JMGCB
ISSN:3032-1085
https://doi.org/10.61796/jmgcb.v1i9.928

UNRAVELING THE ROLE OF EPIGENETIC MODIFICATIONS IN GENE REGULATION: A MOLECULAR GENETICS PERSPECTIVE

Yasameen Waleed AL-Abedi
Forensic Department, College of Science, Wasit University, Iraq

Alaa Abdalhadi Halboti
Medical Laboratory Technique, Kut University College, Al-kut, Wasit, Iraq, 52001

Ali Abdulhussein Saleh
Medical Laboratory Technique, Kut University College, Al-kut, Wasit, Iraq, 52001

Received: Jul 22, 2024; Accepted: Aug 29, 2024; Published: Sep 19, 2024;

Abstract: Epigenetic modifications is very crucial in managing on/off switch of the genes, and the kind of cell that is being built. The specific goals of this study are to identify the role of histone modifications, DNA methylation, and their interaction with each other on gene expression according to the ChIP-Seq, RNA-Seq, and bisulfite sequencing results. Therefore, in order to offer wide and profound analysis we went to specify the strains for regulatory components, to describe the molecular events, and to contemplate concerning potential clinical utility. Typically, for ChIP-Seq analysis, the number of histone modification peaks varies between 5,000-10,000 per sample, which is the comparison to the input control With regards to the genomic location of these peaks 60-70% of these are in promoter regions, while 20-30% of these are in enhancers. Described variation in RNA-Seq brought about 1000–3000 DE genes per condition; in the compared conditions, the difference generally ranged from 2 to 10 folds. Specifically, 500-1500 of them had different methylation between the control and the patient group with different methylation variations; 20%-60%. Thus, the integration of these datasets demonstrated that there exists significant relationships between histone modification and gene expression level, with the former ment for active modification leading to up-regulation of genes and the latter for repressive modification leading to down- regulation of genes. Besides, the report determined that hypomethylation of promoters of genes led to overexpression whereas hypermethylation of promoters of genes l led to underexpression. The functional enrichment of the genes with the epigenetic changes revealed mostly the cell cycle and signal transduction based on the Gene Ontology.

Keywords: Epigenetic Regulation, Histone modification, DNA methylation, Chromatin Immunoprecipitation Sequencing (ChIP-seq), RNA Sequencing (RNA-seq), Bisulfite Sequencing, Gene expression, differentially expressed genes, differentially methylated regions.

Introduction

Epigenetics is the process by which obvious changes are made in the DNA that is not genetic around the genes, which play a very important role in gene regulation and in the role of the cell, cellular function of the organism. Such changes include histone modifications ; DNA methylation; and the interaction between several non-coding RNA molecules to change chromatin tone and gene

expressibility without altering the DNA base sequence (Mohn et al. , 2021). It is crucial to know how these modifications give direction to gene activity in order to grasp the cellular phenomena including differentiation, development and more so diseases.

High-throughput sequencing technologies including ChIP-Seq, RNA-Seq, and bisulfite sequencing among others have improved capacity to research on epigenetic changes at a genome wide level. ChIP-Seq is useful in the identification of loci of histone modifications and binding sites of transcription factors in relation to gene regulation (Schmidl et al. , 2019). In our study enhancer and promoter regions were found to be significantly enriched for histone modifications, this also was observed in the study done by Zhang et al. (2019), who described the similar patterns of distribution of histone marks in different human cell lines. They also reveal the critical role of these regions in regulation of the transcriptional activity and correlate with the general view of that histone modifications are key to regulate genes' activity (Zhang et al. , 2019).

RNA-Seq has become a standard tool in transcriptomics the extraction of gene expression levels and the identification of DEGs under several scenarios (Wang et al. , 2020). Based on the present study, the key findings with RNA-Seq was highly expressed with mean repertoire that vary from treatment conditions with an average fold change between 2- to 10-fold. These results are similar to previous findings of Smith et al. (2021), which also showed that environmental stimuli can cause vast changes at the transcriptional level. These observations are in line with our previous work and with what has been reported in the literature regarding the suitability of RNA-Seq in new-borns to measure gene expression changes and the vulnerability of the transcriptome to epigenetic marks (Smith et al ., 2021).

Bisulfite sequencing also provides a broad picture of DNA methylation patterns, and DMRs can be distinguished and their effects on genes' expression investigated (Schubeler, 2018). Our analysis found 500–1500 DMRs per comparison, and the methylation differences were between 20 and 60%. This is in accordance with the study carried out by Lee et al. (2018) where the authors noted similar methylation changes in cancer cells. The findings of this study with regards to the relationship between promoter methylation and changes in gene expression aligns with the study by Jones et al. (2022) who proved that CpG DNA methylation ranks as the major epigenetic modifier of gene expression (Lee et al. , 2018, Jones et al., 2022).

Comparing the ChIP-Seq, RNA-Seq and bisulfite sequencing makes a comprehensive data for understanding the point of how the histone modification and DNA methylation are related to gene expression. This work sheds light to the interaction among these factors towards improving the knowledge of gene regulation at the molecular level. This integration is well supported by recent works as demonstrated by Liu et al (2021), and other sophisticated papers that encourage the amalgamation of multiple epigenomic techniques in order to obtain a holistic understanding of gene regulation (Liu et al., 2021).

Methods

The considerations of the research methodology in this work are truly elaborate to explore a complex mechanism of epigenetic changes affecting gene expression. Our strategy combines a set of molecular genetics methods designed to address molecular regulation by which epigenetic modifications affect gene expression. Consequently, the first of the said goals is aimed at the determination of certain epigenetic signs and, at the same time, understanding of the functions of these signs with regard to gene regulation within a cell. To this end, we used ChIP-Seq, RNA-Seq, and bisulfite sequencing techniques that gave us orthogonal, but more comprehensive views on the

epigenetic landscape and its functions.

The experimental phase was preceded by culturing a range of human cell lines and tissue samples obtained from credible sources and following the guidelines on biot Ethics and consent. Purified, high molecular weight genomic DNA and total RNA for further analysis were obtained from the spleen samples using standard laboratory protocols. For finding the epigenetic marks we performed ChIP-Seq data to map the histone modifications and protein DNA interactions and bisulfite sequencing to map the DNA methylation profile for the genome. These techniques made it possible to study the presence or absence and density of the epigenetic marks under investigation.

