

Replies to Reviewer 2

In the present paper, the authors propose, an integrated analysis by applying a previously proposed tensor decomposition (TD)-based unsupervised feature extraction (FE) method with standard deviation (SD) optimization to microarray-measured gene expression data for diabetes mellitus from multiple tissues. The authors also state that "...found that applying the TD-based unsupervised FE with SD optimization to gene expression profiles from individual tissues can identify diseases associated with diabetes that cannot be identified by the other state-of-the-art methods". The present work is interesting in its conception, yet several issues should be addressed.

In lines 14-15 the authors state "...are singular value matrices and orthogonal matrices". Please explain, why are the described tensors orthogonal.

> Thank you very much for your comments. This is because of algorithm. HOSVD results in orthogonal matrices. In order to clarify this point, we have briefly described the algorithm of HOSVD as a subsection entitled "The brief description of the HOSVD algorithm".

In lines 32-33 it states "Each of these profiles is composed of gene expression profiles from five individuals seen at two hospitals, and they include data for three tissues" The sentence is very confusing. GEO datasets concern the species *Rattus Norvegicus*. How is a hospital and individuals involved?

> Thank you very much for your comments. It was wrongly copied and pasted from the description in the draft of the previous paper. Hospitals are strains.

How are the differentially expressed genes derived?

> Thank you very much for your comments. This is represented by selected u_{11j} , u_{12k} , u_{13m} , and u_{14s} . We added the following sentence "The genes selected are presumed to be coincident with the profiles associated with selected u_{11j} , u_{12k} , u_{13m} , u_{14s} ; this corresponds to the identification of differentially 82 expressed genes."

the p-value adjustment is the method, and how is the amount of data series taken into account?

> Thank you very much for your comments. We are not very sure what the reviewer would like to ask. The point is considered by correction by BH criterion. (Multiple comparisons correction).

How was ttest applied to three sets of data?

> Thank you very much for your comments. We have applied t test to each of three data sets as shown in Table 5.

Figure 1, makes no sense. What does it show, what are the lines and what do they represent? the figure should be better explained.

> Thank you very much for your comments. I have added more description in caption.

In addition, the gene annotation for pathway enrichment makes also no sense since there is no previous information on how the DE genes were derived.

> Thank you very much for your comments. I hope that it makes sense now since we denoted about DEG.

Overall, my general comment to the authors is that their manuscript is very difficult to understand and follow. The use of tensors in gene expression surely looks fancy and interesting, yet it pointless if a reader, especially from the biological discipline has no use of it, because it is impossible to understand. I am myself aware of tensors and its uses as well as their mathematical formulation, yet it was impossible to follow the paper.

Some background should have been given to help the reader connect the dots; why use tensors, what is the benefit and how can they be applied to gene expression data?

> Thank you very much for your comments. We have added some description as follows "There are multiple benefits in using TD to identify DEGs. First, since it is not a supervised method, it can select DEGs that are biologically more plausible than those selected using supervised methods. This can be explained using the following example wherein the aim is to identify DEGs that are distinct between two classes, e.g., patients and healthy controls. Supervised methods attempt to identify DEGs associated with a smaller divergence within individual classes, whereas TD allows to select DEGs with within-class divergence to some extent (since TD tries to identify the representative state of distinction between two classes). If the representative state is associated with within-class divergence that has biological origins, e.g., age and sex, this divergence should not be penalized. However, supervised methods often do so whereas the unsupervised method allows biological within-class divergence. Second, TD can select more stable DEGs; i.e., those independent of specific sets of samples considered in the analysis. This is because TD attempts to identify DEGs coincident with those of the representative state, which should be robust. Since sub-sampling does not change the representative state drastically, the gene set selected by TD is not altered drastically either. Third, TD can deal with multiple conditions. For example, if gene expression is measured in various tissues of several people, it is natural to format them as gene $\times$ person $\times$ tissue, which results in a tensor form. We have listed only a few important advantages here. Readers interested in acquiring information on other advantages of TD can refer to our recent book [1]."

Finally, the least the authors should do is compare their results with more "classical" approaches and compare the efficiency of their algorithm. There is a plethora of methods for the comparison of microarray data and one would suffice to prove their concept.

> Thank you very much for your comments. In addition to t test, sam and limma that we have done in the previous version, we added the integrated analysis using comBat.