

RESEARCH

Open Access



Chromatin remodeling protein BPTF mediates chromatin accessibility at gene promoters in planarian stem cells

Prince Verma¹, John M. Allen¹, Alejandro Sánchez Alvarado² and Elizabeth M. Duncan^{1*}

Abstract

Background The regulation of chromatin accessibility is essential in eukaryotic cells as one of several mechanisms that ensure gene activation occurs at appropriate times and in appropriate cell types. Accordingly, mutations in chromatin remodeling proteins are linked to many different developmental disorders and cancers. One example of a chromatin protein that has been linked to both developmental abnormalities and cancer is BPTF/NURF301, the largest subunit of the Nucleosome Remodeling Factor (NuRF) complex. The BPTF subunit is not only important for the formation of NuRF but also helps direct its activity to particular regions of chromatin by preferentially binding histone H3 lysine four trimethylation (H3K4me3). Notably, defects caused by knockdown of *bptf* in *Xenopus* embryos mimic those caused by knockdown of *wdr5*, a core subunit of all H3K4me3 methyltransferase complexes. However, the mechanistic details of how and where BPTF/NuRF is recruited to regulate gene expression vary between studies and have been largely tested in vitro and/or in cultured cells. Improving our understanding of how this chromatin remodeling complex targets specific gene loci and regulates their expression in an organismal context will provide important insight into how pathogenic mutations disrupt its normal, in vivo, cellular functions.

Results Here, we report our findings on the role of BPTF in maintaining chromatin accessibility and essential function in planarian (*Schmidtea mediterranea*) stem cells. We find that depletion of planarian BPTF primarily affects accessibility at gene promoters near transcription start sites (TSSs). BPTF-dependent loss of accessibility did not correlate with decreased gene expression when we considered all affected loci. However, we found that genes marked by Set1-dependent H3K4me3, but not MLL1/2-dependent H3K4me3, showed increased sensitivity to the loss of BPTF-dependent accessibility. In addition, knockdown of *bptf* (*Smed-bptf*) produces loss-of-function phenotypes similar to those caused by knockdown of *Smed-set1*.

Conclusions The *S.mediterranea* homolog of NuRF protein BPTF (SMED-BPTF) is essential for normal homeostasis in planarian tissues, potentially through its role in maintaining chromatin accessibility at a specific subset of gene promoters in planarian stem cells. By identifying loci that lose both chromatin accessibility and gene expression after depletion of BPTF, we have identified a cohort of genes that may have important functions in stem cell biology.

Keywords Chromatin remodeling, Stem cells, Planarians, Transcription

*Correspondence:

Elizabeth M. Duncan
elizabeth.duncan@uky.edu

¹Department of Biology, University of Kentucky, Lexington, KY, USA

²Stowers Institute for Medical Research, Kansas City, MO, USA



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

Planarians are a robust and powerful model for studying *in vivo* stem cell function. These animals maintain a large population of adult stem cells that they use for both unlimited tissue turnover and to regenerate new organs and tissues after major injury [1]. Unlike most cultured stem cells grown in optimized conditions, the planarian stem cell population is a highly heterogeneous mix of multi and pluripotent cells (also known as neoblasts) that proliferate and differentiate asynchronously in response to injury signals [2]. Unsurprisingly, the planarian stem cell population is also transcriptionally heterogeneous [3–9] with some stem cells expressing various lineage-specific genes (e.g., *myoD* or *foxA*) in addition to common stem cell markers (e.g., *piwi-1*). Studies have shown that these lineage-specific stem cells are required for the maintenance and regeneration of their corresponding tissues [10–12]. Additional research has revealed that positional cues [13], cell cycle phase [14], mitochondrial state [15], and specific signaling pathways [16] are all important regulators of stem cell heterogeneity and function. However, the details about how such changes in cell state and cell signaling are translated into transcriptional changes remain unclear.

One category of proteins with the potential to act as major regulators of transcriptional dynamics in planarian stem cells is chromatin modifying proteins. This group of proteins includes enzymes that add covalent modifications to histone proteins, those that remove them, and complexes that regulate (or “remodel”) chromatin accessibility by shifting and/or removing nucleosomes [17–19]. These modifiers each play important roles in the multilayered, synergistic mechanisms that regulate gene expression in eukaryotic cells. Several studies have identified the planarian homologs of various chromatin proteins and uncovered loss-of-function phenotypes for many of them [20–28]. More recently, as new protocols have been developed, better genomes have been assembled, and the sensitivity of chromatin assays has increased, researchers in the planarian field have made substantial progress in understanding the chromatin state and how it is regulated in planarian cells [25, 26, 28–32]. Together, these studies have revealed both predictable and unexpected findings about the chromatin states of planarian cells and how they are regulated, suggesting that further research in this model will uncover important mechanistic details about how chromatin mechanisms regulate both regeneration and stem cell function.

In this study, we examined both the functional and molecular roles of the chromatin remodeling protein BPTF in planarian stem cells. BPTF is the largest subunit of the Nucleosome Remodeling Factor (NuRF), an ISWI-containing chromatin remodeling complex that executes

its function by shifting nucleosomes bidirectionally in 10 bp increments [33–35]. BPTF is not only an essential subunit of NuRF but also plays a major role in focusing its activity at specific genomic loci through its PHD2 domain, which preferentially binds H3K4me3 [36, 37]. BPTF binding is not only highly specific for H3K4me3 (over other methylated lysines in the H3 tail or other states of H3K4 methylation), but knockdown of *bptf* in *Xenopus* embryos also phenocopies the axial, blood, and gut defects seen in embryos with reduced H3K4me3 (via *wdr5* knockdown) [37]. Together, these data support the hypothesis that BPTF mediates chromatin remodeling and downstream functional effects at genomic loci with specific chromatin features.

In keeping with previous studies that uncovered essential roles for BPTF in mammalian stem cells [38–40], we found that knockdown of *Smed-bptf* by RNA interference (RNAi) caused defects in planarian stem cell function and tissue homeostasis. To identify the specific gene loci affected by the loss of BPTF, we used the Assay for Transposase-Accessible Chromatin with sequencing (ATAC-seq), Chromatin Immunoprecipitation plus sequencing (ChIP-seq), and RNA sequencing (RNA-seq) to identify changes in both the chromatin state and transcriptional output in stem cells isolated from *bptf(RNAi)* planarians. These data revealed that BPTF is important for maintaining chromatin accessibility at the promoters of over 3000 genes, although only a subset of these loci showed corresponding changes in gene expression. We found that genes with both H3K4me3 and ATAC-seq peaks at their promoters had robust and consistent patterns of chromatin accessibility focused near Transcription Start Sites (TSSs), whereas genes with ATAC-seq peaks only (no H3K4me3) were highly accessible but with less consistent patterns of open chromatin. These data support the model in which H3K4me3 recruits NuRF activity to specific loci via BPTF binding.

The *S.mediterranea* genome encodes two H3K4me3 KMTases, Set1 and MLL1/2, and RNAi of these enzymes causes distinct phenotypes: knockdown of *Smed-set1* causes progressive stem cell dysfunction, tissue regression, and death, whereas knockdown of *Smed-ml1/2* induces motility defects due to the gradual loss of epidermal cilia [23, 29]. Importantly, these phenotypes are clearly linked to their distinct genomic targets in planarian stem cells i.e., genes associated with stem cell function and ciliogenesis respectively [29]. Notably, the changes we detected in chromatin accessibility in *bptf(RNAi)* stem cells were more likely to correlate with changes in gene expression if their promoters were marked by Set1-dependent H3K4me3. This finding aligns with the strong similarity in *set1(RNAi)* and *bptf(RNAi)*, but not *ml1/2(RNAi)*, phenotypes. These data strongly support conservation in function across organisms, as

loss of H3K4me3 and loss of BPTF also produce highly similar phenotypes in *Xenopus*. Additionally, our findings identify those H3K4me3-marked genes at which BPTF binding is critical for their stable expression. Given that BPTF has been shown to be essential for stem cell function in both mammals and planarians, our results may point to specific genes and mechanisms that are fundamental to stem cell biology.

Methods

Homolog identification and phylogenetic tree construction

To identify putative BPTF orthologs in the planarian species *Schmidtea mediterranea*, we used validated human (Q12830), mouse (A2A654), and *Drosophila melanogaster* (Q9W0T1.2) BPTF/NURF301 protein sequences in the UniProtKB/Swiss-Prot database [41] as query sequences for protein BLAST (BLASTP) searches against the S2F19 HAP1 PRJNA885486 genome assembly [31] using an evalue cutoff 1.0e-10 (accessed at WormBase ParaSite [42]). The top hit for all queries was h1SMcG0002696. Reciprocal Best Hits (RBH) [43] was used to validate this putative ortholog against all query sequences. Other BPTF orthologs for phylogenetic tree construction were identified from the ENSEMBL database [44] and the associated protein sequence downloaded from UniProtKB/Swiss-Prot [41]. Species chosen for inclusion were model organisms (*Danio rerio*), comparator species (*Crassostrea gigas*, *Ixodes scapularis*, *Ciona intestinalis*) and other invertebrates based on their phylogenetic position (*Hofstenia miamia* and *Macrostomum lignano*). The sequence of the conserved PHD2 domain was identified and isolated from each full-length predicted protein using SMART sequence analysis tool [45]. Both full-length proteins and the isolated PHD2 domains were each aligned by Mafft [46] using default settings and these alignments were subsequently curated using the BMGE software [47]. Phylogenetic trees were constructed in the <https://ngphylogeny.fr/> interface [48] using PhyML with default settings except best of NNI and SPR was used for tree topology search and approximate Bayes was used for branch support [49]. Trees were visualized and re-rooted using the iTOL online tool [50]. Human p300 (Q09472) was used as an outgroup to root both the full-length and isolated PHD-domain trees.

Gene cloning and RNAi

Two separate, non-overlapping regions of the identified *Smed-bptf* gene (h1SMcG0002696) were cloned from a cDNA library using the following gene-specific primers: bptf-1_F = TGAGTAATGTCAATATGAAACC, bptf-1_R = AATAGTCCACATCCGTATATCT, bptf-2_F = AATGTGATAAAACACACGATAC, bptf-2_R = GAAGTTCAGTTAAAAAGTAGGC. We then used established methods [51] to clone each of these sequences into a T4P

vector with flanking T7 binding motifs, transform the resulting plasmids into the bacterial strain Ht115 [52], induce these bacteria to express the cloned sequence as dsRNA (1mM final IPTG, 2 h shaking at 37 C), pellet, rinse, and mix the resulting bacterial pellet with calf liver paste. We then fed animals with this paste as indicated in the results, using a previously generated T4P vector containing a segment of the *Caenorhabditis elegans* unc-22 gene as a control [53]. Each non-overlapping bptf-RNAi construct (1 & 2) induced the same phenotype in multiple independent experiments, after which the bptf-1 RNAi construct was used for all genomic/transcriptomic experiments.

Planarian culture and radiation

An asexual strain (CIW4) of *S.mediterranea* worms were maintained as described [54] with gentamicin supplemented at 50ug/ml. Worms used in experiments were starved ≥ 7 days prior to selection for experiments. Irradiation was done with a Gammacell 40 Exactor (MDS Nordion) at a dose rate of 85 Rad/minute.

Live worm imaging and whole-mount in situ hybridization (WISH)

Live worms images were captured using a Leica M205-FA. Riboprobe for *piwi-1* [55] was made by in vitro transcription with T7 polymerase from the pPR-T4P vector as previously described using DIG labeling [56]. Whole-mount in situ hybridization (WISH), including tyramide amplification of the fluorescent *piwi-1* riboprobe, was performed as previously described [57]. Worms were then mounted in modified ScaleA2 (30% glycerol/4 M urea/0.1% TritonX-100/2.5% DABCO) and imaged using a confocal microscope (LSM 700, Zeiss). Quantitation of *piwi-1*⁺ cells in the confocal images was performed using ImageJ [58]. Briefly, we used the “analyze particles” function to count particles (i.e., *piwi-1*⁺ cells) that met defined size and circularity criteria (10–120 pixels, 0–1 circularity). We then averaged the number of *piwi-1*⁺ cells across the animals in each RNAi condition ($n \geq 9$).