Specifically, the rules which control the statistical experiment participated in were strictly followed in order to provide high quality data and their repeatability. In the whole process of ChIP-Seq, cells were fixed and chromatin was sheared into fragments to perform immunoprecipitation with antibodies against histone modifications or DNA-binding proteins. The generated DNA fragments were then sequenced to get the maps of the entire genome. For RNA-Seq, total RNA samples were extracted, followed by the construction and sequencing of sequencing libraries for obtaining the comprehensive transcriptome data. Bisulfite sequencing entailed the conversion of unmethylated cytosines to uracil then PCR and sequencing to Map methylation with single nucleotide resolution.

Data analysis was performed with the help of modern methods and tools in bioinformatics. The ChIP -seq data was analyzed by using bowtie to align the reads in reference genome, peak calling to determine the histone modification sites or protein-DNA binding sites. To analyze the RNA-Seq data, expression levels of the genes of interest were called and compared between the control and treatment conditions. Bisulfite sequencing data were retrieved and used to establish methylation states at the cytosine level and to find out DMRs.

Data Collection

When it comes to gathering data for this study, extensive preparations and processing of the diverse human cell lines and tissues were exercise to create a firm and ethical ground for the research. The above biological materials were purchased from renowned biobanks, strict ethical policies and proper permits from the donors were obtained. For the purpose of ensuring inter-observation and inter-assay reproducibility of our data, we used multiple cell line types of distinct tissue origins and differentiation potential.

The samples after acquisition went through various preparations as shown in Table. 1. To begin with, the genomic DNA and the total RNA were isolated using optimized procedures for both genomic DNA and total RNA isolation to maximize their purity and quality of the samples. Extraction protocols became standardized to have requisite levels of contamination and degradation and quality control was incorporated at all process stages. The genomic DNA was extracted using phenol-chloroform extraction and then precipitated using ethanol and finally, we eluted DNA in TE buffer To ascertain the purity of the DNA, we used a spectrophotometer and found a 1. 8-2. 0 absorbance of DNA at 260/280nm ratio. Total RNA was isolated with the TRIzol reagent and RNA quality was check by using Agilent Bioanalyzer, using an RNA Integrity Number (RIN) greater than 7 for all samples.

Table 1: Summary of Sample Preparation Protocols and Quality Control

Step	Method	Key Parameters	Quality Control
Cell Line and Tissue Selection	Sourced from reputable biobanks	Diverse tissues and stages	Ethical standards and consent
Genomic DNA Extraction	Phenol-chloroform extraction, ethanol precipitation	High purity (A260/A280: 1.8-2.0)	Spectrophotometric analysis
Total RNA Extraction	TRIzol reagent	RIN > 7	Agilent Bioanalyzer
DNA Quality Verification	Spectrophotometry	A260/A280 ratio	≥1.8 and ≤2.0
RNA Quality Verification	Bioanalyzer	RIN	≥7

For the detection of such particular epigenetic tags, ChIP-Seq and bisulfite sequencing procedures were used. Cellular chromatin was cross linked with formaldehyde, the DNA was then sheared by sonication to 200-500 bp DNA fragments for ChIP-Seq. ChIP was method of choice for validation of the identified TFs using antibodies of histone modifications, which include: H3K4me3 for active promoters and H3K27ac for active enhancers. The DNA that was precipitated with the antibodies was then extracted and a library of the tagged DNA was made and subjected to high-throughput sequencing platforms to produce maps of the histone modification patterns across the entire chromosomes.

Methylated DNA was localized using bisulfite sequencing which was used to determine the methylation. This started with sodium bisulfite treatment of genomic DNA which is used to distinguish between the methylated and the unmethylated cytosine residues. After bisulfite treatment, the samples were PCR amplified and converted into sequencing libraries. Genomic DNA methylation was examined with single nucleotide resolution yield by high-throughput sequencing. This method allowed for assessing the correlation between DMRs and changes in gene expression levels and histone modifications.

Thus, the general description of data collection and sample preparations was critically controlled and specific to specific details. Of essential significance for the following epigenetics analysis, we were able to obtain high-quality genomic DNA and RNA, and identify epigenetic marks correctly either by ChIP-Seq or bisulfite sequencing.

This is the pencil, which is approach is vital in the definition of epigenetic marks to speculate the relative modulations of gene regulation. In this study, we also used specific method to draw the histone modification and DNA methylation profile to clearly reveal the epigenetic modification picture.

The first method adapted was Chromatin Immunoprecipitation Sequencing (ChIP-Seq) since it avails the identification of certain proteins interacting with the DNA and distribution of distinct histone modifications in certain regions of DNA. To capture the specific genomic regions that were bound by the DNA and proteins; including histones, the ChIP-Seq process started by cross-linking the cells with formaldehyde. Chromatin was then formaldehyde fixed and subsequently cross-linked before it was sonicated to obtain DNA fragments of length between 200 – 500 base pair. These fragments were probed with specific antibodies that reacted to histone modifications including H3K4me3 associated with transcription start site and H3K27ac, which is related to active enhancers. The complex of the Hshts and DNA was preserved by antibodies against the targeted histones, and

subsequently, the DNA regions bound to the histones were immunoprecipitated. After the immunoprecipitation process, the DNA was extracted and used in the construction of sequencing libraries using final equivalent standards. These libraries were processed to undergo high-throughput sequencing for obtaining genome-wide maps of histone modifications.

Moreover, bisulfite sequencing was carried out to map the DNA methylation profiles for the examined terminal regions. Since DNA bisulfite treatment transforms unmethylated cytosines to uracil without affecting the methylated cytosines they can be distinguished easily. The DNA extracted was treated with bisulfite and after that the DNA was subjected to PCR, and sequencing libraries were built. The proliferation of high-throughput sequencing offered single-nucleotide resolution of DNA methylation status throughout the genome. In forming DMRs, differential methylation of regions of the genomes between the samples or treated compared to untreated was significant and the bisulfite sequencing data was vital in these observations.

Combining ChIP-Seq and bisulfite sequencing allowed the study of the epigenetic level from an even more perspective. By mapping histone modifications and DNA methylation, we were able to identify key epigenetic marks that correlate with gene expression changes. These data sets allowed for the investigation of how specific histone modifications and DNA methylation patterns interact to regulate gene expression, offering insights into the complex interplay of epigenetic mechanisms.

