

Sheng et al. report in this paper approach of characterizing the risk lncRNA and they develop a lncRNA expression-based signature to predict the overall survival of patients with muscle invasive bladder cancer. They also try to investigate the molecular pathway alteration between the patients stratified by the lncRNA risk scores using gene set enrichment analysis. Following questions/comments were raised during the review of the paper:

1. In the "Statistical and data mining analyses of TCGA BLCA lncRNA profiles" section, the second sentence of the second paragraph, the parameters setting should be clarified: The fold change and P-values were $\text{LogFC} = \log_2(0.5)$ and $\text{P-Value} < 0.05$.
2. Please also clarify the risk score formula description in the method part: $\text{lncRNA}(m) \text{ expression} \times \text{coefficient}(n)$, what is n denoted? This can be written as: the coefficient of each lncRNA was measured by a multivariable Cox regression hazard model with all selected lncRNAs. $\text{lncRNA}(i) \text{ expression} \times \text{coefficient}(i)$ would be better.
3. For the result section "5-lncRNA signature-associated signaling pathways.", please write a few sentences about how the authors conduct the GSEA analysis? High-risk group vs low-risk group? Also, the parameter setting should be mentioned, for instance, the number of permutations, gene sets, and so on.