nature biotechnology

Quantitative comparison of genome-wide DNA methylation mapping technologies

Christoph Bock[1–4,6], Eleni M Tomazou[1–3,6], Arie B Brinkman[5], Fabian Müller[1–4], Femke Simmer[5], Hongcang Gu[1], Natalie Jäger[1–3], Andreas Gnirke[1], Hendrik G Stunnenberg[5] & Alexander Meissner[1–3]

DNA methylation plays a key role in regulating eukaryotic gene expression. Although mitotically heritable and stable over time, patterns of DNA methylation frequently change in response to cell differentiation, disease and environmental influences. Several methods have been developed to map DNA methylation on a genomic scale. Here, we benchmark four of these approaches by analyzing two human embryonic stem cell lines derived from genetically unrelated embryos and a matched pair of colon tumor and adjacent normal colon tissue obtained from the same donor. Our analysis reveals that methylated DNA immunoprecipitation sequencing (MeDIP-seq), methylated DNA capture by affinity purification (MethylCap-seq), reduced representation bisulfite sequencing (RRBS) and the Infinium HumanMethylation27 assay all produce accurate DNA methylation data. However, these methods differ in their ability to detect differentially methylated regions between pairs of samples. We highlight strengths and weaknesses of the four methods and give practical recommendations for the design of epigenomic case-control studies.

DNA methylation is a common mechanism of epigenetic regulation in eukaryotes. It occurs most frequently at cytosines that are followed by guanines (CpG). High levels of DNA methylation in promoter regions are typically associated with robust gene silencing[1]. Twenty-five years of research on cancer epigenetics have firmly established the prevalence of aberrant DNA methylation in cancer cells[2–6]. Moreover, recent studies have investigated the role of DNA methylation for neural and autoimmune diseases, its correlation with physiological conditions and its response to environmental influences[7–9]. Comprehensive mapping of DNA methylation in relevant clinical cohorts is likely to identify new disease genes and potential drug targets, help to establish the relevance of epigenetic alterations in disease and provide a rich source of potential biomarkers[10]. DNA methylation mapping could also facilitate quality control of cultured cells by exploiting the fact that cell states and differentiation potential of stem cells are reflected in their DNA methylation patterns[11].

Several methods have been developed to map DNA methylation on a genomic scale. Most of these methods combine DNA analysis by microarrays or high-throughput sequencing with one of four ways of translating DNA methylation patterns into DNA sequence information or library enrichment. (i) MeDIP-seq uses an antibody that is specific for 5-methylcytosine to retrieve methylated fragments from sonicated DNA[12,13]. (ii) MethylCap-seq employs a methyl-binding domain protein to obtain DNA fractions with similar methylation levels[14–16]. (iii) Bisulfite-based methods use a chemical reaction that selectively converts unmethylated, but not methylated, cytosines into uracils, thus introducing methylation-specific, single nucleotide polymorphisms into the DNA sequence[11,17,18]. (iv) Methylation sensitive digestion uses prokaryotic restriction enzymes to selectively fractionate only methylated or only unmethylated DNA[19–21].

The diversity of methods to map DNA methylation and the absence of an uncontested commercial market leader raise questions about each method's strengths and weaknesses—questions that researchers have to answer for themselves when selecting the most appropriate technology for any given project. The goal of this study was to comprehensively evaluate four popular methods—MeDIP-seq[12], MethylCap-seq[14], RRBS[22] and the Infinium HumanMethylation27 assay[17] with a special emphasis on their practical utility for biomedical research and biomarker development. All four methods are relatively easy to set up because detailed protocols have been published and/or commercial kits are available. We chose RRBS because it targets bisulfite sequencing to a well-defined set of genomic regions with moderate to high CpG density[22], which makes RRBS substantially more cost efficient than genome-wide bisulfite sequencing. The Infinium HumanMethylation27 assay, also a bisulfite-based method, was included because of its wide use and easy integration with existing genotyping pipelines; it is the only microarray-based method in our comparison. Methods that use tiling microarrays were excluded because they have been benchmarked previously[20] and because next-generation sequencing enables higher resolution and/or higher genomic coverage at competitive cost. Methylation-specific digestion was excluded because no algorithm exists that could accurately infer quantitative DNA methylation data from digested read frequencies[19]. An outline of the experimental and analytical procedure of this technology comparison is shown in **Figure 1**.

[1]Broad Institute, Cambridge, Massachusetts, USA. [2]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, Massachusetts, USA. [3]Harvard Stem Cell Institute, Cambridge, Massachusetts, USA. [4]Max Planck Institute for Informatics, Saarbrücken, Germany. [5]Radboud University Department of Molecular Biology, Nijmegen Center for Molecular Life Sciences, Nijmegen, The Netherlands. [6]These authors contributed equally to this work. Correspondence should be addressed to C.B. (cbock@broadinstitute.org) or A.M. (alexander_meissner@harvard.edu).

Published online 19 September 2010; doi:10.1038/nbt.1681

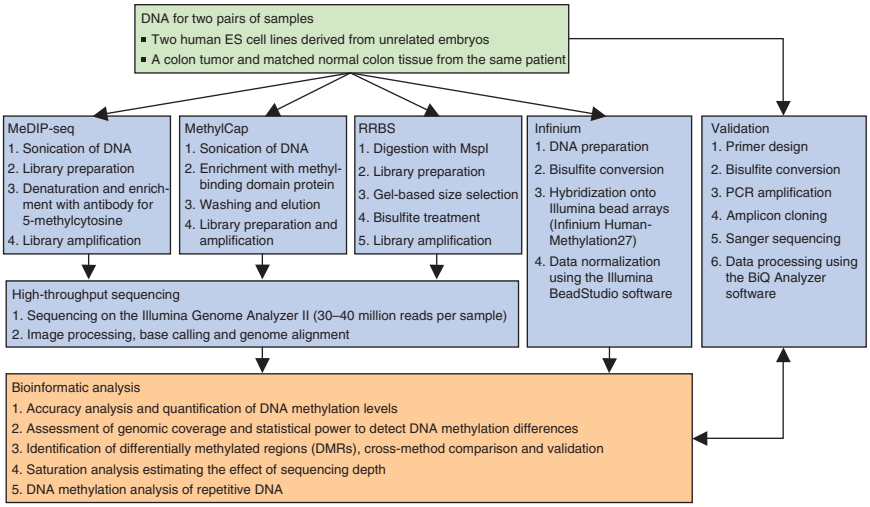

Figure 1 Outline of the DNA methylation technology comparison. Four methods for DNA methylation mapping were compared on two pairs of samples. The resulting 16 DNA methylation maps were bioinformatically analyzed and benchmarked against each other. In addition, clonal bisulfite sequencing was performed on selected genomic regions to validate DNA methylation differences that were detected exclusively by one method.

RESULTS

DNA methylation mapping by four methods

Genome-wide DNA methylation mapping is most commonly used as a discovery tool to identify differentially methylated regions (DMRs) as candidates for further research. Typical examples are cancer-specific DMRs, which are increasingly used as biomarkers for cancer diagnosis and therapy optimization[10]. To emulate the case-control approach that is widely used for epigenetic biomarker development, we focused on sample pairs that we statistically compare with each other. Specifically, we selected two human embryonic stem (ES) cell lines that were derived from genetically unrelated embryos[23], and a matched pair of colon tumor and adjacent normal colon tissue obtained from the same donor. We applied each of the four methods (MeDIP-seq, MethylCap-seq, RRBS, Infinium) to all four samples (HUES6 ES cells, HUES8 ES cells, colon tumor and matched normal colon tissue), generating a total of 16 genome-scale DNA methylation maps. All data were processed with a standardized bioinformatic pipeline, and the technical data quality turned out to be similarly high across all samples and methods (**Table 1**).

When plotting the DNA methylation data as genome browser tracks, we found excellent visual agreement between all four methods (**Fig. 2**; tracks are available online for interactive browsing: http://meth-benchmark.computational-epigenetics.org/). MeDIP-seq and MethylCap-seq gave rise to peaks of methylated DNA that were similar in shape, size and location, indicating that MeDIP-seq's monoclonal antibody and MethylCap-seq's methyl-binding domain enrich for similar DNA fragments. However, MeDIP-seq exhibited higher baseline levels and lower peak heights than MethylCap-seq. This smaller dynamic range is already apparent from **Figure 2** (note the different scale of the *y* axis) and becomes more obvious when plotting MeDIP and MethylCap-seq tracks along an entire chromosome (**Supplementary Fig. 1**). This observation was quantitatively confirmed by plotting the mean read frequency for enriched and depleted fractions of the genome (**Supplementary Fig. 2**). We also observed high visual agreement between RRBS and Infinium, with the limitation that Infinium covers two orders of magnitude fewer CpGs than RRBS (**Table 1**). Finally, the bisulfite-based methods (RRBS, Infinium) generally confirm the results of the enrichment-based methods (MeDIP, MethylCap-seq), although there are deviations in repeat-rich as well as in CpG-poor genomic regions (**Supplementary Fig. 3**).

Accuracy of DNA methylation mapping

For a more quantitative assessment of measurement accuracy, we compared the results of the three sequencing-based methods (MeDIP-seq, MethylCap-seq, RRBS) with the Infinium HumanMethylation27 assay as a common reference (**Fig. 3**). The Infinium assay was used as reference because its quantitative accuracy has been established in previous studies[17,24], which reported correlation coefficients around 0.9 relative to the GoldenGate and MethyLight assays. Note, however, that the probes of the Infinium assay cover only a small percentage of all CpGs in the genome and are preferentially located in unmethylated promoter regions. To compensate for this potential source of bias, we calculated two correlation coefficients, one across the entire spectrum of methylation levels and the other focusing only on those CpGs that exhibit at least 20% methylation according to the Infinium assay.

