scientific reports

OPEN

PRDM6 promotes medulloblastoma by repressing chromatin accessibility and altering gene expression

Christin Schmidt[1,3,4,5,9], Sarah Cohen[1,3,4,5,6], Brian L. Gudenas[8], Sarah Husain[5,9], Annika Carlson[1,3,4,5], Samantha Westelman[5,9], Linyu Wang[5,9], Joanna J. Phillips[1,5,10], Paul A. Northcott[8], William A. Weiss[1,4,5,9✉] & Bjoern Schwer[1,2,3,4,5,6,7✉]

SNCAIP duplication may promote Group 4 medulloblastoma via induction of PRDM6, a poorly characterized member of the *PRDF1 and RIZ1 homology domain-containing* (PRDM) family of transcription factors. Here, we investigated the function of PRDM6 in human hindbrain neuroepithelial stem cells and tested PRDM6 as a driver of Group 4 medulloblastoma. We report that human PRDM6 localizes predominantly to the nucleus, where it causes widespread repression of chromatin accessibility and complex alterations of gene expression patterns. Genome-wide mapping of PRDM6 binding reveals that PRDM6 binds to chromatin regions marked by histone H3 lysine 27 trimethylation that are located within, or proximal to, genes. Moreover, we show that PRDM6 expression in neuroepithelial stem cells promotes medulloblastoma. Surprisingly, medulloblastomas derived from PRDM6-expressing neuroepithelial stem cells match human Group 3, but not Group 4, medulloblastoma. We conclude that PRDM6 expression has oncogenic potential but is insufficient to drive Group 4 medulloblastoma from neuroepithelial stem cells. We propose that both PRDM6 and additional factors, such as specific cell-of-origin features, are required for Group 4 medulloblastoma. Given the lack of PRDM6 expression in normal tissues and its oncogenic potential shown here, we suggest that PRDM6 inhibition may have therapeutic value in PRDM6-expressing medulloblastomas.

Medulloblastoma, a malignant childhood tumor of the cerebellum, represents a clinically and molecularly heterogeneous cancer originating from aberrant hindbrain development[1–4]. The cellular origins and genetic drivers of WNT and SHH medulloblastomas are relatively well characterized but drivers of Group 3 and 4 medulloblastoma remain unclear. Especially Group 4 medulloblastomas show a complex range of molecular aberrations and clinical outcomes[4,5].

One of the most characteristic alterations of Group 4 medulloblastoma is the expression of PRDM6, which occurs in 17% of all Group 4 medulloblastoma cases[1,4]. Moreover, PRDM6-expressing Group 4 medulloblastomas often show other alterations, such as aberrant expression of the *MYCN* oncogene[5]. The upregulation of PRDM6 expression in Group 4 medulloblastoma is thought to occur via enhancer hijacking caused by tandem duplication of *SNCAIP*, which is located approximately 600 kb upstream of *PRDM6*[4]. Although other genes are located within the topologically-associated domains of *SNCAIP*, only *PRDM6* shows robust upregulation[4]. PRDM6 is a poorly characterized member of the *PRDF1 and RIZ1 homology domain-containing* (PRDM) family of transcription

[1]Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA, USA. [2]Department of Cellular & Molecular Pharmacology, University of California, San Francisco, San Francisco, CA, USA. [3]Eli and Edythe Broad Center of Regeneration Medicine and Stem Cell Research, University of California, San Francisco, San Francisco, CA, USA. [4]Weill Institute for Neuroscience, University of California, San Francisco, San Francisco, CA, USA. [5]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA, USA. [6]Bakar Aging Research Institute, University of California, San Francisco, San Francisco, CA, USA. [7]Kavli Institute for Fundamental Neuroscience, University of California, San Francisco, San Francisco, CA, USA. [8]Division of Brain Tumor Research, Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, TN, USA. [9]Department of Neurology, University of California, San Francisco, San Francisco, CA, USA. [10]UCSF Tumor SPORE Biorepository, University of California, San Francisco, San Francisco, CA, USA. ✉email: william.weiss@ucsf.edu; bjoern.schwer@ucsf.edu

factors that possess histone lysine methyltransferase activities through their catalytic SET domain[6]. However, the enzymatic activity and molecular function of human PRDM6 is unclear. Based on findings in other species, PRDM6 is currently classified as a pseudo-methyltransferase that may interact with the methyltransferase G9a and corepressors to indirectly affect the methylation status of histone H3 lysine 9 and histone H4 lysine 20[6–8].

Several PRDM proteins have been implicated in cellular processes related to stem cell maintenance or differentiation in the central nervous system[9], but the role of PRDM6 in the developing and mature brain is not understood. Further, the role of PRDM6 expression in the etiology of Group 4 medulloblastoma or other brain cancers has not been tested. To assess if PRDM6 is a driver of Group 4 medulloblastoma, we set out to elucidate the functional impact and oncogenic potential of sustained PRDM6 expression in human neuroepithelial stem (NES) cells. Human NES cells are multipotent and retain the ability to give rise to cerebellar lineage cells, including granule neuron progenitors in cell culture and after orthotopic implantation in mice[10,11]. Because of their genetic stability and competence as cerebellar progenitors, NES cells are well-suited for studies of medulloblastoma development[11–14].

Results

PRDM6 expression and subcellular localization in human hindbrain neuroepithelial stem (NES) cells

We used human hindbrain neuroepithelial stem (NES) cells (Fig. 1A), a representative of cerebellar progenitors that normally do not express PRDM6 (Fig. 1B), to elucidate the molecular function and tumorigenic potential of PRDM6. We transduced NES cells with lentiviral vectors expressing V5 epitope-tagged human PRDM6 (hereafter referred to as "PRDM6 NES" cells) or the parental, empty lentiviral control vector (hereafter referred to as "EV NES" cells) (Fig. 1A). qRT-PCR analysis and immunoblotting showed robust PRDM6 expression at the transcript and protein levels, respectively, in PRDM6 NES but not in EV NES cells (Fig. 1B and C, Figure S1A).

The subcellular localization of human PRDM6 has not been reported. To gain insight into its molecular function, we first assessed its subcellular localization in NES cells. Consistent with prior reports on mouse PRDM6[7,8], subcellular fractionation (Fig. 1D and Figure S1B) and immunofluorescence microscopy (Fig. 1E) experiments revealed a predominantly nuclear localization, indicating that PRDM6 functions in the nucleus in human NES cells.

PRDM6 represses chromatin accessibility

Given the potential roles of other PRDM proteins in chromatin modification and gene expression[6], we set out to elucidate the impact of human PRDM6 expression on chromatin accessibility. To this end, we performed a genome-wide analysis of chromatin accessibility by *Assay for Transposase-Accessible Chromatin using sequencing* (ATAC-seq) of PRDM6 and EV NES cells, respectively (Fig. 2). First, we performed a qualitative occupancy analysis in PRDM6 and EV NES cells ($n = 3$ biological replicates per group) to elucidate chromatin accessibility regions common or unique to each condition. This analysis revealed 11,076 open chromatin regions in PRDM6 NES cells, of which 10,844 (97.9%) were shared with EV NES cells (Fig. 2A). In contrast, EV NES cells showed 35,273 open chromatin regions, of which 24,429 (69.26%) were absent in PRDM6 NES cells.

Next, we performed a quantitative analysis of differential chromatin accessibility in PRDM6 *vs.* EV NES cells. In total, 35,441 regions showed differential chromatin accessibility between the conditions (false-discovery rate (*FDR*) ≤0.01). Strikingly, only 16 genomic regions (0.045%) were found to be more accessible, whereas 35,425 regions (99.95%) showed reduced chromatin accessibility upon PRDM6 expression, respectively. When considering only regions with ≥2-fold changes in chromatin accessibility, 24,298 regions showed reduced accessibility and only 9 regions showed increased accessibility in PRDM6 NES cells (Fig. 2B). The very few genomic regions that gained accessibility upon PRDM6 expression showed only small increases in accessibility (Fig. 2B and C).

Analysis of the genomic distribution of the 35,441 regions with differential chromatin accessibility between PRDM6 and EV NES cells revealed regions with altered accessibility not only in 5'-UTRs and promoters, but also in intronic and intergenic regions, and in other genome ontologies (Fig. 2D, *left*). Differential accessibility regions were most enriched in 5'-UTRs, promoters, exons, miRNAs, ncRNAs, and pseudogenes (Fig. 2D, *right*), indicating a predominant role of PRDM6 in those regions. However, the presence of differential accessibility regions in intronic and intergenic regions suggests that PRDM6 may also repress chromatin accessibility of enhancer regions located within intronic and intergenic regions.

Notably, gene ontology analysis of all regions with differential chromatin accessibility caused by PRDM6 overexpression in NES cells revealed a strong association with medulloblastoma (Table S1), suggesting that PRDM6 expression may prime NES cells for medulloblastoma development by affecting chromatin accessibility.