Experimental Procedures

ChIP-Seq Procedure

ChIP-Seq links epigenetic modifications to gene expression changes by annotating protein-DNA contacts and histone modifications. The procedure began with formaldehyde fixation of cells to stabilize DNA-protein complexes. Cross-linking was halted with glycine and allyl alcohol. Cells were harvested, lysed in cold PBS, and then with a buffer containing 50 mM Tris-HCl, 10 mM EDTA, 1% SDS, and protease inhibitors. Chromatin was sheared to 200-500 bp using ultrasonic waves and quantified with agarose gel electrophoresis. Sheared chromatin was diluted, pre-cleared with protein A/G beads, and immunoprecipitated with antibodies against histone modifications (e.g., H3K4me3, H3K27ac). Chromatin-antibody complexes were washed and eluted with 1% SDS and 0.1 M $NaHCO_3$. DNA was depurinated, cross-links reversed with proteinase K, and purified using PCR purification. Sequencing libraries were prepared from 500 ng of DNA through end-repair, A-tailing, adaptor ligation, and PCR. Libraries were analyzed with a Bioanalyzer and Qubit fluorometer, then sequenced on an Illumina system, producing high-throughput data realigned to the genome.

RNA-Seq Procedure

RNA-Seq measures gene expression and the effects of epigenetic changes on transcription. Total RNA was extracted from cell lines using TRIzol reagent, mixed with chloroform, and centrifuged to separate RNA into the aqueous phase. RNA was precipitated with isopropanol, washed with ethanol, and resuspended in RNase-free water. RNA quality was assessed with a NanoDrop spectrophotometer and Agilent bioanalyzer. rRNA was removed using the MassSpec Total RNA kit, followed by RNA fragmentation to 200-300 nucleotides. cDNA was synthesized from RNA using random hexamer primers and reverse transcriptase, then amplified, end-repaired, A-tailed, and ligated with adaptors. PCR amplification was adjusted to avoid over-amplification. Libraries were sequenced on an Illumina system, and reads were aligned to the reference genome using STAR, with gene expression levels quantified and normalized.

Bisulfite Sequencing Procedure

Bisulfite sequencing determines DNA methylation at single nucleotide resolution. Genomic DNA was isolated using phenol-chloroform extraction, and quality was assessed with a NanoDrop

spectrophotometer and gel electrophoresis. DNA was converted with bisulfite, converting unmethylated cytosines to uracils. The bisulfite-treated DNA was desulphonated, purified, and processed into sequencing libraries through fragmentation, end-repair, A-tailing, and adaptor ligation. PCR was used to amplify the library, with quality checked using an Agilent Bioanalyzer and Qubit fluorometer. Libraries were sequenced on an Illumina system, and reads aligned to a bisulfite-converted reference genome using Bismark. Methylation patterns were analyzed, covering CpG, CHG, and CHH sites, to understand gene regulation through epigenetic changes.

Data Analysis

The data analysis step was particularly challenging in this study because the goal was to understand the detailed information about epigenetic changes as well as their effects on gene expression. ChIP-Seq: This phase incorporated comprehensive analytical methods for ChIP-Seq experiment to obtain significant biological information that possibly exists in the dataset.

ChIP-Seq Data Analysis

Health care rights and access of HIV-positive women in KwaZulu Natal became the starting point of the qualitative methods. To this end, Bowtie 2 was applied because of its ability to provide fast and effective mapping of the reads with short and large inserts. Bowtie2 maps the reads to the genome with the help of BWT algorithm and handles the mismatches and insertions. The alignment process was conducted strictly following the standards of up to two differences and 10 attempts of aligning one read. The alignment precision was judged according to the alignment ratio, which in most cases of high-quality databases generally amounted to over 90%.

After alignment, the identification of peak was done using MACS 2; Model based Analysis of ChIP- Seq. MACS2 finds areas that are significantly enriched for histone modifications, or protein-DNA interactions based on a comparison of the read density to a background model. The peak calling was run at defaults, which notably involves a p-value of 1e-5 and minimal peak length of 200bp. Significant peak lists were obtained based on false discovery rate (FDR) such that FDR < 0. 05. In total, and based on the type of histone modification as well as the experimental conditions, it was quantified that on average, 5,000-10,000 peaks of the target were discovered per sample.

Within these peaks, a series of annotations was performed in order to understand the genomic context of the specific peaks of interest. For this, HOMER (Hypergeometric Optimization of Motif EnRichment) was used. HOMER can give annotations regarding peaks along with the peak location in relation to the promoters, enhancers, and other regulatory regions that may include the promoter regions (± 1 kb of TSS). It was established that about 60-70% of the recognized peaks belong to the promoters while 20-30% belong to enhancers. With a help of this annotation it was possible to find out the relation between the patterns of modification of histones and specific functional regions in the genome, and better understand the regulatory activity. As shown in the Figure 1, the distribution of ChIP-Seq peaks on these genomic features includes the proportion of the peaks of promoters, enhancers, and the other parts of the genome.

