Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-23-1164

Reviewer A

1. **Comment:** Why the interest in EPSTI1, as this is not specifically associated with lung cancer.
Reply: We thank the reviewer for meticulous observation. CXCL13 and EPSTI1 are among the top 35 immunosensitive genes on the ICBatlas website, and EPSTI1 has recently been reported in lung squamous cancer. At the same time, we added the immunotherapeutic sensitization reasons for EPSTI1 and CXCL13 (around line 362 and 364).

2. **Comment:** If you focus on 3 genes, make explicit WHY these, and why ONLY these genes, i.e. "These genes are relevant or ICI-based immunotherapies...etc."
Reply: We thank the reviewer for this suggestion and we added it in the revised version (around line 73 and 75). Although our screening process was in the context of lung adenocarcinoma, the marker is promising for pan-cancer applications. We hypothesize that it could reflect epithelial, immune, and stromal interactions

3. **Comment:** Clarify sentences such as “Expression profiles were mostly derived from the processed data”
Reply: We apologize for any potential confusions caused, and have deleted the related sentence.

4. **Comment:** Ask for linguistic corrections of native speakers. Especially in the abstract sentences are incorrect, or simply not finished.
Reply: We thank the reviewer for pointing this out. We have revised the the manuscript, especially the abstract selection. We primarily add to the description of 1) the association of CDK1, EPSTI1 and CXCL13; 2) the role of the three-gene signature in pan-cancer and a comparison of previous pathways.

5. **Comment:** You show pan-cancer results and even spatial scRNA results for breast cancer from other sources but it is not clear why, or how this supports your observations for lung cancer. Either make this more explicit or leave it these non-lung cancer related findings.
Reply: We thank the reviewer for this profound question and explained in new main text (around line 218-219 and 328-330). Because [CXCL13, EPSTI1, CDK1] positive may be a marker for pan-cancer, a spatial transcriptome is warranted. On the other hand, the spatial transcriptome can compensate for the limitations of single-cell analysis. Importantly, we have revised the title of the manuscript to make it more relevant to the pan-cancer context, as the subgroups do hold promise for pan-cancer applications.

Despite the fact that the currently collected dataset for our study primarily concentrates on immunotherapy for lung adenocarcinoma, and the identified clusters exhibit higher immunogenicity in this type of cancer.

6. **Comment:** Having said this, it is also not clear what/to-what extent figures are produced by you, for this work and what is fetched from open sources.
Reply: We thank the reviewer for this profound question. Indeed, the inclusion of single-cell processed datasets were downloaded from other studies, and the bulk transcriptome was described in our previous manuscript (https://www.biorxiv.org/content/10.1101/2023.01.18.524544v1).

7. **Comment:** Your introduction reads like an abstract, here you should not explain what you found but rather, what you will do, and why.
Reply: We thank the reviewer for this suggestion, and revised the introduction selection. We have relocated the original elaboration in the results section into the introduction section to make it easier to understand.

8. **Comment:** You use multiple open sets but are only focused on your three genes of interest, make this explicitly clear, e.g. in the introduction.
Reply: We thank the reviewer for mentioning this and explain in new main text (around line 73 and 75). Quantity is a significant advantage. We believe the three-gene signature could reflect key pathways.

9. **Comment:** You use the phrase "CXCL13+ EPSTI1+ CDK1+ subpopulation" repeatedly but you refer to samples where the three genes of interest are all expressed positively, a less ambigous naming would be "[CXCL13, EPSTI1, CDK1] positive samples"
Reply: We agree with the reviewer, and have added the description to the abstract selection and part of the main text.

10. **Comment:** Having said that, your subpopulation definition is not clear, you say, there is positivity if the three genes have an expression > 0, but this does not consider measurement noise, please do include/discuss this.
Reply: We thank the reviewer for this suggestion and explained in new main text (around line 127-128, 300 and 380-381).

11. **Comment:** Please mention in the introduction, and in the abstract, the relevance of identifying [CXCL13, EPSTI1, CDK1] positive samples. I.e. as an indicator to apply ICI treatments more effectively.
Reply: We agree with the reviewer and added this description to the introduction and abstract selections.

12. **Comment:** I believe the writing can be improved a lot and with that the overall understanding of the work, this will however be a major revision in terms of the amount

of work.
Reply: We appreciate the reviewer for recognizing the work, and extensively revising misleading descriptions.

Reviewer B

1. In the text, the references should be cited **numerically and consecutively in the order of appearance**. Therefore, the first citation of ***Ref 33*** should be right after *Ref 32*. Please check and revise.

Reply: We've refined this content.

2. Please check if any reference should be added in the sentences since you mention "studies".

   "Previous studies have shown that CDK1 and CDK inhibitors could regulate the IFN pathway.[46]"
   "Previous studies have shown that O-glycosylation can be modified by proliferative factors (e.g., FOXM1 and EZH2), and it is plausible that transcription factor synergist mediator may regulate the activity of the downstream enzyme GALNT2.[48]"

Reply: We've refined this information. For the first one, we add new literature (ref. 46). Identification of Cyclin-Dependent Kinase 1 as a Novel Regulator of Type I Interferon Signaling in Systemic Lupus Erythematosus (doi: 10.1002/art.39543)

3. Figures
   - All abbreviations in figures and legends should be explained.
   - Please add the scale bars in the y-axis of Figure 1D.
   - There is a typo in Figure 2B.
   - It is suggested to unify the description of the y-axis in Figure 3A and in the main text.
   - It is suggested to double check the number in Figure 4B legend.
     "(B) Waterfall chart showing the top 5 most frequent mutations in SU2C-MARK LUAD cohort in high and low scoring populations, respectively."
   - Please indicate the meaning of "*" "**" in Figure 4C legend.
   - Please check whether it should be "low" in Figure 4C.

Reply: We modified the figures.