Peer Review File

Article information: https://dx.doi.org/10.21037/tcr-21-1162

Reviewer A

Comment 1: Authors have downloaded GBM RNA-seq data from TCGA and then used 5 normal brain expression data from another dataset to analyze expression differences in GBM tumors. It is generally known that RNA-seq platform can significantly affect the results and there are even reports on this problem and how to normalize datasets from different resources before combining data for and gene expression analysis. But it is not clear whether authors have considered any normalization tool to their dataset from two different resources.

Reply 1: It may be that the information of dataset described in the methodology section was easily confused, so the explanation was as follows:

① Data downloaded from TCGA including 169 GBM and 5 non-tumoral brain samples was used for differentially expressed mRNAs (DEmRNAs) and differently expressed lncRNA (DElncRNAs) analysis. The "edgeR" package was used to normalize the dataset before differently expressed analysis.

② While the GSE25631 dataset contained miRNA microarray data of 82 GBM tissues and 5 normal brain tissues was used to identify differentially expressed miRNAs (DEmiRNAs). The "edgeR" package was also used to normalize the dataset before differently expressed analysis.

③ The differential analysis of the two datasets was independent, so there was no need to merge for batch correction.

Changes in the text: No

Comment 2: All significant genes reported in their final conclusion, has already been reported to be important in GBM pathogenesis and there is no new finding in this manuscript.

Reply 2: ①The aim of our study was to identify prognostic biomarkers based on

ceRNAs.

②Beyond identifying the previously reported important prognostic molecules DCBLD2, H19, HMGA2, we also discovered new important prognostic molecules TCF12, ITGB3, C10orf25, LINC00343, FGF12-AS2, LINC00336, HOTAIRM1. Previous research about these key molecules have been descripted in the discussion part.

Changes in the text: no

Comment 3: It is reported that the found H19 expression to correlate with survival of GBM patients. H19 is one of the best-known examples of gene imprinting in brain and it is important to check whether changes in the expression of H19 is correlated to lose of imprinting (LOI) event in these tumors.

Reply 3: We consulted a lot of literature under the advice. As we all known, loss of imprinting (LOI) is one of the most common genetic factors that cause tumors and associated with poor survival in patients. Previous studies have reported that upregulated expression of H19 caused by Loss of imprinting was associated with poor survival of patients or promoted neoplastic development. However, the relationship between imprint loss of H19 and GBM has not been reported, and it is will be a good direction in the next research.

Changes in the text: no

Comment 4: In their first step of expression analysis, they found 1668 upregulated mRNAs and 1872 downregulated mRNAs, 869 upregulated lncRNAs and 1297 downregulated lncRNAs based on TCGA dataset and 2 upregulated miRNA and 64 downregulated miRNAs based on GEO data set, but in the next step they have only considered 56 DEmRNAs, 188 DElncRNA and 8 DEmiRNAs for ceRNA. The exact selection criteria for this selection is not clearly defined. Also it is not shown how was the expression pattern of these selected genes compared to other DE genes in their analysis.

Reply 4: ①Our DEmiRNAs and DElncRNAs/DEmRNAs were identified form two different databases. DERNAs (56 DEmRNAs, 188 DElncRNA and 8 DEmiRNAs) for

ceRNA network were reidentified based on all DERNAs identified from TCGA and GEO database as descripted in the flowchart in figure 1, and the detailed selection criteria was re-described in the methodology and result section.

②These DERNAs in the ceRNA network were significant based on the analysis of identification of differential expression of mRNA, lncRNA, and miRNA in figure 2. Besides, selection criteria of these DERNAs were not based on expression pattern compared to other DE genes. More importantly, we focused on the potential regulatory mechanism of these DERNAs in GBM.

Changes in the text: we have modified our text as advised (see Page 6, line 7-20, page9, line3-8)".

Comment 5: Although the main idea in this manuscript is ceRNA network and authors have reported C10orf25/mir-218/DCBLD2 regulatory axis to be important in GBM. This conclusion in not supported with their data. They have to perform detailed functional studies to confirm the suggested axis to be important in GBM or they should totally remove all their theoretical conclusions out of this manuscript as there is no real evidence for their assumption.

Reply 5: After checking the results repeatedly, we do can't get a direct conclusion that C10orf25/mir-218/DCBLD2 regulatory axis was important in GBM based on existing data. We have combined the last two results section as the lncRNA-miRNA-mRNA regulatory axis of the hub gene in GBM. According to the functional analysis, we further confirm that molecules in our model constructed above were closely related the progression of GBM on one hand. On the other hand, C10orf25/mir-218/DCBLD2 may be an important regulatory axis discussed in the discussion part.

Changes in the text: We have modified our text as advised (see page2/line8,13-14, page12/line 15-17, page13/ line 17-22, Page 17/line11-17, page18/line1).

Comment 6: Discussion part is too long and a lot of the text there should be moved either to introduction or result sections to make it clearer and easier to follow.

Reply 6: We have modified as suggested.

Changes in the text: The current state of lncRNA-mRNA-miRNA-gene analysis in GBM in discussion part was appropriately moved to the introduction. (see page 3, line22, page4, line 1-14).

Comment 7: All the bioinformatics scripts applied to analyze data in this manuscript should be provided as supplementary data.

Reply 7: Like many bioinformatics analysis articles, we have described the key scripts and analysis packages used in each step of the analysis in the corresponding methods and results.

Changes in the text: no

Comment 8: In their pathway analysis, they have found CMV infection and Papilloma virus infection as two important gene sets but there is no discussion of these two pathways at all in their manuscript.

