Title

Nano-MeDIP-seq Methylome analysis using low DNA concentrations

Running Head: Nano-MeDIP-seq

Authors

Lee M. Butcher*† & Stephan Beck*†

*UCL Cancer Institute, University College London, London WC1E 6BT, UK

†Correspondence to LMB (l.butcher@ucl.ac.uk) or SB (s.beck@ucl.ac.uk)

Summary

DNA methylation is an epigenetic mark that is indispensable for mammalian development and occurs at cytosine residues throughout the genome (the "methylome"). Approximately 70% of all CpG dinucleotides are affected by DNA methylation, which serve to "lock in" chromatin states and thus transcriptional programs. The systemic and pervasive occurrence of DNA methylation throughout the genome defines cellular identity and therefore requires genome-wide assays to fully appreciate and discern differential patterns of methylation that influence aspects of phenotypic plasticity including susceptibility to common complex disease.

One method that permits methylome analysis is methylated DNA immunoprecipitation (MeDIP) combined with next generation sequencing (MeDIP-seq). MeDIP uses an antibody raised against 5-methylcytosine to capture methylated fragments of DNA, which are subsequently sequenced to envisage the methylome landscape. The advantageous cost versus coverage balance of MeDIP-seq has made it the method of choice to replace or complement array-based methods for population epigenetic studies. Here we detail nano-MeDIP-seq, which allows methylome analysis using nanogram quantities of starting material.

Key words

Epigenetics, DNA methylation, whole genome, next generation sequencing, low concentration, bioinformatics

1. Introduction

DNA methylation is the presence of a methyl group at the carbon-5 position of the cytosine pyrimidine ring, catalysed by DNA methyltransferases (1) forming 5-methylcytosine. DNA methylation is present throughout the genome in a tissue- and cell-specific fashion and its presence effectively “locks in” transcriptional programs dictated primarily by the local chromatin state. DNA methylation localises to all inter- and intra-genic features and, depending on where it occurs, can alter transcription both quantitatively (gene expression) and qualitatively (e.g., alternative splicing). There are a number of mechanisms by which DNA methylation achieves these effects and these are being discovered at an ever-increasing rate (see Ref. 2 for a review). The sequence context in which DNA methylation affects cytosines can be largely divided into two categories: at CG dinucleotides (‘mCG’) and non-CG dinucleotides (‘mC’). Post implantation, approximately 60–80% (16.5M-22.5M) of CG dinucleotides are methylated in all nucleated human cells, including stem cell populations (3-5). In embryonic pluripotent stem cells however, an additional 7.5M cytosines are methylated at non-CG dinucleotides (3;5), which permits a myriad of additional regulatory controls to achieve diverse cellular phenotypes.

Given that DNA methylation is distributed genome-wide, and that patterns of this distribution vary both within- and between- individuals for matched cell types, comprehensive and systematic assays are required. A number of technologies exist to meet this need (reviewed in Ref. 6), including Methylated DNA immunoprecipitation or ‘MeDIP’. MeDIP involves antibodies directed against mC or mCG to precipitate methylated DNA fragments. MeDIP is able to detect methylated cytosines in both mC and mCG contexts, and a major benefit is that it is capable of targeting the vast majority of the methylome.

Because antibodies used for MeDIP were raised in a way to yield equal specificity against mC and mCG, MeDIP offers a hypothesis-free approach without prior assumptions about which regions of the methylome might be targeted. The most information-rich way of capitalising on the largely balanced qualities of MeDIP has been to combine it with next-generation sequencing, or MeDIP-seq (7). This application provides high-quality methylomes at typically 100- to 300-bp resolution (depending on chosen insert size) at costs comparable to other capture-based techniques (8) and was used to generate the first methylome of any mammalian genome (7).

One limitation of MeDIP-seq concerns genomic resolution. Although single–base pair resolution is desirable, we feel that the resolution offered by MeDIP-seq offsets issues of coverage and cost associated with single-base pair (e.g., bisulfite treatment) sequencing-oriented technologies. We also consider 150–200 bp to be a suitable resolution for most applications, as DNA methylation at adjacent CpGs is correlated for up to approximately 1 kb (9).

Another limitation of MeDIP-seq is that methylated DNA recovery by the antibody is affected by mC/mCG density, such that regions of very low (<1.5%) density may be underrepresented or even interpreted as unmethylated. Furthermore, MeDIP enriches only for methylated portions of the genome, and unmethylated portions can only be inferred by an absence of reads. Consequently, the confidence placed on this inference is highly dependent on sequencing depth; the cost of this however, is continually falling (10) and may cease to become a limiting factor in the future.

In addition to being a cost-effective method for analyzing all currently known forms of mammalian DNA methylation, the MeDIP-seq protocol has the advantage that as little as 50ng DNA can be used (11). This makes MeDIP-seq suitable for studies involving minute clinical samples, microdissected tissues and rare cell types.

In our laboratory, MeDIP-seq libraries can be created in three days; however, inexperienced users might prefer to spread the protocol over five days. The time required for the actual sequencing depends on the model of the sequencer (at the time of writing an Illumina HiSeq 2500 takes ~5.5 days for a ‘high output’ paired-end run of 50 bases). Bioinformatic processing e.g. using MeDUSA(12) takes approximately 10 hours.

2. Materials

2.1. Reagents

1. 1 M Tris-HCl (pH 7.8)
2. 1 M Tris-HCl (pH 8.0)
3. 1 M Tris-HCl (pH 8.5) [Elution buffer]
4. 0.5 M EDTA solution
5. 5 M NaCl: Dissolve 292.2 g of NaCl in 1,000 ml of PCR grade water. Store at room temperature and use within 2 years.

