Reviewer's report

Title: Comprehensive profiling of DNA methylation in colorectal cancer reveals three subgroups with distinct clinicopathological and molecular features

Version: 1 Date: 11 January 2010

Reviewer: Beatriz Carvalho

Reviewer's report:

Ang and collaborators have described in this study a comprehensive profiling of the DNA methylation of colorectal cancer, where they define three subgroups with distinct clinicopathological and molecular features. The research question is valid and interesting and the study is technically well performed. However I do have some major and minor remarks.

Major Compulsory Revisions:

- As the authors state themselves at the end of the discussion, one of the limitations of this study, is that only one CpG site is analysed per gene with this technique, thus how valid are the interpretations of the results? As they also admit, no relation to lack of expression is validated. Shouldn’t they try to validate at least part of the results with other techniques?

- In the same line of thinking, it would strength the study if the authors would have checked whether the Weisenberger CIMP panel genes are included in the 202 set of genes analysed with the Illumina platform. Is the methylation status of these 5 genes the same in both techniques? Are they falling in the CpG loci clusters A and C?

- Also in the same line, when the authors speculate that 2 CIMP-L MSI+ tumours (44 and 60 years old) are probably mutated rather than methylated? Why did not they check the methylation status of hMLH1 in their own data? That would strength the assumption.

- Results, the clustering of the normals do show two clusters although only based in one sample. Is this normal a “real normal”? No history of previous disease or other disease besides colon? Is the adjacent mucosa from all samples real normal? Are these mucosae revised by a pathologist?

- Interpretation of the results in the discussion is very limited, not to say absent sometimes. For example the authors describe the contradictions of the findings regarding mutations of KRAS but then they don’t come up with a possible reasoning. In the light of these results what is now the advice from the authors, concerning the use of a CIMP panel. Which one to use? Do we trust more the “classic” CIMP panels or do we trust more the Illumina CIMP panels? So what is the final message of this study?

Minor Essential Revisions:

- Supplementary files lack detailed legends. One is not able to understand what
is what, only guessing.

Discretionary Revisions:
- Shouldn’t the p-values and frequencies described in table 1 also be present in the body text? One should be able to read the text independently from the tables, and vice-versa.
- It would have been interesting if the authors would have looked at more molecular features besides the ones analysed standard in literature.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**
I declare that I have no competing interests