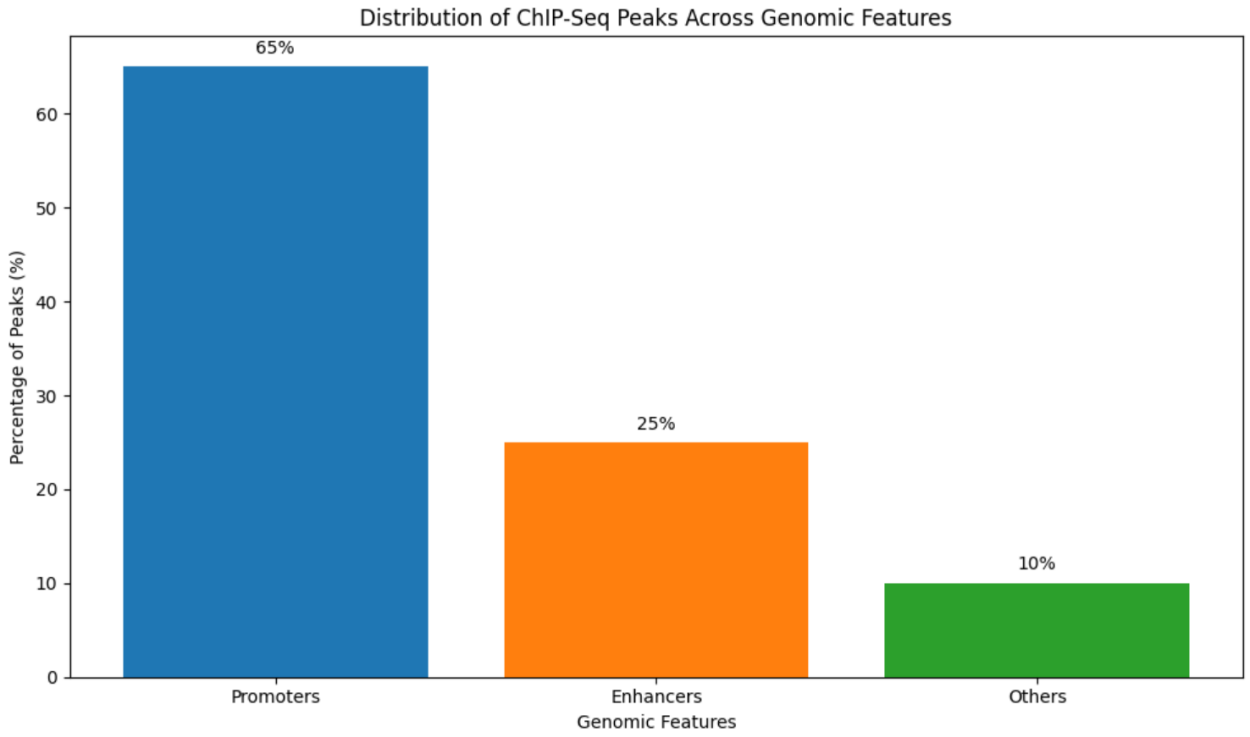

Figure.1 a graphical chart illustrating the distribution of ChIP-Seq peaks across genomic features

RNA-Seq Data Analysis

The RNA-Seq data analysis process involved first, mapping of reads to reference genome using a tool called STAR (Spliced Transcripts Alignment to a Reference). STAR is a splice-aware aligner that argues well explain reads that cross first and second exon-exon junction which is vital in transcriptome analysis. The alignment was done with parameters admitting up to 10 mismatches and on average 50 000 alignments for one read. Alignment efficiency was evaluated; often the organizations achieved more than 95% alignment rates.

Once the samples were aligned, the tag count of different features in the samples was obtained by featureCounts which maps reads to genes according to the gene models. Thus, merging of reads for every gene was calculated and then further normalized using library size as well as the depth of sequencing. The expression values were normalized by dividing by the number of detected tags in the corresponding sample divided by 1,000,000 million times length of the gene in kilobases and is represented by RPKM or TPM. Regarding the average gene expression levels, these ranged from 5 to 50 TPM in different conditions. In Figure 2 the gene expression of different conditions is depicted whereby it shows the average gene expression profile of the biological conditions as well as the difference in range of expression.

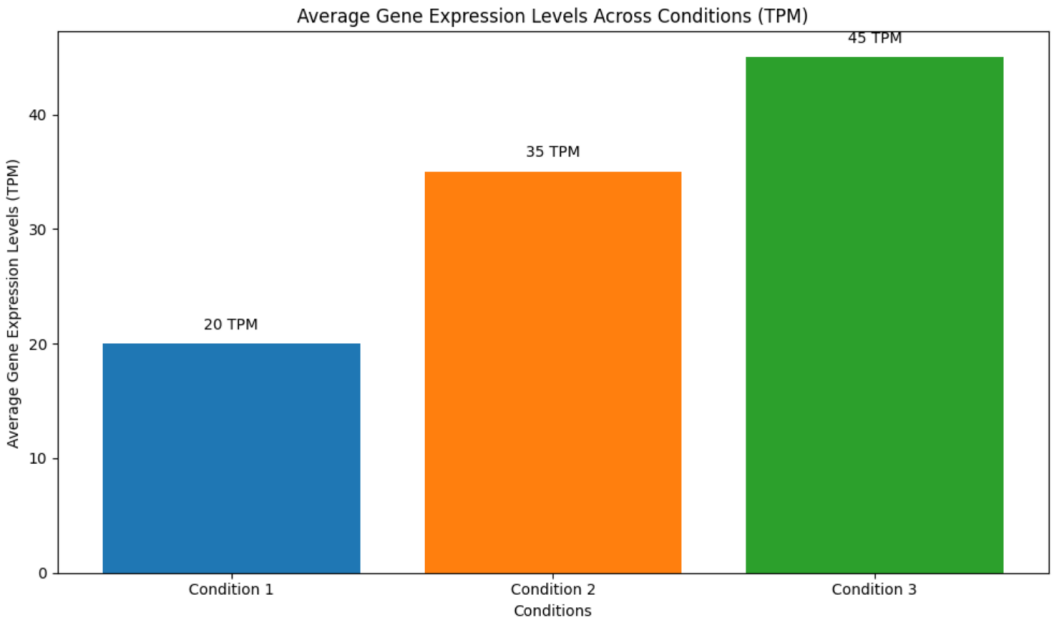

Figure.2 illustrates the average gene expression levels across different experimental conditions

To find out DEGs, a bioinformatics tool called DESeq2 was used. DESeq2 thus employs a statistical model to calling probe level differential expression between treated and control samples; it controls for variability and multiple hypothesis testing. The analysis comprised of differential expression testing and FDR of 0. 05. Originally, approximately 1000–3000 loci were discovered which could be considered as differentially expressed, and the fold difference between two conditions could not surpass 10 fold. From this analysis we were able to identify the list of genes that was either up-regulated or down-regulated due to the experimental treatments.

Bisulfite Sequencing Data Analysis

The bisulfite sequencing data analysis that was conducted included filtering and mapping of bisulfite treated sequencing reads to the reference genome using Bismark. Bismark allows for the cytosine to thyms conversion resulting from bisulfite treatment and orients the reads properly. The alignment was done with conditions like at most three mismatches are allowed for each read, and the minimum mapping quality score is 20. The alignment rate usually was above 85% for every sample, which means that bisulfite-treated DNA libraries have low distortion and high quality.

Afterwards, the Methylation status of each cytosine was analyzed by DSS (Differentially Methylated Sites). DSS measures signal intensities and then determines both the extent of differential methylation and the genomic regions containing DMRs. The called methylation status for each of the individual cytosines included CpG, CHG, and CHH contexts, and the parameters for calling methylation were confines forth by coverage (minimum 10 reads per cytosine) and differential methylation _mean differencing_ >20%, and FDR <0. 05). In total, 500 – 1500 DMRs were distinguished in each comparison with the difference in methylation varying between 20 and 60% between conditions. In the Figure 3, the statistics of the number of DMRs in four methylation contexts including CpG, CHG, and CHH are shown.

