Reviewer's report

**Title:** Identification of valid reference genes for gene expression studies of human stomach cancer by quantitative real-time RT-PCR

**Version:** 1  **Date:** 1 January 2010

**Reviewer:** Ruprecht Kuner

**Reviewer's report:**

The authors describe a study to identify suitable endogenous reference genes for qRT-PCR analysis of gastric cancer. The overall workflow is well designed and acceptable. However, one should be more careful to generalize findings within one sample collective without applying them to independent collectives.

**Minor Compulsory Revisions:**

1. The outcome of this study should be phrased more carefully (for example in the Conclusions: This study indicates ... the most suitable reference gene...)

As mentioned above, the findings have not been applied to independent sample collectives. In gene expression studies, it is a common obstacle that results from smaller (n<100) sample collective have been overestimated. The authors can strengthen their findings in an independent study (public GEO or ArrayExpress gene expression study of gastric cancer), if this is not possible, they should avoid terms like "optimal" or "best". Furthermore, they should clearly discuss this limitation in the discussion.

2. The rules for suitable reference genes are described in the background on page 4.

The authors should discuss their findings with respect to these rules. For example for rule 3, is the abundance of your suitable reference genes similar to your target genes? Does it means that you need other reference genes for lower or higher abundant target genes?

3. Page 13: The authors applied the methods to three categorized groups. Why did they exclude the group of stomach normal samples? The comparison of stomach and non-stomach cell lines is very limited to the selection of specific cell lines. The comparison of cancerous and non-cancerous gastric tissues is the post promising part of the study and should be completed (also Figure 2 and 3).

4. One common qRT-PCR validation output means that a gene of interest is differentially expressed between two sample groups. Please refer in the paragraph page 16/17 the outcome of figure 5 in the following manner: P-value and fold change of GPNMB between normal and cancerous tissues (paired t-Test statistic) when using different reference genes.
5. Page 3: Abstract Result: Please shortly mention that GPNMB is the potential target gene; Page 5: Reference genes have been "described" for qRT-PCR studies....Page 8: Hexamer "or" poly-dT primed total RNA - Did the authors choose different protocols for the sample set? Page 14 first line: ...the identification of the minimal number of suitable reference genes....

6. The language has to be improved, e.g. include necessary adjectives.

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

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

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

'I declare that I have no competing interests'