

# Peer Review File

Article Information: Available at <http://dx.doi.org/10.21037/tau-20-1348>

## Reviewer Comments

Comment 1: Fig.1 lack labels to draw meaningful interpretations. It would be great if the author would add highlights of the 4 prognostic genes.

Reply 1: Thank you for your comment. The labels of Fig.1 have been added in the revised figure. We have also highlighted the 4 prognostic genes according to your valuable suggestion.

Changes in the text: We have modified our figure and added label as advised (see Fig.1, Page 21 and Line 467-474)

Comment 2: After sorting with lasso regression, the author listed 8 immune-related genes that have potential prognostic values. There is no text that explains why only 4 genes were selected to carry on the following studies. From Table S8, the selection was not ranked by HR. The author may consider adding relevant text to explain the rationales of selection.

Reply 2: We appreciate for your valuable comment. After sorting 8 candidates with lasso regression, 4 genes which were statistical significantly differences in in multivariate Cox proportional hazards regression analysis for OS (Table S8). We finally chose the 4 genes with the  $P$ -value  $<0.05$  to construct the prognostic model.

Changes in the text: We have modified our text (see Page 8 and Line 173-175).

Comment 3: The difference of prediction between the training and validation datasets from Fig.3 and 4 could result from cohort differences. Although the ICGC cohort was listed in suppl. Table, it would be great if the author would aggregate the table in a similar form of TCGA cohort to allow a direct comparison between the two.

Reply 3: Considering the reviewer's suggestion, we have aggregated information in Table S3 for comparison.

Changes in the text: We have added the aggregated table (see Table S3).

Comment 4: It would be interesting to compare the Kaplan-Meier study and ROC curve stratified with individual gene expression with the overall score calculated from 4 genes for both training and validation datasets, as the TEK gene expression seems to have a better validation fit than the other three genes (Fig.4).

Reply 4: Thank you for your valuable suggestion. We have conducted the Kaplan-Meier analysis and ROC curve analysis with individual gene expression in both training and validation datasets as below. It is interesting that TEK gene expression have a better predictive accuracy than the other three genes, just as you supposed. In our future study we may focus on the TEK gene for further exploration.

Changes in the text: We have conducted the Kaplan-Meier analysis and ROC curve analysis stratified with individual gene expression (see below).

