Supplementary File

INSTRUCTIONS FOR RUNNING BISMARK IN VIRTUALBOX

1. System requirements

- A minimum of 2GB RAM for VirtualBox guest operating system with a preferred host system total of 8GB RAM or greater.
- A minimum of 10GB disk storage.
- "Virtualization Technology" is supported in the host system BIOS. Set it to Enable.

2. VirtualBox installation

- Download the latest VirtualBox version for your host operating system:

  https://www.virtualbox.org/wiki/Downloads

Follow the instructions and install VirtualBox.

3. Download image file containing Bismark tool-set

- Open any web browser on the host system and go to:

  http://toxicology.uga.edu/resources/dna_methylation_analysis/
- Download the "TGBS.ova" file and save it on the host system.
- In the VirtualBox Manager, go to "File" and "Import Appliance" and browse for the location of the "TGBS.ova" file.

3. Run the VirtualMachine

- Start the "TGBS machine". If error occurs, go to "Settings", select "USB" and disable the "USB controller". Then, start the "TGBS machine" again.
- Use password **TGBSKolli** for the **User** login.
- Click on the "Activities" tab and open "Files".

- The home directory includes the Bismark software version 0.17.0 and the required tools like ActivePearl. It also includes a sample fastq sequence file "SampleSeq_R1.fastq.gz".

4. Download Reference Genome:

- Select "Firefox" from the "Activities" tab and go to https://www.ncbi.nlm.nih.gov/
- In the search bar, select "Nucleotide" from the drop-down menu and then search for "U24170.1"
- Click on "FASTA" and select the option to "Send" the "Complete Record" to the "File" destination and select "Create File" and "Save File".
- The "Sequence FASTA" is now in the "Downloads" folder.
- Using the Files tool again, move sequence.fasta to:

  Home->DNA->bowtie2-2.3.0->bismark_v0.17.0->REF->.

5. DNA Methylation analysis

Following are the command lines for a single-end analysis using R1 forward reads.

- Note: Click on the "Activities" tab to move between the folders.

5.1. Prepare Reference Genome:

- Select "Terminal" from the "Activities" tab.
- The reference genome folder "REF" now contains a fasta file from NCBI with the Accession# U24170.1 for the human *p21* promoter. Use the following command to prepare the reference genome before mapping the bisulfite converted reads from the sample datafile. Note: In the command below there is a single space before each "--" (double hyphen) and a single space after verbose.

bismark_genome_preparation --/home/user/DNA/bowtie2-2.3.0/ --verbose /home/user/DNA/bowtie2-2.3.0/bismark_v0.17.0/REF/

- This creates two folders within the genome folder “REF”, one with C ->T genome index and another with G ->A.

5.2. Run Bismark:

- Read alignment step for sequences in the R1 read file “SampleSeq_R1.fastq.gz” with a single-end approach. Note: In the command below there is a space before “--“ and one before each “/home/user…”.

  bismark --bowtie2 /home/user/DNA/bowtie2-2.3.0/bismark_v0.17.0/REF/ /home/user/SampleSeq_R1.fastq.gz

- This aligns the sequence reads to the reference genome and creates a combined alignment/methylation call output in BAM format, and gives a run statistics report.
- Output files: written to /home

  SampleSeq_R1_bismark_bt2.bam

  SampleSeq_R1_bismark_bt2_SE_report.txt

5.3. Methylation extraction:

- Extracting methylation information out of the “.bam” file created in step 2. Note: In the command below there is a space before the “-s”, “--“, and “/home/user…”

  bismark_methylation_extractor -s --comprehensive /home/user/SampleSeq_R1_bismark_bt2.bam

- This extracts methylation information from the alignment output of the above

NGS file.

- Output files: written to /home

  SampleSeq_R1_bismark_bt2.M-bias.txt

  SampleSeq_R1_bismark_bt2.M-bias_R1.png

  SampleSeq_R1_bismark_bt2_splitting_report.txt

  CHG, CpG and CHH contexts for the SampleSeq_R1_bismark_bt2.txt

5.4. Generate report:

- Generating Bismark processing report on read alignment and methylation extraction.

  bismark2report

- This gives an overall methylation report on the total number of reads and their alignment and methylation.
- Output file: written to /home

  SampleSeq_R1_bismark_bt2_SE_report.html

Supplementary Datafile 1. SampleSeq_R1_bismark_bt2_SE_report.txt

Bismark report for: /home/user/SampleSeq_R1.fastq.gz (version: v0.17.0)

Option '--directional' specified (default mode): alignments to complementary strands (CTOT, CTOB) were ignored (i.e. not performed)

Bismark was run with Bowtie 2 against the bisulfite genome of /home/user/DNA/bowtie2-2.3.0/bismark_v0.17.0/hp21sie1ref/ with the specified options: -q --score-min L,0,-0.2 --ignore-quals

Final Alignment report

==========

Sequences analysed in total: 30135

Number of alignments with a unique best hit from the different alignments: 10571

Mapping efficiency: 35.1%

Sequences with no alignments under any condition: 19564

Sequences did not map uniquely: 0

Sequences which were discarded because genomic sequence could not be extracted: 0

Number of sequences with unique best (first) alignment came from the bowtie output:

CT/CT: 10571 ((converted) top strand)

CT/GA: 0 ((converted) bottom strand)

