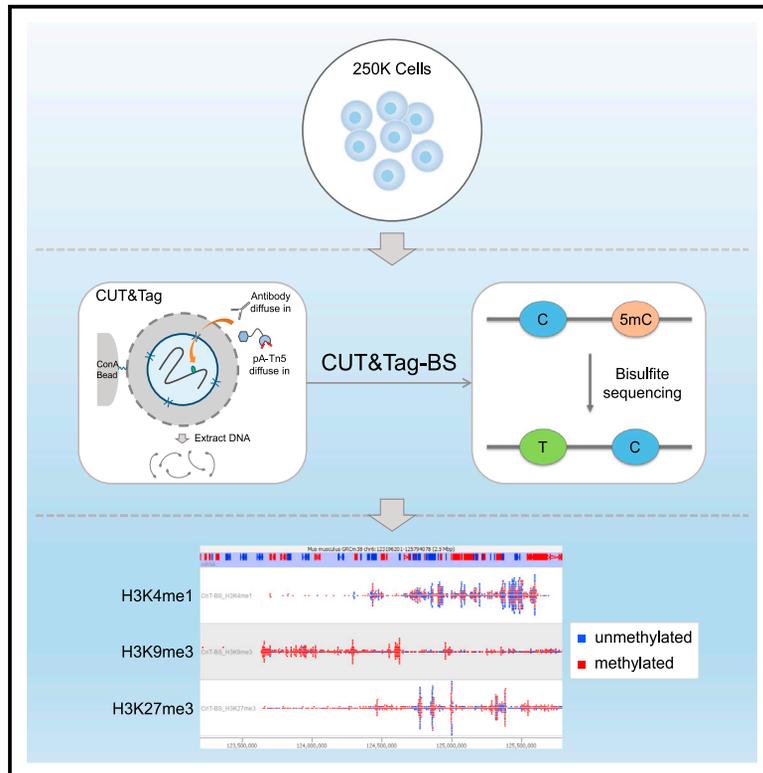


CUT&Tag-BS for simultaneous profiling of histone modification and DNA methylation with high efficiency and low cost

Graphical abstract



Authors

Ruifang Li, Sara A. Grimm, Paul A. Wade

Correspondence

lir1@mskcc.org

In brief

Li et al. present an efficient and low-cost method (CUT&Tag-BS) to simultaneously measure histone modification and DNA methylation on the same DNA molecules using a small number of cells. This provides a unique solution for characterizing the two main epigenetic modifications concurrently with limited sample material in various biological contexts.

Highlights

- CUT&Tag-BS is an efficient, low-input, and low-cost method for dual-omics profiling
- Concurrent profiling of histone modification and DNA methylation in the same cells
- Lack of DNA methylation concordance at H3K4me1- and H3K27me3-marked regions



Article

CUT&Tag-BS for simultaneous profiling of histone modification and DNA methylation with high efficiency and low cost

Ruifang Li,^{1,4,*} Sara A. Grimm,² and Paul A. Wade³¹Epigenetics Innovation Lab, Center for Epigenetics Research, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA²Integrative Bioinformatics Support Group, Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA³Eukaryotic Transcriptional Regulation Group, Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA⁴Lead contact*Correspondence: lir1@mskcc.org<https://doi.org/10.1016/j.crmeth.2021.100118>

MOTIVATION To investigate the relationship between DNA methylation and histone modification, the conventional correlative approach employs separate mono-omics assays and thus requires large amounts of input material and deep sequencing. The recently developed more direct approaches such as BisChIP-seq and ChIP-BS-seq suffer from suboptimal signal-to-noise ratio, poor resolution, and potential inaccuracy of DNA methylation measurement due to formaldehyde cross-linking. Here, we present an efficient, low-input, and low-cost method by coupling CUT&Tag with tagmentation-based bisulfite sequencing (CUT&Tag-BS) to directly interrogate genomic localization of histone modification and DNA methylation from the same DNA molecules with a low number of native cells and dramatically reduced sequencing depth.

SUMMARY

It remains a challenge to decipher the complex relationship between DNA methylation, histone modification, and the underlying DNA sequence with limited input material. Here, we developed an efficient, low-input, and low-cost method for the simultaneous profiling of genomic localization of histone modification and methylation status of the underlying DNA at single-base resolution from the same cells in a single experiment by integrating cleavage under targets and tagmentation (CUT&Tag) with tagmentation-based bisulfite sequencing (CUT&Tag-BS). We demonstrated the validity of our method using representative histone modifications of euchromatin and constitutive and facultative heterochromatin (H3K4me1, H3K9me3, and H3K27me3, respectively). Similar histone modification enrichment patterns were observed in CUT&Tag-BS compared with non-bisulfite-treated control, and H3K4me1-marked regions were found to mostly be CpG poor, lack methylation concordance, and exhibit prevalent DNA methylation heterogeneity among mouse embryonic stem cells (mESCs). We anticipate that CUT&Tag-BS will be widely applied to directly address the genomic relationship between DNA methylation and histone modification, especially in low-input scenarios with precious biological samples.

INTRODUCTION

DNA methylation and histone modification are two main mediators of epigenetic regulation, structurally and functionally coordinating with each other to regulate a multitude of biological processes on chromatin. A wide range of different modifications exist on the N-terminal tails of histones, including acetylation, methylation, phosphorylation, and ubiquitylation (Bannister and Kouzarides, 2011), and their relationship with DNA methylation

differs depending on the type and location of the modification (Fu et al., 2020; Xiao et al., 2012). The interdependent deposition and mutual exclusion of DNA methylation with different histone modifications generate complex chromatin modification patterns that demarcate distinct functional elements in the genome (Roadmap Epigenomics et al., 2015). Traditionally, the spatial relationship of DNA methylation with a given histone modification is determined by parallel genomic mapping of the two marks using whole-genome bisulfite sequencing (WGBS)



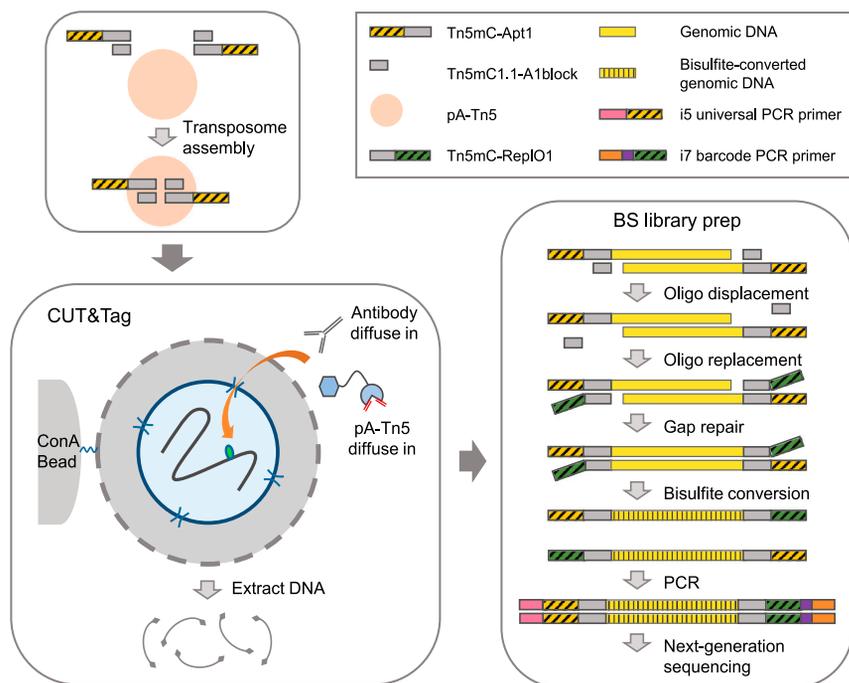


Figure 1. A schematic overview of CUT&Tag-BS workflow

CUT&Tag is adapted to couple with tagmentation-based bisulfite sequencing by using pA-Tn5 assembled with a methylated adapter. The top adapter, Tn5mC-Apt1, and the replacement oligonucleotide, Tn5mC-ReplO1, must be methylated at all cytosines to maintain their identity during bisulfite treatment. CUT&Tag is performed according to the published protocol (Kaya-Okur et al., 2019) with pA-Tn5 loaded with a methylated adapter. Briefly, cells are harvested and bound to concanavalin A-coated magnetic beads. The cell membrane is permeabilized with digitonin (indicated by holes in the membrane) to allow the antibodies and pA-Tn5 to diffuse into the cells to find their targets and then tagment genomic binding sites of the target protein. Tagmented DNA from CUT&Tag is subsequently subjected to tagmentation-based bisulfite sequencing library preparation. The oligonucleotide replacement and gap-repair step covalently attach the methylated adapter Tn5mC-ReplO1 to each DNA strand, followed by bisulfite conversion and PCR amplification to generate the library for sequencing. This figure is adapted from previously published figures in Skene et al. (2018) and Wang et al. (2013) by permission from Springer Nature.

and chromatin immunoprecipitation sequencing (ChIP-seq), followed by integrative overlap analysis of the profiles. However, by overlaying profiles from independent assays on different batches of cells, the traditional approach only permits correlative analysis and incurs prohibitive sequencing costs to obtain adequate depth of coverage for DNA methylation. To more directly and cost-effectively interrogate the interplay between DNA methylation and histone modification, bisulfite sequencing of chromatin-immunoprecipitated DNA (BisChIP-seq) (Statham et al., 2012) and ChIP bisulfite sequencing (ChIP-BS-seq) (Brinkman et al., 2012) were developed to measure DNA methylation specifically from genomic regions marked with a given histone modification through bisulfite sequencing of ChIP-captured DNA (Kagey et al., 2010). Those two methods use a ligation-based bisulfite sequencing library preparation strategy, which requires a large amount of ChIP-ed DNA (~100 ng); thus, they are not suitable for low-input samples. To reduce the amount of ChIP-ed DNA required, EpiMethylTag (Lhoumaud et al., 2019) was developed recently and uses a tagmentation-based library preparation strategy instead. Despite the differences in library construction, the three methods all rely on cross-linked ChIP to capture chromatin fragments associated with the protein of interest. As a result, they suffer from the same limitations as ChIP, such as requiring large amounts of input material, a suboptimal signal-to-noise ratio, and poor resolution. Moreover, formaldehyde cross-linking damages DNA (Do and Dobrovic, 2015), and incomplete reverse cross-linking interferes with bisulfite conversion, thus confounding DNA methylation measurements (Wen et al., 2017).