6. 1× sodium chloride-Tris-EDTA (STE) buffer: Mix 100 µl 1 M Tris-HCl (pH 7.8), 20 µl 0.5 M EDTA, and 100 µl 5 M NaCl in a total volume of 10 ml (fill up with PCR grade water). Store at room temperature and store indefinitely. Final concentrations: 10 mM Tris-HCl, 1 mM EDTA and 50 mM NaCl.
7. 1× Tris-EDTA (TE) buffer: Mix 10 ml 1 M Tris-HCl (pH 8.0), 2 ml 0.5 M EDTA in a total volume of 1 l (fill up with PCR grade water). Store at room temperature and store indefinitely. Final concentrations: 10 mM Tris-HCl, 1 mM EDTA.
8. Lambda (λ)-DNA (NEB, cat. no. N3011S) for methylated control fragments (See Note[1])
9. SssI CpG methyltransferase (NEB, cat. no. M0226S)
10. Taq DNA Polymerase with 10x Buffer containing 15 mM MgCl2 (ABgene, cat no. AB-0192/A)
11. Agencourt Ampure XP (60 ml; Beckman Coulter, cat. no. A63881)
12. Library Preparation End repair module (NEB, cat. no. E6050L)
13. Library Preparation A-tail module (NEB, cat. no. E6053L)
14. Library Preparation Adapter ligation module (NEB, cat. no. E6056L)
15. Auto-MeDIP kit [Diagenode, cat. no. AF-Auto01-0016 (16 reactions) or AF-Auto01-0100 (100 reactions)]. See Note[2].
16. High-fidelity PCR kit (Kapa Biosystems, cat. no. KK2101)
17. Gel DNA Recovery Kit (Zymo Research, cat. no. D4001)
18. 96-100% Ethanol
19. 70% Ethanol
20. Agarose
21. 50bp DNA Ladder (see Note[3])
22. 10× Tris Borate EDTA (TBE)
23. Ethidium bromide (10mg/ml)
24. Gel loading dye containing bromophenol blue, xylene cyanol, and orange G
25. DNA1000 kit (Agilent, cat. no. 5067-1504)
26. High-Sensitivity kit (Agilent, cat. no. 5067-4626)
27. SYBR qPCR master mix
28. PCR grade water
29. Paired-end (PE) sequencing adapters (see Note[4]). Spin down lyophilized adapter oligos in a chilled (4 °C) microcentrifuge at 200g for 5 min. Resuspend each oligo in 1× STE buffer to 100 µM. Add equimolar quantities

of each adapter into a 1.5 ml microcentrifuge tube. Divide into small (e.g., 200 µl) aliquots and incubate at 95 °C for 15 min, and then leave to cool down to room temperature (~1 h). Store aliquots at −20°C until required. Final concentration: 50 µM.

30. Adapter-mediated PCR primers (see Note[5]). Spin down lyophilized primer oligos in a chilled (4 °C) microcentrifuge at 200g for 5 min. Resuspend each oligo in Elution buffer to create 100 µM stocks. Dilute aliquots tenfold in 1 × TE buffer to create 10 µM dilutions. Store dilutions in manageable volumes at − 20 °C until required.
31. Quality Control (QC) primers (see Note[6]).

2.2. Equipment

1. Microcentrifuge tubes (0.5 ml, 1.5 ml & 2 ml)
2. 8-strip, 0.2 ml PCR tubes and caps
3. 12-strip, 0.2 ml PCR tubes and caps
4. IP-Star tips (Diagenode, cat. no. WC-001-1000)
5. qPCR plates and seals
6. Centrifuge with 0.2 ml tube adapter (‘microfuge’)
7. Centrifuge with 2.0 ml adaptors
8. Pipettes
9. Pipette filter tips
10. Vortex
11. Transilluminator and imaging software
12. Microwave oven
13. Magnetic separation rack
14. Scalpels
15. Thermocycler
16. qPCR system
17. Gel electrophoresis system (inc. gel tanks and combs)
18. Weighing scale and weighing boats
19. Sonicator (see Note[7])

20. Heat block or oven
21. Agilent Bioanalyzer 2100
22. Illumina sequencer (see Note[8])
23. SX-8G IP-Star (see Note[9])

3. Methods

Before creating MeDIP-seq libraries, it is worth spending time familiarising oneself with the workflow (see **Figure 1**).

3.1. Creation of methylated and unmethylated controls

The following describes a one-off set of procedures to create a central laboratory resource of methylated and unmethylated control fragments that should be suitable for over a million MeDIP assays.

1. For each control region, combine the following reagents in a 0.2 ml PCR tube (note, it is often more convenient to make a mastermix excluding primers):

Reagent	**×1 reaction**
10× Reaction buffer IV	2.5 µl
$MgCl_2$ (25 mM)	1.0 µl
dNTP mix (20 mM)	1.0 µl
Forward primer (see **Table 1**)	1.25 µl
Reverse primer (see **Table 1**)	1.25 µl
PCR grade water	17.625 µl
λ-DNA (0.0005 ng/µl)	1.0 µl
Taq DNA polymerase	0.125 µl
Total Volume/Reaction	**25.0 µl**

2. Pipette mix and briefly centrifuge in a microfuge to consolidate reactions

3. Setup and run the following PCR program: incubate for 2 min at 94 °C, followed by a total of 40 cycles of (20 s at 94 °C, 30 s with touchdown: 53/65°C, 60 s at 72°C), then 5 min at 72°C
4. Purify amplicons using 1.8 volumes Ampure XP purification beads according to manufacturer's procedure (see Note[10]) and elute in 10 µl Elution buffer.
5. Run the sample on an Agilent Bioanalyzer 2100 using DNA1000 chips and reagents according to manufacturer's instructions. Verify DNA fragment size range and concentration using the instrument's software.
6. Divide each amplicon into two aliquots and label.
7. Set up one in vitro methylation reaction and one control reaction for each aliquot in a 0.2 ml PCR tube as follows and incubate for 2 hrs at 37°C, followed 20 mins at 65°C:

