Global Gene Expression Regulation Mediated by TGFβ Through H3K9me3 Mark

Cancer Informatics
Volume 21: 1–17
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11769351221115135

Ankit Naik, Nidhi Dalpatraj and Noopur Thakur

Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Ahmedabad, Gujarat, India.

ABSTRACT

BACKGROUND: Epigenetic alterations play an important part in carcinogenesis. Different biological responses, including cell proliferation, migration, apoptosis, invasion, and senescence, are affected by epigenetic alterations in cancer. In addition, growth factors, such as transforming growth factor beta (TGFβ) are important regulators of tumorigenesis. Our understanding of the interplay between the epigenetic bases of tumorigenesis and growth factor signaling in tumorigenesis is rudimentary. Some studies suggest a link between TGFβ signaling and the heterochromatinizing histone mark H3K9me3. There is evidence for signal-dependent interactions between R-Smads and histone methyltransferases. However, the effects of TGFβ signaling on genome wide H3K9me3 landscape remains unknown. Our research examines TGFβ -induced genome-wide H3K9me3 in prostate cancer.

METHOD: Chromatin-Immunoprecipitation followed by sequencing was performed to analyze genome-wide association of H3K9me3 epigenetic mark. DAVID Functional annotation tool was utilized to understand the involvement of different Biological Processes and Molecular Function. MEME-ChIP tool was also used to analyze known and novel DNA-binding motifs.

RESULTS: H3K9me3 occupancy appears to increase at intronic regions after short-term (6 hours) TGFβ stimulation and at distal intergenic regions during long-term stimulation (24 hours). We also found evidence for a possible association of SLC transporters with H3K9me3 mark in presence of TGFβ during tumorigenesis. No direct correlation was found between the occupancy of H3K9me3 mark and the expression of various genes. The epigenetic mechanisms-mediated regulation of gene expression by TGFβ was concentrated at promoters rich in SRY and FOXJ3 binding sites.

CONCLUSION: Our results point toward a positive association of oncogenic function of TGFβ and the H3K9me3 mark and provide a context to the role of H3K9me3 in TGFβ-induced cell migration and cell adhesion. Interestingly, these functions of TGFβ through H3K9me3 mark regulation seem to depend on transcriptional activation in contrast to the conventionally known repressive nature of H3K9me3.

KEYWORDS: Prostate cancer, H3K9me3, ChIP-sequencing, TGFβ, epigenetics

RECEIVED: March 10, 2022. **ACCEPTED:** July 2, 2022.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by funding from the Science and Engineering Research Board (SERB), Government of India (EMR/2017/001987) and from Ahmedabad University Start-Up Grant (URB-SG7518) to Dr. Noopur Thakur.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Noopur Thakur, Biological and Life Sciences, School of Arts and Sciences, Ahmedabad University, Commerce Six Roads, Navrangpura, Ahmedabad, Gujarat 380009, India. Email: noopur.thakur@ahduni.edu.in

Introduction

Prostate cancer is the major cause of death in males, and usually it advances with age. However, the incidence of prostate cancer in young men is increasing, posing a higher risk of metastasis and increased mortality in an increasingly young population.[1] Overexpression of transforming growth factor-β (TGFβ) depicts the poor outcome for cancer patients in many cancer types. TGFβ is a multifunctional cytokine involved in many cellular processes including cell growth, cell differentiation, cell proliferation, cellular homeostasis, apoptosis, and other cellular functions. The TGFβ superfamily of ligands includes bone morphogenetic proteins (BMPs), GDFs, AMH, Activin, Nodal, and TGFβs. In the early stages of cancer, TGFβ suppresses the tumor growth while at the later stages of the tumor development, it leads to the metastasis of tumor cells. TGFβ signals through the canonical (small mothers against decapentaplegic (Smad) signaling pathway) and non-canonical pathways (non-Smad pathways).[2] In the canonical TGFβ signaling pathway, TGFβ ligand binds to the TGFβR2 which catalyzes the phosphorylation of the TGFβ receptor 1 (TGFβR1) forming a hetero-tetrameric complex with the ligand. The activated TGFβR1 phosphorylates R-Smads (Smad2 and Smad3), which forms complexes with co-Smad (Smad4). The R-Smad/co-Smad complex enters the nucleus where it binds to the promoters of genes and regulates their expression. The non-canonical/non-Smad pathways include various types of MAP kinase pathways, phosphatidylinositol-3-kinase (PI3)/AKT and Rho-like GTPase signaling pathways.[3]

It is now well understood that epigenetic modifications like DNA methylation, histone acetylation, histone methylation, and others play an important role in carcinogenesis. Epigenetic modifications get altered in cancer cells and can control various cellular processes like cell migration, proliferation, apoptosis, invasion, and senescence.[4]

H3K9me3 is found primarily at the repeat regions of the genome which constitute heterochromatic regions of mammalian genomes. H3K9me3 is mainly enhanced at the major and minor satellite repeats regions of the pericentromeric heterochromatin and also at various specific DNA transposons.[5] H3K9me3 has a unique feature among the "repressive" histone

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

trimethyl lysine modifications as unlike other modifications, it can be distributed at both silent heterochromatin and at active euchromatin genes. H3K9 methylation is not the sole marker of heterochromatin but also of the transcribed regions of mammalian euchromatin.[6] The simultaneous trimethylation of H3K4 and H3K9 occurs in the transcribed region of active genes in mammalian chromatin.[6,7]

Despite the H3K9me3 functional role in cancer, it also plays an important role in normal cellular development. It acts as a repressor of various inappropriate genes which should be expressed in a lineage specific manner and maintains cellular integrity and stability of the genome. H3K9me3 is an effective covalent post-translational histone modification, and its deregulation causes several diseases. The roles for H3K9me3 have been discovered in regulating apoptosis,[8] autophagy,[9] development,[10] DNA repair,[11] splicing,[12] self-renewal,[13] transcriptional elongation,[6,13] viral latency,[6,13,14] imprinting,[15] aging,[16] and cell identity.[17] Global reprograming of epigenetic modifications has been shown in the AML12 cells undergoing epithelial-mesenchymal transition (EMT). Global reduction in the H3K9me2, increase in the H3K4me3, and increase in H3K36me3 were found in the AML12 undergoing EMT in response to TGFβ.[18]

The robust nature of H3K9me3 is due to the number of enzymes that are involved in its regulation. A tightly coordinated collection of readers, writers, and erasers of this histone mark is required to establish and maintain a transcriptional landscape; however, the small changes can also lead to various diseases. The earliest methyltransferases identified were SUV39H1 and SUV39H2. The second family of methyltransferases identified were SET domain, bifurcated 1 (SETDB1) and SET domain, bifurcated 2 (SETDB2)[18] and they specifically target H3K9. SETDB1 methyltransferase methylates and stabilizes the oncogenic p53 mutants, leading to the growth of the cancer cells. Along with the lysine specific methylases, lysine-specific demethylase 4 (KDM4) JmjC-domain-containing demethylase family of proteins were also identified, whose catalytic activity can reverse the H3K9me3 mark.

Very few studies suggest the crosstalk between the TGFβ signaling and H3K9me3. In the TGFβ signaling, Receptor SMADS (R-Smads) has been shown to interact with histone methyltransferase. SMAD1/5 interacts with H3K9 methyltransferase SUV39H thus increasing the levels of H3K9me3, leading to the silencing of target genes by recruiting NURD complexes. The increased SMAD2/3 occupancy at certain loci coincides with reduced H3K27me3 and recruitment of KDM6B to these promoters upon Activin stimulation. This indicates that the expression of developmental genes are regulated by R-SMAD through interfering with the chromatin structure.[19] In mESCs and human bone marrow mesenchymal stem/stromal cells, BMP signaling can induce the expression of KDM6B, which in turn regulates the genes specific for various developmental processes.[20] It has also been demonstrated that the H3K9me3 mark can also be used by the nodal effectors to convert the bivalent state of the regulators of stem cell differentiation to the active state.[21]

The functional role of H3K9me3 in prostate cancer is not well understood. PHF8 demethylase has been demonstrated to erase the H3K9 methylation mark on the promoter of integrin and Rho-Associated Protein Kinase 1 (ROCK) genes in prostate cancer cells. The removal of H3K9me3 on the promoter of these genes induces their expression and promotes migration and invasion, thus being correlated with poor prognosis.[22] Few other studies suggest that Histone methyltransferase SETDB1 is necessary for the proliferation, migration and invasion of prostate cancer cells[23] and Metformin inhibits the migration of prostate cancer cells mediated by SUV39H1.[24] Our study provides a genome wide analysis of TGFβ induced H3K9me3 in prostate cancer.

Material and Methods

Maintenance and culture of prostate cancer cells

Prostate cancer cells (PC3) were cultivated in RPMI-1640 medium with 10% Fetal Bovine Serum (FBS) (Gibco, USA) and 1% antimycotic and antibiotic mixture (HiMedia Labs, India) in the presence of 5% CO_2 at 37°C.

TGFβ treatment

PC3 cells were cultured in tissue culture flask until they reached 70% to 80% confluency. The cells were then serum-starved in serum-free medium for 24 hours. TGFβ (Abcam, USA) stimulation was given to the cells with a final concentration of 5 ng/mL after serum starvation. TGFβ stimulation was given for 0 (Control), 6, and 24 hours.

qRT-PCR

PC3 cells were grown in T-75 flasks until they reached 70% to 80% confluency. The cells were subjected to 24 hours of serum starvation by incubating them in serum-free medium for 24 hours. The cells were stimulated with TGFβ (Abcam, USA) at a final concentration of 5 ng/mL post serum starvation. TGFβ stimulation was given for a duration of 0 (Control), 6, and 24 hours. Total RNA was purified as per the manufacturer's instructions using the RNeasy Mini Kit (Qiagen, Germany). The RNAse free DNAse set (Qiagen, Germany) was used to extract DNA from the samples. With the help of the iScript™ cDNA Synthesis Kit (Bio-Rad, USA), the RNA was reverse transcribed to complementary DNA (cDNA) in a thermal cycler (Eppendorf Mastercycler). Primers (Sigma-Aldrich, USA) specific for XRN1, OR2W5, RAB24, HERC2P4, BGLT3, and GAPDH with the following sequences were used: XRN1: F, 5′-CACAGACAATGACCGTTTGC-3′ and R, 5′-AAAGGCTGGAGGAATGGTTC-3′, OR2W5:

F, 5′-ATCTCCTATGGCGTGATTGC-3′ and R, 5′-TGATGCTGGGAATGACGATG-3′, RAB24: F, 5′-GTGATGACAGAGGACAAGGG-3′ and R, 5′-GACTACCCAAGCCCAGAAAG-3′, HERC2P4: F, 5′-GCACGATGAGTTTGGAGAAGGC-3′ and R, 5′-CTGAGCACAAGTTCACCTACTGG-3′, BGLT3: F, 5′-GTGTGCCCTGTCTATTCCTG-3′ and R, 5′-AGAAGTGTGCTGCTCTGTTC-3′, GAPDH: F, 5′-ATGTTCGTCATGGGTGTGAA-3′ and R, 5′-TGTGGTCATGAGTCCTTCCA-3′. cDNA was quantified in QuantStudio™ 5 Real-Time PCR system (Applied Biosystems, USA) using iTaq Universal SYBR Green Supermix (Bio-Rad, USA) as per manufacturer's instruction. With GAPDH as the reference gene, data was analyzed using QuantStudio™ Design and Analysis Software v1.5.1.

