nature communications

Article

https://doi.org/10.1038/s41467-024-51566-5

PHF6 cooperates with SWI/SNF complexes to facilitate transcriptional progression

Received: 27 September 2023

Accepted: 7 August 2024

Published online: 24 August 2024

Check for updates

Priya Mittal[1], Jacquelyn A. Myers[1], Raymond D. Carter[1], Sandi Radko-Juettner[1,2], Hayden A. Malone[1,2], Wojciech Rosikiewicz[3], Alexis N. Robertson[1], Zhexin Zhu[1], Ishwarya V. Narayanan[4], Baranda S. Hansen[5], Meadow Parrish[1], Natarajan V. Bhanu[6], Robert J. Mobley[1], Jerold E. Rehg[7], Beisi Xu[3,7], Yiannis Drosos[1], Shondra M. Pruett-Miller[5], Mats Ljungman[4], Benjamin A. Garcia[6], Gang Wu[3,7], Janet F. Partridge[1] & Charles W. M. Roberts[1]

Genes encoding subunits of SWI/SNF (BAF) chromatin remodeling complexes are mutated in nearly 25% of cancers. To gain insight into the mechanisms by which SWI/SNF mutations drive cancer, we contributed ten rhabdoid tumor (RT) cell lines mutant for SWI/SNF subunit *SMARCB1* to a genome-scale CRISPR–Cas9 depletion screen performed across 896 cell lines. We identify *PHF6* as specifically essential for RT cell survival and demonstrate that dependency on *Phf6* extends to *Smarcb1*-deficient cancers in vivo. As mutations in either *SWI/SNF* or *PHF6* can cause the neurodevelopmental disorder Coffin-Siris syndrome, our findings of a dependency suggest a previously unrecognized functional link. We demonstrate that PHF6 co-localizes with SWI/SNF complexes at promoters, where it is essential for maintenance of an active chromatin state. We show that in the absence of SMARCB1, PHF6 loss disrupts the recruitment and stability of residual SWI/SNF complex members, collectively resulting in the loss of active chromatin at promoters and stalling of RNA Polymerase II progression. Our work establishes a mechanistic basis for the shared syndromic features of *SWI/SNF* and *PHF6* mutations in CSS and the basis for selective dependency on PHF6 in *SMARCB1*-mutant cancers.

Chromatin regulators function in part to facilitate the precise spatio-temporal gene regulation that underlies cell fate specification. The importance of this function is exemplified by the recurrent mutations observed in genes encoding subunits of the SWI/SNF chromatin remodeler, manifesting in both developmental disorders and cancers[1–3]. The discovery of bi-allelic inactivating mutations in the gene encoding the SWI/SNF subunit *SMARCB1* in nearly all cases of rhabdoid tumor (RT) provided the first link between SWI/SNF complexes and cancer[4,5].

RTs are highly aggressive cancers that typically develop in young children. They arise most frequently in the kidneys, brain, and liver and have poor survival outcomes[6]. Aside from *SMARCB1* inactivation, RTs have remarkably simple diploid genomes, making them an ideal model

[1]Division of Molecular Oncology, Department of Oncology, St. Jude Children's Research Hospital, Memphis, TN, USA. [2]St. Jude Graduate School of Biomedical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA. [3]Center for Applied Bioinformatics, St. Jude Children's Research Hospital, Memphis, TN, USA. [4]Department of Radiation Oncology, Rogel Cancer Center and Center for RNA Biomedicine, University of Michigan, Ann Arbor, MI, USA. [5]Center for Advanced Genome Engineering, Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN, USA. [6]Department of Biochemistry and Biophysics, Smilow Center for Translational Research, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [7]Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA. e-mail: Charles.roberts@stjude.org

to gain mechanistic insights into the functional consequences of SWI/SNF mutations in cancer[7].

SWI/SNF complexes are multi-protein assemblies that hydrolyze ATP to mobilize nucleosomes. There are three families of SWI/SNF complexes, cBAF, PBAF, and gBAF/ncBAF, two of which (cBAF and PBAF) contain SMARCB1[8–10]. All three families contain bromodomains, a motif that binds to acetylated lysines. These include core ATPase subunits SMARCA4/SMARCA2, which are shared across all three families, as well as family-specific members such as PBRM1 (specific to PBAF) and BRD9 (specific to ncBAF). Recent biochemical studies have demonstrated that the bromodomains of SMARCA4 and PBRM1 specifically recognize acetylation of the lysine present at residue 14 of histone 3 (H3K14ac)[11,12]. Notably, of 340 known histone modifications, acetylation of H3K14 alone is sufficient to potentiate the remodeling activity of the BAF, PBAF, and ncBAF families[13].

Loss of SMARCB1 impairs the function of cBAF and PBAF, the two SWI/SNF families that include SMARCB1. This is particularly evident at enhancers, where the absence of SMARCB1 results in reduced accessibility and loss of H3K27ac[14,15]. At promoters, residual members of BAF and PBAF complexes are still recruited in the absence of SMARCB1[16]. Notably, SMARCB1-deficient cancer cells show enhanced dependence on the SWI/SNF ATPase subunit SMARCA4 and members of the ncBAF complex for survival, demonstrating that the absence of SMARCB1 compromises SWI/SNF function but does not render it completely inactive[8,9,17]. These changes culminate in defective activation of transcription programs associated with differentiation and development[14,15]. However, the mechanisms by which SMARCB1 and SWI/SNF complexes contribute to the integrated control of transcriptional regulation and how mutation of *SWI/SNF* subunits disrupts transcription programs remains poorly understood.

To gain insight into these mechanisms, we sought to identify dependencies created by the loss of function of SMARCB1. We, therefore, contributed 10 *SMARCB1*-mutant RT cell lines to a genome-wide CRISPR-Cas9 loss-of-function screen performed in collaboration with the Broad Institute across 896 cancer cell lines[18,19]. Here, we report that SMARCB1-deficient RTs are specifically dependent on the Plant Homeodomain-like Finger protein 6 (PHF6) for survival. PHF6 is an X-chromosome–encoded protein that contains two evolutionarily conserved atypical PHD finger domains. Such zinc-finger-containing PHD domains have been implicated in protein-protein interactions that are often involved in epigenetic regulation[20]. PHF6 was initially linked to ribosomal DNA regulation in the nucleolus[21]; however, nuclear roles in chromatin binding have recently emerged[22,23]. PHF6 itself has been linked to cancer, with approximately 20% of pediatric patients with T-cell acute lymphoblastic leukemia (T-ALL) harboring somatic inactivating mutations in *PHF6*[24,25]. Specifically, the TLX3 subgroup of T-ALL can contain mutation of either *PHF6* or SWI/SNF subunit *SMARCA4* in a mutually exclusive manner. Germline mutations in X-linked *PHF6* have also been identified in females with the neurodevelopmental disorder Coffin–Siris syndrome (CSS). CSS has otherwise been principally attributed to mutations in genes that encode *SWI/SNF* subunits, including SMARCB1[26,27]. The fact that *PHF6* and *SWI/SNF* mutations phenocopy each other in disease suggests the possibility of a previously unrecognized cooperative function.

We establish a functional connection between PHF6 and SWI/SNF complexes. Through various model systems, we demonstrate that PHF6 acts as a central hub for protein interactions at active promoters. Our findings reveal that the absence of PHF6 leads to compromised recruitment of the histone acetyltransferase p300, resulting in decreased acetylation of H3 at residues K14 and K27. Additionally, PHF6 loss causes impaired recruitment of bromodomain-containing SWI/SNF members, resulting in the loss of active chromatin at promoters and the stalling of RNA polymerase II (Pol II). These results collectively demonstrate a pivotal role for PHF6 in facilitating transcriptional activation at promoters. Furthermore, in the setting of impaired enhancer function caused by the absence of SMARCB1, PHF6 becomes essential for maintaining transcriptional activity.

Results

Rhabdoid tumors are specifically dependent on PHF6 for survival

To identify genes that are essential in the absence of the SWI/SNF subunit SMARCB1, we contributed 10 *SMARCB1*-mutant RT cell lines to the Cancer Dependency Map (DepMap), a large-scale collaboration with the Broad Institute, where genome-wide CRISPR-Cas9 knockout screens were performed across 896 adult and pediatric cancer cell lines[18,19,28]. We identified *PHF6* as a specific dependency in RT cell lines compared to other cell lines ($P = 1.12 \times 10^{-8}$) (Fig. 1a and Supplementary Fig. S1a, b). Dependence on *PHF6* did not correlate with its expression levels (Supplementary Fig. S1c).

In addition to RTs, small-cell carcinoma of the ovary, hypercalcemic type (SCCOHT) cell lines were also dependent on *PHF6*. SCCOHT is a cancer of young adults that is driven by biallelic inactivating mutations in the SWI/SNF subunit *SMARCA4 and* has been referred to as 'ovarian rhabdoid tumor' due to histological similarities (Supplementary Fig. S1d). Of note, dependency upon *PHF6* was not observed in other SWI/SNF mutant cancer cell lines (Supplementary Fig. S1e).

We independently validated this dependency by performing CRISPR-based competitive fitness assays to determine whether there was selection against loss-of-function *PHF6* mutations in RT cells. Out-of-frame alleles of *PHF6* were selected against in SMARCB1-deficient RT cells (G401) but not in *SMARCB1*-wildtype Pfeiffer cells (control) (Fig. 1b). The strength of the negative selection was such that we were unable to generate *PHF6*-deficient RT clones, although we were readily able to do so in control cell lines (Supplementary Fig. S1f). To further validate PHF6 dependency, we performed shRNA-mediated knockdown of PHF6 in four SMARCB1-deficient RT lines (G401, TTC549, TTC709, and CHLA266) and two control *SMARCB1*–wildtype cell lines (ES2 and MCF7). Loss of PHF6 impaired proliferation and led to apoptosis in the RT cell lines (Fig. 1c, Supplementary Fig. S1g, h, j) but did not affect the viability of control cells (Fig. 1d, Supplementary Fig. S1i). We also validated *PHF6* as a dependency in the *SMARCA4*-inactivated SCCOHT cell line (BIN67) (Supplementary Fig. S1k). Rescue of PHF6 through inducible expression of a PHF6 construct that was not recognized by the shRNA in the knocked down cell lines rescued proliferation, demonstrating that the effects of the PHF6 shRNAs were on-target (Fig. 1e, Supplementary Fig. S1l). Collectively, these findings show that *PHF6* is a specific dependency and is required for the viability of *SMARCB1*-mutant RT cell lines.

PHF6 loss impairs acetylation of histone H3 at Lysine 14 and 27

Given the mutually exclusive occurrence of *PHF6* and *SWI/SNF* mutations in both leukemia and CSS[25,27], we sought to investigate whether *PHF6* has unknown functions related to SWI/SNF that become preferentially essential in the absence of *SMARCB1*. Previously, we and others demonstrated that SMARCB1 promotes transcriptional activation and facilitates the acetylation of H3K27 to support differentiation programs[14,15]. Accordingly, we tested whether PHF6 served a similar role. We first performed quantitative analysis of histone post-translational modifications by mass spectrometry[29]. Although H3K27 acetylation was reduced in G401 cells that had lost PHF6, the most significant change was a reduction of H3K14 acetylation (Fig. 2a) and a less pronounced reduction in H3K36me2. We confirmed that PHF6 knockdown resulted in a marked reduction in H3K14ac and reductions in H3K27ac and H3K36me2 across three independent RT cell lines by immunoblot analysis (Fig. 2b). Other marks of active transcription, including H3K9ac and H3K4me1, were largely unchanged across the three RT cell lines, with only modest changes in H3K4me3 (Fig. 2b). Histone modifications associated with repression, H3K9me3 and

https://doi.org/10.1038/s41467-024-51566-5

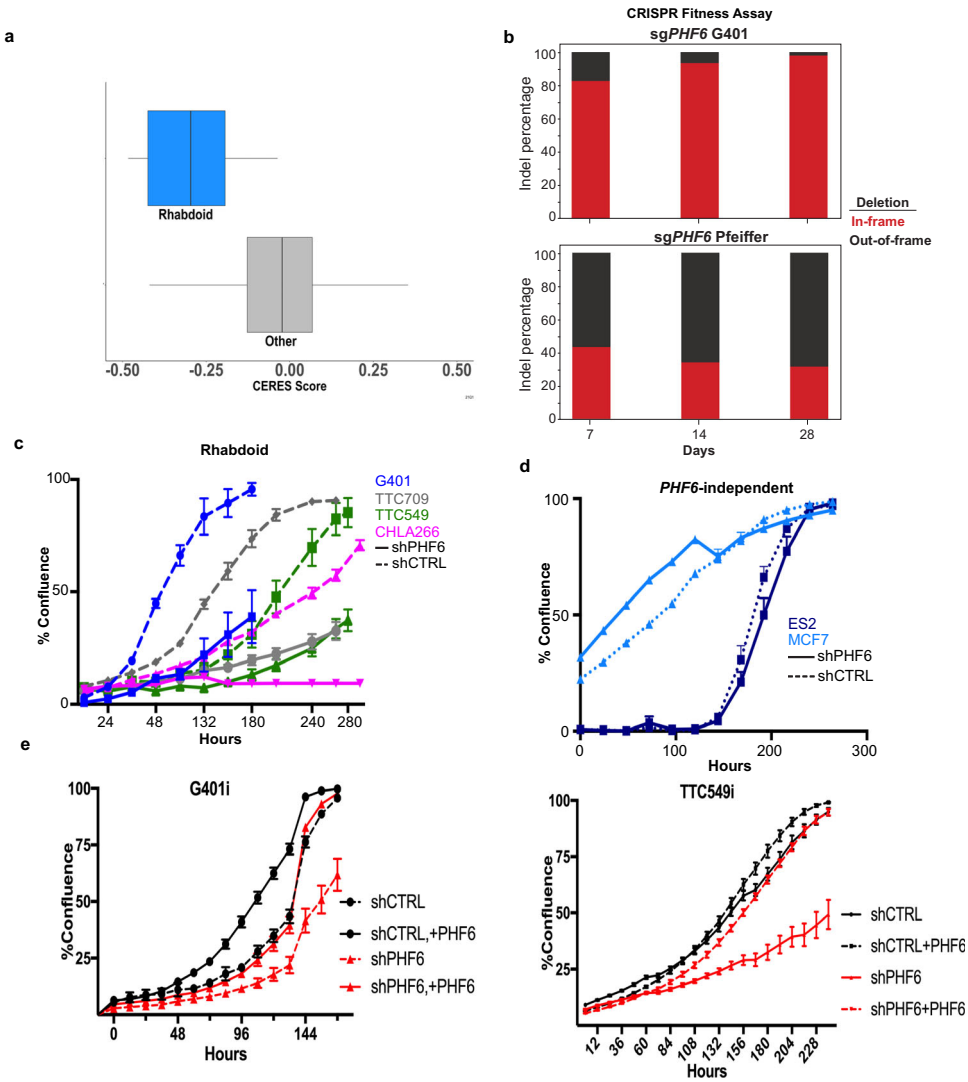

Fig. 1 | Genome-scale CRISPR-Cas9 screen identifies *PHF6* as a dependency in rhabdoid tumor cell lines. a Two-class comparison of 10 biologically independent RT cell lines (blue) versus $n = 886$ biologically independent other cell lines (gray) plotted with CERES score for dependency on PHF6 on the *x*-axis. Negative CERES score demonstrates increased dependency. Statistical analysis was performed using a Benjamini–Hochberg corrected two-tailed Student's *t*-test; $P = 1.12 \times 10 - 8$. **b** Indel frequency assay. PHF6 was targeted with a guide RNA-Cas9 RNP in G401 RT cells and control Pfeiffer cells, and genomic DNA was harvested and sequenced after 7, 14, and 21 days. All indels were binned into in-frame (red) or out-of-frame (black). A non-targeting guide was used as a control. Data are represented from mean of three biological replicates. **c**, **d** Effects of PHF6 knockdown on proliferation of SMARCB1-deficient RT cell lines or (**d**) SMARCB1-expressing ES2 and MCF7 control cell lines. Data are represented from mean of $n = 16$ technical replicates per cell line per condition from one independent experiment. **e** Effects of exogenous re-expression of PHF6 following knockdown of PHF6 in G401 and TTC549 cells. Data are represented from mean of $n = 8$ technical replicates per cell line per condition from one independent experiment. Source data are provided as a Source Data file.

