Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-23-1012

Reviewer A

In the paper by Wang et al, the authors, through a bioinformatic analysis of singles cell mRNA sequencing data derived from normal lung tissues of patients with lung cancer at different stages, including lymph node and brain metastasis, identified different macrophages subtypes and the dynamic changes in this population during lung tumor progression. Although the data suggest a progression in the composition of the macrophage population in the different tumor microenvironments the biological meaning of these changes is not sustained by the analysis performed.

Reply: We would like to thank reviewer A for their summarization of our work, while pointing out the limitations of our study.

Major points

1. There is no description of the statistical methods used to support the presented findings.

Reply: We are sorry for the lack of the statistical section and in this revised version, we have added this information (Page 5/Line 155-159).

2. The number of genes used to identify the different cell populations described in Figure 1 is not mentioned.

Reply: We are sorry for the lack of this information, and in this revised version, we have added this information. (Page 6/Line 168-174)

3. It is unclear whether the data supporting identifying the macrophage populations described is statistically relevant.

Reply: Thank you so much for your comment. The identification of macrophage subgroups is illustrated in **Figure 1A**, and the relative expression status of related markers is shown in **Figure1B**. During the process of single-cell analysis, the identification of different subgroups is mainly based on the expression status of specific markers, and all the involved marker genes have been calculated by Findmarkers() functions of R seurat package, suggesting that all the selected markers are statistically significant. We have involved the selection process of markers genes in the "Materials and Methods" section (Page 4/ Line127-129), and we are sorry for not providing this information in the original version.

4. The data is presented as a descriptive list of the differences between the populations identified but the biological meaning of the data is not discussed.

Reply: We are sorry for not discussing the biological meanings of the results in the original version, and in this new version, we have added the summaries in the result section (Page 9/Line202-209，Page 10/306-309，Page 13/407-420)

5. There is no rational explanation for the formation of the different groups considering the macrophage's biology, why these groups were considered? based on what functions?

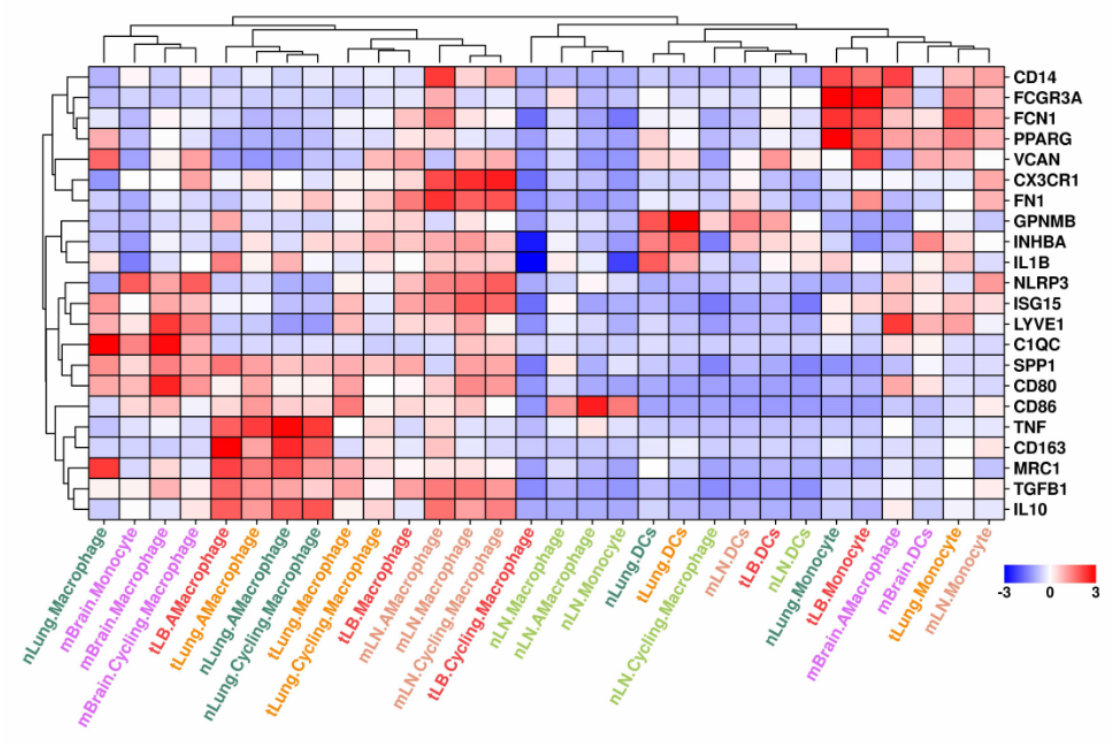
Reply: In the current version, we have clustered all myeloid into 5 different subgroups including macrophages, alveolar macrophages, cycling macrophages, monocytes and DCs based on the subgroup types described in Sorin et al and Travaglini et al. (Page 7/Line217-220). All the markers used to determine these subgroups are illustrated in Figure 1B.

