

## Peer Review File

Article information: <https://dx.doi.org/10.21037/tlcr-23-357>

### Reviewer A

Major comments:

1) Please review the document carefully, there are many misspellings and grammatical errors. (These are just a few examples:)

“ We successfully constructed a microbiota-geneset network to characterrize the”

“This finding provide a strongly corelation network for LUSC with lymph nodem metastasis and improve understanding of interaction between mocrobiota “

“To the best of our knowledge, this is the first study to identify the relationship between themicrobiome and lymph node metastasis of squamous cell lung carcinoma.”

Reply 1: First, we would like to thank the reviewer for constructive comments and suggestions, which were of great help in improving the quality of our manuscript. We have further revised our manuscript according to these comments.

Changes in the text:

We revised “microbiota-geneset network to characterrize” to “microbiota-geneset network to characterize” in line 80.

We revised “a strongly corelation network for LUSC with lymph nodem metastasis and improve understanding of interaction between mocrobiota” to “a strongly correlation network for LUSC with lymph node metastasis and improve understanding of interaction between microbiota” in line 83-84.

We revised “themicrobiome and lymph node metastasis of squamous cell lung carcinoma” to “the microbiome and lymph node metastasis of squamous cell lung carcinoma” in line 352.

2) “representing a higher risk of death with LN metastasis (hazard ratio [HR]: 1.16, 95% CI: 0.86 – 1.55; P = 0.331)(Figure 1C).”... please re-state this sentence... the p-value is not significant, so you cannot conclude that there is higher risk of death.

Reply 2: We agree with the reviewer. In addition, we performed the survival analysis for progression-free survival (PFS), and found a higher risk of progression with LN metastasis (hazard ratio [HR]: 1.55, 95% CI: 1.10-2.19; P = 0.011). Considering the PFS analysis better confirms poor prognosis with LN metastasis, we used the PFS (Figure 1B,C) to replace the results of OS in the supplement (Figure S1).

Changes in the text:

We have revised “representing a higher risk of death with LN metastasis (hazard ratio [HR]: 1.16, 95% CI: 0.86-1.55; P = 0.331)(Figure 1C).” to “representing a similar trend to the PFS” in line 243-244.

We have add the results of PFS survival analysis in line 239-241. We have add the results of OS survival analysis in Figure S1 and include figure legends. The original Figure S1, S2, and

S3 were changed to Figure S2, S3, and S4 in turn.

3) Figure 2 legend “PCoA shows a difference in  $\beta$ -diversity between LN+ and LN-.”... please change to “does not show a difference...”

Also, there are no significant comparison in figure 2, please remove this sentence “Levels of significance:  $*p < 0.05$  (Wilcoxon rank sum test),”

Reply 3: Thanks for the constructive comments.

Changes in the text:

We have changed “PCoA shows a difference in  $\beta$ -diversity between LN+ and LN-.” to “PCoA does not show a difference in  $\beta$ -diversity between LN+ and LN-.”, and we have deleted the sentence “Levels of significance:  $*p < 0.05$  (Wilcoxon rank sum test)” in Figure 2 legend.

4) Figure 4 and 5 “Asterisks are labeled according to the P values calculated”... please write out what one \*, two \*, three \* mean for p-value; this is not stated

Reply 4: Thanks for the constructive comments.

Changes in the text:

We have revised “Asterisks are labeled according to the P values calculated” to “Asterisks are labeled according to the P values calculated; \* mean for p-value 0.05, \*\* mean for p-value 0.01, \*\*\* mean for p-value 0.001” in the Figure 4 and 5 legends.

5) Figure 6- what is blue node mentioned in the legend. There is no blue node in the figure. How does listeria fit into this figure?

(Overall the figure legends are poorly done with little description of the data, please include better description which summarize the data)

Reply 5: Thanks for the comments. We include a better figure legends which summarize the Figure 6. We sorry to mistakenly marked *listeria* as a marker, which should be changed to *Bacteroides* and *Proteus*, the bacterial markers of LN+ group, which are represented by the red nodes, not blue node.