Chromatin immunoprecipitation (ChIP)

Histone proteins were crosslinked to DNA at a final concentration of 1% using a 37% formaldehyde solution (Sigma-Aldrich, USA), and cells were kept for 10 minutes at 37°C in a CO_2 Incubator. The cross-linking reaction was stopped by quenching the formaldehyde with glycine at a final concentration of 125 mM and incubation at room temperature with shaking. After quenching, cells were washed in ice-cold phosphate buffered saline (PBS) (0.137 M NaCl, 0.0027 M KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) containing Protease Inhibitor Cocktail (Sigma-Aldrich, USA), scraped, and pelleted by centrifugation. Liquid-nitrogen was used to flash-freeze the pellets, which were then stored at −80°C until needed.

The cell pellet was resuspended in sodium lauryl sulfate (SLS) lysis buffer with protease inhibitors (1% SLS, 0.01 M EDTA, and 0.05 M Tris-Cl, pH 8.1) and incubated for 10 minutes on ice. Bioruptor Pico (Diagenode, USA) was used to sonicate the lysate. The supernatant was recovered after centrifuging the samples. The supernatant was diluted 10-times in ChIP-Dilution Buffer (0.01% SLS, 1.1% Triton X-100, 0.0012 M EDTA, 0.0167 M Tris-Cl, 0.167 M NaCl, pH 8.1). Immunoprecipitation antibody specific for H3K9me3 (SantaCruz, USA) was added, followed by Protein A Sepharose beads (Abcam, USA), and incubated at 4°C overnight. For each sample, a no-antibody sample was stored aside to be used as Input-DNA. The Histone-DNA-Antibody complex was washed with Low Salt Immune Complex Wash Buffer (0.1% SLS, 1% Triton X-100, 0.002 M EDTA, 0.020 M Tris-Cl, pH 8.1, 0.15 M NaCl), High Salt Immune Complex Wash Buffer (0.1% SLS, 1% Triton X-100, 0.002 M EDTA, 0.02 M Tris-Cl, pH 8.1, 0.5 M NaCl), LiCl Immune Complex Wash Buffer (250 mM LiCl, 1% IGEPAL CA630, 1% deoxycholic acid (Na-salt), 0.001 M EDTA, 0.01 M Tris-Cl, pH 8.1), and 1X TE buffer (0.01 M Tris-Cl, pH 8.0, 0.001 M EDTA) in that respective order. By vortexing and incubating at room temperature for 15 minutes with rotation, the DNA was eluted using an elution buffer (1% SLS, 100 mM $NaHCO_3$). The supernatant (eluate) was obtained after centrifugation of the mixture. Reverse crosslinking was performed using 5 M NaCl, which was heated at 65°C for 4 hours before being treated with RNase overnight at 60°C and Proteinase K for 1 hour at 45°C. A DNA purification kit was used to retrieve and purify the DNA.

Library construction and sequencing:

The NEBNext® Ultra™ II DNA Library Prep Kit was used to create libraries for ChIP-sequencing. In a nutshell, ChIP-DNA and Input-DNA were treated to a series of enzymatic processes that included end repair, tailing using dA-tail, and adapter sequence ligation. SPRI beads were used to size-select these adapter-ligated fragments. Following that, the size selected DNA fragments were indexed using limited cycle PCR to produce final libraries for paired-end sequencing. Before being sequenced on the Illumina HiSeq X system to yield 2 × 150 bp sequencing reads, the libraries were quantified. On the Illumina HiSeq X, prepared libraries were sequenced to obtain 100 M, 2 × 150 bp reads per sample. A FASTQ file was created from the sequenced data.

Data analysis

Assessing Read Quality: We analyzed the following specifications from the FASTQ file: (a) Per Base Sequence quality, (b) Per Sequence quality score distribution, (c) Per Base Sequence content, (d) Frequency distribution of GC content in the reads, (e) PCR amplification errors, and (f) overrepresented sequences.

Trimming of adapter sequences and low-quality reads: Based on quality reports of FASTQ files, sequence reads were trimmed to retain only high-quality sequence for downstream analysis. Further, the low-quality reads were removed from the analysis. The trimming of adapter sequences and read filtering was done with the use of Trimmomatic (*Ver-0.36*)[25] using default settings.

Alignment of Reads to Reference Genome: The filtered paired-end reads were aligned to the Human Reference genome (hg19) Feb. 2009 release downloaded from UCSC database (GRCh37/hg19). The sequence mapping to the reference genome was accomplished using BWA MEM using default settings and the chromosome FASTA file was obtained from the following URL (http://hgdownload.soe.ucsc.edu/goldenPath/hg19/bigZips/chromFa.tar.gz) (Ver-0.7.12).

Peak Calling and annotation: Model-based Analysis of ChIP-Seq (MACS 2.1.3 version)'s callpeak function was utilized for peak calling[26] using default settings. ChIPseeker, an R package[27] was utilized for the annotation of the ChIP peaks obtained from the MACS2 callpeak function using default settings.

Functional annotation: Database for Annotation, Visualization, and Integrated Discovery (DAVID) v6.8 was

used to analyze the enriched biological themes, particularly GO terms, discover enriched functional-related gene groups and cluster redundant annotation terms.

Functional analysis of gene lists

Annotated Genes which were present in H3K9me3-6h and H3K9me3-24h but absent in H3K9me3-Control were used for the DAVID Functional Analysis. The gene list was imported to the David Analysis tool. "OFFICIAL_GENE_SYMBOL" was used as the Identifier with List type set to "Gene List." Default parameters on the DAVID platform were used to analyze the list. From the DAVID Functional annotation tools, a Functional annotation chart was downloaded and analyzed for enriched GO terms.[28,29]

MEME-ChIP motif analysis

250 bp sequences flanking the peak summits on both sides in the Promoter Region were retrieved. We brought the sequences into the MEME-ChIP (v5.3.3).[30] Motif Discovery was accomplished using the Discriminative Mode by using Sequences from TGFβ stimulated samples as Primary sequences and sequences from Control sample as Control sequences in the MEME-ChIP using default parameters. HOCOMOCO Human(v11)[30,31] was used as a Known Motif Database. Motifs that came up in H3K9me3-6h and H3K9me3-24h through the MEME-ChIP were compared against the HOCOMOCO Human(v11) motifs database using Tomtom using default settings.[32]

Dataset availability

The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus[33] and are accessible through GEO Series accession number GSE191096 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE191096).[34]

Statistical analyses

Statistical Analysis was performed using 2-way ANOVA with Tukey Multiple Correction Test using GraphPad Prism (v9.3.1) (GraphPad Software, USA).

Results

TGFβ treatment validation and outline of ChIP-sequencing -quality scores, alignment, and peak calling

Our approach relied on molecular changes initiated in PC3 cells upon acute TGFβ stimulation. We needed to make sure that TGFβ stimulation was effective in our system first. TGFβ stimulation was applied to the cells for 6 and 24 hours. Following the stimulation, the cells were processed for either ChIP-seq (described later) or gene expression analysis using RT-PCR to establish the effect of TGFβ stimulation on known target genes. The RNA from these cells was purified at time intervals that matched the harvesting of cells for ChIP-Seq. Two known TGFβ target genes, p21 and N-cadherin, were tested using qRT-PCR (Supplemental Figure 1). qRT-PCR results showed that TGFβ stimulation caused an increase in expression of p21, and N-Cadherin as compared to unstimulated controls. According to the qRT-PCR data, TGFβ stimulation was efficient in causing marker gene expression changes as expected.

At the same time points as stated above for RT-PCR, cells were collected for ChIP. Paired-end ChIP-sequencing was performed with and without TGFβ stimulation. ChIP protocol was followed as mentioned in our published article.[35] The Illumina HiSeq × platform was used to sequence the ChIP-DNA samples. The essential details of the ChIP-Seq run and sequencing output are listed in Table 1.

The aligned read sets were subjected to DNA sequence property analyses. There was an interesting drop in the GC content in the aligned reads of the Control (H3K9me3-Control) and TGFβ treated samples, 6 hours (H3K9me3-6h) and 24 hours (H3K9me3-24h). The average GC content of H3K9me3-Control samples was 38.98% (Figure 1A), while the average GC content of H3K9me3-6h and H3K9me3-24h samples was 38.44% (Figure 1B) and 36.71% (Figure 1C), respectively.

The peaks in the aligned readings were called using MACS2 callpeak (Table 1). In the H3K9me3-Control, H3K9me3-6h, and H3K9me3-24h samples, we detected 1334, 1220, and 272 peaks (q-value 0.05), respectively. The influence of TGFβ stimulation was shown in the decrease in the number of peaks. The average peak pile-up for H3K9me3-Control, H3K9me3-6h, and H3K9me3-24h samples was 174, 79.45, and 637.72, respectively (Table 1). With the 3 samples, we acquired good quality ratings and significant alignment, and we used this data for further analysis.

TGFβ stimulation alters H3K9me3 occupancy

We looked at the proportions of total regions named in the Control sample-H3K9me3-Control and TGFβ stimulated samples-H3K9me3-6h and H3K9me3-24h. The MACS2 callpeak function was used to identify and call the total regions (q-value < 0.05) related with H3K9me3, of which 47.2% were associated with Control, 43.2% with H3K9me3-6h, and 9.6% with H3K9me3-24h (Figure 2A). The total number of peaks found across all samples was set to 100%. It highlights the fact that when TGFβ is induced, the number of regions linked with H3K9me3 decreases indicating that TGFβ reduces H3K9me3 mark at a few regions across the genome and in many regions, it completely removes the H3K9me3 mark, thereby reducing the repressive conditions and possibly promoting gene expression.

Table 1. ChIP-sequencing run specifications on the Illumina Hi-Seq X. Paired-end sequencing of the Control and stimulated samples was performed. Alignment was greater than 80% across all samples.

SAMPLE	READ ORIENTATION	MEAN READ QUALITY (PHRED SCORE)	NUMBER OF READS	%GC	% READS WITH Q > 30	NUMBER OF BASES (MB)	MEAN READ LENGTH (BP)	# READS ALIGNED (% ALIGNMENT)	# PEAKS IDENTIFIED (Q-VALUE < 0.05)	AVERAGE PEAK PILEUP	AVERAGE PEAK LENGTH
H3K9me3-Control	R1	39.3	6.2E+07	38.88	95.47	9317.67	151	96 916 815 (80.62)	1334	174	311.5
	R2	38.27	6.2E+07	39.08	91.71	9317.67	151				
H3K9me3-6h	R1	39.14	6.1E+07	38.32	94.91	9158.54	151	97 995 246 (83.99)	1220	79.45	308
	R2	37.71	6.1E+07	38.56	89.72	9158.54	151				
H3K9me3-24h	R1	39.32	6.3E+07	36.62	95.53	9558.54	151	111 936 597 (90.40)	272	637.72	287.45
	R2	38.44	6.3E+07	36.8	92.31	9558.54	151				

The intersecting and non-intersecting zones across the 3 datasets, H3K9me3-Control (a), H3K9me3-6h (b), and H3K9me3-24h (c), are depicted in a Venn diagram (Figure 2B). After comparing the intersecting and unique regions we observed that a major proportion of H3K9me3 occupied regions uniquely belonged to H3K9me3-6h and H3K9me3-Control indicating 2 distinct groups of regions associated with Control conditions and short duration (6 hours) of TGFβ stimulation. The group of regions unique to Control conditions may point toward those biological processes and molecular functions which are activated upon TGFβ stimulation as the repressive H3K9me3 mark is removed. It will be important to evaluate these regions where the loss of the H3K9me3 mark could lead to activation or repression of pathways that promote tumorigenesis.