RRBS and Infinium data can be compared directly and without normalization, because both methods measure absolute DNA methylation levels. For a total of 5,088 single CpGs that were covered by both an Infinium probe and at least five RRBS reads, we observed a Pearson correlation of 0.92 across all DNA methylation levels and a Pearson correlation of 0.83 when we excluded unmethylated CpGs. Because neighboring CpGs tend to exhibit highly correlated DNA methylation levels[18,25], we also evaluated the correlation for RRBS measurement averages over a 200-base pair (bp) sequence window around each Infinium probe. Again, we observed excellent agreement between the two methods (**Fig. 3c**), with an overall Pearson correlation of 0.92 across all DNA methylation levels and a Pearson correlation of 0.84 when we excluded unmethylated CpGs. This second comparison supports the hypothesis that a single-CpG measurement can often act as an indicator of the DNA methylation levels at neighboring, unmeasured CpGs.

Comparison with MeDIP-seq and MethylCap-seq is less straightforward because both methods measure the relative enrichment of methylated DNA rather than absolute DNA methylation levels. When we correlated the number of sequencing reads per 1-kb region with the DNA methylation measurements of the Infinium assay, the Pearson correlation did not exceed 0.6 across all DNA methylation levels and 0.4 when we excluded unmethylated CpGs (**Supplementary Fig. 3a**,**b**). High density of repetitive DNA was identified as a major source of spurious read enrichment in regions with low absolute DNA methylation levels. In contrast, low CpG density gave rise to low read numbers in regions with high levels of DNA methylation (**Supplementary Fig. 3c**,**d**). The confounding effect of DNA sequence is also visible in **Figure 2**. Low read counts can indicate either the relative absence of CpGs (e.g., region 1 in **Fig. 2**) or the absence of DNA methylation in the presence of CpGs (**Fig. 2**, region 2); and strong peaks can occur in genomic regions that are incompletely methylated if the CpG density is sufficiently high to give rise to substantial read enrichment (**Fig. 2**, region 3).

Table 1 Summary of DNA methylation mapping experiments

Run no.	Method	Sample name	Number of lanes[a]	Number of reads (total)	Number of reads (aligned)	Alignment rate	Number of reads (unique)	Number of reads (duplicates)	Unique read rate[b]
1	MeDIP-seq	HUES6 ES cell line	2	37,086,239	22,798,831	61.5%	12,849,623	9,949,208	56.4%
2	MeDIP-seq	HUES8 ES cell line	2	36,078,308	24,266,670	67.3%	12,287,174	11,979,496	50.6%
3	MeDIP-seq	Primary colon tumor	2	33,453,797	18,582,183	55.5%	7,006,484	11,575,699	37.7%
4	MeDIP-seq	Matched normal colon tissue	2	37,789,936	21,793,567	57.7%	10,360,103	11,433,464	47.5%
5	MethylCap-seq	HUES6 ES cell line	3	38,436,495	23,401,511	60.9%	21,712,433	1,689,078	92.8%
6	MethylCap-seq	HUES8 ES cell line	3	38,735,596	21,670,301	55.9%	19,585,988	2,084,313	90.4%
7	MethylCap-seq	Primary colon tumor	3	37,718,830	23,206,054	61.5%	21,600,129	1,605,925	93.1%
8	MethylCap-seq	Matched normal colon tissue	3	38,330,519	22,724,002	59.3%	21,290,282	1,433,720	93.7%
							Number of CpGs (total)	Number of CpGs (unique)	Mean CpG coverage
9	RRBS	HUES6 ES cell line	2	30,004,147	12,150,905	40.5%	22,181,147	2,181,128	10.2x
10	RRBS	HUES8 ES cell line	2	28,395,040	12,670,034	44.6%	29,704,332	2,185,751	13.6x
11	RRBS	Primary colon tumor	4[c]	40,015,958	9,545,423	23.9%	16,891,325	1,297,296	13.0x
12	RRBS	Matched normal colon tissue	4[c]	32,072,287	6,214,732	19.4%	10,190,227	1,134,963	9.0x
			Number of arrays	Number of CpGs (total)	Number of CpGs (valid)	Number of CpGs (unique)	Valid probe rate		
13	Infinium	HUES6 ES cell line	1	27,578	27,192	27,192	98.6%		
14	Infinium	HUES8 ES cell line	1	27,578	27,090	27,090	98.2%		
15	Infinium	Primary colon tumor	1	27,578	27,561	27,561	99.9%		
16	Infinium	Matched normal colon tissue	1	27,578	27,478	27,478	99.6%		

[a]All sequencing was performed in 2009 using the Illumina Genome Analyzer II (36-bp, single-end reads). As of June 2010, we routinely observe total read numbers per lane averaging ~40 million for MeDIP-seq and MethylCap-seq and close to 30 million for RRBS. Current alignment rates range from 60% to 80% for all three methods.[b]The unique read rate was calculated by dividing the number of reads that map to a unique position in the genome (defined by chromosome, read start position and strand) by the total number of aligned reads.[c]Samples 11 and 12 were part of a sequencing-optimization run that resulted in lower sequencing yield and reduced alignment rates. Four lanes were sequenced to reach the target of 30–40 million reads per sample and method.

It has previously been reported that statistical correction for CpG density can improve the quantification of DNA methylation levels based on MeDIP-seq data[12,26]. We therefore constructed a linear regression model that corrects for the confounding effect of DNA sequence, and we observe substantially improved results (**Fig. 3a**,**b**). Across all DNA methylation levels the correlation between the statistically corrected read counts and the DNA methylation measurements of the Infinium assay amounted to 0.84 for MeDIP-seq and to 0.88 for MethylCap-seq. However, the correlations dropped to 0.57 (MeDIP-seq) and 0.66 (MethylCap-seq) when we excluded unmethylated CpGs. These results indicate that MeDIP-seq and MethylCap-seq can distinguish between methylated and unmethylated regions almost as precisely as RRBS, but are less accurate for quantifying the DNA methylation levels in partially methylated genomic regions.

Genomic coverage of DNA methylation mapping

The single-bp resolution of the two bisulfite-based methods comes at the cost of reduced genomic coverage compared to the two enrichment-based methods. RRBS reads cover less than 10% of the 28 million CpGs in the human genome and Infinium is by design restricted to 27,578 promoter-associated CpGs (**Table 1**). In contrast, MeDIP-seq and MethylCap-seq are theoretically able to identify methylated genomic regions located anywhere in the genome, although they too are subject to intrinsic limitations[27]. To assess the empirical genomic coverage of each method, we calculated the number of reads (MeDIP-seq, MethylCap-seq) or CpG methylation measurements (RRBS, Infinium) for each of the following genomic regions: (i) CpG islands, (ii) gene promoters and (iii) a 1-kb tiling of the genome. The results are shown in **Figure 4**, and coverage details for a total of 13 types of genomic regions are available online (http://meth-benchmark.computational-epigenetics.org/).

As expected, MeDIP-seq and MethylCap-seq provide broad coverage of the genome, whereas RRBS and Infinium are more restricted to CpG islands and promoter regions. However, the practically relevant differences in genomic coverage are lower than **Figure 4** may suggest. This is because a minimum number of reads are required in at least one sample to reliably detect differential methylation among a given pair of samples. We illustrate this point by two statistical power calculations, which were performed with G*Power 3 (ref. 28). Assume that a genomic region is covered by five MeDIP-seq or MethylCap-seq reads in one sample. Then it has to contain at least 20 reads in the second sample to be detected as hypermethylated (assuming a statistical power of 80% and a *P*-value of 5% without multiple-testing correction). Similarly, RRBS would detect a DNA methylation increase from 30% to 70% only when at least 25 measurements are available in each sample (again assuming a statistical power of 80% and a *P*-value of 5% without multiple-testing correction).

Identification of differentially methylated regions

Genome-wide DNA methylation mapping is most commonly used for detecting DNA methylation differences, for example, between diseased and healthy tissue or between genetically modified and unmodified control cells. To assess how well MeDIP-seq, MethylCap-seq and RRBS perform on this task, we developed a bioinformatic method that identifies statistically significant DMRs from multiple types of sequencing data (the Infinium assay requires a different approach and is discussed in a separate section below). For a predefined set of genomic regions we count the numbers of sequenced reads (for MeDIP-seq and MethylCap-seq) or, alternatively, the numbers of methylated versus unmethylated CpGs (for RRBS), and we test for statistically significant differences between two samples using Fisher's exact test. When applied to a complete tiling of the human genome, this method performs genome-wide DMR detection. Alternatively, it can be targeted to specific region types such as CpG islands, gene promoters or putative enhancers, which often leads to more sensitive detection of small differences because the multiple-testing burden is reduced compared to genome-wide DMR detection. We pursued both the unbiased and the annotation-guided approach in parallel, focusing our comparison on three types of genomic regions: (i) CpG islands, (ii) gene promoters and (iii) a 1-kb tiling of the genome (**Fig. 5** and **Supplementary Figs. 4–8**).

