

Dear editors,

Please find attached our revised manuscript "Dynamics of sex-biased gene expression during development in the stick insect *Timema californicum*". We thank you and the reviewers for the constructive comments which helped us to improve the manuscript. Please find our detailed responses to each point below.

Yours sincerely,

Jelisaveta Djordjevic (on behalf of all authors)

Dear authors,

Thank you for submitting your preprint " Dynamics of sex-biased gene expression over development in the stick insect *Timema californicum*" to PCI Evol Biol. Your manuscript has been read by two reviewers, whose comments are enclosed. As you will see, the reviews are largely positive, and, based on these reviews as well as my own evaluation, I would recommend your manuscript to be eventually included in PCI Evol Biol. However, before reaching a final decision, I would ask you to revise your manuscript according to the recommendations by the reviewers. Please address the main issues underscored by the reviewers, including the following:

- Addressing the issue of comparing only one species of each developmental type and interpreting it in the context of evolution in a more explicit manner.
- Discussing more explicitly how a whole animal transcriptome approach may affect the data
- Respond to the reviewer's comment about the rationale for not using a fold change cut-off in addition to a statistical significance threshold in the case of studying development while using whole-animal transcriptomes.
- Presenting the statistics supporting some of the claims and modifying some figures to insure that your message is clearly understood by the readers.

I also encourage you to revise your manuscript according to the more minor suggestions from the reviewers, which will certainly improve it.

1.

This manuscript addresses an important knowledge gap in the field of sexual selection, namely how do sex differences in gene regulation manifest through development. Despite the fact that sexually dimorphic phenotypes are likely a product of processes acting through development, gene regulation is typically studied without an ontogenetic perspective. In this regard, this paper is very timely as it focuses on three developmental stages in the stick insect *Timema californicum*. Furthermore, this insect has hemimetabolous development, where phenotypic sex differences amplify gradually through development, and so provides an important contrast to better studied holometabolous species such as *Drosophila*. However, there are a number of bold claims made throughout the manuscript that unfortunately I do

not think are supported by the results in their current form. I have a number of suggestions, detailed below:

My primary concern focuses on the use of whole body in these analyses. It is likely that tissue composition varies both between males and females, but also through development. This can result in shifts in gene regulation that are then falsely attributed to differential expression in this study (see Montgomery and Mank 2016).

Montgomery SH, Mank JE. Inferring regulatory change from gene expression: the confounding effects of tissue scaling. *Mol Ecol.* 2016 Oct;25(20):5114-5128

First, this limitation should be fully acknowledged in the discussion. Currently, it is not mentioned at all. Second, Montgomery & Mank recommend using a strict 2x fold change to exclude any genes that might exhibit patterns of regulatory variation arising from allometric shifts. Currently, I believe only a p-value is used to identify sex-biased genes and so many weakly sex-biased genes are included in the analyses that are likely a product of allometric shifts. Although this will obviously reduce the number of sex-biased genes in the study, I would strongly urge the authors to implement this measure to increase confidence in their results. Other studies with this problem have used publicly available data to identify genes post hoc with tissue enriched patterns of expression (Immonen et al 2014). I assume that isn't possible here but may be mistaken.

Immonen, Snook and Ritchie *Ecology and Evolution* 2014; 4( 11): 2186– 2201

The previous version of our manuscript addressed the caveat of analyses based whole-body transcriptomes briefly. We have now extended it for increased emphasis, see lines [400-404].

Regarding the fold change cutoff for identifying sex-biased genes, it is important to note that many of our analyses are designed to describe broad patterns of gene expression changes, rather than to quantify the numbers of sex-biased genes. For example, we are interested to know whether genes that are sex-biased at early developmental stages remain sex-biased in the same direction at later stages. For such analyses, applying a 2-fold-change cutoff could bias results towards very stable patterns. Note however that we do distinguish between groups of genes with different FC levels in the figures (i.e., slight sex-bias, strong sex-bias, sex-limited), for the exact reasons pointed out by the reviewer. For the subset of analyses where we quantify the portion of sex-biased genes, or where false positives in the category of sex-biased genes could bias our results [line 263], we now also apply a threshold of 2 fold-change ( $>1 \log_2FC$ ) in all cases (e.g., Fig. 8. Page. 16). When this threshold is applied, the patterns of gene expression we describe remain similar both over ontogeny and between species, showing our results are robust to tissue allometry.