Reagent	***in vitro* methylation reaction**	**control reaction**
PCR grade water	10.0 µl	11.0 µl
10× NEBuffer 2	2.0 µl	2.0 µl
S-adenosylmethionine (1.6 mM)	2.0 µl	2.0 µl
Amplicon DNA (from step 4)	5.0 µl	5.0 µl
SssI methyltranserferase (4 U/µl)	1.0 µl	--
Total Volume/Reaction	**20.0 µl**	**20.0 µl**

8. Purify each reaction using 1.8 volumes Ampure XP purification beads according to manufacturer's procedure and elute in 10 µl Elution buffer.
9. Each amplicon contains at least one 5'-ACGT-3' motif, which allows the methylation status of each region to be verified using the methylation sensitive enzyme, *HpyCH4*IV. To do so, dilute 1 µl of each amplicon from each reaction 1 in 10 with PCR grade water
10. Set up the following reaction for each diluted DNA incubate for 1 hr at 37°C then 20 min at 65°C:

Reagent	**×1 reaction**
10× NEBuffer 1	2.5 µl
*HpyCH4*IV (10 U/µl)	1.0 µl
diluted DNA	10 µl

PCR grade water	11.5 µl
Total Volume/Reaction	**25.0 µl**

11. Purify each sample using 1.8 volumes Ampure XP purification beads according to manufacturer's procedure and elute in 10 µl Elution buffer.
12. Run the sample on an Agilent Bioanalyzer 2100 using DNA1000 chips and reagents according to manufacturer's instructions. Validate DNA fragment size range and concentration using the instrument's software. See Note[11].
13. Create 1 nM working dilutions of methylated and unmethylated controls based on Bioanalyzer readings
14. Use the 1 nM dilutions to create a equimolar cocktail of control fragments consisting of methylated (0.5%, 3.7%, 5.6%, 6.5% and 10% CpG density) and unmethylated (14.5% CpG density) fragments
15. Store stock solutions and working dilutions (e.g., 10µM) of in vitro methylated and unmethylated fragments at -20°C indefinitely
16. Run the sample on an Agilent Bioanalyzer 2100 using DNA High Sensitivity chips and reagents according to manufacturer's instructions. Validate DNA fragment size range and concentration using the instrument's software. See Note[12].
17. Dilute the cocktail in 1×TE buffer to create a 100pM (each) stock cocktail
18. Dilute aliquots of the 100 pM stock in 1× TE buffer to create working dilutions of 0.19 pM (or ~ 114,000 copies per µl). See Note[13].

3.2. DNA Fragmentation

1. In a 0.5 ml or 1.5 ml microcentrifuge tube, resuspend gDNA in 85 µl Elution buffer.
2. Pre-chill the sonicator water bath by filling with ice and leave for 15 min.
3. After 15 min, replace the ice in the water bath with ice-cold water.
4. Place microcentrifuge tube containing DNA in the appropriate tube holder, making sure the holder is balanced.
5. Set sonicator to 'High' and sonicate for 15 min consisting of 30-sec on/off periods.
6. Following 15 min cycle, briefly centrifuge DNA samples and store on ice.
7. Cool the sonicator for 5-10 min by filling with crushed ice.
8. Repeat steps 5-7 up to another six times (see Note[14])

9. Run the sample on an Agilent Bioanalyzer 2100 using High Sensitivity DNA chips and reagents according to manufacturer's instructions; validate DNA fragment size range and concentration using the instrument's software[15].

3.3. Library Preparation (1 of 2)

1. *DNA end repair.* Mix the following on ice in a sterile PCR tube and incubate for 30 mins at 20°C. Scale accordingly.

Reagent	×1 reaction
NEBNext End Repair Enzyme mix	5 µl
NEBNext End Repair Reaction Buffer (10X)	10 µl
Fragmented DNA	85 µl
Total Volume	**100 µl**

2. Place the reaction on ice after incubation, spin briefly in a microfuge for ~3 s to collect condensate.
3. Purify DNA sample using 1.8 volumes Ampure XP purification beads according to manufacturer's procedure and elute in 37 µl Elution buffer.
4. *dA-tailing*. Mix the following on ice in a sterile PCR tube and incubate for 30 mins at 37°C (see Note[16]). Scale accordingly.

Reagent	×1 reaction
Klenow Fragment (3' → 5' exo$^-$)	3 µl
NEBNext dA-tailing Reaction Buffer (10X)	5 µl
PCR grade water	5ul
Purified End-Repaired, blunt-ended DNA from Step 3	37 µl
Total Volume	**50 µl**

5. Purify DNA sample using 1.8 volumes Ampure XP purification beads according to manufacturer's procedure and elute in 25 µl Elution buffer.

6. Calculate the reaction concentration of DNA adapters to be used in the adapter ligation reaction (see Note[17])

7. *Adapter ligation.* Mix the following on ice in a sterile PCR tube and incubate in a thermocycler for 2 h at 18°C. Scale accordingly.

Reagent	**×1 reaction**
Quick Ligation Reaction Buffer (5x)	10 µl
Adapters	*Variable*
PCR grade water to 50 µl	*As required*
End Repaired, Blunt, dA- Tailed DNA from Step 5	25 µl
Quick T4 DNA Ligase	5 µl
Total Volume	**50 µl**

8. Purify DNA sample using 1.8 volumes Ampure XP purification beads according to manufacturer's procedure and elute in 40 µl Elution buffer.

9. Run 1 µl of eluted sample on an Agilent Bioanalyzer 2100 using High Sensitivity DNA chips and reagents according to manufacturer's instructions; determine the concentration of the full range of DNA fragments (see Note[18]).