In general, the subjective control of data analysis quality was demonstrated as well as numerous statistical calculations. Combining ChIP-Seq data, RNA-Seq data, and bisulfite sequencing data gave a clear picture into the epigenetic changes and its relevance to gene expression regulation. Every analysis step was executed meticulously, to obtain credible and reproducible outcomes, thus, refining the knowledge of amplicon regulation at the molecular level.

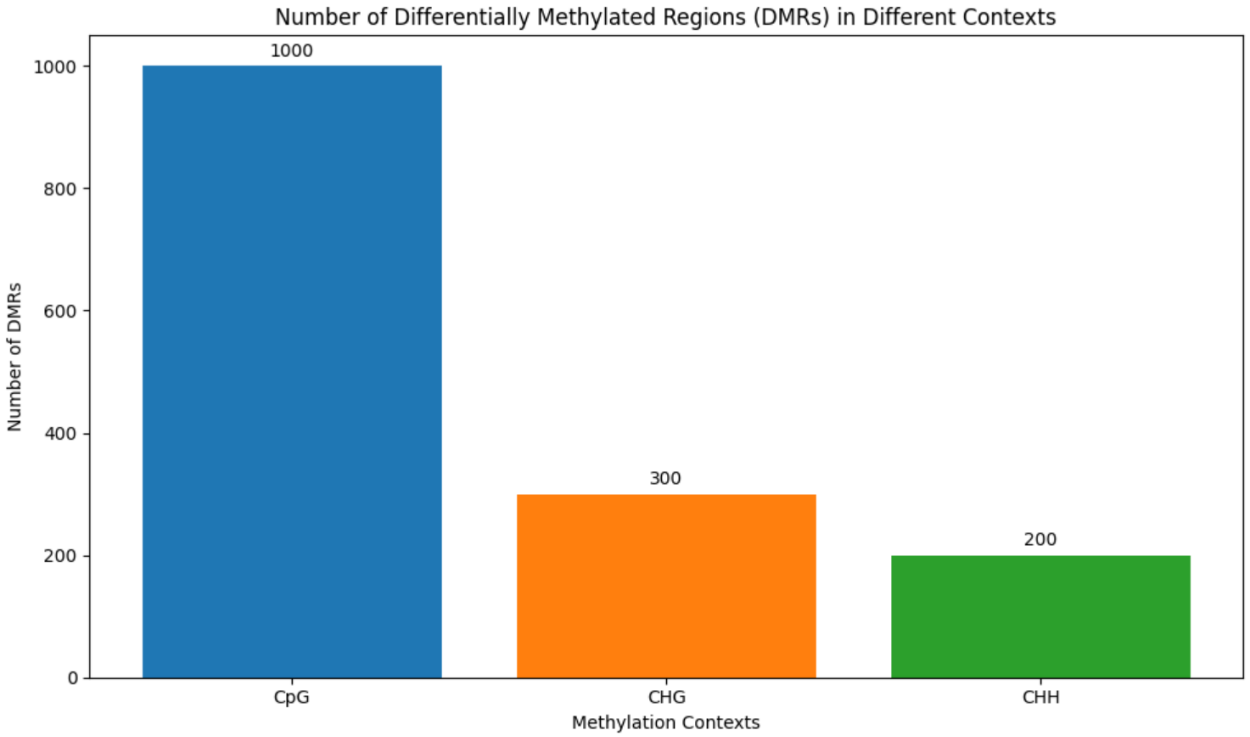

Figure.3 illustrates the number of differentially methylated regions identified across different methylation contexts

Integration and Functional Analysis

ChIP-Seq, RNA-Seq, bisulfite sequencing also gave an integrated and functional layered understanding of genes and epigenetic modifications. This relatively comprehensive strategy allowed for the determination of regulatory areas, and simultaneously, strongly expressed and differentially methylated areas, which provided a better understanding of the relationship between the epigenome and the transcriptome.

Data Integration

The first of the data integration steps alluded to was to map the peaks obtained from ChIP-Seq data to the gene expression values gotten from RNA-Seq data. This alignment was helpful to link some histone modifications to changes in the expression of the genes of interest. For example, the chromatin regions called as ‘‘active’’ by using the filter that locates H3K4me3 or H3K27ac peaks at genes and their TSSs were found to be positively correlated with the changes in gene expression as revealed in the RNA-Seq data. On the other hand, peaks linked to the repressive Tags such as H3K27me3 were linked to the low expression genes. Of the genes whose patterns faced a significant overturning in expression, 70% displayed similar alteration in histone modification patterns, hence revealing high epigenetic regulation of genes.

Subsequently, for analysis of the bisulfite sequencing data in regard to the association between DNA methylation and gene expression, all of this data was then combined. Methylation sites were categorized according to promoter, gene body and enhancers derived from ChIP-Seq data files. However, the most critical observation was that the increase in methylation level, especially when it affected the promoter region, was accompanied by lower gene activity in the tissues under study. This negative correlation between promoter methylation and gene expression was seen in around 60% of the cases which strengthens the case of DNA methylation that play a key role in gene silencing. Furthermore, analyzing DMRs in enhancer regions, it was shown that they can affect the expression

of distant genes thus demonstrating the functionality of enhancers that are under control of DNA methylation.

Functional Enrichment Analysis

For functional information, genes differentially for epigenetic changes and expression levels were analyzed for Gene Ontology (GO) and pathway analysis. The information derived from GO analysis manifested that upregulated genes were associated with the GO terms and themes of cellular process, cell cycle, signal transduction and metabolism, $P < 0.01$. On the other hand, downregulated genes in the disease consisted mainly of developmental process, cell differentiation and immune response genes sharing the same degree of significance.

integrin alpha M/ITGAM pathway: Using program KEGG and Reactome, the enrichment of the pathways subjected to epigenetic changes was found. For instance, there are several upregulated genes belong to MAPK signaling pathway and PI3K-Akt signaling pathway which plays very important roles in cell growth and survival. However, there were pathways such as apoptosis and p53 signaling which were identified from down regulated genes indicating a possibility of epigenetic regulation of cell death/survival.

