Dear editor,

Thank you for the positive feedback. We have carefully reviewed the valuable comments from the reviewers and have tried our best to revise the manuscript. Revised portion are marked in red in the paper. Our point by point responses to the reviewer's comments are as follow:

Response to Reviewers' comments:

Reviewer #1 (ID**: 03724953)**:

1. In this study, Long Pan and colleagues analyzed The Cancer Genome Atlas (TCGA) in the aspect of genes associated with tumor microenvironment and prognosis of hepatocellular carcinoma. With the usage of ESTIMATE algorithm method, it would be possible to deeply investigate the gene expression according to the cell types in target tissue. Generally, the results are interesting and the methods are reasonable. The manuscript is well prepared and the figures are well illustrated. The major question, which might also be the limitation, is the method in identifying the genes for predicting prognosis.

Response: Thank you very much for the positive comments on our work and all suggestions for improvement. In this study, we only pay attention to tumor microenvironment-related genes and aim to screen the genes associated with the prognosis. LASSO is a popular method for regression of high-dimensional predictors, which uses an L1 penalty to shrink some regression coefficients to exactly zero[1-4]. The penalty parameter λ, called the tuning parameter, controls

the amount of shrinkage. The larger the λ value, the fewer the number of predictors selected. LASSO has been extended and broadly applied to the Cox proportional hazard regression model for survival analysis with high-dimensional data. LASSO can also be used for optimal selection of markers in high-dimensional data with a strong prognostic value and low correlation among each other to prevent overfitting. Therefore, we adopted the penalized Cox regression model with LASSO penalty to simultaneously achieve shrinkage and variable selection.

2. It seems that the clinical stage and other wellknown risk factors are not adjusted in this model. Would is possible to do the multivariate analysis?

Response: Thank you for your suggestion and we completely agree with your advice. We included several factors such as TNM stage, age, gender, race, Child-pugh classification, ECGO, AFP, Fibrosis ishak score, vascular invasion, hepatitis virus and alcohol consumption and performed the univariate Cox regression. Factors whose P value < 0.05 in the univariate analysis were selected for multivariate analysis. We have revised our manuscript and the results are shown in the Table 3 (Line 23-27, Page 7; Line 7-10, Page 11).

Reviewer #2 (ID**: 00054255)**:

This is pure bioinformatics/biostatistics work using data opened publically. To pick out significant genes from the database, authors should verify each gene in vitro/in vivo using their own samples.

Response: Thank you for your advice. Our study initially used TCGA database to screen the promising genes, which were further validated by four independent HCC-cohorts including GEO datasets (GSE76427, GSE14520 and GSE10143) and Japan cohort from International Cancer Genome Consortium. The results were shown in the table 1. Besides, ten key genes were identified as the most powerful prognostic markers via LASSO algorithm. Among them, two genes (NLRC3 and CD5L) have been shown to be involved in the physiopathologic mechanism of liver cancer by other scholars[5,6], which further increase the credibility of our study. Specifically, Ma et al. have found that high expression levels of NLRC3 correlated with a favorable clinical prognosis in a Chinese HCC population, and down-regulation of NLRC3 expression inhibited cell apoptosis and contributed to cell proliferation in vitro[6]. Another gene, CD5L, called apoptosis inhibitor of macrophage, was investigated by Maehara et al. and they found that CD5L can prevent the hepatocellular carcinoma through complement activation[5]. Besides, other two studies also found that CD5L is in favor of overall HCC survival, which was consistent with our study[7,8]. The related section of the manuscript is also revised (Line 10-12, Page 13). To make the results more visual and clearer, we add K-M survival curves of 10 key genes to the manuscript as Figure 5 (Line 24-26, Page 10). Of course, it is true that the genes our study screened are not verified in vitro/vivo, which is the limitation of our study (Line 27-28, Page 13). As for other 8 genes, we will investigate their roles in HCC in future studies.

Reviewer #3 (ID**: 00053659)**:

1. Pan L et al. investigated to analyze several databases to identify a prognostic profile of the gene expressions using the ESTIMATE algorithm. They identified that ten key genes (STSL2, TMC5, DOK5, RASGRP2, NLRC3, KLRB1, CD5L, CFHR3, ADH1C, and UGT2B15) constructed a prognostic gene signature. They claimed that one of the ten key genes might be a candidate for targeted molecular therapy in the future. The analysis is quite unique and seems to be interesting.

Response: Thank you so much for your encouraging review.

2. The only concern is that the clinical information of the data set was unclear. The etiology of the HCC is very unique, depending on viral status. I have no idea how the data provided in detail. However, the essential clinical background should be presented, such as viral status, liver functional status, and tumor status.

Response: Thank you so much for the insightful review. We have added the baseline information of HCC patients to the manuscript as supplementary table 1, where we present the difference of the clinical and pathologic data between high and low immune/stromal score groups (Line 15-16, Page 8).

3. The AUC of the model seems to be similar even though the value was higher than the conventional TNM stage. It is very hard to evaluate any clinical benefit of this study without clinical characters. Therefore, I am quite skeptical about their conclusion was

that the results presented the future molecular target under the current results. A comparison of the AUC among their model and TNM stage should analyze multiple comparisons of the ROC analysis.

Response: Thank you so much for your encouraging review and suggestions. According to your advice, we have added the baseline information of HCC patients to the manuscript as supplementary table 1, where we present the difference of the clinical and pathologic data between high and low immune/stromal score groups (Line 15-16, Page 8). In addition, we put the ROC curves of gene signature and several clinicopathologic features on one graph (Figure 4H), which can visually display the performance difference between the 10-gene-based signature and TNM stage (Line 13-17, Page 11).

Response to Editors' comments:

Response: Thank you so much for your suggestions. We have revised the manuscript and the figures according to the requirement.

References:

1. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med* 1997; 16(4):385-95.
2. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society Series B-Statistical Methodology* 2011; 73:273-282.
3. Zhang JX, et al. Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. *Lancet Oncol* 2013; 14(13):1295-306.
4. Jiang YM, et al. ImmunoScore signature: a prognostic and predictive tool in gastric cancer. *Ann surg* 2018; 267(3): 504-513.

5. Maehara N, et al. Circulating AIM prevents hepatocellular carcinoma through complement activation. *Cell reports*, 2014, 9(1): 61-74.
6. Ma YY, et al. The correlation of NLRC3 expression with the progression and prognosis of hepatocellular carcinoma. *Human pathology*, 2018, 82: 273-281.
7. Zhang X, Kang C, Li N, Liu X, Zhang J, Gao F, Dai L. Identification of special key genes for alcohol-related hepatocellular carcinoma through bioinformatic analysis. *PeerJ* 2019; 7: e6375.
8. Brunner SM, et al. Tumor-infiltrating B cells producing antitumor active immunoglobulins in resected HCC prolong patient survival. *Oncotarget* 2017; 8(41): 71002-71011