PHF6 interacts with the LMO2 complex in T-cell acute lymphoblastic leukemia

Vesna S. Stanulović,[1,2]* Sarah Binhassan,[1,2]* Budoor A. Jaber,[1] Shimaa Alazmi,[1,2] Fatma M.B. Saleman,[1,2] Sandeep Potluri,[1,2] Guy Pratt,[1,3] Christian Ludwig,[4] Douglas G. Ward[1] and Maarten Hoogenkamp[1,2]

[1]Department of Cancer and Genomics Sciences, School of Medical Sciences, College of Medicine and Health, University of Birmingham; [2]Birmingham Center for Genome Biology, University of Birmingham; [3]Center for Clinical Hematology, Queen Elizabeth Hospital Birmingham and [4]Department of Metabolism and Systems Sciences, School of Medical Sciences, College of Medicine and Health, University of Birmingham, Birmingham, UK

*VSS and SB contributed equally as first authors.

Abstract

The transcriptional mediator LIM domain only 2 (LMO2) forms a large multi-protein complex together with TAL1/LYL1, HEB/E2A, LDB1 and GATA. This complex regulates transcription from the onset of hematopoietic development and during differentiation. Chromosomal re-arrangements involving LMO2 and TAL1 are causative for T-cell lymphoblastic leukemia (T-ALL). We have identified Plant Homeodomain (PHD)-like Finger 6 (PHF6) as a new LMO2-interacting factor. Somatic mutations in *PHF6* have been found to occur in several types of leukemia. We show that PHF6 interacts with LMO2 as a part of the TAL1, GATA2, LDB1 complex in T-ALL and binds to the DNA. These findings show that PHF6 associates with the TAL1/LMO2/LDB1/GATA2 complex and regulates genes that have a major role in blood development, such as *SPI1* (PU.1).

Introduction

LIM domain only 2 (LMO2) is a protein that mediates interactions between the transcription factors GATA and a heterodimer composed of TAL1 or LYL1 and HEB or E2A. The essential role of the LMO2 complex in hematopoiesis is facilitated through the ability of LDB1 to trimerize and facilitate long range interactions between distal regulatory DNA elements.[1,2] This transcriptional complex regulates gene expression of many hematopoietic and non-hematopoietic genes.[3,4] In hematopoietic development, LMO2 is expressed from the haemangioblast stage, in hematopoietic stem and progenitor cells and downregulated during differentiation, except in the erythroid lineage.[5-7] During T-cell development, *LMO2* and *TAL1* are downregulated in double-negative (CD4⁻ CD8⁻) thymocytes.[8] Recurrent chromosomal aberrations involving *LMO2* and *TAL1* (45-50% of T-cell lymphoblastic leukemia [T-ALL]) result in their overexpression and lead to the development of T-ALL.[9-11] Gene expression profiling analysis on pediatric T-ALL patient samples showed that cases with TAL/LMO rearrangements constitute a separate cluster which accounts for 45-50% of T-ALL.[12,13] Human PHD finger protein 6 (PHF6) is a nuclear and nucleolar protein involved in suppression of rRNA expression,

chromatin remodeling, DNA repair, H2B ubiquitination, and transcriptional regulation.[14-20] Germline mutations in *PHF6* cause X-linked Börjeson-Forssman-Lehmann syndrome, a rare form of syndromic intellectual disability (BFLS; OMIM301900). Somatic *PHF6* mutations are found in 18% of pediatric and 36% of adult T-ALL, 2-3% of acute myeloid leukemia (AML), 16-55% of mixed phenotype acute leukemia (MPAL) and small numbers of chronic myeloid leukemia (CML) and high-grade B-cell lymphoma (HGBL).[14,21-26] While BFLS mutations are missense mutations occurring throughout the coding region or C-terminal nonsense mutations, the mutations observed in leukemia are predominantly nonsense and frame shifts spread throughout the coding region, leading to a loss of PHF6 protein expression.[16,21] Several *Phf6* knock-out studies in mouse hematopoietic stem cells (HSC) have shown that a lack of PHF6 leads to increased reconstitution capability and reduced number of double-negative (DN) T cells.[27-30]

Within T-ALL cases, PHF6 mutations were found to be positively associated with the TLX1/TLX3 subgroup and negatively-associated with the TAL/LMO subgroup of T-ALL, suggesting a possible functional relationship between the LMO2-complex and PHF6.[21,29,31] This study reveals the direct interaction between the LMO2 complex

and PHF6 in T-ALL and its importance for transcriptional regulation.

Methods

Cell culture
T-ALL cell lines were grown in RPMI1640 (Merck; Darmstadt, Germany) supplemented with 10% fetal calf serum (FCS), 5 U/mL penicillin/streptomycin, 2 mM GlutaMax (Thermo Scientific; Waltham, Massachusetts, USA) and 0.075 mM 1-Thyoglycerol (Merck). Knockout of *PHF6* in the ARR cell line was performed using commercially available Cas9, guideRNA and HDR plasmids (sc-413618) (Santa Cruz; Dallas, Texas, USA).

Patient samples
Patients' cells were from diagnostic samples from presentation cases, with ethical approval from the NHS Research Ethics Committee (reg.: 10/H1206/58). Blast cells were isolated from mononuclear cells using anti-hCD34 (T-ALL_1) or ant-hCD7 (T-ALL_2) MACS microbeads (Miltenyi; Bergisch Gladbach, Germany).

Immunoprecipitation
Nuclear extracts were prepared as described.[4] Protein-G Dynabeads (Thermo Scientific) were bound to immunoglobulin (Ig)G or antibody (*Online Supplementary Table S1*) in phosphate-buffered saline (PBS) containing 3% bovine serum albumin (Merck) for 30 minutes (min), at 4°C. Antibody-beads were washed in PBS and incubated with nuclear extract (100 μg NE per 10 μL beads) for 2 hour (h). After PBS washes, samples were analyzed by western blotting as described.[4] Input controls were 100 μg nuclear extract.

Mass spectrometry
Immunoprecipitated proteins were separated on polyacrylamide gels. Lanes were divided into 12 slices and processed with acetonitrile/ammonium bicarbonate (ACN/AB). Slices were destained in 50%/50 mM ACN/AB, equilibrated in 10%/50 mM ACN/AB and alkylated with 100 mM iodoacetamide. After washing in 10%/50 mM ACN/AB the gel slices were lyophilized and in-gel digested with sequencing-grade trypsin (Promega; Madison, Wisconsin, USA) in 10%/50 mM ACN/AB. Peptides were extracted with 1% formic acid/10% acetonitrile, lyophilized, dissolved in 1% formic acid and analyzed by LC-MS/MS using an Impact ESI-Q-TOF-MS (Bruker; Billerica, Massachusetts, USA). Data were searched against Swiss-Prot human databases using MASCOT with required peptide scores >25 and protein false discovery rate <1%.

Expression analysis
RNA isolation, cDNA preparation and quantification were performed as reported.[4] RNA sequencing (RNAseq) libraries were prepared using the TruSeq Stranded mRNA Kit (Illumina; San Diego, California, USA) and run on an Illumina NextSeq500 sequencer obtaining 150 nt paired-end reads.

Chromatin immunoprecipitation
Chromatin immunoprecipitation (ChIP) assays all targeted endogenous protein and were performed as described before.[32] Briefly, cells were crosslinked with 1% formaldehyde for 12 min at room temperature. Nuclei were isolated and chromatin was sonicated for 20 cycles of 30 seconds (s) on/30 s off on ice, using a Bioruptor (Diagenode; Liege, Belgium). Immunoprecipitation (IP) was carried out using protein-G Dynabeads, using 1 μg antibody (*Online Supplementary Table S1*) per 10 μL of beads. After elution, crosslinks were reversed overnight at 65°C, after which the DNA was isolated using Ampure XP beads (Beckman-Coulter; Brea, California, USA). Resulting material was analyzed by quantitative polymerase chain reaction (qPCR) (primers in *Online Supplementary Table S1*) on an ABI7500 real-time PCR machine, or ChIPseq libraries were analyzed on an Illumina NextSeq500 obtaining 75 nt single-end reads.

Cellular staining
Cells were incubated with 1 μg/mL primary antibody, followed by the appropriate Alexa-conjugated secondary antibody (Molecular Probes; Eugene, Oregon, USA). Confocal images were obtained on a LSM880 Confocal Microscope (Zeiss; Oberkochen, Germany).

Proximity ligation assay
Proximity ligation assay (PLA) was conducted using a Duolink-PLA kit (Merck) according to the manufacturer's protocol.

Nuclear magnetic imaging resonance spectroscopy
Nuclear magnetic imaging resonance (NMR) spectroscopy data acquired as described before[33] and analyzed using MetaboLabPy.[34]

Bioinformatics
Next-generation sequencing analysis was performed on usegalaxy. Detailed information is available as *Online Supplementary Appendix*.

Results

PHF6 is an LMO2 binding partner
To identify novel LMO2-interacting partners we chose four T-ALL cell lines, which express LMO2 and its binding partners TAL1, GATA2 and LDB1. The ARR cell line was chosen as its development is blocked at the initial stages of the T-cell development when LMO2 and its partners have an important role in regulating gene expression.

DU.528, HSB2 and CCRF-CEM cell lines were selected because they belong to the TAL/LMO subgroup of T-ALL and carry the SIL-TAL1 chromosomal rearrangement but are arrested at different stages of T-cell development (*Online Supplementary Figure S1*).

IP using αLMO2 antibody and nuclear protein extracts of four T-ALL cell lines (ARR, DU.528, HSB2, CCRF-CEM), followed my mass spectrometry (MS) found PHF6 in three of four LMO2-IP samples and not in the negative control IgG samples, suggesting that PHF6 interacts with LMO2 (*Online Supplementary Table S2*).