6. What is the biological meaning of the findings regarding the tumor microenvironment in each tumor sample?

Reply: In the revised version, we have discussed the impact on tumor microenvironment, especially tumor immune microenvironment, such as the discussion of IL10 (Page 10/Line 310-313), overexpressed genes in SPP1+ macrophages and their potential involvement in the promotion of M2 polarization (Page 13/Line 411-424); LR interactions on tumor immune microenvironments (Page 14/Line 449-461)

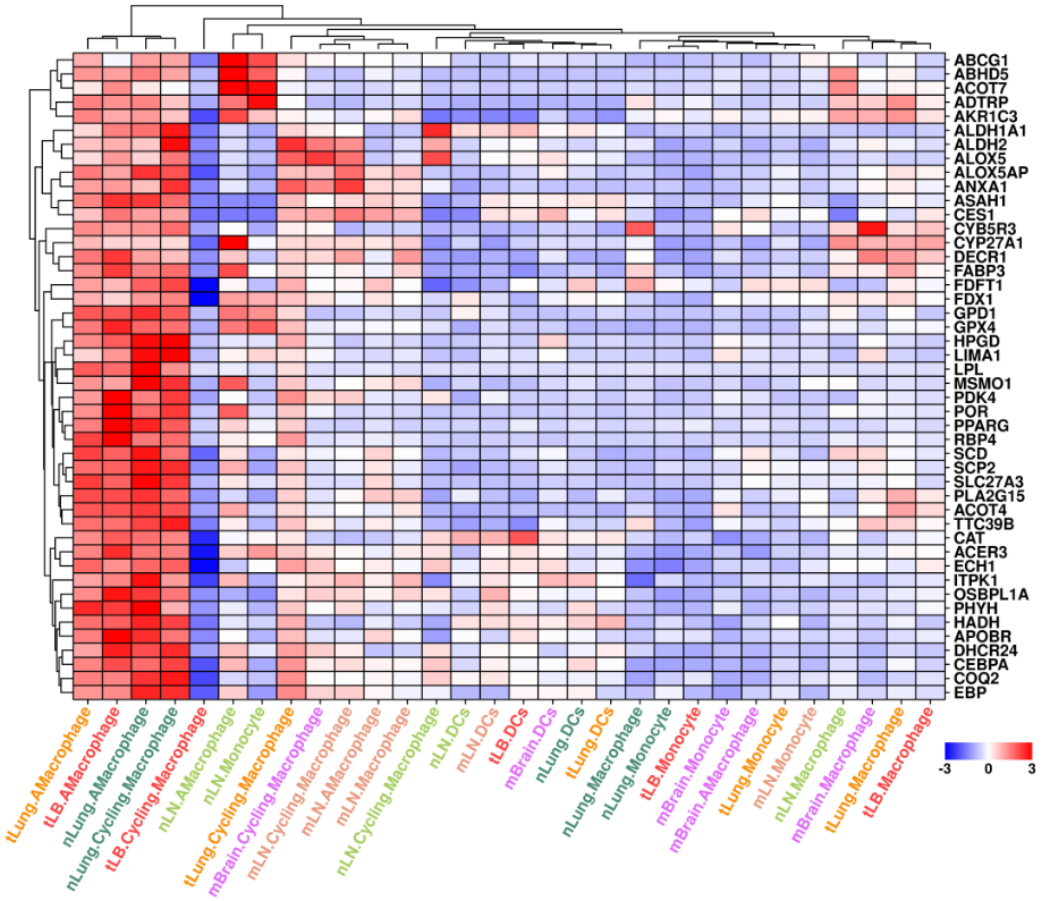
8. The M1 and M2 classification of macrophage function is obsolete, there are other functional subpopulations. Here the authors based this subpopulation using the differential expression of only 4 genes, how relevant is this?