H3K27me3, showed little change in the RT cells (Supplementary Fig. S2a). To rule out off-target effects, we performed rescue experiments using an inducible PHF6 expression construct that was not recognized by the shRNA against PHF6. Induced expression of PHF6 in shPHF6 G401 and TTC549 RT cells rescued H3K14ac, H3K27ac, and H3K36me2, indicating that the changes in histone modifications were caused by PHF6 knockdown (Fig. 2c, Supplementary Fig. S2b).

As acetylation of H3K14 is a comparatively little-studied histone modification, we first sought to characterize its genomic localization. We, therefore, performed chromatin immunoprecipitation followed by sequencing (ChIP-Seq) for H3K14ac and several other well-established histone modifications in RT cells (Supplementary Fig. S2c) and found that H3K14ac is enriched at a subset (40–50%) of active promoters and regions flanking promoters (Supplementary Fig. S2d).

Next, we evaluated the contributions of PHF6 to the localization of histone modifications by performing chromatin immunoprecipitation followed by sequencing (ChIP-Seq) for H3K14ac, H3K27ac, and H3K4me3, H3K4me1, H3K36me2 upon PHF6 knockdown. PHF6 loss resulted in a widespread and pronounced reduction of H3K14ac (Fig. 2d, Supplementary Fig. S2e). Further, H3K27ac was substantially reduced at transcription start sites (TSS) (Fig. 2d), enhancers, and super-enhancers, whereas the effect on H3K4me3 and H3K4me1 was more modest (Supplementary Fig. S2f, g, h). Loss of PHF6 also resulted in a global reduction in H3K36me2 (Supplementary Fig. S2i).

Lastly, we generated an inducible PHF6 gain-of-function system utilizing PHF6-deficient DND41 T-ALL cells. Following re-expression of *PHF6*, we observed increased levels of H3K14ac and H3K27ac (Fig. 2e), while the repressive marks H3K27me3 and H3K9me3 remained unchanged (Supplementary Fig. S2j). Together, these results suggest

https://doi.org/10.1038/s41467-024-51566-5

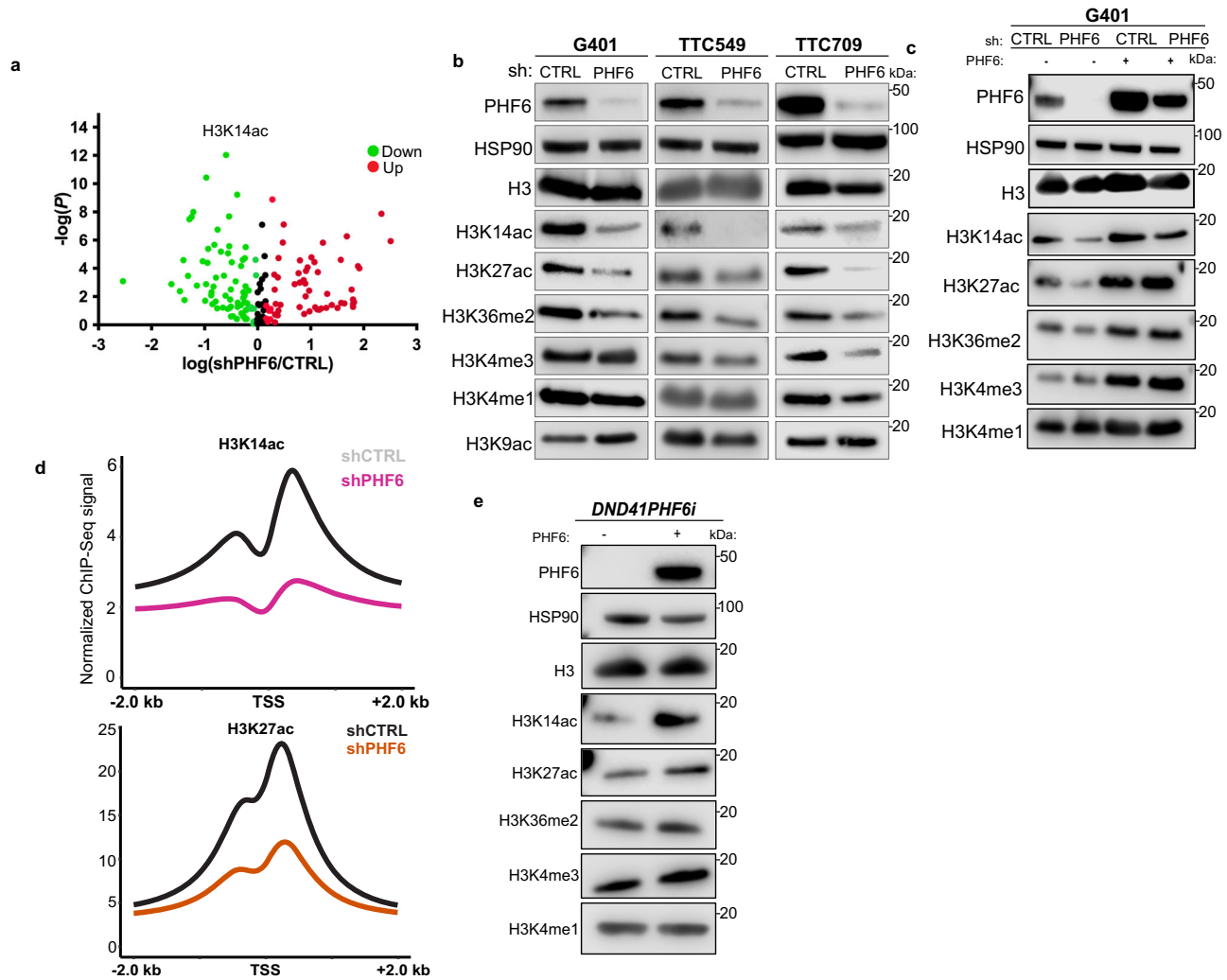

Fig. 2 | PHF6 is required for maintenance of H3K14 and H3K27 acetylation. **a** Volcano plot displaying histone post-translational modification changes as determined by mass spectrometry of histones extracted from G401 shPHF6 versus shCTRL cells. Two-tailed homoscedastic Student's *t*-test was performed. Data are represented from mean of three biological replicates. **b** Western blots of histone modifications in histone extracts prepared from RT cells 72 hr post selection with shCTRL or shPHF6 knockdown ($n = 3$ independent biological replicates). **c** Western blots of histone modifications in G401 cells re-expressing inducible exogenous PHF6 (72 h) after knockdown of endogenous PHF6 ($n = 2$ independent biological replicates). **d** Metagene plots of H3K14ac, H3K27ac centered at the TSS ± 2 kb in shCTRL and shPHF6 G401 cells ($n = 3$ independent biological replicates). **e** Re-expression of doxycycline-inducible exogenous PHF6 (72 h) in DND41 cells that naturally lack PHF6 expression ($n = 2$ independent biological replicates). Source data are provided as a Source Data file.

that PHF6 serves a role in the establishment of an active chromatin landscape marked by H3K14ac and H3K27ac.

PHF6 is essential for the recruitment of p300 to maintain H3K14 acetylation and chromatin accessibility

Noting the impact of PHF6 on histone modifications associated with active promoters, we next investigated the localization of PHF6 using ChIP-Seq in RT cells (Supplementary Fig. S3a, b). PHF6 was strongly enriched at active promoters and flanking regions. Although these regions constitute less than 1% of the genome, they represented 44% of the PHF6 bound sites in RT cells (Fig. 3a). We established reproducibility of PHF6 binding events with additional antibodies as well as demonstrated specificity using both our RT and PHF6 inducible DND41 cell lines (Supplementary Fig. S3a–d).

Since PHF6 does not contain an acetyltransferase domain, we evaluated whether PHF6 interacted with acetyltransferases that modify H3K14 and H3K27. CREB-binding protein (CBP/p300) is the acetyltransferase that writes H3K27ac in mammals and has additionally been shown to be capable of acetylating H3K14[30,31]. Moreover, p300 is known to interact with SWI/SNF complexes[32,33] and is also a preferential genetic dependency in RTs (Supplementary Fig. S3e). We first evaluated whether p300 was responsible for acetylating H3K14 in RTs via treatment with the p300 inhibitor A-485. Treatment with A-485 resulted in marked reductions in the levels of both H3K14ac and H3K27ac (Supplementary Fig. S3f). Next, we asked whether PHF6 interacts with p300. Immunoprecipitation of either PHF6 or p300 resulted in co-immunoprecipitation of the other (Fig. 3b, Supplementary Fig. S3g). To evaluate the potential role of PHF6 in supporting p300 function, we tested the effect of PHF6 loss on the levels of p300. Knockdown of PHF6 resulted in reduced levels of p300 in RT, and rescue of PHF6 expression in DND41 cells resulted in a modest increase of p300 levels (Fig. 3c, Supplementary Fig. S3h, Supplementary Table S1). PHF6 was also required for the recruitment of p300 to chromatin as the loss of PHF6 caused a global reduction in occupancy of p300 in G401s (Supplementary Fig. S3i). Importantly, loss of p300 occupancy corresponded to sites that also lost H3K14ac and H3K27ac (Fig. 3d). In addition to p300, KAT7/HBO1 has been reported to write H3K14acetylation[34,35]. Therefore, we evaluated both the levels and

https://doi.org/10.1038/s41467-024-51566-5

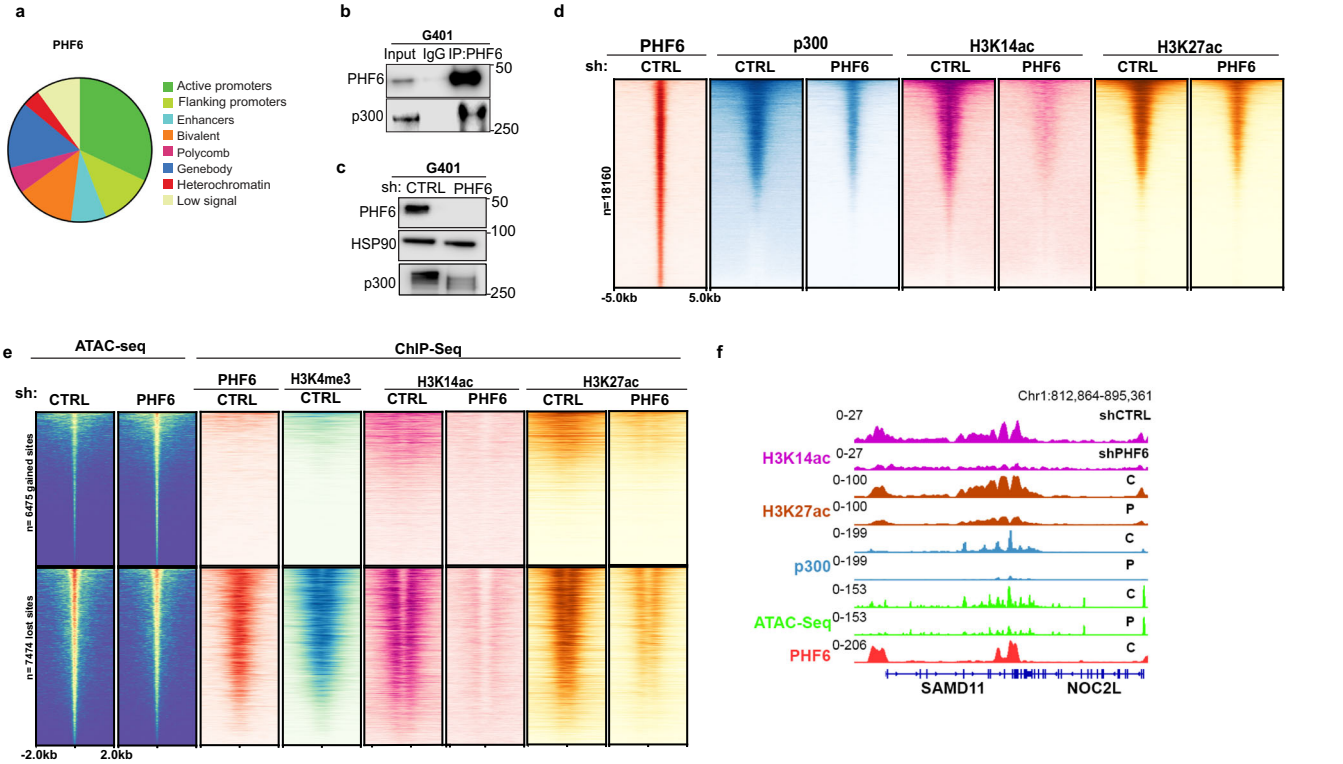

Fig. 3 | PHF6 regulates histone acetyltransferase p300 and maintains chromatin accessibility. a Distribution of PHF6 occupancy across chromatin states in G401 RT cells ($n = 2$ independent biological replicates). **b** Co-immunoprecipitation of PHF6 with p300 from nuclear extracts of G401 RT cells ($n = 3$ independent biological replicates). **c** Western blot analysis p300 protein levels in nuclear extracts from G401 RT cells after knockdown of PHF6 ($n = 3$ independent biological replicates). **d** Density heatmap of chromatin occupancy of PHF6, p300, H3K14ac, and H3K27ac in shCTRL and shPHF6 conditions in G401 (ordered by PHF6 occupancy). **e** Density heatmap of significant losses ($\log_2FC < 0$, FDR < 0.05) and gains ($\log_2FC > 0$, FDR < 0.05) in accessibility and chromatin occupancy changes of H3K14ac and H3K27ac upon loss of PHF6 ranked based on PHF6 binding are shown in G401 cells ($n = 3$ independent biological replicates). Promoter regions are designated by the presence of H3K4me3. **f** Example loci reflecting loss of H3K14ac, H3K27ac, p300, and resultant accessibility at PHF6-bound promoters. Source data are provided as a Source Data file.

chromatin occupancy of HBO1 upon loss of PHF6. While we observed marked effects upon loss of PHF6 on CBP/p300, we observed minimal effects on either the levels or recruitment of HBO1 (Supplementary Fig. S3j, k, l). This demonstrates that p300 is the dominant acetyltransferase for H3K14 and H3K27 in RTs and that the reductions in H3K14ac and H3K27ac upon PHF6 loss are due to disruption in the recruitment of p300.