A negative association between the TAL/LMO oncogenic subtype of T-ALL and the occurrence of PHF6 mutations has previously been reported.[21,31] Our analysis of a large data set of 1309 T-ALL cases recently reported in Pölönen *et al.* (2024) confirms this (*Online Supplementary Figure S2*). The results show that while 26% of total T-ALL patient samples harbor a mutated *PHF6* gene, a statistically significant underrepresentation of *PHF6* mutations was observed in the TAL1 αβ-like (5%), and TAL1 DP-like subtypes (6%), which in addition to TAL-rearrangements, together also harbor >90% (303/333) of samples with *LMO2* alterations. The identified LMO2 γδ-like and SATG2/LMO2 subgroups also showed an underrepresentation of *PHF6* mutations, however, these subgroups were too small to confer statistical significance. Underrepresentation of *PHF6* mutations was additionally observed in the KMT2A and MLLT10 subtypes. Conversely, a positive correlation with mutated *PHF6* genes was observed for the TLX1 (59%), TLX3 (69%) and HoxA9 TCR (62%) subtypes, as reported in the publication.[35]

RNAseq data and qPCR showed that *PHF6* mRNA is expressed in all four T-ALL cell lines, while western blotting confirmed the presence of PHF6 protein (Figure 1A-C). Additionally, immunofluorescent staining showed overlapping nuclear localization of PHF6 and LMO2 (Figure 1D). T-ALL cell lines Molt4 and Jurkat were also analyzed by western blotting, which showed that both express PHF6, while Molt4 did not express TAL1 and Jurkat did not express LMO2 (Figure 1C).

Co-IP of LMO2 and PHF6, followed by western blotting, confirmed the PHF6-LMO2 interaction (Figure 1E). Additional validation of the interaction was achieved in proximal ligation assays combining αLMO2 and αPHF6 antibodies in HSB2 and CCRF-CEM cells (Figure 1F). The combination of αLMO2 and αLDB1 was used as a positive control, whereas αLMO2 and IgG was used as a technical negative control. As a biological negative control, we performed the same assay on Jurkat T-ALL cells, which do not express LMO2. Taken together, we conclude that PHF6 interacts with LMO2.

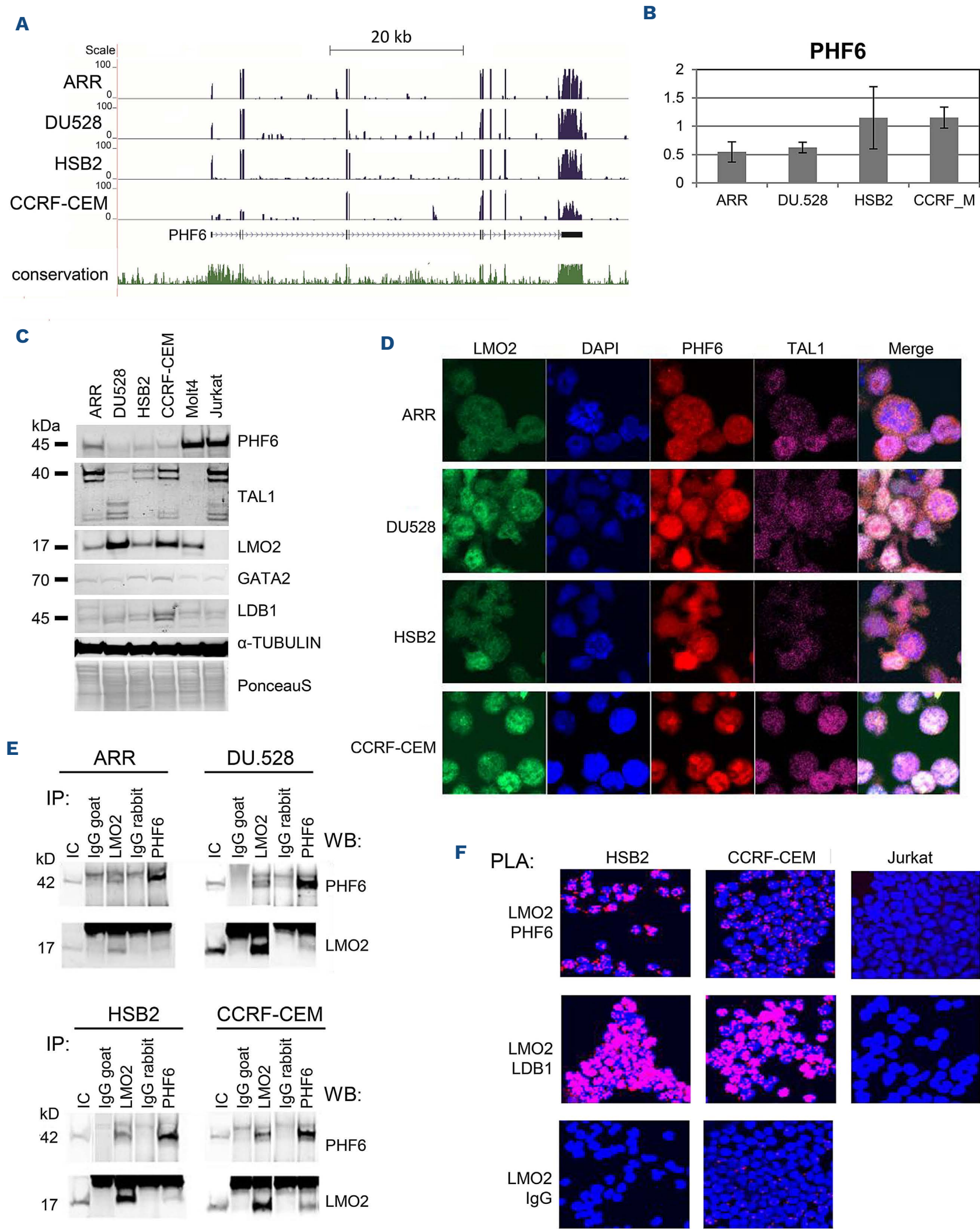### PHF6 binds LMO2/TAL1/GATA2/LDB1 target sites in T-cell lymphoblastic leukemia cells

To identify DNA elements that PHF6 associates with, we performed PHF6 ChIPseq and identified PHF6 peaks in all four T-ALL cell lines. Analyses showed that 169 regions

were occupied by PHF6 in all four cell lines (ADHC-PHF6 peaks) and an additional 2,183 regions overlapped between at least two (Figure 2A). *De novo* motif analysis of the ADHC-PHF6 peaks found enrichment of binding sites for ETS/IKZF, GATA and RUNX transcription factors (Figure 2B). Distribution analyses identified that 78% resided in known enhancers and only 9% in promoter regions, with the majority of ADHC-PHF6 peaks associated with two genes within 1 Mb distance up and down from the peaks (Figure 2C). Among the genes, 293 have a known function and annotation of biological processes for these genes found the greatest enrichment for the genes that are involved in regulation of transcription (15%, 43 genes), followed by differentiation, signal transduction, endocytoses, as well as differentiation of T-helper type 2 (Th2) and regulatory T (Treg) cells (Figure 2D; *Online Supplementary Table S3*). Next, we investigated the interaction of PHF6 with the LMO2 complex by performing further ChIPseq experiments. Intersection of the PHF6 peaks with those for LMO2, TAL1, GATA2 and LDB1 revealed that for each cell line many peaks showed an overlap between all five ChIPseq samples (*Online Supplementary Figure S3A*) and that at least 44% of PHF6 peaks co-localized with at least one of the members of the LMO2-complex (*Online Supplementary Figure S3A, B*). Furthermore, of the above-mentioned ADHC-PHF6 peaks, >90% were also bound by LMO2, TAL1, LDB1 and GATA2 (154 of 169) in all four cell lines (Figure 3A). In contrast to this, DNA elements that were only occupied by PHF6 were cell line specific, as none was common for all four cell lines (Figure 3B). The RUNX1 intronic enhancer is one of the DNA fragments found to be bound by the PHF6-LMO2 complex in all four T-ALL cell lines and this result was confirmed in independent ChIP experiments quantified by qPCR (Figure 3C; *Online Supplementary Figure S4*). Heat maps illustrating the ChIPseq overlap with PHF6 peaks, ranked according to the PHF6-peak intensity, showed the co-localization and that the intensity of PHF6 binding to the DNA elements correlates with the intensity of LMO2, TAL1, LDB1 and GATA2 binding (Figure 3D; *Online Supplementary Figure S5*). Correlation analysis showed that TAL1 and LMO2-bound regions cluster together while PHF6-occupied regions have a higher correlation with GATA2 and LDB1 (*Online Supplementary Figure S6*). Together, these results confirm that PHF6 interacts with the DNA-bound LMO2/TAL1/LDB1/GATA2 transcriptional complex.

