Whole-genome bisulfite sequencing for the analysis of genome-wide DNA methylation and hydroxymethylation patterns at single-nucleotide resolution

Magali Kernaleguen, Christian Daviaud, Yimin Shen, Eric Bonnet, Victor Renault, Jean-François Deleuze, Florence Mauger, Jörg Tost

▶ To cite this version:

HAL Id: cea-04407166
https://cea.hal.science/cea-04407166v1

Submitted on 24 Jan 2024

Whole genome bisulfite sequencing for the analysis of genome-wide DNA methylation and hydroxymethylation patterns at single-nucleotide resolution

Magali Kernaleguen[1*], Christian Daviaud[1*], Yimin Shen[1], Eric Bonnet[2], Victor Renault[3], Jean-François Deleuze[1,3], Florence Mauger[1] and Jörg Tost[1]

[1] Laboratory for Epigenetics and Environment, Centre National de Recherche en Génomique Humaine, CEA-Institut de Biologie Francois Jacob, Evry, France

[2]Laboratory for Bioinformatics, Centre National de Recherche en Génomique Humaine, CEA-Institut de Biologie Francois Jacob, Evry, France

[3] Laboratory for Bioinformatics, Fondation Jean Dausset – CEPH, Paris, France

[*] These authors have contributed equally and should be considered as joint first authors

Corresponding Author: Jörg Tost, Centre National de Recherche en Génomique Humaine, CEA-Institut de Biologie François Jacob, Bâtiment G2, 2 rue Gaston Crémieux, CP 5721, 91057 Evry Cedex, France. Tel.: +33-(0)1 6087 8423. Fax: +33-(0)1 6087 848. E-mail: tost@cng.fr

Running Title: Methyl-seq and OxBs-Seq

Abstract

The analysis of genome-wide epigenomic alterations including DNA methylation and hydroxymethylation has become a subject of intensive research for many biological and disease-associated investigations. Whole-genome bisulfite sequencing (WGBS) using next generation sequencing technologies is currently considered as the gold standard for a comprehensive and quantitative analysis of DNA methylation throughout the genome. However, bisulfite conversion does not allow distinguishing between cytosine methylation and hydroxymethylation requiring an additional chemical or enzymatic step to identify hydroxymethylated cytosines. Here we provide two detailed protocols based on commercial kits for the preparation of sequencing libraries for the comprehensive whole genome analysis of DNA methylation and/or hydroxymethylation. If only DNA methylation is of interest, sequencing libraries can be constructed from limited amounts of input DNA by ligation of methylated adaptors to the fragmented DNA prior to bisulfite conversion. For samples with significant levels of hydroxymethylation such as stem cells or brain tissue, we describe the protocol of oxidative bisulfite sequencing (OxBs-Seq), which in its current version uses a post-bisulfite adaptor tagging (PBAT) approach. Two methylomes need to be generated: a classic methylome following bisulfite conversion and analyzing both methylated and hydroxymethylated cytosines and a methylome analyzing only methylated cytosines, respectively. We also provide a step-by-step description of the data analysis using publicly available bioinformatic tools. The described protocols have been successfully applied to different human samples and yield robust and reproducible results.

1   Introduction

Epigenetic modifications add additional layers of information on top of the bare genomic sequence thereby dramatically extending the information-encoding potential of the DNA. The analysis of epigenetic modifications has gained great momentum in recent years as it has become clear that epigenetics plays a key role in normal development as well as in disease. The technical revolution of massively parallel sequencing, which allows the interrogation of multiple epigenetic modifications using the same analytical technologies, as well as the development of advanced high-throughput epigenotyping technologies has enormously spurred our knowledge on gene regulatory mechanisms. DNA methylation occurring at CpG dinucleotides is probably the best-studied epigenetic modification due to the extensive mapping of DNA methylation patterns in different diseases. DNA methylation based biomarkers bear the promise to provide valuable information for early diagnosis, prognosis, disease classification and might assist in the prediction of response to therapy [1]. While initially epigenetic research has focused on the analysis of epigenetic alterations in cancer [2], research in recent years has demonstrated that alterations are also present in nearly all complex diseases including autoimmune and inflammatory diseases, allergic disorders, metabolic as well as neurodegenerative and psychological disorders [3-9].

Whole-genome bisulfite sequencing (WGBS) or MethylC-seq can be considered as the current gold standard for the genome-wide identification of differentially methylated CpGs (DMCs) and differentially methylated regions (DMRs) at single nucleotide resolution.

The whole genome read-out after bisulfite conversion overcomes the limitations of cloning and Sanger sequencing, limited to a small number of loci of interest, in which the quantitative resolution was limited by the number of clones analyzed (in most studies < 20), or pyrosequencing. However, the unprecedented quantitative and spatial resolution that is currently transforming DNA methylation analysis still comes at a high cost requiring substantial sequencing to obtain a proper and even coverage and specialized bioinformatic expertise and resources. The most widely used WGBS protocol consists of the fragmentation of the genomic DNA, adapter ligation, bisulfite conversion and limited amplification using adapter-specific PCR primers. While initially several micrograms of DNA were required to perform WGBS, the replacement of electrophoretic steps and gel extraction by magnetic beads has enabled to create libraries suitable for sequencing from ~ 100 ng of input material [10]. Libraries have been reported to be constructed from even lower amounts of input material, but require in most cases a high number of PCR cycles (up to 25 cycles [11]) inducing potentially a large bias for the estimation of the DNA methylation levels and a much more substantial sequencing effort to obtain a homogeneous and sufficient coverage.

To assess the bisulfite conversion efficiency, DNA of the bacteriophage λ is spiked in the reaction. Mapping the reads against the bisulfite-converted genome of the phage and counting any remaining cytosines allows estimation of the conversion rates and identification of problems during bisulfite conversion. A large number of programs have been developed to perform the quality control, pre-processing steps (such as adaptor, barcode and quality score trimming), mapping of the reads to a bisulfite converted reference genome and scoring of DNA methylation levels (count statistics) and identification of DMCs and DMRs [12]. WGBS has been widely used for the methylome-wide analysis of a large number of organisms and plants as well as human tissues and ~90-95 % of the cytosines present in the genome are generally covered [13-19].

4

Commonly used alternative approaches to the described MethylC-seq WGBS approach are tagmentation-based WGBS [20] and post bisulfite adaptor tagging (PBAT)- based WGBS [21].

The discovery that 5-methylcytosines can be oxidized to 5-hydroxymethylation (5hmC) has attracted substantial interest in the role of 5hmC in gene regulation [22,23]. Oxidation is catalyzed by the Ten-eleven translocation methylcytosine dioxygenase (TET) family of enzymes, consisting of three mammalian subtypes, TET1-3 [23-25]. 5hmC is an intermediate in an active demethylation process [26], where it is further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) again by the TET enzymes. The latter two modifications are present at barely detectable levels in the human genome [27]. However, 5hmC is now considered to be not only an intermediate in oxidative DNA demethylation, but also to constitute a distinct layer in the complex process of epigenetic regulation with its own distribution and regulatory functions. 5hmC is most abundant in human brain tissue and embryonic stem cells, but at levels approximately 10-fold lower than those of 5-methylcytosine [28]. TET enzymes are expressed in a tissue/cell-type and developmental stage dependent manner with 5hmC decreasing during cell differentiation [29]. 5hmC levels do not correlate with 5mC levels of the respective tissue and 5hmC was found enriched at specific active functional elements of the genome, in particular enhancers, promoters and gene bodies associating 5hmC with open chromatin and transcriptional activity [28,30-32]. 5hmC levels are globally reduced in cancer and alterations of the TET enzymes have been reported for various cancers [24,33-36]. This observation suggests that 5hmC alterations may have a distinct role in the development and progression of malignancies.

The discovery of 5hmC has made the understanding of gene regulation through DNA methylation more difficult, especially as the most widely used method for analysis of DNA methylation, bisulfite conversion, cannot distinguish between 5mC and 5hmC. The DNA

methylation degree measured following a conventional bisulfite conversion is a mixture of both marks and specialized approaches such as OxBS-seq, TAB-seq or chemical labeling are required to distinguish methylcytosine from its oxidative derivative [37-40].

In the present chapter, we describe two protocols used in routine in our laboratory, one classical WGBS protocol and one OxBs protocol. Both protocols allow the comprehensive analysis of the methylome and hydroxymethylome of biological samples. While the OxBs protocol can be probably further improved, it is currently the only protocol allowing to analyze hydroxymethylation in samples with limited amounts of DNA available.

1.1 Summary of the WGBS protocol

The WGBS protocol using conventional bisulfite sequencing is based on the commercial Ovation® UltraLow Methyl-seq DR Multiplex System (NuGen, San Carlos, CA) and has been updated and modified from a previous publication [41]. The workflow of the library preparation protocol is depicted in Figure 1 and follows the classical library preparation protocol, in which methylated adaptors are ligated to the fragmented DNA prior to bisulfite conversion. Fifty to 100 ng of genomic DNA is fragmented to a size of approximately 200 base pairs (bp), purified and methylated adaptors compatible with sequencing on an Illumina HiSeq instrument are ligated. The resulting DNA library is purified and bisulfite converted. A qPCR assay determines the optimal number of PCR amplification cycles (between 10 to 15 cycles) required to obtain a high diversity library with minimal duplicate reads prior to the actual library amplification. Due to the specific proprietary structure and sequence of the adaptors employed in the protocol, a dedicated sequencing primer has to be added during cluster generation on the cBOT of Illumina sequencers, which will be used as the sequencing primer for read 1. The index read as well as the second read in pair-end modus will be using standard Illumina primers.

Following sequencing using 2 x 101 or 150 base pairs reads, the data is quality controlled

and mapped to the genome to obtain the percentage of DNA methylation at all cytosines

using Bismark [42]. With the protocol described in this chapter, we obtain per Hiseq 4000

lane routinely 250 M reads, of which 75-80 % can be mapped to a bisulfite treated human

genome resulting in a genome-wide coverage of 15-20 x. Duplicate rates are ~ 12 %

(Picard) and a conversion efficiency of ~ 99 % is routinely achieved as measured by using

either unmethylated lambda phage DNA or methylation outside CpG context in the

human genome. Coverage was found to be comparable to a standard bisulfite sequencing

protocol starting with 4 µg of DNA and quantitative values correlated well with data

obtained from Illumina Infinium Human Methylation 450K BeadArrays (Renault et al.,

submitted; **Figure 2**). The presented protocol has been found to be very robust and

applicable without modifications to DNA extracted from various human, mammalian and

plant tissues.

1.2 Summary of the OxBS-seq protocol

The second approach analyzing both DNA methylation and hydroxymethylation is based

on the commercial TrueMethyl® WholeGenome (Version 3.1) kit (CEGX, Cambridge, UK).

