Published in final edited form as:

Methods Mol Biol. 2015 ; 1238: 273–287. doi:10.1007/978-1-4939-1804-1_14.

Single base resolution analysis of 5-methylcytosine and 5-hydroxymethylcytosine by RRBS and TAB-RRBS

Maria A. Hahn¹, Arthur X. Li², Xiwei Wu³, and Gerd P. Pfeifer¹

¹Department of Cancer Biology, Beckman Research Institute, City of Hope, Duarte, CA, USA

²Department of Information Sciences, Beckman Research Institute, City of Hope, Duarte, CA, USA

³Department of Molecular and Cellular Biology, Beckman Research Institute, City of Hope, Duarte, CA, USA

Summary

Sodium bisulfite-assisted deamination of cytosine forms the basis for conducting single base resolution analysis of 5-methylcytosine in DNA. The TET family of proteins represents a group of enzymes that can oxidize 5-methylcytosine to 5-hydroxymethylcytosine. A modification of the bisulfite-based DNA methylation mapping technique employs TET1-mediated oxidation of 5-methylcytosine (TET-assisted bisulfite sequencing) for single base analysis of 5-hydroxymethylcytosine. Whole genome analysis of cytosine modifications with bisulfite sequencing techniques still is challenging and expensive. Reduced representation bisulfite sequencing (RRBS) has been used to limit the complexity of the analysis to mostly CpG-rich genomic fragments flanked by restriction enzyme cleavage sites, for example MspI (5'CCGG). In this chapter, we describe detailed methods used in our laboratory for analysis of 5-methylcytosine and 5-hydroxymethylcytosine combined (RRBS) and for specific analysis of 5-hydroxymethylcytosine (TAB-RRBS).

Keywords

DNA methylation; 5-methylcytosine; 5-hydroxymethylcytosine; CpG islands; high throughput sequencing

1. Introduction

Epigenetic changes such as the establishment of cytosine modification patterns play a crucial role in development and cancer. Recently, in addition to 5-methylcytosine (5mC), a new cytosine modification, 5-hydroxymethylcytosine (5hmC) was discovered (1, 2). This base is thought to be an intermediate formed during DNA demethylation processes (3). For example, loss of 5mC and gain of 5hmC occurs in the paternal genome immediately after

Correspondence: Dr. Gerd P. Pfeifer or Dr. Maria A. Hahn, Beckman Research Institute, City of Hope, Duarte, CA 91010, USA, Phone: 626-301-8853, gpfeifer@coh.org or mhahn@coh.org.

Conflict of Interest Statement:

G.P.P. is a paid consultant of Zymo Research; Irvine, CA.

fertilization (4, 5). The Tet protein family of 5mC dioxygenases, consisting of the mammalian members Tet1, Tet2 and Tet3, is responsible for the oxidation 5mC to 5hmC (2, 6). Recent studies suggest that 5hmC, due to its relatively high abundance, is not only a transient intermediate formed when 5mC is converted to unmodified cytosine but that it rather plays its own specific role in genome regulation. 5hmC is particularly abundant in neuronal cells of the brain (1, 7). According to recent studies, 5hmC is viewed as a stable base modification and substitutes a substantial portion of 5mC in the genome during neuronal differentiation (8) and during the aging process (9). Similar to 5mC, changes in 5hmC are strongly linked to cancer development. The level of 5hmC is dramatically reduced in many cancer types compared to corresponding normal tissue (10). It has been suggested that the loss of 5hmC in cancer is associated with DNA hypermethylation. In order to understand the role of 5hmC in cancer and during development and cell differentiation, this epigenetic mark needs to be mapped in mammalian tissues at high resolution. However, there are technical and cost limitations that have hampered these approaches. We can compare three major types of 5hmC detection methodologies for genome-wide applications: (i) methods which are based on restriction enzymes which, under certain conditions, are able to distinguish 5hmC from 5mC (11), (ii) methods based on oxidation of modified cytosines (12, 13) and (iii) methods based on enrichment of 5hmC or modified 5hmC (7, 14, 15). 5hmC mapping in embryonic stem cells revealed that 5hmC is frequently found asymmetrical at the CpG sites of complementary DNA strands, which is in contrast to 5mC (12). This fact suggests that some 5hmC detection methodologies based on restriction enzyme usage may provide limited information since efficient differential DNA digestion with enzymes such as MspI and HpaII requires presence of 5hmC on both DNA strands. At the same time, methods based on affinity capture of 5hmC or modified 5hmC are dependent on 5hmC density and in addition do not provide single DNA base resolution. In contrast to those approaches, methods based on modified cytosine oxidation provide single base resolution and are very sensitive. There are two types of 5hmC profiling methods involving oxidation: chemical oxidation of 5hmC, the so-called oxidative bisulfite sequencing (oxBS-Seq) (13) and 5mC oxidation by Tet1 protein, the so-called Tet-assisted bisulfite sequencing (TAB) (12). Whole genome TAB sequencing of 5hmC requires high coverage due to the low frequency of 5hmC in many tissues and is associated with very high cost. In contrast to whole genome sequencing based on bisulfite conversion, the RRBS technique, which focuses on specific compartments of the genome, notably CpG islands, allows significant cost reduction due to sequencing of only a part of the genome (16, 17). The combination of 5mC oxidation approaches with the RRBS method results in a high number of reads per single cytosine at reasonable cost (12).

