CANCER

DUX4-r exerts a neomorphic activity that depends on GTF2I in acute lymphoblastic leukemia

Daniele Campolungo[1], Mara Salomé[1], Beatrice Biferali[1], Anna Sofia Tascini[2], Davide Gabellini[1]*

Translocations producing rearranged versions of the transcription factor double homeobox 4 (DUX4-r) are one of the most frequent causes of B cell acute lymphoblastic leukemia (B-ALL). DUX4-r retains the DNA binding domain of wild-type DUX4 but is truncated on the C-terminal transcription activation domain. The precise mechanism through which DUX4-r causes leukemia is unknown, and no targeted therapy is currently available. We found that the rearrangement leads to both a loss and a gain of function in DUX4-r. Loss of CBP/EP300 transcriptional coactivator interaction leads to an inability to bind and activate repressed chromatin. Concurrently, a gain of interaction with the general transcription factor 2 I (GTF2I) redirects DUX4-r toward leukemogenic targets. This neomorphic activity exposes an Achilles' heel whereby DUX4-r–positive leukemia cells are exquisitely sensitive to GTF2I targeting, which inhibits DUX4-r leukemogenic activity. Our work elucidates the molecular mechanism through which DUX4-r causes leukemia and suggests a possible therapeutic avenue tailored to this B-ALL subtype.

INTRODUCTION

B cell acute lymphoblastic leukemia (B-ALL) is the most common pediatric malignancy and the most frequent cause of death from cancer at a young age (*1*–*3*). B-ALL is a highly heterogeneous disease driven by recurrent chromosomal alterations giving rise to multiple disease subtypes. Therefore, the characterization of the leukemic pathways associated with specific B-ALL drivers could allow the discovery of new therapeutic vulnerabilities from a personalized medicine perspective (*4*).

Recurrent translocations of the *double homeobox 4* (*DUX4*) gene to the *immunoglobulin heavy chain* (*IGH*) locus have been reported in up to 10% of patients with B-ALL and define a new B-ALL subtype, characterized by deregulation of *ETS-related gene* (*ERG*), positivity for CD371, a distinctive gene expression profile and widespread hypomethylation compared to healthy B cells or other ALL subtypes (*5*–*9*).

DUX4 is a sequence-specific transcription factor whose expression is normally confined to early embryonic development, where it activates the expression of cleavage-stage genes (*10*–*12*). DUX4 has an N-terminal DNA binding domain and activates gene expression by recruiting CREB-binding protein/E1A binding protein p300 (CBP/EP300) coactivators through its C-terminal transcription activation domain (CTD) (*13*).

All *DUX4* translocations in B-ALL maintain the region encoding for the DNA binding domain but result in random truncations of the CTD, which, in most cases, is substituted by random amino acids encoded by the *immunoglobulin heavy* (*IGH*) locus giving rise to DUX4-IGH chimeras (*5*–*7*). The resulting rearranged DUX4 (DUX4-r) is expressed specifically in B cell precursors (*5*–*7*), where wild-type (wt) DUX4 is never expressed.

The *DUX4* rearrangement is a clonal event occurring early in leukemogenesis and maintained in relapsed leukemia. DUX4-r silencing in patient-derived xenograft or the NALM6 cell line derived from DUX4-r B-ALL patients supports an essential role of DUX4-r in the development and maintenance of this leukemia subtype (*5*, *14*).

Nevertheless, it was recently proposed that wt DUX4 and DUX4-r function interchangeably and that leukemia development requires "just right" wt DUX4 or DUX4-r levels. Too much wt DUX4 or DUX4-r would result in apoptosis and too little would just block cell proliferation (*15*).

In addition to *DUX4* rearrangements, *ERG* alterations are observed almost exclusively in the DUX4-r B-ALL subtype (*16*–*18*). However, these alterations are absent in up to 40% of DUX4-r patients, are often subclonal, and inconsistent between diagnosis and relapse, thus indicating that *ERG* alterations do not drive leukemogenesis but represent a secondary event in disease progression (*16*, *19*). Hence, additional factors in addition to *ERG* alterations may be required for DUX4-r–induced leukemogenesis.

The proposed models for DUX4-r–driven B-ALL, including cooperation with ERG alterations or just right DUX4 levels, are not fully supported by the literature (*5*, *6*, *10*, *15*). Additional studies suggest a novel activity of the DUX4-r oncogene (*7*, *8*, *20*), but are based on very different genetic backgrounds or cellular systems. Functional enrichment analyses on DUX4-r transcriptomic data so far failed to identify specifically dysregulated pathways (*7*–*9*). RNA sequencing (RNA-seq) has been performed upon DUX4-r overexpression in NALM6 cells, which nevertheless already express it endogenously, or 72 hours after DUX4-r knockdown (*9*). Moreover, chromatin immunoprecipitation sequencing (ChIP-seq) data for endogenous DUX4-r in NALM6 cells have been compared to those of ectopically overexpressed wt DUX4 in pluripotent stem cells or human myoblasts (*6*, *9*, *10*, *21*, *22*). Thus, it is very difficult to discern primary DUX4-r transcriptional targets from secondary events and there remains a need to compare the activity of wt DUX4 and DUX4-r in a relevant system to define the molecular mechanism through which DUX4-r promotes B-ALL and identify vulnerabilities that could be exploited for therapeutic purposes.

[1]Gene Expression and Muscular Dystrophy Unit, Division of Genetics and Cell Biology, IRCCS San Raffaele Scientific Institute, 20132 Milano, Italy. [2]Center for Omics Sciences, IRCCS Ospedale San Raffaele, 20132 Milano, Italy.
*Corresponding author. Email: gabellini.davide@hsr.it

Here, we compared side by side the molecular and functional activity of wt DUX4 and representative patient-derived DUX4-r variants in a B-ALL context. Despite sharing the same DNA binding domain, we found that wt DUX4 and DUX4-r regulate different gene sets and control very different cell behavior. Only DUX4-r activates genes belonging to the DUX4-r B-ALL patient gene signature and stimulates B-ALL cell proliferation. We found that DUX4-r acquires these abilities through a gain of interaction with the transcription factor GTF2I. We further demonstrated that the activity of DUX4-r strictly depends on the cellular availability of GTF2I. Accordingly, genetic or pharmacological GTF2I targeting inhibits DUX4-r leukemogenic activity in vitro and in vivo. Collectively, our data clarify the molecular mechanism responsible for DUX4-r B-ALL and identify a possible therapeutic option for this leukemia subtype.

RESULTS

DUX4 and DUX4-r associate to and activate different gene sets

To compare the activity of wt DUX4 and DUX4-r variants, we used REH cells, a near-diploid B-ALL cell line lacking endogenous expression of either wt DUX4 or its rearrangements (*5*, *6*). We focused on patient-derived DUX4-r variants representing the distinct types of mutations documented in patients with B-ALL: the DUX4-IGH chimera mostly used in the literature (*5*, *6*, *22*) and a DUX4-r lacking the whole CTD without any extra amino acid (DUX4-del50) (*6*). We generated doxycycline (Dox)-inducible REH cells for the empty vector (EV), wt DUX4, DUX4-IGH, or DUX4-del50.

Upon Dox induction, wt DUX4 was expressed at consistently lower levels compared to DUX4-r variants (fig. S1, A and B). Unexpectedly, only wt DUX4 was able to induce apoptosis (fig. S1C), in contrast with the "just right" model, which posits that wt DUX4 or its variants induce apoptosis when expressed at high levels (*15*).

We then decided to compare the transcriptional ability of wt DUX4 and DUX4-r variants. To this aim, we performed RNA-seq upon 12 hours of Dox induction, to avoid the toxic effect of prolonged wt DUX4 expression and enrich for primary targets of the transcription factors. Principal components analysis showed very tight clustering of the biological replicates and that the gene expression profiles of cells expressing DUX4-IGH and DUX4-del50 cluster nearby and distantly from those of EV or DUX4 samples (fig. S1D). Despite differential gene expression (DGE) analysis indicated that wt DUX4 or DUX4-r variants are mainly associated with the activation of gene expression, and DUX4-r variants regulate a very different set of genes compared to wt DUX4 (Fig. 1A and table S1). Notably, DUX4-IGH and DUX4-del50 regulated genes are qualitatively similar, even if DUX4-del50 appears to have a milder transactivation ability with respect to DUX4-IGH (Fig. 1A), which could be possibly explained by lower protein levels as compared to DUX4-IGH (fig. S1B). Accordingly, we found a nearly complete overlap between DUX4-IGH and DUX4-del50 regulated genes, which was instead minimal with wt DUX4 (fig. S1E).

Previous works failed to identify significantly deregulated pathways on the gene signature specific to DUX4-r B-ALL patients (*8*, *20*). Using the publicly available datasets (*6*, *8*), we found a highly significant overlap of the DUX4-r B-ALL patient gene signature with our DUX4-IGH or DUX4-del50 datasets, which was instead negligent with our wt DUX4 dataset (Fig. 1B). We next performed gene set enrichment analysis (GSEA) on the list of genes commonly deregulated in DUX4-r B-ALL patients and our DUX4-IGH and DUX4-del50 datasets. We will refer to these commonly deregulated genes as the DUX4-r core gene set from here onward. GSEA revealed that the up-regulated DUX4-r core gene set is mostly enriched in molecules involved in cell-cell/cell-matrix interaction, cell migration, and cancer-related pathways (Fig. 1C). The down-regulated DUX4-r core gene set is instead mostly involved in pre-B cell receptor activity, B cell activation, and immune-related processes (Fig. 1C).

Using the previously reported wt DUX4-associated gene expression signature (*23*), we obtained opposite results: a significant overlap for only our wt DUX4 dataset (fig. S1F), matching previously reported (*23*) GSEA pathways including regulation of transcription, RNA processing, and apoptosis (fig. S1G).

Previous studies demonstrated that wt DUX4 activates the expression of several classes of repetitive elements (REs) including endogenous retroviruses and pericentromeric human satellite II repeats (*10*, *12*, *21*, *24*, *25*). To investigate the activity of DUX4-r variants in the context of the REs genome, we reanalyzed our RNA-seq results and performed DGE analysis focused on the REs transcriptome. In line with previous reports (*10*, *12*, *21*, *24*, *26*), we found that wt DUX4 activates the expression of hundreds of REs (Fig. 1D and table S2). Instead, the REs transcriptome of both DUX4-r variants mostly resembles that of control EV cells (Fig. 1D and table S2). Hence, contrary to wt DUX4, DUX4-r variants minimally affect the REs transcriptome.

To identify direct transcriptional targets of DUX4-r variants, we mapped the chromatin occupancy of wt DUX4 and DUX4-r variants by performing Cleavage Under Targets and Tagmentation (CUT&Tag) (*27*) using an antibody (Ab) recognizing a DUX4 epitope retained in DUX4-r. We found that wt DUX4 and its rearranged versions associate with largely nonoverlapping genomic regions, while DUX4-IGH and DUX4-del50 mostly bind to the same genomic loci (Fig. 1E and table S3). This result is unexpected considering that all DUX4-r variants described to date maintain the DNA binding domain of wt DUX4 (*5*, *6*, *9*). Wt DUX4 recognizes the consensus motif 5′-TAAYBBAATCA-3′ (*24*, *28*). Notably, de novo motif identification analyses indicated that, for both wt DUX4 and DUX4-r variants, the most enriched sequence motif under the CUT&Tag peaks was the DUX4 consensus (Fig. 1F).

Wt DUX4 functions as a pioneer transcription factor able to bind condensed chromatin and increase its accessibility, by recruiting the histone acetyltransferases CBP and EP300 via its CTD to promote histone H3 lysine 27 acetylation (H3K27Ac) (*13*). We evaluated H3K27Ac by CUT&Tag and chromatin accessibility by Cleavage Under Targeted Accessible Chromatin (CUTAC) (*29*). In line with previously reported results (*10*, *11*, *13*), wt DUX4 induction was associated with significantly increased H3K27Ac and chromatin accessibility of its direct targets. In contrast, DUX4-IGH induction did not cause significant changes in H3K27Ac or chromatin accessibility on DUX4-r direct targets (Fig. 1, G and H). Most wt DUX4 direct targets were not expressed in its absence, while nearly half of DUX4-r direct targets were already expressed, albeit at lower levels, also in its absence (fig. S1H and table S1). Hence, DUX4-r fails to bind and activate repressed chromatin.

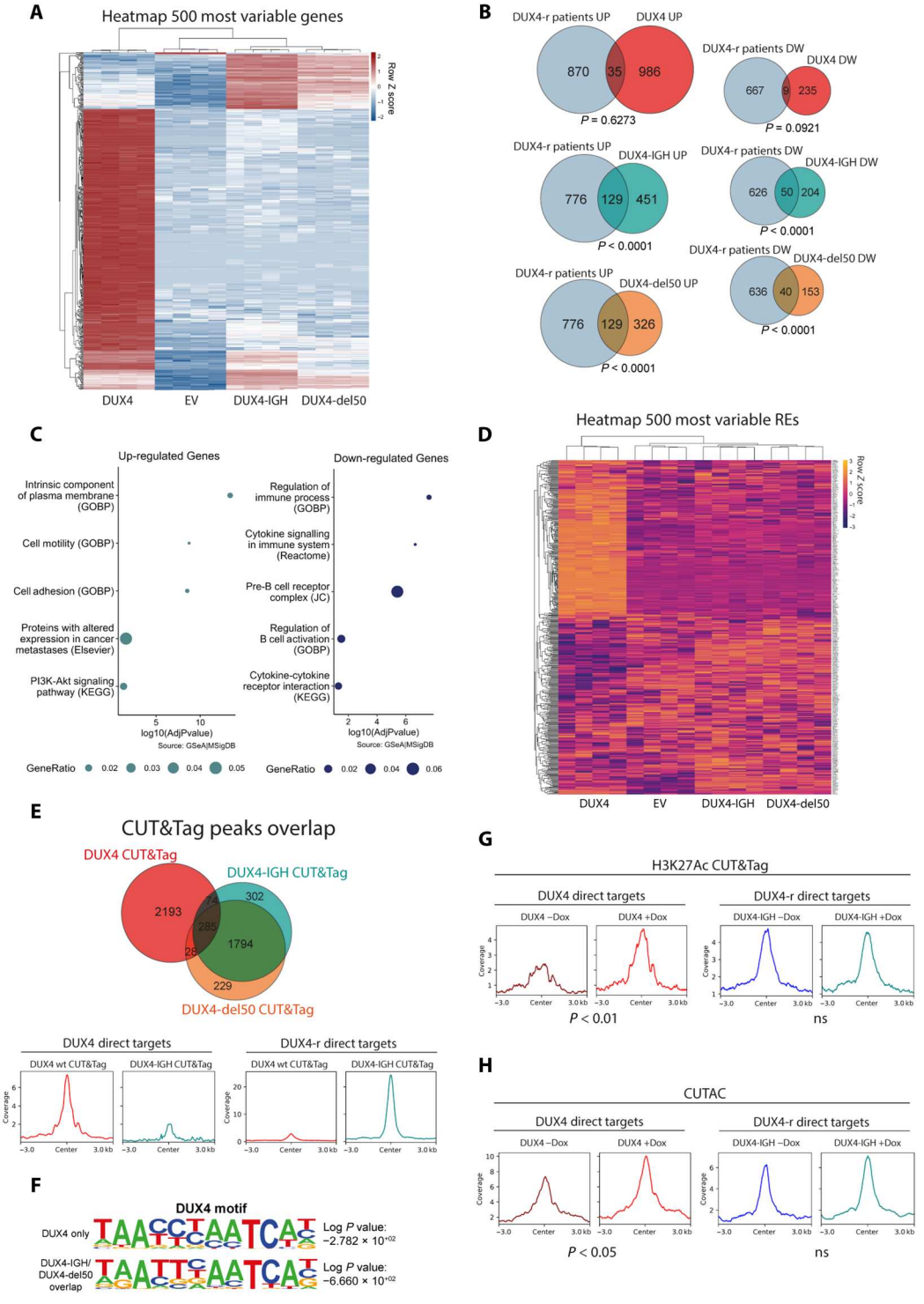

Fig. 1. DUX4 and DUX4-r associate to and activate different gene sets. (**A**) Heatmap showing the 500 most variable genes in REH cells expressing EV, wt DUX4, DUX4-IGH, or DUX4-del50. (**B**) Venn diagram displaying (left) overlap between genes up-regulated in the present study and in DUX4-r B-ALL patients (EGAS00001000654) and (right) overlap between genes down-regulated in the present study and in DUX4-r B-ALL patients (EGAS00001000654). Two-tailed Fisher's exact test. (**C**) Functional families enriched in the overlap between: genes positively (left) or negatively (right) regulated in DUX4-r B-ALL patients and by DUX4-r in REH cells. Dot size is proportional to the percentage of modulated genes in each pathway. (**D**) Heatmap showing the 500 most variable repetitive elements in REH cells expressing EV, wt DUX4, DUX4-IGH, or DUX4-del50. (**E**) Top: Venn diagram displaying the overlap between CUT&Tag peaks in wt DUX4–, DUX4-IGH–, or DUX4-del50–expressing REH cells. Bottom: CUT&Tag profile plots of wt DUX4 (red) and DUX4-IGH (cyan) peaks on wt DUX4 (left) or DUX4-r (right) direct transcriptional targets. (**F**) De novo motif calling of peaks selectively associated to wt DUX4 (DUX4 only) or DUX4-IGH and DUX4-del50 (DUX4-IGH/DUX4-del50 overlap). (**G**) CUT&Tag profile plot of H3K27Ac at wt DUX4 (maroon/red) or DUX4-IGH (blue/teal) direct targets in the absence (left) or presence (right) of Dox. Wilcoxon rank sum test to compare the signal of H3K27Ac signal between −Dox and +Dox. (**H**) Cleavage Under Targeted Accessible Chromatin (CUTAC) profile plot at wt DUX4 (maroon/red) or DUX4-IGH (blue/teal) peaks on direct targets in the absence (left) or presence (right) of Dox. Wilcoxon rank sum test to compare the CUTAC signal between −Dox and +Dox conditions. See also fig. S1.

