Published in final edited form as:

Methods Mol Biol. 2018 ; 1708: 645–663. doi:10.1007/978-1-4939-7481-8_33.

Tet-Assisted Bisulfite Sequencing (TAB-seq)

Miao Yu, Dali Han, Gary C. Hon, and Chuan He

Abstract

5-Hydroxymethylcytosine (5hmC) is a modified form of cytosine, which has recently been found in mammalian cells and tissues. 5hmC is derived from 5-methylcytosine (5mC) by Ten-eleven translocation (TET) family protein-mediated oxidation and may regulate gene expression. Numerous affinity-based profiling methods have been developed to help understand the exact function of 5hmC in the genome. However, these methods have a relatively low resolution (~100 bp) without quantitative information of the modification percentage on each site. Here we demonstrated the detailed procedure of Tet-Assisted Bisulfite Sequencing (TAB-Seq), which can detect 5hmC at single-base resolution and quantify its abundance at each site. In this protocol, the genomic DNA is first treated with β GT and recombinant mTet1 consecutively to convert 5hmC to 5gmC and 5mC to 5caC, respectively. The treated genomic DNA can be directly applied to bisulfite treatment to detect 5hmC on specific loci or applied to whole-genome bisulfite sequencing as needed.

Keywords

TAB-Seq; 5-Hydroxymethylcytosine; Bisulfite sequencing; DNA methylation; Glycosylation

1 Introduction

The significance of 5-hydroxymethylcytosine (5hmC), which was first reported in 1972, did not receive much attention until its presence in the mammalian genome was recently confirmed in Purkinje neurons and mouse embryonic cells [1-3]. 5hmC is the direct oxidative product of 5-methylcytosine (5mC) through Ten-eleven translocation (TET) family protein-mediated oxidation in a Fe^{II} - and 2-oxoglutarate-dependent manner [3]. 5hmC has been suggested to be an intermediate in active and passive demethylation pathways, whereas the uneven distribution of 5hmC in different cell types and tissues indicates its potential roles in gene regulation [4-11]. Although the exact function of 5hmC is still elusive, studies on 5hmC and TET proteins demonstrated that 5hmC was closely related to embryonic stem cells maintenance, brain development, and zygote development [12-20]. 5hmC can be further oxidized to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), both can be excised by thymine DNA glycosylase (TDG) and converted back to unmodified cytosine through base excision repair (BER) [5, 21-24].

Antibodies specific against 5hmC have been developed and applied to profile 5hmC in genome (hMeDIP) shortly after the discovery of 5hmC [12, 25-27] (*see also* Chapter 35). Other affinity-based profiling methods were also reported, which take advantage of enzymes or chemicals to reduce the pull-down bias toward high-density hydroxymethylated regions

[11, 28-31] (*see also* Chapter 35). However, profiling methods cannot give base resolution information and the quantitative information on the modification percentage is lost during enrichment. Bisulfite sequencing has been widely used to read out 5mC quantitatively by deaminating only unmodified cytosines to uracils [32-34]. Unfortunately, the same strategy is not feasible for 5hmC detection because both 5mC and 5hmC resist bisulfite-mediated deamination and are read as C, which also indicates that the previously identified 5mC sites are the sum of 5mC and 5hmC [35, 36] (Fig. 1a). To differentiate these two modifications, we modified the traditional bisulfite sequencing method and developed Tet-Assisted Bisulfite Sequencing (TAB-Seq), in which two enzymatic treatment steps were introduced before bisulfite conversion [37]. First, β -glucosyltransferase (β GT) is utilized to convert 5hmC in genomic DNA to β -glucosyl-5-hydroxymethylcytosine (5gmC) which cannot be further oxidized to 5fC or 5caC by the Tet proteins [38]. Next, the glucosylated genomic DNA is treated with an excess of highly active recombinant mouse Tet1 (mTet1) to oxidize most 5mC to 5caC, which can be readily deaminated upon standard bisulfite treatment [5]. After bisulfite treatment and amplification, 5hmC is sequenced as C, whereas both C and 5mC are sequenced as T, therefore distinguishing 5hmC sites from C and 5mC sites (Fig. 1b).

Besides providing the precise location of 5hmC, TAB-Seq can also provide information about the abundance of 5hmC at each site. To accurately estimate 5hmC abundance, three key parameters are required: (1) the conversion rate of C to T; (2) the conversion rate of 5mC to T; (3) the protection rate of 5hmC. To assess these three parameters in each TAB-Seq sample, proper controls bearing C, 5mC, or 5hmC need to be spiked into the genomic DNA prior to the enzymatic treatment steps. There are several rules to follow to determine the control sequence: (1) complex enough and have multiple modification sites (5mC or 5hmC); (2) at least 1 kb long, allowing for PCR duplicate removal during data analysis in order to accurately calculate the conversion and protection rate; (3) do not align to the target genome after bisulfite conversion. In this protocol, we use lambda DNA and a pUC19 vector as template to generate the spike-in controls, which are compatible with human and mouse genomes. If samples from other species are sequenced, proper controls with a different sequence may be needed and the PCR primers should be redesigned for the conversion/protection test described in Subheadings 3.2 and 3.5.

