Author’s response to reviews

Title: Novel secretome-transcriptome integrated or secreto-transcriptomic approach to reveal liquid biopsy biomarkers for predicting individualized prognosis of breast cancer patients

Authors:

Astor Ankney (jaankney@email.unc.edu)
Ling Xie (xiel@email.unc.edu)
John Wrobel (john_wrobel@email.unc.edu)
Li Wang (liwang@email.unc.edu)
xian chen (xiang@email.unc.edu)

Version: 1 Date: 01 Apr 2019

Author’s response to reviews:

March 27, 2019

Dear Dr. Matteo Pasini:

Thank you very much for providing us this reviewing summary and helpful editor’s comments. We highly appreciate the reviewers’ recognition of highly significant nature of our manuscript. We also thank the reviewers for their thoughtful and constructive comments on our manuscript. We appreciate the concerns of the Reviewer #2. Accordingly we have performed additional analysis and clarified some misunderstandings. We believe the revisions made with these concerns in mind have strengthened our manuscript.

Below are our point-to-point responses to the reviewers’ critiques:

Reviewer 1. “Authors should improve the quality of figure 1 as several panels are not clear. Also, Authors should clearly describe this figure in the manuscript.”
Response: We appreciate the reviewer’s alert attention and have redesigned the schematic of our workflow. We also rewrote the panels for clarification. Because of space limit our strategy is described in detail through the text.

Reviewer 2.

Critique 1. “authors begin by showing a displeasure towards the 50-gene expression model and by detailing how this does not adequately represent all the different BC subtypes. Then they take just 5 cell lines - 4 cancer cell lines representing TNBC and luminal subtypes (just two subtypes) and one normal. How do these 5 cell lines represent all the major BC subtypes? In continuation of my above concern, why did not the authors consider HER2 overexpressing cell lines as a necessary model to be included here?”

Response: We apologize for this reviewer’s misunderstanding that was probably caused by a lack of clarification in our presentation. Basically, we intended to point out the limitation of the 50-gene expression model or PAM50 that can be used to classify different breast cancer subtypes or PAM50 subtypes. However, this 50-gene or PAM50 expression model cannot discriminate the distinct patient subpopulations having different clinical outcomes (e.g., survival, response to therapy, etc) within each PAM50 subtype. Therefore, we did not intend to express “displeasure toward the PAM50 model or suggest that the PAM50 does not adequately represent all the different BC subtypes.” In fact, our secreto-metabolome integrated or secretome-transcriptomics is a novel method to identify new additional gene expression/secretion patterns that identify the distinct subpopulations of breast cancer (BC) patients with distinct clinical outcomes (e.g., survival) within each single PAM50 subtype (e.g., luminal, or basal, or HER2+). Basically, our work does not abolish the PAM50 subtype model, but instead further stratifies patients within the existing PAM50 subtypes who will have worse prognosis than their peers. Hence, the five cell lines we selected do not, and are not intended to, represent all major BC subtypes. Due to the presence of a few highly abundant proteins in human blood, e.g. albumin, immunoglobulins, complement factors, etc., discovery proteomics using patient samples is inefficient and unreliable. Therefore we use the selected cell lines as a model system to identify potential markers and use the data of large patient cohorts to reinforce these identifications. These cell lines were selected because they represent both the most common BC subtype (luminal) and the most aggressive (basal-like) which account for approximately 85% of BC diagnoses. In consideration of the gap the reviewer has noted, we further analyzed the patient data and discovered that hierarchical clustering of TCGA patients using the basal and luminal SeCEP gene sets interspersed HER2 patients among luminal A/B patients with few exceptions. This suggests that our phenotype-based sorting does not regard the HER2-enriched subtype as a distinct group. Notably, the HER2-enriched subtype is the most histologically and clinically diverse group within the PAM50 indicating that HER2-enriched as it is currently defined may be less descriptive than other subtypes. We were therefore encouraged to analyze HER2 patient data
using our luminal SeCEP genes. These new results are now added in the Results section of the manuscript to reflect this analysis (Lines 411-430, Figure 3, and Additional Figure 7) and caveats are noted in the Discussion.

Critique 2. “Given the inherent differences between cell lines, even the 2 cell lines representing same BC subtype can be very different. Obviously, this is reflective of the variations observed in the patient populations. But my concern is how can similarly expressed/secreted proteins only represent the relevant proteins secreted by a given BC subtype tumor. In this era of personalized therapy, this is a very dangerous approach.”

Response: We’d like to clarify the novelty of our method: To avoid the difficulties in analyzing the mass spectrometry (MS)-incompatible plasma/serum (also see our response to the Critique 1) we first used established breast cancer PAM50-subtype cell lines. Taking an advantage of TCGA/METABRIC data with > 2600 patients and associated clinical outcomes our SeCEP method is able to identify additional genes showing altered mRNA expression patterns in correlation with distinct clinical outcomes in each single PAM50 subtype. The Discussion section has been revised to clarify our scope and claims. We did not intend to suggest that our work identifies all of the secreted proteins relevant to the characteristics of any given tumor. In order to efficiently connect patient clinical outcomes to the potential markers identified from the model system cell line secretomes, we must select practical criteria. The high expression/high secretion correlation provides a reasonable and straightforward link between cell line secretome and patient transcriptome. Alternate expression/secretion patterns are observed, however these are harder to quantify and correlate. Importantly, it is not necessary to identify every gene combination of relevance in order to identify some specific patient subpopulations with worse outcomes. Further, we recognize that the limitations of the present study preclude the ability to identify all such subgroups. Our work names several noteworthy combinations but more importantly provides a template for the further identification of combinations defining other subgroups.

Critique 3. “The statement in the Abstract ‘There is a pressing need for inexpensive and minimally invasive biomarker tests to easily and accurately predict individuals’ clinical outcomes, stage disease progression, and monitor response to treatments’ is a little too far reaching and clearly not supported well by the results. Need to tone it down.”

Response: Although the statement of pressing need is accurate, the abstract has been updated to better reflect the specific issues our manuscript attempts to address.
Critique 4. “The Introduction is a little too long. Can it be shortened to be more focused? The information deleted from here can be moved to Discussion.” And “The Results section is too elaborated but no efforts have been made to draft a compelling Discussion. Its just too short. This section needs much more writing.”

Response: We appreciate the suggestion by the reviewer and have moved a portion of the Background section to the Discussion section. As noted above, the Discussion section has also been expanded to address the issues raised by the reviewers.