Author’s response to reviews

Title: Tumor classification and biomarker discovery based on the 5'isomiR expression level

Authors:

shengqin wang (wzsqwang@gmail.com)
Zhihong Zheng (zhihongzheng@zju.edu.cn)
Peichao Chen (chenpeichao@wzu.edu.cn)
Mingjiang Wu (wmj@wzu.edu.cn)

Version: 2 Date: 28 Jan 2019

Author’s response to reviews:

Dear Catherine Rice, Linda Gummlich and Reviewers,

On behalf of my colleagues, I wish to offer our appreciation to you and the reviewer for all the work done on our paper “Tumor classification and biomarker discovery based on the 5'isomiR expression level” (Ms. No.: BCAN-D-18-02009). We now respond all comments in detail (in blue text and “#” marked) as follows and the comments made to us are in plain text.

Editor Comments:

1 - Please move the methods to follow the Background.

# Yes. We also added “methods:” in the abstract section.

2 - Please include in the Declarations an Availability of Data and Materials statement which should reflect accurately one of the formats indicated in our submission guidelines: https://www.biomedcentral.com/getpublished/editorial-policies#availability+of+data+and+materials

# Yes. We added “All data generated or analysed during this study are included in this published article”.

3 - In the Funding section of the Declarations please indicate the role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the
manuscript. If no specific funding was received for this study, please clearly indicate this in the Funding section.

# Yes. We added “The funders played no role in study design, data collection, analysis or interpretation of the data” at the end of Funding section.

4 - At this stage, please upload your manuscript as a single, final, clean version that does not contain any tracked changes, comments, highlights, strikethrough or text in different colours. All relevant tables and figures should also be clean versions. Figures (and additional files) should remain uploaded as separate files. Should you wish to respond to these revision requests, please include the information in the designated input box only.

# Yes.

BMC Cancer operates a policy of open peer review, which means that you will be able to see the names of the reviewers who provided the reports via the online peer review system. We encourage you to also view the reports there, via the action links on the left-hand side of the page, to see the names of the reviewers.

Reviewer reports:

Hong-Qiang Wang (Reviewer 1): N/A

# Thank you for your kind review!

Wenying Yan (Reviewer 2): The current version has revised according my concerns, so my recommendation is accepted.

# Thank you for your kind review!

Russell S. Schwartz (Reviewer 3): The authors have been fairly responsive to the critiques and the paper has been improved in the latest version. They provide a reasonable justification for focusing on the task of classifying cancer types, given the data available to them. They likewise provide a reasonable defense based on prior literature for the use of the chosen wrapper method for feature selection, although I do maintain that a comparison to alternatives on this dataset would be beneficial. They provide updated results from random forests in comparison to support vector machines. They provide some response to the question of how isomiR classification results compare to those from other data such as mRNA expression, although I again think the paper would be improved by actually applying the same classification algorithm to the corresponding TCGA mRNA expression data and comparing results. They clarify a number of prior points of confusion, showing that they have reasonably handled some methodological
issues for which that was previously unclear, and have added a requested flowchart. The clarification suggests that they handled the term enrichment analysis reasonably, even if the end results are not greatly enlightening. Given all of this, I believe the paper is clear and technically sound even if the analysis could be taken further in some respects.

# Thank you for your kind review!

I have no new substantive points to raise, although there are a number of minor text errors in need of correction. The paper would benefit from a thorough proofreading, but I noted the following specific errors:

# Thanks. Please find below a point by point response to each comment.

p. 2: ``and the cell-cell adhesion'' should be ``and cell-cell adhesion''

# Revised. Abstract section, line 19, page 2, and Results section, line 24, page 9.

p. 3: ``integrate the comprehensive'' should be ``integrate comprehensive''

# Revised. Background section, line 3, page 3.

p. 3: ``good at deal with data'' should be ``good at dealing with data''

# Revised. Background section, line 24, page 3.

p. 3: ``which contained'' should be ``which contain''

# Revised. Background section, line 24, page 3.

p. 4: ``by combined the machine learning algorithm'' should be perhaps ``the combined machine learning algorithm''

# Revised. Background section, line 3, page 4.

p. 4: ``Besides the machine learning algorithms, such as SVM, SR, SRC, RF, KNN, has'' should perhaps be ``Beside the machine learning algorithms --- such as SVM, SR, SRC, RF, and KNN --- have''. Also, each acronym should be spelled out the first time it is used (e.g., support vector machine (SVM))
# Revised. Background section, line 9, page 4.

p. 4: ``GA is based on the Charles Darwin's theory of natural evolution" would read better as ``GA is based on Darwin's theory of natural selection"

# Revised. Background section, line 13, page 4.

p. 4: ``evolution, the GA has proven" should be ``evolution. GA has proven"

# Revised. Background section, line 13, page 4.

p. 4: ``GA are frequently" should be either ``GA is frequently" or ``GAs are frequently"

# Revised. Background section, line 15, page 4.

p. 4: ``Furthermore, the 5' isomiR may" should be ``Furthermore, 5' isomiRs may"

# Revised. Background section, line 20, page 4.

p. 4: ``with their canonical miRNA" should be ``with their canonical miRNAs"

# Revised. Background section, line 21, page 4.

p. 5: ``will be left only one" should be ``will be left with only one"

# Revised. Background section, line 2, page 5.

p. 5: ``300 generations GA" should be ``300 generations of GA"

# Revised. Results section, line 13, page 7.

p. 5: ``each of which comprised of 50" should be ``each of which is comprised of 50"

# Revised. Results section, line 14, page 7.

p. 5: ``the average sensitivity were" should be ``the average sensitivities were"
# Revised. Results section, line 15, page 7.

p. 8: ``samples from some cancer that'' should be ``samples from some cancers that''

# Revised. Discussion section, line 13, page 10.

p. 8: ``It is note that'' should be ``It is noted that''

# Revised. Discussion section, line 17, page 10.

p. 8: ``whether the detected isomiRs are the tissue-specific miRNA'' should be ``whether the detected isomiRs correspond to the tissue-specific miRNA''

# Revised. Discussion section, line 21, page 10.

p. 9: ``are also list in'' should be ``are also listed in''

# Revised. Discussion section, line 17, page 11.

p. 11: ``300 "generation" of GA/RF'' should be ``300 "generations" of GA/RF''


Figure 1: I am unclear what is meant by ``digible''

# It was revised as “eligible”.

Figure 1: ``Combing'' should be ``Combining''

# Revised.

Figure 1: ``Chromosome :50'' should be ``Chromosome: 50'' and would look better in the same font size as the text below it

# Revised.

Figure 1: ``Functional enrichment Analysis'' should be ``Functional enrichment analysis'' to be consistent with other capitalization in the figure
# Revised.

# And we also revised some other minor test errors as below.

“may” was revised as “might”. Background section, line 4, page 5, Background section, line 20, page 2, and Conclusions section, line 7, page 12.

“suggested GA/RF model” was revised as “suggested that GA/RF model”. Background section, line 21, page 2, and Conclusions section, line 8, page 12.

“was selected” was revised as “were selected”. Methods section, line 4, page 7.

“genes” was revised as “items”. Results section, line 9, page 8.

“the the activity” was revised as “the activity”. Results section, line 24, page 9.

“detected by the presence” was revised as “detected by the method using the presence”. Discussion section, line 16, page 11.

# Again, we would like to thank the reviewer and editors for their valuable suggestions and comments, which have led to the improvement of review articles.