

Peer Review File

Article information: <https://dx.doi.org/10.21037/tlcr-23-231>

Reviewer A

Although there is good data in this article, the following points are worth mentioning:

Title and Abstract

1. Although the title of the manuscript is about “Early lung cancer patients”, but there is no discussion about this stage of cancer in the manuscript.

Response: Thank you for pointing this out and we should address this issue in the discussion. In the Discussion section (see lines 344-347), we added the reasons why we choose to study the microbiome in early lung cancer patients and the implications of microbiome in early lung cancer patients.

2. The fourth sample is bronchus, but in the title of the manuscript, all samples are considered as lung tissue, which I am not sure is correct (though both are parts of respiratory tree).

Response: Technically, the bronchus should not be a part of lung tissue. Sorry for this confusing expression. We have changed the description in the title to “lower respiratory tract” (from “lung tissues”) (see line 2). To keep consistency, we also changed the expression throughout the whole article to avoid misunderstanding.

3. Abbreviations LL, LP, LB do not look right and should be corrected based on the corresponding words.

Response: We have changed the abbreviations throughout the article as follows: lung tumor tissues (change LC to TT), para-carcinoma tumor tissues (change LP to PT), distal normal lung tissues (change LL to DN), and bronchial tissues (change LB to BT).

4. In the Methods section, there is no explanation about the method of data analysis and prediction of metabolites.

Response: We added how the PICRUSt works to predict the metabolites in the methods (lines 210-213).

5. For keywords, it is recommended that “16s rRNA sequencing” or “metagenomic analysis” is used instead of “Alpha diversity”.

Response: Thank you for your suggestion. The “16S rRNA sequencing” should be more appropriate to be a keyword. We have changed it in the manuscript (see line 78-79).

Introduction

1. As a general note, regarding to metagenomic analyses, it should be better to use “microbiome” instead of “microbiota” in the manuscript.

Response: Thank you, and we have revised the expression throughout the manuscript according to your suggestion.

2. Line 134: “Two bacterial biomarkers, Capnocytophaga and Veillonella, have shown good efficacy in predicting squamous cell carcinoma (SCC) and adenocarcinoma (AC), which can facilitate lung cancer screening good efficacy in predicting squamous cell carcinoma (SCC) and adenocarcinoma (AC), which can facilitate lung cancer screening (16)”. This sentence does not relate to reference 16. It is related to the article titled: the microbiome and lung cancer (Maddi et al, 2018).

Response: Thank you for pointing this out, and we have corrected the reference as you suggested (see line 141).

3. Line 137 the word “Veillonellaparvula” should be written separately as “Veillonella parvula”.

Response: Thank you for pointing this out, and we have corrected the word (line 139).

Methods

6. There are no references in this section.

Response: We have added some references for the methods used in this article (see lines 199, 201, 209).

7. The most important problem of the article is the small sample size, and based on this small number of samples, the obtained results are not very reliable. For this reason, according to the level of articles in prestigious journal, this manuscript is rejected in my opinion.

Response: The small sample size is indeed the limitation of this article and the result of this article should be verified in larger sample set.

However, the sampling of four sites is much more difficult, and to our knowledge no previous study has reported such comprehensive analysis of the microbiome in the lower respiratory tract.

8. The level of statistical significance is not mentioned anywhere in the method.

Response: Sorry for this neglect, and we added it in the method of statistical analysis (see lines 220).

9. Regarding the characteristics of the patients, such as smoking, gender, and stage of cancer, since there are only 1 or 2 cases as opposed to 14 or 15 cases, it cannot be said that the statistical analysis is not significant, but it must be said that it cannot be analyzed and compared at all.

Response: Thank you for pointing this out. We made revisions according to your suggest to make the description of our results more stringent (line 244-246).

10. In line 237 (the method section), the order and class levels are mentioned, but there are no results for these levels in the result section.

Response: Sorry for this confusing expression. We deleted the “class, order, family” in the sentence (see line 259).

11. For AMOVA and HOMOVA in lines 190 and 191, the full name should be written.

Response: Sorry for this neglect, and we added the full names in the corresponding site (see lines 204-205).