The PHF6/LMO2/TAL1/LDB1/GATA2-complex is associated with different genes than the *LMO2/TAL1/LDB1/GATA2-complex* in T-cell lymphoblastic leukemia cells

To explore the role of PHF6 in the regulation of gene expression, we separately analyzed the regions occupied by *PHF6/LMO2/TAL1/LDB1/GATA2* (PHF6-LTLG) and LMO2/TAL1/LDB1/GATA2 without PHF6 (LTLG) and integrated these results with our previously published RNAseq data.[33] PHF6-LTLG peaks were associated with 375 transcriptionally active

Figure 1. PHF6 interacts with LMO2 and binds LMO2/TAL1/GATA2/LDB1 target sites in T-cell lymphoblastic leukemia cell lines.
(A) RNA sequencing (RNAseq) analyses of PHF6 expression in T-cell lymphoblastic leukemia (T-ALL) cell lines. Screenshot from the UCSC browser illustrating the distribution of reads over the *PHF6* gene. Uniform y-axis scales were used. (B) *PHF6* mRNA expression level assessed by quantitative polymerase chain reaction (qPCR) and relative to *GAPDH* mRNA level. Data points are

the mean of at least 3 independent samples measured in duplicate ± standard deviation (StDev). (C) Western blots showing protein levels of PHF6, TAL1, LMO2, GATA2 and LDB1 in T-ALL cell lines with αtubulin and PonceauS illustrating equal loading. (D) Confocal microscopy of immuno-fluorescently stained T-ALL cell lines, showing DNA (DAPI-blue), LMO2 (green), PHF6 (red) and TAL1 (magenta); magnification 65x. (E) PHF6 and LMO2 co-immunoprecipitation (IP) using nuclear extracts from T-ALL cell lines. Indicated immunoglobulin (Ig)G were used as a control. IC: input control. Presence of PHF6 and LMO2 was visualized by western blotting (WB). (F) Proximal ligation assay (PLA) showing LMO2-PHF6 interaction in HSB2, CCRF-CEM cells but not in the LMO2-negative T-ALL cell line Jurkat. Confocal microscopy images show the nuclei visualized by DAPI (blue), while magenta staining indicates places of interaction between the indicated antibodies. LMO2-LDB1 PLA was used as a positive control, while LMO2-rabbit IgG was a technical negative control, and Jurkat was a biological negative control.
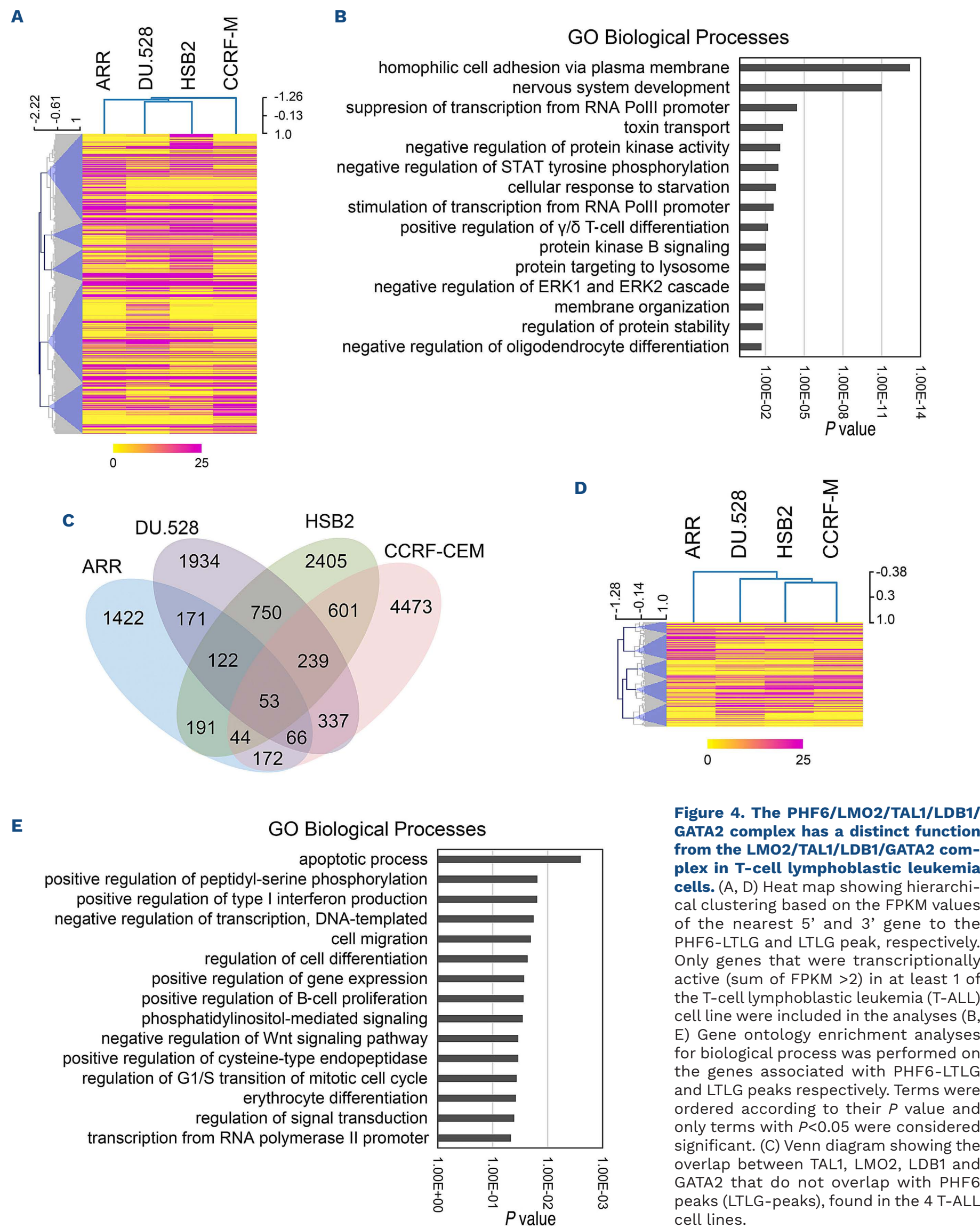
A

B

Motif	*P* value	Name	N
	8.6e-042	ETS / IKZF	92/169
	5.8e-025	GATA	72/169
	1.4e-017	RUNX	74/169
	4.5e-002		37/169

