signatures, were affected, likely reflecting a differentiation block escape of these cells (Figure S2D).

Interestingly, even in overt leukaemia a subset of cells remain slow-cycling and are associated with resistance to therapy (Lutz et al., 2013, Ebinger et al., 2016). We examined the impact of ETV6-RUNX1 on a gene signature associated with dormant, label-retaining cells (LRCs) in B-ALL (Ebinger et al., 2016). Principal component analysis (PCA) limited to the LRC signature genes separated samples according to ETV6-RUNX1 status. The greatest impact was on ETV6-RUNX1 expressing proB cells which clustered with the IL7R$^+$ population showing that, with respect to the LRC signature, ETV6-RUNX1$^+$ proB cells are similar to more immature progenitors (Figure 2D and S2E). In wild-type cells the LRC signature was more highly expressed in HSPC/IL7R+ populations, becoming downregulated in proB cells. Notably, this downregulation was delayed in ETV6-RUNX1$^+$ cells until the preB stage (Figures 2E). These results imply that the fusion causes a cell cycle impairment in early progenitors that appears to resolve in the more mature compartments.

In summary, we demonstrate that ETV6-RUNX1 acts predominantly as a transcriptional repressor, both as a first-hit and in the context of full blown leukemia with acquired second hits. In addition, our transcriptional analysis shows that ETV6-RUNX1 represses cell cycle-related signatures and that proB cells, corresponding to the point of differentiation arrest in precursor B-ALL, are most significantly impacted.

Mass cytometry reveals an accumulation of ETV6-RUNX1 expressing progenitor cells in S phase

Our transcriptional analysis strongly suggests that ETV6-RUNX1 has an impact on the cell cycle of B-cell progenitors. Given the limited numbers of differentiated cells obtained with the

hIPSC *ETV6-RUNX1* model, to directly assess whether cell cycle is altered upon ETV6-RUNX1 expression, we used mass cytometry time-of-flight (CyTOF) which enables the simultaneous interrogation of surface and intracellular markers across multiple barcoded samples (Table S2, see the Methods) (Behbehani et al., 2012). Two *ETV6-RUNX1* knock-in hiPSC clones and two control lines, the parental MIFF3 and an ETV6-RUNX1 clone that has been reverted to wild-type, were subjected to a 30-31 day differentiation protocol that results in a mixture of immature B-cells, progenitors and other haematopoietic cell types (Figure S2C) (Boiers et al., 2018, Carpenter et al., 2011). Cell populations were clustered and visualized with SOM (Self-organizing map) and UMAP (Uniform Manifold Approximation and Projection) dimensionality reduction tools (see the Methods). Non-haematopoietic cells ($CD45^-$) were excluded and the remaining clusters manually annotated according to surface marker expression levels, including HSPCs ($CD34^+CD19^-$) and preB ($CD34^-CD19^+$) cells (Figures 3A, 3B and S3A). Consistent with our previous work, our analysis showed that both fusion-expressing lines have markedly decreased B-cell output while an increased proportion of HSPCs and CD19-low cells was observed, indicative of a partial block in differentiation (Figures 3A and 3C).
To assess the cell cycle status of the different, immuno-phenotypically defined cell populations, we used clustering to assign each cell to a cell cycle phase based on the markers pRb, IdU, CycB1 and pHisH3 (Figures 3D, S3B and S3C) (Behbehani et al., 2012). Strikingly, *ETV6-RUNX1*-expressing HSPC and CD19-low populations had an altered cell cycle profile with an increased proportion of cells in S-phase as compared to controls (Figure 3E). This difference was not observed in preB cells suggesting that the cell cycle is not altered in those cells which progress beyond the differentiation block.

Taken together, our data show, in a developmentally relevant “first-hit” pre-leukemic model, that ETV6-RUNX1 alters the cell cycle profile of early compartments within the B cell differentiation hierarchy.

ETV6-RUNX1 binds DNA through the Runt domain, competing with native RUNX1

Our data suggest that ETV6-RUNX1 hijacks the native RUNX1-driven network by binding to RUNX1 targets through the runt domain. Exploring this in an ETV6-RUNX1$^+$ context is difficult, in that a pan-RUNX1 antibody will recognize both the fusion protein and native RUNX1. To directly interrogate their relationship, we used the NALM-6 pre-B ALL cell line which expresses wild-type RUNX1, generating a derivative expressing V5-tagged ETV6-RUNX1, facilitating independent immuno-precipitation of RUNX1 and the fusion. To directly assess the requirement for the Runt and helix-loop-helix (HLH) pointed domains for ETV6-RUNX1

function, we expressed two mutant versions of the fusion carrying a loss-of-function point mutation (R139G) and a HLH deletion (ΔHLH) respectively (Figure S4A).

Comparison of RUNX1 and V5-*ETV6-RUNX1* ChIP-seq datasets revealed a strong correlation of the binding affinities of the two transcription factors, consistent with their having similar DNA-binding properties (Figure 4A, 4B (red dashed line), S4B, S4C and S4D). The most significantly enriched motifs across all binding sites were RUNX and ETS as observed for native ETV6-RUNX1 ChIP in Reh and patient samples (Figures 4C, 1A and Table S3).

ChIP-seq in the R139G mutant line showed an almost complete loss of binding, formally proving that DNA interaction is dependent on the Runt domain (Figures 4A, 4B (blue dashed line), S4B, S4C, S4D). The fusion inherits the HLH domain from ETV6, which mediates dimerization between members of the ETS family of transcription factors. Interestingly, the ΔHLH exhibited increased binding affinity, suggesting that the HLH domain may in fact inhibit interaction with DNA (Figures 4B, blue dashed line, S4B, S4C and S4D). To explore this further we categorised ChIP peaks into those more strongly bound by the fusion (ER) or RUNX1 (R1) or similarly bound by both (R1_ER) (Figures 4D, 4E). ER peaks tended to be further from the TSS and a slightly higher proportion had one or more RUNX motifs, reflecting the increase in RUNX1 motif frequency with distance from the TSS (Figure S4E, S4F). All three categories of peaks were strongly bound by the HLH mutant suggesting that the pointed domain is responsible for the diminished binding affinity of the fusion protein at R1 peaks (Figure 4E). Taken together these data confirm that ETV6-RUNX1 is able to bind to the same target sites as RUNX1 through the Runt domain, although its distribution across those peaks is altered by the HLH domain.

As a first hit ETV6-RUNX1 is expressed together with wild-type RUNX1 and the two proteins must compete for binding and regulation of target genes. In the NALM-6 model genes associated with sites bound by both RUNX1 and the fusion (R1_ER) are enriched in GO-terms relating to the cell cycle suggesting that the cell cycle phenotype we observe in our pre-leukaemia model results from disruption of the RUNX1-regulated transcriptome (Figure 4F).

Defining a core RUNX1 program in B-ALL reveals antagonism between ETV6-RUNX1 and native RUNX1 in cell cycle regulation

Since RUNX1 has been shown to regulate cell cycle progression, we hypothesized that the cell cycle phenotype induced by ETV6-RUNX1 as a first-hit is largely due to dysregulation of the

RUNX1 program (Friedman, 2009). We adopted an integrative approach to define a core RUNX1-driven transcriptional network in this disease and extended our analysis from NALM-6 to include RUNX1 ChIP-seq from two additional cell lines representing major B-ALL subtypes. RUNX1 binding was highly correlated across the three cell lines with 10178 common sites identified, enriched for RUNX and ETS motifs, and mapping to 6843 genes (Figure S5A and S5B). GO term analysis of shared targets confirmed "cell cycle" as the most enriched biological process, suggesting it is a core RUNX1 function across B-ALL subtypes (Figure S5C).

While childhood and adult B-ALL share some molecular features, they are distinct in their development, prognosis and outcome. To assess the extent to which RUNX1 regulome is conserved across B-ALL subtypes associated with distinct age groups, we examined the transcriptional program engaged upon RUNX1 inactivation by depleting RUNX1 in five cell lines, representing the most common childhood and adult disease subtypes, namely ETV6-RUNX1 (Reh), MLL-AF4 (RS4;11), E2A-PBX1 (RCH-ACV), ETV6-PDGFRB (NALM-6) and BCR-ABL (TOM-1) (Table S4). To minimize off-target effects and identify high confidence targets we used multiple shRNAs against RUNX1 and performed RNA-seq on biological triplicates for each condition (Figure 5A and S5D, red box). DEGs were defined for individual cell lines (see the methods) and overlapped to reveal a core set of 321 common DEGs. Of these 247 were up- and 74 downregulated, enriched in GO terms relating to apoptosis and cell cycle/DNA replication respectively (Figures 5B, 5C and S5E). We next ranked genes according to their response to RUNX1 knockdown across all five cell lines. The most highly differential genes had a significantly greater association with RUNX1 binding, regardless of expression directionality, indicating that RUNX1 can act as both repressor and activator (Figure S5F). GSEA revealed p53 among the most significantly upregulated pathways, while E2F and MYC targets were significantly enriched in downregulated genes (Figure 5D). This is consistent with wild-type RUNX1 positively regulating cell cycle progression, in direct contrast to ETV6-RUNX1 (Figures S5G, 2 and 3).

To better understand the apparent competition between RUNX1 and the fusion in cell cycle regulation, we made use of a meta-analysis that categorizes genes according to their involvement in the cell cycle (Fisher et al., 2017). As the number of studies in which a gene is implicated in the cell cycle increases so does the likelihood that it is associated with a RUNX1 binding event (Figure S5H). Correlated to this is a decrease in gene expression upon RUNX1 knockdown showing that cell cycle genes are enriched for direct targets of, and positively regulated by, RUNX1 (Figure S5H). In contrast, ETV6-RUNX1 negatively regulates the same

target genes, clearly demonstrating its opposition to RUNX1 with respect to cell cycle regulation (Figure S5H).

We next analyzed targets of the DREAM (Dimerization partner, RB-like, E2F and Multi-vulval class B) complex, which mediates gene repression during quiescence; RB/E2F, which regulate the G1/S transition and DNA synthesis; and FOXM1, implicated in gene regulation throughout the cell cycle, particularly in G2/M phases. All three sets of genes were significantly overrepresented among ETV6-RUNX1-repressed genes while DREAM and RB/E2F were significantly enriched in RUNX1-activated genes (Figure 5E). The strong association with DREAM and RB/E2F implies opposing roles in the regulation of quiescence and in G1/S progression and is in agreement with the GSEA results which highlighted "E2F targets" as counter-regulated by RUNX1 and ETV6-RUNX1. Furthermore, these observations are consistent with our first-hit iPSC model, where the most marked impact of ETV6-RUNX1 is observed in S-phase (Figure 3).

Overall our data are consistent with ETV6-RUNX1 opposing RUNX1, resulting in altered cell cycle. We note however, that RUNX1 knockdown also leads to upregulation of genes associated with p53 activation and apoptosis, suggesting a dependence on RUNX1 not only for proliferation, but for survival of B-ALL cells.

B-ALL cell lines and primary patient cells are dependent on RUNX1 activity for survival *in vitro* and *in vivo*

To functionally assess the requirement for RUNX1 in leukemia, we depleted RUNX1 *in vitro* in four B-ALL cell lines and a control CML cell line and monitored proliferation of leukemic cells over nine days (Figures S5D, red box, and 6A, Table S4). B-ALL cells transduced with any of the RUNX1 targeting shRNAs showed greatly reduced proliferative capacity compared to scrambled control, while the CML cell line K562 was unaffected, consistent with previous reports (Goyama et al., 2013). While differences in sensitivity between shRNAs reflected the degree of RUNX1 knockdown, we noted that even small perturbations of RUNX1 levels caused proliferation arrest (Figures S5D, 5'UTR shRNA and 6A).

Concurrently, we subjected RUNX1 depleted cells to flow cytometry analysis to assess cell cycle distribution and apoptosis in these cells. RUNX1 depleted cells showed an increase in G1 and S-phases and an increase in AnnexinV^{+}Hoechst33258^{+} population, demonstrating that

diminished RUNX1 activity causes a delay in cell cycle progression, as well as increase in apoptosis (Figures 6B, 6C and S6A).

To assess the impact of RUNX1 depletion on the engraftment and proliferative capacity of B-ALL cells in vivo we used bioluminescence imaging to follow leukaemic propagation of RUNX1 shRNA or control scr-shRNA-transduced cells in a competitive setting (Figure 6D). NOD SCID gamma (NSG) mice injected intra-tibia with NALM-6 cells (see methods) showed initial engraftment of all samples at the injection site but while control shRNA-transduced cells were observed at secondary sites, such as the spleen and brain, RUNX1-depleted cells failed to disseminate (Figures 6E and S6C). In addition, splenic size was visibly decreased in RUNX1-depleted mice, consistent with a diminished leukemic burden in those animals (Figure S6D).

To determine if primary, patient-derived B-ALL cells are similarly dependent on RUNX1 in vivo, we transduced two clinically aggressive B-ALL patient samples with the two most effective RUNX1 shRNAs and engrafted them into NSG mice (Table S5). FACS analysis of total bone marrow at the end point (3 months) showed that shRNA-RUNX1 cells are outcompeted by the non-transduced control cells and were barely detectable by FACS (Figure 6F and S6E), while in animals engrafted with scrambled shRNA cells both GFP^+ and GFP^- populations were present.

In sum, our results clearly demonstrate dependence on RUNX1 for survival across B-ALL, offering a promising therapeutic target.

An allosteric CBFβ inhibitor mimics RUNX1 depletion phenotype and offers a targeted treatment for B-ALL

Our results thus far confirm a critical role for RUNX1 in B-ALL, with leukemic cells being strikingly sensitive to even small perturbations of RUNX1 levels. This observation led us to explore if RUNX1 can serve as a novel therapeutic target for treatment of B-ALL.

RUNX1 is part of the CBF complex and its binding to DNA is stabilized by the CBFβ subunit. We therefore assessed if disruption of their interaction would phenocopy the effects of loss of RUNX1. We first depleted CBFβ using an shRNA (Figure S5D, blue line) and compared transcriptional changes to RUNX1 knockdown across the five B-ALL cell lines described above (Table S4). For significantly dysregulated genes the direction of change was highly correlated, demonstrating that depletion of CBFβ results in reduced RUNX1 transcriptional activity

(Figure S7A). As observed for RUNX1, CBFβ knockdown resulted in reduced growth in vitro of the four B-ALL cell lines tested, while t(9;22) K562 cells were unaffected (Figure 7A). Strikingly, CBFβ-depleted NALM-6 cells and two primary B-ALL patient samples failed to engraft and proliferate in vivo (Figures 7B, 7C, S7B and S7C). This genetic proof-of-principle suggested that pharmacological disruption of the CBF complex would be sufficient to induce cell cycle arrest and apoptosis in B-ALL. We therefore tested the effect of AI-14-91, an allosteric inhibitor of the CBFβ-RUNX1 interaction (Illendula et al., 2015, Oo et al., 2018). ChIP-seq analysis of Reh cells treated with AI-14-91 showed reduction of global binding patterns of RUNX1 (Figure 7D). Notably, ETV6-RUNX1 binding was similarly reduced showing that interaction with DNA remains CBFβ-dependent in the fusion. Transcriptional changes observed following treatment with AI-14-91 were highly correlated to those observed upon RUNX1 knockdown (Figures S7D and S7E), with a similar impact on transcriptional pathways relating to Myc and p53, demonstrating that treatment with the compound results in inhibiting RUNX1-dependent transcription (Figure 7E). In addition, the LRC signature used previously showed a significant enrichment in the CBFi upregulated genes (Figure 7F).

To test the functional impact of the inhibitor we treated four B-ALL cell lines with increasing concentrations. Strikingly, all B-ALL cells exhibited a dose-dependent reduction in viability and colony formation, corroborating our genetic data and consistent with a pan-B-ALL dependency on CBF activity (Figures S7F and S7G). Furthermore, CyTOF analysis revealed an increase in cell death, accompanied by increased expression of apoptotic markers cPARP and Caspase3, and an increase in G0 and/or G1 cells consistent with CBFi retarding G1/S transition (Figure 7G).

Finally, to directly assess the contribution of CBF complex to cell cycle progression we arrested cells in G1 or G2 – through pharmacological inhibition of CDK4/6 or CDK1 respectively – and monitored re-entry to the cell cycle in the presence or absence of AI-14-91. We observed delayed progression from G1 through S phase (Figure S7H) but no impact on cell cycle progression from G2 (data not shown), demonstrating a role for CBF-regulated transcription in G1/S phases of the cell cycle consistent with reduced expression of E2F targets.

Taken together, our data demonstrate that allosteric inhibition of the CBF complex phenocopies genetic depletion of either of the two subunits, thereby opening a novel therapeutic avenue in B-ALL.