Results

1. The resolution and quality of image 1 is not good and needs to be improved.

Response: Sorry, we have changed image 1 in the manuscript by a larger copy version (see line 406). If it still can not meet the requirement, please use the pictures that upload separately.

2. The figures captions do not have explanations and should be added.

Response: Thank you for pointing this out. We have added the necessary explanations in every figure caption.

3. The method section does not have references.

Response: Sorry for this neglect. We have added some references for the methods used in this article (see lines 199, 201, 209).

4. Line 193, add the version and reference of PICRUSt.

Response: Sorry for this neglect. We have added the version and reference at the corresponding site (see line 209).

5. Line 186, add the version and reference of QIIME2.

Response: Sorry for this neglect. We have added the version and reference at the corresponding site (see line 201).

6. Line 184, add reference of Silva.

Response: Sorry for this neglect. We have added the reference at the corresponding site (see line 199).

7. Line 240: “common among the four tissue group”, should be: “common among the four tissue groups”. It is advised to check the whole text for grammar.

Response: Thank you for pointing this out. We are sorry for the errors in grammar. We have corrected it (see line 262) and checked the whole text again for grammar and make some corrections.

Discussion

1. Overall, the discussion is poorly written. Most of the discussion parts (such as lines 300-314) are discussed about topics that are not directly related to the results of this study and explain about other unrelated topics, including therapeutic intervention with antibiotics or probiotics. The discussion should be focused on the results of the present study and comparison with the results of other similar articles.

Response: Thank you for your advice. We have added two paragraph that are more related to the results of the present study (lines 322-347), and deleted the parts that you mentioned (To keep the accordance of the reference inserted in the articles and the paragraph list, we accepted the deletion of these two paragraphs so we cannot see them at this step).

2. In the introduction and discussion, a little explanation has been given about the importance of para-carcinoma tissue microbiome content and its comparison with normal tissue, and thus more explanation and discussion are needed.

Response: Thank you for pointing this out. As we include both the para-carcinoma tissue and distant normal tissue, the importance of para-carcinoma tissue should be address in the introduction or discussion. We have added a sentence to explain it in the introduction part as follows (see lines 150-153):

“We included the para-cancerous tissues because they may exhibit a transitional state between cancer and normal tissues [20], including in the microbial community [21], potentially reflecting the underlying processes of carcinogenesis.”

3. There is no discussion about the obtained metabolites results in the discussion section and should be added.

Response: Thank you for pointing this out. We have added the discussion about the metabolic functional analysis in the discussion (see lines 331-340).

4. It should be explained more about Rubellimicrobium and Fictibacillus obtained in the study and its possible role in lung cancer in the line 280.

Response: Thank you for your suggestion. We have added more explanation about this two taxa in the discussion (see lines 304-307).

4. It is said at the beginning of the discussion section: “This study investigated the microbiota composition of lung tissues from different sites of lung cancer patients and analyzed their associations with clinical characteristics. From the taxonomic profiles, we demonstrated that different species were enriched at specific lung sites of lung cancer patients.”Firstly, in most parts, no statistically significant difference has been found in the results. Second, what do you mean by different species? Did you mean the genus level?

Response: Sorry for the confusing expression. Indeed, no statistically significant difference has been found in diversity and functional profiles of the microbiome between different tissue sites. However, we found the enrichment or shrinkage of specific taxa in cancerous tissue or other tissue sites. To avoid misunderstanding, we revised the sentence as follows (see lines 282-287):

“This study investigated the microbiome composition of lung tissues from different sites of lung cancer patients and analyzed their associations with clinical characteristics. From the taxonomic profiles, we demonstrated that specific species at both the phylum and genus levels were enriched at particular lung sites of lung cancer patients.”

Reviewer B

It is a very interesting manuscript please add recent review at the introduction Int J Mol Sci. 2021 Sep 27;22(19):10429.

Response: Thank you for your suggestion. We have added this recent review in the first paragraph of introduction (see line 108).