Flow cytometry

Animals were first minced with a clean razor blade, then dissociated in cold calcium- and magnesium-free buffer with 1% BSA (CMFB) over the course of 15 min with occasional pipetting using a P1000 filter top. Dissociated worms were then filtered through 100 μ m followed by 40 μ m cell strainers to remove non-dissociated tissue. Remaining cells were pelleted by low-speed centrifugation at 300rcf (5 min) and resuspended in room temperature CMFB + Hoechst 33,342 (10 μ g/ml). Dissociated cells were incubated with Hoechst 33,342 at room temperature in the dark for 1.5 h. Propidium iodide was

added (5 µg/ml) for the last 5 min to assess cell viability. Viable >2n “X1” stem cells [59] were then collected using a MoFlo Legacy (Beckman Coulter) or BD Influx (BD Biosciences). Both control (*unc-22*) and *bptf(RNAi)* animals were processed in parallel, using both available FACS machines in the Stowers Institute Core facility concurrently to minimize variation between control and *bptf(RNAi)* cells. Cells were stored on ice briefly until all replicates were isolated and then processed immediately (see below). Post-sort analysis was conducted with X1-gated cells to ensure purity before proceeding with downstream applications.

Assay for transposase-accessible chromatin with sequencing (ATAC-seq)

Two replicates of 50 K stem cells per *control(RNAi)* and *bptf(RNAi)* worms were isolated by flow cytometry (see above) and subjected to ATAC according to the original published protocol [60] and using Tn5 transposase enzyme purchased from Illumina as part of the Nextera DNA library prep kit (#FC-121-1030). A total of ten amplification cycles were used in the generation of these ATAC-seq libraries, a number determined by the mid-amplification qPCR assessment recommended in the protocol [60]. The resulting ATAC-seq libraries were sequenced on the Illumina NextSeq 500 (paired-end reads) and quality check (QC) was done using fastqc (version 0.11.9) [61]. Fragment size distributions were generated using ATACseqQC (version 1.26.0) with the fragSizeDist function [62]. Bowtie2 (version 2.4.4) [63] was used to align sample reads to the S3h1 *S.mediterranea* genome [31] using the parameters --very-sensitive -X 1000. After alignment, samtools [64] was used to create BAM files for each sample with those alignments that are properly paired (samtools view -h -b -f 3). Samples were then subsampled based on library complexity and processed as previously described [65]. PCR duplicated reads were removed using the MarkDuplicates function in Picard (version 2.26.4) (<http://broadinstitute.github.io/picard>).

To compensate for Tn5 insertion, a Tn5 shift (+4 and -5 bp) was applied using the bedpeTn5shift.sh script [65]. ATAC-seq peaks were then called using the MACS2 [66] callpeak function (version 2.2.7.1) and parameters -g 8.40e8 -q 0.05 --keep-dup all. Consensus peaks present in both replicates were identified using BEDtools [67] and the annotation of MACS2 peaks to gene models was done using CHIPseeker (version 1.38.0) [68, 69]. deepTools (version 3.5.1) [70] was used to generate normalized coverage files (bigWig files) and profile plots, using parameters bamCoverage --binSize 5 --smoothLength 60 --normalizeUsing RPKM. The Integrative Genomics Viewer (IGV version 2.16.2) was used to visualize bigWig files [71]. The identification of differentially accessible

ATAC-seq peaks was done using csaw (version 1.36.0) [72] as recommended in a previous study [65]. Further quantitative and statistical analysis was performed using R (version 4.3.1) and figures were created using ggplot2 [73].

Gene expression analysis

Four replicates of 100 K X1 stem cells per *control(RNAi)* and *bptf(RNAi)* condition per time point were isolated by flow cytometry (see above) and collected directly into Trizol LS reagent (a version of Trizol optimized for liquid samples, cat #10296010). The isolated cells were then homogenized and total RNA extracted per the Trizol manufacturer protocol. RNAseq libraries were then generated using the Illumina TruSeq kit and sequenced in 50 bp single reads using the Illumina HiSeq 2500. Single-end reads were then aligned to the S3h1 *S.mediterranea* genome using Hisat2 (version 2.2.1) [31, 74] and transcript quantification was performed using StringTie (version 2.1.7). Differential gene expression analysis was then performed using DESeq2 (version 1.42.0) [75, 76]. Genes were considered differentially expressed if they had $p\text{-adj} < 0.05$. Further quantitative and statistical analysis was performed using R (version 4.3.1) and figures were created using ggplot2 [73]. For correlation analyses, regression lines were fitted using the linear model (lm method) in geom_smooth() from ggplot2.

Chromatin-immunoprecipitation plus DNA sequencing (ChIP-seq)

Chromatin-immunoprecipitation was performed as previously described [29, 77]. For each replicate/condition, 300 K planarian X1 stem cells were isolated by FACS (see above), cross-linked in 4% PFA, and mixed with 10 M (cross-linked) *Drosophila* S2 cells before proceeding to the next steps. Chromatin was sheared in the S220 Focused Ultrasonicator (Covaris) with silica beads and immunoprecipitation was performed with an antibody to H3K4me3 (Cell Signaling #9751). Isolated genomic DNA from both input and ChIP samples were used to generate libraries using the Illumina TruSeq kit and then sequenced (single end) on the Illumina HiSeq 2500. Quality checks of all resulting reads (QC) were done using fastqc (version 0.11.9) [61] followed by alignment to the S3h1 *S.mediterranea* genome with bowtie2 (version 2.4.4) [76, 78]. PCR duplicated reads were removed using MarkDuplicates function in Picard (version 2.26.4) (<http://broadinstitute.github.io/picard>). Peak calling was done using the MACS2 [66] callpeak function (version 2.2.7.1) with -g 8.40e8 --nomodel and -q 0.01 parameters. Differential analysis between *set1(RNAi)*, *mll1/2(RNAi)*, and *bptf(RNAi)* H3K4me3 ChIP libraries with their individual matching *control(RNAi)* libraries was performed using diffReps (version 1.55.6), parameters --btr --bco

--nsd --frag 250 [79]. Consensus peaks present in both replicates were identified using BEDtools [67] and peaks annotation to gene models was done using ChIPSeeker (version 1.38.0) [68, 69]. deepTools (version 3.5.1) [70] was used to generate normalized coverage files (bigWig files) and for creating profile plots (bamCoverage --normalizeUsing RPKM). Integrative Genomics Viewer (IGV version 2.16.2) was used to visualize bigWig files [71]. The GO (gene ontology) enrichment analysis was done using enricher function from package clusterProfiler (version 4.10.0) [80, 81]. Further quantitative and statistical analysis was performed using R (version 4.3.1) and figures were created using ggplot2 [73].

Results

The genome of *Schmidtea mediterranea* encodes a homolog of NuRF protein BPTF

BPTF is a H3K4me3-binding protein and the largest subunit of the NuRF chromatin remodeling complex (Fig. 1A). To characterize its functional and molecular role in planarian stem cells, we first identified the likely homolog of *bptf* in the *S. mediterranea* genome using a reciprocal BLAST strategy [43]. This analysis identified a single gene encoding a protein with 27% sequence identity to human BPTF and 28% identity to the *Drosophila* homolog NURF301 (calculated using Clustal Omega; DmNURF301 has 34.5% identity with human BPTF in the same matrix). Despite the modest conservation of overall sequence between these three BPTF homologs, their predicted protein domains are highly conserved (Fig. 1B). Moreover, after aligning their PHD2 motifs and examining these sequences more closely, we found that all four aromatic residues essential for H3K4me3 binding [82] are conserved across human, *Drosophila*,

and SMED-BPTF (Fig. 1C, asterisks). To address the question of SMED-BPTF homology further, we used this sequence and the BPTF homologs of nine other species to construct phylogenetic trees (Fig. 1D, E). Trees constructed with either full-length BPTF sequences (Fig. 1D) or their PHD2 motifs only (Fig. 1E) suggested that human BPTF, *Drosophila* NURF301 and SMED-BPTF are orthologous to each other.

Knockdown of chromatin remodeling protein BPTF mimics knockdown of KMTase Set1

Having identified a homolog of BPTF in *S. mediterranea*, we then asked if it is required for in vivo function. To address this question, we used RNAi to knock down *Smed-bptf* in adult planarian worms. In parallel, we also knocked down *Smed-set1*, *Smed-mll1/2*, and a control gene (*C. elegans unc-22*) using previously reported RNAi constructs [29] for comparison. As shown in previous studies [23, 29], *set1(RNAi)* animals showed defects in tissue homeostasis, including progressive head regression

(Fig. 2A), within two weeks of dsRNA treatment and were all dead within three weeks (Fig. 2B). In contrast, *mll1/2(RNAi)* worms did not show signs of head regression or death even after several weeks (Fig. 2A, B) but did develop motility defects due to loss of ventral cilia [23, 83]. RNAi of *Smed-bptf* caused a penetrant and morphologically similar phenotype to that of *set1(RNAi)*, although its progression developed more slowly (Fig. 2A, B). Notably, this was true for two separate, non-overlapping *bptf* RNAi constructs (Fig. 2B), supporting the conclusion that this phenotype is specific to *Smed-bptf* knockdown and not an off-target effect. However, we note that because the tissue regression phenotypes of *set1(RNAi)* and *bptf(RNAi)* are relatively quick (2–3 weeks) compared to the slower onset of the *mll1/2(RNAi)* motility defect (3–4 weeks), we cannot conclude from these data that SMED-BPTF preferentially operates at Set1 target genes.

Head and tissue regression in planarians can be caused by the failure of many different cellular processes, including those that are critical to planarian stem cells (as these are the only cells that divide and replenish planarian tissues). To assess whether the expression pattern of *Smed-bptf* showed enrichment in a particular tissue or cell type, we performed in situ hybridization using a riboprobe for *Smed-bptf* (Fig. 2C). *Smed-bptf* was detected in a broad pattern of expression with some darker staining outlining the cephalic ganglia and the stem cell dense “tail stripe” (Fig. 2C). Published RNA-sequencing datasets generated from FACS-isolated stem cells [85] and single cell RNA-sequencing (scRNA-seq; Fig. 2D) [7] both indicate that *Smed-bptf* expression is moderately enriched in planarian stem cells. However, it is still possible that the *bptf(RNAi)* phenotype is due to its function in other cells or a combination of cell types.