The workflow of the library preparation protocol is depicted in Figure 3 and follows a

PBAT protocol for library preparation, in which adaptors are ligated after bisulfite

conversion. 400 ng of genomic DNA is best used as input material, but libraries can be

constructed from as little as 100 ng, although duplicate rates will substantially increase. In

a first step sequencing and digestion spike-in controls are added to the sample that allow

for quality monitoring of both the sequencing as well as the conversion efficiency. The

digestion spike-in control allows assessing the efficiency of the oxidation and bisulfite

conversion prior to sequencing using an amplification of a control template followed by a

restriction digestion with the endocnuclease TaqI. The enzyme will only cleave the

template, which underwent bisulfite conversion, while the oxidation leads to a sequence

change of the recognition sequence (TC^{5hm}AG→ TTAG) abolishing the digestion. The sequencing spike-in controls provide a more detailed picture and are designed to assess the performance of oxidation and bisulfite conversion using the read1 of the sequencing results. After addition of the standards, the sample is split into two parts, and one part is oxidized converting hydroxymethylated cytosines to formylcytosines while the other part undergoes a mock reaction in the absence of oxidating reagents. Both aliquots are subsequently bisulfite-converted. In the conventional bisulfite reaction, the remaining cytosine signal reflects a mixture of both methylated and hydroxymethylated cytosines, with the first exceeding largely in most tissues. By comparing the conventional bisulfite-converted sample with the sample that has undergone the additional oxidation step, 5mC and 5hmC can be distinguished (see Table below) as the oxidation-induced 5fCs are deaminated to uracils in the bisulfite reaction, and only original 5mC bases are read as C in the final sequencing. Hydroxymethylated cytosines are thus indirectly determined as the difference in methylation between the two reactions.

Base	Sequence	Bisulfite conversion	Oxidative bisulfite conversion
C	C	T	T
5mC	C	C	C
5hmC	C	C	T
5fC	C	T	T
5caC	C	T	T

Libraries are constructed by ligating a first adapter (end activation step), second strand synthesis (strand synthesis step) followed by ligation of the second adaptor (library finishing step) and introduction of an index in the subsequent amplification step allowing for multiplexed sequencing (library indexing). Between the library finishing and library

8

indexing steps, a qPCR assay determines the optimal number of PCR amplification cycles (between 10 to 14 cycles) required to obtain sufficient material to sequence a high diversity library with minimal duplicate reads. Of note, oxBS samples are usually less concentrated compared to BS samples requiring more amplification cycles. Therefore the use of a standard number of amplification cycles as presented in the original protocol is not recommended. After quality control of the library through the digestion control and shallow sequencing of multiple libraries in parallel on a MiSeq instrument, sequencing is performed using conventional Illumina sequencing primers on a HiSeq 4000 or a HiSeq X5 in a multiplex of four to eight samples per lane on multiple lanes to obtain sufficient coverage. Only libraries with a similar size distribution should be sequenced as a multiplex to ensure a similar number of reads for each sample on the flowcell. Data is quality controlled using BSExpress (CEGX) and mapped to the corresponding genome to obtain the percentage of DNA methylation and cytosine methylation at all cytosines using Bismark [42]. The levels of hydroxymethylated cytosines are obtained by subtracting methylated cytosines from the overall DNA methylation levels determined without oxidation for each CpG. With the protocol described in this chapter, we obtain per samples sequenced on e.g. 1/4 of 4 HiSeq lanes ~ 400 M reads, of which 60-70 % can be mapped to a bisulfite treated human genome. Rate of duplicate reads are ~ 25 % resulting in a genome-wide coverage of 7-15 x with generally lower coverage for the oxidized sample.

2   Material

2.1 Reagents and consumables for Ovation Methyl-seq workflow

1. 50-100 ng genomic DNA (*see* **Note 1**).

2. Ovation Ultralow Methyl-Seq DR Multiplex System 1-8 32rx (NuGEN, cat#0335-32) containing: End Repair Buffer Mix (ER1 ver 8), End Repair Enzyme Mix (ER2 ver 4), End Repair Enhancer (ER3 ver 2), Ligation Buffer Mix (L1 ver 4), Ligation Adaptor Mix (L2V11DR-BC1) to Ligation Adaptor Mix (L2V11DR-BC8), Ligation Enzyme Mix (L3 ver 4), Final Repair Buffer Mix (FR1 ver 4), Final Repair Enzyme (FR2), Amplification Primer Mix (P2 ver 8), Amplification Enzyme Mix (P3 ver 3), 25 µM Sequencing Primer (MetSeq Primer 1), nuclease-free water (D1), and Agencourt RNAClean XP Beads.

Ligation Adaptor Mix	6 nt Barcode Sequence as Read by the Sequencer	Barcode Pairing (2-plex)	Barcode Pairing (> 2-plex)
L2V11DR-BC1	AACCAG	Duplex Set 1	One of the duplex sets from the column to the left should be used in combination with any of the other remaining six individual barcodes.
L2V11DR-BC2	TGGTGA		
L2V11DR-BC3	AGTGAG	Duplex Set 2	
L2V11DR-BC4	GCACTA		
L2V11DR-BC5	ACCTCA	Duplex Set 3	
L2V11DR-BC6	GTGCTT		
L2V11DR-BC7	AAGCCT	Duplex Set 4	
L2V11DR-BC8	GTCGTA		

3. Unmethylated cl857 Sam7 Lambda DNA (*e.g.* Promega).

4. EpiTect FAST DNA Bisulfite kit 50rx (Qiagen) containing: Bisulfite Solution, DNA Protect Buffer, RNase-Free Water, MinElute® DNA Spin Columns, collection Tubes (2 mL), BL Buffer, BW Buffer, BD Buffer, EB Buffer, and Carrier RNA.

5. DNA 1000 kit (Agilent).

6. 100 % and 70 % Ethanol.

7. 1X TE Buffer: 10 mM Tris-HCL, 1 mM EDTA, pH 8.0.

8. EvaGreen® 20X.

9. Covaris tube: 6x16 mm Round bottom glass, AFA fiber and cap system (100 µL) (*e.g.* Kbioscience).

10. 0.2 mL reaction tube « PCR clean » with lid.

11. 1.5 mL reaction tube « PCR clean » with lid.

12. LightCycler$^{®}$ 480 Multiwell Plate 384.

13. LightCycler $^{®}$480 Sealing Foil.

14. KAPA Library Quantification Kit (Roche) containing : KAPA SyBR qPCR Master Mix, standards.

2.2 Equipment for Ovation Methyl-seq workflow

1. DiaMag02 magnetic rack (Diagenode) or similar.

2. ThermoCycler ProS (Eppendorf) or similar.

3. LightCycler480$^{®}$ (Roche) or a similar qPCR instrument.

4. Qiacube (Qiagen, optional).

5. Ultrasonicator E210 (Covaris).

6. Bioanalyser 2100 instrument (Agilent).

7. Vortex.

8. Centrifuge.

9. Laminar Flow Cabinets (optional, *see* **Note 2**).

2.3 Reagents for TrueMethyl® WholeGenome workflow

1. TrueMethyl® WholeGenome (Version 3.1) kit from CEGX including ultra-pure water, Denaturing Solution, Stop Solution, Bisulfite Reagent, Diluent and Elution buffer (store at room temperature); Binding buffer 1, 2, 3, Magnetic Bead Solution, Desulfonation Buffer all to be stored at 4°C; Sequencing and Digestion Spike-in Control, Cutting Control, PCR primers, Oxidant Solution, Adapter and Adapter Additive, Enzymes and Buffers for end activation, strand synthesis, library finishing and indexing and 24 index primers to be stored at -20°C.

2. Quant-it HS kit (Thermo Fisher Scientific).

3. dsDNA HS Qubit® Kit (Thermo Fisher Scientific).

4. Qubit® tubes (Thermo Fisher Scientific).

5. 100% Ethanol.

6. 100% Acetonitrile.

7. EvaGreen® 20X or similar.

8. High Sensitivity DNA kit and HS DNA chip (Agilent).

9. KAPA Library Quantification Kit (ROCHE) containing: KAPA SyBR qPCR Master Mix, standards.

10. 2X Phusion U Hot start Master Mix (Thermo Fisher Scientific).

11. QIAquick PCR purification kit (Qiagen) or similar.

12. TaqI restriction enzyme (New England Biolabs or similar).

13. Agarose.

14. TBE Buffer 10X.

15. SyBRsafe (Thermo Fisher Scientific) or similar.

16. Agarose Gel loading buffer and size ladder 100 base pairs.

17. PhiX (Illumina).

18. Elution Buffer (Qiagen) or similar.

19. Tween20.

20. 0.2 mL, 1.5 mL, 2 mL and 5 mL microtubes (PCR clean and low-bind).

21. 15 mL and 50 mL tubes.

22. ABgene 96 –well plate (Thermo Fisher Scientific) or similar.

23. Hard-Shell® PCR plate 96 (BioRad) or similar.

24. Adhesif PCR film.

25. LightCycler®480 Multiwell Plate 384 and LightCycler®480 Sealing Foil.

2.4 Equipment for TrueMethyl® WholeGenome workflow

1. A pre-PCR room with Laminar Flow Cabinets (*see* **Note 2**) and a post-PCR room.

2. Pipettes and filter tips in pre-PCR room (*see* **Note 2**).

3. Vortexer.

4. Centrifuge for 2 mL microtubes and adapter for microtubes.

5. Magnetic separation rack for 0.2 mL microtubes.

6. Refrigerant block for 96 plate and ice bucket.

7. Laboratory timer.

8. Thermocycler with heated lid.

9. Thermomixer.

10. Molecular Devices SpectraMAX Gemini XPS and Molecular Devices SpectraMax M3.

11. Qubit® Fluorometer (Thermo Fisher Scientific).

12. LightCycler480® thermocycler (Roche Applied Science).

13. BioAnalyser 2100 instrument (Agilent).

14. Miseq and HiSeq 2000, 2500, 4000 or X series sequencers (Illumina).

2.5   Data analysis

1. One linux / unix multi-core machine with at least 16GB of RAM and 5 cores (see Bismark user guide referenced in http://www.bioinformatics.babraham.ac.uk/projects/download.html#bismark).

2. FastQC: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

3. Cutadapt for read trimming and Illumina adapter removal: https://cutadapt.readthedocs.org/en/stable/. Requires python and module python-pip for installation.

4. Bismark: http://www.bioinformatics.babraham.ac.uk/projects/bismark/ [42].

5. Bowtie2: http://bowtie-bio.sourceforge.net/bowtie2/index.shtml.

6. Samtools: http://samtools.sourceforge.net/ [43].

7. Trim Galore: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

8. BsExpress: https://bitbucket.org/cegx-bfx/cegx_bsexpress.

3 Method

3.1 Ovation MethylC-seq workflow

3.1.1 DNA fragmentation

1. Dilute 50 to 100 ng of genomic DNA into 50 µL of 1X TE buffer.

2. Add 0.25% w/w of unmethylated Lambda DNA (e.g. 0.25 ng of Lambda DNA for 100

   ng DNA)

3. Transfer the sample to a Covaris tube.

4. Fragment using the settings of the Covaris focused ultrasonicator listed below for a

   200 bp target length.

Parameters	
DNA Volume	50 µL
Duty Cycle	10%
Intensity	5.0
Cycles per Burst	200
Duration	60 seconds (x2) (Spin between the two runs of fragmentation)
Power	36W
Temperature	8°C

5. After fragmentation, briefly spin the Covaris tube and transfer the entire sample to

   a new 0.2 mL tube.

3.1.2 DNA purification

1. Thirty min before proceeding take the Agencourt RNA Clean XP beads and

   nuclease-free water out of the fridge and leave to warm up at room temperature.

2. Resuspend the RNAClean XP beads by vortexing the tube. Ensure beads are fully

   resuspended before adding to sample.

