Supplemental Figure Legends

Supplemental Figure 1

Schematic of M-NGS library generation

Supplemental Figure 2

Regression analysis of M-NGS mapped reads and HMM output shows high correlation between sequencing runs. (**A**) Reads that mapped to chromosome 21 in LNCaP400bp-1 and -5, and PrEC400bp-1 and -5 runs were compared using the window size of 25bp. In LNCaP samples, a total of 33,627 reads were present at 25 bp windows with R^2 value of 0.9508, and in PrEC, 37,406 reads with R^2 value of 0.8556 was observed. (**B**) Linear regression analysis of all DNA methylation that occurred on CGIs showed high correlation (R^2 value = 0.9398 and 0.9819, n=5,734 and 4,966, respectively).

Supplemental Figure 3

Correlation between M-NGS vs Methylplex-Array and M-NGS vs MeDIP-Seq results. (A) Methylplex-array libraries made from LNCaP (Cy5) and PrEC (Cy3) cells were hybridized to Agilent human CGI microarray. Array results are displayed on the left in heatmap form (*Yellow*: hypermethylated in LNCaP; *Blue:* hypermethylated in PrEC), and were compared to M-NGS results (*Yellow:* methylated regions) on right. PrEC and LNCaP 200/400, indicates data obtained by M-NGS from size-selected bands excised at 200 and 400 bp during sample preparation. Gene names are displayed on the right, and genes previously identified as methylated in prostate cancer are indicated in blue, with genes known to be methylated in other cancers indicated in pink. The dataset can be found in Supplemental Table 2. (B) Overlap in methylated CpG islands located within or outside gene promoters (1500bps flanking the transcription start site) in LNCaP cells identified by M-NGS and MeDIP-Seq.

Supplemental Figure 4

Promoter methylation in LNCaP and PrEC cells identified by M-NGS. Each row represents a unique promoter region, ±1,500 bp from the transcription start site (white dotted line) at 100 bp window size. CpG island location is indicated in red in the first column. The methylation (yellow) observed in the corresponding location in each cell line is indicated (LNCaP second column and PrEC third column). Among 3,496 promoters represented, three distinct patterns of methylation were found. Methylation occurred either (1) on CpG islands, (2) in regions flanking the island (5' or 3') and (3) in promoters without any CpG island and the regions were methylated in either both or in a sample specific manner.

For example, regions I, IV, VII, X are specifically methylated in LNCaP, while II, V, VIII and XI are predominantly methylated in PrEC. The dataset is provided in Supplemental Table 3.

Supplemental Figures 5, 6 and 7

Representations of Methylplex NGS sequencing data used for nomination of methylated candidate gene promoters. UCSC Genome Browser visualizations of the M-NGS sequencing data (PrEC and LNCaP), as well as LNCaP MeDIP-Seq data, is shown for 18 genomic regions validated by bisulfite sequencing (shown in Figure 1D and Supplemental Figure 8). Methylated regions are a HMM output and shown as peaks (PrEC - light green; LNCaP M-NGS – light blue; LNCaP MeDIP-Seq - dark blue). A schematic of the target gene is shown below the peaks (exons as solid boxes lines connected by blue lines representing introns). CpG islands are represented by dark green horizontal bars. The red horizontal bars in the "BS-Seq-Amplicon" track displayed at the top indicate the regions validated by bisulfite sequencing. Additional information on the primer sequences, genomic location and amplicon size can be found in Supplemental Table 5.

Supplemental Figure 8

DNA methylation in *AOX1, C9orf125, NTN4, AMT, PPP1R3C and NAP1L5* gene promoters in LNCaP (L), PrEC (P), Universally methylated control DNA (MC-DNA, ZymoResearch Inc) and Unmethylated control Fetal DNA (UMC-DNA, Millipore Inc) were validated by bisulfite sequencing. Methylation status of each CG residue from 10 clones sequenced on both strands was analyzed using the BIQ Analyzer (Bock et al. 2005) program, where the height of the blue bar indicates percent methylation at a given position, yellow indicates no methylation, grey indicates not measured and numbers indicate the distance between analyzed CG dinucleotides. Two overlapping regions were monitored in *NAP1L5*, using the primers F1R1 and F2R2

Supplemental Figure 9

Significance Analysis of Microarray (SAM) identified re-expression of LNCaP methylated genes after 5-Aza treatment. LNCaP cells treated with either vehicle (DMSO) or 5-Aza were profiled by Agilent gene expression microarray. Genes methylated in LNCaP (n=973, identified by M-NGS and depicted in Supplemental Fig 4) were chiefly overexpressed (red points, n=246 at 5% false discovery rate) after 5'Azacytidine treatment in the microarray data as depicted in the SAM plot.

Supplemental Figure 10

Genes hypermethylated in LNCaP cells are enriched for biologically-significant concepts. (A) Molecular Concept Map (MCM) analysis of LNCaP methylated genes (black) revealed enrichment of gene signatures (red) repressed in prostate cancer and over-expressed in benign prostate tissues from multiple studies. Histone modification concepts (green), gene ontology concepts such as tumor suppressor genes (blue), genes previously known to be methylated from the Pubmeth database (Ongenaert et al. 2008) (pink) and genes methylated in prostate cancer by differential methylation hybridization (DMH, yellow) were also enriched. EF = embryonic fibroblast, ES = embryonic stem cells. (B) MCM analysis of PrEC methylated genes (black) shows the enrichment only for histone modification concept.

