**Peer Review File**

Article information: https://dx.doi.org/10.21037/jtd-23-1712

**Reviewer A**

The manuscript presented by Peiquan Zhu and co-authors, titled "Systematic Analysis of Apoptosis-Related Genes in the Prognosis of Lung Squamous Cell Carcinoma: A Combined Single-Cell RNA Sequencing Study," I believe presents interesting to this journal readers. Following an in-depth review, I acknowledge the substantial contribution this work offers to the field of lung squamous cell carcinoma (LUSC) and its implications for understanding apoptosis within this clinical context. However, some changes will improve the quality and reproducibility of this work.

Given the difficulty of the bioinformatic analysis described in this work, I suggest methodological enhancements to the authors to strengthen the reproducibility of results and facilitate comprehension within the scientific community. Specifically, some modifications are required in the methodology section to specify the precise versions of software packages employed in the bioinformatics analysis and the parameters utilized in each instance. Additionally, I want to emphasize the necessity of providing public access to the R scripts used for the calculations, as this transparency is pivotal for result reproducibility. These modifications will enhance the clarity and reproducibility of results, thereby strengthening the impact and validity of this work.

In any case, I doubt how the high-risk and low-risk genes have been grouped, just like why the mutations were more significant in the low-risk group compared to the high-risk group.

On the other hand, I am concerned regarding the translational application of these results into clinical practice due to the intricacies of the technique involved. While the findings hold promise, there exists a potential gap between the complexity of the methodology and its practical implementation in a clinical setting. Addressing this aspect significantly bolsters the significance of the study's applicability in real-world clinical scenarios.

**Comment 1:** Improvement of the methodology in the article.

**Reply 1:** we appreciate your thorough review of our research and the valuable feedback you provided. In response to your specific suggestions, we have made improvements in the revised manuscript.

**Change in the text:** We have modified our text as advised (see Page 6，line 178-180), (see Page 6，line 185-187) , (see Page 7，line 203-205) , (see Page 7，line 230-232).

**Comment 2:** Regarding how to group the high and low-risk groups and how mutations become significant in the low-risk group compared to the high-risk group.

**Reply 2:** Regarding your question about the grouping of high-risk and low-risk genes, I will

provide a detailed explanation.

1. The division of patients with lung squamous cell carcinoma (LUSC) into high-risk and low-risk groups was based on a prognostic model constructed to predict overall survival (OS). The least absolute shrinkage and selection operator (LASSO) penalized Cox proportional hazards regression model was employed to refine the model and address the issue of overfitting.
2. For both the training and testing sets, we calculated the risk score for each patient based on the regression coefficients of the ARGs. The risk score formula is as follows: risk score = expression level of gene a × coefficient a + expression level of gene b × coefficient b + expression level of gene c × coefficient c + ... + expression level of gene n × coefficient n. Patients with LUSC were then divided into high-risk and low-risk groups based on the median risk value.

Therefore, the division of high-risk and low-risk gene groups was based on the risk scores calculated using the method described above. The risk score is a quantitative indicator that integrates the expression levels of multiple genes and their coefficients in the prognosis of LUSC patients. This approach allows us to quantitatively categorize patients into different risk groups based on their tumor's gene expression profile and prognostic risk.

Thank you for your insightful question regarding the observed differences in mutation significance between the low-risk and high-risk groups. If mutations appear more significant in the low-risk group compared to the high-risk group, it may seem counterintuitive; however, it is important to consider the following points:

1. The prognostic model is based on the expression levels of selected genes, which may not directly correlate with mutation frequency. Gene expression can be influenced by a variety of factors, including but not limited to genetic mutations. Regulatory elements, epigenetic modifications, and post-transcriptional changes can also significantly impact gene expression.
2. The risk score derived from the prognostic model is a composite measure that integrates multiple aspects of gene expression. It is possible that patients with fewer or less severe mutations might still exhibit gene expression patterns that are associated with a lower risk of adverse outcomes.
3. Tumor heterogeneity is a well-recognized phenomenon in cancer biology. Within the low-risk group, there may be subpopulations of tumor cells with distinct mutational profiles that do not necessarily confer a high risk of poor prognosis. Conversely, in the high-risk group, there may be mutations that, while less frequent, are more deleterious.
4. Lastly, the observed pattern may be a result of the specific cohort studied, and these findings could vary with a different population or larger sample size. In light of these considerations, the finding of more significant mutations in the low-risk group does not necessarily contradict the validity of the prognostic model.

