Professor Rui Henrique

Editor

*Frontiers in Genetics*      July 9th, 2018

Dear Professor Henrique,

We appreciate for your letter about the information of our manuscript entitled “Identification of hyper-methylated tumor suppressor genes-based diagnostic panel for esophageal squamous cell carcinoma (ESCC) in a Chinese Han population”. The revised manuscript has benefited greatly from consideration and incorporation of the constructive and insightful comments from you and the reviewers. As for the reviewers’ comments, a point-by-point description of our responses is as follows, where the reviewers’ comments are in italics and our responses in Times New Roman type.

Enclosed is the revised version of the article with the title “Identification of hyper-methylated tumor suppressor genes-based diagnostic panel for esophageal squamous cell carcinoma (ESCC) in a Chinese Han population”.

Thank you so much for handling with our manuscript!

Sincerely,

Minghua Wang, Ph.D.

Professor,

Department of Biochemistry and Molecular Biology

Soochow University

**Reviewer 1:**

We appreciate your thorough review and thoughtful comments for our submitted manuscript. Your suggestions and critiques have significantly improved our manuscript and we hope we have addressed all concerns adequately.

*The authors used four public DNA methylation microarray datasets and the literature searching to identify novel aberrantly methylated Candidate Tumor Suppressor Genes in esophageal squamous cell carcinoma (ESCC). In a second phase they validated by bisulfite sequencing selected TGS (ADHFE1, EOMES, SALL1 and TFPI2) in an independent cohort of 94 pairs of ESCC and normal tissues from a Chinese Han population. All four genes were significantly hyper-methylated in ESCC cases and they have shown a good overall performance (Sensitivity = 66%, Specificity = 87%, AUC = 0.81) that improved in the non-alcohol use subgroup (Sensitivity = 66%, Specificity = 96%, AUC = 0.84). Based on these results the authors proposed the methylation panels of hypermethylated genes ADHFE1, EOMES, SALL1 and TFPI2 as an effective assay for ESCC diagnosis.*

*Question 1: I could not find the Suppl table 1 in the word file named “Data Sheet 2”. There is only the title. Please provide the missing table.*

Answer 1: We thank review for the notice to this confusion. Due to the large size of supplementary table 1, we have made a separate file for it titled as “Data Sheet 1.csv” in the “Supplementary Material” section. To make it clearer, we have changed the file name of “Data Sheet 1.csv” to supplementary table 1 and added the legend for it.

*Question 2: Pag 4 Lane 23 please correct “methylaiton” with “methylation”.*

**Answer 2**: Thanks for your correction. We have corrected the spelling mistakes in Pag 4 Lane 24, and we also checked and corrected the other spelling or syntax mistakes carefully in the whole manuscript. The corrections were all highlighted in red.

*Question 3: Please clarify the literature search criteria used in the text of the materials and methods.*

Answer 3: We thank the reviewer for the suggestion. We extracted all the gene symbols occurred in the abstracts together with the keyword ‘tumor suppressor gene’ by our own Perl script and then check the gene list manually. We add this description in the manuscript between line 9 and 11 in the Page 4 as the following:

Candidate tumor suppressor genes were collected through the keyword matching (“tumor suppressor gene”) with custom script among 91,225 abstract downloaded from Pubmed database and manually re-checked (listed in Supplementary Table 1).

*Question 4: Pag 5 Lane 9 please change “Patients, samples and DNA” to “Patients and samples” since in this paragraph there is no mention to DNA processing.*

**Answer 4**: We thank the reviewer for the suggestion. In the revised manuscript, we have changed the “Patients, samples and DNA” to “Patients and samples” in Page 5, Line 13.

*Question 5: Fig 1 please improve the graphics, space between the words in the first lane “Candidate Tumor Suppressor Gene (65) screening”.*

**Answer 5**: We thank the reviewer for the suggestion. We have adjusted the graphics and space between the words in the first lane of Figure 1 in the revised manuscript.

*Question 6: Please mention the fig 1 in the materials and methods or in the result paragraph.*

*.*

**Answer 6**: We thank the reviewer for the suggestion. We have added the sentences in Page 7, Line 15-16 as following “The detailed biomarker identification procedure was shown in Figure 1.”

*Question 7: Please mention the Suppl fig 1 and Suppl fig 2 in the materials and methods or in the results paragraphs.*

**Answer 7**: Thanks for your suggestion. In the revised manuscript, we have added the following sentences in page 4, Line 16-18 as following: Due to the similarities which were shown through PCA analysis between adjacent control tissues from ESCC and EAC, the 13 normal tissues of EAC were included in our combined dataset as controls equally (Supplementary Figure 1).

Meanwhile, we also mentioned the Supplementary Figure 2 in Page 8, Line 15-16 as following: PCA analysis revealed that a significant distinction between ESCC samples and control samples (Supplementary Figure 2).