A balance between CBF complex and ETV6-RUNX1 acting on the P53 regulated cell cycle-apoptosis axis promotes a silent pre-leukemic state

Our data demonstrate that in a pre-leukemic setting, ETV6-RUNX1 expression inhibits cell cycle progression, primarily by downregulating E2F and Myc signatures, to achieve a dormant-like preleukemic state (Figures 2C and 2D). Conversely, genetic and chemical inactivation of RUNX1, or its partner CBFβ, leads to cell cycle arrest, supporting a clear transcriptional antagonism of ETV6-RUNX1 and native RUNX1 relating to cell cycle. Yet, while inactivation of the CBF complex leads to cell death of B-ALL cells *in vitro* and *in vivo*, ETV6-RUNX1 does not induce apoptosis in a preleukemic setting, reflected in the lack of p53-signature enrichment in our GSEA analysis. To address this disparity, we made use of a publicly available p53 meta-analysis which ranks the p53 regulome broadly into two classes: I) p53 downregulated genes, associated with cell cycle regulation; and II) p53 upregulated genes, enriched for direct transcriptional targets of p53, associated with DNA damage and induction of apoptosis (Fisher et al., 2017). We compared the transcriptomes of our two ETV6-RUNX1 models as well as the gene expression profiles of RUNX1-depleted and CBFi-treated cells (together referred to as CBF complex) to the p53 regulome. All datasets displayed high correlation, with a definite inverse correlation between ETV6-RUNX1 and CBF complex activities (Figure 8A). However, while the p53 downregulated cell cycle class was overrepresented in both CBF-activated and ETV6-RUNX1-repressed gene sets, the apoptotic class was found to be significantly enriched only in genes up-regulated upon CBF inhibition (Figures 8B and S8A).

These data show that a fine balance exists between ETV6-RUNX1 and the wild-type CBF complex, allowing the fusion to alter cell cycle without inducing apoptosis, establishing a clinically covert preleukemic state.

Discussion

Our systems level analysis shows that the role of ETV6-RUNX1 in initiation of leukaemia is primarily mediated through modulation of the native RUNX1 program. This is achieved through competition for canonical RUNX1 binding sites with the dominant effect being transcriptional repression of target genes. The competition between the oncofusion and its wild-type counterpart led us to explore the role of the CBF complex in leukemia, establishing that its function is required for the proliferation and survival of B-ALL cells. This observation extends to all genetic subtypes tested, including representatives of both childhood and adult B-ALL, where a small molecule inhibitor of the CBF complex was effective in establishing

proof-of-principle for targeted therapy. Moreover, since RUNX1 is consistently retained in B-ALL, targeting CBF may provide a means to overcome intratumor heterogeneity and therapy resistance in subclones (Dobson et al., 2020).

ETV6-RUNX1 is the most common genetic aberration in BCP-ALL and as such has been a subject of extensive research. However, two key elements of ETV6-RUNX1 function remain unaddressed: 1. How ETV6-RUNX1 and the native CBF complex interact; and 2. The role of the fusion as a first-hit in the establishment of a pre-leukemic state. Comparison of ChIP-seq data revealed that ETV6-RUNX1 retains canonical DNA binding properties of the RUNX family, binding the same target sites through the Runt domain. A minority of binding sites, tending to be more distant from the TSS, were bound with greater affinity by the fusion and these will require further analysis to understand their role but the main function of the fusion is to compete with RUNX1 for targets and modulate transcription. Although counter-regulation was observed for both activated and repressed targets of the CBF complex direct targets were particularly enriched among the CBF-activated genes suggesting that the major contribution of ETV6-RUNX1 is repression of RUNX1 targets.

Since ETV6-RUNX1 arises *in utero* as an initiating lesion, it is important to understand its function as a “first-hit” prior to acquisition of additional mutations. We previously developed an iPSC model that recapitulates fetal lymphopoiesis and shows that ETV6-RUNX1 imposes a partial block on B-cell differentiation (Boiers et al., 2018). Here, we show that ETV6-RUNX1, through repression of RUNX1 targets, downregulates transcriptional pathways associated with cell cycle progression. RUNX1 has previously been described to be a cell cycle regulator and analogous findings in AML show that the onco-fusion RUNX1-ETO compromises cell cycle and high levels of RUNX1-ETO induce apoptosis (Friedman, 2009, Martinez-Soria et al., 2019, Ptasinska et al., 2014, Mandoli et al., 2016).

It is perhaps counterintuitive that an oncogene should oppose cell cycle. However, ETV6-RUNX1 alone is not sufficient to establish leukemia. Rather, by impeding differentiation and cell cycle progression it may generate a pool of progenitors arrested in their development which may be vulnerable to second hit mutations. Common second hits include loss of B-cell transcription factors such as PAX5 and EBF1, which compounds the differentiation arrest, and loss of CDKN2A – encoding p16 and p14 from alternative reading frames – which promotes cell cycle due to loss of a CDK inhibitor and survival through MDM2-mediated suppression of p53, respectively (Heltemes-Harris et al., 2011, Sulong et al., 2009). The negative effects of ETV6-RUNX1 on cell cycle may create positive selection for loss of p16.

The t12;21 translocation that creates *ETV6-RUNX1* fuses one *RUNX1* allele but spares the other. Given that the fusion functions by opposing the native CBF complex, it might be expected that loss of function mutations in the second RUNX1 allele would be observed in B-ALL. However, the opposite appears to be the case with retention or amplification of the wild-type allele. We show here that loss of RUNX1 in B-ALL cells results in cell cycle exit and apoptosis showing that B-ALL cells are RUNX1 addicted, consistent with the requirement for Runx1 for survival of immature B-cells in mice (Niebuhr et al., 2013). Runx1 heterozygosity is well tolerated in the B-cell lineage although compound haploinsufficiency with EBF1 reveals a role in induction of the B cell transcriptional program (Lukin et al., 2011). Our results suggest that loss of the second allele of *RUNX1* would not be tolerated as indeed was the case in a mouse model where ETV6-RUNX1 expression, together with homozygous loss of RUNX1, resulted in severe anemia and death of the animals (Schindler et al., 2009). We suggest therefore that ETV6-RUNX1 is only tolerated in pre-leukaemia in the continued presence of wild-type RUNX1 and may even create selective pressure for additional copies. While ETV6-RUNX1 recapitulates the cell cycle effects of RUNX1 knock-down, it does not lead to cell death as inhibition of CBF does, allowing it to be tolerated as a first hit. This is explained by the fact that CBF inhibition results in activation of the p53 apoptotic pathway but expression of ETV6-RUNX1 does not. A link between ETV6-RUNX1 and p53 has previously been suggested, with forced expression of the fusion inducing MDM2, which represses p53 activity (Kaindl et al., 2014). Although we observed ETV6-RUNX1 binding at the *MDM2* promoter we did not observe any change in its expression.

p53 is rarely lost in B-ALL suggesting that CBF inhibition may provide a route to tumor cell killing without the collateral damage associated with systemic p53 activation. Moreover, if RUNX1 is selectively retained, targeting the CBF complex may provide a means to overcome chemo-therapy resistance associated with subclonal heterogeneity (Dobson et al., 2020). However, current CBF inhibitors are not sufficiently potent to be used in the clinic. Future efforts should see more potent compounds developed and will explore combination therapies as a possible route to less toxic treatments.

Acknowledgements

This work was generously funded by grants from the UK: Blood Cancer UK (Grant #16001), Children with Cancer UK (Grant#14-173) and the Medical Research Council UK (Grant#MR/N000838/1). We thank Cure Cancer@UCL and Mrs Sandra Hamilton for the generous support. We thank John Brown for input on the RNA-sequencing experiments, Mathew Robson for help with in vivo animal imaging, Chela James and Javier Herrero for input

on bioinformatic analysis and George Morrow from the UCL Cancer Institute Flow Cytometry Facility for help with Mass Cytometry. We further thank Gill May for critical input on the manuscript.

Author Contributions

J.P.W., E.M.D., R.N. and T.E. conceived the experiments. J.P.W., E.M.D., C.B., J.B.C., S.G., Y.G., M.R. performed the experiments. J.P.W. performed the bioinformatics analysis. S.E.R. made the ETV6-RUNX1 iPS lines. A.I. and J.H.B. provided the CBFβ inhibitor. J.H.A.M. and H.G.S. performed the ChIP-seq of NALM-6 mutant lines. J.P.W., E.M.D. and T.E. wrote the manuscript.

Declaration of Interests

The authors declare no competing interests

Figure Legends

Figure 1. Delineating the ETV6-RUNX1 regulome in childhood B-ALL

(A) Integrative Genomics Viewer (IGV) screenshot of ETV6-RUNX1 binding at the IER2 locus in an ETV6-RUNX1$^+$ cell line (Reh, upper panel, red dashed line) or patient-derived ETV6-RUNX1$^+$ leukaemia samples (bottom panel, blue dashed line). HPA and PAS, two independent ETV6 antibodies; DNaseI, DNaseI hypersensitivity sequencing; control, signal following immunoprecipitation with an IgG antibody.

(B) MEME enrichment of motifs in sites bound by ETV6-RUNX1. See also Table S1 (enrichment motifs).

(C) Correlation matrix for normalized signal across all ETV6-RUNX1 binding sites.

(D) Distribution of peaks relative to their nearest genes (see methods for analysis).

(E) ETV6-RUNX1 binding across genes binned according to relative expression in ETV6-RUNX1$^+$ B-ALL as compared to all other subtypes. **** $p<0.0001$, hypergeometric test.

Figure S1; related to Figure 1

(A) Dotplot showing normalised read counts (counts per million) for a 500bp window centred on the summit for all ETV6-RUNX1 peaks identified for two independent ETV6 antibodies (HPA and PAS). Colours represent peaks called in one (1, red dot) or both (2, green dot) ChIP-seq datasets.

(B) Heatmap showing ChIP-seq and DNaseI-hypersensitivity read density in a 5kb window centred on peak summits for all ETV6-RUNX1 peaks identified.

(C) Boxplot showing distribution of normalised read counts for each of 8 ChIP-seq samples across all ETV6-RUNX1 peaks identified. Colour represents the number of ChIP datasets a peak was identified in.
(D) Motif counts in peaks binned according to the number of ETV6-RUNX1 ChIP datasets in which they were identified.
(E) Density plot of RUNX and ETS motif distribution across a 500bp window centred on ETV6-RUNX1 ChIP peak summits.

Figure 2. ETV6-RUNX1 induces transcriptional changes indicative of cell cycle repression in pre-leukemia
(A) Bar plot showing percentage of genes associated with a RUNX1 binding event for genes significantly up- or down (dn)-regulated or not significant (ns, $p>0.05$) upon ETV6-RUNX1 knock down in Reh or ETV6-RUNX1 expression in proB cells derived from IPSCs. **** $p<0.0001$, *** $p<0.001$, ** $p<0.01$, * $p<0.05$.
(B) Dotplot with density contours showing log fold change for ETV6-RUNX1 vs wild-type iPSC-derived proB cells and ETV6-RUNX1 knockdown in Reh. Genes displayed have padj<0.1 in both datasets.
(C) GSEA for ranked gene lists from ETV6-RUNX1 vs wild-type iPSC-derived proB cells and ETV6-RUNX1 knockdown in Reh.
(D) Principal component analysis using label-retaining cell (LRC) signature genes for wild-type and ETV6-RUNX1 expressing iPSC-derived cell populations.
(E) Plot showing normalized gene expression (mean +/- 95% confidence intervals) in wild-type and ETV6-RUNX1 expressing iPSC-derived cell populations for label-retaining cell leading edge genes derived from GSEA (see Figure S2G).

Figure S2; related to Figure 2.
(A) Western blot and quantitative PCR showing ETV6-RUNX1 expression in Dox-induced or control Reh cells for vector control or ETV6-RUNX1 shRNAs (251/818) after 72h (western) or at the times indicated.
(B) Schematic of hPSC differentiation protocol
(C) Principal component analysis of RNA-seq data from wild-type and ETV6-RUNX1 expressing iPSC-derived cell populations
(D) Barplot showing normalized enrichment scores (NES) from GSEA for the indicated genesets against pre-ranked gene lists for ETV6-RUNX1 vs wild-type for each of the four populations indicated. * $p<0.05$

(E) GSEA for label retaining cell (LRC) signature genes against pre-ranked gene list for ETV6-RUNX1 vs wild-type in iPSC-derived proB cells

Figure 3. Mass cytometry reveals an accumulation of *ETV6-RUNX1* expressing progenitor cells in S phase

(A) Dotplots of UMAP dimensionality reduction for cell surface markers CD45, CD45RA, CD34, and CD19. Colours represent manually annotated SOM clusters. 10000 cells are shown in each plot. See also Figure S3A.

(B) Heatmap showing median scaled expression of cell surface markers in the annotated clusters represented in (A).

(C) Barplot showing proportion of cells in each population annotated in (A)

(D) UMAP dimensionality reduction for cell cycle markers pRb, IdU, CycB1 and pHisH3. Colours represent manually annotated SOM clusters. See also figures S3B and C.

(E) Barplot showing proportion of cells in each phase of the cell cycle for each population annotated in (A).

Figure S3; related to Figure 3.

(A) UMAP dimensionality reduction for cell surface markers CD45, CD45RA, CD34, and CD19. Colour represents scaled expression of the indicated markers.

(B) UMAP dimensionality reduction for cell surface markers pRb, IdU, CycB1 and pHisH3. Colour represents scaled expression of the indicated marker.

(C) Heatmap showing median scaled expression of cell cycle markers in the annotated clusters represented in figure 3D.

Figure 4. ETV6-RUNX1 binds DNA through the Runt domain, competing with native RUNX1

(A) Integrative Genomics Viewer (IGV) screenshots of RUNX1 and ETV6-RUNX1 binding and DNase1 hypersensitivity (DNaseHS) in NALM6 at the *RUNX1* and *SPI1* loci. R139G: loss-of-function point mutation in the Runt domain of ETV6-RUNX1. ΔHLH: deletion of helix-loop-helix pointed domain of ETV6-RUNX1. CTL: V5 ChIP in NALM6 lacking V5-tagged protein.

(B) Heatmaps showing ChIP-seq signal across all identified RUNX1/ETV6-RUNX1 binding sites in a 3kb window centred on peak summits.

(C) MEME enrichment of motifs in sites bound by RUNX1 and/or ETV6-RUNX1.

(D) Dotplot of CPM for RUNX1 vs ETV6-RUNX1 ChIP. Colours show peaks classified as more highly bound by RUNX1 (R1), ETV6-RUNX1 (ER) or similarly bound by both (R1_ER).

(E) Boxplot of CPM for peaks classified in (D) for RUNX1 or ETV6-RUNX1 ChIP datasets.

(F) Bar plot showing enrichment of GO-terms relating to cell cycle in genes mapped to R1_ER peaks as classified in (D).

Figure S4; related to Figure 4.

(A) Schematic of ETV6-RUNX1 and R139G/ΔHLH mutants.

(B) Correlation matrix for ChIP datasets based on normalised counts across all RUNX1/ETV6-RUNX1 peaks identified.

(C) Dotplots of CPM for RUNX1 vs each ETV6-RUNX1 dataset.

(D) Average profile plot of ChIP-seq signal across all RUNX1/ETV6-RUNX1 peaks identified in a 3kb window centred on peak summits.

(E) Proportion of peaks in the indicated bins relative to TSS for each of the peak sets classified in figure 4D.

(F) Proportion of peaks with >= 1 RUNX1 motif mapped to a 500bp window centred on the peak summits for peaks binned according to classifications in figure 4D (left panel) or distance from the TSS (right panel).

Figure 5. Defining a core RUNX1 program in B-ALL reveals antagonism between ETV6-RUNX1 and native RUNX1 in cell cycle regulation

(A) Schematic showing shRNAs targeting RUNX1 and/or ETV6-RUNX1 and western blot showing RUNX1 knock down following transfection with the indicated shRNAs.

(B) Heatmap showing relative expression of the 25 most significantly up- or down-regulated genes following RUNX1 knock-down in five B-ALL cell lines.

(C) Bar plots showing significantly enriched GO-terms in the "core" up- or down-regulated genes.

(D) Plots of GSEA results for the indicated gene sets against a list of genes ranked from the most significantly up- to the most significantly down-regulated following RUNX1 knock-down.

(E) Bar plots showing the percentage of genes significantly up- (UP) or down- (DN) regulated in ETV6-RUNX1 knock-in iPSC-derived proB cells, or following ETV6-RUNX1 or RUNX1 knock-down which were defined as DREAM, Rb/E2F, or FOXM1 targets. NS = not significant

**** $p<0.0001$, *** $p<0.001$, ** $p<0.01$, * $p<0.05$, hypergeometric test.

Figure S5; related to Figure 5.