PRDM6 expression deregulates gene expression in NES cells

As indicated above, regions with differential chromatin accessibility induced by PRDM6 expression were most strongly enriched in promoters and 5'-UTRs (Fig. 2D, *right*). This prompted us to investigate the impact of PRDM6 on gene expression. Comparative RNA-seq analysis of PRDM6 and EV NES cells ($n = 3$ biological replicates per group) revealed that the expression of a total of 1,588 genes was altered ≥2-fold (adjusted *P*-value < 0.01) (Fig. 3A). Of these, 1,039 (65.43%) genes were upregulated and 549 genes (34.57%) were downregulated in PRDM6 NES cells. Gene ontology analysis of genes with ≥2-fold upregulation indicated a significant association with processes related to extracellular matrix (ECM) organization, ECM glycoproteins, collagens, and proteoglycans (Fig. 3B, *top*), whereas genes that showed ≥2-fold downregulation were most significantly associated with processes related to neuronal and synaptic processes (Fig. 3B, *bottom*).
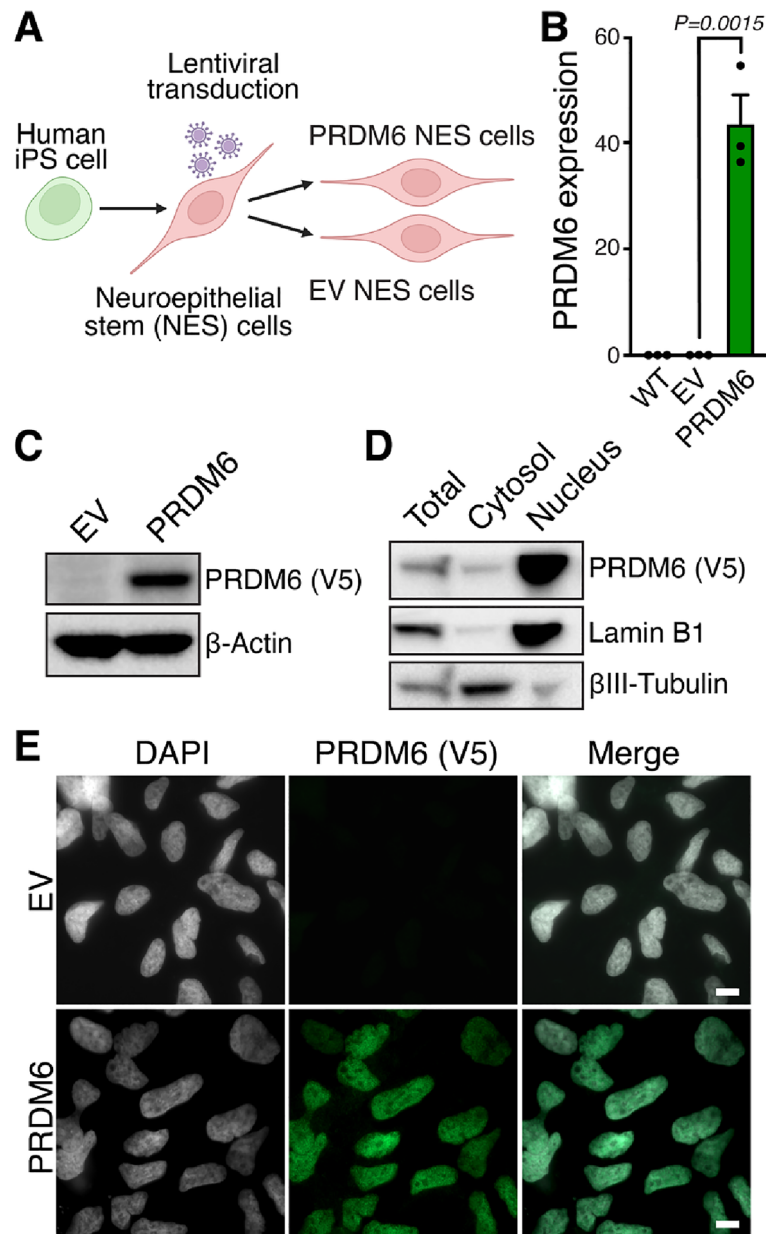
Figure 1. Expression and subcellular localization of PRDM6 in human neuroepithelial stem cells. (**A**) Schematic showing the generation of PRDM6-expressing NES cells from human iPS cells. (**B**) Wild-type (WT) NES cells do not express PRDM6. qRT-PCR analysis of *PRDM6* expression in wild-type, empty vector (EV), and PRDM6-expressing (PRDM6) NES cells. Expression levels are fold changes relative to *GAPDH*. Error bars denote SEM. *P*-values were determined by an unpaired, two-tailed *t*-test. (**C**) PRDM6 protein levels in EV and PRDM6 NES cells. Equal amounts of whole cell extract from empty vector (EV) or PRDM6-V5-transduced (PRDM6) NES cells were probed with anti-V5 antibodies to assess PRDM6 expression and antibodies to β-Actin as a loading control. (**D**) Subcellular fractionation of PRMD6 NES cells. Equal amounts of whole cell extract (total), cytosol, and nuclear extract from NES cells expressing V5-epitope-tagged, human PRDM6 were probed with antibodies against V5 and the indicated proteins of known subcellular localization. (**E**) Representative confocal microscopy images of anti-V5-immunostained EV and PRDM6 NES cells. Nuclei were counterstained with DAPI. Scale bars, 10 μm.

PRDM6 affects transcription factor binding in open chromatin regions in NES cells

Chromatin accessibility modulates transcription by impacting the physical access of transcription factors and other regulatory factors, thus affecting cellular function and identity. To assess the potential mechanistic implications of the observed changes in gene expression and chromatin accessibility between PRDM6 and EV NES cells, we analyzed all genomic regions with differential accessibility for known transcription factor binding motifs via
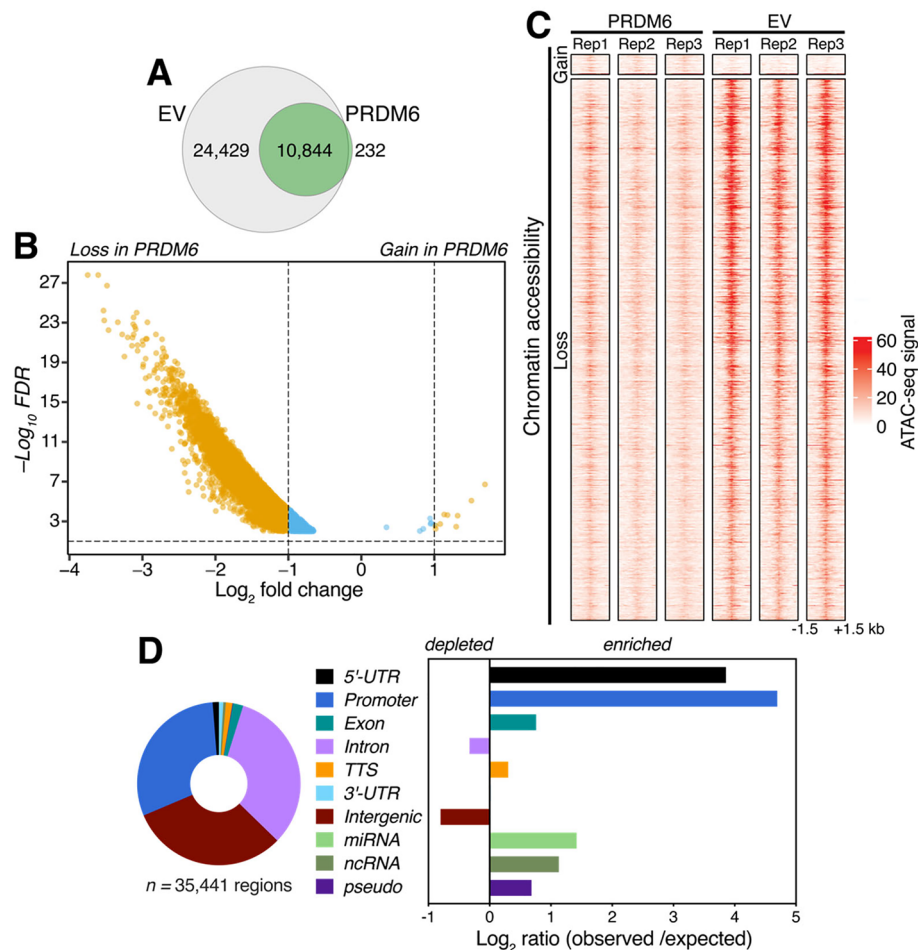
Figure 2. PRDM6 represses chromatin accessibility in neuroepithelial stem cells. (**A**) Venn diagram showing the numbers and overlap of regions with accessible chromatin in EV (gray) and PRDM6 (green) NES cells. (**B**) Volcano plot of results from a quantitative analysis of differential chromatin accessibility in PRDM6 *vs.* EV NES cells (*FDR < 0.01*). Regions with a *Log2* fold change ≥1 in chromatin accessibility are highlighted in yellow. Blue dots correspond to regions with significant (*FDR < 0.01*) but < 1 *Log2* fold change. (**C**) Heatmaps showing ATAC-seq signal within 1.5-kb of all sites with differential chromatin accessibility (*n* = 35,441 sites, *FDR < 0.01*) across three replicate experiments (Rep1-3) in PRDM6 and EV NES cells. (**D**) *left,* Distribution of regions with differential accessibility (*FDR < 0.01*) across the indicated genomic annotations. *Right*, distribution of differential accessibility regions in PRDM6 NES relative to the expected genomic distribution.

HOMER. This analysis revealed significant enrichment of binding motifs for transcription factors known to be involved in the regulation of genome organization, development, cell fate, and tumor suppression (Table S2).