The group of regions unique to the H3K9me3-6h sample also holds key insights into how H3K9me3 occupancy affects the expression of genes in these regions and the role played by these genes. Since these regions did not show the H3K9me3 mark during control conditions as well as during 24 hours stimulation with TGFβ, they could be associated with acute gene expression responses to TGFβ. These regions need to be studied further to ascertain the role of the H3K9me3 mark and its downstream effects as their expression may be affected by H3K9me3 occupancy since this mark is involved in transcriptional activation and mRNA elongation by RNA pol II in association with HP1γ. Followed by set (a) and (b) in terms of the number of H3K9me3 occupied regions, set (ab) had the most regions (110) marked by H3K9me3. These regions lost the H3K9me3 mark upon longer stimulation (24 hours) of TGFβ which could be potentially linked to activation of pathways associated with TGFβ mediated signaling. Further investigations needs to be conducted in order to understand the effects of the removal of the H3K9me3 mark. There were fewer regions that had the H3K9me3 mark unique to the sample stimulated with TGFβ for 24 hours. This group includes those regions where the H3K9me3 mark was present only after 24 hours of TGFβ stimulation. The subsequent effects on the addition of H3K9me3 mark post 24 hours of TGFβ stimulation need to be studied so the cells can potentially be linked to processes that require a longer duration of active TGFβ signaling. Set (bc) consists of regions containing the H3K9me3 mark exclusive to the treated samples-6 and 24 hours indicating an association of H3K9me3 modification with TGFβ simulation. The list of all genes and set wise sorted list of genes have been mentioned in Supplemental File 1.

The list of H3K9me3 bound regions was produced after peak calling using MACS2 callpeak function from the aligned output sequences. H3K9me3 genomic distribution in response to TGFβ was investigated. The peaks' positions were annotated based on genetic characteristics. The genomic annotations were assigned using the Annotate Peak function from the ChIPseeker package. Peak annotations included promoters,

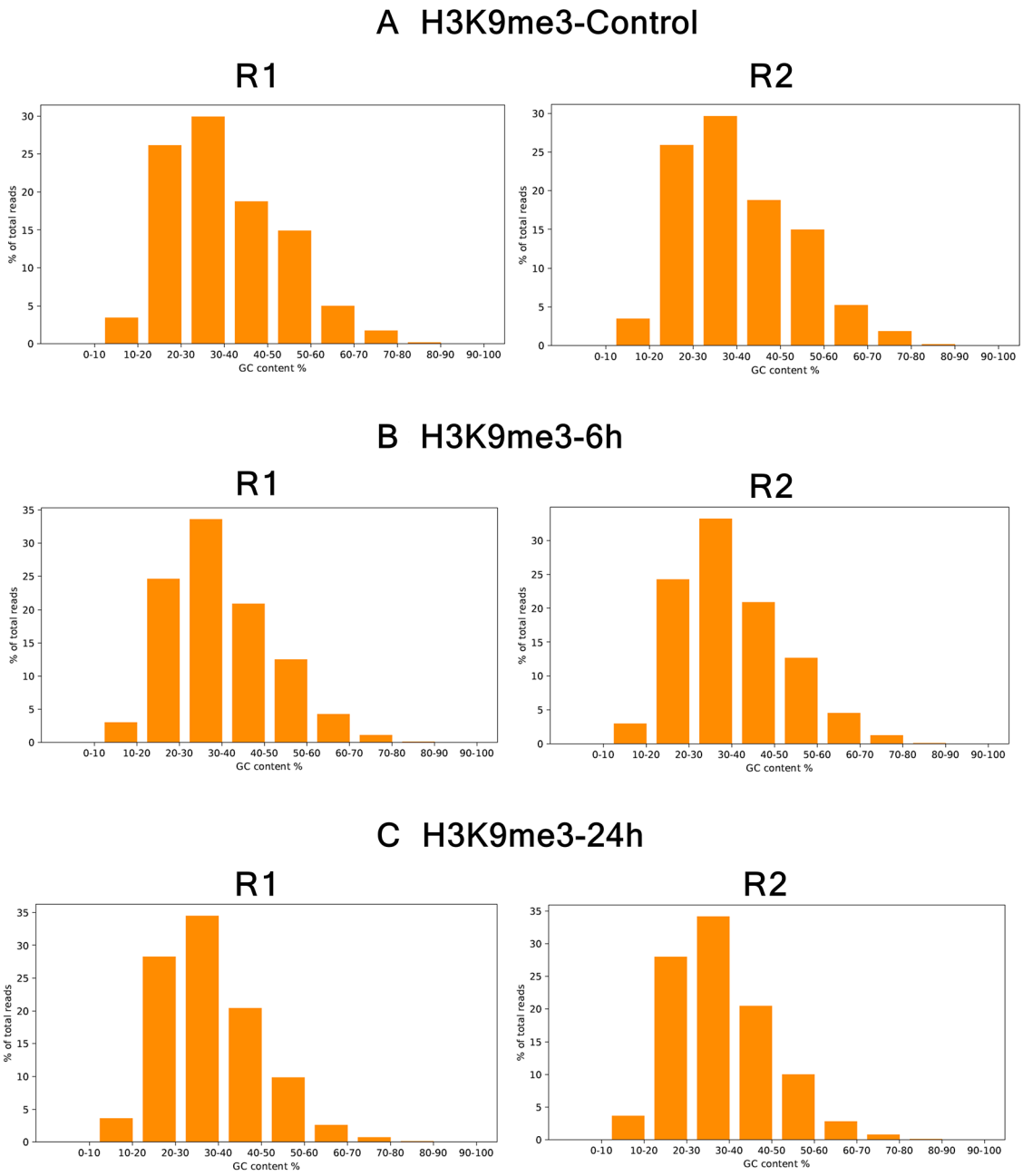

Figure 1. Frequency distribution of GC content. The average percentage of GC content was highest in: (A) H3K9me3-Control (38.98%), followed by (B) H3K9me3-6h (38.44%), and (C) H3K9me3-24h (36.71%).

5′UTRs, 3′UTRs, first introns, other introns, first exons, other exons, and downstream or distal intergenics (Figure 3).

The majority of H3K9me3 occupied regions (45.3%) in the H3K9me3-Control Samples belonged to the Distal intergenic and Intron categories (32.19%). The Promoter Region category accounted for 18.83% of all regions. Similar to the Control samples, the majority of the H3K9me3 occupied regions in the 6 and 24 hours TGFβ stimulated samples H3K9me3-6h and H3K9me3-24h respectively corresponded to the Distal Intergenic region (31.09% and 52.76% respectively) and Intronic regions (43.97% and 26.94% respectively). We saw a small increase in the percentage of areas in the Promoter Category (20.26%) in the H3K9me3-6h sample, but a marginal drop in the H3K9me3-24h sample (17.34%) when compared to the control sample. We observed increased H3K9me3 occupancy at Intronic regions during TGFβ stimulation for 6 hours, which again reduced in 24 hours treated sample. Studying the short-term effects of increased H3K9me3 occupancy around intronic regions during TGFβ stimulation will be interesting. Contrasting effects were observed in terms of H3K9me3 occupancy at Distal Intergenic regions wherein it was highest in 24 hours treated sample whereas 6 hours treated sample had the lowest number of H3K9me3 occupied distal intergenic regions as compared to control. Exonic, 5′UTR and 3′UTR regions showed poor H3K9me3 occupancy in Control as well as treated sample. In conclusion,

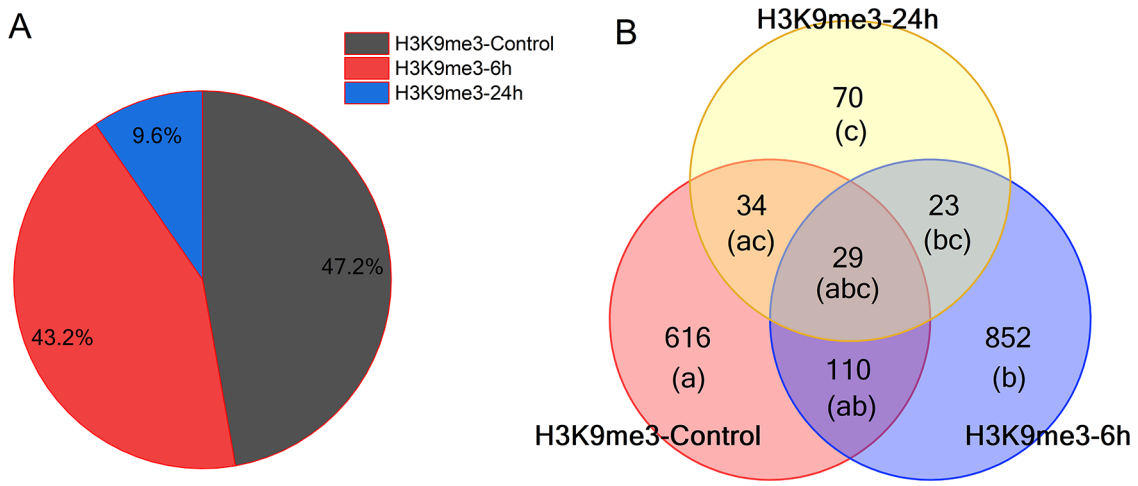

Figure 2. Peaks obtained from MACS2 callpeak function were identified using the ChIPseeker Bioconductor package. (A) Pie-chart distribution of annotated genes obtained after removal of pseudogenes and unidentified transcripts from total regions associated with H3K9me3. (B) Venn diagram of comparison of 3 gene lists—H3K9me3-Control, H3K9me3-6h, and H3K9me3-24h obtained by annotation of MACS2 peaks using ChIPseeker package. The comparative list was created by identifying common and unique genes in MS Excel and using conditional formatting to visualize common and unique regions for each sample data set.

it appears that H3K9me3 occupancy increases at intronic regions during short-term stimulation (6 hours) of TGFβ and it increases at distal intergenic regions during long-term stimulation(24 hours).

Genes associated with short-term (6 hours) stimulation of TGFβ show extensive enrichment in biological processes and molecular function

During 6 hours of TGFβ stimulation, we compiled a list of genes that were selectively enriched. We studied the regions of H3K9me3 occupancy by subjecting the gene list to functional enrichment analysis using DAVID (v6.8). A total of 55 GO categories (Biological Processes) came up in the enrichment analysis out of which 27 were found to be significant (P-value $< .05$) (Figure 4A) (Supplemental File 2). Simultaneous enrichment of GO:0045892~negative regulation of transcription, DNA-templated and GO:0051091~positive regulation of DNA-binding transcription factor activity indicated toward a possible dual role of the H3K9me3 mark, wherein it could be associated with both gene repression and gene activation in a TGFβ-context-dependent manner. The role of plasma membrane transporters in cancer has received a lot of attention in recent years. In cancer, several essential nutrient transporters are up-regulated and act as tumor promoters. Transporters may have the ability to suppress tumor growth. So far, 4 transporters belonging to the SLC gene family have been identified as tumor suppressors. Enrichment of GO:0055085~transmembrane transport along with GO:0006820~anion transport with most genes contributing to these biological processes was found to be a part of the Solute carrier family (SLC) of transporters. Our results establish a potential link among the expression of these transporters and the H3K9me3 mark which could be influenced by TGFβ signaling. Further investigation needs to be carried out to identify the role of SLC transporters in tumorigenesis in coordination with TGFβ signaling and the H3K9me3 mark.

