MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing

Mark A Urich[1], Joseph R Nery[1], Ryan Lister[2], Robert J Schmitz[3] & Joseph R Ecker[1,4]

[1]Genomic Analysis Laboratory, The Salk Institute for Biological Studies, La Jolla, California, USA. [2]Australian Research Council Center of Excellence in Plant Energy Biology, The University of Western Australia, Perth, Australia. [3]Department of Genetics, University of Georgia, Athens, Georgia, USA. [4]Howard Hughes Medical Institute, The Salk Institute for Biological Studies, La Jolla, California, USA. Correspondence should be addressed to J.R.E. (ecker@salk.edu).

Published online 18 February 2015; doi:10.1038/nprot.2014.114

Current high-throughput DNA sequencing technologies enable acquisition of billions of data points through which myriad biological processes can be interrogated, including genetic variation, chromatin structure, gene expression patterns, small RNAs and protein–DNA interactions. Here we describe the MethylC-sequencing (MethylC-seq) library preparation method, a 2-d protocol that enables the genome-wide identification of cytosine DNA methylation states at single-base resolution. The technique involves fragmentation of genomic DNA followed by adapter ligation, bisulfite conversion and limited amplification using adapter-specific PCR primers in preparation for sequencing. To date, this protocol has been successfully applied to genomic DNA isolated from primary cell culture, sorted cells and fresh tissue from over a thousand plant and animal samples.

INTRODUCTION

Cytosine DNA methylation is a covalent base modification that can be stably transmitted through mitotic and meiotic cell divisions[1–3]. DNA methylation has the capacity to alter proximal chromatin structure and transcriptional activity of the genome, depending on the location and sequence context of the methylated base. Base-resolution determination of methylation status is important for understanding the cellular pathways by which the genome modification is established and maintained. In plant cells, multiple molecular pathways mediate the methylation of cytosines in distinct sequence contexts (CG, CHG, CHH, where H = A, C, T)[4]. In most mammalian cell types profiled to date, the vast majority of DNA methylation is present in the CG context[5]. However, base-resolution studies have identified widespread DNA methylation in the CH context in mammalian pluripotent cells and in the brain, particularly in neurons[6–9]. In plant genomes, genic CG methylation is associated with constitutively expressed loci[10,11], whereas regions of the genome targeted by CG and non-CG methylation are under active silencing by the RNA-directed DNA methylation pathway[4,12].

The gold-standard method for determining DNA methylation states of individual cytosines is to combine sodium bisulfite conversion with PCR and Sanger sequencing[13–18]. Treatment of genomic DNA with sodium bisulfite converts unmethylated cytosine into uracil, which is subsequently converted to thymine during PCR. Cytosines present in the bisulfite-converted sequences indicate that the cytosine in the original fragment of genomic DNA was methylated, as both 5-methylcytosine and 5-hydroxymethylcytosine are protected from this conversion reaction. Through accumulation of sufficient genomic sequence coverage, this method can also enable quantification of the aggregate level of DNA methylation at each covered position in the population of genomes sampled. Although this approach has been a cornerstone for studying DNA methylation states of individual loci, it requires primer design that often introduces biases, it is limited to surveying a few loci from each bisulfite-treated sample and it is of low throughput.

Overview of MethlyC-seq

To survey the methylation states of cytosines at single-base resolution on a genome-wide scale, we developed a whole-genome bisulfite sequencing approach called MethylC-seq (**Fig. 1**). This method uses high-throughput DNA sequencing of genomic DNA subjected to sodium bisulfite conversion[15–18]. After deep sequencing of a library generated from fragments of sodium bisulfite–treated DNA, the basecall at each cytosine reference position indicates the original methylation status of the cytosine in each genomic DNA (gDNA) fragment, where a thymine indicates that it was unmethylated and a cytosine indicates that it was methylated. The frequency of DNA methylation at any cytosine with sufficient sequence coverage can be estimated for the population of genomes that comprised the genomic DNA sample. Typical methylomes generated by MethylC-seq for mouse, human and *Arabidopsis* genomes achieve coverage of >90–95% of the cytosines in the genome[6,19]. This protocol is largely framed around standard protocols designed to construct DNA sequencing libraries, but substantial modifications have been made such as eliminating all electrophoresis and gel extraction steps, adding the sodium bisulfite conversion reaction and making modifications to the number of PCR cycles. Briefly, purified genomic DNA (50 ng–2 μg) is fragmented, end repaired, 3′-adenylated and ligated to sequencing adapters in which all cytosines are methylated. Adapter-ligated DNA is then subjected to bisulfite conversion, after which limited amplification of the library is performed by PCR using primers specific for the sequencing adapters. The resulting library is then ready for sequencing after library quantification.