GA/CT: 0 (complementary to (converted) top strand)

GA/GA: 0 (complementary to (converted) bottom strand)

Number of alignments to (merely theoretical) complementary strands being rejected in total: 0

Final Cytosine Methylation Report

==========

Total number of C's analyzed: 829618

Total methylated C's in CpG context: 18365

Total methylated C's in CHG context: 647

Total methylated C's in CHH context: 2294

Total methylated C's in Unknown context: 0

Total unmethylated C's in CpG context: 13334

Total unmethylated C's in CHG context: 189351

Total unmethylated C's in CHH context: 605627

Total unmethylated C's in Unknown context: 0

C methylated in CpG context: 57.9%

C methylated in CHG context: 0.3%

C methylated in CHH context: 0.4%

Can't determine percentage of methylated Cs in Unknown context (CN or CHN) if value was 0

Bismark completed in 0d 0h 0m 19s

Supplementary Datafile 2. SampleSeq_R1_bismark.bt2.M-bias.txt

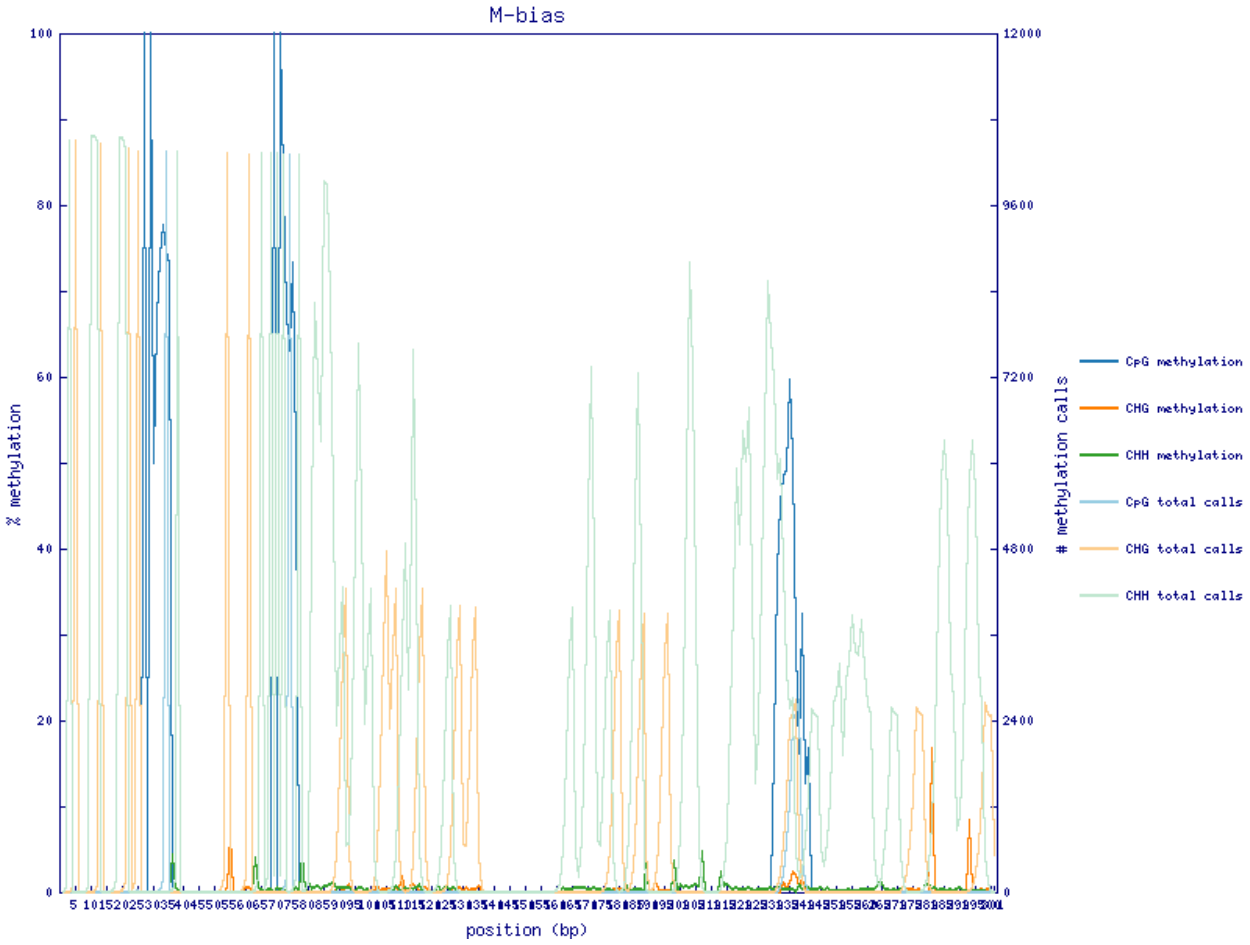
CpG context

==========

position		count methylated		count unmethylated	% methylation	coverage
1	0	0		0		
2	0	0		0		
3	0	0		0		
4	0	0		0		
5	0	0		0		
6	0	0		0		
7	0	0		0		
8	0	0		0		
9	0	0		0		
10	0	0		0		
11	0	0		0		
12	0	0		0		
13	0	3	0.00	3		
14	0	0		0		
15	0	0		0		
16	0	0		0		
17	0	0		0		
18	0	0		0		
19	0	0		0		
20	0	0		0		
21	0	0		0		
22	0	3	0.00	3		
23	0	0		0		
24	0	0		0		

