Supporting Information

Microfluidic MeDIP-seq for low-input methylomic analysis of mammary tumorigenesis in mice

Yan Zhu¹, Zhenning Cao², Chang Lu^{1,*}

- 1. Department of Chemical Engineering, Virginia Tech, Blacksburg, VA 24061, USA
- 2. Department of Biomedical Engineering and Mechanics, Virginia Tech, Blacksburg, VA

24061, USA

Corresponding author: *Email: changlu@vt.edu Phone: (540) 231-8681

- 1. Figure S1-S6
- 2. Table S1-S7



Figure S1. DNA fragment size profiles. 10 ng DNA sample from GM12878 cell line was tested at different points of the process. (a) Sonicated DNA profile before MeDIP. (b) DNA profile after library preparation.



Figure S2. Microfluidic MeDIP-seq data on GM12878 cell line at various loci of chromosome (a) 3, (b) 5, (c) 7, (d) 12. Genome browser tracks for our microfluidic MeDIP-seq using various amounts of DNA (100-0.5 ng). Two replicates were profiled for each sample.



Figure S3. Saturation analysis on GM12878 MeDIP-seq data taken with different input amounts (0.5-100 ng) indicate the reproducibility of the genome wide coverage at regular genomic intervals given an increasing sequencing depth.



Figure S4. Calibration plots. RPKM is MeDIP Reads Per Kilobase Million. CG Coupling Factor is CpG density within given genomic widow. MeDIP-seq data taken with (a) 0.5 ng DNA from GM12878 cell line. (b) 100 ng DNA from GM12878 cell line.



Figure S5. Enriched GO terms on (a) Biological Process, (b) Mouse Phenotype, (c) Disease Ontology in the two periods of 6-16 weeks and 16-23 weeks for mouse mammary tumor development. Annotation terms were in different shades that represent their statistical enrichments (black for the highest statistical significance, and white for the lowest statistical significance).



Figure S6. Percent input data by MeDIP-qPCR. GEMIN4 and ZC3H13 are two known positive loci for human, GRICK3 and ZC3H13 are two known positive loci for mouse. N1 and N2 are two known negative loci for human and mouse. The error bars were calculated as standard deviation from 3 MeDIP-qPCR replicates. (a) Percent input data on human GM12878 cell line. 10 ng DNA was used as starting material. (b) Percent input data on 6 weeks mouse mammary gland tissue. 10 ng DNA was used as starting material.

Table S1. DMRs identified between 6 and 16 week mouse mammary samples. See attached "Table S1.xlsx".

Table S2. DMRs identified between 16 and 23 week mouse mammary samples. See attached "Table S2.xlsx".

Table S3. DMR-associated genes identified between 6 and 16 week mouse mammary samples. See attached "Table S3.xlsx".

Table S4. DMR-associated genes identified between 16 and 23 week mouse mammary samples. See attached "Table S4.xlsx".

Table S5. DMRs identified between normal tissue and adenoma intestinal tumor using data from ref. 26. See attached "Table S5.xlsx".

Table S6. DMR-associated genes identified between normal tissue and adenoma intestinal tumor using data from ref. 26. See attached "Table S6.xlsx".

Table S7. DMR-associated genes identified in both mammary tumor development (6 to 16 weeks) and early stage intestinal tumor (normal to adenoma) development. See attached "Table S7.xlsx".