3.4 Methylated DNA Immunoprecipitation

1. Perform MeDIP on purified adapter-ligated DNA from Step 8 (previous section), using the Diagenode Auto-MeDIP and SX-8G IP-star (see Note[2] and Note[19]).

2. *Antibody (Ab) Dilution.* Mix the following on ice and scale accordingly (see next step):

Reagent	**vol**
Antibody	1 µl
PCR grade water	15 µl
Total Volume	16 µl

3. *Antibody Mix*. Mix the following on ice:

Reagent	**×1 IP**
Diluted Ab (from Step 2)	2.40 µl
MagBuffer A (5X)	0.60 µl
MagBuffer C	2.00 µl
Total Volume	5.00 µl

4. *1 × Magbuffer A*. Gently mix the following on ice and scale accordingly:

Reagent	×1 IP
MagBuffer A (5X)	30 µl
PCR grade water	120 µl
Total Volume	150 µl

5. *Incubation Mix*. Mix the following (sufficient for 1× IP and 1× Input) on ice and scale accordingly:

Reagent	×1 IP
MagBuffer A (5X)	24 µl
MagBuffer B	6 µl
Lambda Spike cocktail (0.2 µM each; see Note[1])	3 µl
Purified Adapter Ligated DNA	*Variable*
PCR grade water up to 90 µl	*As required*
Total Volume/reaction	90 µl

6. Incubate *Incubation Mix(es)* at 99°C for 10 min.
7. Snap cool on ice for 10 min.
8. Switch on the IP Star (see Note[20]).
9. Mix the *Incubation Mix(es)* by pipetting up and down and spin for 3 sec on microfuge.

The following steps detail the loading of the IP-Star (a description of the automated process is presented in Note[21]).

10. Load reagents into a 12-strip tube as per **Figure 2** (1 strip tube per IP):
11. Place the filled 12-strip on the Peltier block(s) (see **Figure 3,** black-boxed zone and Note[22], Note[23], and Note[24]).
12. Place IP Star tips into tip rack (see **Figure 3,** red-boxed zone; see Note[25]).
13. Double click a tip rack in the GUI until it shows tips are present. The clicked tip rack should correspond to the tip rack just filled.
14. Select **8IPs or 16IPs** protocol for 8 and 16 MeDIP reactions respectively.
15. Click 'Modify' to set IP incubation time and temperature (see Note[26]).
16. Press *Start* and OK the confirmation screen that appears: The automated MeDIP reaction starts. After the IP Star has eluted the mDNA-Ab-bead complex in DIB solution, a dialog box appears on the screen.
17. Add 1 µl proteinase K (from AutoMeDIP kit) to well 1 and well 12 for each 12 strip.
18. Add 7.5 µl Input DNA into well 1 of the 12-strip.
19. Cap the 12-strip.
20. OK the dialog box and close the IP Star door.

21. OK the dialog box; the proteinase K digestion reaction will start.
22. Once completed, a confirmation message appears in the GUI: Remove samples from the IP Star and transfer the contents of well 12 (proteinase K digested IP reaction) into clean 0.2 ml PCR tubes and collect the magnetic beads on a magnetic rack (see Note[27]).
23. Shut down the IP Star software and turn off the IP Star off.

3.5 Quality Control: test for recovery of spiked-in λ-DNA and specificity of the MeDIP reaction (QC 1)

1. Prepare the following on ice in a qPCR plate: perform each reaction in triplicate for methylated and unmethylated control regions (see **Table 1**).

Reagent	**×1 reaction**
2x SYBR qPCR master mix	6.25 µl
Primer Pair (10 µM)	0.625 µl
PCR grade water	4.375 µl
MeDIP or Input DNA (from Step 22 of previous section)	1.25 µl
Total Volume/Reaction	**12.5 µl**

2. Seal plate and centrifuge briefly to ensure no bubbles remain at the bottom the wells.
3. Run the following program on the real time PCR machine: 1) 95°C for 5 min, 2) 40 cycles of 95°C for 15 s, 60°C for 1 min, C) meltcurve.
4. Calculate Recovery and Specificity (see Note[28]).

For troubleshooting problems associated with the MeDIP reaction see **Table 2.**

3.6 Library Preparation (2 of 2)

1. Purify DNA sample (from Step 22 of section 3.4) using 1.8 volumes Ampure XP purification beads according to manufacturer's procedure and elute in 25 µl Elution buffer
2. *Adapter-mediated PCR (amPCR).* On ice, mix and scale accordingly the following in a sterile PCR tube. Pipette up and down to mix.

Reagent	×1 reaction
5X KAPA GC Buffer	10 µl
10 mM dNTP Mix	1.5 µl
Forward primer (10 µM)	5 µl
Reverse primer (10 µM)	5 µl
DMSO	2.5 µl
Purified MeDIP DNA (from Step 1)	25 µl
1 U/µl KAPA HiFi Polymerase	1 µl
Total Volume	**50 µl**

3. Split each reaction into two 25 µl aliquots and perform PCR in a preheated thermocycler, using the following cycling conditions: 1) 95°C for 2 min, 2) 8-13 cycles of: 98°C for 20 s → 60°C for 15 s → 72°C for 15 s, C) 72°C for 5 min.
4. Combine the two PCR MeDIP reactions and purify DNA sample with 1.8 volumes Ampure XP purification beads and elute in 15µl Elution buffer (see Note[29]).
5. Run 1 µl of eluted sample on an Agilent Bioanalyzer 2100 using DNA1000 chips and reagents according to manufacturer's instructions; validate DNA fragment size range using the instrument's software. For troubleshooting see **Table 2**.
6. Prepare a 2% TBE agarose gel, with 0.5 µl/ml Ethidium Bromide (see Note[30]).
7. Mix the remaining PCR amplified MeDIP reactions (from Step 4) with an appropriate volume of loading dye and load on gel, leaving space either side for ladders (see Note[30]).
8. Load 1 µg of a 50 bp DNA ladder in wells flanking the DNA sample.
9. Carry out gel electrophoresis in freshly prepared 1 × TBE buffer at 100 volts until orange G dye runs off gel.
10. Following gel electrophoresis, carefully transfer gel onto a UV transilluminator.
11. With a clean scalpel, excise the desired 50 bp library size range[31]. Ensure the gel slice is cut as close to DNA smear as possible (see Note[32]). See Note[33].
12. Purify DNA libraries using Gel DNA Recovery Kit (Zymo Research) according to manufacturer's protocol and elute in 10 µl Elution buffer.
13. Assess 1 µl of size selected DNA on an Agilent Bioanalyzer using High Sensitivity DNA chips and reagents according to manufacturer's instructions to determine gel extraction has worked.