C

D
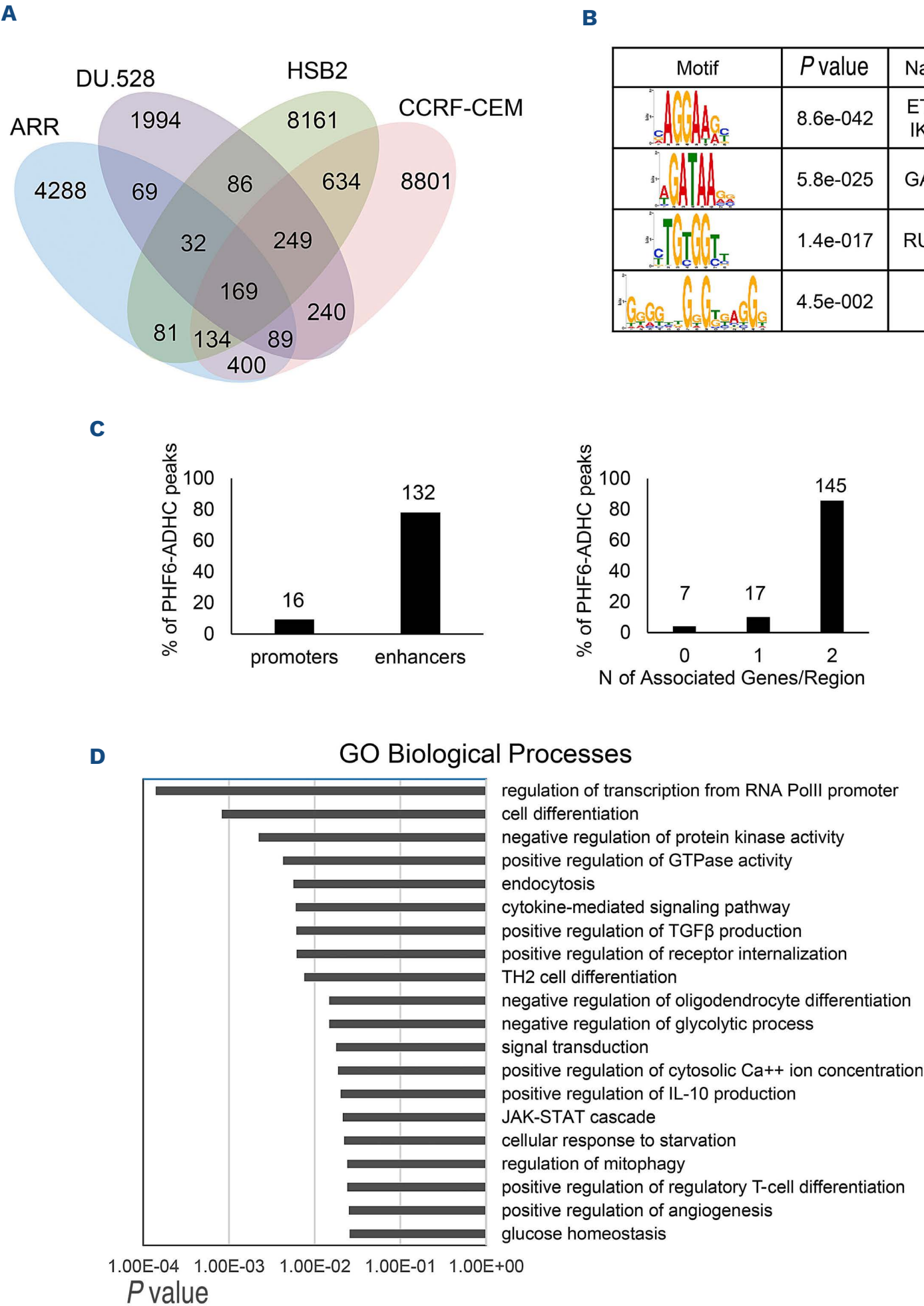
Figure 2. PHF6 co-localizes with the DNA-binding LMO2 complex. (A) Venn diagram showing the overlap between PHF6 chromatin immunoprecipitation sequencing (ChIPseq) peaks in the four T-cell lymphoblastic leukemia (T-ALL) cell lines. (B) *De novo* motif analysis showing enriched motifs within the 169 ADHC-PHF6 ChIPseq peaks. Transcription factors that are predicted to bind these motifs, *P* values showing the significance of the enrichment and the number of peaks found to contain the motif, are indicated. (C) Distribution of ADHC-PHF6 peaks over known promoters and enhancers, and the number of genes that these peaks associate with within 1 Mb up and down from the peak coordinates, as identified by GREAT software. (D) GREAT gene ontology enrichment analysis for biological process performed on genomic regions of the PHF6 ChIPseq peaks.
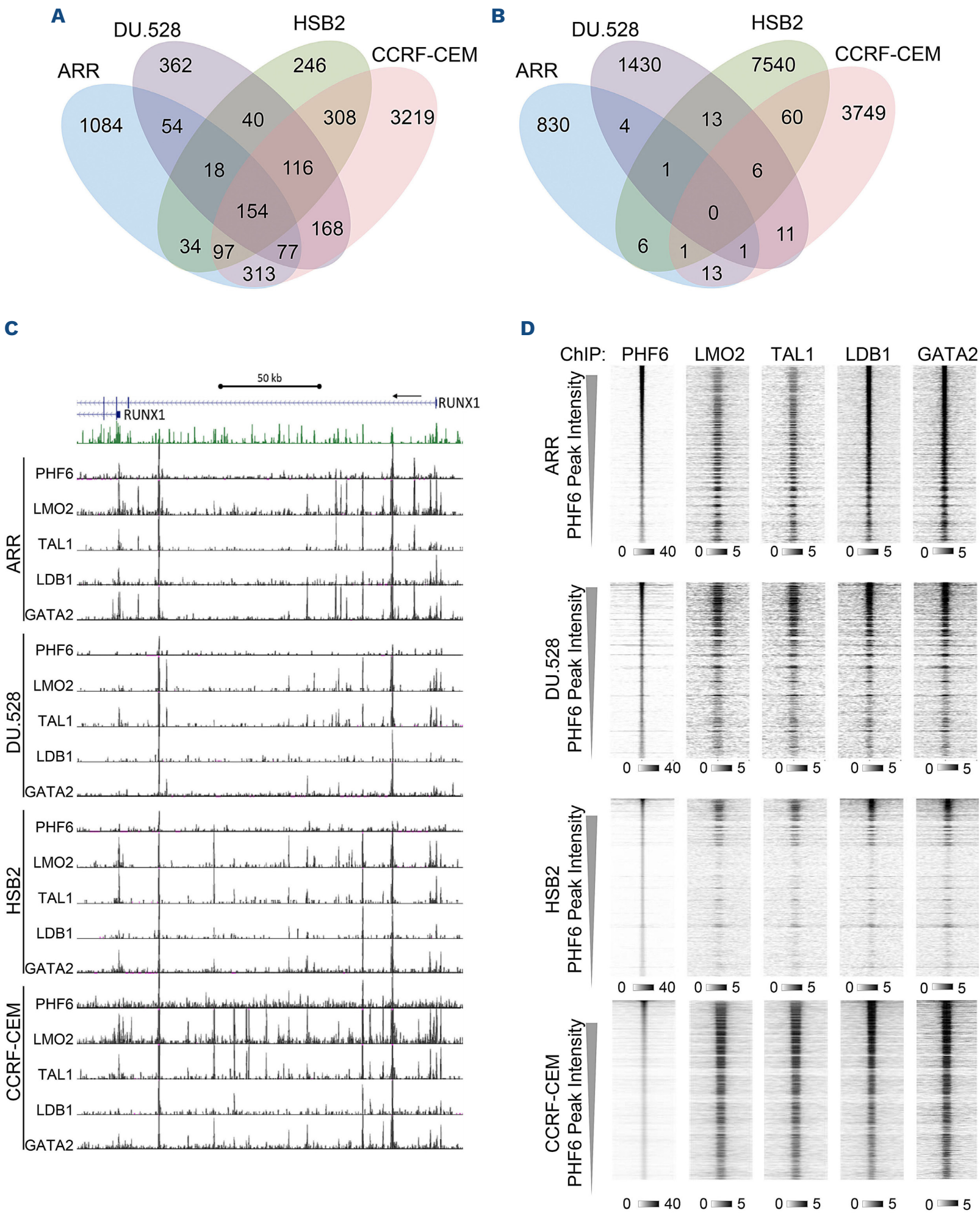
Figure 3. PHF6 co-localizes with the LMO2 complex. (A) Venn diagram showing the overlap of PHF6 co-localization with LMO2, TAL1, LDB1 and GATA2 chromatin immunoprecipitation sequencing (ChIPseq) peaks between the four T-cell lymphoblastic leukemia (T-ALL) cell lines. (B) Venn diagram showing the overlap of PHF6-only peaks between the four T-ALL cell lines. The first column indicates the name of the cell line and the total number of PHF6-only peaks and subsequent columns indicate the overlaps between cell lines and the number of times they occurred. (C) Screenshot from the UCSC browser showing PHF6, LMO2, TAL1, LDB1 and GATA2 binding profiles at the *RUNX1* locus in T-ALL cell lines. Uniform y-axis scales were used for all ChIPseq tracks. (D) Heat maps showing PHF6, LMO2, TAL1, LDB1, GATA2 ChIPseq results ranked according to the intensity of PHF6 binding, depicting windows from −2 kb to +2 kb around the center of the PHF6 peaks. The intensity of the greyscale is indicated below the heat maps.

A

B

GO Biological Processes

homophilic cell adhesion via plasma membrane
nervous system development
suppresion of transcription from RNA PolII promoter
toxin transport
negative regulation of protein kinase activity
negative regulation of STAT tyrosine phosphorylation
cellular response to starvation
stimulation of transcription from RNA PolII promoter
positive regulation of γ/δ T-cell differentiation
protein kinase B signaling
protein targeting to lysosome
negative regulation of ERK1 and ERK2 cascade
membrane organization
regulation of protein stability
negative regulation of oligodendrocyte differentiation

P value

C

D

E

GO Biological Processes

apoptotic process
positive regulation of peptidyl-serine phosphorylation
positive regulation of type I interferon production
negative regulation of transcription, DNA-templated
cell migration
regulation of cell differentiation
positive regulation of gene expression
positive regulation of B-cell proliferation
phosphatidylinositol-mediated signaling
negative regulation of Wnt signaling pathway
positive regulation of cysteine-type endopeptidase
regulation of G1/S transition of mitotic cell cycle
erythrocyte differentiation
regulation of signal transduction
transcription from RNA polymerase II promoter

P value

Figure 4. The PHF6/LMO2/TAL1/LDB1/ GATA2 complex has a distinct function from the LMO2/TAL1/LDB1/GATA2 complex in T-cell lymphoblastic leukemia cells. (A, D) Heat map showing hierarchical clustering based on the FPKM values of the nearest 5' and 3' gene to the PHF6-LTLG and LTLG peak, respectively. Only genes that were transcriptionally active (sum of FPKM >2) in at least 1 of the T-cell lymphoblastic leukemia (T-ALL) cell line were included in the analyses (B, E) Gene ontology enrichment analyses for biological process was performed on the genes associated with PHF6-LTLG and LTLG peaks respectively. Terms were ordered according to their *P* value and only terms with *P*<0.05 were considered significant. (C) Venn diagram showing the overlap between TAL1, LMO2, LDB1 and GATA2 that do not overlap with PHF6 peaks (LTLG-peaks), found in the 4 T-ALL cell lines.

genes with a known function, of which the majority was expressed in all four T-ALL cell lines (Figure 4A; *Online Supplementary Table S4*). Functional annotation of these genes identified 45 terms with the highest enrichment found for homophilic cell adhesion, neural development, regulation of transcription and T-cell development (Figure 4B). Amongst the genes that were involved in transcriptional regulation were many hematopoietic transcription factors, such as: *CEBPA*, *RUNX1*, *GATA3*, *ETS2*, *STAT3*, *STAT5A*, *MYBL2*, *JARID2*, *ERG*, *ID2* and *NFE2* (*Online Supplementary Table S4*). Analyses of the LTLG-binding sites without PHF6 found 53 regions occupied by the four factors in all four T-ALL cell lines and associated with 135 genes (Figure 4C; *Online Supplementary Table S5*). Expression clustered into ARR and a separate DU.528, HSB2 and CCRF-CEM (SIL-TAL1) group, which suggests distinct influence of the LTLG-complex on the gene expression in ETP than in SIL-TAL1 leukemia (Figure 4D). Functional annotation showed that LTLG-associated genes were mainly involved in apoptosis, signaling, interferon production, regulation of transcription, development and differentiation. In line with the already known function of LMO2, B-cell proliferation and erythrocyte differentiation were also amongst the significantly enriched terms. (Figure 4E; *Online Supplementary Table S5*). Based on this we conclude that the PHF6-LTLG and LTLG-complexes are located in the vicinity of two different sets of genes and that these genes have distinct functions. This may indicate that the two complexes have different roles in the regulation of gene expression.