3. Prepare a fresh 70% ethanol wash solution (1.5 mL per sample).

4. Add 90 µL (1.8 volumes) of the bead suspension to the 50 µL of fragmented DNA and mix by pipetting 10 times.

5. Incubate at room temperature for 10 min.

6. Place the tube into the magnetic rack and incubate for 5 min at room temperature to completely clear the solution.

7. Carefully remove 130 µL of the binding buffer and discard it (*see* **Note 3**).

8. Add 200 µL of freshly prepared 70% ethanol, wash and leave to stand for 30 s.

9. Remove and discard the 70% ethanol using a pipette.

10. Repeat steps 8 and 9 one more time, for a total of two 70 % ethanol washes (*see* **Note 4**).

11. Air-dry the bead pellets on the magnetic rack for 5–10 min at room temperature. Inspect each tube carefully to ensure that all the ethanol has evaporated.

12. Add 14 µL room temperature nuclease-free water (provided in the kit) to the dried beads. Mix thoroughly to ensure all the beads are resuspended.

13. Place the tube into the magnetic separation rack and incubate for 5 min at room temperature to completely clear the solution.

14. Carefully transfer 13 µL of the eluate into a new 0.2 mL tube ensuring to carry over as few beads as possible and place on ice.

The following paragraphs (3.1.3-3.1.5) describe details of the different steps required for the preparation of the sequencing library prior to bisulfite conversion including repair (fill-up) of 5' and 3' extremities, ligation of indexed adaptors and a final end repair step. Figure 4 provides a schematic overview of the different steps.

3.1.3   End Repair

1.   Thaw the End Repair Buffer (ER1) of the Ovation Ultralow Methyl-Seq DR Multiplex System.

2.   Prepare the End-Repair master mix by combining in a 0.2 mL tube 2 µL of End Repair Buffer (ER1), 0.5 µL of End Repair Enzyme mix (ER2 ver 4) and 0.5 µL of End Repair Enhancer (ER3 ver 2).

3.   Add 3 µL of the End Repair master mix to the fragmented DNA sample (from section 3.1.2). Mix by pipetting, cap and briefly spin down the tubes and place on ice.

4.   Incubate the sample in a thermal cycler at 25°C for 30 min followed by an incubation at 70°C for 10 min and cool down to 4°C.

5.   Remove the tubes from the thermal cycler, briefly centrifuge and place on ice.

3.1.4   Ligation

1.   Choose and thaw one adaptor per sample from the Ligation Adaptor Mixes (L2V11DR-BC1 through L2V11DR-BC8). For multiplex sequencing refer to the table below for possible combinations as shown in 2.1.2. Thaw the Ligation Buffer Mix (L1 ver4) on ice. Thaw the nuclease-free water (D1) at room temperature.

2.   Add 3 µL of L2 Ligation Adaptor Mix to each sample (from step 5 of section 3.1.3). Mix thoroughly by pipetting.

3.   Prepare a Ligation master mix by combining in a 0.2 mL tube 4.5 µL of nuclease-free water (D1), 6 µL of Ligation Buffer Mix and 1.5 µL of Ligation Enzyme Mix. Mix by pipetting slowly, without introducing bubbles, briefly spin down and place on ice (*see* **Note 5**).

4.   Add 12 µL of Ligation Master Mix to each reaction tube. Mix thoroughly by pipetting slowly, spin and place on ice. Proceed immediately to the next step.

5.   Incubate the sample in a thermal cycler at 25°C for 30 min followed by an incubation at 70°C for 10 min and cool down to 4°C.

6.   Remove the tubes from the thermal cycler, briefly spin down and place on ice.

7.   Purify the ligation products as described in steps 1 to 11 of section 3.1.2 with the exception of adding 45 µL (1.5 volumes) of the bead suspension to the 30 µL of sample and mix by pipetting 10 times.

8.   Add 16 µL nuclease-free water (green: D1) equilibrated to room temperature to the dried beads. Mix thoroughly to ensure all the beads are resuspended.

9.   Transfer the tube into the magnetic separation racks and incubate for 5 min at room temperature to completely clear the solution.

10.   Carefully transfer 15 µL of the eluate into a new PCR tube ensuring to carry over as few beads as possible and place on ice.

3.1.5   Final Repair

1.   Thaw the Final Repair Buffer Mix (Purple: FR1 VER 4) on ice.

2.   Make a Final Repair master mix by combining 4.5 µL of Final Repair Buffer Mix (FR1) and 0.5 µL of Final Repair Enzyme Mix (FR2) in a 0.2 mL capped tube. Mix by pipetting, spin down briefly and place the tube on ice.

3.   Add 5 µL of Final Repair master mix to each sample (from section 3.1.4).

4.   Mix thoroughly by pipetting slowly, spin and place on ice.

5.   Incubate the sample in a thermal cycler at 60°C for 10 min and cool down to 4°C.

6.   Remove the tubes from the thermal cycler, spin down briefly and place on ice.

3.1.6   Bisulfite conversion (*see* Note 6)

1.   Thaw the DNA to be used in the bisulfite reactions. Make sure the Bisulfite Solution is completely dissolved (*see* Note 7).

2. Add 85 μL of Bisulfite Solution and 35 μL of DNA Protect Buffer to each sample (20 μL from section 3.6). Mix the bisulfite reaction thoroughly and gently spin down. Place the tube at room temperature (15–25°C) *(see* **Note 8**).

3. Place the reaction in a thermal cycler and incubate the sample use the following program:

Temperature	Time
95 °C	5 min
60 °C	20 min
95 °C	5 min
60 °C	40 min
95 °C	5 min
60 °C	45 min
20 °C	HOLD

3.1.7 Cleanup of bisulfite converted DNA

1. Briefly centrifuge the tubes containing the bisulfite reactions, and then transfer the totality of the bisulfite reactions into new 1.5 mL microcentrifuge tubes.

2. Add 310 μL of freshly prepared BL Buffer containing 10 μg/mL carrier RNA to each sample. Mix the solutions by vortexing and then centrifuge briefly.

3. Add 250 μL ethanol (96–100%) to each sample. Mix the solutions by pulse vortexing for 15 s, and centrifuge briefly.

4. Place a MinElute DNA spin column and a collection tube into a suitable rack and transfer the entire mixture from each tube into the corresponding MinElute DNA spin column.

5. Centrifuge the spin column at 15,000 x *g* for 1 min at room temperature. Discard the flow-through, and place the spin column back into the collection tube.

6. Add 500 μL of BW Buffer (wash buffer) to each spin column, and centrifuge at maximum speed for 1 min at room temperature. Discard the flow-through, and place the spin column back into the collection tube.

7.　　　　Add 500 μL of BD Buffer (desulfonation buffer) to each spin column (*see* **Notes 9 and 10**), and incubate for 15 min at room temperature.

8.　　　　Centrifuge the spin columns at 15,000 x *g* for 1 min at room temperature. Discard the flow-through, and place the spin column back into the collection tube.

9.　　　　Add 500 μL of BW Buffer to each spin column and centrifuge at 15,000 x *g* for 1 min at room temperature. Discard the flow-through and place the spin column back into the collection tube.

10.　　　　Add 500 μL of BW Buffer to each spin column and centrifuge at 15,000 x *g* for 1 min at room temperature. Discard the flow-through and place the spin column back into the collection tube.

11.　　　　Add 250 μL of ethanol (96–100%) to each spin column and centrifuge at 15,000 x *g* for 1 min at room temperature.

12.　　　　Place the spin column into new 2 mL collection tube and centrifuge at 15,000 x *g* for 1 min to remove any residual liquid.

13.　　　　Place the spin column into a new 1.5 mL tube. Add 22 μL of the EB Buffer (elution buffer) directly onto the center of each spin-column membrane and close the lid gently.

14.　　　　Incubate the spin column at room temperature for 1 min.

15.　　　　Centrifuge for 1 min at 15,000 x *g* at room temperature to elute the DNA.

16.　　　　Transfer 20 μL of the eluate into a new 0.2mL tube and place on ice (*see* **Note 11**).

3.1.8　Determination of optimal amplification cycles by qPCR

1.　　　　Thaw the Amplification Primer Mix (P2 ver 5) and Amplification Enzyme Mix (P3) on ice.

2.　　　　Transfer 4 μL of converted DNA (from step 16 of 3.1.7) into a LightCycler® 480 Multiwell Plate 384.

3.   Prepare a qPCR master mix by combining 1 µL of Amplification Primer Mix (P2 ver 5), 0.5 µL of EvaGreen® 20X and 4.5 µL of Amplification Enzyme Mix (P3) in a 0.5 mL capped tube. Mix by pipetting, spin down briefly and place on ice.

4.   Add 6 µL of qPCR master mix to 4 µL of the library preparation or a standard (included in the KAPA Library Quantification kits). Seal the LightCycler® Plate 384 with LightCycler® 480 Sealing Foil.

5.   Centrifuge at 4,000 *g* for 1 min at room temperature and place the tube on ice.

6.   Place the reaction into a LightCycler® 480 using the following program:

Temperature	Time	Nb of cycles
95 °C	2 min	1
95 °C	15 s	
60 °C	1 min	35
72 °C	30 s *Optical reading (SyBRgreen channel)*	
40 °C	HOLD	

7.   Determine the optimal number of cycles needed for sufficient amplification without over-amplifying the library (N): plot linear Rn (*see* **Note 12**) versus cycle and determine the cycle number that corresponds to 1/2 of the maximum fluorescent intensity (**Figure 5**).

3.1.9   PCR amplification

1.   Thaw the Amplification Primer Mix (P2 ver 5) and Amplification Enzyme Mix (P3) on ice.

2.   Make a PCR master mix by combining 4 µL of Amplification Primer Mix (P2 ver 5) and 20 µL of Amplification Enzyme Mix (P3) in a 1.5 mL capped tube. Mix by pipetting, spin down tube and place on ice.

3.   Add 24 µL of PCR master mix to each sample (16 µL from section 3.1.7)

4.   Place the reaction in a pre-heated thermal cycler using the following program:

Temperature	Time	Nb of cycles
95 °C	2 min	1
95 °C	15 s	N = as determined in section 3.1.8
60 °C	1 min	
72 °C	30 s	
4 °C	HOLD	

5.   Remove the tube from the thermal cycler, spin down briefly and place on ice.

6.   Purify the amplification products as described in steps 1 to 11 of section 3.1.2 with the exception of adding 50 µL (1 volume) of the bead suspension to the 50 µL of sample and mix by pipetting 10 times.

7.   Add 20 µL of the nuclease-free water (green: D1) equilibrated to room temperature to the dried beads. Mix thoroughly to ensure all the beads are resuspended.

8.   Transfer the tube into the magnetic separation racks and incubate for 5 min at room temperature to completely clear the solution.

9.   Carefully transfer 18 µL of the eluate into a new PCR tube ensuring to carry over as few beads as possible and place on ice.

3.1.10 Quantitative and Qualitative Assessment of the Library

1.   Thirty min before proceeding, take the DNA1000 kit reagents out of the fridge and leave to equilibrate to room temperature.

2.   Run 1 µL of the libraries on a Bioanalyzer DNA1000 Chip.

3.   Profiles should display a peak at approximately 300 bp (270 and 420 bp) as shown in Figure 5 corresponding to 150–200 base pair inserts (*see* **Note 13**).

4. Perform a qPCR quantification of the library preparation using the KAPA Library Quantification Kit. Prepare a solution of Elution Buffer/0.001% Tween20 and two dilutions of each sample at 1/5000 and 1/50000.

5. In a qPCR plate, add 6 µL of KAPA SyBR qPCR Master Mix and 4 µL of the library preparation or standard (included in the KAPA Library Quantification Kit). Each sample should be analyzed in triplicate. Seal the plate and centrifuge it at 4000 x *g* for 1 min at room temperature.