Reply: We are sorry for these non-rigorous analysis regarding the separation of M1/M2 macrophage functions, and in this revised version, we have involved a macrophage-specific geneset and compared the expression status of these genes across different subgroups, as shown in Figure 2B. (Page 8-9/Line 248-269)

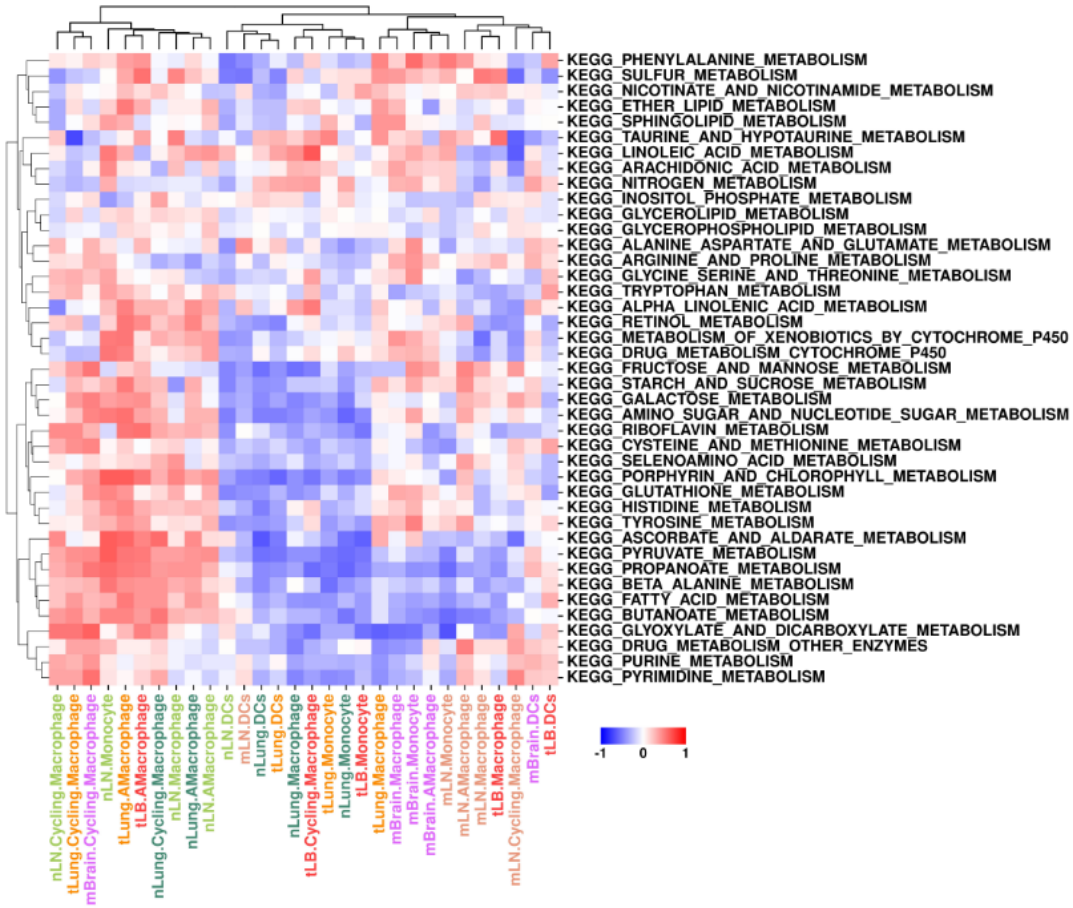


9. What about genes involved in metabolism? macrophages metabolism is correlated with its functional state.

Reply: Thank you for this suggestion, in this new version, we have examined the expression status of metabolic related genes (within group3) and the results are illustrated in Supplemental Figure 3. (Page 11/Line 346-350)



we also performed GSVA (Gene Set Variation Analysis) analysis of KEGG metabolism-related terms on subgroups from different origins, as shown in the below figure (Supplementary Figure 4)



10. It is necessary to describe all abbreviations used the first time it is used; some are not described at all.

Reply: We are sorry for these errors, and in the revised version, we have described all abbreviations.