Together, our results indicate that wt DUX4 and DUX4-r regulate mutually exclusive gene sets despite displaying the same DNA binding domain and sequence specificity, with only wt DUX4 behaving as a pioneer factor. Notably, only DUX4-r is capable of activating the expression of genes involved in leukemia, with an extensive overlap with the DUX4-r B-ALL patient gene signature.

DUX4-r activates cell adhesion and migration, promoting cell proliferation

Several direct DUX4-r target genes (27%) encode for surface proteins involved in cell adhesion or migration (table S3). Adhesion molecules are critical regulators of leukemia development and chemoresistance, being involved in cell survival, differentiation, proliferation, and migration (*30*, *31*). Bone marrow (BM) is considered a

protective niche to ALL cells (*32*), which contributes to chemotherapy resistance (*33*). To directly address the functional relevance of our findings, we tested the adhesive and migratory properties of DUX4-r inducible REH cells using human BM stroma cells (HS-5). To facilitate cell tracking, REH cells were modified by introducing a constitutively expressed enhanced green fluorescent protein (eGFP) reporter (Fig. 2A). As shown in Fig. 2 (B and C), induction of both DUX4-r variants significantly increased REH cell adhesion and transwell migration to HS-5 stromal cells as compared to the uninduced control. Adhesion and migration were unaffected by inducing EV or wt DUX4 (Fig. 2, B and C). Accordingly, the knockdown of endogenous DUX4-r in NALM6 cells, a B-ALL cell line that expresses DUX4-r and depends on it for its proliferation (*5*), decreased cell adhesion and migration as compared to control

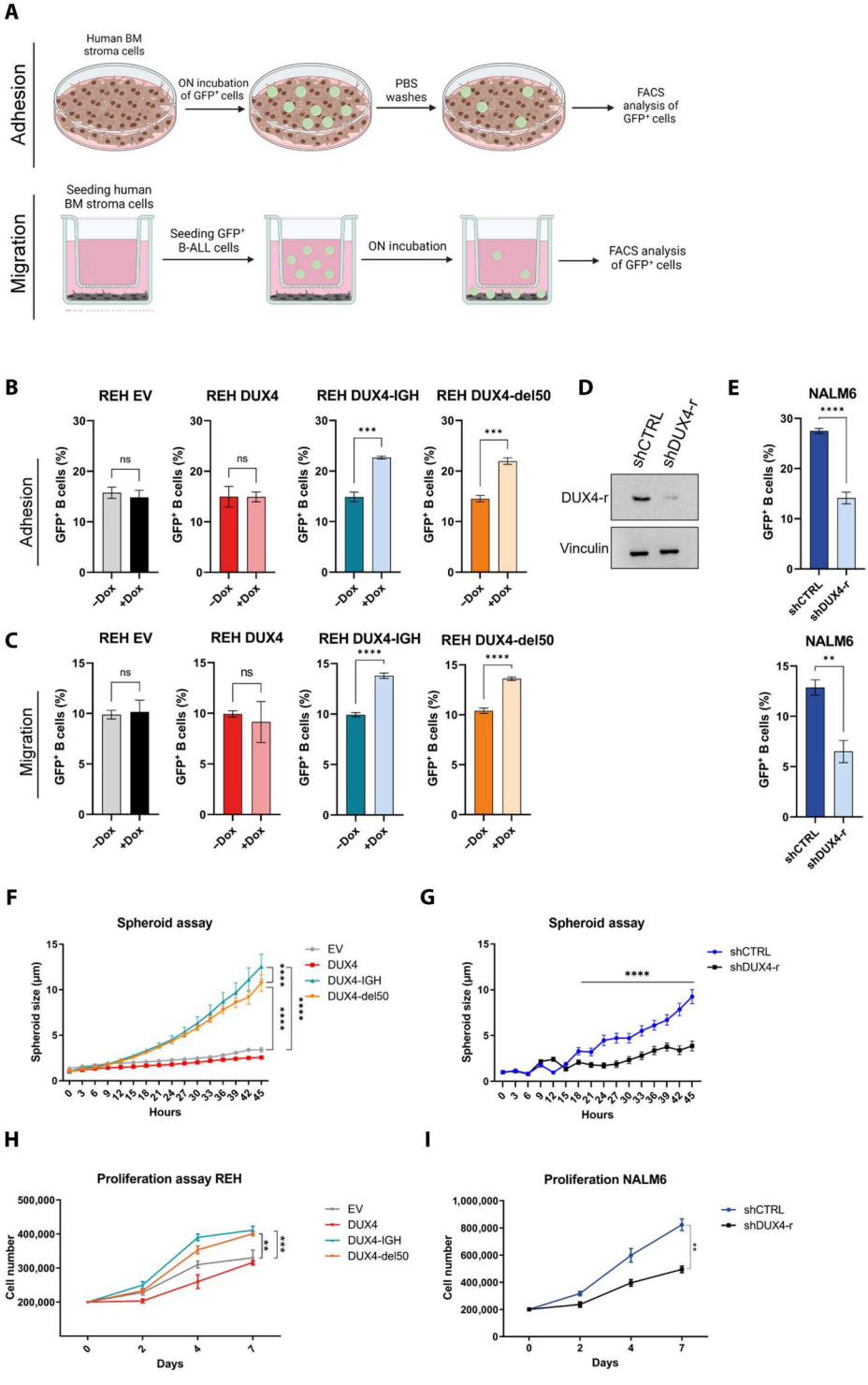

Fig. 2. DUX4-r activates cell adhesion and migration, promoting cell proliferation. (**A**) HS5 bone marrow stroma adhesion (top) and transwell migration assays (bottom) schematics. (**B**) Adhesion of REH cells expressing EV, wt DUX4, or DUX4-r variants in −Dox/+Dox. Bar plots represent the means ± SD of three biological replicates ($n = 3$). Student's t test. ***$P \leq 0.001$; ****$P \leq 0.0001$. (**C**) Migration of REH cells expressing EV, wt DUX4, or DUX4-r variants in −Dox/+Dox. Bar plots represent the means ± SD of three independent biological replicates ($n = 3$). Student's t test. **$P \leq 0.01$ and ****$P \leq 0.0001$. (**D**) Western blot validation of DUX4-r knockdown in NALM6 cells using vinculin as a loading control. (**E**) Adhesion (top) and migration (bottom) of control (shCTRL) or DUX4-r (shDUX4-r) knockdown NALM6 cells. Bar plots represent the means ± SD of three independent biological replicates ($n = 3$). Student's t test. **$P \leq 0.01$; ****$P \leq 0.0001$. (**F**) Real-time tumor spheroid quantification of REH cells expressing the indicated transgene. Curves represent the means ± SD of six biological replicates ($n = 6$). Two-way analysis of variance (ANOVA) with Bonferroni's correction to compare curves at the 45-hour time point. ****$P \leq 0.0001$. (**G**) Real-time tumor spheroid quantification of control (shCTRL) or DUX4-r (shDUX4-r) knockdown NALM6 cells. Curves represent the means ± SD of six biological replicates ($n = 6$). Two-way ANOVA with Bonferroni's correction to compare curves at the 45-hour time point. ****$P \leq 0.0001$. (**H**) Proliferation of REH cells expressing EV, wt DUX4, DUX4-IGH, or DUX4-del50. Curves represent the means ± SD of three independent biological replicates ($n = 3$). One-way ANOVA. **$P \leq 0.01$; ***$P \leq 0.001$. (**I**) Proliferation assay of control (shCTRL) or DUX4-r (shDUX4-r) knockdown NALM6 cells. Curves represent the means ± SD of three independent biological replicates ($n = 3$). One-way ANOVA. *$P \leq 0.05$. See also fig. S2. ns, not significant; PBS, phosphate-buffered saline; FACS, fluorescence-activated cell sorting.

knockdown cells (Fig. 2, D and E), strengthening our conclusion that DUX4-r promotes cell adhesion and migration.

Tumor spheroids are a useful tool for preclinical cancer research (*34*). Usually, spheroid formation is induced by the use of extracellular matrix/scaffold components and/or treatments forcing the cells to aggregate. Unexpectedly, we found that DUX4-r–expressing REH cells migrate toward each other and perform homotypic adhesion, generating areas of high cellular density which result in the spontaneous formation of spheroids in liquid cell culture (fig. S2A). Real-time quantification of live cells demonstrated significantly enhanced tumor spheroid formation by DUX4-IGH– and DUX4-del50–expressing cells compared to EV or wt DUX4 (Fig. 2F). Tumor spheroid formation by NALM6 cells was significantly reduced upon the knockdown of endogenous DUX4-r (Fig. 2G and fig. S2B).

When grown in 5% fetal bovine serum (FBS) instead of the usual 20% FBS, we found that DUX4-r–expressing cells proliferated significantly more than EV or wt DUX4 cells (Fig. 2H). Accordingly, the proliferation of NALM6 cells was significantly reduced by the knockdown of endogenous DUX4-r (Fig. 2I) in agreement with (*5*).

Collectively, these results indicate that DUX4-r directly activates the expression of cell adhesion molecules mediating the ability of DUX4-r–expressing cells to migrate and perform homotypic adhesion and stimulates cell proliferation.

DUX4-r transcriptional activity is prominent in a B cell setting

To validate our RNA-seq results, we selected representative DUX4-r direct targets consistently up-regulated in DUX4-r B-ALL patients. Reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis using RNA extracted from induced REH cells confirmed the ability of DUX4-r and wt DUX4 to activate mutually exclusive target genes (Fig. 3A). To further characterize the transcriptional activity of DUX4-r variants, we generated Dox-inducible human embryonic kidney 293 (HEK) cells with EV, wt DUX4, DUX4-IGH, or DUX4-del50. RT-qPCR analysis revealed that, despite similar transgene induction levels, DUX4-r variants were unable to activate the expression of their direct targets in HEK cells, while wt DUX4 efficiently drove the expression of its direct target genes (Fig. 3B). Similar outcomes were also obtained using human T-ALL Jurkat cells (Fig. 3C).

To begin exploring the molecular determinants of DUX4-r transcriptional inefficiency in non–B cells, we performed CUT&Tag for wt DUX4 and DUX4-IGH in HEK cells. In line with its comparable transcriptional proficiency in different cellular settings, wt DUX4 was associated with its direct targets similarly in HEK and REH cells (Fig. 3D). Instead, DUX4-IGH association to its target genes was significantly higher in REH compared to HEK cells (Fig. 3D). Overall, our results indicate that DUX4-r genomic association and transactivation ability are prominent in B cells.

GTF2I is a DUX4-r–selective interactor, which associates to DUX4-r direct targets

Given the above results, we hypothesized that the activity of DUX4-r requires a cofactor preferentially expressed in REH cells. To test this, we took advantage of the fact that our inducible DUX4 versions are fused to a streptavidin-binding peptide and a hemagglutinin (Strep-HA) double affinity tag. To identify factors selectively associated with DUX4-r variants, we performed tandem affinity purification using nuclease-treated and precleared nuclear extracts under high stringency followed by quantitative mass spectrometry (TAP-MS) to reduce nonspecific background binding and identify tight interactors. This resulted in the identification of 70 proteins specifically interacting with wt DUX4 and/or DUX4-r (table S4). Of importance, wt DUX4 and DUX4-r signals were comparable across the different TAP-MS, indicating that the differential protein association with DUX4 or DUX4-r variants was not due to different efficiency of wt DUX4 or DUX4-r affinity purification (fig. S3A). To determine selective DUX4-r interactors preferentially expressed in REH cells, we analyzed by quantitative MS the nuclear proteomes of REH, HEK, and Jurkat cells. Subsequently, we filtered our results based on the following: (i) enrichment of DUX4-r associated proteins over the EV control, (ii) enrichment of proteins in DUX4-r over wt DUX4, (iii) DUX4-r interactor enrichment in REH over HEK and Jurkat cells. This approach identified 12 proteins selectively interacting with DUX4-IGH and DUX4-del50, which were more abundant in REH cells (fig. S3B and table S4). In line with the fact that the DUX4 CTD is deleted in DUX4-r variants, the histone acetyltransferase EP300 was identified as wt DUX4–selective interactor (fig. S3C), thus providing a possible molecular explanation for DUX4-r lack of pioneer activity (Fig. 1, G and H). Among the selective DUX4-r interactors preferentially expressed in REH cells, the most enriched in both DUX4-IGH and DUX4-del50 TAP-MS was general transcription factor IIi (GTF2I/TFII-I) (Fig. 4A). GTF2I is a transcription factor regulated by several signaling pathways, including the pre-B cell receptor (*35*, *36*). *GTF2I* translocations are associated with ALL and acute promyelocytic leukemia (*37*, *38*). Moreover, a recurrent *GTF2I* mutation is the most frequent oncogenic driver in thymic epithelial tumors (*39*).

New affinity purifications followed by immunoblotting confirmed GTF2I-selective interaction with both DUX4-r variants, but not wt DUX4 (Fig. 4B). Moreover, immunoblotting of total protein extracts confirmed significantly higher levels of GTF2I in REH compared to HEK and Jurkat cells (Fig. 4, C and D). Coimmunoprecipitation (co-IP) confirmed the interaction between the endogenous DUX4-r and GTF2I in NALM6 cells (fig. S3D).

We performed CUT&Tag for GTF2I in HEK and REH cells lacking expression of DUX4-r and in REH cells expressing DUX4-IGH. For direct DUX4-r targets, we observed a significantly higher GTF2I signal in REH compared to HEK cells, which was further increased upon DUX4-IGH expression (Fig. 4E). No significant difference in GTF2I chromatin association was observed in REH cells expressing DUX4-IGH as compared to EV-expressing REH cells, indicating that DUX4-IGH expression enhances GTF2I binding selectively to DUX4-r target genes (fig. S3E).

Intrigued by the correlation between GTF2I expression levels and DUX4-r transcription activation ability in cell lines, we re-analyzed DUX4-r B-ALL patient datasets (*6*) to determine the expression levels of *GTF2I* and the level of activation of the DUX4-r core gene set in each patient. Despite all patients analyzed overexpressed the DUX4-r core gene set compared to healthy B cells or other ALL subtypes (*6*), they do so at variable levels (Fig. 4F). Notably, we found a significant and positive correlation between the expression levels of *GTF2I* and the relative activation of DUX4-r core genes (Fig. 4, F and G) strongly supporting the relevance of our findings.

