Reviewer’s report

Title: MMP-9 expression varies according to molecular subtypes of breast cancer

Version: 4 Date: 23 May 2014

Reviewer: Evette Radisky

Reviewer’s report:

In this manuscript from Yousef et al., the authors demonstrate by in silico analyses of publically available microarray data that MMP-9 transcripts are elevated in breast cancers in association with high histological grade, and are particularly elevated in triple-negative and HER2-positive breast cancers. They also demonstrate similar correlations with MMP-9 protein expression by IHC analysis of TMAs representing 300 human breast tumors, and find that MMP-9 protein is associated primarily with tumor cells rather than surrounding stroma. Finally, using clinical outcome data from 200 of the patients represented in the TMAs, they demonstrate that MMP-9 protein expression is associated with higher rates of local and distant metastasis, with shorter disease-free intervals and with shorter survival after relapse.

This is a timely study on a topic of high importance, and contributes a significant advance to the literature in the breast cancer field. As triple-negative and HER-2 positive breast cancers tend to be aggressive subtypes that advance rapidly, it is important to unravel the mechanisms that drive progression of these tumors, and to identify biomarkers that may help to further stratify tumor subtypes to guide clinical management.

On the whole the manuscript is clearly written and compelling; however, a few minor points would benefit from clarification:

(Minor essential revisions)

1. It would be very helpful if the statistical tests used to determine significance were identified in the figure legends where appropriate (in addition to the Methods).

2. In Figure 1, it is not clear to what the numbers 388, 103, 443, 116, 210 correspond (in the row between molecular sub-type and MMP-9 expression). Are these the total number of patients in each molecular sub-type category? This does not seem consistent with the text indicating that over 3400 microarrays were analyzed.

3. Figure 6 is derived from the review of 200 patient charts, but it is not indicated how many of these patients had some versus no metastasis. It is also not reported what proportion of the 200 patients had high versus low MMP-9. Can the authors show on the plot the distribution of high and low MMP-9 expression in the patients without metastasis as well as those with metastasis? Also, can the legend more clearly describe what comparison is being made that leads to the determination of statistical significance for LN, L.V., and Lung?
4. The description of the multivariate analysis model displayed in Table 2, and its interpretation, should be more clearly described. If I understand correctly, it appears that adjusting for the presence of metastasis eliminates the significant association of the HER-2 positive subtype with MMP-9 expression, but strengthens the significant association of the triple-negative subtype with MMP-9 expression. What are the possible reasons for this? What are the implications? (Discretionary Revisions)

5. In Figure 7, on the Kaplan-Meier plots it would be more usual to report time in months or years on the x-axis.

6. At the bottom of page 14 the authors state, “multivariate analysis could not identify any other significant factors involved in case of relapse.” The meaning of this statement is not clear.

7. In the Discussion the authors state that there has been no previous report in the literature that specifically correlated MMP-9 expression with individual breast cancer molecular sub-types. It would be appropriate for the authors to update their literature survey, as our group has very recently published a new study in Oncotarget that is highly complementary to the present work, and corroborates their observations on the significance of MMP-9 upregulation in triple-negative breast cancers.

**Level of interest:** An article of importance in its field

**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.