Reviewer C

The paper titled “A comparison of the microbiota composition in lung tissues at different sites in early lung cancer patients” is interesting. A non-significant result was obtained from the microbiota diversity comparison between different tissues. However, we demonstrated that lung tumors were enriched with specific bacterial species, which might contribute to tumorigenesis. Moreover, we found an inverse relationship between BMI and alpha diversity in these tissues, providing a new clue for deciphering the mechanisms of lung carcinogenesis. However, there are several minor issues that if addressed would significantly improve the manuscript.

1) The abstract is not sufficient and needs further modification. The research background did not indicate the clinical needs of the research focus.

Response: Thank you for pointing this out. We added a sentence in the background to indicate the clinical needs of the research focus (see lines 54-56).

2) What are the different roles of the lung microbiota in promoting carcinogenesis and maintaining homeostasis under different conditions? It is recommended to include relevant content in the discussion.

Response: Thank you for your suggestion. We added relevant content to discuss the different role of lung microbiota under different conditions (lines 332-344).

3) Can the lung microbiota predict clinical outcomes in critically ill patients? It is recommended to add relevant content.

Response: A study reported the predictive role of specific microbiota for RFS and DFS, and we added this reference in the discussion (see lines 339-342).

4) What is the correlation between microbial composition and functional potential and the status and severity of lung cancer? It is recommended to add relevant contents.

Response: The functional potential were inferred by PICRUSt with the OTU abundance which is also the upstream origin of the taxonomy analyses. The work principle of PICRUSt was added in the method (see lines 210-213). In addition, to discuss the correlation between microbial composition and stage of the lung cancer, we added relevant content to discuss the different role of lung microbiota under different conditions in the discussion (see lines 332-344).

5) The number of patient samples in this study is too small, and a large sample study should be added for verification.

Response: The small sample size is indeed the limitation of this article and the result of this article should be verified in larger sample set.

However, the sampling of four sites is much more difficult, and to our knowledge no previous study has reported such comprehensive analysis of the microbiome in the lower respiratory tract. Therefore, there is no public database that can be used to validate the current study.

6) The introduction part of this paper is not comprehensive enough, and the similar papers have not been cited, such as “Microbiota dysbiosis in lung cancer: evidence of association and potential mechanisms, Transl Lung Cancer Res, PMID: 32953527”. It is recommended to quote the articles.

Response: Thank you for your suggestion. The reference was added in the introduction (see line 131).

7) This study is based on bioinformatics analysis. It is recommended to increase in vivo and in vitro experimental studies, which may be more meaningful.

Response: Indeed, the current study is based on bioinformatics analysis of the microbiome obtained from four different tissue sites. While it would be more meaningful to complement these findings with in vivo and in vitro experimental studies, the scope of such experiments is currently beyond the capabilities of our research team. It is important to note that sampling from four sites presents greater challenges compared to previous studies, and to the best of our knowledge, no prior study has reported such a comprehensive analysis of the microbiome in the lower respiratory tract. Despite these limitations, we believe that the results obtained from this study still hold certain reference value for future experimental or clinical investigations.

Reviewer D

1. There're **two references lists** in your paper, one contains 43 references while 52 are included in the second list.

Please also check if the first author's name match to its reference cited.

167 and PI3K signaling pathways [15]. According to Lee **et al.**, two phyla (*Firmicutes* and TM7) and
168 two genera (*Veillonella* and *Megasphaera*) are abundant in bronchoalveolar lavage fluid (BALF) in
169 lung cancer patients [16]. Two bacterial biomarkers, *Campytophaga* and *Veillonella*, have shown

Response: The second list is the right one. It is confusing due to the word processing program we used. It can't delete the references when the text (in which the references were inserted) were deleted under the revision mode. After I accepted the deletion of the two paragraphs in the discussion section, the total reference number of the second list turned to 46.

We have checked the reference you mentioned, and the first author's name can match to its reference cited.