Overall, we observed high correlation for each of the two sample pairs, but also outliers suggesting the presence of DMRs. Based on the

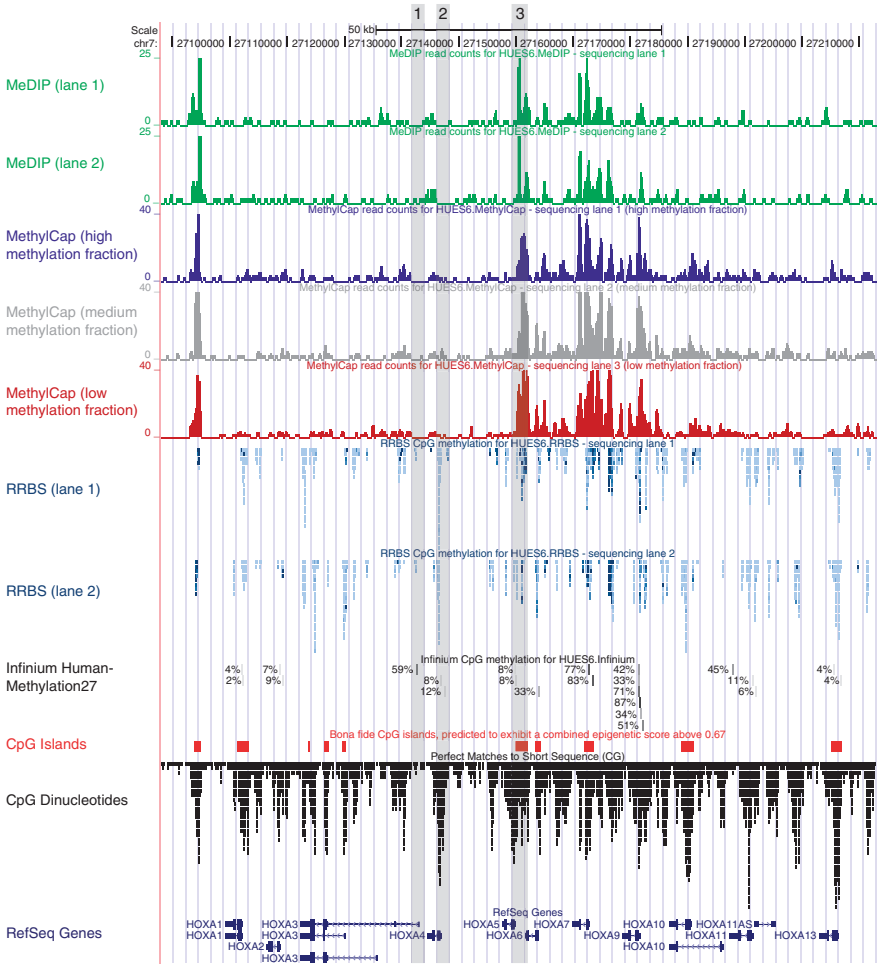

Figure 2 Comparison of DNA methylation maps obtained with four different methods. The screenshot shows genome browser tracks for MeDIP-seq (first two tracks, in green), MethylCap-seq (three tracks in blue, gray and red), RRBS (stacked light blue tracks) and Infinium (single black track with percentage values) across the *HOXA* cluster in a human ES cell line (HUES6). Each track represents data from a single sequencing lane (MeDIP-seq, MethylCap-seq, RRBS) or microarray hybridization (Infinium). MeDIP-seq and MethylCap-seq data are visually similar to ChIP-seq data, with peaks in regions that show high density of the target molecule (5-methylcytosine) and troughs in regions with low density of methylated cytosines. The heights of the peaks represents the number of reads in each genomic interval, for each track normalized to the same genome-wide read count. RRBS gives rise to clusters of CpGs with absolute DNA methylation measurements, separated by regions that are not covered due to the reduced-representation property of the RRBS protocol. Each data point corresponds to the methylation level at a single CpG, and dark blue points indicate higher methylation levels than light blue points. Infinium data is represented in a similar way to the RRBS data, and the methylation levels at single CpGs are shown as percentage values. For reference, the CpG density is indicated by stacked points (black) at the bottom of the diagram, and CpG islands (red) as well as known genes (blue) are listed as described previously[55,56].

RRBS data, we obtained Pearson correlations around 0.9 for all three region types, both between the two ES cell lines (HUES6 and HUES8) and between the colon tumor and matched normal colon tissue. For MethylCap-seq and MeDIP-seq, the correlations were somewhat lower and ranged from 0.75 to 0.92 (**Fig. 5** and **Supplementary Figs. 4–8**). Using the DMR detection algorithm (Online Methods), we identified several hundred to several thousand DMRs in both sample pairs. There was substantial, but by no means perfect, overlap between the DMRs identified by all three methods. For the two human ES cell lines, 277 out of 44,440 CpG islands were detected as differentially methylated by each of the three methods (**Fig. 5d**). Pairwise comparisons for each sample and region type (**Supplementary Figs. 4–8**) confirmed that the agreement between the three methods was statistically significant in all cases ($P < 0.01$, Fisher's exact test). In total, we observed that up to 1,000 CpG islands, 405 promoter regions or 1,924 of the 1-kilobase tiling regions (that is, <0.1% of the genome) were detected as differentially methylated by at least two methods. Note, however, that it is not possible to combine these values into a single sum of DMRs because many CpG islands overlap with promoter regions and every CpG island and promoter region overlaps with at least one tiling region. Nor does the number of differentially methylated tiling regions provide an accurate estimate of the 'true' number of DMRs because a sizable number of DMRs are no longer statistically significant when split into 1-kb regions. Despite these conceptual difficulties, which preclude us from giving a single 'true' number of DMRs for each sample pair, our data clearly indicate that—on average—MethylCap-seq identifies more DMRs than RRBS, and MeDIP-seq identifies the fewest DMRs. This order was observed not only when we focused on the total number of DMRs per method, but also when we considered only those DMRs that were detected by at least two methods, indicating that the comparison is not distorted by high numbers of method-specific artifacts.

Validation of method-specific DMRs

To pinpoint potential problems of MeDIP-seq, MethylCap-seq or RRBS, we manually inspected a large number of regions that were identified as significant DMRs by only one method. The most common reasons why DMRs identified by one method were missed by the other methods were insufficient genomic coverage (RRBS, Infinium) and low read numbers conferring insufficient statistical power to detect differential DNA methylation (MeDIP-seq, MethylCap-seq). No cases were identified in which the RRBS and Infinium data were in direct contradiction with each other. However, we could identify a few cases in which MeDIP-seq or MethylCap-seq were inconsistent with RRBS and/or Infinium data. These were almost exclusively located in repetitive regions, indicating that high copy-number repeats can amplify minor differences in the efficiency of methylated DNA enrichment and give rise to a small number of spurious DMRs. In contrast, RRBS seems more robust toward such fluctuations because it measures DNA methylation based on the DNA sequence of the reads in a given region, rather than on their read frequency. We also assessed whether copy-number variation was a major confounding factor for DMR discovery. This does not seem to be the case for our data. The vast majority of DMRs were shorter than 10 kb (**Supplementary Fig. 9**), whereas it is not uncommon for cancer-specific as well as germline-transmitted copy-number variations to extend for much longer distances[29,30].

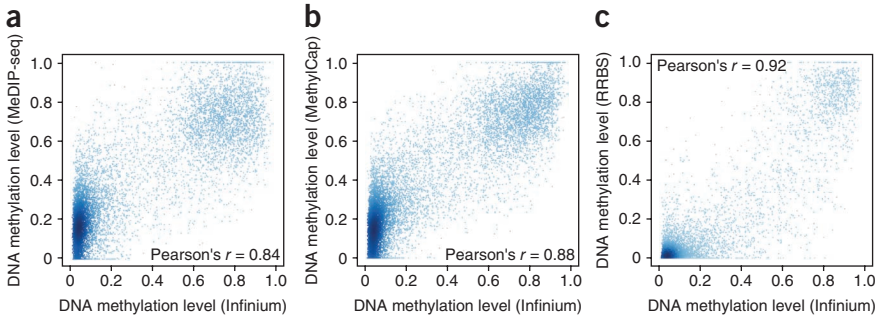

Figure 3 Quantification of DNA methylation with MeDIP-seq, MethylCap-seq and RRBS. (**a**–**c**) Absolute DNA methylation levels were calculated from the data obtained by MeDIP-seq (**a**), MethylCap-seq (**b**) and RRBS (**c**), respectively, and compared to DNA methylation levels determined by the Infinium assay. For MeDIP-seq and MethylCap-seq, sequencing reads were counted in 1-kb regions surrounding each CpG that is interrogated by the Infinium assay, and a regression model was used to infer absolute DNA methylation levels. Scatter plots and correlation coefficients were calculated on a test set that was not used for model fitting or feature selection. For RRBS, the DNA methylation level was determined as the percentage of methylated CpGs within 200 bp surrounding each CpG that is interrogated by the Infinium assay. Data shown are for the HUES6 human ES cell line, and regions that did not have sufficient sequencing coverage were excluded.

As an additional validation, we selected eight method-specific DMRs based on the ES cell comparison, and we investigated DNA methylation patterns in the two ES cell lines by clonal bisulfite sequencing (**Table 2**). These genomic regions were handpicked such that one method clearly identified them as DMRs whereas the two other methods did not show a trend in either direction. Note that this preselection makes the validation substantially harder than confirming randomly selected DMRs, because the magnitude of the DNA methylation difference tends to be lower for method-specific DMRs than for DMRs that are detected by multiple methods. As an additional complication, some of the selected DMRs are highly repetitive or overlap with known copy-number variations. Sequencing an average of 11 clones per sample and region we were able to confirm three out of three MethylCap-seq–specific DMRs and two out of two RRBS-specific DMRs. In contrast, two MeDIP-seq–specific DMRs could not be confirmed, and for the third region the agreement was marginal (**Table 2** and **Supplementary Data 1**).