Advantages and applications of MethylC-seq

By using cytosine-methylated universal adapters, the need for targeted primer design is eliminated, thereby eliminating associated biases. This protocol is also capable of surveying >90% of cytosines in most genomes studied to date, which is far more comprehensive than reduced-representation bisulfite sequencing

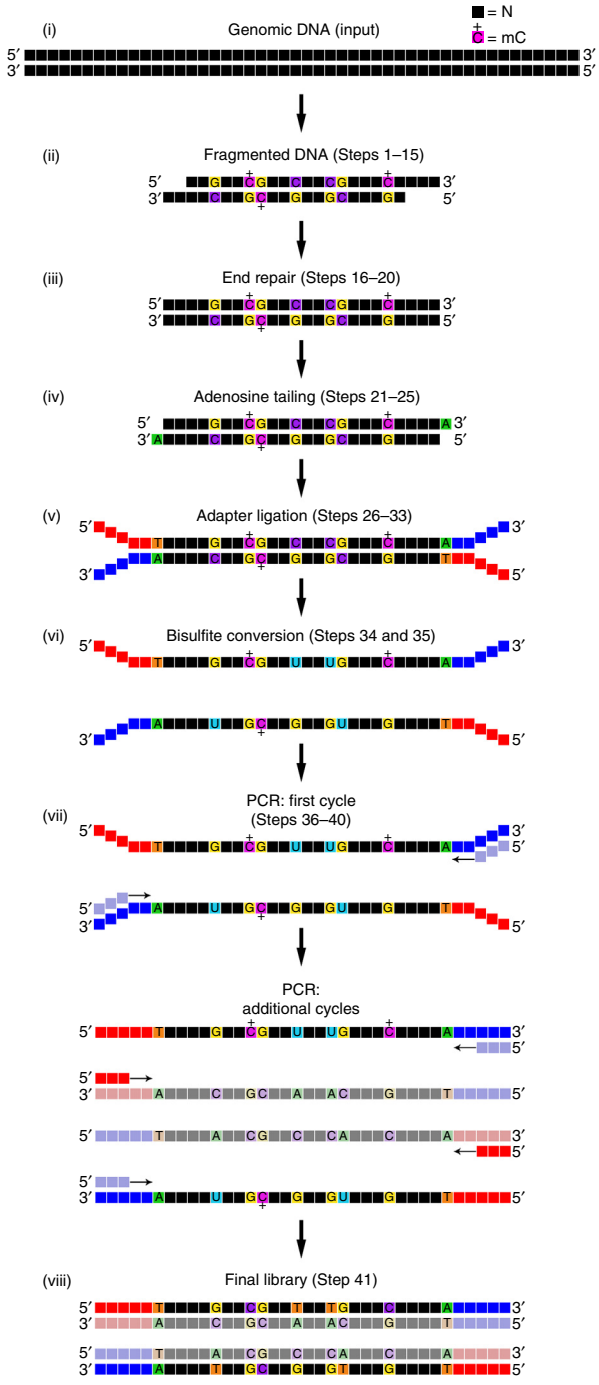

Figure 1 | MethylC-seq library preparation protocol overview. gDNA (i) is fragmented to ~200 bp by sonication (ii). DNA fragments containing damaged or incompatible 5′- and/or 3′-protruding ends are converted to 5′-phosphorylated, blunt-ended DNA (iii). Blunt-ended DNA fragments are converted to DNA with 3′-dAMP overhangs (iv). Methylated Y-shaped adapters are ligated to the dA-tailed DNA fragments (v). All cytosines in the adapters must be methylated to allow for primer binding and amplification after bisulfite conversion. Adapter-ligated DNA fragments are denatured, and unmethylated cytosine is converted to uracil during sodium bisulfite treatment (vi). Bisulfite-treated DNA fragments remain single-stranded as they are no longer complementary. Low-cycle PCR amplification is performed with a polymerase that can tolerate uracil residues (vii). The final library fragments contain thymines and cytosines in place of the original unmethylated cytosines and methylated cytosines, respectively (viii).

(RRBS)[20] and DNA methylation arrays[21]. However, it should be noted that these latter techniques are more powerful approaches for surveying a larger number of samples owing to the added costs of sequencing entire genomes associated with MethylC-seq.

By using MethylC-seq, features of DNA methylation have been discovered that were not readily apparent from traditional bisulfite-PCR methods, such as the identification of non-CG methylation in human embryonic stem cells and brain tissues and the identification of large partially methylated domains (up to megabases in length) in animal genomes[6,7]. By applying this method to unique populations of plants, the inheritance of DNA methylation states has been documented, revealing the spontaneous epimutation rate for single-methylation polymorphisms (SMPs), the prevalence of spontaneous epialleles and widespread evidence for population-wide association of genetic variants (methylQTL) with methylation variants[3,22,23]. This approach has also been broadly applied to understanding the patterns of DNA methylation throughout development[2,24–29], between species[7,12,30–35], within hybrids[36–40] and between mutants[41–43].