The sequencing depth required for each whole-genome TAB-Seq depends on both the abundance of 5hmC in the sequenced sample and the 5mC conversion rate. Less sequencing is required for samples with higher 5hmC abundance and a higher 5mC conversion rate. For example, if the median abundance of 5hmC at 5hmC sites in the genomic DNA is ~20%, to resolve base-resolution 5hmC at this level, a depth of ~25 times per cytosine, or ~50 times the haploid genome size is needed when the 5mC conversion rate is 97.8%. The number will increase to ~30 times per cytosine, or ~60 times the haploid genome size if the 5mC conversion rate drops to 96.5%. On the other hand, a depth of ~15 times per cytosine is sufficient to resolve base-resolution 5hmC with 30% abundance when 5mC conversion is 97.8%. With a 5mC conversion rate of 99%, a depth of ~16 times per cytosine is sufficient to resolve base-resolution 5hmC with 20% abundance.

2 Materials

2.1 Expression and Purification of Recombinant mTet1

1. Bac-to-Bac Baculovirus Expression System (Thermo Fisher Scientific).
2. pFastBac Dual-mTet1 plasmid (available upon request from the corresponding author).
3. Adherent High Five cells.
4. Grace's insect medium, supplemented.
5. Fetal Bovine Serum (FBS).
6. Penicillin/Streptomycin/L-Glutamine (Pen/Strep/L-Gln) mixture 25,000 U/25000 U/200 mM.
7. Insect cell culture medium: Add 55 mL FBS, 2.25 mL Pen/-Strep/L-Gln to 500 mL of Grace's insect medium, supplemented and filter (0.22 µm filter unit). Store at 4 °C.
8. Binding buffer: 20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), 1 mM Phenylmethylsulfonyl fluoride (PMSF), 1 µg/mL Leupeptin, 1 µg/mL Pepstatin (*see Note 1*).
9. Econo-Column chromatography column.
10. Anti-flag M2 Affinity Gel (Sigma-Aldrich).
11. Econo-Column chromatography columns (Bio-Rad, 2.5 × 20 cm).
12. 3× Flag peptide (Sigma-Aldrich).
13. GF running buffer: 20 mM HEPES, pH 8.0, 150 mM NaCl, 1 mM DTT (*see Note 2*).
14. Elution buffer: Dissolve 3× Flag peptide in precooled GF running buffer at 0.2 mg/mL. Add 1 mM PMSF, 1 µg/mL Leupeptin and Pepstatin just before use.
15. Amicon Ultra centrifugal filters (4 mL, 50K cutoff).
16. Superdex-200 columns.
17. Glycerol.

2.2 Activity Test of Recombinant mTet1

1. Qubit 2.0 Fluorometer (Thermo Fisher Scientific).
2. Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific).
3. Tet oxidation reagent 1: 1.5 mM Fe(NH₄)₂(SO₄)₂. Add 14.7 mg Fe(NH₄)₂(SO₄)₂·6H₂O to 1 mL of Milli-Q water, then make a 24-fold dilution (*see Note 3*).
4. Tet oxidation reagent 2: 333 mM NaCl, 167 mM HEPES (pH 8.0), 4 mM ATP, 8.3 mM DTT, 3.3 mM α-KG, 6.7 mM L-Ascorbic acid (*see Note 4*).

5. Proteinase K (20 mg/mL).
6. Micro Bio-Spin 30 Columns (Bio-Rad).
7. QIAquick PCR purification kit (Qiagen) or DNA purification kits from other suppliers.
8. Epitect Bisulfite kit (Qiagen) or MethylCode Bisulfite Conversion kit (Thermo Fisher Scientific) or similar.
9. *Pfu*Turbo C_x Hotstart DNA polymerase.
10. 5mC test primer-F: 5'-TTTGGGTTATGTAAGTTGATTTTATG.
11. 5mC test primer-R: 5'-CACCTACTTACTAAAATTACACC.
12. dNTP mix.
13. Zero Blunt TOPO PCR Cloning Kit.
14. Agarose.

2.3 Generation of Spike-In Controls

1. *Sss*I CpG methyltransferase with associated buffer and S-adenosylmethionine (e.g., New England Biolabs).
2. Unmethylated lambda DNA from *E. coli* strain lacking both the *dam* and *dcm* methylase activities.
3. 5hmC dNTP mix.
4. pUC19 vector.
5. Zymo *Taq*TM DNA polymerase.
6. 5hmC control primer-F: 5'-GCAGATTGTACTGAGAGTGC.
7. 5hmC control primer-R: 5'-TGCTGATAAATCTGGAGCCG.
8. Agarose
9. Gel extraction kit.