**Change in the text:** None

**Comment 3: Concerns about the translational application of results in clinical practice**

**Reply 3:** The concerns you mentioned about the translational application of research findings in clinical practice are indeed crucial. We understand that the complexity of these technologies may limit their use in routine clinical settings. It is recognized that the current application of single-cell RNA sequencing (scRNA-seq) technology in clinical practice is relatively limited. However, our study, by revealing the specificity and heterogeneity of different cell populations within the tumor microenvironment, provides essential biomarkers and targets for future personalized therapies. To bridge the current gap between the complexity of the technology and its clinical application, we believe that standardizing the scRNA-seq workflow to make it more suitable for clinical environments is crucial. Additionally, it is of utmost importance to validate our prognostic models and biomarkers in larger-scale multicenter clinical samples.

I hope the above explanation clarifies your concerns.

**Change in the text:** None

**Reviewer B**

1. The study relies on RNA-seq data from TCGA and scRNA-seq data from GEO. The quality and consistency of data from different sources could impact the reliability of the results. A detailed quality control assessment and potential batch effect correction should be considered.
2. The criteria for selecting cells for scRNA-seq analysis (gene expression, mitochondrial content) are briefly mentioned. A more thorough justification of these thresholds and potential impact on downstream analysis is needed. Additionally, the handling of missing clinical information from TCGA should be explicitly discussed.
3. The differential expression analysis of ARGs is conducted using the "limma" package, but there is limited discussion on the biological relevance of the selected fold change and FDR thresholds.
4. The steps for developing the prognostic model are outlined, but there is a lack of discussion on how the model's performance is evaluated. Additional details on model validation metrics, beyond Kaplan-Meier curves and ROC analysis, would enhance the robustness of the prognostic model.
5. The nomogram development is outlined, but the rationale behind integrating various clinical risk factors and apoptosis prognostic models needs more explanation. The choice of factors and their weightings in the nomogram should be justified.
6. The pathway enrichment analysis is performed using DAVID, but there is no mention of corrections for multiple testing. A discussion on how multiple testing correction is handled would strengthen the validity of the enrichment results.

7. TMB analysis is conducted, but there is limited discussion on the clinical relevance of TMB in LUSC. The rationale for choosing TMB as a metric and its significance in the context of the study should be addressed.
8. More references on bioinformatics workflow should be added to attract a broader readership i.e., PMID: 36936815, PMID: 35851932.
9. The association between the risk score and immune and stromal components is explored, but the biological interpretation of these associations is not discussed. Elaborating on the implications of the molecular patterns in the tumor microenvironment (TME) would enhance the biological context.
10. The preprocessing steps for single-cell sequencing data are described, but the choice of specific parameters (mitochondrial content, gene count thresholds) requires justification. Additionally, the annotation of cell clusters using singleR and CellMarker should be discussed in more detail.
11. The pseudotemporal analysis using Monocle 2 is outlined, but the biological significance of pseudotemporal trajectories and their relevance to the study's objectives should be clarified.
12. The statistical methods used (Cox regression, LASSO regression, etc.) are mentioned, but additional details on assumptions, model diagnostics, and potential confounders should be addressed for a more comprehensive understanding of the analyses.