Supplemental Figure 11

Gene set enrichment analysis (GSEA) shows the association between gene repression and promoter methylation. (A) Methylated genes from LNCaP and PrEC cells were tested for their corresponding ranked gene expression in next generation transcriptomic sequencing (RNA-Seq). Both LNCaP and PrEC methylated promoters show enrichment with gene repression in LNCaP (p-value <0.0013) and PrEC (p-value <0.03) respectively. (B) While no significant association was observed between gene body methylation and gene expression in LNCaP (p-value <0.623), candidates with gene promoter methylation with and without the presence of CGIs in LNCaP cells are enriched with under-expressed genes (p-value <0.0039 and 0.0015, respectively).

Supplemental Figure 12

DNA methylation is associated with gene repression. Examples of promoter methylated targets (identified by M-NGS) associated with gene repression (RNA-Seq data). Methylation peaks (blue) in HMM output for PrEC and LNCaP cells are depicted in the first two rows respectively, followed by the gene promoter schematic from UCSC genome browser, where location of CpG islands are indicated in dark-green. Expression values for the selected genes in RNA-Seq data for PrEC (light green bars) and LNCaP (Blue bars) and displayed in the last row. Genes with promoter methylation in LNCaP cells (*TIG1*, *GSTP1*, *CALML3*, *TACSTD2*, *KCTD1*) and PrEC cells (*SPON2*, *GAGE* genes) had low transcript expression in their corresponding cell lines. *HIC1*, which is methylated in both LNCaP and PrEC, was minimally expressed in both.

Supplemental Figure 13

Methylprofiler PCR- bisulfite sequencing comparison in PCa panel. Samples that showed *WFDC2* promoter methylation (4, 5 and 6) by MethylProfiler qPCR analysis were independently confirmed by bisulfite sequencing in the corresponding region. Samples 1, 2 and 3 did not show any methylation by

both methods. Five clones were sequenced from each sample and circles represent CpG position monitored. Open circles represent unmethylated CpG dinucleotides. Closed circles represent methylated CpG dinucleotides. ",-" indicates a CpG island that was not measured. The Methylprofiler data (identical to Figure 5) is displayed alongside for comparison with bisulfite sequencing analysis.

Supplemental Figure 14

Cancer-specific DNA methylation enables switching of alternative transcriptional start sites (TSS) leading to transcript isoform regulation. Figure details characterization of *APC* gene similar to the data presented for *RASSF1* and *NDRG2* (Fig.6). **(A)** In contrast to the two examples presented in Fig.6, *APC* exhibits the reverse, with H3K4me3 on the longer variant 1 and CpG methylation at the shorter variants 2 and 3. **(B)** Preferential silencing and 5-Aza-induced re-expression of CpG-methylated variants of *APC* in LNCaP cells. Variants exhibiting CpG methylation on their TSSs show preferential silencing compared to variants with H3K4me3 marks in LNCaP cells. These variants were preferentially re-expressed upon treatment of LNCaP cells with 6µM 5-Aza. qRT-PCR data is normalized to variant expression levels in PrEC prostate primary epithelial cells or DMSO-treated LNCaP cells in the respective panels. **(A)** 5'RACE results validated *APC* variant-1 expression in LNCaP cells. **(C)** Isoform-specific expression patterns of *NDRG2* in the prostate tissue cohort (n=12). qRT-PCR results on variants 1-4 (red) and 5-8 (yellow) of *NDRG2* in 12 prostate tissue samples are shown as box plots. Variants 5-8, which share a common transcriptional start site, were significantly over-expressed compared to variants 1-4, which share a different common transcriptional start site, in adjacent normal (p-value = 0.012) and localized prostate cancer (p-value = 0.034) samples.

List of Supplemental Tables

Supplemental Table 1: Summary of M-NGS results from LNCaP and PrEC samples.

Supplemental Table 2: Agilent CpG array and M-NGS data overlap represented in Supplemental Figure 3A.

Supplemental Table 3: Promoter methylation analysis heatmap data from LNCaP and PrEC samples represented in Supplemental Figure 4.

Supplemental Table 4: List of gene promoters methylated in LNCaP cells (n=813).

Supplemental Table 5: Oligonucleotide sequences used in this study.

Supplemental Table 6: One-class Significance Analysis of Microarray (SAM) shows upregulation of genes targeted by methylation after 5-Aza treatment in LNCaP (Supplemental Figure 9).

Supplemental Table 7: MCM analysis of LNCaP hypermethylated genes (Supplemental Figure 10).

Supplemental Table 8: M-NGS results summary for the prostate tissue samples.

Supplemental Table 9: Promoter methylation analysis heatmap data represented in Figure 3.

Supplemental Table 10: Methylation status of previously characterized prostate cancer gene promoters and known imprinted genes in M-NGS and MeDIP-Seq data.

Supplemental Table 11: Oncomine Meta-analysis.

Supplemental Table 12: DNA methylation and H3K4Me3 ChIP-Seq data integrative analysis heatmap represented in Figure 7.

Supplemental Table 13: DNA methylation marks and alternate transcription start site (TSS)