



Article Multiplexed Methylated DNA Immunoprecipitation Sequencing (Mx-MeDIP-Seq) to Study DNA Methylation Using Low Amounts of DNA

Inam Ridha ^{1,2}, Chenxi Xu², Yining Zhang², Yunro Chung^{2,3}, Jin G Park², Joshua LaBaer² and Vel Murugan^{2,*}

- School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ 85281, USA; iridha@asu.edu
- ² Center for Personalized Diagnostics, Biodesign Institute, Arizona State University, Tempe, AZ 85281, USA; chenxix1@asu.edu (C.X.); yzhan937@asu.edu (Y.Z.); yunro.chung@asu.edu (Y.C.); jin.park.1@asu.edu (J.G.P.); joshua.labaer@asu.edu (J.L.)
- ³ College of Health Solutions, Arizona State University, Phoenix, AZ 85004, USA
- * Correspondence: vel.murugan@asu.edu

Abstract: Background/Objectives: DNA methylation is a key epigenetic mark involved in regulating gene expression. Aberrant DNA methylation contributes to various human diseases, including cancer, autoimmune disorders, atherosclerosis, and cardiovascular diseases. While whole-genome bisulfite sequencing and methylated DNA immunoprecipitation (MeDIP) are standard techniques for studying DNA methylation, they are typically limited to a few samples per run, making them expensive and low-throughput. Therefore, an automation-friendly method is needed to increase throughput and reduce costs without compromising data quality. Methods and Results: We developed a novel method called Multiplexed Methylated DNA Immunoprecipitation Sequencing (Mx-MeDIP-Seq), which can be used to analyze many DNA samples in parallel, requiring only small amounts of input DNA. In this method, 10 different DNA samples were fragmented, purified, barcoded, and pooled prior to immunoprecipitation. In a head-to-head comparison, we observed a 99% correlation between MeDIP-Seq performed individually or combined as Mx-MeDIP-Seq. Moreover, multiplexed MeDIP led to more than 95% normalized percent recovery and a 25-fold enrichment ratio by qPCR, like the enrichment of the conventional method. This technique was successfully performed with as little as 25 ng of DNA, equivalent to 3400 to 6200 cells. Up to 10 different samples were processed simultaneously in a single run. Overall, the Mx-MeDIP-Seq method is cost-effective with faster processing to analyze DNA methylome, making this technique more suitable for high-throughput DNA methylome analysis. Conclusions: Mx-MeDIP-Seq is a cost-effective and efficient method for high-throughput DNA methylation analysis, offering faster processing and reduced sample requirements. This technique makes DNA methylome analysis more accessible for large-scale studies.

Keywords: methylated DNA immunoprecipitation; epigenetics; DNA methylation

1. Introduction

Epigenetics studies heritable phenotypic changes, including gene expression, that do not involve changes to DNA sequences [1,2]. DNA methylation (DNAm) is one of the most widely studied epigenetic marks, as it plays a crucial role in gene silencing, cell fate decisions, and disease development [2,3]. It involves a covalent modification of the pyrimidine ring of cytosine nucleotide at the C-5 position through the addition of a methyl group. DNAm is mainly found as symmetrical CpG dinucleotides in mammalian cells [4,5]. It is established during gametogenesis and embryo development by recruiting de novo methyltransferase enzymes (DNMT1) and preserved by maintenance methyltransferases like DNMT3 [6,7]. Changes in DNA methylation alters gene expression during adulthood by the recruitment of the methyl-CpG-binding domain family of proteins, like methyl-CpG-binding domain proteins such as MeCP2 in concert with histone deacetylases [8–10].



Citation: Ridha, I.; Xu, C.; Zhang, Y.; Chung, Y.; Park, J.G.; LaBaer, J.; Murugan, V. Multiplexed Methylated DNA Immunoprecipitation Sequencing (Mx-MeDIP-Seq) to Study DNA Methylation Using Low Amounts of DNA. *DNA* **2024**, *4*, 397–416. https://doi.org/10.3390/ dna4040028

Academic Editor: Darren Griffin

Received: 6 September 2024 Revised: 22 October 2024 Accepted: 24 October 2024 Published: 29 October 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). This epigenetic modification is essential for the normal function of cellular processes such as the differentiation and preservation of genomic stability [11,12]. Aberrant DNA methylation may cause cancer and other diseases such as autoimmune disorders [13,14] or Alzheimer's disease [15,16]. Up-regulation or down-regulation of the methyltransferase enzymes [17], nutrient intake [18,19], environmental stress such as air pollution, infectious agents, and toxic elements [20–22] are considered among the factors that can cause DNA methylation alteration.

Several techniques study changes in methylation patterns, [23] transforming that information into quantitative and measurable signals [24,25]. This includes methods to investigate methylation in detail at the level of a single base, such as whole-genome bisulfite sequencing (WGBS) or enrichment of methylated DNA fragments through immunoprecipitations using antibodies or methyl binding domain proteins (MBD) [26]. WGBS combines sodium bisulfite conversion of input DNA and high-throughput DNA sequencing [27] to create a comprehensive measurement of all methyl additions to the DNA. Although bisulfite sequencing leads to DNA methylation analysis at a single-base resolution [28], the treatment causes substantial DNA degradation and needs excessive purification to remove the sodium bisulfite, thus requiring at least 50 to 1000 ng of input DNA to start the process [29]. Other limitations of WGBS are the high cost of the technology, especially on large sample sizes [30–32], and the inability of this method to distinguish between 5-methylcytosine (5mC) and hydroxy-methylcytosine (5hmC) [33]. Recent advances allow single cell bisulfite sequencing with some limitations and even in parallel with scRNA-Seq [34–38].

Methylated DNA immunoprecipitation followed by sequencing (MeDIP-Seq) is another widely used method to study DNA methylation profiles. MeDIP enriches methylated DNA fragments by using monoclonal antibodies against 5-methyl cytosine. The immunoprecipitated sample can then be analyzed by high-throughput sequencing to identify the methylated regions and compared to the total input DNA that did not undergo immunoprecipitation, thus assessing the enriched fraction representing the methylated regions. This technique was also adapted to detect 5-hydroxymethylcytosine (5hmC) DNA [39]. 5hmC, as a different epigenetic modification, is an oxidized form of 5mC [40]. Unlike 5mC, which is associated with transcriptional repression, 5hmC is associated with gene transcription [40]. It is reported that 5hmC may increase gene expression through increasing chromatin accessibility or inhibiting repressor binding [40,41].

Although MeDIP does not provide information at single-base resolution, it has a sufficient resolution to study differentially methylated regions (DMRs), which are functionally more important than single methylation polymorphisms (SMPs) [42–44]. Most CpG dinucleotides are methylated, and there are about 30 million CpGs in the human genome, where 7% of CpGs are located within CpG islands (CGIs). About 70% of annotated genes' promoters are associated with CGIs, mostly with housekeeping functions [45]. The MeDIP method can target CpG islands and show enrichment for the entire length of the genes [46].

Early MeDIP-Seq methods also required one microgram of input DNA [47]. Several recent protocols use less fragmented DNA per immunoprecipitation (IP) reaction (100 ng to 1 μ g) [48–51] but typically require 3–5 days to process [49] each sample. In those protocols, DNA isolation is performed on individual samples, and DNA is then sheared by sonicators. This process can take ~1–2 days (depending on the sample size and desired fragment distributions) [51]. Sonication is followed by immunoprecipitation and library preparation for sequencing, which approximately takes an additional ~2–3 days [51]. The ability to process samples in high throughput is limited currently due to the single-sample procedure of the MeDIP step, an issue we address here. We have developed a MeDIP-Seq multiplexing strategy that circumvents these challenges and allows the analysis of many samples in parallel, which reduces processing time, cost, and the amount of DNA needed. The new platform generated consistent results with as little as 25 ng of DNA, corresponding to the DNA content of 3400 to 6200 cells [52].

In standard MeDIP-Seq, the DNA library construction occurs after immunoprecipitation (IP). In the Mx-MeDIP-Seq approach, the library was prepared before IP, i.e., the IP reaction was sandwiched between indexed adapter ligation and PCR amplification followed by sequencing. In Mx-MeDIP-Seq, unique barcodes were ligated to DNA from different samples. The prepared DNA libraries were pooled, immunoprecipitated, sequenced, and then demultiplexed. We confirmed the reliability of this Multiplexed MeDIP-Seq (Mx-MeDIP-Seq) protocol by comparing it to standard MeDIP for the same samples individually, including the DNA fold enrichment and the correlation between the two groups after immunoprecipitation. Performing IP reactions with barcode-ligated libraries and the ability to combine as many as 10 samples in a single run and optimizing the amount of 5mC antibody could reduce the amount of required starting DNA per sample and reduce time and costs, all prerequisites for developing an automated platform to perform high-throughput MeDIP-Seq.