To assess the practical relevance of the method-specific differences, we asked whether biologically interesting hits were missed by any of the three methods. For this analysis we focused on the colon samples because of the large number of genes with a known or suspected role in colon cancer. Our results show that several interesting DMRs are detected by all methods, including tumor-specific hypermethylation in the promoters of *GATA2* (ref. 31) and *GATA5* (ref. 32). However, a considerable number of interesting DMRs were missed by MeDIP-seq, whereas MethylCap-seq and RRBS both detected those regions; these include tumor-specific hypermethylation in the promoter regions of *SOX17* (ref. 33), *POU2AF1* (ref. 34) and *SEPT9* (ref. 35). Somewhat more rarely, we also observed interesting DMRs being missed by MethylCap-seq or RRBS. For example, MethylCap-seq overlooked tumor-specific hypermethylation at the promoter of *SFRP1* (ref. 36), and RRBS missed tumor-specific hypermethylation at the promoter of *DKK2* (ref. 37).

The effect of sequencing depth on mapping performance

MeDIP-seq, MethylCap-seq and RRBS use DNA sequencing as a way of counting DNA fragments to determine the percentage of methylation-enriched reads that align to specific genomic regions (MeDIP-seq, MethylCap-seq) or to calculate the ratio of methylated and unmethylated cytosines at single CpGs (RRBS). Conceptually, sequencing can be thought of as random sampling from a large pool of DNA fragments. It is therefore expected that the performance of these methods increases when sequencing more DNA fragments, until it levels off as the sequencing depth approaches saturation. To quantify

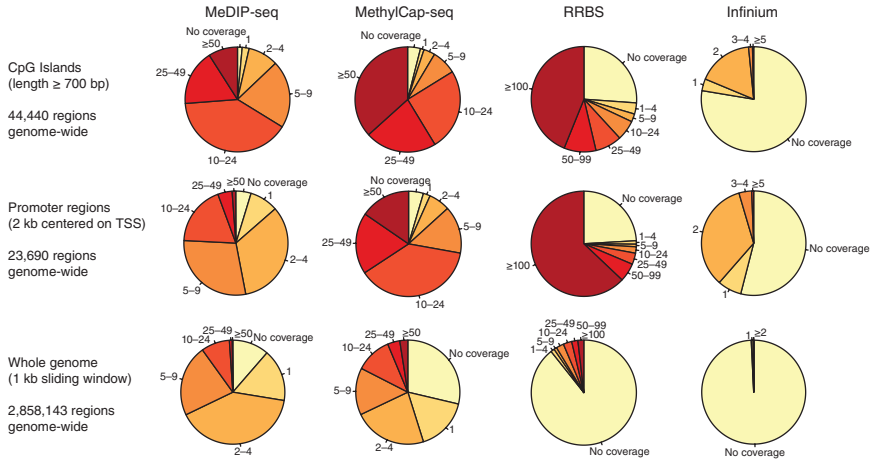

Figure 4 Genomic coverage of MeDIP-seq, MethylCap-seq, RRBS and Infinium. Genomic coverage was quantified by the number of DNA methylation measurements that overlap with CpG islands (top row), gene promoters (center row) and a 1-kb tiling of the genome (bottom row). For MeDIP-seq and MethylCap-seq, the number of measurements is equal to the number of unique sequencing reads that fall inside each region. For RRBS, it refers to the number of valid DNA methylation measurements at CpGs within each region (one RRBS sequencing read typically yields one measurement, but can also give rise to more than one measurement if it contains several CpGs). For Infinium, the number of measurements is equal to the number of CpGs within each region that are present on the HumanMethylation27 microarray. CpG islands were calculated using CgiHunter (http://cgihunter.bioinf.mpi-inf.mpg.de/), requiring a minimum CpG observed versus expected ratio of 0.6, a minimum GC content of 0.5 and a minimum length of 700 bp[55]. Promoter regions were calculated based on Ensembl gene annotations, such that the region starts 1 kb upstream of the annotated transcription start site (TSS) and extends to 1 kb downstream of the TSS. The genomic tiling was obtained by sliding a 1-kb window through the genome such that each tile starts at the position where the previous tile ends. No repeat-masking was performed for any of the three types of genomic regions. Data are shown for the HUES6 human ES cell line.

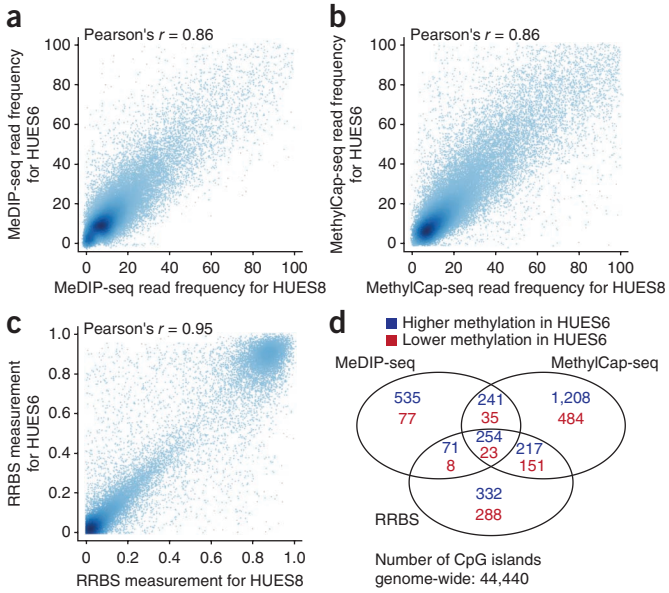

Figure 5 Detection of DMRs with MeDIP-seq, MethylCap-seq and RRBS. Average DNA methylation measurements were calculated for each CpG island and compared between two human ES cell lines (HUES6 and HUES8). (**a**–**c**) Total read frequencies are shown for MeDIP-seq (**a**) and MethylCap-seq (**b**), and mean DNA methylation levels are shown for RRBS (**c**). Regions with insufficient sequencing coverage were excluded. (**d**) The Venn diagram displays the total number and mutual overlap of differentially methylated CpG islands that could be identified by each method. CpG islands were classified as hypermethylated or hypomethylated (depending on the directionality of the difference) if the absolute DNA methylation difference exceeded 20 percentage points (for RRBS) or if there was at least a twofold difference in read number between the two samples (for MeDIP-seq and MethylCap-seq)—but only if Fisher's exact test with multiple-testing correction gave rise to an estimated false-discovery rate of differential DNA methylation that was <0.1.

this effect, we repeated the accuracy analysis (**Fig. 3**) and the DMR detection (**Fig. 5**) on randomly sampled subsets of sequencing reads. First, we benchmarked each method against the Infinium data, assessing their ability to quantify DNA methylation levels based on reduced read numbers (**Supplementary Fig. 10**). The results show that all three methods give rise to accurate DNA methylation measurements based on as little as 20% of the total read coverage, and almost no improvement was observed between 50% and 100% sequencing depth. Although these data suggest that relatively low sequencing depths are often sufficient for obtaining accurate DNA methylation levels, this cannot be generalized to the entire genome. Infinium probes tend to be located in CpG-rich genomic regions, which are also preferentially covered by MeDIP-seq, MethylCap-seq and RRBS measurements (**Fig. 4**), such that saturation is reached earlier in the vicinity of Infinium probes than in CpG-poor genomic regions.

Second, we tested how many DMRs were still detected among the two sample pairs when the number of sequencing reads in each of the samples was reduced (**Supplementary Fig. 11**). For MeDIP-seq, the number of detected DMRs dropped to less than half when the sequencing depth was reduced to 50%, and there was little indication that the number of MeDIP-seq DMRs approaches saturation even at the highest sequencing depth. For MethylCap-seq the decrease in the number of detected DMRs is less dramatic and there is a trend toward saturation. RRBS quickly approaches saturation especially for the ES-cell comparison (**Supplementary Fig. 11**). Overall, the saturation analysis reinforced a conceptual difference between RRBS on the one hand and MeDIP-seq and MethylCap-seq on the other hand. In RRBS, all sequencing is focused on a well-defined, CpG-rich 'reduced representation' of the genome, which leads to relatively early saturation but limited coverage of DMRs in CpG-poor genomic regions. In contrast, MeDIP-seq and MethylCap-seq reads are widely distributed over the genome (albeit with a significant tendency toward high coverage in CpG-rich regions), and deep sequencing increasingly uncovers weak DMRs located in CpG-poor genomic regions.

DNA methylation mapping of repetitive DNA

DNA methylation differences in repetitive regions have frequently been ignored by genome-wide studies, due to technical difficulties such as ambiguous read alignment (for sequencing) and cross-hybridization (for microarrays). This is unfortunate given that loss of DNA methylation in repetitive DNA was the first epigenetic alteration shown to play a role in cancer[4] and has been an area of active research ever since[38]. In the current study, we explored two complementary approaches to test for repeat-associated DNA methylation differences. First, we included repetitive regions alongside nonrepetitive regions in the DMR detection described above (**Fig. 5** and **Supplementary Figs. 4**–**8**), rather than discarding all sequencing reads that map to repetitive portions of the genome. It was thus possible to identify repeat-associated DMRs in a similar way as nonrepetitive DMRs, and we could validate several such cases by clonal bisulfite sequencing (**Table 2**). However, the focus on specific genomic regions makes it difficult to detect global trends that affect certain repeat classes independent of their exact location in the genome. We therefore developed a second approach, which was motivated by the common origin of many repetitive regions from a small number of retrotransposons. The basic concept was to align sequencing reads to prototypic sequences (e.g., of Alu and L1 elements) to obtain DNA methylation measures per repeat class rather than per repeat instance.