2.4 Glycosylation and Oxidation of Genomic DNA

1. 10× βGT protection buffer: 500 mM HEPES, pH 8.0, 250 mM MgCl₂.
2. T4 Phage β-glucosyltransferase (T4-βGT, New England Biolabs)
3. 10 mM UDP-Glucose: Dissolve UDP-Glucose in Milli-Q water and store at -20 °C.
4. MinElute PCR purification kit (Qiagen) or equivalent.

All other required reagents can be found in Subheading 2.2.

2.5 Verification of Oxidized Genomic DNA

1. 5hmC test primer-F: 5'-GTAGATTGTATTGAGAGTGT.

2. 5hmC test primer-R: 5'-TACCCAACCTTAATCGCCTTG.

Other required reagents can be found in Subheading 2.2.

2.6 Sequencing of Oxidized Genomic DNA

1. Self-designed primer sets for loci-specific TAB-Seq.
2. End-It™ DNA End-Repair Kit (Lucigen).
5. Klenow Fragment (3' → 5' exo-).
6. dATP.
7. Quick Ligation Kit (New England Biolabs).
8. Agarose (super fine resolution, APEX).
9. MinElute Gel Extraction Kit (Qiagen).
10. KAPA HiFi HotStart ReadyMix PCR Kit.
11. TruSeq® DNA Sample Prep Kit (Illumina).
12. Library primer-F: 5'-AATGATACGGCGACCACCGAGATCTACAC.
13. Library primer-R: 5'-CAAGCAGAAGACGGCATACGAGAT.
14. Illumina Hi-Seq 2000 or 2500.
15. Bismark software (<http://www.bioinformatics.babraham.ac.uk/projects/bismark/>) [39].

Other required reagents can be found in Subheading 2.2.

3 Methods

3.1 Expression and Purification of Recombinant mTet1

Recombinant mTet1 needs to be carefully expressed and purified to ensure high oxidation efficiency (over 97% conversion of 5mC to 5caC) for the TAB-Seq assay. Perform all the purification steps on ice or at 4 °C and try to finish in 1 day, if time allows. The TAB-Seq kit is also commercially available from Wisegene and it provides required enzymes and buffers for the glycosylation and oxidation of genomic DNA in Subheading 3.4.

1. Generate the bacmid and produce baculovirus P1, P2, and P3 with the pFastBac Dual-mTet1 plasmid using the Bac-to-Bac Baculovirus Expression System following the manufacturer's instructions.
2. Infect adherent High Five cells with 1 mL of P3 baculovirus for each 150 mm culture dish and incubate at 27 °C for 48 h.
3. Harvest 20–40 dishes of infected cells by centrifuging at $3800 \times g$ for 10 min at 4 °C. Discard the medium and resuspend the cell pellets in 30–50 mL of binding buffer (see **Note 5**).

4. Break cells by passing through the homogenizer twice and collect the supernatant by centrifuging at $30,000 \times g$ for 40 min. Clarify the cell lysate by passing through a 0.45 μ M filter.
5. Add 6 mL of 50% anti-flag M2 affinity gel into the Econo-Column (2.5 $\times$ 20 cm) and allow the gel bed to drain. Add 50 mL of binding buffer to wash out the glycerol and drain the remaining buffer until it reaches the top of the gel bed.
6. Load the clarified cell lysate onto the equilibrated gel with the bottom closed and resuspend the gel using a pipette. Incubate for 10 min.
7. Let the cell lysate draw through the column under gravity at the rate of one drop per second and collect the flow-through.
8. Load the flow-through from **step 7** back to the column and repeat **steps 6** and **7**.
9. Wash the gel with 200–300 mL of binding buffer under gravity flow. Periodically resuspend the gel with a pipette during the washing.
10. Resuspend the gel in 3 mL of elution buffer and collect the eluate with the flow rate of one drop per second after incubating for 5 min (*see Note 6*).
11. Repeat **step 10** twice (*see Note 7*).
12. Combine the eluate and concentrate to less than 2 mL with an Amicon Ultra 50K filter. Load it onto a 120 mL Superdex-200 column equilibrated with precooled GF running buffer.
13. Pool fractions and concentrate purified mTet1 with an Amicon Ultra 4 mL 50K filter to about 7 mg/mL as quantified using a Bradford assay.
14. Add glycerol to the final concentration of 30% and aliquot the protein to a smaller volume as needed. Freeze with liquid nitrogen and store at -80°C (*see Note 8*).

3.2 Activity Test of Recombinant mTet1

The amount of DNA is quantified with a Qubit Assay unless indicated otherwise.

