Dear Editors and Reviewers:

Thank you for your letter and for the reviewers'comments concerning our manuscript entitled "Identification of a five-lncRNA signature to improve prognosis prediction for patients with hepatocellular carcinoma". Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our researches. We have studied comments carefully and have made correction which we hope meet with approval. Besides, we have our paper re-edited by AJE again for language polishing. The corrected parts according to the reviewers' suggestions are in red. Here are answers to the questions by reviewers.

Reviewer #1:

Minor comments:

Q1 In the introduction section, it is helpful if authors should clearly state what the purpose of the study is.

A1: Thank you for your suggestion and we then added the purpose of the study in the introduction section and they are in red.

Q2 On page 9, line24, 'LnRNA signatures,,,havebeen' must be'LnRNA signatures,,,have been'.

A2: We feel very sorry for this mistake and we have corrected it.

Q3 Abstract. Does the abstract summarize and reflect the work described in the manuscript?The authors have over a word limit. They should shorten their abstract.

A3: We have already shorten the abstract according to your suggestion.

Q4 Key words. Do the key words reflect the focus of the manuscript?the key words reflect the focus of the manuscript, but authors should choose prognosis or survival.

A4: We chose "prognosis" and deleted "survival".

Q5 Background. Does the manuscript adequately describe the background, present status and significance of the study? In the introduction section, it is helpful if authors should clearly state what the purpose of the study is.

A5: Thank you for your suggestion and we add it in this part.

Q6 Ethics statements should be included in the manuscript.

A6: Because TCGA data are a community resource project, additional ethical approval was not acquired, and the present study adhered to TCGA publication guidelines and data access policies.

Reviewer #2:

Major comments:

Q1. As indicated by Figure 6, the new 5-lncRNA signature was better than two existing methods. Although the manuscript has claimed that the data were coming from TCGA, the origins of data have not been clearly stated. Therefore, it is not clear whether they are totally independent. For the development of signature for prediction, this is not acceptable. The signature obtained should be tested in several totally independent datasets so as to ensure that the signature could be generalized.

A1: In this study, we downloaded Level 3 RNA-seq data (HTSeq-counts) from 374 HCC tumor specimens and 50 peritumoral liver specimens from The Cancer Genome Atlas (TCGA) project (https://cancergenome.nih.gov/). A total of 370 HCC patients with complete follow up data were enrolled in our study and randomly divided into a training set (n=184) and test set (n=186) using SPSS software (version 24.0). We acquired the 5-lncRNA signature from the training set and further validated its prognostic value in the test set, which is totally independent from the training set. We also attempted to validate the prognostic value of the 5 lncRNA in other HCC cohorts, so we downloaded data from several GEO datasets. Unfortunately, we failed because most of the 5 lncRNA probes could not be found. In future, we would like to further validate the value of this 5-lncRNA in other independent cohorts.

Q2. In Figure 7, although a lot of GO terms related to cell division and cell cycle, they are quite redundant. Moreover, no enrichment of GO terms has been shown. On the other hand, more of the pathways identified by KEGG were metabolism related. Consequently, the functions of these lncRNAs were not clearly defined. More work should be done to look at the functions of these lncRNAs.

A2: The GO terms related to these 5 lncRNAs seemed redundant since they are derived from more than 300 correlated protein coding genes. To simplify it, at this time we pick up 200 correlated protein coding genes with highest correlation coefficient for GO and KEGG pathway analyses. (We also attempted to pick up 150 or 100 correlated genes with highest coefficient for GO analysis, but it turned out that there were too few genes (only 2-3) enriched in one GO term). Functional enrichment analysis revealed that these 200 genes were primarily enriched in 32 GO terms (Benjamin P value <0.1, Figure 7A) and 23 KEGG pathways (P<0.001, Figure 7B). Further analysis using EnrichmentMap plugin in Cytoscape revealed that these enriched GO functional terms are mostly involved in metabolic processes, fibrinolysis and complement activation (Figure 7A). The KEGG pathways enriched by these 200 genes were also metabolic related. The roles of lncRNAs in cancer are complicated. In future, we should do much more to testify and elucidate the functions of these 5 lncRNAs.

Q3: In Wang Z et al. 2017 (PeerJ), similar signature has been constructed. The aim of this piece of work was similar. I am curious whether there are any overlaps of the lncRNAs of that work with the 5 lncRNAs of this study. If there is no overlap, is there any rationale behind this phenomenon? Could the two signatures be combined to construct an even better signature?

A3: In Wang Z et al. 2017 (PeerJ), they developed a 4-lncRNA signature for prognosis prediction from TCGA LIHC cohort. They performed univariate and multivariate survival analysis with expression value of lncRNAs as independent variable based on the entire TCGA cohort including 371 HCC patients. In our present study, we randomly divided the entire TCGA LIHC set into two independent set. We developed the 5-lncRNA signature from the training set including 184 patients and further validated it in the test set including 186 patients. During our investigation, we acquired many different signatures and finally picked up this 5-lncRNA signature because its high reproducibility and robustness compared with other signatures. Besides, there is one overlapping lncRNA (AC015908.3, also named CTC-297N7.9) in our 5-lncRNA signature and Wang's 4-lncRNA signature.

Minor comments:

Q1. In P. 6, I found this sentence "The median survival time for the high-risk group and the low-risk group was 6.811 and 2.096 years, respectively."

However, a high risk group should have a shorter survival.

A1: We feel very sorry for this mistake and it should be “The median survival time for the high-risk group and the low-risk group was 2.096 and 6.811 years, respectively.”

Q2. Figure 6A and 6B were redundant and therefore unnecessary.

A2: Figure 6A represents survival analysis with 3-gene signature by Binghua Li and its comparison with our 5-lncRNA signature while Figure 6B represents survival analysis with the 4-lncRNA signature by Zhonghao Wang and it comparison with our signature.

Q3. In Figure 6C, 5LncSig is only slightly, or even not, better that than the other signature.

A3: As shown in Figure 6C, the AUC for our 5-lncRNA signature is 0.769, higher than 0.721 for the 4-lncRNA signature by Zhonghao Wang. Besides, the 4-lncRNA signature by Zhonghao Wang was directly derived from the entire TCGA LIHC cohort while our 5-lncRNA signature was derived from the training set. The AUC for our 5-lncRNA signature in the training set is 0.857. As shown in Figure 6B, Kaplan-Meier survival analyses and log-rank test indicated that our 5-lncRNA signature proved better than Zhonghao Wang.