6. Place the plate into a LightCycler® 480 and perform the following program :

Temperature	Time	Nb of cycles
95 °C	5 min	1
95 °C	30 s	
60 °C	45 s *Optical reading* *(SyBRgreen channel)*	40
40 °C	HOLD	

7. Determine the Ct (Cycle Threshold) using the Abs Quant/2nd Derivative Max for each sample. The Ct should be between 8 and 26 cycles. Determine the concentration of each sample using the Ct and the standard curve. The concentration of each sample should be in the range of the standard curve, i.e. between 0.0002 pM and 20 pM. Calculate the library concentration using this formula: Library concentration in nM = [Conc qPCR] x dilution factor x 452 (corresponding to the length of the standard included in the KAPA Library Quantification Kit) /average size of the library preparation (*see* **Note 14).**

3.2 TrueMethyl® WholeGenome workflow

Figure 3 provides an overview of the OxBs-Seq protocol.

Day 1 (*see* **Note 15**)

3.2.1 Digestion and sequencing spike-in controls

1. Perform the steps in a pre-PCR room in Laminar Flow Cabinets and use pipettes with filter tips (*see* **Note 2**).

2. Quantify samples with the Quant-it HS kit according to the manufacturer's instructions.

3. Add 400 ng or less (down to 100 ng) of high molecular weight DNA into a 0.2 mL microtube (*see* **Notes 16** and **17)**. We recommend to process between six and 12 samples maximum in one experiment (*see* **Note 18**).

4. Calculate the volume for the digestion (5 ng/μL) and sequencing (8 ng/μL) spike-in controls to add at 2% w/w in a final volume of 50 μL (*see* **Note 19**) and add the appropriate volume to each sample.

5. Add water to achieve a final volume of 50 μL.

3.2.2 Purification and DNA denaturing

1. Equilibrate Magnetic Bead Solution and Binding Buffer 1 to room temperature for 30 min.

2. Prepare a fresh MBBS1 solution: vortex Magnetic Beads for 1 min. Add 14.4 μL of Magnetic Beads and 720 μL of binding buffer 1 (for 6 reactions) to a new 1.5 mL microtube (*see* **Note 20).**

3. Prepare a fresh 80% acetonitrile wash solution (600 μL per sample).

4. Vortex MBBS1 (*see* **Note 20**) and add 100 μL to each sample. Vortex and briefly spin down. Incubate at room temperature for 20 min.

5. Place samples into a magnetic rack and incubate at room temperature for 5 min until a completely clear solution is obtained. Carefully remove the supernatant and discard it (*see* **Note 3**).

6. Add 200 μL of 80 % acetonitrile. Remove and discard the supernatant.

7. Repeat step 6 one more time.

8. Air-dry bead pellets on the magnetic rack at room temperature for 5–10 min (*see* **Note 4**).

9. Add 20 µL of denaturing solution to the dried beads. Vortex thoroughly to ensure that all beads are resuspended and briefly spin down (*see* **Note 21**).

10. Incubate at 37°C for 5 min. Briefly spin down and place sample back into the magnetic separation rack. Incubate at room temperature for 2 min until the solution is completely clear.

11. Carefully transfer 9 µL of the eluate into a new 0.2 mL microtube for bisulfite conversion (Bs) and 9 µL into a second, new 0.2 mL microtube for oxidative bisulfite conversion (OxBs) (*see* **Note 22**).

3.2.3 Oxidation

1. Thaw the oxidant solution on ice (*see* **Note 23**).

2. Add 1 µL of Ultra-Pure Water to the Bs sample and 1 µL of oxidant solution to the OxBs sample. Vortex and spin down.

3. Incubate at 40°C for 10 min in pre-heated thermal cycler (*see* **Note 24**).

4. During the oxidation reaction, prepare the Bisulfite Reagent Solution by adding 700 µL of Bisulfite Diluent to an aliquot of the Bisulfite Reagent (1 aliquot for 6 reactions). Incubate at 60 °C for 15 min mixing at 1400 rpm. Ensure that the Bisulfite Reagent is completely dissolved. Equilibrate Bisulfite reagent to room temperature (*see* **Note 25**).

5. At the end of oxidation reaction, centrifuge samples for 10 min at 14000 x *g*.

6. Transfer very carefully the supernatant without disturbing the black pellet into a new 0.2 mL microtube (*see* **Note 24**).

Proceed immediately to the next step.

3.2.4 Bisulfite Conversion

1. Add 30 µL of the Bisulfite Reagent Solution to each sample. OxBs samples become transparent (*see* **Note 26**). Vortex and spin down.

2. Place samples in a thermal cycler with the lid heated to 110°C and perform the following program:

Temperature	Time
95 °C	5 min
60 °C	20 min
95 °C	5 min
60 °C	40 min
95 °C	5 min
60 °C	45 min
20 °C	HOLD

Stop point: the bisulfite conversion can be made overnight and samples can be stored at 20°C without any loss of performance (*see* **Note 15**).

3.2.5 Purification and Desulfonation.

1. At the first use of the kit, add 18.75 mL of absolute ethanol to the Desulfonation Buffer Concentrate. Vortex and store until use at 4°C.

2. Equilibrate the Desulfonation Buffer, Magnetic beads and Binding Buffer 2 at room temperature for 30 min.

3. Centrifuge samples at 14000 x *g* for 10 min at room temperature.

4. Prepare fresh MBBS2: vortex magnetic beads for 1 min. Add 28.8 µL of Magnetic Beads to 2400 µL of binding buffer 2 to a 5 mL microtube (for 12 reactions) (*see* **Note 20**).

5. Prepare a fresh 70% Ethanol wash solution (600 µL per sample).

6. Transfer 40 µL of Bs supernatant into a new 0.2 mL microtube and 40 µL of OxBs supernatant without disturbing the white pellet into a new 0.2 mL microtube (*see* **Note 27**).

7. Vortex MBBS2 and add 160 µL to each sample (*see* **Note 20**). Mix by pipetting thoroughly and briefly spin down.

8. Incubate at room temperature for 5 min. Place samples into the magnetic separation rack and incubate at room temperature for 15-30 min until the solution is completely clear.

9. Carefully remove 190 µL of the supernatant and discard (*see* **Note 3**).

10. Add 200 µL of 70% ethanol to the beads. Vortex to resuspend the beads and briefly spin down. Place samples into the magnetic separation rack and incubate at room temperature for 3 min until the solution is completely clear.

11. Carefully remove 190 µL of the supernatant and discard it.

12. Add 200 µL of desulfonation buffer to the beads. Vortex to resuspend beads and briefly spin down. Place samples into the magnetic separation rack and incubate at room temperature for 5 min until completely clear solution. Discard 190 µL supernatant.

13. Repeat the wash step with 70% ethanol (step 10) one more time.

14. Air-dry beads pellets on magnetic rack at room temperature for 5–10 min (*see* **Note 4**).

15. Add 14 µL of ultra-pure Water to the dried beads. Remove samples from the magnetic rack and mix thoroughly with a pipette to ensure all beads are resuspended (*see* **Note 21**). Briefly spin down.

16. Incubate at room temperature for 5 min. Transfer samples back into the magnetic separation rack at room temperature for 2 min until the solution is completely clear.

17.   Carefully transfer 12 μL of the eluate into a new 0.2 mL microtube (*see* **Note 22)**.

      Transfer 11 μL of the 12 μL into a new 0.2 mL microtube and store 1 μL for the

      digestion PCR control (*see* **Note 20**).

Stop point: samples can be stored at 4°C for 24h or frozen at -20°C for long term (*see* **Note 15**).

Day 2 (*see* **Note 15**).

3.2.6   End activation

1.   Thaw Buffer 1 on ice.

2.   Add to each sample, in the following order, 2 μL of buffer 1 and 2 μL of enzyme A

      without prior mixing the two reagents. Mix by pipetting 10 times and spin down.

3.   Incubate samples in a thermal cycler at 37°C for 20 min, followed by heating at

      95°C for 3 min to achieve denaturation.

4.   Immediately after the end of the program, incubate on ice for 5 min.

5.   Thaw Adapter 1 and Additive Adapter 1 on ice.

6.   Add to each sample, in the following order, 2 μL of Adapter 1, 1 μL of Additive

      Adapter 1 and 2 μL of Enzyme B without prior mixing. Mix by pipetting 10 times

      and spin down.

7.   Incubate samples in a thermal cycler at 37°C for 30 min.

8.   Equilibrate Magnetic Bead Solution and Binding Buffer 3 to room temperature for

      30 min.

9.   During the end activation program, prepare fresh MBBS3 (*see* **Note 28**). Vortex

      Magnetic Beads for 1 min. Add 61.2 μL of Magnetic Beads to 3000 μL of binding

      buffer 3 (for 12 reactions) to a 5 mL microtube (*see* **Note 20**).

10.   At the end of the program, add 2 μL of Stop Solution to each sample. Vortex

      MBBS3 and add 66 μL to each sample (*see* **Note 20**). Vortex and briefly spin down.

11. Incubate at room temperature for 15 min.

12. Prepare a fresh 70% ethanol wash solution (1.6 mL per reaction).

13. Place samples into the magnetic separation rack and incubate at room temperature for 5 min until a completely clear solution is obtained. Discard the supernatant (*see* **Note 3**).

14. Repeat the 70% ethanol wash step one more time.

15. Air-dry beads pellets on the magnetic rack at room temperature for 5–10 min (*see* **Note 4**).

16. Add 23.5 µL of ultra-pure water to the dried beads. Remove samples from the magnetic rack, mix thoroughly with a pipette to ensure that all beads are resuspended (*see* **Note 21**) and briefly spin down.

17. Incubate at room temperature for 2 min and place samples into the magnetic separation rack at room temperature for 2 min until a completely clear solution is obtained.

18. Transfer 22 µL of the eluate into a new 0.2mL microtube (*see* Note **22**).

Stop Point: samples can be stored overnight at -20°C (*see* **Note 15**).

3.2.7   Strand synthesis

1. Thaw Buffer 2 on ice and equilibrate MBBS3 to room temperature for 30 min.

2. Add to each sample, in the following order, 6 µL of Buffer 2, 1 µL of enzyme A and 1 µL of enzyme C (Enzymes can be mixed prior to addition). Mix by pipetting 10 times and briefly spin down.

3. Incubate samples in a thermal cycler at 37°C for 30 min.

4.   After the end of the program, add 3 μL of Stop Solution to sample. Vortex MBBS3 containing the magnetic beads for 1 min and add 66 μL to each sample and (*see* **Note 20**). Vortex and briefly spin down.

5.   Purify samples as described in section. 3.2.6 steps 11 to 17 adding 19.5 μL of ultra-pure water to the dried beads for elution.

6.   Transfer 18 μL of the eluate into a new 0.2 mL microtube (see **Note 10**).

Stop point: samples can be stored overnight at -20°C (*see* **Note 15**).

3.2.8   Library Finishing

1.   Thaw Buffer 3 and Adapter 2 on ice and equilibrate MBBS3 to room temperature for 30 min.

2.   Add to each sample, in the following order, 22.5 μL of Buffer 3, 3.5 μL of adapter 2 (Buffer 3 and Adapter 2 can be mixed prior to addition) and 1 μL of enzyme D. Mix by pipetting 10 times and spin down.

3.   Incubate samples in a thermal cycler at 25°C for 15.

4.   After the end of the program, add 5 μL of Stop Solution to each sample. Vortex MBBS3 for 1 min and add 50 μL to each sample (*see* **Note 20**). Vortex and briefly spin down.