As loss of PHF6 has substantial effects upon several histone modifications associated with the active chromatin landscape, we next asked whether PHF6 is required for chromatin accessibility by performing an assay for transposase-accessible chromatin using sequencing (ATAC-Seq). Upon PHF6 loss, we identified substantial changes, with 7474 regions losing and 6475 regions gaining accessibility (Supplementary Fig. S3m). Notably, PHF6 was present solely at promoters that lost accessibility (3310/7474), indicating that the promoters that displayed a modest gain in accessibility represented secondary effects (Fig. 3e, f). Consequently, PHF6 is required both for the recruitment of p300 and maintenance of open chromatin, establishing PHF6 as a critical regulator of active chromatin at transcription start sites.

PHF6 is required for Pol II pause release

The presence of a positioned nucleosome at the +1 position relative to the TSS impedes the progression of RNA Polymerase II (Pol II)[36,37]. Considering the diminished accessibility consequent to PHF6 loss and our observation that PHF6 binding extends over the +1 nucleosome (Supplementary Fig. S4a), we asked whether PHF6 played a role in Pol II licensing. Pol II is phosphorylated on multiple sites during transcription, and different phosphorylation patterns are associated with distinct stages of transcription. Phosphorylation of Serine 5 (Ser5P) at the Pol II C-terminus occurs 20–60 bp downstream of the TSS and correlates with pausing, while Serine 2 (Ser2P) phosphorylation occurs during the elongation phase and is associated with productive RNA synthesis[38–40].

We first evaluated for potential physical interactions between PHF6 and Pol II and found that PHF6 co-immunoprecipitates both forms of Pol II (Ser5P and Ser2P) (Fig. 4a). While depletion of PHF6 in RT cells had no effect on the overall expression levels of Pol II (Supplementary Table S1), it resulted in a marked increase in the levels of Ser5P-modified Pol II and a decrease in Ser2P-modified Pol II (Fig. 4b, Supplementary Fig. S4b). These data suggested that PHF6 deficiency results in a failure to release paused Pol II. To assess the effects of PHF6 depletion on Pol II occupancy, we first performed ChIP-Seq for Pol II using an antibody that recognized both phosphorylated and unphosphorylated forms of Pol II. Depletion of PHF6 revealed only modest overall changes in Pol II signal (Supplementary Fig. S4c). To gain resolution on how the loss of PHF6 specifically influenced modified Pol II, we performed ChIP-Seq for active forms of Pol II, i.e., RNA Pol II Ser5P and Ser2P. We observed accumulation of Ser5P Pol II at promoter-proximal sites and the depletion of Ser2P-Pol II within gene bodies (Fig. 4c, d). Genes lacking PHF6 occupancy did not exhibit changes in Ser5P or Ser2P modified Pol II (Supplementary Fig. S4d). Consistent with a direct role of PHF6 in Pol II pause-release, re-expression of PHF6 in PHF6-deficient DND41 cells reduced Ser5P levels (Supplementary Fig. S4e).

To more directly measure the effect of PHF6 on Pol II stalling, we performed Bromouridine-sequencing (Bru-seq), a type of

https://doi.org/10.1038/s41467-024-51566-5

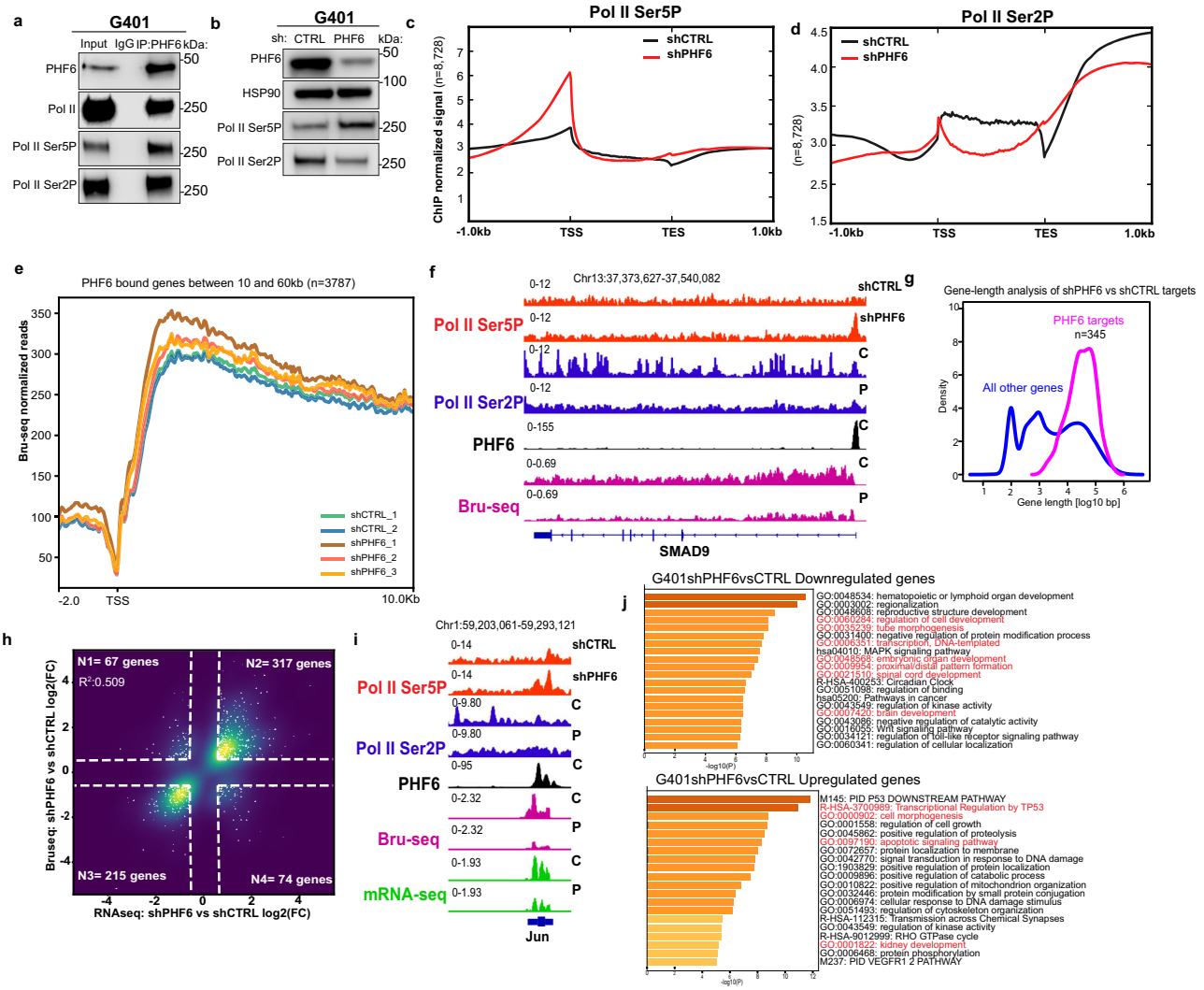

Fig. 4 | PHF6 is required for promoter-proximal pause release of RNA polymerase II. a PHF6 co-immunoprecipitation of Pol II and its phosphorylated forms in G401 RT cells ($n = 2$ independent biological replicates) **b** Western blot analysis of the levels of Pol II Ser5P (promoter proximal) and Pol II Ser2P (elongating) forms in shCTRL and shPHF6 G401 cells ($n = 3$ independent biological replicates). Metagene plots of (**c**) Pol II Ser5P (promoter proximal) and **d** Pol II Ser2P (elongating) in shCTRL (black) and shPHF6 (red) conditions ($n = 2$ independent biological replicates) at PHF6-bound genes ($n = 8728$). **e** Metagene nascent transcriptional profiles (Bru-seq read density) of PHF6 bound genes stratified based on gene-length (10–60 kb) at the 5′ (TSS) of protein-coding genes in shCTRL and shPHF6 conditions in G401 cells ($n = 3$ independent biological replicates). **f** Example locus depicting Pol II Ser5P, Pol II Ser2P, and nascent transcription upon loss of PHF6 at PHF6-bound SMAD9. **g** Gene length distributions (log10 transformed) for all hg19 genes (blue) and 345 significantly downregulated PHF6 target genes defined in RT cell lines (TTC709, TTC549, G401) in magenta. **h** Scatter plot displaying correlation between differentially expressed genes from Bru-seq and mRNA-seq upon PHF6 knockdown in G401 cells. **i** Representative locus depicting Pol II Ser5P, Pol II Ser2P, nascent transcript, and mRNA upon loss of PHF6 at PHF6-bound Jun locus. **j** Gene ontology analysis by Metascape of significantly downregulated and upregulated genes upon PHF6 knockdown (adjusted $p < 0.05$ and $\log_2FC < 0$ or $\log_2FC > 0$) in G401 cells. Source data are provided as a Source Data file.

metabolically labeled nascent RNA sequencing[41]. Since longer genes are most susceptible to elongation stalling[42,43], we incorporated gene length into our analysis. Transcription start site (TSS)-focused metagene-analysis of PHF6 target genes (>10 kb) indeed revealed nascent signal accumulation at the 5′ end upon loss of PHF6 (Fig. 4e, f). Conversely, transcription end site (TES) plots revealed a decrease in read density, specifically at longer genes (>60 kb) (Supplementary Fig. S4f). Taken together, we demonstrate that without PHF6, Pol II stalls, and transcriptional elongation is impaired at PHF6-bound genes.

To characterize the effects of loss of PHF6 on steady-state transcription, we performed mRNA-seq in three RT cell lines, G401, TTC549, and TTC709, in the presence and absence of PHF6. Depletion of PHF6 resulted in downregulation of its target genes ($p = 8.24 \times 10^{-23}$) (Supplementary Fig. S4g, h) with significant overlap in affected genes, particularly those of above-average length, among the three RT cell lines (Fig. 4g–i, Supplementary Fig. S4i,j). Metascape analyses[44] of differentially expressed genes upon loss of PHF6 enriched for terms related to development, apoptosis, and cell cycle. Specifically, the downregulated genes (primary targets) revealed enrichment for pathways associated with RUNX1 targets, mitotic cell cycle, and E2F-6 targets known to be essential for mitotic G1 transition (Fig. 4j, Supplementary Fig. S4k, l, m) reminiscent of recently described SWI/SNF functions[45–47].

PHF6 facilitates SWI/SNF function

Recently, a link has been established between the BAF complexes and Pol II to aid in the eviction of nucleosomes[48]. Additionally, the bromodomain-containing SWI/SNF members (Supplementary Fig. S5a) are known to specifically recognize and be potentiated by H3K14acetylation[11,13,49,50]. Given our discovery of PHF6's role in the

https://doi.org/10.1038/s41467-024-51566-5

regulation of H3K14acetylation, maintaining accessible chromatin, and facilitating polymerase pause-release, we sought to investigate whether PHF6 cooperates with SWI/SNF complexes. We found that PHF6 indeed co-localizes with PBAF (PBRM1, SMARCA4) and ncBAF (BRD9, SMARCA4) families of SWI/SNF complexes, predominantly at active promoters enriched for H3K14ac (Fig. 5a, b, Supplementary Fig. S5b).

