

I'd like to thank reviewers for providing comments on my WQE work. I have revised my paper according to reviewers' comments. Below are my responses to faculty comments:

*1. The methods section is difficult to understand, and many technical details necessary to understand the methods are lacking.*

I addressed this general concern regarding Methods section in following ways:

First, for the description of adopted methods, like 'mutation matrix smoothing' or 'consensus clustering', I've now rewrote the entire paragraph to elaborate the details, including how it works, the reason of using it, and the parameters I used.

Second, for the evaluation metrics I used in this paper, like Rand Index/Adjusted Rand Index, Cophenetic correlation coefficient, etc. now are described thoroughly by the not only the definition but also what does it means related to this paper.

Third, I fixed some definition of mathematic notations in the text.

Last, the use of terms is unified.

*2. Every mathematical notation used should be clearly defined before the use.*

Fixed now. Sorry about missing some of definitions of mathematical notations. Now I have revised the two sections in Methods describing the proposed methods, to make sure the definitions of notations are adequately provided.

*3. For every method cited and used, please clearly explain their use and how they work. For example, the mutation matrix smoothing is unclear. Why is it good? How does it work? What is the output like?*

Thanks reviewers for bringing up this issue and providing guidance. For mutation matrix smoothing, I was hesitated to spend too much space on it since it was not the focus of this work and I just merely cited and used it. But I do agree it is necessary to provide a well-rounded description if the method was used in my work. Thus, now I have completely revised the 'mutation matrix smoothing' section and other sections like 'consensus clustering' under Methods. The cited and used techniques/pipelines are elaborated clearly with the explanation why this is used, how it works, what is the parameters (if applicable) and outputs.

*4. Jargons should be avoided as much as possible. The paper should be understandable to a large group of audience who work on genomic data.*

I was trying to avoid jargons in this manuscript but as reviewer pointed out, there are still some left causing difficulties for large group of audience to understand this work. So in the revised version the changes have been made throughout the paper, including terms in Methods section, Results section and even in figure legends. Words like 'H' or 'W' matrix, 'binary matrix', 'inputs' were replaced by more concrete terms like 'gene-patient matrix'; measurements 'ARI'

Yang Pan  
UID: 504567911

or 'Cophenetic correlation coefficient' were annotated by term reflecting their actual meaning, like 'accuracy (ARI)' or 'stability (Cophenetic correlation coefficient)'.

*5. Please define the evaluation metrics clearly before using them in figures.*

I've now provided clear and direct sentences in the Result section linking evaluation metrics to their definition, so that readers can understand figures more easily. I also revised figure legends, instead of using 'ARI' or 'Cophenetic coefficient', I use accuracy (ARI) or 'stability (Cophenetic coefficient)' to make it easier to understand for a large group of audience.