Reviewer's report

Title: NeuroGeM, a knowledgebase of genetic modifiers in neurodegenerative diseases

Version: 1 Date: 2 October 2013

Reviewer: Sheng Zhang

Reviewer's report:

As the authors stated, this work described a first reported searchable database (NeuroGEM) for the large number of toxicity and aggregation modifiers on several neurodegenerative diseases from a number of existing studies in different model organisms. Given the complexity of the earlier studies, whose results sometimes are confusing and conflicting, the availability of this database is informative and timely to the field.

Overall the NeuroGEM database is well-designed and relatively easy to navigate, and contains appropriate amount of information as well as many useful links. The feature that allows a user to quickly compile and compare the effect of a given gene on different disease models in different species is especially useful. The manuscript is clearly written.

Overall the study is well-designed and should be of great interest to neurodegenerative field, and fit the publication criteria of the BMC Genomics.

Revision/clarification on the following areas will help improve the manuscript and its application.

- Major Compulsory Revisions

1. Page 5, line 2: “allow users to upload their own data”:
   This could be a very useful feature for the other investigators on new studies. However, the manuscript did not elaborate on this at all and no example was given. A more detailed description of this function, with example, should be provided. For example, will this feature allow another user to upload his screen data set and quickly compare his hits with that from existing studies?

2. A clear definition of “homologous genes and orthologs” for modifiers in different species.
   For a given gene, several homologues are often found in another species (e.g., fly DnaJ-1 and Hsf and its homologues in mouse and human, as in Table 3). It will be informative to provide more details on their sequence similarity at amino acid level, so that it will be easier for users to make their own judgment as to the significance of a homology.

3. Systematic comparison of data set from different species.
The meta-analysis presented is interesting. However, unless I missed something, otherwise the analyses were mostly on data from the same species such as the fly or worm only. It is not clear whether the authors have compared the data set from different species, such as the screen results between the yeast, worm and the fly. It will be very interesting to make such comparison and find out to what extent the modifiers on the same disease genes from different model organisms are conserved, and whether the effect of a conserved modifier on a disease gene is the same across species.

This question is also, in some way, related to the comment above as to a better detailed definition of “homologous genes” in different species.

4. Reference in Table 3.
References for the studies that isolated these fly modifiers should be provided. Table 3 only listed the references for “Evidence” but not original fly studies.

- Discretionary Revisions

1. Page 10, paragraph 2, “Below the report summary, detailed information of the gene is displayed on the left side (Figure 3b)… the left half of each node (protein) is colored according to the evidence for it being a modifier. The right side of each node is colored according to the………”, on the layout information of a displayed gene.

It will be helpful to the reader and NeuroGEM users by including a brief description of the above coding system in the figure legend and also on its web link.

2. Page 13, sentence “that modifiers across species are often involved in protein folding (Figure 4a). However, they account for only 3% of all genetic modifiers, suggesting that researchers should focus their efforts more on genes involved in cell cycle or splicing, biological processes that are also often enriched in modifiers.”

It is not clear of the link between “However, they account for only 3% of all genetic modifiers” and the conclusion that “researchers should focus their efforts more on genes involved in cell cycle or splicing, biological processes that are also often enriched in modifiers.” A better clarification of this sentence will be helpful.

3. In supplemental, the paragraph on HSPH1 (HSP110)
-It should be noted that the fly HSP110 is also named as dhsp110 in one of the Drosophila study.
- It should also be noted that recent studies have shown that Hsp110 does not act as classic chaperone, but as a co-chaperone, or more specifically, a nucleotide exchange factor (NEF) for Hsp70 chaperones to catalyze the replacement of ADP-bound Hsp70 with ATP, thus completing the chaperone cycle.
- On the sentence “….Hsp70, which is the most prominent modifier family”. Is this statement accurate?

4. In supplemental, the paragraph on SEC61A1 and A2: “SEC61A1 and A2 are orthologous genes of Drosophila’s Sec61alpha and components of SEC61 complex. The ER-associated degradation process (ERAD) ensures that misfolded polypeptides are retro-translocated to the cytosol for proteasomal degradation. The SEC61 complex is involved in the translocation of polypeptides across the ER membrane; thus SEC61A1 and A2 could be implicated in SCA3 [70, 71].”

- It is not clear from the above statement why “SEC61A1 and A2 could be implicated in SCA3”. A more detailed explanation of the refereed literature would be helpful.

**Level of interest**: An article of importance in its field

**Quality of written English**: Acceptable

**Statistical review**: No, the manuscript does not need to be seen by a statistician.