To further interrogate the implications of the PRDM6-driven, altered chromatin accessibility landscape in NES cells, we performed footprinting of transcription factor binding in all open chromatin regions via *Transcription factor Occupancy prediction By Investigation of ATACseq Signal* (*TOBIAS*)[15]. Whereas *HOMER*-based motif enrichment analysis can reveal enriched transcription factor binding motifs within genomic regions, *TOBIAS* quantifies the occupancy of transcription factor motifs and infers which transcription factor is bound according to the sequence of the actual footprint. Footprinting analysis of all open chromatin regions revealed 841 transcription factor binding sites (Fig. 4 and Table S3), of which 73 showed a differential binding score of ≥0.1 (*P < 0.0001*) in PRDM6 *vs.* EV NES cells. Among these, 15 transcription factors (20.55%) showed more binding in PRDM6 NES cells, whereas 58 transcription factors (79.45%) showed less binding in PRDM6 NES cells (Fig. 4A and Table S3). The binding of RFX and SOX family transcription factors was most strongly impaired by PRDM6 expression in NES cells (Fig. 4B). Notably, RFX transcription factors, including RFX1[16], have been implicated in tumor suppression[17], suggesting that reduced binding of these transcription factors in the presence of PRDM6 may have oncogenic potential in NES cells. Moreover, the binding of other transcription factors with roles in genome organization, development, and cell fate was affected (Table S3).

To assess the potential implications of altered transcription factor binding in PRDM6 NES cells (Fig. 4A and Table S3), we queried the *TRRUST v2* reference database[18] to identify the target genes of transcription factors with a differential binding score of ≥ 0.1 (*P < 0.0001*) in PRDM6 *vs.* EV NES cells. This yielded 910 genes, many of which are targeted by more than one transcription factor (Table S4). Filtering the list of 910 target genes to
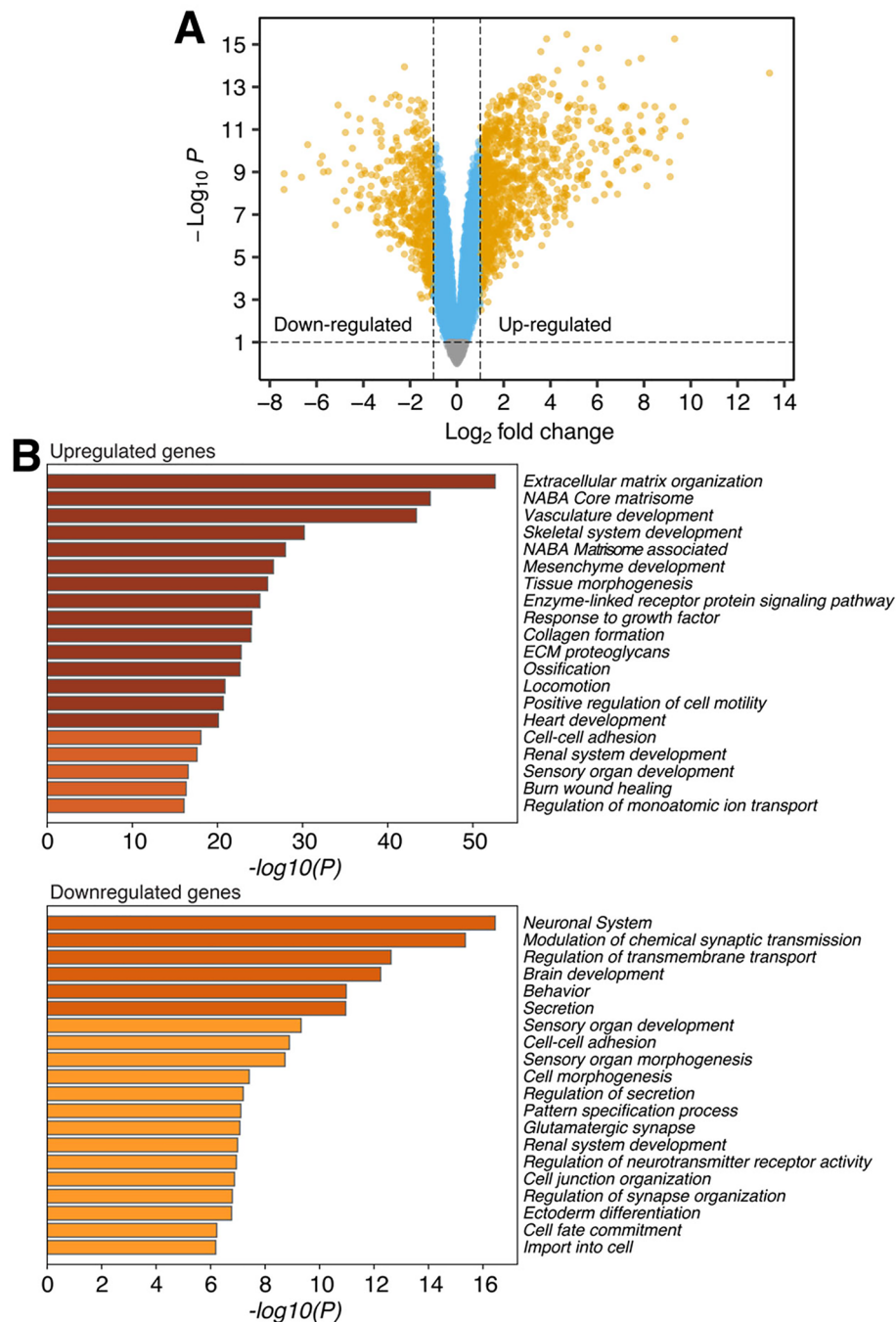
Figure 3. PRDM6 expression causes widespread gene expression changes in neuroepithelial stem cells. (**A**) Volcano plot of gene expression changes in PRDM6 *vs.* EV NES cells (*n* = 3 biological replicates). Dots correspond to individual genes. Genes with significant (adjusted *P*-value < 0.01) up- or downregulation (*Log2* fold change ≥1) are highlighted in yellow; genes with significant *Log2* fold changes of < 1 are in blue. Gray dots indicate genes without significant changes (adjusted *P*-value > 0.01). (**B**) Gene ontology (GO) analysis of upregulated (*top*) and downregulated (*bottom*) genes in PRDM6 NES cells. The bars correspond to significantly enriched GO terms and are colored by *P*-values in *log* base 10.

keep unique target genes yielded a list of 592 target genes. Notably, the transcription factor's impact is unclear for many of these target genes, as both "activation" and "repression" have been reported in the literature (Table S4). Therefore, deducing the implications of altered transcription factor binding on gene expression is not straightforward. Consequently, we queried the RNA-seq data from EV and PRDM6 NES cells to determine how many of these putative target genes are differentially expressed. We found that 122 of the 592 predicted target genes showed a ≥ 1.5-fold change in expression (*P < 0.05*) in our RNA-seq data (Figure S2A). Next, we performed gene set enrichment analysis[19,20] to compute overlaps between the set of 122 genes and all 10,461 Gene Ontology gene
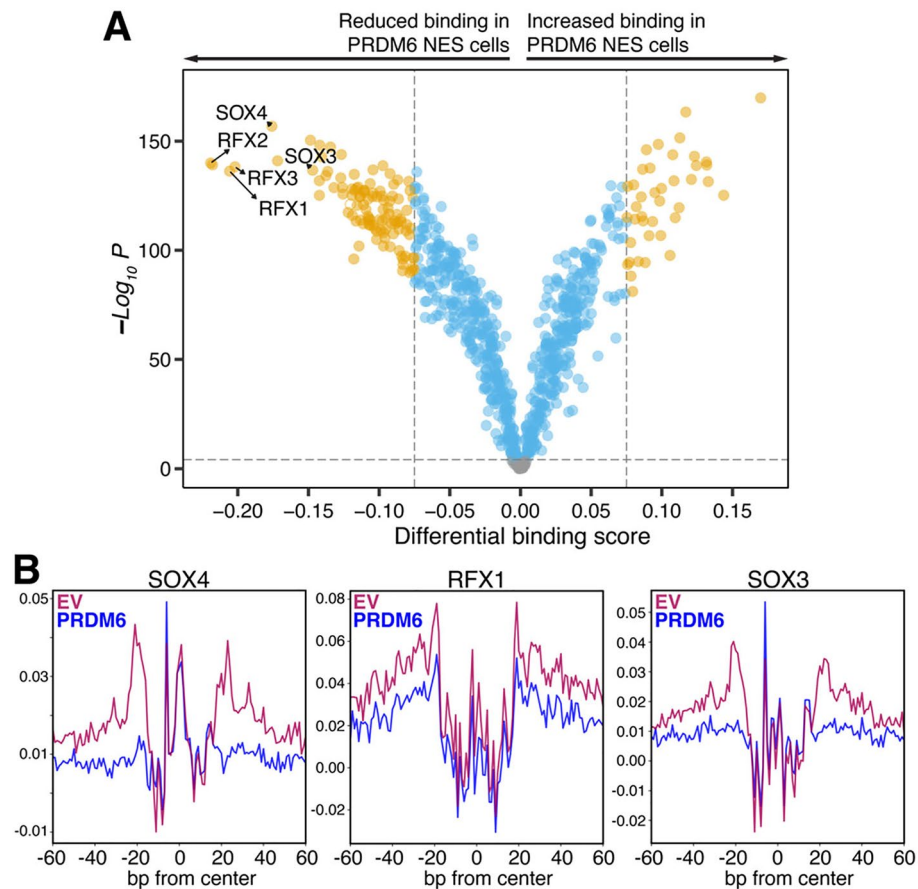
Figure 4. Impact of PRDM6 on transcription factor binding in open chromatin regions. (**A**) Volcano plot showing the quantification of differential transcription factor binding based on digital footprinting analysis of *JASPAR* transcription factor motifs ($n = 841$) in open chromatin regions of PRDM6 and EV NES cells. Transcription factor motifs with significantly differential binding scores ($P < 0.0001$) are indicated in blue; transcription factor motifs with high (≥ 0.075) changes in differential binding score ($P < 0.0001$) are highlighted in orange. Gray dots represent transcription factor motifs without significantly altered binding scores ($P \geq 0.0001$). Footprinting was performed in $n = 3$ biological replicates per group. (**B**) Aggregate plots of footprinting signal across all SOX4, RFX1, and SOX3 motifs in EV (purple) and PRDM6 (blue) NES cells.

sets in the Molecular Signatures Database (*Human MSigDB* version *v2023.2.Hs*). Results from this analysis suggest that PRDM6 expression may affect aspects of programmed cell death, cell motility, and cell differentiation via altered transcription factor binding (Figure S2B).