Our study also found H3K9me3 occupancy at genes that play a role in GO:0007155~cell adhesion indicating that process of cell adhesion and the expression of genes associated with the process is altered due to H3K9me3 occupancy under the influence of short-term stimulation with TGFβ. A set of 8 genes were enriched for GO:0046777~protein autophosphorylation indicating an association of TGFβ mediated phosphorylation event through the H3K9me3 mark. Enrichment of GO:0031532~actin cytoskeleton reorganization indicates an alteration in expression of genes playing a role in actin cytoskeleton reorganization by TGFβ mediated marking of certain genes with H3K9me3 modification. We also found exclusive H3K9me3 occupancy during the 6 hours TGFβ stimulation at genes that play a role in directly or indirectly regulating the GTPase activity. The significant enrichment of GO:0043087~regulation of GTPase activity pointed out that short-term TGFβ stimulation may modulate GTPase activity via H3K9me3 mark to indirectly influence DNA methylation activity which can upregulate or downregulate gene expression depending on the context. A total of 16 GO: Molecular Functions were enriched out of which 5 were found to be significantly enriched (Figure 4B) (Supplemental File 3). Significant Molecular Function enriched included protein binding, calcium ion binding, ligase activity, hydrolase activity, and ion channel binding.

Different DNA-binding motifs associated with H3K9me3 were discovered at different time points of TGFβ stimulation

We used MEME-ChIP (https://meme-suite.org/meme/tools/meme-chip) to discover DNA-binding motifs, utilizing the sequences from H3K9me3-6h and H3K9me3-24h samples.

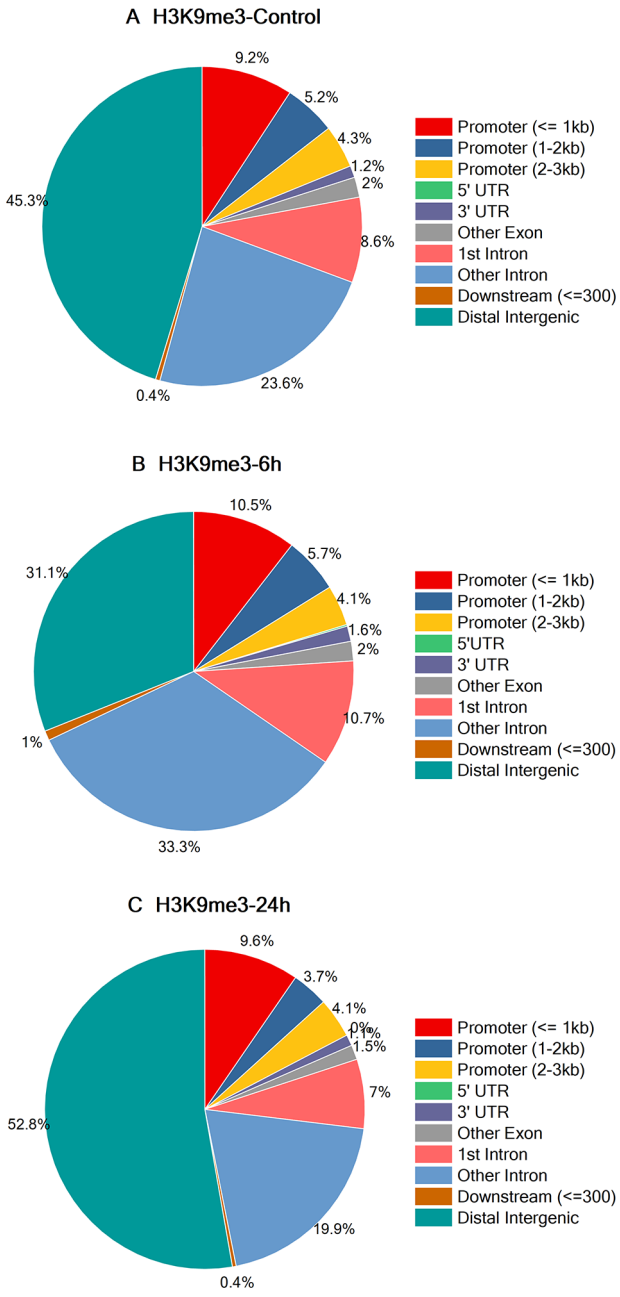

Figure 3. Peak annotations using annotatePeak from the ChIPseeker R/Bioconductor package. Regions with H3K9me3 mark during 0 hour (H3K9me3-Control) (A), 6 hours-H3K9me3-6h (B), and 24 hours-H3K9me3-24h (C) of TGFβ stimulation. The groupings included were Promoter (≤1 kb), Promoter (1-2 kb), Promoter (2-3 kb), 5′UTR, 3′UTR, first exon, other exons, first intron, other introns, Downstream (≤300), and distal intergenic.

The MEME-ChIP results were used as input in Tomtom to match our motifs to known binding motifs from the HOCOMOCO v11 database. In H3K9me3-6h, we discovered 11known DNA-binding Motifs (P-value < .001) viz. CPEB1, HXC10, PRDM6 (Figure 5A-C) along with HMX1, FOXG1, FOXL1, FOXJ3, NFAC1, and MNX1 (not in Figure, Supplemental File 4). In the H3K9me3-24h sample, we discovered 17 known DNA-binding motifs (P-value < .01). These included ZN680, GLI1, NF2L2 (Figure 5D-F) along with ZNF41, BRCA1, TAL1, STAT3, SOX2, STA5B, PPARD, PO5F1, EVI1, FOXJ3, BCL6B, KAISO, SRY, and OLIG2 (not in Figure, Supplemental File 4). SRY and FOXJ3 (Figure 6) were the common DNA motifs discovered in H3K9me3-6h and H3K9me3-24h. DNA motifs are shown to regulate histone modifications in humans and mice by providing locus-specific guidance for the histone-modifying enzymatic functions.[36] Our results point out the possibilities that these different motifs can translate the time-dependent effects of TGFβ through the regulation of the H3K9me3 mark. Also, we report that these discovered motifs are novel in terms of their association with the H3K9me3 mark after comparing with the dataset provided by Ngo et al[36] in their study and thus warrant extensive investigation as to how these motifs can modulate gene expression and regulate H3K9me3 mark by providing locus-specific guidance for the enzymatic functions, in a TGFβ stimulated environment. The complete list of discovered motifs for H3K9me3-6h and H3K9me3-24h can be found in Supplemental File 4.

Relative gene-expression analysis of randomly selected genes associated with H3K9me3 mark

We randomly selected genes from different combination categories of samples, for example, genes associated with H3K9me3 marked only in untreated samples or only associated with H3K9me3 in treated samples or exclusive to one sample only and measured their relative expression using qRT-PCR (Figure 7). No direct correlation was found between the occupancy of the H3K9me3 mark and the expression of various genes. XRN1 was associated with the H3K9me3 mark in all the samples but H3K9me3 occupancy of the gene was found decreasing in a time-dependent manner in the TGFβ treated sample. With decreasing H3K9me3 occupancy and H3K9me3's repressive property, the increase in relative fold-change in 24 hours sample was warranted for but contrasting results were observed in 6 hours where the decrease in H3K9me3 occupancy was associated with a decrease in expression. Similarly, we observed a contrasting result in OR2W5 expression, wherein H3K9me3 occupancy was observed in the control and 6 hours treated sample with no significant change in H3K9me3 enrichment between the 2 samples, yet the expression decreased in 6 hours treated sample and increased in 24 hours treated sample. RAB24 showed H3K9me3 occupancy in the control sample only. Its expression decreased in the 6 hours treated sample and increased in the 24 hours sample. BGLT3 showed H3K9me3 occupancy in the 6 hours treated sample only and in accordance with H3K9me3 as a repressive mark, we observed a decrease in expression in the 6 hours sample and elevation in the 24 hours sample. HERC2P4 showed H3K9me3 occupancy in 24 hours treated sample only and to our surprise gene expression was highly upregulated in the 24 hours treated sample. From this we can possibly conclude a fact that a gene's expression is not directly linked to one histone modification, instead,

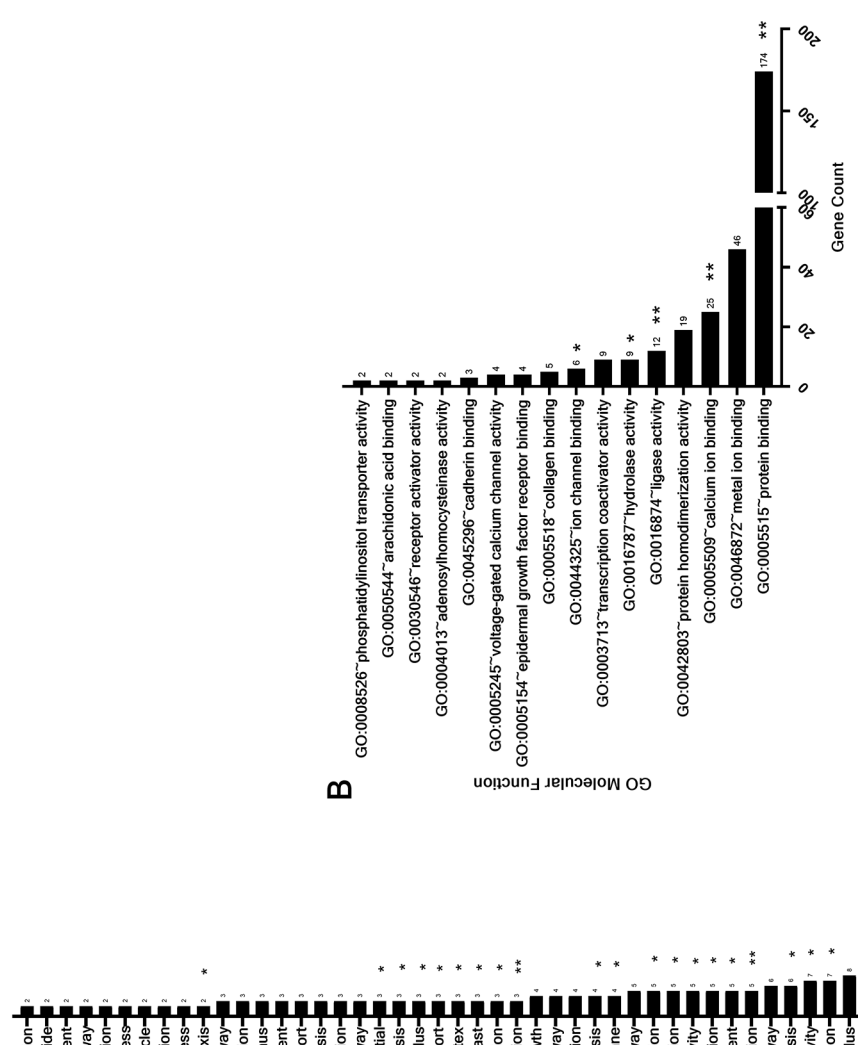

Figure 4. Using DAVID (v 6.8), enriched GO: Biological Processes (A) and GO: Molecular Functions (B) associated with TGFβ-stimulation were identified using the annotated genes specific to H3K9me3-6h. (**P*-value < .01, ***P*-value < .05).