Although MethylC-seq enables comprehensive identification of both 5-methylcytosine and 5-hydroxymethylcytosine through one approach, it may be desirable to distinguish between these two modifications. Fortunately, techniques such as Tet-assisted bisulfite sequencing (TAB-seq)[44] or oxidative bisulfite sequencing (oxBS-seq)[45] enable whole-genome base-resolution identification of 5-hydroxymethylation, which when used in conjunction with MethylC-seq enable comprehensive mapping and discrimination of 5-methylcytosine and 5-hydroxymethylcytosine. Furthermore, a technique termed reduced bisulfite sequencing (redBS-seq)[46] has recently been reported; it enables whole-genome base-resolution identification of 5-formylcytosine, an oxidized derivative of 5-hydroxymethylcytosine formed by the ten-eleven translocation (TET) enzymes. RedBS-seq uses a specific chemical reduction reaction to convert 5-formylcytosine to 5-hydroxymethylcytosine, which are sensitive and resistant to sodium bisulfite conversion, respectively. Thus, by comparison of a redBS-seq profile with a MethylC-seq profile of the same genomic DNA, 5-formylcytosine bases can be identified as those that are resistant to conversion in redBS-seq but sensitive to conversion in MethylC-seq[46].

Limitations of MethylC-seq and alternative methods

This protocol is optimized for generating single-base-resolution DNA methylome libraries from samples that yield >50 ng of DNA. Successful sequencing libraries have even been prepared from input material as low as 1 ng, depending on the genome size. Sequencing libraries can be prepared from low-input samples, but they require additional cycles of PCR. These additional cycles can increase the frequency of duplicate reads generated from the PCR, which need to be dealt with computationally when analyzing the sequencing data. In general, a lower number of PCR cycles is preferred, as this increases the total number of sequencing reads that are useful in downstream analysis, decreases the costs per usable sequencing read and reduces the influence of any possible sequence-dependent read amplification bias. When sample yield is limiting, increasing PCR cycles will

generate libraries that can be sequenced, but it will result in a higher cost per sequenced read, as duplicate reads from PCR will be sequenced more frequently. Given the whole-genome nature of MethylC-seq, experimental cost scales linearly with genome size and sample number. Consequently, population-scale studies for organisms with very large genomes may currently be prohibitively costly. However, population-scale studies of organisms with moderate-sized genomes, such as *Arabidopsis*, have already been conducted[23], and given the continually decreasing cost of DNA sequencing the scope of large-scale MethylC-seq-based experiments is progressively increasing.

Alternative methods for whole-genome base-resolution identification of methylcytosine have been developed, including BS-seq[47], which through the use of an alternative sequencing adapter strategy produces all four possible distinct DNA sequences that are generated through copying a bisulfite-converted duplex DNA strand. Consequently, the bioinformatics strategy for analysis of BS-seq data is distinct and more complex than that for MethylC-seq. More recently, a method termed post-bisulfite adapter tagging (PBAT)[48] has been developed, in which the bisulfite conversion process itself is used to fragment the genomic DNA, followed by adapter tagging via two rounds of random priming extension. Thus, PBAT avoids bisulfite conversion–induced nicking of adapter-ligated DNA, and it does not use PCR-based amplification after adapter ligation. However, the use of random primer extension with bisulfite-converted DNA has the potential to induce sequence-specific biases in the resulting data.

Experimental design

Quality of genomic DNA. Before construction of MethylC-seq libraries, efforts should be made to ensure that the sample contains high-quality genomic DNA. Failure to do so can result in uneven coverage of cytosines in the genome, and it can unnecessarily bias the sequencing results. However, lower-quality genomic DNA tends to have similar coverage bias profiles across the genome, so results from such samples can still be used for analysis with the understanding that not all of the possible regions of the genome are being surveyed.

Bead-based size selection. We highly recommend testing the Agencourt AMPure XP bead solution used for purification before beginning this protocol, as recommended by the manufacturer. Typically, the bead solution can be tested on a 50-bp DNA ladder at varying ratios of bead solution to ladder solution. To evaluate the size-selection effectiveness of the beads, run the samples on a 2% (wt/vol) agarose gel.

Unmethylated control sequence. It is essential to have genomic DNA sequence in the sample that does not contain any methylated cytosines. These unmethylated sequences provide a control sequence from which to determine the efficiency of the sodium bisulfite conversion reaction. For plant genomes, simply using sequencing reads that align to the chloroplast genome sequence will suffice, as they are unmethylated; however, for most other methylomes we recommend adding unmethylated λ-DNA to the genomic DNA sample before library preparation as an internal control.

MATERIALS

REAGENTS

- Genomic DNA at a concentration of at least 0.5 ng/μl. We have successfully used gDNA samples prepared from *Arabidopsis thaliana*[3,12,19,23,43], *Mus musculus*[6,34] and *Homo sapiens*[6–8] by using the Qiagen plant DNeasy and the Qiagen blood and tissue kits
- Unmethylated λ phage DNA (Promega, cat. no. D1501)
- Agencourt AMPure XP (Beckman Coulter, cat. no. A63880)
- Ethyl alcohol (200 proof, for molecular biology; Sigma-Aldrich, cat. no. 32205)
- UltraPure nuclease-free water (Life Technologies, cat. no. AM9937)
- End-It DNA end-repair kit (Epicentre, cat. no. ER81050)
- dA-tailing buffer (NEB, cat. no. B6059)
- Klenow (3′–5′ exo⁻; NEB, cat. no. M0212)
- T4 DNA ligase with buffer (NEB, cat. no. M0202)
- Methylated adapters and corresponding PCR primers (NEXTflex bisulfite-seq barcodes −12, Bioo Scientific, cat. no. 511912)
- MethylCode bisulfite conversion kit (Life Technologies, MECOV-50) or EZ DNA methylation-Gold kit (Zymo, D5005); although other kits could be used, we use the MethylCode or EZ DNA methylation-Gold kits because of their ease of use and short time to completion compared with lengthy conversion reactions required for some kits from other vendors
- KAPA HiFi HotStart Uracil+ ReadyMix (Kapa Biosystems, cat. no. KK2801)