Together, these findings support the notion that PRDM6 affects the developmental and oncogenic potential of NES cells by altering transcription factor binding.

Genome-wide mapping of PRDM6 binding reveals colocalization with genes in H3K27me3-rich regions

To further elucidate the molecular function of PRDM6, we used *Cleavage Under Targets & Release Using Nuclease* (CUT&RUN)[21] to map the genome-wide distribution of PRDM6 in NES cells. In the absence of reliable antibodies against endogenous human PRDM6, we used V5 antibodies for CUT&RUN in EV and PRDM6 NES cells. We used the *nf-core* best-practices CUT&RUN analysis pipeline[22,23] to identify peak regions and performed differential peak calling by using EV NES cells as controls. This resulted in the detection of 1,083 peaks in PRDM6 NES cells, which were absent in EV NES cells (Fig. 5B). PRDM6 binding sites in the NES genome were enriched in 5'-UTRs, promoters, and other genomic features but were depleted in intergenic regions (Fig. 5A). 952 of 1083 PRDM6 peaks (87.90%) overlapped or were located within two kb of an annotated gene, which implicates PRDM6 in gene regulation (Table S5).

705 of 1083 (65.10%) PRDM6-bound genomic regions overlapped with regions showing significantly reduced chromatin accessibility in PRDM6 NES cells (Fig. 5B), suggesting a predominantly repressive effect of PRDM6 binding to chromatin. However, given the relatively small number of PRDM6 binding sites detected by CUT&RUN, most regions with differential accessibility did not show PRDM6 binding. There are at least two, not mutually exclusive, interpretations of this finding: first, most of the effects of PRDM6 on chromatin accessibility occur through indirect mechanisms; second, binding of PRDM6 to chromatin is indirect or very
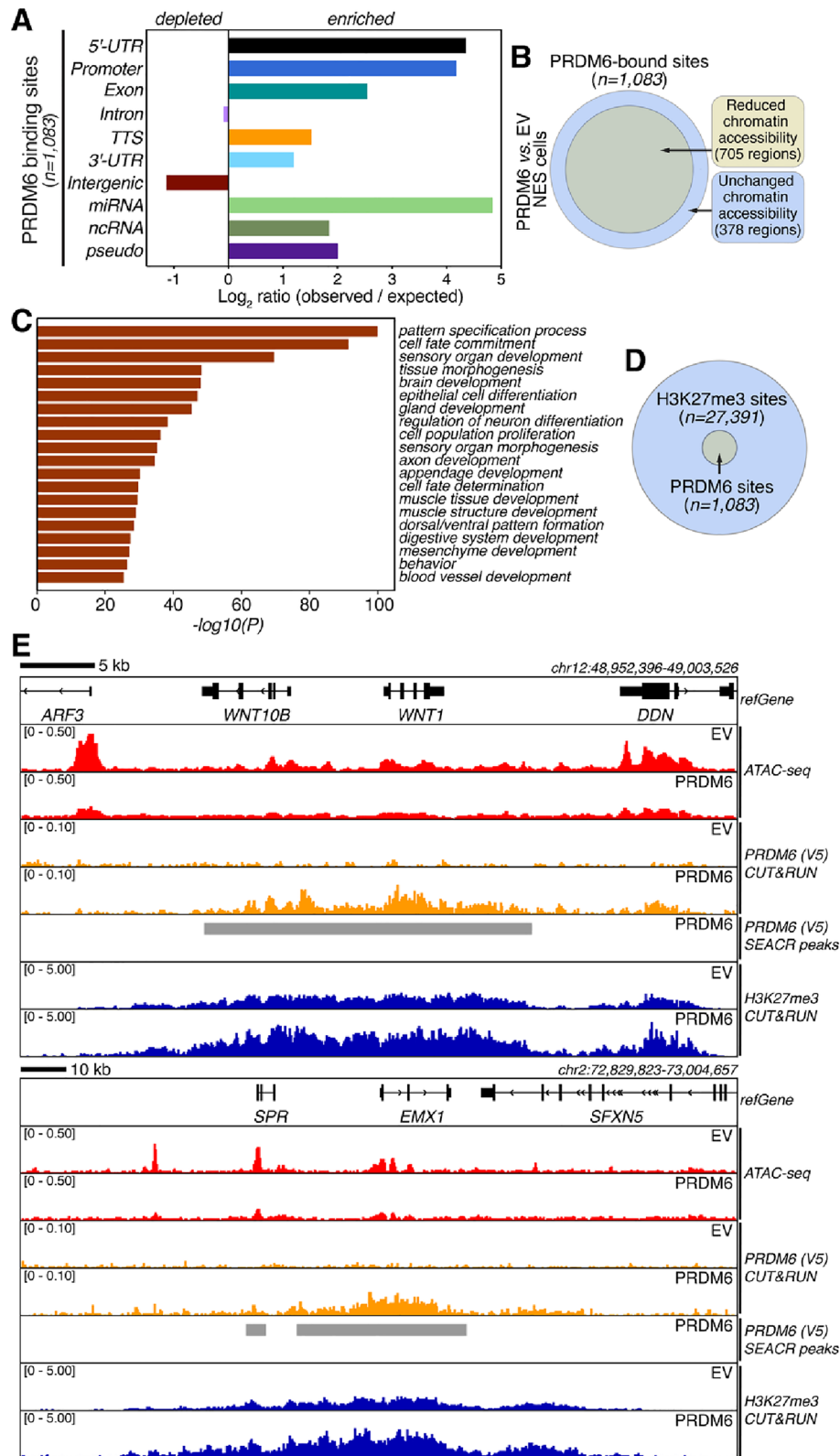
Figure 5. Elucidation of genome-wide PRDM6 binding in neuroepithelial stem cells. (**A**) Genomic distribution of PRDM6 binding sites determined by CUT&RUN. (**B**) Venn diagram showing the overlap of PRDM6-bound sites with regions showing altered chromatin accessibility (*FDR* < 0.01). (**C**) Gene ontology (GO) analysis of genes overlapping with or located within two kb of a PRDM6 binding site. Bars are colored by *P*-values in *log* base 10 and indicate significantly enriched GO terms. (**D**) Venn diagram showing the overlap of PRDM6-bound sites with H3K27me3-marked regions identified by CUT&RUN in PRDM6 NES cells. (**E**) Examples showing PRDM6 enrichment in relation to H3K27me3-marked regions in EV and PRDM6 NES cells, respectively. Bins per million (BPM)-normalized CUT&RUN and ATAC-seq signals are shown. SEACR peak calling results for PRDM6 CUT&RUN signal are indicated by gray bars. Genomic region coordinates and *refGene* annotations are shown.

transient and hence difficult to detect by CUT&RUN, which results in underestimation of the true number of PRDM6 binding sites in the genome.

Gene ontology analysis of genes overlapping with or located within two kilobases (kb) of a PRDM6 binding site revealed a strong enrichment of terms related to developmental and neural processes (Fig. 5C). Further, *Molecular Signatures Database* (MSigDB)[24] analysis, suggested a significant association with processes related to histone H3 lysine 27 trimethylation (H3K27me3) (Table S6). To follow up on this latter observation, we determined the H3K27me3 landscape in PRDM6 NES cells via CUT&RUN. Bioinformatic analysis revealed 27,391 H3K27me3 binding sites. Notably, all 1,083 PRDM6 binding sites overlapped H3K27me3 binding sites (Fig. 5D), suggesting a functional interplay between H3K27me3-related processes and PRDM6. In the latter regard, the overlap of PRDM6 binding sites with H3K27me3 regions (Fig. 5D-E) is consistent with the predominantly repressive effect of PRDM6 binding reported above. We did not detect changes in global H3K27me3 levels in PRDM6 *vs.* EV NES cells (Figures S3A and S1C). This is perhaps not surprising given that we detected PRDM6 binding in only approximately 3.95% of H3K27me3 regions (Fig. 5D). Notably, PRDM6 binding sites showed significantly higher H3K27me3 signal in PRDM6 NES cells than in EV NES cells (Figure S3B-C). Together, the finding that PRDM6 binding completely overlaps with chromatin marked by H3K27me3, which affects spatiotemporal control of gene expression in stem cells, further implicates PRDM6 as a modulator of chromatin structure and gene expression in NES cells.