3.7 Quality Control: test for enrichment of methylated regions (QC 2)

1. Dilute 1 µl of size selected DNA to 9 µl with PCR grade water and validate the enrichment of genomic regions of known methylation status (see **Table 1** and Note[6]) by qPCR using reaction setup as detailed in Steps 1-2 of section 3.6.
2. Calculate enrichment as described Note[34]. For troubleshooting see **Table 2**.

3.8 Next generation sequencing and bioinformatic processing

1. Subject sample to next generation sequencing with an Illumina machine that uses flowcells compatible with the primer sequences used during amPCR according to manufacturer's instructions.
2. Use FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) to perform quality control of the reads.
3. Align MeDIP-seq reads to the relevant reference genome using BWA (see Note[35] and Note[36]).
4. Use BWA to run 'bwa aln' on both read ends to generate two .sai output files.
5. Use BWA to run 'bwa sampe' using the .sai files from previous Step as input; the output is a SAM formatted alignment file.
6. Use SAMtools to discard reads from the SAM formatted alignment file that fail to map as a proper pair.
7. Use MeDUSA to discard read pairs in which neither read scored a BWA alignment score ≥ 10.
8. Use MeDUSA to identify groups of non-unique reads (i.e. 'clonal reads' – reads that align to the exact same start and stop position on the same chromosome) to discard all but one read from the group.
9. Perform MeDIP-seq specific QC on the filtered data from previous Step using MeDIPs.

4. Notes

---

[1] The construction of six λ-DNA fragments of varying CpG content is described in Box 1 of Ref. (11)

[2] MeDIP can also be performed manually using, e.g., the MagMeDIP kit (Diagenode, cat. no. mc-magme-A10) or other kits/reagents. In this case, we advise taking care with respect to accurate and consistent timing of incubation steps, as well as performing lengthy and thorough bead washing steps.

[3] To increase the accuracy of gel excision we advise using a ladder with molecular-weight standards ranging from 50 to 700 bp in 50-bp size intervals.

[4] Depending on the experimental design, single-plex or multiplex Illumina adapters can be created and ligated to fragmented DNA. Table 1 presents adapter sequences for single-plex and multiplex applications, although other adapter systems could be used (e.g., see seqanswers.com) providing they DO NOT contain methylated cytosines.

[5] Depending on the experimental design (e.g., single-plex or multiplex library construction), primers complementary to the appropriate Illumina adapters should be used. Table 1 presents primer sequences for single-plex and multiplex applications, although other primers compatible with other adapter systems can be used (e.g., see seqanswers.com). If multiplexing, always pool > 3 indexed samples for sequencing.

[6] For QC1, nested qPCR primers were designed for each λ-DNA fragment; sequences are presented in Table 1. For QC2, primers were designed to amplify regions of assumed methylation status in appropriate reference genomes, e.g., for previous work in humans see Ref. (13); we present example sequences for mouse and human in Table 1. Prepare and treat oligos as detailed in Note 5.

[7] We recommend using the Bioruptor (Diagenode) sonicator for DNA fragmentation; however, other methods can be applied provided these can generate the specified DNA fragment range and peak.

[8] The protocol detailed in these pages describes applications compatible with MiSeq, GAIIx and HiSeq instruments.

[9] We use the SX-8G IP-Star for our automated MeDIP-seq protocol, as it was specially designed and programmed for this purpose; however, generic liquid handlers can be programmed to carry out the MeDIP step as detailed in the MagMeDIP protocol (Diagenode).

[10] After the second ethanol wash, place the keep the tubes on the magnetic rack and heat in an oven at 55°C for about 7 minutes to evaporate all the ethanol. When the bead pellet starts to crack, this is a good time to elute.

[11] In vitro methylated fragments will not be digested by HpyCH4IV and will result in intact, full length fragments that correspond to the length of the original amplicon; unmethylated fragments will be digested at the restriction site and show multiple fragment lengths

[12] The molarity of each fragment in the cocktail is expected to be 166 pM. If not, correct accordingly by adding more of the required fragment(s) to normalize concentrations and re-check on a Bioanalyzer 2100 with High Sensitivity chips

[13] For the vast majority of QC purposes we find that qPCR of just two regions [one methylated (10% CpG density) and one unmethylated (14% CpG density)] is sufficient to evaluate MeDIP. However, the full range of fragments in the cocktail can be assayed if desired

[14] Fragmented DNA can be stored at -20°C for up to 1 month.

[15] We aim to create fragment the sample so the majority of fragments are below 500bp and that 25-30% of fragments are between 150-250bp.

[16] If using a thermocycler, use a heated lid (100°C) during all incubations involving temperatures over 30°C to prevent evaporation of samples.