The RRBS methodology was developed in Rudolf Jaenisch's laboratory and became popular due to its cost effectiveness and a relative high coverage of CpG sites in the genome (16). This method became particularly applicable for samples with limited cell numbers such as germ cells or oocytes (18). The method is based on DNA digestion with MspI (5'CCGG recognition site) followed by end repair, A-tailing, linker ligation, DNA bisulfite conversion and library amplification (Fig. 1A). This methodology provides information on DNA fragments located between two MspI sites with distance of up to 350 bp. According to our data, the method can represent approximately 82% of the CpG islands and approximately

70% of the promoters in the human genome. It is important to note, however, that standard sodium bisulfite treatment does not distinguish 5mC from 5hmC (19, 20).

The TAB sequencing method was developed in Chuan He's laboratory (12). This technique is based on the oxidation properties of the Tet1 protein, which is able to oxidize 5mC, 5hmC and 5-formylcytosine (5fC) to 5-carboxylcytosine (5caC), and on inhibition of Tet1 oxidation by glycosylation of 5hmC (Fig. 1B). The 5caC formed by Tet1 oxidation is then decarboxylated to cytosine and deaminated in a solution of sodium bisulfite, and therefore can be read as cytosine in the sequencing data. The 5mC to C conversion rate by Tet1 oxidation and bisulfite treatment can be above 96% (12). The protocol includes a glycosylation protection step of 5hmC by T4 phage β -glucosyltransferase, which prevents further oxidation of 5hmC by the Tet1 protein fragment. As an end result, unmodified cytosine, 5mC, 5fC, and 5caC in the genome are all deaminated and read as T, whereas only 5hmC will be read as C. The method is outlined in Fig. 1. A 5hmC TAB-Seq Kit (Wisegene; Chicago, IL) has become available for general use. This kit provides researchers with high quality Tet1 protein fragment and eliminates difficulties of Tet1 purification in the lab.

In our laboratory, we combined and optimized the TAB and RBBS techniques, which resulted in increased 5mC to C conversion efficiency of over 99% and a dramatic decrease of cost for 5hmC profiling. This approach allowed us to reach a 20 read coverage on average for each analyzed cytosine located on each separate DNA strand by using approximately 50–60 million paired-end sequence reads. The combined protocol consists of MspI digestion, DNA glycosylation, two consecutive rounds of oxidation by Tet1 fragment, end repair, A-tailing, linker ligation, DNA bisulfite conversion and library amplification (Fig. 1C).