(A) Heatmaps showing RUNX1 ChIP-seq signal across all identified RUNX1 binding sites for the indicated cell lines in a 3kb window centred on peak summits.

(B) Venn diagram showing overlap of RUNX1 ChIP peaks from Reh, RCH-ACV and Nalm6.

(C) Bar plot showing significantly enriched biological processes from GO term analysis (selected biological processes shown).
(D) Western blots of Reh, NALM-6 and RCH-ACV cells transduced with control (empty and scrambled), RUNX1 or CBFβ shRNAs. Protein levels are measured 48h after transduction.
(E) Dot plot of principal components 1 and 2 showing RUNX1 shRNA and control samples for 5 cell lines. Ellipses show 80% confidence intervals for the indicated groupings.
(F) Bar plot showing percentage of genes bound by RUNX1 for genes binned according to response to RUNX1 knock-down. **** $p<0.0001$, hypergeometric test.
(G) Bar plot showing normalized enrichment scores (NES) for GSEA analysis of the indicated gene sets against ranked lists from ETV6-RUNX1 knock-in iPSC-derived proB cells, ETV6-RUNX1 knock-down and RUNX1 knock-down. * padj < 0.05.
(H) Comparison of RUNX1 and ETV6-RUNX1 regulomes to cell cycle meta-analysis. Genes were binned according to the number of data sets in which they were defined as cell cycle genes. Bar plot shows the percentage of genes associated with a RUNX1 binding event while boxplots show log fold change following RUNX1 or ETV6-RUNX1 knock-down or in ETV6-RUNX1 knock-in iPSC-derived proB cells. **** $p<0.0001$, *** $p<0.001$, ** $p<0.01$, * $p<0.05$, ns = not significant, t-test vs Cell cycle datasets = 0.

Figure 6. B-ALL cell lines and primary patient cells are dependent on RUNX1 activity for survival *in vitro* and *in vivo*
(A) Growth curves for the indicated cell lines over 9 days following transduction with shRNAs targeting RUNX1 (Figure 5A) or a scrambled control.
(B) Bar plot showing cell cycle distribution of t(12;21) and t(1;19) cell lines transduced with scrambled or RUNX1-shRNAs, 5 days after transduction. G1, S and G2 populations are shown. (n=3).
(C) Apoptosis analysis of cell lines transduced with scrambled or RUNX1-shRNAs, 5 days after transduction. % of $GFP^+AnnexinV^+Hoechst33258^+$ is shown; (n=3).
(D) Schematic of competitive engraftment experiments. NALM-6 cells transduced with shRNAs (GFP^+) and stably expressing a luciferase/RFP reporter were mixed 1:1 with non-transduced (GFP^-) cells and injected into NSG mice. Mice were imaged weekly for luciferase activity and culled for end-point analysis (4 weeks) of leukaemic engraftment.
(E) Bioluminescence imaging at 2 and 4 weeks (end point) after injection of samples transduced with shRNAs targeting RUNX1 or control shRNA (see Figure 6C).
(F) Plots showing end-point analysis of leukaemic engraftment of two patient samples transduced with scrambled or RUNX1-shRNAs. Human ($CD19^+$), gated on human $CD45^+$ cells,

were assessed for the proportion of shRNA-transduced (GFP^+) vs wild-type GFP^- cells (Figure S7C). **** p<0.0001, unpaired t-test.

Figure S6; related to Figure 6.

(A) Representative FACS plots of apoptosis gating in B-ALL cell lines

(B) Summary of bioluminescence imaging of NSG mice engrafted with a 1:1 mix of NALM-6 cells transduced with shRNAs (GFP^+) and stably expressing a luciferase/RFP reporter were mixed 1:1 with non-transduced (GFP^-) over four weeks. Relative bioluminescence signal is estimated as the sum of measured signal of ventral and dorsal positions at each time point (n=3). ****, p<0.0001, 2-way Anova test.

(C) Representative FACS plot of end point (4 weeks) analysis of NSG mice engrafted with NALM-6 cells as in Figure 6C.

(D) Representative image of splenic size obtained from NSG mice engrafted with a 1:1 mix of NALM-6 cells transduced with shRNAs (GFP^+) and stably expressing a luciferase/RFP reporter were mixed 1:1 with non-transduced (GFP^-) at week four (end point).

(E) Representative FACS plot of endpoint (3 months) analysis of NSG mice competitively engrafted with primary B-ALL cells. Human ($CD19^+$) engraftment and the proportion of shRNA-transduced (GFP^+) cells for two ALL patient samples with the indicated genotypes.

Figure 7. An allosteric CBFβ inhibitor mimics RUNX1 depletion phenotype and offers a targeted treatment for B-ALL

(A) Growth curves for the indicated cell lines over 9 days following transduction with an shRNA targeting CBFβ (n=3) or a scrambled control (plotted values from).

(B) In vivo bioluminescence imaging of NOD/SCID mice competitively engrafted with $shGFP^+Luc^+$ and wild-type GFP^-Luc^- NALM-6 cells (see methods and Figure 6C). Ventral and dorsal views of mice at weeks 2 and 4 (end point); control (scrambled shRNA); CBFβ - CBFβ-targeting shRNA.

(C) Summary endpoint FACS analysis of leukaemic engraftment of Pt#1 and Pt#2 cells transduced with a scrambled or CBFβ shRNA. Human ($CD19^+$), gated on human $CD45^+$ cells were assessed for the proportion of shRNA-transduced (GFP^+) vs wild-type GFP^- cells (Figure S7C). **** p<0.0001, unpaired t-test.

(D) Heatmap (left) and average profiles (right) of ChIP-seq signal across all RUNX1 and ETV6-RUNX1 binding sites in a 3kb window centred on peak summits from control (DMSO) or CBFi (AI-14-91) treated Reh (ETV6-$RUNX1^+$ ALL) following immunoprecipitation with RUNX1, ETV6 or control (IgG) antibodies.

(E) GSEA plots for p53 and Myc against genes ranked from most significantly up- to most significantly down-regulated following CBFβ knockdown.
(F) GSEA plot of LRC signature genes ranked from most significantly up- to most significantly down-regulated following CBFβ knockdown
(G) Mass cytometry analysis of three B-ALL cell lines Reh, t(12;21), RCH-ACV t(1;19) and RS4;11 t(4;11) treated with 15uM AI-14-91 for 48h. See the methods for details.

Figure S7; related to Figure 7.
(A) Dotplot comparing log fold change following RUNX1 knockdown (x-) and CBFβ knockdown (y-axis) in ALL cell lines. "Core" RUNX1 target genes defined in Figure 5 are shown.
(B) Representative FACS plot of endpoint (4 weeks) analysis of NSG mice engrafted with NALM-6 cells as in Figure 6C.
(C) Representative FACS plot of endpoint (3 months) analysis of NSG mice competitively engrafted with primary B-ALL cells. Human ($CD19^+$) engraftment and the proportion of shRNA-transduced (GFP^+) cells for two ALL patient samples with the indicated genotypes.
(D) Dotplot (left) comparing log fold change following RUNX1 knockdown (x-) and CBFi treatment (y-axis) in ALL cell lines. "Core" RUNX1 target genes defined in Figure 5 are shown.
(E) Heatmap of all DEGs after RUNX1 or CBFβ knockdown and treatment with CBFi.
(F) Colony formation assay of B-ALL cell lines treated with increasing concentrations of AI-14-91 (n=2)
(G) Celltox green assay B-ALL cell lines treated with increasing concentrations of AI-14-91. Each dataset is normalised to the respective DMSO control. (n=3).
(H) Histograms of DNA content (Hoechst 33342, left panel) showing the impact of AI-14-91 (CBFi) on cell cycle re-entry following palbociclib (Pal) wash-out. Following wash-out cells were incubated for a further 18h in the presence of pal, CBFi or DMSO. Quantification (right panel) of cell cycle distribution base on DNA content (G1=2N, 2N<S<4N, G2=4N). Error bars show SD (n=3). Note that Nalm6 transit through the cell cycle more rapidly, hence a high proportion of control (DMSO) cells have re-entered G1.

Figure 8. A balance between CBF complex and ETV6-RUNX1 on the P53 regulated cell cycle-apoptosis axis promotes a silent pre-leukemic state
(A) Plot of mean log fold change for genes binned according to their p53 score for ETV6-RUNX1-expressing ProB cells, ETV6-RUNX1 knock-down, RUNX1 knock-down, and CBFi treatment. Negative p53 scores are associated with cell cycle genes while positive scores are associated with direct p53 targets involved in apoptosis and DNA damage (DD) response (Fischer et al, 2016). See figure S8A.

(B) Boxplot of log fold change for genes grouped into cell cycle (p53 score <= -17), apoptosis/DNA damage (DD) (p53 score >= 17) or control (-17 < p53 score < 17) for ETV6-RUNX1-expressing ProB cells, ETV6-RUNX1 knock-down, RUNX1 knock-down, and CBFi treatment. ns = not significant, **** p<0.0001, ***p<0.001, ** p<0.01, * p<0.05, hypergeometric test.

Figure S8; related to Figure 8.

(A) Bar plots showing proportion of direct p53 targets (Fisher et al., 2017) in significantly up (UP) or down-regulated (DN) genes for each of the indicated datasets. NS = snot significant.

Methods

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Tariq Enver (t.enver@ucl.ac.uk).

Materials Availability

Plasmids used in this study are available upon request from the Lead Contact.

Data and Code Availability

Data relating to IPSC ETV6-RUNX1 knock-in model: ArrayExpress, E-MTAB-6382.

St Judes pediatric B-ALL: PanALL, SJC-DS-1009

Cell lines

For full details (age and cytogenetics) please refer to Table S4. Reh, NALM-6, RCH-ACV, RS4;11 and TOM-1 were purchased from DSMZ. K562 were a kind gift from Prof Asim Khwaja, UCL Cancer Institute, Department of Hematology, London, UK. 293T cells were purchased from ATCC. All leukemic cells were grown in RPMI 10% FBS at 37°C and 5% CO_2. 293T cells were grown in DMEM 10% FBS at 37°C and 5% CO_2. iPSC cells were maintained and differentiated as in (Boiers et al., 2018).

Primary patient material

Diagnostic childhood B-ALL bone marrow (BM) samples were obtained upon informed consent and approval of relevant research ethics from patients at the Great Ormond Street Hospital (GOSH). Mononuclear cells were isolated by Ficoll gradient centrifugation and frozen at -80° until further use. For full details (age and cytogenetics) please refer to Table S5. Prior

to experiment samples were thawed at 37° and transduced with the respective lentiviral constructs (MOI of 10), incubated for 48h and flow cytometry sorted prior to injection.

Bone marrow reconstitution assay

All *in vivo* experiments were performed in strict accordance with the United Kingdom Home Office regulations. Animals were housed in individually ventilated cages (IVC) under specific pathogen free (SPF) conditions. Primary childhood B-ALL cells were transplanted into 8-12 weeks old sub-lethally irradiated NOD/SCID IL2Rγ^{null} (NSG) mice (males and females, obtained from Dominique Bonnet's Lab, The Francis Crick Institute) by an intravenous (IV) injection. Sub-lethal irradiation was achieved with a single dose of 375 cGy. Prior to the procedure, mice were administered acid water for a week and Baytril (resuspended at 25.5 mg/kg in the drinking water) for the 2 weeks following it. Each mouse received a total of $1x10^5$ leukemia cells resuspended in 100µl PBS/0.5% FBS. For FACS analysis following antibodies were used: Human CD45 (Biolegend, AF700, clone HI30, 1/100 dilution), mouse CD45 (Biolegend, PE, clone 30-F11, 1/100 dilution), CD19 (BD, APC, clone HIB19, 1/100 dilution).

In vivo bioluminescence imaging

For *in vivo* imaging studies the lentiviral TWR vector (Firefly Luciferase-Red fluorescence protein), kindly donated by Prof Dominique Bonnet, was used (Gandillet et al., 2011). NALM-6 cells were co-transduced with TWR and LLX3.7-shRNA vectors and RFP^+GFP^+ double positive cells were FACS sorted and injected intra tibia. IVIS®Lumina was used to image animals. Prior to imaging animals were anaesthetized with 3% isoflurane and D-Luciferin (Caliper Life Science) was injected intraperitoneally (IP). Anterior and posterior images were taken at different exposure times post injection to determine peak signal. Total bioluminescent signal was assessed as a sum of anterior and posterior signals.

Flow cytometry

In general, for flow cytometry analysis or sorting cells were centrifuged at 300rcf for 5min and washed in 1ml PBS/2%FBS. Staining was done in 100µl volume. A master mix with appropriate antibodies was prepared and added to each sample. Unstained and single stain controls were included for each experiment and Hoechst 33258 (SigmaAldrich) was used as a viability dye. Prior to staining cells obtained from animals were lysed with 1ml of Red Blood Cell (RBC) lysis buffer (SigmaAldrich) and incubated for 5-10 min at room temperature.

Cell Cycle Assay

Ki67 staining was used to assess cell cycle status. Cells were washed with 1ml of PBS/2%FBS, resuspended in PBS/2%FBS and fixed with freshly made 3.2%PFA (paraformaldehyde)/PBS for 10min at room temperature in the dark. After washing cells were permeabilized in 300μl cold 90%Methanol/PBS on ice for 30min and then washed twice as before. 25μl of Ki67 antibody (Biolegend, PE-350503, APC-350514). was added and samples were incubated at RT for an hour. DAPI (0.5μg/ml) staining was done for 40 min on ice. Cells were washed, resuspended in 300μl PBS/2%FBS and FACS analysed on Gallios® (Beckman Coulter) or ARIAIII (BD). Unstained, isotype and single stain controls underwent the same procedure without addition of the respective antibodies.

Apoptosis Assay

AnnexinV (BD; PE-65875X, APC-550474) and Hoechst 33258 dye were used to detect apoptotic cells. Cells were stained following manufacturer's protocol, resuspended in 200μl 1xBinding buffer (BD) containing Hoechst58 (1:10000 dilution) and analysed on the Gallios® (Beckman Coulter).

Knockdown using shRNAs

Small hairpin RNAs targeting RUNX1, CBFβ and ETV6-RUNX1 were designed using several prediction tools or The RNA Consortium (TRC) website. In general, a sequence of 21nt in length ting with G was chosen where possible. Sequences were ordered as primers from Integrated DNA Technologies (IDT) website and resuspended in DNase–free water as a 100μM stock concentration. For oligo annealing 1μl of 100μM forward and reverse primers for each shRNA construct were mixed in a PCR tube with 48μl of IDT Duplex Buffer (Integrated DNA Technologies) and cloned into the LLX3.7 lentiviral vector. For TRIPZ vector shRNA were designed as in (Fellmann et al., 2013) and cloned following manufacturer's instructions. To induce knockdown doxycycline was added at a final concentration of 0.5μg/ml and cells were collected for western blot analysis 72h later. For RNA-seq cells were collected 7 days post induction with Doxycycline.

Lentiviral production

Lentiviral particles were packaged using the second generation packaging system and for each lentiviral construct three 175cm^2 flasks of HEK293T cells were used. HEK293T cells were grown to 70-80% confluency and media was changed before adding the DNA mix. For each lentiviral construct a mix of 4.47μg lentiviral vector, 2.98μg of psPAX2 and 2.98μg of pMD2.G was made in a sterile TE Buffer. Opti-MEM® media was added to the DNA mix. FuGENE®6 was used as a transfection reagent. The DNA/media/FuGENE®6 mix was incubated at RT for 15-

30 min and added dropwise to the flask containing HEK293T cells avoiding contact with the plastic. The cells were then incubated for 48h without changing the media and at that point supernatant was collected from each flask. Further 18ml of DMEM were added to each flask and cells were incubated for another 24h. To collect viral particles, supernatants were filtered through a sterile 0.45μm surfactant-free, non-pyrogenic filter into Oak Ridge centrifuge tubes with a seal cap (Nalgene). The tubes were centrifuged at 50 000g for 3h at 4°C, supernatant was discarded and the pellet was resuspended in 100μl of IMDM (FBS-free, volume for one 175cm^2 flask), aliquoted into screw-cap tubes and kept at -80° for further use.

Lentiviral transduction

Cell lines were washed with PBS/2%FBS and resuspended in complete medium at a concentration of 1*10^6-3*10^6/ml. 1ml of cells was added into a well of a 6 well plate and virus was added dropwise (volume of virus depending on MOI). 24h later 1ml of media was added, cells were incubated for further 24h and washed twice with PBS/2%FBS prior to FACS analysis or sorting. Primary patient material was aliquoted into 24- or 48-well plates in 200μl StemSpan™ Serum-Free Expansion Medium (SFEM) (STEMCELL Technologies) supplemented with 20% FCS and the following cytokines: 50ng/ml human SCF, 20ng/ml IL-7, 10ng/ml IL-3, /ml FLT-3.