[17] To achieve 10-fold excess molarity of adapters to DNA fragments in a 50 µl reaction, the following heuristic provides a good estimate:

$$adapters_{rxn_conc}(nM) = \frac{606.1 \times r^2 \times ng}{bp},$$

where *ng* is starting concentration of DNA, *bp* is average size of DNA fragments and *r* is estimated recovery from beads. For example, a starting concentration of 200ng DNA with average size of 180bp, assuming 0.8 (80%) recovery from beads will require a reaction concentration of adapters of approximately 431 nM.

[18] If multiple samples are being run, determine the concentration across a set size range of fragments (e.g., 240-290bp) using the 'Smear Analysis' feature of the Bioanalyzer software. Sample input (ng) can then be normalized across samples during MeDIP to maintain reaction stoichiometry.

[19] The following six steps deviate from the manufacturer's instruction and represent optimized steps.

[20] The IP Star takes approximately 10 minutes to equilibrate to the appropriate temperature (4°C), so now is a convenient time to switch it on.

[21] The IP-Star (Diagenode) is capable of automating up to 16 MeDIPs per run. Each run is split into two parts with a manual intervention step in between. In the first part, the IP-Star washes the magnetic beads (to remove any preservatives), immunoselects and captures the methylated fraction, washes away unbound (unmethylated) DNA and elutes in a proteinase K buffer. The second part requires the manual addition of input DNA well 1 and proteinase K to

wells 1 and 12; this is followed by two 15 min incubations at 55°C and 95°C, respectively; the Peltier block then cools to 4°C until the user removes the samples. A small proportion of these products are subject to quality control by qPCR (see below) and the rest are purified. The result is a pure methylated fraction ready for amPCR.

[22]Peltier blocks and tip racks of the IP Star should be conceptualised as 96-well plates, i.e., 90° clockwise to the user. We call a row of 12 wells a 'lane' and the MeDIP protocol proceeds in the direction of the arrow so align Well 1 of the 12-strip tube with the beginning of the arrow.

[23] If running 8 or fewer IPs, load the 12-strips in the right-hand Peltier block. If running > 8 IPs, load the remaining 12-strips in the left-hand Peltier block.

[24] If running > 8 IPs it is best to balance the workload across both Peltier blocks and mirror the layout between blocks. Therefore, if running an odd number of IPs > 8 we recommend placing an empty 12-strip on the Peltier block containing fewer samples, to mirror the formation of the Peltier block with more samples.

[25] Each IP lane in the Peltier Block should have a corresponding lane of IP Star tips in a tip rack. When performing eight or fewer IPs, load 6 IP Star tips per lane; if performing between 8-16 IPs, load 9 IP Star tips per lane. Tip lanes must correspond to sample lanes.

[26] IP incubation times are for each block. We recommend at least 7.5 hours incubation time (15 hrs is preferable) and 4°C incubation temperature

[27] MeDIP and Input DNA can be stored at -20°C for up to 3 months and longer at -80°C

[28] From the qPCR results, 'Recovery' and 'Specificity' is calculated to evaluate MeDIP efficiency. Recovery, which is a measure of the MeDIP efficiency – indexed as percentage input – is calculated using the cycle threshold (Ct) of MeDIP and Input fractions from the qPCR reaction. Since only 10% of the DNA used in MeDIP is used for Input, the Ct value obtained for the Input is adjusted prior to calculating Recovery, using the formula:

$$Adjusted\,Input\,Ct = Input\,Ct - \frac{log_{10}(10)}{log_{10}(2 \times AE)}$$

Where AE is the % amplification efficiency expressed as a decimal (e.g., 100% AE = 1) and 10 represents the dilution factor of Input to MeDIP (i.e., 75 µl IP vs. 7.5 µl Input = x10).

Recovery is then calculated for each set of primers as:

$$Recovery\,(\%) = 2^{AE \times (Ct_{input} - Ct_{MeDIP})}\ x\ 100$$

Specificity of MeDIP, indexed as the ratio of methylated DNA recovery to unmethylated DNA is calculated as:

$$Specificity = 1 - \left\{ \frac{recovery_{unmeth}}{recovery_{meth}} \right\}$$

We consider MeDIP successful when specificity is ≥ 95% and that this has been derived from unmethylated recovery < 1%.

[29] Purified PCR products can be stored at -20°C for up to 3 months and longer at -80°C

[30] Multiple samples can be run on a single gel. However, maintain a minimum of 2cm (3 wells) between ladder and sample wells.

[31] We excise and purify 300-350bp libraries, which correspond to 180-230bp insert sizes. We also excise 250-300bp- and 350-400bp 'contingency' libraries.

[32] To prevent cross contamination of libraries, use a fresh scalpel for each library.

[33] Because UV exposure can be carcinogenic, wear protective clothing and UV resistant face shield.

[34] The aim of QC 2 is to verify the library insert remains enriched for methylated fragments and depleted for unmethylated fragments following amPCR and gel excision. In contrast to QC 1, QC 2 involves the qPCR of fragments from just the sample DNA. This is because the λ-DNA fragments are not ligated to adapter sequences and evade amplification during amPCR. Calculate the fold-enrichment ratio for methylated vs. unmethylated regions:

$$Fold\,Enrichment\,Ratio = 2 \times AE^{(Ct_{MeDIP_unmeth} - Ct_{MeDIP_meth})}$$

We recommend sequencing libraries if the fold-enrichment ratio exceeds 25, that is, the methylated fragments tested are 25-fold more enriched than the unmethylated fragments; we routinely see however, fold-enrichment ratios greater than 100 and higher scores are better.

[35] Prior to alignment, BWA should be used to index the reference genome using 'bwa index', using fasta files as input (the genomic fasta files are available from numerous public repositories including the UCSC Genome Browser (http://genome.ucsc.edu/))

[36] Other alignment tools such as Bowtie (14) and Novoalign (Novocraft Technologies) may also used

5. Acknowledgements

The authors wish to acknowledge funding support from IMI-JU OncoTrack (115234), EU-FP7 BLUEPRINT (282510) and a Royal Society Wolfson Research Merit Award (WM100023).