2. You've mentioned “studies”, while only one reference was cited in the below sentences. Please check.

163 hotly investigated [13]. **Studies** have documented the existence of *Staphylococcus spp.*, *Bacillus*
164 *spp.*, *Haemophilus influenza*, and *Candida albicans* in lung cancer patients [14]. The microbial

451 Previous studies have shown that microbiomes play specific roles under different conditions.
 452 In the case of the healthy lung, it has been confirmed that diverse communities of bacteria play an
 453 important role in maintaining an immune-tolerant environment and preserving lung homeostasis
 454 [44]. However, exposure to certain taxa such as *Streptococcus* and *Veillonella* has been shown to

Response: The first one was corrected into “Study” as there is only one reference mentioned; the second one was added by three references for the sentence.

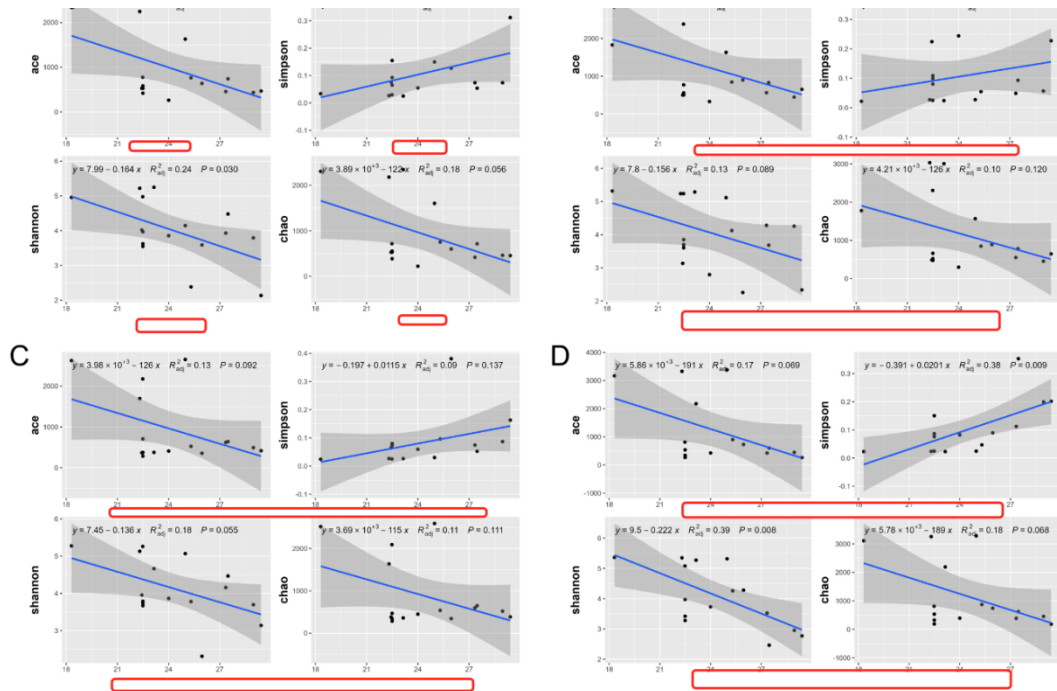
3. Figure 1-4: Please check all the figure legends, the full names of those abbreviations (TT, PT, DN, BT) are not the same with those in Abstract. You should unify them in the whole text.

680 PT (B), DN (C), and BT (D). BMI, body mass index; TT, lung tumor; PT, adjacent normal tissue;
 681 DN, normal distal lung tissue; BT, bronchus tissue.

62 **Methods:** A total of 16 patients with non-small lung cancer (NSCLC) were recruited for this study.
 63 Samples were obtained from four sites, including lung tumor tissues (TT) para-tumor tissues (PT)
 64 distal normal lung tissues (DN) and bronchial tissues (BT). The DNA was isolated from the tissues.

