Innovative Journal of Pediatrics
Volume 31 · Issue 6 · e114730

OPEN PEER REVIEW

Screening Key Differentially Expressed Genes in Kawasaki Disease via Integrated Analysis

Keliang Li, Yan Jiao, Jingjing Liang, Pingping Pan, Yanji Zhu, Yilei Li

SUBMITTED	REVISED	ACCEPTED	DOI
22 Apr 2021	18 Oct 2021	20 Oct 2021	10.5812/ijp.114730

About this report

This open peer review (OPR) report shows the public comments made on this article during peer review, by the reviewers, the associate editor and the editor-in-chief, for every revision round, together with the authors' replies. Comments appear only with the consent of the people who wrote them. Reviewers who chose to sign their report are named; the others are shown as "Referee". Comments are edited for language only, and confidential remarks to the editors are not included.

More about open peer review: brieflands.com/knowledgebase/opr

Contents

1 Original submission Elsayed Abdelkreem 2
2 Revision 1 Elsayed Abdelkreem 3
Authors' reply to reviewers 4
Reviewer and editor comments 11
3 Revision 2 12
Authors' reply to reviewers 13
Reviewer and editor comments 15

1 Original submission

1 comment · 17 May 2021

Elsayed Abdelkreem Reviewer

MAJOR REVISION 17 May 2021

The authors in this study investigated differentially expressed genes associated with Kawasaki Disease (KD) through integrated bioinformatics analysis and qRT-PCR validation on a small sample. This study contributes to improving our understanding of the pathophysiology of KD and opens the door for future research to confirm these findings. This is an interesting study, but some issues must be addressed.

Major comments:

• Searching in the GEO NCBI at https://www.ncbi.nlm.nih.gov/geo using the terms Kawasaki Disease and Homo sapiens revealed 329 results (on 17 May 2021). The authors have to justify why they chose these three datasets (GSE73464, GSE109351 and GSE68004) in particular. Would including other datasets produce different results? Of note, a recently published and relevant study (Yazdan Rahmati, Soodabeh Shafiee, Sajjad Esmaeili, Identification of 7 key age-related genes involved in Kawasaki disease, an integrated study by metaDE and weighted gene co-expression network analysis(2020)Journal of Computational Chemistry & Molecular Modeling 4(2) pp:392-402) used three datasets (GSE73464, GSE18606, and GSE68004) and found different results (only one changed dataset resulted in different results!).

• The authors should discuss whether the increased/decreased expression of certain genes detected in this study could be associated with the stage of KD (acute, sub-acute, convalescent). The authors could identify the stage of KD in the seven cases included in the study, but what about the data obtained from databases? The study design couldn't allow follow-up.

• The authors should explain that their study design can't confirm a causal relationship.

• If data are available, it would be interesting if the authors could investigate the effect of some factors on the DEG, such as gender, age at onset, ethnicity, and outcome.

Minor comments

• In the abstract, all abbreviations must be spelled out when first mentioned, such as GEO, GO, KEGG, PPI, and ROC. The same applies to the manuscript.

• In the datasets collection ("After filtering, three datasets, including GSE73464, GSE109351 and GSE68004, were included in this study"), please delete "including".

• It would be better to discuss the study limitations and future research suggestions at the end of the discussion rather than the conclusion.

2

Revision 1

AUTHORS' RESPONSE TO REVIEWERS

Authors' reply to reviewers

The next pages contain the authors' point-by-point reply to the comments of this revision round, as submitted with the revised manuscript. It is published as the authors wrote it.

Dear editors and reviewers,

Many thanks for your high efficiency of work and good suggestions. We have revised the manuscript according to the comments and the recommendations suggested by the editors and reviewers. We use ‘Track changes’ model to show the part that we revised in the manuscript. In addition, in order to prevent confusion, we deleted the original figures and attached the new ones in the revised version.

The responses to the editors' and reviews' comments are as follows:

Reviewers' Comments:

Reviewer 1:

This article is well-written and I have no comment.

Response: Thanks for your high efficiency of work.

Reviewer 2:

The authors in this study investigated differentially expressed genes associated with Kawasaki Disease (KD) through integrated bioinformatics analysis and qRT-PCR validation on a small sample. This study contributes to improving our understanding of the pathophysiology of KD and open the door for future research to confirm these findings. This is an interesting study, but some issues must be addressed.