Together, the above results identify GTF2I as the strongest, selective DUX4-r interactor, which discriminates DUX4-r targets from non-target genomic regions and whose expression levels

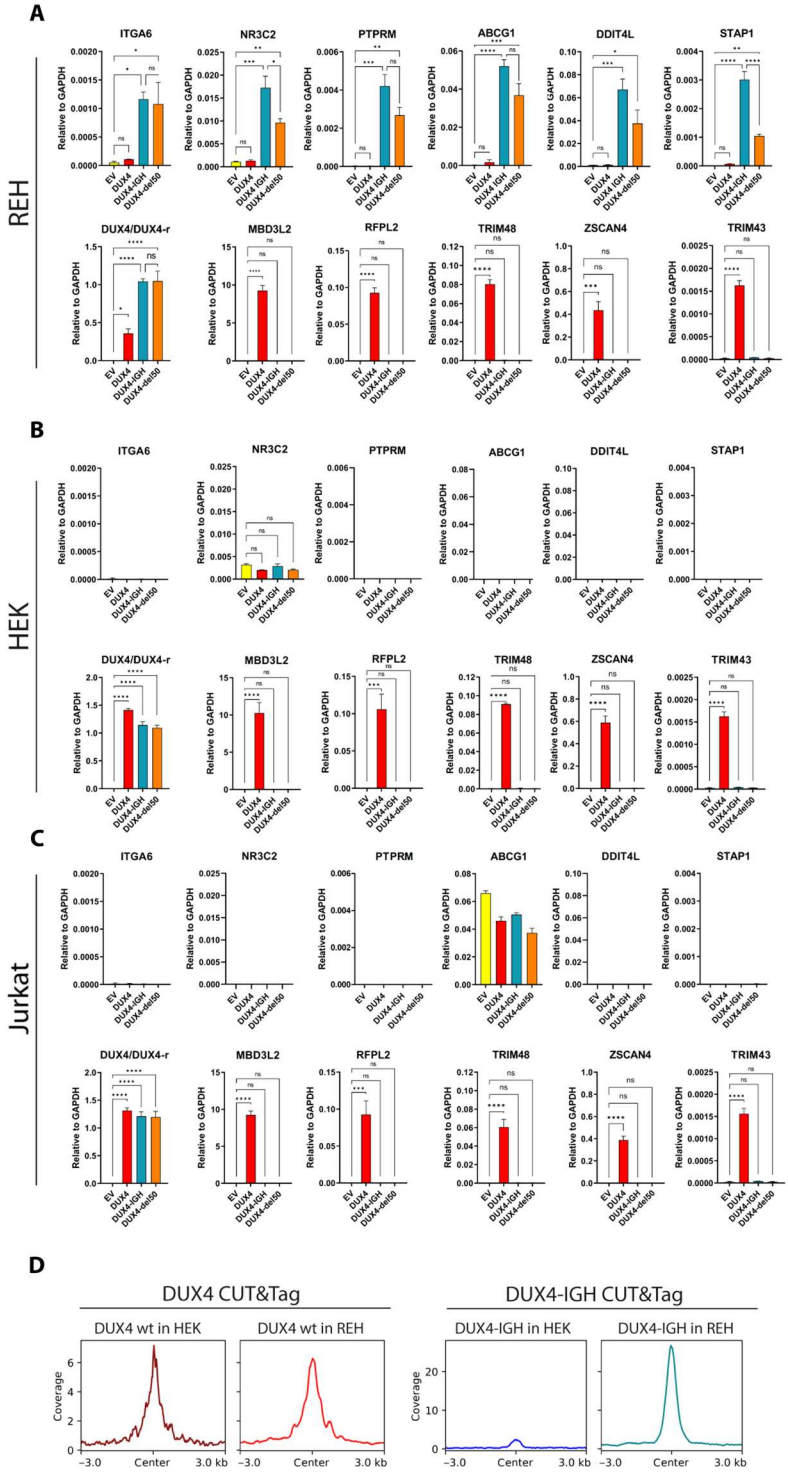

Fig. 3. DUX4-r activity is prominent in B-ALL cells. (**A**) RT-qPCR expression analysis of the indicated DUX4-r (top) or wt DUX4 (bottom, right) direct target genes in REH cells expressing EV, wt DUX4, DUX4-IGH, or DUX4-del50. Relative levels of wt DUX4/DUX4-r are also shown (bottom, left). Relative mRNA levels were normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Bar plots represent the average of three independent biological replicates ($n = 3$). Error bars represent means ± SD. One-way ANOVA with multiple comparisons. *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$. (**B**) RT-qPCR analysis of the same genes as in (A) in HEK cells expressing EV, wt DUX4, DUX4-IGH, or DUX4-del50. Relative mRNA levels were normalized to GAPDH. Bar plots represent the average of three independent biological replicates ($n = 3$). Error bars represent means ± SD. One-way ANOVA with multiple comparisons. ***$P \leq 0.001$; ****$P \leq 0.0001$. (**C**) RT-qPCR analysis of the same genes as in (A) and (B) in Jurkat cells expressing EV, wt DUX4, DUX4-IGH, or DUX4-del50. Relative mRNA levels were normalized to GAPDH. Bar plots represent the average of three independent biological replicates ($n = 3$). Error bars represent means ± SD. One-way ANOVA with multiple comparisons. ***$P \leq 0.001$; ****$P \leq 0.0001$. (**D**) CUT&Tag profile plot of genomic enrichment at (left) wt DUX4 on its direct transcriptional targets in human embryonic kidney 293 (HEK) (dark red) and REH (red) cells and (right) DUX4-IGH on its direct transcriptional targets in HEK (blue) and REH (teal) cells. Wilcoxon rank sum test was performed to compare the signal of wt DUX4 or DUX4-IGH between HEK and REH cells.

significantly correlate with the level of activation of DUX4-r targets in patients with B-ALL.

DUX4-r biological activity depends on the cellular availability of GTF2I

Given the above results, we next asked whether genetic GTF2I targeting could affect the ability of DUX4-r to activate its direct target genes. To test this, we performed GTF2I knockdown in REH cells expressing ectopic DUX4-IGH or in NALM6 expressing endogenous DUX4-r (Fig. 5A). In line with a GTF2I requirement for DUX4-r transcriptional activity, we observed significant down-regulation of DUX4-r target genes (Fig. 5B). GTF2I silencing had no significant effect on the activation of wt DUX4 targets (fig. S4A).

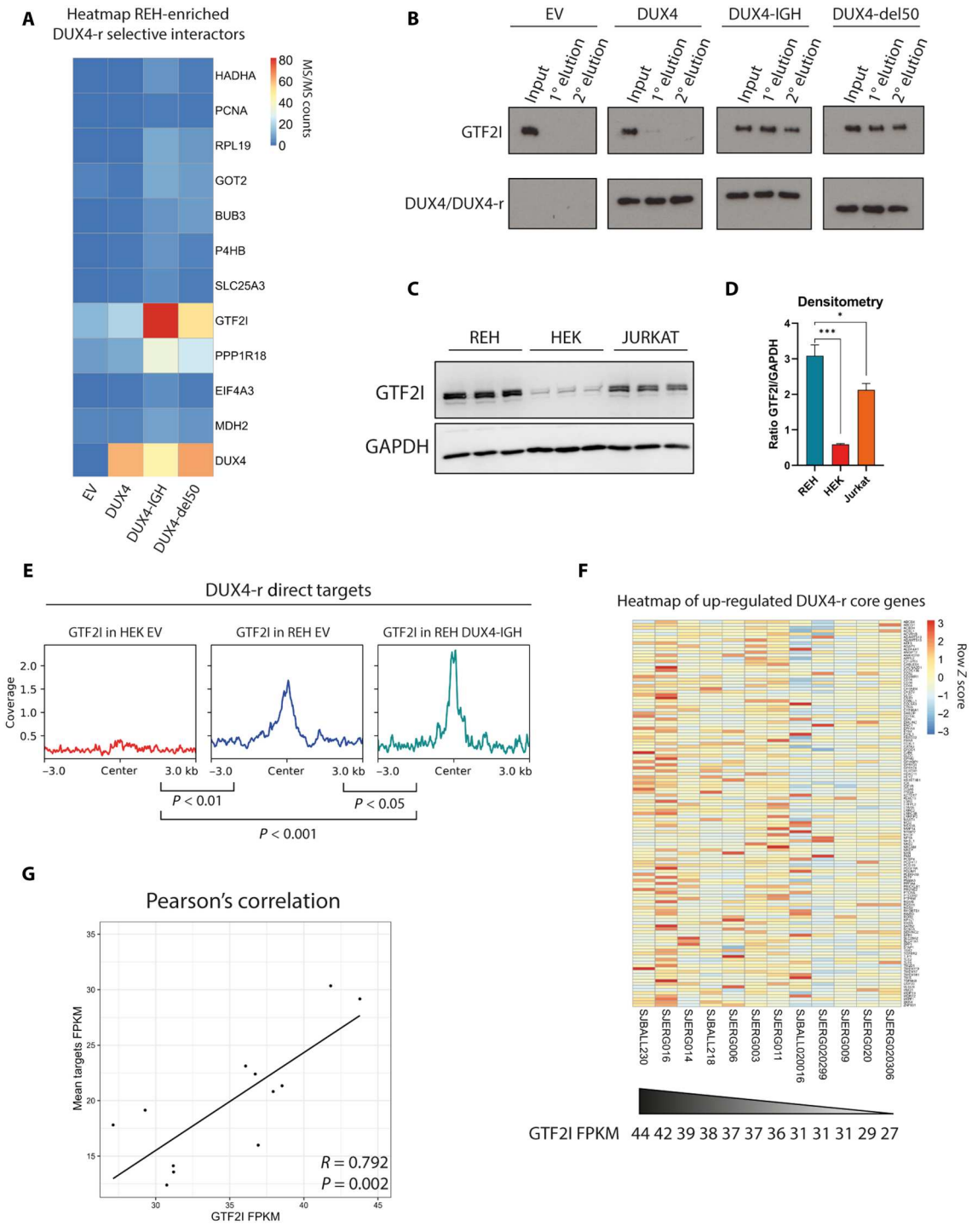

Fig. 4. DUX4-r selectively interacts with GTF2I. (**A**) Heatmap showing DUX4-r–selective interactors which are more expressed in REH cells. Protein levels were expressed as MS spectral counts. The heatmap reports unscaled MS/MS counts. (**B**) Tandem affinity purification followed by immunoblotting validation of selective GTF2I interaction with DUX4-IGH and DUX4-del50. Five percent of input and first elution and 10% of second elution were loaded and incubated with antibodies specific for GTF2I or wt DUX4 and DUX4-r. (**C**) Western blot validation of nuclear protein levels of GTF2I in REH, Jurkat, or REH cells using GAPDH as a loading control. (**D**) Densitometric analysis of signals in Fig. 3D using the ImageLab software (Bio-Rad, ver. 6.1). Bar plots represent the mean of three independent biological replicates ($n = 3$). Error bars represent ± SD. One-way ANOVA with multiple comparisons against REH. *$P \leq 0.05$; ***$P \leq 0.001$. (**E**) CUT&Tag profile plot of GTF2I enrichment on DUX4-r direct targets in HEK EV (red), REH EV (blue), or REH DUX4-IGH (teal) cells. Wilcoxon ranks sum test was performed to compare the signal of GTF2I between HEK EV, REH EV, and REH DUX4-IGH cells. (**F**) Heatmap showing the expression levels (FPKM, fragments per kilobase of transcript per million fragments mapped) of DUX4-r core genes in different DUX4-r B-ALL patients ordered based on decreasing GTF2I expression levels. The heatmap reports *Z* scores scaled by row. (**G**) Pearson's correlation plot showing the positive correlation between GTF2I expression levels and the mean activation of DUX4-r core genes in different DUX4-r B-ALL patients. The correlation coefficient (*R*) and *P* value are indicated. See also fig. S3.

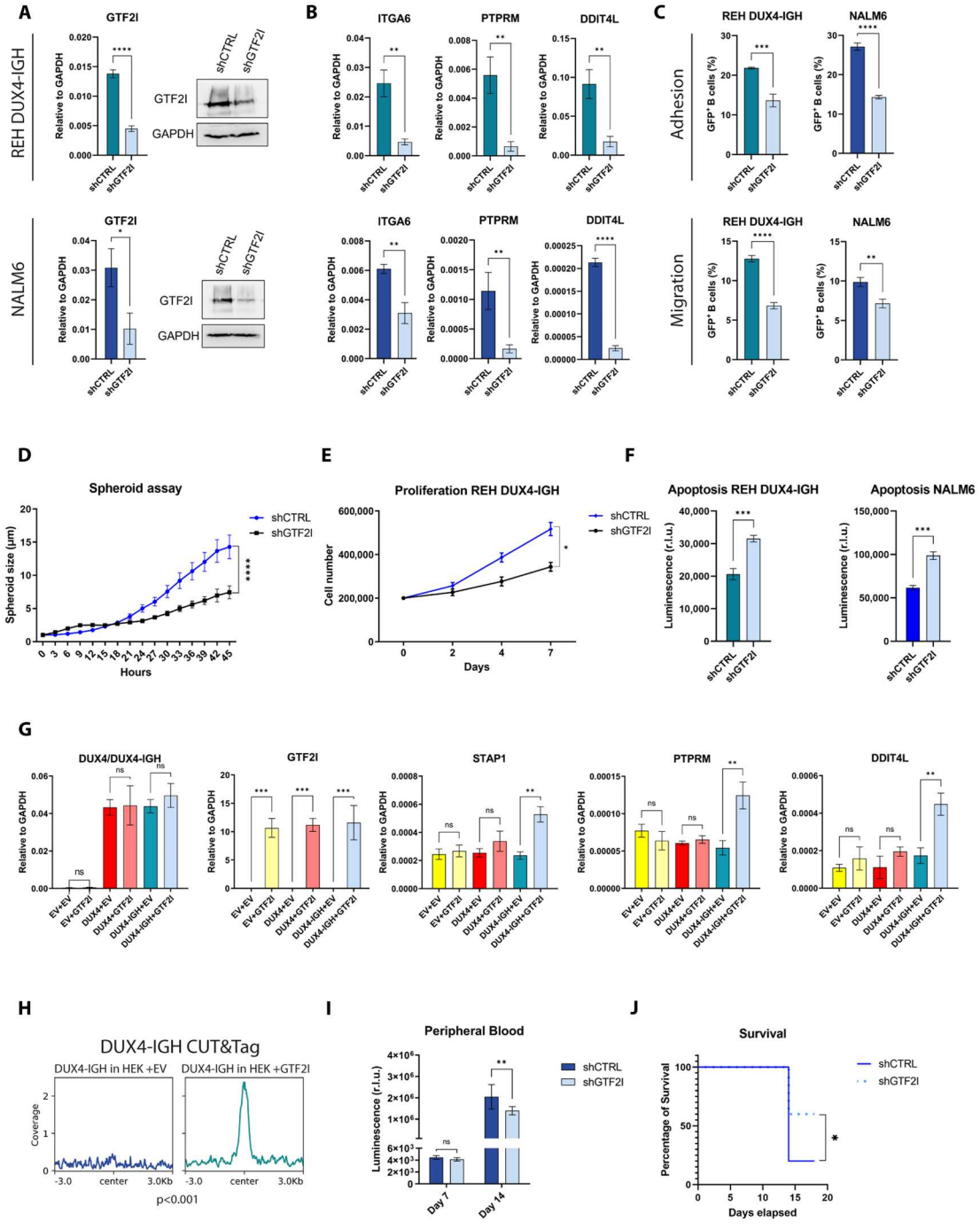

Fig. 5. GTF2I is required for DUX4-r biological activity in vitro and in vivo. (**A**) RT-qPCR (left) and immunoblot (right) of GTF2I knockdown in REH DUX4-IGH (top) or NALM6 cells (bottom). Bar plots show means ± SD ($n = 3$). Student's t test. *$P \leq 0.05$ and ****$P \leq 0.0001$. (**B**) RT-qPCR of the indicated direct DUX4-r targets in REH DUX4-IGH (top) and NALM6 (bottom) CTRL or GTF2I knockdown. Bar plots show means ± SD ($n = 3$). One-way ANOVA with multiple comparisons. **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$. (**C**) Adhesion (top) and migration (bottom) of REH DUX4-IGH (left) and NALM6 (right) in shCTRL or shGTF2I cells. Bar plots show means ± SD ($n = 3$). Student's t test. **$P \leq 0.01$; ***$P \leq 0.001$; ****$P \leq 0.0001$. (**D**) Real-time tumor spheroid quantification of REH DUX4-IGH plus CTRL or GTF2I knockdown. Curves show means ± SD ($n = 6$). Two-way ANOVA with Bonferroni's correction to compare curves at the 45-hour time point. ****$P \leq 0.0001$. (**E**) Proliferation of REH DUX4-IGH plus CTRL or GTF2I knockdown. Curves represent the means ± SD ($n = 3$). One-way ANOVA. *$P \leq 0.05$. (**F**) Apoptosis in CTRL or GTF2I knockdown REH DUX4-IGH (left) and NALM6 cells (right). Bar plots show means ± SD ($n = 4$). Student's t test. ***$P \leq 0.001$ (r.l.u., relative luminescence unit). (**G**) RT-qPCR for GTF2I and the indicated DUX4-r direct targets in HEK cells expressing EV, wt DUX4, or DUX4-IGH plus EV or GTF2I. Bar plots show means ± SD ($n = 3$). One-way ANOVA with multiple comparisons was performed. **$P \leq 0.01$; ***$P \leq 0.001$. (**H**) CUT&Tag profile plot of DUX4-IGH on its direct targets in HEK DUX4-IGH cells transfected with EV (blue) or GTF2I (teal). Wilcoxon ranks sum test to compare the signal of DUX4-IGH between EV and GTF2I conditions. (**I**) Leukemia expansion in peripheral blood of NSG mice 7 and 14 days after NALM6-Lucia cell transplantation. Bar plots show the mean Luciferase signal ± SD of shCTRL or shGTF2I NALM6-Lucia mice. ($n = 5$). Two-way ANOVA. **$P \leq 0.01$. (**J**) Kaplan-Meier survival analysis of the above mice. *$P \leq 0.05$. See also fig. S4.