Several steps were done in order to achieve the highest conversion rate of 5mC and to keep high library diversity. We use the EZ DNA Methylation-Gold Kit (Zymo Research; Irvine CA), which leaves no more than 0.5% unconverted cytosines in the CHH context after regular RRBS. We used two rounds of Tet1 oxidation in order to achieve the highest oxidation rate. In order to evaluate Tet1 oxidation efficiency we used non-glycosylated control samples for each oxidation batch. Absence of glycosylation leads to de-protection of 5hmC from Tet1 oxidation. Theoretically, all modified cytosines should be converted after Tet1 oxidation in absence of glycosylation. By using LC-MS/MS quantification of 5mdC in genomic DNA, we identified a sample with the highest 5mC level in the genome and performed regular TAB-RRBS and non-glycosylated TAB-RRBS with this sample. Our analysis of the non-glycosylated control revealed that Tet1 protein is inefficient at the ends of MspI restriction fragments (Fig. 2). The data indicate that the first six DNA bases next to the MspI site after digestion have a low conversion rate. The lowest Tet1 oxidation rate, only 81%, was observed at the third DNA base next to the MspI site. This fact indicates a potential weakness of the Tet1 oxidation methodology in whole genome 5hmC mapping since Tet1 may oxidize poorly also DNA fragment ends after sonication, which may cause a decreased 5mC conversion rate (thus erroneously scoring as 5hmC) in samples. However, the RRBS approach allows us to exclude the first 6 bases located next to MspI sites in the genome. This is reflected in very high oxidation rates of 5mC, more than 99%, when moving more than 6 bases away from the MspI sites (Fig. 2). If we take into account that

bisulfite conversion efficiency is approximately 99.6% then the number of unconverted cytosines remaining after Tet1 oxidation is extremely low, less than 0.5% [Number of unconverted cytosines in the CHH context after regular RRBS is ~0.4% to 0.5% (bisulfite conversion efficiency) and the number of unconverted cytosines in the CG context is ~0.9% in the nonglycosylated control. 0.9%–0.4%≈0.5%]. Further inspection of the data indicated very few specific genomic DNA regions, which have a poor Tet1 oxidation rate. In order to eliminate this potential artifact, we subtracted levels of unconverted cytosines in non-glycosylated control from the level of unconverted cytosines for each analyzed cytosine. Incorporation of all steps in the TAB-RRBS protocol resulted in a >99% oxidation rate of 5mC. An example of final profiling of 5hmC in colon cancer specimens is shown in Fig. 3. Our protocol provides an opportunity for cost-efficient profiling of 5hmC in DNA samples with relatively low 5hmC levels (~0.1% of all cytosines are 5hmC). Here we describe our protocol for combination RRBS mapping of 5mC plus 5hmC and for TAB-RRBS mapping specific for 5hmC.

2. Materials

- 2.1 500 ng double-stranded genomic DNA.
- 2.2 MspI 100 units/μl.
- 2.3 TE buffer: 10 mM Tris-Cl, pH 8.1, 1 mM EDTA.
- 2.4 Klenow Fragment (3'→5' exo-).
- 2.5 Deoxynucleoside triphosphates.
- 2.6 20 mg/ml glycogen from mussels.
- 2.7 Phenol/chloroform/isoamyl alcohol (25:24:1 mixture, pH 6.7/8.0).
- 2.8 Chloroform.
- 2.9 Sodium acetate (NaOAc), 3 M solution, pH 5.5.
- 2.10 T4 DNA ligase.
- 2.11 Agencourt AMPure XP beads (Beckman Coulter Inc.; Fullerton, CA)
- 2.12 EZ DNA Methylation-Gold kit (Zymo Research; Irvine, CA).
- 2.13 MinElute PCR purification kit (Qiagen; Valencia, CA).
- 2.14 PfuTurbo Cx Hotstart DNA polymerase (Agilent; Santa Clara, CA; supplied with 10x reaction buffer).
- 2.15 Novex® TBE gel, 6% (Life Technologies).
- 2.16 Novex® Hi-Density TBE sample buffer (5X) (Life Technologies).
- 2.17 SYBR® Safe DNA gel stain (Life Technologies).
- 2.18 Amicon Ultra-0.5, Ultracel-30 membrane, 30 kDa (Millipore; Billerica, MA).
- 2.19 5hmC TAB-Seq kit (Wisegene; Chicago, IL, cat. no. K001).

2.20 5hmC Tet1 Oxidation kit (Wisegene, cat. no. K003).

3. Methods

3.1. RRBS protocol for mapping of 5mC plus 5hmC

3.1.1. MspI digestion—Set up a 30 µl digestion reaction as follows:

	Amount	Final concentration
10x reaction buffer	3 µl	1x
MspI 100 u/µl	1 µl	3.3 u/µl
DNA, 500 ng in TE buffer	26 µl	16.7 ng/µl

Mix the sample. Incubate overnight at 37°C.