PRDM6 expression in NES cells promotes medulloblastoma formation

Having revealed the molecular impact of PRDM6 expression in NES cells, we asked whether the observed changes in chromatin landscape and gene expression establish a cellular environment prone to tumorigenesis. To test the tumorigenic potential of PRDM6 in NES cells, we performed in vitro proliferation assays and orthotopic intracranial implantation experiments in mice (Fig. 6A). In vitro, PRDM6 NES cells showed no change in proliferation rate when compared to NES cells without PRDM6 expression (Figure S4A). Next, we orthotopically implanted 3 $\times 10^5$ cells into the cerebellum of adult, female NSG mice and monitored tumor growth for a year. About a third (8/30) of mice implanted with PRDM6-expressing NES cells developed tumors with a latency of 158 ± 16 days (mean ± SEM) after transplantation (Fig. 6B). In contrast, implantation of NES cells without PRDM6 expression (EV) did not cause tumor formation (0/15) within a year of implantation, the endpoint in our study. These findings indicate that PRDM6 expression can indeed establish a pro-tumorigenic milieu.

Next, we performed a histopathological analysis of tumors derived from PRDM6 NES cells. Microscopic examination of hematoxylin and eosin (H&E)-stained tumor tissue revealed a highly cellular neoplasm with mitotic activity and apoptotic bodies (Fig. 6C, *left panel,* Figure S5). Tumors lacked glial components and showed an overall morphology reminiscent of primitive neuroectodermal tumors, which is consistent with human medulloblastoma histopathology (Fig. 6C**,** *left panel*). Increased mitotic activity was further supported by strongly positive Ki67 immunostaining (Fig. 6C**,** *right panel*). Moreover, immunostaining for PRDM6 via detection of its V5 epitope tag showed that NES-derived tumors had maintained PRDM6 expression, suggesting that PRDM6 may function in both tumor initiation and maintenance (Fig. 6C**,** *middle panel*).

Because a subset of Group 4 medulloblastomas shows high expression of PRDM6 in conjunction with amplification of the *MYCN* oncogene[1,4], we decided to investigate potential synergies between PRDM6 and MYCN expression on tumor development from NES cells. Notably, MYCN expression alone is sufficient to drive tumor formation from NES cells, resulting in SHH medulloblastoma[11]. We generated human NES cells with either MYCN overexpression ("MYCN" cells) or combined MYCN and PRDM6 overexpression ("MYCN-PRDM6" cells), respectively (Figure S4B-C). MYCN-overexpressing NES cells showed significantly faster proliferation in vitro, irrespective of PRDM6 expression (Figure S4D).

To evaluate the impact of combined PRDM6 and MYCN overexpression on tumor formation, we performed orthotopic transplantation experiments in mice. Mice implanted with MYCN- or MYCN-PRDM6 NES cells showed 100% penetrance of tumor growth in both groups (Figure S4E). Latencies of tumor development from MYCN- *vs.* MYCN-PRDM6 NES cells were not different (67 ± 7 days and 68 ± 8 days (median ± SEM), respectively; Figure S4E). Notably, these latencies were significantly shorter than those observed after the transplantation of PRDM6 NES cells (Fig. 6D). Overall, these findings suggest that PRDM6 does not modulate the tumorigenic effects of MYCN overexpression in NES cells.

To molecularly characterize tumors arising from PRDM6- or MYCN-PRDM6 NES cells, we performed RNA-seq analysis. We then carried out a comparative gene expression clustering analysis with medulloblastomas and other pediatric central nervous system (CNS) tumors (*i.e.*, "non-medulloblastoma" tumors) from human patients (Fig. 6E). For medulloblastoma subgroup clustering, we used a classifier trained on 1,112 medulloblastomas (WNT, SHH, Group 3, Group 4) and 1,473 primary, non-medulloblastoma tumors corresponding to 10 distinct CNS tumor types (Table S7). Tumors derived from PRDM6 NES cells clustered most closely with Group 3 medulloblastomas and showed no similarity with pediatric non-medulloblastoma CNS tumors (Fig. 6E). In contrast, tumors derived from MYCN NES cells clustered most closely with SHH medulloblastomas, consistent with previous findings (Fig. 6E)[11]. Tumors derived from MYCN-PRDM6 NES cells clustered with SHH medulloblastomas as well, suggesting that MYCN has a dominant effect on tumor phenotype (Fig. 6E). These latter findings suggest that PRDM6 expression does not modulate the onset, progression, or molecular phenotypes of tumors derived from MYCN-overexpressing NES cells.

Together, our findings show that PRDM6 expression in NES cells has oncogenic potential and promotes medulloblastoma formation. However, we conclude that additional factors are required for Group 4 medulloblastoma formation as PRDM6 NES cell-derived tumors most closely resemble Group 3 medulloblastoma.
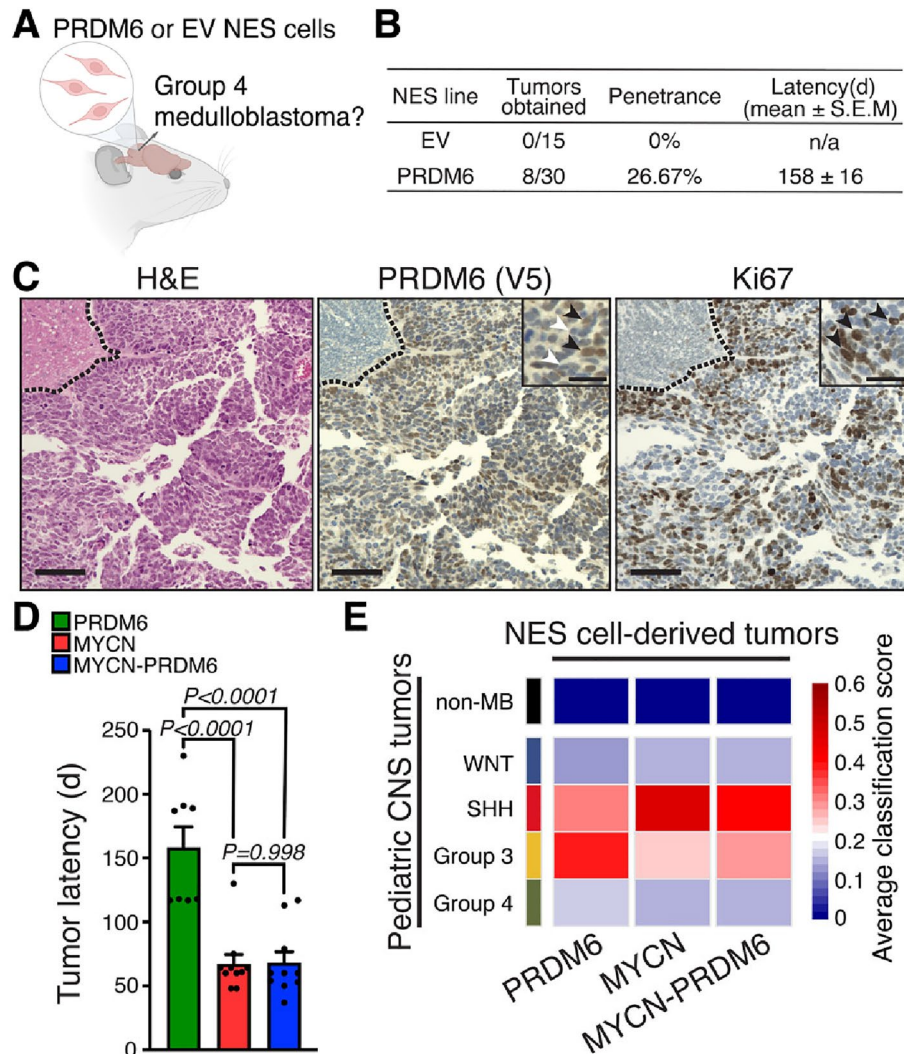
Figure 6. PRDM6 expression in neuroepithelial stem cells induces medulloblastomas in mice. (**A**) Schematic of orthotopic human NES transplantation into the mouse cerebellum. (**B**) Outcomes of PRDM6 or EV NES cell transplantations in mice. d, days post-injection. (**C**) H&E staining of representative PRDM6 NES-derived tumor (*left*). Immunohistochemistry staining of PRDM6 via the V5 epitope tag (*middle*) and Ki67 (*right*) in PRDM6 NES-derived tumor. Dashed lines denote tumor margins. Main scale bars correspond to 100 μm; scale bars in insets correspond to 25 μm; black arrowheads in the middle panel inset denote anti-V5-reactive (PRDM6-expressing) nuclei; white arrowheads indicate PRDM6-negative nuclei; black arrowheads in the right panel inset denote Ki67-positive, proliferating cells. (**D**) Latency of tumor development after implantation of PRDM6-, MYCN- or MYCN-PRDM6 NES cells into the cerebellum of mice. Error bars represent SEM. *P*-values were determined by one-way ANOVA with Tukey's post hoc correction. (**E**) Clustering of tumors derived from PRDM6-, MYCN-, or MYCN-PRDM6 NES cells with human medulloblastomas and pediatric, non-medulloblastoma (non-MB) CNS tumors.

