Reviewer's report

Title: Expression of marker genes in the lymph nodes predicts the recurrence of squamous cell vulvar carcinoma.

Version: 1 Date: 1 December 2011

Reviewer: Marta Ewa Polak

Reviewer's report:

Kowalewska et al. seek to determine the candidate marker genes suitable for prediction of the recurrence of vulvar carcinoma. They utilise valuable material coming from human lymph node biopsies, and employ microarray gene analysis to examine them, which give a unique opportunity to investigate the molecular mechanisms beyond the VC progression. They perform the data analysis using freely available software enabling analysis of complex microarray datasets, and they properly correct for uncertainty and multiple comparison errors. The microarray results are validated against wider range of biologically relevant specimens. Authors succeed in identifying 5 novel candidate genes, two of which may possibly correlate with VC recurrence. Authors clearly acknowledge work upon which they had built, and cite the references adequately in the text.

Major concerns:

Major compulsory revisions

1) Results: Identification of expressed genes associated with LN metastasis:

Authors perform microarray analysis on LN pairs coming from 4 patients, but the selection criteria for the specific pairs/patients have not been presented. Similarly, authors choose to analyse data from one LN pair only, in the first instance. It is not entirely clear what the decision is based on, and why they did not perform analysis of all available microarrays to determine the candidate gene set? It is conceivable, that by doing so, they might have missed other (more suitable?) potential marker gene candidates.

2) Results: Validation of gene expression levels by real-time RT-PCR

Expression of seven candidate genes was further verified by qPCR in 71 LN samples. Authors present the Mean of group and p values for these data. It would be of interest for the readers to see the complete expression data, either in a table or in a figure, including the group medians and ranges.

3) Results: Correlations between gene expression and TTR

Finally the authors compare the levels of expression of selected candidate genes with time to groin recurrence (TTR). Using the Cox-model coefficient they identify two correlating genes, while using the Kaplar-Meier model authors model the effect of four candidate gene expression on TTR. It is not clear why the authors
present the correlations of these particular four genes, and what the correlations for the remaining three candidate genes were. It should also be explained in detail, what the cut-off values for each candidate gene expression are, how the authors divided the samples into high and low expressing of each candidate gene, how many LN fell into each category and what is presented on y-axis. It would also be of interest to determine whether the expression of any of the candidate genes in histologically uninvolved LN correlated with the TTR.

The results of Kaplan-Meier analysis should also be appropriately addressed in the discussion

2. Minor essential corrections:

1) Table 2: please explain what the term: Exp. No. - the expected number of probe refers to

2) Supplementary Table 1: please explain where the FC is derived from, if you simply divide the expression levels from positive and negative LN (="expression ratio"), the results are quite different to listed in the table.

3) Correlations between gene expression and TTR – section title missing (?)

4) Please add the details about the Cox test in the main manuscript, not only the table legend

5) Figure 1: p values: digits should be separate by a point, to a comma.

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