Peer Review File

Article information: http://dx.doi.org/10.21037/tcr-22-177

Reviewer Comments

Dear Authors, an interesting research article. Pyroptosis is indeed interesting aspect which should be further investigated in carcinogenesis. Although the overall manuscript preparation is good, I see a room for improvement. The are also few other aspects for which I need clarification. Please find them all below:

(1). First, please clarify why OS was used in methodology instead of DFS outcome? Events caused by disease recurrence occur earlier than death from the disease and moreover DFS also include tumors that do not necessarily lead to death, which are not included in OS.

Reply (1):
In our study, OS was used in methodology to filter the PRGs. To our knowledge, OS was used in methodology instead of DFS outcome in many articles (PMID: 34150627, PMID: 32655621, PMID: 31419958, PMID: 33602233). Actually, considering events caused by disease recurrence occur earlier than death from the disease and moreover DFS also include tumors that do not necessarily lead to death, we will investigate the association between DFS and pyroptosis related gene in further study.

(2). No https link is provided for TRIPOD checklist (line 83). Please add the link or refer directly to the checklist which is at the end of PDF file (maybe make it as a supplementary/additional file?).

Reply (2):
The TRIPOD checklist has been provided as a supplement file.

(3). If you used FPKM values of transcriptome data and at the same time “limma” R-package, did you log2-transform them prior to analysis? Please check helpful links such as: https://support.bioconductor.org/p/56275/, https://support.bioconductor.org/p/66233/, https://www.biostars.org/p/389815/.

Reply (3):
The FPKM values of transcriptome data have been log2-transformed prior to analysis. We added the description (see Page Methods, line 92-93)

(4). At the beginning of “Methods” section you wrote that 19 adjacent normal and 411 tumor tissues were examined in this research. However, in Fig1A there are 11 adjacent and 409

tumor samples – does this mean that some patients were excluded due to lack of corresponding clinical data? If yes, then it should be mentioned somewhere in the "Methods", ideally close to the above sentence I referred to.

Reply (4):
The number of adjacent normal tissues was 19, and tumor sample was 411. We have revised in Figure Legend (Fig.1, line 505-506).

(5). In "Methods" you wrote that GSEA was employed to explore signal pathways – does it mean that you used collection C2:CP (canonical pathways) from MSigDB? If yes, then please add such information. I suspect that not the entire C2:CP was used as the enrichment plots in Fig6 include KEGG gene sets only, so it would be C2:CP:KEGG subset that was used in the study.

Reply (5):
We have modified our text as advised (see Page Methods, line 119-120)

(6). Which software was used to present the graph visible in Fig1C? Maybe I misunderstood some part of methodology, but I could not find such description (I really like the final effect, though). Also, am I correct that this correlation network does not include statistical significance?

Reply (6):
The "igraph" package of R language was used to present the graph visible in Fig1C. Red line indicates positive correlation, whereas blue line indicates negative correlation in Fig1C. This correlation network doesn't include statistical significance.

(7). I think that in Fig2B, the end of description in figure legends should be "related 8 PRGs" because at this stage there were still 8 coefficient values but two turned out to be positive. Thanks to this stage of methodology, you excluded two out of eight PRGs and subsequently analyzed only six of them. Moreover, correct "RPGs" to "PRGs" in figure legend.

Reply (7):
We have modified our text as advised (see Page Figure lenends, line 514)

(8). Consider improving the order of introducing subsequent subfigures – initially, I thought that subfigures such as Fig2E, 2G, 2H, and the graphical equivalents in Figures 3 and 4, are not mentioned in the main text but it turned out to be mentioned later in subsequent sections. Also, the PCA plots from Figures 2, 3, 4 are not mentioned at all in the main text.

Reply (8):
We have modified our text as advised (see Results, line 183-188, 201-204,206-211)
We have the description for PCA plots from Figures2, 3, 4 as advised (see Page Results, line 172-174, line 200-201).

(9). This might be a small detail but when describing Figure 2H (lines 183-186) the lack of statistical significance for age in multivariate analysis could be mentioned, even if in other Cox analyses it was found significant.

Reply (9):
We have modified our text as advised (see Page Results, line 188)

(10). I would swap Fig7C and D with E and F, so that you can present subfigures from left to right in two rows. Also, in Fig7C and D there is an error on Y-axis, probably "tumor tmbation burden" instead of "tumor mutation burden".

Reply (10):
We have modified Fig7 as advised (see Fig7).
In addition, we also have modified the ""tumor tmbation burden" to "tumor mutation burden" in Fig7C and D as advised (see Fig7).

(11). Figures 7, 8 and 9 should be entirely visible on one page, currently they are partly at page 26-27, 27-28 and 28-29, respectively. Make sure you upload the complete figures in subsequent submission stages.

Reply (11):
We have modified as advised.

(12). In "Discussion" you wrote that "we identified GSDMB as the most significant biomarker associated with BLCA prognosis, where a high expression level of GSDMB predicted a good clinical prognosis in BLCA patients". Can I ask how GSDMB was the most significant in this study, compared to other five PRGs?

Reply (12):
We have modified (see discussion, line 347-348)

(13). A very minor aspect but maybe add some information about correlation network in the title of Figure 1? The title mostly concerns Fig1A, therefore it could be more detailed to include also correlation network.

Reply (13):
We have modified as advised (see Figure legends, Fig1, line 504)

(14). Since you named Figure 5 as “Boxplots of correlation”: If you would like to draw boxplots with correlation instead of p-values, then this might be helpful -> https://stackoverflow.com/questions/51810586/display-spearman-correlations-between-pairs-of-boxplots-in-ggplot2-ggpubr

Reply (14):
We have modified as advised (see Figure legends, Fig5, line 562-563)

(15). Delete the graphical artefact in Figure 6 between Fig6A-B and Fig6C (there is a half-cut “Gene Ratio” remainder).

Reply (15):
We have modified as advised (see Fig6)

(16). Please correct obvious typos e.g. “he correlation” -> “the correlation” (line 79), “compered” -> “compared” (line 139), “based ton” -> “based on” (line 159), “fgures” -> “figures” (page 20). Also, consider changing “subsequent conducted” to “subsequently conducted” (line 33), “training cohorts” or “validation cohorts” to “training/validation cohort” (without “s”; lines 148/149), “datas” to “data” (line 149), “biology process” to “biological processes” (figure legend of Fig6A), “expression of level” to “expression level” or just “expression”.

Reply (16):
We have modified as advised (see line 80, 147, 168, 503, 33, 157-158, 158, figure legend of Fig6A, 314).

(17). Please make sure to put space before in-text citation like in ref1 (line 50). I spotted that this is not a case for ref2 (line 54) or ref10-12 (line 90).
Reply (17):
We have modified as advised (see line 50, line 54, line 77, line 320, line 330, line 344, line 366, line 375).

(18). Explain abbreviations on first use, examples: FPKM (line 86), DEGs (line 109/110), ESTIMATE (line 122), DRGs (line 195; it means “differently regulated genes”?) Please also double-check the rest of the paper.

Reply (18):

We have modified as advised (see Page Methods, line 87-88, line 115-116, Results, line 212).