**Reply 1:** In response to your concerns regarding the reliance of this study on RNA-seq data from TCGA and scRNA-seq data from GEO, with the potential for differences in data quality and consistency that may impact result reliability, we have taken the following measures to ensure the accuracy and reliability of our study results:
1. Batch Effect Correction: We recognize that data generated under different platforms and laboratory conditions may have batch effects. To mitigate this potential impact, we employed advanced statistical methods to correct for batch effects. Additionally, we used separate training and testing sets from TCGA to ensure the generalizability of our model.
2. Integrative Analysis: During integrative analysis, we paid special attention to differences between datasets and employed appropriate normalization and standardization methods to reduce biases due to technical variability.
3. Biological Interpretation: Our analysis is not solely based on statistical methods but also incorporates biological knowledge to interpret the results. In summary, we have taken multiple measures to ensure the quality and consistency of the different source data used and to minimize the impact of these factors on the reliability of our results.
**Change in the text:** None

**Reply 2:** Thank you for your meticulous review and valuable comments on our study. In response to your questions regarding the cell selection criteria for scRNA-seq analysis and the handling of TCGA clinical information, we provide the following detailed replies:

1. Cell selection criteria for scRNA-seq analysis:

In our study, the criteria for selecting cells for scRNA-seq analysis included gene expression and mitochondrial DNA content. These criteria were set based on existing research experience. Specifically, we excluded cells with a high proportion of mitochondrial gene expression to total gene expression, as this is typically a sign of cells being under stress or in the process of apoptosis. Additionally, we set a minimum detection threshold for gene expression to ensure the quality of the data. The establishment of these thresholds aimed to select active cells to more accurately depict the heterogeneity of the tumor microenvironment.

2. The potential impact of threshold settings on downstream analysis:

We acknowledge in our paper that the threshold settings might have an impact on downstream analysis, such as the potential omission of certain cell subgroups. However, we believe that this impact is limited because our goal is to depict biologically meaningful cell populations, not cells in a state of apoptosis or damage. In future research, we will consider more comprehensive criteria to further explore the biological significance of different cell states.

3. Handling of TCGA clinical information:

1. Data collection: We collected gene expression data and related clinical information from the TCGA database for patients with lung squamous cell carcinoma (LUSC). This information includes important clinical parameters such as patient survival time, survival status, age, gender, tumor staging, etc.

2. Data organization: We thoroughly cleaned and organized the collected data to ensure its completeness and consistency. We addressed missing data and standardized various clinical parameters to facilitate subsequent statistical analysis.

3. We calculated the risk score for each patient and divided patients into high-risk and low-risk groups based on the median risk value. Survival between the two groups was compared using Kaplan-Meier survival curves.

The significance of handling clinical information in our study is to develop an integrated nomogram that incorporates various clinical risk factors and apoptotic prognostic models to predict the 3-year, 5-year, and 7-year survival probabilities for patients with lung squamous cell carcinoma (LUSC).

**Change in the text:** None

**Reply 3:** Thank you for your valuable comments. In response to the discussion on the biological significance of using the "limma" package for differential expression analysis of ARGs and the selection of specific fold change (FC) and false discovery rate (FDR) thresholds, we provide the following reply: In our study, the "limma" package was used to identify genes related to the prognosis of lung squamous cell carcinoma (LUSC). The "limma" package is a widely used tool in the field of bioinformatics, capable of effectively processing microarray and high-throughput sequencing data to detect statistically significant differences in gene expression. The selection of FC and FDR thresholds was based on common biostatistical standards.

Typically, the FC threshold is used to identify biologically meaningful expression differences, while the FDR threshold is used to control the false positive discovery rate, enhancing the reliability of the results. In this study, we chose thresholds with biological and statistical significance to ensure that the screened ARGs play an important role in the development and prognosis of LUSC. We recognize that suitable FC and FDR thresholds may vary for different biological studies and datasets. Therefore, in selecting thresholds, we considered precedents in the literature, the distribution characteristics of the data, and the objectives of our study. The thresholds we selected aimed to balance the opportunity to discover potentially important genes with the need to reduce the false positive rate. Ultimately, our study results indicate that the selection of these thresholds was reasonable, as our risk scoring model demonstrated good predictive performance in predicting the overall survival (OS) of patients with LUSC.

