Peer Review File

Article information: http://dx.doi.org/10.21037/gs-20-571.

Review Comments

Comment: *This manuscript by Le et al. proposes that Th2 cell infiltrations predict neoadjuvant chemotherapy response of estrogen receptor (ER)-positive breast cancer. The topic is important as many patients with BC do not respond to neoQT, and predictive biomarkers are sparse. Particularly, ER+ BC is less responsive, which has been related to decreased proliferation rates. The hypothesis is that Th2 ER+ BC may be more proliferative, thus responsive to neoQT.*
A cohort of 1069 BC cases from TCGA was used, and immune infiltration assessed with CIBERSORT. There was a high correlation with Ki-67 and proliferation score in ER-positive subtypes. High Th2 tumors achieved pathological complete response (pCR) significantly higher in ER-positive BC.
Overall, the paper is a little confusing and less convincing because Thorsson and xCell analysis are not in accordance, and by using Thorsson to infer Th2 and immune associations the authors failed to do so. Authors should focus on Th2 in ER+BC and its association with proliferation, which is the strength of the paper.

Reply: First of all, we would like to thank Reviewer A for taking his/her time and effort to review our manuscript and provide us with constructive criticism. We are delighted to learn that the Reviewer found our topic to be important.

Major aspects:

Comment 1: *Table 1 presents confusing data. In the all cohort analysis Th2 high is enriched in stage II but not III or IV; T2 tumors are enriched in Th2 high but T3 and 4 are associated with Th2 low. Therefore, bigger tumors seem to be Th2 low. We would expect ER+ analysis, the scope of the paper, which is missing here. These data will appear later, but Figure 3A is repetitive with Table 1. 3B and C make more sense and data is very significant.*

Response 1: We agree with the Reviewer that Table 1 may be confusing and somewhat appear repetitive with Figure 3A. Main purpose of Table 1 was to demonstrate the confounding factors between Th2 high and Th2 low group and we found that there was distribution difference in Stage and T category. On the other hand, the purpose of Figure 3A was to show Th2 levels by Stage and TNM categories in the whole cohort, where we found that there was no clear trend such as bigger tumors have lower Th2 levels, although there were statistically significant. This led to the further analyses by the subtypes as Figure 3B and 3C, where we see significant association of Th2 levels and cell proliferation in ER positive, but not in TNBC. With that said, we agree with the Reviewer that ER+ analysis is the highlight of the paper, and that Table 1 may mislead the readers that it is repetitive with Figure 3A.

Changes in the text: Table 1 was moved to Supplementary data as Supplementary Table S1 and text revised accordingly (Line 159-161).

Histological subtypes such as infiltrating ductal and tubular carcinoma showed statistical significance by Th2 expression levels (Supplementary Table S1).

Comment 2: *Thorsson analysis shows that tumor promoting M2 macrophages were significantly infiltrated in Th2 high tumors, but so were M1, which suggests that the microenvironment specificities are not being captured by this method. This was in fact revealed in the xCell analysis, where in fact tumor promotion/regression signatures are mutually exclusive. These differences should be explored and discussed. Importantly, subsequent analysis was based on Thorsson analysis that failed to associate the microenvironment with Th2.*

Response 2: We agree with the Reviewer that both M2 and M1 macrophages were significantly infiltrated in Th2 high tumors, which was consistent in both Thorsson and xCell analyses that suggest that both tumor promotion and regression immune microenvironment coexisted in Th2 high tumors. Given that this result disagrees with previous reports, we agree with the Reviewer that this should be explored and discussed.

First of all, it is not our intention to persuade the readers and Reviewer that Th2, defined by the Thorsson method, is identical to Th2 cells detected by gold standard, which is flow cytometry. Our intention is to introduce this modern technology of computational biology to the field that allows analyses of multiple large patient cohorts and allow investigation on clinical relevance when we define cells by transcriptome. One of the reasons we chose this method was because the original article that reported the Thorsson method (16) has been cited 766 times in 2 years after its publication and is now becoming a new standard. To this end, our main goal was to show the clinical relevance of Th2 cell levels defined by Thorsson method rather than to validate whether the Thorsson analysis reproduces the gold standard. That is the very reason why we pursued subsequent analysis using the Thorsson method. Additionally, we have recently reported that tumor promotion/regression signatures are not always mutually exclusive that although M1 and M2 macrophages commonly show the opposite trend (Oshi M. Cells 2020, PMID 32650578; Oshi M. Int J Mol Sci 2020, PMID 32331421) we found that both were significantly decreased in intratumoral adipocyte-high breast cancer (Tokumaru Y, Int J Mol Sci 2020, PMID 32796516).

