A novel methyl-binding domain protein enrichment method for identifying genome-wide tissue-specific DNA methylation from nanogram DNA samples
Verity F Oliver, Jun Wan, Saurabh Agarwal, Donald J Zack, Jiang Qian, Shannath L Merbs

To cite this version:
Verity F Oliver, Jun Wan, Saurabh Agarwal, Donald J Zack, Jiang Qian, et al.. A novel methyl-binding domain protein enrichment method for identifying genome-wide tissue-specific DNA methylation from nanogram DNA samples. Epigenetics and Chromatin, BioMed Central, 2013, 6, pp.17. <10.1186/1756-8935-6-17>. <hal-01549207>

HAL Id: hal-01549207
https://hal.sorbonne-universite.fr/hal-01549207
Submitted on 28 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
A novel methyl-binding domain protein enrichment method for identifying genome-wide tissue-specific DNA methylation from nanogram DNA samples

Verity F Oliver¹, Jun Wan¹, Saurabh Agarwal²³, Donald J Zack¹⁴⁵⁶⁷, Jiang Qian¹ and Shannath L Merbs*¹

Abstract

Background: Growing evidence suggests that DNA methylation plays a role in tissue-specific differentiation. Current approaches to methylome analysis using enrichment with the methyl-binding domain protein (MBD) are restricted to large (≥1 μg) DNA samples, limiting the analysis of small tissue samples. Here we present a technique that enables characterization of genome-wide tissue-specific methylation patterns from nanogram quantities of DNA.

Results: We have developed a methodology utilizing MBD2b/MBD3L1 enrichment for methylated DNA, kinase pre-treated ligation-mediated PCR amplification (MeKL) and hybridization to the comprehensive high-throughput array for relative methylation (CHARM) customized tiling arrays, which we termed MeKL-chip. Kinase modification in combination with the addition of PEG has increased ligation-mediated PCR amplification over 20-fold, enabling >400-fold amplification of starting DNA. We have shown that MeKL-chip can be applied to as little as 20 ng of DNA, enabling comprehensive analysis of small DNA samples. Applying MeKL-chip to the mouse retina (a limited tissue source) and brain, 2,498 tissue-specific differentially methylated regions (T-DMRs) were characterized. The top five T-DMRs (Rgs20, Hes2, Nfic, Cckbr and Six3os1) were validated by pyrosequencing.

Conclusions: MeKL-chip enables genome-wide methylation analysis of nanogram quantities of DNA with a wide range of observed-to-expected CpG ratios due to the binding properties of the MBD2b/MBD3L1 protein complex. This methodology enabled the first analysis of genome-wide methylation in the mouse retina, characterizing novel T-DMRs.

Keywords: DNA methylation, Tissue-specific differentially methylated regions, Retina, MBD, Nanogram

Background

DNA methylation is an epigenetic modification known to be important in many cellular processes, including tissue-specific gene expression. There is no standardized method for DNA methylation analysis and most array-based methods for genome-wide profiling require input DNA in microgram quantities, making them impractical for small tissue samples (Table 1). Sodium bisulfite modification is traditionally considered the gold-standard technique for assessing methylation [1]. Alternatives to bisulfite conversion include enrichment of methylated/ unmethylated DNA by the use of methylation-sensitive restriction enzymes (MREs, for example, HpaII), immunoprecipitation with an anti-5-methylcytosine antibody (MeDIP) or affinity enrichment with methyl-binding domain proteins (MBDs) (Table 1). After bisulfite conversion or enrichment, DNA methylation can be quantified via next-generation sequencing or hybridization to a microarray. While whole genome bisulfite sequencing (WGBS) can quantify DNA methylation at single-nucleotide resolution using nanogram amounts of DNA, it requires robust bioinformatics resources, and the current cost-per-sample is still relatively high; thus, alternative methods to perform targeted bisulfite sequencing, such as reduced-representation bisulfite sequencing (RRBS) have been developed [2].
RRBS comes at the cost of reduced genomic coverage compared to MeDIP and MBD enrichment-based methods [3,4]. The most common MBD enrichment uses MBD2b, which shows high affinity for double-stranded, methylated DNA, in contrast to the MeDIP antibody, which binds methylated single-stranded DNA. The binding affinity of MBD2b is sequence independent, and MBD enrichment can be further enhanced by the addition of MBD3L1. Methylated-CpG island recovery assay (MIRA) [5] uses this protein complex and is now commercially available as a kit. Direct next-generation sequencing comparison has shown that MBD enrichment can identify more differentially methylated regions (DMRs) than either RRBS or MeDIP [3], but it is currently an underutilized methodology, particularly in combination with the more economical array-based analyses.

Methylation data obtained by hybridization to promoter, CpG island or CpG site-specific microarrays is biased and restricted by the design of the array platform. To overcome this bias, the comprehensive high-throughput array-based analyses.

Table 1 Comparative DNA requirements for genome-wide methylation analysis platforms

| Methodology          | Starting DNA required (μg) | Platform          |
|----------------------|-----------------------------|-------------------|
| WGBS [6-8]           | 0.01–0.1                    | Illumina          |
| MeDIP-Seq [3,4]      | 0.05–0.3                    | Illumina          |
| MED-Seq [3]          | 1*                          | Illumina          |
| RRBS-Seq [3,9]       | 0.01–0.05                   | Illumina          |
| MRE-Seq [10]         | 3                           | Illumina          |
| MeDIP-Chip [11]      | 4                           | Affymetrix promoter array |
| MBD-Chip [11]        | 1                           | Affymetrix promoter array |
| MRE-Chip [12]        | 2                           | User defined      |
| Infinium Methylation | 0.8                         | Illumina\[^{a}\]  |
| BeadChip [13]        |                             |                   |
| CHARM [14]           | 10                          | NimbleGen         |
| MeKL-Chip            | 0.02                        | NimbleGen/any     |

\[^a^]{Pooled four to eight 1-μg fractions.}
\[^b^]{Not available for the mouse genome.}

be excluded from promoter and CpG island arrays. However, applicability of the McrBC-based CHARM protocol is limited because the enrichment protocol requires large amounts (10 μg) of starting genomic DNA (Table 1).