**Change in the text:** None

**Reply 4:** You are correct in pointing out that the description of the performance evaluation of the model in our study is relatively brief. To enhance the robustness of the prognostic model, we actually employed multiple methods to assess the performance of the model. We used ROC curves and the area under the curve to evaluate the model's predictive accuracy at different time points. These metrics help us understand the model's ability to predict short-term and long-term survival probabilities. In addition to C-index analysis, we also used calibration curves to assess the consistency between the predicted risk scores and the actual observed outcomes. This helps us to test the reliability of the model predictions. We also conducted internal validation by repeating the model construction and testing process on different subsets of the same cohort to assess the model's internal stability. Furthermore, in our study, we used Kaplan-Meier survival curves and ROC curve analysis to evaluate the predictive performance of the model. In addition, we employed several other methods to further verify the robustness of the model. Among them, the C-index is a commonly used survival analysis assessment metric that can quantify the consistency between the model's predicted risk score and the patient's actual survival outcome. In our study, C-index analysis indicated that the model's risk score is an independent prognostic indicator and has better predictive value than other traditional clinical parameters. Secondly, multivariate Cox regression analysis assesses the relative importance and independence of each gene in the model, thereby ensuring that the model's predictive performance is not solely dependent on the expression of a single gene.

**Change in the text:** None

**Reply 5:** The development of the nomogram was aimed at providing an intuitive tool for integrating different prognostic factors to predict the prognosis of patients with Lung Squamous Cell Carcinoma (LUSC). The factors we selected include prognostic-related genes screened from multifactorial and LASSO Cox regression analyses, as well as clinical features such as age, sex, pathological staging, etc. These factors were included in the model based on their

significance in the statistical model and clinical relevance. The weight of each factor in the nomogram is determined by the beta coefficients of the Cox regression model. These coefficients reflect the relative impact of each factor on patient prognosis. This means that the length of each factor in the nomogram is allocated according to its statistical weight in the prognostic analysis, ensuring the accuracy and clinical utility of the model. Our nomogram was developed based on rigorous statistical methods and has undergone multiple validations, proving its effectiveness in predicting the prognosis of LUSC patients.

**Change in the text:** None

**Reply 6:** We have carefully considered the issue of multiple testing that you mentioned. We acknowledge that not including multiple testing correction is a limitation in our current analysis. Due to a lack of in-depth expertise in statistical analysis within our team, particularly in conducting multiple testing, we face certain challenges. We understand the importance of multiple testing, especially when performing enrichment analyses, to control the false positive rate and ensure the reliability of the results. However, due to limitations in experience and resources, we are currently unable to complete this step independently. We are also considering collaborating with statistical experts in future research to ensure that all necessary statistical tests are properly addressed. Thank you again for your valuable comments, and we look forward to your further guidance.

**Reply 7:** You asked about the rationale for choosing Tumor Mutational Burden (TMB) as an indicator in this study and its importance in the context of LUSC. Here we elaborate on our reasoning and analysis. Firstly, TMB refers to the number of non-synonymous mutations per megabase of tumor cells. In recent years, TMB has been considered an important biomarker for predicting the response to tumor immunotherapy. A higher TMB is associated with the generation of more neoantigens, which may increase the likelihood of the immune system recognizing and attacking tumor cells. In our study, we found significant differences in TMB between the high- and low-risk groups. Previous research has indicated that blood TMB can be used to evaluate the efficacy of camrelizumab in combination with chemotherapy in patients with advanced LUSC. During treatment, blood TMB levels are positively correlated with treatment efficacy, suggesting that higher TMB leads to better treatment outcomes and longer OS and PFS. Therefore, our selection of TMB as an indicator for this study is based on its potential value in predicting immune therapy responses and patient prognosis. Our study results are consistent with previous research, further confirming the clinical relevance of TMB in LUSC. These findings not only support the choice of TMB as a component of the prognostic model but also provide valuable directions for future immunotherapy targeting LUSC.