25	0	13	0.00	13	
26	0	0		0	
27	1	0	100.00		1
28	0	0		0	
29	1	0	100.00		1
30	2	2	50.00		4
31	4	2	66.67		6
32	17	6	73.91		23
33	136	39	77.71		175
34	7696	2639	74.47		10335
35	11	4	73.33		15
36	0	2	0.00		2
37	0	0		0	

Supplementary Datafile 3. SampleSeq_R1_bismark_bt2.M-bias_R1.png



Supplementary Datafile 4: CpG_context_SampleSeq_R1_bismark_bt2.txt

Bismark methylation extractor version v0.17.0

M02849:171:000000000-ANHND:1:1101:12957:2144_1:N:0:102 + gi|902576|gb|U24170.1|HSU24170 499 Z

M02849:171:000000000-ANHND:1:1101:12957:2144_1:N:0:102 + gi|902576|gb|U24170.1|HSU24170 539 Z

M02849:171:000000000-ANHND:1:1101:12957:2144_1:N:0:102 + gi|902576|gb|U24170.1|HSU24170 703 Z

M02849:171:000000000-ANHND:1:1101:17290:2154_1:N:0:102 + gi|902576|gb|U24170.1|HSU24170 499 Z

M02849:171:000000000-ANHND:1:1101:17290:2154_1:N:0:102 + gi|902576|gb|U24170.1|HSU24170 539 Z

M02849:171:000000000-ANHND:1:1101:17290:2154_1:N:0:102 - gi|902576|gb|U24170.1|HSU24170 703 z

M02849:171:000000000-ANHND:1:1101:19751:2237_1:N:0:102 - gi|902576|gb|U24170.1|HSU24170 499 z

M02849:171:000000000-ANHND:1:1101:19751:2237_1:N:0:102 + gi|902576|gb|U24170.1|HSU24170 539 Z

Note: Methylation data on human *p21* SIE1 at CpG site -691 is an average of methylation call values at nucleotides -689, -690, -691 and -692. This spread might be due to a possible secondary structure formation at this transcription binding site.

Chromatin Immunoprecipitation (ChIP)

1. Preparation of Assay Plate and Cells

Prior to the assay, strip wells were washed with 150 µl of wash buffer, followed by the antibody buffer (100 µl) and then the Anti-Acetyl-Histone H3 antibody (1 µl). The wells were coverered with Parafilm M and incubated at room temperature for 2 hrs. Cell extracts were than prepared by washing HEK293 or NRK cells (3×10^6) cells with 10 ml of phosphate buffered saline (PBS), followed by centrifugation at 1,000 rpm for 5 min. The resulting cell pellet was resuspended in fresh culture medium (9 ml) containing 1% formaldehyde (final concentration) and incubated at room temperature for 10 min on an orbital shaker (50-100 rpm).

2. Cell lysis and DNA shearing

1.25 M Glycine solution (1 ml) was added to fixed cells, mixed and centrifuged at 1,000 rpm for 5 min. The supernatant was removed, and the cells washed with ice-cold PBS (10 ml). The cell pellet was lysed using the pre-lysis buffer (600 µl). The cell suspension was transferred to a 1.5 ml vial and incubated for 10 min on ice, vortexed vigorously for 10 sec and centrifuged at 5,000 rpm for 5 min. The supernatant was carefully removed and the lysis buffer (300 µl containing 3 µl protease inhibitor cocktail) added. The sample was incubated on ice for 10 min and vortexed occasionally. DNA was

sheared by sonication using a Branson Microtip probe sonicator, with 8 sets of 10 sonication pulses for 1 sec each were performed at level 3, followed by 1 set of a 10 sec pulse. The samples were placed on ice between each pulse. Cell debris was pelleted by centrifugation at 14,000 rpm for 10 min.

3. Protein/DNA Immunoprecipitation

Clear supernatant was transferred to a new 1.5 ml vial (often supernatant was stored at –80ºC at this step). The supernatant (60 µl) was diluted at a 1:1 ratio with ChIP dilution buffer. The incubated antibody solution was aspirated, and the strip wells washed three times with the antibody buffer (150 µl) by pipetting. The diluted supernatant (100 µl) was transferred to each strip well, covered and incubated at room temperature for 2.5 hrs on a rocking platform (50-100 rpm). The supernatant was aspirated, and the wells washed six times with the wash buffer (150 µl) at 100 rpm for 2 min. The wells were then washed once with 1X Tris-EDTA buffer (150 µl).