3.1.2. End repair and A-tailing

1. Set up an end-repair reaction as follows:

	Amount
DNA from 3.1.1	30 µl
dNTPs (10 mM dATP, 1mM dCTP, 1mM dGTP)	1 µl
Klenow Fragment (3'→5' exo-) 5 u/µl	1 µl

2. Incubate 20 min at 30°C.
3. Incubate 20 min at 37°C.
4. Perform phenol/chloroform purification followed by ethanol precipitation as follows:
5. Add 370 µl of 1x TE buffer. Add 400 µl phenol/chloroform. Mix for 3 min.
6. Centrifuge for 3 min at 14,000×g in an Eppendorf centrifuge at room temperature. Transfer supernatant to a new tube.
7. Add 400 µl chloroform to the supernatant. Mix for 3 min.
8. Centrifuge for 3 min at 14000×g at room temperature. Transfer supernatant into new tube.
9. Add 0.5 µl 20 mg/ml glycogen and 40 µl of 3 M sodium acetate pH 5.5 to the supernatant. Vortex for 10 sec.
10. Add 1.1 ml of 96% ethanol. Mix sample. Incubate 15 min at −70°C.
11. Spin down the DNA at 14000×g for 20 min at room temperature.
12. Wash pellet twice with 700 µl 75% ethanol.
13. Remove residual ethanol. Dry DNA pellet.

14. Add 12 µl TE buffer.

3.1.3. Linker ligation

1. Set up a ligation reaction as follows:

	Amount	Final Concentration
DNA from 3.1.2.	12 µl	
10x T4 ligase reaction buffer	1.5 µl	1x
7.5 µM linker (Illumina)	0.5 µl	0.25 µM
T4 ligase 400 u/µl	1 µl	26.7 u/µl

2. Incubate overnight at 16°C.
3. Incubate 10 min at 65°C.
4. Add 37.5 µl Agencourt AMPure XP beads. Mix well.
5. Incubate for 15 min at room temperature.
6. By using a magnet, separate beads and remove all solution.
7. Wash beads by using a magnetic rack. Add 150 µl of 80% ethanol. Incubate 1 min at room temperature. Remove ethanol.
8. Repeat step 7.
9. Dry beads.
10. Add 20 µl TE buffer. Vortex. Incubate for 2 min at room temperature. Transfer DNA into a new tube.

3.1.4. Bisulfite conversion

1. To DNA from 3.1.3, add 130 µl of freshly prepared CT- mix (from EZ DNA Methylation-Gold Kit).
2. Incubate at 98°C for 10 min, then 64°C for 2.5 hours, then put on hold at 4°C (maximum for 24 hours).
3. Convert DNA according to the manufacturer's instructions (EZ DNA Methylation-Gold Kit).
4. Elute with 20 µl elution buffer (EZ DNA Methylation-Gold Kit).
5. Add 100 µl of PB buffer (MinElute PCR Purification Kit) and purify DNA by using MinElute columns according to the manufacturer's instructions.
6. Elute DNA with 20 µl EB buffer (MinElute PCR Purification Kit).

3.1.5. Library amplification

1. Set up a PCR reaction as follows:

	Amount	Final Concentration
DNA from 3.1.4	20 µl	
10x PfuTurbo Cx Hotstart DNA polymerase reaction buffer	20 µl	1x
10 µM primers R + F (Illumina)	20 µl	1 µM
2.5 mM dNTPs	20 µl	0.25 mM
PfuTurbo Cx Hotstart DNA Pol. 2.5u/µl	4 µl	0.05 u/µl
Water	116 µl	

Primer F: 5'CAAGCAGAAGACGGCATACG

Primer R: 5'AATGATACGGCGACCACCGA

- Mix well. Amplify samples by using the following conditions:

95°C	5 min	
98°C	30 sec	
98°C	10 sec	}
65°C	30 sec	
72°C	30 sec	
		8–9 cycles
Hold at 4°C		

- Purify PCR products by using MinElute PCR purification kit according to the manufacturer's instructions and elute with 10 µl EB buffer.

3.1.6. Gel purification

- Add 2.5 µl Novex® Hi-Density TBE sample buffer (5X) to the samples from 3.1.5. Mix well.
- Load samples into wells of a Novex® TBE Gel, 6%.
- Run gel for 40 min at 200 V.
- Dye the gel with SYBR® Safe DNA gel stain according to the manufacturer's instructions.
- Under UV light, cut DNA fragments ranging from 150 to 500 bp (insert size range should be from 30–380 bp).
- Add 500 µl of EB (MinElute PCR Purification Kit). Smash gel in buffer.
- Shake samples overnight at room temperature.
- Concentrate the eluted DNA with an Amicon Ultra-0.5, Ultracel-30 membrane according to the manufacturer's instructions.
- Quantify library with QPCR.

