We have attached our detailed responses to the comments and reviews to our paper, titled 'Tissue Enrichment Analysis for *C. elegans* Genomics'. We hope our answers are satisfactory.

1 Editor's Comment

1: My own comment is that you may want to do a second check in the literature for previous ontology-trimming methods. For example, Garrido et al. Different approaches to build brief ontologies (Volume 348 of the series Communications in Computer and Information Science pp 232-246, Springer, 2013) presents one such methodology.

Reply: We thank Dr. Setubal for pointing us towards the book chapter by Garrido *et al*. We have included the citation in the text for completeness. Originally we had not included it because the book chapter in question, like much of the ontology pruning literature, explains methods for generation of a complete ontological structure that is a logical subset of a greater ontology. Our approach does not generate an ontological structure, but rather identifies terms in the ontology that are balanced with regards to specificity (tree depth) and evidence (gene annotations). Once the terms are selected, these terms are not rebuilt into a formal 'brief ontology'. However, we recognize the similarity of our work with the Computer Science literature and have included this citation as suggested as a consequence.

2 Review 1

1: Authors havent discussed how often the background database for the enrichment analysis will be updated. Every time, when new GO terms are added by the GO consortium does the trimming algorithm takes into account these terms and refine the method?

Reply: The background dataset for the enrichment analysis will be updated with every version of WormBase every two months. As part of a new Worm-Base release, the tissue ontology is annotated with new gene expression data, and our trimming pipeline is used to generate a new background dataset. Due to the tight integration of this tool with WormBase, we can guarantee that our database will continuously remain up to date, which is a drawback other popular tools very often suffer from.

We have updated the text to reflect this by adding the following text at the end of the subsection **Filtering greatly reduces the number of nodes used for analysis**:

"Our trimming pipeline is applied as part of each new WormBase release. This ensures that the ontology database we are using remains up-to-date with regards to both addition or removal of specific terms as well as with regard to gene expression annotations."

2: I am skeptical about the usage of this tool as it is only confined and tested with C. elegans *genome. Such tools should be always part of the big enrichment analysis tool in the form of plugins rather than as standalone.*
For example, the recently published 'FunRich', a functional enrichment analysis tools has an option of using custom database. Using this custom database, any species genome can be used as background database for functional enrichment analysis which also serves the purpose of identifying GO terms enriched in the input gene lists.

Reply: While we understand the concerns regarding the usage of standalone software, we would like to point out that the cell and tissue ontology referred to in the text has been developed independently of the GO. The ontology was developed and is maintained, annotated and extended independently by WormBase. Moreover, due to the species-specific nature of cell and tissue ontologies, at this moment we cannot allow users to input any species genome as a background database. To do this, we would require cell and tissue ontologies for every species in question (available only for some species); ortholog mapping between species (available) and ontology mapping between species (non-existent).

While we do intend to expand TEA to major model systems in the future, we cannot follow the approach that PANTHER takes of identifying orthologs in a background set and using pre-existing GO annotations to identify enrichment. Doing this would yield uninterpretable results, for example if a researcher input a Drosophila background and enrichment set, he/she would at best receive a list of C. elegans tissues. To add tissue enrichment analysis for other species in the future, more cell and tissue ontologies must be developed. Other tissue ontologies do exist at this point in time, such as a Zebrafish anatomy ontology.

3: Majority of the figures is slightly distorted and needs to be replaced with high resolution clear images.

Reply: We have remade all of the figures using vector-formatting and saved

each figure as a PDF for maximum resolution. We thank the reviewer for pointing this out.

4: Citing FunRich tool mainly used for the functional enrichment analysis is worth mentioning in the Background section of the manuscript.

Reply: We have cited 'FunRich' and thank reviewer 1 for making us aware of this tool.

5: In the web GUI (http://www.wormbase.org/tools/enrichment/tea/tea.cgi), please include standard list of C.elegans gene names by default in the search box.

Reply: We have added example genes in the search box.

6: In the results table, include Hypergeometric p-values in addition to q-values.

Reply: We have added p-values to the results table. However, we feel that these values add no relevant information for the user, since the table only contains statistically significant results as assessed by a q-value cutoff of 0.1.

3 Review 2

1: Introduction needs to be re-organized and partially re-written because it feels a bit strange that details about pruning and trimming are discussed in the last sentences. Instead, the authors need to outline the ideas behind the software and its implementation. A brief outline of what was achieved with the development of this software should also be included.

Reply: We have re-written the introduction to reflect this. At the end of our section titled **Background**, we have included the following paragraph:

"We have developed a tool that tests a user-provided list of genes for term enrichment using a nematode-specific tissue ontology. This ontology, which is not a module of Gene Ontology, is verbose. We trim our ontology using an algorithmic approach, outlined below, that reduces multiple hypothesis testing issues by limiting testing to terms that are well-annotated. The results are provided to the user in a GUI that includes a table of results and an automatically generated bar-chart. This software addresses a previously unmet need in the *C. elegans* community for a tool that reliably and specifically links gene expression with changes in specific cells, organs or tissues in the worm."

2: The sources of data need to be better described. This may be obvious for the worm researchers but may be harder to understand for those outside.

Reply: We would like to thank Reviewer 2 for pointing this out to us, as it had not ocurred to us that non-worm researchers may find the sources of data confusing. We have added the following paragraph explaining the sources of data more extensively in our section **Reducing term redundancy through a similarity metric**:

"For our tool, we employ a previously generated cell and tissue ontology for *C. elegans*[6], which is maintained and curated by WormBase."

3: The significance and the way to determine the parameter S in the 'avg' vs 'any' is pretty hard to understand. More illustrations and explanations need to be provided

Reply: We have expanded our verbal explanation of the similarity criterions to make it more approachable to readers. We did not include additional figures because we did not know how to graphically show the difference between a threshold based on an average score and one based on a supremum score. We added the following text to the subsection **Reducing term redundancy through a similarity metric**:

"Intuitively, a set of sisters can be considered very similar if they share most gene annotations. Within a given set of sisters, we can calculate a similarity score for a single node by counting the number of unique annotations it contains and dividing by the total number of unique annotations in the sister set. Having assigned to each sister a similarity score, we can identify the **average** similarity score for this set of sisters, and if this average value exceeds a threshold, all of the sisters are removed from the ontology. An alternative method is check whether **any** of the scores exceeds a predetermined threshold, and if so remove this sister set from the ontology. We referred to these two scoring criteria as '**avg**' and '**any**' respectively."

4: Hypergeometric tests. What was the benefit of setting alpha to 0.1 vs 0.05 or another value? Did you find this using some of the established datasets?

Reply: An alpha of 0.1 is standard for many applications that employ FDR. We didnt fine-tune this parameter cutoff. We have added a sentence to explicitly point this out in the text, and we have also made it clear that this parameter can be tuned by users in the command-line, but not the web, version.