To that end, we obtained a manually curated list of 1,267 prototypic repeat sequences that spans the spectrum of repetitive DNA present in the human genome[39], and we aligned the sequencing reads of all three methods to this collection of repeat sequences. Approximately 20% of all MeDIP-seq, MethylCap-seq and RRBS reads could be aligned with high confidence, enabling us to estimate the global DNA methylation levels for 553 prototypic repeat sequences. The results of the three methods were in excellent agreement with each other (**Supplementary Data 2**) and detected substantial differences in the DNA methylation levels of different repeat classes. Among Alu, SVA (SINE-VNTR-Alus) and satellite repeat sequences we observed consistently high levels of DNA methylation, whereas most LINE (long interspersed nuclear elements), LTR (long terminal repeat) and DNA repeat sequences exhibited low levels of DNA methylation in the four samples that we investigated. However, we found that the repeat sequences with the highest copy-number throughout the genome were highly methylated for all repeat classes.

When we compared the DNA methylation levels in the two sample pairs (**Supplementary Data 3**), we observed widespread but relatively moderate hypomethylation in the colon tumor relative to matched normal colon tissue. The most common targets were Alu, SVA and satellite repeat sequences, consistent with previous reports about cancer-specific hypomethylation[38]. A notable difference was identified between the two ES cell lines on the one hand and the two colon samples on the other hand: the only human-specific LINE repeat sequence in our collection (L1HS_5end) exhibited high levels of DNA methylation in the two colon samples, but was largely unmethylated and even marked by histone H3K4 trimethylation in the two ES cell

Table 2 Validation of method-specific DMRs for MeDIP-seq, MethylCap-seq and RRBS

DMR location	Description	Experimental validation	MeDIP-seq	MethylCap-seq	RRBS
MeDIP-seq–specific DMR chr10:88,149,016-88,149,732	Intergenic CpG island ~30 kb upstream of GRID1, partial overlap with degenerate L1 element	HUES6: 38/56 (68%) methylated CpGs HUES8: 26/44 (59%) methylated CpGs → **insignificant** (P = 0.41)	Hypermethylated (Q = 1.1E-04)	Insignificant (Q = 0.59)	Insignificant (Q = 0.43)
MeDIP-seq–specific DMR chr16:31,142,904-31,143,799	CpG island overlapping with the terminal exon of TRIM72	HUES6: 342/362 (94%) methylated CpGs HUES8: 466/523 (89%) methylated CpGs → marginally **hypermeth.** (P = 0.0051)	**Hypermethylated** (Q = 1.2E-05)	Insignificant (Q = 0.73)	Insufficient coverage
MeDIP-seq–specific DMR chr1:211,290,079-211,290,896	CpG island overlapping with the putative promoter region of *RPS6KC1*	HUES6: 53/60 (88%) methylated CpGs HUES8: 45/50 (90%) methylated CpGs → **insignificant** (P = 1.0)	Hypermethylated (Q = 3.0E-06)	Insignificant (Q = 0.97)	Insignificant (Q = 0.29)
MethylCap-seq–specific DMR chr20:29,526,646-29,527,380	CpG island overlapping with the putative promoter region of *REM1*	HUES6: 5/72 (7%) methylated CpGs HUES8: 78/84 (93%) methylated CpGs → **hypomethylated** (P = 1.4E-30)	Insufficient coverage	**Hypomethylated** (Q = 1.8E-09)	Insufficient coverage
MethylCap-seq–specific DMR chr2:151,825,938-151,826,902	CpG island overlapping with the putative promoter region of *RBM43* and a known copy-number variation	HUES6: 161/208 (77%) methylated CpGs HUES8: 9/104 (9%) methylated CpGs → **hypermethylated** (P = 3.3E-33)	Insignificant (Q = 0.18)	**Hypermethylated** (Q = 7.3E-09)	Insufficient coverage
MethylCap-seq–specific DMR chr13:44,348,934-44,349,700	Intergenic CpG island ~60 kb upstream of *NUFIP1*, partial overlap with degenerate Alu element	HUES6: 80/88 (91%) methylated CpGs HUES8: 41/79 (52%) methylated CpGs → **hypermethylated** (P = 1.2E-08)	Insignificant (Q = 0.40)	**Hypermethylated** (Q = 8.3-07)	Insufficient coverage
RRBS-specific DMR chr3:186,889,821-186,890,200	CpG island overlapping with an internal exon and intron of *IGF2BP2*	HUES6: 5/90 (6%) methylated CpGs HUES8: 88/90 (98%) methylated CpGs → **hypomethylated** (P = 4.3E-42)	Insufficient coverage	Insignificant (Q = 0.18)	**Hypomethylated** (Q = 3.5E-40)
RRBS-specific DMR chr3:32,609,320-32,609,612	Intergenic CpG island ~20kb upstream of *DYNC1LI1*	HUES6: 41/121 (34%) methylated CpGs HUES8: 130/143 (91%) methylated CpGs → **hypomethylated** (P = 3.5E-23)	Insufficient coverage	Insignificant (Q = 0.52)	**Hypomethylated** (Q = 2.9E-26)

Experimental validation of method-specific DMRs between two ES cell lines (HUES6 and HUES8). The table summarizes the results of clonal bisulfite sequencing for eight regions that showed clear-cut DNA methylation differences according to one method but not according to the other two. The P values in column 3 were calculated from the clonal bisulfite sequencing data using Fisher's exact test, based on the DNA methylation levels of individual CpGs. The Q values in columns 4–6 were derived from the DNA methylation maps as described in the Online Methods. One out of three MeDIP-seq–specific DMRs, three out of three MethylCap-seq–specific DMRs and two out of two RRBS-specific DMRs could be confirmed by clonal bisulfite sequencing data (bold print). All genomic coordinates are relative to the NCBI36 (hg18) genome assembly and refer to the amplicon on which the validation was performed. A detailed documentation of the validation experiments is available in **Supplementary Data 1**.

lines (**Supplementary Data 2**). These data suggest that young retrotransposons find ways to evade silencing by DNA methylation in pluripotent cells, which may contribute to their ability to maintain activity in spite of an elaborate epigenetic genome defense[40].

DMR discovery using the Infinium assay

Our study used the Infinium HumanMethylation27 assay as a common reference for evaluating the accuracy of the sequencing-based methods, which was justified by prior studies showing high quantitative accuracy of the Infinium assay[17,24]. However, no prior study investigated the Infinium HumanMethylation27 assay's power to detect DMRs on a genome-wide scale, hence we could not use the Infinium assay as reference when evaluating DMR discovery by the sequencing-based methods. In fact, its low genomic coverage is expected to limit the utility of the Infinium assay for DMR discovery in spite of its well-established accuracy (**Fig. 4**). To empirically address this question, we initially performed statistical testing in much the same way as was done for **Figure 5**. However, most CpG islands were covered by only two Infinium probes, which resulted in low statistical power to detect significant differences. Specifically, paired-samples *t*-tests identified just three significant DMRs among the ES cell lines and two DMRs between the colon tumor and matched normal colon tissue (data not shown).

Thus, we reformulated our question and asked how many true DMRs exhibited suggestive (albeit insignificant) DNA methylation differences in the Infinium data. As an approximation of true DMRs, we focused on those CpG islands that were detected by at least two sequencing-based methods (which are unlikely to contain a high number of technical artifacts according to the comparative validations described above). Between the two ES cell lines a total of 1,000 consensus DMRs were identified (corresponding to the sum of all center fields in **Fig. 5**), of which 251 were covered by at least one Infinium probe. Similarly, we identified 463 consensus DMRs between the colon tumor and matched normal colon tissue, of which 177 were covered by at least one Infinium probe. In most cases, the directionality of the difference was consistent between the consensus DMRs and the Infinium data (**Supplementary Fig. 12**). But when we imposed a minimum threshold of 20 percentage points DNA methylation difference in the same way as for RRBS, the number of Infinium-detected DMRs dropped to 162 (ES-cell comparison) and 95 (colon cancer comparison). In other words, the Infinium assay detected approximately one-fifth of the consensus DMRs that we identified by the sequencing-based methods.

DISCUSSION

Over the last decade, DNA methylation mapping has played an important role in establishing the prevalence of altered DNA methylation in cancer cells[41,42]. More recently, researchers have also started to systematically study the role of DNA methylation in a wide range of non-neoplastic diseases[43]. This is indeed a good time to probe for epigenetic alterations that contribute to human diseases. Genome-wide association studies have been completed for all common diseases and point to a major role of nongenetic factors in the etiology of most diseases[44]. Furthermore, it has been suggested that epigenetic events could provide a tractable link between the genome and the environment, with the epigenome emerging as a biochemical record of relevant life events[45,46]. Systematic investigation of these topics requires powerful, accurate and cost-efficient methods for identifying DNA methylation differences across many samples.