We next evaluated whether PHF6 physically interacts with SWI/SNF complexes. Immunoprecipitation of PHF6 in RT cells resulted in the co-

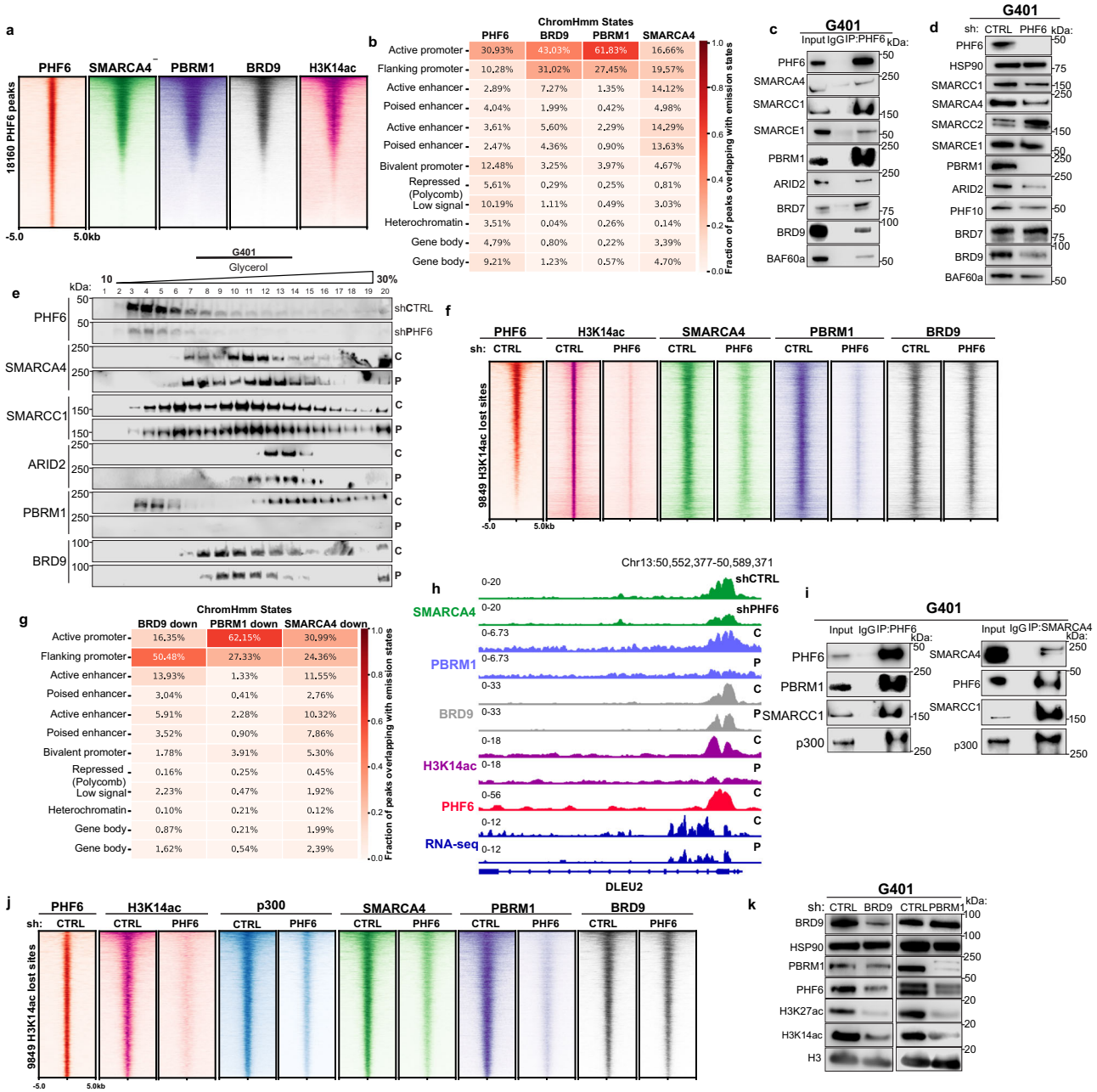

Fig. 5 | PHF6 facilitates SWI/SNF function. a Density heatmap of chromatin occupancy (ordered by PHF6) for PHF6, SMARCA4, PBRM1, BRD9, and H3K14ac in G401 cells ($n = 2$ independent biological replicates). **b** Chromatin occupancy peaks of PHF6, BRD9, PBRM1, and SMARCA4 overlapping with chromatin states defined based on ChromHMM. **c** PHF6 immunoprecipitation of SWI/SNF subunits from G401 nuclear extracts ($n = 3$ independent biological replicates) **d** Western blot analysis of SWI/SNF complex members from shCTRL and shPHF6 G401 nuclear lysates ($n = 3$ independent biological replicates). **e** Density sedimentation assays using 10–30% glycerol gradients performed on G401 lysates in shCTRL and shPHF6 conditions ($n = 2$ independent biological replicates). **f** Density heatmap in shCTRL and shPHF6 conditions (rank ordered by PHF6) at sites where H3K14ac is lost upon knockdown of PHF6 in G401 cells ($n = 2$ independent biological replicates). **g** Overlap of sites where BRD9, PBRM1, and SMARCA4 peaks are lost upon PHF6 knockdown in G401 with chromatin states defined based on ChromHMM. **h** Example locus reflecting reduced occupancy of SMARCA4, PBRM1, BRD9, H3K14ac, and RNA levels at PHF6-bound DLEU2. **i** Co-immunoprecipitation of PHF6 with p300, PBRM1, and SMARCC1 from nuclear extracts of G401 RT cells ($n = 3$ independent biological replicates). Co-immunoprecipitation of SMARCA4 with PHF6, SMARCC1, and p300 ($n = 3$ independent biological replicates). **j** Density heatmap of p300, SMARCA4, PBRM1, and BRD9 at sites (rank ordered by PHF6) of H3K14ac occupancy loss upon knockdown of PHF6 in G401 cells ($n = 2$ independent biological replicates). **k** Western blots of histone modifications in histone extracts prepared from RT cells 72 h post selection with shCTRL or shBRD9 or shPBRM1 knockdown ($n = 2$ independent biological replicates). **l** Western blot analysis of knockdown of BRD9 and PBRM1 in G401 cells reveals a role for BRD9 and PBRM1 in controlling PHF6 levels ($n = 3$ independent biological replicates). Source data are provided as a Source Data file.

immunoprecipitation of ncBAF complex subunits, as well as residual subunits of PBAF complexes despite the absence of SMARCB1 (Fig. 5c, Supplementary Fig. S5c). Reciprocally, the pulldown of the SWI/SNF ATPase subunit SMARCA4 resulted in co-immunoprecipitation of PHF6 (Supplementary Fig. S5d). The pulldown of the ncBAF-specific subunit BRD9 and the PBAF-specific subunit PBRM1 similarly resulted in the co-immunoprecipitation of PHF6 (Supplementary Fig. S5d). Depletion of PHF6 in RT cells reduced the abundance of several SWI/SNF subunits (Fig. 5d, Supplementary Fig. S5e, Supplementary Table S1). The effect was particularly notable for the bromodomain-containing subunits, such as SMARCA4, BRD9, and the six-bromodomain-containing subunit PBRM1.

To evaluate the effects of PHF6 loss on SWI/SNF complex assembly, we performed density sedimentation assays. Of note, SWI/SNF complexes are large and heterogeneous and are, therefore, present in multiple consecutive fractions. We identified modest effects of PHF6 loss upon SWI/SNF complex assembly, as evidenced by molecular mass shifts in PBAF and ncBAF-associated components. Specifically, loss of PHF6 resulted in a profound reduction of PBRM1, accompanied by a shift of both ARID2 (another PBAF-specific subunit) and SMARCA4 (an ATPase subunit presents in all three SWI/SNF families) into lower molecular mass fractions (Fig. 5e). Consistent with these findings, re-expression of PHF6 in DND41 cells resulted in a substantial increase of PBRM1 and a modest increase of BRD9 protein abundance (Supplementary Fig. S5f).

To assess the impact of PHF6 depletion on the genomic targeting of bromodomain-containing SWI/SNF subunits, we performed ChIP-Seq for SMARCA4, PBRM1, and BRD9 in G401 RT cells in the presence and absence of PHF6. Depletion of PHF6 resulted in a decrease in the occupancy of PBRM1, SMARCA4, and BRD9 at active promoters typically bound by PHF6, which also exhibit a loss of H3K14ac (Fig. 5f and Supplementary Fig. S5g, h).

We have previously demonstrated that the absence of SMARCB1 in RT cells results in the ncBAF complex (the only SWI/SNF complex family of which SMARCB1 is not a member) becoming essential[8,9]. Given that we found co-localization of PHF6 and ncBAF, we asked whether loss of PHF6 resulted in effects upon transcription that are similar to those that occur following loss of ncBAF subunit BRD9. Gene set enrichment analysis revealed a strong correlation between genes downregulated following loss of PHF6 and those downregulated upon BRD9 knockdown, including pathways for cell cycle regulation, chromatin organization, and developmental processes (Supplementary Fig. S5i, j, k). These findings establish a direct, collaborative interplay between PHF6 and SWI/SNF at promoters, exerting precise transcriptional control.

PHF6 mediates cooperation between p300 and SWI/SNF complexes

Given the pronounced changes of H3K14ac and SWI/SNF subunit recruitment upon PHF6 knockdown, we next sought to investigate whether the loss of SWI/SNF binding was a direct consequence of PHF6 absence or instead mediated via the reduced acetylation of H3K14 causing reduced affinity of bromodomain-containing SWI/SNF subunits.

Immunoprecipitation of the SWI/SNF ATPase-SMARCA4 resulted in co-immunoprecipitation of both PHF6 and p300 (Fig. 5i). Likewise, immunoprecipitation of p300 co-IP'd SWI/SNF subunits and PHF6 (Supplementary Fig. S5l). Furthermore, depletion of PHF6 resulted in reduced occupancy of p300 at the sites co-bound by PHF6 and the SWI/SNF subunits (Fig. 5j, Supplementary Fig. S5m).

To evaluate the role of p300 and H3K14ac in the stability of SWI/SNF subunits, we employed the p300 inhibitor A-485. Despite the reduction in H3K14 acetylation upon p300 inhibition, immunoblots revealed no changes in BRD9, PBRM1, or SMARCA4 levels (Supplementary Fig. S5n). This contrasts with the effects observed upon PHF6 depletion, where changes in the abundance of bromodomain-containing SWI/SNF subunits were readily detected alongside reduced H3K14acetylation. This indicated that the reduction in SWI/SNF subunits was caused by the loss of PHF6 and not by reductions in histone acetylation that accompany PHF6 loss.

To further investigate inter-relationships between PHF6, SWI/SNF, and histone acetylation, we asked the converse: whether depletion of SWI/SNF subunits impacted H3K14acetylation, similar to PHF6 depletion. Knockdown of either BRD9 or PBRM1 (both bromodomain-containing SWI/SNF subunits) resulted in the reduction of H3K14ac and H3K27ac levels (Fig. 5k). Additionally, PHF6 protein was also reduced upon knockdown of either BRD9 or PBRM1 (Fig. 5l). To specifically assess the impact of BRD9 on the regulation of H3K14acetylation, we ectopically over-expressed BRD9 in G401 control and PHF6 knockdown cells. Over-expression of BRD9 resulted in increased acetylation of H3K14 and H3K27 (Supplementary Fig. S5o); interestingly, this effect was predominantly dependent on the presence of PHF6. The reciprocal changes observed upon loss of PHF6 and the depletion of SWI/SNF subunits upon the levels of each other, as well as their collective impact on H3K14 acetylation, support a model whereby these components interact as part of a transcriptional hub at promoters to license transcriptional elongation.

Dependency on *PHF6* is a consequence of *SMARCB1* loss

We next sought to determine whether it is the absence of *SMARCB1* that confers dependency on *PHF6*. As re-expression of SMARCB1 in RT cell lines not only results in differentiation and cell cycle arrest but also triggers downregulation of PHF6 (Supplementary Fig. S6a), we turned to our 293 T model in which we can independently control the expression of both PHF6 and SMARCB1. Knockdown of PHF6 impaired the proliferation of SMARCB1-deficient 293 Ts but had no effect upon the proliferation of 293Ts in which expression of *SMARCB1* had been restored (Fig. 6a), thus demonstrating dependency upon *PHF6* to be a specific consequence of *SMARCB1* loss.

Seeking to understand why the loss of SMARCB1 results in a dependency upon PHF6, we tested how the absence of SMARCB1 impacted the roles of PHF6 utilizing three independent model systems: the conditional 293T cells, the SMARCB1-wildtype cancer cell line HCT116, and primary *Phf6*$^{fl/fl}$ mouse embryonic fibroblasts. In all cases, PHF6 loss reduced H3K14acetylation (Fig. 6b,c, Supplementary Fig. S6b, c) and led to the accumulation of RNA Pol II Ser5P (Fig. 6d, e, Supplementary Fig. S6d) irrespective of SMARCB1 status. Consequently, these are intrinsic functions of PHF6 and not aberrant functions gained following SMARCB1 loss. Notably, however, the absence of SMARCB1 amplified the effects of PHF6 loss upon SWI/SNF subunit abundance such that PBRM1 levels were ablated and levels of BRD9, SMARCA4, and BAF60a substantially reduced (Fig. 6f).

Finally, to evaluate the relationship between *PHF6* and *SMARCB1* in vivo, we leveraged genetically engineered mouse models. We had previously demonstrated that conditional inactivation of *Smarcb1* with Lck-Cre gives rise to CD8+ mature T-cell lymphomas with rapid onset and full penetrance[51,52]. We interbred these *Smarcb1* floxed mice with *Phf6* floxed mice[53]. Deletion of *Phf6* (*Phf6*$^{fl/y}$ *Smarcb1*$^{fl/fl}$ *Lck-Cre*) significantly delayed the onset of lymphomas driven by *Smarcb1* deletion (Fig. 6g, Supplementary Fig. S6e). Synthesizing these in vivo results with our in vitro observations, we establish a critical role for PHF6 in modulating the survival and progression of SMARCB1-deficient tumors.