While MIRA can enrich DNA samples in the nanogram range [5], hybridization to the 2.1M NimbleGen microarray requires 1 μg of each DNA sample post-enrichment, necessitating high-magnitude whole genome amplification (WGA). There are two categories of commonly used WGA methodology: PCR amplification and isothermal DNA amplification. Ligation-mediated PCR (LM-PCR) [17] falls into the former category and involves the ligation of a unidirectional, double-stranded oligonucleotide universal adapter to blunted DNA fragments. The universal primer sequence can then be used to amplify all ligated fragments via PCR. LM-PCR enables amplification of a range of small PCR fragments (up to 2,000 bp in length), irrespective of the genomic sequence.

Currently, no protocols are available for MBD-chip with low amounts of starting DNA. With the goal of establishing a user-friendly method for assessing genome-wide DNA methylation in small (nanogram) DNA samples, we developed a new protocol based upon MBD2b/MBD3L1 enrichment [5] followed by amplification using modified LM-PCR, which we call MeKL-chip. The development of an array-based method for small tissue samples is advantageous for global methylation analysis because it does not have the computational burden of sequencing-based methods. Here we report on our MeKL-chip assay, its application to the mouse retina (a limited tissue source), and highlight its ability to detect novel T-DMRs with a wide range of CpG densities.

Results

Improved amplification efficiency by modified ligation-mediated PCR

Using the published LM-PCR protocol as a starting point [17], we explicitly tailored our amplification for nanogram levels of input DNA. As LM-PCR can be limited by the efficiency of the ligation reaction, we made two modifications to improve the ligation reaction. Firstly, we increased the efficiency by adding a pre-ligation kinase treatment (kinase-modified ligation-mediated PCR, KLM-PCR) to repair DNA damage induced during fragmentation. Sonication can result in DNA breaks with 5′ hydroxyl and 3′ phosphate; treatment with T4 polynucleotide kinase reinstates the 5′ phosphate and 3′ hydroxyl moieties. This increased the amount of the template for ligation, improving the amplification 14-fold over the published method (Figure 1A). Additionally, we incorporated higher polyethylene glycol (PEG-8000) concentrations (12.5% or 15% vs 5%) in the ligation reaction buffer, as a previous report showed improved ligation efficiency under these conditions [18]. PEG-8000 concentrations of 12.5% and 15% increased
the amplification two- to threefold, respectively (Figure 1A). Combined, these modifications increased the amplification 24- to 38-fold. Due to the difficulty of working with the viscous 15% PEG solution, we elected to use 13% PEG for subsequent experiments. To maximize PCR efficiency, a palindrome within the universal adapter sequence was disrupted to prevent the formation of an adapter-only template (Additional file 1: Figure S1A). We termed this combination of MBD2b/MBD3L1 enrichment [5] of DNA followed by KLM-PCR and subsequent hybridization to a microarray platform MeKL-chip (Figure 1B). MeKL can be used successfully to amplify sufficient DNA for microarray labeling and hybridization from as little as 10 ng of pre-enrichment, fragmented DNA (Additional file 1: Figure S1B).

Identification of T-DMRs using MeKL-chip

MeKL-chip was used to perform genome-wide profiling of adult mouse retinas to identify retinal T-DMRs, using the brain as a comparison neuronal tissue. As a quality control to demonstrate successful MBD2b/MBD3L1 enrichment of methylated DNA, quantitative PCR (QPCR) was performed at regions known to be differentially methylated between the retina and brain. The retinol-binding protein 3 (Rbp3) and rhodopsin (Rho) genes are specifically expressed in retinal photoreceptors and have been shown to be hypomethylated in a cell-specific manner in the retina compared to non-retinal tissues [19]. In addition to QPCR validation of Rho and Rbp3, the imprinted gene H19 (which is fully methylated on the paternal allele [20]) was also examined to ensure equal enrichment of methylated DNA in the retina and brain samples. After MBD enrichment, samples showed differential amplification of Rho and Rbp3 in the brain compared to the retina, and a lack of differential enrichment of H19 between the retina and brain (Figure 2A). The differential enrichment pattern between the retina and brain was maintained after KLM-PCR (Figure 2B). The amplified methylation-enriched and unenriched DNA from each sample were labeled and co-hybridized to the mouse CHARM array. In total, 2,498 novel T-DMRs were identified of which 1,449 were hypermethylated in the brain (Additional file 2: Table S1).

Validation of potential T-DMRs by pyrosequencing

The top five T-DMRs were validated by pyrosequencing (Figure 2C and Additional file 3: Figure S2). The highest ranking T-DMR, which was hypermethylated in the retina, covered Exon 3 of Rgs20 and its flanking introns (Figure 2C). An alternative transcription start site is located at Exon 3, an exon which is included in the brain-specific isoform of Rgs20 and excluded from the retina-specific isoform [21]. The remaining top T-DMRs (associated with genes Hes2, Nfic, Cckbr and Six3os1) overlapped transcription start sites or were located intragenically (Additional file 3: Figure S2). The directionality of the differential methylation at all T-DMRs identified by MeKL-chip was completely consistent with the methylation levels observed in an independent mouse cohort by bisulfite pyrosequencing.