Cleavage under targets and tagmentation (CUT&Tag) (Kaya-Okur et al., 2019) is a cutting-edge technique recently developed for genome-wide mapping of protein-DNA interactions. As an alternative to ChIP, CUT&Tag employs protein A-Tn5 (pA-Tn5)

for antibody-guided *in situ* tagmentation of target-bound DNA in native cells. It has several advantages over ChIP, including a faster and simpler workflow, lower input requirement, higher signal-to-noise ratio, and fewer sequencing reads needed. Tagmentation-based WGBS (T-WGBS) (Wang et al., 2013) utilizes Tn5 assembled with a methylated adapter to tagment genomic DNA for subsequent bisulfite sequencing library preparation. It has a shorter workflow and requires less input DNA, while exhibiting higher reproducibility and robustness than conventional ligation-based bisulfite sequencing (Wang et al., 2013). We reasoned that, if one is conducting a CUT&Tag experiment using pA-Tn5 loaded with a methylated adapter adopted from T-WGBS, the resulting CUT&Tag-DNA can be further subjected to bisulfite sequencing following the same library preparation procedure as T-WGBS. CUT&Tag was thus adapted to couple with bisulfite sequencing (CUT&Tag-BS) for simultaneous profiling of histone modification and DNA methylation from the same cells. As a proof of concept, we performed H3K4me1-, H3K9me3-, and H3K27me3-CUT&Tag-BS on mouse embryonic stem cells (mESCs) to demonstrate the utility of our method in directly addressing the relationships of DNA methylation with histone modification at euchromatin and constitutive and facultative heterochromatin, respectively.

RESULTS

CUT&Tag-BS for simultaneous profiling of histone modification and DNA methylation

To enable simultaneous measurement of DNA methylation and histone modification on the same DNA molecule, we developed CUT&Tag-BS by integrating the CUT&Tag procedure with tagmentation-based bisulfite sequencing (Figure 1). In our method, CUT&Tag is employed to generate tagmented DNA in native

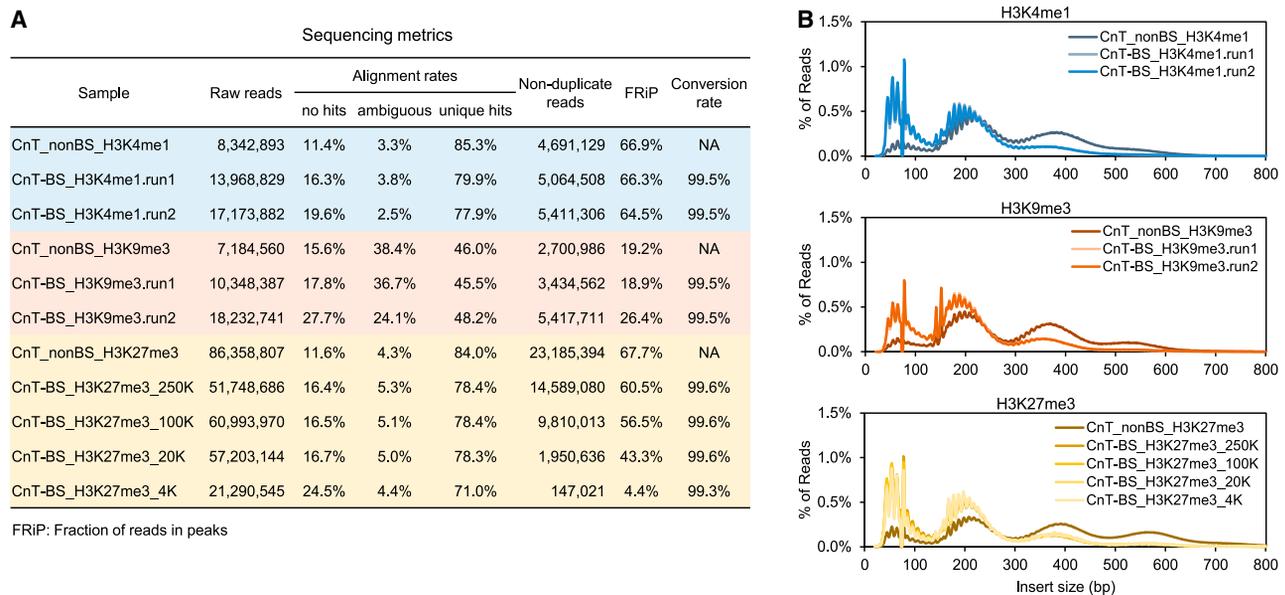


Figure 2. CUT&Tag-BS shows library quality control metrics comparable to those of non-bisulfite-treated control
(A and B) (A) Sequencing metrics and (B) insert size distribution of CUT&Tag-BS libraries compared with non-bisulfite-treated control libraries.

cells specifically at genomic binding sites of a given histone modification using pA-Tn5 under the guidance of target-specific antibody, and the CUT&Tag DNA is then subjected to tagmentation-based bisulfite sequencing, resulting in a base-resolution DNA methylation map at genomic regions bound by the given histone modification (Figure 4A). Our method utilizes pA-Tn5 loaded with a methylated adapter (Table S1) to make the resultant CUT&Tag-DNA compatible with subsequent tagmentation-based bisulfite sequencing library preparation. The library preparation procedure includes an oligonucleotide replacement and gap-repair step to covalently append methylated adapter sequences to each single strand of tagmented DNA fragments, followed by bisulfite conversion and PCR amplification. During the gap-repair step, unmethylated nucleotides are used to fill in the nine-base gap. Those bases (the first nine bases of read 2 and the last nine bases before the adapter on read 1) serve as an internal control to determine the efficiency of bisulfite conversion, but they must be excluded from downstream methylation analysis. As a combination of two techniques, both characterized by low-input requirement and simple workflow, CUT&Tag-BS is thus suitable for efficient simultaneous profiling of histone modification and DNA methylation with a low number of cells.

As a proof of principle, we performed CUT&Tag-BS for representative histone modifications of euchromatin and constitutive and facultative heterochromatin (H3K4me1, H3K9me3, and H3K27me3, respectively) using 250,000 mESCs. Since bisulfite treatment damages DNA extensively, to assess whether it alters CUT&Tag enrichment signals, we also generated non-bisulfite-treated CUT&Tag libraries as controls, which were sequenced and analyzed in parallel with CUT&Tag-BS libraries. CUT&Tag-BS libraries were sequenced with both paired-end 36-basepair (bp) reads (run-1) and paired-end 76-bp reads (run-2) to deter-

mine the proper sequencing read length for this technique. We further evaluated the robustness and input requirement of CUT&Tag-BS by using different numbers of input cells.

CUT&Tag-BS shows similar quality control metrics and target enrichments compared to non-bisulfite-treated control

CUT&Tag-BS libraries exhibited alignment rates similar to those of corresponding non-bisulfite-treated controls, with ~80% of H3K4me1 reads, ~46% of H3K9me3 reads, and ~78% of H3K27me3 reads uniquely mapped to the reference genome (Figure 2A). Of note, H3K9me3 is enriched at repetitive regions (Martens et al., 2005); therefore, a relatively high percentage of H3K9me3 reads (~37% of 36-bp reads and ~24% of 76-bp reads) were ambiguously mapped (Figure 2A), leading to a lower percentage of uniquely mapped reads. Similar to non-bisulfite-treated controls, CUT&Tag-BS libraries showed a nucleosomal ladder pattern in fragment size distribution, although the proportion of smaller fragments increased after bisulfite treatment (Figure 2B), likely due to its detrimental effects on DNA integrity. Lastly, CUT&Tag-BS libraries exhibited a high bisulfite conversion rate (>99%; Figure 2A). Overall, based on library quality control metrics, all the CUT&Tag-BS libraries were of high quality, irrespective of the type of chromatin being profiled.

