Charles A Whittaker, Ph.D.
Koch Institute at MIT
77 Massachusetts Avenue 76-189B
Cambridge, MA 02139
charliew@mit.edu
Tel: 617-324-0337
Fax: 617-324-2238

Ron Orlando, Ph.D.
Complex Carbohydrate Research Center
The University of Georgia
315 Riverbend Road
Athens, Georgia 30602-4712
rorlando@uga.edu
Tel: 706-542-4429
Fax: 706-542-4412

13 November 2022

Dear Dr. Orlando,

We are writing to submit the revised version of our article titled "Functional Annotation Routines Used by ABRF Bioinformatics Core Facilities - Observations, Comparisons and Considerations". We are grateful to the reviewers for their suggestions and comments. Our article is substantially improved because of their efforts.

The material that follows this letter contains point-by-point responses to the reviewers and a catalog of the changes that we have made.

Sincerely,

Charles A. Whittaker (on behalf of GBIRG)

Reorganizing for clarity. Both reviewers suggest editing for improved organization clarity. We agree, some of the specific reviewer comments address this point and our revisions are discussed in detail below. In summary, we have:

1) Edited the section title "Concept Mapping" beginning on page 13 to improve the wording.
2) The results section covering the experimental datasets was edited to remove information redundant with the methods section.
3) We added Table 1 to summarize the datasets used in the study
4) We added Supplemental Figure 1 that provides a graphical representation of our workflow and referenced that figure at the start of the manuscript.

Reviewer 1 - comments to be addressed

Need improvement on clarity and overall organization of the manuscript to make it easier to follow.

See Above

The reviewer would like that the authors:

1) add some more real data examples; and

We have adapted the Introduction and Discussion sections of the manuscript to expand our consideration of the functional annotation literature. Specifically, we expanded the discussion of Xie et al., 2021 highlighting their analysis approach and results and emphasizing their shiny app resource that summarizes tool and benchmarking papers. This resource provides readers with an extensive listing of the field. We also added a discussion of Wijesooriya et al., 2022, a paper that documents misuse of functional annotation tools in the literature. In our discussion, we have added a section that compares our study to examples of large-scale and sophisticated comparative studies (Tarca et al., 2015, Geistlinger et al., 2021 and Ihnatova et al., 2018). Finally, Timmons et al., 2015 described different types of bias that can impact functional annotation, particularly in the context of background set size and added reference to this work in our discussion section.

2) improve the clarity of the manuscript to make it easier to follow.

See Above

Reviewer 2 - comments to be addressed

The presentation of some of the study details could be improved somewhat – see specific comments below.

See Above

In section "Materials and Methods – Data processing and Differential expression analysis":

This section is a bit confusing. There is no explanation of what the various labels (G1vRev, EvC, etc) are. I assumed that they are names of the four experiments being studied, but it would be helpful to state that explicitly. Even better would be to present the tables showing the list of all datasets (with a one-line description of what they are) and the list of programs tested at this point of the paper, instead of further down.

We have added Table 1 to list the datasets considered in our study and shifted the original Table 1 to Table 2 in the revised manuscripts. References to both these tables were added to the methods at the start of the relevant sections.

The sentence "V4ersion 1.24.0 and apeglm log fold change shrinkage was used for G1vRev, EvC and hpm3AvC and version 1.16.0 and normal log fold change shrinkage was used for SvC" is unclear. Apart from the typo (V4ersion), do the authors mean that they used two different versions of DESeq2 for their experiments? If so, why?

The "V4ersion" typo has been repaired. Yes, a different version of DESeq2 was used for the SvC experiment. Our study combined analyses done by different authors at different times and because the SvC differential expression workflow closely matched the analyses done for other experiments and we used this processing in our manuscript and reported the version difference.

Page 10:

The sentence starting with "The biological replicates..." is not clear. In "it compares", it seems that "it" refers to "the difference", but it sounds weird to say that a difference compares two things. Also, in the same sentence, it's not obvious that comparing a mixed population with one enriched for a specific cell type should lead to a small difference between the two conditions. This is explained better further down the paper, so maybe there could be a reference to that in the text.

The information presented in the sentence discussed in this comment is partially redundant with both the previous sentence and the subsequent section referenced by the reviewer. Its inclusion does not add to the presentation and is also unclear. It has been removed and this and other sections covering the 4 experiments have been edited for consistency and clarity.