Changes in the text:

We have included a better Figure 6 legends which summarize the data.

6) RNA-seq and microbial data from LUSC database which were used in this study should be shared with the readers. Either identify which samples were used and share in a file; or share the sequencing that were used and upload to a share drive.

Reply 6: Thanks for the constructive comments. We identify which samples were used and share clinical data in a file.

Changes in the text:

We have shared patients’ clinical data in Table S3, and added it in line 433.

7) Author has to state in the discussion that they are not able to identify bacterial contaminate

through this database.

Reply 7: Thanks for the constructive comments.

Changes in the text:

The sentence “Finally, low levels of the bacterial contaminate are not be able to exclude through this database perfectly, the precise experiments are needed to verify microbe-host interactions.” have been added in the revised manuscript (line 437-439).

## **Reviewer B**

1)The authors provide the various results using several existing pipelines. Many results are obtained, but it is difficult to summarize in one conclusion. For example, in figure4, the authors investigated not the difference with or without LN metastasis, but common co-mutation in lung squamous cell cancer. Finally, the result derives from a result of Figure 6 after all.

Reply 1: We thank the reviewer for the constructive comments and suggestions which are of great help to improve the quality of our manuscript. We have further revised our manuscript according to these comments.

Firstly, we confirmed that lymph node metastasis was associated with lung squamous cell cancer poor prognosis. We then investigated the difference with or without LN metastasis in multi level, including diver mutants, exposure signatures, genes expression, and microbiome. The up-regulated genesets and microbial markers were identified as prominent features in the lymph node metastasis. Finally, we worked on a closer correlation network between the microbial markers and up-regulated genesets in the LN+ group (Figure 6). It is important to note that the sample in LN+ (n=165) was fewer than the LN- group (n=296), therefore the correlation was not due to large statistics. The evidence for the microbial profile was limited to abundance, the role of microbial metabolism in tumor progression can serve as a potential strategy in the future.

Changes in the text:

The sentence “The up-regulated genesets and microbial markers were identified as significant differences between the LN- and LN+” have been added in the revised manuscript (line 328-329).

2)The clinical data such as the degree of the lymph node dissection, surgical procedures, or chemotherapy and/or radiation, which is really affected by lung cancer prognosis, are not introduced in this manuscript.

Reply 2: Thanks for the constructive comments. We have updated the clinical data including the radiation, neoadjuvant, and new tumor event in Table 1. We found no significant differences in these three factors between metastatic and non-metastatic patients. Due to the lack of degree of the lymph node dissection and surgical procedures, we could not added these to the baseline. The new table also support our conclusion.

Changes in the text:

We have updated the clinical data including the radiation, neoadjuvant, and new tumor event in Table 1.

3) In addition, it is impossible to accept for a result that lymph node metastases are not associated with lung cancer prognosis against TNM classification. The authors should arrange the data.

Reply 3: Thanks for the comment and the honesty from reviewer. By updating and reanalyzing the prognostic data, we confirmed that patients with lymph node metastases were subjected to progression-free survival (PFS). The overall survival (OS) did not show significant differences, however, it also displayed a similar trend with PFS. Due to the absence of surgical procedures, chemotherapy or immunotherapy, as well as the deletion in patients with long-term follow-up, which cannot correct for these factors that clearly affect survival time. Instead, we included overall survival information in supplementary materials, which did not affect the conclusion on the relationship between microbiota and lymph nodes metastasis. In recent years, patients' survival time has benefited from immunotherapy. The results showed that the lymph node metastasis is more likely to lead to tumor progression, which also suggests that the process of lymph node metastasis is a good entry point to explore the etiology.

Changes in the text:

We have added the results of PFS survival analysis in line 240-241. We have added the results of OS survival analysis in Figure S1 and line 241-244. The original Figure S1, S2, and S3 were changed to Figure S2, S3, and S4 in turn.