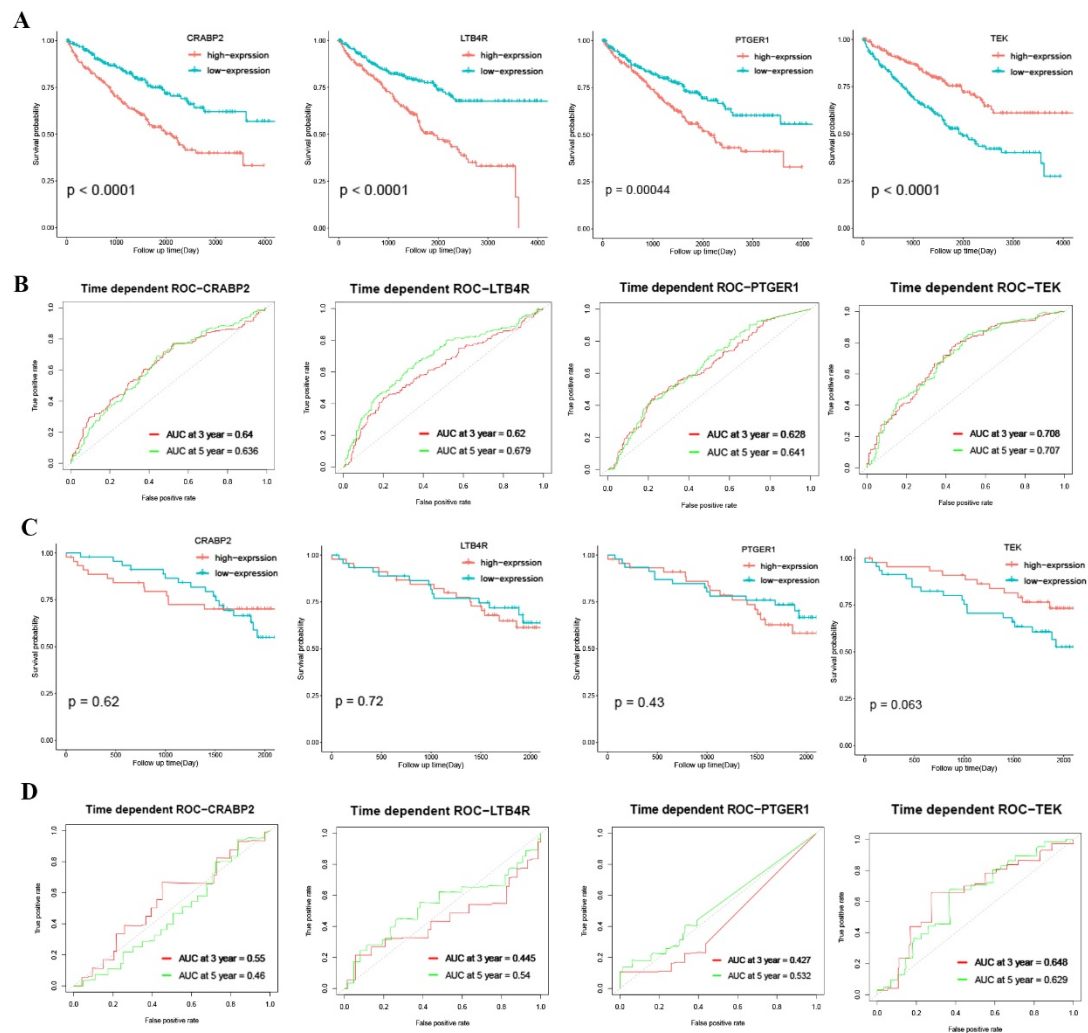


Figure (A) The Kaplan–Meier curves of individual gene for predicting OS in training set. (B) the AUC of 3, 5-year in training set. (C) The Kaplan–Meier curves of individual gene for predicting OS in validation set. (D) the AUC of 3, 5-year in validation set.

Comment 5: The nomogram in Fig.7A provided a straightforward overall picture of the prediction matrix. However, the actual score for each element can be

subjective to interpret from the figure. It would be necessary to provide a table clearly listing the corresponding score of each element in supplementary. Also, do the 3-year survival and 5-year survival were also incorporated in the prediction scoring? If so, it would diminish the prediction value although will surely increase the accuracy.

Reply 5: Thank you for your considerate comment. We have added Table S11 listing the corresponding point of each variable. The 3-year and 5-year survival do not incorporate in the prediction scoring.

Changes in the text: We have added table listing the corresponding point of each variable (see Table S11).

Comment 6: The author could potentially compare their matrix with the widely-used risk score system from MSKCC and IMDC.

Reply 6: Thanks for your comment. Because of the limitation of clinical information in public datasets, currently we are not able to obtain data for serum calcium, hemoglobin, Lactate dehydrogenase, platelet level, neutrophils level which are the parameters for MSKCC and IMDC. Therefore, we couldn't perform the comparison in this moment. Further prospective investigation for this issue would be concerned in our future study.

Comment 7: The interesting observation of the protective effect of TEK expression may relate to the unique poor prognostic value of high CD8

infiltration in RCC. TEK was identified as an anti-inflammatory factor to prevent lymphocyte transmigration from vessels. An in-depth discussion is preferred.

Reply 7: Thanks a lot for your valuable comment. We have made some discussion focus on the inflammatory events mediated by TEK. We hope our work can promote better understanding of the role of TEK in tumor microenvironment.

Changes in the text: We have modified our text (see Page 13-14, Line 289-298).