PHF6 binds different DNA elements in different T-cell lymphoblastic leukemia subtypes

To investigate the importance of the presence of LMO2 and TAL1 for the interaction of PHF6, we performed ChIPseq in Molt-4 (lacking TAL1) and Jurkat (lacking LMO2) cell lines. Additionally, we included two non-SIL-TAL1 patient samples (T-ALL1, 46 XY normal karyotype; T-ALL2, 46XY t(10,14) involving TLX1) from which we isolated the T-ALL blasts and performed ChIPseq. Correlation analysis of all the obtained PHF6 ChIPseq data showed distinctive grouping of MOLT-4, Jurkat, T-ALL1 and T-ALL2 (MJT1T2), separate from the ARR and SIL-TAL1 cell lines DU.528, HSB2 and CCRF-CEM (Figure 5A). Intersection of the PHF6 peaks found that 514 regions were occupied in all four cell lines of the MJT1T2 group (*Online Supplementary Table S6*). Reflecting the correlation analyses, none of the MJT1T2-PHF6 peaks overlapped with the ADHC-PHF6 peaks, showing that PHF6 has a different binding pattern in the MJT1T2 group. In line with this, motif analysis found distinctly different motifs compared to the ADHC-PHF6 peaks, with enrichment for LHX8, SOX, NFAT, IRF and STAT (Figure 5B). Analysis of the function of the MJT1T2-PHF6 peaks found that almost half of these DNA elements did not associate with any gene within 1 Mb distance, in contrast to ADHC-PHF6 peaks where the majority associated with three genes (Figure 5C;

Figure 2C). Functional annotation of the MJT1T2-PHF6-associated genes found significant enrichment for oxidative respiration, apoptosis and mitochondrial function (Figure 5D; *Online Supplementary Table S6*).
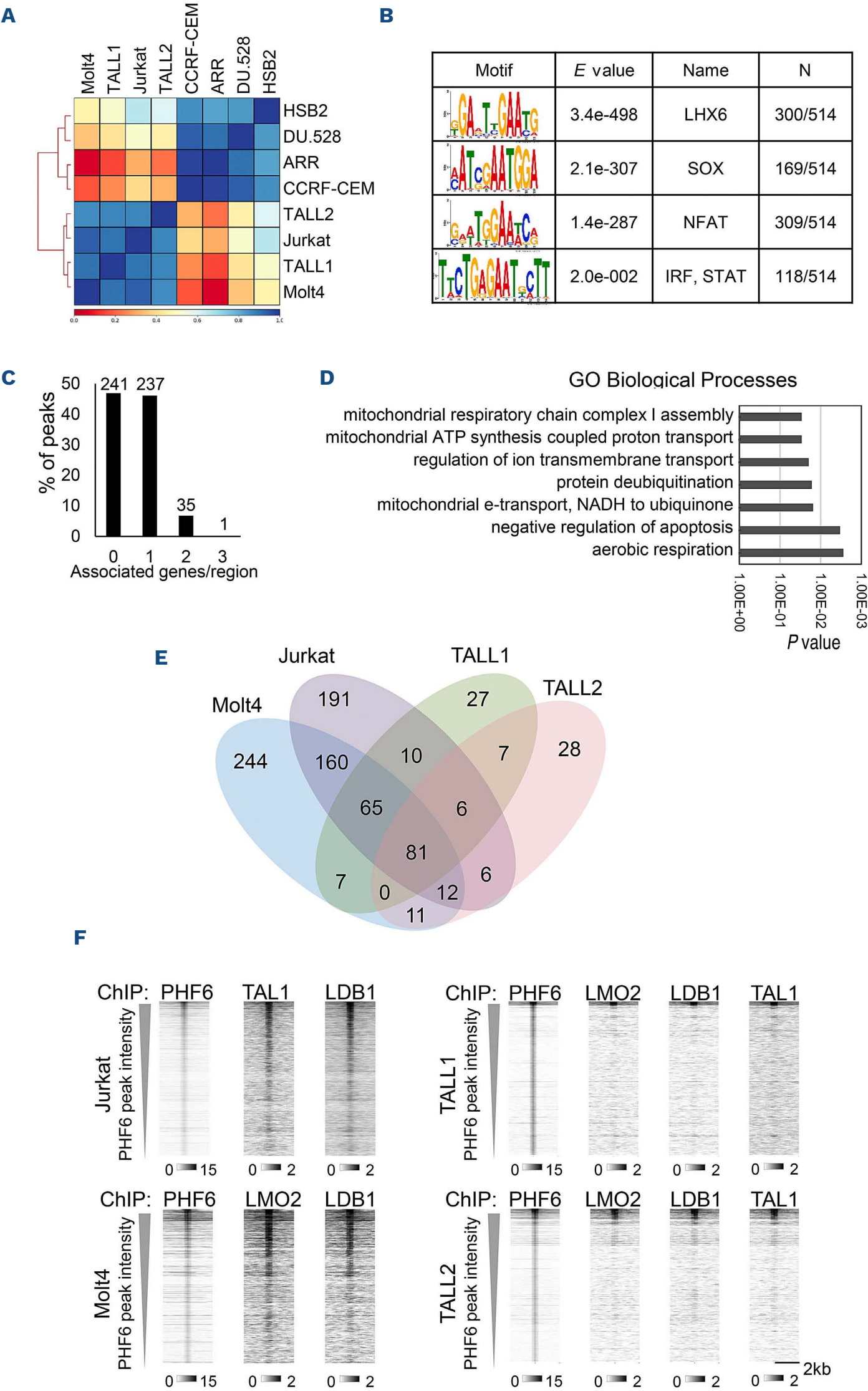
To establish if PHF6 co-localized with LMO2, TAL1 and LDB1 in the MJT1T2 group, ChIPseq experiments were performed for these proteins. The analysis showed overlap, but to a much lower extent than observed in the ADHC group, with only 81 peaks shared between MOLT-4, Jurkat, T-ALL1 and T-ALL2 (Figure 5E). The number of PHF6/LMO2/TAL1/LDB1 peaks was higher in Jurkat and MOLT-4 than in T-ALL1 and T-ALL2. The overlays showed that the intensity of the PHF6 peaks correlated with the strength of the TAL1, LMO2 and LDB1 binding, even in the patient samples where only a small number of overlapping peaks were identified (Figure 5F). These results show that PHF6 is associated with LMO2 and its binding partners in all tested T-ALL models but the regions that are occupied by the PHF6 are different between the ADHC and MJT1T2 groups, suggesting both a difference in prevalence of the interaction and a distinct role for PHF6 in ETP and SIL-TAL1 T-ALL compared to other T-ALL.

Loss of *PHF6* leads to changes in gene regulatory networks and chromatin accessibility

To further explore the function of *PHF6* in T-ALL, we generated PHF6 knock-out ARR cells using CRISPR-Cas9 technology (Figure 6A). RNAseq analysis performed on the *PHF6*[-/-] ARR and Cas9 control ARR cells resulted in 597 statistically significant differentially expressed genes (DEG) (Figure 6B). Gene ontology analysis found that the loss of *PHF6* led to downregulation of genes involved in transcription and hematopoiesis, and upregulation of genes with a function in signal transduction, endocytosis, cell migration, erythropoiesis and T-cell differentiation (Figure 6C). One of the most affected genes by the loss of *PHF6* was *SPI1*, encoding transcription factor PU.1,[36] which expression was reduced 32-fold (from 227 to 7, average feature counts) at the mRNA level, followed by a reduction in PU.1 protein levels (Figure 6A, D). Suppression was also observed for transcription factors CEBPα, GATA1, and JUN that directly interact with PU.1 to regulate erythroid *versus* myeloid lineage determination.[37-39] Additionally, we observed reduced expression of *CSF2RB* (53-fold, from 147 to 2.7), which encodes a common subunit for the GM-CSFR, IL3R and IL5R receptors.[40] Among other genes that were affected by the loss of *PHF6*, we found *ITGA2B* (CD41), *MPO, IL1B, CD1D, CD5* and *CD96*. All these genes have essential roles in myeloid/lymphoid differentiation and function (*Online Supplementary Table S7*).

To identify high-confidence *PHF6*-target genes we integrated *PHF6* ChIPseq with the RNAseq data and we found that *PHF6*-peaks were located in the vicinity of 228 DEG with 213 of these, located near a *PHF6*-peak that was also bound by at least one TAL1/LMO2/LDB1/GATA2 complex member (Figure 7A). Of these, 22 were genes encoding transcription

A

B

| Motif	*E* value	Name	N
	3.4e-498	LHX6	300/514
	2.1e-307	SOX	169/514
	1.4e-287	NFAT	309/514
	2.0e-002	IRF, STAT	118/514

C

D

E

F

Figure 5. PHF6 binds to distinct DNA elements in the presence and absence of the LMO2-containing transcriptional complex. (A) Clustered heatmap based on the Pearson Correlation analysis for the PHF6 chromatin immunoprecipitation sequencing (ChIPseq) bound DNA elements found in indicated cell lines and two T-cell lymphoblastic leukemia (T-ALL) patient samples, T-ALL1 and T-ALL2. Dendrogram on the right illustrates the hierarchical relationships between the samples while the color scale shows the relation between the color and the correlation value. (B) Motifs enriched in the 514 PHF6 peaks found in Jurkat, MOLT-4, T-ALL1 and T-ALL2. Transcription factors known to bind to the motifs, *P* values showing the significance of the enrichment and the number of peaks found to contain the motif, are indicated. (C) Percent of MJT1T2-PHF6 bound DNA elements that associates with 0, 1, 2 or 3 genes within 1 Mb from the peak coordinates, identified by GREAT software. (D) Enriched biological functions for the MJT1T2-PHF6 peaks, based on the gene list identified by GREAT and analyzed in DAVID database. Enriched terms for the biological processes are listed according to the *P* value. (E) Venn diagram showing the number of common PHF6, TAL1, LMO2, LDB1 peaks that are shared between Jurkat, MOLT-4, T-ALL1 and T-ALL2. (F) Heat maps showing overlays of the PHF6, LMO2, TAL1, LDB1, GATA2 ChIPseq results. Matrix was derived from the coordinates of the PHF6 peaks ±2 kb around the center of the peaks, ranked according to the score of the intensity. Aligned reads from PHF6, LMO2, TAL1, LDB1, GATA2 ChIPseq were overlayed and the intensity of pixel color corelates with the number of reads found at each genomic position. The intensity of the scale is indicated below the heat maps.