▲ **CRITICAL** Although there are multiple enzymes that could potentially be used for this protocol, in our experience the enzyme from KAPA produces libraries that are at least double the final concentration compared with alternatives.

- Qubit dsDNA high-sensitivity (HS) assay (Life Technologies, cat. no. Q32851)
- Tris-Cl, 10 mM, pH 8.5

EQUIPMENT

- S-series focused ultrasonicator (Covaris)
- microTUBE holder (Covaris, cat. no. 500114)
- Snap-Cap microTUBEs (Covaris, cat. no. 520045)
- Thermocycler
- Low-binding DNase-/RNase-free microcentrifuge tubes, 1.5 ml (Fisher, cat. no. 02681320)
- PCR tubes, 0.2 ml
- DynaMag-2 magnet (Life Technologies, cat. no. 12321D)
- Qubit 2.0 fluorometer (Life Technologies, cat. no. Q32866)

REAGENT SETUP

Ethanol, 80% (vol/vol) To 40 ml of 200 proof ethanol, add 10 ml of nuclease-free water. Keep it sealed when it is not in use. Prepare the solution freshly before each library preparation.

AMP-/ATP- and dA-/NTP-containing buffer aliquots Before the first use, split the entire contents of the dA-tailing buffer and the 10× T4 DNA ligase buffer into aliquots. Store the aliquots at −20 °C. Avoid freezing and thawing a single aliquot more than three times, as repeated freeze-thaw cycles can result in ATP-depleted buffers.

EQUIPMENT SETUP

Creating a '200-bp target peak size' protocol Create the protocol on the Covaris focused-ultrasonicator with the following specifications: Duty Cycle, 10%; Intensity, 5; Cycles/Burst, 200; Time, 60 s; Number of Cycles, 3.

PROCEDURE

gDNA fragmentation ● TIMING ~5 min per sample

1| Dilute 2 µg of high-quality gDNA in 115 µl of nuclease-free water, and add 10 µl of 1 ng/µl λ-DNA stock for a total of 125 µl. The λ-DNA lacks cytosine methylation, and it is used as an internal spike-in control at 0.5% (wt/wt) for measuring the conversion rate of unmethylated cytosines. The λ-DNA can be excluded when using gDNA from plants, as the unmethylated chloroplast DNA can be used as the control.
▲ CRITICAL STEP 50 ng–2 µg of gDNA is optimal. For lower input, spike in λ-DNA at 0.5% (wt/wt; for example, 5 ng of λ-DNA per 1 µg of gDNA). It is recommended that the gDNA be RNase-treated to ensure accurate quantification. The Qubit 2.0 dsDNA HS assay is strongly preferred for DNA quantification rather than 260/280-nm spectrophotometry, as only dsDNA but not RNA would be detected by the dsDNA HS assay.

2| Transfer the 125-µl sample into a Covaris microTUBE, and sonicate using the '200-bp target peak size' protocol (see Equipment Setup).

Upper-size-cutoff bead purification of fragmented gDNA ● TIMING ~20 min

3| Pulse-spin the microTUBE for 5 s at 3,000*g*, and then transfer 120 µl of sonicated DNA to a 1.5-ml tube.

4| Add 72 µl of AMPure XP bead solution to the sample, vortex it for 10 s and incubate it at room temperature (23 °C) for 10 min.
▲ CRITICAL STEP A bead solution volume of 0.6× DNA solution binds DNA fragments that are ~600 bp and larger. All ratios of bead solution to sample volume are important to successful completion of this protocol.

5| Place the tube on the magnetic stand for 5 min. The beads will collect near the magnet and the solution will become clear.

6| With the tube still on the magnetic stand, transfer 185 µl of the supernatant to a new 1.5-ml tube, and then discard the tube containing the bead pellet. The supernatant will contain DNA fragments smaller than 600 bp.
▲ CRITICAL STEP The beads at this step contain large DNA fragments (600 bp and larger) that can interfere with later steps and reduce library yield. It is important to not transfer any beads to the new tube along with the supernatant.

Lower-size-cutoff bead purification of fragmented gDNA ● TIMING ~30 min

7| Add 105 µl of AMPure XP bead solution to the sample, vortex it for 10 s and incubate it at room temperature for 10 min.
▲ CRITICAL STEP This step increases the overall bead solution volume to 1.4× that of the DNA solution. At this ratio, fragments of ~100 bp and larger will bind to the beads.