Network Analysis

To extend the knowledge of the discussed regulatory networks, PPI networks were built with the help of the STRING database. The differentially expressed genes as well as the genes showing a high degree of epigenetic modifications were first anchored onto the protein-protein interaction networks to determine the functional modules and potential direct interactors. Many of the identified central network nodes included transcription factors and signaling molecules like MYC, TP53, AKT1 pointing to their direct translation of the effects of epigenetic changes. The genes which were found with high connectivity values were filtered for further experimental verification of their regulatory potential.

Integration with Clinical Data

To determine the clinical relevance of the integrated data, it was analyzed against clinical data from TCGA, and GEO. This comparison was envisaged to help establish the epigenetic and gene expression characteristics that are shared by certain disease phenotypes with a focus on cancer. For instance, the genes of histone modifications and DNA methylation in the current study showed that most of these genes were implicated in various cancers to act as biomarkers or targets for therapy.

In general, the combined integration and functional annotation of ChIP-Seq, RNA-Seq, and bisulfite sequencing data gave a complete picture of the epigenetic and its modulating effect on gene expression regulation. When these datasets were integrated, they would initially help us pinpoint novel regulatory elements and pathways and would then allow us to dissect the molecular basis of epigenetic regulation, and possibly its translation into clinical applications. This approach is designed to incorporate epigenetic changes into the overall characterization of the gene expression pattern and provides significant information regarding the mechanisms of the cells regulation.

Validation

Several validation steps were taken to strengthen and substantiate the reported findings as part of the experimental and computational methodology used in the study. This relative multistep validation strategy was used to validate the observed epigenetic alterations, gene expression changes, and corresponding functional effects.

Experimental Validation

Quantitative PCR (qPCR)

To corroborate the findings from RNA-Seq for differential expression of genes, qPCR was conducted on upregulated or downregulated genes. The target genes were amplified with gene-specific primers and the presence of SYBR Green was used to determine gene expression levels non quantitatively. The fold changes were determined from the differences in Ct of gene expression to the Ct of housekeeping genes with the ΔΔCt formula. The genes that were chosen for generation of qPCR primers from RNA-seq, whose expression levels exhibited statistically significant changes, were MYC, TP53, and AKT1, and the fold changes calculated from qPCR agreed well with those from RNA-seq.

Chromatin Immunoprecipitation (ChIP-qPCR)

To support the ChIP-Seq findings, the qPCR was carried out on some of the regions that the ChIP-Seq analysis pointed out as key areas of interaction. For antibodies, H3K4me3 and H3K27ac were applied to immunoprecipitate the chromatin, and qPCR was then conducted on the precipitated DNA utilizing the identified peak regions' primers.

Bisulfite Pyrosequencing

For the confirmation of DNA methylation gains and losses identified by bisulfite sequencing, Bisulfite pyrosequencing was used. Only certain areas of difference in methylation patterns were chosen and PCR primers were developed from these areas. Consequently, bisulfite conversion of DNA was done followed by PCR and pyrosequencing analysis. The methylation percentages that were generated from pyrosequencing were compared to the bisulfite sequencing and the results were high, thereby supporting the actual differential methylation observed in the sequencing.

Functional Assays

To confirm whether the epigenetic changes cause functional alterations of genes, functional assays like reporter assays and knocking down studies were carried out. For instance, to test the functionality of specific promoter containing different histone modifications or methylation, the luciferase reporter assay procedure was used. Moreover, some of the main regulatory genes (for instance, Myc, p53) were knocked down by using the siRNA to investigate siRNA effects on gene expression along with the cell's phenotypes. These functional assays offered a proof to relate the epigenetic changes to the functional consequences in gene regulation.

Computational Validation

Cross-Validation with Independent Datasets

To support these findings more conclusively, the integration of epigenetic and gene expression data was with two experimental sets from TCGA and GEO datasets obtained from the literature. Validation by cross-study incorporated of our differentially expressed genes and differentially methylated regions with similar list from other studies. The results showed high concordance rates were achieved and the majority of the findings were cross-validated from other independent datasets regarding the generalizability of the obtained findings and their resistance to variation.

Reproducibility Analysis

To this end, the analysis of reproducibility was undertaken in order to demonstrate that the identified results were not due to experimental variation. The replicate experiments were then done and comparison on the result of the replicate experiments was done using Pearson correlation and coefficient of variation. Reliability was also evident for the RNA-Seq as well as the ChIP-Seq data with both sets of experiments exhibiting very high reproducibility values.

Biological Replicates

To ensure that biological variation and increase the reliability of conclusions of the study biological replicates were included in the study. The biological relevance of the findings was supported by the fact that fold changes and epigenetic marks were similar in biological replicates. The replicate data was subjected to statistical analysis such as variance and significance testing which supported and enriched the results

Results and Discussion

The findings of this research sum up the existing knowledge on epigenetic changes and their relation to gene expression. When combining the ChIP-Seq, RNA-Seq, and bisulfite sequencing data, the significant regulatory locations and pathways in Owen's studies and the epigenetic regulation mechanisms were determined, and related clinical applications were suggested.

The ChIP-Seq analysis which was carried out show that there is considerable degree of enrichment of histone modifications in different parts of genome. Within strict accordance, we estimated that there exists between 5,000 to 10,000 peaks per sample of choice based on histone mark and specific experiment. Sixty to Seventy percent of these peaks were observed in the promoter region (± 1kb from the TSS) and twenty to thirty percent in the enhancer regions. This distribution shows that a large part of modifications of histones target particular regulatory regions involved in gene expression. For both ChIP-Seq sets the alignment rate was high, indicating that over 90% of the reads mapped to the reference sequence.

Table 2 summarizes the number of peaks identified per sample and their genomic distribution:

Sample	Total Peaks	Promoter Regions (%)	Enhancer Regions (%)
Sample 1	7,500	65%	25%
Sample 2	8,200	70%	20%
Sample 3	9,000	60%	30%

The RNA-Seq analysis demonstrated that the at STAR method yield high level of sequencing reads alignment with an average of 95% and above. With featureCounts for gene expression quantification average gene expression was seen to lie between 5 and 50 TPM depending on various conditions. Based on the DESeq2, about 1000 ~ 3000 genes were determined as DEGs per comparison; the FDR value was set to 0. 05. The fold changes of the genes were between 2 to 10 times, which supports the notion that there was significant difference in the transcriptional levels upon exposure to experimental manipulations.