11. For clarity it will be better to present the results and the discussion simultaneously in a single results and discussion section.

Reply: We agree with the reviewer that the combination of result and discussion sections would be clearer, however, the journal requires distinct sections of result and discussion, hence we have to keep them this way.
Change in the text: None.

We would like to appreciate review A for these constructive suggestions, and we hope that we have addressed all these concerns in this revised version.

Reviewer B

This manuscript described a secondary analysis from the Kim et al., Nat Comm 2021scRNAseq study on normal lung and lung cancer. The authors focused on changes in macrophage populations during tumor progression. The study is descriptive, with no hypothesis being tested or mechanistic insight into the disease process. The markers used to define the sub-populations of macrophages are not cited from the literature, although some markers are cited for the M1/M2 polarization states. It is unclear what this study adds to the field or to the original analysis.

Major points include, while the data in Figure 1 is somewhat interesting given that is shows the heterogeneity among the samples. However, the Kim et al., study already did an analysis of the macrophage subtypes, with more precision and depth than done here in Figure 4 of their paper. For example, they used a classifier for microglia, which are brain-specific macrophages and this current study did not. Furthermore, the fact that there are no neutrophils in this myeloid cell analysis suggests an issue with the tissue processing and is a limitation of the previous study that is now perpetuated by a secondary analysis (described well here https://doi.org/10.1016/j.ccell.2022.10.008). One analysis that was added and not done in the original paper was the trajectory inference, but it does not make sense. Alveolar macrophages are thought to predominately arise from lung resident cell populations, while the other populations are BM-derived. The authors may have been able to computationally 'see' transition between these cell states, but the cells are likely to evolve separately given their distinct origins.

Reply: We would like to thank reviewer B for pointing out the shortcomings of our study, and we do agree with the reviewer that the originality of this study is limited. The purpose of this study is to explore in depth of the polarization/transition status of monocytes and macrophages during the progression and metastasis processes of lung cancer, which is not detailly elaborated in the original paper. Moreover, through LR pairing analysis, we have identified potential interactions that might contribute to the polarization/transition process of these myeloids. Finally, through the analysis of SPP1+ macrophages, we have identified a set of 8 genes that might involve in the progression of LUAD. All these findings might co-contribute to understandings of LUAD progression mechanism.

The results of PeudoTime trajectory analysis do not represent true evolutionary process, and the purpose of this analysis is to explore the potential models explaining the biological progression changes among these different cell types. In this new version, we have added more analysis on enrichment items through the perspective of all subgroups to supplement and further explain the pseudoTime results, and we hope this could help to address the concerns (Line 351-363, Figure 5, supplemental Figure 3; supplemental Figure 4).

Minor points include that t-tests are not appropriate for the comparisons (one way ANOVA with multiple comparisons and a post-hoc correction would be much better) and that the plotting of the 'ratios' of macrophages is confusing and the samples are in different orders in each graph.

Reply: We would like to thank reviewer for this suggestion, and in this revised version, we have performed one-way anova with "bonferroni" corrections for all the analysis (Figure2A, Figure3B and Figure6). We also changed the orders of subgroups in Figure 2A to keep consistent.

Once again, we would like to thank reviewer B for the time and effort in reviewing our work.

