TCR TRANSLATIONAL CANCER RESEARCH
ADVANCES CLINICAL MEDICINE TOWARD THE GOAL OF IMPROVING PATIENTS' QUALITY OF LIFE
IMPACT FACTOR 1.241

Peer Review File

Article information: https://dx.doi.org/10.21037/tcr-21-1145

Reviewer Comments

In this study entitled 'Characterization of diagnostic and prognostic significance of cell cycle-linked genes in hepatocellular carcinoma' Wang et al describe using a prediction model to generate risk signatures of cell cycle genes associated with poor prognosis in HCC patients. Furthermore, they tested the accuracy of the prediction model in two independent HCC patient cohorts. They have also explored the link between risk signatures with clinicopathological parameters, immune status and mutation status. I believe this study is very important and will contribute towards generation of robust biomarkers for HCC prognosis in the future that are currently lacking in the field.

Comment 1. Are the authors able to segregate HCC patients into early or advanced stages prior to testing differential expression of cell cycle gene signatures? As these are public databases, I can understand if this segregation is not possible before the expression analyses.

Reply 1: Thank you very much for your valuable comments. Before confirming the differentially expressed cell cycle-related genes, we can divide the patients into early and advanced stages based on their clinicopathological stages. According to most of the literature, clinicopathological stage I and II are classified as early stage, while clinicopathological stage III and IV are classified as advanced stage. Although, TNM stage is the main cancer staging system and the basic determinant of disease prognosis. However, patients in the same clinicopathological stage often have large differences in prognosis. Therefore, it is urgent to find some reliable biomarkers to further reflect the individual biological differences of patients. In our study, we constructed a prognostic model based on cell cycle function in order to supplement the shortcomings of the TNM staging system. Therefore, before performing gene expression analysis, we can assess whether a patient is early or advanced cancer

according to the classic TNM staging system; however, to assess the difference in prognosis of patients with the same stage, reliable molecular markers need to be explored and confirmed. Thank you again for your review.

Comment 2. Please clarify that the 50 cell cycle genes were identified in TCGA dataset in line 179.

Reply 2: Thank you very much for your valuable comments. Here, we apologize for not being able to clearly indicate which data set these 50 genes come from. As suggested by the reviewer, we have clarified that these 50 genes are confirmed based on the TCGA dataset. Thank you again for your review. (see Page 9, line 181)

Changes in the text: Page 9, line 181

Comment 3. Figure 2A, C and F, the names of the genes are not clear. Can the resolution be improved or can the names of the genes be mentioned in a separate table? The values on the bar graph in figure 2A is hard to see as it is in blue and the color of the bar graph is blue as well.

Reply 3: Thank you very much for your valuable comments. Here, we apologize for the confuse caused by this issue. As suggested by the reviewer, we have redrawn Figure 2A and improved the resolution of Figure 2, especially the size of the gene names in Figure 2A, Figure 2C, and Figure 2F. Thank you again for your review. (see revised Figure 2)

Changes in the text: revised Figure 2

Comment 4. 2B venn diagram- number missing on the yellow/prognostic genes. In figure legend please mention what N and T stand for in figure 2C.

Reply 4: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. Please allow us to explain the number missing on the yellow. In this study, based on the results of the PPI network, we screened the top 50 genes from 1875 genes related to the cell cycle for subsequent analysis. Among these 50 genes, a total of 35 genes are significantly differentially expressed between HCC tissues and adjacent normal tissues. At the same time, we also performed Univariate Cox regression analysis on these 50 genes, and found that 30 genes are significantly related to the overall survival rate of HCC patients, and they happen to

be differentially expressed genes. Therefore, in the yellow area in the Venn diagram, there is a result of missing numbers.

As suggested by the reviewer, we have added the sample types represented by "N" and "T" in the figure legend. Thank you again for your review. (see Page 10, line 201)

Changes in the text: Page 10, line 201

Comment 5. Figure 2F – in which HCC patient dataset was this generated from? TCGA or ICGC?

Reply 5: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have clarified that Figure 2F is obtained based on the TCGA database. Thank you again for your review. (see Page 9, line 190)

Changes in the text: Page 9, line 190

Comment 6. Why were 6 genes selected from the 30 gene list?