Local CpG density and observed-to-expected CpG ratio analysis at T-DMRs

To assess the properties of the MBD2b/MBD3L1 enrichment within the MeKL-chip assay, the CpG density and the observed-to-expected CpG ratios (CpGO/E) were determined for all T-DMRs identified. T-DMRs with a wide range of local CpG densities (1.7 to 46.2 in a 300
bp window) and CpGO/E (0.10 to 1.05) were isolated (Figure 3A,B). The majority of the T-DMRs were located in regions with relatively low CpG content (mean local CpG density of 9.2, mean CpGO/E of 0.39). An example of the diversity of CpG densities detectable within a genomic locus was found at the Rbp3 gene, which is hypermethylated in brain tissue [19]. Three T-DMRs were identified within a 3,000 bp region of Rbp3 with local CpG densities of 4.6, 9.4 and 13.2 and CpGO/E of 0.31, 0.38 and 0.47, respectively (Figure 3C).

**MeKL-chip for nanogram DNA samples**

To determine the effectiveness of MeKL-chip on low-input DNA samples, differing amounts of total starting DNA were fragmented, enriched and amplified from mouse retina and brain. Successful fragmentation was achieved with 50, 125 and 250 ng of sample DNA (data not shown). MBD2b/MBD3L1 enrichment was achieved using as little as 10 ng of fragmented sample DNA (Additional file 4: Figure S3A). Enrichment was maintained after KLM-PCR amplification in all low-input
amounts examined (Additional file 4: Figure S3B). MeKL-chip data of the Rgs20 region from the 10-ng input samples from the retina and brain revealed tissue-specific methylation differences comparable to those detected in the 250-ng arrays (Additional file 5: Figure S4). Array hybridization of the 10, 25 and 50 ng fragmented and MBD2b/MBD3L1-enriched DNA from retina and brain samples yielded analyzable results that were combined to provide a low-input MeKL-chip group. The difference in methylation between retina and brain ($\Delta M$) at T-DMRs identified within the high-input samples (250 ng) was compared to the corresponding $\Delta M$ of these same T-DMRs from the low-input samples (10, 25 and 50 ng; Figure 4A). A strong correlation was observed between the high- and low-input results (correlation coefficient [c.c.] = 0.63, 0.59 and 0.75, respectively). In comparison, there was no correlation between high- and low-input experiments when the probe positions were randomly shuffled (Figure 4B). Over 61% of the 2,498 T-DMRs identified in the high-input samples had the same directionality of $\Delta M$ for all three low-input samples, which were termed ‘verified’ T-DMRs. The remaining T-DMRs showed inconsistent $\Delta M$ directions between the high- and low-input groups (visible in the upper left and lower right quadrants of Figure 4A), and were referred to as ‘dropouts.’ Notably, the majority (72%) of these dropout T-DMRs had an opposite $\Delta M$ direction in only one of the three low-input samples and less than 11% of all 2,498 T-DMRs had a conflicting $\Delta M$ direction with two or more low-input samples. The dropouts tended to have lower CpGO/E and lower CpG density (Figure 4C). Integrative analysis of the T-DMR data with exon array data in the same mouse tissues [22] revealed that only 6% of these dropouts were associated with retina-/brain-enriched genes when filtering for those dropouts absent from at least two low-input arrays (Figure 4D). These results indicated that the low-input arrays were able to detect differential methylation within biologically relevant regions and validated the use of the MeKL-chip method on low-input ($\leq$50 ng) DNA samples.

**Discussion**

We have demonstrated a novel combination of MBD2b/MBD3L1 enrichment, modified ligation-mediated PCR amplification and microarray hybridization, in a method termed MeKL-chip. Using our unique protocol, we have shown that MeKL-chip permits genome-wide DNA
methylation analysis with as little as 20 ng of DNA (comprising 10 ng for enrichment and 10 ng input for comparison). The ability to study genome-wide methylation in small samples enabled us to identify novel T-DMRs in a limited tissue sample: the mouse retina. The MeKL-chip protocol fills a void in the currently available methylation profiling technologies, enabling regional examination of methylation patterns with a low computational requirement.

Ligation-mediated PCR (LM-PCR) is an approach well suited to the amplification of small DNA fragments [17]. In LM-PCR, universal linkers are ligated to blunt-ended DNA fragments, which are then used to amplify the ligated DNA via PCR [17]. We modified the LM-PCR protocol to include a pre-ligation kinase step to repair the fragmented DNA, included PEG during the ligation to increase ligation efficiency and modified the linker sequence to remove a palindrome. After these modifications, KLM-PCR was able to produce sufficient DNA for microarray hybridization (a minimum of 70-fold amplification) from post-enrichment samples derived from as little as 10 ng.

The top five T-DMRs identified by MeKL-chip and validated by pyrosequencing were all genes with known biological relevance in the retina. The top T-DMR at Rgs20 (regulator of G protein signaling 20, Gene ID: 58175) overlapped an alternative transcription start site [21], which suggests that DNA methylation may play an essential role in the tissue-specific expression of Rgs20 transcripts in the retina and brain. The location of the T-DMR within Rgs20 is consistent with existing results that have shown a role for intragenic DNA methylation in tissue-specific expression mediated by alternative transcriptional start sites [23]. The transcription factor Hes2 (hairy and enhancer of split 2, 15206) has been previously implicated in cell fate determination of the Xenopus retina [24], and Six3os1 (100043902), a long non-coding RNA on the opposite strand to the sine oculis-related homeobox 3 transcription factor, has also been described as important during retinal development [25]. Cckbr (cholecystokinin B receptor, 12426) mRNA has been reported in the rat retina [26]. Nfic (nuclear factor I/C, 18029) has been described as overexpressed in the retina of patients with proliferative vitreoretinopathy, a disease where aberrant wound healing occurs in the retina [27]. The ability to analyze global DNA methylation in the mouse retina has identified T-DMRs at biologically relevant genes in addition to novel
targets that warrant further study. The vast majority of these top T-DMRs identified in the high-input (250 ng) arrays were also identified in the low-input (≤50 ng) arrays.