Next, we investigated whether GTF2I is required for the cellular phenotypes induced by DUX4-r. GTF2I knockdown significantly reduced both the adhesion and transwell migration selectively of DUX4-r–expressing B-ALL cells (Fig. 5C and fig. S4B), leading to a significantly decreased tumor spheroid formation selectively in DUX4-r–expressing cells (Fig. 5D and fig. S4, C to E). GTF2I genetic down-regulation significantly reduced the proliferation of REH expressing ectopic DUX4-IGH (Fig. 5E) or NALM6 expressing endogenous DUX4-r (fig. S4F), significantly increasing their cell death (Fig. 5F).

The above results indicate that GTF2I is required for DUX4-r activity. To evaluate sufficiency, we then investigated the effects of supplementing HEK cells with GTF2I. We introduced vectors encoding for *GTF2I* or a control empty vector (*EV*) in HEK cells together with *EV*, *wt DUX4*, or *DUX4-IGH*. *GTF2I* was expressed at comparable levels in all cell lines and did not affect *wt DUX4* or *DUX4-IGH* expression levels (Fig. 5G). Notably, in HEK cells, ectopically expressed GTF2I reached levels comparable to endogenous GTF2I in REH cells (fig. S4G). GTF2I expression in HEK cells significantly increased target gene activation by DUX4-IGH (Fig. 5G), while wt DUX4 activity was unaffected (fig. S4H).

To investigate the molecular determinants of the observed DUX4-r transcriptional "rescue" by GTF2I, we performed CUT&Tag in HEK cells expressing DUX4-IGH in the presence/absence of transfected GTF2I. In the absence of GTF2I, DUX4-IGH displayed negligent chromatin association genome-wide (fig. S4I). Instead, upon GTF2I supplementation, DUX4-IGH displayed a significantly enhanced ability to associate with its target gene loci in HEK cells (Fig. 5H). In line with the fact that we identified GTF2I as a selective DUX4-r interactor, wt DUX4 association to its target loci was unaffected by GTF2I overexpression (fig. S4J).

Given the above results, we explored the implications of the DUX4-r/GTF2I transcriptional axis in leukemia. To this aim, we took advantage of a modified version of NALM6 cells expressing Gaussia luciferase Lucia, which is actively secreted by cells and detectable in the blood facilitating tumor progression monitoring in mice (*40*). NALM6-Lucia cells were transduced with lentiviral vectors encoding for a short hairpin RNA (shRNA) against *GTF2I* or a control shRNA (sh*CTRL*). NSG (NOD.Cg-Prkdcscid Il2rgtm1WjI/SzJ) mice were then transplanted with knockdown cells and disease progression was followed over time. One week after transplantation, we observed a similar luciferase signal between groups, indicating a comparable leukemia cell engraftment (Fig. 5I). Nonetheless, after 2 weeks, we detected significantly lower leukemia cell expansion in peripheral blood of *GTF2I* knockdown animals compared to controls (Fig. 5I). This translated into significantly higher survival rates in mice transplanted with *GTF2I* shRNA compared to control shRNA leukemia cells (Fig. 5J). At sacrifice, *GTF2I* knockdown transplants showed also a significantly lower leukemia burden in both BM and spleen compartments (fig. S4K). We found that *GTF2I* shRNA cells harvested from the mice at sacrifice displayed a milder reduction in the expression level of *GTF2I* and DUX4-r targets compared to the one that we observed in vitro (compare Fig. 5, A and B, with fig. S4L). These results are in line with the possibility of a negative selection against cells with a strong decrease with GTF2I expression in vivo. All in all, our findings indicate that GTF2I is required for DUX4-r transcriptional and leukemogenic activities.

GTF2I pharmacological targeting inhibits DUX4-r leukemogenic activity

Recently, a high-throughput screening identified the histone deacetylase inhibitor vorinostat/suberylanilide hydroxamic acid (SAHA) as a selective repressor of GTF2I expression (*41*). By performing a SAHA dose-response analysis in REH cells, we found that treatment with as low as 50 nM SAHA was sufficient to down-regulate GTF2I consistently and significantly at both mRNA and protein levels (fig. S5A). Similarly to GTF2I knockdown, SAHA treatment significantly reduced the ability of DUX4-IGH to activate its target genes (Fig. 6A). SAHA caused no significant effect on the ability of wt DUX4 to activate its targets (fig. S5B). In line with a GTF2I requirement for DUX4-r genomic association, by CUT&Tag, we found a significantly lower DUX4-IGH enrichment at its direct targets in SAHA-treated compared to control-treated cells (Fig. 6B).

We next assessed the impact of SAHA treatment on the cellular phenotypes caused by DUX4-r. Similar to what was observed with DUX4-r or GTF2I genetic silencing, SAHA significantly reduced cell adhesion, transwell migration, tumor spheroid formation, and cell proliferation of both REH cells expressing ectopic DUX4-IGH and NALM6 cells expressing endogenous DUX4-r (Fig. 6, C to E, and fig. S5, C to E). Consistently, SAHA significantly increased cell death both of REH expressing ectopic DUX4-IGH and NALM6 expressing endogenous DUX4-r (Fig. 6F). Proliferation and survival of REH cells expressing EV control or wt DUX4 were unaffected by SAHA treatment (Fig. 6, E and F, and fig. S5F), in line with the possibility that DUX4-r expression sensitizes leukemia cells to SAHA treatment.

To test the therapeutic potential of SAHA treatment for DUX4-r B-ALL, NALM6-Lucia cells were transplanted in NSG mice. One week after transplantation, when comparable NALM6 engraftment was observed in all mice (Fig. 6G), animals were randomly divided into two cohorts: one was treated with SAHA and the other with dimethyl sulfoxide (DMSO), as a control. After 1 week (2 weeks from cell transplantation), we detected significantly lower leukemia cell expansion in the peripheral blood of SAHA-treated compared to control mice (Fig. 6G). This translated into significantly higher survival rates in mice treated with SAHA compared to controls (Fig. 6H). Notably, also in this case, we found a milder reduction in the expression level of *GTF2I* and DUX4-r targets in the cells harvested from the mice at sacrifice compared to the one that we observed in vitro (compare Fig. 6A with fig. S5G), further supporting the possibility that cells with a strong decrease in GTF2I expression are negatively selected in vivo. Collectively, our results indicate that SAHA recapitulates the effects of GTF2I knockdown on DUX4-r leukemogenic activity.

DISCUSSION

By combining transcriptomics, genomics, and proteomics with functional assays in vitro and in vivo, our results elucidate the molecular mechanism of B-ALL caused by DUX4-r. We propose that the chromosomal rearrangement producing DUX4-r leads to both a loss and a gain of function in the leukemogenic transcription factor. Deletion of the CTD causes a loss of interaction with CBP/EP300 transcriptional coactivators leading to an inability to bind and activate repressed chromatin. On the other hand, DUX4-r gains the ability to interact with the transcription factor GTF2I, which

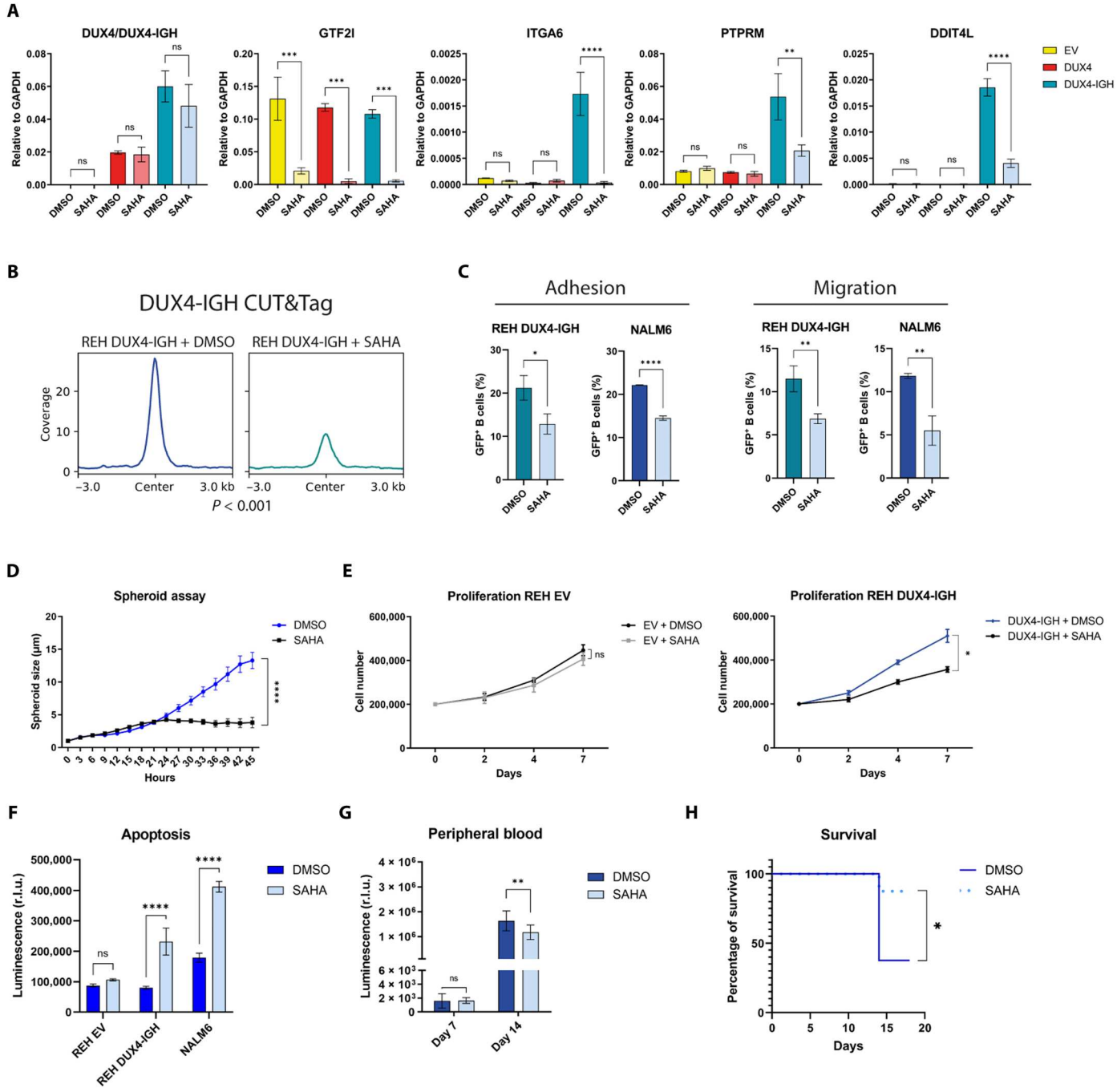

Fig. 6. GTF2I pharmacological targeting inhibits DUX4-r leukemogenic activity. (**A**) RT-qPCR expression of wt DUX4, DUX4-IGH, GTF2I, and the indicated DUX4-IGH direct targets upon suberylanilide hydroxamic acid (SAHA) treatment in EV, wt DUX4, or DUX4-IGH REH cells. Bar plots show means ± SD of three independent biological replicates ($n = 3$). One-way ANOVA with multiple comparisons. $^{**}P \leq 0.01$; $^{***}P \leq 0.001$; $^{****}P \leq 0.0001$. (**B**) CUT&Tag profile plot of DUX4-IGH on its direct targets in REH DUX4-IGH cells treated with dimethyl sulfoxide (DMSO) (teal) or SAHA (blue). Wilcoxon rank sum test to compare the DUX4-IGH signal between DMSO and SAHA. (**C**) Adhesion (left) and migration (right) performed in REH DUX4-IGH (teal) or NALM6 (blue) cells treated with DMSO or SAHA. Bar plots show means ± SD of three independent biological replicates ($n = 3$). Student's t test. $^{*}P \leq 0.05$; $^{**}P \leq 0.01$; $^{****}P \leq 0.0001$. (**D**) Real-time tumor spheroid quantification in REH DUX4-IGH cells after DMSO or SAHA treatment. Curves show means ± SD of six biological replicates ($n = 6$). Two-way ANOVA with Bonferroni's correction to compare curves at the 45-hour time point. $^{****}P \leq 0.0001$. (**E**) Proliferation of control REH cells expressing EV (left) or DUX4-IGH (right) treated with DMSO or SAHA. Curves show means ± SD of three biological replicates ($n = 3$). One-way ANOVA. $^{*}P \leq 0.05$. (**F**) Apoptosis in REH cells expressing EV or DUX4-IGH, and NALM6 cells treated with DMSO or SAHA. Bar plots show means ± SD of three independent biological replicates ($n = 3$). Student's t test. $^{****}P \leq 0.0001$. (**G**) Leukemia expansion in peripheral blood of NSG mice at 7 and 14 days after NALM6-Lucia transplantation. Bar plots show means ± SD signal of DMSO- or SAHA-treated mice ($n = 8$). Two-way ANOVA. $^{**}P \leq 0.01$. (**H**) Kaplan-Meier survival analysis of the above mice. $^{*}P \leq 0.05$. See also fig. S5.

redirects DUX4-r toward the binding and activation of targets involved in leukemogenesis.

We found that wt DUX4 and DUX4-r bind to and activate highly different gene sets, although they share the same DNA binding domain and recognize the same DNA consensus. Our results differ from those of a previous study (*9*), which reported a large overlap between wt DUX4 and DUX4-r genomic targets. Those authors performed ChIP-seq only for DUX4-IGH in NALM6 cells and compared their results to a previously published ChIP-seq dataset of wt DUX4 performed by others in induced pluripotent stem cells (*9*). To perform a side-by-side comparison and enrich for primary targets, we instead used REH cells lacking endogenous wt DUX4/DUX4-r expression, induced wt DUX4/DUX4-r at levels comparable to those of DUX4-r B-ALL patients and performed CUT&Tag early upon DUX4/DUX4-r induction. The technical differences between the previously published and our study are probably responsible for the different results.

While wt DUX4 shows essentially the same transcriptional activity in every cell type that we and others have tested (*10*–*12*, *21*, *23*, *24*, *42*, *43*), we found that the activity of DUX4-r is intimately linked to GTF2I cellular availability. In DUX4-r B-ALL patients, the magnitude of activation of the DUX4-r signature positively correlates with GTF2I expression levels. At the genomic level, DUX4-r enrichment at its targets is significantly higher in REH compared to HEK cells, in line with GTF2I expression levels. Also, GTF2I down-regulation leads to significantly lower DUX4-r chromatin association and target activation. Similarly, in HEK cells, which express relatively low GTF2I endogenous levels, DUX4-r displays negligent chromatin association genome-wide. Notably, ectopic GTF2I expression in HEK cells is sufficient to drive DUX4-r recruitment selectively to its target genes. It thus is tempting to speculate that the DNA binding activity of DUX4-r is not solely determined by its DNA binding domain, but that it relies on GTF2I for target gene definition.