To assess whether bisulfite treatment distorts CUT&Tag enrichment signals, peak calling was performed on both CUT&Tag-BS and non-bisulfite-treated control samples. Despite bisulfite conversion, CUT&Tag-BS detected a similar number of peaks, largely overlapping with those identified in the control sample (Figure 3A); the non-overlapped peaks generally had low enrichment signals in both datasets (Figure S1A). Furthermore, the peak signals exhibited a high correlation between CUT&Tag-BS and the corresponding control (Pearson's $r = 0.994, 0.978,$

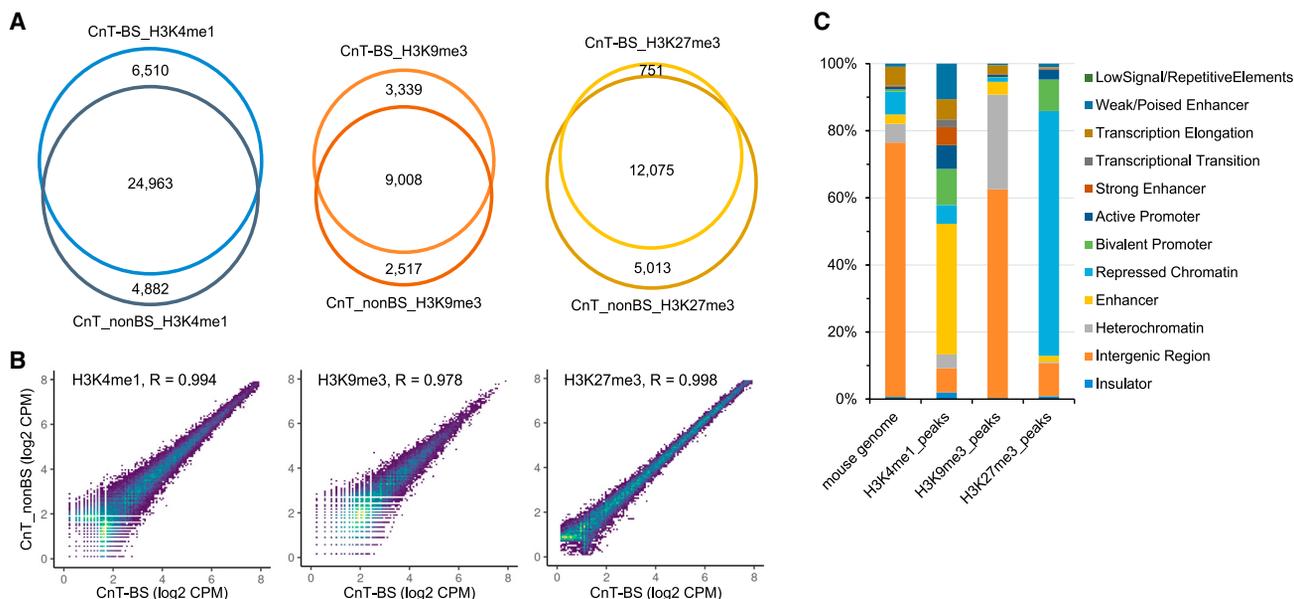


Figure 3. CUT&Tag-BS shows similar target enrichments to non-bisulfite-treated control

(A) Venn diagrams showing overlap of peaks identified in CUT&Tag-BS and corresponding non-bisulfite-treated control.

(B) Density scatter plots displaying correlation of peak signals between CUT&Tag-BS and the corresponding non-bisulfite-treated control. Each dot represents an individual peak in the unified peak set with viridis color scale indicating density. Pearson's r value is shown at the top of each plot.

(C) Bar graph showing the proportion of detected peaks and the reference mouse genome falling into each chromatin state of mESCs.

and 0.998 for H3K4me1, H3K9me3, and H3K27me3, respectively) (Figure 3B). Moreover, CUT&Tag-BS showed a fraction of reads in peaks (FRiP) score similar to that observed in the corresponding control (Figure 2A), indicating a comparable signal-to-noise ratio despite bisulfite conversion. Collectively, these analyses suggest that no obvious signal bias was observed after bisulfite treatment. Since different histone modifications are associated with distinct functional elements in the genome, to further verify that CUT&Tag-BS correctly identifies the expected functional elements, we conducted overlap analysis of detected peak regions with chromatin states of mESCs defined by the ChromHMM model (Pintacuda et al., 2017). H3K4me1 peaks were enriched preferentially at enhancers (~55% peaks) and promoters (~18% peaks), while H3K9me3 peaks were mainly located at heterochromatin (~28% peaks) and intergenic regions (~62% peaks), and H3K27me3 peaks were primarily at repressed chromatin (~73% peaks) and bivalent promoters (~10% peaks) (Figure 3C), consistent with the ChIP-seq histone modification patterns used to define the chromatin states. Furthermore, the peak signals of CUT&Tag-BS were highly correlated with those of ChIP-seq, using the same antibodies (Figure S1B). Taken together, compared with conventional CUT&Tag and ChIP-Seq, CUT&Tag-BS achieved similar histone modification enrichments at expected genomic regions despite bisulfite treatment.

CUT&Tag-BS simultaneously measures DNA methylation at target-binding sites

In addition to histone modification enrichments, CUT&Tag-BS simultaneously measures DNA methylation at the enriched genomic regions (Figure 4A). M-bias plots show an average

CpG methylation rate of ~32.5% for H3K4me1 reads, ~81.5% for H3K9me3 reads, and ~32.9% for H3K27me3 reads across the read length, except for the first nine bases of read 2 (Figure S2A), which correspond to the nine-bp gap region filled in with unmethylated nucleotides. Since a CpG site is either methylated or unmethylated in a given single cell, when DNA methylation is profiled from a cell population, the methylation level at a given CpG site actually reflects the percentage of cells methylated at that site, with intermediate methylation indicating methylation heterogeneity among the cells. The composite methylation level averages across all profiled CpG sites; thus, the observed intermediate methylation level may result from averaging across CpGs with different methylation levels or may simply represent a similar intermediate methylation state across all profiled CpG sites. To ascertain which is the case, we further investigated the distribution of methylation levels of individual CpG sites in peak regions. Irrespective of the minimum coverage required at the CpG sites, we observed consistently skewed distributions of CpG methylation levels, with H3K4me1-marked CpGs peaking at 0% methylation and H3K9me3-marked CpGs peaking at 100% methylation, in contrast to H3K27me3-marked CpGs exhibiting bimodal distribution with 0% methylation being more prevalent; however, a substantial number of CpGs exhibited intermediate methylation levels in all three cases (Figure S2B). To quantitatively determine the proportion of CpGs with different methylation status, we classified CpG methylation levels into four categories: less than 10% methylation (unmethylated), 10%–50% methylation (low-intermediate), 50%–90% methylation (high-intermediate), and more than 90% methylation (fully methylated). CpG sites covered by at least five reads at peak regions were then categorized based on their methylation levels.

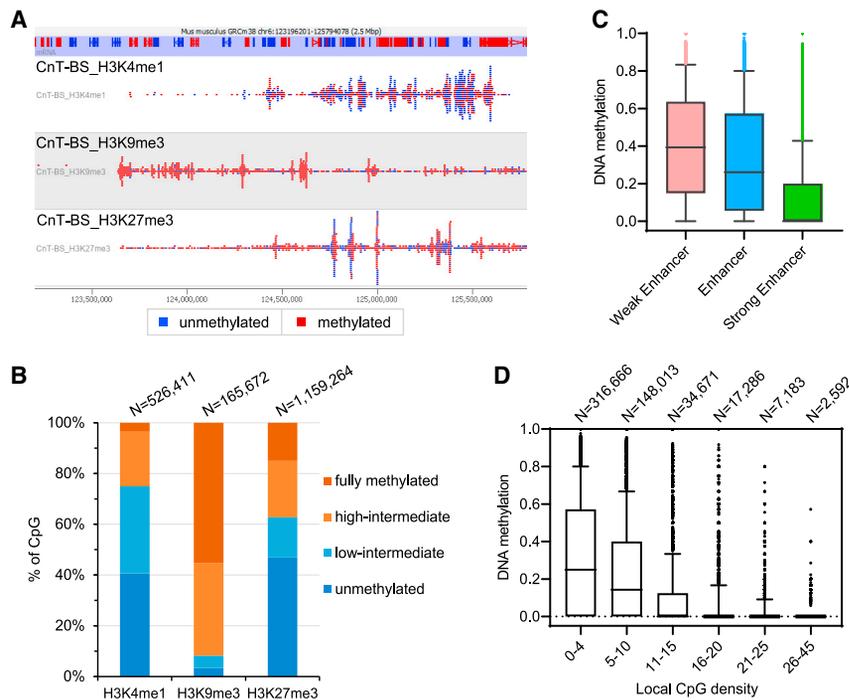


Figure 4. CUT&Tag-BS simultaneously measures DNA methylation at target-binding sites

(A) Example of CUT&Tag-BS tracks at chromosome 6 (chr6):123,196,201–125,794,078, with CpG methylation status indicated by color (red, methylated; blue, unmethylated).

(B) Bar graph showing the percentage of H3K4me1-, H3K9me3-, or H3K27me3-marked CpGs in each methylation category (unmethylated, <10% methylation; low-intermediate, 10%–50% methylation; high-intermediate, 50%–90% methylation; fully methylated, >90% methylation).

(C) Box-and-whisker plot showing methylation levels of H3K4me1-marked CpGs (minimum of 5 \times coverage) at different types of enhancers. The whiskers represent the 10th and 90th percentiles.

(D) Box-and-whisker plot showing methylation levels of H3K4me1-marked CpGs (minimum of 5 \times coverage) with different local CpG density, which is defined as the number of CpG sites within the region ± 100 bp of a given CpG. The whiskers represent the 10th and 90th percentiles. The number of CpGs in each group is shown at the top of the plot.

