April 6, 2013

Dear Dr. Cooke and Reviewers,

Thank you all for your comments on our manuscript. We greatly appreciate your input and we have revised the manuscript in accordance with your suggestions. We re-analyzed our data using NSAF, as suggested by Reviewer 2. We have also worked to clarify the methods and expand on the discussion as requested by the reviewers. Lastly, language has been added to the manuscript to clarify how this study relates to the field of conservation physiology.

Below are our responses to each of the editor’s and reviewers’ comments. The original comments are provided after “COMMENT” and our reply follows directly below. We have supplied the line numbers for where the revisions can be found in the revised manuscript.

Again, thank you for your consideration of our manuscript and for your comments. We look forward to further correspondence regarding our revisions.

Sincerely,

Emma Timmins-Schiffman

**Editor’s Comments**

*COMMENT: I will also add that it is not intuitive from the introduction or conclusion as to how the data relate to conservation physiology.  I will email you a proof of our introductory paper and I would ask that you give it a read and consider how your work fits into our vision for the discipline and journal.  Doing so will be of benefit to the readers.*

REPLY: Thank you for sending me the article describing the meaning of “conservation physiology” for this journal. We have added text clarifying how our work fits into this definition in both the introduction [revision version lines 141-142, 176-178] and discussion [revised version lines 407-411].

**Reviewer: 1**
*COMMENT:*  *The author states clearly that this trial analyses the proteins within the gill tissue and the reason for this should be stated in the introduction (there is a brief mention in the abstract).*

REPLY: This information has been added to the introduction: “The gill is the interface between bivalves and their environment, necessitating that the tissue performs a variety of physiological functions in response to the environment (e.g. David et al. 2007; Wang et al. 2010). The identification of proteins that are expressed in gill tissue supports the development of tools can help to guide future research on the molecular physiology of molluscs faced with stresses such as climate change and disease.” [revised version lines 173-178]

*COMMENT: In addition, the use of other tissue or haemolymph should not be discounted as, although the gills are at the organism/environmental interface, there is the possibility that proteins may be present elsewhere and not in the gills.  A sentence relating to this should be added to the discussion where other experimental design and target tissue is discussed (original version lines 324-338).*

REPLY: On further investigation into proteomics studies on oyster hemolymph we did find that other studies report a larger number of proteins identified and these references have been included [revised version lines 356-359]. The authors of the hemolymph proteomics paper (Muralidharan et al.) state in their own paper that the reason for identifying relatively few proteins is due to limited cell types in hemolymph and limited genomic resources, so we have included this explanation in our discussion since we believe it is relevant [revised version lines 360-361].

**Reviewer: 2**
*COMMENT: Proteins abundance was compared using the spectral counts as a proxy for relative expression. This is not accurate. The data should be re-analyzed using the Normalized spectral abundance factor (NSAF) as demonstrated previously (Zhu et al., 2010. J Biomedicine and Biotechnology). Then, to consider whether NSAF are consistent across biological replicates, the log(NSAF) data should be compared and a R2 can be calculated for each comparison.*
*For each sample, the calculated NSAF should be provided in Supplementary table 1.*

REPLY: Thank you for bringing our attention to the application of NSAF in this study. Analysis using NSAF has been added to the methods [revised version lines 266-270]. We have re-analyzed the expression values of the data using NSAF for each oyster and the data are included in Supplementary Data 4. The correlations between oysters with R2 values are visualized in Supplementary Data 1.

*COMMENT: lines 247: A minimum of four total tandem mass spectral assignments in the combined technical replicates were used... A protein can be considered identified only when at least 2 unique peptides have been characterized from the same sample. If that is not the case, the protein identification is not reliable. Please review your dataset in order to conserve only proteins for which two unique peptides have been identified in at least one of the technical and biological replicates.*

REPLY: We have re-analyzed the data as suggested by the reviewer. A minimum of 2 unique peptide hits had to be present in a technical replicate for a protein to be included in the analysis.. Additionally, across the entire data set, only the proteins that had at least 4 total spectral counts were maintained for further analysis. [revised version lines 255-257]

*COMMENT: [In the Discussion] the authors should further discuss the different proteins that were found in abundance in gills and their potential role. Over 1500 proteins were identified but the function of only about 10 of them was discussed.
A comparison with transcriptomic data such as gills-enriched Pacific oyster genes could also provide interesting information (Dheilly et al., 2011 BMC Genomics).*

REPLY: We agree that the discussion could use more breadth and depth in terms of discussing the proteins identified in this study. We have added a comparison to the Dheilly et al. 2011 manuscript [revised version lines 374-377]. We have also added more explanations regarding the proteins identified and their significance [revised version lines 377-390].