Reviewer C

In this manuscript, the authors reanalyze a previously published scRNAseq dataset from lung cancer patients. While the original dataset included tumor, stromal, and immune compartments, all of which were investigated in the original article, the authors of the current manuscript focus on the myeloid compartment of the data, with the goal of dissecting changes in different myeloid subsets and interactions amongst the myeloid components.
Overall, the authors have done a good job of focusing on the myeloid compartments of the original dataset and confirming the conclusions drawn in the original paper. Their effort to enhance our understanding of how these cells may behave differently at various disease stages adds to our knowledge about the roles of alveolar macrophage and monocyte-derived macrophage in human lung cancer. This especially applies to their ligand-receptor analysis between monocytes and macrophages, suggesting dynamic and yet varying interactions between these cells at different sampling sites.
However, the authors' approach mirrors that of the original paper, necessitating additional revisions to improve the novelty and completeness of their findings. Please see the following sections for major and minor issues that I recommend the authors to address.

Major Issues:
1. In Figure 1, the authors present the distribution pattern of all previously identified subsets and the relative abundance by patients. However, considering the main objective of this manuscript, it would be more suitable to focus on the myeloid compartment in this effort, specifically the distribution of myeloid cells with different tissue origins (as presented in Figure 2). Therefore, Figure 1 may be moved to the supplemental material instead of being extensively discussed in the main text.

Reply: We would like to thank reviewer for this suggestion, and in the new version, we have put Figure 1 into supplemental section.

2. It may be worth examining the data by disease state to see whether there are significant changes in the distribution pattern or relative abundance of each myeloid subset. Undetermined cells should be excluded from the analysis.

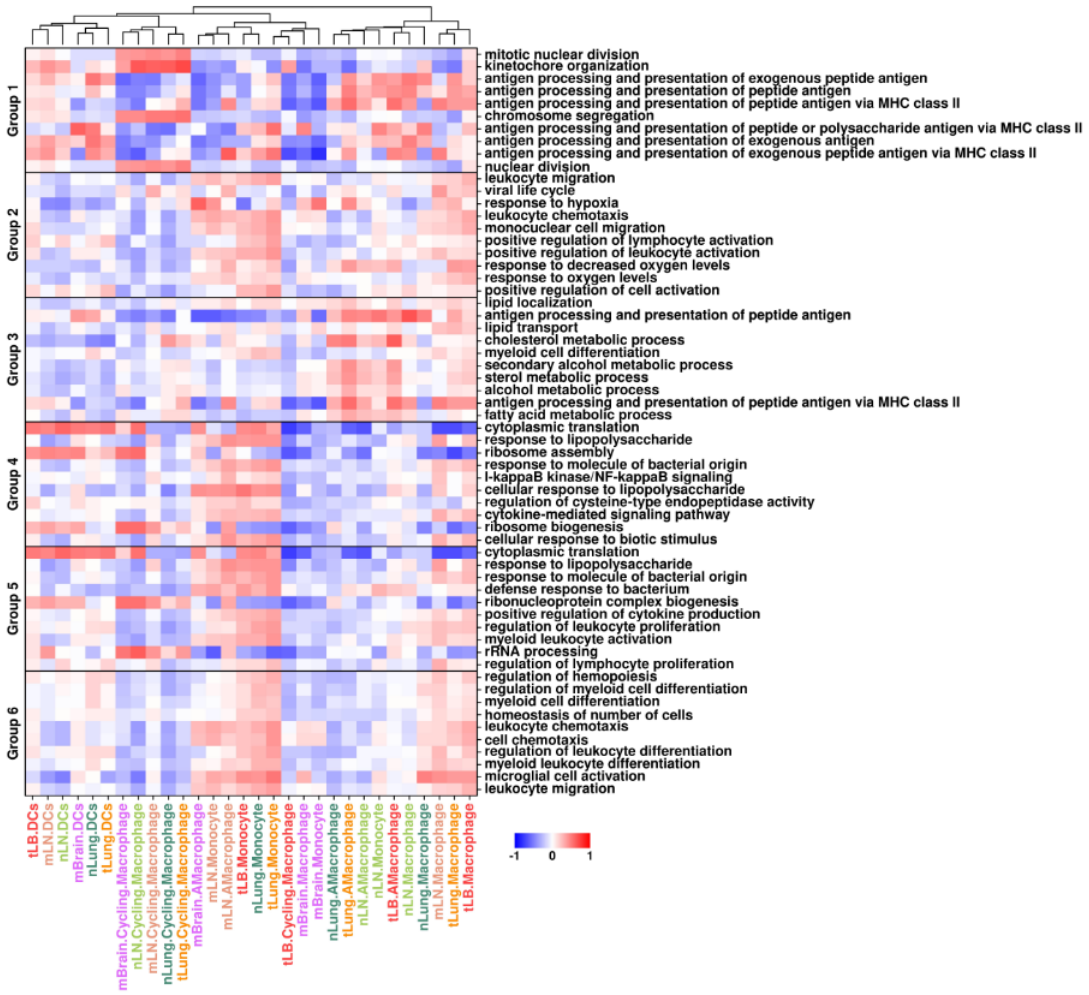
Reply: Thank you for this suggestion, and in the revised version, we have performed the analysis of abundance comparison among different myeloid components with undetermined cells excluded from the analysis. The component was also compared between different lung cancer stages (from normal/nLung to lung cancer/tLung, and to advanced stage of lung cancer tL/B), as shown in Figure2