6. References

1. **Goll M.G., Bestor T.H. (2005) Eukaryotic cytosine methyltransferases. *Annu Rev Biochem* 74, 481-514**
2. **Ehrlich M., Lacey M. (2013) DNA methylation and differentiation: silencing, upregulation and modulation of gene expression. *Epigenomics* 5, 553-568**
3. **Lister R., Pelizzola M., Dowen R.H., Hawkins R.D., Hon G., Tonti-Filippini J., Nery J.R., Lee L., Ye Z., Ngo Q.M. et al. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462, 315-322**
4. **Li Y., Zhu J., Tian G., Li N., Li Q., Ye M., Zheng H., Yu J., Wu H., Sun J. et al. (2010) The DNA Methylome of Human Peripheral Blood Mononuclear Cells. *PLoS Biol* 8, e1000533**
5. **Lister R., Pelizzola M., Kida Y.S., Hawkins R.D., Nery J.R., Hon G., Antosiewicz-Bourget J., O'Malley R., Castanon R., Klugman S. et al. (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* 471, 68-73**
6. **Bock C. (2012) Analysing and interpreting DNA methylation data. *Nat Rev Genet* 13, 705-719**
7. **Down T.A., Rakyan V.K., Turner D.J., Flicek P., Li H., Kulesha E., Graf S., Johnson N., Herrero J., Tomazou E.M. et al. (2008) A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat Biotechnol* 26, 779-785**
8. **Beck S. (2010) Taking the measure of the methylome. *Nat Biotechnol* 28, 1026-1028**
9. **Eckhardt F., Lewin J., Cortese R., Rakyan V.K., Attwood J., Burger M., Burton J., Cox T.V., Davies R., Down T.A. et al. (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* 38, 1378-1385**
10. **Pettersson E., Lundeberg J., Ahmadian A. (2009) Generations of sequencing technologies. *Genomics* 93, 105-111**

11. **Taiwo O., Wilson G.A., Morris T., Seisenberger S., Reik W., Pearce D., Beck S., Butcher L.M. (2012) Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat Protoc* 7, 617-636**
12. **Wilson G.A., Dhami P., Feber A., Cortazar D., Suzuki Y., Schulz R., Schar P., Beck S. (2012) Resources for methylome analysis suitable for gene knockout studies of potential epigenome modifiers. *Gigascience* 1, 3**
13. **Rakyan V.K., Down T.A., Thorne N.P., Flicek P., Kulesha E., Graf S., Tomazou E.M., Backdahl L., Johnson N., Herberth M. et al. (2008) An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res* 18, 1518-1529**
14. **Langmead B., Trapnell C., Pop M., Salzberg S.L. (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10, R25**

Table legends

Table 1 Details of the oligos used in the protocol. CpG% is defined as 2×number CpGs / fragment length [bp]. For fragments derived from the fragmented sample DNA, CpG% was estimated by simulation. Enterobacteria phage lambda genomic Start and Stop coordinates were derived from accession number NC_001416; human genomic start and stop coordinates were derived from reference assembly GRCh37.p2; mouse genomic start and stop coordinates were derived from reference assembly MGSCv37. [*] = phosphothiolate; [Phos] = 3'-phosphate. †Adapters and amPCR primers for single-plex library prep; ‡Adapters and amPCR primers for multiplex library prep.

Table 2 Troubleshooting guide.

Figure legends

Figure 1 The MeDIP-seq workflow. MeDIP-seq libraries can be created in three days. The time required for the actual sequencing depends on the model of the sequencer (at the time of writing, an Illumina HiSeq 2500 takes ~5.5 days for a 'high output' paired-end run of 50 bases). Bioinformatic processing takes approximately 10 hours.

Figure 2 Reaction mixes for a single AutoMeDIP run. Reagents are mixed and dispensed into 12-strip tubes prior to starting an AutoMeDIP run, at which point the IP-Star performs MeDIP sequentially across the wells.

Figure 3 Layout of the IP-Star bed. The red-boxed zone illustrates the location of tip-holders; the black-boxed zone shows the two 96-well Peltier blocks. The arrows demonstrate the direction of workflow of the IP-Star. Twelve consecutive wells are called 'lanes'; when <=8 IPs are being performed, only the right Peltier block (lanes 1-8) is used; when >8 IPs are being performed, both blocks are used. It is important to match tips in lanes of tip racks with samples in lanes of Peltier blocks.

Tables

Table 1

Oligo name	Genome	Sequence (5'→3')	Length (bp)	CpG (%)	amplicon size (bp)	Start	End
Adapter PE 1.0†	N/A	[Phos]GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG	33	N/A	N/A	N/A	N/A
Adapter PE 2.0†	N/A	ACACTCTTTCCCTACACGACGCTCTTCCGATC*T	32	N/A	N/A	N/A	N/A
Adapter InPE 1.0‡	N/A	[Phos]GATCGGAAGAGCACACGTCT	20	N/A	N/A	N/A	N/A
Adapter InPE 2.0‡	N/A	ACACTCTTTCCCTACACGACGCTCTTCCGATC*T	33	N/A	N/A	N/A	N/A
PCR.primer.PE.1.0 (F)†	N/A	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T	58	N/A	N/A	N/A	N/A
PCR.primer.PE.2.0 (R)†	N/A	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATC*T	61	N/A	N/A	N/A	N/A
PCR Primer InPE 1.0 (F)‡	N/A	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATC*T	58	N/A	N/A	N/A	N/A
PCR Primer InPE 2.1 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.2 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.3 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATGCCTAAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.4 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATTGGTCAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.5 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATCACTGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.6 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATATTGGCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.7 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATGATCTGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.8 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATTCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.9 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.10 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATAAGCTAGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.11 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATGTAGCCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
PCR Primer InPE 2.12 (R)‡	N/A	CAAGCAGAAGACGGCATACGAGATTACAAGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC*T	64	N/A	N/A	N/A	N/A
1CpG_qPCR_F	λ-DNA	ACAAGTTGTTTGATCTTTGC	20	N/A	145	24004	24148
1CpG_qPCR_R		CCTATGAGCAACGTGTTAG	19				
5CpG_qPCR_F		CACTTGAATCTGTGGTTCAT	20	N/A	130	25942	26071
5CpG_qPCR_R		TAGAAAAAGACAACTCTGGC	20				
10CpG_qPCR_F		GAACTCACACACAACACCA	19	N/A	125	33662	33786
10CpG_qPCR_R		ACTCTGAATACCGACTCAAT	20				
15CpG_qPCR_F		TATCACTGTTGATTCTCGC	19	N/A	121	28669	28789
15CpG_qPCR_R		GGTAAAGAGTTTGGATTAGG	20				
20CpG_L_qPCR_F		GGTGAACTTCCGATAGTG	18	N/A	108	24202	24309
20CpG_L_qPCR_R		CAGTCATAGATGGTCGGT	18				