*Question 8: Pag 5 lane 26 please correct “AIIperp DNA/RNA Mini Kit.*

**Answer 8**: Thanks for your correction. In the revised manuscript, we have corrected the misspelling to “AllPrep DNA/RNA Mini Kit” in Page 6, Line 1.

*Question 9: Fig. 2B please improve the graphic using for “LINE 1” the same font used for the title of the other panels C, D, E and F.*

**Answer 9**: Thanks for your suggestion. In the revised figure, we have adjusted the font of “*LINE-1*” to the same as the title of other panels.

*Question 10: Legend figure 3, please clarify the statistical test used and the p value related to each gene.*

**Answer 10**: Thanks for your suggestion. In the revised Figure 3, we have added the p value for each gene. Meanwhile, we have added the explanation for the statistical test used in the legend of figure 3 as following: P value is calculated through the Wilcoxon rank-sum test and the Benjamini-Hochberg procedure was applied for multiple test correction.

*Question 11: Pag 11 lane 6 fig 3, please clarify the statistical test to which the plots of fig 3 refers to.*

Answer 11: We thank the reviewer for the suggestion. We modified our description for the Figure 3 between Line 22-26 in page 8 as the following so that the readers can understand it clearly.

“To better characterize the methylation status of the four genomic regions as well as the four candidate genes, we averaged the methylation status of all the CpG sites in each genomic region and conducted the DMR analysis with the same approach. We found all these 4 genes are significantly differentially methylated between ESCC and normal samples (Figure 3)”

*Question 12: Pag 2 lane 14 please clarify the statistical test used.*

**Answer 12**: We thank the reviewer for the suggestion. We have revised the abstract and clarified the statistical test used in our analysis.

*Question 13: Pag 4 lane 7 please change “Patients and Methods” into “Materials and Methods”.*

**Answer 13**: We thank the reviewer for the suggestion. We have changed the “Patients and Methods” into “Materials and Methods” in the revised manuscript.

*Question 14: Supplem table 4: please use the same font size for the whole title.*

**Answer 14**: We thank the reviewer for the correction. In the revised version of supplementary table 4, the font size of the whole title was same.

*Question 15: Please provide some explanation related to the content of the “data sheet 1”. It does represent the Suppl table 1?*

**Answer 15**: Thanks for your suggestion. The data sheet 1 was actually the supplementary table 1. Due to the fact that the size of supp table 1 is relatively large, so we made a separate file for it. In the revised version, we have added the explanation in this table for a clearer understanding.

*Question 16: Supplem table 7 please correct “Smoked” into “Smokers” and “non-smoked” into “non-smokers”.*

**Answer 16**: Thanks for your correction. In the revised version of supplementary table 7, we have corrected the “Smoked” into “Smokers” and “Non-smoked” into “Non-smokers”.

**Reviewer 2:**

We appreciate your thorough review and thoughtful comments for our submitted manuscript. Your suggestions and critiques have significantly improved our manuscript and we hope we have addressed all concerns adequately.

*This study describes methylation profiling of esophageal cancer to determine a panel of candidate tumor suppressor genes that may provide a methylation-based diagnostic approach in esophageal squamous cell cancer (ESCC). The candidate genes were selected using the TCGA and GEO datasets and methylation of the candidate genes was analyzed by NGS using bisulfite-converted DNA from the patients in the validation set. The authors suggest that an assay consisting of a panel of four genes may be useful for ESSC diagnosis.*

***Question 1****:* *A proper discussion is almost non-existent. A comprehensive, detailed discussion with evaluation and interpretation of the results in view of the literature data is warranted. The first paragraph repeats previous statements and can be almost completely omitted.*

**Answer 1**: Thanks for your suggestion. In the revised manuscript, we have deleted the repeatments about the previous statements in the first paragraph of the discussion. Meanwhile, we also rewrote the whole discussion to give a more comprehensive and detailed evaluation of our results.

***Question 2****: Further, analysis of the correlation between methylation and gene expression (5-Aza treatment) is not mentioned anywhere in the results or in the methods. These data belong to the results section and the experimental details should be added to the Methods.*

Answer 2: We thank the reviewer for the correction. In the revised manuscript, we have adjusted the contents concerning the 5-Aza treatment experiments to the Results part and added the experimental details to the Methods in Page 6, Line 16-28. Meanwhile, we changed the Supplementary Figure 5 to Figure 4 to show the reverse relationship between DNA methylation and gene expression so that the readers can have a clear understanding to this result.

***Question 3****: Likewise, the sensitivity of EOMES and specificity of ADHFE1 (as mentioned in the conclusion) are not mentioned in the results.*

**Answer 3**: Thanks for your correction. In the revised version of manuscript, we have added the sentence in Page 9, Line 1-2 as following: Of these four candidates, *EOMES* showed the highest sensitivity (0.69) and AUC (0.78), while the *ADHFE1* showed the best specificity (0.94).

