Comprehensive DNA Methylation Analysis on the Illumina® Infinium® Assay Platform

Contributed by Daniel J. Weisenberger, David Van Den Berg, Fei Pan, Benjamin P. Berman, and Peter W. Laird, University of Southern California, Keck School of Medicine, USC/Norris Comprehensive Cancer Center, Los Angeles, CA 90033

INTRODUCTION

DNA methylation and chromatin modifications have been shown to be important markings in gene expression regulation, development, genetic imprinting, and the suppression of repetitive elements[1]. In addition, DNA methylation alterations are ubiquitous in human cancers, as promoter CpG island hypermethylation occurs in the context of a global decrease in methylation[2]. These widespread phenomena make access to fast and accurate methylation analysis methods imperative for discovering the underlying role that methylation plays in the course of normal vs. diseased development.

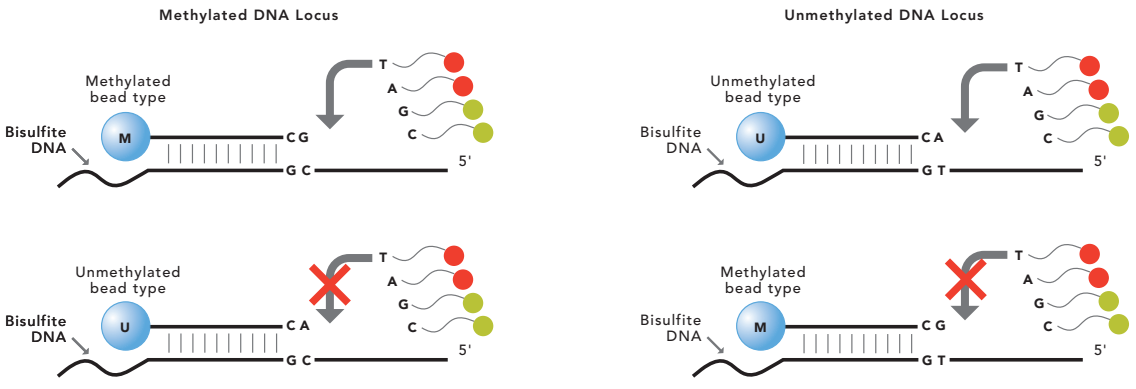
Current methods for genome-wide DNA methylation studies include the use of tiling microarrays[3,4] or bisulfite-genomic DNA sequencing of selected regions[5]. Although effective, these platforms require large amounts of sample and labor, and rely heavily on bioinformatic analyses, making them difficult to use in large-scale studies where sample amounts may be limited.

In an effort to overcome these challenges, we recently evaluated Illumina's Infinium HumanMethylation27 BeadChip. In this assay, quantitative measurements of DNA methylation are determined for 27,578 CpG dinucleotides spanning 14,495 genes. Only 1 µg of genomic DNA is required, giving the Infinium Assay a strong advantage over other methods. With the ability to provide genome-wide coverage for 12 samples concurrently, the Infinium platform is easily capable of high-throughput analyses. This throughput along with whole-genome DNA methylation analysis capabilities makes the Infinium DNA methylation platform highly suitable for detailing the biological role of DNA methylation in both normal and diseased cells, and for novel DNA methylation marker discovery.

DESCRIPTION OF THE INFINIUM METHYLATION ASSAY

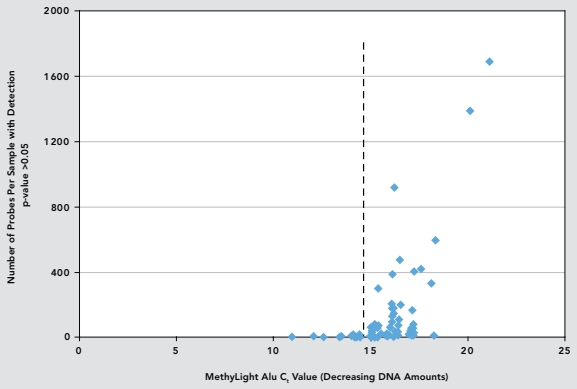
The HumanMethylation27 BeadChip uses Infinium technology, previously described for SNP genotyping[6], to perform genome-wide screening of DNA methylation

FIGURE 1: THE INFINIUM ASSAY FOR METHYLATION



The Infinium Assay for Methylation detects methylation status at individual CpG loci by typing bisulfite-converted DNA. Methylation protects C from conversion (left), whereas unmethylated C is converted to T (right). A pair of bead-bound probes is used to detect the presence of T or C by hybridization followed by single-base extension with a labeled nucleotide.