A

C

GO for Genes Supressed in PHF6$^{-/-}$

GO for Genes Upregulated in PHF6$^{-/-}$

B

D
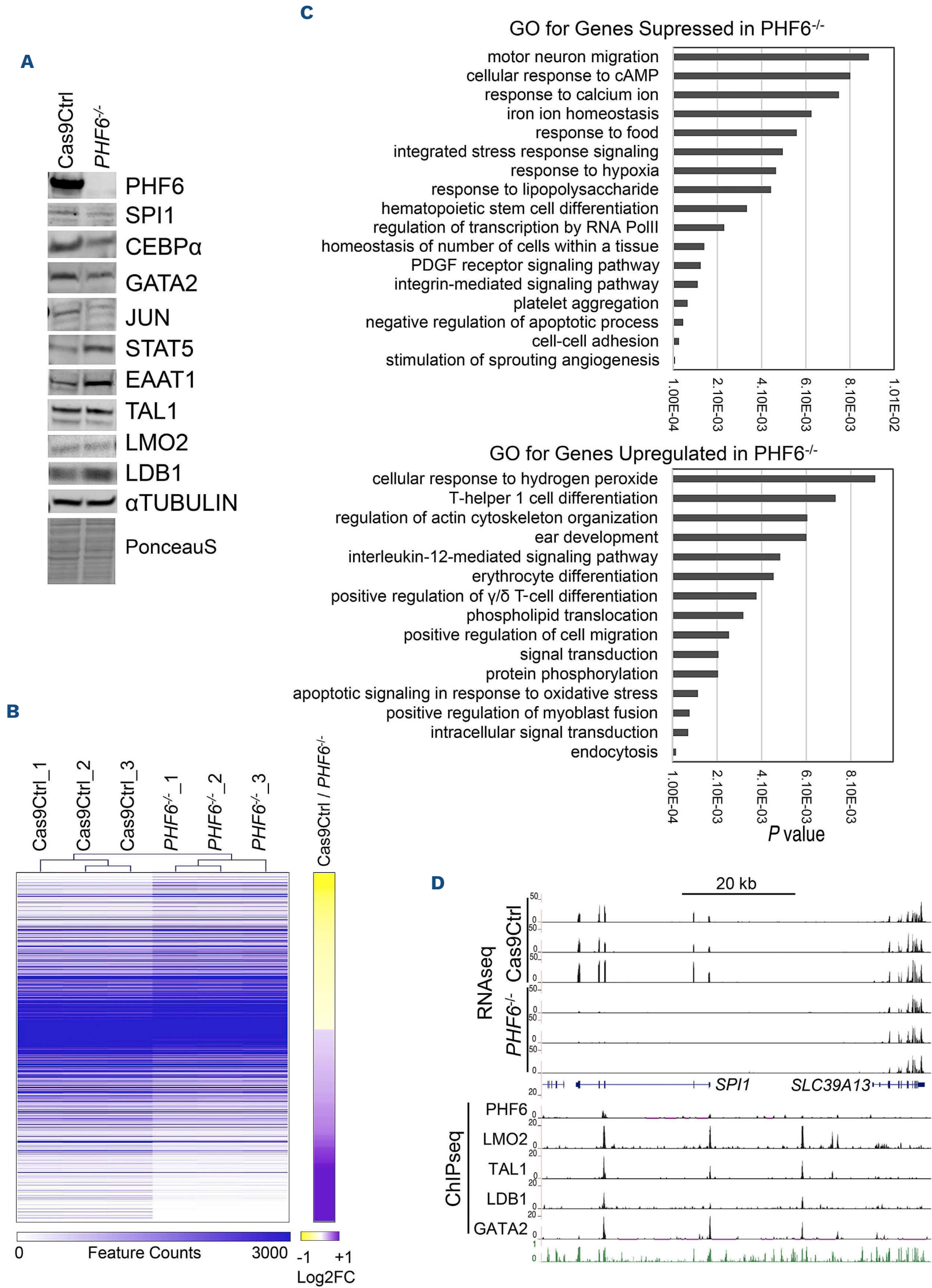
Continued on following page.

Figure 6. Loss of PHF6 leads to changes in gene expression of nearby genes. (A) Western blots showing protein levels of several proteins that showed down-regulation (PHF6, SPI1, C/EBPα, GATA2, JUN), up-regulation (STAT5, EAAT1), or no change (TAL1, LMO2, LDB1, αTubulin) in Cas9 control ARR and *PHF6*⁻/⁻ ARR cell lines. (B) Heat map showing gene expression levels of the differentially expressed genes, ranked according to the fold change between the Cas9 control ARR and *PHF6*⁻/⁻ ARR cell, as shown by the color bar alongside. Each genotype was measured in triplicate by RNA sequencing (RNAseq). (C) Gene ontology enrichment analysis for biological process was performed on the downregulated genes (top) or upregulated genes (bottom) in *PHF6*⁻/⁻ cells as compared to the Cas9 control cells. Terms were ordered according to their *P* value and only terms with *P*<0.01 are shown. (D) Screenshot from the UCSC browser at the *SPI1* locus showing RNAseq expression data of Cas9 control ARR and *PHF6*⁻/⁻ ARR cells, and PHF6, LMO2, TAL1, LDB1 and GATA2 binding profiles in the parental ARR cell line. Uniform y-axis scales were used within the RNAseq and chromatin immunoprecipitation sequencing (ChIPseq) tracks.
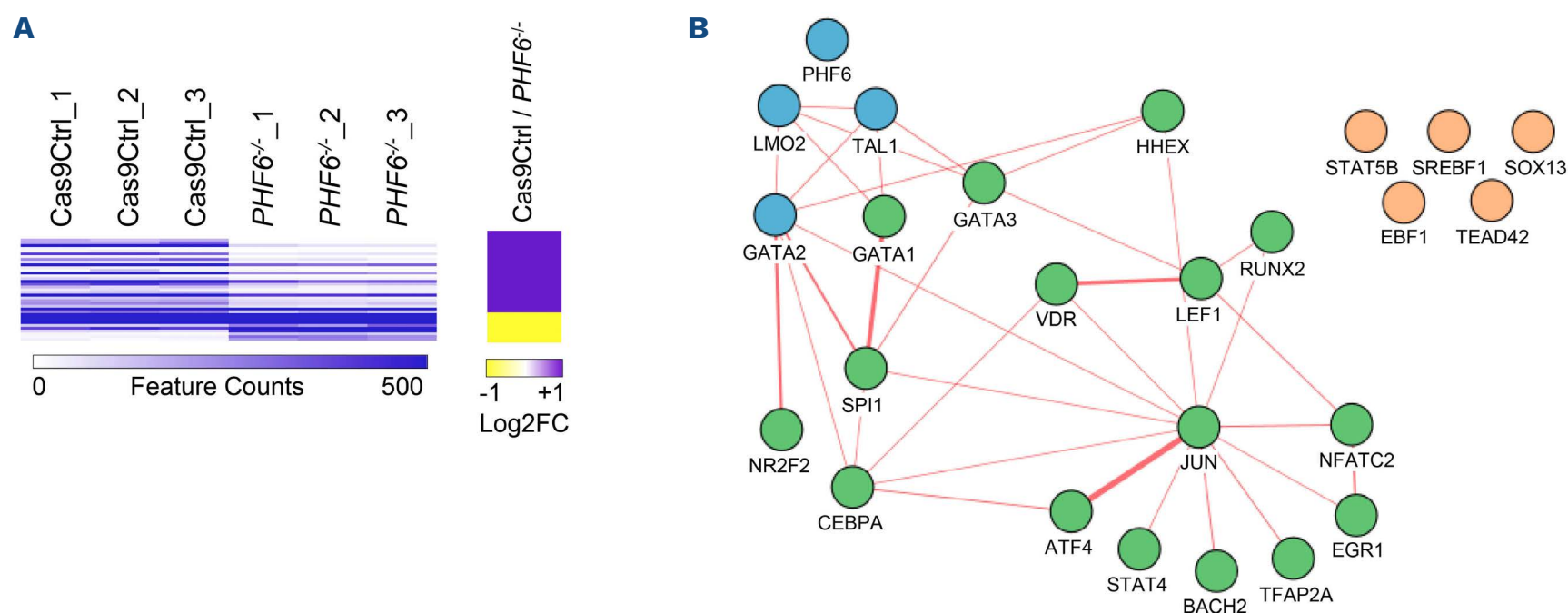
Figure 7. Loss of PHF6 leads to changes in the gene regulatory network. (A) Heat map showing gene expression levels of differentially expressed genes (log2 fold change >1), ranked according to the fold change between the Cas9 control ARR and *PHF6*⁻/⁻ ARR cell, as shown by the color bar alongside. Each genotype was measured in triplicate by RNAseq. (B) High confidence and significantly differentially expressed PHF6-target genes encoding transcription factors (nodes) were used to conduct the transcriptional regulatory architecture analysis, mapping physical protein-protein interactions (edges), using Cytoscape with the stringApp. PHF6, LMO2, and TAL1 were manually added to the list.

factors, including *SPI1*, *CEBPα* and *GATA2* (Figures 6A and 7B; *Online Supplementary Table S8*). The majority of the identified transcription factors are known to interact and/or regulate each other during hematopoietic development. These findings show that PHF6 is part of the TAL1/LMO2/LDB1/GATA2 complex, regulating genes that have a major role in blood development.

