Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-24-827

Reviewer A

The authors indicate that, based on information from the TCGA database, they analyze the combined expression of mRNA, microRNA, and alternative splicing (AS), and its association with survival in lung cancer patients. The analysis of differentially expressed genes was performed in a cohort of LUSC patients that was separated into sensitive and resistant groups according to reported clinical response to cisplatin. To define risk scores in the two groups of patients, information derived from the gene signature was applied in a univariate and multivariate Cox analysis. In addition, survival was assessed for the different clinical parameters of the patients by Kaplan-Mier analysis. The signature proposed by the authors was compared with that previously reported by other groups, indicating that their combined signature has better predictive for the survival of LUSC patients. In addition, the authors associate tumor infiltrating immune cells in both risk groups using the ESTIMATE database and validate the expression of CTTN and CPM markers in LUSC tissues according to the coefficient of LUSC-splicing factors. Finally, the authors decide to study the cytotoxic effect of two natural compounds in the SKMES1 cell line and suggest that these compounds act on cisplatin-resistant tumors cells.

The manuscript is interesting. However, there are some concerns that need to be attended.

Major concerns:

1. Regarding Figure 1.1, Overview of the study. The authors indicate that patients with LUSC after treatment can be divided into cisplatin-sensitive and cisplatin-resistant. However, the composition depicts viable tumor cells in the sensitive patients and dead cells in the resistance patients.

Re: Thank you very much for your suggestion. We have modified Figure 1.1 so that the tumor cells in sensitive patients are dead cells, while the cells in resistant patients are viable tumor cells.

2. Figure 4. In the Forest plots (E and F), 8 parameters are evaluated. However, in Fig F only 7 hazard ratio data are shown.

Re: We are grateful for you pointing out our mistake. Figure 4F shows that only pathological parameters that are significant to prognosis in the univariate COX ($p<0.05$), while N-stage is meaningless to the prognosis of patients in the univariate

analysis (p>0.05), so no multivariate COX analysis is performed. There is an error and we have deleted the N stage in Figure 4F.

3. Regarding the correlation of the integrated signature and immune infiltration in the low-risk and high-risk groups, there are some cell populations that show significant differences. However, the manuscript only lists these alterations and does not indicate more information on this aspect in the two patient groups, the authors could suggest how the indicated alterations of the immune cells could impact on the prognosis of the clinical response.

Re: Thank you very much for your valuable suggestion. We have added the effect of immune cells on patient outcomes on page 14, lines 467-470.

4. Data from the COX analysis shown in Table S3 describes several genes with a risk ratio and increased significance. Could you clarify what was the authors' decision to validate cortactin (CTTN) and carboxypeptidase M (CPM) as the markers proposed in the manuscript?.

Re: Thank you very much for your suggestion. Table S3 showed the genes that are significant for patients' prognosis by univariate Cox analysis. We further performed LASSO cox analysis on these genes to obtain the final several genes with prognostic value, including the two genes CTTN and CPM. Therefore, these two genes were selected as possible markers.

Changes in the text: None.

5. In the text and Fig 10, the authors indicate that the tested drugs (Sulforaphane and Parthenolide) have the potential to reverse cisplatin resistance. Decrease in the proliferation rate of the SKMES1 by these drugs are shown. To analyze the effect suggested by the authors, it is required to induce acquired resistance to cisplatin in the cell line and subsequently expose these cells to the proposed small-molecule drugs.

Re: Thank you very much for your valuable suggestion. Your suggestion is very correct, but due to the limited conditions, we cannot get these data in time, and we will improve this part in the future.

Minor concerns:

1. Before the abbreviation indicate the meaning. For example: Serine and arginine-rich factor 1 (SRSF1), Decreased bridging integrator 1 (BIN1), etc.

Re: Thanks a lot for your advice. We have added the meaning before the abbreviation in the full text.

Reviewer B

The authors have written the paper well; however, a few minor revisions must be addressed before accepting the paper.

1) The authors are requested to cross-check the occurrence of all abbreviations at their first place of occurrence. E.g., once lung cancer has been mentioned in the Introduction section it must be abbreviated as LC throughout (Line 65, Page 3).
Re: Thank you very much for your suggestion. We have checked the occurrence of all abbreviations at their first place of occurrence.

2) The Introduction and Discussion sections lack the novelty, aim, future prospects, limitations, and strength of current study.
Re: Thank you very much for your suggestion. We have added aim and future prospects in the introduction section on page 5, lines 141-145, adding the following "The purpose of this study is to explore the roles of mRNA, miRNA, and AS in predicting the prognosis of cisplatin-resistant LUSC patients, identify a combination of prognostic biomarkers, and lay the foundation for finding potential therapeutic targets and exploring the mechanism of cisplatin resistance in LUSC patients". Moreover, we have added strength of current study in the discussion section on page 15 lines 483-486, adding the following "In this study, we have constructed an integrated signature and identified several potential compounds, providing possibilities for future treatment of cisplatin-resistant LUSC patients. Furthermore, the predictive performance of this integrated signature is superior to that of individual mRNA, miRNA, and AS biomarkers". Limitations of this study are shown on page 15, lines 486-493.
Changes in the text: We have added aim and future prospects in the introduction section on page 5, lines 141-145. We also have added strength of current study in the discussion section on page 15 lines 483-486.