Previous studies from our lab and others [23, 29, 86] have shown that *set1(RNAi)* worms show defects in planarian stem cell function, including an inability to recover from low doses of radiation. To assess more specifically if knockdown of *Smed-bptf* impacts stem cells, we performed a previously established assay of stem cell function [24, 26, 84] (Fig. 2E): first, we depleted specific gene transcripts (*bptf*, *set1*, *mll1/2*, and a control gene) with RNAi. Next, we divided each group of RNAi worms into two groups: one that we treated with 1250 rads (12.5 Gy) ionizing radiation and a second that was not irradiated (0 rad). As shown previously [24, 26, 84], this dose of radiation reduces the number of stem cells in each animal, challenging those remaining to repopulate and maintain tissue homeostasis. On days 3 and 7 post radiation, we fixed a subset of *bptf(RNAi)* and control(RNAi) animals and then monitored the remaining (live) worms for several weeks to score survival and phenotype (Fig. 2E, F). Notably, both *set1(RNAi)* and

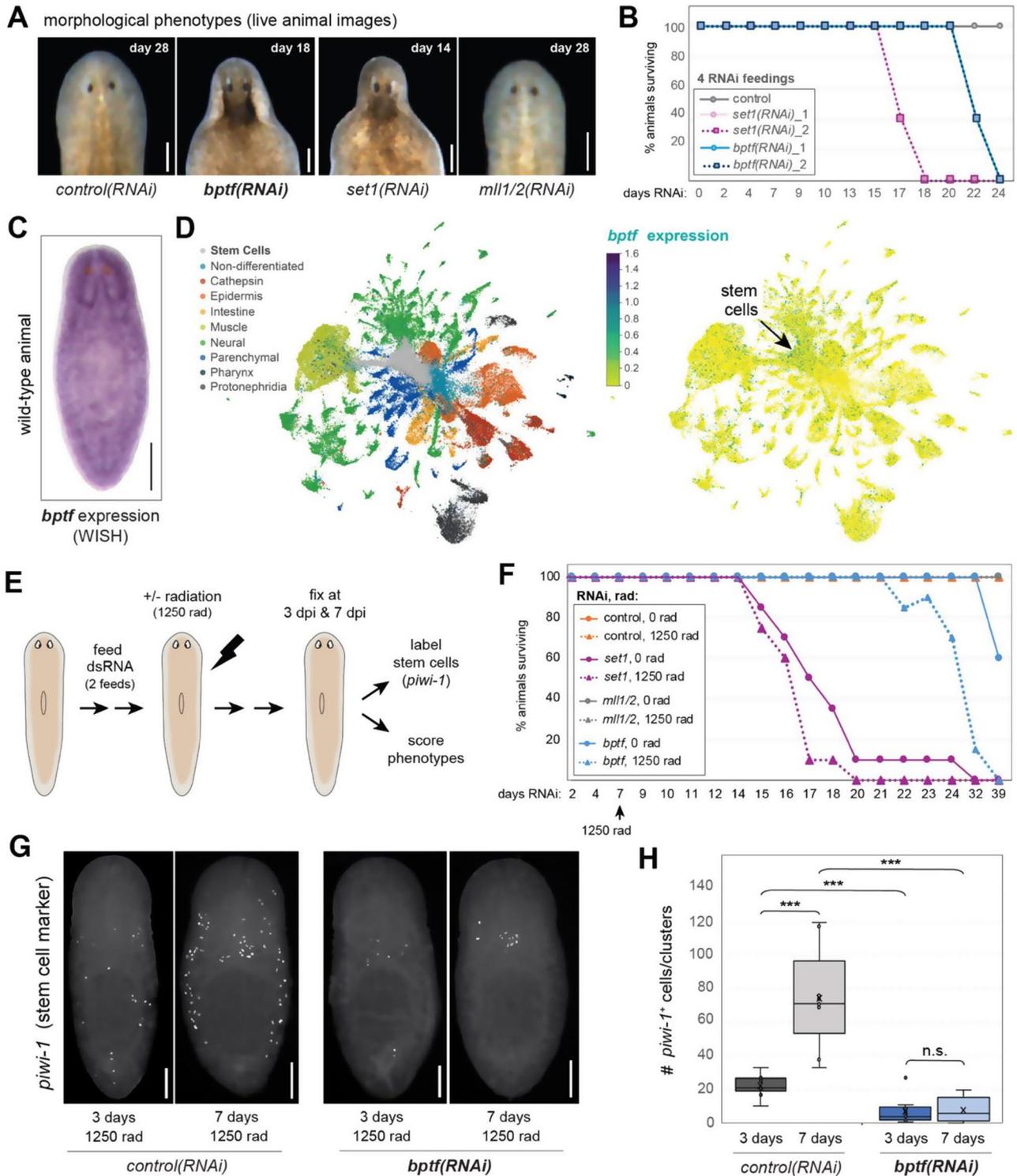


Fig. 2 (See legend on next page.)

Notably, *set1(RNAi)* stem cells also show the same loss-of-function effects in this radiation assay [29], suggesting that the *set1(RNAi)* and *bptf(RNAi)* phenotypes may be due to disruption of similar cellular and/or molecular mechanisms.

Knockdown of *Bptf* leads to loss of chromatin accessibility at gene promoters

After observing that knockdown of *Smed-bptf* caused significant stem cell dysfunction, we then asked if *bptf(RNAi)* causes loss of chromatin accessibility at

(See figure on previous page.)

Fig. 2 RNAi knockdown of *Smed-bptf* phenocopies that of *Smed-set1*, but not *Smed-mll1/2*. **A**) Live images showing the morphological phenotypes of *bptf*(RNAi) planarians (*S.mediterranea*) compared to control(RNAi), *set1*(RNAi), and *mll1/2*(RNAi) animals. Scale bars = 200 μ m. **B**) Survival plot of *set1*(RNAi) and *bptf*(RNAi) animals compared to control RNAi worms. Two non-overlapping constructs are shown for each gene, *set1* and *bptf*, to control for off target effects. **C**) Whole mount In situ hybridization (WISH) in wild-type worm using a riboprobe to *bptf*. **D**) Plots generated from previously published single cell RNA-seq data [7] using the publicly available Shiny App https://simrcompbio.shinyapps.io/bbp_app/ to show *bptf* expression in specific cell types. **E**) Schematic of the experimental setup used to test planarian stem cell function in RNAi worms. In normal conditions, a small but significant number of planarian stem cells will survive 1250 rad (12.5 Gy) γ -radiation (3 days post-irradiation, or dpi), then resume proliferating (7 dpi) to restore the population [24, 26, 84]. **F**) Survival plot of *set1*(RNAi), *mll1/2*(RNAi), *bptf*(RNAi) and control(RNAi) animals with (dotted line) and without (solid line) radiation treatment. **G**) Representative Fluorescence In situ hybridization (FISH) images of control(RNAi) and *bptf*(RNAi) planarians stained with the stem cell marker *piwi-1*. Scale bars = 250 μ m. **H**) Quantitation of *piwi-1*+ cells per animal for all animals included in E. For each condition (RNAi treatment, +/- radiation, time point) $n = 9-12$. Statistical significance determined using student's t-test, *** = p -value < 0.001

specific loci in these stem cells due to its predicted association with NuRF. We tested this prediction by performing ATAC-seq (60) on isolated X1 stem cells (Fig. 3A) and aligning the data to the S3h1 *S.mediterranea* genome assembly [31] (Supplemental Figure S1A). We detected thousands of ATAC-seq peaks in both samples, including those with measurable differences between *control*(RNAi) and *bptf*(RNAi) and those with no significant change (Fig. 3B). After mapping these peaks to the nearest annotated gene model [68, 69], we found that most ATAC-seq peaks in *control*(RNAi) stem cells map near promoters and a significant number map to distal intergenic regions (Fig. 3C, Supplemental Figure S1B). We then used *csaw* [87] to identify those peaks with significant changes in ATAC-seq signal in *bptf*(RNAi) stem cells.

More than half of the differentially-accessible (DA) peaks we detected in *bptf*(RNAi) stem cells mapped near gene promoters, with a relatively smaller fraction mapping to distal intergenic regions (Fig. 3D; Supplemental Figure S1C). The majority of these DA peaks lose ATAC-seq signal versus gain (Fig. 3E, Supplemental Figure S1C), as predicted if SMED-BPTF functions in chromatin remodeling via NuRF. This difference was also unambiguously observed when we averaged ATAC-seq signal across all annotated gene models and centered it at Transcription Start Sites (TSSs): chromatin accessibility was significantly lower in *bptf*(RNAi) stem cells +/- 1 kb of the TSS (Fig. 3F, Supplemental Figure S2A). This was also the case when we compared ATAC-seq signal at defined promoter peaks (Supplemental S2E, F). However, chromatin accessibility was not detectably changed at distal intergenic loci (Fig. 3G, Supplemental Figure S2B) or peaks at "other introns" (Supplemental Figure S2G, H) in *bptf*(RNAi) stem cells. Additionally, we did not find significant overlap between these BPTF-dependent DA peaks and those previously shown to be differentially accessible after knockdown of the BAF chromatin remodeling complex (also in X1 stem cells [28]). Together, these data support the hypothesis that SMED-BPTF is a conserved subunit of the NuRF complex that helps direct NuRF activity to specific genomic loci in planarian stem cells. However, further biochemical studies will be needed to confirm this hypothesis more conclusively.

Knockdown of *Bptf* leads to significant changes in gene expression

As changes in chromatin accessibility have the potential to affect transcription, we then asked if knockdown of *bptf* also impacted gene expression in planarian stem cells. We again isolated X1 stem cells by flow cytometry from *control*(RNAi) and *bptf*(RNAi) animals and then extracted their total RNA for RNA-sequencing. We isolated stem cells at two time points post-RNAi treatment (day 9 and day 12 post first RNAi feeding) to assess the timing and consistency of gene expression changes after loss of chromatin accessibility. Indeed, we detected many significant changes at both time points (Fig. 4A), although overall the changes in gene expression were consistent between time points for both up- and down-regulated genes (Fig. 4B, Supplemental Figure S3A). Interestingly, we saw a relatively even distribution of up and down-regulated genes in *bptf*(RNAi) stem cells (Figs. 4A; 1603 up-regulated and 1886 down-regulated genes at day 9, 1288 up-regulated and 1332 down-regulated genes at day 12), which was somewhat unexpected given the overall loss of chromatin accessibility at gene promoters in these cells (Fig. 3E-F). However, it is possible that many of these DEGs are indirect effects of *bptf*(RNAi).

We then asked if there was a correlation between BPTF-dependent changes in chromatin accessibility and gene expression. We first generated Venn diagrams to determine how many gene loci had both DA (Differentially Accessible) ATAC-seq peaks at their promoters and differential expression by RNA-seq (Fig. 4C). The overlap between DA promoters and differentially expressed genes (DEGs) was significantly greater than what is predicted by chance, as calculated with hypergeometric tests using the total number of annotated gene models as the universe. However, these comparisons do not assess the relationship between changes in chromatin accessibility and gene expression at individual loci. To address this question, we compared the \log_2 Fold Change (\log_2 FC) in chromatin accessibility (*csaw*) at gene promoters with their \log_2 FC in gene expression (RNA-seq; Supplemental Figure S3C). When we included all genes with a calculated value in both datasets (i.e., any p Adj), we found no correlation between changes in chromatin accessibility