Among H3K9me3-marked CpGs, 55.3% were fully methylated, 36.4% exhibited high-intermediate methylation, and the rest showed low-intermediate methylation (4.8%) or were unmethylated (3.4%) (Figure 4B), consistent with the previous notion that H3K9me3 is associated with hypermethylated DNA in mESCs (Brinkman et al., 2012). In contrast, 46.9% of H3K27me3-marked CpGs were unmethylated, 15.8% showed low-intermediate methylation, 22.2% exhibited high-intermediate methylation, and 15.0% were fully methylated (Figure 4B), in agreement with the finding of ChIP-BS-seq that the DNA methylation pattern at H3K27me3-marked regions was bimodal in mESCs (Brinkman et al., 2012). Lastly, 40.7% of H3K4me1-marked CpGs were unmethylated, 34.3% showed low-intermediate methylation, 21.5% exhibited high-intermediate methylation, and 3.5% were fully methylated (Figure 4B). This suggests that even though H3K4me1-marked histones generally bind to unmethylated DNA, they also bind to methylated DNA at some genomic locations in a subset of the cells, as reflected by intermediately methylated H3K4me1 CpGs, in line with previously reported cell-to-cell DNA methylation heterogeneity at enhancers in mESCs (Angermueller et al., 2016; Song et al., 2019). Considering epigenetic heterogeneity among the cells, we expected that DNA methylation levels at enhancers measured by CUT&Tag-BS would not be perfectly aligned with those obtained with WGBS (Lu et al., 2014) (Figure S3), which neglects cell heterogeneity and includes all cells regardless of whether they carry H3K4me1 or not at a given CpG site.

H3K4me1 is a chromatin hallmark of enhancers (Bulger and Groudine, 2011); however, the role of DNA methylation at enhancers is poorly understood. We thus investigated DNA methylation at different types of enhancers. Although “weak/poised enhancer,” “enhancer,” and “strong enhancer” were all marked

with H3K4me1, they showed different DNA methylation levels, diminishing gradually with the increase in enhancer activity (Figure 4C), suggesting a negative association of DNA methylation with enhancer activity in general. However, DNA methylation was highly variable at “weak/poised enhancer” and “enhancer” (Figure 4C). To determine other influencing factors of DNA methylation at H3K4me1-marked CpG sites, we further investigated the relationship between DNA methylation status and local CpG density around the site. H3K4me1-marked CpGs were mostly located at low-CpG-density regions, and their methylation levels were inversely correlated with the local CpG density (Figure 4D), in concordance with previously observed global conflicts of CpG density and DNA methylation in human and mouse tissues (Chen et al., 2018).

CUT&Tag-BS reveals lack of methylation concordance at enhancers

To determine the proper read length for CUT&Tag-BS, we first explored the necessity of sequencing through the library insert. We reasoned that if CpG methylation status is consistent across the insert, it would be valid to use CpGs present on the partially sequenced portion to represent all CpGs on a given insert; otherwise, reading through the insert would be necessary to obtain more accurate information on DNA methylation at target-binding sites. We then assessed the concordance of methylation calls across the sequencing read containing three or more CpGs. More than one-third of the assessed H3K4me1, H3K9me3, and H3K27me3 reads (42.1%, 33.7%, and 35.3%, respectively) showed mixed methylation (Figure 5A), indicating inconsistent methylation status of CpGs on the same insert DNA. It is seemingly contrary to the previous findings that DNA methylation levels were strongly correlated at nearby CpG sites, particularly when they

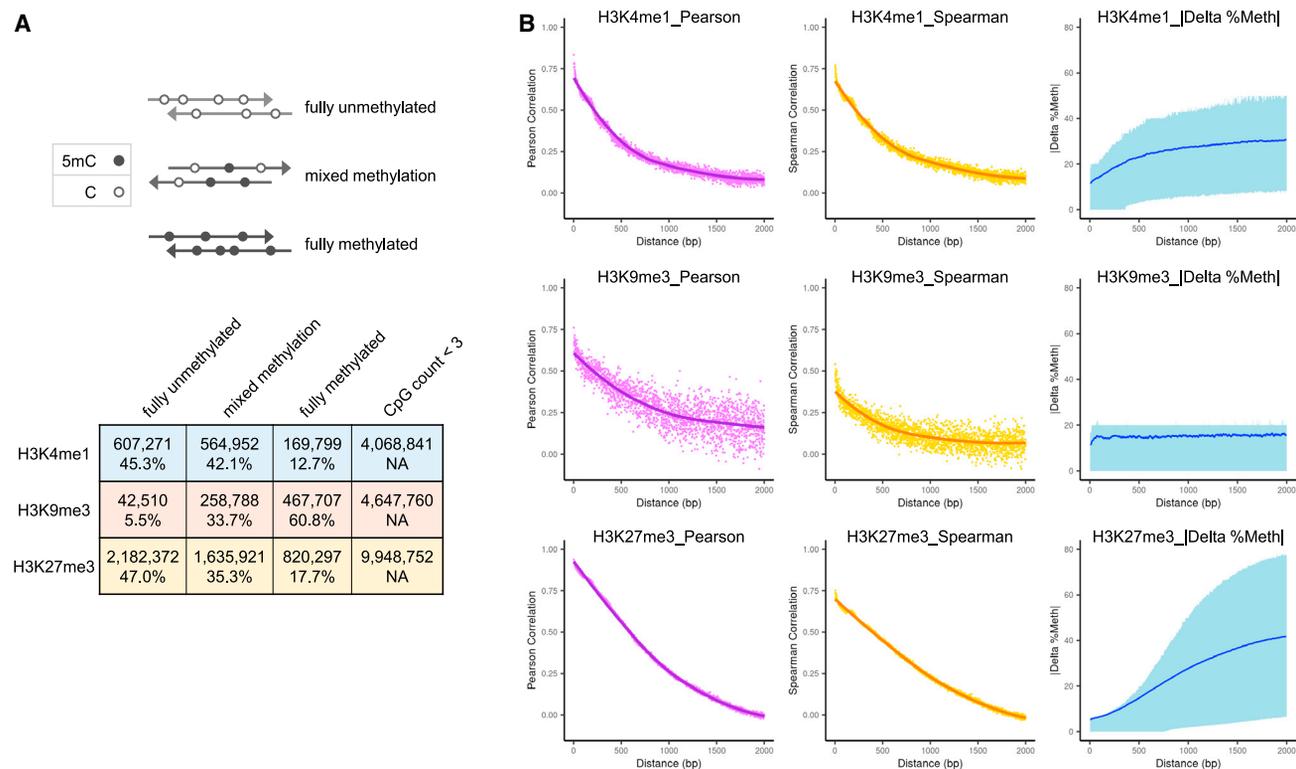


Figure 5. CUT&Tag-BS reveals a lack of methylation concordance at enhancers

(A) Concordance of CpG methylation calls on individual sequencing reads of CUT&Tag-BS. Only reads with three or more CpGs were assessed. The “lower” and “upper” thresholds were set at 10% and 90% methylation, respectively, to split the reads into three categories (fully unmethylated, mixed methylation, and fully methylated).

(B) Correlation of (Pearson, left; Spearman, middle) and difference in (right) methylation level between neighboring CpG sites as a function of their genomic distance. The x axis represents genomic distance between pairs of CpG sites, with at least one CpG per pair located in the peak regions (H3K4me1, top; H3K9me3, middle; H3K27me3, bottom). Loess curves were added to the correlation plots (left and middle) in ggplot2 v.3.3.2 with geom_smooth. The absolute methylation differences between neighboring CpGs are shown (right), with the dark blue line and the light blue ribbon indicating the average and interquartile range, respectively.

were within 1–2 kb of each other (Bell et al., 2011; Eckhardt et al., 2006). However, our analysis of CpG methylation concordance within single sequencing reads only assessed one allele of a given cell, while previous findings were based on CpG methylation levels of the cell population. To rule out the possibility that the discrepancy was due to methodology difference, we further measured the correlation and difference of methylation levels between pairs of neighboring CpG sites as a function of their genomic distance. Correlation of methylation levels rapidly dropped below 0.5 within 200 bp of distance for both H3K4me1- and H3K9me3-marked CpGs and within 500 bp of distance for H3K27me3-marked CpGs (Figure 5B). The absolute methylation difference between neighboring CpGs exhibited different patterns in the three marks; H3K9me3-marked CpGs showed a constant small difference in methylation levels at neighboring CpGs, irrespective of their distance, while H3K4me1- and H3K27me3-marked CpGs displayed larger methylation differences that positively correlated with the genomic distance between the sites, with H3K27me3-marked CpGs showing more dramatic methylation differences with the increase in distance (Figure 5B), suggesting that methylation concordance is highly dependent on the genomic regions interrogated. Nevertheless, considering that longer reads greatly

increased CpG coverage (Figure S4), sequencing through the library insert would be beneficial for achieving higher accuracy in DNA methylation. Therefore, the proper read length for CUT&Tag-BS is subject to the library insert size pertinent to the target protein.