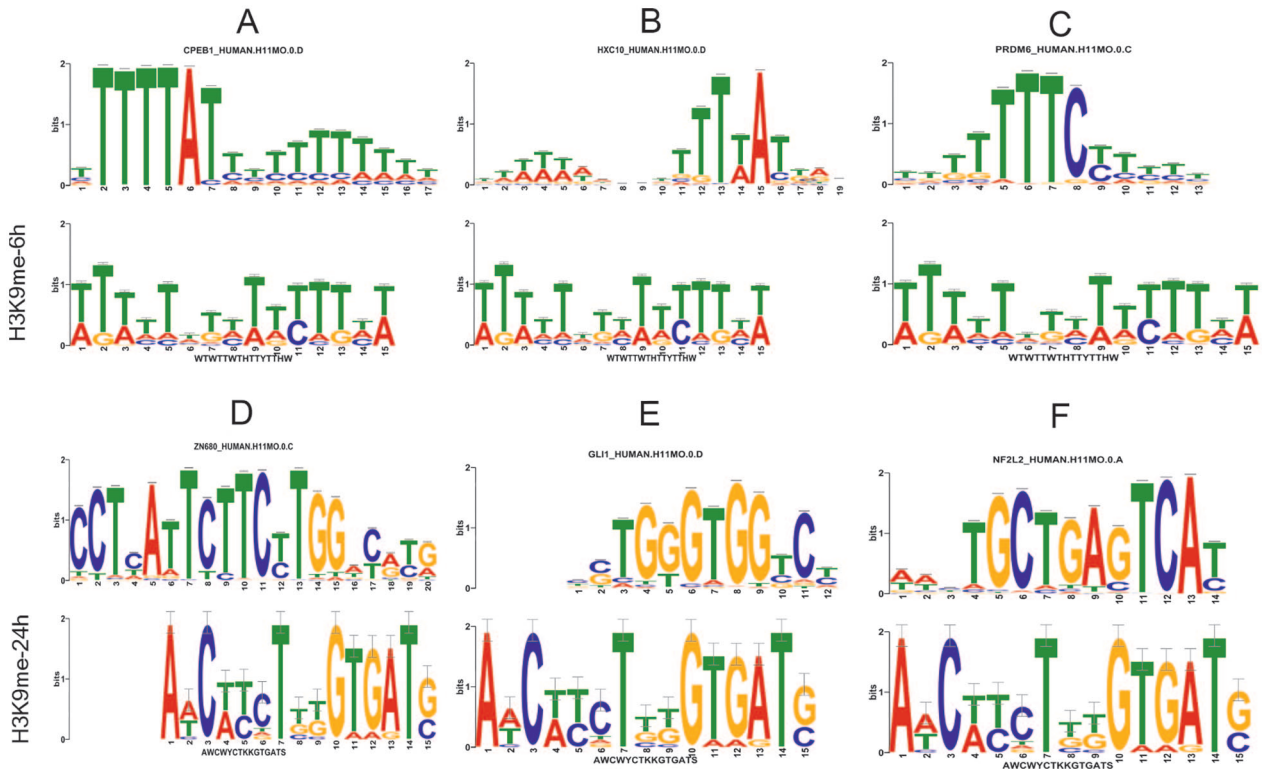

Figure 5. Comprehensive motif analysis on sequences centrally located in the Promoter Regions using MEME-ChIP (v5.3.3). The Motifs were discovered by importing 250bp sequence flanking on both side of the summit of MACS peaks into the MEME-ChIP tool (v5.3.3). Using Tomtom (v5.4.1), the discovered motifs were compared against a database of known human motifs that is, HOCOMOCO (v11). The top 3 DNA-binding motifs discovered in the H3K9me3-6h sample were (A) CPEB1, (B) HXC10, and (C) PRDM6, and in the H3K9me3-24h were (D) ZN680, (E) GLI1, and (F) NF2L2.

it is the outcome of a complex interplay between 2 or more histone modifications. The gene expression can also be regulated depending upon the reader of the histone mark where one reader can upregulate while another reader can downregulate the gene expression at the same histone mark.

Comparison of biological processes and molecular functions associated with genes showing H3K9me3 occupancy between RPWE2 (Normal cell line) and PC3 (Cancer cell line)

We prepared 2 lists of genes with H3K9me3-one exclusive for RPWE2 (a normal prostate cell line) and another exclusive for PC3 (a cancerous prostate cell line). The RWPE2 dataset was obtained from NCBI GEO Accession no: GSE175152.[37] The 2 gene lists were put through Functional Enrichment using DAVID Functional Annotation tool to identify and compare the Biological Processes (Figure 8) and Molecular Functions (Figure 9) among the 2 cell lines. Enrichment of GO:0006468~protein phosphorylation was common in both the cell lines indicating the association of protein phosphorylation process with H3K9me3 is generalized and not specific to any specific cell condition viz normal and cancerous. Enrichment of GO:0000086~G2/M transition of the mitotic cell cycle, GO:0045893~positive regulation of transcription, DNA-templated and GO:0010628~positive regulation of gene expression pointed out toward the association of H3K9me3 mark with active gene expression and progression of the cell cycle in the normal cell line. Also, enrichment GO: 0033169~H3-K9 demethylation is indicative of the removal of the repressive mark increasing gene expression. The positive association between the H3K9me3 mark and normal cell functioning is further confirmed by the enrichment of GO:0071320~cellular response to cAMP. The detailed list of GO Biological Processes enriched are mentioned in the Supplemental File 5. Surprisingly, in the cancer cell line, H3K9me3 was found to be associated with genes involved in processes such as GO:0000082~G1/S transition of cell cycle and GO:0008284~positive regulation of cell proliferation, indicating that H3K9me3 could also act as an active mark, but further studies are needed to confirm this fact. In contrast to this, we also observed non-significant enrichment of GO:0007050~Cell cycle arrest. The detailed list of GO Biological Processes enriched are mentioned in the Supplemental File 6.

In the Molecular Function category, we observed enrichment of several significant functions in the normal cell line (Figure 9A). A major proportion of enriched molecular functions belonged to protein kinase and binding activity such as JUN kinase (GO:0004705), MAP kinase (GO:0004707) indicating the positive association of genes with protein phosphorylation activity and H3K9me3 mark. Enrichment of

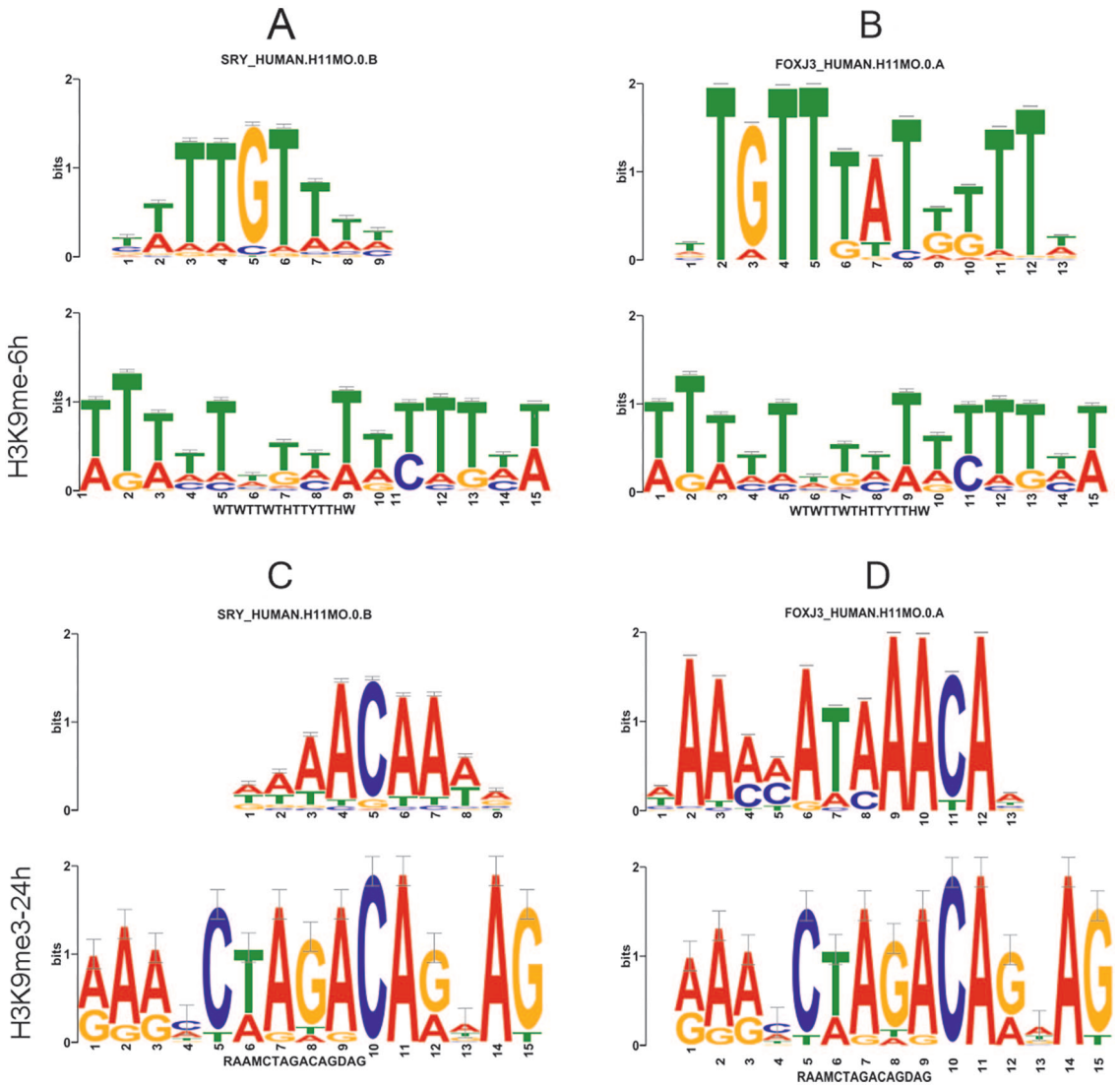

Figure 6. DNA-binding Motifs of SRY discovered in H3K9me3-6h (A) and H3K9me3-24h sample(C) and FOXJ3 discovered in H3K9me3-6h (B) and H3K9me3-24h sample(D).

GO:0005524~ATP binding indicated the association of H3K9me3 mark with ATP mediated enzyme regulation in the normal cell line. The detailed list of exclusive GO Molecular Function enriched in RPWE2 cell line is mentioned in the Supplemental File 7. In the cancer cell line, genes with H3K9me3 occupancy showed enrichment of Molecular Function (Figure 9B) such as GO:0001085~RNA polymerase II transcription factor binding and GO:000979~RNA polymerase II core promoter sequence-specific DNA binding indicated the possible role of genes with H3K9me3 occupancy in initiating and regulating transcription factor binding, thus modulating the transcription process of certain genes. The detailed list of exclusive GO Molecular Function enriched in PC3 cell lines are mentioned in the Supplemental File 8.