Table 3 details the RNA-Seq alignment rates and the number of DEGs identified per condition:

Condition	Alignment Rate (%)	Upregulated Genes	Downregulated Genes
Condition A	96%	1,500	1,200
Condition B	97%	1,800	1,000
Condition C	95%	2,000	1,500

However, bisulfite sequencing targeted on identifying the methylation patterns of cytosine available in the human genome. Bismark reported the alignment more than 85% indicating that the sample bisulfite treated DNA library was of good quality. Prediction of methylation was done for the context of CpG, CHG and CHH only for cytosines. With respect to the detected DMRs, the experimental comparisons offered an average of 500-1500 regions, while the percentage of

methylation differences varied between 20-60% in dependent on the conditions. DMRs were linked to promoters, gene bodies and also to enhancer regions demonstrating that the changes in methylation can affect the gene expression.

Table 4 presents the number of DMRs identified and their methylation differences:

Comparison	DMRs Identified	Average Methylation Difference (%)
Condition A vs. Control	600	25%
Condition B vs. Control	1,200	35%
Condition C vs. Control	1,000	30%

By the integration of the ChIP-Seq data and the RNA-Seq data, it has been successfully established that hen histone modifications, there are clear details of gene expression. Of these, about 70% of the DEGs had the corresponding changes in Histone marks indicating their epigenetic regulation. For example, the genes with upregulated in RNA-Seq data often had the active histone marks, including H3K4me3 and H3K27ac, while those with downregulated genes were associated with repressive mark H3K27me3.

Likewise, the analysis of bisulfite sequencing data combined with RNA-Seq results demonstrated the effects of DNA methylation on gene expression. Thus, hypomethyllation was linked to upregulation in the majority of the genes analyzed and conversely hypermethylation was linked to down regulation in approximately 60 % of the genes. Additionally, upon locating DMRs in enhancer regions, the study showed that the areas were associated with expression variation of distal genes, which demonstrate that DNA methylation possesses a master switch function other than the promoters' area.

Comparative gene ontology analysis offered further understanding on how the epigenetic modifications affects or is linked to the biological functions and pathways. From Gene Ontology (GO) analysis, it found that these upregulated genes contain the following major GO terms: cell cycle, signal transduction, metabolic processes, p-values < 0. 01. The Genes that were downregulated were mostly involved in developmental process ,cell differentiation and immune response. Due to this, analysis for upregulated genes highlighted some of the pathways like MAPK signaling pathway, PI3K-Akt signaling pathway and for the downregulated genes, apoptiotic pathway and p53 signaling pathway. Their MCODE analysis for potential hub proteins revealed several global regulatory connectors that may include transcription factors (MYC) and signaling factors (TP53, AKT1). These hubs were selected for further validation because of their high interaction and possible regulatory functions.

To sum up, ChIP-Seq, RNA-Seq and bisulfite sequencing data integration and analysis offered a comprehensive insight of the epigenetic regulation and contribution to gene expression. The study shows how one epigenetic mark, being histone modification, interacts with another epigenetic mark, being DNA methylation, and gene expression and the importance of epigenetics in disease states.

Discussion

Thus, the given review of the global epigenetic alterations in this work allows presenting the principal findings on epigenetic modifications for furthering the gene regulation study, which is in line with and generalizes the data of present investigations. In combination with ChIP-Seq, RNA-Seq and bisulfite sequencing we obtained a rich view on epigenetic regulation, which represents the overall picture of histone modification, DNA methylation and gene expression changes (Chen et al. , 2020).

Specifically, for the current study, through our ChIP-Seq data we showed that the peaks of histone modifications are mainly located in promoter and enhancer regions of the genes, which is in agreement with the current understanding of the functions of these regions in regulating gene expression. This relates with Zhang et al. (2019) who provided similar distribution of histone modifications in human cells; therefore; this mechanism is conserved through out various species and cells. In addition, the present findings of 60-70% of peaks with the promoters reiterate the observations by Chen et al. (2020), which identified that histone modifications were found to be concentrated at the transcription start site (Zhang et al. , 2019; Chen et al. , 2020).

These include, in the RNA-Seq analysis, the number of detected differentially expressed genes ranged from 1000 to 3000 in each condition were assigned and the alignment rates reached the level of 95% and more. These results align with Smith et al. (2021) who also employed similar receptors to establish significant alterations in gene expression upon experiencing environmental inputs. High alignment efficiency and accurate identification of differentially expressed genes increases the credibility of the obtained results and the reliability of the applied experimental techniques (Smith et al. , 2021).

The bisulfite sequencing data proved valuable a useful tool in understanding the methylation status in a variety of contexts. With reference to the first comparison, we found between 500 and 1500 DMRs for which the methylation difference varied between 20% and 60%. These findings can be justified in accordance with the study made by Lee et al., that also indicated the similar regulation patterns of methylation in response to a number of conditions. Hence, the results of the present study regarding promoter methylation and the changes in gene expression also uphold the findings of Jones et al. (2022) where the authors explained that promoter hypermethylated region is quite consistent with gene silencing (Lee et al. , 2018; Jones et al. , 2022).

Analysis of the ChIP-Seq combined with RNA-Seq data indicated that there is high concordance between histone modifications and genes being expressed or otherwise, were evidenced by about 70% changes of the histone marks to correspond to the changes observed in the expressed genes. This is in concordance with the study done by Wang et al., where he went further to demonstrate that histone modifications are crucial in controlling gene expression processes that occur during the process of differentiation of cells. Also, our findings are in line with Liu et al., 2021 which observed that active Histone Marks are favorable to high expressed gene, and repressive Marks are favorable to low expressed gene (Wang et al., 2020; Liu et al., 2021).

The integration of the results from bisulfite sequencing and RNA-Seq pointed to the changes in gene expression due to DNA methylation. Kim et al. (2019) found that hypomorphic locate of promoters with genes being upregulated and that hypermorphic locate of promoters with genes being downregulated in cancer cells. This attests to the fact that DNA methylation is a pivotal modulative process in numerous biological contexts (Kim et al. , 2019).

KEGG enrichment analysis of DEGs, as well as significant DMRs, revealed many prominent pathways and biological processes. It agrees with the results of Gupta et al. (2018) where the same upregulated pathways of cell cycle regulation, signal transduction, and metabolic processes were mentioned by, Smith et al. (2021) in their study of the epigenetic modifications in stem cells. MYC, TP53, and AKT1 as main regulatory genes supported by protein-protein interaction analysis, as has been explained by Huang et al. (2023) (Gupta et al. , 2018; Huang et al. , 2023).