We observed that the overall difference in methylation at the T-DMRs was greater by pyrosequencing than by MeKL-chip. This difference in percentage methylation between array hybridization and pyrosequencing has been noted previously [15], suggesting that hybridization to CHARM arrays using either unmethylated-enriched or methylated-enriched DNA may underestimate the relative difference in methylation between samples as a consequence of the fixed dynamic range of microarrays.

One of the main advantages of using MBD2b/MBD3L1 enrichment, due to the unique binding properties of the MBD2b/MBD3L1 complex, is the ability to detect methylation within a wide range of CpG densities [11]. Although MeKL-chip is unable to measure site-specific CpG methylation compared to bisulfite-based sequencing methods, we successfully identified robust T-DMRs throughout the mouse genome. We were able to isolate T-DMRs with a broad range of observed-to-expected CpG ratios and CpG densities unlike the more frequently used MeDIP technique, which is known to be most efficient at low CpG density ranges [11]. Although the CpGO/E of our T-DMRs ranged from 0.10 to 1.05, the majority of our T-DMRs had CpGO/E less than that of a CpG island (CpGO/E > 0.6). This observation supports previously published data indicating that the majority of T-DMRs are identified in regions outside of CpG islands with relatively low CpG density, for example, CpG shores and CpG shelves [15,16,28,29]. While we selected the NimbleGen CHARM 2.1M platform because regions of low CpG density were incorporated into the array design, the MeKL-chip methodology is amendable to any array platform. The incorporation of PEG during ligation should facilitate next-generation sequencing library preparation from nanogram amounts of methylation-enriched DNA.

Conclusions
We have demonstrated that the KLM-PCR method of WGA results in greater than 400-fold amplification of low-input DNA samples. This protocol has many potential downstream applications, including hybridization to custom arrays as well as next-generation sequencing-based platforms. By combining KLM-PCR WGA and MBD-affinity methylation enrichment with hybridization to a custom CHARM microarray (MeKL-chip), we were able to achieve robust identification of T-DMRs between the retina and brain within biologically relevant genes using nanogram quantities of input DNA. The MeKL-chip method enables genome-wide assessment of methylation in samples previously considered below the threshold for array-based, global methylation analyses. This methodology will be particularly useful for the identification of regional methylation differences between small tissue samples, for example, laser-capture microdissection collected DNA, and in detecting disease-associated methylation differences within affected cell layers.

Methods
Animals
Eight-week-old C57BL/6J male mice (n = 5) (Jackson Laboratories) were euthanized using IsoSol™ (VEDCO) exposure followed by neck dislocation. A second cohort of eight-week-old C57/B6J male mice (n = 5) was processed in an identical manner for validation using bisulfite pyrosequencing. All procedures were approved by the Johns Hopkins University Institutional Animal Care and Use Committee (IACUC) and were performed in accordance with guidelines in the National Research Council’s Guide for the Care and Use of Laboratory Animals.

Samples
After euthanasia, the mouse eyes were immediately enucleated and placed into 1X PBS buffer. The cornea and lens were discarded and the eyecup placed into 500 µL of 1X Hanks’ balanced salt solution (HBSS, Invitrogen). The eyecup was incubated for 15 min at 37°C and then microdissected in 1X HBSS. Surrounding sclera was removed from the retina, and any remnants of retinal pigment epithelium were removed by gentle scraping. Three samples of 25 mg of brain cortex were removed from each mouse. DNA extraction was performed on each sample using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer’s protocols with addition of RNase A. DNA was eluted in 400 µL Buffer AE. Ethanol precipitation of the DNA samples was performed [30] and the DNA resuspended in 40 µL 1X TE pH 8.0. Retina/brain DNA samples originating from the same mouse were pooled. The amount of DNA in each sample was quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Invitrogen).

MBD enrichment for methylated DNA
Prior to enrichment, 2.5 µg of DNA in 100 µL 1X TE pH 8.0 was fragmented to an average target size of 300 bp (duty cycle 10%, intensity 4, cycles per burst 200, time 60 sec) using a Covaris™ S220 Ultrasonicator. For smaller amounts of DNA, 50, 125 or 250 ng of DNA was fragmented in 50 µL (duty cycle 10%, intensity 5, cycles per burst 200, time 50 sec). The accuracy of the fragmentation was checked using an Agilent 2100 Bioanalyzer with the DNA 1000 kit. Enrichment was performed using the MethylCollector Ultra Kit (Active Motif) according to the manufacturer’s protocols (version C1) using Low-Salt Binding Buffer AM12. DNA cleanup after enrichment was performed using the MinElute CR
Purification kit (Qiagen). The enriched DNA was eluted in 10 μL Buffer EB. To enable quality control experiments, fragmented DNA was enriched in duplicate and the duplicate samples combined.