In REH cells, we found that GTF2I is already associated with DUX4-r targets in its absence. Hence, GTF2I pre-marking in the vicinity of DUX4 binding sites could contribute to DUX4-r target gene definition. Our results suggest that also GTF2I is affected by DUX4-r. GTF2I enrichment at DUX4-r targets is significantly enhanced upon DUX4-r expression. While the precise mechanism at the basis of the apparent synergy between DUX4-r and GTF2I remains to be determined, it is tempting to speculate that DUX4-r binding tethers GTF2I to target loci and/or that the two factors stabilize each other chromatin association.

Previous work reported that treatment with the HDAC inhibitor SAHA down-regulates GTF2I expression (*41*). We confirmed and extended this result by showing that SAHA treatment in the nanomolar range is sufficient to down-regulate GTF2I at both mRNA and protein levels. Notably, SAHA blocks the activation of DUX4-IGH targets and the downstream phenotypic effects as effectively as direct GTF2I or DUX4-r knockdown. We found that DUX4-r confers SAHA sensitivity to expressing leukemia cells in line with the possibility that DUX4-r B-ALL cells are addicted to the DUX4-r/GTF2I circuit.

Despite that our results provide a rationale for the development of a pharmacological treatment for DUX4-r B-ALL, GTF2I targeting was not completely "curative." Future studies are required to establish if interfering with GTF2I is not enough to completely block DUX4-r leukemic progression or if negative selection against cells with a strong decrease in GTF2I expression was responsible for the partial in vivo efficacy of GTF2I targeting.

While SAHA is already approved for the treatment of cutaneous T cell lymphoma (*44*) and is in a clinical trial for several other types of cancers (*45*, *46*), further studies concerning the molecular mechanisms of GTF2I down-regulation by SAHA could identify additional targets to develop therapeutic strategies personalized to DUX4-r B-ALL. Overall, our findings indicate that DUX4-r acquires a neomorphic oncogenic activity depending on GTF2I that can be targeted for therapeutic purposes in B-ALL.

MATERIALS AND METHODS

Plasmids and cloning

Lentiviral vectors encoding inducible, N-terminally tagged DUX4, DUX4-IGH, and DUX4-del50 were generated using Gateway technology (Invitrogen) starting from the destination vector pCW57.1, a gift from D. Root (Addgene, plasmid #41393; http://n2t.net/addgene:41393; RRID:Addgene_41393). The N-terminal *Twin-Strep-2xHA* (Strep-HA) double affinity tag was recovered by PCR (table S5) from pCDNA5/FRT/TO STREP-HA vector, kindly provided by G. Superti-Furga (CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria), and inserted in the *EcoRI* site of pCW57.1 generating the pCW57.1 Strep-HA vector, which was subsequently used as an EV control. Codon-altered *DUX4* open reading frame (ORF) was recovered by PCR (table S5) from pCW57.1-DUX4-CA, a gift from S. Tapscott (Addgene, plasmid #99281; http://n2t.net/addgene:99281; RRID:Addgene_99281). *DUX4-IGH* ORF was recovered by PCR (table S5) from pMXs-DUX4-IGH JALSG-003, a gift of H. Mano (*5*), using the same 5′ primer but different 3′ ones (table S5). *Wt DUX4* ORF was recovered by PCR (table S5) from pCS2-mkgDUX4, a gift from S. Tapscott (Addgene, plasmid #21156; http://n2t.net/addgene:21156; RRID:Addgene_21156). Primers were designed to carry attB1/2 recombination sites. The *DUX4-del50* construct was generated through site-directed mutagenesis starting from the *wt DUX4* construct, by introducing a premature stop codon 150–base pair (bp) upstream of the *DUX4 wt* stop codon.

The shRNA vector targeting *DUX4-r* was generated by the GeneScript gene synthesis service by introducing it into pLKO.1-hPGK-Puro lentiviral vector, a gift from B. Weinberg (Addgene, plasmid #8453; http://n2t.net/addgene:8453; RRID:Addgene_8453), the sh*DUX4-r* sequence described in (*14*). The nontarget shRNA control plasmid pLKO.1-hPGK-Puro sh*CTRL* vector was obtained from Sigma-Aldrich (catalog no. SHC016). The *GTF2I* mammalian expression vector was a gift from A. Roy (Addgene, plasmid #22148; http://n2t.net/addgene:22148; RRID:Addgene_22148). The lentiviral vector encoding for the constitutively expressed GFP reporter was a gift from E. Campeau and P. Kaufman (Addgene, plasmid #17447; http://n2t.net/addgene:17447; RRID:Addgene_17447). The constitutive shRNA vectors targeting *GTF2I* (pLV[shRNA]-EGFP:T2A:Neo-U6 > hGTF2I[shRNA#1]) and control (pLV[shRNA]-EGFP:T2A:Neo-U6 > Scramble_shRNA#1) were sourced from VectorBuilder.

Lentivirus production

Preparation of lentiviral particles was performed by calcium phosphate cotransfection of HEK293T cells with the lentiviral vector of

interest in combination with a pCMV-VSV-G lentiviral envelop, a gift from B. Weinberg (Addgene, plasmid #8454; http://n2t.net/addgene:8454; RRID:Addgene_8454), and the pCMV-dR8.2-dvpr lentiviral packaging plasmids, a gift from B. Weinberg (Addgene, plasmid #8455; http://n2t.net/addgene:8455; RRID:Addgene_8455). For each transfection, 9×10^6 HEK293T cells were plated 18 hours before transfection. Culture medium was replaced with a fresh Iscove′s modified Dulbecco′s medium (IMDM) medium (Lonza-LOBE12726F), supplemented with 10% FBS (Gibco-10270106), or tetracycline-negative FBS when transfecting inducible constructs, 5 mM L-glutamine, and 1% penicillin/streptomycin antibiotic cocktail (P/S), 2 hours before transfection. Respectively, 32, 9, and 12.5 μg of lentiviral, VSV-G, and packaging vectors were diluted in 0.06× tris-EDTA and brought to a final volume of 1125 ml. Ice-cold $CaCl_2$ (125 μl) was added and the mix was allowed to incubate at room temperature (RT) for 10 min, following the dropwise addition of 1250 ml of in-house–prepared 2× Hanks′ balanced salt solution while vortexing the mix at full speed. The transfection mix (2.5 ml) was gently added dropwise to HEK293T cells. After 16 hours, the medium was replaced with a fresh IMDM medium, and cells were incubated for 30 hours before collection of the viral supernatant. Medium was then replaced and, after 24 hours of further incubation, a second round of virus collection was performed. Viral collections were then pooled and ultracentrifuged at 20,000*g* for 2 hours at 4°C. Viral particles were then resuspended in 75 μl of Opti-MEM (Thermo Fisher Scientific, 31985-047) for each tube and stored at −80°C.

Cell culture, transfection, and transduction

REH (DSMZ-ACC22), NALM6 (DSMZ-ACC128), and JURKAT (DSMZ-ACC282) cells were maintained in RPMI 1640 (Euroclone-ECM2001L) medium supplemented with 20, 10, and 10% tetracycline-negative FBS (Euroclone, ECS0182L), respectively, and 1% P/S. HEK293 [American Type Culture Collection (ATCC), crl1573], HEK293T (ATCC, crl3216), and HS-5 (ATCC, crl11882) cells were maintained in Dulbecco′s modified Eagle′s medium supplemented with 10% FBS and 1% P/S.

Transient transfection experiments were performed in six-well plates using Lipofectamine LTX&PLUS reagent (Invitrogen, 15338-100) according to the manufacturer′s instructions. *Wt DUX4*, *DUX4-IGH*, and *GTF2I*-encoding and EV control plasmids were transfected into HEK cells using 2 μg of inducible *EV*, *wt DUX4*, or *DUX4-IGH* vectors and 250 ng of *EV* or *GTF2I* vectors plus 750 ng of control vector to balance the transfection reaction. After approximately 18 hours from transfection, Dox (1 μg/ml; Sigma-Aldrich, D9891-10G) was added to a final concentration of 1 μg/ml. RNA was collected 8 hours after Dox induction.

Transduction of suspension cells was performed by using 1×10^6 cells in 1 ml of RPMI 1640 medium containing polybrene (8 μg/ml; Sigma-Aldrich, TR-1003-G) and 100 μl of viral particles by spinoculation to increase transduction efficiency. Spinoculation was performed at 1290*g*, for 1.30 hours at 34°C, and cells were returned to the cell culture incubator for 48 hours before changing the medium.

Inducible REH cells were generated by transduction with pCW57.1 Strep-HA (EV), pCW57.1 Strep-HA codon-optimized DUX4 (wt DUX4), pCW57.1 Strep-HA DUX4-IGH (DUX4-IGH), or pCW57.1 Strep-HA DUX4-del50 (DUX4-del50) lentiviruses as described above, followed by selection with puromycin (0.25 μg/μl; Thermo Fisher Scientific, A1113803) for 2 weeks. Then, cells were maintained at puromycin (0.1 μg/μl). Versions of these cells constitutively expressing GFP were generated by transduction with the pCMV-GFP-Neo lentivirus and selected with G418 sodium salt (500 μg/ml; Sigma-Aldrich, A1720-5G) dissolved in water for 1 month. Then, cells were maintained at G418 (100 μg/ml) and puromycin (0.1 μg/ml).

NALM6 cells expressing sh*CTRL* or sh*DUX4-r* were generated by transduction with pLKO.1-hPGK-Puro lentiviruses described above. Versions of these cells constitutively expressing GFP were generated by transduction with the pCMV-GFP-Neo lentivirus as above.

Inducible REH, NALM6, or NALM6 LUCIA cells expressing sh*CTRL* or sh*GTF2I* were generated by transduction with pLV[shRNA]-EGFP:T2A:Neo-U6 lentiviruses described above. Cells were selected with G418 sodium salt (500 μg/ml) for 1 month.

RNA extraction and RT-qPCR

For $>2 \times 10^6$ cells, total RNA was extracted using the PureLink RNA Mini Kit (Life Technologies, 12183025), following the manufacturer′s instructions. For processing fewer cells, total RNA was extracted using the NucleoSpin RNA XS kit (Macherey-Nagel-FC140955N), following the manufacturer′s instructions. Reverse transcription was performed with SuperScript III First-Strand Synthesis SuperMix (Life Technologies, 11752-250), following the manufacturer′s instructions. As input, 1 μg of DNA-free RNA was used, or the same amount of RNA for each sample in case of less abundant samples. Quantitative real-time PCR was performed with the iTaq Universal SYBR Green Supermix (Bio-Rad, 1725122), using 5 or 3 μl of a 1:10 dilution of cDNA and 0.4 or 0.2 μM concentration of primers in the 96- or 384-well PCR plates, respectively. The real-time PCR program used was 10 min at 95°C, 30 s at 95°C, 30 s at 58°C, 30 s at 72°C, repeated 39 times, 5 s at 65°C, and 30 s at 65°C to terminate the reaction. To evaluate the effect on gene expression, inducible REH cells or NALM6 cells were treated with DMSO or 50 nM SAHA dissolved in 100% DMSO for 12 hours before induction of transgenes (EV, wt DUX4, and DUX4-IGH) in REH cells and additional 12 hours during transgene induction. RNA was collected after a total of 24-hour SAHA treatment and 12-hour transgene induction.

Protein extraction and immunoblotting

Total proteins were extracted either via direct lysis by the addition of 50 μl of 2× Laemmli buffer (Bio-Rad, #1610747) to ~10^6 cells followed by boiling for 10 min at 95°C. Alternatively, protein extracts were performed in radioimmunoprecipitation assay (RIPA) buffer [25 mM tris-HCl (pH 8), 150 mM NaCl, 10 mM NaF, 20 mM Na3VO4, 0.1% SDS, and 1% NP-40) by incubation for 10 min in ice, followed by centrifugation of the extract at 16,000*g*, for 10 min at 4°C, and the supernatant containing proteins was collected. RIPA-extracted proteins were quantified with the Protein Assay Dye Reagent Concentrate (Bio-Rad, 5000006) using gamma-globin (1 μg/μl) as a standard. Absorbance was read at 595 nm using a spectrophotometer. For immunoblotting analyses of REH, HEK, and Jurkat cells, 1 μl of whole cell extract was diluted to 10 μl with 1× Laemmli buffer. For the other immunoblotting analyses, 15 μg of proteins was diluted in 4× Laemmli buffer to a concentration of 1× and boiled for 5′ at 95°C.

For immunoblotting analysis of tandem affinity purification material, 5, 5, and 10% fractions of input, first elution, and second

elution were loaded, respectively. For immunoblotting analysis of HEK cells supplemented with EV or GTF2I, cells were transfected as indicated in the "Cell culture, transfection, and transduction" section of Materials and Methods, and proteins were extracted via direct lysis. Standard SDS–polyacrylamide gel electrophoresis was performed at 10% polyacrylamide concentration, followed by wet-transferred onto a 0.45-μm nitrocellulose membrane using an 80% tris-HCl and 10% methanol solution for 1 hour and 30 min. Membranes were stained with Ponceau solution, and then incubated with a 5% milk-blocking solution in tris-buffered saline supplemented with 0.1% Tween 20 for >1 hour. Primary Abs (table S6) were incubated overnight (O.N.), oscillating at 4°C, and secondary, horseradish peroxidase–conjugated Abs were incubated for 1 hour at RT. The SuperSignal West Pico PLUS Chemiluminescent Substrate (Life Technologies, 34580) was used for the visualization of proteins.

Apoptosis assay

Apoptosis was measured through Caspase-Glo 3/7 luminescent assay (Promega, #G8090), following the manufacturer's instructions. Luminescence was quantified by Wallac 1420 multilabel Victor3 microplate reader (PerkinElmer). In inducible REH cells, assays were performed 24 hours after Dox administration. In sh*CTRL*/sh*GTF2I* NALM6 cells, assays were performed 72 hours after lentiviral infection. In sh*CTRL*/sh*GTF2I* REH DUX4-IGH cells, Dox was added 48 hours after lentiviral transduction, and apoptosis was measured 24 hours after Dox administration. For SAHA treatment, NALM6 cells were treated with 1 μM of SAHA (dissolved in 100% DMSO) or DMSO, as a control, for 72 hours before measurement of apoptosis. For inducible REH cells, 1 μM of SAHA or DMSO was administered at the start of the experiment, while Dox was supplemented 24 hours after the start of the experiment. Apoptosis was measured after 72 hours of SAHA treatment.

Cell-stroma adhesion assay

HS-5 cells were maintained in a fully supplemented RPMI 1640 medium for at least 72 hours before starting the experiment. Next, 250,000 HS-5 cells were plated into a 12-well plate 8 to 12 hours before coculturing with 100,000 REH-GFP or NALM6-GFP cells for 18 hours. Cells were then washed twice with phosphate-buffered saline (PBS). Subsequently, adhering cells were recovered with trypsin, washed, resuspended in PBS, and analyzed by fluorescence-activated cell sorting (FACS) for the percentage GFP-expressing cells.

To evaluate adhesion in knockdown cells, NALM6-GFP cells or GFP^+ (GFP-positive)-inducible REH cells were transduced with sh*CTRL*, sh*DUX4*-r, or shGTF2I lentiviral particles 48 hours before the start of the experiment.

To evaluate adhesion in SAHA-treated cells, NALM6-GFP cells or GFP^+-inducible REH cells were treated with DMSO or 1 μM SAHA dissolved in 100% DMSO before coculture with HS-5 cells. At the same time, GFP^+ REH cells were supplemented with 1 μg of Dox to induce DUX4-IGH expression.

Transwell migration assay

HS-5 cells were maintained in a fully supplemented RPMI 1640 medium for at least 72 hours before starting the experiment. A total of 100,000 HS-5 cells were plated into the bottom compartment of a 24-Transwell plate (Corning, CLS3399) 18 to 24 hours before the start of the coculture. Once HS-5 cells were fully attached and well-elongated, 50,000 inducible REH-GFP cells, sh*CTRL* or sh*DUX4-r* NALM6-GFP, sh*CTRL* or sh*GTF2I* inducible REH-GFP, were seeded in the upper compartment and Dox or vehicle (H_2O) was added. After 16- to 18-hour incubation, cells in the bottom compartment were trypsinized and the amount of GFP^+ cells that had migrated from the upper compartment was determined by flow cytometry.