5.   Purify samples as described in section 3.2.6 steps 11 to 17 adding 20 μL of ultra-pure water to the dried beads for elution.

6.   Transfer 18.75 μL of the eluate into a new 0.2 mL microtube (*see* Note **22**).

Stop point: samples can be stored overnight at -20°C (*see* **Note 15**).

Day 3 (*see* **Note 15**)

3.2.9 Optimization of PCR cycles (*see* Note 29)

1. Thaw Buffer 4 on ice.

2. Add to each sample in the following order, 5 μL of Buffer 4, 1.25 μL of enzyme E without prior mixing of the two reagents. Mix by pipetting 10 times and spin down.

3. Incubate samples in a thermal cycler at 37°C for 20 min, followed by incubation at 95°C for 3 min.

4. Immediately at the end of the program, incubate for 5 min on ice.

5. Thaw indexes on ice (*see* **Note 30**).

6. For the optimization of the PCR cycles: add 0.2 μL of Enzyme F, 0.5 μL of EvaGreen® 20X, 0.5 μL of Buffer 4 and 2.75 μL of ultra-pure water to a new 1.5 mL microtube. Mix by pipetting 10 times and spin down. Place mix on ice.

7. In a qPCR plate, add 3.95 μL of mix, 4.8 μL of the selected index and 1.25 μL of sample. Seal the plate and centrifuge it at 4000 x *g* for 1 min at room temperature.

8. Place plate into a LightCycler® 480 and perform the following program:

Temperature	Time	Nb of cycles
95 °C	2 min	1
95 °C	30 s	
60 °C	30 s	
	90 s	20
	Optical reading	
72 °C	*(SyBRgreen channel)*	
40 °C	HOLD	

9. Determine the optimal number of cycles needed (N) (*see* **Note 31**): Use N-1 cycle for the amplification of the indexed libraries. The number of cycles should be between 10 and 14.

3.2.10 Library Finishing

1.   Library Indexing: to 23.75 µL of each sample add 22.8 µL of the selected index (*see* **Note 30**) and 0.95 µL of enzyme F. Mix by pipetting 10 times and spin down.

2.   Place samples in a thermal cycler with the lid heated to 110 °C using the following program:

Temperature	Time	Nb. of cycles
95 °C	2 min	1
95 °C	30 s	N as determined in section 3.2.9
60 °C	30 s	
72 °C	90 s	
72 °C	5 min	1
4 °C	HOLD	

12.   Transfer aliquots of stop solution (5 µL per sample), MBBS3 (38 µL per sample) and 70% ethanol (400 µL per sample) to a post-PCR laboratory and equilibrate MBBS3 to room temperature for 30 min.

13.   Vortex MBBS3 for 1 min (*see* **Note 20**) and add 38 µL (0.8 X) to each sample. Vortex and briefly spin down.

14.   Purify samples as described in section 3.2.6 steps 11 to 17 adding 11 µL of ultra-pure water to the dried beads for elution.

15.   Transfer 10 µL of the eluate into a new 0.2 mL micro tube (*see* **Note 22**).

Sample can be stored at -20°C (see **Note 15**).

Day 4 (*see* **Note 15**)

3.2.11 Quality control of the library preparation (*see* **Note 32**)

Figure 6 shows representative results from the quality controls of OxBs-Seq Libraries.

1.   Equilibrate a High Sensitivity DNA kit to room temperature for 30 min.

2.   Transfer 1 µL of the sequencing library into a BioAnalyzer HS DNA Chip.

3. Profiles of library preparation should display a peak at approximately 500 bp (range 200 - 1000 bp) as shown in Figure 6 corresponding to 200–350 base pair inserts. This analysis also allows estimating the quantity of the library.

4. Quantify the sequencing library using qPCR as described in steps 4 to 11 of 3.1.10 (*see* **Note 33**).

5. As an additional quality control, we recommend to analyze the library preparation in a multiplex experiment on a Miseq before sequencing on a HiSeq X instrument. This analysis also allows determining insert size of the library on a sequencer to choose the best combination of samples (*see* **Note 34**).

6. Prepare a 12-plex at 4 nM in EB/0.001% tween20 (20 μL final volume minimum) and add 20% of PhiX to determine the size of the library preparation.

7. Sequence samples using 2*75 bp protocols on a Miseq.

3.2.12 Digestion Control (optional)

The optional digestion control step assesses the efficiency of the oxidation and bisulfite conversion using spike-in digestion controls (*see* **Note 19**).

1. Add 23 μL of ultra-pure Water, 25 μL of Phusion Hot Start Master Mix and 0.5 μL of each PCR digestion control primers in a new 0.2 mL microtube.

2. Add 49 μL of the mix to 1 μL of each sample from section 3.2.5. Vortex and briefly spin down.

3. Place samples in a thermal cycler with the lid heated to 110 °C using this program:

Temperature	Time	Nb. of cycles
95 °C	15 min	1
95 °C	30 s	
60 °C	30 s	40
72 °C	15 s	
72 °C	5 min	1

4 °C	HOLD	

Stop point: reaction can be done overnight and samples stored at 4 °C (*see* **Note 15**).

4. After the end of the program, vortex and spin down samples.

5. Store 1 µL of the PCR product to control the efficiency of the PCR on a 2% agarose gel.

6. Add 245 µL Buffer PB and the remaining 49 µL of each sample to a new Qiaquick column.

7. If the pH is too high (solution orange/purple) add 10 µL of 3 M sodium acetate (solution yellow).

8. Mix by reversing the column and spin at 17000 x *g* for 1 min.

9. Discard eluate and place column in a new collector tube.

10. Add 740 µL buffer PE and spin at 17000 x *g* for 1 min.

11. Discard the eluate and place the column in a new collector tube.

12. Spin at 17000 x *g* for 3 min with the cap open to dry the column.

13. Place the column in a new 1.5 mL PCR tube and add 30 µL of Elution Buffer to the center of the column.

14. Incubate at room temperature for 1 min and spin at 17000 x *g* for 1 min to recover the eluate and discard the column.

15. Equilibrate a dsDNA HS Qubit® kit to room temperature for 30 min.

16. Quantify 1 µL of the purified PCR product with the Qubit® Fluorometer.

17. Prepare these reactions in a 96 plate:

Digestion Control mix	Taq+ Volume / µL	Taq- Volume / µL
purified PCR (10 ng/µL)	10	10
Ultra-pure water	6	8

Cutting Control mix	Taq+ Volume / µL	Taq- Volume / µL
Cutting Control (20 ng/µL)	5	5
Ultra-pure water	11	13

10X Cut Smart Buffer	2	2		10X Cut Smart Buffer	2	2
TaqI RE 20 U/µL	2	0		TaqI RE 20 U/µL	2	0
Total	20	20		Total	20	20

18. Vortex and spin down. Incubate samples in a thermal cycler at 65°C for 18h, followed by heat inactivation at 80°C for 20 min.

Stop point: reaction can be made overnight and stored at 4 °C (*see* **Note 15**).

19. Prepare a 2% agarose gel and add 10 µL of SyBRsafe per 100 mL of gel.

20. Transfer 5 µL of digestion product and 5 µL of loading buffer to the gel.

21. Transfer 1 µL of the digestion control PCR product (from section 3.2.5.) and 9 µL of loading buffer to the gel.

22. Transfer 5 µL DNA size standard (DNA ladder) to the gel.

23. Migrate PCR products for 1h at 120 V.

24. Analyse restriction patterns of samples: Bs samples are digested by TaqI: the corresponding lane on the gel should show one band less than 100 pb. OxBs samples are not digested by TaqI and the corresponding lane on the gel should display one band around 140 pb. Cutting control (positive control) should resemble the BS sample and the reaction without Taq (negative control) the OxBs sample.

3.3 Sequencing

1. Sequencing is performed using standard Illumina sequencing protocols in 2 x 101 or 2 x 150 base pairs following the instructions' provided by Illumina.

2. For sequencing TrueMethyl® WholeGenome libraries on a HiSeq 4000/X instrument, select samples with very similar insert size (*see* **Note 34**). Prepare a 4-plex equimolar pool standardized with 5 % of PhiX at 4.2 nM in EB/0.001 % Tween 20 into 30 µL final volume (*see* **Note 35**).

3. For Ovation MethylC-seq, prepare a multiplex equimolar pool of samples standardized with 5 % of PhiX at 2.6 nM in EB/0.001 % Tween 20 into 30 μL final volume (*see* **Note 35**).

4. The design of the Ovation Ultralow Methyl-Seq DR Multiplex System requires the use of a custom Read 1 sequencing primer, called MetSeq Primer 1, that is included in this kit at a concentration of 100 μM. 2.4 μL of this primer are added manually during the Cluster Generation Process on a cBOT for sequencing on a HiSeq 4000. The standard primers provided in the Illumina sequencing kit are appropriate for Read 2 and for sequencing the DR barcodes.

3.4 Data analysis

The following paragraph is based on the assumption that the requirements defined in the Materials section are met. Each required software should have its associated bin folder in the PATH environment variable (for example for Bismark, export PATH=$PATH:PATH_TO_BISMARK_BIN). As current sequencers tend to generate longer reads (≥ 50 bp), bowtie2 will be used in the protocol below with Bismark (see Bismark user guide at http://www.bioinformatics.babraham.ac.uk/projects/download.html#bismark for more details). If lambda DNA has been spiked into the sample, steps 5 to 9 should be performed successively for both the lambda genome and the sequenced genome (e.g. human, mouse, plant, …). Step 8 could be omitted for the lambda genome to save time and computational resources.

1. Run fastQC on read 1 and 2:

   fastqc FASTQ_READ1

   fastqc FASTQ_READ2.

2. Check that % of the GC content and the estimated % PCR duplicates have consistent values for both reads i.e. %GC should be ~20% and % PCR duplicates < 20%.

3. Read quality trimming and Illumina adapter removal.

   For the **NUGEN** protocol:

   trim_galore --length 36 -q 20 --illumina --paired -a2 AAATCAAAAAAAC FASTQ_READ1 FASTQ_READ2

   --length 36: set minimum length to 36 nt.

   -q 20: set quality threshold to 20.

   --illumina: set adapter type to Illumina.

   -a2 AAATCAAAAAAAC: remove sequence insert specific to the NUGEN protocol on READ 2.

   NB: trimmed reads will be generated with the suffix "val_1.fq" and "val2.fq" respectively (named here TRIMMED_FASTQ_READ1 and TRIMMED_FASTQ_READ2).

   For the **CEGX** protocol, we first perform the quality control of the conversion using CEGX BsExpress, analyzing the tailed CEGX spike-in controls only in Read1:

   bsExpress -i FASTQ_READ1

   -r PATH_TO_BSEXPRESS_CONTROLS/control_reference/oxBS_controls-v1.0.fa

   -p OUT_FOLDER --maxlen 60

   --trim_galore_opts='--quality 20 --three_prime_clip_R1 10' (*see* **Note 36**)

   A successful experiment should show for the spike-in controls (Figure 7):

   1) ≥90% for hmC conversion in the oxBS reaction

   2) <10% for hmC conversion in the BS reaction (hmC/BS over-conversion error rate)

3) <5% for mC conversion in both the oxBS and BS reactions (over-conversion

error rate)

Perform read quality trimming with the trim_galore command and then cut out

10nt with a specific cutadapt command:

trim_galore --length 36 -q 20 --illumina --paired FASTQ_READ1 FASTQ_READ2

cutadapt -m 15 -u -10 -U 10 -o TRIMMED_CUT_FASTQ_READ1 -p

TRIMMED_CUT_FASTQ_READ2 TRIMMED_FASTQ_READ1

TRIMMED_FASTQ_READ2

-m 15: set minimum length to 15 bp.