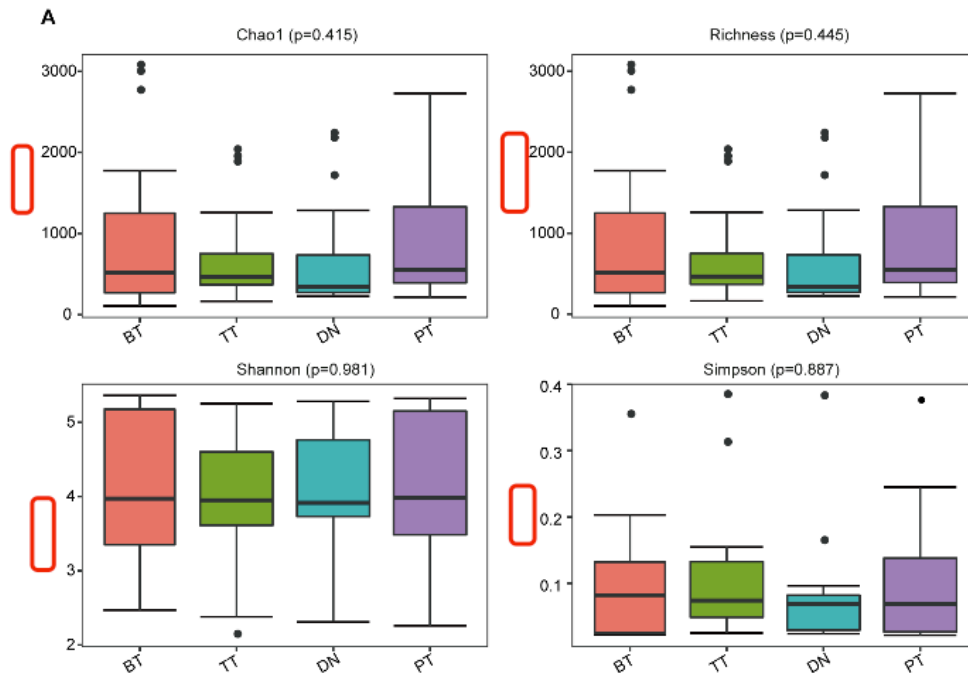
Response: We unified the full names of four site tissues in the whole text.

3. Figure 1: The descriptions of X-axis are missing, please supplement them and resend us updated figure.



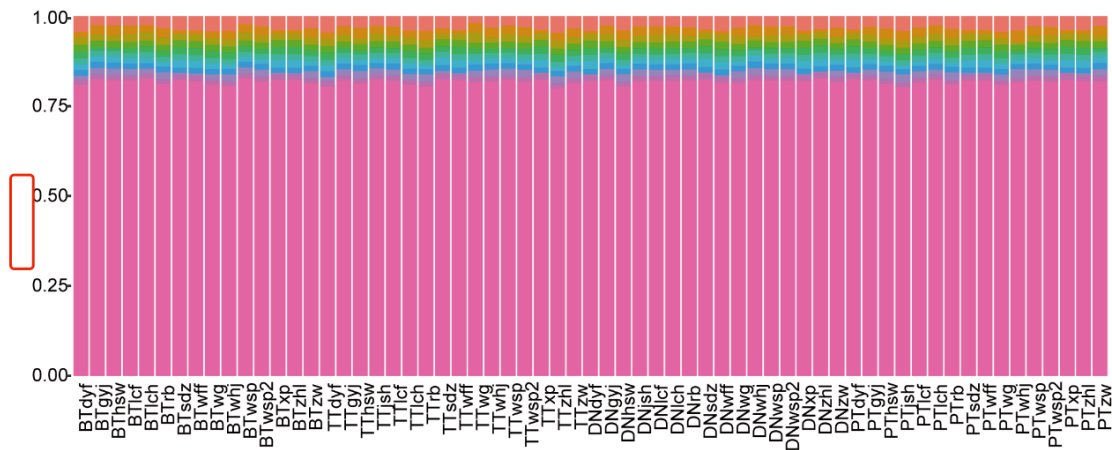
Response: The X-axis all represent BMI in this figure. We have added the description.

4. Figure 2: The descriptions of Y-axis are missing, please supplement them and resend us updated figure.



Response: Each Y-axis represents the corresponding index score of alpha diversity metrics. We supplemented the descriptions of Y-axis in this figure. If this can not meet your requirement, we are pleased to revised it further.

5. Figure 4: The descriptions of Y-axis are missing, please supplement them and resend us updated figure.



Response: The Y-axis represents the relative abundance of the predicted function. We supplemented the description of Y-axis in this figure.