***Question 4:*** *The reference list is too extensive, contains studies not directly related to the particular study but does not mention multiple publications which are more closely associated with the reported data and therefore should also be evaluated in the discussion. A representative (non-complete) list is indicated: [ Xi and Zhang Pathol. Res Pract. 2017; Tsunoda et al. Oncol. Rep. 2009; Chettouh et al. Gut 2017; Corrie et al. Mol. Biosyst. 2009; Liu et al. Anticancer Drugs 2016].*

**Answer 4**: We thank the reviewer for the great suggestions. In the revised manuscript, we have deleted the references which were not so closely associated with the reported data and added more references including the papers you mentioned, along with the others which are colsely associated with the reported data. Meanwhile, we revised the discussion section thoroughly to give a better evaluation on the previous studies and our results of these four candidate biomarkers.

***Question 5:*** *An issue to consider is the specificity of the suggested panel for ESSC. In contrast to Asia and Far East a significant proportion of the esophageal cancers in the rest of the world are adenocarcinomas. Inclusion of a group of adenocarcinomas and data on the performance of the panel in this cohort would have been useful in indicating the validity of the suggested panel. Otherwise, this drawback should be indicated.*

**Answer 5**: Thanks for your suggestion. We are also trying to see if our panel could also be applied to the esophageal adenocarcinomas (EACs) as well. However, as you mentioned, more than 90% of the esophageal cancers in China are ESCCs instead of EACs. Unfortunately, we could not collect enough EAC samples for our evaluation right now. Therefore, we have added this drawback in the discussion part of our revised manuscript in Page 12, Line 20-22 as following: In addition, the diagnostic ability of our panel was only validated in ESCC samples but not in EAC samples due to our limited samples, and further studies based on EAC samples should be conducted.

***Question 6:*** *The CpG sites selection criteria are not sufficiently explained in the Methods. A more detailed description would be appropriate. For example, observing Table1 I could not figure out why the RUNX gene was not selected. What were the “technical limitations”, PCR conditions, etc.?*

**Answer 6**: Thanks for your suggestion. The technical limitations were mainly the PCR conditions. When we constructed a multiplex PCR system, we found that the addition of these six candidate primers (*ADHFE1, EOMES, RUNX1, SALL1, TFPI2, WT1*) could not generate enough high quality reads for *RUNX1* and *WT1*. It may be caused by the GC percent, PolyT and the number of SNPs in the primers, or caused by the mutual interferences of these primers. Therefore, we only selected part of these six genes for further analysis.

In the revised version of manuscript, we have added the detailed description of CpG sites selection criteria in Page 5, Line 6-12 as following: After that, we designed the primers for these six genes separately and then applied for multiplex PCR system. Due to the GC percent, PolyT and the number of SNPs in the primers of our targeted regions, we only obtained the multiplex PCR system consisting of the four genes including *ADHFE1*, *EOMES*, *SALL1*, *TFPI2* but could not generate enough high quality reads for *RUNX1* and *WT1*. Therefore, these two genes were then discarded for further analysis. Finally, we validated the methylation of these four candidate genes with 94 pairs of Chinese ESCC and control samples (Table 1).

***Question 7****: What was the predictive performances of the triple gene combinations?*

Answer 7: Thanks for your suggestion. We have assessed the predictive performances of the triple gene combinations and were shown below. It is shown that the combination of *ADHFE1*, *SALL1* and *TFPI2* showed the highest sensitivity, while the combination of *ADHFE1*, *EOMES* and *TFPI2* showed the highest specificity. Meanwhile, and the AUC in two of the four combinations reached 0.80, while the other two combinations reached 0.79. In summary, the combinations of the four genes yield a slightly better diagnostic performance (AUC = 0.81, Sensitivity = 0.66, Specificity = 0.87) than the triple gene combinations, and may be more stable in early diagnosis of ESCC due to the increased features

**The diagnostic ability of the triple gene combinations**

|  |  |  |  |
| --- | --- | --- | --- |
| **Combinations** | **Sensitivity** | **Specificity** | **AUC** |
| *ADHFE1, EOMES, SALL1* | 0.65 | 0.86 | 0.79 |
| *ADHFE1, SALL1, TFPI2* | 0.73 | 0.73 | 0.79 |
| *ADHFE1, EOMES, TFPI2* | 0.65 | 0.87 | 0.80 |
| *EOMES, SALL1, TFPI2* | 0.70 | 0.83 | 0.80 |

***Question 8:*** *The language of the manuscript needs some revision. Vague statements like “flexible stability and diagnosis accuracy of DNA methylation” should be replaced.*

**Answer 8**: Thanks for your suggestion. In the revised version of manuscript, we have deleted these vague statements and the manuscript has been thoroughly edited by a native English speaker who has a PhD degree in Biology and is an editor of a biological journal. In the revised version, we have highlighted the corresponding changes with red color.

Again, we thank the reviewer for the brilliant comments and suggestions. We hope that these changes could improve the readability of our manuscript.