1. Prepare the reaction as follows. Add each component in the order listed and mix well.

Component	Amount
Milli-Q water	To final volume of 50 μ L
Sheared genomic DNA of interest (from Subheading 3.4)	300 ng
Sheared CpG methylated lambda DNA (from Subheading 3.3.1)	2.5 ng
Tet oxidation reagent 2	15 μ L
Tet oxidation reagent 1	3.5 μ L
mTet1 protein (5 mg/mL)	4 μ L

2. Incubate at 37°C for 80 min.
3. Add 1 μ L Proteinase K (20 mg/mL) and incubate at 50°C for 1h.

4. Clean up with a Micro Bio-Spin 30 Column and purify with the QIAquick PCR purification kit following the manufacturer's instructions. Elute DNA in 30 μ L of Milli-Q water (*see Note 9*).
5. Use 20 μ L of the oxidized DNA from the previous step for the EpiTect or MethylCode Bisulfite Conversion kit following the manufacturer's instructions (*see Note 10*).
6. Prepare the following PCR in a final volume of 50 μ L:

Component	Amount
10 $\times$ <i>Pfu</i> Turbo C _x reaction buffer	5 μ L
dNTP Mix (10 mM each dNTP)	1 μ L
5mC test primer-F (10 μ M)	1 μ L
5mC test primer-R (10 μ M)	1 μ L
Bisulfite-treated DNA (from step 5)	2–4 μ L
<i>Pfu</i> Turbo C _x Hotstart DNA polymerase (2.5 U/ μ L)	1 μ L (<i>see Note 11</i>)
Milli-Q water	To a final volume of 50 μ L

7. Run the PCR program as provided below.

Initial denaturation	95 °C	2 min
Denaturation	95 °C	30 s
Annealing	57 °C	30 s
Extension	72 °C	1 min
		40 cycles
Final extension	72 °C	7 min
Hold	4 °C	

8. Verify the PCR product by running 5 μ L of the reaction on a 2% agarose gel. There should be a single, clear band at ~300 bp.
9. TOPO clone 2 μ L of the reaction using the Zero Blunt TOPO PCR Cloning Kit following the manufacturer's instruction and isolate plasmids from at least 20 clones for Sanger sequencing.
10. Calculate the 5mC conversion rate. The ideal conversion rate of 5mC is over 97% (*see Note 12*).

Conversion rate of 5mC equals:

number of 'T' reads at CpG cytosines/(number of CpGs on amplicon $\times$ number of clones sequenced)

3.3 Generation of Spike-In Controls

3.3.1 C/5mC Spike-In Control

1. Prepare an in vitro methylation reaction. Add each component in the order listed and mix well (*see Note 13*).

Component	Amount
Milli-Q water	44 μ L

10× NEBuffer 2	10 µL
SAM (32 mM)	2 µL
Unmethylated lambda DNA (450 µg/ mL)	40 µL
SssI methylase (20 U/µL)	4 µL
Final volume	100 µL

2. Incubate at 37 °C for 2 h followed by heating at 65 °C for 20 min to stop the reaction.
3. Clean-up DNA with phenol chloroform extraction and ethanol precipitation, or use proper commercial DNA purification kits (*see Note 14*).

3.3.2 5hmC Spike-In Control

1. Prepare the following PCR in a final volume of 50 µL:

Component	Amount
2× Reaction buffer	25 µL
5hmC dNTP Mix (10 mM each dNTP)	1 µL
5hmC control primer-F (10 µM)	2 µL
5hmC control primer-R (10 µM)	2 µL
pUC 19 vector (1 ng/–L)	1 µL
Zymo <i>Taq</i> TM DNA Polymerase (5 U/µL)	0.4 µL
Milli-Q water	18.6 µL
Final volume	50 µL

2. Run the PCR program as provided below.

Initial denaturation	95 °C	2 min
Denaturation	95 °C	30 s
Annealing	57 °C	30 s
Extension	72 °C	1.5 min
		40 cycles
Final extension	72 °C	7 min
Hold	4 °C	

3. Verify the PCR product by running 5 µL of the reaction on a 1.2% agarose gel. There should be a single, clear band at ~1.6 kb.
4. Run the remaining reaction on a 1.2% agarose gel. Cut the gel and purify the PCR product with a commercial gel extraction kit. Elute DNA in Milli-Q water (*see Note 15*).

3.4 Glycosylation and Oxidation of Genomic DNA

1. Add 0.5% (w/w) CpG methylated lambda DNA (C/5mC spike-in control) and 0.25% (w/w) 5hmC control (generated from Subheading 3.3) to the genomic DNA of interest (*see Note 16*).
2. Shear the mixed genomic DNA to desired size range. For whole-genome bisulfite sequencing, the size range depends on the downstream sequencing parameters.

We typically shear to an average of 400 bp. For loci-specific analysis, shear the genomic DNA to average 1 kb or longer (*see Note 17*).