4. Reverse Cross-link and DNA Purification

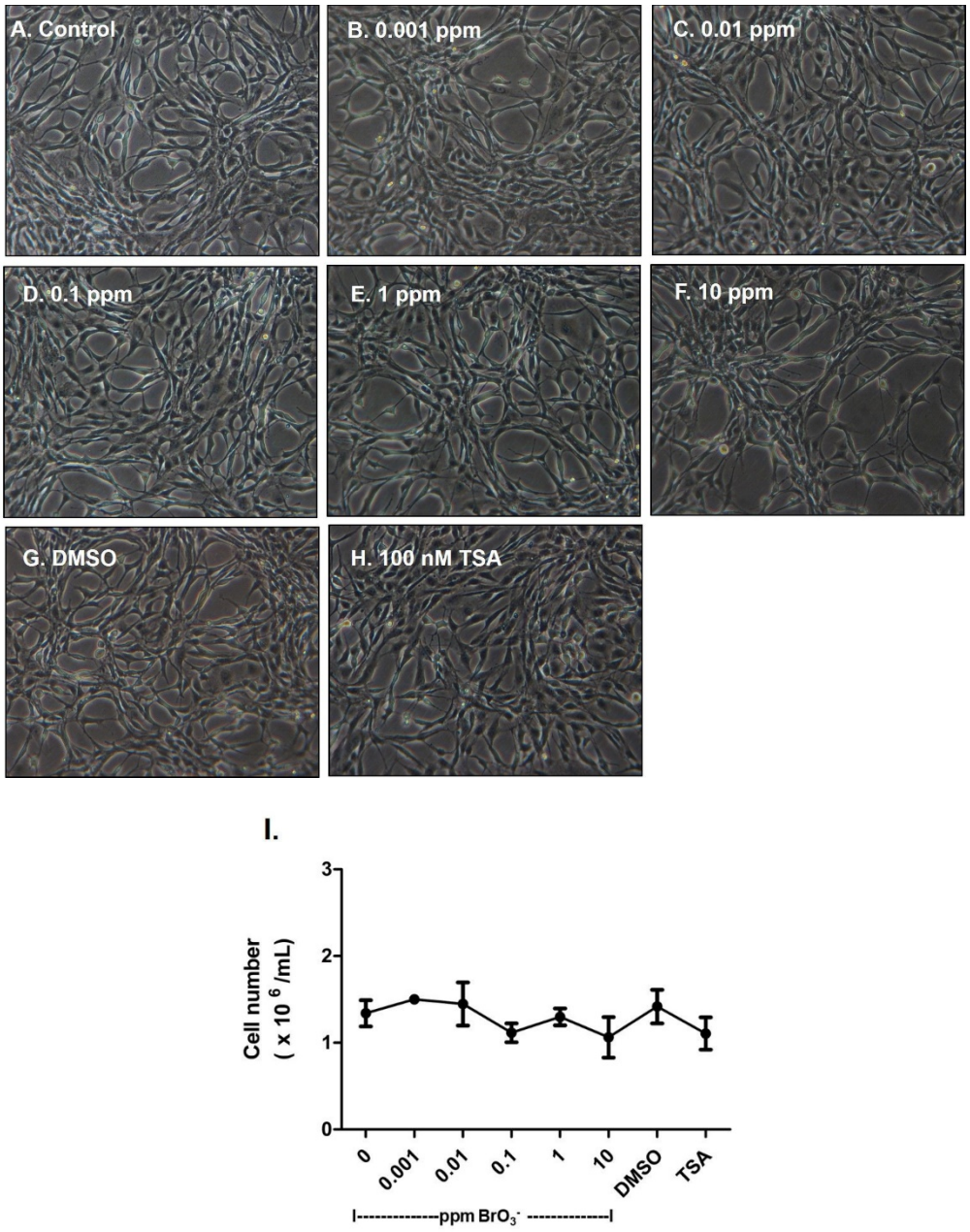
DNA release buffer (40 µl containing 1 µl proteinase K) was added to each well and the sample wells were covered with strip caps and incubated at 65ºC in a water bath for 15 min. Reverse buffer (40 µl) was added to the samples, which were mixed, re-covered and incubated in a65ºC water bath for 1.5 hr. Binding buffer (150 µl) was added to the samples and the solution transferred to a spin column placed in a 2 ml collection tube and centrifuged at 12,000 rpm for 20 sec. The column was washed with 70% ethanol

(200 µl) followed by 95% ethanol (200 µl) twice. The column was placed in a new 1.5 ml vial and purified DNA was eluted using the elusion buffer (15 µl).