Reply 8: As we all known, many viral infections are related to the occurrence and development of cancer. They can induce the transformation of normal cells into cancer cells and this may be the underlying cause of carcinogenesis in many different types of cancer. Previous studies have found the relevance between Glioblastoma and CMV infection or Papillomavirus infection. In our study, we found CMV infection and Papilloma virus infection as two important gene sets in our pathway analysis, which may indicate that such viral infections were related to the progression of GBM in our data. The two gene sets were recognized as cancer-related pathways, but we didn't list all in the article. And further research will be discussed in the next experiments, so there is no exact discussion in the article.

Changes in the text: human cytomegalovirus and human papillomavirus infection were added in the text. (see page10, line 3)

Reviewer B

Comment 1: The introduction needs significant work. The last paragraph detailing the

study findings belongs in the conclusions section, not here. Conversely, a paragraph is missing on the current state of mRNA-miRNA-gene analysis in GBM and other cancers. This paragraph should describe the current state of the art (e.g., DOI: 10.1038/s41598-019-56983-x, 10.2147/CMAR.S314011, 10.1093/nar/gkt1054, etc.] My search found 66 citations in web of science under miRNA-mRNA AND glio*. I suggest the authors mine some of these papers to write an appropriate current state of the art paragraph and add it to the introduction.

Reply 1: It is a good advice and we have modified our text as advised.

Changes in the text: we have modified our text as advised (see page 3, line 22, page4, line1-14).

Comment 2: Fig 1. Additional definitions are needed (e.g., GSE 25631, miRcode). It would also be helpful to add links to all of the databases and packages used here or in the supplemental. Also, the DE prefix should be defined here.

Reply 2: Fig 1 is the flowchart of this study, and all of the databases and packages used here were described in detail in the material method section. DE prefix also defined in figure 2 legend and material method section.

Changes in the text: no

Comment 3: Figure 2A, B. Normally volcano plots are shown with the data rotated 90 deg. Counterclockwise (i.e., data forms a V visually). Usually these are FC vs. -log(p value) with the significant data indicated in colors and shown as the top of the V. Please adjust.

Reply 3: we have adjusted our figure 2 A, B as advised.

Changes in the text: we have adjusted our figure 2 A, B as advised. (see the profile of figure +table).

Comment 4: Construction of the ceRNA network in GBM. P. 8. Please provide more detail on how the list from the previous section was narrowed down for analysis here. Some of this is in the methods, but this should be explained in the results section as

well.

Reply 4: Our DEmiRNAs and DElncRNAs/DEmRNAs were identified form two different databases. DERNAs (56 DEmRNAs, 188 DElncRNA and 8 DEmiRNAs) for ceRNA network were reidentified based on all DERNAs identified from TCGA and GEO database as descripted in the flowchart at figure1, and the detailed selection criteria was re-descripted in the methodology and result section.

Changes in the text: We have modified our text as advised (see Page 6, line 7-20, page9, line3-8).

Comment 5: Fig 3, 4. Please provide at higher resolution and larger. It is difficult to read the RNA/pathway names even when the figure is zoomed in.

Reply 5: We have adjusted our figure 4 as advise. However, we tried to modify figure 3, but the font still cannot be displayed very clearly when the figure is small. We have confirmed that when the figure is saved in PDF format, the font is clearly visible.

Changes in the text: We have adjusted our figure 4 as advised (see the profile of figure +table)

Comment 6: Similarly, in the following section on the GO and KEGG analysis, how was this narrowed down list of 56 DEmRNAs selected?

Reply 6: The GO and KEGG analysis were to further elucidate underlying biological functions and main signaling pathways of key genes in the ceRNA network. 56 DEmRNAs in the ceRNA network was selected as described in flow chart of this study and described in detail in the methodology and result section.

Changes in the text: We have modified our text as advised (see Page 6, line 7-20, page9, line3-8).

Comment 7: Fig 5. Please move the p values in the caption and increase the figure size. It is hard to read.

Reply 7: We have adjusted our figure 5 as advised.

Changes in the text: We have adjusted our figure 5 as advised (see the profile of figure

+table)

Comment 8: It is unclear how info in the "prognostic markers" section differs from the following section. I think this looks at the genes impacted by the DElncRNA? This is very unclear. Both contain DElncRNA analysis. The latter section seems more extensive (uses all DElncRNAs) versus the smaller set in the previous section. Can these sections be combined?

Reply 8: ①Prognostic markers were identified as independent prognostic factors. Univariate cox regression followed by multivariate cox was performed to construct risk prediction models for GBM.

②prognostic markers and risk prediction models were independent analysed based on 188 DElncRNA and 56 DEmRNAs in the ceRNA network. All DElncRNAs refer to 188 DElncRNAs, All DEmRNAs refer to 56 DEmRNAs.

Changes in the text: All DElncRNAs changed to All DElncRNAs in the ceRNA network, All DEmRNAs changed to All DEmRNAs in the ceRNA network, (see page 10/line 19-20, page 11 /line 21).

Comment 9: Table 3 is very confusing. There is no definition of what the different * levels are. I suspect these are differences in significance, but there are 7 so labeled items and only 5 mentioned in the discussion part. I think the other 2 TCF12 and DCBLD2 are genes and not lncRNAs. However, the column label is deceptively given as "gene". The authors should potentially split this into two lists or at least designate which are genes and which are lncRNAs.

Reply 9: We have modified our table as advised.

Changes in the text: We have modified our table as advised (see the profile of figure +table).

Comment 10: P. 10. I think there is a typo in the Figure 6C call-out. It says $p < 0.00$, which is not possible.

Reply 10: We have corrected the typo. It should be $p=0.00$.

Changes in the text: We have corrected the typo as advised (see page 11, line 11)