2. Materials and Methods

2.1. Peripheral Blood Mononuclear Cells (PBMCs) Isolation

To isolate the peripheral blood mononuclear cells (PBMCs), peripheral whole blood was drawn from healthy patients, with local IRB approval (STUDY00009580), into CPTTM tubes (BD Vacutainer, cat no 362761, Becton Dickenson, Franklin Lakes, NJ, USA). The tubes were centrifuged according to the manufacturer's directions. The plasma upper layer was removed, and the remaining buffy coat diluted into an equivalent volume of PBS + 2% FBS in 15 mL conical centrifuge tubes and respun. After each centrifugation, the pellet was resuspended into PBS + 2% FBS. Following two washes and centrifugations, the PBMCs were resuspended in BambankerTM freezing media and stored in liquid nitrogen for future use.

2.2. Micrococcal Nuclease Digestion

The Micrococcal Nuclease Digestion method was based on published methods [53] with some modifications. Frozen PBMCs were centrifuged twice at $600 \times g$ for 10 min at 20 °C and washed with RPMI + 20% FBS. The cell pellet was resuspended in PBS to achieve a final volume of 7500 cells per 1 µL of suspension. An equal volume of the MNase digestion buffer (0.02 U/µL MNase (Thermo Scientific, Cat. No: 88216, Waltham, MA, USA) in 2× lysis buffer (Table 1; from stock reagents as per Table S1 in Supplementary Materials) was added to the cell suspension and gently vortexed. The cell suspension was then incubated on ice for 20 min, followed by incubation at 55 °C water bath for 10 min. The digestion reaction was terminated using 30 mM EGTA. Then, the digested chromatin was protease k (*Ambion*TM) treated at 55 °C and then pre-cleared with 1:2 KAPA Pure beads (Roche Cat. No: 07983298001, Basel, Switzerland). The quality and quantity of fragmented DNA were verified using agarose gel (1.5%) electrophoresis and Qubit[®] 4.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA), respectively. The resulting purified DNA was utilized for library construction.

Table 1. $2 \times$ Lysis buffer stock preparation contacting MNase. This mixture is used to lyse the cells and isolate and fragment the DNA strand in one step.

Reagents	Volume (mL)
100 mM Tris-HCl (1 M, pH 8.0)	20
300 mM NaCl (5 M)	12
2% Triton X-100 (25%)	16
0.2% sodium deoxycholate (DOC, 12.5%)	3.2
$10 \text{ mM CaCl}_2 (1 \text{ M})$	2
10 mM sodium butyrate (800 mM)	2.5
H ₂ O	144.3
$1 \text{ U}/\mu \text{L}$ MNase stock solution	4
Total volume	200

2.3. End-Repair, A-Tailing, and Adapter Ligation

Illumina KAPA LTP Library Preparation Kit was employed for library construction (Cat. No: KK8233, San Diego, CA, USA). Purified, fragmented DNA was end-repaired to produce 5'-phosphorylated, blunt-ended dsDNA fragments. Unique Dual Indexes (UDIs) were ligated to 3'-A-tailed library fragments holding appropriate adapter sequences on both ends. Indexed DNA fragments were subjected to methylated DNA immunoprecipitation, followed by amplification.

2.4. Multiplex Methylated DNA Immunoprecipitation Sequencing (Mx-MeDIP-Seq)

Adapter ligated DNA fractions were pooled together at indicated amounts (ng), and methylated DNA immunoprecipitation was performed on the pooled, denatured DNA samples. This method comprises immunocapturing methylated DNA fragments using methylated cytosine-specific antibodies. The affinity of the antibody used in MeDIP enables the detection of any methylated cytosine and is not restricted to the analysis of CpG island. MeDIP was performed according to the manufacturer's instructions (Active Motif Cat. No: 55009). The procedure was optimized based on the input DNA and the quantity of required antibodies. Fragmented DNA was used (50 ng, 100 ng, 200 ng, 500 ng, 1 µg) with different amounts of antibody (0.5 to 8 μ g). The reaction was carried out in special low DNA LoBind tubes (Eppendorf; Catalog No. 022431048, Hamburg, Germany) to minimize DNA loss during the MeDIP procedure. The enriched fraction of the immunoprecipitated (IP) sample DNA was amplified, followed by a real-time quantitative polymerase chain reaction (qPCR) technique to analyze the final enrichment for methylated DNA. To confirm the enrichment of methylated DNA in the immunoprecipitated samples, four primer sets specific for the highly methylated region were used as positive controls (TSH2B, BRDT, ZC3H13, NBR2 genes). TSH2B and BRDT genes are transcribed exclusively in testis, and CpG sites of this gene are methylated in all somatic tissue. ZC3H13 is a known positive methylated locus in human cells with no expression in PBMCs [54]. NBR2 is located within a large CpG island and is found to be methylated in most somatic cells [55]. Similarly, one primer for the unmethylated region (GAPDH gene) was selected as the negative control. Normalize Percent Recovery (NPR) was calculated as per Equations (1) and (2), adapted and modified according to this study [49]. Moreover, the fold change in gene expression and fold change ratio were calculated as per Equations (3) and (4).

$$Percent Recovery = \frac{Amount of Enriched sample DNA (ng) from IP}{Amount of Enriched Input DNA (ng) from input \times \frac{sample}{input} ratio} \times 100$$
(1)

Normalized Percent Recovery
$$= 1 - \frac{Percent Recovery (Unmethylated)}{Percent Recovery (Methylated)}$$
 (2)

 $Fold Change = 2^{Input Ct - MeDIP Ct}$ (3)

Fold Change Ratio =
$$\frac{Fold \ Change \ (Methyated)}{Fold \ Change \ (Unmethyated)}$$
(4)

2.5. Size Selection

Using KAPA pure beads, library size selection was performed in a two-step process based on manufacturer recommendations. The library fragments ranging between 200 bp and 1000 bp were selected, and the rest were eliminated. First, DNA fragments larger than 1000 bp were removed by incubating the library with $0.5 \times$ volume of KAPA pure beads for 5 min at room temperature. Second, to capture the fragments above 200 bp, the beads were magnetized, and the supernatant was incubated with $0.5 \times$ volume of KAPA beads. After 5 min of binding at room temperature, the beads were washed $2 \times$ with 80% ethanol and resuspended in 10 mM Tris for DNA elution. The supernatant contains the size selected library DNA between 200 and 1000 bp.

2.6. Sequencing

Library quality was analyzed on an Agilent Tapestation and quantified by qPCR (KAPA KK4835- Thermo Fisher Sci Quantstudio 5, Waltham, MA, USA) at The Arizona State University (ASU) genomic core facility, before pooling and sequencing (2×150 flow cell—Illumina NovaSeq 6000) at the University of Colorado Anschutz Medical Campus Genomics Core Facility. Read quality and other quality metrics are included in Supplementary Materials (Table S2).

2.7. Bioinformatics Analysis

Cutadapt (v 1.18) [56] was used to remove the adapter sequence and low-quality ends from raw sequencing reads.

cutadapt --cores=40 --no-trim --quality-cutoff=15,10 -o .fastq.gz fastq.gz > .fastq.gz.txt

Cleaned reads were aligned to the human reference genome (GRCh38.p12, hg38) using Burrows-Wheeler Aligner (BWA v 0.7.17 [57]) to generate BAM files.

bwa mem -t 45 reference_genome input.fastq.gz input.fastq.gz

samtools sort -O bam -o fastq.gz.bam -T .fastq.gz_aln

MethylQA (v 0.2.1) [58] was used to generate genome density for MeDIP-Seq data and reads mapping and CpG coverage statistics.

methylQA medip -o file -m file.bed hg38_lite.size file

Aligned reads were used to call peaks with MACS2 (v 2.2.6) [59] using input as a control to identify enriched areas in the genome based on performed methylated immuno-precipitation.

```
macs2 callpeak \setminus
```

-B \

-t trimmed_M192-A18_S11_ME_L001.extended.bed \

-c trimmed_M206-A18_S11_ME_L001.extended.bed \

-f BEDPE \setminus

-g hs \

```
-n M192-A18_vs_M206-A18_q01_macs2 \
-q 0.01 > M192-A18_q01_macs2.log
```

Next, the bdgcmp utility from the MACS2 package was used to deduct noise by comparing two signal tracks (IP and input) in bedGraph containing the log10 value of fold enrichment. The percentage of CG coverage of cleaned reads was compared to assess the similarity between individual and pooled samples. Next, the correlation test of read counts on common peaks regions for all different samples was performed. Before computing the correlation, all overlapping peaks regions between individual and pooled samples were first merged into one bed file. The bed file contains all regions that should be considered for the coverage analysis. Then, multiBamSummary from deeptools 2.0 computed the read coverages (the number of unique reads mapped at a given nucleotide) over the merged peaks by using BAM files (individual and pooled samples). Finally, deepTools were used to plot correlation (Pearson) to visualize the multiBamSummary file.

To calculate the enrichment correlation between individual and pooled processed samples on CpG island, we implied ComputeMatrix from deepTools to calculate the log10 value of fold enrichment scores of each base in the CpG islands. In this method, gene body length was divided by the bin size, which was considered 50 bp in this study. The fold enrichment average in each bin across the CpG islands was then calculated. The common regions overlapping two files were selected per genome region (CpG island here). Pearson correlation was used to calculate the correlation score between the average enrichment scores from the common region in two files.