Major comments:

• Searching in the GEO NCBI at https://www.ncbi.nlm.nih.gov/geo using the terms Kawasaki Disease and Homo sapiens revealed 329 results (on 17 May 2021). Authors have to justify why did they choose these three datasets (GSE73464, GSE109351 and GSE68004) in particular? Would including other datasets produce different results? Of note, a recently published and relevant study (Yazdan Rahmati, Soodabeh Shafiee, Sajjad Esmaeili, Identification of 7 key age-related genes involved in Kawasaki disease, an integrated study by metaDE and weighted gene co-expression network analysis (2020)Journal of Computational Chemistry & Molecular Modeling 4(2) pp:392-402) used three datasets (GSE73464, GSE18606, and GSE68004) and found different results (only one changed dataset resulted in different results!).

Response: Thanks for your comment. There are indeed lots of results when searching in the

GEO NCBI at https://www.ncbi.nlm.nih.gov/geo using the terms Kawasaki Disease and Homo sapiens. But not all datasets will be included in our study. From 2016 to 2018, datasets meet the following criteria would be included in our study: (1) selected datasets should be whole-genome mRNA expression profile by array; (2) these data were derived from blood samples of patients with KD and normal controls; (3) datasets were normalized or original. After filtering, three datasets, GSE73464, GSE109351 and GSE68004, were included in this study. In addition, this is not surprising that the analysis included different datasets would produce different results, which probably due to the datasets recruited based on different platforms and different technical settings, different analysis methods and screening criteria, the different samples handled on different days, in different groups or by different people. Then, we validated the expression levels of six DEGs, including ADM, S100A12, ZNF438, MYD88, FCGR2A and FCGR3B, in GSE18606 and performed ROC analysis as well. The results revealed that these six DEGs exhibited the same patterns with our integrated analysis and except ADM (0.684) and FCGR2A (0.695), the AUC of these six DEGs, including S100A12 (0.811), ZNF438 (0.758), MYD88 (0.843) and FCGR3B (0.775), was more than 0.7 (Figure 5). These findings further demonstrate the reliability of our results.

• Authors should discuss whether the increased/decreased expression of certain genes detected in this study could be associated with the stage of KD (acute, sub-acute, convalescent)? The authors could identify the stage of KD in the seven cases included in the study, but what about the data obtained from databases? The study design couldn't allow follow-up.

Response: Thanks for your good suggestion. We totally agree that we should explore the association between the expression of genes and the stage of KD. However, there is no available clinical information (the stage of KD) for these three datasets (GSE73464, GSE109351 and GSE68004). The seven cases included in this study were all during the acute phase at the time of sampling, which has been indicated in the "Quantitative real-time polymerase chain reaction (qRT-PCR) validation" of Methods section. Then, we retrieved two datasets from the GEO database, GSE18606 and GSE63881, to explore the association between the expression of genes and the stage of KD. GSE18606 consists of 48 whole blood samples from 9 healthy age-

appropriate controls, 8 IVIG non-responding and 12 IVIG-responding patients at acute and convalescent stages, and GSE63881 consists of the acute and convalescent whole blood from 171 KD subjects. The results revealed that the expression levels of these six DEGs exhibited significantly increased trend in acute phase when compared with healthy controls in GSE18606, and exhibited significantly decreased trend in convalescent phase when compared with acute phase in GSE18606 and GSE63881. These results may indicate that the expression levels of these key genes decrease with the remission of KD symptoms. However, whether the expression of these key genes is related to the stage of KD needs to be further verified in a large sample containing different stages. In addition, as a pilot study, we are currently focusing on the key genes closely related to KD. Hence, follow-up studies were not designed.

• Authors should explain that their study design can't confirm causal-relationship.

Response: Thanks for your comment. This study design indeed can't confirm causal-relationship. As a preliminary study, this study mainly focuses on the key genes closely related to KD. As for the causal-relationship between these genes and the disease, it needs to be confirmed in larger samples and more in-depth functional studies.

• If data are available, it would be interesting, if the authors could investigate the effect of some factors on the DEG, such as gender, age at onset, ethnicity, and outcome.

Response: Thanks for your comment. We totally agree that it would be interesting to investigate the effect of some factors (such as gender, age, ethnicity, and outcome) on the DEGs. However, there is no available clinical information for these three datasets (GSE73464, GSE109351 and GSE68004). By searching the original articles corresponding to these datasets, we found the following information: GSE68004 included 76 pediatric patients and age- and sex-matched healthy controls; there were no significant difference was observed in age or gender between normal control and KD subjects in GSE109351; and no relevant information was found in GSE73464. Similarly, there were no significant differences in age (p = 0.274) and sex (p = 0.643) in GSE18606.

