Peer Review File

Article information: http://dx.doi.org/10.21037/tcr-22-607

Review comments

Comment 1: Ubiquitination of substrate proteins is regulated by the action of the E1 ubiquitin-activating enzyme, E2 ubiquitin-conjugating enzymes, E3 ubiquitin-protein ligases, and deubiquitinase. The authors need to present the data and results for these genes separately based on their molecular characteristics.

Reply 1: Thank you for your comment. As the Reviewers suggested, we have presented the data and results for these genes separately based on their molecular characteristics and modified our text as advised (see Pages 11-12, lines 411-449).

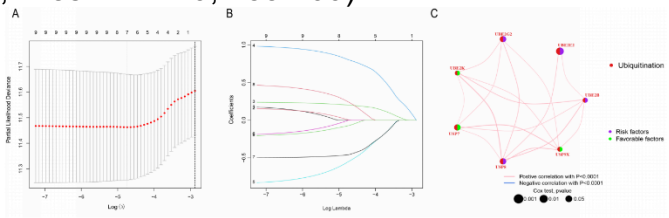
Changes in the text: Among the seven genes, *USP7*, *USP9X*, and *USP8* are all deubiquitinating enzymes (DUBs). Quite a few studies have shown that *USP7* promotes tumor cell development, progression, and metastasis by deubiquitinating the *p53* negative regulatory protein *MDM2*, which decreases intracellular levels of *p53*(1-4). However, *USP7* also prevents genetic alterations in various ways independent of *p53*. It promotes telomere maintenance, repairs broken double-stranded DNA, and regulates DNA damage checkpoint 1-mediated protein stability, which may reduce genomic instability and gene amplification, leading to tumorigenesis(5, 6). These functions probably make it a favorable factor in our results(7). The crystal structure of *USP9X* was shown to be close to *USP7*, with a canonical USP-fold comprised of fingers, palm, and thumb subdomains, as well as an unusual β-hairpin insertion that may have significant effects on its function. Therefore, we speculate that the structural similarity between *USP9X* and *USP7* may have led to their functional similarity, ultimately making them together as favorable factors. Although *USP8* is a DUB, it was classified as a risky factor in our study, which may be related to its remodeling of TME. It has been suggested that inhibition of *USP8* increases *PD-L1* protein abundance and activates NF-κB signaling to trigger innate immune responses and MHC-I expression. Combining *USP8* inhibitors with *PD-1*/*PD-L1* blockers inhibited tumor growth and improved survival in colon cancer mouse model(8). Conversely, high levels of *USP8* may lead to T cell dysfunction, consistent with our results that *USP8* is a risky factor and its expression is negatively correlated with $CD8^+$ T cells(8, 9). *UBE2B*, *UBE2G2*, *UBE2E2* and *UBE2K* are all E2 ubiquitin-conjugating enzymes and E2s are classified into four different types(10). Studies have suggested that downregulation of *UBE2B*, also known as *Rad6B*, attenuates the expression of cancer stem cell markers(11). *Rad6B* protein promotes proliferation and invasion of normal human hepatocytes and has been associated with tumorigenesis and platinum resistance in breast cancer(12-14) and ovarian cancer(15). Meanwhile, Menezes et al. identified deleterious in Blastic Plasmacytoid Dendritic Cell Neoplasm *UBE2G2*(16) mutation. These

findings are consistent with our results that *UBE2B* and *UBE2G2* are risky factors. *UBE2E2* is a class II E2 whose effects in tumors are currently unknown. *UBE2K*, a class III E2, is a favorable factor containing a unique C-terminal. It has been demonstrated that *UBE2K* is involved in mediating polyglutamine aggregate formation and cell death(17). This indicates the importance of ubiquitination for cellular clearance or storage of toxic proteins beyond the proteasome itself, possibly qualifying it as a favorable factor.

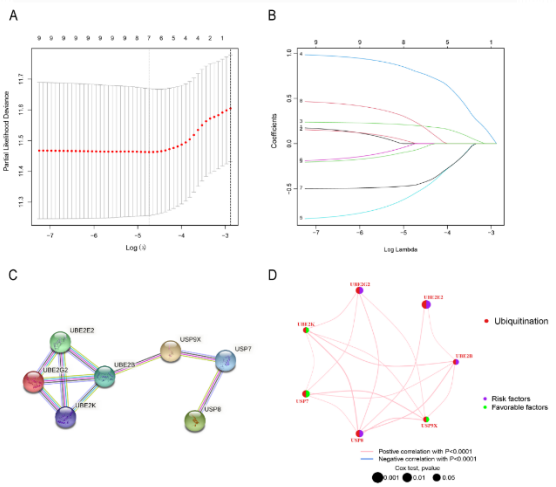
Comment 2: Although the authors identified the 7 survival-related genes useful for predicting patient prognosis, the function and relationship between genes are still unclear. The authors need to analyze the functional connection based on molecular properties.
Do the molecules interact with each other? What kinds of cells express these molecules?