Collectively, our results demonstrate that while loss of PHF6 alone impacts histone acetylation, polymerase licensing, and the bromodomain-containing SWI/SNF subunits, it is in the context of loss of SMARCB1 where the impact on these functions leads to cellular lethality.

https://doi.org/10.1038/s41467-024-51566-5

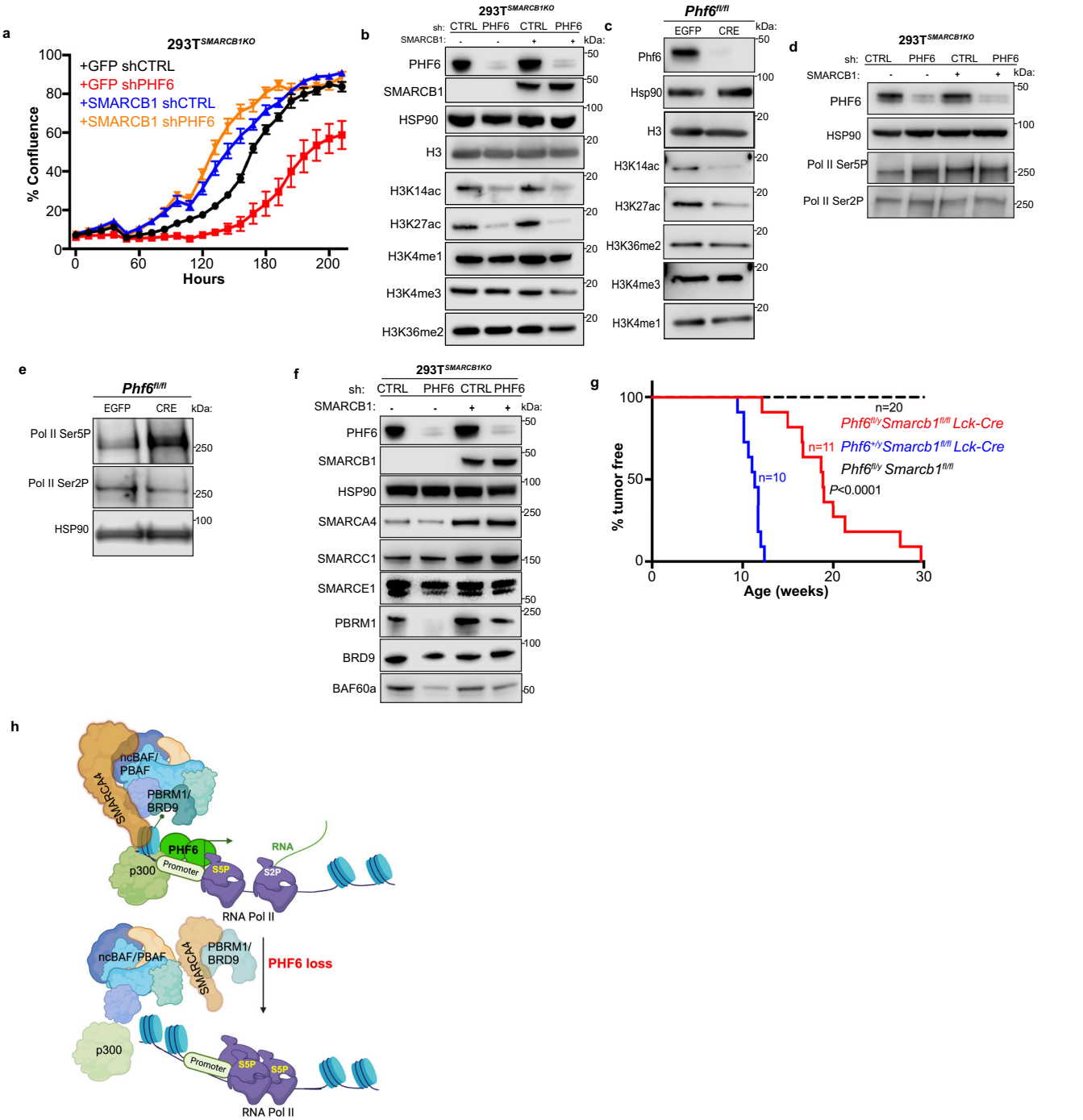

Fig. 6 | Dependency on PHF6 is a specific consequence of SMARCB1 loss. **a** Effects of shPHF6 vs shCTRL in SMARCB1-knockout 293 T cells treated with either GFP (+GFP) or SMARCB1 rescue (+SMARCB1) expression constructs. Data are represented from mean of $n = 16$ technical replicates per condition from one independent experiment. **b** Western blot analysis of histone modifications in 293T^{SMARCB1KO} cells following knockdown of PHF6 and/or re-expression of SMARCB1 (48 h) or GFP (control) ($n = 3$ independent biological replicates). **c** Western blots of histone modifications in mouse embryonic fibroblasts derived from *Phf6*$^{fl/y}$ mice treated with either EGFP (control) or Cre to delete *Phf6* ($n = 2$ independent biological replicates). Western blot analysis of the levels of Pol II Ser5P and Pol II Ser2P forms in (**d**) 293T^{SMARCB1KO} following knockdown of PHF6 and/or re-expression of SMARCB1 or GFP (48 h) ($n = 3$ independent biological replicates) (**e**) in mouse embryonic fibroblasts derived from *Phf6*$^{fl/fl}$ mice treated with either EGFP (control) or Cre to delete *Phf6* ($n = 2$ independent biological replicates). **f** Western blot analysis of SWI/SNF complex members in 293T^{SMARCB1KO} cell line following knockdown of PHF6 and/or re-expression of SMARCB1 or GFP (48 h) ($n = 3$ independent biological replicates). **g** Kaplan–Meier survival curves demonstrating percent mice that are tumor free among cohorts of *Phf6*$^{fl/y}$ *Smarcb1*$^{fl/fl}$ Lck-Cre mice, *Smarcb1*$^{fl/fl}$ Lck-Cre mice, and *Phf6*$^{fl/y}$ *Smarcb1*$^{fl/fl}$ mice. $n = 11$ mice per group. Statistical analysis was performed using two-sided Mantel-Cox test; $P = 0.0148$. **h** Schematic illustration of PHF6 at the interface of chromatin (p300, H3K14ac modified nucleosome, ncBAF, and PBAF complexes) and transcriptional regulation (proximal promoter pause release; Pol II Ser5P). Figure 6h created with BioRender.com released under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International license. Source data are provided as a Source Data file.

Discussion

Genetic syndromes can reveal surprising molecular connections among genes that initially seem unrelated. One example is Coffin-Siris Syndrome (CSS), where mutations in either *SWI/SNF* complex subunits or *PHF6* have been implicated[26,27]. Notably, mutually exclusive mutations in *SWI/SNF* subunits or *PHF6* also occur in T-ALL[24]. These findings piqued our interest when, through a robust CRISPR inactivation screen, we identified *PHF6* as specifically essential for the viability of SMARCB1-deficient RT cells. PHF6 was originally shown to influence RNA Pol I activity in the nucleolus[21], with more recent research highlighting its role in chromatin accessibility[22,23]. However, understanding its regulatory function in chromatin and transcription has been incomplete, and the mechanistic link between PHF6 and SWI/SNF complexes elusive.

Our study addresses these gaps by leveraging a model cancer in which inactivation of SWI/SNF subunit *SMARCB1* revealed specific dependence upon *PHF6*. The loss of SMARCB1 in RTs is notable for its impact on enhancers and transcriptional regulation. In the absence of SMARCB1, enhancer-bound SWI/SNF complexes are degraded, resulting in loss of accessibility and reduction of H3K27 acetylation, leading to a loss of enhancer function[14,15,54]. In contrast, promoters retain SWI/SNF complexes such as PBAF and ncBAF, and histone acetylation remains intact[16]. The necessity of these promoter-bound SWI/SNF complexes for the survival of *SMARCB1*-mutant cancer cells is demonstrated by residual SWI/SNF components such as BRD9 and SMARCA4 becoming preferentially essential[8,17].

Using rhabdoid cells, other cancer models, and normal cells, we discovered that PHF6 operates as a central player in an interaction network at active promoters. PHF6 is instrumental in recruiting SWI/SNF complexes and p300 to promoters, where it cooperates with these proteins to facilitate an active chromatin state and the progression of poised RNA polymerases (Fig. 6h). When considered in the context of enhanced dependence of RT cells upon BRD9 and SMARCA4, dependency upon PHF6 reveals a theme that select proteins involved in promoter activation become synthetic lethal upon loss of the enhancer regulatory function of SMARCB1. This synthetic lethality is conceptually reminiscent of the dependence of *BRCA*-mutant or *SWI/SNF*-mutant cancers upon PARP and EZH2, respectively, in which related proteins become essential upon mutation of a tumor suppressor[55,56].

It seems notable that loss of PHF6 has particularly strong effects on bromodomain-containing SWI/SNF subunits. Structurally, the bromodomains are the most outer-facing components of SWI/SNF complexes and are thought to be incorporated at the last stages of complex assembly[57–60]. This raises the possibility that appropriate recruitment of these bromodomain-containing subunits by PHF6 facilitates targeting and complete functionality of SWI/SNF complexes.

Taken together, our discovery that PHF6 serves key roles in facilitating SWI/SNF function in transcriptional regulation provides a mechanistic explanation for the phenocopy of *PHF6* and *SWI/SNF* mutations in causing Coffin-Siris syndrome and by the mutual exclusivity of their mutations in T-ALL. These insights suggest that targeting PHF6 may offer a new therapeutic opportunity in *SMARCB1*-mutant cancers.

Methods

Ethics statement

All animal experiments were reviewed and approved by the St. Jude Children's Research Hospital Institutional Animal Care and Use Committee (Protocol number: 595-100562). We complied with all relevant ethical guidelines while conducting this study.

Cell culture

G401 (ATCC-CRL1441), Jurkat (ATCC-TIB152), ES-2 (ATCC-CRL1978), Pfeiffer (ATCC-CRL2632), HCT116 (ATCC-CCL-247), 293 T(ATCC-CRL-3216), and MCF7 (ATCC HTB-22) cell lines were purchased from ATCC. TTC549 and TTC709 cell lines were a kind gift from Tim Triche at the University of California Los Angeles. CHLA-266 was obtained from the Children's Oncology Group. DND41 cells were provided by Dr. Charles Mullighan at St. Jude Children's Research Hospital. BIN67 cells were provided by Bernard E. Weissman at the University of North Carolina. G401 and ES-2 cells were grown in culture in McCoy's medium with 10% FBS (Sigma-Aldrich), and 1% GlutaMAX (Gibco); Jurkat, Pfeiffer, DND41, BIN67, TTC709, and TTC549 cells were grown in cultured in RPMI medium with 10% FBS and 1% GlutaMAX. MCF7, 293 T, and HCT116 were maintained in DMEM with 10% FBS and 1% GlutaMAX. CHLA-266 cells were grown in IMDM with 20% FBS, 1% P/S, and 1% ITS (insulin, transferrin, selenium). All cells were grown in culture at 37 °C with 95% humidity and 5% CO_2 and were regularly tested for mycoplasma by PCR (Genlantis). Cells were transduced at an MOI of 8–10 with shRNAs in the presence of Polybrene (Santa Cruz) (8 μg/mL) and were selected for 72 h with 1 μg/mL puromycin (Thermo Fisher). Tet-inducible re-expressing cells were maintained in Tet-System Approved FBS and induced with doxycycline (10 μg/mL, Clontech) for the time indicated.

Animals

Phf6$^{fl/fl}$ mice (C57BL/6) were a kind gift from Dr. Adolfo Ferrando at Columbia University. Lck-Cre mice (012837, C57BL/6) were purchased from the Jackson Laboratory. Animal maintenance and procedures were conducted in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals. DNA from tail biopsies of *Phf6*$^{fl/fl}$, *Smarcb1*$^{fl/fl}$, and Lck-Cre mice were harvested and genotyped as previously described[53,61]. The animals were maintained on a 12 h light/dark cycle, housed at 70 °F, 30–70% humidity levels and provided with food and water ad libitum.

Phf6$^{fl/y}$*Smarcb1*$^{fl/fl}$ *Lck-Cre* mouse lines

Mice with heterozygous alleles of *Smarcb1* (6-week) and homozygous alleles of *Phf6* Lck-Cre (6-weeks) were interbred to generate *Phf6*$^{fl/y}$*Smarcb1*$^{fl/fl}$ *Lck-Cre* mice. As *Phf6* is present on the X-chromosome, female mice were excluded from this study due to variable effects associated with random X-inactivation. *Phf6*$^{fl/y}$*Smarcb1*$^{fl/fl}$ *Lck-Cre* mice were monitored daily for signs of illness. Moribund mice were euthanized, and necropsy was performed on all animals. Additionally, spleen, thymus, liver, and bone marrow were harvested from moribund animals for subsequent downstream analysis. DNA-based allelic recombination of *Phf6* was confirmed by PCR (Supplementary Table S2). Allelic recombination of *Smarcb1* was confirmed as previously described[61].

Mouse embryonic fibroblast collection and culture. Mouse embryonic fibroblasts (MEFs) were prepared from *Phf6*$^{fl/fl}$ E13.5 embryos. Briefly, embryos were harvested, their internal organs were removed, and the remaining embryonic material was minced. The minced tissues were then trypsinized for 45 min, after which the cells were seeded in T-75 flasks containing MEF medium (DMEM containing 10% FBS, 1% GlutaMAX, 1% Pen/Strep, 0.1 mM MEM NA amino acids, and 55 mM β-mercaptoethanol). The cells were sub-cultivated at a 1:3 ratio. For Cre-mediated recombination experiments, 2.0×10^6 *Phf6*$^{fl/fl}$ MEFs were seeded in 10-cm^2 dishes and transduced with either G0169pacAd5CMVCre-eGFPpA or a G0169pacAd5CMV-eGFPpA control (from the University of Iowa Viral Vector Core) at an MOI of 8. The cells were harvested 72 h later, and the efficiency of recombination was assessed by Western blot analysis.

Lentiviral generation and infection. Lentiviral vectors were produced and titrated as described previously[62] and were generated by the St. Jude Vector Production and Development Core. Briefly, a 50-mL culture of SJ293TS cells was transfected with the lentiviral vector and with the helper plasmid pCAG-kGP1-1R, pCAG-VSVG, or pCAG4-RTR2, using PEIpro (Polyplus Transfection). Next day, transfected cells were

diluted with an equal volume of fresh medium containing 12.5 U/mL of Benzonase (Sigma-Aldrich). Vector supernatants were collected 48 h post-transfection, clarified by centrifugation at 330 × *g* for 5 min, and passed through a 0.22-μm filter. Lentiviral vector–containing supernatants were adjusted to 300 mM NaCl, 50 mM Tris pH 8.0 and loaded onto an Acrodisc Mustang Q membrane (Pall Life Sciences) in accordance with the manufacturer's instructions, using an Akta Avant chromatography system (GE Healthcare Bio-Sciences). The column was washed with 300 mM NaCl, 50 mM Tris pH 8.0 and the viral particles were eluted from the column with 2 M NaCl, 50 mM Tris pH 8.0, directly onto a PD10 desalting column (GE Healthcare). Vector-containing flow-through was diluted with an equal volume of phosphate-buffered saline containing 1% human serum albumin (Grifols Biologicals), passed through a sterile 0.22-μM filter, aliquoted, and stored at −80 °C.

Lentiviral vectors were titrated by transducing HOS cells with serially diluted vector preparations in the presence of Polybrene (5–8 μg/mL). Four days post-transduction, genomic DNA was isolated from the transduced HOS cells with a Quick-DNA Miniprep Kit (Zymo Research). Vector titers were determined by calculating the ratio of the copies of HIV *psi* to every two copies of *RPP30* via a QX200 Droplet Digital PCR System (Bio-Rad), then multiplying it by the number of cells transduced and, if necessary, by the dilution factor.