8| Place the tube on a magnetic stand for 5 min, until the solution is clear. With the tube still on the magnetic stand, carefully pipette out and discard the supernatant, leaving behind 5–10 µl so as not to remove any beads. The supernatant contains the unwanted DNA fragments (<100 bp), whereas the beads contain the appropriately sized fragments (100–600 bp).

9| Add 500 µl of 80% (vol/vol) ethanol to the tube. Immediately remove the tube from the magnetic stand, rotate it 180° and place it back on the magnetic stand. The beads will eventually jump across the ethanol solution, removing any residual binding buffer trapped within the bead pellet. Repeat the rotation 6–10 times. The beads should eventually appear to separate and move across the solution as a cloud instead of a solid bead pellet.
? TROUBLESHOOTING

10| With the tube on the magnetic rack, let the beads gather for 1 min, and then carefully pipette out and discard the supernatant. In this step, the beads should stay against the side of the tube when the ethanol solution is removed completely.

11| Repeat Steps 9 and 10 to wash the beads a second time.

12| Remove the tube from the magnetic stand and use a mini-centrifuge to pulse-spin any remaining ethanol to the bottom, for 5 s at 3,000*g* at room temperature. Place the tube back on the magnetic stand and remove any remaining liquid.

13| Remove the tube from the magnetic stand and let it sit with the cap open for 5 min or until the beads are dry. Small cracks can be observed in the dried bead pellet.

14| Add 34 µl of 10 mM Tris-Cl directly to the pellet. Mix it at least 15 times using a pipette, let it sit for 10 min and then mix it again.
? TROUBLESHOOTING

15| Place the tube on the magnetic stand for at least 2 min to allow complete capture of the beads. When the suspension has cleared, transfer the entire supernatant to a new 1.5-ml tube.

End repair of fragmented gDNA ● TIMING ~60 min
16| Repair the sticky ends of the fragmented gDNA to create blunt-ended fragments by using the End-It DNA end repair kit. Add 16 µl of the following components to the sample, for a total reaction volume of 50 µl.

Component	Amount per reaction (µl)	Final amount/concentration
10× End-It buffer	5	1×
10 mM dNTP mix	5	1 mM
10 mM ATP	5	1 mM
End-It enzyme mix	1	–

17| Mix the tube gently, pulse-spin it for 5 s at 3,000*g* and then incubate it at room temperature for 45 min.

Bead purification of the end-repair reaction ● TIMING ~30 min
18| Add 70 µl of AMPure XP bead solution to the end-repair reaction. Vortex the tube for 10 s and then incubate it at room temperature for 10 min.
▲ CRITICAL STEP The use of a 1.4× volume of bead solution to DNA solution in this step will slightly raise the lower cutoff to continue to exclude fragments that are <100 bp.

19| Repeat Steps 8–13 to wash and dry the beads.

20| Elute the DNA from the beads as in Steps 14 and 15, by using 42 µl of 10 mM Tris-Cl.
■ PAUSE POINT Samples can be stored at 4 °C overnight or at –20 °C for 7 d.

Addition of 'A' bases to the 3′ end ● TIMING ~45 min
21| Perform the A-tailing reaction. Add 8 µl of the following components to the sample, for a total reaction volume of 50 µl.

Component	Amount per reaction (µl)	Final amount/concentration
10× dA-tailing buffer	5	1×
Klenow (3′–5′ exo$^-$; 5 U µl^{-1})	3	15 U

22| Mix the sample gently by flicking or pipetting, pulse-spin it for 5 s at 3,000*g* at room temperature if necessary and then incubate it at 37 °C for 30 min.

Bead purification of the A-tailing reaction ● TIMING ~30 min
23| Add 70 µl of AMPure XP bead solution to the A-tailing reaction. Vortex the tube for 10 s, and then incubate it at room temperature for 10 min.
▲ CRITICAL STEP The use of a 1.4× volume of bead solution to DNA solution in this step will slightly raise the lower cutoff to exclude fragments that are <100 bp.

24| Repeat Steps 8–13 to wash and dry the beads.

25| Elute the DNA from the beads as in Steps 14 and 15, with 16.25 µl of 10 mM Tris-Cl.
■ PAUSE POINT Samples can be stored at 4 °C overnight or at –20 °C for 7 d.

Ligation of methylated adapters ● TIMING ~10 min (plus overnight incubation)

26| Place the sample tube on ice and add the following components in order. If multiple samples are being prepared, it is best to add components individually instead of making a master mix.

Component	Amount per reaction (μl)	Final amount/concentration
Methylated adapter (25 μM)	2	2 μM
10× DNA ligase buffer with ATP	2.5	1×
T4 DNA ligase (400 U μ^{-1})	1.25	500 U
Nuclease-free water	3	–

27| Mix the tube gently, pulse-spin it for 5 s at 3,000*g* at room temperature if necessary and then incubate it at 16 °C overnight.

Bead purification of the ligation reaction ● TIMING ~60 min

28| Add 25 μl of nuclease-free water to the ligation reaction to bring the volume to 50 μl.