3. Purify the sheared genomic DNA with MinElute PCR purification kit and elute in 10 μ L of Milli-Q water (*see Note 18*).
4. Prepare the glucosylation reaction as follows:

Component	Amount
Sheared DNA from step 3	1–3 μ g
10 mM UDP-Glucose	0.4 μ L
10 $\times$ β GT protection buffer	2 μ L
T4- β GT (10 U/ μ L)	2 μ L
Milli-Q water	To a final volume of 20 μ L

5. Pipette up and down to mix and incubate at 37 °C for 1 h.
6. Clean up the DNA using a QIAquick PCR purification kit and elute in 30–50 μ L of Milli-Q water. After purification, 60–70% of DNA can be recovered.
7. Prepare the oxidation reactions as follows. Add each component in the order listed and mix well (*see Note 19*).

Component	Amount
Milli-Q water	To a final volume of 50 μ L
Glucosylated DNA from step 6	300 ng
Tet oxidation reagent 2	15 μ L
Tet oxidation reagent 1	3.5 μ L (<i>see Note 20</i>)
mTet1 protein (5 mg/ ml)	4 μ L (<i>see Note 21</i>)

8. Incubate at 37 °C for 80 min.
9. Add 1 μ L of Proteinase K (20 mg/mL) per reaction and incubate at 50 °C for 1 h.
10. Clean up the DNA with Micro Bio-Spin 30 Columns followed by the QIAquick PCR purification kit. Elute DNA in 50 μ L of Milli-Q water. Normally, 40–60% of DNA can be recovered.

3.5 Verification of Oxidized Genomic DNA

3.5.1 5mC Conversion Test

1. Convert 10 ng treated of oxidized genomic DNA from Subheading 3.4. with the EpiTect or the MethylCode Bisulfite Conversion kit following the manufacturer's instructions.
2. Follow **steps 6–10** in Subheading 3.2 using bisulfite-treated DNA from **step 1** as the PCR template (*see Note 22*).

3.5.2 5hmC Protection Test

1. Prepare the following PCR in a final volume of 50 μ L:

Component	Amount
-----------	--------

10 × <i>Pfu</i> Turbo C _x reaction buffer	5 µL
dNTP Mix (10 mM each dNTP)	1 µL
5hmC testprimer-F (10 µM)	1 µL
5hmC test primer-R (10 µM)	1 µL (<i>see Note 23</i>)
Bisulfite-treated DNA (from step 1 in Subheading 3.5.1)	2–4 µL (<i>see Note 24</i>)
<i>Pfu</i> Turbo C _x Hotstart DNA polymerase (2.5 U/µL)	1 µL (<i>see Note 25</i>)
Milli-Q water	To a final volume of 50 µL

2. Run the PCR program as follows:

Initial denaturation	95 °C	2 min
Denaturation	95 °C	30 s
Annealing	51 °C	30 s
Extension	72 °C	1 min
		40 cycles
Final extension	72 °C	7 min
Hold	4 °C	

- Verify the PCR product by running 5 µL of the reaction on a 2% agarose gel. There should be a single, clear band at ~200 bp.
- TOPO clone 2 µL of the reaction using a Zero Blunt TOPO PCR Cloning Kit following the manufacturer's instructions and isolate plasmids from at least 20 clones for Sanger sequencing.
- Calculate the 5hmC protection rate. The ideal protection rate of 5hmC is over 80% (*see Note 26*).

Protection rate of 5hmC equals:

number of 'C' reads at cytosines/(number of cytosines × number of clones sequenced)

3.6 Sequencing of Oxidized Genomic DNA

3.6.1 Loci-Specific Sequencing

- Convert 150 ng of treated genomic DNA from Subheading 3.4. using the EpiTect or MethylCode Bisulfite Conversion kit following the manufacturer's instructions.
- Prepare the following PCR in a final volume of 50 µL:

Component	Amount
10× <i>Pfu</i> Turbo C _x reaction buffer	5 µL
dNTP Mix (10 mM each dNTP)	1 µL
Self-designed primer-F (10 µM)	1 µL (<i>see Note 27</i>)
Self-designed primer-R (10 µM)	1 µL
Bisulfite-treated DNA (from previous step)	1–2 µL
<i>Pfu</i> Turbo C _x Hotstart DNA polymerase (2.5 U/µL)	1 µL
Milli-Q water	To a final volume of 50 µL

3. Run the PCR program as follows:

Initial denaturation	95 °C	2 min
Denaturation	95 °C	30 s
Annealing	Primer dependent (see Note 28)	30 s
Extension	72 °C	1 min
	40 cycles	
Final extension	72 °C	7 min
Hold	4°C	

4. Verify the PCR product by running 5 µL of the reaction on a 2% agarose gel. Purify the PCR product with a QIAquick PCR purification kit if a single band is observed. Otherwise, cut the band with the correct size and purify with a gel extraction kit.
5. Apply the purified PCR product to Sanger sequencing or perform colony picking with a Zero Blunt TOPO PCR Cloning Kit if quantitative information is needed. The signal of 'C' represents the position of 5hmC.