Discussion

Chromatin immunoprecipitation was performed on PC3 cells stimulated with TGFβ for 6 and 24 hours using H3K9me3 antibody to find out the regions associated with H3K9me3 in response to TGFβ. Out of the total 100% of peaks obtained, 47.2% were associated with control, 43.2% were associated with H3K9me3-6h and 9.6% were associated with H3K9me3-24h. So, the overall percentage of regions associated with H3K9me3 decreases with the increasing time of TGFβ stimulation indicating it reduces H3K9me3 mark at a few regions across the genome and in many regions, it completely removes the H3K9me3 mark, thereby reducing the repressive conditions and possibly promoting gene expression. Very few regions were found in H3K9me3-24h suggesting a sudden decrease (as compared to H3K9me3-6h) in the repressive H3K9me3 mark in response to TGFβ. Further experiments are required to confirm the role of decreased H3K9me3 in TGFβ induced migration and invasion of prostate cancer cells. Few studies have shown a positive correlation between SUV39H1 (H3K9 methyltransferase) and prostate cancer pathological stages. It would be interesting to study the 70 genes that are unique to the

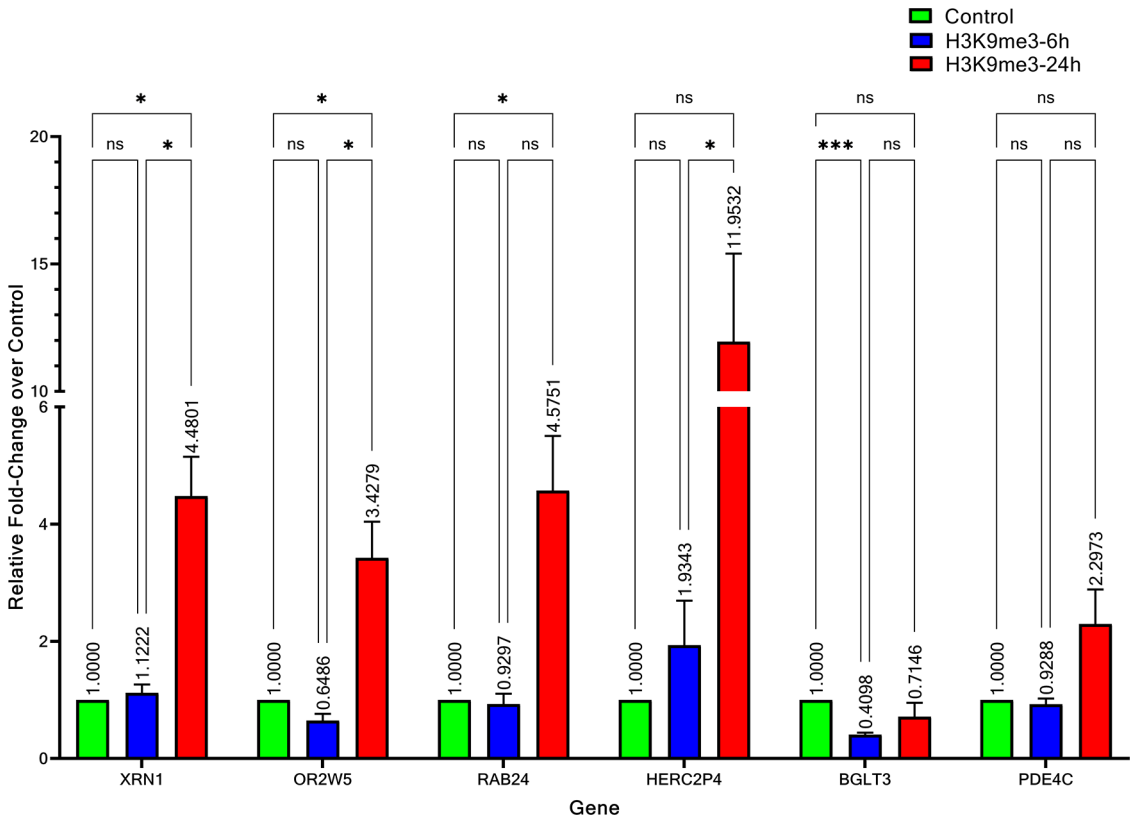

Figure 7. Relative fold-change in expression of randomly selected genes-having decreasing H3K9me3 occupancy during 6 and 24 hours TGFβ Stimulation (XRN1) as compared to Control, having decreased H3K9me3 occupancy in 6 hours sample & no H3K9me3 occupancy in 24 hours sample (OR2W5) with respect to Control, having no H3K9me3 occupancy in treated sample (RAB24), having H3K9me3 occupancy exclusively in 6 hours TGFβ stimulated sample (BGLT3) and having H3K9me3 occupancy exclusive to 24 hours TGFβ stimulated sample (HERC2P2).

sample H3K4me3-24h and did not come up in other time points of TGFβ stimulation. Also, the group of regions unique to Control conditions may point toward those biological processes and molecular functions which are activated upon TGFβ stimulation as the repressive H3K9me3 mark is removed. It will be important to evaluate these regions where the loss of the H3K9me3 mark could lead to activation or repression of pathways that promote tumorigenesis. Further investigations also need to be conducted to understand the effects of the removal of the H3K9me3 mark. The regions unique to 6 hours-TGFβ stimulated sample also needs to be explored further to ascertain the role of the H3K9me3 mark and its downstream effects as their expression may be affected by H3K9me3 occupancy since this mark is involved in transcriptional activation and mRNA elongation by RNA pol II in association with HP1γ.[6,7] There were fewer regions that had the H3K9me3 mark unique to the sample stimulated with TGFβ for 24 hours. This group includes those regions where the H3K9me3 mark was present only after 24 hours of TGFβ stimulation. The subsequent effects on the addition of H3K9me3 mark post 24 hours of TGFβ stimulation need to be studied so the cells can potentially be linked to processes that require a longer duration of active TGFβ signaling.

Surprisingly, 23.6 regions belong to the introns in the control sample, while it increases to 33.3% in response to TGFβ for 6 hours and it decreases to 19.9% in response to TGFβ for 24 hours. To our surprise, occupancy of H3K9me3 at distal intergenic regions follows a completely different pattern. The percentage occupancy at the distal intergenic region is 45.3%, it decreases to 31.1% in 6 hours of TGFβ stimulation and increases to 52.8% in 24 hours of TGFβ stimulation. So, the occupancy at distal intergenic regions decreases while at the intronic regions it increases in response to TGFβ for 6 hours. In 24 hours of TGFβ stimulation, the occupancy at the distal intergenic region increases while the intronic region decreases. It would be of great importance to understand the functional significance of this increase and decrease in the occupancy at different regions of the genome in response to TGFβ.

Data were analyzed to categorize the regions based on their biological processes involved and molecular functions. Very few regions were found to be associated with H3K9me3 in the sample stimulated with TGFβ for 24 hours, thereby cannot be categorized based on their biological process and molecular function. Only 6 hours of TGFβ stimulated samples were analyzed and categorized based on the biological process and molecular function. Simultaneous enrichment of negative regulation of transcription and positive regulation of DNA-binding transcription factor activity suggests the dual role of H3K9me3 as a transcriptional activator and repressor in response to TGFβ for 6 hours. The other biological processes

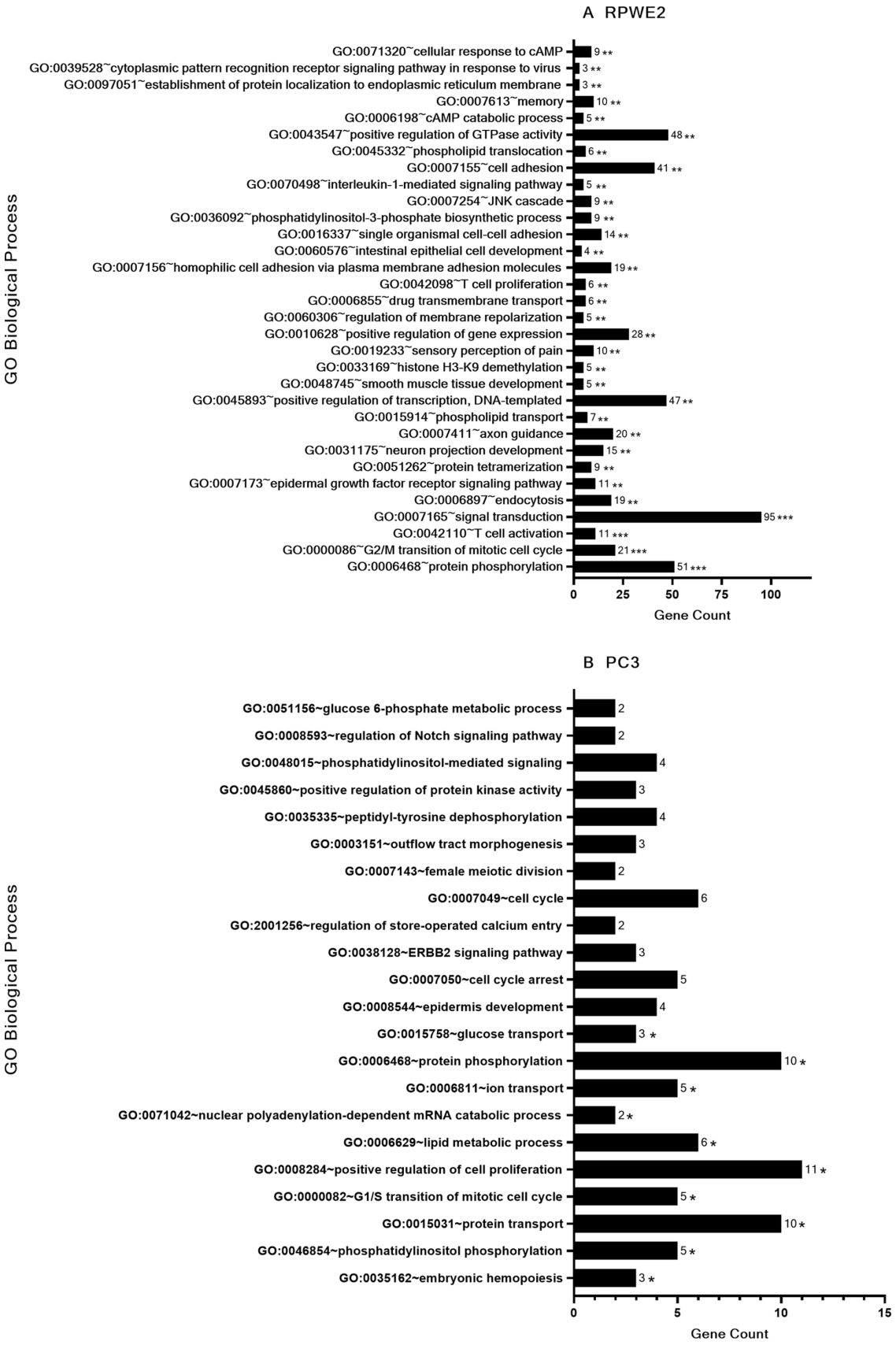

Figure 8. Genes exclusive to RPWE2 cell line dataset and exclusive to PC3 cell line dataset obtained after comparison of untreated H3K9me3-RPWE2 dataset with untreated H3K9me3-PC3 dataset were put through functional enrichment using DAVID Functional Annotation tool to identify biological processes associated with H3K9me3 mark in (A) RPWE2 cell line and (B) PC3 cell.