Changes in the text: We added the following sentences in the discussion section, Line 318-338.

Further, we have recently reported that tumor promotion/regression signatures are not always mutually exclusive and although M1 and M2 macrophages commonly show the opposite trend (19, 36), we found that both were significantly decreased simultaneously in intratumoral adipocyte-high breast cancer (37).

The gold standard to analyze tumor immune microenvironment is flow cytometry of fresh tissue samples. Occasionally immunohistochemistry of fixed slides has been used, however, quantification of pathological analyses is known to be inaccurate. Although it proves value in basic science research, their utility in a large sample size for clinical patients can be challenging given the limited access to fresh samples, cost, and labor. Thus, to overcome this difficulty we have defined Th2 following the methods reported by Thorsson (25) and validated the results using xCell. It was not our intention to prove that Th2 defined by the Thorsson

method, is identical to Th2 cells detected by gold standard. Our intention was to introduce this modern technology of in Silico computational biology to the field that allows analyses of multiple large patient cohorts and allow investigation on clinical relevance when we define cells by transcriptome. One of the reasons we chose this method was because the original article that reported the Thorsson method (25) has been cited 766 times in 2 years after its publication and is now becoming a new standard. To this end, our main goal was to show the clinical relevance of Th2 cell levels defined by Thorsson method rather than to validate whether the Thorsson analysis reproduces the gold standard.

Comment 3: *Likewise, authors failed to associate Th2 high with specific Th2-related cytokines, and association with TGFb and IFN signatures was the opposite of expected.*

Response 3: We agree with the Reviewer and were also surprised to find that Th2 high tumors were not associated with Th2-related cytokines and were associated with TGF-beta and IFN gamma signatures in the opposite direction from the previous report. These results were not what we expected. But considering that our study is the first to investigate this association in large human breast cancer patient cohort and are not repeating the same experiments of the earlier reports, it is possible that the mechanisms proved in cell culture and animal experimental settings may not be applicable in the patients. As we have mentioned in the text and in Response 2, we were only able to analyze large patient cohorts because we used in Silico computational biological approach, which is not a gold standard; however, we believe it does shed light on what is going on in the patients. To this end, we believe that these findings that do not align with previous reports are also worth reporting.

Changes in the text: We added the following sentences in the discussion section, Line 313-318.

We were surprised to find that Th2 high tumors were not associated with Th2-related cytokines and were not associated with TGF-beta and IFN gamma signatures in the way of the earlier reports. However, considering that our study

is the first to investigate this association in large human breast cancer patient cohort and are not repeating the same experiments of the earlier reports, it is possible that the mechanisms that were proved in cell cultures and animal experimental settings may not be applicable in the patients.

Comment 4: *The focus of the paper comes only at Figure 3B onwards, with more interesting and convincing data.*

Response 4: We are very happy to learn that the Reviewer found our results from Figure 3B onwards are interesting and convincing. At the same time, as we described above, we believe showing what we found inconsistent from earlier publications is important since this is the first report that analyzed large patient cohorts.

Comment 5: *Figure 6A and 6B: top and bottom panel are repetitive, consider eliminate bottom panel or move to Supp Data.*

Response 5: We agree with the Reviewer that upper and lower panel of Figure 6, A and B, may appear repetitive, although they stand for different analysis and different meaning. The upper panels demonstrate the difference of expression of Th2 between RD and pCR group, which means that the tumor that achieve pCR has high amount of Th2 cells. The lower panels, on the other hand, demonstrate the pCR rate in Th2 low and Th2 high groups, which means that Th2 high is a predictive biomarker of neoadjuvant chemotherapy for ER-positive breast cancer that it is more likely to achieve pCR. With that said, upper and lower panels are demonstrating the same data from the opposite directions, therefore, we moved the upper figure to supplementary data and added the description below in the Results section Line 268-274.

Changes in the text: We added the following sentences in the Result section, Line 268.