Vectors and stable cell line generation

PHF6 was knocked down in cells by lentiviral infection with PLKO Mission (Millipore Sigma) shRNA TRCN 0000367877 or TRCN0000020120. PBRM1 was knocked down with TRCN0000015994. BRD9 was knocked down by lentiviral infection with pGIPZ RHS4430-200159861 (Horizon Discovery). The cells were then selected with 1 μg/mL puromycin (Gibco). Isogenic 293T^{SMARCB1KO}SMARCB1i or 293T^{SMARCB1KO}GFPi cells and G401SMARCB1i cells were generated as described previously[8]. 293T^{SMARCB1KO}SMARCB1i or 293T^{SMARCB1KO}GFPi were then subjected to infection with shPHF6 or a non-targeting control, and this was followed by selection in puromycin (1 μg/μL) for 72 h and simultaneous induction of SMARCB1 or GFP (for 48 h).

All doxycycline-inducible cell lines were generated using the two-vector Tet system (Vectorbuilder). G401, TTC549, and DND41 cells were first infected with pLV-Bsd-CMV > Tet3G or pLV-Bsd-EF1A > Tet3G (VB180621-1194qmk) and selected with blasticidin (5 μg/mL) (Gibco). Subsequently, the cells were infected with TRE3G-driven expression vectors. Doxycycline-regulated temporal re-expression of exogenous PHF6 in G401 and TTC549 cells was achieved by using pLV-mCherry-TRE3G > hPHF6*mut (VB180124-1019kwn). Synonymous mutations (AAGCGATACGCGGCCGAAGTG, GCACTACAAATGTATGCTCTT) were inserted in the seed-code region of PHF6 cDNA to avoid the recognition of exogenous PHF6 cDNA by shRNAs (Sigma TRCN 0000367877) targeting endogenous PHF6. The cells were then sorted to isolate mCherry-positive cells. Temporal re-expression of BRD9 in G401 cells was achieved by infecting the cells with pLV-EGFP-TREhBRD9 (VB220425-1156xae) and then sorting for GFP-positive cells. Flow cytometry sorting was performed by the Flow Cytometry and Cell Sorting Shared Resource at St. Jude Children's Research Hospital. To knock down PHF6 in G401 and TTC549 Tet3G-mCherry-TRE3GhPHF6 cells or in G401 Tet3G-EGFP-TREhBRD9 cells, cells were infected with shPHF6 or a non-targeting control and selected in puromycin (1 μg/μL) for 72 h. To temporally re-express PHF6 in DND41 cells, the cells were infected with pLV-Puro-TRE3G > V5/hPHF6 (VB180621-1180vyt) and selected in puromycin (1 μg/mL) for 72 h.

Proliferation assays

Cells were transduced with shPHF6 or shCTRL at an MOI of 8 for 48 h and were then subjected to 72 h of selection in puromycin (1 μg/μL). After selection, the optimum numbers of G401 cells (3000 cells), TTC549 cells (4000), CHLA266 cells (8000), TTC709 cells (8000), MCF7 cells (5000), ES2 cells (2000), BIN67 cells (4000), G401PHF6i cells (3000), TTC549 PHF6i cells (3000), and 293T^{SMARCB1KO} cells (4000) were seeded into 96-well plates, using eight wells per condition tested. Cell proliferation was recorded every 24 h by an Incucyte Live-Cell Analysis System (Essen BioScience), used in accordance with the manufacturer's recommendations. The recorded cell confluence data were analyzed using the Incucyte Zoom software and plotted using GraphPad Prism software.

Treatment with the p300 inhibitor A-485

Cells were seeded at a density of 200,000 per well and treated with A-485 (Tocris, cat. no. 6387) at the indicated concentration for 72 h.

Protein extraction

Whole-cell protein extracts. Cells were lysed in NP-40 lysis buffer (150 mM NaCl, 0.5% NP-40, 50 mM Tris-HCl pH 8.0) for 10 min on ice, then centrifuged at 16,000 g for 10 min. The protein was then quantified by a Bradford assay (Bio-Rad).

Nuclear extracts

Nuclear extracts were prepared for protein immunoprecipitation. In brief, 20–40 × 10^6 cells were resuspended in Buffer A (20 mM HEPES pH 7.4, 10 mM KCl, 0.2 mM EDTA, 1× protease inhibitors), then lysed by adding 10% NP-40 and centrifuged for 1 min at 16,000 g to concentrate the nuclei in a pellet. The nuclei were resuspended in IP Buffer (20 mM phospho-buffer pH 7.5, 150 mM NaCl, 10% glycerol, 0.2% Triton X-100, 1× protease inhibitors) and 2.5 μL of nuclease (Thermo Fisher). The lysate was then incubated for 20 min at 4 °C followed by rotation at room temperature for 30 min. The nuclear supernatant was collected by centrifugation at 20,000 g for 10 min. The protein concentration was determined by a Bradford assay.

Histone extraction for immunoblotting. The EpiQuik Total Histone Extraction Kit (Epigentek) was used for histone extraction. For immunoblotting experiments, 7–10 μg of histone extracts from various cell lines and 15 μg of histones from mouse embryonic fibroblasts were loaded onto the gel.

Histone extraction for histone mass spectrometry. Histones were extracted for mass spectrometry according to the quantitative analysis protocols described previously[62]. Briefly, cell pellets were resuspended in nuclear isolation buffer (15 mM Tris-HCL, pH 7.5, 60 mM KCL, 15 mM NaCl, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 250 mM sucrose) in a 10:1 ratio and mixed by vortex mixing. Cell solutions were centrifuged at 1000 × *g* for 5–10 min at 4 °C. The supernatants were discarded, and the nuclear pellets were resuspended in 0.4 N H_2SO_4 at a 5:1 ratio. The cell lysates were placed on a rotary mixer at 4 °C for a period ranging from 1 h to overnight, depending on the number of cells. Subsequently, the protein lysates were centrifuged, and TCA was added to the supernatants to a concentration of 20%. The lysates were centrifuged at 3400 × *g* for 5 min, then the pellets were resuspended in acetone with 0.1% HCl. Lysates were then centrifuged at 4000 × *g* for 5 min, and the supernatants were discarded. After being washed twice with 100% acetone, the pellets were air dried and finally resuspended in 50 μL of water. The histone extracts were quantified by Bradford assays.

Immunoprecipitation

For each nuclear extract, 30 μL of DynaBeads (Thermo Fisher) were washed and resuspended in 200 μL of immunoprecipitation (IP) buffer. The beads were then incubated with rotation at room temperature for 10 min with 3 μg of antibody or rabbit IgG. Next, 200 μL of bead/antibody slurry was added to 500 μg of nuclear lysate, and the mixture was incubated at 4 °C overnight with rotation. Next day, the beads were washed three times with IP buffer. Protein was eluted from the

beads with 4× LDS buffer containing 10% 2-mercaptoethanol, with heating at 70 °C for 10 min.

Immunoblotting

For whole-cell extracts, 25 μg of protein was subjected to electrophoresis on NuPAGE 4-12% Bis-Tris or 3–8% Tris-Acetate gels (Invitrogen). The proteins were then transferred to a nitrocellulose membrane. Membranes were blocked in 5% milk or in 5% BSA in TBS-T (for phosphorylated proteins) and were then incubated with primary antibody diluted in blocking buffer (Supplementary Data File) at 4 °C overnight. Next day, the membranes were washed three times with TBS-T and incubated with secondary rabbit or mouse antibody (Jackson ImmunoResearch Laboratories) (diluted 1:10,000) at room temperature for 1 h. The membranes were then washed three times and imaged using a LI-COR Odyssey Fc system and LI-COR Image Studio software (LI-COR Biosciences).

Density sedimentation assays

Nuclear fractions were extracted from G401 and G401SMARCB1i cells by using Buffer A (20 mM HEPES pH 7.9, 10 mM KCl, 0.2 mM EDTA, and 1× Protease Inhibitors) to remove the cytoplasmic fraction, then the nuclear fraction was isolated with Buffer C (20 mM HEPES pH 7.9, 150 mM NaCl, 10 mM EDTA, 1 mM EGTA, and 1× Protease Inhibitors). For each cell line, 1 mg of nuclear protein (measured by a Bradford assay) was added to a 15-mL 10–30% glycerol gradient and centrifuged at 200,000 g in an SW40 rotor for 16 h at 4 °C. To ensure the reproducibility of the 10–30% glycerol gradients, a BioComp Gradient Master piston gradient fractionator was used to collect 20 fractions for analysis. Immunoblot analysis was then performed on the 20 fractions obtained from each cell line to assess the positioning of the SWI/SNF subunits or PHF6.

Quantitative histone mass spectrometry

Approximately 20 μg of acid-extracted histone was chemically derivatized to propionylate the unmodified lysine residues. The derivatized histones were then digested with trypsin (1 μg of trypsin per 20 μg of histone) at 37 °C overnight. The resulting peptides were then desalted and eluted in 70% acetonitrile/0.1% formic acid. Each peptide sample was directly injected for mass spectrometry without liquid chromatography[29].

Direct-injection mass spectrometry. Samples were sprayed directly into the mass spectrometer by using a TriVersa NanoMate (Advion) supported by ChipSoft software (Advion). Samples were programmed for analysis and acquired by contact closure in an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Samples were injected with a voltage of 1.7 kV and a gas pressure of 0.5 psi in the NanoMate. All scans were acquired in the Orbitrap, at 240,000 resolution for the full MS and at 120,000 resolution for tandem mass spectrometry (MS/MS, MS^2). MS^2 scans were performed in HCD mode for accurate identification of approximately 44 isobaric peptides. The AGC target for the tSIM-MSX scans was 10E6; detailed acquisition parameters are as described as described in reference[29]. Briefly, the parameters for the database search using the pFind search engine version 3.1 include a precursor m/z tolerance of ±10 ppm, a fragment m/z tolerance of ±0.02 Th for HCD and ±0.4 Th for CID, fully enzymatic trypsin cleavage after arginine, a maximum of 2 missed cleavages, a fixed modification of Propionyl[Peptide N-term] with a mass addition of +56.026, and variable modifications including Propionyl[K]/ + 56.026, Acetyl[K]/ + 42.011, Methyl_Propionyl[K]/ + 70.042, Dimethyl[K]/ + 28.031, Trimethyl[K]/ + 42.047, and Phospho[ST]/ + 79.966.

Histone peptide quantification. The in-house software EpiProfileLite was customized for direct injection-mass spectrometry and is available on GitHub at https://github.com/zfyuan/EpiProfileLite (the user guide is included). Raw files were searched for intensities of histone peptides in MS scans and were used to calculate the relative ratios of all forms+PTM for each histone peptide. Targeted pre-set MS/MS scans were analyzed to identify unique fragment ions in the MS/MS scans and were extracted to discriminate isobaric peptide intensities from the MS scans. In all, quantification was performed for 29 peptide sequences with 45 PTMs (methylations, acetylations, and phosphorylations) for a total of 151 histone marks plus 16 unmodified histone peptides for relative quantification of histone variants.

PHF6 CRISPR–Cas9 indel fitness assay

For the CRISPR fitness assay, 1×10^6 cells were transiently transfected with either the PHF6 guide (5′-cauuguccuggagcaacaau-3′) or a non-targeting control guide (5′-caaguagucggggaugucgg-3′). The guides were delivered as a precomplexed ribonuclear protein (RNP) consisting of 150 pmol of chemically modified sgRNA (Synthego) and 50 pmol of Cas9 protein (St. Jude Protein Production Facility) via nucleofection with a Lonza 4D-Nucleofector™ X-unit, using solution P3 and program EH-100 in a small (100-mL) cuvette, according to the manufacturer's recommended protocol. A portion of cells was collected for gDNA and total protein extraction at 7, 14, and 21 days post nucleofection. Genomic DNA was sequenced via targeted deep sequencing using gene-specific primers with partial Illumina adapter overhangs (hPHF6.F – 5′-accaatttgttttccttgacagaa-3′ and hPHF6R – 5′-attccttgtgaaggtttctctcgt-3′) (overhangs not shown). Next-generation sequencing (NGS) analysis of edited cell pools was performed using CRIS.py [63]. For the growth fitness assay, all indels were binned into in-frame, out-of-frame, or 0-bp indels Protein extracts obtained from each time point were assessed for PHF6 levels, and HSP90 was used as a loading control.

Chromatin immunoprecipitation and sequencing

Cells were trypsinized, washed twice with PBS, and apportioned into aliquots of 20 million cells each. The cells were fixed in 1.1% formaldehyde for 10 min then quenched with 0.125 M glycine for 5 min. For the PHF6 ChIP, this protocol was slightly modified: 2×10^7 cells were harvested and fixed for 30 min with 0.2 mM DSG (Thermo Fisher), followed by incubation in formaldehyde for 10 min and quenching with 0.125 M glycine for 5 min. Quenched cells were washed with PBS and stored at −80 °C until further use.

Cells were thawed on ice and nuclear lysates were extracted with a Covaris truChIP Chromatin Shearing Kit (Covaris) in accordance with the manufacturer's guidelines. Chromatin was then sheared with a Covaris E220 Focused-Ultrasonicator. The sonicated chromatin was centrifuged at 16,000 × *g* for 10 min at 4 °C. Five million cells per epitope were used for subsequent ChIPs. The chromatin was then incubated overnight with 5 μg of the indicated antibodies (Supplementary Data File) at 4 °C. All ChIP-Seq samples received 60 ng of spike-in (Active Motif) *Drosophila* chromatin and 4 μg of spike-in antibody (Active Motif).

Next day, antibody-bound chromatin was incubated with protein G DynaBeads (Thermo Fisher Scientific) for 2 h at 4 °C. The beads were then washed, and the chromatin was eluted from the beads for 30 min at 65 °C. The eluted chromatin was subjected to de-crosslinking after treatment with RNase A and Proteinase K (Thermo Scientific) at 65 °C overnight. The DNA was purified with a PCR Purification Kit (Qiagen) in accordance with the manufacturer's instructions.

DNA was quantified using Qubit HS Assay Kits (Thermo Fisher Scientific). ChIP-Seq libraries were prepped with 1–10 ng of DNA, using KAPA HyperPrep library prep kits (Roche) and following standard protocols, and was subsequently sequenced (50-bp single-end reads) on the Illumina NovaSeq 6000 system at the Hartwell Center for Biotechnology at St. Jude Children's Research Hospital.