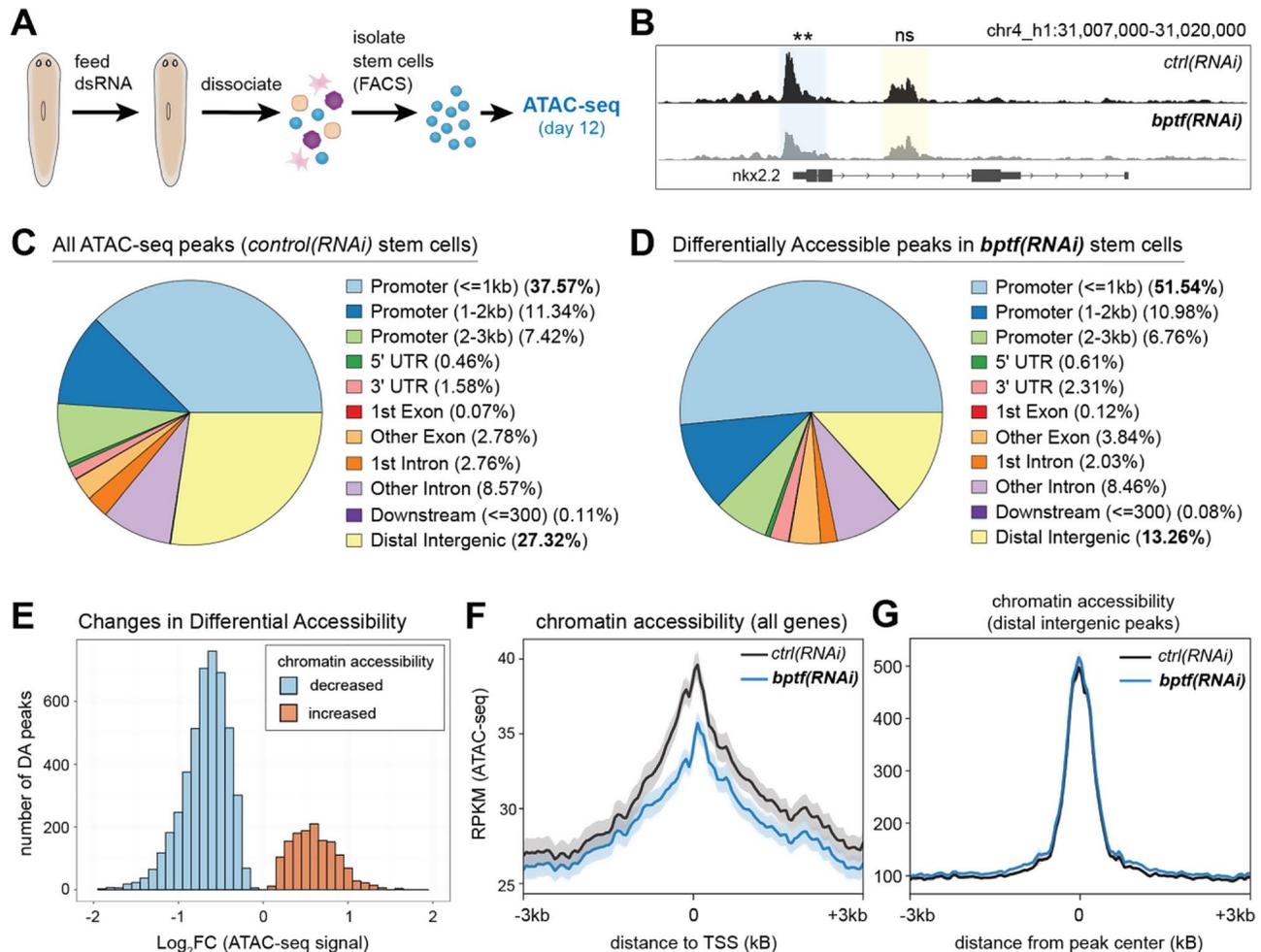


Fig. 3 Knockdown of *bptf* leads to loss of chromatin accessibility at gene promoters in planarian stem cells. **A**) Schematic of the experimental setup used to assay the chromatin state of planarian stem cells isolated from *control(RNAi)* and *bptf(RNAi)* animals. **B**) A representative locus at which there are multiple ATAC-seq peaks, including one that shows significant loss of accessibility after *bptf(RNAi)*. Gene *nkx2.2*=h15McG0021724. **C**) Pie chart summarizing the locations of ATAC-seq peaks across the planarian genome (relative to their nearest gene models) in *control(RNAi)* stem cells. **D**) Pie chart summarizing the locations of differentially accessible (DA) peaks in *bptf(RNAi)* stem cells (compared to *control(RNAi)* stem cells). **E**) Histogram summarizing the \log_2FC of all differentially accessible (DA) ATAC-seq peaks in *bptf(RNAi)* stem cells (binwidth=0.1 \log_2FC). 4652 peaks have decreased accessibility (blue bars), 1411 have increased accessibility (orange bars). **F**) Profile plot comparing ATAC-seq signal in *control(RNAi)* and *bptf(RNAi)* stem cells averaged across all gene models and centered at transcription start site (TSS). Signal is represented as Reads Per Kilobase per Million mapped reads (RPKM). The shaded area represents the standard error. **G**) Profile plot comparing ATAC-seq signal at distal intergenic peaks (identified in C). The shaded area represents the standard error

and gene expression (Supplemental Figure S3C). When we restricted our analysis to genes with significant ($pAdj < 0.05$) changes in both datasets (Fig. 4D), we still found no significant correlation at day 9 post-RNAi and a marginally significant positive correlation at day 12. We also did not find a significant correlation when comparing differential expression with differential accessibility using peaks in other mapping categories (e.g., distal intergenic, 5' UTR, etc.; data not shown). However, we note that those categories contain relatively fewer genes in each and it is more challenging to assign those peaks to the gene(s) they may be regulating.

Importantly, when we binned genes into four broad expression categories based on their average TPM

(transcripts per million) in *control(RNAi)* cells and then compared the ATAC-seq signal at their promoters (± 1 kb), we did observe a significant correlation between average gene expression and promoter accessibility (Supplemental Figure S3B). Moreover, RNAi of *bptf* had a significant impact on accessibility at genes of all expression levels except the highest bin (> 100 TPM). These data suggest that BPTF-dependent loss of chromatin accessibility can impact gene expression, but not at all gene loci and potentially not at those genes expressed at very high levels.

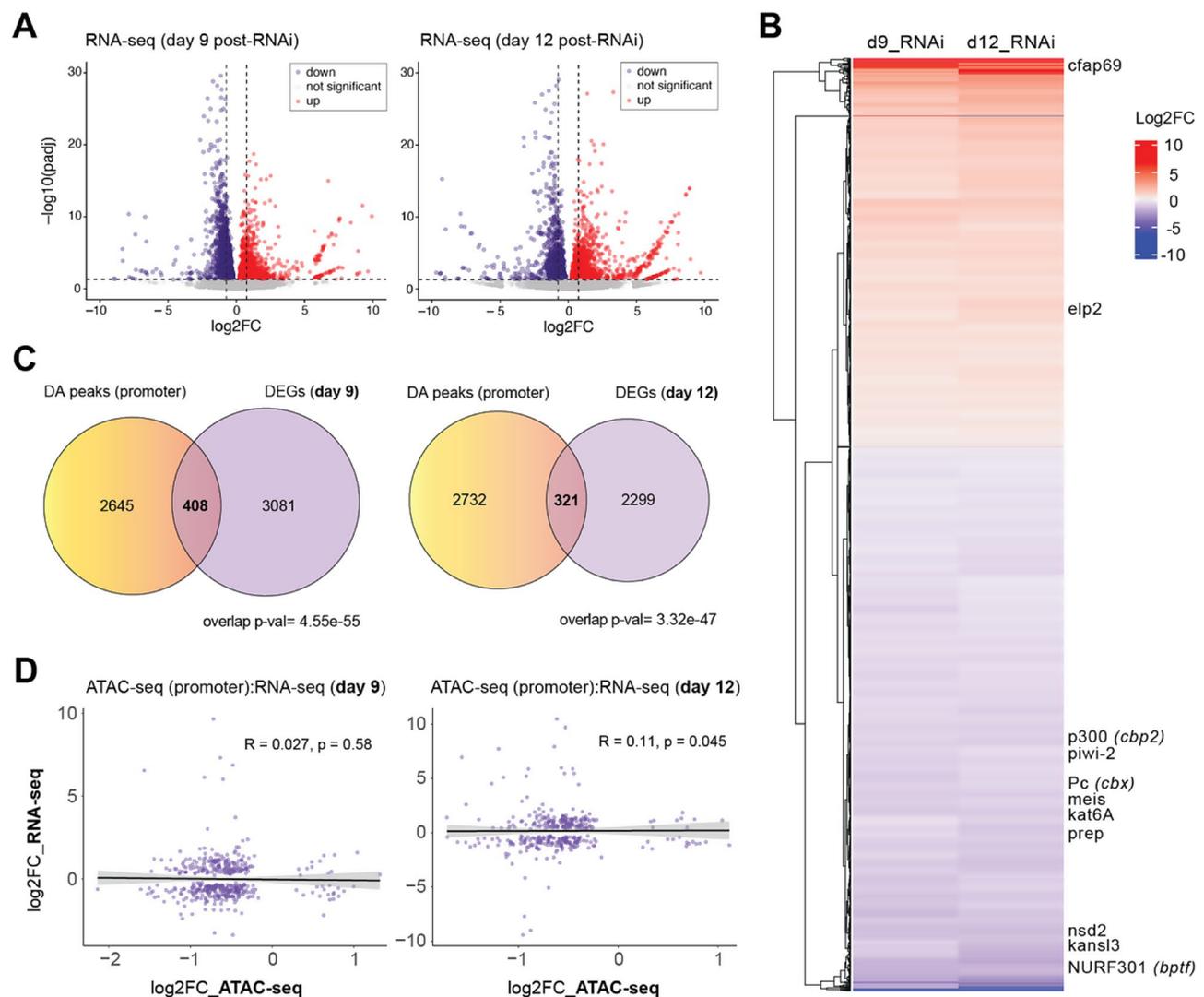


Fig. 4 Loss of BPTF leads to significant changes in gene expression that indicate the dysregulation of transcription and chromatin regulation. **A**) Volcano plots showing transcript changes in *bptf(RNAi)* stem cells versus *control(RNAi)* as detected by RNA-seq. Left plot shows data from stem cells isolated 9 days post-RNAi; right plot shows data from 12 days post-RNAi. Blue dots = significantly down-regulated ($p_{Adj} < 0.05$; 1886 at day 9, 1332 at day 12), red dots = significantly up-regulated ($p_{Adj} < 0.05$; 1603 at day 9, 1288 at day 12). Vertical dashed lines = $\log_2FC \pm 0.75$. Horizontal dashed line = linear $p_{Adj} < 0.05$. **B**) Clustered heat map of day 9 and day 12 RNA-seq data in **B**; only genes with significant differential expression ($p_{Adj} < 0.05$) at both time points are included. Changes in expression between time points correlate strongly (Supplemental Figure S3A). **C**) Venn diagrams comparing genes identified as differentially accessible (Fig. 3; csaw) and differentially expressed (**A**; DESeq2). Genes included have $p_{Adj} < 0.05$. Overlap p -values were calculated using the Hypergeometric test. **D**) Plots comparing changes in chromatin accessibility (ATAC-seq) and gene expression (RNA-seq) for those genes identified as significantly differential in both in **C**. Left plot shows correlation with RNA-seq data from day 9 post-*bptf(RNAi)* and right plot shows data from day 12. Correlation coefficients (R) and p -values were calculated using Spearman Rank correlation. The grey shaded areas around the regression lines represent 95% confidence intervals

BPTF-mediated accessibility is enhanced at promoters with H3K4me3

Because BPTF is known to bind H3K4me3 and this modification is highly enriched at gene promoters [29, 88, 89], we wondered if its presence might distinguish those loci at which BPTF affects gene expression from those at which it does not. To address this question, we first examined H3K4me3 patterns in both *control(RNAi)* and *bptf(RNAi)* stem cells. We again isolated X1 stem cells from both *control(RNAi)* and *bptf(RNAi)* animals,

isolated their chromatin, and performed ChIP-seq with an antibody specific to H3K4me3. As shown previously in *S.mediterranea* [29, 31] and many other organisms [88–90], nearly all MACS2-called H3K4me3 peaks were detected at gene promoters (Fig. 5A, B). Importantly, we did not observe a detectable change in H3K4me3 after knockdown of *bptf* (Fig. 5B), indicating that the chromatin accessibility changes we detected in *bptf(RNAi)* stem cells were not due to loss of its binding site. We then asked how many gene promoters have both a