2.8. Statistical Analyses

Statistical analysis for qPCR data was carried out using one-way ANOVA to compare pooled processed samples and individually processed with different amounts of starting DNA library with respect to Normalize Percent Recovery and fold change ratio, respectively. Tukey's method was used for multiple comparisons to determine whether there is a difference between the mean of all possible pairs. Statistical significance is indicated by * for *p*-values < 0.05, ** for *p*-values < 0.01, and *** for *p*-values < 0.001.

3. Results

3.1. Development of Mx-MeDIP-Seq

The designed workflow to expedite the MeDIP process and minimize the required input DNA is illustrated in Figure 1. We elected to fragment the chromatin using micrococcal nuclease to avoid both the time and sample requirements inherent to mechanical fragmentation. The conditions for enzymatic digestion were chosen to produce about 30% mono-nucleosomal DNA and an average of 25% di-nucleosomal DNA and minimize DNA loss during library construction. The band percentage was calculated by ImageJ. Bead-based size selection was used to exclude DNA over 1000 bp and under 200 bp; the remaining digested products were included in the library construction. Removing large fragment sites reduced the noise and increased resolution after sequencing.

Before pooling different samples, unique patient-specific barcodes were included in the adaptors ligated to the DNA fragments. Unique dual (UD) index adapters from Illumina (Cat# 07141530001, San Diego, CA, USA) were used, containing 96 different pairs. Adapter ligation was confirmed by the size shift in the mono-nucleosomal and di-nucleosomal bands in the e-gel image and electrogram, as shown in Figure 1C. A key aspect of this design is that all DNA manipulations steps like end-repair, A-tailing, and index ligation are amenable to automation using a standard liquid handling robot.

Multiplexing was achieved by combining several purified DNA libraries after appending unique barcode adapters (up to 10 in this study). In one tube, the mixed library was subjected to methylated DNA immunoprecipitation. Performing this process on all samples simultaneously simplified the workflow by reducing hands-on time, reagents, error, and complexity. It shortened our processing time for 100 samples from one week to one day (Figure 2). The library DNA, magnetic beads, and 5mc-antibody were incubated overnight. The amount of antibody used was based on the total amount of library DNA (Figure S2). The enriched DNA was then amplified and purified by PCR, removing DNA over 1000 bp by size selection with Ampure XP beads. Larger fragments (>1000 bp) were shown to have lower efficiency and yield during amplification on the flow cell surface while sequencing [60,61]. After IP, we confirmed that the resulting DNA contained only fragments between 300 bp and 1000 bp on an Agilent 1200 bioanalyzer (Figure 1B).

Several primer pairs were selected for enrichment estimation based on qPCR, which amplifies either methylated or unmethylated DNA fragments (see Table 2, primer pairs 2 for sequences). *TSH2B, BRDT, ZC3H13,* and *NBR2* genes were selected as the positive controls since they are highly methylated in PBMCs [54,55,62–64]. The human *GAPDH* primer was considered a negative control that recognizes the human *GAPDH* promoter as an unmethylated region in all somatic cells [62,63,65]. The efficiency of the immunoprecipitation in the MeDIP was assessed based on Ct qPCR data using normalized percent recovery and the fold change ratio formula adapted from the manufacturer's protocol and published studies [49,62,63] (Equations (1)–(4)). A MeDIP reaction was passed on for DNA sequencing when there was a high recovery (normalized percent recovery \geq 95%) of known methylated over unmethylated fragments and a fold change ratio >5 [66,67].

Sample comparisons were made by first establishing sample quality by examining %CG coverage after demultiplexing and adaptor sequence removal (Figure 3). Then, reads on common peaks and enrichment of CpG islands were correlated and compared.



Figure 1. Mx-MeDIP workflow and DNA fragment analysis throughout the experiment. (**A**) Automation friendly high-throughput DNA fragmentation, patient specific adapter ligation DNA, immunoprecipitation, and sequencing. (**B**) Representative image of agarose gel (1.5%) electrophoresis for DNA fragmented by micrococcal nuclease (MNase) digestion. The chromatin has been isolated from peripheral blood mononuclear cells (PBMCs). The left lane is a 100 bp marker. MNase digestion conditions yielded about 30% mono-nucleosomes and 25% di-nucleosomes (analyzed by ImageJ). (**C**) Left: Pattern of fragments after library prep and adapter ligation. The fragments shifted by 150 bp due to adding the adapter indexes during library prep. (**C**) Right: Electropherograms of High-Sensitivity chips on an Agilent Bioanalyzer 2100 and pattern of fragments after purification and before library prep (top). The bioanalyzer gel image illustrates the electropherograms result for DNA fragments before (top) and after library prep (bottom). DNA library has proceeded for multiplexing and MeDIP. (**D**) Left: The fragment pattern of DNA after methylated DNA immunoprecipitation and input samples obtained from High-Sensitivity chips on an Agilent Bioanalyzer 2100, the fragment pattern of DNA after methylated DNA immunoprecipitation and input samples obtained from High-Sensitivity chips on an Agilent Bioanalyzer 2100, the fragment pattern of DNA after methylated DNA immunoprecipitation and input samples obtained from High-Sensitivity chips on an Agilent Bioanalyzer 2100, the fragment pattern of DNA after methylated DNA immunoprecipitation for DNA immunoprecipitation and input samples obtained from High-Sensitivity chips on an Agilent Bioanalyzer 2100, the fragment pattern of DNA after methylated DNA immunoprecipitation and input samples.

Table 2. Primers used in qPCR to assess the enrichment of pooled down DNA. The first four primers were considered positive controls, and *GAPDH* was considered a negative control.

Target Name	Oligonucleotide Sequence	
Human Testis/Sperm-Specific Histone H2B (TSH2B)	Forward: CAGACATCTCCTCGCATCAA Reverse: GGAGGATGAAAGATGCGGTA	
Bromodomain Testis Associated (BRDT)	Г) Forward: CCCTTTGGCCTTACCAACTT Reverse: GCCCTCCCTTGAAGAAAAAC	
Zinc Finger CCCH-Type Containing 13 (ZC3H13) Forward: TCTCGGTCCACTCGT Reverse: CCGGGATTCTTCTGGA		
Neighbor of BRCA1 Gene 2 (NBR2)	Forward: TGTTATTTTTCGGGGTTCAGCTT Reverse: GATTGGCTCTTACCACTTGTCC	
Glyceraldehyde-3-Phosphate Dehydrogenase Peptidyl-Cysteine S-Nitrosylase (GAPDH)	Forward: - TCGACAGTCAGCCGCATCT Reverse: CTAGCCTCCCGGGTTTCTCT	



Figure 2. Mx-MeDIP-Seq workflow in comparison to standard workflow. Mx-MeDIP-Seq combines two steps for DNA isolation and fragmentation into one step. Immunoprecipitation is sandwiched between two steps of library preparation. In the first step, each sample receives a unique adapter. Several samples can be put together, followed by immunoprecipitation. The isolated methylated DNA is amplified and sent to sequencing, followed by demultiplexing and analysis.



Figure 3. Stages and component steps of bioinformatics data analysis. Sequence data from MX-MeDIP-Seq and MeDIP-Seq were compared in three steps by comparing %CG, genome-wide read coverage correlation, and fold enrichment correlation after peak calling.

3.2. Sonication vs. Enzyme Fragmentation on Methylation Data

In order to compare fragmentation by MNase and sonication on the methylation result, three MeDIP-seq data were downloaded from Gene Expression Omnibus under accession number GSE39604 [68]. In this study, PBMCs were isolated from freshly collected blood and sonicated to approximately 100–500 bp. Methylated DNA was enriched by immunoprecipitation with an anti-5-methyl-cytosine monoclonal antibody like our study. The intersect of each MedIP-seq data set on CpGIsland was calculated using bedtools/2.30.0 intersect command. Out of 30,344 CpGIsland listed in the human genome, the enzymatically fragmented MeDIP-seq matched 26,234 \pm 3765 sites, while groups with sonicated fragments matched 24,941 \pm 4591 sites. These two samples matched on 89.8% \pm 7.6% of sites on CpG islands.

3.3. Individual MeDIP and Mx-MeDIP Resulted in Similar Enrichment and Correlation

To investigate whether multiplexing affected the enrichment and sequencing results, a set of ten different library DNA with unique adapters were tested individually or as a single pool using standard MeDIP-Seq [48,51,69] (Figure 4A). The enrichment for every sample was assessed with qPCR vs. input samples (Figure 4B,C).

Table 3. Summary of GC content and correlation coefficients between ten samples that were processed individually (conventional MeDIP-Seq) and the same sample processed in the pool (Mx-MeDIP-Seq).

