Reviewer's report

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

Version: 2 Date: 27 December 2011

Reviewer: Rene Cortese

Reviewer's report:

Major compulsory revisions

- Results, Methods and Figure Legends section. The authors incorporated bisulfite sequencing data on the candidate region at the Spp1 locus. The data presented certainly support a differential methylation in this locus. However, the authors do not mention how the bisulfite sequencing data was quantified. How many clones were counted in each case? Are the detailed differences (100% and 70% in controls and CCl4-treated mice, respectively) the average across all CpG positions? Were the differences statistically significant?

- 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.

The authors answered:

“To indicate more clearly this pointed-out result, we calculated the relative number of MBP peaks for all chromosomes (Revised Fig. 2D). This statistic information can support our statement “The chromosomal distribution of the MBP peaks revealed that the number of control-specific methylation peaks was greater than that of CCl4-specific peaks in most chromosomes (Figure 2B-D)” in the revised Result section (Page 13, line 14-16).”

In my opinion this addition does not answer the question on whether the number of control-specific methylation peaks is significantly higher in each chromosome. An appropriate statistical test per chromosome should be used.

Minor essential revisions

- Results section: Epigenetic features and functional validations of the hypomethylated region upstream of Spp1. second paragraph. Spelling and grammar.

The authors wrote:

“Considering with the technical limitation of MBD-IP assay, a bisulfite modification assay was carried out to verify at a single-locus-based resolution
Complementary using these assays, we confirmed that the site annotated by the sequence database was actually hypomethylated by CCl4 treatment (Figure 4B and C).

Correct to:

“Considering the technical limitations of the MBP-IP assay, a bisulfite modification assay was carried out to verify the methylation values of both samples at a single-locus-based resolution (See Methods for details). Using these assays, we confirmed that the site annotated by the sequence database was actually hypomethylated by CCl4 treatment (Figure 4B and C).”

- Background section, third paragraph. Stylistic modification.

Upon my recommendation, the authors modified the original paragraph as follows:

“HSC activation is inhibited by 5'- Azacytidine (5'-Aza), a DNA methylation inhibitor, resulting in the transdifferentiation of HSCs to myofibroblasts [7].”

For the sake of clarity I would change to:

HSC activation resulting in the transdifferentiation of HSCs to myofibroblasts is inhibited by 5'- Azacytidine (5'-Aza), a DNA methylation inhibitor [7].

- Results section, paragraph: In silico functional analysis of genes annotated by MBD peaks.

The authors wrote: “We corrected the title name to” In silico functional analysis of genes annotated by MBD-seq” in Results section (Page 13, line 20).”

Change to MBP-seq

- 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.

The authors wrote:

“We apologize for our insufficient description. We clearly described the reviewer’s pointed-out sentence in Results section as follow; “Then, the functions of the genes with methylated sites, which are assigned to the gene body and the transcription regulatory region, are classified using in silico analysis software, IPA (See for details Methods for IPA analysis, and Discussion for the functional aspect of methylated sites within intergenic region).” (Page 13, line 21-24).”

I would suggest modifying also the sentence following this paragraph stating that “cirrhosis”, “fibrosis” and “HCC” are categories from the IPA software.

Discretionary revisions
I suggested that the authors present the verification of some loci, other than Spp1 by a single-locus-based method. The authors answered the following:

“We then attempted to investigate DNA methylation status of several other loci by using methylation-sensitive restriction PCR, which is more suitable for screening number of methylation locus. However, we could not find adequate enzymes since our identified loci consist of few CpG sites. Therefore, we further performed MBP-IP and a bisulfite modification assay, which require quite a time to perform. As a result, we confirmed that hepatocellular growth factor (Hgf) locus was hypomethylated in the CCl4 treated mice liver, and determined that Hgf mRNA levels were induced by CCl4 treatment (Fig. 1 in this letter). In this study, we would like to focus on the relationships between DNA methylation and gene transcription (This is mentioned in revised manuscript at Page 14, line 3-9). However, the methylation site in Hgf was located in 3'-flanking region, which leaves the discussion of the relationship with transcriptional activity. Therefore we decided not to include these data in the revised manuscript but instead we included the data in the Editor letter (Fig. 1 in this letter).”

I would suggest adding this information to the results and discussion section, accordingly, perhaps as supplementary information. The manuscript will certainly profit of it.

The authors wrote:

“We would like to thank the reviewer for this important comment. To answer this comment, we further analyzed our sequence data in detail, and attached annotation of CpG sites in each seq-peak in revised Additional file 1.”

I would suggest incorporating the main finding of this analysis into the text.

The authors wrote:

“We would like to thank the reviewer for this important comment. To answer this comment, we further analyzed our sequence data in detail, and attached annotation of CpG sites in each seq-peak in revised Additional file 1.”

I would suggest incorporating the main finding of this analysis into the text.

The authors wrote:

“We would like to thank the reviewer for this important comment. To answer this comment, we further analyzed our sequence data in detail, and attached annotation of CpG sites in each seq-peak in revised Additional file 1.”
“We apologize for our insufficient description. MBP peaks in promoter, 3’-flanking region, and gene body were analyzed in silico by using IPA software. Then, we selected and described the functions related to liver diseases in the revised manuscript (Revised Fig. 3). However, we attached all functions, which were significantly over-represented in IPA function (Fig. 2 in this letter). To avoid confusion, we did not include this IPA data in the revised manuscript.”

I would suggest mentioning clearly this in the text. The Fig.2 in the cover letter could be incorporated as supplementary material.

**Level of interest:** An article of importance in its field

**Quality of written English:** Needs some language corrections before being published

**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