20CpG_S_qPCR_F		GTTAGAGCCTGCATAACG	18	N/A	109	44747	44855
20CpG_S_qPCR_R		GAAAGAGCACTGGCTAAC	18				
1CpG_F		GAGGTGATAAAATTAACTGC	20	0.5	197	23967	24163
1CpG_R		GGCTCTACCATATCTCCTA	19				
5CpG_F		CATGTCCAGAGCTCATTC	18	4.0	270	25869	26138
5CpG_R		GTTTAAAATCACTAGGCGA	19				
10CpG_F		CTGACCATTTCCATCATTC	19	5.6	360	33638	33997
10CpG_R		GTAACTAAACAGGAGCCG	18				
15CpG_F		ATGTATCCATTGAGCATTGCC	21	6.4	462	28488	28949
15CpG_R		CACGAATCAGCGGTAAAGGT	20				
20CpG_L_F		GAGATATGGTAGAGCCGCAGA	21	8.0	275	24148	24643
20CpG_L_R		TTTCAGCAGCTACAGTCAGAATTT	24				
20CpG_S_F		CGATGGGTTAATTCGCTCGTTGTGG	25	14.6	403	44589	44863
20CpG_S_R		GCACAACGGAAAGAGCACTG	20				
MM_meth_qPCR_F	Mouse	CATGGCCCACAAAGTAATAAAA	22	16.0	93	Chr1:19718938	Chr1:19719030
MM_meth_qPCR_R		AACGACTTACAACGAGCTCAAA	22				
MM_unme_qPCR_F		GGCTAGAACTGACCAGACAGAC	22	16.0	86	Chr4:16751777	Chr4:16751862
MM_unme_qPCR_R		ATCTGTAGCCAATCCTAGAGCA	22				
HS_meth_qPCR_F	Human	GGGAATATAAGGAGCGCACA	18	10.0	114	Chr22:29607176	Chr22:29607289
HS_meth_qPCR_R		TCGGTTAAAACGGTCAGGTC	18				
HS_unme_qPCR_F		CGAGGCGTGAGTTATTCCTG	20	10.0	110	Ch6:4011057	Ch6:4011166
HS_unme_qPCR_R		CTCTTGTGGCTGAGCTCCTT	20				

Table 2

Step	Problem	Possible reason	Solution
31	No recovery of MeDIP fraction	IP incubation time too short	Increase IP incubation time
		Ab AND/OR beads	Check all reagents have been correctly dispensed
		Poor construction of control cocktail	Check control cocktail sequences are methylated using methylation specific enzyme
			Check control cocktail sequences amplify independently of MeDIP-seq experiment
			Try using species-specific qPCR primers to validate recovery of MeDIP (expect 10-20% recovery)
		None of the above?	Repeat MeDIP on adapter-ligated DNA
	Low (< 10%) recovery of MeDIP fraction and low (< 90%) specificity	IP incubation time too short	Increase IP incubation time
		Aspiration of bead complex following IP (manual only)	If beads remain in tips between MeDIP and purification, pipette mix the solution until no beads remain in the tip
		None of the above?	Repeat MeDIP on adapter ligated DNA
	Low (< 10%) recovery but high (> 95%) specificity	IP incubation time too short but MeDIP should not be considered failed	Proceed to amPCR (step 41) and evaluate
	Ct values for both QC regions in MeDIP fraction are similar	IP Incubation temperature too high	Reduce IP incubation temperature to 4°C
		Incorrect PCR product has been amplified	Check qPCR melt curve
37	No or very low post-amPCR yield	Poor or very inefficient adapter ligation	Check that peak size (assessed using Bioanalyzer) of fragments increases between sonication and adapter ligation
			dilute post-adapter ligated samples 1 in 10 and perform a single amPCR reaction; should see a smear on a gel
			Ensure samples have been purified prior to amPCR
47	Little or no enrichment of methylated regions in MeDIP fraction	PCR failure	Check gel for DNA smear
		Gel purification failure	Ensure adequate volumes of Buffer QG are used
			Ensure pH of QG & sample is at correct pH (< 7.5); if not add 10 µl 3 M NaAc (pH 5.0)
		Inaccurate DNA quantification	Use Agilent High Sensitivity DNA chips. Avoid spectrophotometry
		MeDIP and input sample concentrations not normalized	Normalize MeDIP and Input sample concentrations to > 0.1 ng/µl
		Poor starting DNA quality	Ensure approximately 25-30% of starting DNA falls into the desired excision range
56	Low proportion of aligned reads	Low quality base score at 3' end of read	Trim bases from the sequence prior to alignment