Samples	Shifted Peak in %CG (Processed Individually)	Shifted Peak in %CG (Processed in Pool)	Pearson Correlation Coefficient (of BAM Files Between Two Groups)	Pearson Correlation Coefficient (of Peaks Called on CpG Island)
S1	59	59	0.99	0.998
S2	60	59	0.99	0.996
S3	59	59	0.90	0.996
S4	59	60	0.89	0.994
S5	59	60	0.89	0.998
S6	58	59	0.88	0.996
S7	60	59	0.87	0.999
S8	59	58	0.87	0.995
S9	58	59	0.89	0.993
S10	60	59	0.99	0.993
Mean	59.10	59.10	0.91	0.995
standard deviation (SD)	0.70	0.53	0.05	0.002

As shown in Figure 4B, the average normalized percent recovery was over 95% for all the assessed targets. Samples that were processed individually showed similar normalized percent recovery compared to samples in the pool. There was no significant difference between individual and pool processing, indicating that multiplexing did not affect the measured normalized percent recovery. The fold change ratio vs. input was also similar for each target (*NBR2*, *ZC3H13*, *TSH2B*, *BRDT*) for samples in both groups. The plot shows that the fold change ratio was between 15 ± 1.6 for *BRDT* and 24 ± 1.8 for *ZC3H13* (mean \pm SD; Figure 4C).

The sequencing of the methylated enriched DNA yielded approximately 43 ± 9.9 (mean \pm SD) million reads per sample with the mean Phred scores above 35 in both pooled and individual assays. The percent of GC content on obtained sequencing reads was measured using FastQC. Since the samples were immunoprecipitated based on GC-rich DNA (i.e., CpG islands), we expected to see a shifted distribution to the right side of the theoretical genomic GC distribution. Figure 4D shows that there was a clear enrichment of GC-rich reads in the MeDIP samples, and the shifted distribution was similar for both individual and pooled groups. The quality control analysis also revealed that both groups (individual and pooled) had a similar % GC content (40.3% \pm 0.24% vs. 40.2% \pm 0.8%



(mean \pm SD)) and shifted the distribution of %GC (59.1% \pm 0.70% vs. 59.1% \pm 0.53% (mean \pm SD); Table 3 and Figure S1, in Supplementary Materials).

Figure 4. Comparison of samples processed in a pool and same samples processed individually before and after sequencing. (**A**) Illustration of the difference between processed individually group (MeDIP) and processed in pool group (Mx-MeDIP). B and C, qPCR result: normalized percent recovery (**B**) and fold change ratio (**C**) for samples that were processed in the pool (Mx-MeDIP-Seq) and the same samples processed individually (standard MeDIP-Seq). Ten different library DNA went through MeDIP individually. Also, a pool containing all those ten samples went through IP simultaneously. The plots show a qPCR result analysis for individual (n = 10) and pooled samples. (**D**) Empirical GC content in the MeDIP enriched samples (orange line) compared to the theoretical distribution (blue and green line). (**E**) Representative images of correlation of read coverage between individual and pooled samples (rest of correlation plots illustrated in Figure S1 (Supporting Information). (**F**) Representative heatmap plot of peaks called on CpG islands for a sample processed individually and samples processed in the pool (detailed information in Table 3).

The overall similarity of read coverages between the pooled and the individually processed samples was then assessed. When the Pearson correlation coefficients (R-squared) of the read counts (i.e., the number of unique reads mapped at a given nucleotide) across the genome and within the called peaks were calculated, the read coverage between these two groups was highly correlated, as shown in a representative image in Figure 4E and Figure S2 (Supplementary Materials) illustrates the correlation plots for other samples. Overall, the average correlation coefficients between the read coverages for individual and pool processing were 0.91 \pm 0.05 (mean \pm SD; Table 3), demonstrating a high similarity of genome-wide read coverage correlation and read coverage on merge of all samples' peak regions for both samples of both groups.

Fold enrichment per CpG island was also evaluated between individual and pool processing methods. The Pearson correlation coefficient between the fold enrichment values for every bin in the gene body for the two groups (individual vs. pool possessing, Figure 4F) was calculated and summarized in Table 3.

For all the samples, the result of ComputeMatrix over CpG island showed a strong correlation (Pearson correlation coefficient always >0.95), indicating that each sample either processed through Mx-MeDIP or standard MeDIP have peaks called on a similar region or CpG island.

3.4. Optimizing Sample Needs for Mx-MeDIP-Seq

As it is not always possible to obtain identical yields from each sample, a potential concern is whether a greater or lesser representation of individual samples in the pool could affect the outcome. To test this notion, several pools were prepared containing 200 ng of input DNA but containing different ratios of constituent samples. One pool contained an equal distribution (25%) of each sample (50 ng each, Figure 5A). The second pool contained different amounts of each DNA sample; the first sample was 10% of the pool (20 ng), the second sample was 20% (40ng), and samples 3 and 4 were 30% (60 ng) and 40% (80 ng), respectively (Figure 5A). After three Mx-MeDIP repeats, the normalized percent recovery and fold enrichment were assessed using qPCR, and read depth and fold enrichment per CpG islands were assessed using sequencing. Enrichment efficacy evaluation through qPCR (Figure 5B) showed that normalized percent recovery was >95% for all four positive targets. These two pools had similar fold enrichment ratios of 21.6 ± 2.1 vs. 21.2 ± 2.2 (mean \pm SD, equally distributed and differentially distributed pools, respectively). The results suggest that since each pool contained the same amount of DNA (200 ng), the final enrichment and normalized percent recovery were similar, independent of the amount and proportion of each DNA sample.

Similarly, the percentage of sequencing reads for each sample corresponded directly to the fraction that sample comprised in the total pool (Figure 5C). For Pool_1, comprising four samples equally distributed, the percentage of total reads after sequencing ranged from 21% to 27% for each sample. On the other hand, pool_2 comprised a different fraction of each sample; the sequencing read percentage (11.6%, 21.9%, 27.7%, and 38.8%) corresponded closely to the percentage of each sample in the pool (10%, 20%, 30%, and 40%) (Figure 5C).

To assess the similarity between read coverages of samples in the equally distributed pool and the differentially distributed pool, scatter plots were acquired using the deeptools package (Figure 5D). The read coverage correlations between equal and differentially distributed pools were 97% to 99% for all four samples, indicating that the qPCR and sequencing results were insensitive to differences in the fraction that a sample comprises in the pool. Next, the peaks located on CpG islands were evaluated. The Pearson correlation coefficient between the sample making 10% of the pool (S11_10%) and the sample making 25% of the pool (S11_25%), as well as 20%, 30%, 40% vs. 25%, were calculated. There was more than 88% agreement between the fold change in the peaks called on CpG island while the samples ranged from 10% to 40% compared to constant with as little as 20 ng of DNA (Figure 5E). As the availability of DNA might vary in different situations, Mx-MeDIP outcomes such as the read coverage and fold changes in peaks remained resilient to different relative fractions of DNA material in the pool over a 4-fold range (10–40%).



Figure 5. Comparison of Mx-MeDIP-Seq pools with equal distribution and different distribution of DNA. (**A**) Pools that contain equal amounts (25% each, equally distributed samples) vs. different amounts of DNA samples (ranging from 10% to 40%). (**B**) qPCR result: normalized percent recovery (left) and fold change ratio (right) for a pool of equal amounts of starting DNA vs. the different amounts of starting DNA. For these studies, we carried out n = 3 for each condition. (**C**) Total number of reads (million) and percentage of reads of each sample in the equally distributed and differentially

distributed pool after sequencing. (**D**) Correlation of reading coverage between equally and differentially distributed pool. Left to right (sample_11 (S11) with 10% in a differentially distributed pool and 25% in an equally distributed pool. S12 with 20% in the differentially distributed pool and 25% in the equally distributed pool. S13 with 30% in the differential pool and 25% in the equally distributed pool. S14 with 40% in the differential pool and 25% in the equally distributed pool. (**E**) Comparing the peak correlation among samples with different amounts of initial input DNA for Mx-MeDIP-Seq (left to right: Heatmap plot of peaks called on CpG islands for S11 with 20 ng initial DNA compared to 50 ng of initial DNA. S12 with 40 ng initial DNA compared to 50 ng of initial DNA. S12 with 40 ng initial DNA. K. S14 with 80 ng initial DNA compared to 50 ng of initial DNA).

3.5. Mx-MeDIP-Seq Can Be Performed Using Small Amounts of DNA

We also investigated how much DNA was required for consistent results. The protocol was tested on three replicates of different amounts of pooled DNA pool, 4 ng, 20 ng, 40 ng, 100 ng, 200 ng, and 400 ng. Each pool contained four different DNA samples with an equal amount of uniquely indexed DNA libraries. After immunoprecipitation, the products were analyzed quantitatively using qPCR to determine the normalized percent recovery and fold enrichment ratios. Mx-MeDIP-Seq was considered successful when normalized percent recovery the fold change ratio was >5.