**Change in the text:** None

**Reply 8:** Regarding your suggestion to add a bioinformatics workflow diagram, we have had

an in-depth discussion and considered the overall structure of the article and the target audience. We understand that adding a flowchart may be helpful to some readers, but given that the article has already detailed the key steps and methods used in data analysis, we are concerned that adding a flowchart may lead to an overly lengthy article and could potentially distract readers from focusing on the research results and conclusions. Therefore, we prefer not to include a workflow diagram in the article. We hope that readers will be able to fully understand our research design and analysis process through the detailed method description in the text. Thank you again for your attention to and review of our work.

**Reply 9:** Regarding the issue you raised about the insufficient biological explanation for the association between risk scoring and immune and stromal components, we have mentioned some key points in the article, such as the role of natural killer (NK) cells and macrophages in the tumor microenvironment, and how they interact with the apoptosis process. In addition, we have also explored how immune cells and stromal cells in the tumor microenvironment affect tumor growth and metastasis by clearing apoptotic cells and promoting inflammatory responses. The tumor microenvironment is a complex network composed of tumor cells, immune cells, stromal cells, and other cell types, which interact with each other through intercellular signal transmission, collectively affecting tumor development and treatment response. Nevertheless, given the complexity of the TME, it may be difficult for us to conduct an in-depth analysis of every molecular pattern of the TME within the current scope of research and dataset. However, we are confident that our research provides a foundation for future studies in this field, and we encourage other researchers to continue to explore in depth based on our work.
**Change in the text:** None

**Reply 10:** Regarding the rationale behind the selection of specific parameters in the preprocessing steps of single-cell sequencing data and a detailed discussion on cell cluster annotation using singleR and CellMarker, we provide the following explanations and supplements:Firstly, concerning the choice of mitochondrial content and gene count thresholds, we based our decisions on existing literature and prior research experience. High mitochondrial gene expression proportions may indicate cells in a necrotic or damaged state. In our study, we set the threshold for mitochondrial gene expression proportion at 10%, aiming to exclude damaged cells and ensure the quality of cells for subsequent analysis. As for the gene count thresholds, we established both low and high thresholds to remove outliers with excessively low or high gene expression levels, which helps reduce noise and enhance the overall data quality. The chosen threshold ranges were based on considerations of cell type and tissue characteristics, as well as practices from similar studies in the field.Secondly, for cell cluster annotation, we utilized two tools: singleR and CellMarker. SingleR is an automated cell type identification method based on a reference dataset, while CellMarker provides an extensive database of marker genes for human and mouse cells. We initially used singleR for preliminary

annotation of unknown cell types and then employed marker gene information from the CellMarker database for cross-validation and refinement. This dual annotation strategy enhances the accuracy of our cell type identification process.