-u -10 -U 10: remove the 10 last nt of read1 (3' end) and the 10 first nt of read2 (5'

end).

4. Download a genome reference file in fasta format with a .fa or .fasta extension

and put it into a specific folder (e.g. "REF_FOLDER")

5. Create the bisulfite converted reference sequences (C to T for forward strand and

G to A for reverse strand) and generate bowtie2 indexes for each sequence:

bismark_genome_preparation --bowtie2 --verbose REF_FOLDER

6. Download the phage lambda reference genome in fasta format (.fa or .fasta

extension) and move it to a specific folder (e.g.: "PHAGE_FOLDER").

7. Create the bisulfite converted reference sequences and generate bowtie2 indexes

for each sequence of the phage lambda genome: bismark_genome_preparation --

bowtie2 --verbose PHAGE_FOLDER

8. Map trimmed paired reads to the reference genome with bismark.

For the **NUGEN** protocol:

bismark --bowtie2 -D 50 -R 3 -X 2000 -N 0 REF_FOLDER -1

TRIMMED_FASTQ_READ1 -2 TRIMMED_FASTQ_READ2

For the **CEGX** protocol:

bismark --bowtie2 -D 50 -R 3 -X 2000 -N 0 REF_FOLDER -1

TRIMMED_CUT_FASTQ_READ1 -2 TRIMMED_CUT_FASTQ_READ2

-X 2000: set the maximum allowed insert size to consider for a paired-end

alignment to 2000 nt.

-N 0: set the number of mismatches to be allowed in seed alignment to 0.

-R 3: set the maximum number of times for "re-seed" to 3.

-D 50: set the number of consecutive seed extension attempts to 50.

9. Check the Bismark report file to make sure that the mapping efficiency is sufficient

    (i.e. > 70%) and the methylation percentage in the different contexts are

    consistent with what is expected (>60% for cytosines in the CpG context for

    human genome and close to 0 for the other sequence contexts, close to 0 for all

    sequence contexts for the lambda phage) (*see* **Note 37**).

10. If the phage lambda genome was spiked in the samples, map the reads on the

    phage lambda genome:

    For the **NUGEN** protocol:

    bismark --bowtie2 -D 50 -R 3 -X 2000 -N 0 PHAGE_FOLDER -1

    TRIMMED_FASTQ_READ1 -2 TRIMMED_FASTQ_READ2

    For the **CEGX** protocol:

    bismark --bowtie2 -D 50   -R 3 -X 2000 -N 0 PHAGE_FOLDER -1

    TRIMMED_CUT_FASTQ_READ1 -2 TRIMMED_CUT_FASTQ_READ2

    The bisulfite conversion rate can be calculated as:

    1 – (methylated_positions / methylated_positions + unmethylated positions)

    The methylated and unmethylated position counts can be found in the bismark

    report (text file). The counts from all the contexts (CpG, CHG, CHH, Unknown)

    should be taken into account.

11. Remove duplicated reads:

deduplicate_bismark -p --bam BAM_FILE

The deduplicated bam file will be automatically created by the software with the extension ".deduplicated.bam" and is noted here BAM_FILE_DEDUPLICATED. The report text file can be analysed to check that the number of duplicates is not too high.

12. Extract methylation values from paired-end reads in bam (the .cov.gz file will contain the methylation information at CpG sites) (*see* **Note 38**):

    bismark_methylation_extractor --multicore NB_CORES -p --no_overlap --gzip --merge_non_CpG --comprehensive --no_header --bedGraph --o OUTPUT_DIR

    BAM_FILE_DEDUPLICATED

    -p indicates BAM_FILE is paired-end

    --no_overlap: for overlapping reads, count methylation only once on read 1.

    --merge_non_CpG: produce two output files (CpG context and non-CpG context).

    --no_header: suppresses the bismark version header line in all output files for more convenient batch processing.

    --bedGraph --counts creates CpG context bedGraph (chr positions + methylation values) and coverage file (same as bedGraph but with #Cs and #Ts).

    --gzip: compress methylation extractor files.

    --comprehensive: creates one methylation file per methylation context.

    Representative results for the Methyl-seq protocol are shown in Figure 2.

13. Optional: For the analysis of DMS or DMRs, a number of freely available tools can be used including MethylKit [44] or RADMeth (http://smithlabresearch.org/software/ radmeth/).

4 Notes

1. Between 50 and 100 ng of input DNA allow obtaining high diversity libraries with comprehensive genomic coverage. We have successfully sequenced samples with input material as limited as 30 ng. However, below this amount the duplicate rates increase significantly and for sufficient coverage extensive sequencing is required. We therefore recommend a minimum input of 50 ng and ideally 100 ng for routine use.

2. For all procedures prior to PCR amplification, it is recommended to work in Laminar Flow Cabinets and use filter tips. After PCR, it is recommended to work in a physically separated post-PCR room.

3. The beads should not disperse, and stick to the walls of the tubes to avoid loss during the wash process. Significant loss of beads at this stage will impact the amount of purified DNA, so ensure beads are not removed with the binding buffer or during the wash steps.

4. With the final wash, it is critical to remove as much of the ethanol as possible. If necessary, use extra pipetting steps and allow excess ethanol to collect at the bottom of the tubes after removing most of the ethanol in the first pipetting step.

5. The L1 Ligation Buffer Mix is very viscous. Pipet this reagent slowly.

6. We have evaluated several commercial bisulfite conversion kits and all kits tested have been found to be compatible with the library preparation protocol. However, the kit cited within this protocol allowed us obtaining the most consistent results and the best yields after bisulfite conversion with excellent conversion rates.

7. If necessary, heat the Bisulfite Solution to 60°C and vortex until all precipitates are dissolved. Do not place dissolved bisulfite solution on ice.

8. DNA Protect Buffer should turn from green to blue after addition to the DNA–Bisulfite Solution mixture, indicating sufficient mixing and correct pH for the bisulfite conversion reaction.

9. If there are precipitates in the BD Buffer, avoid transferring them to the spin column.

10. To avoid acidification from carbon dioxide in the air, it is important to close the bottle containing the BD Buffer immediately after use and close the lids of the spin columns before incubation.

11. Proceed immediately or store purified DNA at -20°C.

12. Rn is the fluorescence of the reporter dye divided by the fluorescence of a passive reference dye.

13. Slightly bigger inserts have not been found to negatively impact the sequencing of the library. For significantly larger inserts, this might however have an influence on library diversity and as inserts are penalized during cluster generation on the cBot.

14. 452 bp is average size of the standard included in the KAPA library quantification kit. The average size of the library preparation is obtained in the Bioanalyzer analysis.

15. For optimal duration of the protocol, *day 1*: 3.2.1 to 3.2.5 and 3.2.12 steps 1-3, *day 2*: 3.2.6 to 3.2.8 and 3.2.12 steps 4-19, *day 3*: 3.2.9, *day 4*: 3.2.10 and 3.2.12 steps 20-25. The TrueMethyl® WholeGenome of CEGX kit (Version 3.1) offers several stop points after: conversion bisulfite, desulfonation, end activation, strand synthesis, library finishing and indexing, and between the PCR for digestion control and the actual digestion. However, we recommend incorporating the control of the digestion within the workflow of the experiment and perform the PCR for the digestion step at the end of day 1 and the purification and digestion steps at the end of day 2.

16. The OxBSseq protocol is optimized using 400 ng of DNA as starting material, but it is possible to start with 100 ng of DNA if limited amounts are available. Determine

the optimal number of PCR cycles to increase concentration of library preparation. However, duplicate rates are generally higher when starting with less material.

17. For library preparation, the workflow can start with high molecular weight or 800 bp fragmented DNA. For sequencing on a HiSeq X instrument, we recommend to use high molecular weight DNA to obtain libraries, which are in the optimal size range for clustering and sequencing.

18. To optimize the amount of bisulfite reagent solution, samples should be treated in multiples of six. Moreover, to obtain reproducible results especially in the purification steps, we do not recommend preparing a large number of samples simultaneously. We recommend starting with 12 samples maximum and use a multichannel pipette for purification steps.

19. The digestion control is an optional step to control the efficiency of oxidation and bisulfite step. As it is very helpful to take a decision on whether proceeding to sequencing, we recommend performing this control. Sequencing spike-in controls are in addition used to quality control the data at the analysis step.

20. Ensure that Magnetic Beads are fully resuspended before adding MBBS solutions to the samples.

21. For the elution step, resuspend the beads by pipetting thoroughly to avoid of beads sticking to the wall of tube.

22. Ensure that the eluate does not contain beads.

23. There should be no ethanol and dry ice handled in the laboratory 2 h before oxidation to ensure complete oxidation of samples.

24. After the oxidation reaction, the color of the sample should be orange/brown. An orange color indicates the end of oxidation. However, if no orange/brown color is obtained, continue nonetheless with the protocol. As can be judged later from the digestion control, oxidation might have worked and does in our experience in

most cases. The pellet of the oxidation reaction is black and might inhibit the next step.

25. For more efficiency, add 700 µL of Bisulfite Diluent to an aliquot of the Bisulfite Reagent, vortex and transfer with maximum of powder into a new 2 mL microtube and incubate at 60 °C mixing at 1400 RPM for 5 min. When the solution is clear, transfer the warm solution in the original tube and incubate at 60°C for 10 min under agitation. Ensure that all Bisulfite Reagent is fully dissolved. Equilibrate Bisulfite Reagent at room temperature.

26. A grey color does not affect the reaction.

27. If the white pellet is aspirated, centrifuge again sample.

28. The MBBS3 can be stored at room temperature for the day or stored at 4°C for one week.

29. We recommend adding this step compared to the standard TrueMethyl® WholeGenome CEGX Version 3.1 protocol to make sure to have a sufficient amount of library to sequence.

30. For multiplexing libraries, choose samples with different compatible indexes (see Illumina recommendations)

Sequences of 24 indexes of Illumina V2

Index	Sequence	Index	Sequence
1	ATCACG	13	AGTCAA
2	CGATGT	14	AGTTCC
3	TTAGGC	15	ATGTCA
4	TGACCA	16	CCGTCC
5	ACAGTG	18	GTCCGC
6	GCCAAT	19	GTGAAA
7	CAGATC	20	GTGGCC
8	ACTTGA	21	GTTTCG
9	GATCAG	22	CGTACG
10	TAGCTT	23	GAGTGG
11	GGCTAC	25	ACTGAT
12	CTTGTA	27	ATTCCT

31. Plot linear Rn versus cycle number and determine the optimal cycle number that corresponds to 1/3 of the highest fluorescent intensity (Rn is the fluorescence of the reporter dye divided by the fluorescence of a passive reference dye).

32. Quality control includes analysis of the library by BioAnalyser, qPCR quantification, control of the conversion using the digestion spike-in controls, sequencing spike-in controls and insert size analysis.

33. Library preparation must be quantified by qPCR for optimal sequencing. PhiX should also be quantified using qPCR. Prepare successive dilutions (1/50-1/500-1/5000-1/50000).

34. Library preparation should be sequenced on HiSeq X series instrument in 4-plex in four lanes using maximum 20 bp size difference of libraries to ensure an equally efficient clustering process on the flowcell. The four indexes have to be different and compatible following the recommendations from Illumina.