Discussion

Molecular function of PRDM6

The PRDM family is implicated in mammalian neurodevelopment and control of cell proliferation in cancer[25] but the role of PRDM6 in these processes is unclear. We found that human PRDM6 functioned in the nucleus, exerting a strongly repressive effect on chromatin accessibility in hindbrain neuroepithelial stem (NES) cells. Moreover, we found a complex dysregulation of gene expression patterns in PRDM6 NES cells, with substantial downregulation and upregulation of genes. These findings are consistent with the notion that human PRDM6 exists in protein complexes containing both transcriptional activators and repressors. Such binary roles have been shown for PRDM1, another PRDM family member, which recruits the histone demethylase KDM4a to remove repressive H3K9me3 marks[26] and also interacts with histone deacetylases to repress transcription[27]. Moreover, mouse PRISM/PRDM6 has been implicated in both transcriptional activation and repression[8], consistent with our findings in human NES cells. It will be important to elucidate the composition of human PRDM6-containing protein complexes to fully reveal the mechanisms by which PRDM6 affects chromatin structure and gene

expression. In this regard, we found that PRDM6 is bound to H3K27me3-rich regions and that these regions show higher levels of H3K27me3 in PRDM6-expressing NES cells. H3K27me3 affects the spatiotemporal control of gene expression in complex ways, so it will be important to interrogate interplays between H3K27 methylation and PRDM6 in NES cells and medulloblastoma.

Oncogenic potential of PRDM6

Sequencing of medulloblastomas has revealed many molecular aberrations. A key challenge is to distinguish "driver" from "passenger" alterations. Our findings that PRDM6 expression in hindbrain neuroepithelial stem cells promotes medulloblastoma indicate that PRDM6 can indeed function as a medulloblastoma driver. PRDM6, however, appears to have a much lower oncogenic potential than MYCN. In this regard, MYCN-overexpressing NES cells caused much more aggressive tumors at a much higher penetrance than PRDM6-expressing NES cells. Moreover, in the context of combined MYCN and PRMD6 overexpression, we did not observe a modulation of tumor onset or progression in comparison to MYCN overexpressing NES cells (Fig. 6D and Figure S2E). This suggests that MYCN expression is dominant regarding tumorigenesis.

Several, not mutually exclusive reasons may underlie why PRDM6 NES cell-derived medulloblastomas do not recapitulate Group 4 medulloblastomas. Additional factors such as a specific cell-of-origin, co-occurring genomic alterations, or the specific microenvironment of the developing cerebellum may contribute. Correspondingly, it will be important to test the oncogenic impact of PRDM6 expression in other cerebellar progenitors, *e.g.,* in unipolar brush cells[1,28], ideally in the immature embryonic hindbrain[1,2], or in the context of other genomic alterations that occur in PRDM6-expressing Group 4 medulloblastomas[4].

Our observations may provide clues as to how PRDM6 expression promotes medulloblastoma formation from NES cells. Our footprinting analysis shows significantly altered binding of transcription factors to open chromatin regions in PRDM6 NES cells. These include RFX1, a transcriptional repressor of tumor suppressors, and several other RFX transcription factors[16]. Notably, multiple RFX1 target genes, including the oncogene *FGF1*, *COL1A1*, and *COL1A2*, are upregulated in PRDM6 NES cells. Moreover, RFX1 dimerizes with RFX2 and RFX3[29], which also show reduced binding in PRDM6 NES cells (Fig. 4). These findings may provide a starting point for how PRDM6 establishes a pro-oncogenic environment. In the latter context, we find that PRDM6 expression results in substantial upregulation of processes related to extracellular matrix (ECM) biology (Fig. 3B). The ECM is important for cell adhesion, growth factor binding, and intracellular signaling. Importantly, tumors use ECM remodeling to promote tumorigenesis and metastasis[30]. In future studies, it will be important to address the functional implications of PRDM6-driven ECM alterations as it may yield insights into the tumorigenic effects of PRDM6.

Given the oncogenic potential of PRDM6 we report here, its inhibition may be of benefit in PRDM6-expressing Group 4 medulloblastomas. PRDM6 expression is shut off postnatally[25,31], which would make it an attractive drug target as PRDM6-selective inhibition should not affect the homeostasis of normal tissues.

Methods
Cell culture

Human neuroepithelial stem (NES) cells were generated from established human WTC10 iPS cells[32] (gift from Bruce R. Conklin, Gladstone Institutes, UCSF) as previously described[11,12]. Experimental work involving iPS cells was approved by the UCSF Stem Cell Research Oversight Committee. All experiments were performed in accordance with all relevant guidelines and regulations using protocols approved by the UCSF Institutional Review Board. All cell lines used here were regularly screened for mycoplasm and were mycoplasm-negative by PCR-based testing. NES cells were cultured on poly-L-ornithine/laminin-coated wells in NES cell medium (DMEM/F-12 with GlutaMAX, $1 \times$ N2 supplement, $0.05 \times$ B27 without vitamin A supplement, 1.6 g/L Glucose, 20 ng/mL EGF, 20 ng/mL FGF2) in a humidified 37 °C incubator with 5% CO_2. NES cells were fed daily, passaged every three to four days using TrypLE Express, and plated at $5–6 \times 10^5$ cells per well of a 6-well plate in 2 mL of NES cell medium. For lentiviral transductions, NES cells were plated on poly-L-ornithine/laminin-coated wells, and PEG-concentrated lentiviruses were added to the cells. Lentiviruses were washed out 24 h after transduction, cells were expanded, and subjected to either antibiotic selection starting 72 h after transduction or sorted by flow cytometry.

Mice

All studies followed ethical regulations established by the Institutional Animal Care and Use Committee (IACUC) and Institutional Biosafety Committee (IBC) at the University of California, San Francisco (UCSF). All methods were performed in accordance with relevant guidelines, regulations, and protocols approved by the UCSF IACUC and IBC. The reporting in this manuscript follows the ARRIVE guidelines[33]. 8–12-week-old female NSG (*RRID:IMSR_JAX:005557*) mice were used for transplantation experiments. For transplantation, 3×10^5 cells in 3 μL of NES cell medium were implanted per mouse. Implantations were performed using a stereotactic frame, starting from lambda 2 mm right, 2 mm down, and 2 mm deep. Mice were euthanized at endpoint, which was either signs of tumor growth (*e.g.,* hunched back, weight loss, head tilt, etc.) or one year after transplantation.

Lentiviral constructs

The coding sequence of human PRDM6 (*NCBI* Reference Sequence NM_001136239.4) was fused to a C-terminal V5-tag separated by a 2× glycine linker and inserted into a lentiviral vector to generate pCDH-Lenti-CAGGS-hPRDM6-2×Gly-V5-EF1-Luc2-Blast (Addgene plasmid #213,826; RRID:Addgene_213826). The lentiviral MYCN expression plasmid pCDH-CAG3×FLAG-MYCN-mScarlet-Luciferase was previously published[11].

Protein extraction and immunoblotting

Protein extracts were prepared using RIPA buffer (150 mM NaCl, 1% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50 mM Tris, pH 8.0; Millipore Sigma) supplemented with protease inhibitors (*cOmplete* protease inhibitor cocktail; Roche) and quantified by BCA assay. Proteins were separated through a 4–12% (w/v) Bis–Tris SDS poly-acrylamide gel and transferred onto nitrocellulose membranes. Blots were incubated with primary antibodies in 5% (w/v) BSA in TBS-T at 4 °C overnight, washed in TBS-T, and incubated with secondary antibodies in 5% (w/v) BSA in TBS-T for 2 h at room temperature. Membranes were washed in TBS-T and developed using ECL chemiluminescence reagent.

Subcellular fractionation

Whole-cell, nuclear, and cytosolic extracts from NES cells were prepared by using the Nuclear Extract Kit (Active Motif) as per the manufacturer's instructions. In brief, cells were scraped into ice-cold PBS supplemented with protease inhibitors (*cOmplete* protease inhibitor cocktail, Roche) and phosphatase inhibitors, sedimented by centrifugation, and resuspended in hypotonic lysis buffer (Nuclear Extract Kit, Active Motif). Cytosolic and nuclear fraction were separated by centrifugation (14,000 × g for 10 min at 4 °C). Nuclei were extracted in Complete Lysis Buffer (Active Motif AM1 Lysis buffer containing 1 mM DTT and protease inhibitors). Whole-cell fractions were prepared by direct extraction of cells in Complete Lysis Buffer. Subcellular fractions were normalized by volume and analyzed by immunoblotting.

Histone extraction

Histones were extracted by H_2SO_4 extraction as previously described[34]. Briefly, cell pellets were lysed in hypotonic lysis buffer (10 mM Tris–HCl, pH 8.0, 1 mM KCl, 1.5 mM $MgCl_2$, 0.5% (v/v) NP40, 1 mM DTT, *cOmplete* protease inhibitor cocktail). Nuclei were sedimented at 600 × g for 5 min at 4 °C. The nuclear pellet was incubated in 0.4 N H_2SO_4 for 16–18 h at 4 °C. Histones were precipitated with trichloroacetic acid, washed with acidified acetone, resuspended in water, and processed for immunoblotting.

Cell proliferation assays

In vitro proliferation assays of NES cells were performed using either *Cell Titer Glo* (Promega) or *CyQUANT NF* (Thermo Fisher) according to the manufacturer's instructions. In brief, 5×10^3 NES cells were plated in 100 μL of NES media in triplicates for each condition. Proliferation was measured at timepoint day 0 (d0, *i.e.*, 2 h after cell plating) and every 24 h thereafter, for seven days. Levels of proliferation are represented as fold changes over d0 values.