**Change in the text:** None

**Reply 11:** In this study, we utilized scRNA-seq data and employed pseudotime analysis to unveil the developmental trajectory and differentiation states of cells in non-small cell lung cancer (NSCLC). Pseudotime analysis is a crucial bioinformatics tool that infers the developmental order of cells based on their single-cell gene expression patterns, thereby constructing a trajectory of cellular differentiation. In this study, we used the Monocle 2 software package to perform this analysis. The biological significance of pseudotime trajectory analysis lies in its ability to help us understand tumor heterogeneity and dynamic changes in cell states at the single-cell level. Through this analysis, we could track the behaviors and functional changes of different cell subpopulations during tumor progression, as well as how these changes correlate with the biological characteristics of the tumor and patient prognosis. Specifically, we identified trajectories of different differentiation states through pseudotime analysis and localized prognostic-related genes along these trajectories, observing the variations of these genes in different cell subpopulations and within the same cell type.In the scope of this study, our goal was to construct a prognostic model for lung squamous cell carcinoma (LUSC) related to apoptosis-related genes (ARG) and predict overall survival (OS) of patients. By combining the expression patterns of prognostic-related genes and pseudotime trajectories, we not only identified key genes influencing patient prognosis but also gained deeper insights into the roles of these genes in tumor development and changes in cell states. For example, we found that BMP2 was highly expressed in epithelial cells and may play a significant role in these cells. In summary, the application of pseudotime trajectory analysis in this study is crucial for understanding cell heterogeneity in LUSC, the molecular mechanisms during tumor progression, and the roles of prognostic genes in different cell states. These analytical results not only provide more precise information for personalized treatment of LUSC but also offer potential biomarkers and targets for future research. We acknowledge that despite the significant conceptual validation value of these findings, our study still has certain limitations and requires further functional validation. We look forward to future research validating our discoveries on larger sample sizes and more diverse datasets, and exploring the potential applications of pseudotime trajectory analysis in LUSC treatment and prognosis assessment. Once again, we appreciate your review and suggestions.

**Reply 12:** Below, we will provide a detailed explanation of the assumptions, model diagnostics, and potential confounding factors associated with these methods. We employed a Cox proportional hazards regression model to identify prognostic-related genes in LUSC. The fundamental assumptions of this model include the proportionality assumption, where the

hazard ratio for different individuals remains constant over the follow-up time, and the assumption of multiplicative effects of each covariate on the hazard function. In the event of violations of the proportional hazards assumption, we would consider the use of time-dependent variables or stratified Cox models. LASSO regression was utilized for variable selection and regularization to enhance the predictive capacity and interpretability of the model. We employed LASSO regression with the purpose of reducing the number of genes in the model to limit model complexity and address overfitting. A key assumption of LASSO regression is the presence of highly correlated predictor variables, with some regression coefficients being shrunk to zero by imposing penalty terms. When using LASSO regression, we employed cross-validation to select the optimal penalty parameter, λ. In the analysis, we considered potential confounding factors such as patient age, gender, pathological staging, and treatment regimens, all of which could impact the prognosis of LUSC patients. By introducing these clinical features as covariates in the Cox regression model, we assessed their independent effects on prognosis. In the final model construction, we included only covariates that demonstrated statistical significance in the multivariable analysis. In summary, we have given careful consideration and evaluation to the assumptions, diagnostics, and potential confounding factors associated with the model during the analysis. We believe that the application of these methods enhances the reliability and effectiveness of our research results. However, we acknowledge that any statistical model has its limitations, and we will continue to explore more precise methods in future work to improve the model's predictive capabilities.

**Reviewer C**

Zhu et al. aimed to construct an apoptosis prognostic model associated with LUSC survival. Through bioinformatics analysis 4 apoptosis-related genes (ARGs) that were associated with LUSC prognosis were identified. Later on, risk-score and prognostic model construction was done. The accuracy of this model was validated in a test set, and the potential role of the risk score in guiding immunotherapy for patients with LUSC was investigated. Finally, they further elucidated the role of the prognostic-related ARGs at the cellular level in the occurrence and development of LUSC with the aim of improving the treatment outcomes and prognosis of patients with LUSC at the single-cell level.

The topic and analysis are interesting but the manuscript has a lot of flaws, has not been written with proper flow and description, figures are very poor. The manuscript needs a substantial revision and needs to be reviewed again. The more points follow as:

1) Please re-write the key findings in bullet points for more clear understanding.

2) The introduction section lacks the aims and objectives of the present study.

3) Introduction and Discussion sections lack novelty, strength, limitations, and future prospects of the present study.