35. Depending on the software used on the sequencing machine, higher amounts of PhiX might be required. Prior to the HiSeq Control software (HCS) V3.4.0. a minimum of 20 % PhiX was required.

36. Make sure your TMPDIR has enough space as these analyses generate large files. Otherwise you may need add this option in BsExpress for the redirection to the TMPDIR of bismark:

    --bismark_opts='--temp_dir YOUR_BIGTMP_PATH'

    and before launching BsExpress for redirection of the trim_galore TMPDIR :

    export TMPDIR= YOUR_BIGTMP_PATH.

37. Plants might have in some species high methylation levels for cytosines in all sequence contexts. In the same line, cellular systems such as iPS cells might show cytosine methylation in other sequence context. In this case only the spike-in DNA can be used to calculate reliably the bisulfite conversion efficiency.

38. Methylation values can also be extracted using MethylKit [44] from the whole bam i.e. paired + single reads. For further details see the read.bismark function in the methylKit R package.

5. Acknowledgements

References

1. How Kit A, Nielsen HM, Tost J (2012) DNA methylation based biomarkers: practical considerations and applications. Biochimie 94:2314-2337
2. Baylin SB, Jones PA (2011) A decade of exploring the cancer epigenome - biological and translational implications. Nat Rev Cancer 11:726-734
3. Lardenoije R, Iatrou A, Kenis G et al (2015) The epigenetics of aging and neurodegeneration. Prog Neurobiol 131:21-64
4. Nielsen HM, Tost J (2012) Epigenetic changes in inflammatory and autoimmune diseases. Subcell Biochem 61:455-478
5. Harb H, Renz H (2015) Update on epigenetics in allergic disease. J Allergy Clin Immunol 135:15-24
6. Zhang Z, Zhang R (2015) Epigenetics in autoimmune diseases: Pathogenesis and prospects for therapy. Autoimmun Rev 14:854-863
7. Shorter KR, Miller BH (2015) Epigenetic mechanisms in schizophrenia. Prog Biophys Mol Biol 118:1-7
8. Abdolmaleky HM, Zhou JR, Thiagalingam S (2015) An update on the epigenetics of psychotic diseases and autism. Epigenomics 7:427-449
9. Ronn T, Ling C (2015) DNA methylation as a diagnostic and therapeutic target in the battle against Type 2 diabetes. Epigenomics 7:451-460
10. Urich MA, Nery JR, Lister R et al (2015) MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. Nat Protoc 10:475-483
11. Kobayashi H, Sakurai T, Imai M et al (2012) Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. PLoS Genet 8:e1002440
12. Adusumalli S, Mohd Omar MF, Soong R et al (2015) Methodological aspects of whole-genome bisulfite sequencing analysis. Brief Bioinform 16:369-379
13. Lister R, Pelizzola M, Dowen RH et al (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. Nature 462:315-322
14. Lister R, O'Malley RC, Tonti-Filippini J et al (2008) Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell 133:523-536
15. Lister R, Pelizzola M, Kida YS et al (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. Nature 471:68-73
16. Li Y, Zhu J, Tian G et al (2010) The DNA methylome of human peripheral blood mononuclear cells. PLoS Biol 8:e1000533
17. Chalhoub B, Denoeud F, Liu S et al (2014) Plant genetics. Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science 345:950-953
18. Lyko F, Foret S, Kucharski R et al (2010) The honey bee epigenomes: differential methylation of brain DNA in queens and workers. PLoS Biol 8:e1000506
19. Guo JU, Su Y, Shin JH et al (2014) Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nat Neurosci 17:215-222
20. Adey A, Shendure J (2012) Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. Genome Res 22:1139-1143
21. Miura F, Enomoto Y, Dairiki R et al (2012) Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res 40:e136
22. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science 324:929-930
23. Tahiliani M, Koh KP, Shen Y et al (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324:930-935

24. Tan L, Shi YG (2012) Tet family proteins and 5-hydroxymethylcytosine in development and disease. Development 139:1895-1902

25. Ludwig AK, Zhang P, Cardoso MC (2016) Modifiers and Readers of DNA Modifications and Their Impact on Genome Structure, Expression, and Stability in Disease. Front Genet 7:115

26. Wu SC, Zhang Y (2010) Active DNA demethylation: many roads lead to Rome. Nat Rev Mol Cell Biol 11:607-620

27. Ito S, Shen L, Dai Q et al (2011) Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. Science 333:1300-1303

28. Nestor CE, Ottaviano R, Reddington J et al (2012) Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. Genome Res 22:467-477

29. Langemeijer SM, Aslanyan MG, Jansen JH (2009) TET proteins in malignant hematopoiesis. Cell Cycle 8:4044-4048

30. Pastor WA, Pape UJ, Huang Y et al (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature 473:394-397

31. Ficz G, Branco MR, Seisenberger S et al (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature 473:398-402

32. Jin SG, Wu X, Li AX et al (2011) Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic Acids Res 39:5015-5024

33. Ono R, Taki T, Taketani T et al (2002) LCX, leukemia-associated protein with a CXXC domain, is fused to MLL in acute myeloid leukemia with trilineage dysplasia having t(10;11)(q22;q23). Cancer Res 62:4075-4080

34. Mercher T, Quivoron C, Couronne L et al (2012) TET2, a tumor suppressor in hematological disorders. Biochim Biophys Acta 1825:173-177

35. Putiri EL, Tiedemann RL, Thompson JJ et al (2014) Distinct and overlapping control of 5-methylcytosine and 5-hydroxymethylcytosine by the TET proteins in human cancer cells. Genome Biol 15:R81

36. Jin SG, Jiang Y, Qiu R et al (2011) 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. Cancer Res 71:7360-7365

37. Booth MJ, Branco MR, Ficz G et al (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science 336:934-937

38. Yu M, Hon GC, Szulwach KE et al (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell 149:1368-1380

39. Skvortsova K, Zotenko E, Luu PL et al (2017) Comprehensive evaluation of genome-wide 5-hydroxymethylcytosine profiling approaches in human DNA. Epigenetics & Chromatin 10:16

40. Song CX, Szulwach KE, Fu Y et al (2011) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat Biotechnol 29:68-72

41. Daviaud C, Renault V, Mauger F et al (2017) Whole-genome bisulfite sequencing using the Ovation® Ultralow Methyl-Seq protocol. Methods Mol Biol:in press

42. Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics 27:1571-1572

43. Li H, Handsaker B, Wysoker A et al (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079

44. Akalin A, Kormaksson M, Li S et al (2012) methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. Genome Biol 13:R87

Figure legends

Figure 1: Workflow of the whole-genome bisulfite sequencing protocol using the Ovation® Low Input Methyl-seq Kit. See the text for a detailed description.

Figure 2: Comparative analysis of the coverage of a locus in the *Hox* gene cluster by WGBS. Upper panel sequence coverage obtained by sequencing two lanes on a HiSeq2000 starting from 4 µg of DNA (average coverage 15 x); middle panel: 2 lanes sequenced of a library prepared from 100 ng of input DNA using the protocol described in this chapter (average coverage 16 x); lower panel: 4 lanes sequenced of a library prepared from 100 ng of input DNA using the protocol described in this chapter (average coverage 33 x). Libraries prepared from 4 µg or 100 ng of DNA yield nearly indistinguishable results. At the recommended coverage of 30 x, a sufficient coverage for accurate estimation of the DNA methylation degree is obtained. Reproduced from [41].

Figure 3: Workflow of the whole-genome oxidative bisulfite sequencing protocol using the TrueMethyl® WholeGenome kit. See the text for a detailed description.

Figure 4: Schematic outline of the different steps of the library preparation for the Ovation® Low Input Methyl-seq: protocol including fragmentation of genomic DNA, repair of 5' and 3' extremities followed by the ligation of indexed adaptors and the final end repair prior to bisulfite conversion.

Figure 5: Left: A representative Bioanalyzer profile using the DNA 1000 Chips. The x axis shows the size distribution of the amplification products in base pairs, and the y axis shows units of

fluorescence. The red and blue curve represent libraries made from 50 ng and 100 ng DNA, respectively. The average library size is approximately 290 base pairs. **Right:** Determination of the optimal number of PCR cycles. Relative fluorescence is plotted against the number of PCR cycles. For 10, 20, 50 and 100 ng of input DNA, the optimal number of cycles is respectively, 13, 12, 10 and 9 cycles. 50 and 100 ng input DNA do allow for a high quality and diversity library, with low duplicate reads. Modified from [41].

Figure 6: Quality-control of OxBS library preparation: analysis of library size, quantification and determination of insert size. **Upper left panel:** A representative Bioanalyzer profile using DNA 1000 Chips. The Bs library is shown in red, the OxBs library in blue. The x axis shows the size distribution of the amplification products in base pairs, and the y axis shows units of fluorescence. The average library size is approximately 460 (Bs) and 550 (OxBs) base pairs. **Upper right panel:** Determination of the optimal number of PCR cycles. Relative fluorescence is plotted against the number of PCR cycles for a Bs and the corresponding OxBs library showing similar amplification efficiency. **Lower panel:** Distribution of insert size as determined from mappable reads in a multiple sequencing experiment on a Mi-seq.