3.1.7. Illumina sequencing—Perform Illumina paired-end sequencing.

3.2. TAB-RRBS protocol for mapping of 5hmC

3.2.1. MspI digestion

1. Set up a digestion reaction as follows:

	Amount	Final Concentration
10x NEBuffer 2	3 µl	1x
MspI 100 u/µl	1 µl	3.3 u/µl
500 ng DNA in TE buffer	26 µl	16.7 ng/µl

2. Mix the sample. Incubate overnight at 37°C.
3. Perform phenol/chloroform purification followed by ethanol precipitation in presence of glycogen as follows:
4. Add 370 µl of TE buffer. Add 400 µl phenol/chloroform. Mix for 3 min.
5. Centrifuge for 3 min at 14000×g at room temperature. Transfer supernatant into a new tube.
6. Add 400 µl chloroform to the supernatant. Mix for 3 min.
7. Centrifuge for 3 min at 14000×g at room temperature. Transfer supernatant into a new tube.
8. Add 0.5 µl 20 mg/ml glycogen and 40 µl 3 M sodium acetate pH 5.5 to the supernatant. Vortex for 10 sec.
9. Add 1.1 ml of 96% ethanol. Mix the sample. Incubate for 15 min at –70°C.
10. Collect the DNA by centrifugation at 14000×g for 20 min at room temperature.
11. Wash pellet twice with 700 µl 75% ethanol.
12. Remove residual ethanol. Dry DNA pellet.
13. Add 3 µl TE buffer.

3.2.2. Glycosylation

1. Set up a glycosylation reaction as follows (Please see Note 1):

	Amount	Final Concentration
Genomic DNA from 3.2.1	3 µl	25 ng/µl
10 mM UDP-Glucose (Wisegene, K001)	1 µl	500 µM
10 × βGT protection buffer (Wisegene, K001)	2 µl	1 x

Note 1 In order to evaluate the 5mC conversion rate, we include a non-glycosylated control sample in every sample batch. This sample should have the highest DNA 5mC level among all oxidized samples in the batch. In contrast to the rest of the samples, the non-glycosylated control does not undergo the glycosylation step (section 3.2.2).

	Amount	Final Concentration
T4-βGT (40 μM) (Wisegene, K001)	0.5 μl	1 μM
Water	13.5 μl	

2. Mix the sample. Incubate for 1 hour at 37°C.
3. Perform phenol/chloroform purification followed by ethanol precipitation in presence of glycogen as described under 3.2.1.

3.2.3. First oxidation

1. Set up an oxidation reaction as follows:

	Amount
Genomic DNA from 3.2.2	3 μl
Tet1 oxidation reagent 1 (Wisegene, K001)	3.5 μl
Tet1 oxidation reagent 2 (Wisegene, K001)	15 μl
Tet1 protein (Wisegene, K001)	9 μl*
Water	19.5 μl

* see Note 2.

2. Incubate 1 hour at 37°C.
3. Add 1 μl 20 mg/ml proteinase K. Mix the sample. Incubate for 10 min at 50°C.
4. Perform phenol/chloroform purification followed by ethanol precipitation in presence of glycogen as described in 3.2.1.

3.2.4. Second oxidation

1. Set up an oxidation reaction as follows (see Note 3):

	Amount
Genomic DNA from 3.2.3.	3 μl
Tet1 oxidation reagent 1, (Wisegene, K001)	3.5 μl
Tet1 oxidation reagent 2, (Wisegene, K001)	15 μl
Tet1 protein (Wisegene, K001)	9 μl
Water	19.5 μl

2. Incubate 1 hour at 37°C.
3. Add 1 μl proteinase K (20 mg/ml). Mix the sample. Incubate 10 min at 50°C.

Note 3 The second oxidation step is necessary to convert the leftover unconverted 5mC. This step is required to achieve a 5mC conversion efficiency of over 99%.

4. Perform phenol/chloroform purification followed by ethanol precipitation in presence of glycogen as described under 3.2.1. After ethanol precipitation, dissolve the DNA in 18 µl of TE buffer.