3.6.2 Whole-Genome Sequencing

1. Take 300–600 ng of oxidized genomic DNA (from Subheading 3.4) and prepare the end repair reaction as follows.

Component	Amount
Treated genomic DNA (300–600 ng)	Variable depending on its concentration
dNTP mix (10 mM each dNTP)	5 µL
10× End-Repair Buffer	5 µL
ATP (10 mM)	5 µL
End-Repair Enzyme Mix	1 µL
Milli-Q water	To a final volume of 50 µL

2. Mix well and incubate at room temperature for 45 min.
3. Purify DNA with the MinElute PCR purification kit and elute in 16 µL of EB buffer twice (32 µL in total).
4. Prepare 3' end adenylation reaction as follows and mix well.

Component	Amount
DNA from step 3	32 µL
10× NEBuffer 2	5 µL
dATP (1 mM)	10 µL
Klenow fragment (3' → 5' exo-, 5 U/µL)	3 µL

5. Incubate at 37 °C for 30 min.
6. Purify DNA with the MinElute PCR purification kit and elute in 10 µL of EB buffer twice (20 µL in total), then speed vac to 7 µL.
7. Prepare ligation reaction as follows and mix well.

Component	Amount
-----------	--------

DNA from step 6	7 μ L
2 $\times$ Quick ligase reaction buffer	10 μ L
TruSeq adapter (Illumina)	2 μ L
Quick ligase	1 μ L

8. Incubate at room temperature for 15 min.
9. Purify the DNA with a MinElute PCR purification kit and elute in 10 μ L of EB buffer twice (20 μ L in total), then speed vac to 10 μ L.
10. Run the ligation product on a 2% super fine resolution agarose gel and excise the band between 250 and 650 bp.
11. Purify DNA fragments from the gel using the MinElute gel extraction kit and elute in 10 μ L of EB buffer twice (20 μ L in total).
12. Perform bisulfite conversion using the EpiTect or Methyl-Code Bisulfite Conversion kit following the manufacturer's instructions.
13. Prepare the following PCR to amplify the bisulfite converted DNA.

Component	Amount
Bisulfite-converted DNA (from step 12)	7 μ L
Library primer-F (10 μ M)	3 μ L
Library primer-R (10 μ M)	3 μ L
2 $\times$ KAPA HiFi HotStart ReadyMix	25 μ L
Milli-Q water	To a final volume of 50 μ L

14. Run the PCR program as follows:

Initial denaturation	95 $^{\circ}$ C	1 min
Denaturation	98 $^{\circ}$ C	15 s
Annealing	60 $^{\circ}$ C	30 s
Extension	72 $^{\circ}$ C	30 s
	10 cycles	
Final extension	72 $^{\circ}$ C	1 min
Hold	4 $^{\circ}$ C	

15. Purify the DNA with a MinElute PCR purification kit and elute in 10 μ L of EB buffer twice (20 μ L in total), then speed vac to 10 μ L.
16. Run the purified library on a 2% super fine resolution agarose gel and excise the band between 250 and 650 bp.
17. Purify the final library from the gel using the MinElute gel extraction kit and elute in 10 μ L of EB buffer twice (20 μ L in total).
18. The libraries generated are sequenced using an Illumina Hi-Seq 2000 or 2500 platform according to manufacturer's instruction.

3.6.3 Data Analysis for Whole-Genome Sequencing—TAB-Seq sequencing analysis is performed following a pipeline modified from a previously published paper [37].

1. Trimming: Trim sequencing reads for low-quality reads and adaptors sequences.
2. Mapping: Align sequencing reads with the bismark software. First, prepare a bismark genome index with “bismark_genome_preparation”. Then, align the trimmed sequencing data with the following parameters: “bismark -n 1 -l 40 -chunkmbs 512 -q/yourpath/genome yourfastqdata”.
3. Calling 5hmC: For each C site in the genome, count the number of ‘C’ reads as hydroxymethylated (denoted N_C) and the number of ‘T’ bases as not hydroxymethylated (denoted N_T). To calculate the probability of observing N_C or greater cytosines by chance, use the binomial distribution with parameters N as the sequencing depth ($N_C + N_T$) and p as the 5mC nonconversion rate.