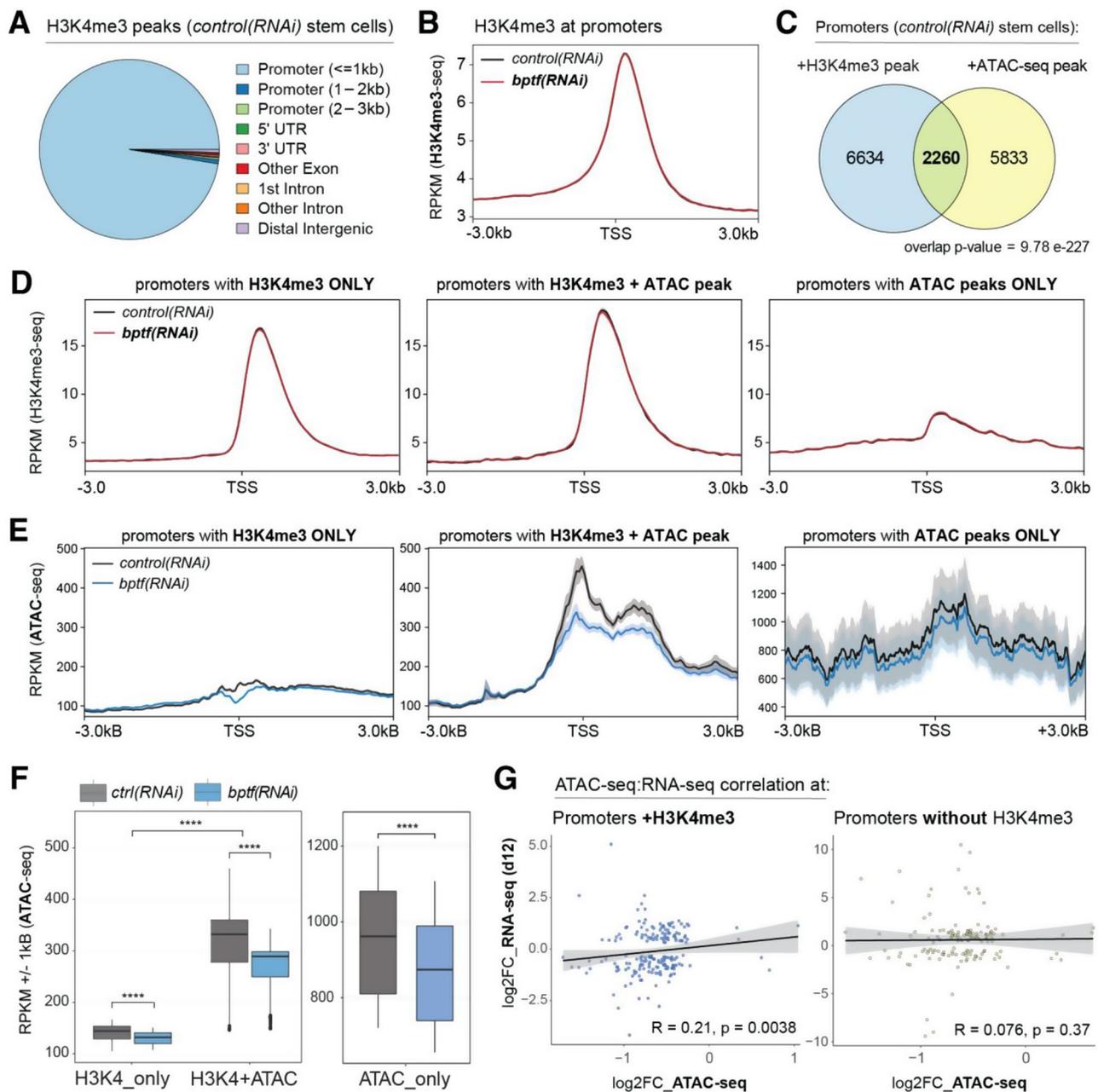


Fig. 5 H3K4me3 concentrates chromatin accessibility at promoters. **A**) Pie chart showing the distribution of mapped H3K4me3 peaks in *control(RNAi)* stem cells. **B**) Profile plot comparing H3K4me3 signal in *control(RNAi)* and *bptf(RNAi)* stem cells averaged across all genes. The shaded area represents the standard error. **C**) Venn diagram showing the number of promoters in planarian stem cells with an H3K4me3 peak, an ATAC-seq peak, or both. Overlap p-value was calculated using the hypergeometric test. **D**) Profile plot comparing H3K4me3 signal in *control(RNAi)* and *bptf(RNAi)* stem cells at the gene promoters in each group of the Venn in C (H3K4me3 only, H3K4me3 + ATAC peak, and ATAC peak only). The shaded area represents the standard error. **E**) Profile plot comparing ATAC-seq signal in *control(RNAi)* and *bptf(RNAi)* stem cells for gene promoters in each group of the Venn shown in C. The shaded area represents the standard error. **F**) Box plots measuring changes in ATAC-seq signal at the promoters of genes in each group in E. Statistical significance was determined using a Wilcoxon test (p-value with Bonferroni correction, **** = $p \leq 0.0001$). **G**) Scatter plots comparing changes in chromatin accessibility (ATAC-seq; csaw) with changes in gene expression (RNA-seq; DESeq2) in *bptf(RNAi)* X1 stem cells at loci with an H3K4me3 peak at their promoters (left plot) versus those without an H3K4me3 peak (right plot). Data points shown have $p_{Adj} < 0.05$ in both datasets. Correlation coefficients (R) and p-values were calculated using Spearman Rank correlation. The grey shaded areas around the regression lines represent 95% confidence intervals

MACS2-called H3K4me3 peak and a MACS2-called ATAC-seq peak in *control(RNAi)* stem cells. We found that 28% of all promoters with a MACS2-called ATAC-seq peak \pm 1 kb also had an H3K4me3 peak within 1 kb (Fig. 5C). Conversely, 25% of all promoters with an H3K4me3 peak also had an ATAC-seq peak (Fig. 5C).

The number of genes with both ATAC-seq and H3K4me3 peaks at their promoters may be fewer than expected based on other studies. However, we note the following: first, when we examined recently published H3K4me3 ChIP-seq and ATAC-seq data generated from whole planarian worms [31], it also showed a large number of ATAC-seq peaks at promoters that did not have H3K4me3 (Supplemental Figure S4C). Second, when we individually inspected several "ATAC-seq only" peaks using IGV, we found they had robust and specific signal near promoters (Supplemental Figure S5) and looked like true ATAC-seq peaks. Lastly, most comparable genomic datasets were generated from mammalian cell lines grown in culture conditions that were optimized for homogeneity [91, 92]. In contrast, our data was generated from isolated planarian stem cells that are known to be transcriptionally heterogeneous and include a mix of pluri- and multipotent stem cells.

To determine whether the presence of an H3K4me3 peak impacts BPTF-mediated chromatin accessibility, we then examined the average H3K4me3 (Fig. 5D) and ATAC-seq signal (Fig. 5E) at genes annotated with: (1) an H3K4me3 peak only; (2) both an H3K4me3 peak and an ATAC-seq peak; and (3) an ATAC-seq peak only (groups based on Fig. 5C). This analysis revealed that *bptf(RNAi)* induces the clearest loss of accessibility at genes with both H3K4me3 and ATAC-seq peaks (Fig. 5E, middle), although it has a statistically significant impact on chromatin accessibility at loci in all three categories (Fig. 5F). Noticeably, genes with "ATAC-seq only" promoters had relatively higher levels of ATAC-seq signal overall (Fig. 5E, right) and the signal was highly variable between genes in this group (as indicated by the shaded standard error). We noted that this third group included significantly more "h1SMnG" gene models (i.e., putative non-coding) and wondered whether the chromatin state at these loci was different from that of "high confidence" gene models (defined in [31]). However, when we split the "ATAC-seq only" group into "high confidence" models versus "putative non-coding", we did not see a major difference in the profile plot (Supplemental Figure S4D). We further noted that the average ATAC-seq signal at the "ATAC peaks only" loci was less concentrated at the TSS than that of gene loci with both ATAC and H3K4me3 peaks (Fig. 5E). Yet, as discussed above, the ATAC-seq only peaks did show robust signal to background (i.e., distinct peaks) when inspected at individual gene loci (Supplemental Figure S5). These data suggest

that H3K4me3 may focus NuRF activity at more precise nucleosome positions relative to the TSS.

We then asked if genes with BPTF-dependent ATAC-seq peaks are more likely to exhibit changes in gene expression if their promoters are marked by H3K4me3 peaks. First, we used Venn diagrams to compare any significant change in accessibility (increased or decreased) with any significant change in gene expression (up or down) after *bptf(RNAi)* in X1 stem cells (Supplemental Figure S4F, G). Based on hypergeometric tests, changes in chromatin accessibility at any promoter (i.e., those with or without H3K4me3) have significant overlap with changes in expression. To determine if these correlations were also directional and at individual genes, we then performed correlation analyses. When we included all genes with a calculated log2FC (any pAdj) for both ATAC-seq and RNA-seq, neither genes with H3K4me3 at their promoters nor those without showed a statistically significant correlation (Supplemental Figure S4H). However, when we restricted our analysis to genes with significant changes (pAdj < 0.05) in both ATAC-seq and RNA-seq after *bptf(RNAi)*, there was only a significant correlation for those genes with H3K4me3 peaks at their promoters but not those without (Fig. 5G).

Set1 targets are significantly regulated by BPTF/NuRF

We then wondered if the type of H3K4me3 peak at a gene promoter impacted the role of BPTF. To address this question, we first realigned our previously published H3K4me3 ChIP-seq data generated from *set1(RNAi)* and *mll1/2(RNAi)* X1 stem cells on the new *S.mediterranea* genome assembly (Supplemental Figure S6 and Supplemental Table 3). After identifying the new coordinates of Set1 and MLL1/2-dependent target loci, we then compared the ATAC-seq signal at these distinct sets of gene loci (Fig. 6A). Unsurprisingly given our previous findings [29], Set1 target genes had considerably more ATAC-seq signal than MLL1/2 target genes (Fig. 6A, B). At the same time, both Set1 and MLL1/2 target genes lost a statistically significant amount of ATAC-seq signal in *bptf(RNAi)* stem cells (Fig. 6B), even though the chromatin at MLL1/2 target genes was overall less accessible.