3. In Figure 3, the authors examine the expression levels of selected M1 and M2 markers. However, the M1/M2 classification is increasingly viewed as an oversimplified model with inherent limitations, as reviewed by Dr. Mantovani and others in their 2022 review paper. The complexity of the tumor microenvironment (TME) complicates describing these cells using only the two extremes of the macrophage phenotypic spectrum. Moreover, the number of markers that the authors included is relatively small, which could lead to an unintentionally biased representation of these cells' phenotypes. scRNAseq data enables us to characterize these cells by their predicted functions, such as angiogenesis-associated and phagocytosis-associated populations. Therefore, I recommend the authors employ a more comprehensive gene signature that focus on the functions of these cells (for example, the signatures reported by Cheng et al., Cell, 2021), and present a heatmap of these myeloid subsets and tissue origins instead of inspecting each gene separately with box plots.

Reply: Thank you for this suggestion, and in this new version, we have involved more genes from Cheng et al. and compared the expression levels of these genes among different subgroups from all sample types, as shown in Figure 2B.

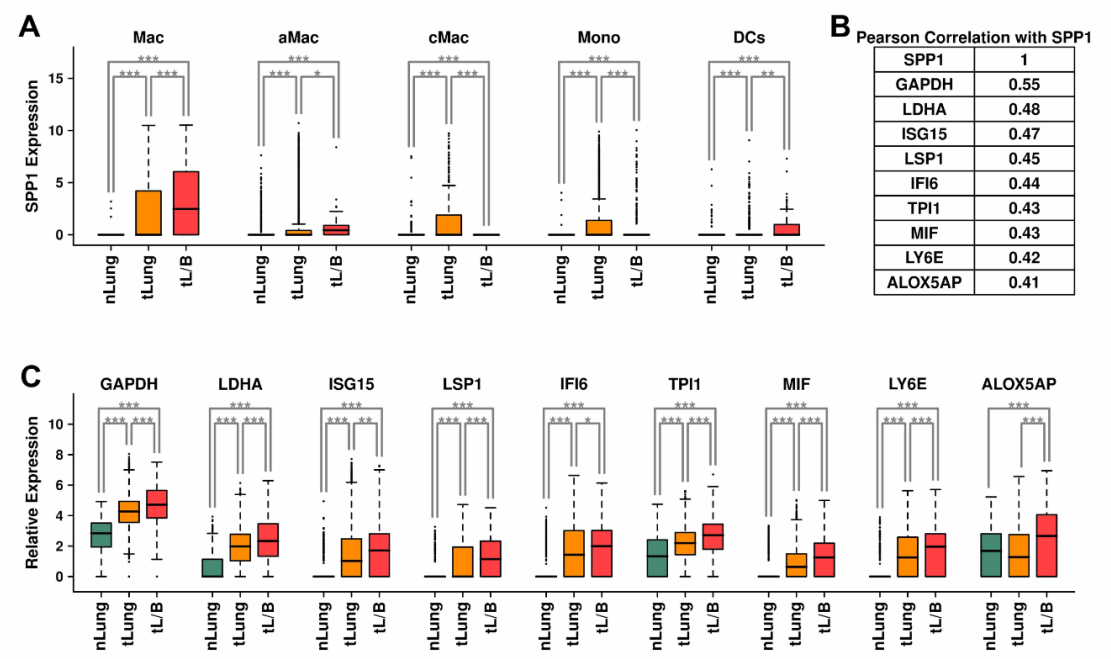
4. In Figure 5, the authors present a GO biological processes enrichment analysis with their pseudotime analysis. The authors describe their results in a way that equates state 1 cells with Monos, state 2 cells with aMacs, and state 3 cells with Macs (Line 274-286). However, this is an overly simplified statement, as for example, aMacs appear not only in state 2 but also in states 1 and 3 depending on the tissue origin (as the authors showed in Figure 4E). Therefore, please revise these sentences by mentioning only the pseudotime state and not a specific subset of cells. It would also be interesting to run the same enrichment analysis with cells from different tissue origins due to the different evolutionary patterns as shown in Figure 4E. This will complement the gene expression analysis in the original paper. Please also include implications of the findings and the authors' interpretation in the main text.

