We used BWA alignment tools [12] with the default settings to map the 36 bp unpaired reads to the hg19 human genome reference assembly [13]. After removing PCR duplicates with Picard, Samtools [14] and Picard (http://picard.sourceforge.net) were used to convert, sort, and index the aligned data.

DMR (differential methylated region between cancer and normal) was identified with two methods, MACS [15] and BALM [16], to increase the detective power of MethylCap-seq. Methylation peaks (hypermethylated regions) were identified using MACS in cancer and normal, respectively, same with our previous setting [Here our previous paper]. For BALM, to decrease the false positive detection of DMR (differential methylated region), dual-threshold strategy was applied with BALM. High-confidence threshold (threshold=0.975) in cancer hyper-methylated region screening and low-confidence threshold (threshold=0.95) in normal hyper-methylated region screening were set. Then cancer specific methylation peaks were defined as hyper-methylated regions. Similar, normal sepecfic methylation peaks were defined as hypo-methylated regions with reverse setting.

Whole genome methylation (methylation of each CpG) was inferred with BALM, which was processed to make Pearson correlation analysis among all the samples under R environment. The refSeq genes (UCSC genes) and corresponding CpG islands (CGIs) were downloaded from Table Browser of the UCSC database [11]. Bed files operation was treated with Bedtools [17] and some other Perl scripts.