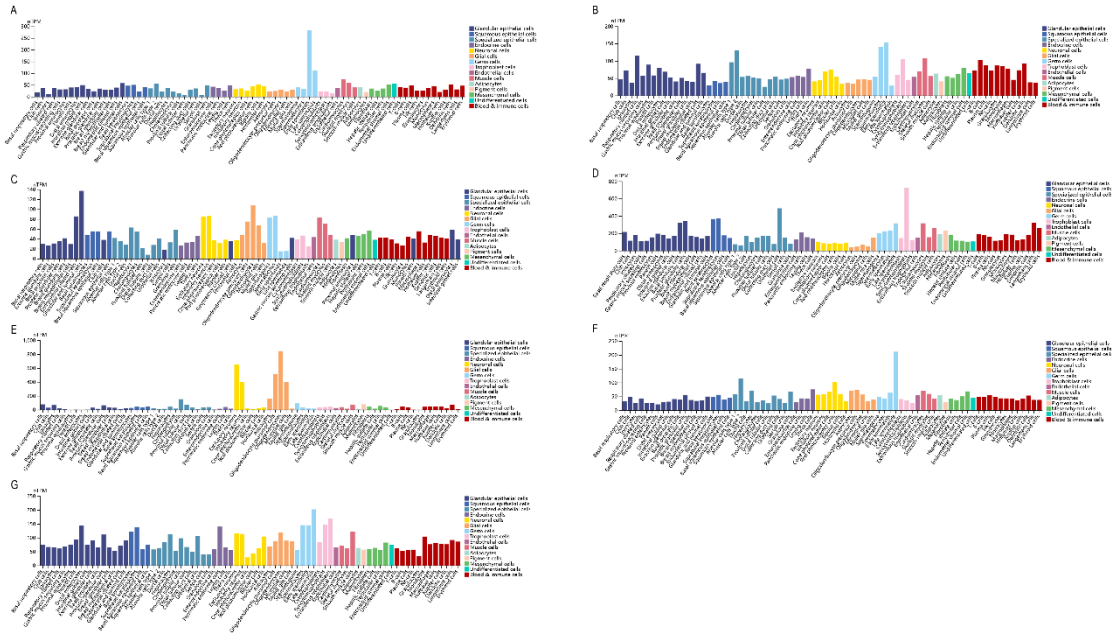
Reply 2: Considering the reviewer's suggestion, we analyzed the molecular properties of 7-URGs based on their functional association. We added the PPI of 7-URG in Figure 3 as shown below. in addition, we downloaded the statistical plot of 7-URG-expressing cells from the Human Protein Atlas and placed it in the supplemental section. We modified our text and added the methods in section 2. (see Page 5, 8, lines 172-179, 288-299).

