Peer Review File

Article Information: http://dx.doi.org/10.21037/jgo-20-208

Review Comments:

Comment 1. The authors collected data from a retrospective cohort. Therefore, it would be useful to better clarify the methodological statistical approach used in order to limit and even avoid statistical biases. I would strongly recommend taking into account those interesting networks while pinpointing to the strengths and weakness of the study.

Reply 1: Thank you very much for your suggestions; we apologize that some of the statistical methods were not clearly described in the previous manuscript. We have corrected these issues by clearly describing the statistical methods in all cases where they were involved to avoid statistical bias. Secondly, we have considered the networks in the article in more detail in the Discussion section and have stated the possible implications of these networks for future research.

Changes in the text: About the statistical methods used in the article (see Page 4, line 84; Page 5, line 93-96; Page 5, line 111-114; Page 6, line 132-135; Page 7, line 145-149). Additional information on networks (see Page 14, line 320-323; Page 20, line 470-473; Page 22, line 513-518).

Comment 2. Specifically, multivariate analysis can show the prognostic impact of several variables. Did the authors check for a statistical association between CBX family proteins and the other variables with significant impact within the uni- and multivariate analysis? Those are fundamental details and information in this regard should be added to the results.

Reply 2: In the original manuscript, we based our statistical analysis between CBX family proteins and other variables on the UALCAN database. We analyzed all

prognosis-related variables in this database, but the multivariate analysis was not provided, and we apologize for the missing data in this regard. To compensate for this lack of information, we downloaded the transcriptomic and clinical information of gastric cancer from the TCGA database, preprocessed and extracted the clinicopathological information, and performed single-factor and multifactor Cox analysis on the remaining 349 tumor samples to show the effect of multiple variables on prognosis.

Changes in the text: Page 5, line 105-114; Page 10, line 215-226.

Comment 3. Regarding the methods declared, I would point out that the biostatistical tests performed may be statistically significant but biologically less relevant if placed into a more complex context, such as a statistically powered prospective study. To compensate for these limits, the multivariate Cox's proportional hazard regression models is a worthy tool. Nevertheless, a mandatory assumption needs to be taken into account in order to apply such a model: hazards proportionality. This assumption has to be made in order to proceed with Cox model. If the answer is affirmative, this should be better highlighted in materials and methods. If it is not, it is necessary to motivate and discuss the use of alternative models.

Reply 3: We have made changes and additions to address the issues you raised. The specific methods and changes are as described in our response under the second proposal (see Page 5, line 105-114; Page 10, line 215-226). We analyzed the CBX family proteins with other included clinicopathological features using Cox proportional hazard regression models and pooled the results in Figure S1.

Comment 4. In order to increase the impact for a broader spectrum of readers from the oncologic landscape, it would be interesting to deeper navigating the implications and correlation with the existing data, while introducing and discussing the authors' results. Do original preclinical/clinical data exist about the translational relevance of the mentioned result? If yes, these elements should be presented, at least in the form

of discussion and / or additional figure from a short literature meta-analysis. One suggestion might be the role of CBX family proteins potential in modulating GC dissemination. Indeed, CBX family proteins are already investigated in Gastric Cancer, but its biological role remains partially obscure. Those are fundamental information in order to deeper validate the pieces of evidence discussed in the manuscripts. Xiao-Wei Zhang et al. summarized the mechanism underlying the transcriptional regulatory network in GC and how this can help researchers to further clarify the underlying regulatory mechanisms of gastric cancer tumorigenesis and impact on cell migration-invasiveness. The authors should provide insights in this regard, with a particular biological focus discussing CBX1 role in aggressive phenotype acquisition. Indeed, they already allude to the interaction with HMGA2 to trigger the Wnt/β-Catenin signaling pathway in HCC (ref. 12, 16). Nonetheless, I personally miss some important evidence recently published in gastric cancer and other paradigmatic cancer models (PMID: 31277479; PMID: 23135750).

Reply 4: We are fortunate to have received such professional and constructive comments from you, and we are sorry that our previous article was not detailed enough, and some important points were not shown in the article. We have meticulously analyzed the literature you mentioned and supplemented some of the article content with a focus on the role and molecular mechanisms of the CBX family proteins in gastric cancer and the studies relied upon to reach these conclusions.
Changes in the text: Page 12, line 276-284; Page 13, line 296-309.

Comment 5. Anh Tuan Nguyen et al. uncovered CBX family proteins and CBX7, in particular, to be relevant in KRAS dependent oncogenic driving mechanisms (PMID: 21729876). Since Sorafenib has been found to hold great promise in specific. clinical scenario (PMID: 31640191; PMID: 20458043). Moreover, CBX7 regulates stem cell-like properties of gastric cancer cells in immunodeficient models. Since immunodeficiency disorder with a high incidence of gastrointestinal manifestations

and an increased risk of gastric carcinoma and lymphoma (including genetics, immune dysregulation and chronic infections by Helicobacter pylori - PMID: 29393912), it would be relevant to highlight those translational aspects pointing towards a potential Achilles' heel of gastrointestinal cancers that might be exploited therapeutically in the future.

Reply 5: Thank you very much for your revisions to the content of our article. We are sorry that the previous content is insufficient enough, and some valuable information has not been captured. We have carefully read the beneficial literature you have provided and have revised it by synthesizing the content of our article.
Changes in the text: Page 19, line 438-465.

Comment 6. The aforementioned literature and Medline review could improve a lot the translational relevance of the exposed results. I'd suggest expanding. i.e. Has the CBX family role been investigated in other GC patient's subgroup in a controlled-statistically powered study? If not, I would recommend highlighting this topic, in order to better corroborate the translational relevance of the discussed data.

Reply 6: We have reviewed the above literature and found no report on the differential role of CBX family in gastric cancer patient subtypes. We tried to refine the study along your lines, but we were limited by the lack of subgroup data in TCGA, so it was difficult for us to refine the study through the database. However, we plan to collect a sufficient number of gastric cancer patients and distinguish subgroups in future studies to systematically study the role of CBX family proteins in gastric cancer, which will take a long enough time and investment.

Comment 7. There are some linguistic gleanings that require a careful revision, a professional linguistic editing might be advisable.

Reply 7: We apologize for the linguistic shortcomings of our article and thank you

very much for this suggestion. We have considered your comments and have revised and improved the manuscript accordingly before sending it to a professional English editing agency in China to ensure that no more linguistic mistakes are made.

Comment 8. Figure beautification to make the network and bioinformatic enrichment analyses clear would be useful (i.e. The GO analysis exposed it is difficult to be clearly distinguished and need some graphic improvement).

Reply 8: We have optimized some information that is not clear enough on the protein interaction network. Furthermore, we have distinguished the GO enrichment analysis chart from three levels: biological process, cellular component, and molecular function, and each of them shows the top 10 most significant terms. The revised Figure number is Figure7&8.