Regarding the first reviewer’s comments

We thank you for your helpful comments. We respond to each comment, in turn, below. Changes to the manuscript has been made according to your suggestion.

1. First, unsuccessful or incomplete capture reactions might have resulted in the sequencing of non-methylated DNA fragments, leading to inconsistencies in or the absence of methylation enrichment in their sample.

Reply:

Yes, it is really good question and you are quite right. The capture of methylated DNA fragments is indeed a major issue, as suggested by the reviewer. We address this issue in 2 aspects of the study. In the technical aspect, we used spike DNA provided in the kit (an exogenous DNA fragment approximately 70nt in length containing 8 CpGs that are either fully methylated or unmethylated) in the MBD DNA enrichment steps to ensure that the methylated DNA fragments were retained while the unmethylated fragment were eluted (Additional File 1, Figure S1). In the biological aspect, BSP or MSP was performed to validate the methylation status of the gene targets identified via library data mining, which is challenging but manageable in our study (Figures 2A and Figure 2B).

2. Second, poor sequencing library complexity and CpG coverage often limits the statistical power to call differential methylation, affecting ultimately the reproducibility of the dataset. Explain.

Reply:

it is really good question and you are quite right. the sequencing library complexity and CpG coverage are critical for our present study. (1) MethylCap-seq by MBD is a reliable method to collect methylation information at the genomic level ([1] see ref. of Meissner A.). (2) Methodologically, exogenous spike DNA and internal DNA targets are used to confirm the accuracy of methylated CpG-containing fragments. (3) Our results show that the CpG coverage was approximately 64%, in the range of a similar study by others (66%-68% by Lin et al, see ref). Moreover, the 64% CpG coverage is a measurement of methylated CpGs, so the investigated CpGs, which include both the methylated and unmethylated CpGs, are much more than 64% of the total CpGs. Furthermore, the PC and PN samples were processed in parallel, which facilitates the detection of differential methylation between the PC and PN samples. Based on the above reasons, the authors consider the experimental data of methylated DNA enrichment to be reliable.

3. How was the quality control experiment performed?

Reply:

it is really good suggestion. The quality control experiments were performed in two aspects: fully methylated and unmethylated spike DNA was used to ensure that DNA enrichment was appropriately performed during the library establishment, and internal gene targets displaying a gradient of methylation statuses, such as GAPDH (unmethylated), CFTR (moderate methylated) and TP63 (highly methylated), were used to evaluate the methylated DNA enrichment process (see Additional File 1: Figure S1).

4. A schematic representation of work-flow adopted for primary and secondary data analysis must be included.

Reply:

This is a very good suggestion. We provide such a schematic representation in Figure 1A.

5. I suggest the authors to provide a clear-cut account on two sections of the manuscript:

First, the principal objective and aim of the study is ill defined in the introduction. I strongly believe that inclusion of the same in the last paragraph of introduction will strengthen the rationale of the study.

Reply:

As suggested, we revised the introduction paragraph (please see the Introduction section.(Page 3)

6. Second, what is translational utility of the work? This has to go a step-beyond explaining the importance of using the methylational profile as biomarker. Specifics must be provided for a few clusters only. A couple of lines on which specific experimental maneuvers will be necessary before these can be utilized in a clinical setting, must be included in the end segment of discussion. This will help to significantly elevate the impact of the article.

Reply:

As suggested, we added a short discussion of the translational utility of the present work. Please see the relevant text in the Discussion section (page 11, 12).

Regarding the second reviewer’s comments:

We thank you for your helpful comments. We respond to each comment, in turn, below. Changes to the manuscript has been made according to your suggestion.

1. In Figs. 1A-G, all data should be shown for both PC and PN to see if there is a clear difference in the methylation pattern.

Reply:

It is really good suggestion. To clearly demonstrate the comparison between PC and PN, we changed the manner in which the PC and PN data are displayed. Originally, we provided the PC-specific hypermethylation and hypomethylation sites without mentioning the corresponding PN-specific sites. However, we now provide the sites of PC-specific hypermethylation and PN-specific hypermethylation (alternatively, the PC-specific hypomethylation sites). We hope that the information provided demonstrates the comparison between the PC and PN samples. (Figure1A-H and all related tables)

2. In Fig. 2A, it is difficult to see which part of the gene locus is methylated or not. The authors should modify this figure.

Reply:

It is really good suggestion. We altered this figure to show the BSP region relative to the gene locus position, as shown in the figures.

3. As shown in table S15, change in methylation pattern is seen in only a subset of PC. Therefore, the authors should analyze whether the change in the methylation pattern is related to clinical profiles of PC.

Reply:

It is really good suggestion. Yes, as shown in Table S15, a portion of the PC samples were MSP-positive. As the PC tissue samples are challenging to obtain, we performed an association analysis between MSP and clinical parameters (such as gender, age, tumor location, tumor differentiation, stage, size and lymph node metastasis etc), but no relationship has been detected.

4. The authors state that ‘the methylation status of DLX4, ELAVL2, IRX1, PITX2, SIM2, TBX5, and TFAP2C was subsequently validated in tissue samples’. However, the data is not shown and the authors should show the data.

Reply:

It is really good suggestion. As suggested, the data are now provided (Figure 2B).

5. In table 2, the authors listed gene ontology enriched in differentially methylated genes in PC. Although it is interesting, it does not show whether genes involved in regulation of tumorigenesis are identified or not. The authors should also make a list of genes that may be involved in tumor suppression or promotion.

Reply:

It is really good suggestion. First, we collect data in the database of website which containing the item of oncogeneand tumorsuppressor gene, We found 742genes which met the above criterion. Then,We have analyzed the this study’s genes that were aberrantly hypermethylated in their promoter regions. Compare two set of genes,a portion of these genes are tumor promoters or tumor suppressors, which are listed in Additional File 14: Table S13.

http://atlasgeneticsoncology.org/index.html

http://210.46.85.180:8080/fcensus/

http://www.binfo.ncku.edu.tw/TAG/GeneFinder\_chr.php

http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/

http://www.tumor-gene.org/Oral/oral.html

6. In Fig. 3C, Orphan7-1 data is not clear. The authors should modify this picture.

Reply:

It is really good suggestion.As suggested, we referred to the original data and modified this figure.

Minor comments

There are numbers of typos (for example, p. 6 line 37, ‘Four candidate DMR genes which included…’, the number should be ‘six’). The authors should carefully check and correct them.

Reply:

Thank you for these comments. We read through the manuscripts again and did our best to correct all errors.

顺利

余坚

Cancer Epigenetics Program,

The State-key Laboratory for Oncogenes and Relatead Genes,

Shanghai Cancer Institute,

Shanghai Jiaotong University,

LN 2200/25, Xietu Road,

Shanghai 200032, China