4 Notes

1. The binding buffer should be precooled on ice before use. Add fresh TCEP, PMSF, Leupeptin, and Pepstatin before use.
2. The GF running buffer should be precooled on ice before use. Add fresh DTT before use.
3. We recommended making small aliquots of Tet oxidation reagent 1 as needed. This buffer can be stored at -80°C for up to 3 months without affecting the oxidation efficiency. Avoid freeze-thaw cycles.
4. Tet oxidation reagent 2 can be prepared by mixing 2 M NaCl, 1 M HEPES (pH 8.0), 24 mM ATP, 50 mM DTT, 20 mM α -KG, 40 mM L-Ascorbic acid stock solution at 1:1:1:1:1:1 ratio. We recommended making small aliquots of Tet oxidation reagent 2 as needed. This buffer can be stored at -80°C for up to 3 months without affecting the oxidation efficiency. Avoid freeze-thaw cycles.
5. The cell pellets can be frozen with liquid nitrogen and stored at -80°C for future purification. However, fresh cells are always recommended. Since recombinant mTet1 is prone to degrade in cell lysate, make sure TCEP, PMSF, Leupeptin, and Pepstatin are added to all binding buffers used during purification.
6. Make sure PMSF, Leupeptin, and Pepstatin are added to the elution buffer before use.
7. Or repeat **step 10** until no more protein is eluted (determined by protein gel) to recover as much protein as possible.
8. Storage of recombinant mTet1 above -80°C or in less than 30% glycerol may lead to decreased oxidation efficiency. Avoid repeated freeze-thaw cycles. Each aliquot can be used up to three times.

After adding glycerol, the concentration of mTet1 is ~ 5 mg/mL.
9. The Micro Bio-Spin 30 Column is used to remove ATP in the reaction. This step may be skipped, but please note there will be ATP left in the purified DNA and will affect the concentration measure if a NanoDrop spectrophotometer is used.

10. Use the regular thermal cycling program described in the manual for the MethylCode Bisulfite Conversion kit. Alternative thermal cycling programs 1 and 2 may result in incomplete 5caC to 5caU/U conversion. If using milder bisulfite conversion conditions, an additional test of 5caC deamination efficiency is needed.
11. A different hot start polymerase may be used. Change the concentration of dNTPs, primers, polymerase, and thermal cycling program accordingly.
12. Increasing the amount of recombinant mTet1 used in **step 1** may enhance the conversion rate by up to ~1%. However, it is not recommended to add more than 7 μ L of mTet1 per 50 μ L reaction since the glycerol in the protein will affect the oxidation efficiency.

There are 22 CpGs on the amplicon.
13. Use fresh SAM for the methylation reaction.
14. Generally, over 97% of CpG sites can be methylated. The methylation percentage on CpG sites can be determined with the same procedures described in Subheading 3.2. (**Steps 5–9**, apply the methylated lambda DNA to bisulfite conversion). If the methylation rate is not good, repeat the methylation reaction and purification steps.

Lambda DNA is 48,502 base pairs in length.
15. Since there is always dCTP contamination in commercial 5hmdCTP, the percentage of 5hmC in the 5hmC control is not 100% (less than 95%). To estimate the real protection rate of 5hmC in the TAB-Seq sample, the precise percentage of 5hmC on this control needs to be determined by conventional bisulfite sequencing performed in parallel.
16. Use RNA-free genomic DNA for treatment. Include a RNA digestion step during extraction of the genomic DNA.

For whole-genome bisulfite sequencing, starting from 3 μ g is recommended but 1 μ g is also acceptable. For loci-specific analysis, 1 μ g is sufficient for screening 8–9 loci.
17. For loci-specific analysis, longer fragments can yield better results in PCR amplification after bisulfite treatment. However, unsonicated DNA is not recommended because of the tedious purification steps and low recovery rate.
18. Typically 10–20% of sheared DNA will be lost after purification.
19. If more than 300 ng of glucosylated DNA is applied, prepare several oxidation reactions in parallel.
20. After mixing Tet oxidation reagents 2 and 1, the solution should slowly change from transparent to light brown. For a large number of samples, a master solution of reagents 1 and 2 can be prepared. In that case, the color change will be much

more obvious. If no color change is observed, the oxidation reagents need to be repurchased.

21. The amount of mTet1 used here is determined by the activity test results in Subheading 3.2 and may vary from batch to batch.
22. A 5mC conversion test is strongly recommended for each different genomic sample.
23. For the reverse primer design, assume those C's on the control are 100% of 5hmC and will not be deaminated (read as C) after bisulfite treatment.
24. Increase the amount of template if it is difficult to get the PCR product.
25. Other polymerases may have a bias toward protected 5hmC (5gmC) during PCR amplification and need a separate test before using.
26. Here 80% refers to the protection rate without normalization with the real percentage of 5hmC on the control.
There are 37 cytosines on the amplicon (except for the primer part).
27. Follow the general rules for bisulfite sequencing based PCR to design the primers.
Test whether self-designed primers works well on standard bisulfite-treated genomic DNA before applying to TAB-Seq samples to save the template.
28. The annealing temperature is primer dependent and we usually start screening from 3 °C below the T_m of the primers.

Acknowledgments

This work is supported by National Institutes of Health 1R01 HG006827 to C.H.