Figure 7: Analysis of sequencing spike-ins after sequencing a Bs and OxBs library on a HiSeq X5 instrument. In the BS experiment, 5hmC spike-ins (red) are read as C, while they are mostly absent in the OxBS experiment. Red arrows indicate suboptimal conversion of some spike-ins compared to the theoretical distribution. However, in our experience, the same positions of the spike-ins yield similar and reproducible conversion efficiencies in independent experiments.
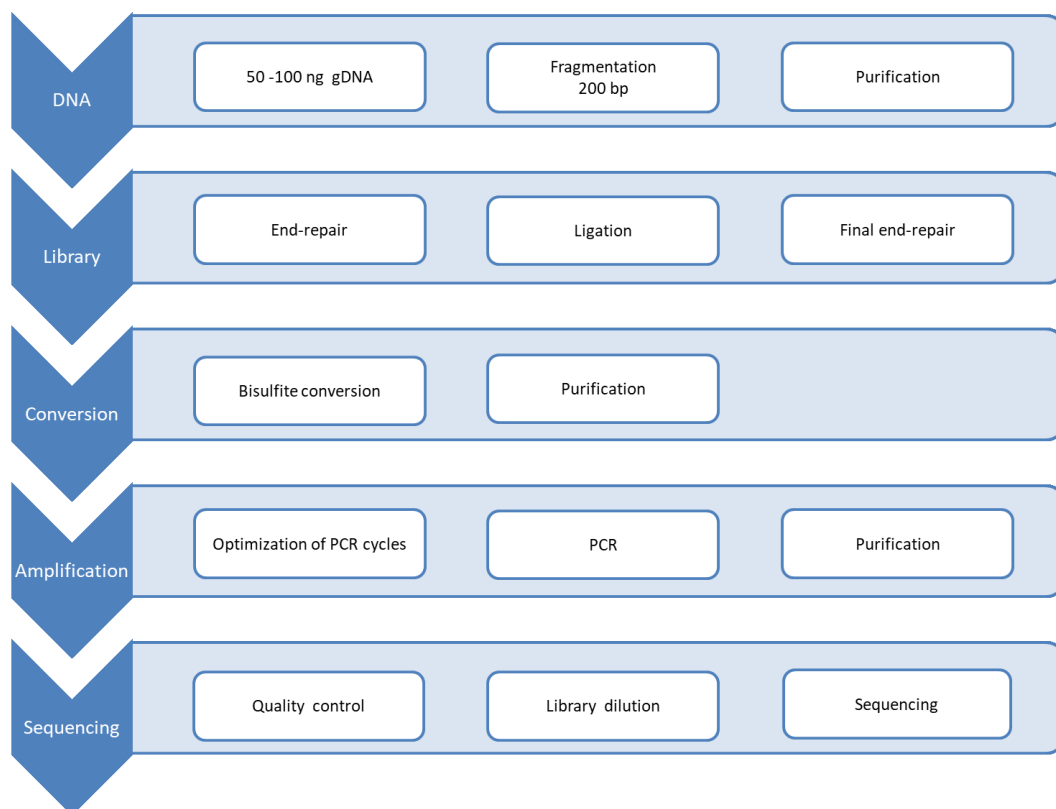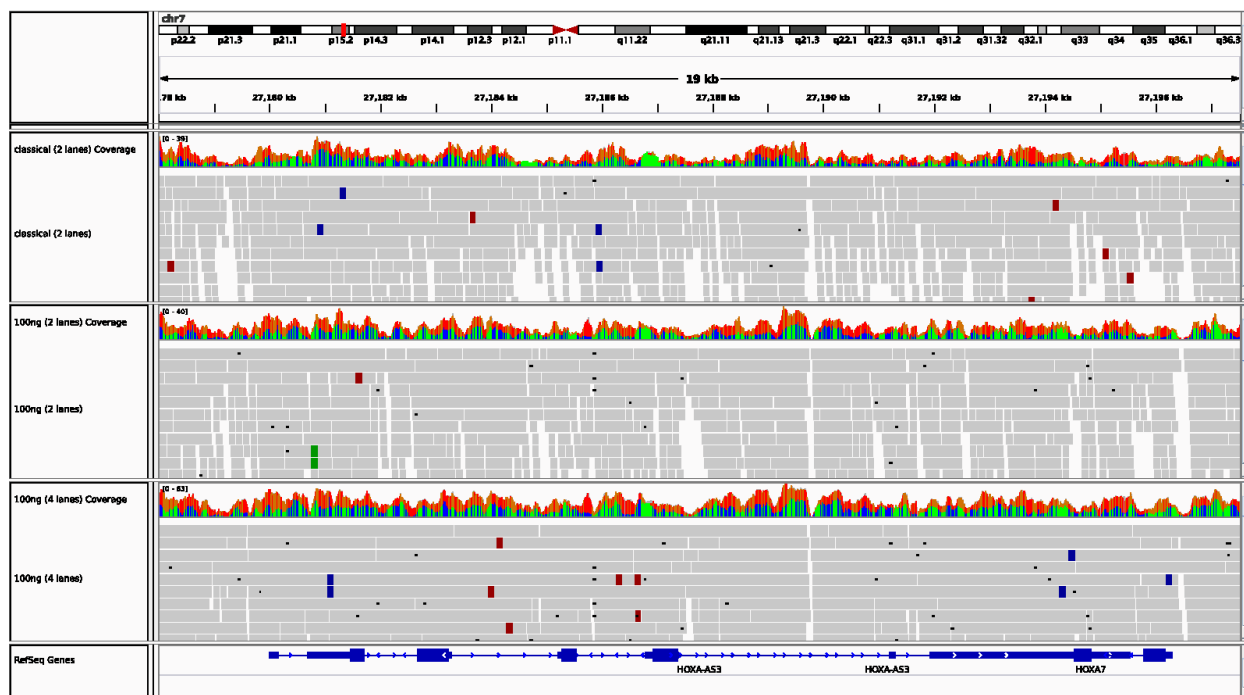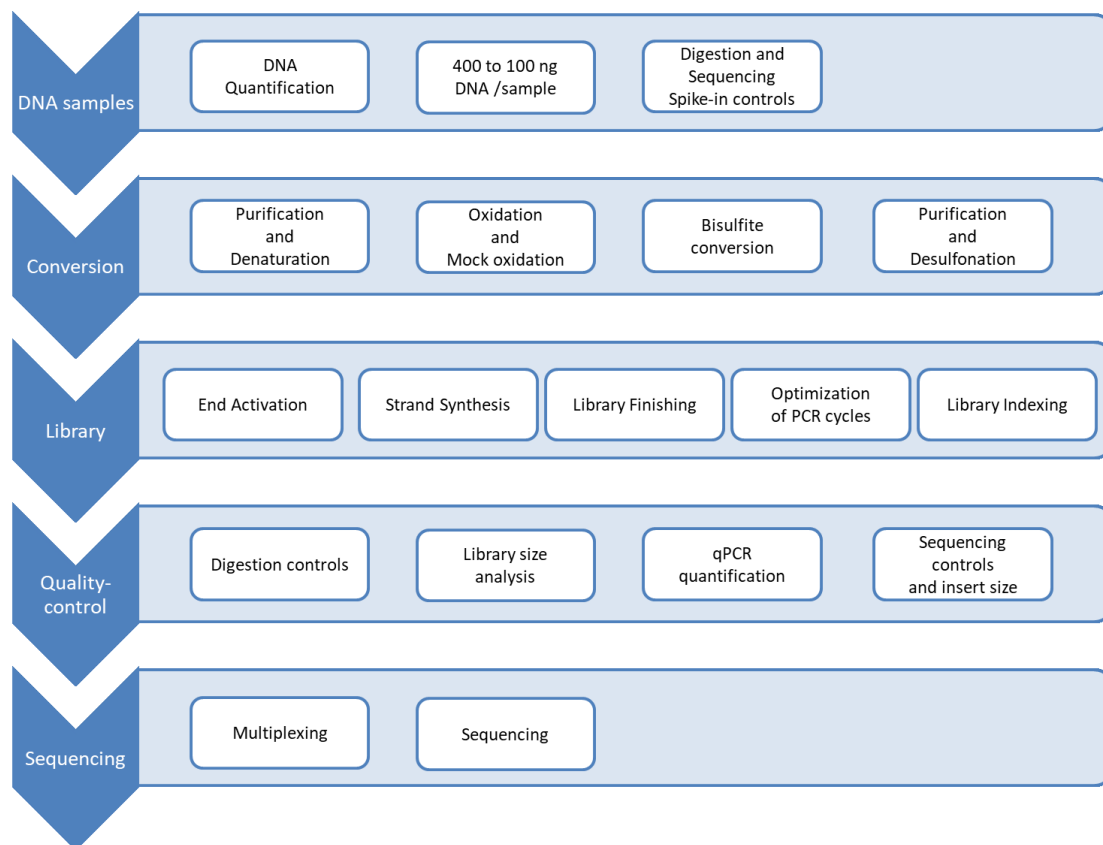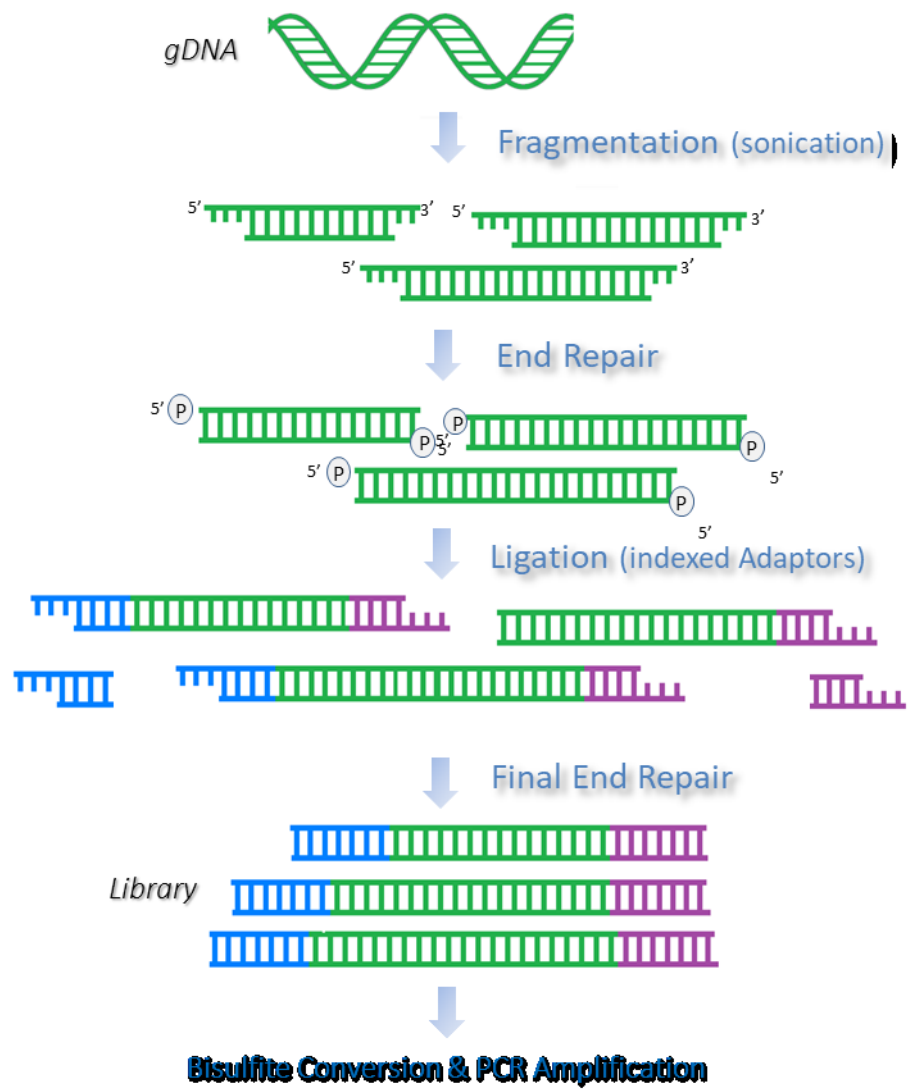
Figure 1

Figure 2

Figure 3

Figure 4
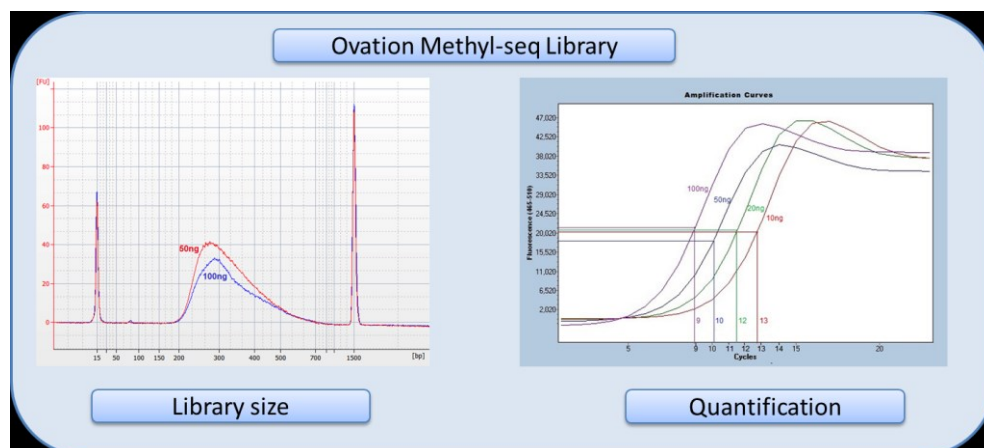
Figure 5

Figure 6

Figure 7