Finally, we investigated whether the absence of PHF6 in ARR cells impacts their response to therapeutic treatments. We cultured the cells in the presence or absence of dexamethasone, doxorubicin, vincristine, or asparaginase (Figure 8A). The vehicle controls showed that there was no difference in proliferation between the two cells lines without treatment. The only difference we observed was that *PHF6*⁻/⁻ ARR were more resistant to asparaginase than Cas9 control ARR cells. This is interesting, because we previously found that the ARR cell line does not take up asparagine.[33] However, asparaginase also has glutaminase activity (Figure 8B) and we showed that T-ALL cells are dependent on glutamine as a substrate for nucleotide production. Furthermore, most T-ALL aberrantly express the

amino acid transporter EAAT1 (gene name *SLC1A3*), which is required for this process.[33] In *PHF6*⁻/⁻ ARR, we observed that EAAT1 is upregulated at both the mRNA level and protein level (Figure 6A), explaining why these cells were able to retain proliferation under reduced glutamine availability.

Discussion

In addition to the common binding partners, the LMO2/LDB1-complex has been shown to interact with SP1, ELK1, NFATC1, LEF1 and SSBP.[41-43] In this paper, we identified PHF6 as a novel LMO2-interacting partner and show that PHF6 and LMO2/TAL1/LDB1 bind the same DNA fragments in T-ALL. In line with this, motif analysis following PHF6 ChIPseq in the ETP and SIL-TAL1 cell lines showed clear overlap with those expected from LMO2 or TAL1 ChIPseq data.[4,44] The finding that PHF6 interacts with the LMO2 complex likely explains why PHF6 mutations are underrepresented in the TAL/LMO subgroup of T-ALL patients, as functional PHF6 is important for TAL/LMO driven oncogenesis.
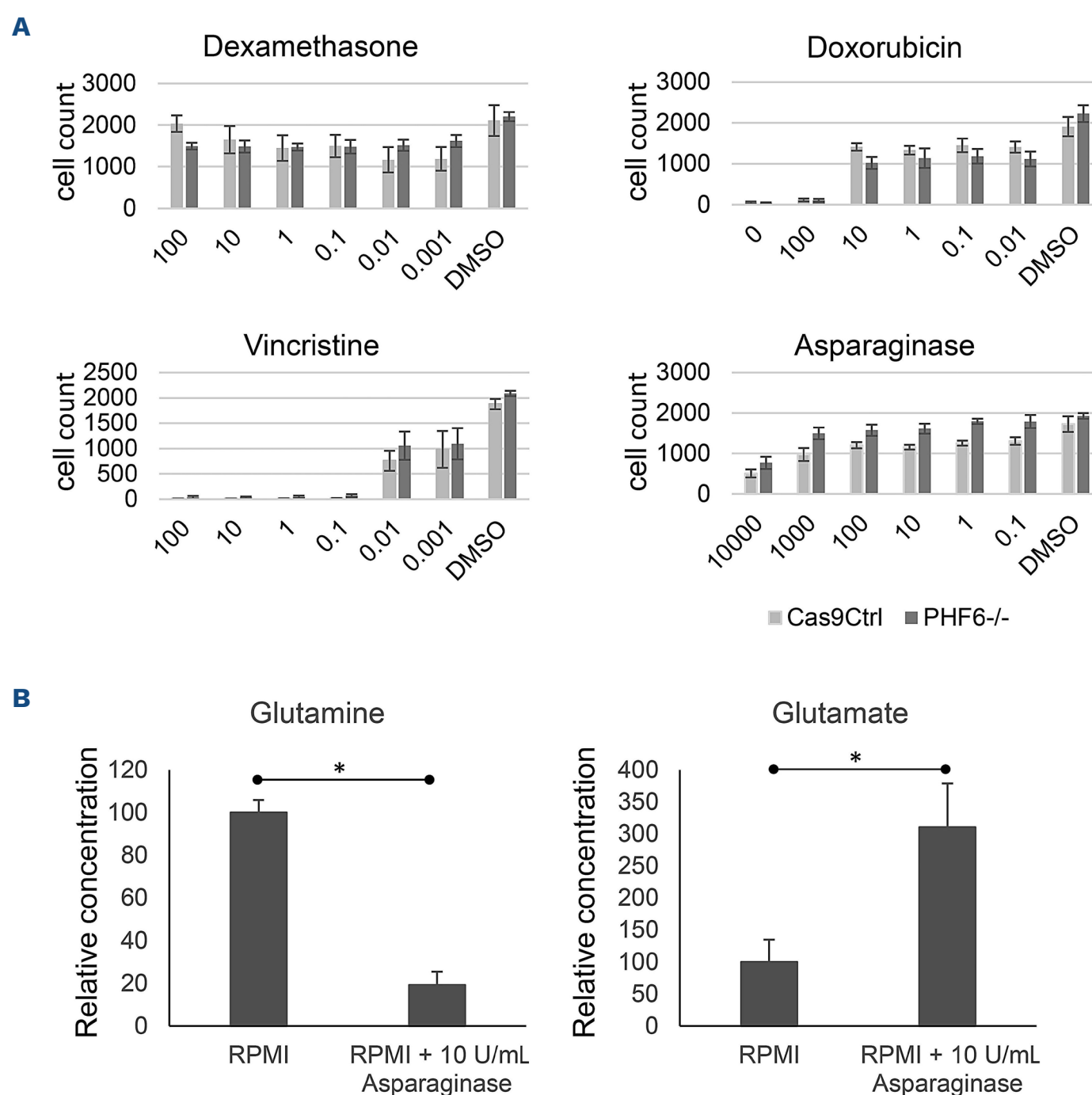
Figure 8. Loss of PHF6 reduces response to asparaginase. (A) Cas9 control ARR and *PHF6⁻ᐟ⁻* ARR cells were grown in the presence of dexamethasone, doxorubicin, vincristine, and asparaginase, or vehicle control (dimethyl sulfoxide [DMSO]). X-axes show µM drug concentrations, except for asparaginase, which shows Units/mL. Data points are the mean of 3 independent samples. (B) RPMI-1640 medium was incubated at 37°C for 24 hours (h) in the presence or absence of 10 U/mL asparaginase. Glutamine and glutamate (product of glutaminase) concentrations were measured by nuclear magnetic resonance spectroscopy. Data points are the mean of 3 independent samples and statistically significant differences ($P<0.05$) are indicated by an asterisk.

Several publications reported different roles for PHF6, including the interaction with the NuRD complex, which facilitates ATP-dependent chromatin remodeling and histone deacetylation, leading to suppression of gene expression and is composed of CHD3/4 (Mi-2a/b), RBBP and HDAC1/2.[16,45,46] We confirmed the HDAC1 interaction with PHF6 and LMO2 by co-immunoprecipitation in T-ALL cell lines (*data not shown*) but our HDAC1 ChIPseq were not successful, and we could not confirm the PHF6-LMO2 interaction with the NuRD complex in T-ALL.

PHF6 was also shown to interact with UBF and bind to the rDNA-regulatory elements, suppressing the expression of rRNA.[15,47] Our PHF6 and LDB1 ChIPseq experiments indicated enrichment for the rDNA sequences but not specifically for the regulatory elements that are located upstream and

downstream of transcribed rDNA. Additionally, our PHF6 ChIP followed by qPCR, using primers specific for rDNA regulatory regions did not find enrichment. Therefore, our results failed to confirm PHF6 binding to the regulatory regions of rDNA, but it is possible that this is a cell-type specific function of PHF6 and not one it assumes in T-ALL.

Our findings indicate that although loss of PHF6 is associated with a particular group of T-ALL cases, PHF6 can have an important function in other T-ALL. This is in line with data from AML. Although PHF6 is mutated and therefore lost in 2-3% of AML, *PHF6* expression was found to be upregulated in AML.[22,48] A recent study conducted in AML found that PHF6 supported the development of RUNX1-ETO9a and MLL-AF9 AML.[49] It is known that both RUNX1 and RUNX1-ETO are associated with the LMO2 complex.[50] It is therefore

possible that a similar mechanism to the one described in this manuscript plays a role in AML. As the interaction we describe here will likely have its function in healthy cells, more knowledge of its DNA-binding characteristics during normal hematopoietic differentiation processes may be required for full understanding of the role of PHF6.

Disclosures
No conflicts of interests to disclose.

Contributions
Original concept, project planning and supervision by VSS and MH. Experimentation, data acquisition, processing and analysis by VSS, SB, BAJ, SA, FMBS and MH. Genome-wide data acquisition and bioinformatical analysis by VSS, SB, BAJ and MH. Statistical analysis by VSS and MH. Mass spectrometry data acquisition and analysis by VSS, SB and DGW. NMR spectroscopy data acquisition and analysis by VSS and CL. Provision of patient samples by SP and GP. Figures by VSS, SB, BAJ and MH. Writing of the manuscript by VSS and MH.

Data-sharing statement
Genome wide sequencing datasets are available have been deposited at the NCBI Gene Expression Omnibus repository, GEO:GSE154675.