The yield of captured DNA increased as expected with the increased DNA starting material. As Figure 6 indicates, performing Mx-MeDIP-Seq using 25 ng, 50 ng, and 100 ng led to a normalized percent recovery \geq 95% and a fold change ratio >5 for all targets. Moreover, there was no significant difference in normalized percent recovery using 50 ng and 100 ng of starting DNA. Lowering the DNA amount to 1 ng reflected the limitation in distinguishing methylated and unmethylated DNA.



Figure 6. Effect of minimizing the amount of input DNA for Mx-MeDIP-Seq. Normalized percent recovery and fold change ratio for a pool of DNA carried through Mx-MeDIP. A triplicate of six pools

of DNA was prepared. Each pool contains four different DNA samples (the final amount of DNA were 4 ng, 20 ng, 40 ng, 100 ng, 200 ng, 400 ng). Starting DNA amounts were reduced from 100 ng to 1 ng in each different pool. The final enrichment efficacy was assessed using qPCR data. Normalized percent recovery and fold change ratio hold their high value while the DNA amount was reduced to 25 ng. A significant decrease in normalized percent recovery and fold change ratio was observed when the DNA amount was reduced to 10 ng. Statistical analysis between groups was determined using one-way ANOVA and Tukey's test; *p*-values < 0.05 were considered statistically significant. For these studies, we carried out *n* = 3 for each condition, and in the figure, * *p*-value < 0.05, ** *p*-value < 0.01, *** *p*-value < 0.001.

In another similar experiment (Figure S2), the Mx-MeDIP technique with different amounts of DNA pools was compared to standard MeDIP using 50 ng of individual DNA samples. The pools contained 50 ng, 250 ng, and 500 ng of DNA composed of ten samples with equal distribution. There was no significant difference using 50 ng of DNA processed individually or in the pool based on enrichment efficiency measurement using the fold change ratio and normalized percent recovery. The reduction in starting materials from 50 ng to 25 ng did not show a statistically significant change in the normalized percent recovery and fold change ratio. The normalized percent recovery was \geq 95%, and the fold change ratio was >10 for both groups indicating that lowering the DNA amounts to 25 ng did not compromise the enrichment result (Figure 6). These two experiments indicated that using 25 ng and 50 ng of individual DNA samples in the pool resulted in successful Mx-MeDIP regardless of the number of samples in the pool in our experimental setting. Additionally, there was a high read coverage correlation and peak correlation of sequence data acquired from 25 ng and 50 ng of DNA (Figure 5D,E).

4. Discussion

Epigenetic phenomena are heritable changes in cellular phenotype that are not due to mutations in DNA sequence. DNA methylation is a key epigenetic mark that is associated with transcriptional repression during development, the maintenance of homeostasis, and disease [70–72]. Many technologies have been successfully used over the years to measure DNA methylation, each with its own advantages and disadvantages [73]. Bisulfite sequencing method is a treatment that distinguishes unmethylated from methylated cytosines. Common affinity enrichment strategies include methyl-DNA immunoprecipitation using methylation specific antibodies (MeDIP [61,74,75]) and methyl-CpG-binding domain protein (MBD [76,77]). Both MeDIP- and MBD-based methods have the advantage of being low-cost and straightforward for laboratories that are already skilled in chromatin immunoprecipitation sequencing (ChIP-seq) or related techniques. The disadvantages include relatively low resolution, susceptibility to copy number variation bias, and a bias toward observations of methylated DNA compared with bisulfite-based protocols [73]. One caveat with affinity approaches is that methylated CpG-rich sequences may give higher enrichments than methylated CpG-low sequences. Improving the capacity and efficiency of the methods available for studying the methylome can expand the study of epigenetics.

Although methylated DNA immunoprecipitation followed by sequencing (MeDIP-Seq) was introduced to reduce the cost and complexity of the data associated with wholegenome bisulfite sequencing methods as the primary method to analyze global methylation information [78], its cost and low throughput remain a barrier to researchers wishing to carry out large-scale genome-wide methylation studies. There are three main barriers for converting low-throughput MeDIP-Seq into a high-throughput method: (1) DNA purification; (2) un-biased DNA fragmentation; and (3) labor-intensive process of individual immunoprecipitations. In this study, we addressed all three impediments by coupling DNA fragmentation with cell lysis/purification and barcoding individual samples and multiplexing by pooling samples; Mx-MeDIP-Seq allowed an individual researcher to process up to sixty samples simultaneously while they could process six samples using historical methods. In Mx_MeDIP-Seq, the use of enzymatic DNA digestion, instead of mechanical fragmentation, provided a uniform size distribution for many samples at once and minimized sample loss during library preparation. DNA over 1000 bp was removed by size selection with Ampure XP beads. Larger fragments (>1000 bp) were shown to have lower efficiency and lower yield during amplification on the flowcell surface while sequencing [60,61]. MNase digestion based MeDIP-Seq method produce similar results in identifying CpG islands compared to DNA fragmentation by sonication (Figure S3, Supplementary Materials). Although there is a clear link between nucleosome occupancy and DNA methylation [79–83], MNase used for DNA fragmentation did not disproportionately affect the identification of DNA methylation comparted mechanical shearing such as sonication.

Following purification, the efficiency of the MeDIP reaction was assessed using quantitative PCR of known hypermethylated regions. To define a reasonable criteria threshold, we have found the normalized percent recovery to be more than 90% while the fold change ratio ranged from 4.6 to 100 in independent studies [49,66,67]. Samples that passed this QC threshold were taken forward for sequencing. Primer pairs were selected for enrichment estimation based on qPCR, amplifying either methylated or unmethylated DNA fragments. *TSH2B, BRDT, ZC3H13,* and *NBR2* genes were selected as the positive controls since they were known to be highly methylated in PBMCs [54,55,62–64]. The human *GAPDH* primer was considered a negative control that recognizes the human *GAPDH* promoter as an unmethylated region in all somatic cells [62,63,65]. Our results show that both approaches (standard MeDIP-Seq and Mx-MeDIP-Seq) resulted in a near-identical enrichment efficiency using the fold-change ratio and normalized percent recovery criteria (Figure 4B,C) based on the known hypermethylated region. It is possible that not every methylated region could yield similar enrichment between conventional and Mx-MeDIP-Seq methods, or between Mx-MeDIP-Seq and bisulfite sequencing.

Poor or inefficient adapter ligation, low starting DNA quality, and quantity, the short incubation time for IP reactions, PCR failure, and the inaccurate calculation of normalized concentrations of samples may result in a low normalized percent recovery and a low fold change ratio for both MeDIP-Seq or Mx-MeDIP-Seq. However, in the current study, optimizing each step, the normalized percent recovery and fold change ratio remain at a successful level. Additionally, the manual library preparation step in Mx-MeDIP-Seq can be automated to reduce errors during pipetting, reduce processing time and labor-intensiveness, and produce consistent results for multiple samples.

After purification and amplification, the samples were sent to a next-generation sequencing platform to investigate the DNA methylome. The sequence reads were then demultiplexed based on their unique adapters. This positively affects cost, such that MeDIP-seq becomes at least $10 \times$ more cost-effective.

Looking at the sequence results, the read coverage strong correlation (Pearson value of 0.9; Figure 4) was achieved for the multiplexing approach compared to standard MeDIP. A similar correlation analysis was performed across the CpG islands to look at the genome-wide concordance between Mx-MeDIP and the standard one, which resulted in a high correlation on a genome-wide level, (0.99 for Pearson correlation coefficient) for all ten samples taken through the two approaches (Table 3; Figure S4 in Supplementary Materials). This strong correlation indicates that Mx-MeDIP can produce the same data from large numbers of samples simultaneously and relatively easier. This approach cuts hands-on time by approximately two-thirds compared with conventional MeDIP.

The available yield of individual DNA for Mx-MeDIP-Seq may vary in different situations due to sample availability or due to differences in DNA extraction efficiency. We tested the tolerance of Mx-MeDIP for varying amounts of individual DNA by pooling DNA at different amounts. Mx-MeDIP outcomes such as enrichment efficacy and sequence data showed resilience to using different fractions of DNA material in the pool (Figures 5 and 6). A significant correlation ($R^2 > 0.9$) for CpG islands between using 20 ng of DNA in a 200 ng pool compared to using 50 ng in a 200 ng pool indicates that Mx-MeDIP produced similar data considering this change. In the case of limited availability of clinical samples, a flexible amount of DNA as small as 20 ng will still give a similar methylome result and an excellent normalized percent recovery and fold change ratio. This amount of DNA can be isolated from as few as 3400 to 6200 cells, depending on the tissue type [52], making the method more applicable to precious samples. In terms of throughput, we currently and routinely perform Mx-MeDIP for 100 samples in 24 h with one technician, whereas standard methods [48,51,69] can take weeks and cost substantially more.

Although this method can be automated to perform hundreds of samples simultaneously using a simple liquid handling robots such as Beckman i7, we have not shown any results using such methods. We are optimizing methods to perform large number of samples using Beckman i7 robots equipped with different temperature incubation stations, magnetic bead retrieving capabilities, and shakers. Experiments covering broad range of DNA concentrations, 25 ng to 2 μ g, will also be necessary to mitigate the varying DNA extraction efficiency of hundreds of patient samples in a typical clinical specimen pool. We recommend optimizing DNA extraction efficiency for each specimen type.