The goal of this study was to evaluate current methods for global DNA methylation mapping and to compare their performance when applied under real-world conditions. To mimic a typical disease-centered case-control study, we worked with primary patient material (colon samples) and used lower amounts of input DNA than in most previous studies (MeDIP-seq: 300 ng; MethylCap-seq: 1 µg; RRBS:

50 ng; Infinium: 1 µg). We focused on cell types that are known to exhibit relatively moderate DNA methylation differences[31,47], in contrast to the massive DNA methylation alterations that are frequently observed in cultured somatic cells[11] and cancer cell lines[48]. Finally, because all four methods included in the current study are widely available and not excessively costly, there are few obstacles to using this technology comparison as a blueprint for individual laboratory efforts as well as large-scale epigenomic case-control studies investigating the epigenetics of human diseases.

Overall, our data confirmed that all four methods provide accurate DNA methylation measurements and can be used to detect DMRs in clinical samples. In terms of accuracy, the bisulfite-based methods (RRBS, Infinium) performed slightly better than the enrichment-based methods and did not require any statistical correction of CpG bias. Furthermore, the genomic coverage was moderately higher for MethylCap-seq than for MeDIP-seq, RRBS coverage was by design focused on CpG-rich regions and the Infinium assay covered a relatively small number of preselected genomic regions.

Despite the striking differences in genomic coverage, a substantial fraction of DMRs detected by MeDIP-seq or MethylCap-seq were also identified by RRBS, and vice versa. This somewhat counterintuitive observation can be explained by the role of region-specific read coverage for the ability to identify statistically significant DMRs. If a genomic region is CpG poor and thus rarely sequenced by MeDIP-seq or MethylCap-seq, both methods have low statistical power to detect differential DNA methylation. In contrast, CpG-rich genomic regions tend to be more amenable to DMR detection by MeDIP-seq and MethylCap-seq and are also frequently covered by RRBS measurements. Finally, we observed that MethylCap-seq was able to detect roughly twice as many DMRs as MeDIP-seq at comparable sequencing depths, RRBS detected more DMRs than MeDIP-seq but fewer DMRs than MethylCap-seq, and the Infinium assay detected only 20% of the consensus DMRs identified by the sequencing-based methods. These differences could be reproduced in two independent pairwise comparisons, providing strong indication that they are robust across biological replicates and cannot be explained by random experimental variation. On the other hand, we used one specific protocol for each method, and it is quite possible that protocol variations (e.g., different antibody for MeDIP-seq, different elution procedure for MethylCap-seq or different size selection for RRBS) would produce different results.

Our study also reinforces the importance of sequencing depth as a key parameter determining the power to detect differential methylation with any of the sequencing-based methods. To allow for a fair and practically relevant comparison, we sequenced ~30–40 million reads for each sample and method. However, it became evident that deeper sequencing would identify further DMRs, especially for MeDIP-seq and MethylCap-seq (**Supplementary Fig. 11**). For disease-centered studies it is therefore necessary to make an informed decision about how to distribute the available resources between sequencing fewer samples more deeply and sequencing more samples less deeply. Such a decision can be guided by statistical power calculations when some prior knowledge exists about the characteristics of expected DMRs (e.g., magnitude of difference, location in CpG-rich versus CpG-poor genomic regions), or they can be dictated by practical considerations such as the number of available samples. In our experience and at current sequencing costs, a range of ~30–60 million reads per sample for MeDIP-seq and MethylCap-seq, and a range of ~10–20 million reads per sample for RRBS constitute a viable compromise between breadth and depth of sequencing. In contrast, whole-genome bisulfite sequencing[49] provides comprehensive genomic coverage at the cost of having to sequence over a billion reads per sample. On the other end of the spectrum, low sequencing depths are often sufficient to detect strong differences such as global loss of DNA methylation but do not provide reliable locus-specific information[50].

Genome-wide studies tend to ignore repetitive regions due to technical difficulties, and the few studies that focused specifically on mapping DNA methylation in repetitive regions did so at relatively low coverage[51–53]. The current data set was well-suited to analyze DNA methylation in repetitive regions because the joint results obtained by three different experimental methods helped us to control for technical artifacts that can burden the analysis of repetitive DNA. We observed that repeat sequences are most highly methylated when they are CpG rich and highly prevalent in the human genome (**Supplementary Data 2**). In contrast, the DNA methylation levels varied widely among repeat sequences that are either CpG poor or infrequent in the genome. These results lend support to the hypothesis that DNA methylation provides a mechanism for keeping active retrotransposons in check[54]. They also argue for a highly specific mechanism of repeat repression, which targets DNA methylation mostly to those repeat sequences that threaten genome integrity, whereas many 'benign' repeat sequences may remain unmethylated.

In summary, we benchmarked four methods for genome-scale DNA methylation mapping in terms of their accuracy and power to detect DNA methylation differences. These results will facilitate the selection of suitable methods for studying the role of DNA methylation in disease and development.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturebiotechnology/.

Note: Supplementary information is available on the Nature Biotechnology website.

ACKNOWLEDGMENTS

We thank A. Crenshaw and M. Parkin (Broad Institute) for assistance with the Infinium assay and K. Halachev (Max Planck Institute for Informatics) for the provision of genome annotation files. C.B. is supported by a Feodor Lynen Fellowship from the Alexander von Humboldt Foundation. A.B.B. is supported by the Dutch Cancer Foundation (KWF, grant KUN 2008-4130). A.M. is supported by the Massachusetts Life Science Center and the Pew Charitable Trusts. The described work was in part funded by the Pew Charitable Trusts, the US National Institutes of Health Roadmap Initiative on Epigenomics (U01ES017155) and the European Union's CANCERDIP project (HEALTH-F2-2007-200620).

AUTHOR CONTRIBUTIONS

C.B., E.M.T. and A.M. conceived and designed the study; E.M.T., A.B.B., F.S. and H.G. performed the experiments; C.B., F.M. and N.J. analyzed the data; C.B., A.G., H.G.S. and A.M. interpreted the results; and C.B. and A.M. wrote the paper.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturebiotechnology/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Bird, A. DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6 (2002).
2. Baylin, S.B. & Ohm, J.E. Epigenetic gene silencing in cancer—a mechanism for early oncogenic pathway addiction? *Nat. Rev. Cancer* **6**, 107–116 (2006).
3. Esteller, M. Epigenetics in cancer. *N. Engl. J. Med.* **358**, 1148–1159 (2008).
4. Feinberg, A.P. & Tycko, B. The history of cancer epigenetics. *Nat. Rev. Cancer* **4**, 143–153 (2004).
5. Issa, J.P. CpG island methylator phenotype in cancer. *Nat. Rev. Cancer* **4**, 988–993 (2004).
6. Jones, P.A. & Laird, P.W. Cancer epigenetics comes of age. *Nat. Genet.* **21**, 163–167 (1999).