We then asked if *bptf(RNAi)* had different impacts on gene expression at Set1 versus MLL1/2 target genes. First, we compared promoters at which the chromatin state was affected by depletion of either Set1, MLL1/2, or BPTF (Fig. 6C). As expected, Set1 and MLL1/2 target distinct gene loci and a subset of each are affected by loss of BPTF. We then compared all common Set1 + BPTF target genes (217 genes) with genes differentially expressed in *bptf(RNAi)* stem cells (Fig. 6D, left). The overlap of these gene sets was significantly greater than what is predicted by chance (hypergeometric test). This was also true for the parallel comparison of genes targeted by both

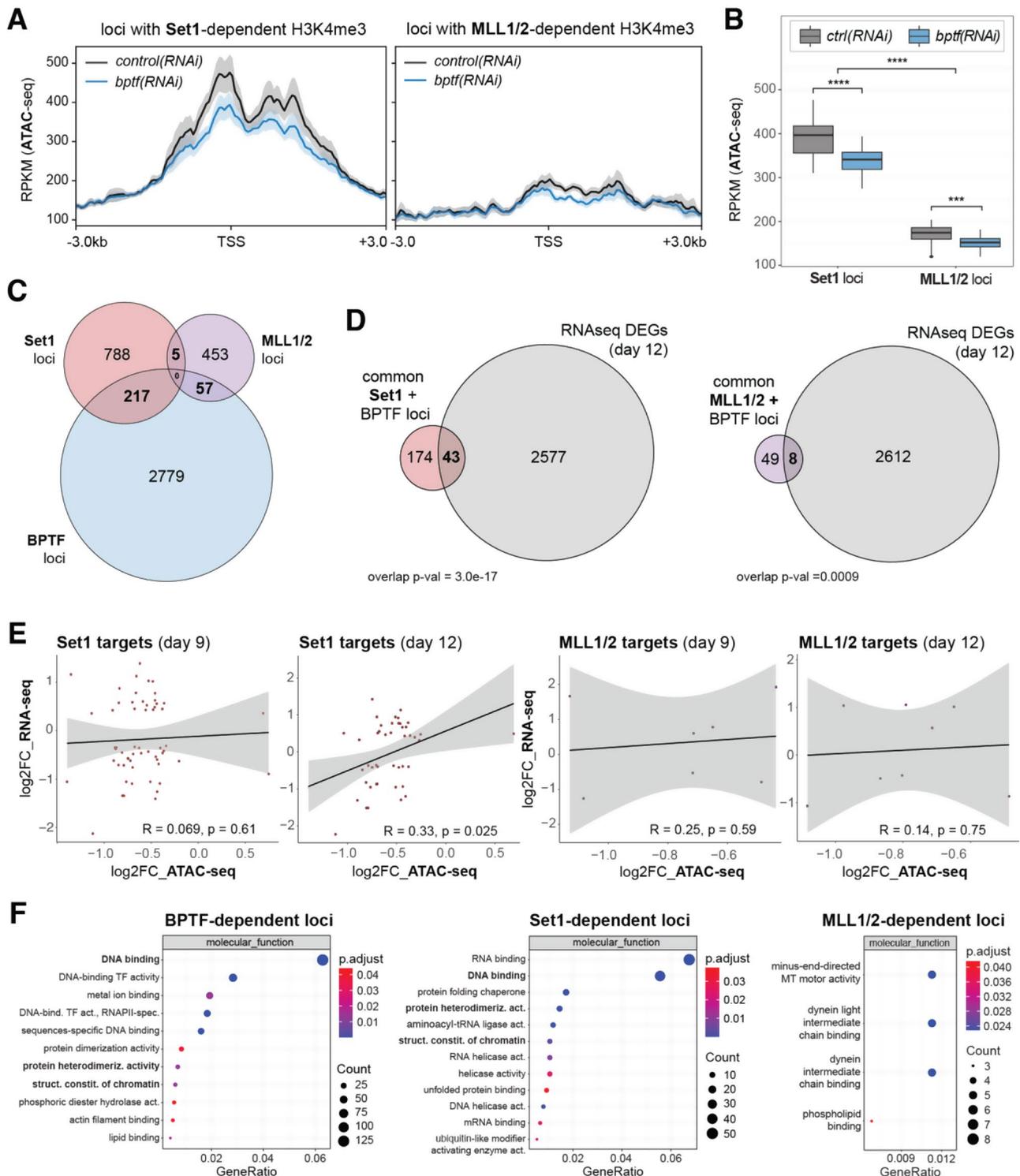


Fig. 6 (See legend on next page.)

MLL1/2 and BPTF (57 genes; Fig. 6D, right). However, these analyses did not consider the degree or direction of change (i.e., increased or decreased accessibility or expression). When we compared significant ATAC-seq changes with significant RNA-seq changes at individual

genes, only genes targeted by both Set1 and BPTF (but not MLL1/2 + BPTF) showed a significant, positive correlation (Fig. 6E). Notably, this was only true for RNA-seq changes at day 12 post RNAi, suggesting the effects of changing chromatin accessibility do not immediately

(See figure on previous page.)

Fig. 6 BPTF shares functional genomic targets with Set1. **A)** Profile plot comparing ATAC-seq data from *control(RNAi)* and *bptf(RNAi)* stem cells at Set1 (left) and MLL1/2 (right) gene targets. The shaded area represents the standard error. **B)** Quantitation and statistical analysis of ATAC-seq data in A; statistical significance was determined using a Wilcoxon test (p-value with Bonferroni correction). * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$, **** = $p \leq 0.0001$. **C)** Venn diagram comparing genes identified as likely targets of Set1, MLL1/2, and BPTF in X1 stem cells. Targets = genes with loss of H3K4me3 or ATAC-seq at promoters (+/- 1 kb of TSS) in their respective RNAi conditions. p-val = $8.65e-74$ for BPTF targets overlap with Set1 targets; p-val = $1.04e-75$ for BPTF targets overlap with MLL1/2. **D)** Venn diagrams comparing common Set1 and BPTF targets (217 genes) with DEGs at day 12 (left) as well as common MLL1/2 and BPTF targets with DEGs at day 12 (right). P-values were calculated using a hypergeometric test. **E)** Plots comparing differential changes in chromatin accessibility (ATAC-seq) with differential changes in gene expression (RNA-seq) in *bptf(RNAi)* stem cells at Set1 versus MLL1/2 target genes (from C, D). Data points shown have pAdj < 0.05 in both datasets. Correlation coefficients (R) and p-values were calculated using Spearman Rank correlation. The grey shaded areas around the regression lines represent 95% confidence intervals. **F)** GO Term enrichment analysis for genes with reduced chromatin accessibility in *bptf(RNAi)* X1 stem cells ("BPTF-dependent loci"), reduced H3K4me3 in *set1(RNAi)* X1 stem cells ("Set1-dependent loci"), and reduced H3K4me3 in *ml1/2(RNAi)* X1 stem cells ("MLL1/2-dependent loci")

impact stable transcript levels. It was also only true if we restricted the analysis to significant values (i.e., pAdj < 0.05 in both RNA-seq and ATAC-seq datasets; Fig. 6E) but not if we included log2FC values with any pAdj (Supplemental Figure S6C, D). Separately, when we performed Gene Ontology (GO) term analysis for genes targeted by BPTF, Set1 and MLL1/2 (Fig. 6F), we found that several molecular function terms were enriched in both BPTF and Set1 lists (e.g., DNA binding, structural constituent of chromatin). In contrast, terms enriched in the descriptions of MLL1/2 target genes were largely related to cilia.

We previously showed that Set1 and MLL1/2-dependent loci are marked by measurably different H3K4me3 patterns; Set1 catalyzes peaks with significantly wider widths than those added by MLL1/2 [29, 86] (Supplemental Figure S6A, B). To test if BPTF/NuRF activity is affected by different H3K4me3 peak widths, irrespective of the specific KMTase creating them, we divided all H3K4me3 peaks from *control(RNAi)* stem cells into three categories: "narrow", "medium", and "broad" (Supplemental Figure S7A-C). After plotting the average ATAC-seq signal in each group (Supplemental Figure S7D), we found that the overall ATAC-seq signal increased with H3K4me3 peak width (Supplemental Figure S7E, F). Unexpectedly, when we analyzed the reduction of BPTF-dependent chromatin accessibility across these three groups, the loss of ATAC-seq signal at genes with broad H3K4me3 peaks was not statistically significant (Supplemental Figure S7E, F). This was unexpected, as the average width of these H3K4me3 peaks was similar to those marked by Set1 (Supplemental Figure S6B, S7B). These findings suggest that Set1 and MLL1/2 target loci have additional features, beyond the H3K4me3 peaks they catalyze, that recruit, enhance and/or restrict BPTF/NuRF function (Fig. 7).

Discussion

Planarian stem cells have many fascinating cellular characteristics and are required for two extraordinary features of their organism: (1) planarians can repeatedly and reproducibly regenerate complex tissues and organs, and

(2) they are effectively immortal. Although many genes have been identified as essential regulators of planarian stem cell function, including genes that encode chromatin proteins [21–24, 26, 28, 29], the molecular mechanisms and genetic targets through which they accomplish this regulation are not well understood. Here, we examined the role of the chromatin remodeling complex NuRF in planarian stem cell biology. By knocking down its largest and specific subunit, BPTF, we both targeted the NuRF component that provides its chromatin-state specificity [37, 93] and tested its effects on transcription in vivo. Our findings show that BPTF, likely through the remodeling activity it recruits to specific gene loci, is critical for planarian stem cell function. We are currently working on analyses to examine the homology and function of these BPTF-sensitive genes, many of which encode proteins with identifiable domains (e.g., zinc fingers) but unclear homology.

Limitations of the study

When interpreting our data, several points must be considered. First, we acknowledge that the method for isolating stem cells from planarians is based on DNA content, i.e., cells with >2n DNA [59]. This means that most cells in our study, and in most molecular/genomic studies of planarian stem cells, are in the late S/G2/M phases. It is possible that this characteristic may skew our chromatin accessibility data (ATAC-seq) toward a less accessible state, as chromosomes become highly compacted in M phase and studies using synchronized cells have reported that ATAC-seq signal decreases as cells progress through these phases of cell cycle [94]. In this study, we are comparing cells in which we have depleted an integral subunit of a chromatin remodeling complex, BPTF, to control cells. Despite the possible dampening of ATAC-seq signal due to cell cycle-induced compaction, we still detect many significant changes in chromatin accessibility after its depletion. However, it is possible that we would detect many more if we could isolate all stem cells, including those with 2n DNA content (in G1 or G0).

Second, knockdown of *bptf* may cause a small, but consistent, loss of accessibility at many gene promoters,

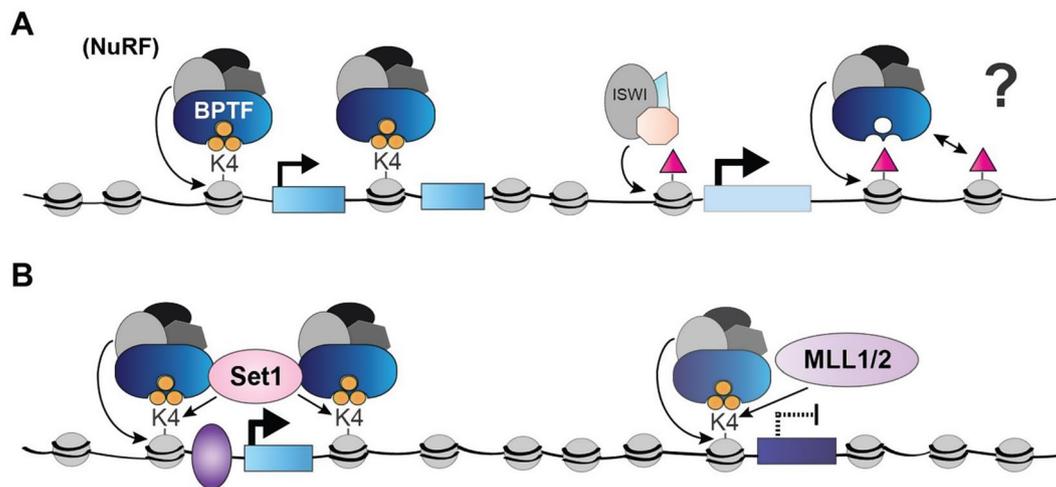


Fig. 7 Model of BPTF/NuRF function in planarian stem cells. **A**) SMED-BPTF is predicted to bind histone H3 lysine 4 trimethylation (H3K4me3; orange circles), recruiting the rest of the NuRF complex and its ATP-dependent chromatin remodeling activity to transcription start sites. Many other genes in the planarian genome are not marked by H3K4me3, but do have BPTF-dependent regions of chromatin accessibility. It is possible that other chromatin modifications, such as acetylation (magenta triangles), are responsible for recruiting and/or stabilizing BPTF/NuRF. **B**) Genes marked by Set1-dependent H3K4me3 have more open chromatin and higher expression than those marked by MLL1/2-dependent H3K4me3

but these changes may not reach the threshold required by csaw to identify them as differentially accessible (DA) peaks. Related, some gene loci may have highly dynamic chromatin states, making it difficult to assay their accessibility reproducibly across replicates. Both scenarios could depress our ability to detect significant changes in chromatin accessibility in *bptf(RNAi)* stem cells, which then also affects our correlation analyses with gene expression. Third, we are using standard RNA-sequencing measurements, which measure all mRNA transcripts in our cells. As these methods largely detect steady-state cytoplasmic mRNA transcripts, they are not an ideal readout of transcriptional activity. Assays such as GRO-seq or NAC-seq would be better measurements of dynamic transcriptional changes, but they are not currently tractable in the planarian model. Nonetheless, these issues are not unique to planarian studies; similarly weak correlations between changes in ATAC-seq and RNA-seq have been reported in other species and contexts, suggesting they may suffer from similar limitations [95].