**QPCR validation of enrichment**

QPCR was performed for the known brain/retina differentially methylated genes *Rho* (Gene ID 212541) and *Rbp3* (19661) [19] and a known equally methylated gene, *H19* (14955) [20]. QPCR was performed on an iQ™ instrument (Bio-Rad) using 1X EvaGreen® dye (Biotium), Fermentas Maxima™ Hot Start Taq DNA Polymerase (Thermo Fisher Scientific) and 1.2 μL of 5 μM forward and reverse primer mix in a final 20 μL reaction volume. Next, 10 ng of fragmented, unenriched DNA, 1 μL of post-enrichment DNA or 10 ng post-amplification DNA were amplified in triplicate for the same sample at all three genes. The primer sequences were (5’ to 3’): *Rho* forward: AA GCACGCTTGGTCTCTGTC, *Rho* reverse: CCCTCTTGTC CCGTTTCATGG, *Rbp3* forward: GCTCGCTCACCTCTCTGG6, *H19* forward: TGTGTAAAGACCAGGGTG and *H19* reverse: GGAGAAAACCTCAATCTGTGC. The QPCR cycling conditions were: 95°C for 3 min, 50 × (95°C for 10 sec, 66.4°C for 30 sec, 72°C for 30 sec) followed by the standard dissociation steps. The mean threshold cycle (Ct) for each sample was used to calculate ΔCt (retina-brain) and ΔCt was then used to calculate the fold enrichment for methylation (2ΔCt). A lower Ct for the brain sample and a fold enrichment greater than 2 should be observed post-enrichment as the brain is hypermethylated at *Rho* and *Rbp3* [19].

**Whole genome amplification by kinase ligation-mediated PCR**

All buffers and enzymes used for amplification were from New England Biolabs (NEB) unless otherwise stated. Based on the LM-PCR protocol developed by the Ren laboratory [17], 5 μg of sonicated DNA was treated with RNase A followed by phenol:chloroform:isoamyl alcohol extraction and purification with the MinElute PCR Purification kit. Next 500 ng of the sonicated DNA was treated with either 60 U of T4 Polynucleotide Kinase or mock treated (no kinase added) in a final volume of 300 μL of 1X T4 DNA ligase buffer and 100 μg/mL BSA for 2 h at 37°C. To fill in the fragmented ends, 200 μL of cold blunting mix containing 25 μL NEBuffer 3, 25 μL of dNTP mix (10 mM each dNTP), 2 μL 10 mg/ml BSA and 7.5 U of T4 DNA polymerase were added to each reaction and incubated at 12°C for 30 min. Reactions were extracted with phenol:chloroform:isoamyl alcohol, precipitated with 1 volume of isopropanol in the presence of 300 mM sodium acetate (pH 5.2) at -80°C overnight, washed twice with 70% ethanol and resuspended in 500 μL of H2O.

Modified linkers (KLM-PCR Oligo_1: 5’ GCG GTG ACC CGG GAG ATC TGA GTT C 3’, Oligo_2: 5’ GAA CTC AGA TC 3’) [17] included a single base change (GAATTC to GAGTTC) in KLM-PCR Oligo_1, which disrupted the GAATTC palindrome thus improving the efficiency of the PCR. To anneal the linkers, 510 μL of 40 μM Oligo_1 and 490 μL of 40 μM Oligo_2 were combined with 250 μL of 5X Duplex Buffer (100 mM Tris–HCl pH 8.0, 0.1 mM EDTA, 250 mM NaCl, 25 mM β-mercaptoethanol) and aliquoted into 100 μL in PCR tubes. To test the efficiency of ligation under different concentrations of PEG-8000 [18], ligations were set up using 20 ng DNA from T4 Polynucleotide Kinase treated/untreated samples with 5%, 12.5% and 15% PEG-8000 using T4 DNA Ligase Buffer, 50% PEG-8000 solution, BSA (15 μg), T4 DNA Ligase (400 U) and modified, annealed linkers (75 μM), and were incubated at 16°C for 16 h followed by purification using the MinElute PCR Purification kit. The following steps were performed in a thermocycler: 95°C for 3 min, 55°C for 1 min, 0.1°C/sec to 48°C, 48°C for 3 min followed by 2°C decrements in temperature, holding for 3 min at each step until 4°C was reached. Real-time PCR of the ligation reactions was performed using iQ™ SYBR® Green Supermix (Bio-Rad). Taq DNA Polymerase and PfuTurbo® DNA polymerase (Agilent) were added to the supermix to reproduce the LM-PCR conditions, as LM-PCR required a non-hot-start DNA polymerase for initial fill-in of ligated products. QPCR was performed in triplicate using KLM-PCR Oligo_1 and in duplicate using a gene specific (Rbp3) primer pair as a loading control for DNA in each reaction. The cycling conditions were 72°C for 3 min for initial fill-in, 94°C for 2 min for initial denaturation, followed by 35 cycles of (94°C for 15 sec, 95°C for 15 sec, 60°C for 30 sec, 72°C for 30 sec and 72°C for 30 sec for image capture) and a dissociation curve. The relative fold-change between the various ligation conditions (+kinase vs kinase at 5% PEG, 12.5% PEG vs 5% PEG, 15% PEG vs 5% PEG, 12.5% PEG/+kinase vs 5% PEG/-kinase and 15% PEG/+kinase vs 5% PEG/-kinase) were calculated and plotted for five replicate experiments, along with the standard error of the mean (Figure 1A). For the microarray samples, the KLM-PCR conditions included phosphorylation of DNA by T4 Polynucleotide Kinase prior to the blunting reaction (see above), modified linkers and 13% PEG-8000 instead of 5% PEG-8000. Either 10 ng of fragmented DNA (input) or 10 μL of methylation-enriched DNA in 100 μL Buffer EB (Qiagen) was treated with 10 U of T4 polynucleotide kinase in 150 μL of 1X T4 DNA ligase buffer. After 1 h incubation at 37°C, 1 U of T4 DNA polymerase, 10 μL of 10X NEBuffer3, 10 μL 10 mM dNTPs and 2 μL 10 mg/ml BSA in a volume of 50 μL were added to a total volume of 200 μL and reactions were incubated at 12°C for 20 min. Amplified DNA was quantified using a NanoDrop (Additional file 1: Figure S1).
Microarray labeling and hybridization
Fragmented (input) DNA was labeled with Cy3, and enriched (methylated) DNA was labeled with Cy5, using the NimbleGen Dual-Color DNA Labeling kit (Roche) according to the manufacturers’ instructions. Samples were hybridized to the custom NimbleGen 2.1M feature mouse CHARM microarray at the Johns Hopkins Medical Institutions Deep Sequencing & Microarray Core Facility or the Johns Hopkins Bloomberg School of Public Health Genomic Analysis and Sequencing Core Facility.