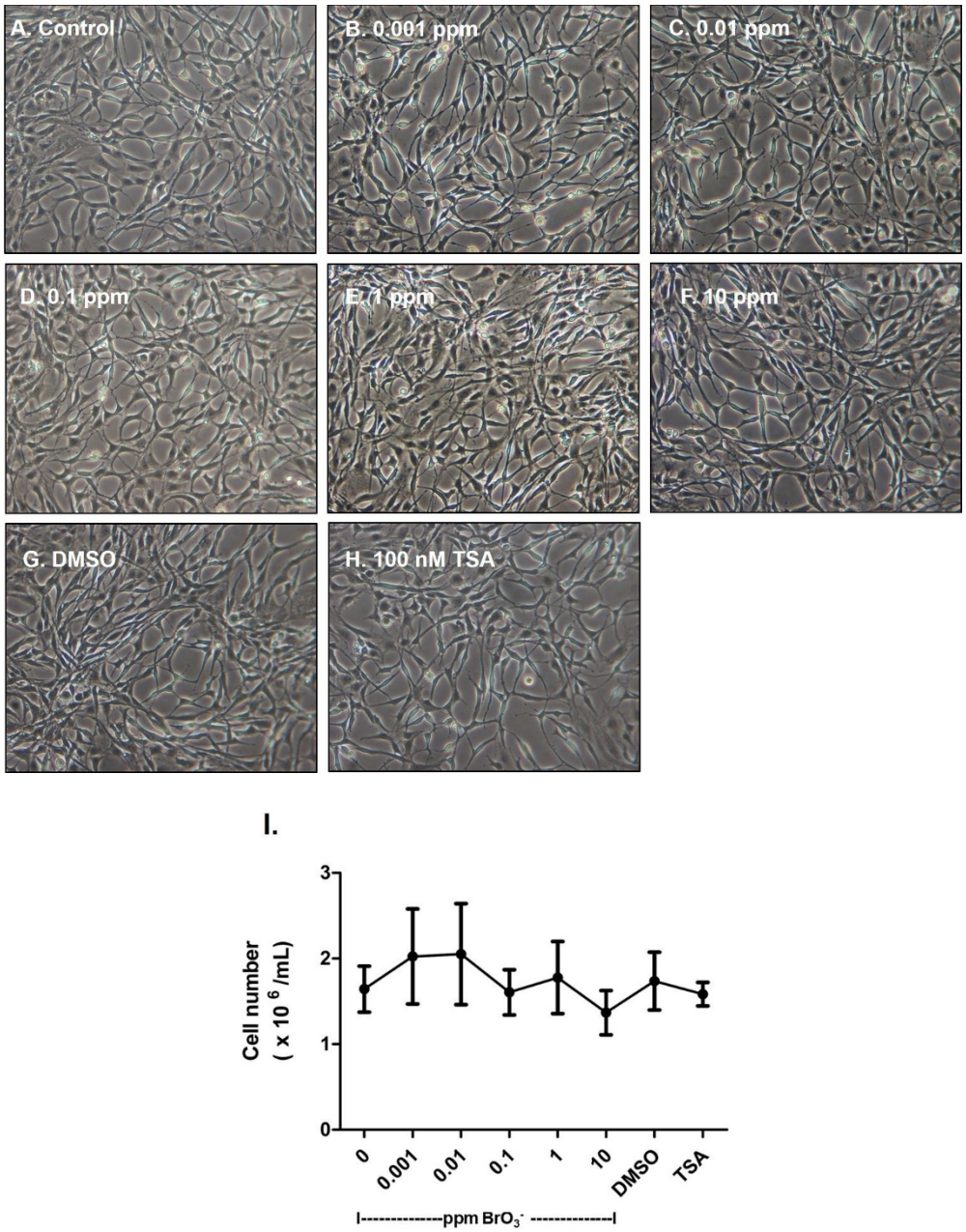
5. Polymerase Chain Reaction

Purified DNA linked to acetylated lysines 9 and 14 on H3 histones were subjected to PCR using primers described in **Table 2**. These include primers for the human and rat *p21* promoter region, as schematically shown in **Figure 2B**, and the housekeeping gene, glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). The 25 µl of PCR reaction mix contained 1X Promega master mix, 0.4 µM each of the forward and reverse primers and 40 ng of the immunoprecipitated DNA template. PCR was performed under the following conditions: 95^0C for 5 min, 35 cycles of 95^0C for 30 sec, respective T_a (**Table 2**) for 45 sec for the regions rp21-ChIP, rGAPDH, hp21-ChIP and hGAPDH, 72^0C for 45 sec and a final 72^0C for 10 min. The PCR products were then separated by electrophoresis on a 1% (w/v) agarose gel and visualized with ethidium bromide under a UV trans-illuminator. The band intensities were then quantified by densitometry using an Alpha Innotech FluorChem HD2 system (ProteinSimple, Santa Clara, CA) and normalized to the respective *GAPDH*.

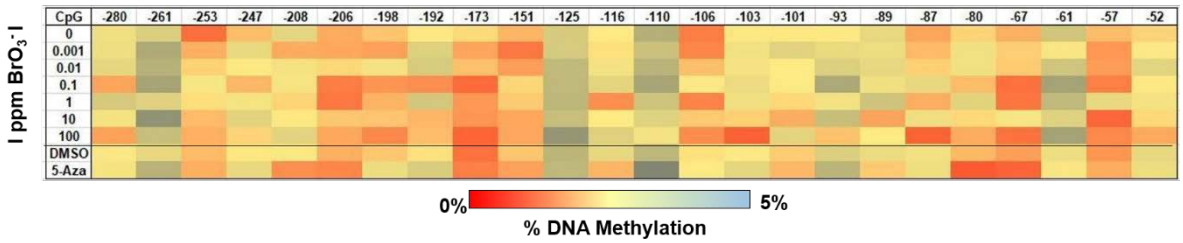
Supplementary Figure S1. TruSeqHT fusion primers for Illumina MiSeq platform. The forward and reverse portions of the primers are the locus-specific primers that are complimentary to the flanking regions. iTru5 and iTru7 primers with unique index combinations allow identifying the source samples in the sequence pool.