CUT&Tag-BS is a low-cost method requiring fewer sequencing reads

We further assessed the sequencing depth required for CUT&Tag-BS. Owing to its high signal-to-noise ratio, CUT&Tag needs only several million sequencing reads (Kaya-Okur et al., 2019). To verify that the advantage also extends to CUT&Tag-BS, we calculated the FRiP score as a measure of signal-to-noise ratio, comparing H3K4me1-, H3K9me3-, and H3K27me3-CUT&Tag-BS data with publicly available corresponding ChIP-seq data in mESCs (Feldmann et al., 2020; Gao et al., 2018; Muller et al., 2021), generated using the same antibodies as CUT&Tag-BS. The FRiP score of CUT&Tag-BS nearly reached a plateau (~0.6), with just 5 million sequencing reads for H3K4me1 and H3K27me3 (Figure 6), similar to the number of reads claimed to be sufficient in CUT&Tag for histone modifications, while H3K9me3 is an exception due to the nature of its enrichment

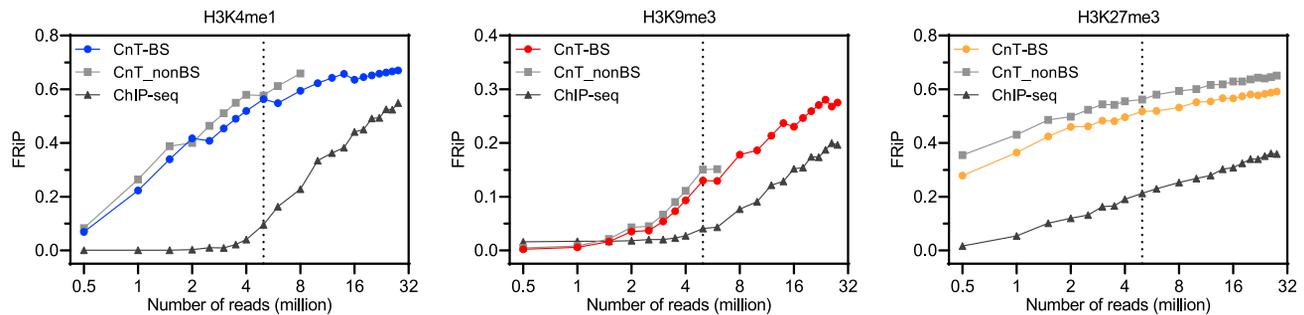


Figure 6. CUT&Tag-BS is a low-cost method requiring fewer sequencing reads

Fraction of reads in peaks (FRIp) is calculated based on HOMER peak calls, using different numbers of randomly downsampled sequencing reads in CUT&Tag-BS compared with CUT&Tag-non-BS and ChIP-seq. The same antibodies were used for CUT&Tag and ChIP on mESCs. The dashed line marks 5 million reads.

pattern. By contrast, ChIP-seq data were less suitable for peak calling when downsampled to 5 million reads, as indicated by the much lower FRIp score (~ 0.1 and ~ 0.2 for H3K4me1 and H3K27me3, respectively) (Figure 6). Due to its intrinsic high background, ChIP-seq would require significantly more sequencing reads than CUT&Tag-BS to reach a similar FRIp score. CUT&Tag-BS allows one to reach reasonable coverage of CpGs at target-binding sites with a greatly reduced amount of sequencing. For example, with 5.4 million usable reads (paired-end 76-bp) in H3K4me1-CUT&Tag-BS, 92.1% and 43.0% of the CpGs at H3K4me1 peaks were covered with a minimum depth of 1 \times and 5 \times , respectively (Figure S4). In conclusion, CUT&Tag-BS shows a higher signal-to-noise ratio than ChIP-seq with the same antibody and is inherently more cost effective than methods relying on ChIP to capture chromatin fragments associated with the histone modification of interest.

CUT&Tag-BS profiles low-cell-number samples

Lastly, we evaluated the input requirement of the method by performing H3K27me3-CUT&Tag-BS with different input cell numbers (250K, 100K, 20K, and 4K). We observed very similar H3K27me3 profiles at highly enriched regions from all experiments (Figure S5A). However, due to sampling smaller repertoires of H3K27me3-marked regions when using lower cell numbers, the library complexity dropped with the reduction in input cell numbers, as indicated by smaller numbers of non-duplicate reads (Figure 2A). Accordingly, the number of peaks and CpG coverage also decreased correspondingly with the decrease in input cell numbers (Figures 7A, 7B, and S5B). Nevertheless, the CUT&Tag signals of samples using lower input cell numbers showed a high correlation with that of the 250K sample (Figure 7C), suggesting that high data quality is still maintained with lower cell numbers. Collectively, CUT&Tag-BS is suitable for profiling lower numbers of cells, but computational tools developed for analyzing bulk data from a large population of cells may not perform well with sparse data associated with a very low cell number (i.e., 4K sample).

DISCUSSION

We developed CUT&Tag-BS to simultaneously profile genomic localization of histone modification and methylation status of

the underlying DNA from the same cells by coupling CUT&Tag with bisulfite sequencing (Figure 1). The conventional bisulfite sequencing strategy requires relatively large amounts of high-quality DNA for optimal conversion (Clark et al., 1994, 2006), thus imposing challenges on the methods that rely on cross-linked ChIP to capture chromatin fragments for subsequent bisulfite sequencing. CUT&Tag-BS overcomes some of the limitations of existing methods. First, CUT&Tag-BS uses native cells without formaldehyde fixation, thus avoiding the side effects of cross-linking/de-cross-linking, as reflected by high bisulfite conversion efficiency ($\sim 99.5\%$) and improved overall alignment frequency in CUT&Tag-BS ($\sim 80\%$) (Figure 2A) compared with EpiMethylTag ($\sim 65\%$) and ChIP-BS-seq ($\sim 33\%$) (Lhoumaud et al., 2019). Second, CUT&Tag-BS adopts a tagmentation-based bisulfite sequencing library preparation strategy, which is faster and requires substantially less DNA than conventional ligation-based bisulfite sequencing library construction (Wang et al., 2013). This, in combination with efficient CUT&Tag, makes CUT&Tag-BS suitable for much lower cell numbers than is practical with existing methods relying on ChIP. Third, CUT&Tag-BS exhibits high signal-to-noise ratio (Figure 6). The inherent low background and high read mappability translate to almost an order-of-magnitude reduction in the amount of sequencing required, thus dramatically lowering sequencing cost. In addition, low-background CUT&Tag-BS ensures DNA methylation measurements derive from DNA truly bound by the given histone modification, thus providing more confidence in the observed relationship of the marks than using methods relying on high-background ChIP.

Notably, bisulfite treatment did not alter the pattern of CUT&Tag enrichment in terms of FRIp score, number of detected peaks, genomic locations of the peaks, and signal strength at peaks (Figures 2A and 3). In addition to well-preserved enrichment signals, CUT&Tag-BS simultaneously measured DNA methylation at enriched regions with single-base resolution. Prevalent DNA methylation heterogeneity was observed at enhancers among the cells, as shown by the intermediate methylation levels of more than half (55.8%) of the H3K4me1-marked CpG sites in mESCs (Figure 4B). Since cell-to-cell differences are always present to some degree in any population of cells, potential epigenetic heterogeneity should be considered when deciphering the relationship between DNA

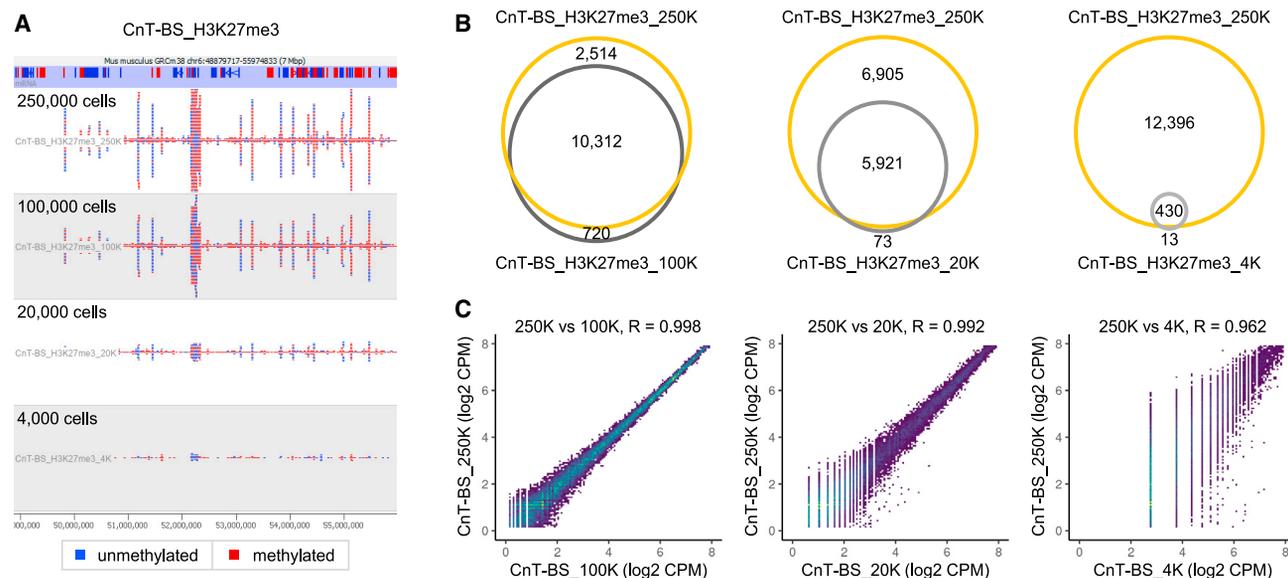


Figure 7. CUT&Tag-BS profiles low-cell-number samples

(A) H3K27me3-CUT&Tag-BS tracks at chr6: 48,879,717–55,974,833 generated with different numbers of input cells (250,000, 100,000, 20,000, and 4,000). CpG methylation status is indicated by color (red, methylated; blue, unmethylated).

(B) Venn diagrams showing overlap of peaks identified in 250,000 sample with those detected in 100,000, 20,000, or 4,000 sample.

(C) Density scatterplots displaying correlation of peak signals between 250,000 sample and samples with lower input cell numbers (100,000, 20,000, and 4,000). Each dot represents an individual peak in the unified peak set, with viridis color scale indicating density. Pearson's r value is shown at the top of each plot.

methylation and histone modification in a cell population. Direct bisulfite sequencing of CUT&Tag-DNA enables more accurate and sensitive analysis of the spatial relationship between the marks than the traditional correlative approach, which neglects cell heterogeneity due to integrating data from independent assays on different cells. DNA methylation was previously reported to be characterized by spatial correlation within 1–2 kb (Bell et al., 2011; Eckhardt et al., 2006), providing a rationale for assigning the methylation state of single-CpG measurements to all CpGs within a genomic interval. However, CUT&Tag-BS analysis indicated that DNA methylation concordance was highly dependent on the genomic regions interrogated, with H3K4me1-marked regions lacking methylation concordance (Figure 5), in line with recently reported faster methylation concordance decay at enhancers (Hui et al., 2018). In contrast to a previously oversimplified relationship of the marks as either coexisting or mutually exclusive, CUT&Tag-BS revealed a more complex relationship between DNA methylation and histone modification, which remains to be further elucidated in a region-specific manner.

