Reviewer's report

**Title:** Combined use of expression and CGH arrays pinpoints novel candidate genes in Ewing sarcoma family of tumors

**Version:** 1  **Date:** 8 September 2008

**Reviewer:** Javed Khan

**Reviewer's report:**

In this manuscript the authors perform aCGH on 31 ESFT using the Agilent 44k platform and in parallel perform gene expression analysis on 19 of these samples using the Affymetrix U133 plus 2 platform. They perform RT-PCR on one of the DNA-RNA correlated genes, HDGF.

Several regions of gains and losses are found. They find some evidence of better outcome in patients with low copy number changes. They also find evidence in three samples of possible gain of the derivative chromosome. Finally they report that HDGF gene is over expressed in ESFT.

**Major Compulsory Revisions.**

It is interesting to perform parallel CGH and gene expression analysis, however several questions need to be addressed.

1. Table 1 should contain stage information.
2. It can be problematic to include both pre and post treatment samples for the genomic studies. This may impact the profiles. The authors should comment on this in the manuscript.
3. The authors should comment as to the percentage necrosis in the tumor samples they profiled that were post treatment, since in a lot of cases one can get >95% necrosis.
4. How did they make sure that the material post resection contained Ewing’s tumors? Please comment.
5. The legend for table 2 should be rewritten because it contains abbreviations that are not in the table.
6. I was not clear if the tumor samples were taken from the primary or the secondary, please clarify in the table 1.
7. Table 1 and 3 include REC in the stage column, Recurrence is not a stage, please change.
8. There should be a list of samples for which CDKN2A is deleted included in table 2.
9. Why did the authors not perform RT-PCR on HEATR3 since it is the most correlated, please comment on this.
10. Figure 2 and figure 4D should include the number of patients in each group.
11. The mean age for these patients is 20.7, which is higher than the median age of presentation for Ewing's, this may skew the data, the authors should comment on this.
12. Page 9 on commenting on the copy number change and relation to prognosis, the authors should comment as to whether it can independently (to stage, site, lung versus bony mets) predict outcome.
13. Page 9 talks about micordeletions, the authors should make a list of this and tabulate.
14. Figure 4 should be made clearer. What is the axis in figure 4A and 4B? I presume it is log2. This means the highest gain for this gene is 1.2 fold, this seems very low. I am not convinced that this is a real finding given the very low level of gains. Pooled CD34 and muscle cells seem a strange combination as a control since these are so divergent. The fold change also seems low (<2 fold). The slightly higher expression of this gene in ESFT compared with muscle or CD34 cells does not prove that this gene “seems important for tumorigenesis for ESFT...”. This language should be changed.

Level of interest: An article of limited interest

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