FIGURE 2: INFLUENCE OF DNA QUANTITY ON DNA METHYLATION DETECTION P-VALUE



In order to determine the influence of DNA quantity on the Infinium HumanMethylation27 BeadChip Assay, data points with p>0.05 for each sample were counted and plotted as a function of the MethyLight Alu C_t value. Higher C_t values indicate decreasing DNA amounts. As Alu C_t values increase (over a C_t=15), the number of insignificant data points dramatically increases (indicated by the vertical dashed line threshold), suggesting that a high yield of DNA methylation data in each experiment is strongly dependent on the input DNA amounts after bisulfite conversion.

patterns. A single BeadChip accommodates 12 samples. With 27,578 DNA methylation measurements per sample, there are 330,936 possible measurements per experiment. A small amount of genomic DNA (1 µg) is required for the initial bisulfite conversion step prior to performing the automated Infinium Assay. Unmethylated cytosines are chemically deaminated to uracil in the presence of bisulfite, while methylated cytosines are refractory to the effects of bisulfite and remain cytosine. A thermocycling program with a short denaturation step included for bisulfite conversion (16 cycles of 95°C for 30 seconds followed by 50°C for 1 hour), improved bisulfite conversion efficiency over the traditional 16-hour, 50°C incubation.

After bisulfite conversion, each sample was whole-genome amplified (WGA) and enzymatically fragmented. The bisulfite-converted WGA-DNA samples were purified and applied to the BeadChips. During hybridization, the WGA-DNA molecules anneal to locus-specific DNA oligomers linked to individual bead types. The two bead types correspond to each CpG locus—one to the methylated (C) and the other to the unmethylated (T) state. Allele-specific primer annealing is followed by single-base extension using DNP- and Biotin-labeled ddNTPs. Both bead types for the same CpG locus will incorporate the same type of labeled nucleotide, determined by the base preceding the interrogated "C" in the CpG locus, and therefore will be detected in the same color channel (*Figure 1*). After extension, the array is fluorescently stained, scanned, and the intensities of the unmethylated and methylated bead types measured. DNA methylation values, described as beta values, are recorded for each locus in each sample via BeadStudio software. DNA methylation beta values are continuous variables between 0 and 1, representing the ratio of the intensity of the methylated bead type to the combined locus intensity. The Infinium Assay can be completed in less than one week.

RESULTS AND DATA ANALYSIS

Samples Analyzed

A total of 96 samples were analyzed over eight BeadChips. The sample set included 24 control samples provided by Illumina (eight processed entirely by Illumina before hybridization). The remaining DNA samples were derived from primary colorectal cancer tissues, histologically normal colon tissues, normal and tumor breast primary tissues, ovarian cancer primary specimens, human embryonic stem cells, white blood cells, and sperm specimens. Finally, we included a DNA sample after treatment with M.*Sss* I as a methylated control sample and a WGA-DNA sample as a negative control for DNA methylation.

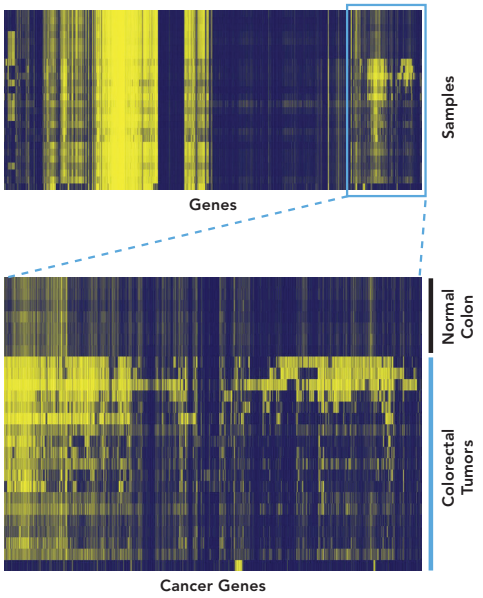
DNA Quantity

We used 1 µg genomic DNA for each sample; however, the DNA samples were collected from different sources and, as a result, there were differences in the absolute DNA amounts analyzed. To determine whether samples with low DNA amounts generated more uncertainty in the DNA methylation measurements, data points with a detection p-value>0.05 were counted and plotted as a function of the amount of bisulfite-converted DNA present using the Alu-based methylation-independent MethyLight reaction[7] (*Figure 2*). As the Alu C_t values increased, the amount of bisulfite-converted DNA decreased, and the number of data points with p>0.05 increased. These data suggest that bisulfite conversion of 1 µg of genomic DNA is critical for obtaining a high yield of DNA methylation data per assay.