MeKL-chip data analysis
Analysis of the MeKL-chip data was performed using the R/Bioconductor software for CHARM as previously published [14,31]. In brief, this method used genome-weighted smoothing of probes within genomic regions to identify T-DMRs. Results from each NimbleGen CHARM array contained two sets of raw data: input (untreated) DNA and methyl-enriched DNA. Hybridization quality was assessed by a signal score, which examined the number of untreated channel signal probes that ranked above the background (anti-genomic control) probes. Successful hybridization was indicated by a higher signal score (usually greater than 0.85). After Loess normalization within samples for all control probes [31] and quantile normalization between samples had been performed, the relative methylation level for each probe was calculated as the ratio of the methylated probe to the input probe signal. As previously described [14], a t-test was adopted to identify differentially methylated probes between brain and retina samples (from triplicate arrays). Triplicate arrays consisted of either high-input (250 ng) or low-input (≤50 ng) retina and brain groups. The t-statistic cutoff in this study was set as \( P < 0.005 \). Consequently, T-DMRs were constituted by neighboring differentially methylated probes. T-DMRs with less than three probes were excluded from further analysis.

The definition of local CpG density by Pelizzola et al. [32] was adopted to analyze the identified T-DMRs. In general, the local CpG density of one nucleotide represents a weighted \( 0 \) at the boundary and 1 at the center) count of CpG sites surrounding the nucleotide, that is, 300 bp upstream and downstream from the nucleotide of interest (a window size selected based on the average fragment size). The local CpG density of a T-DMR was then obtained as mean value of local CpG densities of all nucleotides within the T-DMR. In addition, the observed-to-expected CpG ratio (CpGO/E) of a T-DMR was calculated using the standard formula:

\[
\text{CpGO/E} = \frac{N_{C} C P G \times N_{G}}{N_{C} \times N_{G}} \times L,
\]

where \( N_{C}, N_{G} \) and \( N_{C} \) are the number of CpGs, and nucleotide Cs and Gs, respectively, and \( L \) is the length of the T-DMR sequence.

Pyrosequencing
Using the EZ DNA Methylation-Gold™ Kit (Zymo), 1 μg of genomic DNA was bisulfite converted. Bisulfite-converted DNA was eluted twice in 10 μL M-Ellution buffer and stored as 5 μL aliquots at −80°C. Genomic sequences surrounding the RefSeq genes were obtained using the UCSC Genome Browser for Rgs20 (Gene ID: 58175), Hes2 (15206), Nfic (18029), Cckbr (12426) and Six3os1 (100043902). Pyrosequencing primers were designed (Additional file 2: Tables S2 and S3) within the identified DMR locations using the PyroMark Assay Design Software (Qiagen). PCR was performed using 1 μL of bisulfite-converted DNA and HotStarTaq DNA Polymerase (Qiagen) under the following cycling conditions: 95°C for 15 min; 45 cycles of (94°C for 30 s, annealing temperature from Additional file 2: Table S2 for 30 s, 72°C for 60 s); 72°C for 3 min; 4°C hold followed by storage at −20°C. Amplicons were analyzed on a PyroMark Q24 Pyrosequencer as per the manufacturer’s protocols and methylation at the CpG sites was quantified using the PyroMark Q24 software version 2.0.6.

Additional files
Additional file 1: Figure S1. The KLM-PCR protocol. (A) Modification of the universal adapter oligo sequence. The original LM-PCR oligo contained a palindromic at the 3’ end [17]. Prevention of dimerization through the disruption of the oligo palindrome increases the amount of available oligo for KLM-PCR amplification. dNTPs are now the limiting factor in the amplification. (B) NanoDrop quantification of the mean micrograms of DNA produced after KLM-PCR. Bars represent the mean of duplicate experiments from amplification of 10, 25, 50 and 250 ng amounts of pre-enriched starting DNA, or 10 ng of unenriched (UE) DNA. Error bars show standard deviation (n = 2).

Additional file 2: Table S1. Top 20 differentially methylated regions identified between mouse retina and brain from 250 ng of starting DNA. Table S2. PCR amplification primers used for pyrosequencing the top five DMRs identified using MeKL-chip. Table S3. Sequencing primers used for pyrosequencing validation of the top five DMRs identified using MeKL-chip.

Additional file 3: Figure S2. Relative CpG methylation in the retina (red) and brain (blue) of the four other top T-DMRs (black boxes) evaluated using MeKL-chip (top plots) and pyrosequencing validation of the differential methylation (bottom graphs). See Figure 2C for description of MeKL-chip results. Pyrosequencing of CpGs within the T-DMR confirmed differential methylation (\( P < 0.001 \), Student's two-tailed, paired t-test) between the retina (red bars) and brain (blue bars) in a second cohort of mice. Error bars show the 95% CI (\( n = 5 \)).