In vitro proliferation assay

B-ALL cell lines were transduced with control, RUNX1 or CBFβ shRNAs, FACS sorted 48h later and cultured in 6-well plates in 4ml of media. Cells were counted daily for a period of 9 days.

AI-14-91 treatment

AI-14-91 was provided as powder and resuspended in DMSO to a stock concentration of 40mM. Prior to experiment respective dilutions were made in media and added to the cells. 0.25% of DMSO was used as control.

CellTox Green assay

Cells were resuspended in media with CellTox™ Green (Promega) in 96-well plates (100 000 cells/well). Fluorescence intensity was measured at (Ex)485/(Em)520nm on Varioscan™ Lux microplate reader (ThermoFischer Scientific) 72h later.

Colony Formation Assay

B-ALL cells were treated with CBFi AI-14-91, washed and seeded into 1.5 mL of Methocult H4230 (Stem Cell Technologies) supplemented with 50ng/ml human SCF, 20ng/ml IL-7,

10ng/ml IL-3, /ml FLT-3 and plated into 35 mm non-coated plates (430,588, Corning Incorporated, Corning, NY, USA). Plates were incubated for 10–14 days at 37 °C, 5% CO_2. Colonies produced were counted and classified.

Palbociclib treatment

Cells were synchronised by incubation with palbociclib (500nM) for 48h. Cells were then washed with PBS and returned to culture in the presence of AI-14-91 (20uM) for 18h. Cells were incubated with Hoechst 33342 (0.5μg/ml) for 30mins and subjected to FACS analysis. For cell cycle analysis sub-G0 cells were excluded. G0/G1 (2N), G2 (4N) and S-phase (2N < S < 4N) cells were gated based on DNA content and quantified.

Western Blotting

Cells were pelleted, washed with PBS Buffer and lysed (30 min on ice with occasional vortexing) in RIPA Buffer (Sigma Aldrich) supplemented with 1x Protease Inhibitor Cocktail (Roche). Lysates were cleared by centrifugation at 4°C for 15 min at 14000g and total protein extracts were transferred to a new pre-chilled Eppendorf tube. For nuclear extracts (ETV6-RUNX1) the remaining pellet was resuspended in 50μl RIPA supplemented with 1x Protease Inhibitor Cocktail and sonicated for 3-5 min with Picoruptor® (Diagenode). 10μg of proteins were then separated through precast 4-12% NuPAGE® Bis-Tris gels (Invitrogen) and transferred onto a PVDF membrane (Millipore). Protein membrane was blocked and incubated with primary antibody O/N at 4°C. Following 1h incubation with secondary antibodies, membranes were developed using the ImageQuant LAS 4000 System (GE Healthcare). Antibodies used: RUNX1 (Abcam, ab23980), ETV6: (Sigma, HPA000264), CBFβ (Abcam, ab33516), GAPDH (14C10, Cell Signaling #2128).

Chromatin Immunoprecipitation

Cells were cross-linked with 1% formaldehyde and sonicated to yield chromatin of 100–500 bp. ChIP was performed by as previously described (May et al., 2013) using antibodies: RUNX1 (Abcam, ab23980), ETV6: (Sigma, HPA000264), V5 (Abcam, ab9116), and nonspecific rabbit IgG (Abcam, ab171870). Libraries were gel-purified, 10ng of DNA was amplified and single end sequenced at 36 bp using Illumina NextSeq500/550 High Output Kitv2.5 (75 cycles).

Mass Cytometry

For antibodies used, please refer to Table S2. iPS or B-ALL cell lines were incubated with IdU at a final concentration of 50μM for 30min at 37°. Cells were then washed with MaxPar Staining Buffer and incubated in 1ml of MaxPar PBS with Cell-ID Cisplatin (final concentration

of 1uM) for 5min at room temperature. Following a wash with 5ml of MaxPar Staining buffer, samples were barcoded following manufacturer's instructions. Up to 20 barcoded samples were then pooled and the sample was counted to determine the amount of surface antibody necessary. As a rule, 1ul of each antibody was used to stain $3x10^6$ cells. Pool was incubated with the surface antibody cocktail (see Table S1 for details of all antibodies used) for 30min at RT with vortexing in between. After washing with MaxPar Staining Buffer cells were incubated with 1x Nuclear Antigen Staining Buffer for 30min, followed by a wash with Nuclear Antigen Staining Perm. Cells were counted prior to addition of a cocktail of intracellular antibodies and incubated at RT for 45min with gentle vortexing in between. Cells were then fixed with 1.6% freshly prepared PFA for 10min at RT. After washing cells were incubated for 1h at RT or overnight at 4°C in 1ml MaxPar Fix and Perm Buffer with 1ul of Cell-ID Intercalator-Ir. Cells were washed with MaxPar Cell Staining Buffer, resuspended in 1ml of H_2O/2mM EDTA with 0.5xEQ beads at a concentration of $0.5*10^6$ cells/ml and 35μM-filtered prior to analysis.

Bulk RNA-sequencing

Cell lines were FACS sorted, washed and resuspended in 1ml of TRIzol™. RNA extraction was performed following manufacturer's instructions. RNA samples with a RIN $\geq$ 8 were processed further following Illumina protocols. TruSeq RNA Library Prep Kit (Illumina) was used for samples with high ting cell numbers and 300 ng total RNA was used for library preparation. Libraries were prepared following manufacturer's instructions with the sole modification of PCR cycle number, which was reduced to 12 instead of the recommended 15 cycles. Libraries were eluted in 30μl of Resuspension Buffer provided with the TruSeq Kit and final concentration was measured by Qubit. 1ng/μl library DNA was analysed on the Bioanalyzer as described above and libraries were diluted based on the corrected Bioanalyzer concentration to 10ng/μl. Libraries were pooled, final pool concentration was re-measured by Qubit and libraries were denatured following Illumina's protocol. Libraries were diluted to 4nM in 0.2N NaOH and further diluted to 20pM in Tris-HCL, pH7. Libraries were loaded on NextSeq500.

Bioinformatics

All sequencing data was assessed to detect sequencing failures using FASTQC and lower quality reads were filtered or trimmed using TrimGalore (https://github.com/FelixKrueger/TrimGalore). Outlier samples containing low sequencing coverage or high duplication rates were discarded.

ChIP-sequencing analysis

Reads were mapped to the human genome (hg38) using Bowtie (Langmead et al., 2009). Peaks were detected against rabbit IgG control, or in the case of NALM-6 experiments a mock V5 pulldown, using MACS (Zhang et al., 2008). Heatmaps and average signal plots were generated using NGSPlot (Shen et al., 2014). Read counts were generated for peaks using featureCount*s* from the Rsubread package (Liao et al., 2019). Tracks were visualized using the Integrative Genomics Viewer (IGV) (Robinson et al., 2011). Motif enrichment was determined for 501bp windows centred on peak summits using MEME-ChIP (Machanick and Bailey, 2011).

Processing of RNAseq

Bulk RNAseq samples were mapped to the human reference GRCh38 using tophat2 (https://ccb.jhu.edu/software/tophat/index.shtml). Analyses were performed within the R statistical computing framework, version 4.0.2 using packages from BioConductor version 3.11 (https://Bioconductor.org). Data was combined into a per-gene count matrix using *featureCounts* from the Rsubread package.

Differential gene expression and principal component analysis

The DEseq2 package was used for outlier detection, normalization and differential gene expression analyses (Love et al., 2014). Pairwise comparisons were carried out using the Wald test whereas core signatures across multiple cell lines were derived using a likelihood ratio test (LRT) with full model ~ *cell_line + shRNA_target*, reduced model ~ *cell_line*. "Core" RUNX1 target genes (Figure 5) are defined as padj<0.1 and direction of change the same in pairwise comparisons (shRNA vs control) for each of the 5 cell lines and LRT padj<0.05. Transformed, normalized counts were obtained by variance stabilizing transformation (VST) for downstream analysis. Principal components were obtained using the *prcomp* function from R base. Plots were generated using the ggplot2 package and pheatmap for heatmaps.

Gene sets and Gene Set Enrichment Analysis

Gene signatures of potential biological interest were retrieved from MSigDB version 7.0. In addition, gene signatures derived from label retaining cells were used (Ebinger et al., 2016). Gene set enrichment analysis (GSEA) was performed using the R package Cluster Profiler, function GSEA with default arguments. Preranked lists of genes were derived from DESeq2 analysis – genes were ranked according to their Wald-statistic or, in the case of LRT, the rank was calculated as sign(LFC)*-$\log_{10}$(pvalue).

CyTOF analysis

Single cell events were gated and exported using Cytobank (https://mrc.cytobank.org/). Clustering (SOM) and dimensionality reduction (UMAP) were carried out using the R package CATALYST.

References

Attarbaschi, A., Mann, G., Konig, M., Dworzak, M. N., Trebo, M. M., Muhlegger, N., Gadner, H., Haas, O. A. & Austrian Berlin-Frankfurt-Munster cooperative study, g. 2004. Incidence and relevance of secondary chromosome abnormalities in childhood TEL/AML1+ acute lymphoblastic leukemia: an interphase FISH analysis. *Leukemia,* 18**,** 1611-6.

Behbehani, G. K., Bendall, S. C., Clutter, M. R., Fantl, W. J. & Nolan, G. P. 2012. Single-cell mass cytometry adapted to measurements of the cell cycle. *Cytometry A,* 81**,** 552-66.

Ben-Ami, O., Friedman, D., Leshkowitz, D., Goldenberg, D., Orlovsky, K., Pencovich, N., Lotem, J., Tanay, A. & Groner, Y. 2013. Addiction of t(8;21) and inv(16) acute myeloid leukemia to native RUNX1. *Cell Rep,* 4**,** 1131-43.

Boiers, C., Richardson, S. E., Laycock, E., Zriwil, A., Turati, V. A., Brown, J., Wray, J. P., Wang, D., James, C., Herrero, J., et al. 2018. A Human IPS Model Implicates Embryonic B-Myeloid Fate Restriction as Developmental Susceptibility to B Acute Lymphoblastic Leukemia-Associated ETV6-RUNX1. *Dev Cell,* 44**,** 362-377 e7.

Carpenter, L., Malladi, R., Yang, C. T., French, A., Pilkington, K. J., Forsey, R. W., Sloane-Stanley, J., Silk, K. M., Davies, T. J., Fairchild, P. J., et al. 2011. Human induced pluripotent stem cells are capable of B-cell lymphopoiesis. *Blood,* 117**,** 4008-11.

Choi, A., Illendula, A., Pulikkan, J. A., Roderick, J. E., Tesell, J., Yu, J., Hermance, N., Zhu, L. J., Castilla, L. H., Bushweller, J. H., et al. 2017. RUNX1 is required for oncogenic Myb and Myc enhancer activity in T-cell acute lymphoblastic leukemia. *Blood,* 130**,** 1722-1733.

Dobson, S. M., Garcia-Prat, L., Vanner, R. J., Wintersinger, J., Waanders, E., Gu, Z., McLeod, J., Gan, O. I., Grandal, I., Payne-Turner, D., et al. 2020. Relapse-Fated Latent Diagnosis Subclones in Acute B Lineage Leukemia Are Drug Tolerant and Possess Distinct Metabolic Programs. *Cancer Discov,* 10**,** 568-587.

Ebinger, S., Ozdemir, E. Z., Ziegenhain, C., Tiedt, S., Castro Alves, C., Grunert, M., Dworzak, M., Lutz, C., Turati, V. A., Enver, T., et al. 2016. Characterization of Rare, Dormant, and Therapy-Resistant Cells in Acute Lymphoblastic Leukemia. *Cancer Cell,* 30**,** 849-862.

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D. Y., Barbosa, I. A., Kwon, J. S., Guan, Y., et al. 2013. An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep,* 5**,** 1704-13.

Fisher, O. M., Lord, S. J., Falkenback, D., Clemons, N. J., Eslick, G. D. & Lord, R. V. 2017. The prognostic value of TP53 mutations in oesophageal adenocarcinoma: a systematic review and meta-analysis. *Gut,* 66**,** 399-410.

Friedman, A. D. 2009. Cell cycle and developmental control of hematopoiesis by Runx1. *J Cell Physiol,* 219**,** 520-4.

Fuka, G., Kauer, M., Kofler, R., Haas, O. A. & Panzer-Grumayer, R. 2011. The leukemia-specific fusion gene ETV6/RUNX1 perturbs distinct key biological functions primarily by gene repression. *PLoS One,* 6**,** e26348.

Gandillet, A., Park, S., Lassailly, F., Griessinger, E., Vargaftig, J., Filby, A., Lister, T. A. & Bonnet, D. 2011. Heterogeneous sensitivity of human acute myeloid leukemia to beta-catenin down-modulation. *Leukemia,* 25**,** 770-780.

Goyama, S., Schibler, J., Cunningham, L., Zhang, Y., Rao, Y., Nishimoto, N., Nakagawa, M., Olsson, A., Wunderlich, M., Link, K. A., et al. 2013. Transcription factor RUNX1 promotes survival of acute myeloid leukemia cells. *J Clin Invest,* 123**,** 3876-88.

Greaves, M. 2018. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat Rev Cancer,* 18**,** 471-484.

Greaves, M. F. & Wiemels, J. 2003. Origins of chromosome translocations in childhood leukaemia. *Nat Rev Cancer,* 3**,** 639-49.

Gu, Z., Churchman, M. L., Roberts, K. G., Moore, I., Zhou, X., Nakitandwe, J., Hagiwara, K., Pelletier, S., Gingras, S., Berns, H., et al. 2019. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat Genet,* 51**,** 296-307.

Guidez, F., Petrie, K., Ford, A. M., Lu, H., Bennett, C. A., MacGregor, A., Hannemann, J., Ito, Y., Ghysdael, J., Greaves, M., et al. 2000. Recruitment of the nuclear receptor corepressor N-CoR by the TEL moiety of the childhood leukemia-associated TEL-AML1 oncoprotein. *Blood,* 96**,** 2557-61.

Heltemes-Harris, L. M., Willette, M. J., Ramsey, L. B., Qiu, Y. H., Neeley, E. S., Zhang, N., Thomas, D. A., Koeuth, T., Baechler, E. C., Kornblau, S. M., et al. 2011. Ebf1 or Pax5 haploinsufficiency synergizes with STAT5 activation to initiate acute lymphoblastic leukemia. *J Exp Med,* 208**,** 1135-49.

Hiebert, S. W., Lutterbach, B. & Amann, J. 2001. Role of co-repressors in transcriptional repression mediated by the t(8;21), t(16;21), t(12;21), and inv(16) fusion proteins. *Curr Opin Hematol,* 8**,** 197-200.

Hiebert, S. W., Sun, W., Davis, J. N., Golub, T., Shurtleff, S., Buijs, A., Downing, J. R., Grosveld, G., Roussell, M. F., Gilliland, D. G., et al. 1996. The t(12;21) translocation converts AML-1B from an activator to a repressor of transcription. *Mol Cell Biol,* 16**,** 1349-55.

Illendula, A., Pulikkan, J. A., Zong, H., Grembecka, J., Xue, L., Sen, S., Zhou, Y., Boulton, A., Kuntimaddi, A., Gao, Y., et al. 2015. Chemical biology. A small-molecule inhibitor of the aberrant transcription factor CBFbeta-SMMHC delays leukemia in mice. *Science,* 347**,** 779-84.

Kaindl, U., Morak, M., Portsmouth, C., Mecklenbrauker, A., Kauer, M., Zeginigg, M., Attarbaschi, A., Haas, O. A. & Panzer-Grumayer, R. 2014. Blocking ETV6/RUNX1-induced MDM2 overexpression by Nutlin-3 reactivates p53 signaling in childhood leukemia. *Leukemia,* 28**,** 600-8.

Liao, Y., Smyth, G. K. & Shi, W. 2019. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res,* 47**,** e47.

Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J. P. & Tamayo, P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst,* 1**,** 417-425.

Linka, Y., Ginzel, S., Kruger, M., Novosel, A., Gombert, M., Kremmer, E., Harbott, J., Thiele, R., Borkhardt, A. & Landgraf, P. 2013. The impact of TEL-AML1 (ETV6-RUNX1) expression

in precursor B cells and implications for leukaemia using three different genome-wide screening methods. *Blood Cancer J,* 3**,** e151.

Love, M. I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol,* 15**,** 550.

Lukin, K., Fields, S., Guerrettaz, L., Straign, D., Rodriguez, V., Zandi, S., Mansson, R., Cambier, J. C., Sigvardsson, M. & Hagman, J. 2011. A dose-dependent role for EBF1 in repressing non-B-cell-specific genes. *Eur J Immunol,* 41**,** 1787-93.