Finally, the planarian field is still in the early stages of understanding the genomes of these animals and how they are regulated. Chromosome-level genome assemblies for *S.mediterranea* were only recently released [31, 96], and work remains to be done to understand its regulatory networks. For example, although the *S.mediterranea* genome is relatively large (~800 MB) with sizeable regions of repetitive sequence, the genic regions are often very dense. These details are relevant to any genomic assay, as they impact steps like mapping ATAC-seq peaks to their nearest gene model. Nevertheless, we anticipate that our data will provide a meaningful

contribution to the characterization of gene loci that regulate planarian stem cell function.

Conclusions

Using ATAC-seq, we identified many specific regions of the planarian genome at which chromatin accessibility changed significantly after *bptf* knockdown. We confirmed that, as expected, the chromatin at Set1 target genes is significantly more accessible than that at MLL1/2 targets. Moreover, genes with both a BPTF-dependent loss of accessibility and a Set1-dependent H3K4me3 peak showed the strongest overall correlation with transcription. Finally, we observed that knockdown of *bptf* replicates many of the physiological, cellular, and molecular changes seen in *set1(RNAi)* planarians, including the inability of their stem cells to repopulate after treatment with ionizing radiation.

Our study is also important because it likely informs how BPTF operates in vivo. In humans, BPTF appears to have dose-dependent functions: haploinsufficiency of BPTF causes abnormal craniofacial and brain developmental [97, 98] whereas over-expression of BPTF is linked to oncogenic behavior in various cancers [99–101]. These data suggest that some cell types and/or genomic loci are more sensitive to the amount of BPTF-dependent NuRF activity than others. However, the exact mechanisms driving these functional outcomes in different in vivo settings remain unclear. Because we robustly reduced, but did not eliminate, *bptf* expression, we identified the loci and genes that are most sensitive to precise levels of NuRF activity in planarian stem cells. Despite this incomplete loss, *bptf(RNAi)* animals have severe and penetrant phenotypes, showing that normal NuRF

activity levels are critical for in vivo cellular function in planarians as in humans. Future studies will be aimed at understanding the functions and gene regulatory networks of those most sensitive target genes.

Abbreviations

H3K4me3	Histone H3 lysine 4 trimethylation
NuRF	Nucleosome Remodeling Factor
ATAC-seq	Assay for Transposase-accessible Chromatin with sequencing
DA	Differentially accessible (ATAC-seq peak)
ChIP	Chromatin Immunoprecipitation
MLL1/2	Mixed Lineage Leukemia 1 or 2
BPTF	Bromodomain PHD finger transcription factor
S.mediterranea	Schmidtea mediterranea
RNAi	RNA interference
KMT	Lysine methyltransferase
KDM	Lysine demethylase
WISH	Whole mount in situ hybridization
Rad	Radiation unit (1 rad = 0.01 Gy)
GO	Gene Ontology

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-025-11405-3>.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Acknowledgements

The authors would like to thank the Cytometry, Molecular Biology, Imaging, and Tissue Culture cores at the Stowers Institute for Medical Research for all their hard work in maintaining such well-run facilities and in assisting E.M. Duncan as she collected and/or processed samples. We thank Chi Wang, Jinpeng Liu, and Hua Li for their helpful discussions and statistical expertise. We would also like to thank Laura Banaszynski, Ann Morris and Eric Ross for their valuable feedback on the manuscript. This work was funded by NIGMS grants 5R35GM142679 (E.M.D.) and 5R37GM057260 (A.S.A.).

Author contributions

E.M.D. designed and performed all experimental work and wrote the manuscript; P.V. performed all bioinformatic analyses, contributed extensively to discussions of results and interpretation of findings, and helped revise the manuscript. J.M.A. analyzed BPTF homology, constructed phylogenetic trees, and helped revise the manuscript. A.S.A. contributed significantly to discussions of results, interpretation of findings, and revising of the manuscript.

Funding

This work was funded by NIGMS grants 5R35GM142679 (E.M.D.) and 5R37GM057260 (A.S.A.).

Data availability

The Bioproject number for all new genomic sequencing data reported here is: PRJNA1093065. Reanalyzed H3K4me3 ChIP-seq data from set1(RNAi) and mll1/2(RNAi) stem cells (and their matching control stem cells) was previously published and can be found under GEO accession number GSE74153. Additionally, code used in the analyses performed here can be found at: https://github.com/PrinceVerma-23/BPTF_analysis.

Declarations

Ethical approval

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 19 June 2024 / Accepted: 25 February 2025

Published online: 11 March 2025

References

1. Elliott SA, Alvarado AS. Planarians and the history of animal regeneration: paradigm shifts and key concepts in biology. *Methods Mol Biol.* 2018;1774:207–39.
2. Eisenhoffer GT, Kang H, Sanchez Alvarado A. Molecular analysis of stem cells and their descendants during cell turnover and regeneration in the planarian *Schmidtea mediterranea*. *Cell Stem Cell.* 2008;3(3):327–39.
3. Swapna LS, Molinaro AM, Lindsay-Mosher N, Pearson BJ, Parkinson J. Comparative transcriptomic analyses and single-cell RNA sequencing of the freshwater planarian *Schmidtea mediterranea* identify major cell types and pathway conservation. *Genome Biol.* 2018;19(1):124.
4. Garcia-Castro H, Kenny NJ, Iglesias M, Alvarez-Campos P, Mason V, Elek A, et al. ACME dissociation: a versatile cell fixation-dissociation method for single-cell transcriptomics. *Genome Biol.* 2021;22(1):89.
5. Zeng A, Li H, Guo L, Gao X, McKinney S, Wang Y, et al. Prospectively isolated Tetraspanin(+) neoblasts are adult pluripotent stem cells underlying Planaria regeneration. *Cell.* 2018;173(7):1593–608. e20.
6. van Wolfswinkel JC, Wagner DE, Reddien PW. Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell.* 2014;15(3):326–39.
7. Benham-Pyle BW, Brewster CE, Kent AM, Mann FG Jr., Chen S, Scott AR, et al. Identification of rare, transient post-mitotic cell states that are induced by injury and required for whole-body regeneration in *Schmidtea mediterranea*. *Nat Cell Biol.* 2021;23(9):939–52.
8. Fincher CT, Wurtzel O, de Hoog T, Kravarik KM, Reddien PW. Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*. *Science.* 2018;360:6391.
9. Plass M, Solana J, Wolf FA, Ayoub S, Misios A, Glazar P et al. Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science.* 2018;360(6391).
10. Adler CE, Seidler CW, McKinney SA, Sanchez Alvarado A. Selective amputation of the pharynx identifies a FoxA-dependent regeneration program in planaria. *Elife.* 2014;3:e02238.
11. Cowles MW, Omuro KC, Stanley BN, Quintanilla CG, Zayas RM. COE loss-of-function analysis reveals a genetic program underlying maintenance and regeneration of the nervous system in planarians. *PLoS Genet.* 2014;10(10):e1004746.
12. Tu KC, Cheng LC, Lange HTKV, McKinney JJ, Seidler SA. Egr-5 is a post-mitotic regulator of planarian epidermal differentiation. *Elife.* 2015;4:e10501.
13. Aristotle CR. *History of animals*, in 10 books. Translated by Richard Cresswell. London; G. Bell. 1878. ix, 326 p. p.
14. Raz AA, Wurtzel O, Reddien PW. Planarian stem cells specify fate yet retain potency during the cell cycle. *Cell Stem Cell.* 2021;28(7):1307–22. e5.
15. Mohamed Haroon M, Lakshmanan V, Sarkar SR, Lei K, Vemula PK, Palakodeti D. Mitochondrial state determines functionally divergent stem cell population in planaria. *Stem Cell Rep.* 2021;16(5):1302–16.
16. Molinaro AM, Lindsay-Mosher N, Pearson BJ. Identification of TOR-responsive slow-cycling neoblasts in planarians. *EMBO Rep.* 2021;22(3):e50292.
17. Soshnev AA, Josefowicz SZ, Allis CD. Greater than the sum of parts: complexity of the dynamic epigenome. *Mol Cell.* 2016;62(5):681–94.
18. Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol.* 2007;14(11):1025–40.
19. Wang Y, Wysocka J, Perlin JR, Leonelli L, Allis CD, Coonrod SA. Linking covalent histone modifications to epigenetics: the rigidity and plasticity of the marks. *Cold Spring Harb Symp Quant Biol.* 2004;69:161–9.
20. Stelman CR, Smith BM, Chandra B, Roberts-Galbraith RH. CBP/p300 homologs CBP2 and CBP3 play distinct roles in planarian stem cell function. *Dev Biol.* 2021;473:130–43.