To evaluate migration in SAHA-treated cells, NALM6-GFP cells or GFP^+-inducible REH cells were treated with DMSO or 1 μM SAHA dissolved in 100% DMSO before incubation in the transwell upper compartment. At the same time, GFP^+ REH cells were supplemented with 1 μg of Dox to induce DUX4-IGH expression.

Cell proliferation assay

A total of 200,000 inducible REH or NALM6 cells were seeded in 2 ml of RPMI 1640 medium supplemented with 5% of FBS in a six-well plate to get a final concentration of cells of 100,000/ml. Transgenes were induced with Dox (1 μg/ml; for inducible REH cells) and cell proliferation was measured by automatic cell counting (Bio-Rad, TC20-1450102) every 48 hours, accompanied by replacement of the medium with fresh Dox for up to 7 days.

Live-cell spheroid assay

A total of 40,000 inducible REH cells were seeded in 800 μl (50,000/ml) of RPMI 1640 medium containing 5% of FBS and Dox (1 μg/ml) into a 48-well plate (Costar, #3548). Cells were then incubated for 72 hours in the Incucyte Live-Cell Analysis Systems (Sartorius) for real-time analysis of cellular aggregation into spheroids. Clustering analysis was performed according to the Incucyte Handbook, Chapter 5a, "Kinetic Assays for Immune Cell Activation and Proliferation". Label-free, phase-contrast imaging was used to detect morphological differences between cells expressing different DUX4 variants. Quantification of the average spheroid area was calculated using the following parameters: Analysis of Phase Area Object Average normalized to T0, cell/background ratio: −1, fill holes: −5.

RNA sequencing

Total RNA was extracted as described above from ~3×10^6 inducible REH cells per sample, after 12 hours of Dox induction in 2 ml (1.5×10^6/ml) of a fully supplemented RPMI 1640 medium. RNA concentration, purity, and integrity were determined through the BioAnalyzer 2100 system. All samples used for RNA-seq library preparation showed an RNA integrity number greater than 9. RNA-seq libraries were generated from biological quadruplicates using the TruSeq Stranded mRNA library prep kit, using 100 ng of total RNA as input. Libraries were run with a NovaSeq 6000 instrument and sequenced in single-end generating 75-bp reads, for a total number of 50 million clusters per sample.

Transcriptomic analyses

Sequencing adapters were removed using Trimmomatic (v0.39) (*47*) and fastq files were then aligned to the human genome assembly GRCh38 (hg38) using the STAR aligner (v2.5.3a) (*48*). Annotation of genomic features was performed using the featureCounts tool (v1.6.4) (*49*), using the GENCODE v31 gene transfer format (GTF). DGE analysis was carried out with DeSeq2 (v1.3) package (Bioconductor) (*50*).

To generate lists of DUX4, DUX4-IGH, and DUX4-del50 up- and down-regulated genes, DGEs showing $Log_2FC \geq 1$ or $Log_2FC \leq -1$ for each condition have been compared with their relative expression levels [reads per kilobase per million mapped reads (RPKM)] in the given condition. Only genes with RPKM ≥ 0.5 were used for further analyses.

GSEA was performed using the GSeABase software (1.57.0) package (Bioconductor) (*51*), using the Canonical Pathways, Gene Ontology, from the Molecular Signatures Database MSigDB (v7.5.1) (*52*) and setting a false discovery rate (FDR) value to 0.05. To check the expression of DUX4/DUX4-IGH direct transcriptional targets in EV control cells, direct target lists have been compared with the list of genes expressed in EV REH cells with RPKM > 0.5. To test the percentage of surface proteins activated by DUX4-IGH, the list of direct DUX4-IGH transcriptional targets has been compared with the list of surface proteins identified by Bausch-Fluck *et al.* (*53*).

CUT&Tag

Starting from >5 × 10^6 REH/HEK cells induced with Dox for 12 hours, nuclei from two independent replicates were prepared by resuspending cells in NE1 buffer [20 mM Hepes-KOH (pH 7.9), 10 mM KCl, 0.5 mM spermidine, and 1% Triton X-100] at a cellular density of 10^6 cells/ml, followed by 10-min incubation in ice. Samples were spun for 4 min at 600*g* at 4°C and the nuclear pellet was resuspended in 1 volume (V) of PBS. A light cross-linking was performed by adding formaldehyde to 0.1% final concentration for 2 min, followed by quenching with a double molar concentration of glycine. Nuclei were spun and resuspended in ½ V of wash buffer 150 [20 mM Hepes-KOH (pH 7.9), 150 mM NaCl, and 0.5 mM spermidine]. Nuclei were then manually counted using the trypan blue exclusion assay, as described (*27*). For each CUT&Tag assay against different DUX4 versions and GTF2I, 100,000 nuclei were immobilized to 4 μl of Concanavalin A–conjugated beads (ConA-EpiCypher-21-1401) for 10 min at RT. For each CUT&Tag against H3K27Ac, 50,000 nuclei were immobilized to 4 μl of Concanavalin A for 10 min at RT.

The primary (first) Ab was diluted 1:100 in Ab buffer [0 mM Hepes-KOH (pH 7.9), 150 mM NaCl, 0.5 mM spermidine, BSA (bovine serum albumin; 0.5 mg/ml), and 0.4 mM EDTA (pH7.5)] and incubated with the ConA-bound nuclei ON at 4°C while rotating. The following day, ConA-nuclei-1stAb were spun 100*g*, 1 s, the supernatant was removed, nuclei were resuspended in 100 μl of wash buffer 150 and incubated with a secondary Ab (second) diluted 1:100 in wash buffer 150 for 1 hour at RT. After removal of the second Ab, nuclei were washed once with 200 μl of wash buffer 150 and incubated with the protein AG-conjugated transposase (pAG-Tn5, EpiCypher, #23615-1117) preloaded with sequencing adapters, diluted 1:20 in wash buffer 300 [20 mM Hepes-KOH (pH 7.9), 300 mM NaCl, and 0.5 mM spermidine] for 1 hour at RT. Nuclei were then washed twice in wash buffer 300 and a tagmentation reaction was performed by the addition of 10 mM MgCl2 following incubation at 37°C for 1 hour. After a wash with TAPS buffer (10 mM TAPS in H_2O), tagmented DNA was released by the addition of 5 μl 0.1% SDS release buffer (0.1% SDS and 10 mM TAPS in H_2O) following incubation at 58°C for 1 hour. Then, 15 μl of 0.67% Triton X-100 was added to allow PCR for library preparation. For the library prep, 2 μl of 10 μM i5 universal adapter primer (Nextera) and 2 μl of 10 μM i7 uniquely barcoded adapter primers were added to released, tagmented DNA, followed by the addition of 25 μl of the Q5 High-Fidelity 2X Master Mix (New England Biolabs, M0492L).

PCR was performed as follows: cycle 1: 72°C for 5 min (gap-filling); cycle 2: 98°C for 30 s, cycle 3: 98°C for 10 s; cycle 4: 63°C for 10 s, 72°C for 1 min and hold at 8°C. For H3K27Ac CUT&Tag and CUTAC, we repeat cycles 3 and 4 13 times; for CUT&Tag against different DUX4 versions and GTF2I, we repeat cycles 3 and 4 19 times. Amplified DNA was then recovered through size exclusion purification by addition of 1.3 V of AMPureXP magnetic beads (Beckman, A63881), followed by two washes in 80% ethanol and elution in 22 μl of 10 mM tris-HCl (pH 8). Quantification and average fragment size of libraries were performed via the 4200 TapeStation system (Agilent), using the D5000 high sensitivity kit (Agilent 5067-5592/5067-5593). The average library fragment size (450 bp) was selected for library quantification. A 5 nM pool of the libraries was then run with the NovaSeq instrument and sequenced in paired-end (PE), generating 2× 150-bp reads. A total of 4 million to 10 million reads per sample for H3K27Ac CUT&Tag and CUTAC and at least 10 million reads per sample for TF (DUX4/DUX4-r/GTF2I) were obtained.

CUTAC

Starting from >5 × 10^6 REH cells induced with Dox for 12 hours, nuclei from two independent replicates were prepared by resuspending cells in NE1 buffer [20 mM Hepes-KOH (pH 7.9), 10 mM KCl, 0.5 mM spermidine, and 1% Triton X-100] at a cellular density of 10^6 cells/ml, followed by 10-min incubation in ice. Samples were spun for 4 min at 600*g* at 4°C and the nuclear pellet was resuspended in 1 volume (V) of PBS. Nuclei were then manually counted using the trypan blue exclusion assay, as previously described (*27*). For each CUTAC assay, 50,000 nuclei were immobilized to 4 μl of Concanavalin A-conjugated for 10 min at RT.

The primary (first) Ab targeting Phospho_Rbp1 CTD (Ser^5) was diluted 1:100 in Ab buffer [20 mM Hepes-KOH (pH 7.9), 150 mM NaCl, 0.5 mM spermidine, BSA (0.5 mg/ml), and 0.4 mM EDTA (pH 7.5)] and incubated with the ConA-bound nuclei ON at 4°C while rotating. The following day, ConA-nuclei-1stAb were spun 100*g*, 1 s, the supernatant was removed, and nuclei were resuspended in 100 μl of wash buffer 150 and incubated with a secondary Ab (second) diluted 1:100 in wash buffer 150 for 1 hour at RT. After removal of the second Ab, nuclei were washed once with 200 μl of wash buffer 150 and incubated with the protein AG–conjugated transposase (pAG-Tn5, EpiCypher, #23615-1117) preloaded with sequencing adapters, diluted 1:40 in wash buffer 300 [20 mM Hepes-KOH (pH 7.9), 300 mM NaCl, and 0.5 mM spermidine] for 1 hour at RT. Fifty microliters of prewarmed CUTAC buffer (5 mM MgCl2 and 10 mM TAPS) was added to each tube followed by incubation at 37°C for 20 min. After a wash with TAPS buffer (10 mM TAPS in H_2O), tagmented DNA was released by the addition of 5 μl 0.1% SDS release buffer (0.1% SDS and 10 mM TAPS in H_2O) following incubation at 58°C for 1 hour. Then, 15 μl of 0.67% Triton X-100 was added to allow PCR for library preparation. For the library prep, 2 μl of 10 μM i5 universal adapter primer (Nextera) and 2 μl of 10 μM i7 uniquely barcoded adapter primers were added to released, tagmented DNA, followed by the addition of 25 μl of the Q5 High-Fidelity 2X Master Mix (New England Biolabs, M0492L).

PCR was performed as follows: cycle 1: 72°C for 5 mini (gap-filling); cycle 2: 98°C for 30 s, cycle 3: 98°C for 10 s; cycle 4: 63°C

for 10 s, 72°C for 1 min and hold at 8°C. We repeat cycles 3 to 4 13 times. Amplified DNA was then recovered through size exclusion purification by the addition of 1.3 V of AMPureXP magnetic beads (Beckman, A63881), followed by two washes in 80% ethanol and elution in 22 μl of 10 mM tris-HCl (pH 8). Quantification of the average fragment size of libraries was performed via the 4200 TapeStation system (Agilent), using the D5000 high sensitivity kit (Agilent, 5067-5592/5067-5593). The average library fragment size (175 to 650 bp) was selected for library quantification. A 5 nM pool of the libraries was then run with the NovaSeq instrument and sequenced in PE, generating 2× 150-bp reads.

CUT&Tag and CUTAC analyses

Sequencing adapters were removed via Cutadapt (v4.1) and then trimmed fastq files were aligned with Bowtie2 (v2.2.1) (*54*) to the human genome assembly GRCh38 (hg38) from Genome Reference Consortium, using the following parameters: --local --very-sensitive --no-mixed --no-discordant --phred33 -I 10 -X 700. Bam files containing uniquely mapped reads were created using SAMtools v1.15.1 (*55*).

CUT&Tag signals were called on using MACS (v2.2.7.1) (*56*) as normalized to immunoglobulin G (IgG) control signal, using the following settings: FDR cutoff of 1 $\times 10^{-3}$, BAMPE format, keep all duplicates.

The MACS --bdg setting was used to generate IgG control scaled, normalized bigwig files, which were used to visualize normalized binding signals. The same bigwig files were also used to generate profile plots using deepTools plotprofile (v3.5.0) (*57*).

Genes proximal to peaks were annotated against the hg38 genome using annotatePeaks.pl from HOMER v4 (*58*). DUX4, DUX4-IGH, and DUX4-del50 binding motifs were identified using findMotifsGenome.pl from HOMER v4, using standard parameters.

To identify regions associated by GTF2I both in REH and HEK cells, intersectBed from the BEDtools suite was used with default parameters.

DUX4 and DUX4-IGH direct targets lists were generated as follows: nonredundant, annotated genes associated with DUX4 or DUX4-IGH peaks were crossed with $Log_2FC \geq 1$ and RPKM ≥ 0.5 DGEs in the respective lists. Only genes present in both peaks and DGEs lists were considered as DUX4 or DUX4-IGH direct target genes.

To assess the difference between genomics signals (coverage) for selected regions sets (wt DUX4 or DUX4-IGH direct targets), per-bin corrected Wilcoxon rank sum tests were performed using the dsCompareCurves tool from deepStats (v0.4) (*59*).

To analyze genome-wide GTF2I DNA association in REH cells expressing EV or DUX4-IGH, reads per genome coverage (RPGC) normalized counts were generated using the deepTools bamCoverage function with standard parameters. Coverage files were then used as input for the deepTools plotHeatmap function and were plotted against the human reference genome GENCODE v43. To analyze DUX4-IGH association to chromatin in HEK cells in the presence/absence of GTF2I, RPGC-normalized counts of DUX4-IGH CUT&Tag in HEK + EV and HEK + GTF2I, as well as RPGC-normalized counts of IgG control CUT&Tag were generated using the deepTools bamCoverage function with standard parameters. The generated coverage files were then used as input for the deepTools plotHeatmap function and were plotted against the human reference genome GENCODE v43. Furthermore, to generate boxplots of normalized counts signals, the generated coverage files were used as input for the multiBigWigSummary function of deepTools.

Tandem affinity purification

Tandem affinity purification from REH cells was performed 12 hours after Dox induction. A total of 2.5×10^8 cells were resuspended in 20 ml of buffer N [10 mM Hepes-KOH (pH 7.9), 300 mM sucrose, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 0.1 mM dithiothreitol (DTT), 0.75 mM spermidine, 0.15 mM spermine, 0.1% NP-40, and 50 mM NaF] and incubated in ice for 10 min to isolate intact nuclei. After one wash in the same volume of buffer N, nuclei were resuspended in the same volume of high-salt nuclear extraction buffer C420 [20 mM Hepes-KOH (pH 7.9), 420 mM NaCl, 25% glycerol, 1 mM EGTA, and 0.1 mM DTT, and 50 mM NaF] and vigorously shaken in a thermomixer at 1400 rpm for 30 min at 4°C to extract nuclear proteins. Samples were then centrifuged at 16,000*g* for 10 min at 4°C, and the supernatant, which contains nuclear proteins, was collected. Nuclear extracts were then adjusted to 150 mM salt by addition of Hepes buffer [20 mM Hepes-KOH (pH7.9) and 50 mM NaF] and incubated at 4°C while rotating with avidin (1 μg/ml; IBA, #2-0204-015) to chelate biotinylated proteins, and benzonase nuclease (20 U/ml; Sigma Aldrich, E8263-5KU) and ribonuclease (RNase) A (50 ng/ml; Thermo Fisher Scientific, EN0531) to remove any remaining nucleic acids. After clearing at 16,000*g* for 10 min at 4°C, supernatants were precleared with 2.5 ml (bead volume) of cross-linked Sepharose (Sigma-Aldrich, DCL6B100) for 1 hour while rotating at 4°C. Extracts were then quantified with the Protein Assay Dye Reagent Concentrate (Bio-Rad, #5000006) using gamma-globin (1 μg/μl) as standard. Absorbance was read at 595 nm using a spectrophotometer. For each purification, 80 mg of precleared nuclear extracts was incubated with 2 ml (bead volume) of StrepTactin sepharose beads (IBA, #2-1201-010), for 4 hours at 4°C while rotating. Then, three washes in TNN-HS [10 mM Hepes-KOH (pH 7.9), 150 mM NaCl, 0.1% NP-40, and 50 mM NaF] buffer were performed, and beads-bound proteins were specifically eluted with three subsequent incubations with 2 V of 2.5 mM biotin-containing TNN-HS buffer. Next, eluted proteins were incubated with 2 ml (bead volume) of anti-HA agarose beads (Sigma-Aldrich, A2095) while rotating at 4°C O.N. The following day, two washes in TNN-HS were performed, followed by two additional washes in TNN buffer [10 mM Hepes-KOH (pH 7.9), 150 mM NaCl, and 50 mM NaF] to get rid of any leftover NP-40 detergent in the buffer. Last, HA beads-bound proteins were eluted with elution buffer [50 mM Hepes-KOH (pH 7.9), 150 mM NaCl, and 1% SDS] and precipitated with acetone. For this, 4 V of ice-cold 100% acetone was added to 1 V of eluted proteins, followed by vortexing and incubation at −20°C for 1 hour. Samples were then centrifuged at 15,000*g* for 10 min. The whole procedure was repeated twice to get rid of all the remaining SDS. One aliquot of each protein sample was kept for quantification via silver staining.