In summary, CUT&Tag-BS was established successfully as a combination of CUT&Tag and T-WGBS to decipher the relationship between histone modification and DNA methylation across the genome, as demonstrated by representative histone modifications at euchromatin and constitutive and facultative heterochromatin (H3K4me1, H3K9me3, and H3K27me3, respectively). It is straightforward and requires minimal adaptation of previous CUT&Tag (Kaya-Okur et al., 2019) and T-WGBS (Wang et al., 2013) protocols. As an efficient and low-cost method, CUT&Tag-BS needs only a small number of input cells and sequencing reads. Hence, we anticipate that

CUT&Tag-BS will be widely applied to directly interrogate the complex interplay between DNA methylation and histone modification, especially in low-input scenarios with precious biological samples. Moreover, we speculate that CUT&Tag-BS could be used to investigate the relationship between DNA methylation and other important epigenetic regulators, such as chromatin modifiers/remodelers and transcription factors. By simultaneously providing information on both chromatin context and strand-specific DNA methylation at base resolution, CUT&Tag-BS would enable better understanding of epigenetic regulation in various biological contexts.

Limitations of the study

Similar to other antibody-targeted chromatin profiling, the success of CUT&Tag-BS heavily relies on the affinity and specificity of the primary antibody for its target under the conditions used for binding. Antibody-specific problems such as low affinity and epitope masking will necessitate deeper sequencing, while inferior specificity of the antibody would lead to misinterpretation of the genomic profile of target protein. We expect that the antibodies successfully tested for specificity and applicability in CUT&Tag will probably work in CUT&Tag-BS as well. Currently, the pA-Tn5 transposome commercially available is loaded with unmethylated adapters, so users need to prepare a pA-Tn5 transposome with methylated adapters. Last, the stringent conditions used to avoid the binding of pA-Tn5 to accessible DNA can also affect target binding on chromatin, thus making CUT&Tag-BS inapplicable for profiling weakly bound targets in native cells, which can be potentially overcome by using formaldehyde-cross linked cells.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
 - Assemble pA-Tn5 with methylated adapter
 - CUT&Tag-BS
 - CUT&Tag-BS data analysis
 - Peak calling and peak overlap between samples
 - Peak signal quantification
 - Overlap of peaks with chromatin states
 - Methylation concordance and methylation correlation
 - ChIP-seq data processing
 - WGBS data processing
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.crmeth.2021.100118>.

ACKNOWLEDGMENTS

We thank Dr. Hua Wang and Dr. Chang Huang in the Helin lab for providing the E14 mESCs and pA-Tn5 fusion protein, respectively. This work was supported by donations to the Center for Epigenetics Research from The Metropoulos Family Foundation and The Ambrose Monell Foundation and by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences (ES101965 to P.A.W.).

AUTHOR CONTRIBUTIONS

R.L. conceived the project and performed the experiments. R.L. and S.A.G. analyzed and interpreted the data. R.L. wrote the manuscript, and S.A.G. and P.A.W. reviewed and edited the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 19, 2021

Revised: September 16, 2021

Accepted: October 25, 2021

Published: November 15, 2021

REFERENCES

Angermueller, C., Clark, S.J., Lee, H.J., Macaulay, I.C., Teng, M.J., Hu, T.X., Krueger, F., Smallwood, S., Ponting, C.P., Voet, T., et al. (2016). Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat. Methods* **13**, 229–232.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res* **21**, 381–395.

Bell, J.T., Pai, A.A., Pickrell, J.K., Gaffney, D.J., Pique-Regi, R., Degner, J.F., Gilad, Y., and Pritchard, J.K. (2011). DNA methylation patterns associate

with genetic and gene expression variation in HapMap cell lines. *Genome Biol.* **12**, R10.

Brinkman, A.B., Gu, H., Bartels, S.J., Zhang, Y., Matarese, F., Simmer, F., Marks, H., Bock, C., Gnirke, A., Meissner, A., and Stunnenberg, H.G. (2012). Sequential ChIP-bisulfite sequencing enables direct genome-scale investigation of chromatin and DNA methylation cross-talk. *Genome Res.* **22**, 1128–1138.

Bulger, M., and Groudine, M. (2011). Functional and mechanistic diversity of distal transcription enhancers. *Cell* **144**, 327–339.

Chen, F., Zhang, Q., Deng, X., Zhang, X., Chen, C., Lv, D., Li, Y., Li, D., Zhang, Y., Li, P., et al. (2018). Conflicts of CpG density and DNA methylation are proximally and distally involved in gene regulation in human and mouse tissues. *Epigenetics* **13**, 721–741.

Clark, S.J., Harrison, J., Paul, C.L., and Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic Acids Res.* **22**, 2990–2997.

Clark, S.J., Statham, A., Stirzaker, C., Molloy, P.L., and Frommer, M. (2006). DNA methylation: bisulphite modification and analysis. *Nat. Protoc.* **1**, 2353–2364.

Do, H., and Dobrovic, A. (2015). Sequence artifacts in DNA from formalin-fixed tissues: causes and strategies for minimization. *Clin. Chem.* **61**, 64–71.

Eckhardt, F., Lewin, J., Cortese, R., Rakyan, V.K., Attwood, J., Burger, M., Burton, J., Cox, T.V., Davies, R., Down, T.A., et al. (2006). DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* **38**, 1378–1385.

Feldmann, A., Dimitrova, E., Kenney, A., Lastuvkova, A., and Klose, R.J. (2020). CDK-Mediator and FBXL19 prime developmental genes for activation by promoting atypical regulatory interactions. *Nucleic Acids Res.* **48**, 2942–2955.

Fu, K., Bonora, G., and Pellegrini, M. (2020). Interactions between core histone marks and DNA methyltransferases predict DNA methylation patterns observed in human cells and tissues. *Epigenetics* **15**, 272–282.

Gao, Y., Gan, H., Lou, Z., and Zhang, Z. (2018). Asf1a resolves bivalent chromatin domains for the induction of lineage-specific genes during mouse embryonic stem cell differentiation. *Proc. Natl. Acad. Sci. U S A.* **115**, E6162–E6171.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589.

Hui, T., Cao, Q., Wegrzyn-Woltosz, J., O'Neill, K., Hammond, C.A., Knapp, D., Laks, E., Moksa, M., Aparicio, S., Eaves, C.J., et al. (2018). High-resolution single-cell DNA methylation measurements reveal epigenetically distinct hematopoietic stem cell subpopulations. *Stem Cell Rep.* **11**, 578–592.

Kagey, J.D., Kapoor-Vazirani, P., McCabe, M.T., Powell, D.R., and Vertino, P.M. (2010). Long-term stability of demethylation after transient exposure to 5-aza-2'-deoxycytidine correlates with sustained RNA polymerase II occupancy. *Mol. Cancer Res.* **8**, 1048–1059.

Kaya-Okur, H.S., Wu, S.J., Codomo, C.A., Pledger, E.S., Bryson, T.D., Henikoff, J.G., Ahmad, K., and Henikoff, S. (2019). CUT&Tag for efficient epigenomic profiling of small samples and single cells. *Nat. Commun.* **10**, 1930.

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

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359.

Lhoumaud, P., Sethia, G., Izzo, F., Sakellaropoulos, T., Snetkova, V., Vidal, S., Badri, S., Cornwell, M., Di Giammartino, D.C., Kim, K.T., et al. (2019). EpiMethylTag: simultaneous detection of ATAC-seq or ChIP-seq signals with DNA methylation. *Genome Biol.* **20**, 248.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079.

Lu, F., Liu, Y., Jiang, L., Yamaguchi, S., and Zhang, Y. (2014). Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev.* **28**, 2103–2119.

