Reviewer's report

Title: Global analysis of DNA methylation in early-stage liver fibrosis

Version: 1 Date: 31 October 2011

Reviewer: Rene Cortese

Reviewer's report:

General assessment:

In the manuscript by Komatsu et al. the authors performed genome-wide DNA methylation analysis on an animal model for liver fibrosis (CCl4-treated mice). In general, the experimental question is well defined: the authors enriched the methylated fraction of the genome using methyl-binding protein precipitation and interrogated it by deep-sequencing. Using bioinformatics tools, they detected 281 peaks corresponding to statistically significant differentially methylated regions between treated and control mice. Next, authors selected the differential methylation at the Spp1 gene and verified the methylation and expression status of this gene in treated and untreated animals.

The method used (MBP-Seq) corresponds to one of the more advanced technologies available for genome-wide DNA methylation analysis. Bioinformatics tools are standard. The method used for the verification of the methylation difference on the Spp1 locus (MBP-IP followed by PCR) is redundant with the discovery method. This reviewer would suggest using a method for verifying the methylation difference in which the methylation detection is based on a different principle (i.e. pyrosequencing of bisulfite-treated DNA or methylation-sensitive restriction PCR).

The data is sound and has potential interest in the areas of epigenetics of liver diseases. However, the fact that the authors limited the downstream to a single-locus reduces the impact of the findings. Data reporting is correct. Authors do not say if the data is publicly available, which might be a desirable feature for a large-scale study.

In general, the discussion and conclusion are adequately supported by the data. The discussion is rather short and focuses on the Spp1 gene. This reviewer would suggest discussing further on general trends of the methylation profiles in the animal model for liver fibrosis (i.e. discuss the causes and consequences of global hypomethylation. Is this trend present rather in repetitive or unique sequences? Do the hypomethylated areas correspond to any particular functional feature in the genome? ). Furthermore, the authors do not indicate any limitation. A paragraph on the technical limitations of MBP-Seq would be a valuable addition to the discussion.

References are correctly cited across the manuscript and the manuscript is well written. The title and abstract are correct. However, this reviewer strongly
recommends adding further analysis on genome-wide profiles and not only the follow up experiments on Spp1 differential methylation.

- Discretionary Revisions

1. The reports on the association of DNA hypermethylation and liver fibrosis are not clearly explained (Background section, third paragraph). The second sentence states that 5-aza treatment inhibits HSC activation “resulting in the transdifferentiation of HSCs to myofibroblasts”. The third sentence states that 5-aza treatment “ameliorates renal fibrosis by inhibiting proliferation of myofibroblasts”. As they are written, these sequences appear to contradict each other. It should be stated, for example, that “Activation of HSCs results in the transdifferentiation of these HSCs to myofibroblasts. Such activation is inhibited by 5-aza treatment leading to improvement of renal fibrosis.”

2. Results section, paragraph: Genome-wide DNA methylation profile of the CCl4-treated livers. The authors close the paragraph by saying “We identified 161 and 120 peaks in the CCl4-treated and control samples, respectively.” It would be highly desirable that authors present verification of some of these loci by a single-locus-based technique (i.e. bisulfite pyrosequencing) in addition to the verification of the Spp1 locus.

3. Results section, paragraph: In silico functional analysis of genes annotated by MBD peaks. In the second sentence, the authors refer to promoter, 3’-flanking regions and gene body. As the definition of these terms may vary according to the authors, it would be useful if the authors provide how they define them in the main text.

- Minor Essential Revisions

1. Methods section, paragraph: Methyl-binding protein (MBP)-based high throughput sequencing (MBP-seq). It should be mentioned which R package and version were used for describing the distribution of MBP-seq peaks in Figure 2B.

2. Results section, paragraph: Chromosomal distribution and genomic features of the MBP-seq peaks. The term “hypermethylation” in the first sentence should be clearly defined. The prefix “hyper-” is usually used in a comparative manner. Hence, hypermethylation would refer to an increased methylation found when one group (CCl4-treated?) is compared to another (control?). Given the fact that MBP-seq peaks correspond to putative methylated regions in both, CCl4-treated and control, animals, this reviewer would suggest to name this as such and not as hypermethylated regions in the descriptive analysis in Figure 2A.

3. Results section, paragraph: In silico functional analysis of genes annotated by MBD peaks. The name of the method in the title seems incorrect. Correct to MBP-seq.

4. Results section, paragraph: In silico functional analysis of genes annotated by MBD peaks. The authors mention “we found that cirrhosis, fibrosis, and HCC were identified only in the control sample”. It is not clear what the authors are
referring to: genes with functions related to these diseases, pathways, GO-terms? Authors should clarify this point.

5. Figure legend. Figure 3. Modify “vertical axis” for “horizontal axis”.

6. Figure 4. Label for panel 4C is missing.

- Major Compulsory Revisions

1. Methods section, Genome-wide DNA methylation profile of the CCl4-treated livers. Authors must say how many CCl4-treated and control mice were used for the genome-wide profile by MBP-Seq. Was any pooling involved?

2. Results section, paragraphs: Genome-wide DNA methylation profile of the CCl4-treated livers and Chromosomal distribution and genomic features of the MBP-seq peaks. MBP-Seq offers a high coverage of the genome at a high resolution. The authors correctly describe the method and the criteria for selecting differentially methylated regions and then describe the chromosomal distribution of the differences. To this reviewer the data produced by MBP-Seq can be further interpreted. Authors should describe other important features of the differentially methylated regions (i.e. association to CpG Island and other genome features, number of regions associated to unique and repetitive sequences, etc.)

3. Results section, paragraph: Chromosomal distribution and genomic features of the MBP-seq peaks. Authors should explain further how the functions in Figure 3A were selected. Were those arbitrarily selected? Are these functions significantly over-represented amongst all the functions in the IPA database?

4. Results section, paragraph: Chromosomal distribution and genomic features of the MBP-seq peaks. At the end of the paragraph, the authors mention that “the chromosomal distribution of the MBP peaks revealed that the number of control-specific peaks was greater than that of CCl4-specific peaks in most chromosomes”. Is this difference statistically significant? The authors must perform an appropriate statistical test for supporting this statement.

5. Results section, paragraph: In silico functional analysis of genes annotated by MBD peaks. The last sentence of the first paragraph states that “These results suggest that DNA methylation affects the onset of fibrosis and its progression to post-fibrotic disease.” This sentence must be tuned down. The data show, at most, that differential methylation in CCl4-treated and control samples is located in regions associated to these functional areas, but it does not necessarily mean they are affecting the onset and progression of liver fibrosis. Likewise, the last sentence in the third paragraph must be tuned down. In silico analysis suggests that hypomethylation of Ssp1 enhancer is associated to liver fibrosis, but it does not necessarily mean causality.

6. Results section. Paragraph: Epigenetic features and functional validations of the hypomethylated region upstream of Spp1. Authors performed the verification of the differential methylation at Spp1 by MBP-IP- PCR. Methylation detection in
MBP-Seq and MBP-IP-PCR is based on the same principle (binding of a methyl binding protein); therefore it is not a suitable method for verification. Authors should validate this differential methylation using another technique with a different principle for methylation detection (i.e. bisulfite treatment or methylation-sensitive restriction enzymes).

7. Methods section. Paragraph: Quantitative PCR (qPCR). Authors must detail how the expression data was analyzed. Was the expression of the target genes normalized against reference genes (or a panel of genes) with proven equivalent expression in CCl4-treated and control samples?

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I declare that I have no competing interests