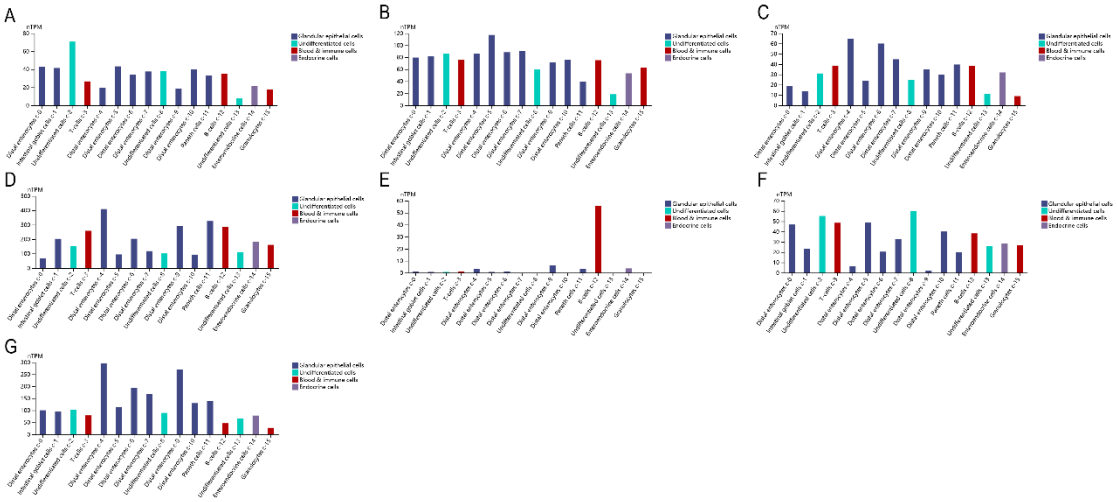
Before



After







Changes in the text: The analysis of PPI network construction can be used to examine the interaction between the seven genes involved in the risk modeling (Figure 3C). And the correlations between the 7-URG were displayed in the network diagram (Figure 3D). The 7-URG expression in the single cell type clusters identified was shown in Figure S1. *USP8*, *USP9X*, and *UBE2K* have low cell-type specificity. At the same time, *USP7* is enhanced in early spermatids, *UBE2B* is enhanced in syncytiotrophoblasts, *UBE2G2* is enhanced in late spermatids, and *UBE2E2* is enhanced in oligodendrocytes, excitatory neurons. Besides, the expression of 7-URG in the single cell-type clusters identified in the colon can be seen in Figure S2.

Comment 3: In Figure 5, the authors show four KEGG pathways were enriched in the high-risk populations. However, it is still unclear what kind of genes were contained in the pathways, and whether there are direct interactions with 7 survival-related genes.

Reply 3: Considering the Reviewer's suggestion, we attempted to analyze 7-

URG with the four KEGG pathways enriched in the high-risk populations in Figure5, but few studies have proven the correlation. However, we found that studies have explored the relationship between four KEGG pathways, ubiquitination, and tumors. We have modified our text as advised (see Pages 12-13, lines 453-480).

Changes in the text: In addition, through GSVA analysis, we found that the high-risk group was significantly enriched for oncogenic activation pathways such as Dilated cardiomyopathy, Glycan metabolism, ECM-receptor interaction and Cytoskeletal regulation, which provides a more comprehensive explanation for its poor prognosis. ARC is an anti-apoptotic protein abundant in cardiomyocytes and plays an important role in mediating apoptosis in dilated cardiomyopathy(18). Roger et al. found that ARC degradation is dependent on the *p53*-induced ubiquitin E3 ligase *MDM2*(19), which *USP7* can deubiquitinate(20). Therefore, we suggest that *USP7* is involved in the development of dilated cardiomyopathy through the *p53*/*MDM2* pathway. Although the relationship between Glycan metabolism(21), ECM-receptor interaction(22-25), Cytoskeletal regulation(26), and 7-URG has not been investigated, ubiquitination is involved in regulating cellular functions by these three pathways. Meanwhile, increasing studies have suggested that these three pathways are closely related to tumor growth and metastasis(27-29), so we speculate that 7-URG is related to the three pathways regulating tumor cell function.

Comment 4: Some of the paper shows that MSI-H patients show a good prognosis (PMID: 15659508; PMID: 33604737). However, this study showed that high-risk patients contain more MSI-H patients than low-risk patients. These results seem to be conflicting, and the authors need to explain and discuss the reason for this contradiction.

Reply 4: Considering the Reviewer's suggestion, we sought to discuss the reasons for the inclusion of more MSI-H patients in the high-risk group than in the low-risk group in our study. We have modified our text as advised (see Pages 10-11, lines 379-391).

Changes in the text: Current studies suggest that CRCs with MSI-H have a better prognosis than MSS, which seems contradictory to our results showing a higher risk score for MSI-H than the MSS group. However, some studies also have highlighted that MSI-H patients have a better sensitivity to 5-fluorouracil than MSS(32-36), making the prognosis of MSI-H patients under 5-fluorouracil therapy potentially favorable. In our study, the treatment regimen of TCGA-COAD patients was unknown. Furthermore, the data of GSE17538 contained a large number of patients without adjuvant chemotherapy(37), making the prognosis of MSI-H patients in our study not necessarily better than that of MSS patients. Therefore, the risk score of MSI may be higher than that of MSS. Although some studies have reported higher survival rates for patients with MSI, estimates of the prognostic value of MSI vary widely among studies(38). The impact of MSI on the prognosis of patients with colon cancer still needs to be further explored.

Comment 5: Line 219, the authors used "vital," an overstatement. The reviewer recommends changing the word to a more neutral one. The authors merely see only a positive correlation.

Reply 5: Considering the Reviewer's suggestion, we replaced the word "vital" with "potential" (see Page 6, line 228).

Changes in the text: …, indicating the potential role of these cryptic URGs in CC development.

Comment 6: In Figure 1B, some genes show frameshift deletion. Did the authors use gene expression data from patients expressing the molecule with loss of function as a protein? If so, the authors need to pay attention to the usage of the patients' data.

Reply 6: Considering the Reviewer's suggestion, we have examined the data from Figure 1B and found that data for this part is missing in the TCGA-COAD dataset. Therefore, we did not modify Figure 1B.

Changes in the text: None.
