**Reviewer's report**

**Title:** Specific Gene Expression Profiles and Unique Chromosomal Abnormalities are Associated with Regressing Tumors Among Infants with Disseminated Neuroblastoma.

**Version:** 1  **Date:** 5 August 2008

**Reviewer:** Benedikt Brors

**Reviewer's report:**

This study describes a gene expression analysis using Affymetrix DNA microarrays on samples from neuroblastoma. The aim is to discriminate between stage 4 and stage 4S tumors and to describe the differentially expressed genes. Some of the genes have been included in a validation experiment that checked gene expression by qRT-PCR on the same samples. The authors claim over-representation of certain genomic regions, which are frequently affected by copy number variations, among their lists of differentially expressed genes.

**Major remarks:**

1. The gene expression profiling was done on the Affymetrix U95 platform (needs to be specified further: 95A, 95Av2, or all A through E arrays?). This platform is outdated and is no more available for some time. Since one of the authors (M. Alaminos) has been involved in a similar study using the same array platform (Alaminos et al., Cancer Res. 63:4538-4546, 2003), it could be suspected that the gene expression profiles presented here are partially or fully identical to the ones reported on earlier. While this would not totally preclude publication of this manuscript, it is still important information that should be given. Furthermore, the authors need to provide public access to their data; it cannot be retrieved from the CAarray database, at least not by public access. Furthermore, the identifier of the data record needs to be given.

2. The section on 'Gene expression data analysis' in the Methods part contains a number of flaws. The authors describe several methods for 'cut-off Family-wise error applied to select significant genes' (p.6, 2nd paragraph). From the methods that are described, however, i.e. step-down permutation, false discovery rate (FDR), and no type-I error adjustment, only the first controls the family-wise error rate (FWER). In fact, FWER and FDR are opposing, mutually exclusive concepts in statistics. Furthermore, analysis without adjustment for multiple testing (the 'raw' method) is not appropriate at all since the number of false positives will be unacceptably high just by random effects. It is also notable that the authors describe three methods but present results only on one, namely the 'raw' method. The 'raw' results should only be included for reference, but lists of differentially expressed genes may only be determined based on thresholds of FDR or FWER. The exact threshold (the authors name three: 0.01, 0.05 and 0.1) should be explicitly named.
3. There are a number of studies that addressed the same questions than the authors, namely Berwanger et al., Cancer Cell 2:377-386, 2002, Schramm et al., Oncogene 24:7902-7912, 2005, and Fischer et al., Clin. Cancer Res. 12:5118-5128, 2006. The findings presented here need to be compared in detail with the results from these previous studies.

4. The authors include tumors with MYCN amplification (MNA) from stage 4 into their analysis. As MNA will affect the expression of thousands of genes, this will significantly distort the gene expression profiles of stage 4 tumors. As a result, it is obvious from Additional File 2 that the lists for the comparisons: 4 (with MNA) versus 4S compared to: 4 (without MNA) versus 4S fail to show large overlap: only 38 probe sets are found in common among all three lists, and substantial numbers (96 probe sets for 4 vs. 4S < 12 months, 90 for 4 vs. 4S < 18 months, and 46 for 4 without MNA vs. 4S < 18 months) are only found in one of the three comparisons. The overlap between these analyses needs to be described, e.g. by a Venn diagram. The lack of overlap might be due to two different effects: first, the number of tumors investigated is small, thus the study lacks power to detect large proportions of the differentially expressed genes, and, second, the presented lists are highly likely to contain large numbers of false positives, see my comment #2 above.

Minor remarks:

1. Were all samples obtained prior to chemotherapy? Samples obtained after initial chemotherapy should be excluded since cytotoxic treatment is likely to considerably affect the gene expression of many genes.

2. The meaning of 'Hierarchical cluster analyses (...) using a multivariate unsupervised method, taking into account the relationships between gene expressions' (p.6, 2nd paragraph) is unclear. The authors need to report the distance metrics used (e.g., Euclidean distance, Manhattan distance or Correlation distance), and the algorithm for hierarchical clustering (single linkage, average linkage, complete linkage or Ward's).

3. How was the adjustment to FDR for the results from Fisher's exact test performed (p.7, 1st par.)? How many genomic regions were tested for association with the gene list?

4. In the discussion, p. 14, the authors speculate on prognostic relevance of their findings. They would need to demonstrate that the found gene expression changes are independent of other prognostic factors. It might turn out, e.g. that the differential expression of NME1 and NME2 is driven by some other factor (e.g. increases with age, or is driven by MNA) and thus is not itself causing tumor progression. In fact, with the current experimental design it is not possible to discriminate accidental changes driven by aberrant expression of some transcription factors from causal dysregulation that actively promotes aggressive behavior of tumors.
Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

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

Declaration of competing interests:

I declare that I have no competing interests.