References

1. Penn NW, Suwalski R, O'Riley C et al. (1972) The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J* 126:781–790 [PubMed: 4538516]
2. Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* 324:929–930 [PubMed: 19372393]
3. 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 [PubMed: 19372391]
4. Inoue A, Zhang Y (2011) Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. *Science* 334:194 [PubMed: 21940858]
5. He YF, Li BZ, Li Z et al. (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* 333:1303–1307 [PubMed: 21817016]
6. Cortellino S, Xu J, Sannai M et al. (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* 146:67–79 [PubMed: 21722948]
7. Guo JU, Su Y, Zhong C et al. (2011) Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* 145:423–434 [PubMed: 21496894]
8. Globisch D, Munzel M, Muller M et al. (2010) Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One* 5:e15367 [PubMed: 21203455]

9. Munzel M, Globisch D, Bruckl T et al. (2010) Quantification of the sixth DNA base hydroxymethylcytosine in the brain. *Angew Chem Int Ed Engl* 49:5375–5377 [PubMed: 20583021]
10. Szwagierczak A, Bultmann S, Schmidt CS et al. (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* 38:e181 [PubMed: 20685817]
11. 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 [PubMed: 21151123]
12. Wu H, D'Alessio AC, Ito S et al. (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev* 25:679–684 [PubMed: 21460036]
13. Ito S, D'Alessio AC, Taranova OV et al. (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* 466:1129–1133 [PubMed: 20639862]
14. Dawlaty MM, Ganz K, Powell BE et al. (2011) Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* 9:166–175 [PubMed: 21816367]
15. Gu TP, Guo F, Yang H et al. (2011) The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* 477:606–610 [PubMed: 21892189]
16. Iqbal K, Jin SG, Pfeifer GP et al. (2011) Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U S A* 108:3642–3647 [PubMed: 21321204]
17. Koh KP, Yabuuchi A, Rao S et al. (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* 8:200–213 [PubMed: 21295276]
18. Szulwach KE, Li X, Li Y et al. (2011) Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet* 7:e1002154 [PubMed: 21731508]
19. Szulwach KE, Li X, Li Y et al. (2011) 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat Neurosci* 14:1607–1616 [PubMed: 22037496]
20. Kriukiene E, Liutkeviciute Z, Klimasauskas S (2012) 5-Hydroxymethylcytosine—the elusive epigenetic mark in mammalian DNA. *Chem Soc Rev* 41:6916–6930 [PubMed: 22842880]
21. 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 [PubMed: 21778364]
22. Pfaffeneder T, Hackner B, Truss M et al. (2011) The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew Chem Int Ed Engl* 50:7008–7012 [PubMed: 21721093]
23. Maiti A, Drohat AC (2011) Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem* 286:35334–35338 [PubMed: 21862836]
24. Zhang L, Lu X, Lu J et al. (2012) Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nat Chem Biol* 8:328–330 [PubMed: 22327402]
25. Ficiz G, Branco MR, Seisenberger S et al. (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* 473:398–402 [PubMed: 21460836]
26. Williams K, Christensen J, Pedersen MT et al. (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* 473:343–348 [PubMed: 21490601]
27. Xu Y, Wu F, Tan L et al. (2011) Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* 42:451–464 [PubMed: 21514197]
28. Pastor WA, Pape UJ, Huang Y et al. (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473:394–397 [PubMed: 21552279]
29. Pastor WA, Huang Y, Henderson HR et al. (2012) The GLIB technique for genome-wide mapping of 5-hydroxymethylcytosine. *Nat Protoc* 7:1909–1917 [PubMed: 23018194]
30. Huang Y, Pastor WA, Zepeda-Martinez JA et al. (2012) The anti-CMS technique for genomewide mapping of 5-hydroxymethylcytosine. *Nat Protoc* 7:1897–1908 [PubMed: 23018193]
31. Robertson AB, Dahl JA, Vagbo CB et al. (2011) A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* 39:e55 [PubMed: 21300643]

32. Cokus SJ, Feng S, Zhang X et al. (2008) Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature* 452:215–219 [PubMed: 18278030]
33. 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 [PubMed: 18423832]
34. Lister R, Pelizzola M, Dowen RH et al. (2009) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* 462:315–322 [PubMed: 19829295]
35. Jin SG, Kadam S, Pfeifer GP (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res* 38:e125 [PubMed: 20371518]
36. Huang Y, Pastor WA, Shen Y et al. (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* 5:e8888 [PubMed: 20126651]
37. Yu M, Hon GC, Szulwach KE et al. (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* 149:1368–1380 [PubMed: 22608086]
38. Josse J, Kornberg A (1962) Glucosylation of deoxyribonucleic acid. III. alpha- and beta-Glucosyl transferases from T4-infected Escherichia coli. *J Biol Chem* 237:1968–1976 [PubMed: 14452558]
39. Krueger F, Andrews SR (2011) Bismark: a flexible aligner and methylation caller for bisulfite-Seq applications. *Bioinformatics* 27:1571–1572 [PubMed: 21493656]

Fig. 1. Scheme of (a) traditional Bisulfite Sequencing and (b) Tet-Assisted Bisulfite Sequencing. CMS cytosine 5-methylenesulfonate, *5caC* 5-carboxylcytosine, *5gmC* β -glucosyl-5-hydroxymethylcytosine