Immunohistochemistry

Formalin-fixed, paraffin-embedded tissue was cut into 5-μm thick sections, deparaffinized with xylene substitute, and rehydrated with decreasing concentrations of ethanol in water. For antigen retrieval, tissue slides were incubated in sodium citrate buffer (pH 6.0), followed by quenching in 0.3% (v/v) hydrogen peroxide. Tissue was permeabilized in blocking buffer (0.4% (v/v) Triton-X-100, 10% (v/v) normal goat serum in phosphate-buffered saline (PBS)). Primary antibodies were applied for 16–18 h at 4 °C in a humidified chamber. Sections were incubated with biotinylated secondary antibody for 1 h at room temperature in a humidified chamber. For signal amplification, slides were incubated with Vectastain ABC reagent and color development was achieved by diaminobenzidine tetrahydrochloride solution. Sections were counterstained with hematoxylin, dehydrated through ethanol and xylene substitute, and cover-slipped using mounting medium. Stained sections were imaged using a digital microscope (Keyence BZ-X) and BZ-X Keyence software at 40× and 100× magnifications.

Immunofluorescence microscopy

NES cells cultured on poly-L-ornithine/laminin-coated coverslips were fixed in 4% (w/v) paraformaldehyde in PBS for 15 min at room temperature. Quenching was performed in 50 mM NH_4Cl in PBS, followed by permeabilization in 0.5% (v/v) Triton-X-100 in PBS. Coverslips were incubated in blocking buffer (5% (v/v) goat serum, 0.1% (v/v) Tween in PBS) for 1 h at room temperature, followed by a 16–18-h incubation at 4 °C with V5 primary antibody (Cell Signaling Technology) diluted 1:250 in blocking buffer. Coverslips were washed in PBS and incubated in blocking buffer containing fluorescent secondary antibodies for 1 h at room temperature. Coverslips were washed in PBS and mounted with Fluoromount G/DAPI. Imaging was performed using a Zeiss Spinning Disk confocal microscope.

ATAC-seq

Preparation of ATAC-seq libraries and sequencing was performed as described[35], with slight modifications. In brief 1×10^5 cells per sample were harvested and lysed in ATAC-seq lysis buffer (10 mM Tris–HCl (pH 7.4), 10 mM NaCl, 3 mM $MgCl_2$, 0.1% (w/v) IGEPAL CA-630). The transposition reaction was performed using Tagment DNA TDE1 Enzyme and Buffer Kit (Illumina) and 4.7 μL TDE1 Tagment DNA enzyme per sample in a 50-μL total reaction volume. Zymo DNA Clean and Concentrator Kit was used for cleanup of the transposition reaction. Quantitive PCR and limited-cycle PCR was performed using NEBNext Ultra II master mix (New England Biolabs), SYBR™ green and 0.5 μM of sample-specific P5/P7 barcoded primer mix (New England Biolabs). Bead clean-up and purification was performed using SPRI beads (Beckman Coulter). Agilent TapeStation 4200 was used to check for quality of sequencing libraries. To support standardization and reproducibility, we used the *nf-core* framework best-practice ATAC-seq analysis pipeline (*nf-core/atacseq* v1.2.1) for sequence processing and data analysis[22,36]. In brief, adapters were trimmed by TrimGalore (v0.6.4_dev)[37] and adapter-trimmed

reads were mapped to the *hg38* reference assembly by BWA (v0.7.17)[38]. Duplicates were marked and removed by Picard (v2.23.1)[39]. ATAC-seq peaks were called by MACS2 (v2.2.7.1)[40] and differential analysis of ATAC-seq peaks present in all three biological replicates per group (PRDM6 or EV NES cells) was performed by DiffBind (v3.6.5)[41,42] and DESeq2[43]. Transcription factor occupancy analysis was performed by TOBIAS (v0.15.1)[15]. Correction for Tn5 transposase insertion bias (*ATACorrect*) and footprint scoring (*ScoreBigwig*) were performed as described[15]. Footprint scores were matched to the *JASPAR 2022 CORE vertebrates (non-redundant)* database[44] and differential binding scores were calculated by *BINDetect*[15].

CUT&RUN

Preparation of CUT&RUN libraries was done with the CUT and RUN assay kit (Cell Signaling Technology). Briefly, 2.5×10^5 NES cells per sample were fixed with 0.1% formaldehyde in PBS for 2 min and quenched with 1 M glycine for 5 min at room temperature. Cells were washed in 1× wash buffer containing spermidine and protease inhibitor cocktail, and incubated in antibody binding buffer containing spermidine, proteinase inhibitor cocktail, digitonin, and 10 μl of activated concanavalin A beads per sample for 16–18 h at 4 °C. Per CUT&RUN reaction, 0.2 μg of H3K27me3 antibody (clone C36B11, #9733, Cell Signaling Technology) or 0.96 μg of V5 antibody (clone D3H8Q, #13,202, Cell Signaling Technology) was used. pAG-MNase in 1× wash buffer containing spermidine, protease inhibitor cocktail, and digitonin was bound to the antibody for 1 h at 4 °C. pAG-MNase was activated by adding 3 μL calcium chloride per sample and incubated for 30 min at 4 °C. Digestion was stopped by adding 1× stop buffer (digitonin solution, RNase A, and 50 pg *S. cerevisiae* spike-in DNA (Cell Signaling Technology) for normalization) and samples were incubated for 10 min at 37 °C. Samples were subjected to SDS and proteinase K treatment for 2 h at 65 °C. Zymo DNA Clean and Concentrator (D4034) was used to clean up the samples. NEBNext Ultra II DNA Library Prep Kit for Illumina (E7645L) was used for library preparation. Library clean-up and size selection was done by using SPRI beads (103778-488, Bulldog Bio). Libraries were quantified and analyzed by Qubit and Tapestation 4200 (Agilent) analysis and sequenced for 38 cycles in paired-end mode on a NextSeq 500 (Illumina). For data analysis, we used the nf-core framework CUT&RUN analysis pipeline (v3.0; https://github.com/nf-core/cutandrun/tree/3.0)[22,23]. In brief, adapter-trimmed reads (TrimGalore v0.6.6) were mapped to the human (*hg38*) and spike-in (*S. cerevisiae*, *sacCer3*) genome reference assemblies by Bowtie2 (v2.4.4). Duplicates were marked by Picard (v2.27.4). Reads were normalized against spike-in. CUT&RUN peaks were called by SEACR (v1.3) using stringent settings, and consensus peaks were identified as peaks that were present in both replicates per condition. The proximity of peaks to genes was assessed by T-Gene (v5.5.3)[45].

RNA isolation and quantitative RT-qPCR

RNA was extracted using an RNA extraction kit (Zymo Research). 500 ng of RNA was converted to cDNA using the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher) and the following settings: 25 °C for 10 min, 37 °C for 2 h, and 85 °C for 5 min. qPCR was performed using PowerUp SYBR green (Thermo Fisher) with the following settings: 95 °C for 2 min, 40 cycles [95 °C for 15 s and 60 °C for 1 min]. Gene expression was normalized to *GAPDH*.

RNA-seq

Total RNA was extracted from cultured cells or flash frozen tumor tissue using an RNA extraction kit (Zymo Research). RNA quality was checked using Agilent TapeStation 4200. Samples were processed for library preparation (polyA enrichment, Illumina TruSeq stranded mRNA kit) and sequenced (NovaSeq PE100 or PE150; 62–112 million total reads per sample). Raw RNA-seq data was pre-processed using *HTStream* (https://github.com/s4hts/HTStream). Pre-processed reads were aligned using STAR (v2.7.10, *GRCh38*). Differentially expressed genes (*Log2* fold change ≥1, adjusted *P*-value < 0.01) were identified by *edgeR* and *limma-voom* analysis[46].

Classification of NES-derived tumors

Raw RNA-seq reads were processed to remove low-quality sequences and adapter sequences. To discriminate between human and mouse reads, *disambiguate* was used to filter out reads with a higher mapping quality score in the mouse genome compared to the human genome[47]. Human-specific reads were mapped to *GRCh38/hg38* using STAR to generate *Fragments Per Kilobase of transcript per Million* mapped reads (FPKM) values[48]. To quantify the resemblance of NES-derived tumors with medulloblastoma subgroups, we trained a machine-learning classification model on a large series of primary medulloblastoma tumors. Specifically, two published gene expression microarray series of medulloblastomas, with medulloblastoma subgroup determined by DNA methylation, were obtained (GEO GSE85218 and EGAS00001001953). After removing duplicate samples this resulted in 1112 unique expression profiles that underwent Robust Multi-array Average normalization. Because gene sets are more robust to platform differences than individual genes, single-sample gene set enrichment (ssGSEA) scores were derived by using all *MSigDB* gene sets that contained ≥ 80% mouse–human orthologs and ≥ 15 and ≤ 400 genes[19]. A random forest classification model was then trained using a 70–30 training-test split of the medulloblastoma expression profiles, using the gene set enrichment scores as features to predict medulloblastoma subgroup. Gene sets with a scaled feature importance score > 8 were used to build the final classification model resulting in a total of 409 features (test set accuracy = 0.956). This model was used to predict the subgroup class probabilities for all NES-derived tumors. The same workflow was utilized to create a binary machine learning model to classify if tumors were medulloblastoma. The dataset used to train the non-medulloblastoma model consisted of 1473 primary CNS tumors, after filtering undefined entities, encompassing 10 distinct CNS tumor types obtained from the R2 pediatric genome portal (accession "Tumor Brain (DKFZ-public)- Kool-1678-MAS5.0-u133p2"). Using a 70–30 training-test split, the non-medulloblastoma model achieved a test set accuracy of 0.987.