7. Richardson, B. Primer: epigenetics of autoimmunity. *Nat. Clin. Pract. Rheumatol.* **3**, 521–527 (2007).
8. Tobi, E.W. *et al.* DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum. Mol. Genet.* **18**, 4046–4053 (2009).
9. Urdinguio, R.G., Sanchez-Mut, J.V. & Esteller, M. Epigenetic mechanisms in neurological diseases: genes, syndromes, and therapies. *Lancet Neurol.* **8**, 1056–1072 (2009).
10. Bock, C. Epigenetic biomarker development. *Epigenomics* **1**, 99–110 (2009).
11. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**, 766–770 (2008).
12. Down, T.A. *et al.* A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. *Nat. Biotechnol.* **26**, 779–785 (2008).
13. Weber, M. *et al.* Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nat. Genet.* **37**, 853–862 (2005).
14. Brinkman, A.B. *et al.* Whole-genome DNA methylation profiling using MethylCap-seq-seq. *Methods* published online, doi:10.1016/j.ymeth.2010.06.012 (11 June 2010).
15. Rauch, T. & Pfeifer, G.P. Methylated-CpG island recovery assay: a new technique for the rapid detection of methylated-CpG islands in cancer. *Lab. Invest.* **85**, 1172–1180 (2005).
16. Serre, D., Lee, B.H. & Ting, A.H. MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res.* **38**, 391–399 (2010).
17. Bibikova, M. *et al.* Genome-wide DNA methylation profiling using Infinium assay. *Epigenomics* **1**, 177–200 (2009).
18. Eckhardt, F. *et al.* DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat. Genet.* **38**, 1378–1385 (2006).
19. Brunner, A.L. *et al.* Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Res.* **19**, 1044–1056 (2009).
20. Irizarry, R.A. *et al.* Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res.* **18**, 780–790 (2008).
21. Oda, M. *et al.* High-resolution genome-wide cytosine methylation profiling with simultaneous copy number analysis and optimization for limited cell numbers. *Nucleic Acids Res.* **37**, 3829–3839 (2009).
22. Gu, H. *et al.* Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat. Methods* **7**, 133–136 (2010).
23. Cowan, C.A. *et al.* Derivation of embryonic stem-cell lines from human blastocysts. *N. Engl. J. Med.* **350**, 1353–1356 (2004).
24. Weisenberger, D.J. *et al. Comprehensive DNA methylation analysis on the Illumina Infinium assay platform* (Illumina, San Diego, California, USA, 2008). ⟨http://www.illumina.com/Documents/products/appnotes/appnote_infinium_methylation.pdf⟩. (2008).
25. Bock, C. *et al.* Inter-individual variation of DNA methylation and its implications for large-scale epigenome mapping. *Nucleic Acids Res.* **36**, e55 (2008).
26. Pelizzola, M. *et al.* MEDME: an experimental and analytical methodology for the estimation of DNA methylation levels based on microarray derived MeDIP-enrichment. *Genome Res.* **18**, 1652–1659 (2008).
27. Robinson, M.D., Statham, A.L., Speed, T.P. & Clark, S.J. Protocol matters: which methylome are you actually studying? *Epigenomics* **2**, 587 (2010).
28. Faul, F. *et al.* G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* **39**, 175–191 (2007).
29. Beroukhim, R. *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature* **463**, 899–905 (2010).
30. Redon, R. *et al.* Global variation in copy number in the human genome. *Nature* **444**, 444–454 (2006).
31. Irizarry, R.A. *et al.* The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* **41**, 178–186 (2009).
32. Hellebrekers, D.M. *et al.* GATA4 and GATA5 are potential tumor suppressors and biomarkers in colorectal cancer. *Clin. Cancer Res.* **15**, 3990–3997 (2009).
33. Zhang, W. *et al.* Epigenetic inactivation of the canonical Wnt antagonist SRY-box containing gene 17 in colorectal cancer. *Cancer Res.* **68**, 2764–2772 (2008).
34. Tenesa, A. *et al.* Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. *Nat. Genet.* **40**, 631–637 (2008).
35. Lofton-Day, C. *et al.* DNA methylation biomarkers for blood-based colorectal cancer screening. *Clin. Chem.* **54**, 414–423 (2008).
36. Caldwell, G.M. *et al.* The Wnt antagonist sFRP1 in colorectal tumorigenesis. *Cancer Res.* **64**, 883–888 (2004).
37. Hirata, H. *et al.* Wnt antagonist gene DKK2 is epigenetically silenced and inhibits renal cancer progression through apoptotic and cell cycle pathways. *Clin. Cancer Res.* **15**, 5678–5687 (2009).
38. Ehrlich, M. DNA hypomethylation in cancer cells. *Epigenomics* **1**, 239–259 (2009).
39. Jurka, J. Repbase update: a database and an electronic journal of repetitive elements. *Trends Genet.* **16**, 418–420 (2000).
40. Bestor, T.H. & Tycko, B. Creation of genomic methylation patterns. *Nat. Genet.* **12**, 363–367 (1996).
41. Esteller, M. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat. Rev. Genet.* **8**, 286–298 (2007).
42. Jones, P.A. & Baylin, S.B. The epigenomics of cancer. *Cell* **128**, 683–692 (2007).
43. Feinberg, A.P. Phenotypic plasticity and the epigenetics of human disease. *Nature* **447**, 433–440 (2007).
44. Manolio, T.A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
45. Foley, D.L. *et al.* Prospects for epigenetic epidemiology. *Am. J. Epidemiol.* **169**, 389–400 (2009).
46. Heijmans, B.T. *et al.* The epigenome: archive of the prenatal environment. *Epigenetics* **4**, 526–531 (2009).
47. Doi, A. *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* **41**, 1350–1353 (2009).
48. Smiraglia, D.J. *et al.* Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. *Hum. Mol. Genet.* **10**, 1413–1419 (2001).
49. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
50. Popp, C. *et al.* Genome-wide erasure of DNA methylation in mouse primordial germ cells is affected by AID deficiency. *Nature* **463**, 1101–1105 (2010).
51. Horard, B. *et al.* Global analysis of DNA methylation and transcription of human repetitive sequences. *Epigenetics* **4**, 339–350 (2009).
52. Rodriguez, J. *et al.* Genome-wide tracking of unmethylated DNA Alu repeats in normal and cancer cells. *Nucleic Acids Res.* **36**, 770–784 (2008).
53. Weisenberger, D.J. *et al.* Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res.* **33**, 6823–6836 (2005).
54. Yoder, J.A., Walsh, C.P. & Bestor, T.H. Cytosine methylation and the ecology of intragenomic parasites. *Trends Genet.* **13**, 335–340 (1997).
55. Bock, C. *et al.* CpG island mapping by epigenome prediction. *PLoS Comput. Biol.* **3**, e110 (2007).
56. Pruitt, K.D., Tatusova, T. & Maglott, D.R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **35** (Database issue), D61–D65 (2007).

ONLINE METHODS

Sample origin and cell culture. Human ES cells were cultured in knockout serum replacement (KOSR) medium according to established protocols[23] and genomic DNA was extracted as described previously[57]. DNA for the colon tumor and matched normal colon tissue was purchased from BioChain (lot number A704198). Both samples originate from the same donor, an 81-year-old male patient diagnosed with moderately differentiated adenocarcinoma.

Methylated DNA immunoprecipitation (MeDIP-seq). MeDIP-seq[12] was performed using the EZ DNA methylation kit (Zymo Research). A total of 300ng DNA per sample was sonicated using Bioruptor (Diagenode) with 8 intervals of 10min (30s on, 30s off). Sonicated DNA was end-repaired and ligated with sequencing adapters as described previously[12]. After gel-based size selection, methylated DNA immunoprecipitation was performed according to the manufacturer's protocol. A total of 1 µg of monoclonal antibody against 5-methylcytosine (included in the EZ DNA methylation kit) was used for immunoprecipitation. The immunoprecipitated DNA was PCR-amplified and the specificity of the enrichment was confirmed by qPCR for selected loci as described previously[58]. Two lanes of 36-bp single-ended sequencing were performed on the Illumina Genome Analyzer II according to the manufacturer's standard protocol. Maq with default parameters was used to align the sequencing reads to the NCBI36 (hg18) assembly of the human genome[59].

Methylated-DNA capture (MethylCap-seq). MethylCap-seq[14] was performed in a robotized procedure using a SX-8G / IP-Star (Diagenode). 2 µg of His6-GST-MBD (Diagenode) was combined with 1µg of sonicated DNA in 200µl of binding buffer (BB, 20mM Tris-HCl pH 8.5, 0.1% Triton X-100) containing 200mM NaCl. This solution was incubated at 4 °C for 2 h. Magnetic GST-beads were prepared by washing 35µl of a well-mixed MagneGST glutathione particle suspension (Promega) with 200 µl of binding buffer plus 200 mM NaCl at 4 °C. Washing was repeated once and the supernatant was removed. The GST-MBD-DNA solution was added to the washed and collected beads, and this suspension was rotated for another hour at 4 °C. After removal of the supernatant (this is the flow-through) the beads-GST-MBD-DNA complexes were eluted by washing. 200 µl of binding buffer with different concentrations of NaCl was added and the suspension was rotated for 10min at 4 °C. Beads were captured using a magnet, and the supernatant was collected. The elution procedure consisted of 1× 300 mM (wash), 2 × 400 mM (wash), 1 × 500 mM ("low" eluate), 1 × 600 mM ("medium" eluate), 1× 800 mM NaCl ("high" eluate). The collected eluates were purified using QIAquick PCR purification spin columns (Qiagen), eluted with 100 µl elution buffer and prepared for sequencing as described previously[14]. A single lane of 36-bp single-ended sequencing on the Illumina Genome Analyzer II was performed for the low, medium and high eluates, respectively. The sequencing reads were aligned to the NCBI36 (hg18) assembly of the human genome using Illumina's analysis pipeline (ELAND) with default parameters. The lanes for each of the three eluates are shown separately in **Figure 2**, and we tested whether the accuracy relative to the Infinium assay could be improved by taking this additional information into account. However, a linear model that was based on the separate read counts of the three lanes did not outperform a model that was based on the sum of the three lanes, which is why we combined the reads from all three libraries per sample for the analyses described in this paper.

Reduced representation bisulfite sequencing (RRBS). RRBS[22] was performed according to a previously published protocol[57] with some optimizations for clinical samples and low amounts of input DNA[22]. The main steps were: (i) A total of 50ng (ES cells) or 1 µg (colon samples) genomic DNA was digested by 5U to 20 U of MspI (New England Biolabs, NEB) for up to 16 h. (ii) End-repair and adenylation of digested DNA were performed in a 20 µl reaction consisting of 10U of Klenow fragments (3′→ 5′ exo-, NEB), 2 µl premixed nucleotide triphosphates (1 mM dGTP, 10 mM dATP, 1 mM 5′ methylated dCTP). The reaction was incubated at 30 °C for 30 min followed by 37 °C for additional 30 min. (iii) Preannealed 5-methylcytosine-containing Illumina adapters were ligated with adenylated DNA fragments in a 20 µl reaction containing of 1 µl concentrated T4 ligase (NEB), 1–2 µl of 15 µM adapters at 16 °C for 16 to 20 h. (iv) Gel-based selection for fragments with insertion sizes of 40 to 120 bp and 120 to 220 base pairs was performed as described previously[22]. (v) Bisulfite treatment with the EpiTect Bisulfite Kit (Qiagen) was conducted following the protocol designated for DNA isolated from formalin-fixed and paraffin-embedded tissues. Two rounds of conversion were performed in order to maximize bisulfite conversion rates. The final bisulfite-converted DNA was eluted with 2× 20µl pre-heated (65 °C) EB buffer. (vi) To determine the minimum number of PCR cycles for final library enrichment, analytical (10 µl) PCR reactions containing 0.5 µl of bisulfite-treated DNA, 0.2 µM each of Illumina PCR primers LPX1.1 and 2.1 and 0.5 U PfuTurbo Cx Hotstart DNA polymerase (Stratagene) were set up. The thermocycler conditions were: 5 min at 95 °C, varied cycle numbers (10–20) of 20 s at 95 °C, 30 s at 65 °C, 30 s at 72 °C, followed by 7 min at 72 °C. PCR products were visualized by running on a 4–20% polyacrylamide Criterion TBE Gel (Bio-Rad) and stained by SYBR Green. The final libraries were generated by 8 of 25 µl PCR reaction with each one containing 2–3µl of bisulfite-converted template, 1.25 U PfuTurbo Cx Hotstart polymerase and 0.2 µM each of Illumina LPX1.1 as well as 2.1 PCR primers. The libraries were PCR amplified and sequenced on the Illumina Genome Analyzer II as described previously[22]. The sequencing reads were aligned to the NCBI36 (hg18) assembly of the human genome using a custom alignment software that was developed for RRBS data[11].