3.2.5. End repair and A-tailing

1. Set up an end-repair reaction as follows:

	Amount
DNA from 3.2.4	18 µl
10x NEBuffer 2	3 µl
dNTPs (10mM dATP, 1mM dCTP, 1mM dGTP)	1 µl
Klenow Fragment (3'→5' exo-)	1 µl
Water	9 µl

2. Incubate 20 min at 30°C.
3. Incubate 20 min at 37°C.
4. Perform phenol/chloroform purification followed by ethanol precipitation as described under 3.2.1. After ethanol precipitation, dissolve the DNA in 12 µl of TE buffer.

3.2.6. Linker ligation

1. Set up a ligation reaction as follows:

	Amount	Final Concentration
DNA from 3.2.5	12 µl	
10X T4 ligase reaction buffer	1.5 µl	1x
7.5 µM linker (Illumina)	0.5 µl	0.25 µM
T4 ligase 400 u/µl	1 µl	26.7 u/µl

2. Incubate overnight at 16°C.
3. Incubate 10 min at 65°C.
4. Add 37.5 µl Agencourt AMPure XP beads. Mix well.
5. Incubate for 15 min at room temperature.
6. By using a magnet, separate beads and remove all solution.
7. Wash beads by using a magnetic rack. Add 150 µl of 80% ethanol. Incubate 1 min at room temperature. Remove ethanol.
8. Repeat step 7.
9. Dry beads.

10. Add 20 µl TE buffer. Vortex. Incubate for 2 min at room temperature. Transfer DNA into a new tube.

3.2.7. Bisulfite conversion

1. To DNA from 3.2.6 add 130 µl of freshly prepared CT- mix (from EZ DNA Methylation-Gold Kit)
2. Incubate at 98°C for 10 min, then at 64°C for 2.5 hours and hold at 4°C.
3. Convert DNA according to the manufacturer's instructions (EZ DNA Methylation-Gold kit).
4. Elute with 20 µl elution buffer (EZ DNA Methylation-Gold kit)
5. Add 100 µl of PB buffer (MinElute PCR purification kit) and purify DNA by using MinElute column according to the manufacturer's instructions.
6. Elute DNA with 20 µl EB buffer (MinElute PCR purification kit).

3.2.8. Library amplification

1. Set up a PCR reaction as follows:

	Amount	Final Concentration
DNA from 3.2.7	20 µl	
10x PfuTurbo Cx Hotstart DNA pol. reaction buffer	20 µl	1x
10 µM Primer R + F	20 µl	1 µM
2.5 mM dNTPs	20 µl	0.25 mM
PfuTurbo Cx Hotstart DNA polymerase 2.5 u/µl	4 µl	0.05 u/µl
Water	116 µl	

2. Mix well. Amplify samples by using the following conditions:

95°C	5min	
98°C	30sec	
98°C	10 sec	}
65°C	30 sec	
72°C	30 sec	
		10–12 cycles
Hold at 4°C		

3. Purify PCR products by using MinElute PCR purification kit according to the manufacturer's instructions and elute with 10 µl EB buffer.

3.2.9. Gel purification

1. Add 2.5 µl Novex® Hi-Density TBE sample buffer (5X) to the samples. Mix well.
2. Load samples into the wells of a Novex® TBE Gel, 6%.
3. Run gel for 40 min at 200 V.
4. Dye the gel with SYBR® Safe DNA gel stain according to the manufacturer's instructions.
5. Under UV light, cut DNA fragments ranging from 150 to 500 bp (insert size range should be from 30–380 bp).
6. Add 500 µl of EB (MinElute PCR purification kit). Smash gel in buffer.
7. Shake samples overnight at room temperature.
8. Concentrate the eluted DNA with an Amicon Ultra-0.5, Ultracel-30 membrane according to manufacturer's instructions.
9. Quantify library with QPCR.

3.2.10. Illumina sequencing—Perform Illumina paired-end sequencing.

Acknowledgments

This work was supported by NIH grants CA084469, CA160965 and MH094599 (recipient G.P.P.).