CUT&RUN

CUT&RUN on DND41 cells was performed with the CUTANA ChIC/CUT&RUN kit (EpiCypher) according to manufacturer's guidelines. Briefly, 500,000 cells per sample were pelleted, rinsed with PBS, centrifuged at 600 g for 3 min. The cell pellet was then washed with the provided wash buffer. Subsequently, cells were resuspended in 100 µl of wash buffer and bound to activated Concanavalin A beads. Antibodies were then added to each sample: PHF6 (Active motif, 1 µg), H3K4me3 (EpiCypher, 0.5 µg), or Rabbit IgG as a negative control (EpiCypher, 1 µg), and samples nutating overnight at 4 °C. Beads were washed in cold permeabilization buffer containing 0.05% Digitonin and incubated with pAG-MNase for 10 min. Subsequently, 100 mM $CaCl_2$ was added to the samples and were incubated for two hours at 4 °C to activate MNase and cleave target chromatin. The reaction was halted with the stop buffer and CUT&RUN-enriched DNA was then purified using the SPRIselect reagents provided in the CUTANA ChIC/CUT&RUN kit (EpiCypher). DNA was quantified using Qubit HS Assay Kits (Thermo Fisher Scientific). Libraries were prepped using KAPA HyperPrep library prep kits (Roche) and with slight modifications indicated in the CUTANA kit above. CUT&RUN libraries were sequenced on the Illumina NovaSeq 6000 instrument (75-bp paired-end reads) at the Hartwell Center for Biotechnology at St. Jude Children's Research Hospital.

RNA collection and sequencing

RNA from 1×10^6 cells was extracted using TRIzol Reagent (Invitrogen) and then purified with a DirectZol Miniprep Kit (Zymo Research). RNA samples were quantified with a Qubit Fluorometer and their quality was assessed with Tape Station. Libraries were generated using an Illumina TruSeq Stranded mRNA Library Prep Kit, and appropriate spike-in was incorporated to avoid experimental errors. RNA-seq libraries were sequenced on an Illumina NovaSeq 6000 system with 75-bp paired-end reads.

Bru-seq. Ten million shCTRL and shPHF6 G401 cells were incubated with 2 mM bromouridine (Bru) (Sigma-Aldrich) at 37 °C for 30 min to label nascent RNA as previously described[41,64]. Cells were then lysed directly in TRIzol (Invitrogen). Total RNA was isolated and subjected to DNase treatment with a TURBO DNA-free Kit (Invitrogen), followed by the addition of Bru-labeled and unlabeled spike-ins. Bru-labeled RNA was then captured using anti-BrdU antibodies (BD Biosciences) conjugated to magnetic beads[65]. Preceding first-strand synthesis, ribosomal RNA was reduced via QIAseq FastSelect (Qiagen). Size-selection was performed using a 3% agarose gel after second-strand synthesis, after which stranded, libraries were prepared using a modified protocol for the TruSeq RNA Library Prep Kit (Illumina). A universal ligation adapter and dual-index, barcoded PCR primers were used for these libraries[65]. The libraries were sequenced at the University of Michigan Advanced Genomics Core on an Illumina NovaSeq 6000 system(-Illumina) to a depth of approximately 80 million reads, generating 150-bp paired-end reads.

ATAC-seq

ATAC-seq experiments were performed using the ATAC-seq Kit (Active Motif) and following the manufacturer guidelines. Briefly, nuclei were isolated from 100,000 G401 cells and were subjected to tagmentation at 37 °C for 30 min with shaking at 800 rpm. Tagmented DNA was purified using the provided columns. Libraries were generated from the purified DNA according to manufacturer's guidelines. Prepped libraries were subsequently sequenced on an Illumina NovaSeq 6000 to generate 200 million 100-bp paired end reads per sample.

Computational methods

ChIP-Seq analysis. Single-end 50-bp reads were mapped to the human genome hg19 (GRCh37-lite) by using bwa (version 0.7.12-r1039, default parameter)[66] and converted to a bam file using SAMtools (v1.2)[67]. biobambam2 (v2.0.87, DOI: 10.1186/1751-0473-9-13) was used to mark duplicated reads. Uniquely mapped reads were extracted by SAMtools (v1.2)[67]. SPP (v1.11)[68] was used for cross-correlation analysis. Uniquely mapped reads were extracted by SAMtools[67] (v1.2) and extended to the estimated fragment size from Cross-Correlation Analysis by BEDtools (v2.24.0)[69], then converted to BigWig track file by UCSC tools(v4)[70]. Tracks were then reviewed using the Integrated Genomic Viewer (IGV_2.4.14)[52].

Peaks were called by MACS2 (version 2.1.1.20160309, parameters "--nomodel --extsize fragment size")[71] for narrow peak analysis (H3K14ac, H3K27ac, PHF6, BRD9, PBRM1, SMARCA4, HBO1, and p300). Each mark had at least two replicates and reproducibility was confirmed by visual inspection. Reproducible peaks were generated using both CTRL and knockdown conditions to generate a reference peak set for each mark. Peaks were then merged using the BEDtools multiIntersectBed and merge functions.

The R package ChIPpeakAnno was used to annotate peaks and to visualize peak overlaps. The TSS (transcript start site) proximal regions were defined as the range from 2000 bp upstream of the TSS to 500 bp downstream of the TSS. The TES (transcription end site) proximal regions were defined as the range from 500 bp upstream of TTS to 200 bp downstream of TTS.

Heatmaps were generated with computeMatrix and plotHeatmap from deepTools[72]. Metaplots for were generated using computeMatrix (deepTools) and plotProfile with the following parameters: -a & -b 2000.

To identify differentially binding peaks, the number of fragments for each reference peak were counted using the intersect command from pybedtools (v0.8.1)[69,73]. The mapped fragments were then converted to an FPKM unit (fragments per kilo base per million mapped reads), and TMM (trimmed mean of M-values) using edgeR[74]. Limma-voom was used to determine significance of differential peak binding[75,76]. For targets for which global changes were predicted, an additional scaling factor (H3K14ac, H3K27ac, p300, PBRM1, SMARCA4) was added to the standard limma-voom approach. The scaling-factor for each target was generated based on Drosophila spike-in or the median enrichment signal derived from the union of high confidence peaks (BigWig files). The median-signal was then used to calculate the pseudo-spike-in (PSI) read counts [(1/median) * 1million], which was then used to determine the scaling factor. The scaling factor was subsequently supplied to the "norm.factors" parameter of the DGEList function from edgeR.

CUT & RUN analysis. Paired end reads were initially aligned against the E. coli reference genome using Bowtie2 (version 1.2.2) with parameters -k 1 –best –un to filter out bacterial sequences and retain unmapped reads. These were then mapped to the human reference genome (hg19) using Bowtie2 in paired-end mode with -p 20 -k 2 -m 2 –best. The resulting alignments were processed using samtools sort -n, followed by bedtools bamToBed -bedpe for conversion to BEDPE and then to BED3 format, with further conversion to BAM for downstream analyses. We extracted fragments smaller than 2000 bp, focusing on the central 80 bp of each fragment to generate bigwig tracks normalized to 10 million fragments for visualization in the IGV browser (version 2.16.2). Peak calling was performed with MACS2, using parameters –keep-dup=auto and a *p*-value cutoff of 1e-9, while excluding ENCODE-identified problematic regions.

Chromatin states. Seven histone marks (H3K14Ac, H3K27Ac, H3K27me3, H3K36me2, H3K4me1, H3K4me3, and H3K9me3) in parent cells have been used for HMM model training by ChromHMM(version 1.10)[9]. Non-duplicated reads from all replicates were first extended to the fragment size (detected by SPP v1.1)[6], combined and then binarized by ChromHMM ("java-jar ChromHMM.jar BinarizeBed -b 200 -colfields

0,1,2,5 -center"). The 12-state model was chosen after reviewing the AUC (area under the curve) and ROC (receiving operating characteristics). Next, genomic regions such as PHF6 peaks or PBRM1, or SMARCA4, or BRD9, or H3K14ac were annotated to chromatin states using intersect function from pybedtools (version 0.8.1). For state assignments, regions of interests were assigned to the state with the highest percentage overlap. Finally, the fraction of the peaks to states annotation was calculated and visualized as a heatmap, as shown in Supplementary Fig. 2d and Figs. 3a, 5b, g.

Super enhancer and enhancers. Super enhancers and traditional enhancers were defined by ROSE[77] using previously defined MACS2 peaks and read coverage for each replicate of H3K27ac and the following parameters -s 12500 -t 2500 -g hg19 -e spp fragment size -l -1. Replicated enhancers and super enhancers were defined using BEDtools intersect.

Metaplots for ROSE defined enhancers and super enhancers were generated using computeMatrix (deepTools) with the scale-regions setting and the following parameters: -a & -b 2000.

Pol II metagene plots. All PHF6-bound regions were annotated with nearby genes using hg19 and HOMER[78] (v4.10). Gene and PHF6-region associations were filtered to only include only those protein-coding genes that have PHF6 binding within ±5 kb of the TSS. To define genes with no PHF6 binding, extended PHF6 reads for each replicate were counted for all 2 kb TSS regions using BEDtools intersect. PHF6 read counts were converted to FPKM, and the average was computed for every TSS. Genes with no PHF6 binding were defined as all TSSs with an average FPKM <1st quartile. For genes (8728) with nearby PHF6 binding and genes (4812) with no PHF6 binding, coordinates defined by biomaRt[79] (v2.42.1) for (TSS) to (TES) were used to plot normalized read coverage for total Pol II, Pol IISer2p, and Pol IISer5p.

Bru-seq nascent RNA metagene plots. Using bamCoverage (deepTools v3.5.1), strand-specific BigWig files were generated for uniquely mapping alignments (parameters used: --binSize 1 --normalizeUsing RPKM --exactScaling --minMappingQuality 60 --filterRNAstrand forward *or* reverse). These BigWig files were then used to generate the metagene plots of the RPKM-normalized Bru-seq signals for different gene lengths (10–60 kb, 60 kb and above), using the TSS and TES as the reference points for genes bound by PHF6 (8728). computeMatrix (deepTools v3.5.1) was used to calculate the region scores with the following parameters: -a = 2000 and -b = 10000 for the TSS, -a = 10000 and -b = 2000 for the TES (additional parameters: --binSize 50 --sortUsing median --averageTypeBins median). The resulting strand-specific matrices were merged using computeMatrixOperations (deepTools v3.5.1) to generate the plots using plotProfile (deepTools v3.5.1).

ATAC-sequencing. Residual adapter sequences and low-quality bases were removed using trim galore (v0.6.3). Quality reads were aligned to GRCh37.p19 (hg19) using pbrun fq2bam (v3.0.0.6) using the following parameters: --bwa-options -K 10000000 -Y -M --num-gpus 2. Quality reads were filtered using SAMtools[67](view -F 1804 -b -q 1 and bamsort). Bam files were converted to bed files (bamToBed) and chrM reads were removed. Coverage was normalized to 20 million uniquely mapped reads for all fragments <2 kb. Peaks were called for NFR for each replicated using MACS2 (--nomodel --keep-dup all -q 0.05 | -q 0.5). Reproducible peaks were defined using a combination of high and low confidence peaks for all replicates using bedtools intersect. Fragments (<109 bp representing the nucleosome free fragments) were counted for each samples using bedtools intersect (-c). Differential accessibility of reproducible peaks across all replicates was performed using the lmFit and eBayes functions of the R package limma (v3.42.2 voom function)[57]. Significant differential accessibility was defined as absolute log2FC > 0 and FDR < 0.05. Accessibility within reproducible peaks was normalized using FPKM for MA plot.

To determine the positioning of PHF6 over the nucleosomes, mononucleosome signal was generated using the ATACseqQC package[80] (v1.10.4) following standard methods (https://bioconductor.org/packages/release/bioc/vignettes/ATACseqQC/inst/doc/ATACseqQC.html) in R (v3.6.1). Mononucleosome signal and PHF6 binding were then visualized at ±1 kb around the TSS on chromosome 1.

RNA-Sequencing. The RNA-seq reads were mapped to the human GRCh37-lite reference genome with STAR[81]. Gene-level counts were quantified by RSEM (v1.3.1) with GENCODE version 19 Ensembl 74 annotation. The read counts were further normalized with the TMM[56] method from the R package edgeR. Differential gene expression analysis was performed with the lmFit and eBayes functions of the R package limma (v3.42.2 voom function)[57]. Genes were considered differentially expressed if their adjusted *P* values were lower than 0.05 and the FC was higher than 0. A heat map was generated using z-score, row scaled, normalized expression for DEG.

Binding and expression target analysis (BETA). The BETA tool[82] was used to predict PHF6 activating/repressive roles of PHF6, using the parameter -da 500 to predict different classes of peaks (activating or repressive). Regulatory potential score was calculated using $Sg = \sum ki = 1e - (0.5 + 4\Delta i)\mathrm{Sg} = \sum_{\mathrm{i}}^{\mathrm{k}} = 1\mathrm{e}{-^{(0.5+4\Delta \mathrm{i})}}$. All peaks (k) within 100 kb of the (TSS) are considered. The distance between a binding site and the TSS is Δ, which is proportional to 100 kb. *P*-values were calculated by Kolmogorov–Smirnov test to measure the significance of the upregulated genes group or downregulated genes group relative to static genes group.

Gene set enrichment analysis. (GSEA) was conducted with GSEApy[83](v0.10.4), run in pre-ranked mode and following the methodology as originally described[84]. For each GSEA run, the listed genes were ranked using a ranking metric. This metric was equivalent to the log2(fold-change) value for each gene, which was derived from differential gene expression analyses. Moreover, as GSEA requires a large collection of gene sets, a merged collection of control gene sets was generated, comprising gene sets from the Molecular Signatures Database (MdigDB), Hallmark gene sets, Reactome, and the KEGG database (version 7.1) and our own internal genesets derived from various differential expression analyses[84,85].

Gene set fragment length analysis. Gene sets were defined as (1) all hg19 genes or (2) PHF6 target genes identified as being significantly down-regulated in all RT cell lines. Gene start and end positions were obtained using the getBM function in biomaRt (v.2.54.0) and were used to calculate gene lengths. Gene lengths were log10 transformed and kernel density estimation was performed using the density function from the stats package (v.4.2.2).

Statistics and reproducibility

GraphPad Prism 7 or R (v3.6.1) packages were used to conduct the statistical analysis. Relevant statistical tests are described in the figure legends.

Power analysis was conducted to determine the sample size for animal cohorts. Experiments were not randomized, and the Investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All NGS data generated have been deposited at the Gene Expression Omnibus (GEO) under accession number: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE180487. MS-based proteomics raw data files are available at MassIVE under the identifier: https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD046064. Source data are provided with this paper.

Materials availability

Plasmids and cell lines generated will be available upon request and upon completion of a material transfer agreement.