21. Onal P, Grun D, Adamidi C, Rybak A, Solana J, Mastrobuoni G, et al. Gene expression of pluripotency determinants is conserved between mammalian and planarian stem cells. *EMBO J*. 2012;31(12):2755–69.
22. Scimone ML, Meisel J, Reddien PW. The Mi-2-like Smed-CHD4 gene is required for stem cell differentiation in the planarian *Schmidtea mediterranea*. *Development*. 2010;137(8):1231–41.
23. Hubert A, Henderson JM, Ross KG, Cowles MW, Torres J, Zayas RM. Epigenetic regulation of planarian stem cells by the SET1/MLL family of histone methyltransferases. *Epigenetics*. 2013;8(1):79–91.
24. Wagner DE, Ho JJ, Reddien PW. Genetic regulators of a pluripotent adult stem cell system in planarians identified by RNAi and clonal analysis. *Cell Stem Cell*. 2012;10(3):299–311.
25. Mihaylova Y, Abnave P, Kao D, Hughes S, Lai A, Jaber-Hijazi F, et al. Conservation of epigenetic regulation by the MLL3/4 tumour suppressor in planarian pluripotent stem cells. *Nat Commun*. 2018;9(1):3633.
26. Poulet A, Kratkiewicz AJ, Li D, van Wolfswinkel JC. Chromatin analysis of adult pluripotent stem cells reveals a unique stemness maintenance strategy. *Sci Adv*. 2023;9(40):eadh4887.
27. Allen JM, Balagtas M, Barajas E, Cano Macip C, Alvarez Zepeda S, Iberkleid I, et al. RNAi screen of RING/U-Box domain ubiquitin ligases identifies critical regulators of tissue regeneration in planarians. *Front Cell Dev Biol*. 2021;9:803419.
28. Wiggans M, Zhu SJ, Molinaro AM, Pearson BJ. The BAF chromatin remodeling complex licenses planarian stem cells access to ectodermal and mesodermal cell fates. *BMC Biol*. 2023;21(1):227.
29. Duncan EM, Chitsazan AD, Seidel CW, Sanchez Alvarado A. Set1 and MLL1/2 target distinct sets of functionally different genomic loci in vivo. *Cell Rep*. 2015;13(12):2741–55.
30. Pascual-Carreras E, Marin-Barba M, Castillo-Lara S, Coronel-Cordoba P, Magri MS, Wheeler GN, et al. Wnt/beta-catenin signalling is required for pole-specific chromatin remodeling during planarian regeneration. *Nat Commun*. 2023;14(1):298.
31. Ivankovic M, Brand JN, Pandolfini L, Brown T, Pippel M, Rozanski A, et al. A comparative analysis of planarian genomes reveals regulatory conservation in the face of rapid structural divergence. *Nat Commun*. 2024;15(1):8215.
32. Dattani A, Kao D, Mihaylova Y, Abnave P, Hughes S, Lai A, et al. Epigenetic analyses of planarian stem cells demonstrate conservation of bivalent histone modifications in animal stem cells. *Genome Res*. 2018;28(10):1543–54.
33. Xiao H, Sandaltzopoulos R, Wang HM, Hamiche A, Ranallo R, Lee KM, et al. Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Mol Cell*. 2001;8(3):531–43.
34. Hamiche A, Sandaltzopoulos R, Gdula DA, Wu C. ATP-dependent histone octamer sliding mediated by the chromatin remodeling complex NURF. *Cell*. 1999;97(7):833–42.
35. Tsukiyama T, Wu C. Purification and properties of an ATP-dependent nucleosome remodeling factor. *Cell*. 1995;83(6):1011–20.
36. Li H, Ilin S, Wang W, Duncan EM, Wysocka J, Allis CD, et al. Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature*. 2006;442(7098):91–5.
37. Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, et al. A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodeling. *Nature*. 2006;442(7098):86–90.
38. Frey WD, Chaudhry A, Slepicka PF, Ouellette AM, Kirberger SE, Pomerantz WCK, et al. BPTF maintains chromatin accessibility and the Self-Renewal capacity of mammary gland stem cells. *Stem Cell Rep*. 2017;9(1):23–31.
39. Xu B, Cai L, Butler JM, Chen D, Lu X, Allison DF, et al. The chromatin remodeler BPTF activates a stemness Gene-Expression program essential for the maintenance of adult hematopoietic stem cells. *Stem Cell Rep*. 2018;10(3):675–83.
40. Landry J, Sharov AA, Piao Y, Sharova LV, Xiao H, Southon E, et al. Essential role of chromatin remodeling protein Bptf in early mouse embryos and embryonic stem cells. *PLoS Genet*. 2008;4(10):e1000241.
41. UniProt C. UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res*. 2023;51(D1):D523–31.
42. Howe KL, Bolt BJ, Shafie M, Kersey P, Berriman M. WormBase ParaSite - a comprehensive resource for helminth genomics. *Mol Biochem Parasitol*. 2017;215:2–10.
43. Bork P, Dandekar T, Diaz-Lazcoz Y, Eisenhaber F, Huynen M, Yuan Y. Predicting function: from genes to genomes and back. *J Mol Biol*. 1998;283(4):707–25.
44. Yates AD, Allen J, Amode RM, Azov AG, Barba M, Becerra A, et al. Ensembl genomes 2022: an expanding genome resource for non-vertebrates. *Nucleic Acids Res*. 2022;50(D1):D996–1003.
45. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. *Nucleic Acids Res*. 2015;43(Database issue):D257–60.
46. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 2013;30(4):772–80.
47. Criscuolo A, Gribaldo S. BMGE (Block mapping and gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol*. 2010;10:210.
48. Lemoine F, Correia D, Lefort V, Doppelt-Azeroual O, Mareuil F, Cohen-Boulakia S, et al. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res*. 2019;47(W1):W260–5.
49. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 2010;59(3):307–21.
50. Letunic I, Bork P. Interactive tree of life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Res*. 2024;52(W1):W78–82.
51. Adler CE, Alvarado AS. Systemic RNA interference in planarians by feeding of DsRNA containing Bacteria. *Methods Mol Biol*. 2018;1774:445–54.
52. Kamath RS, Martinez-Campos M, Zipperlen P, Fraser AG, Ahringer J. Effectiveness of specific RNA-mediated interference through ingested double-stranded RNA in *Caenorhabditis elegans*. *Genome Biol*. 2001;2(1):RESEARCH0002.
53. Reddien PW, Bermange AL, Murfitt KJ, Jennings JR, Sanchez Alvarado A. Identification of genes needed for regeneration, stem cell function, and tissue homeostasis by systematic gene perturbation in planaria. *Dev Cell*. 2005;8(5):635–49.
54. Newmark PA, Sánchez Alvarado A. Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Dev Biol*. 2000;220(2):142–53.
55. Reddien PW, Oviedo NJ, Jennings JR, Jenkin JC, Sanchez Alvarado A. SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science*. 2005;310(5752):1327–30.
56. Pearson BJ, Eisenhoffer GT, Gurley KA, Rink JC, Miller DE, Sánchez Alvarado A. Formaldehyde-based whole-mount in situ hybridization method for planarians. *Dev Dynamics: Official Publication Am Association Anatomists*. 2009;238(2):443–50.
57. King RS, Newmark PA. In situ hybridization protocol for enhanced detection of gene expression in the planarian *Schmidtea mediterranea*. *BMC Dev Biol*. 2013;13:8.
58. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9(7):676–82.
59. Hayashi T, Asami M, Higuchi S, Shibata N, Agata K. Isolation of planarian X-ray-sensitive stem cells by fluorescence-activated cell sorting. *Dev Growth Differ*. 2006;48(6):371–80.
60. Buenrostro JD, Wu B, Chang HY, Greenleaf WJ. ATAC-seq: A method for assaying chromatin accessibility Genome-Wide. *Curr Protoc Mol Biol*. 2015;109:21.9 1–99.
61. Andrews S. FastQC: a quality control tool for high throughput sequence data. Available online at: 2010 [Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>]
62. Ou JH, Liu HB, Yu J, Kelliher MA, Castilla LH, Lawson ND et al. ATACseqQC: a bioconductor package for post-alignment quality assessment of ATAC-seq data. *BMC Genomics*. 2018;19.
63. Langdon WB. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. *BioData Min*. 2015;8(1):1.
64. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and samtools. *Bioinformatics*. 2009;25(16):2078–9.
65. Reske JJ, Wilson MR, Chandler RL. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. *Epigenetics Chromatin*. 2020;13(1):22.
66. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9(9):R137.
67. Quinlan AR. BEDTools: the Swiss-Army tool for genome feature analysis. *Curr Protoc Bioinf*. 2014;47(11 2):1–34.
68. Wang Q, Li M, Wu T, Zhan L, Li L, Chen M, et al. Exploring epigenomic datasets by chipseeker. *Curr Protoc*. 2022;2(10):e585.
69. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for chip peak annotation, comparison and visualization. *Bioinformatics*. 2015;31(14):2382–3.

70. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. DeepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* 2014;42(Web Server issue):W187–91.
71. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011;29(1):24–6.
72. Lun ATL, Smyth GK. Cseq: a bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.* 2016;44(5).
73. Wickham H. ggplot2: elegant graphics for data analysis. Springer, New York; 2016. <https://ggplot2.tidyverse.org>.
74. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods.* 2015;12(4):357–60.
75. Love MI, Huber W, Anders S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12).
76. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* 2015;33(3):290–.
77. Lee TI, Johnstone SE, Young RA. Chromatin Immunoprecipitation and microarray-based analysis of protein location. *Nat Protoc.* 2006;1(2):729–48.
78. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nat Methods.* 2012;9(4):357–U54.
79. Shen L, Shao NY, Liu X, Maze I, Feng J, Nestler EJ. DiffReps: detecting differential chromatin modification sites from ChIP-seq data with biological replicates. *PLoS ONE.* 2013;8(6):e65598.
80. Wu TZ, Hu EQ, Xu SB, Chen MJ, Guo PF, Dai ZH et al. ClusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation-Amsterdam.* 2021;2(3).
81. Yu GC, Wang LG, Han YY, He QY. ClusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* 2012;16(5):284–7.
82. Li H, Fischle W, Wang W, Duncan EM, Liang L, Murakami-Ishibe S, et al. Structural basis for lower lysine methylation state-specific readout by MBT repeats of L3MBTL1 and an engineered PHD finger. *Mol Cell.* 2007;28(4):677–91.
83. Duncan EM, Allis CD. Errors in erasure: links between histone lysine methylation removal and disease. *Prog Drug Res.* 2011;67:69–90.
84. Lei K, Thi-Kim Vu H, Mohan RD, McKinney SA, Seidel CW, Alexander R, et al. Egf signaling directs neoblast repopulation by regulating asymmetric cell division in planarians. *Dev Cell.* 2016;38(4):413–29.
85. Labbe RM, Irimia M, Currie KW, Lin A, Zhu SJ, Brown DD, et al. A comparative transcriptomic analysis reveals conserved features of stem cell pluripotency in planarians and mammals. *Stem Cells.* 2012;30(8):1734–45.
86. Verma P, Waterbury CKM, Duncan EM. Set1 targets genes with essential identity and Tumor-Suppressing functions in planarian stem cells. *Genes (Basel).* 2021;12(8).
87. Lun AT, Smyth GK. Cseq: a bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res.* 2016;44(5):e45.
88. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. *Cell.* 2007;129(4):823–37.
89. Kharchenko PV, Alekseyenko AA, Schwartz YB, Minoda A, Riddle NC, Ernst J, et al. Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature.* 2011;471(7339):480–5.
90. Li JJ, Huang H, Bickel PJ, Brenner SE. Comparison of *D. melanogaster* and *C. elegans* developmental stages, tissues, and cells by ModENCODE RNA-seq data. *Genome Res.* 2014;24(7):1086–101.
91. Mulas C, Kalkan T, von Meyenn F, Leitch HG, Nichols J, Smith A. Defined conditions for propagation and manipulation of mouse embryonic stem cells. *Development.* 2019;146(6).
92. Choi J, Huebner AJ, Clement K, Walsh RM, Savol A, Lin K, et al. Prolonged Mek1/2 suppression impairs the developmental potential of embryonic stem cells. *Nature.* 2017;548(7666):219–23.
93. Badenhorst P, Voas M, Rebay I, Wu C. Biological functions of the ISWI chromatin remodeling complex NURF. *Genes Dev.* 2002;16(24):3186–98.
94. Yu Q, Liu X, Fang J, Wu H, Guo C, Zhang W, et al. Dynamics and regulation of mitotic chromatin accessibility bookmarking at single-cell resolution. *Sci Adv.* 2023;9(4):eadd2175.
95. Kiani K, Sanford EM, Goyal Y, Raj A. Changes in chromatin accessibility are not concordant with transcriptional changes for single-factor perturbations. *Mol Syst Biol.* 2022;18(9):e10979.
96. Guo L, Bloom JS, Dols-Serrate D, Boocock J, Ben-David E, Schubert OT, et al. Island-specific evolution of a sex-primed autosome in a sexual planarian. *Nature.* 2022;606(7913):329–34.
97. Stankiewicz P, Khan TN, Szafranski P, Slattery L, Streff H, Vetrini F, et al. Haploinsufficiency of the chromatin remodeler BPTF causes syndromic developmental and speech delay, postnatal microcephaly, and dysmorphic features. *Am J Hum Genet.* 2017;101(4):503–15.
98. Grinton KE, Hurst ACE, Bowling KM, Cristian I, Haynes D, Adstamomkonkul D, et al. Phenotypic expansion of the BPTF-related neurodevelopmental disorder with dysmorphic facies and distal limb anomalies. *Am J Med Genet A.* 2021;185(5):1366–78.
99. Gong YC, Liu DC, Li XP, Dai SP. BPTF biomarker correlates with poor survival in human NSCLC. *Eur Rev Med Pharmacol Sci.* 2017;21(1):102–7.
100. Pan Y, Yuan F, Li Y, Wang G, Lin Z, Chen L. Bromodomain PHD-finger transcription factor promotes glioma progression and indicates poor prognosis. *Oncol Rep.* 2019;41(1):246–56.
101. Miao J, Zhang M, Huang X, Xu L, Tang R, Wang H, et al. Upregulation of bromodomain PHD finger transcription factor in ovarian cancer and its critical role for cancer cell proliferation and survival. *Biochem Cell Biol.* 2021;99(3):304–12.

Publisher's note

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