29| Add 50 μl of AMPure XP bead solution (for a 1:1 ratio, which eliminates fragments below ~200 bp) to the ligation reaction. Vortex the tube for 10 s, and then incubate it at room temperature for 10 min.

30| Repeat Steps 8–13 to wash and dry the beads.

31| Elute the DNA from the beads as in Steps 14 and 15, with 50 μl of 10 mM Tris-Cl.

32| Purify the adapter-ligated gDNA a second time by repeating Steps 29 and 30.

33| Elute the DNA from the beads as in Steps 14 and 15, with 21 μl of 10 mM Tris-Cl.
■ **PAUSE POINT** Samples can be stored at 4 °C overnight or at –20 °C for 7 d.

Bisulfite conversion of adapter-ligated gDNA ● TIMING ~180 min

34| Use 1 μl of sample to measure the DNA concentration via Qubit. Transfer ≤450 ng or 20 μl (whichever is less) of adapter-ligated gDNA to a 0.2-ml PCR tube. If needed, increase the volume to 20 μl by using nuclease-free water.
? TROUBLESHOOTING

35| Perform conversion according to the instructions given in the MethylCode or EZ DNA methylation-Gold kit. Elute with 23 μl of elution buffer. Alternatively, if multiple PCRs are to be performed on the sample, elute in 23 μl per PCR. Multiple PCRs are helpful in downstream applications, when attempting to acquire a high sequencing depth, by reducing duplicate frequency.
▲ **CRITICAL STEP** When you are mixing the conversion reagent, tape the tube in a vertical position to the vortex mixer and let it mix for 10 min. There should be no visible solids in the conversion reagent. It is crucial that this solution be properly made, as this step will determine the conversion efficiency of unmethylated cytosines in the sample.
▲ **CRITICAL STEP** When you are mixing the reaction with binding buffer by inversion, watch the liquid to ensure that it mixes, as it has a tendency to remain stuck at one end of the tube.

Low-cycle-number amplification of bisulfite-converted gDNA ● TIMING ~60 min

36| Perform a four-cycle PCR to amplify sequenceable fragments from the adapter-ligated gDNA while keeping PCR duplicates low. Transfer the sample to a new 0.2-ml PCR tube and add 27 μl of the following components.

Component	Amount per reaction (μl)	Final amount/concentration
25× PCR primer mix	2	1×
2× KAPA HiFi uracil+ Readymix	25	1×

37| Load the sample on a thermocycler and run the following PCR program.

Cycle number	Denature	Anneal	Extend
1	95 °C for 2 min	—	—
2	98 °C for 30 s	—	—
3–6	98 °C for 15 s	60 °C for 30 s	72 °C for 1 min
7	—	—	72 °C for 10 min

Bead purification of the PCR ● TIMING ~30 min

38| Add 50 µl of AMPure XP bead solution (for a 1:1 ratio, which eliminates fragments below ~200 bp) to the PCR. Vortex the tube for 10 s and then incubate it at room temperature for 10 min.

39| Repeat Steps 8–13 to wash and dry the beads.

40| Elute the completed library from the beads as in Steps 14 and 15, with 16 µl of 10 mM Tris-Cl.

▲ CRITICAL STEP Bead carryover should be minimal during the final elution, as beads can interfere with sequencing. To reduce bead carryover when you are transferring the elution into a new tube, place the empty tube on the magnetic rack and use a p10 pipette with an elongated tip. Slowly dispense the elution from the pipette into the bottom of the tube. The beads should visibly gather in the pipette tip near the magnet. Leaving 0.5 µl of liquid in the tip will prevent transfer of the beads.

? TROUBLESHOOTING

■ PAUSE POINT The completed library can be stored at –20 °C for at least 6 months.

Library quantification ● TIMING ~15 min

41| Use 1 µl of sample to measure the DNA concentration via Qubit.

? TROUBLESHOOTING

? TROUBLESHOOTING

Troubleshooting advice can be found in **Table 1**.

TABLE 1 | Troubleshooting table.

Step	Problem	Possible reason	Solution
9	Beads do not jump across the ethanol solution, or they jump across as a solid pellet	If the tube is not removed from the magnet and rotated soon enough after the addition of ethanol, the beads have a tendency to stick together and to the side of the tube. This can also be caused by large-molecule impurities in the input gDNA, such as starches in plant gDNA	You can attempt to dislodge the beads from the side of the tube by flicking the outside of the tube, near the pellet, with your fingernail, while the tube is on the magnetic rack. If you are unable to dislodge the pellet easily, you can continue with the protocol without much negative impact
14	Beads are stuck to the side of the tube, or they stay in clumps during elution	This is often associated with problems at Step 9 and is most likely to be caused by not using low-binding 1.5 ml tubes or not rotating the tubes soon enough after addition of ethanol	The bead clumps can be manually broken apart with a pipette tip during elution by scraping them from the inside of the 1.5-ml tube and crushing clumps of beads. It is ok to roughly disturb the beads during this step to ensure complete separation and elution

(continued)

TABLE 1 | Troubleshooting table (continued).