References

1. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009; 324:929–930. [PubMed: 19372393]
2. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009; 324:930–935. [PubMed: 19372391]
3. Song CX, He C. Potential functional roles of DNA demethylation intermediates. *Trends Biochem Sci*. 2013; 38:480–484. [PubMed: 23932479]
4. Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG, Iqbal K, Shi YG, Deng Z, Szabo PE, Pfeifer GP, Li J, Xu GL. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*. 2011; 477:606–610. [PubMed: 21892189]
5. Iqbal K, Jin SG, Pfeifer GP, Szabo PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci U S A*. 2011; 108:3642–3647. [PubMed: 21321204]
6. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010; 466:1129–1133. [PubMed: 20639862]
7. Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen CH, Zhang W, Jian X, Wang J, Zhang L, Looney TJ, Zhang B, Godley LA, Hicks LM, Lahn BT, Jin P, He C. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*. 2011; 29:68–72. [PubMed: 21151123]
8. Hahn MA, Qiu R, Wu X, Li AX, Zhang H, Wang J, Jui J, Jin SG, Jiang Y, Pfeifer GP, Lu Q. Dynamics of 5-Hydroxymethylcytosine and Chromatin Marks in Mammalian Neurogenesis. *Cell Rep*. 2013
9. Szulwach KE, Li X, Li Y, Song CX, Wu H, Dai Q, Irier H, Upadhyay AK, Gearing M, Levey AI, Vasanthakumar A, Godley LA, Chang Q, Cheng X, He C, Jin P. 5-hmC-mediated epigenetic

- dynamics during postnatal neurodevelopment and aging. *Nat Neurosci.* 2011; 14:1607–1616. [PubMed: 22037496]
10. Jin SG, Jiang Y, Qiu R, Rauch TA, Wang Y, Schackert G, Krex D, Lu Q, Pfeifer GP. 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res.* 2011; 71:7360–7365. [PubMed: 22052461]
 11. Kinney SM, Chin HG, Vaisvila R, Bitinaite J, Zheng Y, Esteve PO, Feng S, Stroud H, Jacobsen SE, Pradhan S. Tissue-specific distribution and dynamic changes of 5-hydroxymethylcytosine in mammalian genomes. *J Biol Chem.* 2011; 286:24685–24693. [PubMed: 21610077]
 12. Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, Min JH, Jin P, Ren B, He C. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell.* 2012; 149:1368–1380. [PubMed: 22608086]
 13. Booth MJ, Ost TW, Beraldi D, Bell NM, Branco MR, Reik W, Balasubramanian S. Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. *Nat Protoc.* 2013; 8:1841–1851. [PubMed: 24008380]
 14. Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P, Tahiliani M, Daley GQ, Liu XS, Ecker JR, Milos PM, Agarwal S, Rao A. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature.* 2011; 473:394–397. [PubMed: 21552279]
 15. Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res.* 2011; 39:5015–5024. [PubMed: 21378125]
 16. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* 2005; 33:5868–5877. [PubMed: 16224102]
 17. Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat Protoc.* 2011; 6:468–481. [PubMed: 21412275]
 18. Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Segonds-Pichon A, Sato S, Hata K, Andrews SR, Kelsey G. Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nat Genet.* 2011; 43:811–814. [PubMed: 21706000]
 19. Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One.* 2010; 5:e8888. [PubMed: 20126651]
 20. Jin SG, Kadam S, Pfeifer GP. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res.* 2010; 38:e125. [PubMed: 20371518]

Figure 1. Flow charts of RRBS, whole genome TAB sequencing and TAB-RRBS procedures
A. RRBS library preparation. **B.** TAB library preparation. **C.** TAB-RRBS procedure as a combination of TAB and RRBS techniques.

Figure 2. The catalytic Tet1 protein fragment is inefficient at DNA ends generated by MspI digestion

A composite profile for levels of unconverted cytosine around MspI restriction sites was generated by calculating a mean of percentage of unconverted cytosines in both directions from the MspI sites.

Figure 3. Representative snap-shots of TAB-RRBS and RRBS results for colon tumors and adjacent nonmalignant tissues

The 5hmC profiles and 5mC (including 5hmC) profiles of a segment of chromosome 3 in two colon tumors (6T and 8T) and two adjacent nonmalignant colon tissues (6N and 8N) were generated by TAB-RRBS and RRBS, respectively. The analyzed ranges of unconverted cytosines remaining after the procedures are indicated on the left side of the profiles. Data are shown only for those CpG sites where the number of reads for each cytosine in the CpG context was at least 10 in all RRBS samples and at least 20 in all TAB-RRBS samples. Please note the decline of 5hmC in the tumors. Genes are indicated at the bottom.