I also found attempts to link developmental mode to the ontogeny of sex-biased expression weak. Given there is only one hemimetabolous and one holometabolous species, it is impossible to distinguish the effect of developmental mode from species identity. Furthermore, the *Drosophila* analysis is focused on whole body which is subject to the same allometric problems discussed above. But most importantly, this analysis has already been conducted in Perry et al 2014 arguably to a more robust standard as it studies gene expression

at the tissue level. This isn't acknowledged in the paper. I think the authors pose very interesting hypotheses which I understand they want to test. However, I actually think the inclusion of this analysis weakens the manuscript overall. I would limit discussion of the link between expression and developmental mode to the discussion section, where I think the authors can draw conclusions with published data and make some robust hypotheses for future work.

Perry JC, Harrison PW, Mank JE (2014) The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. *Molecular Biology & Evolution* 31: 1206-1219

We agree that it is impossible to draw general conclusions from two species. We state this in our manuscript and also state that we need multiple species of each developmental type to test our hypothesis (lines 388-391). To compare our data to other insects, we needed data similar to our own (whole-body rather than gonad-specific transcriptomes) which is why we re-analyzed published data from *Drosophila*. We agree with the reviewer that the previous study on *Drosophila* by Perry et. al 2014 is extremely interesting, and have made changes to acknowledge their work more clearly. After all, their work was a model for our study, see lines [82-87].

dn/ds is calculated using a pairwise comparison between *T. californicum* and *T. poppensis*. This means that it is impossible to attribute rates of change to either lineage and therefore weakens power to test for the relationship between rates of sequence evolution and sex-bias in *T. californicum*. Ideally, an outgroup should be added so that dn/ds can be calculated along the branch leading to *T. californicum* after the split with *T. poppensis*. This approach obviously relies on having a reference genome for another closely related species.

Pairwise dN/dS calculations between *T. californicum* and *T. poppensis* were done to maximize the number of genes (one-to-one orthologs) for the sequence rate evolution comparison between sex-biased and un-biased genes. We now calculated values of dN/dS along the branch leading to *T. californicum* after the split with *T. poppensis* as suggested by the reviewer. With this approach we had fewer genes for comparison, but results remained largely the same, see lines [226-229], and (Fig. 5). We have also updated our methods section (lines [525-532]) to reflect this change.

The manuscript tests if genes that are more sex-biased have higher rates of coding sequence evolution. There are a number of factors that influence the rate of coding sequence evolution and need to be accounted for. This includes expression level, which the authors already control for, but also tissue-specificity and GC bias which should be included in the analysis.

With the current dataset we are not able to consider tissue-specificity, but we now control for GC bias in addition to average expression levels, see methods lines [540-545], [232 and (Supplemental Table 6), (Fig 6. Lines [246-250]). After correcting for GC, we find similar results, namely that genes with a greater sex-bias have higher rates of coding sequence evolution.

Fidel Botero-Castro, Emeric Figuet, Marie-Ka Tilak, Benoit Nabholz, Nicolas Galtier, Avian Genomes Revisited: Hidden Genes Uncovered and the Rates versus Traits Paradox in Birds, *Molecular Biology and Evolution*, Volume 34, Issue 12, December 2017, Pages 3123–3131

Richard P. Meisel, Towards a More Nuanced Understanding of the Relationship between Sex-Biased Gene Expression and Rates of Protein-Coding Sequence Evolution, *Molecular Biology and Evolution*, Volume 28, Issue 6, June 2011, Pages 1893–1900

L36 I don't think this statement is supported by the results. It is impossible to distinguish species differences from differences in development when only two species are compared.

We toned down the statement to “are consistent with the prediction” lines [37- 39].

L45 Seems inappropriate to cite Chauhan et al here in isolation. I suggest only citing the review paper (Mank 2017) or a more exhaustive list of empirical papers.

We removed the Chauhan et al citation [48].

L49 Seems odd to single out these two papers when the majority of transcriptional studies focus on adults. I would simply cite Mank 2017 here.

Done, lines [51-52].

L51 Missing citations:

Mank JE, Nam K, Brunström B, Ellegren H. Ontogenetic complexity of sexual dimorphism and sex-specific selection. *Mol Biol Evol.* 2010 20142440.

Perry JC, Harrison PW, Mank JE. The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. *Mol Biol Evol.* 2014;31(5):1206-1219.

Hale MC, Xu P, Scardina J, Wheeler PA, Thorgaard GH, Nichols KM. Differential gene expression in male and female rainbow trout embryos prior to the onset of gross morphological differentiation of the gonads. *BMC Genomics.* 2011;12:404.