Additional file 4: Figure S3. Pre-hybridization validation of enrichment from low-input (10, 25 and 50 ng) enriched DNA from retina and brain samples. The mixed effects regression model with a random intercept for the measures from triplicate QPCR of two PCR amplifications of the same sample were used to calculate the mean difference and standard error in fold enrichment. (A) Post-enrichment, Rbp3 (grey bars) and Rho (black bars) were enriched for methylated DNA in the brain samples for all amounts of starting DNA, whereas no enrichment was observed in unenriched (UE) DNA. (B) Post-KLM-PCR, QPCR showed maintenance of the enrichment pattern for methylation in the brain samples for Rho (black bars) and Rbp3 (grey bars) and no enrichment of the UE samples.

Additional file 5: Figure S4. MeKL-chip CpG site methylation profiles of the Rgs20 region identified as a T-DMR (highest ranked, \( P < 10^{-16} \), black box; lower ranked, \( P < 0.00391 \), dashed box) for the 250-ng
High-input samples in the brain (blue) and retina (red) (top plot) as previously shown in Figure 2. The 10 ng low-input sample at the same region of Rgs20 is shown for direct comparison (lower plot) in brain (blue) and retina (red). Each point is the relative percentage methylation for 1 probe in 1 sample. The 250 ng plot contains biological triplicates and blue and red lines show the average methylation. The 10 ng plot contains one biological sample. The T-DMRs are still detectable in the 10 ng low-input sample.

Abbreviations
cC: Correlation coefficient; CHARM: Comprehensive high-throughput array for relative methylation; CpG/OE: Observed-to-expected CpG ratio; Ct: Threshold cycle; DMR: Differentially methylated region; dNTP: Deoxyribonucleotide triphosphates; HSBS: Hanks' balanced salt solution; KLM-PCR: Kinase-modified ligation-mediated PCR; LM-PCR: Ligation-mediated PCR; MBD: Methyl-binding domain protein; MeDIP: Methylated DNA immunoprecipitation; MeKl: MBD2b/MBD3L1 enrichment of DNA followed by KLM-PCR; MIRA: Methylated-CpG island recovery assay; MRE: Methylation-sensitive restriction enzyme; PCR: Polymerase chain reaction; PEG: Polyethylene glycol; QPCR: Quantitative PCR; RRBS: Reduced-representation bisulfite sequencing; T-DMR: Tissue-specific differentially methylated region; TSS: Transcription start site; WGA: Whole genome amplification; WGBS: Whole genome bisulfite sequencing.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
SLM, DJZ, VFO and SA designed the study. VFO performed the experiments. JW and JQ designed and performed the bioinformatics analysis. VFO, JW, SA, DJZ, JQ and SLM wrote the manuscript. DJZ, JQ and SLM supervised the project. All authors read and approved the final manuscript.

Acknowledgements
This work was supported by funding from the National Eye Institute (R21EY018703 to SLM, R01EY009769 to DJZ and the Wilmer Core Grant 5P30EY001765 to JQ), unrestricted funding from the Research to Prevent Blindness and the generosity of A Nixon. Additional biostatistics assistance was provided by J Wang funded by the Wilmer Core Grant. Microarray hybridization was performed by H Hao and L Orzolek at the Johns Hopkins Medical Institutions Deep Sequencing & Microarray Core Facility, and by A hybridization was performed by H Hao and L Orzolek at the Johns Hopkins Blindness and the generosity of A Nixon. Additional biostatistics assistance