BeadStudio Software

A Methylation Module in the BeadStudio software enables researchers to obtain and analyze beta values, Cy3 and Cy5 intensities, and p-value measurements for sample and control probes. The BeadStudio Methylation Module is also used to analyze methylation results achieved us-

FIGURE 3: IDENTIFICATION OF DNA METHYLATION EVENTS



DNA methylation in tumor and normal colonic epithelium using the Infinium DNA Methylation assay. Upper panel: Unsupervised hierarchical clustering of DNA methylation data obtained for 23,039 CpG loci on 20 colorectal tumors and 8 normal colonic mucosa samples using the Illumina Infinium DNA methylation assay. A total of 645,092 data points are represented. Lower panel: Expanded image of DNA methylation events of 2,250 genes for the colorectal samples. Similar to other assay platforms, the Infinium DNA methylation platform also highlights cancer-specific DNA methylation in colorectal tumors in a high-throughput manner.

ing the GoldenGate® Methylation Assay. Data are easily exported into Excel-compatible spreadsheets for further analyses. BeadStudio also has the ability to perform unsupervised two-dimensional clustering and generate heat maps to determine sample- and gene-methylation relationships.

Data Analysis

We evaluated the DNA methylation profiles of 96 samples on the Infinium HumanMethylation27 BeadChip, a subset of which consisted of 20 primary colorectal tumors and eight normal colonic mucosa from apparently healthy individuals. From these samples, a total of 645,092 data points were analyzed. The presence of DNA methylation in these samples is evident from the yellow colored pixels observed in the unsupervised hierarchical data clustering in *Figure 3* (upper panel). Expanding the image of DNA methylation events for 2,250 genes in the colorectal sample set highlights tumor-specific DNA methylation events (*Figure 3*, lower panel). These data demonstrate that the Infinium DNA Methylation Assay robustly identifies DNA hypermethylation events in colorectal cancer.

We identified cancer-specific DNA methylation alterations in breast and ovarian tumor specimens, as well as DNA methylation profiles of human embryonic stem cells (data not shown). Moreover, we observed tissue-specific DNA methylation differences between white blood cells and sperm specimens using the Infinium BeadChip technology (data not shown). The M.*Sss*I- and WGA-DNA samples served as positive and negative controls for DNA methylation, respectively, and the data obtained from these control samples highlighted the specific detection of methylated and unmethylated alleles on the Infinium platform.

Reproducibility

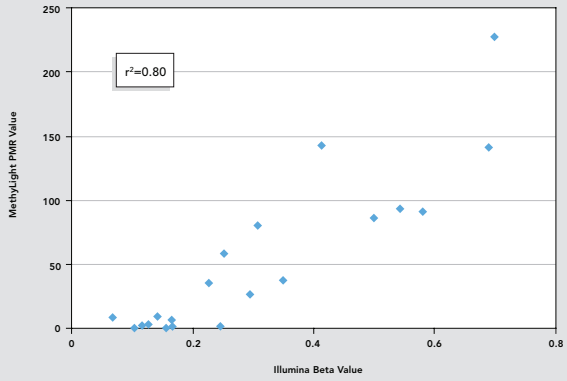
The reproducibility of the Infinium HumanMethylation27 BeadChip was demonstrated by assaying ten samples in duplicate, beginning with the bisulfite conversion step. The mean correlation coefficient was 0.987 for all 10 samples, suggesting that the Infinium platform delivers highly reproducible data.

Data Validation

To validate the data obtained using the HumanMethylation27 BeadChip, several samples were evaluated using the Infinium Assay and either the MethyLight[8] or GoldenGate Assays. Since DNA methylation information can vary based on genomic location, Infinium methylation data were compared to MethyLight data for gene regions with near-identical genomic coordinates. For example, the genomic coordinates of the TFAP2A differ by 23 nucleotides as evaluated by the MethyLight and Infinium platforms. In 20 samples, we found a high level of correlation of TFAP2A DNA methylation between the two platforms with an r^2=0.80 (*Figure 4*).