3) In the data retrieval step, authors must mention the data extraction protocol, pre-processing, and packages used for downstream analysis. The authors can refer to these papers for more information: https://link.springer.com/article/10.1007/s13353-023-00782-8, https://www.nature.com/articles/s41598-023-43484-1, https://www.nature.com/articles/s41598-023-36842-6.
Re: Thank you very much for your suggestion. According to your suggestion, we have referred to these papers in this manuscript on page 6, lines 163.

4) The authors also need to mention the process of batch correction and how they performed using R packages. It would be nice if they could include PCA and scree plots before and after batch correction.

Re: Thank you very much for your constructive suggestion. We have already mentioned the use of R packages in the material method. We analyzed mRNA and miRNAs differential expression by limma package, LASSO cox regression analysis by glmnet 4.0R package, and ssGSEA analysis by GSEAbase package. First, we ensured that all samples came from the same project by downloading RNA sequencing and SpliceSeq data from the TCGA database. Secondly, we utilized the limma package to perform differential expression analysis for genes and miRNAs, employing standardized methods to mitigate technical variations among samples. Lastly, we constructed a predictive model using the LASSO Cox regression model and conducted internal validation. These steps were taken to mitigate the impact of batch effects, ensuring our results possess a certain level of accuracy and reliability.

Changes in the text: None.

5) It would be really interesting to see a comparative analysis between LASSO, RIDGE, and elastic net models instead of only LASSO.

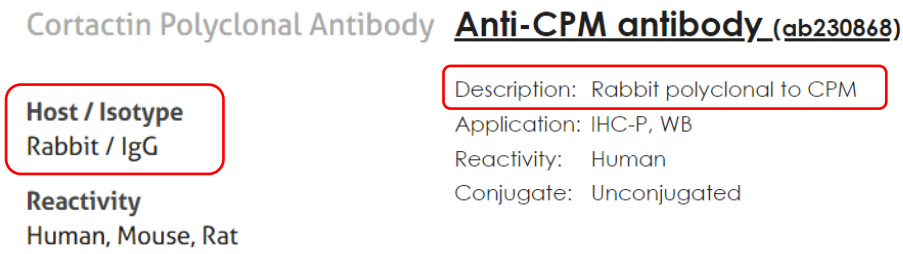
Re: Thank you very much for your constructive suggestion. We will improve this part in the future.

6) Please include a risk-score model of the prognostically significant genes. The authors can refer to this paper for more information: https://www.sciencedirect.com/science/article/pii/S1687157X23015081

Re: Thank you very much for your suggestion. According to your suggestion, we have referred to this paper in this manuscript on page 7, lines 195.

7) Why did the authors choose an anti-rabbit antibody and not an anti-goat or anti-mouse in IHC? Please justify.

Re: Thank you very much for your suggestion. IHC antibodies can be selected to be anti-rabbit, anti-mouse or anti-goat. CTTN antibody comes from proteintch and CPM antibodies comes from abcam. Since both antibodies provided by these two companies are both anti-rabbit, as shown in the figure below, we chose anti-rabbit antibodies.



Changes in the text: None.

8) The authors have not mentioned any information regarding DEGs analysis in the results section.

Re: Thank you very much for your suggestion. The DEGs analysis of mRNA, miRNA and AS has been reflected in the results of Figures 2 and 3.

9) The quality of the figures needs to be rendered at a higher resolution.

Re: Thank you very much for your suggestion. We have increased the resolution of these figures to 600dpi.

10) Please explain how the mutational analysis was performed. Were the sample counts the same as used initially for DEGs analysis?

Re: Thank you very much for your suggestion. In our study, mutational analysis was conducted to identify and analyze genetic mutations in the samples. The specific steps involved in mutational analysis are as follows:

Data acquisition: We downloaded mutation data for LUSC from database, which includes information about gene mutations in the samples.

Protein domain annotation: To further analyze the functional impact of gene mutations, we used the protein domain annotation provided by the maftools package to extract domain annotation information for specific genes. This annotation information helps understand whether the mutations occur in important functional domains.

Extraction of mutation information: From the mutation annotation, we extracted relevant information such as amino acid positions to enable a more detailed analysis of the mutations.

Visualization of mutation data: By visualizing the mutation information, we gained a better understanding of the mutations in different genes within the samples. These visualizations can include bar charts, scatter plots, or other applicable chart types.

It is important to note that both the mutation data and RNA-seq data were downloaded from the TCGA database. Therefore, the sample counts used should be the same as those initially used for the differential expression gene analysis (DEGs analysis). This

ensures that the mutational analysis and differential expression analysis are performed on the same set of samples, allowing for meaningful comparisons and integration.

Changes in the text: None.