References

1. Shain, A. H. & Pollack, J. R. The spectrum of SWI/SNF mutations, ubiquitous in human cancers. *PLoS One* **8**, e55119 (2013).
2. Kadoch, C. et al. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet* **45**, 592–601 (2013).
3. Mittal, P. & Roberts, C. W. M. The SWI/SNF complex in cancer—biology, biomarkers and therapy. *Nat. Rev. Clin. Oncol.* **17**, 435–448 (2020).
4. Versteege, I. et al. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394**, 203–206 (1998).
5. Roberts, C. W., Galusha, S. A., McMenamin, M. E., Fletcher, C. D. & Orkin, S. H. Haploinsufficiency of Snf5 (integrase interactor 1) predisposes to malignant rhabdoid tumors in mice. *Proc. Natl. Acad. Sci. USA* **97**, 13796–13800 (2000).
6. Finetti, M. A., Grabovska, Y., Bailey, S. & Williamson, D. Translational genomics of malignant rhabdoid tumours: Current impact and future possibilities. *Semin Cancer Biol.* **61**, 30–41 (2020).
7. Lee, R. S. et al. A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. *J. Clin. Invest.* **122**, 2983–2988 (2012).
8. Wang, X. et al. BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nat. Commun.* **10**, 1881 (2019).
9. Michel, B. C. et al. A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nat. Cell Biol.* **20**, 1410–1420 (2018).
10. Alpsoy, A. & Dykhuizen, E. C. Glioma tumor suppressor candidate region gene 1 (GLTSCR1) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *J. Biol. Chem.* **293**, 3892–3903 (2018).
11. Liao, L. et al. High affinity binding of H3K14ac through collaboration of bromodomains 2, 4 and 5 is critical for the molecular and tumor suppressor functions of PBRM1. *Mol. Oncol.* **13**, 811–828 (2019).
12. Enriquez, P. et al. Binding specificity and function of the SWI/SNF subunit SMARCA4 bromodomain interaction with acetylated histone H3K14. *J. Biol. Chem.* **297**, 101145 (2021).
13. Mashtalir, N. et al. Chromatin landscape signals differentially dictate the activities of mSWI/SNF family complexes. *Science* **373**, 306–315 (2021).
14. Nakayama, R. T. et al. SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nat. Genet* **49**, 1613–1623 (2017).
15. Wang, X. et al. SMARCB1-mediated SWI/SNF complex function is essential for enhancer regulation. *Nat. Genet* **49**, 289–295 (2017).
16. Wolf, B. K. et al. Cooperation of chromatin remodeling SWI/SNF complex and pioneer factor AP-1 shapes 3D enhancer landscapes. *Nat. Struct. Mol. Biol.* **30**, 10–21 (2023).
17. Wang, X. et al. Oncogenesis caused by loss of the SNF5 tumor suppressor is dependent on activity of BRG1, the ATPase of the SWI/SNF chromatin remodeling complex. *Cancer Res.* **69**, 8094–8101 (2009).
18. Dharia, N. V. et al. A first-generation pediatric cancer dependency map. *Nat. Genet* **53**, 529–538 (2021).
19. Doench, J. G. et al. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).
20. Todd, M. A. & Picketts, D. J. PHF6 interacts with the nucleosome remodeling and deacetylation (NuRD) complex. *J. Proteome Res.* **11**, 4326–4337 (2012).
21. Todd, M. A., Huh, M. S. & Picketts, D. J. The sub-nucleolar localization of PHF6 defines its role in rDNA transcription and early processing events. *Eur. J. Hum. Genet* **24**, 1453–1459 (2016).
22. Meacham, C. E. et al. A genome-scale in vivo loss-of-function screen identifies Phf6 as a lineage-specific regulator of leukemia cell growth. *Genes Dev.* **29**, 483–488 (2015).
23. Oh, S. 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.* **48**, 9037–9052 (2020).
24. Liu, Y. et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat. Genet* **49**, 1211–1218 (2017).
25. Van Vlierberghe, P. et al. PHF6 mutations in T-cell acute lymphoblastic leukemia. *Nat. Genet* **42**, 338–342 (2010).
26. Zweier, C. et al. Females with de novo aberrations in PHF6: clinical overlap of Borjeson-Forssman-Lehmann with Coffin-Siris syndrome. *Am. J. Med Genet C Semin Med Genet* **166C**, 290–301 (2014).
27. Sekiguchi, F. et al. Genetic abnormalities in a large cohort of Coffin-Siris syndrome patients. *J. Hum. Genet* **64**, 1173–1186 (2019).
28. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet* **49**, 1779–1784 (2017).
29. Bhanu, N. V., Sidoli, S. & Garcia, B. A. A workflow for ultra-rapid analysis of histone post-translational modifications with direct-injection mass spectrometry. *Bio Protoc.* **10**, e3756 (2020).
30. Dancy, B. M. & Cole, P. A. Protein lysine acetylation by p300/CBP. *Chem. Rev.* **115**, 2419–2452 (2015).
31. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H. & Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* **87**, 953–959 (1996).
32. Huang, Z. Q., Li, J., Sachs, L. M., Cole, P. A. & Wong, J. A role for cofactor-cofactor and cofactor-histone interactions in targeting p300, SWI/SNF and Mediator for transcription. *EMBO J.* **22**, 2146–2155 (2003).
33. Alver, B. H. et al. The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nat. Commun.* **8**, 14648 (2017).
34. Pena-Hernandez, R. et al. BAZ2A-mediated repression via H3K14ac-marked enhancers promotes prostate cancer stem cells. *EMBO Rep.* **22**, e53014 (2021).
35. Kueh, A. J. et al. HBO1 (KAT7) Does not have an essential role in cell proliferation, DNA replication, or histone 4 acetylation in human cells. *Mol. Cell Biol.* **40**, e00506 (2020).
36. Mayer, A. et al. Native elongating transcript sequencing reveals human transcriptional activity at nucleotide resolution. *Cell* **161**, 541–554 (2015).
37. Weber, C. M., Ramachandran, S. & Henikoff, S. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell* **53**, 819–830 (2014).
38. Heidemann, M., Hintermair, C., Voss, K. & Eick, D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim Biophys. Acta.* **1829**, 55–62 (2013).
39. Komarnitsky, P., Cho, E. J. & Buratowski, S. Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev.* **14**, 2452–2460 (2000).

40. Adelman, K. & Lis, J. T. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet* **13**, 720–731 (2012).
41. Paulsen, M. T. et al. Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods* **67**, 45–54 (2014).
42. Krajewska, M. et al. CDK12 loss in cancer cells affects DNA damage response genes through premature cleavage and polyadenylation. *Nat. Commun.* **10**, 1757 (2019).
43. Zumer, K. et al. Two distinct mechanisms of RNA polymerase II elongation stimulation in vivo. *Mol. Cell* **81**, 3096–3109.e3098 (2021).
44. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
45. Zhu, Z. et al. Mitotic bookmarking by SWI/SNF subunits. *Nature* **618**, 180–187 (2023).
46. Andrades, A. et al. SWI/SNF complexes in hematological malignancies: biological implications and therapeutic opportunities. *Mol. Cancer* **22**, 39 (2023).
47. Leung, J. Y. & Nevins, J. R. E2F6 associates with BRG1 in transcriptional regulation. *PLoS One* **7**, e47967 (2012).
48. Brahma, S. & Henikoff, S. The BAF chromatin remodeler synergizes with RNA polymerase II and transcription factors to evict nucleosomes. *Nat. Genet* **56**, 100–111 (2024).
49. Shen, W. et al. Solution structure of human Brg1 bromodomain and its specific binding to acetylated histone tails. *Biochemistry* **46**, 2100–2110 (2007).
50. Charlop-Powers, Z., Zeng, L., Zhang, Q. & Zhou, M. M. Structural insights into selective histone H3 recognition by the human Polybromo bromodomain 2. *Cell Res.* **20**, 529–538 (2010).
51. Roberts, C. W., Leroux, M. M., Fleming, M. D. & Orkin, S. H. Highly penetrant, rapid tumorigenesis through conditional inversion of the tumor suppressor gene Snf5. *Cancer Cell* **2**, 415–425 (2002).
52. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
53. Wendorff, A. A. et al. Phf6 loss enhances hsc self-renewal driving tumor initiation and leukemia stem cell activity in T-ALL. *Cancer Discov.* **9**, 436–451 (2019).
54. Radko-Juettner, S. et al. Targeting DCAF5 suppresses SMARCB1-mutant cancer by stabilizing SWI/SNF. *Nature* **628**, 442–449 (2024).
55. Lord, C. J. & Ashworth, A. PARP inhibitors: synthetic lethality in the clinic. *Science* **355**, 1152–1158 (2017).
56. Kim, K. H. et al. SWI/SNF-mutant cancers depend on catalytic and non-catalytic activity of EZH2. *Nat. Med* **21**, 1491–1496 (2015).
57. He, S. et al. Structure of nucleosome-bound human BAF complex. *Science* **367**, 875–881 (2020).
58. Yuan, J., Chen, K., Zhang, W. & Chen, Z. Structure of human chromatin-remodelling PBAF complex bound to a nucleosome. *Nature* **605**, 166–171 (2022).
59. Mashtalir, N. et al. Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell* **175**, 1272–1288.e1220 (2018).
60. Mashtalir, N. et al. A structural model of the endogenous human BAF complex informs disease mechanisms. *Cell* **183**, 802–817.e824 (2020).
61. Wang, X. et al. TCR-dependent transformation of mature memory phenotype T cells in mice. *J. Clin. Invest* **121**, 3834–3845 (2011).
62. Sidoli, S. et al. One minute analysis of 200 histone posttranslational modifications by direct injection mass spectrometry. *Genome Res.* **29**, 978–987 (2019).
63. Connelly, J. P., Pruett-Miller, S. M. & CRIS py: a versatile and high-throughput analysis program for CRISPR-based genome editing. *Sci. Rep.* **9**, 4194 (2019).
64. Paulsen, M. T. et al. Coordinated regulation of synthesis and stability of RNA during the acute TNF-induced proinflammatory response. *Proc. Natl. Acad. Sci. USA* **110**, 2240–2245 (2013).
65. Magnuson, B. et al. CDK12 regulates co-transcriptional splicing and RNA turnover in human cells. *iScience* **25**, 105030 (2022).
66. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
67. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
68. Kharchenko, P. V., Tolstorukov, M. Y. & Park, P. J. Design and analysis of ChIP-Seq experiments for DNA-binding proteins. *Nat. Biotechnol.* **26**, 1351–1359 (2008).
69. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
70. Kuhn, R. M., Haussler, D. & Kent, W. J. The UCSC genome browser and associated tools. *Brief. Bioinform.* **14**, 144–161 (2013).
71. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
72. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A. & Manke, T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191 (2014).
73. Dale, R. K., Pedersen, B. S. & Quinlan, A. R. Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. *Bioinformatics* **27**, 3423–3424 (2011).
74. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
75. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
76. Law, C. W., Chen, Y., Shi, W., Smyth, G. K. & voom Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
77. Loven, J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).
78. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
79. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
80. Ou, J. et al. ATACseqQC: a Bioconductor package for post-alignment quality assessment of ATAC-seq data. *BMC Genomics* **19**, 169 (2018).
81. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
82. Wang, S. et al. Target analysis by integration of transcriptome and ChIP-Seq data with BETA. *Nat. Protoc.* **8**, 2502–2515 (2013).
83. Fang, Z., Liu, X. & Peltz, G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **39**, btac757 (2023).
84. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
85. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).

Acknowledgements

We thank members of the Roberts Lab for insightful discussions, especially Dr. Stacy Throm, and Kendall Wills for their contributions. The authors thank Keith A. Laycock for scientific editing of the manuscript. We also thank June Bursi for animal technical services and Michelle Paulsen for performing the Bru-seq experiments. This work was supported by grants from the National Cancer Institute (NCI) (R01 CA113794, R01 CA172152, and R01 CA273455) to C.W.M.R., from National Institute of

Health (NIH) (R01HD106051, P01CA196539) to B.A.G., from CURE AT/RT Now to C.W.M.R, from the Garrett B. Smith Foundation to C.W.M.R., and from the St. Jude Children's Research Hospital Collaborative Research Consortium on Chromatin Regulation in Pediatric Cancer to B.A.G. and C.W.M.R. We thank the following St. Jude core facilities: the Protein Production Facility for Cas9, the Animal Resources Center for animal care, the Veterinary Pathology Core for animal specimen preparation, the Flow Cytometry and Cell Sorting Shared Resource for FACS analysis, The Hartwell Center for Biotechnology for library preparation and sequencing, and Vector Development and Production Core for virus preparation. The St. Jude core facilities are supported by a NCI Cancer Center Support Grant (CCSG 2 P30 CA021765) and by the American Lebanese Syrian Associated Charities (ALSAC). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

P.M. and C.W.M.R. conceived the experiments and the study design. J.A.M., W.R., and B.X. performed all computational and statistical analyses with guidance and oversight from G.W.; P.M., H.A.M., S.R.-J., R.D.C., R.J.M., M.P., A.N.R., and Y.D. contributed to the cell line experiments. P.M., H.A.M., and R.D.C. performed the animal studies. J.E.R. conducted pathological analysis. P.M., H.A.M., S.R.-J., and Z.Z. contributed to the ChIP-Seq experiments. P.M., H.A.M., and S.R.-J. contributed to the RNA-Seq experiments. R.D.C. conducted the ATAC-Seq experiment. S.M.P-M. designed the CRISPR fitness assay and guides. B.S.H. generated the CRISPR-edited pools and analyzed the CRISPR fitness assay results. I.V.N. and M.L. performed and analyzed the Bru-seq experiments. B.A.G. and N.B. contributed to the histone PTM MS analysis. P.M., J.A.M., H.A.M., S.R.-J., R.D.C., M.P., W.R., N.B., A.N.R., B.X., G.W., S.M.P.-M., J.F.P., and C.W.M.R. contributed to the interpretation of experiments. S.R.-J. and R.D.C. contributed to editing and figure generation. P.M. and C.W.M.R. wrote the manuscript with input from all co-authors.

Competing interests

The authors declare no competing interests.

Ethics

All animal experiments were reviewed and approved by the St. Jude Children's Research Hospital Institutional Animal Care and Use Committee (Protocol number: 595-100562). We complied with all relevant ethical guidelines while conducting this study.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51566-5.

Correspondence and requests for materials should be addressed to Charles W. M. Roberts.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.