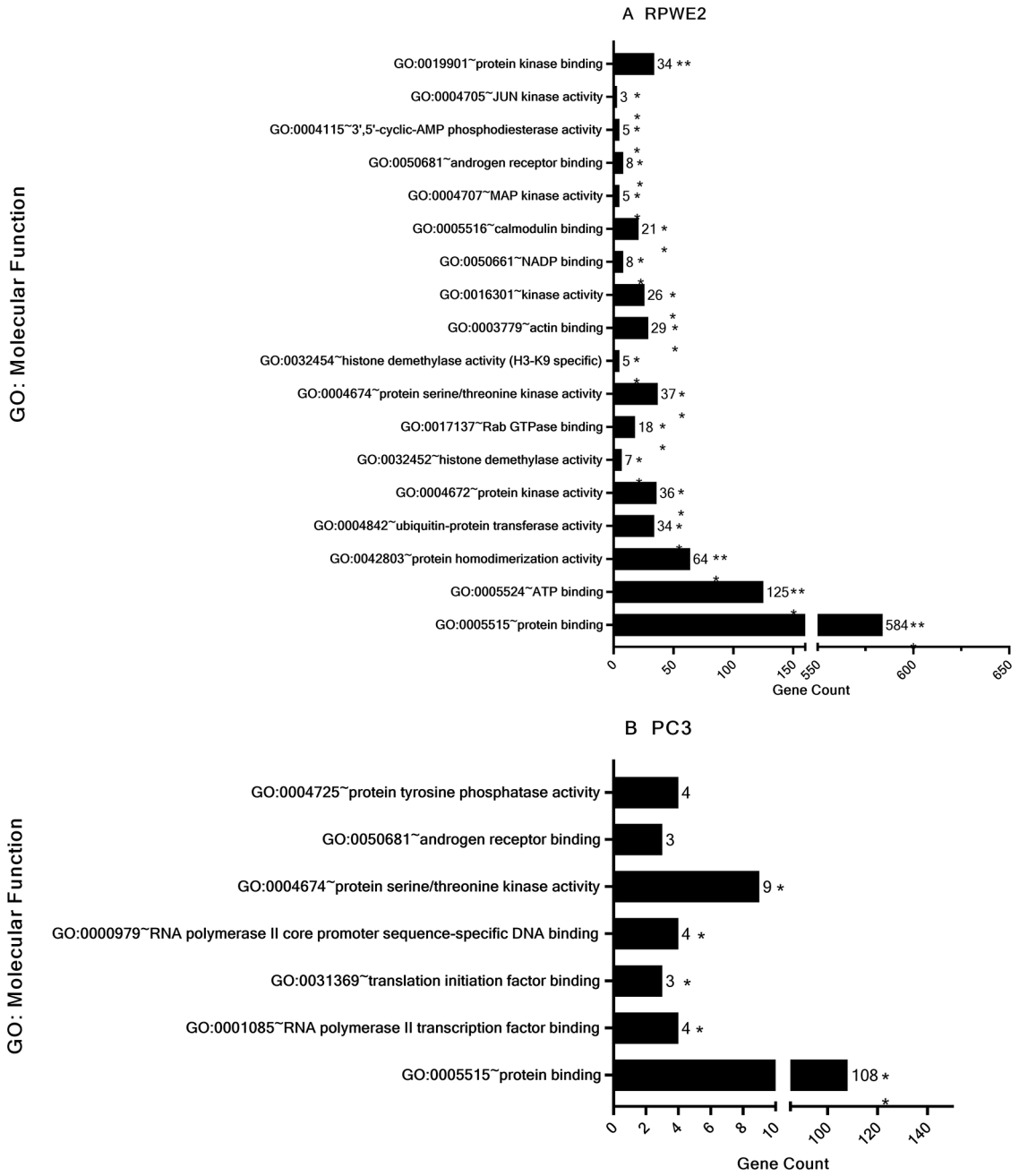

Figure 9. Genes exclusive to RPWE2 cell line dataset and exclusive to PC3 cell line dataset obtained after comparison of untreated H3K9me3-RPWE2 dataset with untreated H3K9me3-PC3 dataset were put through functional enrichment using DAVID Functional Annotation tool to identify molecular function associated with H3K9me3 mark in (A) RPWE2 cell line and (B) PC3 cell.

enriched were actin cytoskeletal rearrangement and cell adhesion. TGFβ plays an important role in cell migration and cell adhesion. Our report suggests the functional role of H3K9me3 in TGFβ induced cell migration and cell adhesion. Also, the role of plasma membrane transporters in cancer has received a lot of attention in recent years. In cancer, several essential nutrient transporters are up regulated and act as tumor promoters. Transporters may have the ability to suppress tumor growth. So far, 4 transporters belonging to the SLC gene family have been identified as tumor suppressors,[38] therefore evidence of enrichment of SLC genes during short-term TGFβ stimulation warrants further studies to identify their role in association with TGFβ signaling.

Random genes were selected from the data to find the correlation between the H3K9me3 occupancy and the gene expression. No direct (positive or negative) correlation was found between the H3K9me3 occupancy and expression suggesting that various histone modifications act together to determine the expression of a gene. BGLT3 promoter was found to be 15-fold enriched with H3K9me3 in only 6 hours of TGFβ stimulation and the expression was decreased to half in 6 hours of TGFβ stimulation. No significant change was found

in the 24 hours stimulated sample. Human LncRNA-BGL3 is identified as a functional long non-coding RNA whose expression is downregulated by Bcr-Abl.[39] Previous studies indicated BGL3 (Beta Globin Locus 3) as a noncoding RNA that may have a role in the regulation of γ-globin expression during development.[40] LncRNA-BGL3 is a target of miR-17, miR-20a, miR-20b, miR-93, miR-106a, and miR-106b, microRNAs that repress messenger RNA of PTEN. PTEN is a tumor suppressor gene that is very commonly lost in many cancer types like prostate cancer, glioblastoma, endometrial, lung, and breast cancer.[41] Over 70% of prostate cancer patients have been observed to have lost the expression of this gene.[42] We found the highest enrichment of H3K9me3 on the promoter of BGLT3 in the prostate cancer cells treated with TGFβ for 6 hours. LncRNA-BGL3 and PTEN are both directly targeted by the shared miRNAs, and they cross-regulate each other probably through sequestration of shared miRNAs. LncRNA-BGL3 acts as a ceRNA to regulate PTEN levels and thus plays an important role in TGFβ induced tumorigenesis.[39] Our results depict the occupancy of H3K9me3 in the BGLT3 promoter region in response to TGFβ for 6 hours. These results suggest the downregulation of BGLT3 in prostate cancer cells in response to TGFβ for 6 hours and thus downregulating the expression of PTEN too in response to TGFβ for 6 hours in prostate cancer cells. It would be interesting to understand the mechanisms involving H3K9me3 on LncRNA-BGL3, in downregulating the expression of PTEN in prostate cancer cells.

HERC2P4 promoter was found to be 5-fold enriched with H3K9me3 in 24 hours of TGFβ stimulation while the expression increases 12-folds as compared to the sample without TGFβ. HERC2P4 (Hect Domain and RLD 2 RLD Domain Containing E3 Ubiquitin Protein Ligase 2 Pseudogene 4) was found to be one of the genes downregulated in organ confined prostate cancer versus benign prostatic hyperplasia in the lean subject group. The 5-fold enrichment of H3K9me3 was found on HERC2P4 promoter in only 24 hours of TGFβ stimulation. This correlated with the 12-fold increase in the expression of HERC2P4 with 24 hours of TGFβ stimulation. It will be very interesting to understand the functional importance of this pseudogene in prostate cancer and the link between the enrichment of H3K9me3 and the increase in expression.

Another gene that was found to be enriched with H3K9me3 on its promoter in response to TGFβ for 6 hours was PDE4C. PDE4 has been shown to play an important role in the mechanism of EMT. It has been demonstrated to attenuate the epithelial-mesenchymal transition by PDE4 inhibition. TGF-β1 has been shown to be a major regulator of EMT. A549 cells are the best characterized alveolar epithelial cells and have been used in studying various aspects of TGFβ induced EMT.[43] Several different techniques were performed to demonstrate the increased role of PDE4A and PDE4D in TGF-β1-induced EMT, as compared to the control A549 cells.[44] PDE4C has not been studied in detail earlier. Our results showed an increased expression of PDE4C in 24 hours of TGFβ stimulated cells suggesting an important role of PDE4C in TGFβ induced biological functions including tumorigenesis.

The occupancy of H3K9me3 decreases on Xrn1 (distal intergenic region) in response to TGFβ for 6 and 24 hours suggesting an increase in the expression of Xrn1 in response to TGFβ. When the expression of Xrn1 was studied by RT-PCR, we found no significant change in expression of Xrn1 in 6 hours and a 4.5-fold increase in the expression in 24 hours. It would be interesting to understand the functional role of decrease in H3K9me3 in the distal intergenic region of Xrn1 gene and how it correlates with its expression. The major 5′-3′ exonuclease Xrn1 has been shown to play an unanticipated role in both activating translation of mRNAs encoding membrane proteins and directing them to the ER which is their translation site.[45] Xrn1 can also activate the translation, transcription, and decay of the specific mRNAs, and these functions are all linked to each other by Xrn1 for a specific group of mRNAs encoding membrane proteins.[46] The hydrophobic domains of these membrane proteins have the tendency to form aggregates. To prevent the aggregation of these proteins which might be toxic, their expression and localization should be strictly controlled.[45] It would be interesting to investigate the mechanism by which Xrn1 would be activating the translation, transcription, and decay of specific RNAs in response to TGFβ for 12 hours. From the overall RT-qPCR analysis, we can possibly conclude a fact that a gene's expression is not directly linked to one histone modification, instead, it is the outcome of a complex interplay between 2 or more histone modifications. The gene expression can also be regulated depending upon the reader of the histone mark where one reader can upregulate while another reader can downregulate the gene expression at the same histone mark.

The 2 known motifs identified for H3K9me3-6h against our submitted motifs were binding sites of CPEB1 and PRDM6. Cytoplasmic polyadenylation element-binding (CPEB) proteins bind to the mRNAs and control their translation. CPEB1 and CPEB4 has been shown to be positively associated with both canonical and non-canonical TGFβ signaling as their knockdown can attenuated TGFβ-activated expression levels of phosphorylated Smad 2, Smad 1/5/8, TAK1, P38, ERK, and JNK in fibroblasts.[47] The protein encoded by PRDM6 (putative histone-lysine N-methyltransferase), a member of the PRDM family, can act as a transcriptional repressor. The PRDM family members are recognized by their PR and multiple zinc-finger domains. The encoded protein has been shown to be involved in the regulation of vascular smooth muscle cells (VSMC) contractile proteins.[48] The regulation of PRDM6 by TGFβ is a novel finding from our experiments. It would be interesting in understanding the crosstalk between these 2 and how they cooperate to induce TGFβ induced tumorigenesis. Interestingly, these 2 hits are both transcriptional

repressors, and it would be very interesting to find out the genes repressed by these motifs and their functional role in relation to TGFβ.

The 3 known motifs identified for H3K9me3-24h against our submitted motifs were binding sites of ZN680, GLI1, and NF2L2. ZN680 belongs to the family of kruppel C2H2-type zinc finger protein and is involved in transcriptional regulation of specific genes. It enables RNA polymerase II cis-regulatory region sequence-specific binding, DNA-binding transcription factor activity, and DNA binding. The increased occupancy of H3K9me3 on this motif suggests a positive transcriptional regulatory function of this histone modification. GLI1 or Glioma-associated oncogene also acts as a transcriptional activator and regulates the transcription of specific genes during normal development. It plays a role in cell proliferation and differentiation. It has been demonstrated that autocrine TGFβ signaling was responsible for high GLI1 expression resulting in the high proliferative capacity of pancreatic carcinoma cell lines. NF2L2 (Nuclear Factor (Erythroid-Derived2)-Like 2 is an NFE2 like BZIP transcription factor 2 that plays an important role in the response to oxidative stress. It binds to antioxidant response elements (AREs) of cryoprotective genes. No research work has been reported on ZN680, NFE2 in the context of its regulation by TGFβ so we will be interested in understanding the connection between these 2. Surprisingly, the motifs that came up at 6 hours of TGFβ stimulation were transcriptional repressors while at 24 hours, they are transcriptional activators.

SRY and FoxJ3 are the only 2 motifs common between H3K9me3-6h versus H3K9me3-Control and H3K9me3-24h versus H3K9me3-Control. Autocrine TGFβ signaling plays an essential role in the retention of stemness of glioma-initiating cells (GICs) and describes the underlying mechanism for it. TGFβ induced [corrected] expression of Sox2, a stemness gene, and this induction was mediated by Sox4, SRY-Box Transcription Factor 4, a direct TGFβ target gene.[49,50] SOX4 has increased expression in prostate cancer as compared with normal prostate tissue. Increased expression levels of SOX4 directly associates with more aggressive tumors.[51] No correlation was found between the TGFβ and FoxJ3.

Comparison of the H3K9me3 ChIP-sequencing data for PC3 (prostate cancer cells) and RWPE2 (normal prostate cells) revealed several interesting genes like *CDK7, PDEC, IL24, TFAMP1, MAP3K14, FGF12, OR2W5, TGFBR2, RNU6-805P, MMP3, MMP20, SLC22A11, RAB24, ACADs*, and *HBG1*, which were only found in the PC3 cells and not in the RWPE2 cells suggesting an important role of H3K9me3 on these genes to drive tumorigenesis. The molecular functions found exclusively in the PC3 cells were the RNA polII transcription factor binding, RNA polII core promoter sequence-specific DNA binding, and translation initiation factor binding suggesting the positive transcriptional role of H3K9me3 in prostate cancer cells as compared to normal prostate cells. H3K9me3 occupancy was higher in the genes involved in the G1/S transition of the mitotic cell cycle in PC3 cells as compared to the G2/M transition in normal RWPE2 cells. The regions found in PC3 cells were involved in the notch signaling while the regions found in normal prostate cells were involved in interleukin mediated and EGFR mediated signaling.