Reply: Thank you for this suggestion, and in this revised version, we have got rid of subgroups information. Meanwhile, we have involved the comparisons of these GO terms among different subgroups from all origins using GSVA (gene set variation analysis), and the results are illustrated in Figure 6.



5. In their 2021 Cell paper, Cheng et al. constructed a pan-cancer single-cell transcriptional atlas, which included lung cancer data, of tumor-infiltrating myeloid cells. They reported relatively consistent dichotomous functional phenotypes of these macrophages (SPP1+ TAMs and C1QC+ TAMs) across different cancer types including NSCLC. I suggest the authors take a closer look at the TAMs to determine whether they share some of the gene signatures of these two functional phenotypes, or if they can further subcluster the cells to allow such differentiation.

Reply: We would like to thank review for this suggestion, and in this revised version, we have compared the expression levels of SPP1 gene in myeloid subgroups, and found that the expression of SPP1 is increased with LUAD progression in macrophages, as shown in Figure 8A. Moreover, through correlation analysis, we have screened 9 genes that correlate with the expression of SPP1 in macrophages (Figure 8B). We have involved the description of these information in Line 396-420.



Minor Issues:

1. The section (Line 158-184) describing Figure 1 is descriptive but lacks data interpretation. Please revise by adding interpretation/conclusion sentences that summarize the findings.

Reply: Thank you for your suggestion, and in this new version, we have involved summaries for Figure 1 findings. (LineXXX)

2. In Figure 3 and Figure 6, the authors used Student's t-test to calculate the statistical significance of each pairwise comparison. However, multiple t-tests without proper adjustment of the Type I error significantly undermines the strength of the conclusion. Therefore, analysis of variance (ANOVA) should be used here, coupled with a suitable multiple comparisons test if the authors wish to evaluate the significance in each pair.

Reply: Thank you for this suggestion, and in this revised version, we have reperformed the analysis using one-way anova and corrected with "bonferroni" method.

3. In the legend of Figure 5, the authors wrote "MHC, XXX". Please revise accordingly.

Reply: Sorry for mistake, and in this revised version, we have corrected this mistake.

4. The statement about the development of alveolar macrophage (Line 324-343) is incorrect, as IFN-γ and LPS are not generally considered the driving factors for the differentiation of monocyte-derived aMacs. Please revise accordingly.

Reply: We are sorry for this mistake, and in this revised version, we have deleted this sentence.

5. I recommend the authors discuss how their results, especially the LR interaction analysis, contribute to the findings of the original publication in the Discussion section.

Reply: In this revised version, we have involved the discussion of LR interactions in Line 445-457.

We would like to thank reviewer C for the time and effort in reviewing our work.