- Martens, J.H., O'Sullivan, R.J., Braunschweig, U., Opravil, S., Radolf, M., Steinlein, P., and Jenuwein, T. (2005). The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.* *24*, 800–812.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* *17*, 10–12.
- Muller, I., Moroni, A.S., Shlyueva, D., Sahadevan, S., Schoof, E.M., Radzish-euskaya, A., Hojfeldt, J.W., Tatar, T., Koche, R.P., Huang, C., and Helin, K. (2021). MPP8 is essential for sustaining self-renewal of ground-state pluripotent stem cells. *Nat. Commun.* *12*, 3034.
- Pintacuda, G., Wei, G., Roustan, C., Kirmizitas, B.A., Solcan, N., Cerase, A., Castello, A., Mohammed, S., Moindrot, B., Nesterova, T.B., and Brockdorff, N. (2017). hnRNPK recruits PCGF3/5-PRC1 to the xist RNA B-repeat to establish polycomb-mediated chromosomal silencing. *Mol. Cell* *68*, 955–969.e910.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* *26*, 841–842.
- Roadmap Epigenomics, C., Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* *518*, 317–330.
- Skene, P.J., Henikoff, J.G., and Henikoff, S. (2018). Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* *13*, 1006–1019.
- Song, Y., van den Berg, P.R., Markoulaki, S., Soldner, F., Dall'Agnese, A., Henninger, J.E., Drotar, J., Rosenau, N., Cohen, M.A., Young, R.A., et al. (2019). Dynamic enhancer DNA methylation as basis for transcriptional and cellular heterogeneity of ESCs. *Mol. Cell* *75*, 905–920.e906.
- Statham, A.L., Robinson, M.D., Song, J.Z., Coolen, M.W., Stirzaker, C., and Clark, S.J. (2012). Bisulfite sequencing of chromatin immunoprecipitated DNA (BisChIP-seq) directly informs methylation status of histone-modified DNA. *Genome Res.* *22*, 1120–1127.
- Wang, Q., Gu, L., Adey, A., Radlwimmer, B., Wang, W., Hovestadt, V., Bahr, M., Wolf, S., Shendure, J., Eils, R., et al. (2013). Tagmentation-based whole-genome bisulfite sequencing. *Nat. Protoc.* *8*, 2022–2032.
- Wen, X., Jeong, S., Kim, Y., Bae, J.M., Cho, N.Y., Kim, J.H., and Kang, G.H. (2017). Improved results of LINE-1 methylation analysis in formalin-fixed, paraffin-embedded tissues with the application of a heating step during the DNA extraction process. *Clin. Epigenetics* *9*, 1.
- Xiao, S., Xie, D., Cao, X., Yu, P., Xing, X., Chen, C.C., Musselman, M., Xie, M., West, F.D., Lewin, H.A., et al. (2012). Comparative epigenomic annotation of regulatory DNA. *Cell* *149*, 1381–1392.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-H3K4me1 antibody (Rabbit monoclonal)	Cell Signaling Technology	Cat# 5326S
anti-H3K9me3 antibody (Rabbit monoclonal)	Abcam	Cat# ab176916
anti-H3K27me3 antibody (Rabbit monoclonal)	Cell Signaling Technology	Cat# 9733S
Guinea Pig Anti-Rabbit IgG	Antibodies-Online	Cat# ABIN101961
Chemicals, peptides, and recombinant proteins		
BioMag®Plus Concanavalin A	Bangs Laboratories	Cat# BP531
cOmplete™, EDTA-free Protease Inhibitor Cocktail	Sigma-Aldrich	Cat# 4693132001
Ethylenediaminetetraacetic acid solution (0.5M EDTA)	Sigma-Aldrich	Cat# 03690-100ML
Spermidine	Sigma-Aldrich	Cat# S2501
Digitonin	EMD Millipore	Cat# 300410
Bovine Serum Albumin solution	Sigma-Aldrich	Cat# A8577-10ML
dNTP Mix (10 mM each)	Thermo Fisher Scientific	Cat# R0191
Protein A–Tn5 (pA–Tn5) fusion protein	Dr. Kristian Helin lab	N/A
T4 DNA Polymerase	New England Biolabs	Cat# M0203S
Ampligase Enzyme and Buffer	Lucigen	Cat# A3202K
AMPure XP beads	Beckman Coulter	Cat# A63881
NEBNext® High-Fidelity 2X PCR Master Mix	NEB	Cat# M0541S
Critical commercial assays		
MinElute PCR Purification Kit	Qiagen	Cat# 28004
EZ DNA Methylation-Lightning Kit	Zymo Research	Cat# D5030
KAPA HiFi Uracil+ Kit	Roche	Cat# KK2801
Qubit™ dsDNA HS Assay Kit	Invitrogen™	Cat# Q32851
Deposited data		
CUT&Tag-BS data	This paper	GEO: GSE179266
Publicly available H3K4me1 ChIP-seq data	NCBI GEO	GEO: GSM4216306, GSM4216314
Publicly available H3K9me3 ChIP-seq data	NCBI GEO	GEO: GSM4561090, GSM4561089
Publicly available H3K27me3 ChIP-seq data	NCBI GEO	GEO: GSM3141530, GSM3141522
Publicly available WGBS data	NCBI GEO	GEO: GSM1372647
Experimental models: Cell lines		
E14 mouse embryonic stem cells	Dr. Kristian Helin lab	RRID: CVCL_C320
Oligonucleotides		
Oligonucleotides for tagmentation-based bisulfite library construction	See Table S1	N/A
Software and algorithms		
Cutadapt v1.12	Martin, 2011	N/A
Bismark v0.23.0	Krueger and Andrews, 2011	N/A
Bowtie2 v2.3.0	Langmead and Salzberg, 2012	N/A
Picard tools v1.11.0	http://broadinstitute.github.io/picard	N/A
HOMER v4.10.3	Heinz et al., 2010	N/A
BEDtools v2.29.2	Quinlan and Hall, 2010	N/A
R 3.5.0	https://www.r-project.org/	N/A
SRA Toolkit v2.11.0	http://ncbi.github.io/sra-tools/	N/A
samtools v1.3.1	Li et al., 2009	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Ruifang Li (lir1@mskcc.org).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The datasets generated during this study are available at NCBI Gene Expression Omnibus (GEO) with accession number GSE179266. Information also listed in the [key resources table](#).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Mouse E14 embryonic stem cells (129/Ola background) were maintained in Glasgow Minimum Essential Medium (GMEM, Sigma) containing 15% fetal bovine serum, supplemented with 1 × Pen-Strep (Gibco), 2 mM Glutamax (Gibco), 50 μM β-mercaptoethanol (Gibco), 0.1 mM nonessential amino acids (Gibco), 1 mM sodium pyruvate (Gibco), and Leukemia Inhibitory Factor (LIF, 1000U/ml, Millipore).

METHOD DETAILS

Assemble pA-Tn5 with methylated adapter

Single-stranded methylated adapter oligos (Tn5mC-Apt1 and Tn5mC1.1-A1block; [Table S1](#)) were dissolved in Annealing Buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM EDTA) to a final concentration of 200 μM. The resuspended oligos were mixed in equal volumes then annealed in a thermocycler using the following program; 95°C for 2 min, followed by decreasing the temperature in 5°C increments to reach 25°C with a ramp rate of −0.1°C/sec and 5 min incubation at each ending, then hold at 8°C. For transposome assembly, 100 μL of 8 μM pA-Tn5 fusion protein was mixed with 20 μL of the annealed adapter, and the mixture was incubated at room temperature (RT) for 1 hour on a rotator then stored at −20°C.

CUT&Tag-BS

CUT&Tag

CUT&Tag of H3K4me1, H3K9me3 and H3K27me3 in mESCs was performed using pA-Tn5 loaded with methylated adapter, following the bench-top CUT&Tag protocol ([Kaya-Okur et al., 2019](#)) with minor modifications. H3K4me1- and H3K9me3-CUT&Tag was conducted with 250,000 cells, while H3K27me3-CUT&Tag was performed with the number of input cells ranging from 250,000 down to 4,000 cells. Harvested mESCs were counted with an automated cell counter. One half million of the cells were transferred into a new 1.5mL Eppendorf tube then pelleted at 600 × g for 3 min at RT. The cell pellet was washed twice with 1 ml Wash Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine (Sigma-Aldrich), 1 × Protease inhibitor cocktail (Sigma-Aldrich)), and then resuspended in 200 μL of Wash Buffer. The target number of cells were transferred into a new Eppendorf tube; when needed, Wash Buffer was added to make the final volume to 100 μL. For the cells to bind to Concanavalin A-coated magnetic beads (Bangs Laboratories), 10 μL activated beads were added into each sample and the cell-bead mixture was rotated at RT for 10 min. The tubes were then placed on a magnet stand to clear and the supernatant was removed. The bead-bound cells were resuspended in 100 μL of Antibody Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 1 × Protease inhibitor cocktail, 0.01% Digitonin (EMD Millipore), 2 mM EDTA, 0.1% BSA (NEB)) with anti-H3K4me1 antibody (Cell Signaling Technology), anti-H3K9me3 antibody (Abcam) or anti-H3K27me3 antibody (Cell Signaling Technology) at 1:100 dilution. Primary antibody incubation was performed at 4°C overnight on a rotator. The tubes were then placed on the magnet stand to clear and unbound primary antibody was removed by discarding the supernatant. Next, the bead-bound cells were resuspended in 100 μL of Antibody Buffer with Guinea Pig anti-Rabbit IgG antibody (Antibodies online) at 1:100 dilution and then incubated at RT for 30 min on a rotator. After secondary antibody binding, the cells on beads were washed three times with 1 mL Dig-Wash Buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.5 mM Spermidine, 1 × Protease inhibitor cocktail, 0.01% Digitonin) to remove unbound antibodies. The bead-bound cells were then incubated with pA-Tn5 loaded with methylated adapter at 1:200 dilution in 100 μL Dig-300 Buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM Spermidine, 1 × Protease inhibitor cocktail, 0.01% Digitonin) for 1 hour at RT on a rotator. After pA-Tn5 binding, the cells on beads were washed three times with 1 mL Dig-300 Buffer to remove unbound pA-Tn5. To activate pA-Tn5 for tagmentation, the bead-bound cells were resuspended in 100 μL Tagmentation Buffer (20 mM HEPES pH 7.5, 300 mM NaCl, 0.5 mM Spermidine, 1 × Protease inhibitor cocktail,

0.01% Digitonin, 10 mM MgCl₂) then incubated at 37°C for 1 hour in a thermomixer with shaking at 1000 rpm. The tagmentation reaction was terminated by mixing with 500 μL Buffer PB from MinElute PCR Purification Kit (Qiagen). The tubes were then placed on the magnet stand to clear, and the supernatant was transferred into the MinElute column (Qiagen) to purify CUT&Tag-DNA according to the manufacturer's instructions.