4) The authors have not mentioned the details of how the DEGs analysis of ARGSs was performed. Please mention how pre-processing such as normalization, log transformation, and batch correction of data values were performed. Please refer to these papers for more clarification:
https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2022.881246/full,
https://link.springer.com/article/10.1186/s43042-023-00401-5,
https://onlinelibrary.wiley.com/doi/full/10.1002/jcb.30213.

5) Authors must validate the enrichment analysis results from other databases.

6) Results and methods sections lack clarity on DEGs analysis. Please re-write properly.

7) The resolution of all figures is very poor and needs to be improved.

8) All figure captions need to be detailed properly.

9) Please explain why the authors chose limma for DEGs analysis instead of DESeq2 for the TCGA dataset.

**Reply 1:**

Here are the key findings of this study for a clearer understanding:

- Identified apoptosis-related prognostic genes and established a risk scoring model with better predictive value than other traditional clinical parameters. Through extensive data analysis, the high predictive value of this model was reconfirmed.
- Through functional enrichment analysis of selected apoptosis differentially expressed genes, critical pathways for their functionality were determined. GO and KEGG pathway analysis indicated that in cancer, cell apoptosis is mediated through mitochondrial membrane changes, involving the BCL-2 family proteins and large molecular complexes.
- In the early stages of apoptosis initiation in mitochondria, various pro-apoptotic signaling pathways or damage pathways are activated. When these signals or pathways converge on the mitochondria, permeability of the inner and outer membranes increases, leading to the execution phase of apoptosis.
- In a comprehensive study of the immune microenvironment, significant differences were found between high and low-risk groups in terms of tumor mutation burden (TMB) and TP53. Higher TMB resulted in better treatment outcomes and longer overall survival (OS) and progression-free survival (PFS) for patients.
- Differences in gene mutations between high and low-risk groups, with TP53 gene mutations being the most common type. These findings provide strategies for researchers to further explore the anti-tumor potential of cell apoptosis and tumor immunity, offering a research direction for apoptosis treatment in squamous cell lung cancer (LUSC).
- Distribution of prognosis genes related to apoptosis in eight cell clusters, where BMP2 is significantly upregulated in epithelial cells, GPX3 has higher expression levels in smooth muscle cells, JUN is expressed in all cell subtypes but relatively higher in T cells.

- Single-cell trajectory analysis was conducted on lung cancer subpopulations, revealing temporal trends and expression profile differences in different cells, as well as biological questions related to cell development, differentiation, and function.
- Temporal analysis on two cell clusters showed the transcriptional states of cell development in these clusters.
- Explored the expression changes of prognostic genes in temporal analysis, revealing a significant decrease in JUN expression with cancer progression, while the expression of the other three genes did not show significant changes.
- In the samples, T cells constitute a significant proportion, and further investigation was conducted on the expression changes of prognostic genes in T cell temporal analysis.

**Change in the text:** None

**Reply 2:** Thank you for your valuable feedback. After careful consideration, we recognize that the introduction section indeed lacked a clear elucidation of the objectives and purposes of this study. We have incorporated these details in the revised manuscript, and we hope that these additions provide a more explicit and comprehensible overview.

**Change in the text:** We have modified our text as advised (see Page 5，line 150-159)

**Reply 3:** Thank you for your review and valuable suggestions. In response to your request, we have supplemented the introduction and discussion sections.

**Change in the text:** We have modified our text as advised (see Page 19，line 629-636)

**Reply 4:** Thank you for your attention to our research and valuable feedback. Regarding your inquiry about the details of the pre-processing steps during the DEGs analysis of ARGS, in this study, we indeed conducted standardized pre-processing procedures, including data normalization, log transformation, and batch effect correction. These steps are crucial for ensuring data quality and the accuracy of the analysis results. Specifically, we initially normalized the raw data to eliminate systematic biases between different samples. Following normalization, we employed log transformation to stabilize variance and make the data closer to a normal distribution, facilitating subsequent statistical analyses. Additionally, recognizing the potential impact of batch effects on the results, we applied appropriate batch correction methods to mitigate this influence. We believe that our analytical approach aligns broadly with the methodologies outlined in the papers you recommended, and our data pre-processing steps are sufficient to ensure the reliability of the analysis results. Furthermore, our research findings have undergone multiple validations, demonstrating the feasibility and accuracy of our model and discoveries.