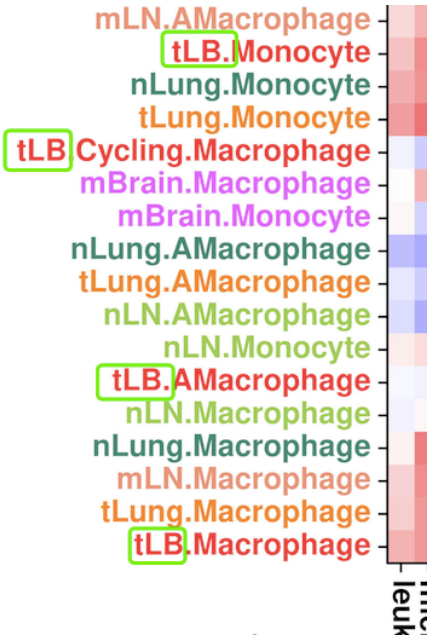
Reviewer D

1. ALL abbreviations used in each figure or figure description should be defined in a footnote below the corresponding table/figure. Please check all figures and provide correspondingly. For example,
nLung, tLung, mBrain, mLN, tL/B, nLN in Figure 2
GSVA in Figure 5

Reply: Thank you. ALL abbreviations have been defined in figure legends.

2. Figure 5

a. It is 'tL/B' in figure legend. Please check which one is correct and unify.



Reply: "tL/B" is correct, and we have attached a revised version of Figure 5

b. Please confirm if the submitted figure 5 match the description in the main text.

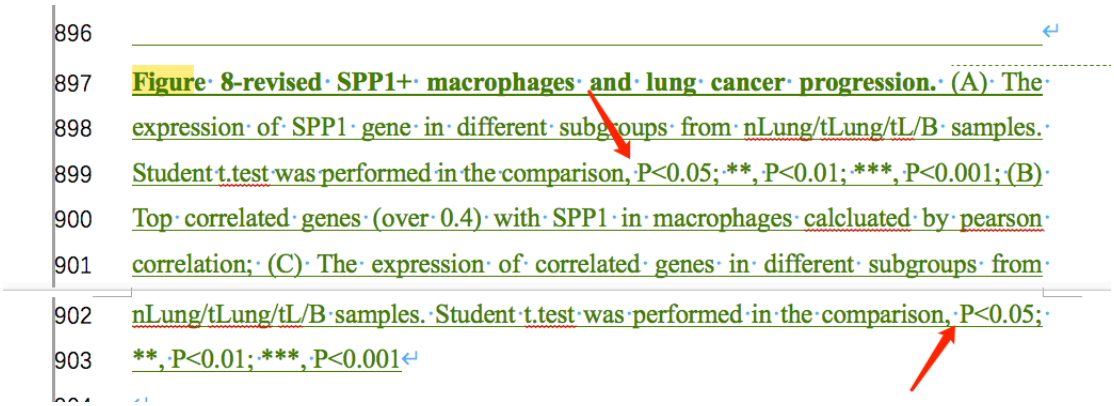
358 are shown in **Figure 5** (left panel). Higher *CCL18* expression was found in Macs from
359 tLung samples than in nLung samples and in cMacs from mLN and mBrain samples
360 compared to nLN samples. Higher *CXCL16* expression was found in Monos from tLung
361 samples than nLung samples and in cMacs from mLN and mBrain samples compared
362 to nLN samples. Higher *CXCL5* expression was found in aMacs from mLN samples
363 than nLN samples. Higher *FAM3B* expression was found in aMacs from mBrain

364 samples than nLung samples and in cMacs from mBrain samples compared to nLN
365 samples. Higher *GRN* expression was found in Macs from tLung samples than nLung
366 samples, in Monos from tLung samples compared to nLung samples, and in cMacs
367 from mLN and mBrain samples compared to nLN samples.
368 The expression levels of receptors corresponding to these ligands were also examined
369 (**Figure** 5 right panel). Regarding *CXCR6* (encoding C-X-C motif chemokine receptor
370 6, receptor of *CXCL16*), increased expression was found in Macs from tLung samples

Reply: These Figure 5 should be changed to “Figure 6” and we have changed these in the revised manuscript

3. Figure 8

a. “*” is missing. Please supplement.



Reply: Thank you. “*” has been added in figure 8 legend.

b. “nLung/tLung/tL/B” should be changed to “nLung, tLung, tL/B” to avoiding any confusion. Please check and revise.

Reply: Thank you. “nLung/tLung/tL/B” has been changed to “nLung, tLung, tL/B.