Review of: "Connexinplexity: The spatial and temporal expression of connexin genes during vertebrate organogenesis"

Wen Wang

Potential competing interests: The author(s) declared that no potential competing interests exist.

The authors analyzed gene expression patterns for the *connexin* gene family in zebrafish under the scope of single-cell resolution during organogenesis. The authors rescued the lacking of expression information in their previously published sing-cell RNA-seq data by checking the gene expression signatures of the connexins genes. It is a somewhat special angle to combine the evolution of gene family with organogenesis during development. However, I have several concerns of the study, including the methods that they used to justify the gene homologous relationship with human, and the methods that they used to integrate single-cell RNA-seq data of three development samples. Please see my detailed comments in the following paragraphs:

1. The novelties in this study are not very obvious given that the single cell data had been published and the current study only picked the connexin genes to look at while the orthology between zebra fish's and human's copies were not well defined.

2. Which kind of sequences did the authors use for reciprocal BLAST? It seems that the authors only use the similarity of genes between zebrafish and human to judge the homology. The synteny relationship of these Connexin genes between zebrafish and human should also be considered to further verify the orthologous relationship.

3. In Fig. Aii, it seems that the three samples were simply merged together but not integrated by removing batch effects. For example, clusters annotated to thymus, neural crest and spleen only emerged at 5 dpf, but not 1 dpf and 2 dpf. How did you integrate single-cell RNA-seq data of development samples? I didn't find the description in the Methods.

4. The author used lengthened 3'UTRs to obtain the full transcripts of Connexin-encoding genes. What 's the criterion of the extended length? If it was based on the bulk RNA-seq data, how did you deal the different isoforms/transcripts belonged to the same genes?

In the section of "Fluorescent RNA in-situ", the tissue hybridized with the 1 x probe overnight at 40C.

"40C" should be "4 ℃"?