Reviewer's report

Title: Co-expression of putative stemness and epithelial-to-mesenchymal transition markers on single circulating tumour cells from patients with early and metastatic breast cancer

Version: 2 Date: 1 October 2013

Reviewer: vera Cappelletti

Reviewer's report:

The manuscript submitted by Maria Papadaki is well written and adequately introduces the field of CTCs in breast cancer and the urgent need to flank the simple quantification of CTCs with a biological characterization able to describe the heterogeneity within the CTC population.

The methods are well described and data are sound

My major concern relates to the message that the manuscript wants to convey to the scientific and clinical community. It is already known (which does not mean that additional studies must not be done) that CTCs can 'swing' between a fully epithelial and a fully mesenchymal phenotype and that epithelial and mesenchymal features can co-exist in the same cells. The paper by dr Papadaki confirms this important observation on an adequate number of breast cancer patients belonging to two different clinical settings. However, unfortunately no data are given to support the clinical meaning of the prevalence of certain CTC-phenotypes over others. Why do not show, at least for some index patients in the metastatic setting, how the CTC phenotypes relate to progression?

The absence of data trying to address the clinical meaning of this well-done and important observations on CTC phenotypes, pushes the entire manuscript from a clinical field into a more methodological area. In such a case however, to mainly convey a technical message the manuscript lacks some important data (eg. on repeatability, on possibility to transfer the method to other laboratories, data on healthy donors, data on yields with spiked cells, etc).

I would therefore suggest the Authors to clarify what is in the area that they want to mainly address (clinical or technical) and to improve/complete the manuscript accordingly.

Minor essential Revisions

Abstract:

In the Methods section please specify which breast cancer cells were used and when mentioning ‘CTC detection’ (line 7) shortly give the criteria for defining a cell as a CTC.

In the Results section avoid using ‘great majority’ without giving a percentage value.
Introduction
Last paragraph claims:
‘Although co-expression of such markers has been shown on CTCs using molecular techniques this has not been demonstrated on individual CTCs’. This is in my opinion a very important statement which deserves a better description of the data obtained in this study, focusing more on the description of CTC heterogeneity.

Methods
Are CTCs obtained from fresh or from frozen PBMCs?
A clear definition of what is considered as a CTC is lacking. It appears to be a cytokeratin positive cell fulfilling additional cytomorphologic criteria which are not detailed.

Discretionary Revisions
Results
I would suggest to the Authors to organize the Results section in a different way to emphasize the description of the heterogeneity of the CTC population. Now Results are organized by marker. This means separate paragraphs are used for each marker (ALDH and TWIST) in the two settings and a third paragraph is addressing the co-expression of the markers (which is the real nodal point about CTC heterogeneity) again in each setting.

Probably a subdivision in paragraphs by clinical setting analyzing the heterogeneous composition of the CTC population, followed by a paragraph making comparisons, would be more efficacious.

It would be probably better to describe CTCs heterogeneity referring to the total number of CTC positive cases rather than the total number of patients (both in the text and in the Figures).

Why not consider a pie chart or a single bar with a different color-code for each subtype, or even more than one bar detailing subdivision by one marker and followed by additional bars going more and more in detail taking into account the other markers? (always referring to CTC positive samples only).

Discussion
Avoid repeating data (and even p values) in the Discussion unless absolutely necessary.

Try to convey a clear message. Now it is mainly a repetition of Results.

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.