Step	Problem	Possible reason	Solution
34	Concentration of adapter-ligated gDNA is <3 ng/μl	Incomplete adapter ligation or adapter ligation failure will result in low concentrations as non-ligated fragments will not be enriched during PCR. This can be caused by incomplete reactions during the end-repair, A-tailing or ligation steps. Low-quality input gDNA might also be the cause	It is recommended to not continue with the protocol. The sample will probably not be recoverable. For precious samples, if you wish to continue, proceed with caution and increase the number of PCR cycles to a minimum of eight. For future preparations, use only high-quality gDNA and fresh buffer aliquots
40	Beads carry over from final elution	Some beads may transfer from the pipette tip to the final tube	Repeat the transfer of the final elution into a new tube on the magnetic rack until bead carry over is minimal
41	Concentration of the final library is too low to measure	Incomplete mixing of binding buffer and conversion solution in the MethylCode columns can reduce binding and recovery from columns. Incomplete adapter ligation may also reduce effectiveness of PCR. Problems in Step 34 will probably lead to problems in Step 41	The library might be able to be rescued by repeating Steps 36–41, and increasing the number of PCR cycles. This will lead to a higher number of clonal reads during sequencing

● **TIMING**

Day 1: ~5 h

Steps 1 and 2, gDNA fragmentation: ~5 min per sample
Steps 3–6, upper-size-cutoff bead purification of fragmented gDNA: ~20 min
Steps 7–15, lower-size-cutoff bead purification of fragmented gDNA: ~30 min
Steps 16 and 17, end repair of fragmented gDNA: ~60 min
Steps 18–20, bead purification of the end-repair reaction: ~30 min
Steps 21 and 22, addition of 'A' bases to the 3' end: ~45 min
Steps 23–25, bead purification of the A-tailing reaction: ~30 min
Steps 26 and 27, ligation of methylated adapters: ~10 min (plus overnight incubation)

Day 2: ~6 h

Steps 28–33, bead purification of the ligation reaction: ~60 min
Steps 34 and 35, bisulfite conversion of adapter-ligated gDNA: ~180 min
Steps 36 and 37, low-cycle-number amplification of bisulfite-converted gDNA: ~60 min
Steps 38–40, bead purification of the PCR: ~30 min
Step 41, library quantification: ~45 min

ANTICIPATED RESULTS

We have successfully used this protocol to prepare over 1,000 MethylC-seq methylome libraries with gDNA derived from plant tissues, animal tissues, primary cell lines and sorted cells[3,6,8,12,19,23,34,43]. Final library concentrations can vary depending on the quality of the input gDNA and the number of PCR cycles. From our experience, high-quality libraries from four cycles of PCR will have final concentrations >5 ng/μl in 16 μl (80 ng), with some as high as 25 ng/μl (400 ng), with a mean of 10 ng/μl (160 ng). We have successfully used this protocol to prepare libraries with an initial gDNA input of 1 ng, although it is recommended to use a larger input quantity of 1 μg when possible. Lower-input libraries require the entire bisulfite-converted DNA to be used in a single PCR with up to 15 cycles. If concentrations of final libraries are <1 ng/μl, it is recommended that additional PCR cycles be used to increase the final concentration. This ensures that optimal cluster density can be achieved on the sequencing platform. Final libraries prepared using this protocol can be sequenced in either a single- or paired-end format. Last, we routinely observe >99% conversion efficiency using this protocol and would caution against using libraries with lower conversion rates for any analysis that performs base-resolution analysis as compared with kilobase-scale analysis.

ACKNOWLEDGMENTS This work was supported by the Howard Hughes Medical Institute and the Gordon and Betty Moore Foundation (GMBF3034) and grants from the National Science Foundation (MCB1344299 and MCB122246) to J.R.E. R.L. was supported with an Australian Research Council Future Fellowship (FT120100862). R.J.S. was supported by the US National Institutes of Health (R00GM100000) and the US National Science Foundation (IOS-1339194).

AUTHOR CONTRIBUTIONS R.L. and J.R.E. conceived and designed the original protocol. M.A.U., J.R.N., R.L. and R.J.S modified and updated the protocol to its current state. M.A.U., R.L., R.J.S. and J.R.E. wrote the manuscript.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests.