References

1. Cross AJ, Jeffries CM, Trewhella J, Matthews JM. LIM domain binding proteins 1 and 2 have different oligomeric states. J Mol Biol. 2010;399(1):133-144.
2. Soler E, Andrieu-Soler C, de Boer E, et al. The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. Genes Dev. 2010;24(3):277-289.
3. Anguita E, Hughes J, Heyworth C, Blobel GA, Wood WG, Higgs DR. Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. EMBO J. 2004;23(14):2841-2852.
4. Stanulovic VS, Cauchy P, Assi SA, Hoogenkamp M. LMO2 is required for TAL1 DNA binding activity and initiation of definitive haematopoiesis at the haemangioblast stage. Nucleic Acids Res. 2017;45(17):9874-9888.
5. Yamada Y, Pannell R, Forster A, Rabbitts TH. The oncogenic LIM-only transcription factor Lmo2 regulates angiogenesis but not vasculogenesis in mice. Proc Natl Acad Sci U S A. 2000;97(1):320-324.
6. Yamada Y, Warren AJ, Dobson C, Forster A, Pannell R, Rabbitts TH. The T cell leukemia LIM protein Lmo2 is necessary for adult mouse hematopoiesis. Proc Natl Acad Sci U S A. 1998;95(7):3890-3895.
7. Warren AJ, Colledge WH, Carlton MB, Evans MJ, Smith AJ, Rabbitts TH. The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. Cell. 1994;78(1):45-57.
8. Herblot S, Steff AM, Hugo P, Aplan PD, Hoang T. SCL and LMO1 alter thymocyte differentiation: inhibition of E2A-HEB function and pre-T alpha chain expression. Nat Immunol. 2000;1(2):138-144.
9. Boehm T, Foroni L, Kaneko Y, Perutz MF, Rabbitts TH. The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 11p15 and 11p13. Proc Natl Acad Sci U S A. 1991;88(10):4367-4371.
10. Begley CG, Aplan PD, Davey MP, et al. Chromosomal translocation in a human leukemic stem-cell line disrupts the T-cell antigen receptor delta-chain diversity region and results in a previously unreported fusion transcript. Proc Natl Acad Sci U S A. 1989;86(6):2031-2035.
11. Bernard O, Guglielmi P, Jonveaux P, et al. Two distinct mechanisms for the SCL gene activation in the t(1;14) translocation of T-cell leukemias. Genes Chromosomes Cancer. 1990;1(3):194-208.
12. Homminga I, Pieters R, Langerak AW, et al. Integrated transcript and genome analyses reveal NKX2-1 and MEF2C as potential oncogenes in T cell acute lymphoblastic leukemia. Cancer Cell. 2011;19(4):484-497.
13. Liu Y, Easton J, Shao Y, et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. Nat Genet. 2017;49(8):1211-1218.
14. Lower KM, Turner G, Kerr BA, et al. Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. Nat Genet. 2002;32(4):661-665.
15. Wang J, Leung JW, Gong Z, Feng L, Shi X, Chen J. PHF6 regulates cell cycle progression by suppressing ribosomal RNA synthesis. J Biol Chem. 2013;288(5):3174-3183.
16. Todd MA, Picketts DJ. PHF6 interacts with the nucleosome remodeling and deacetylation (NuRD) complex. J Proteome Res. 2012;11(8):4326-4337.
17. Zhang C, Mejia LA, Huang J, et al. The X-linked intellectual disability protein PHF6 associates with the PAF1 complex and regulates neuronal migration in the mammalian brain. Neuron. 2013;78(6):986-993.
18. Warmerdam DO, Alonso-de Vega I, Wiegant WW, et al. PHF6 promotes non-homologous end joining and G2 checkpoint recovery. EMBO Rep. 2020;21(1):e48460.
19. Soto-Feliciano YM, Bartlebaugh JME, Liu Y, et al. PHF6 regulates phenotypic plasticity through chromatin organization within lineage-specific genes. Genes Dev. 2017;31(10):973-989.

20. Oh S, Boo K, Kim J, et al. The chromatin-binding protein PHF6 functions as an E3 ubiquitin ligase of H2BK120 via H2BK12Ac recognition for activation of trophectodermal genes. Nucleic Acids Res. 2020;48(16):9037-9052.

21. Van Vlierberghe P, Palomero T, Khiabanian H, et al. PHF6 mutations in T-cell acute lymphoblastic leukemia. Nat Genet. 2010;42(4):338-342.

22. de Rooij JD, van den Heuvel-Eibrink MM, van de Rijdt NK, et al. PHF6 mutations in paediatric acute myeloid leukaemia. Br J Haematol. 2016;175(5):967-971.

23. Van Vlierberghe P, Patel J, Abdel-Wahab O, et al. PHF6 mutations in adult acute myeloid leukemia. Leukemia. 2011;25(1):130-134.

24. Xiao W, Bharadwaj M, Levine M, et al. PHF6 and DNMT3A mutations are enriched in distinct subgroups of mixed phenotype acute leukemia with T-lineage differentiation. Blood Adv. 2018;2(23):3526-3539.

25. Alexander TB, Gu Z, Iacobucci I, et al. The genetic basis and cell of origin of mixed phenotype acute leukaemia. Nature. 2018;562(7727):373-379.

26. Li X, Yao H, Chen Z, Wang Q, Zhao Y, Chen S. Somatic mutations of PHF6 in patients with chronic myeloid leukemia in blast crisis. Leuk Lymphoma. 2013;54(3):671-672.

27. Wendorff AA, Quinn SA, Rashkovan M, et al. Phf6 loss enhances HSC self-renewal driving tumor initiation and leukemia stem cell activity in T-ALL. Cancer Discov. 2019;9(3):436-451.

28. Miyagi S, Sroczynska P, Kato Y, et al. The chromatin-binding protein Phf6 restricts the self-renewal of hematopoietic stem cells. Blood. 2019;133(23):2495-2506.

29. McRae HM, Garnham AL, Hu Y, et al. PHF6 regulates hematopoietic stem and progenitor cells and its loss synergizes with expression of TLX3 to cause leukemia. Blood. 2019;133(16):1729-1741.

30. Hsu YC, Chen TC, Lin CC, et al. Phf6-null hematopoietic stem cells have enhanced self-renewal capacity and oncogenic potentials. Blood Adv. 2019;3(15):2355-2367.

31. Vicente C, Schwab C, Broux M, et al. Targeted sequencing identifies associations between IL7R-JAK mutations and epigenetic modulators in T-cell acute lymphoblastic leukemia. Haematologica. 2015;100(10):1301-1310.

32. Lichtinger M, Ingram R, Hannah R, et al. RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. EMBO J. 2012;31(22):4318-4333.

33. Stanulovic VS, Al Omair S, Reed MAC, et al. The glutamate/aspartate transporter EAAT1 is crucial for T-cell acute lymphoblastic leukemia proliferation and survival. Haematologica. 2024;109(11):3505-3519.

34. Ludwig C. MetaboLabPy-An open-source software package for metabolomics NMR data processing and metabolic tracer data analysis. Metabolites. 2025;15(1):48.

35. Polonen P, Di Giacomo D, Seffernick AE, et al. The genomic basis of childhood T-lineage acute lymphoblastic leukaemia. Nature. 2024;632(8027):1082-1091.

36. Korczmar EA, Bookstaver AK, Ober E, Goldfarb AN, Tenen DG, Trinh BQ. Transcriptional regulation of the lineage-determining gene PU.1 in normal and malignant hematopoiesis: current understanding and therapeutic perspective. Front Biosci (Schol Ed). 2024;16(2):10.

37. Liew CW, Rand KD, Simpson RJ, et al. Molecular analysis of the interaction between the hematopoietic master transcription factors GATA-1 and PU.1. J Biol Chem. 2006;281(38):28296-28306.

38. Behre G, Whitmarsh AJ, Coghlan MP, et al. c-Jun is a JNK-independent coactivator of the PU.1 transcription factor. J Biol Chem. 1999;274(8):4939-4946.

39. Pundhir S, Bratt Lauridsen FK, Schuster MB, et al. Enhancer and transcription factor dynamics during myeloid differentiation reveal an early differentiation block in Cebpa null progenitors. Cell Rep. 2018;23(9):2744-2757.

40. Woodcock JM, Zacharakis B, Plaetinck G, et al. Three residues in the common beta chain of the human GM-CSF, IL-3 and IL-5 receptors are essential for GM-CSF and IL-5 but not IL-3 high affinity binding and interact with Glu21 of GM-CSF. EMBO J. 1994;13(21):5176-5185.

41. Lecuyer E, Herblot S, Saint-Denis M, et al. The SCL complex regulates c-kit expression in hematopoietic cells through functional interaction with Sp1. Blood. 2002;100(7):2430-2440.

42. Xu Z, Meng X, Cai Y, Liang H, Nagarajan L, Brandt SJ. Single-stranded DNA-binding proteins regulate the abundance of LIM domain and LIM domain-binding proteins. Genes Dev. 2007;21(8):942-955.

43. Cubedo E, Gentles AJ, Huang C, et al. Identification of LMO2 transcriptome and interactome in diffuse large B-cell lymphoma. Blood. 2012;119(23):5478-5491.

44. Sanda T, Lawton LN, Barrasa MI, et al. Core transcriptional regulatory circuit controlled by the TAL1 complex in human T cell acute lymphoblastic leukemia. Cancer Cell. 2012;22(2):209-221.

45. Liu Z, Li F, Ruan K, et al. Structural and functional insights into the human Borjeson-Forssman-Lehmann syndrome-associated protein PHF6. J Biol Chem. 2014;289(14):10069-10083.

46. Liu Z, Li F, Zhang B, Li S, Wu J, Shi Y. Structural basis of plant homeodomain finger 6 (PHF6) recognition by the retinoblastoma binding protein 4 (RBBP4) component of the nucleosome remodeling and deacetylase (NuRD) complex. J Biol Chem. 2015;290(10):6630-6638.

47. Todd MA, Huh MS, Picketts DJ. The sub-nucleolar localization of PHF6 defines its role in rDNA transcription and early processing events. Eur J Hum Genet. 2016;24(10):1453-1459.

48. Mousa NO, Gado M, Assem MM, Dawood KM, Osman A. Expression profiling of some acute myeloid leukemia-associated markers to assess their diagnostic/prognostic potential. Genet Mol Biol. 2021;44(1):e20190268.

49. Hou S, Wang X, Guo T, et al. PHF6 maintains acute myeloid leukemia via regulating NF-kappaB signaling pathway. Leukemia. 2023;37(8):1626-1637.

50. Ptasinska A, Assi SA, Martinez-Soria N, et al. Identification of a dynamic core transcriptional network in t(8;21) AML that regulates differentiation block and self-renewal. Cell Rep. 2014;8(6):1974-1988.