H3K9me3 ChIP data were analyzed for the genes belonging to the biological process in the 6 and 24 hours of TGFβ stimulation. Genes belonging to the cell adhesion process were the most prevalent ones for both 6 and 24 hours of TGFβ stimulation suggesting an important role of H3K9me3 in cell adhesion in response to TGFβ. A comprehensive understanding of the oncogenic function of TGFβ by modifying H3K9me3 levels in prostate cancer and the respective epigenetic regulators involved will permit the development of more efficient therapeutic strategies in the future.

Conclusion

This study describes how H3K9me3 landscape changes in response to TGFβ stimulation and correlates it with functional category-specific gene regulation. TGFβ stimulation decreases the highly enriched H3K9me3 occupancy as seen by the peak count reduction. The genomic coordinates of TGFβ-targeted reductions in H3K9me3 occupancy correspond to significantly rich occurrences of specific transcription factors suggesting an epigenetic means of transcription regulation by acute stimulation with TGFβ. Contrary to the canonical transcription repressive function attributed to H3K9me3, our results suggest the dual role of H3K9me3 in transcriptional activation as well as repression. We thus report the specific genomic regions and associated genes that are regulated by TGFβ through an H3K9me3-mediated epigenetic mechanism. Understanding the functional role of H3K9me3 in TGFβ induced tumorigenesis will help in development of new therapeutic strategies against prostate cancer.

Acknowledgements

Authors acknowledge Department of Science and Technology (DST), Government of India for Ankit Naik's INSPIRE-fellowship and Human Resource Development Group, Council of Scientific and Industrial Research (CSIR-HRDG), Government of India for Nidhi D's fellowship. We also acknowledge MedGenome Labs for providing the sequencing facility.

Author Contributions

All the experiments and analysis were performed by AN, ND contributed to the qPCR results and the designing of the project was done by AN and NT. AN and NT designed the experiment. The study was supervised by NT. The manuscript was written by AN and NT. All authors approved the final draft of the manuscript.

Data and Materials Availability

The ChIP-seq data has been deposited in Gene Expression Omnibus (GEO) of NCBI and can be accessed using the GEO accession number GSE191096 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE191096).

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Sartor O. Why is prostate cancer incidence rising in young men? *Cancer*. 2020;126:17-18.
2. Derynck R, Miyazono K. *The TGF-[beta] Family*. CSHL Press; 2008:1114.
3. Zhang YE. Non-Smad pathways in TGF-beta signaling. *Cell Res*. 2009; 19:128-139.
4. Darwiche N. Epigenetic mechanisms and the hallmarks of cancer: an intimate affair. *Am J Cancer Res*. 2020;10:1954-1978.
5. Martens JH, O'Sullivan RJ, Braunschweig U, et al. The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J*. 2005;24:800-812.
6. Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell*. 2005;19:381-391.
7. Eissenberg JC, Shilatifard A. Leaving a mark: the many footprints of the elongating RNA polymerase II. *Curr Opin Genet Dev*. 2006;16:184-190.
8. Olcina MM, Leszczynska KB, Senra JM, Isa NF, Harada H, Hammond EM. H3K9me3 facilitates hypoxia-induced p53-dependent apoptosis through repression of APAK. *Oncogene*. 2016;35:793-799.
9. Biga PR, Latimer MN, Froehlich JM, Gabillard J-C, Seiliez I. Distribution of H3K27me3, H3K9me3, and H3K4me3 along autophagy-related genes highly expressed in starved zebrafish myotubes. *Biol Open*. 2017;6:1720-1725.
10. Magaraki A, van der Heijden G, Sleddens-Linkels E, et al. Silencing markers are retained on pericentric heterochromatin during murine primordial germ cell development. *Epigenet Chromatin*. 2017;10:11.
11. Sun Y, Jiang X, Xu Y, et al. Histone H3 methylation links DNA damage detection to activation of the tumour suppressor Tip60. *Nat Cell Biol*. 2009;11: 1376-1382.
12. Bieberstein NI, Kozáková E, Huranová M, et al. TALE-directed local modulation of H3K9 methylation shapes exon recognition. *Sci Rep*. 2016;6:29961.
13. Pedersen MT, Kooistra SM, Radzisheuskaya A, et al. Continual removal of H3K9 promoter methylation by JMJD2 demethylases is vital for ESC self-renewal and early development. *EMBO J*. 2016;35:1550-1564.
14. Imai K, Kamio N, Cueno ME, et al. Role of the histone H3 lysine 9 methyltransferase suv39 h1 in maintaining epsteinn-barr virus latency in B95-8 cells. *FEBS J*. 2014;281:2148-2158.
15. Fukuda A, Tomikawa J, Miura T, et al. The role of maternal-specific H3K9me3 modification in establishing imprinted X-chromosome inactivation and embryogenesis in mice. *Nat Commun*. 2014;5:5464.
16. Mendelsohn AR, Larrick JW. Stem cell depletion by global disorganization of the H3K9me3 epigenetic marker in aging. *Rejuvenation Res*. 2015;18:371-375.
17. Koide S, Oshima M, Takubo K, et al. Setdb1 maintains hematopoietic stem and progenitor cells by restricting the ectopic activation of nonhematopoietic genes. *Blood*. 2016;128:638-649.
18. Monaghan L, Massett ME, Bunschoten RP, et al. The emerging role of H3K9me3 as a potential therapeutic target in acute myeloid leukemia. *Front Oncol*. 2019;9:705.
19. Wang L, Xu X, Cao Y, et al. Activin/Smad2-induced histone H3 lys-27 trimethylation (H3K27me3) reduction is crucial to initiate mesendoderm differentiation of human embryonic stem cells. *J Biol Chem*. 2017;292:1339-1350.
20. Fei T, Xia K, Li Z, et al. Genome-wide mapping of SMAD target genes reveals the role of BMP signaling in embryonic stem cell fate determination. *Genome Res*. 2010;20:36-44.
21. Xi Q, Wang Z, Zaromytidou A-I, et al. A poised chromatin platform for TGF-β access to master regulators. *Cell*. 2011;147:1511-1524.
22. Maina PK. *Novel oncogenic roles and regulations of histone demethylase PHF8 in prostate cancer*. Dissertation. The University of Iowa; 2017. doi:10.17077/etd.7qoq6vqe
23. Sun Y, Wei M, Ren S-C, et al. Histone methyltransferase SETDB1 is required for prostate cancer cell proliferation, migration and invasion. *Asian J Androl*. 2014;16:319-324.
24. Yu T, Wang C, Yang J, Guo Y, Wu Y, Li X. Metformin inhibits SUV39H1-mediated migration of prostate cancer cells. *Oncogenesis*. 2017;6:e324-e324.
25. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114-2120.
26. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of chip-Seq (MACS). *Genome Biol*. 2008;9:R137.
27. Yu G, Wang L-G, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*. 2015;31: 2382-2383.
28. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4: 44-57.
29. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37:1-13.
30. Ma W, Noble WS, Bailey TL. Motif-based analysis of large nucleotide data sets using MEME-ChIP. *Nat Protoc*. 2014;9:1428-1450.
31. Kulakovskiy IV, Vorontsov IE, Yevshin IS, et al. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res*. 2018;46:D252-D259.
32. Gupta S, Stamatoyannopoulos JA, Bailey TL, Noble WS. Quantifying similarity between motifs. *Genome Biol*. 2007;8:R24.
33. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002; 30:207-210.
34. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res*. 2013;41:D991-D995.
35. Naik A, Dalpatraj N, Thakur N. Global Histone H3 lysine 4 trimethylation (H3K4me3) landscape changes in response to TGFβ. *Epigenet Insights*. 2021;14:25168657211051755.
36. Ngo V, Chen Z, Zhang K, Whitaker JW, Wang M, Wang W. Epigenomic analysis reveals DNA motifs regulating histone modifications in human and mouse. *Proc Natl Acad Sci U S A*. 2019;116:3668-3677.
37. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57-74.
38. Bhutia YD, Babu E, Ramachandran S, Yang S, Thangaraju M, Ganapathy V. SLC transporters as a novel class of tumour suppressors: identity, function and molecular mechanisms. *Biochem J*. 2016;473:1113-1124.
39. Guo G, Kang Q, Zhu X, et al. A long noncoding RNA critically regulates Bcr-Abl-mediated cellular transformation by acting as a competitive endogenous RNA. *Oncogene*. 2015;34:1768-1779.
40. Ivaldi MS, Diaz LF, Chakalova L, Lee J, Krivega I, Dean A. Fetal γ-globin genes are regulated by the BGLT3 long noncoding RNA locus. *Blood*. 2018; 132:1963-1973.
41. Aboushousha T, El-Nahas EI, El-Hindawi A, et al. Implication of miRNA-153 on PTEN expression in prostatic adenocarcinoma. *Eur Rev Med Pharmacol Sci*. 2021;25:6834-6843.
42. Gurbuz V, Sozen S, Bilen CY, Konac E. miR-148a, miR-152 and miR-200b promote prostate cancer metastasis by targeting DNMT1 and PTEN expression. *Oncol Lett*. 2021;22:805.
43. Han ST, Kim JS, Lee JY, et al. The mechanism of attenuation of epithelial-mesenchymal transition by a phosphodiesterase 5 inhibitor via renal klotho expression. *Clin Exp Pharmacol Physiol*. 2018;45:269-277.
44. Kolosionek E, Savai R, Ghofrani HA, et al. Expression and activity of phosphodiesterase isoforms during epithelial mesenchymal transition: the role of phosphodiesterase 4. *Mol Biol Cell*. 2009;20:4751-4765.
45. Blasco-Moreno B, de Campos-Mata L, Böttcher R, et al. The exonuclease Xrn1 activates transcription and translation of mRNAs encoding membrane proteins. *Nat Commun*. 2019;10:1298.
46. Cherry PD, Peach SE, Hesselberth JR. Multiple decay events target HAC1 mRNA during splicing to regulate the unfolded protein response. *eLife*. 2019;8:e42262. doi:10.7554/eLife.42262
47. Cui HS, Joo SY, Cho YS, Kim J-B, Seo CH. CPEB1 or CPEB4 knockdown suppresses the TAK1 and Smad signalings in THP-1 macrophage-like cells and dermal fibroblasts. *Arch Biochem Biophys*. 2020;683:108322.
48. Davis CA, Haberland M, Arnold MA, et al. PRISM/PRDM6, a transcriptional repressor that promotes the proliferative gene program in smooth muscle cells. *Cell Mol Biol*. 2006;26:2626-2636.
49. Ikushima H, Todo T, Ino Y, Takahashi M, Miyazawa K, Miyazono K. Autocrine TGF-beta signaling maintains tumorigenicity of glioma-initiating cells through SRY-related HMG-box factors. *Cell Stem Cell*. 2009;5:504-514.
50. Vervoort SJ, van Boxtel R, Coffer PJ. The role of SRY-related HMG box transcription factor 4 (SOX4) in tumorigenesis and metastasis: friend or foe? *Oncogene*. 2013;32:3397-3409.
51. Fu C, Xin J, Zhang W, Lai J, Huang Z. LINC00992 exerts oncogenic activities in prostate cancer via regulation of SOX4. *Exp Cell Res*. 2021;408:112855.