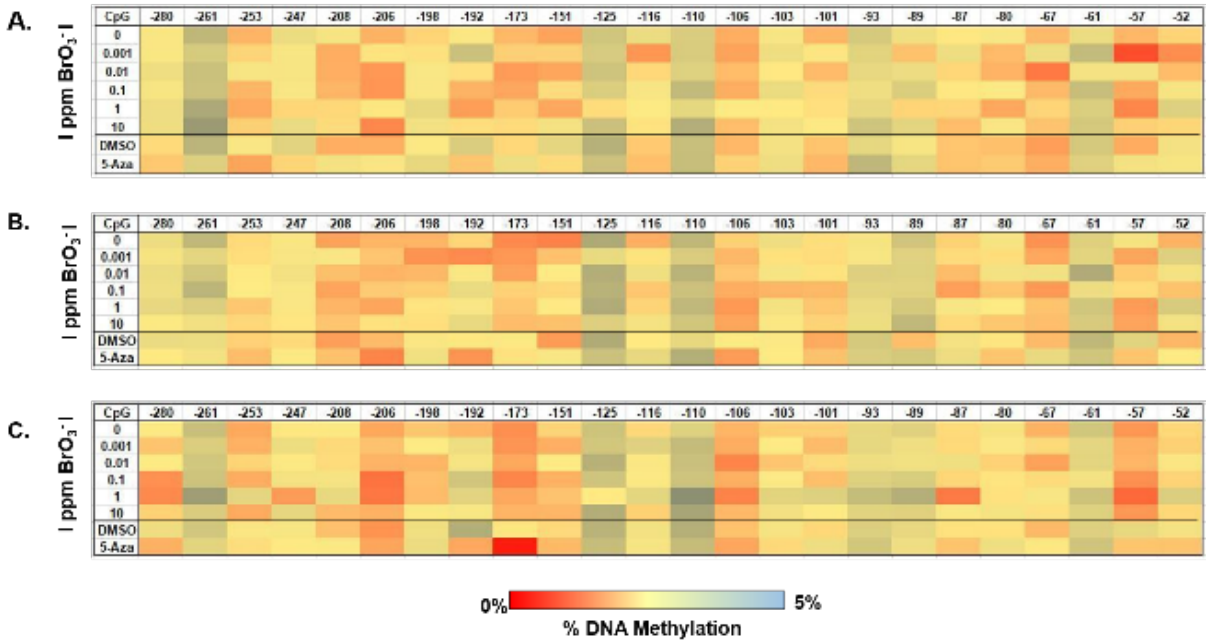
Supplementary Figure S2. Sub-chronic effect of BrO_3^- and epigenetic inhibitors on NRK cell morphology and number. NRK cells were sub-chronically exposed to 0-10 ppm BrO_3^- (**A-F**), vehicle control DMSO (**G**) or 100 nM TSA (**H**) for 18 days (P6). The cell number data in **I** are represented as mean ± SEM of three separate passages (n = 3).



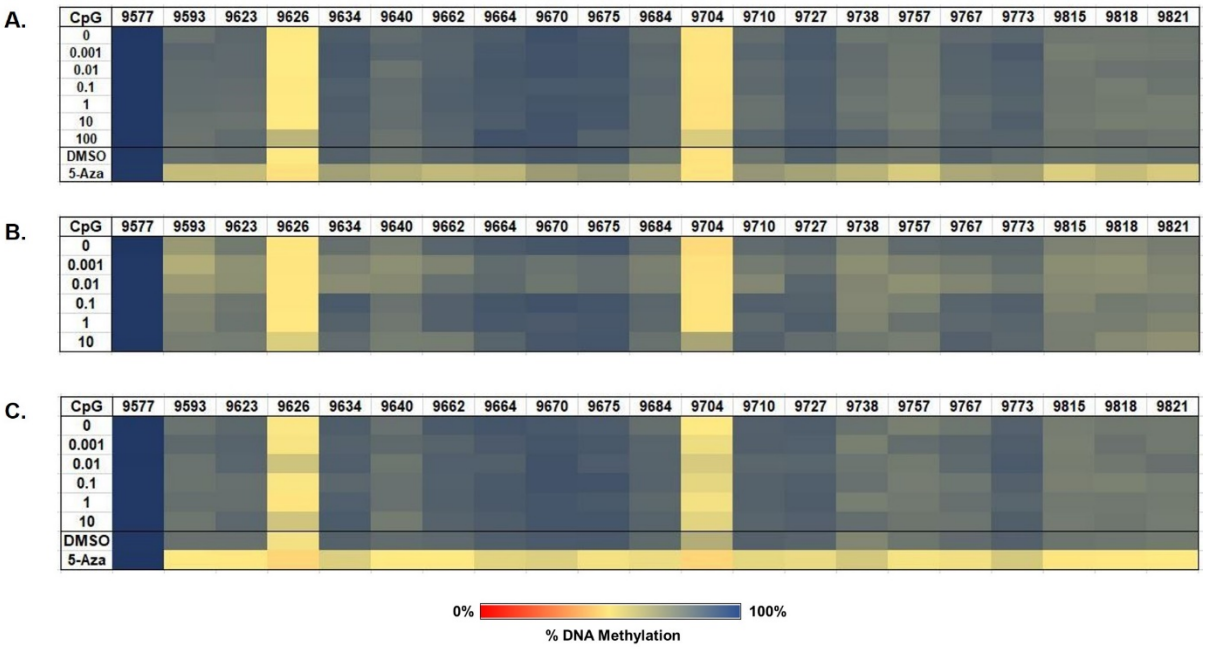
Supplementary Figure S3. Effect of BrO_3^- and epigenetic inhibitors on NRK cell morphology and number after withdrawal. NRK cells were sub-chronically exposed to BrO_3^- for 18 days and then allowed to recover (**A-F**), vehicle control DMSO (**G**) or 100 nM TSA (**H**) for 9 days (P9). The cell number data in **I** are represented as mean ± SEM of three separate passages (n = 3).