References
1. Frommer M, McDonald LE, Millar DS, Collins CM, Watt F, Grigg GW, Molloy PL, Paul CL. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci USA 1992, 89:1827–1831.
2. Gu H, Bock C, Mikkelsen TS, Jager N, Smith ZD, Tomazou E, Gnirke A, Lander ES, Meissner A. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. Nat Methods 2010, 7:133–136.
3. Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, Jager N, Gnirke A, Stunnenberg HG, Meissner A. Quantitative comparison of genome-wide DNA methylation mapping technologies. Nat Biotechnol 2010, 28:1106–1114.
4. Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, Beck S, Butcher LM. Methyleme analysis using MeDIP-seq with low DNA concentrations. Nucleic Acids Res 2012, 40:e136.
5. Rauch T, Pfeifer GP. Methylated-CpG island recovery assay: a new technique for the rapid detection of methylated-CpG islands in cancer. Lab Invest 2005, 85:1172–1180.
6. Miura F, Enomoto Y, Dairiki R, et al. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adapter tagging. Nucleic Acids Res 2012, 40:e136.
7. Adey A, Shendure J. Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. Genome Res 2012, 22:1139–1143.
8. Seisenberger S, Andrews S, Krueger F, Anard J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. Mol Cell 2012, 48:849–862.
9. Gu H, Smith ZD, Bock C, Boyle P, Gnirke A, Meissner A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. Nat Protoc 2011, 6:468–481.
10. Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, et al. Comparison of sequencing-based methods to profile DNA methylation and identification of monosodium urate crystals in biologic samples. Nucleic Acids Res 2010, 38:e107–1105.
11. Nair SS, Coolen MW, Stratzer C, Song JZ, Statham AL, Strebenc A, Robinson MD, Clark SJ. Comparison of methyl-DNA immunoprecipitation (MeDIP) and methyl-CpG binding domain (MBD) protein capture for genome-wide DNA methylation analysis reveals CpG sequence coverage bias. Epigenetics 2011, 6:34–44.
12. Shakinovich R, Figueira ME, Melnick A. HELP (HiAfflial fragment enrichment by ligation-mediated PCR) assay for DNA methylation profiling of primary normal and malignant B lymphocytes. Methods Mol Biol 2010, 632:191–201.
13. Dedeuwenwaarder S, Defance M, Calonne E, Denis H, Sotiriou C, Fuls K. Evaluation of the Infinium Methylation 450K technology. Epigenomics 2011, 3:771–784.
14. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, Sen PN, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J. DNA methylation profiling. Nat Biotechnol 2010, 28:771–780.
15. Lee RS, Tamashiro KL, Ayye MJ, Murakami P, Seifuddin F, Herb B, Hsu Y, Rongione M, Feinberg AP, Moran TH, Pothak SB. Adaptation of the CHARM DNA methylation platform for the rat genome reveals novel brain region-specific differences. Epigenetics 2011, 6:1378–1390.
16. Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, Sen PN, Wyrick JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J. DNA methylation profiling of primary normal and malignant B lymphocytes. Methods Mol Biol 2010, 632:191–201.
17. Dedeuwenwaarder S, Defance M, Calonne E, Denis H, Sotiriou C, Fuls K. Evaluation of the Infinium Methylation 450K technology. Epigenomics 2011, 3:771–784.
18. Ben R, Robert F, Wryan JJ, Aparicio O, Jennings EG, Simon I, Zeitlinger J, Schreiber J, Hannett N, Kanin E, et al. Genome-wide location and function of DNA binding proteins. Science 2000, 290:2306–2309.
19. Pfeiffer BH, Zimmerman SB. Polymer-stimulated ligation: enhanced blunt- or cohesive-end ligation of DNA or deoxyribonucleic acids by T4 DNA ligase in polymer solutions. Nucleic Acids Res 1983, 11:7853–7871.
20. Merbs SL, Khan MA, Hackler L, Jr, Oliver VF, Wan J, Qian J, Zack DJ. Cell-specific DNA methylation patterns of retina-specific genes. PLoS One 2012, 7:e32602.
21. Bartolome MS, Weber BL, Brunkowski ME, Tilghman SM. Epigenetic mechanisms underlying the imprinting of the mouse H19 gene. Genes Dev 1993, 7:1663–1673.
22. Barker SA, Wang J, Sierra DA, Ross EM: RGSZ1 and Ret RGS: two of several splice variants from the gene RGS20. Genomics 2001, 78:232–239.
23. Wan J, Masuda T, Hackler L, Jr, Tones KM, Merbs SL, Zack DJ, Qian J. Dynamic usage of alternative splicing exons during mouse retina development. Nucleic Acids Res 2011, 39:7920–7930.
24. Maunakea AK, Nagarajan RP, Bilenky M, Ballinger TJ, D’Oucaux G, Fouse SD, Johnson BE, Hong C, Nielsen C, Zhao Y, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. Nature 2010, 466:253–257.
25. Soltner M, Locker M, Boy S, Taekema V, Bellefroid EJ, Perum M, Pfeifer T. Characterization and function of the bHLH-O protein XHes2: insight into...
the mechanisms controlling retinal cell fate decision. Development 2006, 133:4097–4108.

25. Rapicavoli NA, Poth EM, Zhu H, Blackshaw S: The long noncoding RNA Six3OS acts in trans to regulate retinal development by modulating Six3 activity. Neural Dev 2011, 6:32.

26. Bragado MJ, Perez-Marquez J, Garcia-Marin LJ: The cholecystokinin system in the rat retina: receptor expression and in vivo activation of tyrosine phosphorylation pathways. Neuropeptides 2003, 37:374–380.

27. Hollborn M, Tenchhoff S, Jahn K, Iandiev I, Biedermann B, Schnurrbusch UE, Limb GA, Reichenbach A, Wolf S, Wiedermann P, et al: Changes in retinal gene expression in proliferative vitreoretinopathy: glial cell expression of HB-EGF. Mol Vis 2005, 11:397–413.

28. Liang P, Song F, Ghosh S, Moriarty E, Qin M, Mahmood S, Fujiwara K, Igarashi J, Nagase H, Held WA: Genome-wide survey reveals dynamic widespread tissue-specific changes in DNA methylation during development. BMC Genomics 2011, 12:231.

29. Davies MN, Volta M, Pidsley R, Lunnon K, Diet A, Lovestone S, Coarfa C, Harris RA, Milosavljevic A, Tzoulis C, et al: Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. Genome Biol 2012, 13:R43.

30. Sambrook J, Russell DW: Molecular Cloning: A Laboratory Manual. 3rd edition. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press; 2001.

31. Aryee MJ, Wu Z, Ladd-Acosta C, Herb B, Feinberg AP, Yegnasubramanian S, Irizarry RA: Accurate genome-scale percentage DNA methylation estimates from microarray data. Biostatistics 2011, 12:197–210.

32. Pelizzola M, Koga Y, Urban AE, Krauthammer M, Weissman S, Halaban R, Molinario AM: MEDME: an experimental and analytical methodology for the estimation of DNA methylation levels based on microarray derived MeDIP-enrichment. Genome Res 2008, 18:1652–1659.

doi:10.1186/1756-8935-6-17

Cite this article as: Oliver et al.: A novel methyl-binding domain protein enrichment method for identifying genome-wide tissue-specific DNA methylation from nanogram DNA samples. Epigenetics & Chromatin 2013 6:17.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at www.biomedcentral.com/submit