Microarray-based epigenotyping (Infinium). Infinium[17] analysis was performed by the Genetic Analysis Platform at the Broad Institute. A total of 1 µg of genomic DNA per sample was bisulfite-treated according to the manufacturer's protocol and hybridized onto Infinium HumanMethylation[27] bead arrays (Illumina). We previously observed almost perfect agreement between technical replicates (Pearson's $r > 0.98$), which is why only a single hybridization was performed for each sample.

Data preparation and quality control. For MeDIP-seq and MethylCap-seq, the aligned reads were extended to the mean fragment length obtained during sonication, and from each group of duplicate reads (that is, reads aligned to the exact same start position on the same chromosome) all but one read were discarded, in order to minimize the impact of PCR bias on downstream analysis. For RRBS, the aligned reads were compared to the reference genome, and the DNA methylation status was determined using custom software as described previously[22]. Infinium HumanMethylation27 data were processed with Illumina's BeadStudio 3.2 software, using the default background subtraction method for normalization. UCSC Genome Browser tracks were constructed by custom scripts implemented in the Python programming language (http://www.python.org/).

Quantification of absolute DNA methylation levels. We used linear regression models to estimate the absolute DNA methylation levels from the MeDIP-seq and MethylCap-seq read counts. Based on a number of different feature selection experiments, we found that the following combination of variables was robustly predictive of DNA methylation levels: (i) the square root of the total number of MeDIP-seq or MethylCap-seq reads within the given region, (ii) the square root of the total number of whole-cell extract (WCE) reads within the region (based on a cross-tissue WCE track that we routinely use for ChIP-seq data normalization), (iii) the logit of the CpG frequency within the region, (iv) the relative GC content of the region, (v) the ratio of Cs relative to CpGs, and (vi) the relative repeat content of the region as determined by RepeatMasker (http://www.repeatmasker.org/). For both MeDIP-seq and MethylCap-seq, we observed that the read frequencies were strongly positively associated with the absolute methylation level obtained using the Infinium assay, whereas the repeat content was moderately positively associated. In contrast, the logit of the CpG frequency was highly negatively associated with DNA methylation, and all other variables as well as the model's intercept exhibited a moderately negative association. For model fitting and performance evaluation, the current data set was split into equally sized training and test sets. All model fitting was performed using the R statistics package (http://www.r-project.org/).

Identification of differentially methylated regions. In our experience, classical peak detection[60,61] is not well-suited for DMR identification because of the high number of spurious hits encountered when borderline peaks are detected in one sample but not in the other (C.B., unpublished observation).

Instead, we used a statistical test to compare two samples directly with each other. For a given region with RRBS data, we count the number of methylated vs. unmethylated CpGs in both samples and perform Fisher's exact test to obtain a *p*-value that is indicative of the likelihood of the region being a DMR. Similarly, for MeDIP-seq and MethylCap-seq we count the numbers of reads that align inside the region for both samples and use Fisher's exact test to contrast these values with the total numbers of reads that align elsewhere in the genome. And for the Infinium assay we use a paired-samples *t*-test to compare the two samples' β-values of all Infinium probes inside the region. These tests are performed on a large number of genomic regions in parallel (e.g., on all CpG islands), and the p-values are corrected for multiple testing using the q-value method[62]. Genomic regions with a q-value of less than 0.1 are flagged as hypermethylated or hypomethylated (depending on the directionality of the difference), but only if the absolute DNA methylation difference exceeds 20 percentage points (for RRBS and Infinium) or if there is at least a twofold difference in the read number (for MeDIP-seq and MethylCap-seq). These thresholds were chosen by their practical utility in a number of comparisons between different cell types and have no further justification. We also mark genomic regions with insufficient sequencing coverage, but do not exclude them from DMR analysis. For MeDIP-seq and MethylCap-seq we require at least ten reads per 10 million total reads for the sample with higher read coverage, and for RRBS we require a minimum of five CpGs with at least five reads each in both samples.

This statistical approach to DMR identification requires us to define sets of genomic regions on which the analysis is being performed. We pursued a two-way strategy to maximize the chances of finding interesting DMRs. One the one hand, we focused specifically on CpG islands and gene promoters, which are prime candidates for epigenetic regulation. This approach provides increased statistical power for regions with well-known functional roles because the relatively low number of CpG islands and gene promoters reduces the burden of multiple-testing correction compared to the genome-wide case. On the other hand, we used a 1-kilobase tiling of the genome to detect DMRs that are located outside of any candidate regions. And to cast an even wider net, we collected a comprehensive set of 13 types of genomic regions, which includes not only CpG islands and gene promoters, but also CpG island shores[31], enhancers[63], evolutionary conserved regions and other types of genomic regions. DMR data for all of these region sets were calculated using a set of Python and R scripts and are available online (http://meth-benchmark.computational-epigenetics.org/).

Experimental validation. Based on the CpG islands that were detected as differentially methylated between the two ES cell lines (**Fig. 5**), we manually selected eight method-specific DMRs for experimental validation. To that end, those CpG islands that were identified as statistically significant DMRs by one method (but not by the other two methods) were visually inspected in the UCSC Genome Browser, and regions were selected for validation only if the data fully supported their classification as method-specific DMRs. In particular, regions were not selected if a second method already picked up a suggestive but insignificant trend in the same direction as the first method, or when the data of the first method already suggested that the DMR was a false-positive hit (e.g., because of contradictory trends in the vicinity of the DMR). Experimental validation was performed by clonal bisulfite sequencing following established protocols[64]. Primers were designed using MethPrimer[65] such that the amplicon overlapped with those CpGs that exhibited the highest levels of differential methylation according to our original data. To prepare for bisulfite sequencing, 1 μg of DNA was bisulfite-converted using the EpiTect kit (Qiagen); 50 ng of bisulfite-converted DNA was PCR-amplified (**Supplementary Data 1** for primer sequences); and purified amplicons were cloned using the TOPO TA cloning kit (Invitrogen). For each region an average of 11 clones were randomly chosen for sequencing. All sequencing data were processed using the BiQ Analyzer software[66], and the results are summarized in **Supplementary Data 1**.

Analysis of repetitive DNA. Repeat sequences were obtained from database version 14.07 of RepBase Update[39], which is publicly available online (http://www.girinst.org/server/RepBase/index.php). From a total of 11,670 prototypic repeat sequences we selected those 1,267 that were annotated either to human or to its ancestors in the taxonomic tree, and we combined these prototypic repeat sequences into a pseudo-genome file. Maq with default parameters was used to align MeDIP-seq, MethylCap-seq, RRBS, ChIP-seq (H3K4me3) and whole-cell extract (WCE) sequencing reads against this pseudo-genome[59]. For RRBS, both the reads and the reference genome were bisulfite-converted in silico before the alignment. The epigenetic status of each prototypic repeat sequence was quantified as follows: (i) For MeDIP-seq, MethylCap-seq and ChIP-seq we calculated the odds ratios relative to the WCE data. (ii) For RRBS we computed the number of methylated CpGs, total number of CpG measurements and percentage of DNA methylation based on the comparison of the aligned reads with the prototypic repeat sequence.

We discarded rare repeats with WCE coverage below 100 aligned reads or RRBS coverage below 25 CpG measurements, resulting in 553 prototypic repeat sequences that were used for further analysis. Among these were 97 LINE class sequences (92 of them from the L1 family), 51 SINEs (48 of them from the Alu family), 6 SVAs, 62 DNA repeats, 15 satellite repeats, 315 LTRs, 1 low-complexity repeat and 6 RNA repeats (**Supplementary Data 2**). To quantify differential methylation between a pair of MeDIP-seq and MethylCap-seq samples, we calculated the pairwise odds ratio of the read coverage for each prototypic repeat sequence. The absolute DNA methylation difference was used in the case of RRBS (**Supplementary Data 3**). The significance of the difference was assessed using Fisher's exact test in the same way as for the nonrepetitive genome (described above).

57. Smith, Z.D. *et al.* High-throughput bisulfite sequencing in mammalian genomes. *Methods* **48**, 226–232 (2009).
58. Rakyan, V.K. *et al.* An integrated resource for genome-wide identification and analysis of human tissue-specific differentially methylated regions (tDMRs). *Genome Res.* **18**, 1518–1529 (2008).
59. Li, H., Ruan, J. & Durbin, R. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res.* **18**, 1851–1858 (2008).
60. Bock, C. & Lengauer, T. Computational epigenetics. *Bioinformatics* **24**, 1–10 (2008).
61. Park, P.J. ChIP-seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* **10**, 669–680 (2009).
62. Storey, J.D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **100**, 9440–9445 (2003).
63. Heintzman, N.D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459**, 108–112 (2009).
64. Hajkova, P. *et al.* DNA-methylation analysis by the bisulfite-assisted genomic sequencing method. *Methods Mol. Biol.* **200**, 143–154 (2002).
65. Li, L.C. & Dahiya, R. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* **18**, 1427–1431 (2002).
66. Bock, C. *et al.* BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* **21**, 4067–4068 (2005).

doi:10.1038/nbt.1681