Minor comments

• In the abstract, all abbreviations must be spelled-out when first mentioned, such as GEO, GO, KEGG, PPI, and ROC. The same applies for the manuscript.

Response: Thanks for your good suggestion, and we have provided the full name of all abbreviations throughout our manuscript.

• In the datasets collection “After filtering, three datasets, including GSE73464, GSE109351 and GSE68004, were included in this study”. Please, delete “including”.

Response: Thanks for your good suggestion, and we have deleted “including” in this sentence.

• It would be better to discuss the study limitations and future research suggestions at the end of discussion rather than conclusion.

Response: Thanks for your good suggestion. We have moved the study limitations and future research suggestions at the end of discussion, before the conclusion.

Associate Editor's Comments:

Associate Editor 1:

Public Comment:

The authors in this study investigated differentially expressed genes associated with Kawasaki Disease (KD) through integrated bioinformatics analysis and qRT-PCR validation on a small sample. This study contributes to improving our understanding of the pathophysiology of KD and open the door for future research to confirm these findings. This is an interesting study, but some issues must be addressed.

Major comments:

• Searching in the GEO NCBI at https://www.ncbi.nlm.nih.gov/geo using the terms Kawasaki Disease and Homo sapiens revealed 329 results (on 17 May 2021). Authors have to justify why did they choose these three datasets (GSE73464, GSE109351 and GSE68004) in particular? Would including other datasets produce different results? Of note, a recently published and relevant study (Yazdan Rahmati, Soodabeh Shafiee, Sajjad Esmaeili, Identification of 7 key age-related genes involved in Kawasaki disease, an integrated study by metaDE and weighted gene co-expression network analysis (2020)Journal of Computational Chemistry & Molecular

Modeling 4(2) pp:392-402) used three datasets (GSE73464, GSE18606, and GSE68004) and found different results (only one changed dataset resulted in different results!).

Response: Thanks for your comment. There are indeed lots of results when searching in the GEO NCBI at https://www.ncbi.nlm.nih.gov/geo using the terms Kawasaki Disease and Homo sapiens. But not all datasets will be included in our study. From 2016 to 2018, datasets meet the following criteria would be included in our study: (1) selected datasets should be whole-genome mRNA expression profile by array; (2) these data were derived from blood samples of patients with KD and normal controls; (3) datasets were normalized or original. After filtering, three datasets, GSE73464, GSE109351 and GSE68004, were included in this study. In addition, this is not surprising that the analysis included different datasets would produce different results, which probably due to the datasets recruited based on different platforms and different technical settings, different analysis methods and screening criteria, the different samples handled on different days, in different groups or by different people. Then, we validated the expression levels of six DEGs, including ADM, S100A12, ZNF438, MYD88, FCGR2A and FCGR3B, in GSE18606 and performed ROC analysis as well. The results revealed that these six DEGs exhibited the same patterns with our integrated analysis and except ADM (0.684) and FCGR2A (0.695), the AUC of these six DEGs, including S100A12 (0.811), ZNF438 (0.758), MYD88 (0.843) and FCGR3B (0.775), was more than 0.7 (Figure 5). These findings further demonstrate the reliability of our results.

• Authors should discuss whether the increased/decreased expression of certain genes detected in this study could be associated with the stage of KD (acute, sub-acute, convalescent)? The authors could identify the stage of KD in the seven cases included in the study, but what about the data obtained from databases? The study design couldn't allow follow-up.

Response: Thanks for your good suggestion. We totally agree that we should explore the association between the expression of genes and the stage of KD. However, there is no available clinical information (the stage of KD) for these three datasets (GSE73464, GSE109351 and GSE68004). The seven cases included in this study were all during the acute phase at the time of sampling, which has been indicated in the "Quantitative real-time polymerase chain reaction

(qRT-PCR) validation" of Methods section. Then, we retrieved two datasets from the GEO database, GSE18606 and GSE63881, to explore the association between the expression of genes and the stage of KD. GSE18606 consists of 48 whole blood samples from 9 healthy age-appropriate controls, 8 IVIG non-responding and 12 IVIG-responding patients at acute and convalescent stages, and GSE63881 consists of the acute and convalescent whole blood from 171 KD subjects. The results revealed that the expression levels of these six DEGs exhibited significantly increased trend in acute phase when compared with healthy controls in GSE18606, and exhibited significantly decreased trend in convalescent phase when compared with acute phase in GSE18606 and GSE63881. These results may indicate that the expression levels of these key genes decrease with the remission of KD symptoms. However, whether the expression of these key genes is related to the stage of KD needs to be further verified in a large sample containing different stages. In addition, as a pilot study, we are currently focusing on the key genes closely related to KD. Hence, follow-up studies were not designed.