Multiple automated methylated DNA immunoprecipitation sequencing protocols, such as AutoMeDIP-Seq and GBS-MeDIP, were proposed and proven to work for low- to medium-throughput sample numbers [67,84]. The advantage of MeDIP is that genomic capture is relatively unbiased and not limited to restriction sites or CpG islands. GBS-MeDIP uses a similar multiplexing strategy to Mx-MeDIP-Seq using *pst1* restriction digestion and adapter ligation [84–86]. GBS-MeDIP has been adapted to multiple tissue/specimen types [85–87]. MxMeDIP provides significantly higher throughput capabilities due to simultaneous DNA fragmentation and purification through MNase digestion. At a resolution determined by DNA fragmentation (110 bp, here), MeDIP-seq offers comparable coverage at much lower cost than MethylC-seq.

In summary, Mx-MeDIP-seq is the method of choice when a cost-effective, unbiased method that requires a small amount of DNA and allows the processing of large numbers of samples in one experiment is needed. This method would be the ideal starting point for an automated or even portable device for monitoring epigenomic methylation.

5. Conclusions

This study proposed a novel technique to perform methylated DNA immunoprecipitation followed by sequencing (Mx-MeDIP-Seq). DNA was simultaneously extracted and fragmented through MNase digestion. After adapters ligation, different samples with unique adapters were pooled together and underwent immunoprecipitation using an anti-5-methylcytosine monoclonal antibody. We performed Mx-MeDIP-Seq for pools containing four to ten different samples in our study. The result showed that pool processing had similar enrichment efficacy compared to individual processing. Additionally, there was a strong correlation between read coverage and a strong correlation between fold change in the peaks on CpG islands of individual and pooled processed samples. Moreover, using a different fraction of DNA in a pool would not affect Mx-MeDIP outcomes such as the read coverage and fold changes in peaks. The method was optimized to use different amounts of starting materials for each sample and pool, and 25 ng of DNA showed comparable results with 100 ng of DNA. Altogether, the proposed approach for methylated DNA immunoprecipitation (Mx-MeDIP-Seq) results in a faster and more cost-effective method. It requires fewer amounts of starting DNA and is less prone to technical errors, making this method more applicable to a broader range of samples and studies.

Supplementary Materials: The following supporting information can be downloaded at https:// www.mdpi.com/article/10.3390/dna4040028/s1, Table S1: Preparation of stock reagents used in lysis buffer, Table S2: Quality metrics in sequencing, Figure S1: Correlation of read coverage between the samples that were processed in the pool (Mx-MeDIP-seq) and the same sample that was processed individually (conventional MeDIP), Figure S2: Effect of minimizing the amount of input DNA for Mx-MeDIP-Seq, Figure S3: Comparisons of MNase treated vs sonicated DNA MeDIP, Figure S4: Shown are peak profiles of 2 representative loci with CpG islands on Chr 21 (top) and Chr 1 (bottom) for 10 samples analyzed by either MeDIP-Seq (in red) or Mx-MeDIP-Seq (in blue). Snapshots were taken by using IGV.

Author Contributions: Conceptualization, V.M., J.L., J.G.P. and I.R.; methodology, V.M. and I.R.; formal analysis, I.R., C.X., Y.Z., Y.C., J.G.P., J.L. and V.M.; writing—original draft preparation, I.R., Y.C. and V.M.; writing—review and editing, J.G.P. and J.L.; visualization, I.R. and V.M.; supervision, V.M.; funding acquisition, J.L. and V.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by DARPA (HR001118S0023, AWD00033689) and ABOR (2950007-01).

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of ARIZONA STATE UNIVERSITY (protocol # STUDY00009580 approved 4 February 2019).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data openly available in a public repository that issues datasets with the GEO accession number GSE224820; approval date: 28 July 2023.

Acknowledgments: We sincerely thank Thiruppavai Chandrasekaran, Candyce McDaniel, and Rowan Johnson, research associates, for their support.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Boyko, A.; Kovalchuk, I. Genome instability and epigenetic modification—Heritable responses to environmental stress? Curr. Opin. Plant Biol. 2011, 14, 260–266. [CrossRef] [PubMed]
- 2. Baylin, S.B. DNA methylation and gene silencing in cancer. Nat. Clin. Pract. Oncol. 2005, 2 (Suppl. S1), S4–S11. [CrossRef]
- Tirado-Magallanes, R.; Rebbani, K.; Lim, R.; Pradhan, S.; Benoukraf, T. Whole genome DNA methylation: Beyond genes si-lencing. Oncotarget 2017, 8, 5629–5637. [CrossRef]
- 4. Zemach, A.; McDaniel, I.E.; Silva, P.; Zilberman, D. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* **2010**, *328*, 916–919. [CrossRef] [PubMed]
- Feng, S.; Cokus, S.J.; Zhang, X.; Chen, P.Y.; Bostick, M.; Goll, M.G.; Hetzel, J.; Jain, J.; Strauss, S.H.; Halpern, M.E.; et al. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. USA* 2010, 107, 8689–8694. [CrossRef] [PubMed]
- Smith, Z.D.; Meissner, A. DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* 2013, 14, 204–220. [CrossRef] [PubMed]
- 7. Greenberg, M.V.C.; Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 590–607. [CrossRef]
- 8. Fuks, F.; Hurd, P.J.; Wolf, D.; Nan, X.; Bird, A.P.; Kouzarides, T. The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J. Biol. Chem.* **2003**, *278*, 4035–4040. [CrossRef]
- 9. Della Ragione, F.; Filosa, S.; Scalabri, F.; D'Esposito, M. MeCP2 as a genome-wide modulator: The renewal of an old story. *Front. Genet.* **2012**, *3*, 181. [CrossRef]
- Ludwig, A.K.; Zhang, P.; Hastert, F.D.; Meyer, S.; Rausch, C.; Herce, H.D.; Muller, U.; Lehmkuhl, A.; Hellmann, I.; Trummer, C.; et al. Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise. *Nucleic Acids Res.* 2017, 45, 2438–2457. [CrossRef]
- Chowdhury, B.; Cho, I.H.; Hahn, N.; Irudayaraj, J. Quantification of 5-methylcytosine, 5-hydroxymethylcytosine and 5-carboxylcytosine from the blood of cancer patients by an enzyme-based immunoassay. *Anal. Chim. Acta* 2014, 852, 212–217. [CrossRef] [PubMed]
- 12. Ito, S.; Shen, L.; Dai, Q.; Wu, S.C.; Collins, L.B.; Swenberg, J.A.; He, C.; Zhang, Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **2011**, *333*, 1300–1303. [CrossRef] [PubMed]
- 13. Moore, J.R.; Hubler, S.L.; Nelson, C.D.; Nashold, F.E.; Spanier, J.A.; Hayes, C.E. 1,25-Dihydroxyvitamin D3 increases the methionine cycle, CD4+ T cell DNA methylation and Helios+ Foxp3+ T regulatory cells to reverse autoimmune neurodegenerative disease. *J. Neuroimmunol.* **2018**, *324*, 100–114. [CrossRef] [PubMed]
- 14. Richardson, B. DNA methylation and autoimmune disease. Clin. Immunol. 2003, 109, 72–79. [CrossRef] [PubMed]
- Heyn, H.; Vidal, E.; Ferreira, H.J.; Vizoso, M.; Sayols, S.; Gomez, A.; Moran, S.; Boque-Sastre, R.; Guil, S.; Martinez-Cardus, A.; et al. Epigenomic analysis detects aberrant super-enhancer DNA methylation in human cancer. *Genome Biol.* 2016, 17, 11. [CrossRef]