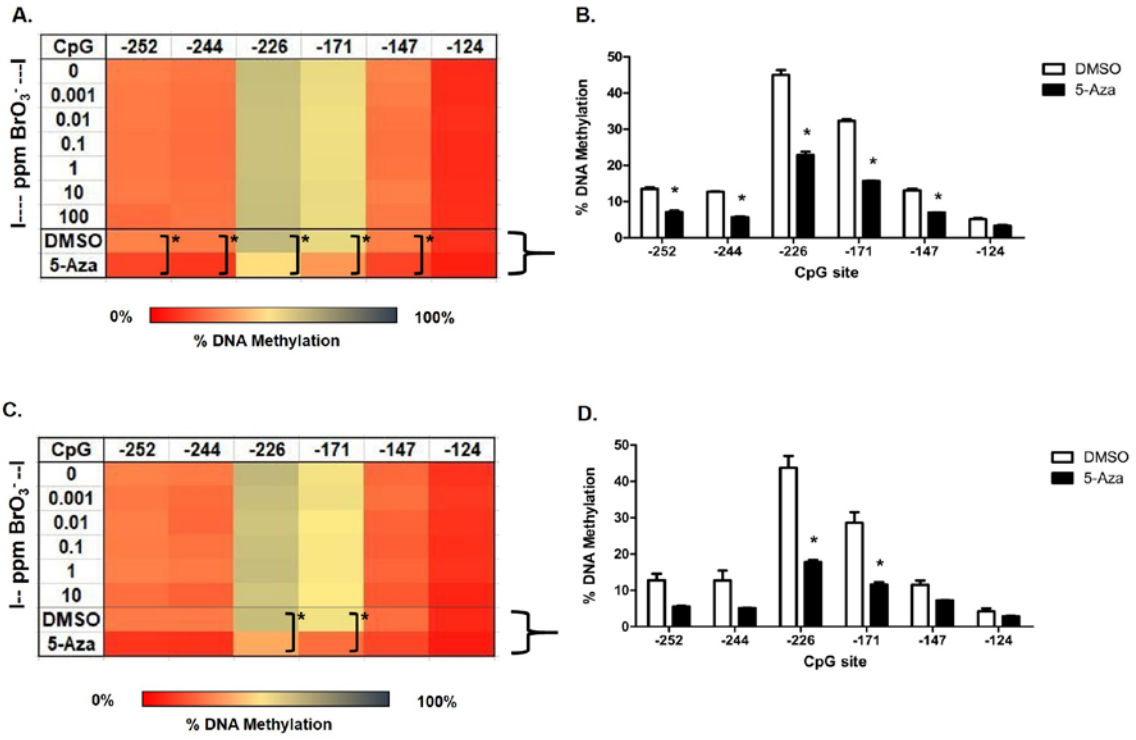
Supplementary Figure S4. Effects of acute BrO_3^- exposure on the percent DNA methylation of cytosine residues at the transcription start site of human *p21* promoter (hp21-TSS). Heat-map of the site-specific percent DNA methylation changes as determined by TGBS in the hp21-TSS. The first row represents the position of the cytosine in the CpG dinucleotide context relative to the TSS. The first column shows the treatments of HEK293 cells water (0), 0.001 to 100 ppm BrO_3^-, DMSO (vehicle control for 5-Aza) or 40 μM 5-Aza (positive control) for 72 hrs. Heat map intensity is shown in the sidebar with deep red indicating percent methylation value towards zero and pale blue indicating relatively higher methylation of 5%.



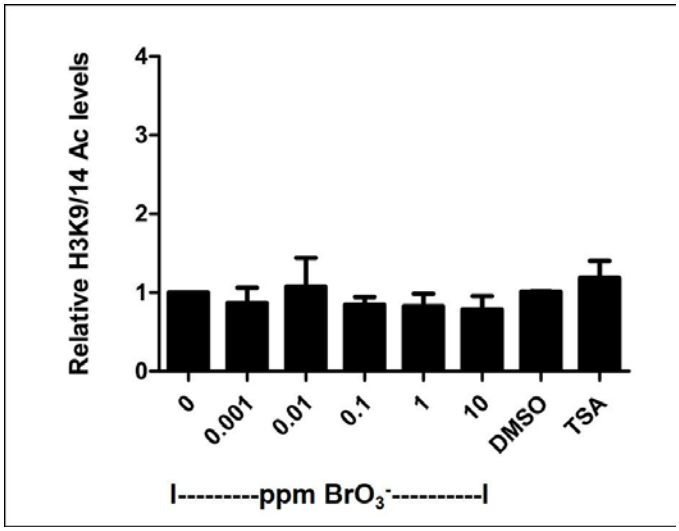
Supplementary Figure S5. Effects of sub-chronic BrO_3^- exposure on the percent DNA methylation of cytosine residues at the transcription start site of human *p21* promoter (hp21-TSS). Heat-map of the site-specific percent DNA methylation changes as determined by TGBS in the hp21-TSS. The first row represents the position of the cytosine in the CpG dinucleotide context relative to the TSS. The first column shows the exposure of HEK293 cells to water (0), 0.001 to 10 ppm BrO_3^-, DMSO (vehicle control for 5-Aza) or 40 μM 5-Aza (positive control) for **A)** 9 days, **B)** 18 days and **C)** 9 days of withdrawal. Heat map intensity is shown in the sidebar with deep red indicating percent methylation value towards zero and pale blue indicating relatively higher methylation of 5%.