Lutz, C., Woll, P. S., Hall, G., Castor, A., Dreau, H., Cazzaniga, G., Zuna, J., Jensen, C., Clark, S. A., Biondi, A., et al. 2013. Quiescent leukaemic cells account for minimal residual disease in childhood lymphoblastic leukaemia. *Leukemia,* 27**,** 1204-1207.

Machanick, P. & Bailey, T. L. 2011. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics,* 27**,** 1696-7.

Mandoli, A., Singh, A. A., Prange, K. H. M., Tijchon, E., Oerlemans, M., Dirks, R., Ter Huurne, M., Wierenga, A. T. J., Janssen-Megens, E. M., Berentsen, K., et al. 2016. The Hematopoietic Transcription Factors RUNX1 and ERG Prevent AML1-ETO Oncogene Overexpression and Onset of the Apoptosis Program in t(8;21) AMLs. *Cell Rep,* 17**,** 2087-2100.

Martinez-Soria, N., McKenzie, L., Draper, J., Ptasinska, A., Issa, H., Potluri, S., Blair, H. J., Pickin, A., Isa, A., Chin, P. S., et al. 2019. The Oncogenic Transcription Factor RUNX1/ETO Corrupts Cell Cycle Regulation to Drive Leukemic Transformation. *Cancer Cell,* 35**,** 705.

May, G., Soneji, S., Tipping, A. J., Teles, J., McGowan, S. J., Wu, M., Guo, Y., Fugazza, C., Brown, J., Karlsson, G., et al. 2013. Dynamic analysis of gene expression and genome-wide transcription factor binding during lineage specification of multipotent progenitors. *Cell Stem Cell,* 13**,** 754-68.

Mikhail, F. M., Serry, K. A., Hatem, N., Mourad, Z. I., Farawela, H. M., El Kaffash, D. M., Coignet, L. & Nucifora, G. 2002. AML1 gene over-expression in childhood acute lymphoblastic leukemia. *Leukemia,* 16**,** 658-68.

Morrow, M., Samanta, A., Kioussis, D., Brady, H. J. & Williams, O. 2007. TEL-AML1 preleukemic activity requires the DNA binding domain of AML1 and the dimerization and corepressor binding domains of TEL. *Oncogene,* 26**,** 4404-14.

Mullighan, C. G. 2012. Molecular genetics of B-precursor acute lymphoblastic leukemia. *J Clin Invest,* 122**,** 3407-15.

Mullighan, C. G., Goorha, S., Radtke, I., Miller, C. B., Coustan-Smith, E., Dalton, J. D., Girtman, K., Mathew, S., Ma, J., Pounds, S. B., et al. 2007. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature,* 446**,** 758-64.

Niebuhr, B., Kriebitzsch, N., Fischer, M., Behrens, K., Gunther, T., Alawi, M., Bergholz, U., Muller, U., Roscher, S., Ziegler, M., et al. 2013. Runx1 is essential at two stages of early murine B-cell development. *Blood,* 122**,** 413-23.

Oo, Z. M., Illendula, A., Grembecka, J., Schmidt, C., Zhou, Y., Esain, V., Kwan, W., Frost, I., North, T. E., Rajewski, R. A., et al. 2018. A tool compound targeting the core binding factor Runt domain to disrupt binding to CBFbeta in leukemic cells. *Leuk Lymphoma,* 59**,** 2188-2200.

Osato, M. 2004. Point mutations in the RUNX1/AML1 gene: another actor in RUNX leukemia. *Oncogene,* 23**,** 4284-96.

Ptasinska, A., Assi, S. A., Martinez-Soria, N., Imperato, M. R., Piper, J., Cauchy, P., Pickin, A., James, S. R., Hoogenkamp, M., Williamson, D., et al. 2014. Identification of a dynamic

core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. *Cell Rep,* 8**,** 1974-1988.

Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. & Mesirov, J. P. 2011. Integrative genomics viewer. *Nat Biotechnol,* 29**,** 24-6.

Schindler, J. W., Van Buren, D., Foudi, A., Krejci, O., Qin, J., Orkin, S. H. & Hock, H. 2009. TEL-AML1 corrupts hematopoietic stem cells to persist in the bone marrow and initiate leukemia. *Cell Stem Cell,* 5**,** 43-53.

Shen, L., Shao, N., Liu, X. & Nestler, E. 2014. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics,* 15**,** 284.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., et al. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102**,** 15545-50.

Sulong, S., Moorman, A. V., Irving, J. A., Strefford, J. C., Konn, Z. J., Case, M. C., Minto, L., Barber, K. E., Parker, H., Wright, S. L., et al. 2009. A comprehensive analysis of the CDKN2A gene in childhood acute lymphoblastic leukemia reveals genomic deletion, copy number neutral loss of heterozygosity, and association with specific cytogenetic subgroups. *Blood,* 113**,** 100-7.

Teppo, S., Laukkanen, S., Liuksiala, T., Nordlund, J., Oittinen, M., Teittinen, K., Gronroos, T., St-Onge, P., Sinnett, D., Syvanen, A. C., et al. 2016. Genome-wide repression of eRNA and target gene loci by the ETV6-RUNX1 fusion in acute leukemia. *Genome Res,* 26**,** 1468-1477.

Tsuzuki, S., Karnan, S., Horibe, K., Matsumoto, K., Kato, K., Inukai, T., Goi, K., Sugita, K., Nakazawa, S., Kasugai, Y., et al. 2007. Genetic abnormalities involved in t(12;21) TEL-AML1 acute lymphoblastic leukemia: analysis by means of array-based comparative genomic hybridization. *Cancer Sci,* 98**,** 698-706.

Wiemels, J. L. & Greaves, M. 1999. Structure and possible mechanisms of TEL-AML1 gene fusions in childhood acute lymphoblastic leukemia. *Cancer Res,* 59**,** 4075-82.

Wilkinson, A. C., Ballabio, E., Geng, H., North, P., Tapia, M., Kerry, J., Biswas, D., Roeder, R. G., Allis, C. D., Melnick, A., et al. 2013. RUNX1 is a key target in t(4;11) leukemias that contributes to gene activation through an AF4-MLL complex interaction. *Cell Rep,* 3**,** 116-27.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome Biol,* 9**,** R137.

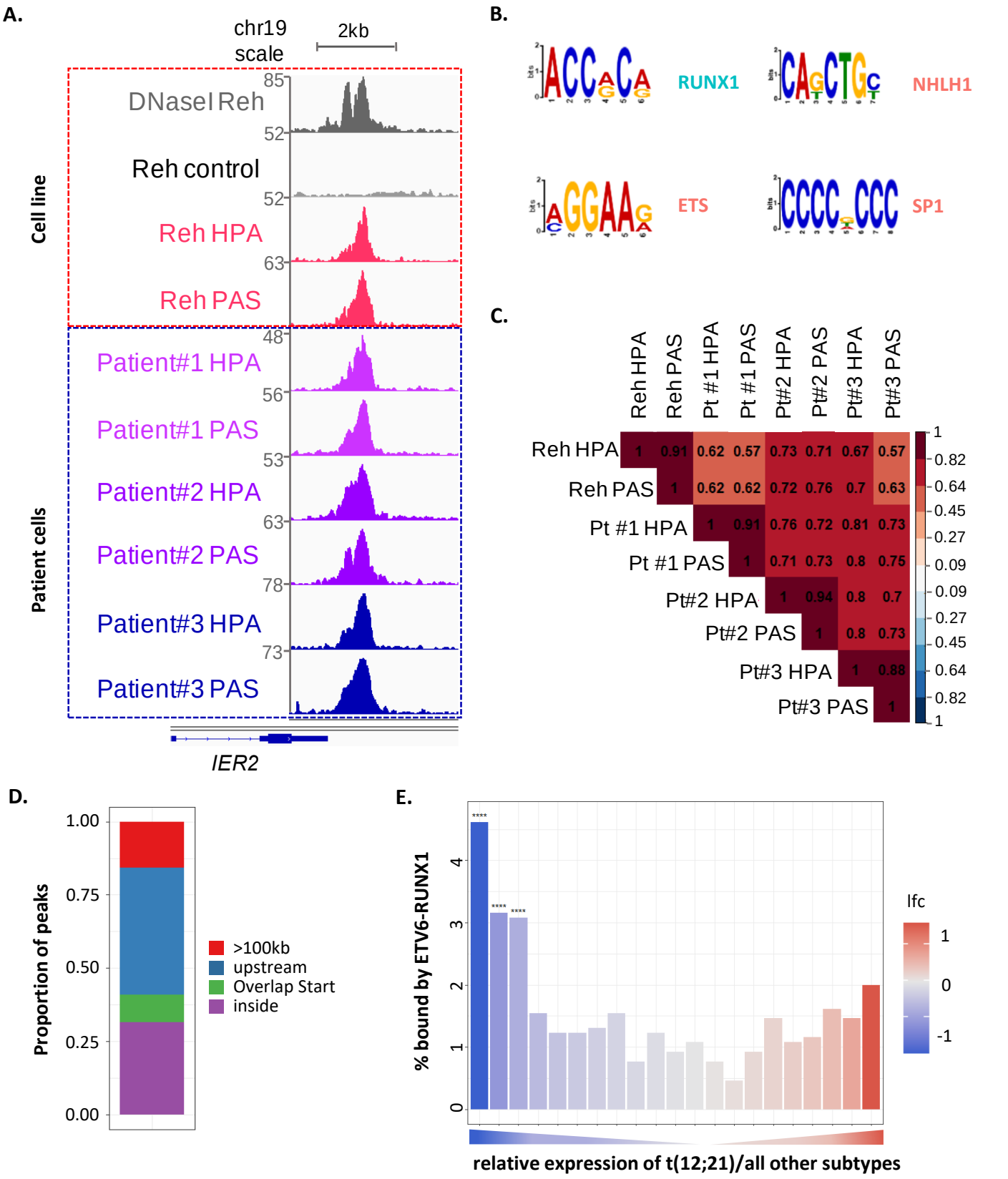

Figure 1. Delineating the ETV6-RUNX1 regulome in childhood B-ALL

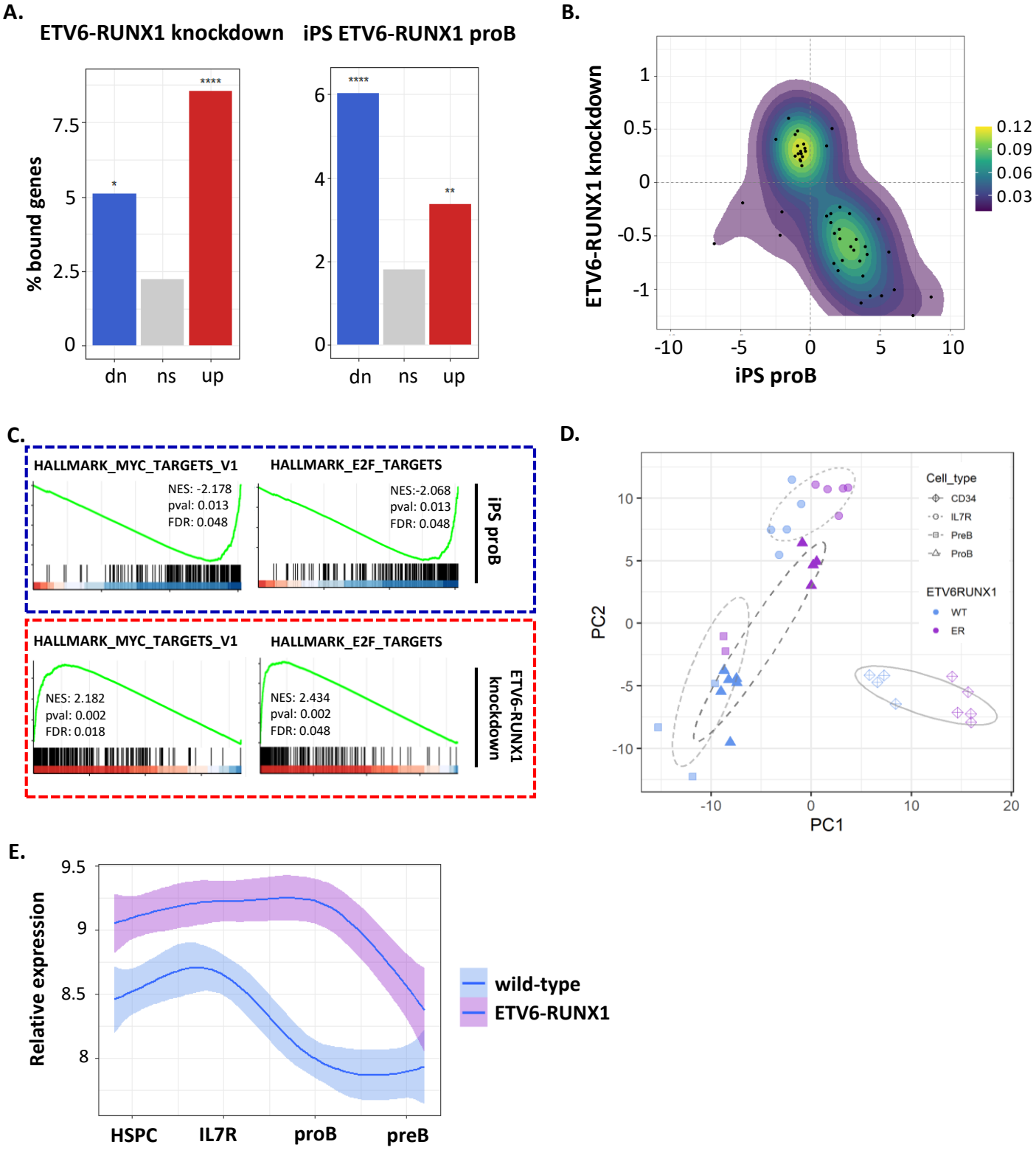

Figure 2. *ETV6-RUNX1* induces transcriptional changes indicative of cell cycle repression in pre-leukemia

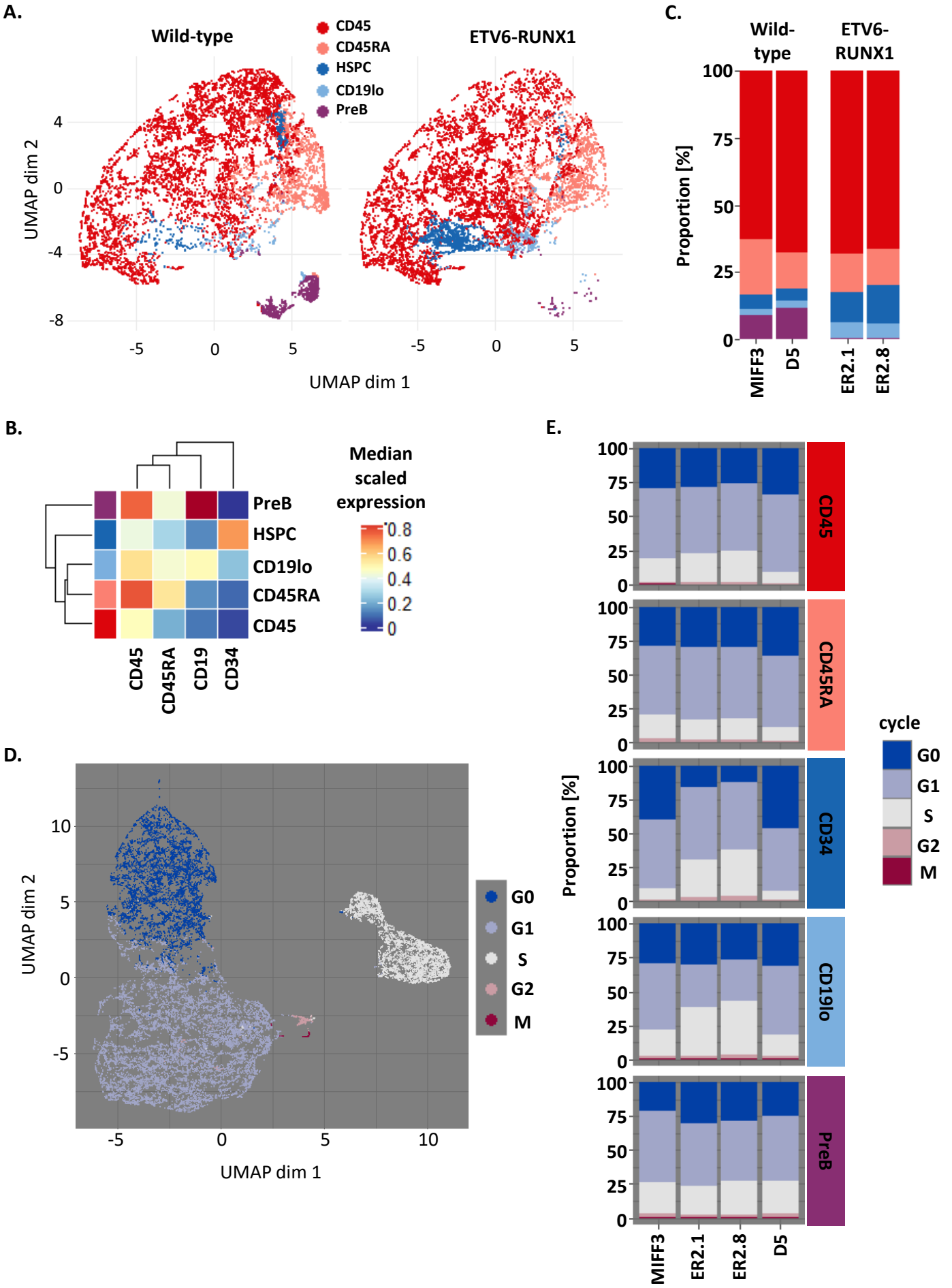

Figure 3. Mass cytometry reveals an accumulation of *ETV6-RUNX1* expressing progenitor cells in S phase