- Cronin, P.; McCarthy, M.J.; Lim, A.S.P.; Salmon, D.P.; Galasko, D.; Masliah, E.; De Jager, P.L.; Bennett, D.A.; Desplats, P. Circadian alterations during early stages of Alzheimer's disease are associated with aberrant cycles of DNA methylation in BMAL1. *Alzheimers Dement.* 2017, 13, 689–700. [CrossRef]
- 17. Mossman, D.; Scott, R.J. Epimutations, inheritance and causes of aberrant DNA methylation in cancer. *Hered. Cancer Clin. Pract.* **2006**, *4*, 75–80. [CrossRef]
- Liu, L.; Wylie, R.C.; Andrews, L.G.; Tollefsbol, T.O. Aging, cancer and nutrition: The DNA methylation connection. *Mech. Ageing Dev.* 2003, 124, 989–998. [CrossRef]
- McCabe, D.C.; Caudill, M.A. DNA methylation, genomic silencing, and links to nutrition and cancer. *Nutr. Rev.* 2005, 63, 183–195. [CrossRef]
- 20. Herceg, Z. Epigenetics and cancer: Towards an evaluation of the impact of environmental and dietary factors. *Mutagenesis* **2007**, 22, 91–103. [CrossRef]
- 21. Feinberg, A.P.; Tycko, B. The history of cancer epigenetics. Nat. Rev. Cancer 2004, 4, 143–153. [CrossRef] [PubMed]
- Klein, C.B.; Conway, K.; Wang, X.W.; Bhamra, R.K.; Lin, X.H.; Cohen, M.D.; Annab, L.; Barrett, J.C.; Costa, M. Senescence of nickel-transformed cells by an X chromosome: Possible epigenetic control. *Science* 1991, 251, 796–799. [CrossRef]
- 23. Illingworth, R.; Kerr, A.; Desousa, D.; Jorgensen, H.; Ellis, P.; Stalker, J.; Jackson, D.; Clee, C.; Plumb, R.; Rogers, J.; et al. A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol.* **2008**, *6*, e22. [CrossRef] [PubMed]
- 24. Weng, Y.I.; Huang, T.H.; Yan, P.S. Methylated DNA immunoprecipitation and microarray-based analysis: Detection of DNA methylation in breast cancer cell lines. *Methods Mol. Biol.* 2009, 590, 165–176. [CrossRef] [PubMed]
- Thu, K.L.; Pikor, L.A.; Kennett, J.Y.; Alvarez, C.E.; Lam, W.L. Methylation analysis by DNA immunoprecipitation. *J. Cell Physiol.* 2010, 222, 522–531. [CrossRef]
- Suzuki, M.M.; Bird, A. DNA methylation landscapes: Provocative insights from epigenomics. *Nat. Rev. Genet.* 2008, 9, 465–476. [CrossRef]
- Ji, L.; Sasaki, T.; Sun, X.; Ma, P.; Lewis, Z.A.; Schmitz, R.J. Methylated DNA is over-represented in whole-genome bisulfite se-quencing data. *Front. Genet.* 2014, 5, 341. [CrossRef]
- Cokus, S.J.; Feng, S.; Zhang, X.; Chen, Z.; Merriman, B.; Haudenschild, C.D.; Pradhan, S.; Nelson, S.F.; Pellegrini, M.; Jacobsen, S.E. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 2008, 452, 215–219. [CrossRef]
- 29. Gu, H.; Smith, Z.D.; Bock, C.; Boyle, P.; Gnirke, A.; Meissner, A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **2011**, *6*, 468–481. [CrossRef]
- Libertini, E.; Heath, S.C.; Hamoudi, R.A.; Gut, M.; Ziller, M.J.; Czyz, A.; Ruotti, V.; Stunnenberg, H.G.; Frontini, M.; Ouwehand, W.H.; et al. Information recovery from low coverage whole-genome bisulfite sequencing. *Nat. Commun.* 2016, 7, 11306. [CrossRef]
- Cao, B.; Luo, H.; Luo, T.; Li, N.; Shao, K.; Wu, K.; Sahu, S.K.; Li, F.; Lin, C. The performance of whole genome bisulfite sequencing on DNBSEQ-Tx platform examined by different library preparation strategies. *Heliyon* 2023, 9, e16571. [CrossRef] [PubMed]
- Wang, S.S.; Hall, M.L.; Lee, E.; Kim, S.-C.; Ramesh, N.; Lee, S.H.; Jang, J.-Y.; Bold, R.J.; Ku, J.-L.; Hwang, C.-I. Whole-genome bisulfite sequencing identifies stage- and subtype-specific DNA methylation signatures in pancreatic cancer. *iScience* 2024, 27, 109414. [CrossRef] [PubMed]
- 33. Huang, Y.; Pastor, W.A.; Shen, Y.; Tahiliani, M.; Liu, D.R.; Rao, A. The behaviour of 5-hydroxymethylcytosine in bisulfite se-quencing. *PLoS ONE* **2010**, *5*, e8888. [CrossRef]
- Cerrizuela, S.; Kaya, O.; Kremer, L.P.M.; Sarvari, A.; Ellinger, T.; Straub, J.; Brunken, J.; Sanz-Morejón, A.; Korkmaz, A.; Mar-tín-Villalba, A. High-throughput scNMT protocol for multiomics profiling of single cells from mouse brain and pancreatic or-ganoids. STAR Protoc. 2022, 3, 101555. [CrossRef] [PubMed]
- Clark, S.J.; Argelaguet, R.; Kapourani, C.A.; Stubbs, T.M.; Lee, H.J.; Alda-Catalinas, C.; Krueger, F.; Sanguinetti, G.; Kelsey, G.; Marioni, J.C.; et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat. Commun.* 2018, *9*, 781. [CrossRef]
- 36. Hu, Y.; Huang, K.; An, Q.; Du, G.; Hu, G.; Xue, J.; Zhu, X.; Wang, C.-Y.; Xue, Z.; Fan, G. Simultaneous profiling of transcriptome and DNA methylome from a single cell. *Genome Biol.* **2016**, *17*, 88. [CrossRef]
- Kremer, L.P.M.; Braun, M.M.; Ovchinnikova, S.; Küchenhoff, L.; Cerrizuela, S.; Martin-Villalba, A.; Anders, S. Analyzing sin-gle-cell bisulfite sequencing data with MethSCAn. *Nat. Methods* 2024, 21, 1616–1623. [CrossRef]
- Smallwood, S.A.; Lee, H.J.; Angermueller, C.; Krueger, F.; Saadeh, H.; Peat, J.; Andrews, S.R.; Stegle, O.; Reik, W.; Kelsey, G. Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods* 2014, *11*, 817–820. [CrossRef]
- 39. Rauluseviciute, I.; Drablos, F.; Rye, M.B. DNA methylation data by sequencing: Experimental approaches and recommendations for tools and pipelines for data analysis. *Clin. Epigenet.* **2019**, *11*, 193. [CrossRef]
- 40. Branco, M.R.; Ficz, G.; Reik, W. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat. Rev. Genet.* **2011**, *13*, 7–13. [CrossRef]
- 41. Kubiura, M.; Okano, M.; Kimura, H.; Kawamura, F.; Tada, M. Chromosome-wide regulation of euchromatin-specific 5mC to 5hmC conversion in mouse ES cells and female human somatic cells. *Chromosome Res.* **2012**, *20*, 837–848. [CrossRef] [PubMed]
- 42. Wardenaar, R.; Liu, H.; Colot, V.; Colome-Tatche, M.; Johannes, F. Evaluation of MeDIP-chip in the context of whole-genome bisulfite sequencing (WGBS-seq) in Arabidopsis. *Methods Mol. Biol.* **2013**, 1067, 203–224. [CrossRef] [PubMed]