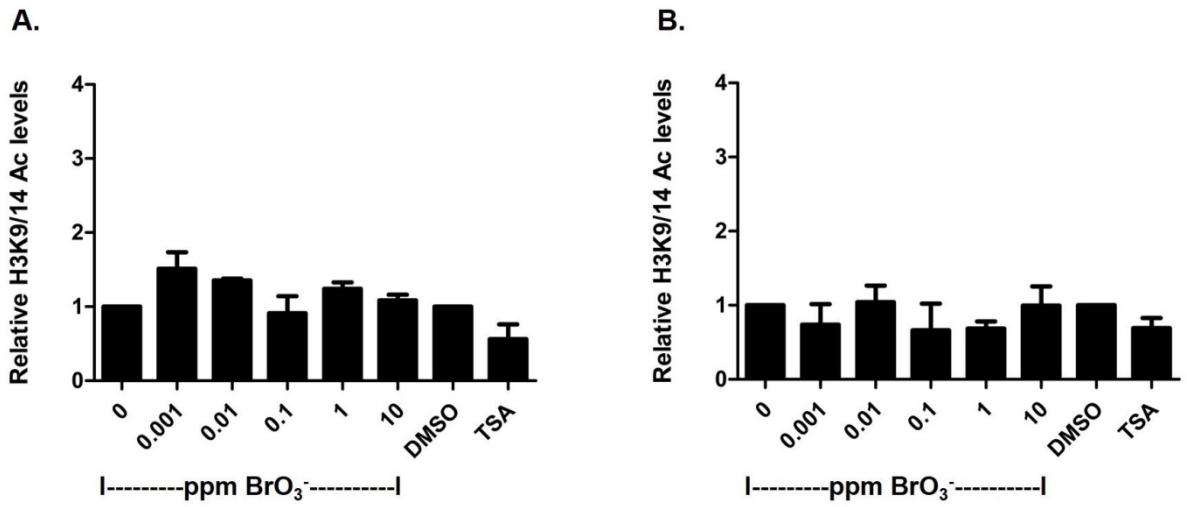
Supplementary Figure S6. Effects of sub-chronic BrO_3^- exposure on the percent DNA methylation of cytosine residues at the coding region of rat *p21* (rp21-coding). Heat-map of the site-specific percent DNA methylation changes as determined by TGBS in the rp21-coding. The first row represents the chromosomal position of the cytosine in the CpG dinucleotide context. The first column shows the exposure of NRK cells to water (0), 0.001 to 10 ppm BrO3-, DMSO (vehicle control for 5-Aza) or 40 μM 5-Aza (positive control) for **A)** 9 days, **B)** 18 days and **C)** 9 days of withdrawal. Heat map intensity is showed in the sidebar with deep red indicating percent methylation value towards zero and pale blue indicating relatively higher methylation of 100%. Data are represented as the mean ± SEM values of three independent experiments (n=3).



Supplementary Figure S7. Sub-chronic effects of BrO_3^- on the percent DNA methylation of cytosines in the TSS site in rat *p21* promoter. NRK cells were treated with water (vehicle control for BrO_3^-), 0.001 to 100 ppm BrO_3^- , DMSO (vehicle control for 5-Aza) or 40 µM 5-Aza (positive control) for 9 days or P3 (**A-B**). The cells were further treated for 9 days and recovered by discontinuing the exposures for 9 days or P9 (**C-D**). The first row in the heat-map represents the position of the cytosine in the CpG dinucleotide context relative to the TSS. Heat map intensity is showed in the sidebar with deep red indicating percent methylation value towards zero and deep blue indicating towards 100%. Data are represented as the average percent DNA methylation of three separate passages (n=3) as determined by TGBS analysis. The effects of 5-Aza are emphasized in the bar graphs for the respective exposure times. Data are represented as mean ± SEM of three different passages (n=3). *P<0.05 compared with with DMSO.



Supplementary Figure S8. Effects of BrO_3^- on the acetylation of lysine 9 and 14 on histone H3 (H3K9/14 Ac) of the rat *p21* promoter region 3 days after withdrawal (P7) from sub-chronic exposure of NRK cells to 0.001 – 100 ppm BrO_3^-. Data represented as relative fold increase in H3K9/14 Ac levels as normalized to *GAPDH* as determined using ChIP assay. Data are expressed as mean ± SEM of three different passages (n=3).



Supplementary Figure S9. Sub-chronic effects of BrO_3^- on the acetylation of lysine 9 and 14 on histone H3 (H3K9/14 Ac) of the human *p21* promoter region. HEK293 cells were exposed to 0.001 – 100 ppm BrO_3^- for 9 days or P3 (**A**) and 18 days or P6 (**B**). Data represented as the relative fold increase in H3K9/14 Ac levels as normalized to *GAPDH* as determined using the ChIP assay. Data are represented as mean ± SEM of three different passages (n=3).