• Authors should explain that their study design can't confirm causal-relationship.

Response: Thanks for your comment. This study design indeed can't confirm causal-relationship. As a preliminary study, this study mainly focuses on the key genes closely related to KD. As for the causal-relationship between these genes and the disease, it needs to be confirmed in larger samples and more in-depth functional studies.

• If data are available, it would be interesting, if the authors could investigate the effect of some factors on the DEG, such as gender, age at onset, ethnicity, and outcome.

Response: Thanks for your comment. We totally agree that it would be interesting to investigate the effect of some factors (such as gender, age, ethnicity, and outcome) on the DEGs. However, there is no available clinical information for these three datasets (GSE73464, GSE109351 and GSE68004). By searching the original articles corresponding to these datasets, we found the following information: GSE68004 included 76 pediatric patients and age- and sex-matched healthy controls; there were no significant difference was observed in age or gender between normal control and KD subjects in GSE109351; and no relevant information was found in GSE73464. Similarly, there were no significant differences in age (p = 0.274) and sex (p =

0.643) in GSE18606.

Minor comments

• In the abstract, all abbreviations must be spelled-out when first mentioned, such as GEO, GO, KEGG, PPI, and ROC. The same applies for the manuscript.

Response: Thanks for your good suggestion, and we have provided the full name of all abbreviations throughout our manuscript.

• In the datasets collection “After filtering, three datasets, including GSE73464, GSE109351 and GSE68004, were included in this study”. Please, delete “including”.

Response: Thanks for your good suggestion, and we have deleted “including” in this sentence.

• It would be better to discuss the study limitations and future research suggestions at the end of discussion rather than conclusion.

Response: Thanks for your good suggestion. We have moved the study limitations and future research suggestions at the end of discussion, before the conclusion.

Thank you again for your great help and attention. I am looking forward to hearing from you about the final decision.

Best regards and wishes!

Yours sincerely,

Jingjing Liang

2 Revision 1

1 comment · 5 Oct 2021

Elsayed Abdelkreem Reviewer

MINOR REVISION 5 Oct 2021

We would like to thank the authors for their great efforts in revising this manuscript. While the authors adequately responded to most reviewers' comments (as replies to reviewers' comments), such responses should be incorporated in the main text, mostly as study limitations and future research suggestions. Specifically, these include the association with the stage of KD (no follow-up in this study) and the effect of some factors on the DEG, such as gender, age at onset, ethnicity, and outcome (which could not be investigated in this study but may be important to investigate in future studies).

3

Revision 2

AUTHORS' RESPONSE TO REVIEWERS

Authors' reply to reviewers

The next pages contain the authors' point-by-point reply to the comments of this revision round, as submitted with the revised manuscript. It is published as the authors wrote it.

Dear editors and reviewers,

Many thanks for your high efficiency of work and good suggestions. We have revised the manuscript according to the comments and the recommendations suggested by the editors and reviewers. We have highlighted all newly added sentences in the revised manuscript. The responses to the editors' and reviews' comments are as follows:

Reviewers' Comments:

Reviewer 1:

We would like to thank the authors for their great efforts in revising this manuscript. While the authors adequately responded to most reviewers' comments (as replies to reviewers' comments), such responses should be incorporated in the main text, mostly as study limitations and future research suggestions. Specifically, association with stage of KD (no follow-up in this study) and the effect of some factors on the DEG, such as gender, age at onset, ethnicity, and outcome (that could not be investigated in this study, but may be important to be investigated in future studies).

Response: Thanks for your comment. According to your good suggestion, we have incorporated the responses in our main text, including association with stage of KD and the effect of some factors on the DEGs.

Associate Editor's Comments:

Associate Editor 1:

The second round of peer review of your manuscript is completed. I now have the reports of the Referees, which suggest that your manuscript needs more correction.

Response: Thanks for your high efficiency of work. We have revised our manuscript according to the suggestions of reviewers and highlighted the part we revised in the manuscript.

EIC Decision:

Thank you for submitting your manuscript to Iran J Pediatr for consideration of publication. The second round of peer review of your manuscript is complete. According to our reviewer and associate editor comments, this version needs some minor corrections.

Response: Thanks for your high efficiency of work. We have revised our manuscript according to the suggestions of reviewers and highlighted the part we revised in the manuscript.

Thank you again for your great help and attention. I am looking forward to hearing from you about the final decision.

Best regards and wishes!
Yours sincerely.

3 Revision 2

0 comments

No public comments in this round.