First, we found that the patients who achieved pCR have significantly higher Th2 levels compared with the ones who had residual disease (RD) (Supplementary

Figure 2). This led us to investigate whether the Th2 high tumors associate with a significantly high ratio of pCR, which was the case (Figure 6). Interestingly, this was only seen in ER-positive and not ER-negative subtypes. This result suggests that high level of Th2 cell can be a predictive biomarker of ER-positive breast cancer to achieve pCR after neoadjuvant chemotherapy.

Minor aspects:

Comment 1: *L90: last sentence is out of context and repeats previous paragraph*

Response 1: We agree with the Reviewer that the last sentence of the third paragraph of the Introduction is out of context and repetitive of the previous paragraph, thus have been removed.

Comment 2: *L93-94: Grade is not necessarily linked to proliferation. Proliferation can be assessed by Ki67 or Mitotic index. Ref 4 must be replaced by a specific topic reference (eg. van Diest PJ, van der Wall E, Baak JP. Prognostic value of proliferation in invasive breast cancer: a review. J Clin Pathol. 2004;57(7):675-681. doi:10.1136/jcp.2003.010777)*

Response 2: We agree with Reviewer A and Ref 4 was replaced by suggested reference in Line 92-93 as follows.

Changes in the text:

Proliferation can be assessed by Ki-67 or mitotic index and increased proliferation has been shown to correlate with poor prognosis (5).

5. van Diest PJ, van der Wall E, Baak JPA. Prognostic value of proliferation in invasive breast cancer: a review. J Clin Pathol. 2004;57(7):675-81.

Comment 3: *L94-97: References are missing in both sentences*

Response 3: We thank the Reviewer for pointing out our oversights. We have included the references for both sentences. This is reflected in the revised manuscript as follows.

Changes in the text: Line 93-97

Proliferation can be assessed by Ki-67 or mitotic index and increased proliferation has been shown to correlate with poor prognosis (5). Triple-negative breast cancer (TNBC), which lacks expression of estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) are highly proliferative. Thus, they are more likely to respond to NAC than the ER-positive subtype (6, 7).

6. Aleskandarany MA, Green AR, Benhasouna AA, Barros FF, Neal K, Reis-Filho JS, et al. Prognostic value of proliferation assay in the luminal, HER2-positive, and triple-negative biologic classes of breast cancer. Breast Cancer Res. 2012;14(1):R3.

7. von Minckwitz G, Untch M, Blohmer J-U, Costa SD, Eidtmann H, Fasching PA, et al. Definition and Impact of Pathologic Complete Response on Prognosis After Neoadjuvant Chemotherapy in Various Intrinsic Breast Cancer Subtypes. Journal of Clinical Oncology. 2012;30(15):1796-804.

Comment 4: *Line 152-153: Section title must reflect better the results*

Response 4: We agree with the Reviewer that the section titles must reflect the results. Thus, we have changed the section title to below.

Changes in the text:

Th2 high breast cancer was associated with T and N category of cancer staging, Her2 receptor positivity, as well as infiltrating ductal and tubular carcinoma histology.

Comment 5: *Table 1: Title must be refined (demographic and clinicopathological xxx).*

Response 5: We agree with the Reviewer and have refined the title of now Supplementary Table S1 as below.

Changes in the text:

Supplementary Table S1. Demographic and Clinicopathological factors of Th2 High and Th2 Low groups in TCGA breast cancer cohort

Comment 6: *Extensive correction for typos and grammar is required.*

Response 6: We would like to thank Reviewer A for pointing out our oversight. All the authors have looked for typos and grammatical errors prior to submission of the revised manuscript.

Changes in the text: Changes can be viewed in multiple sections of the manuscript including the introduction, material and methods, results and a significant portion of the discussion. These significant changes have been highlighted in the revised manuscript.

Comment 7: *Figures 4-6 are called as 5-7 in the text.*

Response 7: We apologize for our sloppiness. We have made the corrections, as reflected in the revised manuscript, L240, L253, L267-268 as follows.

Changes in the text:

Line 240: Enrichment plots for these gene sets are shown in Figure 4.

Line 251-253: Utilizing the precalculated data published by Thorsson et al. we found that intratumoral heterogeneity was significantly higher in Th2 high breast cancer (Figure 5).

Line 264-268: We identified two cohorts (GSE25066, n = 502, treated with Taxane and anthracycline-based regimen and GSE23988, n = 61, treated with FEC (5-Fluorouracil, Epirubicin, Cyclophosphamide) and Taxane regimen) with gene expression data associated with response to NAC (Figure 6).