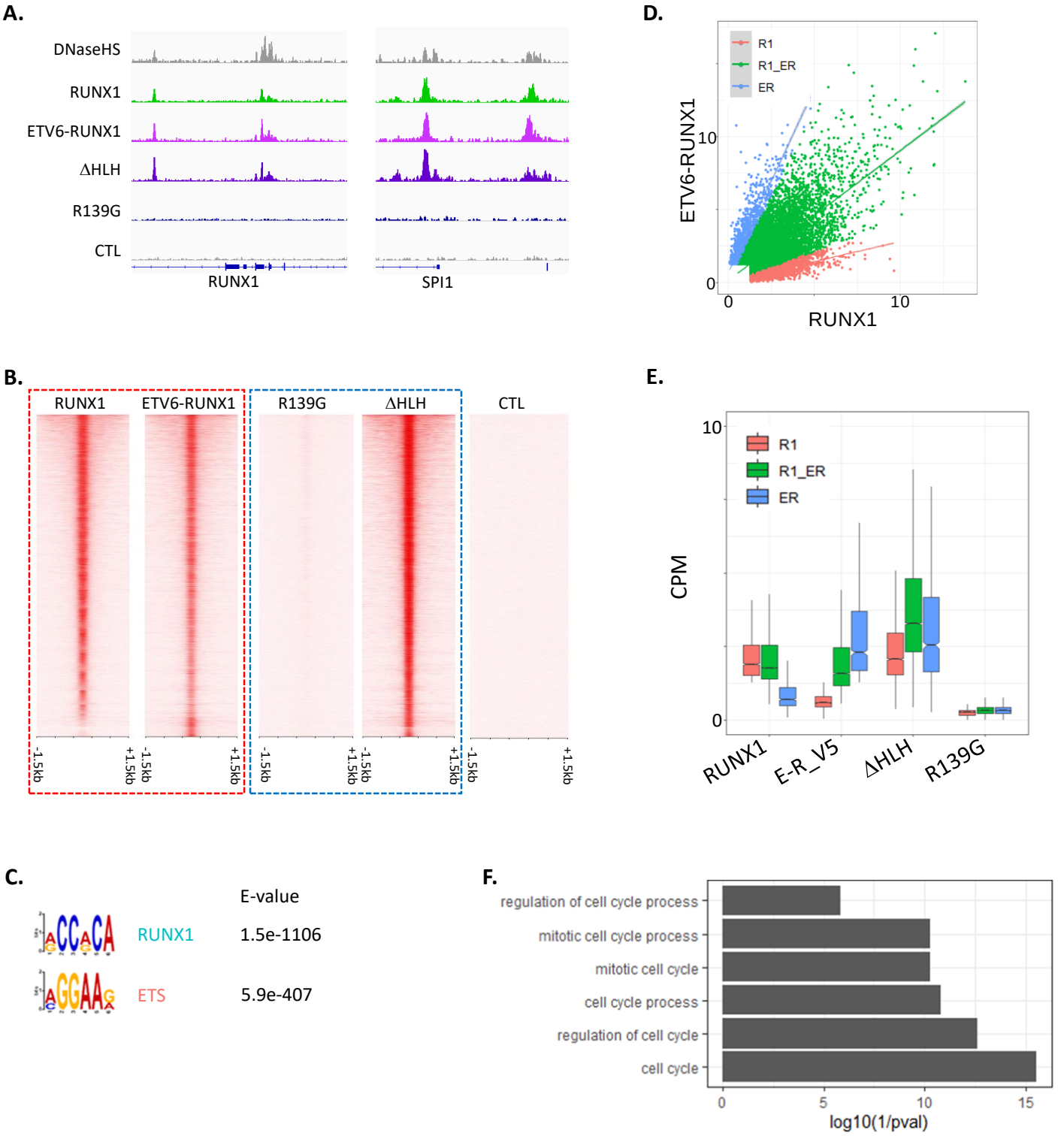

Figure 4. ETV6-RUNX1 binds DNA through the Runt domain, competing with native RUNX1

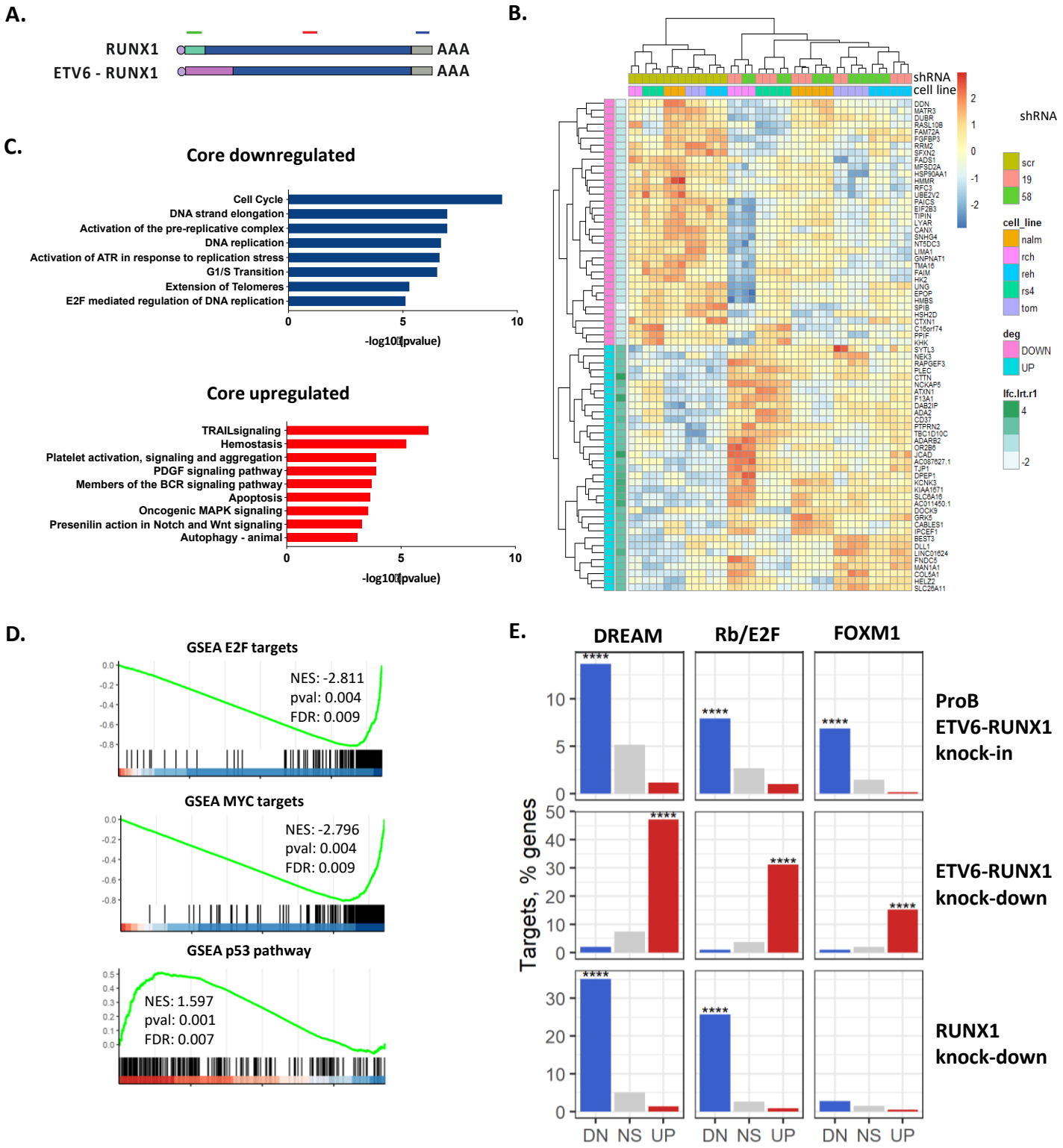

Figure 5. Defining a core RUNX1 program in B-ALL reveals antagonism between ETV6-RUNX1 and native RUNX1 in cell cycle regulation

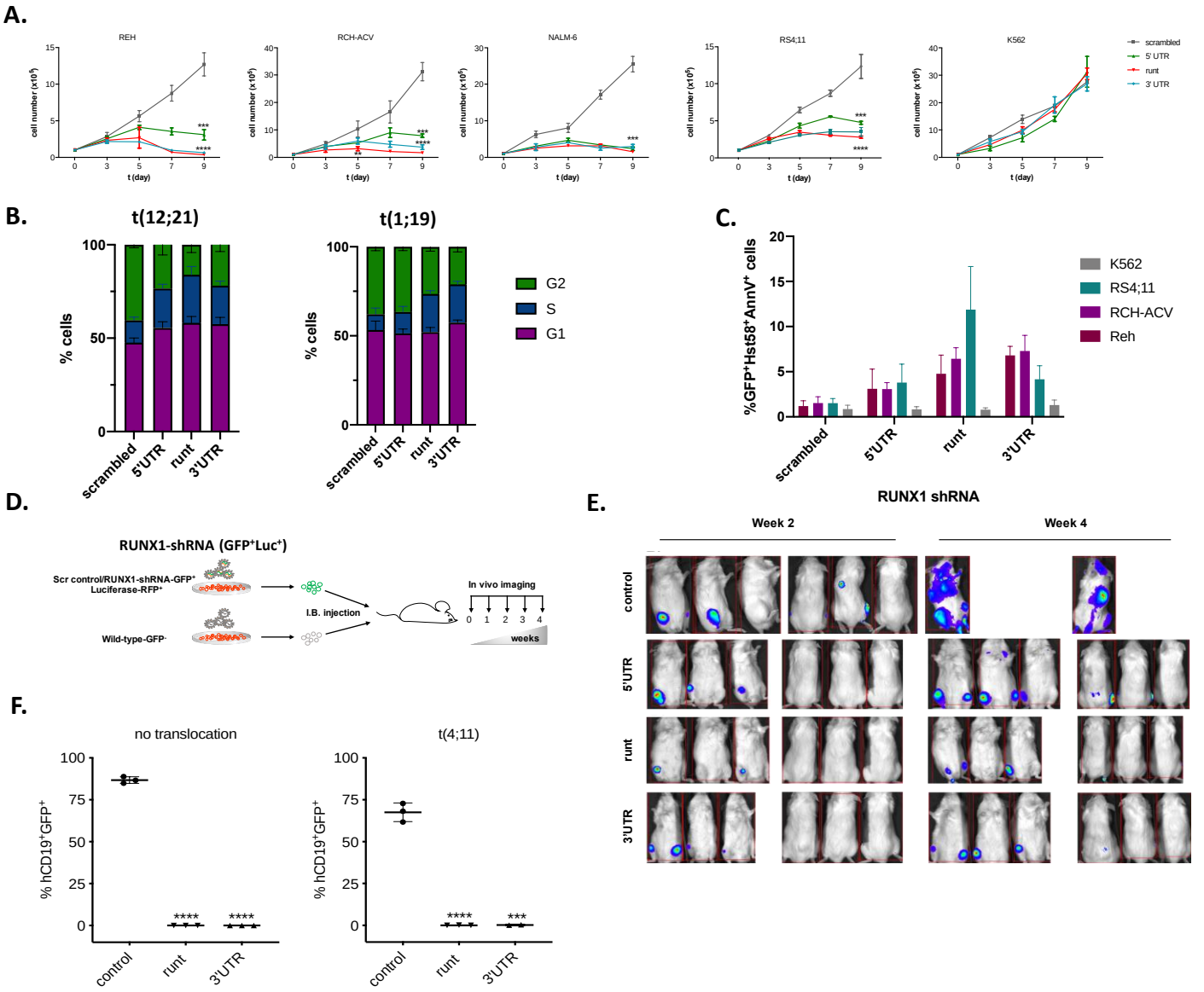

Figure 6. B-ALL cell lines and primary patient cells are dependent on RUNX1 activity *in vitro* and *in vivo*

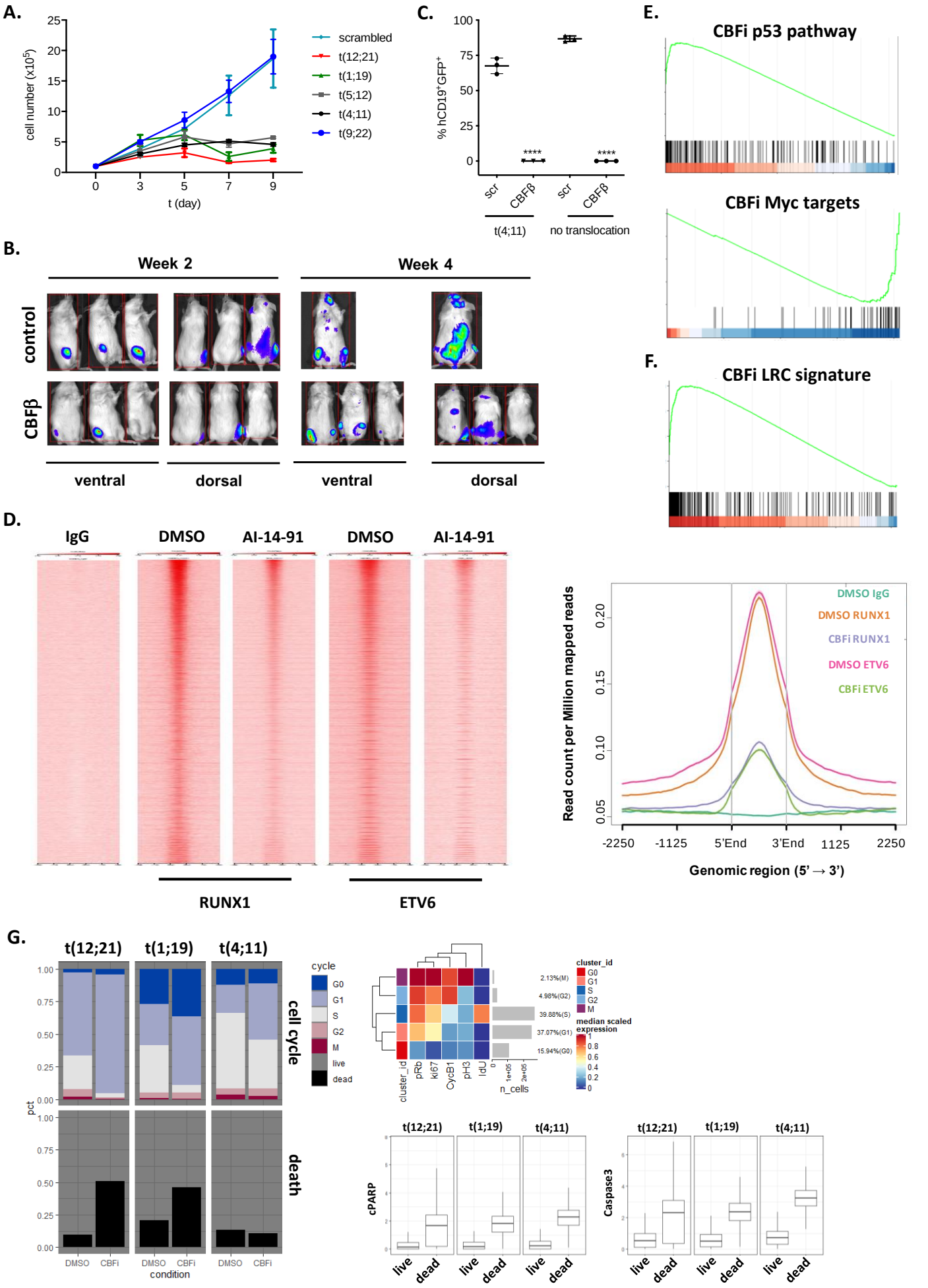

Figure 7. An allosteric CBFb inhibitor mimics RUNX1 depletion phenotype and offers a targeted treatment for B-ALL