Reply 6: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. Please allow us to explain how these 6 genes were selected from these 30 genes to construct prediction models. In this study, we identified 30 differentially expressed genes related to the cell cycle that have independent prognostic value. In order to construct a predictive model, we performed Lasso regression analysis on these 30 genes. Lasso is a penalized regression method that selected variables and calculated the coefficients of selected variables. After performing Lasso regression analysis, we screened out 6 genes from these 30 genes for constructing prediction models. Thank you again for your review.

Comment 7. In methods section, the TCGA patient cohort is 371 individuals, however in figure 3 A risk score analysis only 365 patients were included. Can authors clarify the correct number of patients included in this study?

Reply 7: Thank you very much for your valuable comments. Here, we apologize for failing to explain the data processing process in detail. In this study, we further deleted 6 samples from the 371 samples from the TCGA database, including 5 samples with a survival time of 0 days and 1 sample with a missing survival time. This is because before the Lasso regression analysis, the survival time of these

samples did not meet the prerequisites for the Lasso regression analysis. Therefore, there are a total of 365 samples included in the analysis in Figure 3A. Thank you again for your review.

Comment 8. What is aDC and iDC? Lines 258-9.

Reply 8: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have added the full names of these abbreviations. Thank you again for your review. (see Page 12, line 263)

Changes in the text: Page 12, line 263

Comment 9. It is not clear which patient cohort was utilized in figures 7 to 10.

Reply 9: Thank you very much for your valuable comments. Here, we apologize for not being able to clearly explain that Figures 7 to 10 use data from the TCGA database. As suggested by the reviewer, we have revised some sentences in the manuscript to clearly clarify which data set was used for these analyses. Thank you again for your review. (see Page 7, line 151; Page 8, line 158; Page 8, line 163; Page 13, line 276; Page 14, line 289; Page 14, line 298)

Changes in the text: Page 7, line 151; Page 8, line 158; Page 8, line 163; Page 13, line 276; Page 14, line 289; Page 14, line 298

Minor comments

Comment 1. Abstract section line 41 needs to be amended to ' there was a remarkable association of..'

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have corrected this wrong sentence. Thank you again for your review. (see Page 2, line 41)

Changes in the text: Page 2, line 41

Comment 2. Lines 44 and 412 amend the sentence to ' the six genes are expected..'.

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have revised this sentence. Thank you again for your review. (see Page 2, line 44; Page 19, line 417)

Changes in the text: Page 2, line 44; Page 19, line 417

Comment 3. Line 68 : suggesting rewording histological kind to histological type

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have revised this sentence. Thank you again for your review. (see Page 4, line 68)

Changes in the text: Page 4, line 68

Comment 4. Line 92: change sentence to 'genes may help to estimate'

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have revised this sentence. Thank you again for your review. (see Page 5, line 92)

Changes in the text: Page 5, line 92

Comment 5. Line 95: Full form of TCGA and ICGC needs to be mentioned here

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have added the full form of TCGA and ICGC. Thank you again for your review. (see Page 5, line 95)

Changes in the text: Page 5, line 95

Comment 6. Provide website link for STRING in line 113

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have provided the website link of STRING. Thank you again for your review. (see Page 6, line 114)

Changes in the text: Page 6, line 114

Comment 7. Line 180 change the sentence to 'the potential to be independent prognostic biomarker'

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have corrected this wrong sentence. Thank you again for your review. (see Page 9, line 184)

Changes in the text: Page 9, line 184

Comment 8. Line 184: include cell cycle linked genes instead of cycle linked genes.

Reply 1: Thank you very much for your valuable comments. Here, we apologize for

the confusion caused by this issue. As suggested by the reviewer, we have corrected this wrong sentence. Thank you again for your review. (see Page 9, line 186)

Changes in the text: Page 9, line 186

Comment 9. Line 200: Can this sentence be re worded? Not clear.

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have revised this wrong sentence. Thank you again for your review. (see Page 10, line 204)

Changes in the text: Page 10, line 204

Comment 10. In figure 10 B can authors change the label title from patients with female to female patients and patient with male to male patients?

Reply 1: Thank you very much for your valuable comments. Here, we apologize for the confusion caused by this issue. As suggested by the reviewer, we have revised the label titles. Thank you again for your review. (see revised Figure 10)

Changes in the text: revised Figure 10