Oligonucleotide replacement and gap repair

At this step, the shorter adapter oligonucleotide, which is not covalently linked to the genomic DNA, is replaced by a methylated oligonucleotide (Tn5mC-RepIO1; [Table S1](#)), and the nine-base gap is repaired by combined actions of DNA polymerase and DNA ligase. Briefly, 11 μL purified CUT&Tag-DNA was used in the reaction (11 μL DNA, 2 μL 10 μM Tn5mC-RepIO1 oligo, 2 μL 10× Ampligase buffer (Lucigen), 2 μL dNTP mix (2.5mM each) (Thermo Fisher Scientific)) assembled in a PCR tube. The reaction was first incubated in a PCR thermocycler with the following program; 50°C for 1 min, 45°C for 10 min, ramp down to 37°C at a rate of -0.1°C/sec, then hold at 37°C. Once the program reached 37°C, while the tubes remained in the thermocycler, 1 μL T4 DNA polymerase (NEB) and 2.5 μL Ampligase (Lucigen) were added into each sample separately. The reaction was mixed by pipetting up and down with a P20 micropipette then incubated at 37°C for 30 min. The reaction was stopped by adding 1 μL 0.5 M EDTA (pH = 8.0) then cleaned up with MinElute PCR Purification Kit (Qiagen) following the manufacturer's instructions. The gap-repaired DNA was eluted in 22 μL Buffer EB, and 2 μL of the eluted DNA was saved as a control to assess DNA damage by bisulfite treatment.

Bisulfite conversion

The purified gap-repaired DNA was bisulfite converted with EZ DNA Methylation-Lightning Kit (Zymo Research) according to the manufacturer's instructions. Briefly, 20 μL gap-repaired DNA and 130 μL Lightning Conversion Reagent were mixed then split equally into two PCR tubes (75 μL/tube). The reaction was incubated in a thermocycler as follows: 98°C for 8 min, 54°C for 60 min, then hold at 4°C. Subsequent purification and desulfonation was performed exactly as described in the user manual. Bisulfite converted DNA was eluted in 25 μL M-elution Buffer.

PCR amplification

Bisulfite converted DNA was amplified and barcoded in 50 μL PCR reaction (22 μL bisulfite-converted DNA, 25 μL 2× KAPA HiFi HotStart Uracil+ ReadyMix (Roche), 1.5 μL 10 μM i5 universal PCR primer, 1.5 μL 10 μM i7 barcode PCR primer) with the following PCR program: 98°C for 45 sec; 14 cycles of 98°C for 15 sec, 63°C for 30 sec, 72°C for 30 sec; final extension at 72°C for 2 min; then hold at 4°C. The reserved 2 μL control DNA (without bisulfite conversion) was amplified and barcoded in parallel using the same PCR program in 50 μL PCR reaction (2 μL gap-repaired DNA, 25 μL 2× NEBNext High-Fidelity PCR Master Mix (NEB), 1.5 μL 10 μM i5 universal PCR primer, 1.5 μL 10 μM i7 barcode PCR primer, 20 μL Nuclease-free H₂O). The PCR reactions were cleaned up with AMPure XP beads (Beckman Coulter).

Sequencing

The libraries were quantified with Qubit dsDNA HS Assay Kit (Invitrogen) then pooled and sequenced on Illumina NextSeq550 with paired-end 36-bp or 76-bp reads.

CUT&Tag-BS data analysis

Raw sequencing reads were filtered to remove any read pairs with mean base quality score less than 20 at either read end. Adapters were removed via Cutadapt v1.12 ([Martin, 2011](#)) with parameters “-a CTGTCTCTTATACAC -A CTGTCTCTTATACAC -O 5 -q 0 -m 20 -p”. For any read pairs in which adapter was identified and removed, an additional nine bases were trimmed from the 3' end of read1, as they correspond to the 9-base gap region filled in with unmethylated nucleotides. Genomic alignment was performed by Bismark v0.23.0 ([Krueger and Andrews, 2011](#)) with parameters “-X 1000 -non_bs_mm” (all other parameters as default) against the mm10 reference assembly (GRCm38) with Bowtie2 v2.3.0 ([Langmead and Salzberg, 2012](#)) as the underlying mapper. Positional methylation bias information was reported by the Bismark v0.23.0 bismark_methylation_extractor tool with parameters “-p -include_overlap -mbias_only”. Duplicate mapped fragments were removed by the Bismark v0.23.0 deduplicate_bismark tool with parameter “-p”. The observed insert size distribution per library was determined by Picard tools v1.110 CollectInsertSizeMetrics.jar (<http://broadinstitute.github.io/picard>). Per-residue methylation data was collected by the Bismark v0.23.0 bismark_methylation_extractor tool with parameters “-p -ignore_r2 9 -comprehensive -mbias_off -bedGraph -cytosine_report”. The bisulfite conversion rate per library was calculated by assessing methylated and unmethylated cytosine counts at the first 9 bases of read2.

Peak calling and peak overlap between samples

HOMER v4.10.3 ([Heinz et al., 2010](#)) was used to call peaks per sample as follows. Mapped read data was prepared using the makeTagDirectory function with parameters “-format sam -read1-keepAll -fragLength 200”, followed by peak calls using the findPeaks function with parameters “-size 500 -minDist 1000 -L 0 -region” for H3K4me1 and parameters “-size 1000 -minDist 2500 -L 0 -region” for H3K9me3 and H3K27me3. Peaks overlapping mm10 blacklist regions (<https://github.com/Boyle-Lab/Blacklist/tree/master/lists>) were discarded. For each histone mark, a unified peak set was generated by BEDtools v2.29.2 merge ([Quinlan and Hall, 2010](#)) as the union of peak calls from the CUT&Tag-BS library and the corresponding non-bisulfite-treated control. Analysis of shared peaks between samples was performed with subsets of unified peaks that overlapped the HOMER peak calls from given samples. Briefly, the unified peak set was intersected individually with HOMER peak calls from each sample in comparison, and the subset of unified peaks in the intersection was used to determine peak overlap between samples.

Peak signal quantification

To quantify signal intensity at peaks, the mapped paired-end hits were converted to a single fragment via BEDtools v2.29.2 `bamtoBed` with the “-bedpe” option. The number of fragments overlapping each unified peak was collected via BEDtools v2.29.2 `coverage` with the “-counts” option, and then subsequently converted to CPM (counts per million).

Overlap of peaks with chromatin states

Chromatin states of mESC (Pintacuda et al., 2017) as defined by ChromHMM were downloaded from https://github.com/gui Fengwei/ChromHMM_mESC_mm10. Overlap of the detected peaks with chromatin states was determined by BEDtools v2.29.2 `intersect`. In cases of peaks overlapping multiple chromatin states, each peak was assigned to the chromatin state with which it showed the most overlap.

Methylation concordance and methylation correlation

Methylation concordance across reads was calculated by summing up methylated and unmethylated cytosine counts per fragment identifier from the CpG context output file generated by the Bismark v0.23.0 `bismark_methylation_extractor` tool.

To assess correlation of CpG methylation as a function of genomic distance, all pairs of CpG sites within 2Kb were identified for which at least one CpG was within the unified peak regions of the given histone mark. The set of CpG pairs was further filtered to retain only those with at least 5× coverage for both sites. Pearson and Spearman correlations of methylation level between pairs of CpGs at a given genomic distance (2 bp to 2001 bp) were calculated by `cor.test` in R 3.5.0 (<https://www.r-project.org/>).

ChIP-seq data processing

Publicly available ChIP-seq data in mESCs for H3K4me1 (Feldmann et al., 2020), H3K9me3 (Muller et al., 2021), and H3K27me3 (Gao et al., 2018) were downloaded as raw FASTQ files from GEO with SRA Toolkit v2.11.0 `fasterq-dump` (<http://ncbi.github.io/sra-tools/>). All reads were clipped to 35 bp in length (H3K4me1) or 70 bp in length (H3K9me3 and H3K27me3) to excise regions of questionable nucleotide content or quality. Reads were then filtered to remove those with mean base quality score less than 20. Genomic mapping against the mm10 reference assembly (GRCm38) was performed by Bowtie2 v2.3.0 with parameters “-local -sensitive-local”; additional parameters “-X 1000 -fr” were used with paired-end samples. For samples with multiple alignment files per library, files were combined with Picard tools v1.110 `MergeSamFiles.jar`. Alignments were then filtered by samtools v1.3.1 (Li et al., 2009) to retain only hits with MAPQ score of at least 5; for paired-end samples, retained hits were also required to be properly paired. Duplicate reads were removed by Picard tools v1.110 `MarkDuplicates.jar` with the `REMOVE_DUPLICATES=TRUE` setting.

WGBS data processing

Publicly available WGBS data in mESCs (Lu et al., 2014) was downloaded in raw FASTQ format from GEO with SRA Toolkit v2.11.0 `fasterq-dump` (<http://ncbi.github.io/sra-tools/>). Raw reads pairs were filtered to remove any with mean base quality score less than 20 at either read end. Genomic alignment was performed by Bismark v0.23.0 with parameters “-X 1000 -non_bs_mm” (all other parameters as default) with Bowtie2 v2.3.0 as the underlying mapper; here, the reference genome index includes the mm10 reference assembly (GRCm38) with the genome sequence of enterobacteria phage λ (NC_001416.1) appended for the purpose of calculating the bisulfite conversion rate. Duplicate mapped fragments were removed by the Bismark v0.23.0 `deduplicate_bismark` tool with parameter “-p”. Per-residue methylation data was collected by the Bismark v0.23.0 `bismark_methylation_extractor` tool with parameters “-p -ignore_r2 2 -comprehensive -mbias_off -bedGraph -cytosine_report”.

QUANTIFICATION AND STATISTICAL ANALYSIS

Pearson correlation coefficient (Pearson’s r) were calculated using R (<https://www.r-project.org/>).