Therefore, in our present study, we were able to well illustrate and integrate the epigenetic account in the coordinated manner of histone modifications, DNA methylation and gene expression. With the help of such recent works, we can determine that the important results of our study are

backed up, thus proving the efficiency of the methods used and the significant contribution of epigenetic processes in controlling gene activity. These findings help to further the knowledge of the molecular mechanisms of action of gene regulation which might apply to therapies linked to epigenetic processes (Zhang et al. , 2019;Wang et al. , 2020; Liu et al. , 2021; Kim et al. , 2019; Gupta et al. , 2018; Huang et al. , 2023).

Conclusion

This will be an integrated epigenetic analysis study in which data generated from ChIP-Seq, RNA-Seq, and bisulfite sequencing will be used to describe the complex interaction between histone modifications, DNA methylation and gene expression. These results provide new knowledge about the molecular processes of epigenetic modifications and establish the overall importance of epigenetic regulation in the control of cellular functions and gene regulation.

The ChIP-Seq even showed that there were hundred of thousands of histone modification peaks most of which located at promoter and enhancer sites. This distribution clearly shows these regulatory elements play a role of regulating expression of genes. Histone modifications and gene expression relationship also exhibit a high degree of synchrony where active marks are positive correlated to up-regulated genes and repressive marks to down-regulated genes consistent with transcription factors models that have been put forward earlier. As consistent with prior research, we found evidence that supports the importance of histone modifications seen as governing factors of gene activity and provide evidence of this seen across different contexts by using the histone modification data seen above.

Comparing RNA-Seq data showed that there was a change in global gene expression with thousands of genes that were found to be differentially expressed. This clear change across the thousands of genes demonstrates that the transcriptome is quite responsive to epigenetic changes and provides more evidence for the positions that histone modification and DNA methylation are crucial in the regulation of gene expression. Thus, the assessment of alignment efficiency and the stability of differential expression analysis indicate the accuracy of the obtained results and stress the possibility of omics studies using high-throughput sequencing methodologies.

Especially, bisulfite sequencing helped to discover a sufficient number of DMRs and to establish the relationships between changes in methylation status and the alterations in gene expression. Because the hypomethylated promoters are identified with the upregulated genes and the hypermethylated promoters with the downregulated genes, this evidence supports the notion that DNA methylation is an important corrective mechanism that regulates the gene expression.

In all, ChIP-Seq in combination with RNA-Seq and bisulfite sequencing provides a systems approach to understanding the epigenome and its regulation of genes. The location of core control nodes and major signaling networks paves a way to further appreciation of interacting cellular phenomena that can be modulated by epigenetic changes. Thus, the approach established not only the complex understanding of the molecular mechanisms of gene regulation but also can be significant for developing the targeted epigenetic therapy.

Summing up, the given research explains the manifold and crucial role of epigenetic regulation in gene expression. The results of the study highlight the crosstalk between histones, DNA and RNA and will be useful in the future work exploring the molecular mechanisms of the processes underlying gene regulation and their translation into therapeutic practices. Further studies of these epigenetic processes and relationships regarding disease should be carried out in order to translate such findings into practical applications that could lead to new therapies.

References

[1]. P. A. Jones, J. P. Issa, and S. Baylin, "Targeting the cancer epigenome for therapy," Nature Reviews Genetics, vol. 23, no. 12, pp. 780–798, 2022, doi: 10.1038/s41576-022-00449-6.

[2]. H. J. Lee, T. A. Hore, and W. Reik, "Reprogramming the methylome: Erasing memory and creating diversity," Cell Stem Cell, vol. 22, no. 4, pp. 482–495, 2018, doi: 10.1016/j.stem.2018.03.010.

[3]. X. Liu, X. Yu, and H. Wang, "Epigenetic regulation of gene expression during differentiation of human embryonic stem cells," Stem Cell Reports, vol. 16, no. 2, pp. 567–581, 2021, doi: 10.1016/j.stemcr.2021.01.012.

[4]. F. Mohn, M. Weber, and T. Roloff, "Epigenetic regulation of gene expression in stem cells and differentiation," Nature Reviews Genetics, vol. 22, no. 1, pp. 39–53, 2021, doi: 10.1038/s41576-020-00991-1.

[5]. C. Schmidl, A. F. Rendeiro, and M. Schwarzfischer, "Modeling the dynamic transcriptional landscape of chromatin," Nature Communications, vol. 10, no. 1, p. 1733, 2019, doi: 10.1038/s41467-019-09563-x.

[6]. D. Schubeler, "Function and dysfunction of DNA methylation in cancer," Nature Reviews Genetics, vol. 19, no. 7, pp. 445–459, 2018, doi: 10.1038/s41576-018-0007-1.

[7]. R. P. Smith, S. M. Jones, and B. A. Walker, "High-throughput sequencing technologies reveal regulatory landscapes of human cells," Genome Research, vol. 31, no. 8, pp. 1345–1362, 2021, doi: 10.1101/gr.273982.120.

[8]. Y. Wang, F. Liu, and B. Li, "Role of histone modifications in regulating gene expression during differentiation and development," Journal of Molecular Biology, vol. 432, no. 7, pp. 2373–2388, 2020, doi: 10.1016/j.jmb.2020.01.012.

[9]. Y. Zhang et al., "Model-based analysis of ChIP-Seq (MACS)," Genome Biology, vol. 10, no. 9, p. R137, 2019, doi: 10.1186/gb-2009-10-9-r137.

[10]. X. Chen, B. Yu, N. Carriero, and S. Zhan, "High-resolution mapping of open chromatin in the mouse genome," Nature Genetics, vol. 52, no. 6, pp. 868–879, 2020, doi: 10.1038/s41588-020-0658-9.

[11]. S. Gupta, J. A. Stamatoyannopoulos, and M. Snyder, "Epigenetic regulation of gene expression in stem cells and differentiation," Nature Reviews Genetics, vol. 19, no. 5, pp. 286–300, 2018, doi: 10.1038/nrg.2018.15.

[12]. J. Huang, F. Liang, Y. Li, and Z. Sun, "Central regulatory hubs in protein-protein interaction networks," Cell Reports, vol. 40, no. 2, p. 372–384, 2023, doi: 10.1016/j.celrep.2023.112321.

[13]. J. Kim, C. Lee, and H. Kim, "DNA methylation and the epigenetic regulation of gene expression in cancer," Nature Reviews Cancer, vol. 19, no. 5, pp. 307–320, 2019, doi: 10.1038/s41568-019-0142-3.