- 43. Schmitz, R.J.; Schultz, M.D.; Lewsey, M.G.; O'Malley, R.C.; Urich, M.A.; Libiger, O.; Schork, N.J.; Ecker, J.R. Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **2011**, *334*, 369–373. [CrossRef] [PubMed]
- 44. Becker, C.; Hagmann, J.; Muller, J.; Koenig, D.; Stegle, O.; Borgwardt, K.; Weigel, D. Spontaneous epigenetic variation in the Arabidopsis thaliana methylome. *Nature* **2011**, *480*, 245–249. [CrossRef]
- 45. Deaton, A.M.; Bird, A. CpG islands and the regulation of transcription. *Genes. Dev.* **2011**, *25*, 1010–1022. [CrossRef]
- 46. Shu, C.; Zhang, X.; Aouizerat, B.E.; Xu, K. Comparison of methylation capture sequencing and Infinium MethylationEPIC array in peripheral blood mononuclear cells. *Epigenet. Chromatin* **2020**, *13*, 51. [CrossRef]
- Mohn, F.; Weber, M.; Schubeler, D.; Roloff, T.C. Methylated DNA immunoprecipitation (MeDIP). *Methods Mol. Biol.* 2009, 507, 55–64. [CrossRef]
- Borgel, J.; Guibert, S.; Weber, M. Methylated DNA immunoprecipitation (MeDIP) from low amounts of cells. *Methods Mol. Biol.* 2012, 925, 149–158. [CrossRef]
- 49. Taiwo, O.; Wilson, G.A.; Morris, T.; Seisenberger, S.; Reik, W.; Pearce, D.; Beck, S.; Butcher, L.M. Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat. Protoc.* **2012**, *7*, 617–636. [CrossRef]
- 50. Zhao, M.T.; Whyte, J.J.; Hopkins, G.M.; Kirk, M.D.; Prather, R.S. Methylated DNA immunoprecipitation and high-throughput sequencing (MeDIP-seq) using low amounts of genomic DNA. *Cell. Reprogram.* 2014, *16*, 175–184. [CrossRef]
- 51. Thu, K.L.; Vucic, E.A.; Kennett, J.Y.; Heryet, C.; Brown, C.J.; Lam, W.L.; Wilson, I.M. Methylated DNA immunoprecipitation. *J. Vis. Exp.* **2009**, *23*, e935. [CrossRef]
- 52. Gillooly, J.F.; Hein, A.; Damiani, R. Nuclear DNA Content Varies with Cell Size across Human Cell Types. *Cold Spring Harb. Perspect. Biol.* **2015**, *7*, a019091. [CrossRef] [PubMed]
- 53. Brind'Amour, J.; Liu, S.; Hudson, M.; Chen, C.; Karimi, M.M.; Lorincz, M.C. An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* **2015**, *6*, 6033. [CrossRef] [PubMed]
- Zhu, Y.; Cao, Z.; Lu, C. Microfluidic MeDIP-seq for low-input methylomic analysis of mammary tumorigenesis in mice. *Analyst* 2019, 144, 1904–1915. [CrossRef]
- 55. Auriol, E.; Billard, L.M.; Magdinier, F.; Dante, R. Specific binding of the methyl binding domain protein 2 at the BRCA1-NBR2 locus. *Nucleic Acids Res.* **2005**, *33*, 4243–4254. [CrossRef]
- 56. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10–12. [CrossRef]
- 57. Li, H.; Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **2009**, 25, 1754–1760. [CrossRef]
- 58. Li, D.; Zhang, B.; Xing, X.; Wang, T. Combining MeDIP-seq and MRE-seq to investigate genome-wide CpG methylation. *Methods* **2015**, 72, 29–40. [CrossRef]
- 59. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoute, J.; Johnson, D.S.; Bernstein, B.E.; Nusbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, R137. [CrossRef]
- Bronner, I.F.; Quail, M.A.; Turner, D.J.; Swerdlow, H. Improved Protocols for Illumina Sequencing. *Curr. Protoc. Hum. Genet.* 2014, 79, 18.2.1–18.2.42. [CrossRef]
- Weber, M.; Davies, J.J.; Wittig, D.; Oakeley, E.J.; Haase, M.; Lam, W.L.; Schübeler, D. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* 2005, 37, 853–862. [CrossRef] [PubMed]
- 62. Lisanti, S.; von Zglinicki, T.; Mathers, J.C. Standardization and quality controls for the methylated DNA immunoprecipitation technique. *Epigenetics* **2012**, *7*, 615–625. [CrossRef] [PubMed]
- 63. Clark, C.; Palta, P.; Joyce, C.J.; Scott, C.; Grundberg, E.; Deloukas, P.; Palotie, A.; Coffey, A.J. A comparison of the whole genome approach of MeDIP-seq to the targeted approach of the Infinium HumanMethylation450 BeadChip((R)) for methylome profiling. *PLoS ONE* **2012**, *7*, e50233. [CrossRef]
- 64. Xu, Y.; Xu, C.; Kato, A.; Tempel, W.; Abreu, J.G.; Bian, C.; Hu, Y.; Hu, D.; Zhao, B.; Cerovina, T.; et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell* **2012**, *151*, 1200–1213. [CrossRef]
- 65. Kim, M.C.; Kim, N.Y.; Seo, Y.R.; Kim, Y. An Integrated Analysis of the Genome-Wide Profiles of DNA Methylation and mRNA Expression Defining the Side Population of a Human Malignant Mesothelioma Cell Line. J. Cancer 2016, 7, 1668–1679. [CrossRef]
- 66. Serre, D.; Lee, B.H.; Ting, A.H. MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res.* **2010**, *38*, 391–399. [CrossRef]
- 67. Butcher, L.M.; Beck, S. AutoMeDIP-seq: A high-throughput, whole genome, DNA methylation assay. *Methods* **2010**, *52*, 223–231. [CrossRef]
- 68. Shen, H.; Qiu, C.; Li, J.; Tian, Q.; Deng, H.W. Characterization of the DNA methylome and its interindividual variation in human peripheral blood monocytes. *Epigenomics* **2013**, *5*, 255–269. [CrossRef]
- 69. Jacinto, F.V.; Ballestar, E.; Esteller, M. Methyl-DNA immunoprecipitation (MeDIP): Hunting down the DNA methylome. *Biotechniques* **2008**, *44*, 35–43. [CrossRef]
- 70. Atlante, S.; Mongelli, A.; Barbi, V.; Martelli, F.; Farsetti, A.; Gaetano, C. The epigenetic implication in coronavirus infection and therapy. *Clin. Epigenet.* 2020, *12*, 156. [CrossRef]
- 71. Chlamydas, S.; Papavassiliou, A.G.; Piperi, C. Epigenetic mechanisms regulating COVID-19 infection. *Epigenetics* **2021**, *16*, 263–270. [CrossRef] [PubMed]

- 72. Kandi, V.; Vadakedath, S. Effect of DNA Methylation in Various Diseases and the Probable Protective Role of Nutrition: A Mini-Review. *Cureus* **2015**, 7, e309. [CrossRef]
- 73. Laird, P.W. Principles and challenges of genomewide DNA methylation analysis. Nat. Rev. Genet. 2010, 11, 191–203. [CrossRef]
- 74. Hsu, H.K.; Weng, Y.I.; Hsu, P.Y.; Huang, T.H.; Huang, Y.W. Detection of DNA methylation by MeDIP and MBDCap assays: An overview of techniques. *Methods Mol. Biol.* 2014, 1105, 61–70. [CrossRef]
- 75. Weber, M.; Hellmann, I.; Stadler, M.B.; Ramos, L.; Pääbo, S.; Rebhan, M.; Schübeler, D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **2007**, *39*, 457–466. [CrossRef]
- Aberg, K.A.; Chan, R.F.; Xie, L.; Shabalin, A.A.; van den Oord, E. Methyl-CpG-Binding Domain Sequencing: MBD-seq. *Methods Mol. Biol.* 2018, 1708, 171–189. [CrossRef]
- 77. Jørgensen, H.F.; Adie, K.; Chaubert, P.; Bird, A.P. Engineering a high-affinity methyl-CpG-binding protein. *Nucleic Acids Res.* 2006, 34, e96. [CrossRef]
- Frommer, M.; McDonald, L.E.; Millar, D.S.; Collis, C.M.; Watt, F.; Grigg, G.W.; Molloy, P.L.; Paul, C.L. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA* 1992, 89, 1827–1831. [CrossRef]
- 79. Collings, C.K.; Anderson, J.N. Links between DNA methylation and nucleosome occupancy in the human genome. *Epigenet*. *Chromatin* **2017**, *10*, 18. [CrossRef]
- Fenouil, R.; Cauchy, P.; Koch, F.; Descostes, N.; Cabeza, J.Z.; Innocenti, C.; Ferrier, P.; Spicuglia, S.; Gut, M.; Gut, I.; et al. CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. *Genome Res.* 2012, 22, 2399–2408. [CrossRef] [PubMed]
- Mieczkowski, J.; Cook, A.; Bowman, S.K.; Mueller, B.; Alver, B.H.; Kundu, S.; Deaton, A.M.; Urban, J.A.; Larschan, E.; Park, P.J.; et al. MNase titration reveals differences between nucleosome occupancy and chromatin accessibility. *Nat. Commun.* 2016, 7, 11485. [CrossRef] [PubMed]
- 82. Portela, A.; Liz, J.; Nogales, V.; Setién, F.; Villanueva, A.; Esteller, M. DNA methylation determines nucleosome occupancy in the 5'-CpG islands of tumor suppressor genes. *Oncogene* **2013**, *32*, 5421–5428. [CrossRef] [PubMed]
- 83. Zhang, T.; Zhang, W.; Jiang, J. Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. *Plant Physiol.* **2015**, *168*, 1406–1416. [CrossRef] [PubMed]
- Rezaei, S.; Uffenorde, J.; Gimm, O.; Hosseinpour Feizi, M.A.; Miemczyk, S.; Coutinho, L.L.; Jensen, P.; Guerrero-Bosagna, C.; Pértille, F. GBS-MeDIP: A protocol for parallel identification of genetic and epigenetic variation in the same reduced fraction of genomes across individuals. *STAR Protoc.* 2022, *3*, 101202. [CrossRef]
- 85. Pértille, F.; Alvarez-Rodriguez, M.; da Silva, A.N.; Barranco, I.; Roca, J.; Guerrero-Bosagna, C.; Rodriguez-Martinez, H. Sperm Methylome Profiling Can Discern Fertility Levels in the Porcine Biomedical Model. *Int. J. Mol. Sci.* **2021**, *22*, 2679. [CrossRef]
- Nunes, J.R.S.; Pértille, F.; Andrade, S.C.S.; Perazza, C.A.; Villela, P.M.S.; Almeida-Val, V.M.F.; Gao, Z.-X.; Coutinho, L.L.; Hilsdorf, A.W.S. Genome-wide association study reveals genes associated with the absence of intermuscular bones in tambaqui (Colossoma macropomum). *Anim. Genet.* 2020, *51*, 899–909. [CrossRef]
- 87. Pértille, F.; Da Silva, V.H.; Johansson, A.M.; Lindström, T.; Wright, D.; Coutinho, L.L.; Jensen, P.; Guerrero-Bosagna, C. Mutation dynamics of CpG dinucleotides during a recent event of vertebrate diversification. *Epigenetics* **2019**, *14*, 685–707. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.