**Change in the text:** None

**Reply 5:** Thank you for your valuable feedback and suggestions. Regarding the issue of

validating enrichment analysis results, we appreciate your concern for the rigor of the research. In our study, we employed methods for functional enrichment analysis that have been widely utilized and accepted in the literature. These methods have demonstrated reliability and effectiveness in previous studies, and as such, we did not conduct additional validation using different databases. However, we acknowledge that each study should strive for independent verification of the repeatability and accuracy of its results. While our research did not include validation steps from other databases, we encourage future researchers to utilize our results as a foundation and conduct validations on different databases to enhance the credibility of enrichment analysis outcomes. We agree that validating analysis results is an important aspect of scientific research. Despite our analysis being based on existing databases and methods, we encourage subsequent studies to validate on independent datasets to strengthen the universality and reliability of research findings. We sincerely hope that our response addresses your concerns. Once again, we appreciate your valuable review comments.

**Change in the text:** None

**Reply 6:** Regarding the clarity issue raised on the DEGs analysis, we have made corresponding modifications and additions to the methods section. As for the results section, given the progress of the study and the completion of data analysis, we recognize that modifying existing results can be challenging without compromising the overall study conclusions and quality. Our results have been elaborated upon, and the incorporation of figures and charts aids in providing readers with a better understanding of our findings. In future research, we will further optimize our reporting style to ensure each analytical step is presented clearly to the readers while maintaining the integrity of the study.

**Change in the text:** We have modified our text as advised (see Page 9，line 178-180)

**Reply 7:** Regarding the issue you mentioned about image resolution, I would like to clarify that all the images we submitted are in PDF format, and they were created and exported to jpg format using Adobe Illustrator (AI) software. We have double-checked the quality of the images, and they exhibit high clarity. If the image quality does not meet the standards of the journal, we are willing to re-upload them.

**Change in the text:** None

**Reply 8:** Regarding the issue you raised about the need for detailed descriptions in the figure legends, I have made corresponding modifications to some of the figure legends to provide clearer information and better understanding. For the remaining unmodified parts, I believe the existing descriptions accurately reflect the main content and key information presented in the figures, ensuring that readers can comprehend the data displayed. However, I highly value your professional advice, and if you believe that specific figure legends require further detailed descriptions, please feel free to let me know. I will carefully consider your specific suggestions

and make necessary additions and improvements to the figure legends.

**Change in the text:** We have modified our text as advised (see Page 28，line 742-747), (see Page 28，line 758-760)

**Reply 9:** Regarding the question you raised about our choice to use limma for differential analysis instead of DESeq2 in handling TCGA dataset, our response is as follows:

In our study, we opted to use the limma package for differential gene expression analysis primarily because the TCGA dataset typically provides already normalized expression data. Limma is well-suited for analyzing continuous, standardized expression data, and existing research has shown its efficacy in such analyses. By employing empirical Bayes methods and model fitting, limma can effectively handle small sample sizes, thereby enhancing statistical power.In contrast, DESeq2 is specifically designed for count data, such as raw RNA-seq read counts. DESeq2 estimates differential expression using a negative binomial distribution model and likelihood ratio-based methods, making it well-suited for handling unnormalized count data. However, when data, as provided in TCGA, has already undergone preprocessing and normalization, limma may be a more suitable choice.Additionally, the limma package offers a range of convenient functionalities, such as ease of integrating multiple types of data and subsequent analyses, factors we considered when selecting our analytical tool. Based on these considerations, we believe that limma is a reasonable choice for analyzing differential gene expression in the TCGA dataset. We hope this explanation addresses your query.

**Change in the text:** None