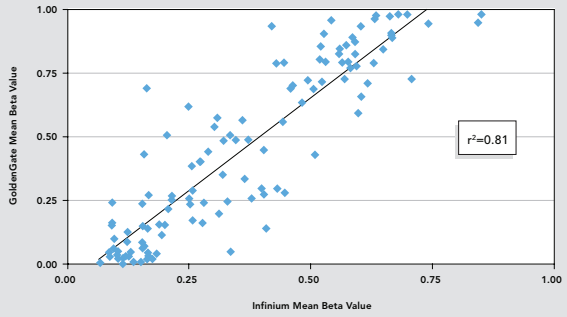
Data from the Infinium Assay were compared with that from the GoldenGate DNA Methylation Cancer Panel I over 117 CpG dinucleotides shared between the two platforms. A high concordance of beta values for these 117 DNA methylation measurements over 21 samples was observed with r^2=0.81 (*Figure 5*). The calculated correlation coefficients for each of the 21 samples between the Infinium and GoldenGate DNA methylation beta were 0.85. In evaluating the correlation coefficient for each sample as a function of its Alu C_t value, there was not a relationship between bisulfite-converted DNA amounts and the correlation coefficient.

FIGURE 4: METHYLIGHT VALIDATION OF TFAP2A DNA METHYLATION



Twenty DNA samples assayed using the Infinium HumanMethylation27 BeadChip were also assayed by MethyLight for the TFAP2A CpG island locus. The MethyLight PMR (Percent of Methylated Reference) value for each sample is plotted against the beta value determined by the Infinium Assay. An r^2=0.80 was determined, suggesting a high correlation between both platforms.

FIGURE 5: HIGH CORRELATION OF INFINIUM AND GOLDENGATE DNA METHYLATION DATA



To compare Illumina's Infinium and GoldenGate platforms, 117 CpG dinucleotides were tested on both assays. Beta values for 21 DNA samples were evaluated for common CpGs on both platforms. Data are plotted for the GoldenGate beta value as a function of the Infinium beta value. An r^2=0.81 was calculated, suggesting a high correlation between platforms.

SUMMARY

We found the Illumina Infinium HumanMethylation27 BeadChip to be a sensitive, reproducible method for genome-wide screening of methylation events. The Infinium platform generates data on a large number of informative loci (27,578 CpG measurements spanning 14,495 genes) for each sample with a throughput that rivals that of microarray platforms. Results achieved with the Infinium Assay are similar to those produced on other DNA methylation platforms, while requiring a lesser amount of input DNA (only 1 µg) and minimal time. This throughput, accuracy, and streamlined workflow enable analysis of the widespread impact of epigenetics in gene regulation.

REFERENCES

(1) Laird PW (2003) The power and the promise of DNA methylation markers. Nat Rev Cancer 3: 253-66.
(2) Jones PA and Baylin S (2007) The epigenomics of cancer. Cell 128: 683–692.
(3) Estécio MR, Yan PS, Ibrahim AE, Tellez CS, Shen L, et al. (2007) High-throughput methylation profiling by MCA coupled to CpG island microarray. Genome Res. 17: 1529–1536.
(4) Weber M, Hellman I, Stadler MB, Ramos L, Pääbo S, et al. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat. Genet. 39: 442–443.
(5) Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, et al. (2006) DNA methylation profiling of human chromosomes 6, 20 and 22. Nat. Genet. 38: 1378–1385.
(6) Steemers FJ, Chang W, Lee G, Barker DL, Shen R, et al. (2006) Whole-genome genotyping with the single-base extension assay. Nat. Methods 3: 31–33.
(7) Weisenberger DJ, Campan M, Long TI, Kim M, Woods C, et al. (2005) Analysis of repetitive element DNA methylation by MethyLight. Nucleic Acids Res 33: 6823-6836.
(8) Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, et al. (2000) MethyLight: a high-throughput assay to measure DNA methylation. Nucleic Acids Res 28: E32.

ADDITIONAL INFORMATION

Visit our website or contact us at the address below to learn more about Illumina's Infinium HumanMethylation27 BeadChip.

Illumina, Inc.
Customer Solutions
9885 Towne Centre Drive
San Diego, CA 92121-1975
1.800.809.4566 (toll free)
1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com

FOR RESEARCH USE ONLY

Pub. No. 270-2008-003 Current as of 25 March 2008