*COMMENT: [original version] lines 147-156: numerous comparative proteomics studies have been carried out in other labs and should be acknowledge. For example, I know of David Raftos lab studies:
Thompson et al., 2012 Environn pollut
Thompson et al., 2012 Aquat toxicol
Thompson et al., 2011 Aquat toxicol
Simonian et al., 2009 J. proteomics
Your introduction implies that you made an exhaustive list of all proteomics studies on oysters but many are lacking. Either re-write or complete the list.*

REPLY: Thank you for bringing these relevant papers to our attention. We have included them in the introduction [revised version lines 154-156].

*COMMENT: lines 165-168: There are examples where this approach has been implemented for marine invertebrates for which the genome has been sequenced. Examples: Dheilly et al., 2012 Dev comp immunol and Dheilly et al., 2013 Dev Comp Immunol.*

REPLY: Again, thank you for bringing these relevant articles to our attention. They should definitely be referenced in this paper due to their importance in advancing the field of proteomics in marine invertebrates [revised version 170-171].

*COMMENT: lines 179-182: what is this for?*

REPLY: The methods in this section outline how the oysters were cared for during the experiment as well as how tissue was sampled for proteomics [revised version lines 188-190].

*COMMENT: Iodoacetamide is usually called IAA, not IAM*

REPLY: We have changed the abbreviations to IAA [revised version lines 200 and 202].

*COMMENT: line 300: Figure 4 does not provide any necessary information. The dataset (number of proteins within each biological function in the gills proteome and in the reference proteome) could be provided as supplementary file. Alternatively, the authors could use these data further to classify the lists of proteins identified and further discuss the proteins identified within each biological function category.*

REPLY: We have added a supplementary data file (Supplementary Data 5), which includes the list of proteins that contribute to the enriched GO terms. As recommended, we have removed Figure 4.

**Reviewer 3**

*COMMENT: The term de novo should not be used here. That implies protein sequencing without a reference database to search against, and that was not the case here [line 5 original version].*

REPLY: The phrasing has been changed to “a non-gel based approach” [revised version line 3].

*COMMENT: Should be a comma after (Crassostrea gigas)[line 77 original version].*

REPLY: The comma has been added [revised version line 77].

*COMMENT: This looks like it says that solid urea was added to each homogenate, which seems very unusual. Is this correct or was a concentrated stock solution used? [lines 189-190 original version]*

REPLY: This is correct. 36 mg of solid urea was added to each sample so that the total sample concentration of urea was 6M [revised version lines 197-198].

*COMMENT: What concentration of trypsin was used? (I assume most likely 1 ug/ul but it needs to say so) [line 198 original version]*

REPLY: That is correct and the information has been added [revised version line 206].

*COMMENT: I was not clear on what was done here on two counts. First, were the digests acidified prior to drying, which is what is commonly used. Second, what volume were the samples reduced to prior to loading on the macrospin columns. [lines 200-203 original version]*

REPLY: Dilute formic acid was added to the samples prior to drying [revised version lines 208-209]. The samples were reduced completely (to residue) and then reconstituted in 200 µl of 5% acetonitrile with 0.1% trifluoroacetic acid [revised version lines 209-210].

*COMMENT: Fluoracetic acid is unusual in this context and is also rather unpleasant to work with.  Could the authors please confirm they meant this and not formic acid or trifluoroacetic acid instead? [line 209 original version]*

REPLY: This was a mistake, formic acid was used and the information has been corrected [revised version lines 220-221].

*COMMENT: The false discovery rate stated to be 0.6% needs to be further explained. Does that mean peptide or protein false discovery rate? [lines 241-245 original version]*

REPLY: This is a peptide discovery false discovery rate. It is the estimated percent of incorrect identifications of MS spectra where multiple identifications could be of the same peptide (not all peptides are unique). It has been clarified in the text [revised version lines 217-218].

*COMMENT: Similarly, the standard deviation is reported for protein numbers, but it would be good to also include that metric for reproducibly identified peptides as well. [line 281 original version]*

REPLY: The standard deviation for peptides identified with high confidence has been included [revised version lines 295-297].

*COMMENT: The discussion of how criteria thresholds affect protein reporting numbers is a little misleading where it states that making criteria more conservative would decrease variability. It would, but it would also decrease the number of proteins involved and that is important to state. Rather than just talk about it, perhaps a table could be included showing how the protein and peptide numbers and variability change when the spectral count threshold is increase from 4 to 5 to 6 and so on. [line 322 original version]*

REPLY: This data is useful information and therefore we have included it as part of the text in the Discussion. The new text details how protein numbers and variability would change if spectral count threshold were to increase to either 5 or 10, compared to the cutoff of 4 [revised version lines 341-344].