1. Becker, C. *et al.* Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. *Nature* **480**, 245–249 (2011).
2. Calarco, J.P. *et al.* Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. *Cell* **151**, 194–205 (2012).
3. Schmitz, R.J. *et al.* Transgenerational epigenetic instability is a source of novel methylation variants. *Science* **334**, 369–373 (2011).
4. Law, J.A. & Jacobsen, S.E. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* **11**, 204–220 (2010).
5. Ziller, M.J. *et al.* Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**, 477–481 (2013).
6. Lister, R. *et al.* Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013).
7. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
8. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
9. Ramsahoye, B.H. *et al.* Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc. Natl. Acad. Sci. USA* **97**, 5237–5242 (2000).
10. Zhang, X. *et al.* Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell* **126**, 1189–1201 (2006).
11. Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T. & Henikoff, S. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat. Genet.* **39**, 61–69 (2006).
12. Lister, R. *et al.* Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. *Cell* **133**, 523–536 (2008).
13. Clark, S.J., Harrison, J., Paul, C.L. & Frommer, M. High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22**, 2990–2997 (1994).
14. Frommer, M. *et al.* A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA* **89**, 1827–1831 (1992).
15. Hayatsu, H. Bisulfite modification of nucleic acids and their constituents. *Prog. Nucleic Acid Res. Mol. Biol.* **16**, 75–124 (1976).
16. Hayatsu, H., Wataya, Y., Kai, K. & Iida, S. Reaction of sodium bisulfite with uracil, cytosine, and their derivatives. *Biochemistry* **9**, 2858–2865 (1970).
17. Shapiro, R., Braverman, B., Louis, J.B. & Servis, R.E. Nucleic acid reactivity and conformation. II. Reaction of cytosine and uracil with sodium bisulfite. *J. Biol. Chem.* **248**, 4060–4064 (1973).
18. Shapiro, R., DiFate, V. & Welcher, M. Deamination of cytosine derivatives by bisulfite. Mechanism of the reaction. *J. Am. Chem. Soc.* **96**, 906–912 (1974).
19. Dowen, R.H. *et al.* Widespread dynamic DNA methylation in response to biotic stress. *Proc. Natl. Acad. Sci. USA* **109**, E2183–E2191 (2012).
20. Gu, H. *et al.* Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **6**, 468–481 (2011).
21. Bibikova, M. *et al.* High density DNA methylation array with single CpG site resolution. *Genomics* **98**, 288–295 (2011).
22. Schmitz, R.J. *et al.* Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. *Genome Res.* **23**, 1663–1674 (2013).
23. Schmitz, R.J. *et al.* Patterns of population epigenomic diversity. *Nature* **495**, 193–198 (2013).
24. Xie, W. *et al.* Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell* **148**, 816–831 (2012).
25. Ibarra, C.A. *et al.* Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. *Science* **337**, 1360–1364 (2012).
26. Zhong, S. *et al.* Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. *Nat. Biotechnol.* **31**, 154–159 (2013).
27. Potok, M.E., Nix, D.A., Parnell, T.J. & Cairns, B.R. Reprogramming the maternal zebrafish genome after fertilization to match the paternal methylation pattern. *Cell* **153**, 759–772 (2013).
28. Jiang, L. *et al.* Sperm, but not oocyte, DNA methylome is inherited by zebrafish early embryos. *Cell* **153**, 773–784 (2013).
29. Stroud, H. *et al.* Plants regenerated from tissue culture contain stable epigenome changes in rice. *eLife* **2**, e00354 (2013).
30. Takuno, S. & Gaut, B.S. Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc. Natl. Acad. Sci. USA* **110**, 1797–1802 (2013).
31. Gent, J.I. *et al.* CHH islands: *de novo* DNA methylation in near-gene chromatin regulation in maize. *Genome Res.* **23**, 628–637 (2013).
32. Regulski, M. *et al.* The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res.* **23**, 1651–1662 (2013).
33. Zemach, A. *et al.* Local DNA hypomethylation activates genes in rice endosperm. *Proc. Natl. Acad. Sci. USA* **107**, 18729–18734 (2010).
34. Vollmers, C. *et al.* Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell Metab.* **16**, 833–845 (2012).
35. Eichten, S.R. *et al.* Epigenetic and genetic influences on DNA methylation variation in maize populations. *Plant Cell* **25**, 2783–2797 (2013).
36. Greaves, I.K. *et al.* Trans chromosomal methylation in *Arabidopsis* hybrids. *Proc. Natl. Acad. Sci. USA* **109**, 3570–3575 (2012).
37. Groszmann, M. *et al.* Changes in 24-nt siRNA levels in *Arabidopsis* hybrids suggest an epigenetic contribution to hybrid vigor. *Proc. Natl. Acad. Sci. USA* **108**, 2617–2622 (2011).
38. Barber, W.T. *et al.* Repeat associated small RNAs vary among parents and following hybridization in maize. *Proc. Natl. Acad. Sci. USA* **109**, 10444–10449 (2012).
39. Li, X. *et al.* High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. *Plant Cell* **20**, 259–276 (2008).
40. Wang, X. *et al.* Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. *Plant Cell* **21**, 1053–1069 (2009).
41. Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V. & Jacobsen, S.E. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. *Cell* **152**, 352–364 (2013).
42. Zemach, A. *et al.* The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**, 193–205 (2013).
43. Kurihara, Y. *et al.* Surveillance of 3′ noncoding transcripts requires FIERY1 and XRN3 in *Arabidopsis*. *G3* (*Bethesda*) **2**, 487–498 (2012).
44. Yu, M. *et al.* Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* **149**, 1368–1380 (2012).
45. Booth, M.J. *et al.* Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* **336**, 934–937 (2012).
46. Booth, M.J., Marsico, G., Bachman, M. & Beraldi, D. Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. *Nat. Chem.* **6**, 435–440 (2014).
47. Cokus, S.J. *et al.* Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* **452**, 215–219 (2008).
48. Miura, F., Enomoto, Y., Dairiki, R. & Ito, T. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res.* **40**, e136 (2012).