Motif, genome ontology, and gene ontology analyses

Genome annotations were determined by *HOMER* version 4.11.1[49] using *HOMER hg38* accession and ontology information with '*annotatePeaks.pl* –annStats' and default settings. *HOMER*-based motif analysis used *findMotifsGenome.pl* with default settings. For gene ontology analysis, gene lists were analyzed for pathway and process enrichment using KEGG Pathways, GO Biological Processes, Reactome Gene Sets, Canonical Pathways, CORUM, and WikiPathways via Metascape[50]. All human genes were used as the enrichment background. Significant terms ($P < 0.01$; minimum count = 3; enrichment factor > 1.5 (enrichment factor is the ratio between the observed counts and the counts expected by chance)) were determined and clustered based on membership similarities (*P*-values were calculated based on the cumulative hypergeometric distribution, and *q*-values were calculated using the Benjamini–Hochberg procedure to account for multiple testing). Kappa scores were used as the similarity metric for hierarchical clustering of enriched terms, sub-trees with a similarity of > 0.3 were considered a cluster, and the most statistically significant term within a cluster was chosen to represent the cluster.

Statistical analysis

Statistical analysis was performed and graphs were generated in GraphPad Prism 10.0.1 and R[51]. qRT-PCR data was analyzed using an unpaired, two-tailed *t*-test. In vitro proliferation data was analyzed using an unpaired, two-tailed *t*-test or one-way ANOVA with Tukey's post hoc correction. In vivo tumor latencies were analyzed using one-way ANOVA with Tukey's post hoc correction.

Data availability

All next-generation sequencing datasets generated as part of this study have been deposited to the *NCBI Gene Expression Omnibus* repository under accession number GSE243558.

References

1. Hendrikse, L. D. *et al.* Failure of human rhombic lip differentiation underlies medulloblastoma formation. *Nature* **609**, 1021–1028 (2022).
2. Smith, K. S. *et al.* Unified rhombic lip origins of group 3 and group 4 medulloblastoma. *Nature* **609**, 1012–1020 (2022).
3. Gibson, P. *et al.* Subtypes of medulloblastoma have distinct developmental origins. *Nature* **468**, 1095–1099 (2010).
4. Northcott, P. A. *et al.* The whole-genome landscape of medulloblastoma subtypes. *Nature* **547**, 311–317 (2017).
5. Hovestadt, V. *et al.* Medulloblastomics revisited: Biological and clinical insights from thousands of patients. *Nat. Rev. Cancer* **20**, 42–56 (2020).
6. Di Tullio, F., Schwarz, M., Zorgati, H., Mzoughi, S. & Guccione, E. The duality of PRDM proteins: Epigenetic and structural perspectives. *FEBS J.* **289**, 1256–1275 (2022).
7. Wu, Y. *et al.* PRDM6 is enriched in vascular precursors during development and inhibits endothelial cell proliferation, survival, and differentiation. *J. Mol. Cell. Cardiol.* **44**, 47–58 (2008).
8. Davis, C. A. *et al.* PRISM/PRDM6, a transcriptional repressor that promotes the proliferative gene program in smooth muscle cells. *Mol. Cell. Biol.* **26**, 2626–2636 (2006).
9. Leszczyński, P. *et al.* Emerging roles of PRDM factors in stem cells and neuronal system: Cofactor dependent regulation of PRDM3/16 and FOG1/2 (Novel PRDM Factors). *Cells* **9**, 2603 (2020).
10. Tailor, J. *et al.* Stem cells expanded from the human embryonic hindbrain stably retain regional specification and high neurogenic potency. *J. Neurosci.* **33**, 12407–12422 (2013).
11. Huang, M. *et al.* Engineering genetic predisposition in human neuroepithelial stem cells recapitulates medulloblastoma tumorigenesis. *Cell Stem Cell* **25**, 433-446.e7 (2019).
12. Koch, P., Opitz, T., Steinbeck, J. A., Ladewig, J. & Brüstle, O. A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3225–3230 (2009).
13. Susanto, E. *et al.* Modeling SHH-driven medulloblastoma with patient iPS cell-derived neural stem cells. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 20127–20138 (2020).
14. Čančer, M. *et al.* Humanized stem cell models of pediatric medulloblastoma reveal an Oct4/mTOR axis that promotes malignancy. *Cell Stem Cell* **25**, 855-870.e11 (2019).
15. Bentsen, M. *et al.* ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat. Commun* **11**, 4267 (2020).
16. Issac, J., Raveendran, P. S. & Das, A. V. RFX1: A promising therapeutic arsenal against cancer. *Cancer Cell Int.* **21**, 253 (2021).
17. Feng, C. & Zuo, Z. Regulatory factor X1-induced down-regulation of transforming growth factor β2 transcription in human neuroblastoma cells. *J. Biol. Chem.* **287**, 22730–22739 (2012).
18. Han, H. *et al.* TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucl. Acids Res.* **46**, D380–D386 (2018).
19. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
20. Mootha, V. K. *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
21. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *elife* https://doi.org/10.7554/eLife.21856 (2017).
22. Ewels, P. A. *et al.* The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.* **38**, 276–278 (2020).
23. Cheshire, C. *et al.* nf-core/cutandrun: nf-core/cutandrun v3.1 Lead Rooster. Zenodo https://doi.org/10.5281/zenodo.7715959 (2023).
24. Liberzon, A. *et al.* The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
25. Kinameri, E. *et al.* Prdm proto-oncogene transcription factor family expression and interaction with the Notch-Hes pathway in mouse neurogenesis. *PLoS One* **3**, e3859 (2008).
26. Prajapati, R. S., Hintze, M. & Streit, A. PRDM1 controls the sequential activation of neural, neural crest and sensory progenitor determinants. *Development* **146**, dev181107 (2019).
27. Yu, J., Angelin-Duclos, C., Greenwood, J., Liao, J. & Calame, K. Transcriptional repression by Blimp-1 (PRDI-BF1) involves recruitment of histone deacetylase. *Mol. Cell. Biol.* **20**, 2592–2603 (2000).

28. Williamson, D. *et al.* Medulloblastoma group 3 and 4 tumors comprise a clinically and biologically significant expression continuum reflecting human cerebellar development. *Cell Rep.* **40**, 111162 (2022).
29. Iwama, A. *et al.* Dimeric RFX proteins contribute to the activity and lineage specificity of the interleukin-5 receptor α promoter through activation and repression domains. *Mol. Cell. Biol.* **19**, 3940–3950 (1999).
30. Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J. & Werb, Z. Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nat. Commun.* **11**, 5120 (2020).
31. Cardoso-Moreira, M. *et al.* Gene expression across mammalian organ development. *Nature* **571**, 505–509 (2019).
32. Hayashi, Y. *et al.* BMP-SMAD-ID promotes reprogramming to pluripotency by inhibiting p16/INK4A-dependent senescence. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13057–13062 (2016).
33. Kilkenny, C., Browne, W. J., Cuthill, I. C., Emerson, M. & Altman, D. G. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *PLOS Biol.* **8**, e1000412 (2010).
34. Shechter, D., Dormann, H. L., Allis, C. D. & Hake, S. B. Extraction, purification and analysis of histones. *Nat. Protoc.* **2**, 1445–1457 (2007).
35. Brunton, H., Garner, I. M., Bailey, U.-M., Upstill-Goddard, R. & Bailey, P. J. Using chromatin accessibility to delineate therapeutic subtypes in pancreatic cancer patient-derived cell lines. *STAR Protoc.* **1**, 100079 (2020).
36. Patel, H. *et al.* nf-core/atacseq: [2.1.2] - 2022-08-07. Zenodo https://doi.org/10.5281/zenodo.8222875 (2023).
37. Babraham Bioinformatics - Trim Galore! https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
38. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
39. Picard Tools - By Broad Institute. https://broadinstitute.github.io/picard/.
40. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
41. Stark, R. & Brown, G. DiffBind: Differential Binding Analysis of ChIP-Seq Peak Data. Bioconductor version: Release (3.17) https://doi.org/10.18129/B9.bioc.DiffBind (2023).
42. Ross-Innes, C. S. *et al.* Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).
43. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
44. Castro-Mondragon, J. A. *et al.* JASPAR 2022: The 9th release of the open-access database of transcription factor binding profiles. *Nucl. Acids Res.* **50**, D165–D173 (2022).
45. O'Connor, T., Grant, C. E., Bodén, M. & Bailey, T. L. T-Gene: Improved target gene prediction. *Bioinformatics* **36**, 3902–3904 (2020).
46. Law, C. W. *et al.* 2016 RNA-seq analysis is easy as 1–2–3 with limma, Glimma and edgeR. *F1000Res* **5**, ISCB Comm J-1408 (2016).
47. Ahdesmäki, M. J., Gray, S. R., Johnson, J. H. & Lai, Z. Disambiguate: An open-source application for disambiguating two species in next generation sequencing data from grafted samples. *F1000Res* **5**, 2741 (2016).
48. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
49. 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).
50. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).
51. R Core Team. *R: A Language and Environment for Statistical Computing.* (R Foundation for Statistical Computing, Vienna, Austria, 2021).

Acknowledgements

Author contributions

CS and BS designed and planned the study; CS, SH, AC, SC, SW, and LW performed research; CS, BLG, JJP, PAN, WAW, and BS analyzed data; WAW and BS supervised the research; CS and BS wrote the manuscript and all authors commented on the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-66811-6.

Correspondence and requests for materials should be addressed to W.A.W. or B.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.