Mass spectrometry sample preparation

Samples were resolubilized in 50 μl of detergent-based buffer [1% sodium deoxycholate, 10 mM tris(2-carboxyethyl)phosphine, 10 mM tris, and 40 mM chloroacetamide] with cOmplete Mini EDTA-free protease inhibitor cocktail (Roche, 04693159001),

sonicated for 15 cycles (30 s on, 30 s off), and heated at 95C for 10 min. Trypsin digestion (1.5 hours) and peptide cleanup were performed using the S-trap system (Protifi) following the manufacturer's specifications. Peptides were eluted and vacuum centrifuged to complete dryness. Purified peptides were resolubilized in 1% formic acid (FA) to a concentration of 1 μg/μl and 1 μl was injected to achieve a peptide load of 1 μg on a column.

Mass spectrometry data acquisition

Spectra were acquired on an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific) in a data-dependent fashion coupled to an Ultimate3000 liquid chromatography system (Thermo Fisher Scientific) and separated on a 50-cm reversed-phase column packed in-house (Poroshell EC-C18, 2.7 μm, 50 cm by 75 μm; Agilent Technologies). Samples were eluted over a linear gradient ranging from 9 to 13% acetonitrile over 2 min and from 13 to 44% over 95 min, 44 to 99% acetonitrile in 3 min, followed by 99% acetonitrile for 5 min with a flow rate of 300 nl/min. MS1 spectra had a resolving power of 60,000 at 200 mass/charge ratio (*m*/z) with the AGC target set to "standard" and the maximum injection time set to "auto." MS/MS spectra were acquired with HCD fragmentation, an HCD collision energy of 28%, a 1.4-m/z-wide isolation window, resolving power of 30,000 at 200 m/z, the AGC target set to standard, and the maximum injection time set to auto.

Mass spectrometry data processing

To quantify proteins, the raw data were loaded into the MaxQuant (*60*) software version 2.2.0 to search the human_proteome 20180425 (93,606 sequences; 37,037,628 residues). Searches were performed using default settings, namely, trypsin as proteolytic enzyme; one missed cleavage allowed; protein N-terminal acetylation. Peptides and proteins were accepted with an FDR of less than 1%. Label-free protein quantification was based on the spectral counts considering only proteins identified with minimum of two peptides in any tandem affinity purification. Specific interactors of wt DUX4, DUX4-IGH, or DUX4-del50 were selected as those being detected in all three biological replicates and showing spectral counts (MS/MS) fold enrichment ≥2 with respect to the control tandem affinity purifications. To generate protein-protein interaction networks, Significance Analysis of INTeractome (SAINT v2.0) (*61*) was performed using as input the list of identified wt DUX4-, DUX4-IGH–, or DUX4-del50–specific interactors. Only proteins showing a fold change score >1.5 and probability score (SAINT SP) >0.9 were included in the protein-protein interaction network.

Endogenous GTF2I-DUX4-r co-IP

A total of 5×10^7 NALM6 cells were resuspended in 4 ml of buffer N [10 mM Hepes-KOH (pH 7.9), 300 mM sucrose, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 0.1 mM DTT, 0.75 mM spermidine, 0.15 mM spermine, 0.1% NP-40, and 50 mM NaF] and incubated on ice for 10 min to isolate intact nuclei. At this step, a 30-μl aliquot was collected and lysed in Laemmli buffer to be used as input. The remaining nuclei, after one wash in the same volume of buffer N, were resuspended in the same volume of high-salt nuclear extraction buffer C420 [20 mM Hepes-KOH (pH 7.9), 420 mM NaCl, 25% glycerol, 1 mM EGTA, 0.1 mM DTT, and 50 mM NaF], and vigorously shaken in a thermomixer at 1400 rpm for 30 min at 4°C to extract nuclear proteins. Samples were clarified by centrifugation at 16,000*g* for 10 min at 4°C, and the supernatant, which contains nuclear proteins, was collected. Nuclear extracts were then adjusted to 150 mM salt by addition of Hepes buffer [20 mM Hepes-KOH (pH 7.9) and 50 mM NaF] and incubated at 4°C while rotating with benzonase (20 U/ml; Sigma-Aldrich, E8263-5KU) and RNase A (50 ng/ml; Thermo Fisher Scientific, EN0531) to remove any remaining nucleic acids. After clearing at 16,000*g* for 10 min at 4°C, supernatants were precleared with 200 μl (bead volume) of cross-linked Sepharose (Sigma-Aldrich, DCL6B100) for 1 hour while rotating at 4°C. Extracts were then quantified with the Protein Assay Dye Reagent Concentrate (Bio-Rad, #5000006) using gamma-globin (1 μg/μl) as standard. Absorbance was read at 595 nm using a spectrophotometer. Two milligrams of nuclear proteins was then incubated with 4 μg of anti-DUX4 Ab (R&D Systems, MAB95351) or ChromPure Rabbit IgG (Jackson ImmunoResearch, #011-000-003), as a negative control, and rotated O.N. at 4°C. The following day, Protein G Dynabeads (Invitrogen, #10004D) were blocked with three washes with 500 μl of 0.5% BSA/PBS solution, and then 30 μl was added to each Ab-protein complex and incubated for 3 hours, rotating at 4°C. Samples were then washed three times by incubation with TNN-HS [50 mM Hepes (pH 8.0), 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 50 mM NaF, and protease inhibitors], followed by 3 min of rotation at 4°C. Last, beads were boiled with 30 μl of 2× Laemmli buffer at 95°C for 10 min, and centrifuged for 5 min at 1200 rpm to remove beads. Five microliters of input sample and 7 μl of either DUX4-r IP or IgG were loaded in a 10% acrylamide gel for Western blot (WB) analysis.

In vivo studies

All mouse experiments were approved by the Institutional Animal Care and Use Committee of San Raffaele and by the Italian Ministry of Health (Rome, Italy). Mice were kept in a specific-pathogen-free (SPF) facility within individually ventilated cages and were given irradiated food and water ad libitum. Eight- to 10-week-old NSG mice (Charles River Laboratories) were infused intravenously (iv) with 250,000 sh*CTRL* or sh*GTF2I* NALM6-Lucia cells. After 1 week, blood was collected, and cell engraftment was evaluated by reading the secreted luciferase using the Berthold Mithras (LB-940) Luminometer with a read time of 10 s. Subsequently, blood was collected every 7 days and mice were euthanized when relative light units (r.l.u.) exceeded the threshold of 1.5×10^6 or when manifesting clinical signs of suffering.

For the pharmacological treatment, NSG mice were transplanted with 250,000 NALM6-Lucia cells. One week after transplantation, once efficient engraftment was reached, mice were randomly split into groups and treated intraperitoneally (ip) with either SAHA or DMSO (50 mg/kg), as vehicle control dissolved in 2% DMSO + 30% PEG300 + 5% Tween 80 + ddH_2O. SAHA or DMSO IP injections were repeated every day, changing sides on every administration. Animals were euthanized when r.l.u. exceeded the threshold of 1.5×10^6 or when manifesting clinical signs of suffering.

To determine leukemia cell engraftment in BMs and spleens of euthanized mice, cells were purified from respective compartments as follows. To purify leukemia cells from BM, both tibias and fibulas were harvested and crushed via mortar and pestle in 5 ml of PBS. The resulting suspension was then passed through a 40-μm cell strainer (Sarstedt, 833945070) to remove debris and centrifuged for 10 min at 300*g* at 4°C to pellet extracted cells. Then, cell pellets were resuspended in 1 ml of 1× red blood cell (RBC) lysis

buffer (Life Technologies, 00433357) for 10 min and centrifuged again to isolate leukemic cells. The final pellet was then resuspended in 100 μl of PBS, and 5 μl of the final cell suspension was used to evaluate cell engraftment using the Berthold Mithras Luminometer with a read time of 10 s. The remaining cells were then pelleted for 10 min at 300*g* at 4°C and used as input for total RNA extraction for subsequent RT-qPCR analysis as described above. To purify splenic leukemia cells, harvested spleens were crushed using a 5-ml syringe plunger in 5 ml of PBS in a 10-cm cell culture dish (Falcon, 353003). The resulting suspension was then passed through a 40-μm cell strainer to remove debris and centrifuged for 10 min at 300*g* at 4°C to pellet extracted cells. Then, cell pellets were resuspended in 1 ml of 1× RBC lysis buffer (Life Technologies, 00433357) for 10 min and centrifuged again to isolate leukemic cells. The final pellet was then resuspended in 100 μl of PBS, and 5 μl of the final cell suspension was used to evaluate cell engraftment using the Berthold Mithras Luminometer with a read time of 10 s.

Quantification and statistical analyses

Statistical analysis was performed with either GraphPad Prism 9 or R. For RNA-seq, CUT&Tag, CUTAC, RT-PCR, WB, FACS, spheroid formation assay, proliferation assay, and in vivo experiments, statistical parameters including *P* value, replicates, and SD are detailed on the figures or annotated in the respective figure legends.

Supplementary Materials

This PDF file includes:
Figs. S1 to S5
Legends for tables S1 to S6

Other Supplementary Material for this manuscript includes the following:
Tables S1 to S6

REFERENCES AND NOTES

1. S. P. Hunger, C. G. Mullighan, Acute lymphoblastic leukemia in children. *N. Engl. J. Med.* **373**, 1541–1552 (2015).
2. G. R. Bunin, E. J. Feuer, P. A. Witman, A. T. Meadows, Increasing incidence of childhood cancer: Report of 20 years experience from the greater Delaware Valley Pediatric Tumor Registry. *Paediatr. Perinat. Epidemiol.* **10**, 319–338 (1996).
3. M. S. Linet, L. A. Ries, M. A. Smith, R. E. Tarone, S. S. Devesa, Cancer surveillance series: Recent trends in childhood cancer incidence and mortality in the United States. *J. Natl. Cancer Inst.* **91**, 1051–1058 (1999).
4. I. Iacobucci, S. Kimura, C. G. Mullighan, Biologic and therapeutic implications of genomic alterations in acute lymphoblastic leukemia. *J. Clin. Med.* **10**, 3792 (2021).
5. T. Yasuda, S. Tsuzuki, M. Kawazu, F. Hayakawa, S. Kojima, T. Ueno, N. Imoto, S. Kohsaka, A. Kunita, K. Doi, T. Sakura, T. Yujiri, E. Kondo, K. Fujimaki, Y. Ueda, Y. Aoyama, S. Ohtake, J. Takita, E. Sai, M. Taniwaki, M. Kurokawa, S. Morishita, M. Fukayama, H. Kiyoi, Y. Miyazaki, T. Naoe, H. Mano, Recurrent DUX4 fusions in B cell acute lymphoblastic leukemia of adolescents and young adults. *Nat. Genet.* **48**, 569–574 (2016).
6. J. Zhang, K. McCastlain, H. Yoshihara, B. Xu, Y. Chang, M. L. Churchman, G. Wu, Y. Li, L. Wei, I. Iacobucci, Y. Y. Y. Y. Liu, C. Qu, J. Wen, M. Edmonson, D. Payne-Turner, K. B. Kaufmann, S. I. Takayanagi, E. Wienholds, E. Waanders, P. Ntziachristos, S. Bakogianni, J. Wang, I. Aifantis, K. G. Roberts, J. Ma, G. Song, J. Easton, H. L. Mulder, X. Chen, S. Newman, X. Ma, M. Rusch, P. Gupta, K. Boggs, B. Vadodaria, J. Dalton, Y. Y. Y. Y. Liu, M. L. Valentine, L. Ding, C. Lu, R. S. Fulton, L. Fulton, Y. Tabib, K. Ochoa, M. Devidas, D. Pei, C. Cheng, J. Yang, W. E. Evans, M. V. Relling, C.-H. H. Pui, S. Jeha, R. C. Harvey, I.-M. M. L. Chen, C. L. Willman, G. Marcucci, C. D. Bloomfield, J. Kohlschmidt, K. Mrózek, E. Paietta, M. S. Tallman, W. Stock, M. C. Foster, J. Racevskis, J. M. Rowe, S. Luger, S. M. Kornblau, S. A. Shurtleff, S. C. Raimondi, E. R. Mardis, R. K. Wilson, J. E. Dick, S. P. Hunger, M. L. Loh, J. R. Downing, C. G. Mullighan; St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project, Deregulation of DUX4 and ERG in acute lymphoblastic leukemia. *Nat. Genet.* **48**, 1481–1489 (2016).
7. H. Lilljebjörn, R. Henningsson, A. Hyrenius-Wittsten, L. Olsson, C. Orsmark-Pietras, S. von Palffy, M. Askmyr, M. Rissler, M. Schrappe, G. Cario, A. Castor, C. J. H. H. Pronk, M. Behrendtz, F. Mitelman, B. Johansson, K. Paulsson, A. K. Andersson, M. Fontes, T. Fioretos, Identification of ETV6-RUNX1-like and DUX4-rearranged subtypes in paediatric B-cell precursor acute lymphoblastic leukaemia. *Nat. Commun.* **7**, 11790 (2016).
8. Y. Marincevic-Zuniga, J. Dahlberg, S. Nilsson, A. Raine, S. Nystedt, C. M. Lindqvist, E. C. Berglund, J. Abrahamsson, L. Cavelier, E. Forestier, M. Heyman, G. Lönnerholm, J. Nordlund, A.-C. Syvänen, Transcriptome sequencing in pediatric acute lymphoblastic leukemia identifies fusion genes associated with distinct DNA methylation profiles. *J. Hematol. Oncol.* **10**, 148 (2017).
9. Y. Tanaka, M. Kawazu, T. Yasuda, M. Tamura, F. Hayakawa, S. Kojima, T. Ueno, H. Kiyoi, T. Naoe, H. Mano, Transcriptional activities of DUX4 fusions in B-cell acute lymphoblastic leukemia. *Haematologica* **103**, e522–e526 (2018).
10. P. G. Hendrickson, J. A. Doráis, E. J. Grow, J. L. Whiddon, J.-W. Lim, C. L. Wike, B. D. Weaver, C. Pflueger, B. R. Emery, A. L. Wilcox, D. A. Nix, C. M. Peterson, S. J. Tapscott, D. T. Carrell, B. R. Cairns, Conserved roles of mouse DUX and human DUX4 in activating cleavage-stage genes and MERVL/HERVL retrotransposons. *Nat. Genet.* **49**, 925–934 (2017).
11. A. De Iaco, E. Planet, A. Coluccio, S. Verp, J. Duc, D. Trono, DUX-family transcription factors regulate zygotic genome activation in placental mammals. *Nat. Genet.* **49**, 941–945 (2017).
12. J. L. Whiddon, A. T. Langford, C. J. Wong, J. W. Zhong, S. J. Tapscott, Conservation and innovation in the DUX4-family gene network. *Nat. Genet.* **49**, 935–940 (2017).
13. S. H. Choi, M. D. Gearhart, Z. Cui, D. Bosnakovski, M. Kim, N. Schennum, M. Kyba, DUX4 recruits p300/CBP through its C-terminus and induces global H3K27 acetylation changes. *Nucleic Acids Res.* **44**, 5161–5173 (2016).
14. M. Carlet, K. Völse, J. Vergalli, M. Becker, T. Herold, A. Arner, D. Senft, V. Jurinovic, W.-H. Liu, Y. Gao, V. Dill, B. Fehse, C. D. Baldus, L. Bastian, L. Lenk, D. M. Schewe, J. W. Bagnoli, B. Vick, J. P. Schmid, A. Wilhelm, R. Marschalek, P. J. Jost, C. Miething, K. Riecken, M. Schmidt-Supprian, V. Binder, I. Jeremias, In vivo inducible reverse genetics in patients' tumors to identify individual therapeutic targets. *Nat. Commun.* **12**, 5655 (2021).
15. L. Tian, Y. Shao, S. Nance, J. Dang, B. Xu, X. Ma, Y. Li, B. Ju, L. Dong, S. Newman, X. Zhou, P. Schreiner, E. Tseng, T. Hon, M. Ashby, C. Li, J. Easton, T. A. Gruber, J. Zhang, Long-read sequencing unveils IGH-DUX4 translocation into the silenced IGH allele in B-cell acute lymphoblastic leukemia. *Nat. Commun.* **10**, 2789 (2019).
16. E. Clappier, M. F. Auclerc, J. Rapion, M. Bakkus, A. Caye, A. Khemiri, C. Giroux, L. Hernandez, E. Kabongo, S. Savola, T. Leblanc, K. Yakouben, G. Plat, V. Costa, A. Ferster, S. Girard, O. Fenneteau, J. M. Cayuela, F. Sigaux, N. Dastugue, S. Suciu, Y. Benoit, Y. Bertrand, J. Soulier, H. Cavé, An intragenic ERG deletion is a marker of an oncogenic subtype of B-cell precursor acute lymphoblastic leukemia with a favorable outcome despite frequent IKZF1 deletions. *Leukemia* **28**, 70–77 (2014).
17. M. Zaliova, M. Kotrova, S. Bresolin, J. Stuchly, J. Stary, O. Hrusak, G. Te Kronnie, J. Trka, J. Zuna, M. Vaskova, ETV6/RUNX1-like acute lymphoblastic leukemia: A novel B-cell precursor leukemia subtype associated with the CD27/CD44 immunophenotype. *Genes Chromosomes Cancer* **56**, 608–616 (2017).
18. C. G. Mullighan, C. B. Miller, X. Su, I. Radtke, J. Dalton, G. Song, X. Zhou, C.-H. Pui, S. A. Shurtleff, J. R. Downing, *ERG* deletions define a novel subtype of B-progenitor acute lymphoblastic leukemia. *Blood* **110**, 691 (2007).
19. M. Zaliova, E. Potuckova, L. Hovorkova, A. Musilova, L. Winkowska, K. Fiser, J. Stuchly, E. Mejstrikova, J. Starkova, J. Zuna, J. Stary, J. Trka, *ERG* deletions in childhood acute lymphoblastic leukemia with *DUX4* rearrangements are mostly polyclonal, prognostically relevant and their detection rate strongly depends on screening method sensitivity. *Haematologica* **104**, 1407–1416 (2019).
20. Y.-F. Liu, B.-Y. Wang, W.-N. Zhang, J.-Y. Huang, B.-S. Li, M. Zhang, L. Jiang, J.-F. Li, M.-J. Wang, Y.-J. Dai, Z.-G. Zhang, Q. Wang, J. Kong, B. Chen, Y.-M. Zhu, X.-Q. Weng, Z.-X. Shen, J.-M. Li, J. Wang, X.-J. Yan, Y. Li, Y.-M. Liang, L. Liu, X.-Q. Chen, W.-G. Zhang, J.-S. Yan, J.-D. Hu, S.-H. Shen, J. Chen, L.-J. Gu, D. Pei, Y. Li, G. Wu, X. Zhou, R.-B. Ren, C. Cheng, J. J. Yang, K.-K. Wang, S.-Y. Wang, J. Zhang, J.-Q. Mi, C.-H. Pui, J.-Y. Tang, Z. Chen, S.-J. Chen, Genomic profiling of adult and pediatric B-cell acute lymphoblastic leukemia. *EBioMedicine* **8**, 173–183 (2016).
21. J. M. Young, J. L. Whiddon, Z. Yao, B. Kasinathan, L. Snider, L. N. Geng, J. Balog, R. Tawil, S. M. van der Maarel, S. J. Tapscott, DUX4 binding to retroelements creates promoters that are active in FSHD muscle and testis. *PLOS Genet.* **9**, e1003947 (2013).
22. X. Dong, W. Zhang, H. Wu, J. Huang, M. Zhang, P. Wang, H. Zhang, Z. Chen, S.-J. Chen, G. Meng, Structural basis of DUX4/IGH-driven transactivation. *Leukemia* **32**, 1466–1476 (2018).
23. S. Jagannathan, S. C. Shadle, R. Resnick, L. Snider, R. N. Tawil, S. M. van der Maarel, R. K. Bradley, S. J. Tapscott, Model systems of DUX4 expression recapitulate the transcriptional profile of FSHD cells. *Hum. Mol. Genet.* **25**, 4419–4431 (2016).