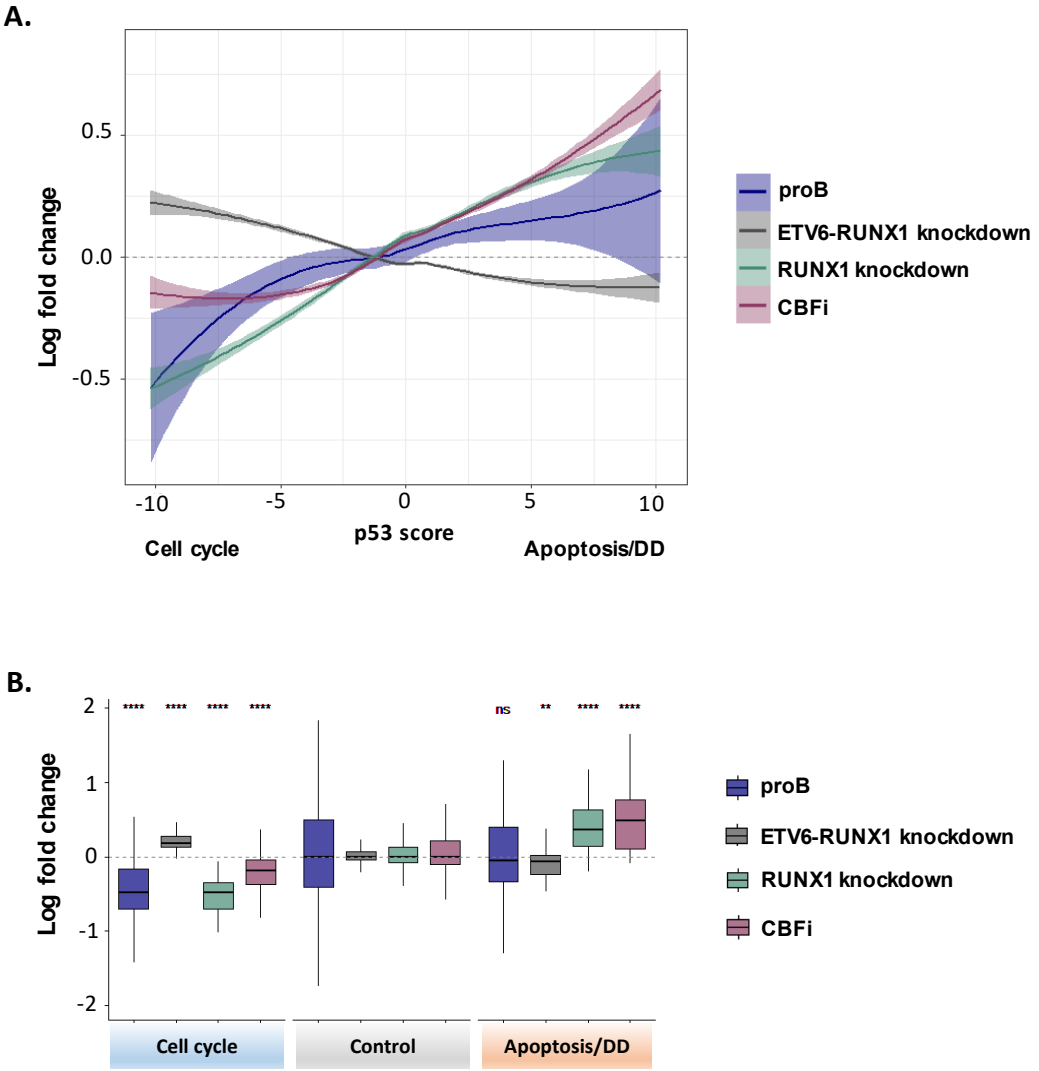

Figure 8. A balance between CBF complex and ETV6-RUNX1 on the P53 regulated cell cycle-apoptosis axis promotes a silent pre-leukemic state

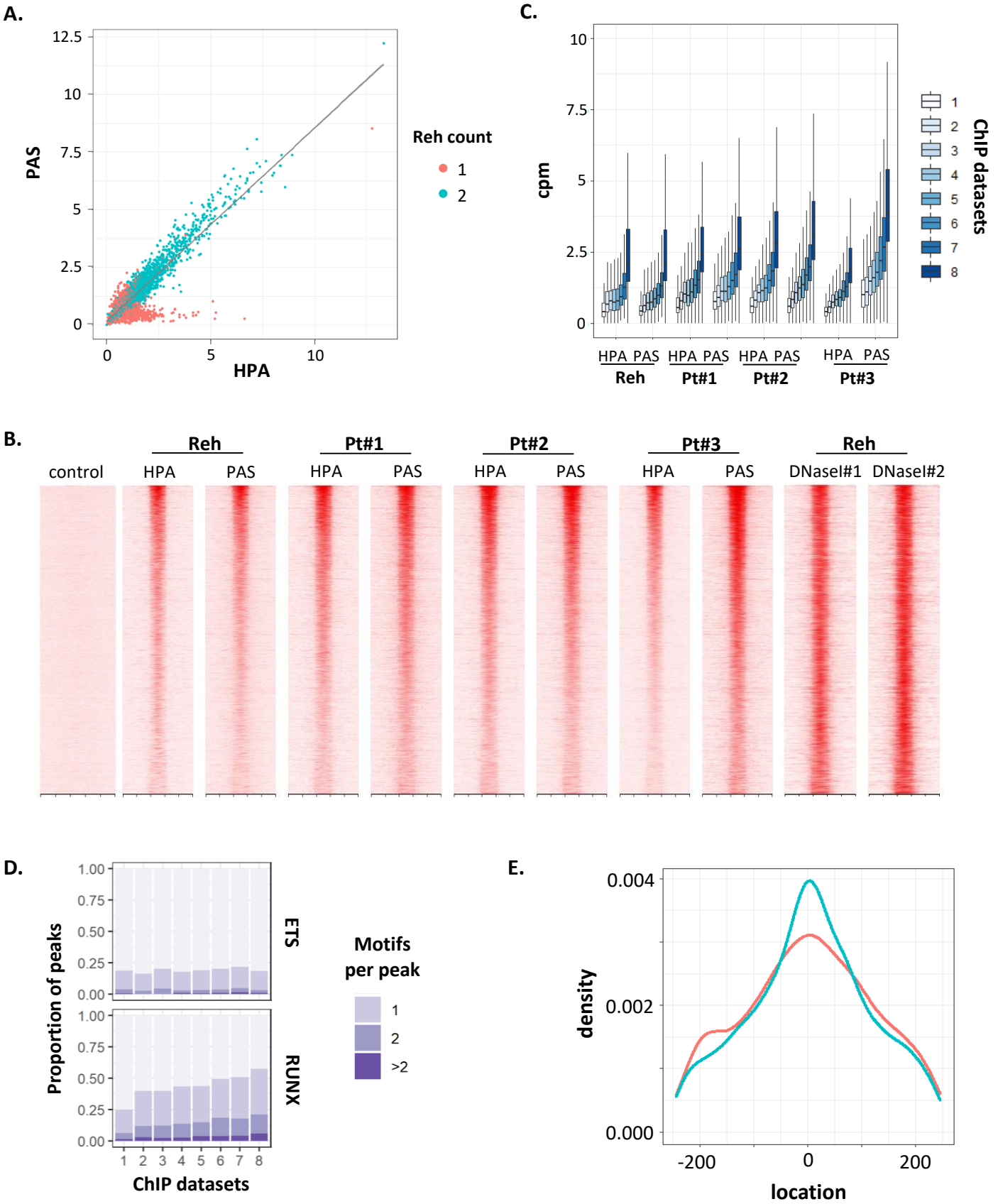

Figure S1. Delineating the ETV6-RUNX1 regulome in childhood B-ALL

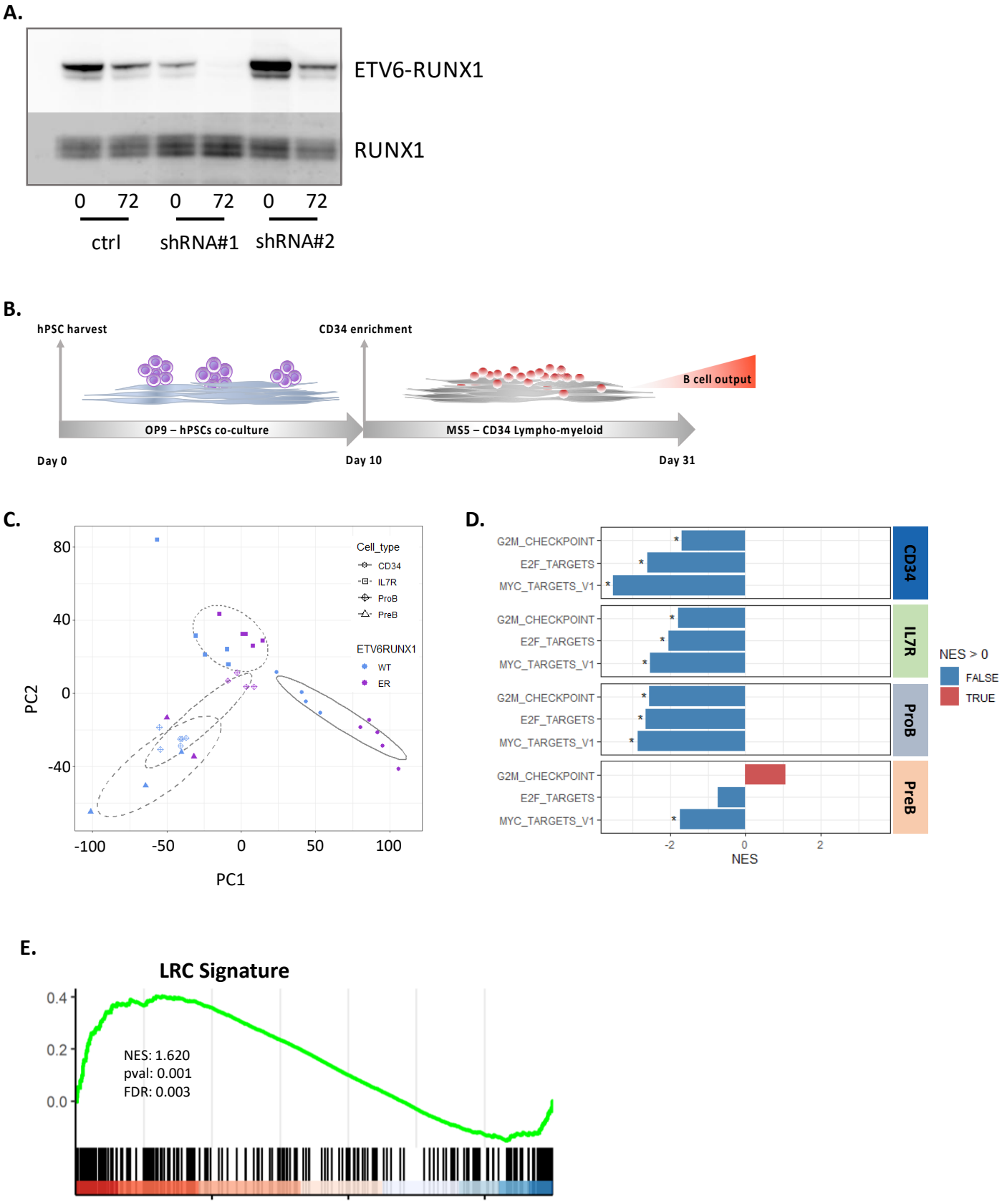

Figure S2. *ETV6-RUNX1* induces transcriptional changes indicative of cell cycle repression in pre-leukemia

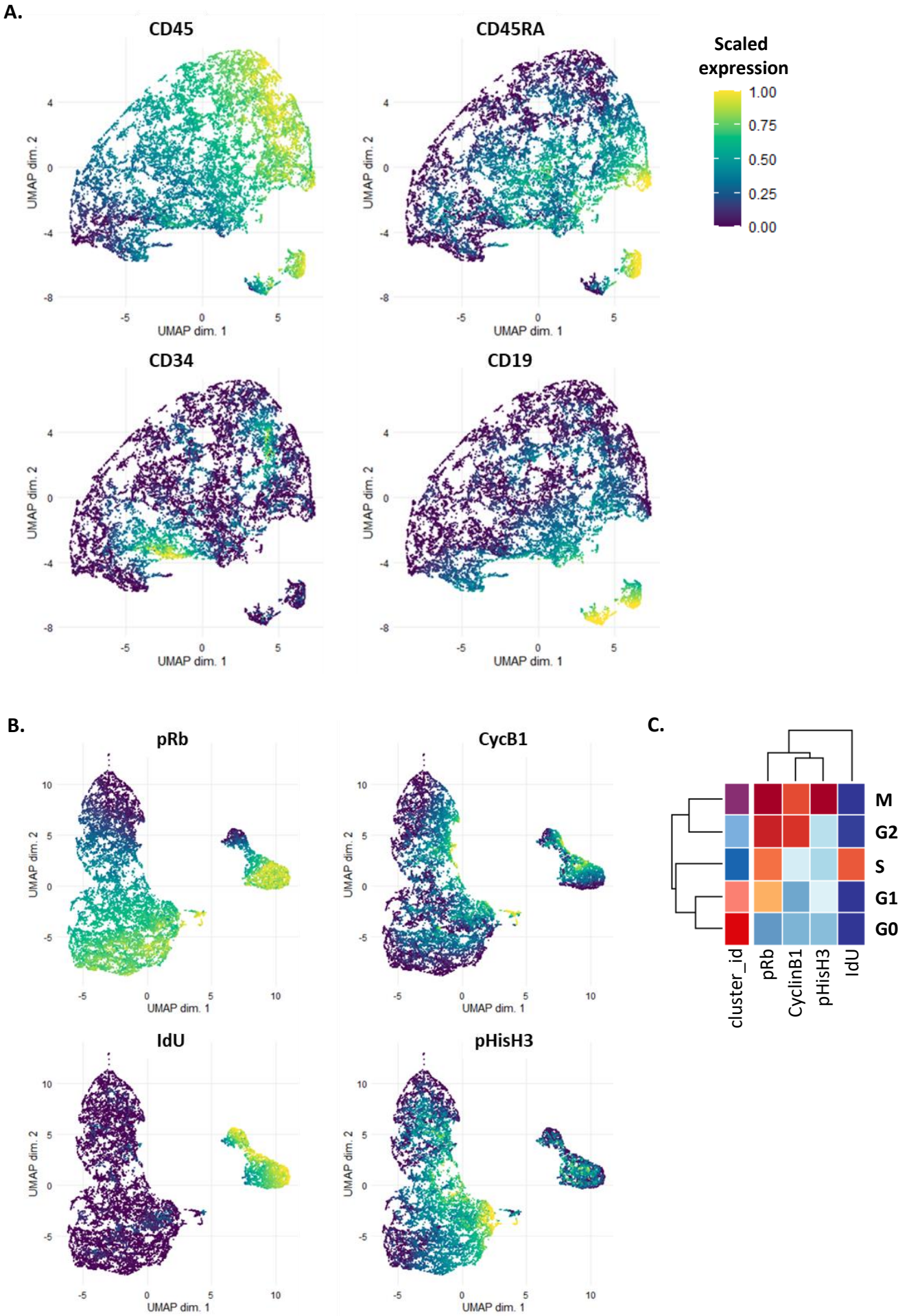

Figure S3. Mass cytometry reveals an accumulation of *ETV6-RUNX1* expressing progenitor cells in S phase