24. L. N. Geng, Z. Yao, L. Snider, A. P. Fong, J. N. Cech, J. M. Young, S. M. van der Maarel, W. L. Ruzzo, R. C. Gentleman, R. Tawil, S. J. Tapscott, DUX4 activates germline genes, retroelements, and immune mediators: Implications for facioscapulohumeral dystrophy. *Dev. Cell* **22**, 38–51 (2012).
25. P. Dmitriev, Y. Bou Saada, C. Dib, E. Ansseau, A. Barat, A. Hamade, P. Dessen, T. Robert, V. Lazar, R. A. N. Louzada, C. Dupuy, V. Zakharova, G. Carnac, M. Lipinski, Y. S. Vassetzky, DUX4-induced constitutive DNA damage and oxidative stress contribute to aberrant differentiation of myoblasts from FSHD patients. *Free Radic. Biol. Med.* **99**, 244–258 (2016).
26. P. Dmitriev, E. Kiseleva, O. Kharchenko, E. Ivashkin, A. Pichugin, P. Dessen, T. Robert, F. Coppée, A. Belayew, G. Carnac, D. Laoudj-Chenivesse, M. Lipinski, A. Vasiliev, Y. S. Vassetzky, Dux4 controls migration of mesenchymal stem cells through the Cxcr4-Sdf1 axis. *Oncotarget* **7**, 65090–65108 (2016).
27. H. S. Kaya-Okur, S. J. Wu, C. A. Codomo, E. S. Pledger, T. D. Bryson, J. G. Henikoff, K. Ahmad, S. Henikoff, CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930 (2019).
28. Y. Zhang, J. K. Lee, E. A. Toso, J. S. Lee, S. H. Choi, M. Slattery, H. Aihara, M. Kyba, DNA-binding sequence specificity of DUX4. *Skelet Muscle* **6**, 8 (2016).
29. S. Henikoff, J. G. Henikoff, K. Ahmad, Simplified epigenome profiling using antibody-tethered tagmentation. *Bio Protoc.* **11**, e4043 (2021).
30. I. C. C. F. E. Fonseca, F. A. C. da Luz, I. A. Uehara, M. J. B. Silva, Cell-adhesion molecules and their soluble forms: Promising predictors of "tumor progression" and relapse in leukemia. *Tumour Biol.* **40**, 1010428318811525 (2018).
31. B. F. S. S. Scharff, S. Modvig, H. V. Marquart, C. Christensen, Integrin-mediated adhesion and chemoresistance of acute lymphoblastic leukemia cells residing in the bone marrow or the central nervous system. *Front. Oncol.* **10**, 775 (2020).
32. R. H. Scheuermann, E. Racila, CD19 antigen in leukemia and lymphoma diagnosis and immunotherapy. *Leuk. Lymphoma* **18**, 385–397 (1995).
33. R. Jacamo, Y. Chen, Z. Wang, M. Wencai, M. Zhang, E. L. Spaeth, Y. Wang, V. L. Battula, P. Y. Mak, K. Schallmoser, P. Ruvolo, W. D. Schober, E. J. Shpall, M. H. Nguyen, D. Strunk, C. E. Bueso-Ramos, S. Konoplev, R. E. Davis, M. Konopleva, M. Andreeff, Reciprocal leukemia-stroma VCAM-1/VLA-4-dependent activation of NF-κB mediates chemoresistance. *Blood* **123**, 2691–2702 (2014).
34. T. Ishiguro, H. Ohata, A. Sato, K. Yamawaki, T. Enomoto, K. Okamoto, Tumor-derived spheroids: Relevance to cancer stem cells and clinical applications. *Cancer Sci.* **108**, 283–289 (2017).
35. T. Chailangkarn, C. Noree, A. R. Muotri, The contribution of GTF2I haploinsufficiency to Williams syndrome. *Mol. Cell. Probes* **40**, 45–51 (2018).
36. A. L. Roy, Pathophysiology of TFII-I: Old guard wearing new hats. *Trends Mol. Med.* **23**, 501–511 (2017).
37. J. Li, H.-Y. Zhong, Y. Zhang, L. Xiao, L.-H. Bai, S.-F. Liu, G.-B. Zhou, G.-S. Zhang, GTF2I-RARA is a novel fusion transcript in a t(7;17) variant of acute promyelocytic leukaemia with clinical resistance to retinoic acid. *Br. J. Haematol.* **168**, 904–908 (2015).
38. I. Panagopoulos, M. Brunetti, M. Stoltenberg, R. A. U. Strandabø, J. Staurseth, K. Andersen, I. Kostolomov, T. S. Hveem, S. Lorenz, T. A. Nystad, T. Flægstad, F. Micci, S. Heim, Novel *GTF2I-PDGFRB* and *IKZF1-TYW1* fusions in pediatric leukemia with normal karyotype. *Exp. Hematol. Oncol.* **8**, 12 (2019).
39. I. Petrini, P. S. Meltzer, I.-K. Kim, M. Lucchi, K.-S. Park, G. Fontanini, J. Gao, P. A. Zucali, F. Calabrese, A. Favaretto, F. Rea, J. Rodriguez-Canales, R. L. Walker, M. Pineda, Y. J. Zhu, C. Lau, K. J. Killian, S. Bilke, D. Voeller, S. Dakshanamurthy, Y. Wang, G. Giaccone, A specific missense mutation in GTF2I occurs at high frequency in thymic epithelial tumors. *Nat. Genet.* **46**, 844–849 (2014).
40. S. Arcangeli, L. Falcone, B. Camisa, F. De Girardi, M. Biondi, F. Giglio, F. Ciceri, C. Bonini, A. Bondanza, M. Casucci, Next-generation manufacturing protocols enriching T_{SCM} CAR T cells can overcome disease-specific T cell defects in cancer patients. *Front. Immunol.* **11**, 1217 (2020).
41. F. Cavallo, F. Troglio, G. Fagà, D. Fancelli, R. Shyti, S. Trattaro, M. Zanella, G. D'Agostino, J. M. Hughes, M. R. Cera, M. Pasi, M. Gabriele, M. Lazzarin, M. Mihailovich, F. Kooy, A. Rosa, C. Mercurio, M. Varasi, G. Testa, High-throughput screening identifies histone deacetylase inhibitors that modulate GTF2I expression in 7q11.23 microduplication autism spectrum disorder patient-derived cortical neurons. *Mol. Autism* **11**, 88 (2020).
42. S. Mitsuhashi, S. Nakagawa, M. Sasaki-Honda, H. Sakurai, M. C. Frith, H. Mitsuhashi, Nanopore direct RNA sequencing detects DUX4-Activated repeats and isoforms in human muscle cells. *Hum. Mol. Genet.* **30**, 552–563 (2021).
43. Z. Yao, L. Snider, J. Balog, R. J. L. F. Lemmers, S. M. Van Der Maarel, R. Tawil, S. J. Tapscott, DUX4-induced gene expression is the major molecular signature in FSHD skeletal muscle. *Hum. Mol. Genet.* **23**, 5342–5352 (2014).
44. O. Khan, N. B. La Thangue, Drug Insight: Histone deacetylase inhibitor-based therapies for cutaneous T-cell lymphomas. *Nat. Clin. Pract. Oncol.* **5**, 714–726 (2008).
45. A. Wawruszak, L. Borkiewicz, E. Okon, W. Kukula-Koch, S. Afshan, M. Halasa, Vorinostat (SAHA) and breast cancer: An overview. *Cancers (Basel)* **13**, 4700 (2021).
46. L. Sanchez, D. H. Vesole, J. R. Richter, N. Biran, E. Bilotti, L. McBride, P. Anand, K. Ivanovski, D. S. Siegel, A phase IIb trial of vorinostat in combination with lenalidomide and dexamethasone in patients with multiple myeloma refractory to previous lenalidomide-containing regimens. *Br. J. Haematol.* **176**, 440–447 (2017).
47. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
48. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
49. Y. Liao, G. K. Smyth, W. Shi, featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
50. J. Costa-Silva, D. Domingues, F. M. Lopes, RNA-seq differential expression analysis: An extended review and a software tool. *PLOS ONE* **12**, e0190152 (2017).
51. M. Morgan, S. Falcon, R. Gentleman, *GSEABase: Gene Set Enrichment Data Structures and Methods* (2021); R package version 1.56.0.
52. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
53. D. Bausch-Fluck, U. Goldmann, S. Müller, M. van Oostrum, M. Müller, O. T. Schubert, B. Wollscheid, The in silico human surfaceome. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E10988–E10997 (2018).
54. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
55. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
56. Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nussbaum, R. M. Myers, M. Brown, W. Li, X. S. Shirley, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
57. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
58. S. Heinz, C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh, C. K. Glass, Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
59. G. Richard, gtrichard/deepStats: deepStats 0.3.1; 10.5281/zenodo.3361799 (2019).
60. J. Cox, N. Neuhauser, A. Michalski, R. A. Scheltema, J. V. Olsen, M. Mann, Andromeda: A peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* **10**, 1794–1805 (2011).
61. D. Mellacheruvu, Z. Wright, A. L. Couzens, J. P. Lambert, N. A. St-Denis, T. Li, Y. V. Miteva, S. Hauri, M. E. Sardiu, T. Y. Low, V. A. Halim, R. D. Bagshaw, N. C. Hubner, A. Al-Hakim, A. Bouchard, D. Faubert, D. Fermin, W. H. Dunham, M. Goudreault, Z.-Y. Lin, B. Gonzalez Badillo, T. Pawson, D. Durocher, B. Coulombe, R. Aebersold, G. Superti-Furga, J. Colinge, A. J. R. Heck, H. Choi, M. Gstaiger, S. Mohammed, I. M. Cristea, K. L. Bennett, M. P. Washburn, B. Raught, R. M. Ewing, A.-C. Gingras, A. I. Nesvizhskii, The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* **10**, 730–736 (2013).

Acknowledgments: We thank M. Casucci for providing NALM6-Lucia cells. We thank F. Voellmy and R. Zenezini-Chiozzi from EPIC-XS for direct support with mass spectrometry analysis. We thank R. Bernardi, G. Tonon, L. Vago, and members of the Gabellini laboratory for helpful discussions. **Funding:** Research in the Gabellini laboratory was supported by Italian Ministry of Health (grant number RF-2018-12366631), European Joint Programme on Rare Diseases (grant number EJPRD20-191), Italian Association for Cancer Research (grant number IG 2017-ID. 19919), and NIH—National Cancer Institute (grant number 1R21CA249378-01). M.S. was supported by Fondazione Umberto Veronesi (FUV Post-doctoral Fellowships 2019 number 2852 and FUV 2023 number 5089) and by Horizon 2020 Marie Skłodowska-Curie Actions (H2020-MSCA-IF-2018 grant number 841957—IDEAL). This work has been supported by EPIC-XS, project number 823839, funded by the Horizon 2020 program of the European Union. **Author contributions:** Conceptualization: D.C. and D.G. Methodology: D.C., M.S., and B.B. Investigation: D.C., M.S., and B.B. Data curation: D.C. and A.S.T. Writing—original draft: D.C. and D.G. Writing—review and editing: D.C., M.S., B.B., A.S.T., and D.G. Funding acquisition: D.G. Supervision: D.G. Project administration: D.G. **Competing interests:** The authors declare they have no competing interests. **Data and materials availability:** All constructs and cell lines generated in this study can be provided by the lead author. Requests for material should be submitted to gabellini.davide@hsr.it. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. RNA-seq, CUT&Tag, and CUTAC data have been deposited in the GEO database with accession number GSE228474. The mass

spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD041273. This paper does not report original codes. Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, D.G. (gabellini.davide@hsr.it).

Submitted 24 April 2023
Accepted 14 August 2023
Published 15 September 2023
10.1126/sciadv.adi3771