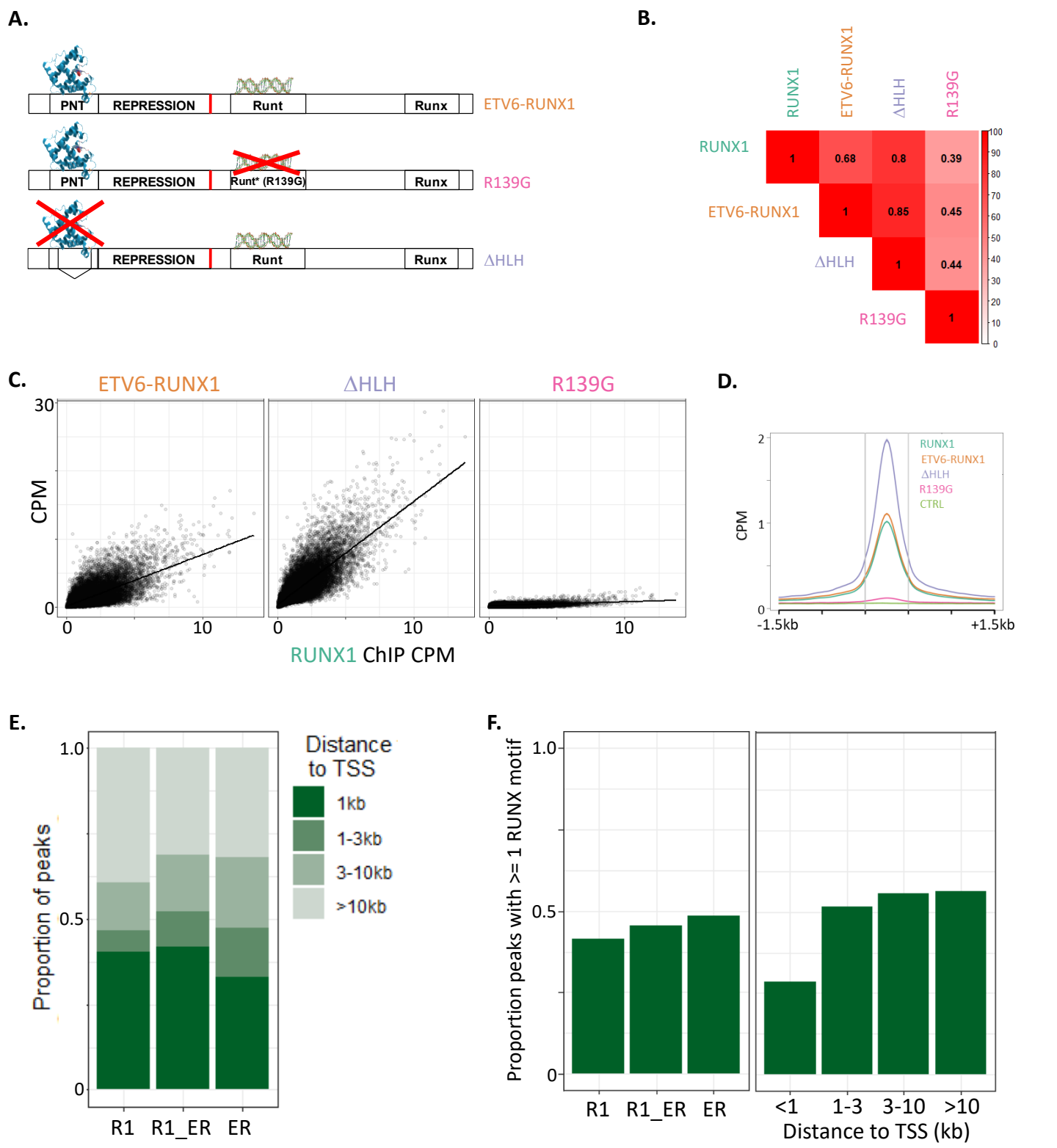

Figure S4. ETV6-RUNX1 binds DNA through the Runt domain, competing with native RUNX1

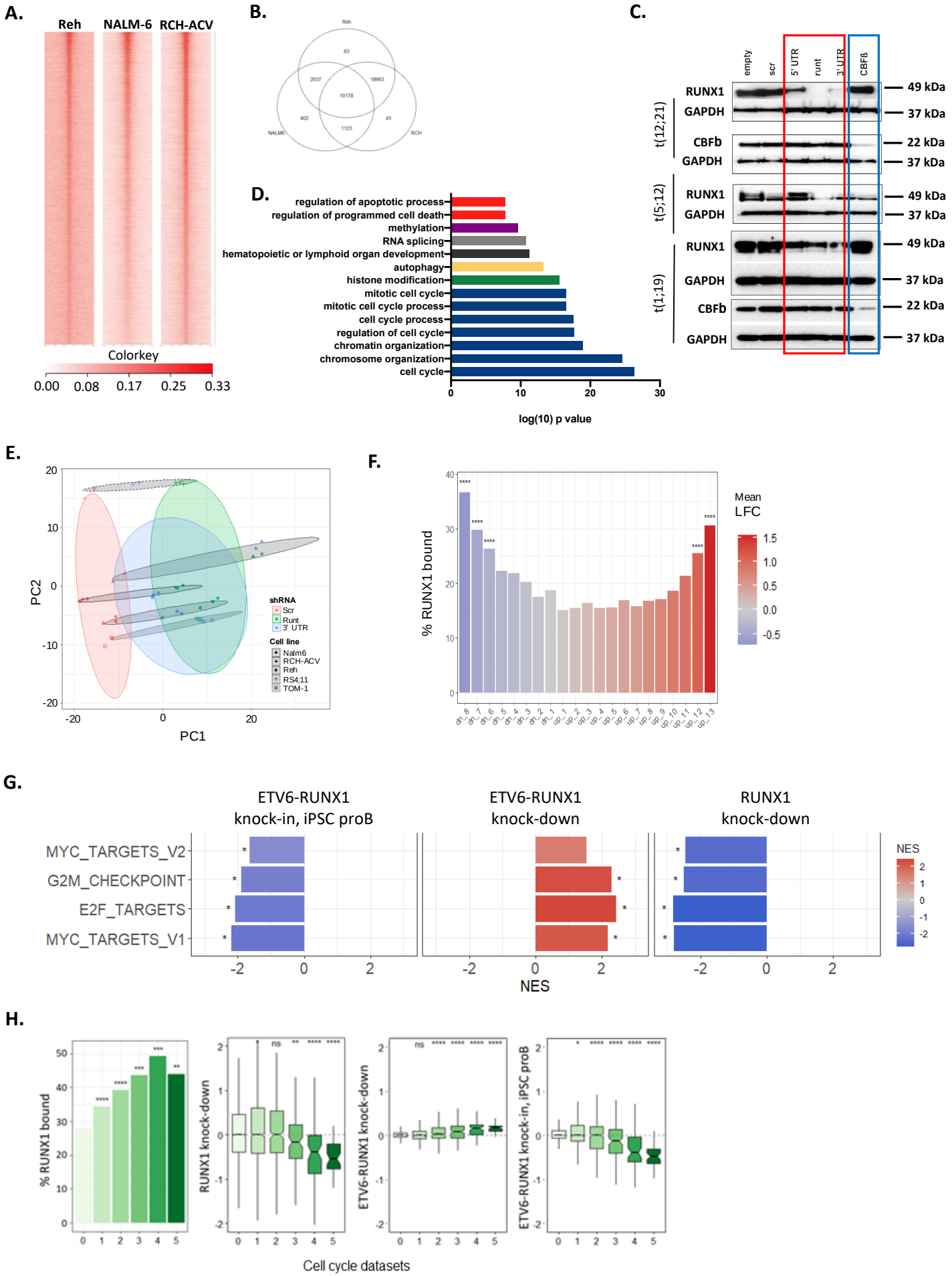

Figure S5. Defining a core RUNX1 program in B-ALL reveals antagonism between ETV6-RUNX1 and native RUNX1 in cell cycle regulation

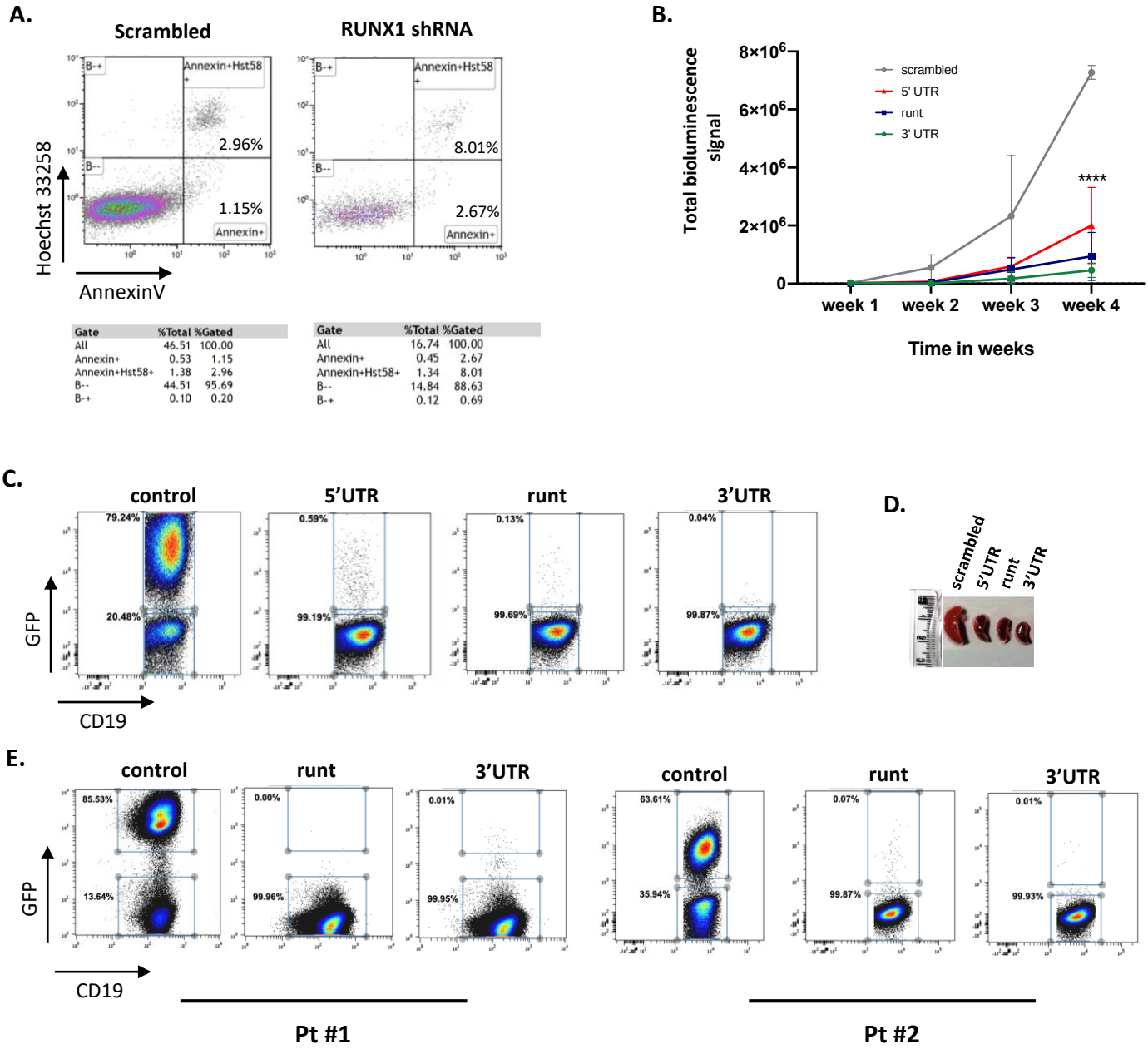

Figure S6. B-ALL cell lines and primary patient cells are dependent on RUNX1 activity *in vitro* and *in vivo*

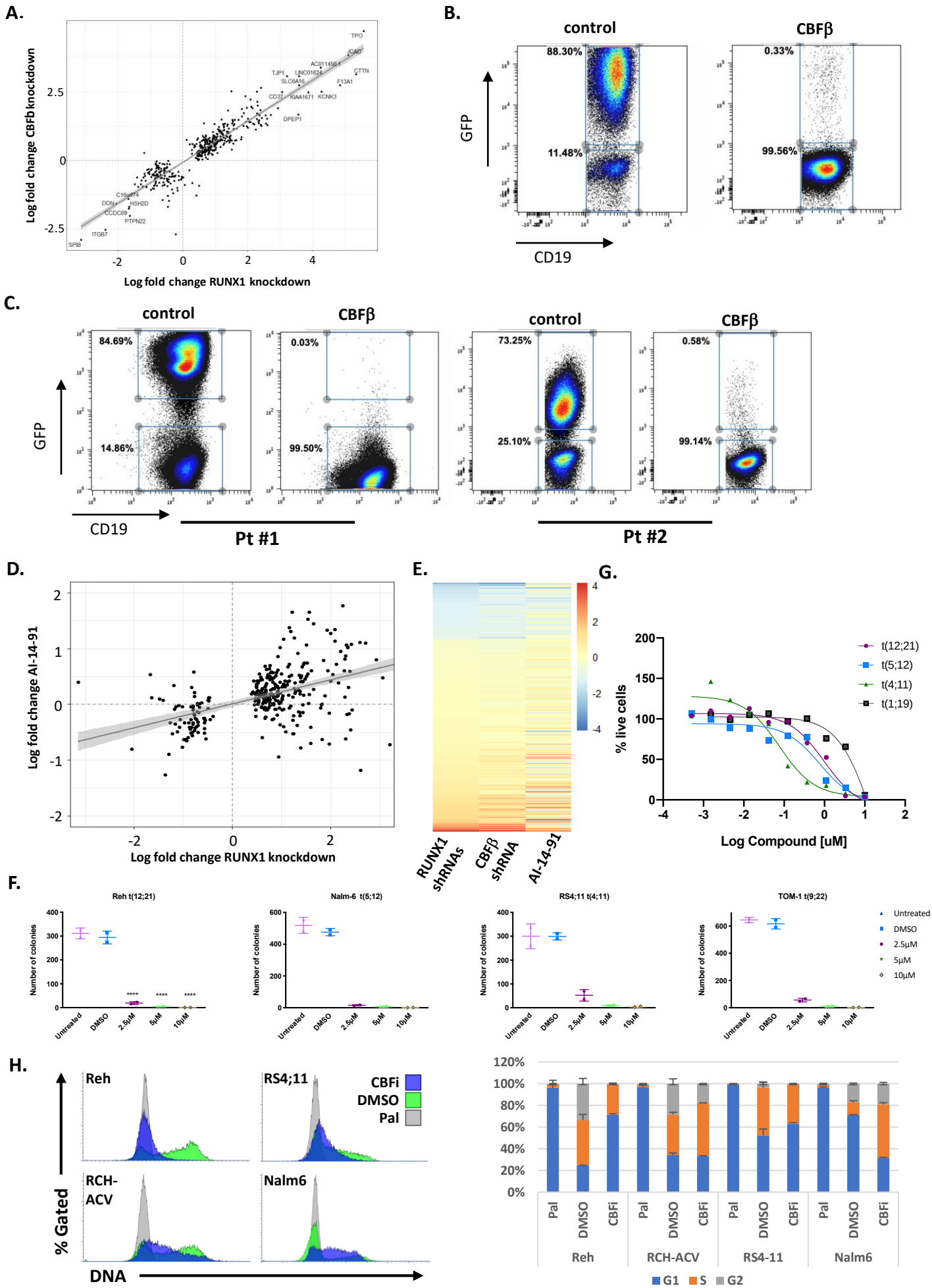

Figure S7. An allosteric CBFb inhibitor mimics RUNX1 depletion phenotype and offers a targeted treatment for B-ALL

A.

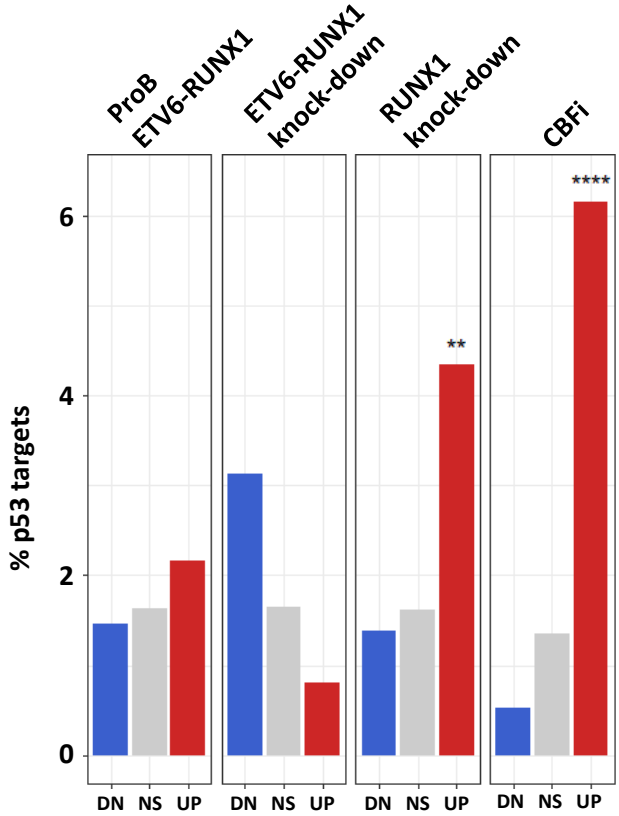

Figure S8. A balance between CBF complex and ETV6-RUNX1 on the P53 regulated cell cycle-apoptosis axis promotes a silent pre-leukemic state