Zhao M, Zha X-F, Liu J, Zhang W-J, He N-J, Cheng D-J, Dai Y, Xiang Z-H, Xia Q-Y. Global expression profile of silkworm genes from larval to pupal stages: toward a comprehensive understanding of sexual differences. *Insect Sci.* 2011;18:607–618.

Added citations, lines [56-57].

L65 Cite Mank 2017

Added citation, line [64].

L75 Cite Zhao M, Zha X-F, Liu J, Zhang W-J, He N-J, Cheng D-J, Dai Y, Xiang Z-H, Xia Q-Y. Global expression profile of silkworm genes from larval to pupal stages: toward a comprehensive understanding of sexual differences. *Insect Sci.* 2011;18:607–618.

We cited here papers that have pre-adult and adult stages analyzed, which is why we did not include this reference, as it only focuses on pre-adult stages.

L372 Similar to points raised above, I do not think there is sufficient evidence to claim this.

We now solely describe our findings from *Timema* in the conclusion and removed our idea that hemi- vs holometabolous development may affect the dynamics of sex biased gene expression lines [405-410].

L485 Specify the model used in PAML. I assume codeml?

Specified. We now use the branch-site model with Godon software, see lines [527-532]. The model we used in the previous manuscript version was M0.

L508 Remove 'the'

Done.

2.

In their pre-print “Dynamics of sex-biased gene expression over development in the stick insect *Timema californicum*”, Djordjevic and colleagues used RNA-sequencing in three developmental stages of the hemimetabolous stick insect, *T. californicum*, as well as previously published RNA-seq data from *D. melanogaster*, which is holometabolous, in order to examine how sex-biased gene expression varies over developmental time and to compare how the dynamics of sex-biased gene expression vary between hemimetabolous and holometabolous insects. The authors found that in *T. californicum* the proportion of sex-biased genes gradually increased over developmental time, with the direction of sex bias generally remaining consistent during developmental progression, but that *T. californicum* had less sex bias than in *D. melanogaster*, where sex-biased gene expression abruptly increased for the adult stage. In general, the pre-print is well-written and well-organized, with the objectives of the study well-outlined and the study itself nicely presented within the context of previous research in the field. However, there are a few issues that I think should be addressed (outlined in the major and minor comments below).

Major comments:

1. Lines 113–115, page 6: Did the authors perform any statistical test to support this statement that “sex-biased gene expression gradually increased over the three developmental stages”?

We did not fit a model to distinguish e.g. between a linear or non-linear increase as this would require more than the 3 points in the development. We now tested for the difference in the proportion of sex-biased genes between neighboring stages. There is a significant difference in the proportion of sex-biased genes in all the comparisons, see lines [131-134]. In addition, the effect sizes (Cohen’s  $h$ ) are similar between neighboring stages, supporting the idea that the increase is linear.;  $h_{(\text{hatchling-juvenile})} = 0.34$ ,  $h_{(\text{juvenile-adult})} = 0.49$ , lines [134-136], see as well method section [511-513].

2. Supp Fig 2 is confusing. In lines 127–128, pages 6–7, the authors claim this figure show “genes sex-biased at earlier stages generally and remaining sex-biased in the same direction at later stages”, but this is not evident from this figure as in several places on the heat map there are genes that clearly switch from red to blue or vice versa, especially between the juvenile and adult stages. I assume these genes were not significant?

It is Supplemental figure 1 which shows that “genes sex-biased at earlier stages generally and remaining sex-biased in the same direction at later stages”. We have clarified this at lines [148-149]. Indeed, genes that switch from red to blue and vice versa in the supplemental figure 2 include genes that are not significantly sex-biased in some of the stages, we now clarified this in the legend, Supplemental fig 2, [line 64].

3. Lines 130–133, page 6 and page 14 lines 245–249: While it is nice to see that there is some overlap in sex-biased genes among stages, I think a more informative way to characterize similarities in sex bias across developmental stages would be to calculate the correlation of sex bias between stages and that the authors should consider adding this as well. This would also allow the author to more directly and thoroughly compare sex-biased gene expression dynamics between the two species.

In both *T. californicum* and *D. melanogaster* there is a significant correlation in sex-bias [ $\log_2FC$ ] between developmental stages, supporting that sex-biased genes in earlier stage largely remain the same in later stages. We added these results, see lines [159-160], lines [277-279], method section; lines [509-511] and Supplemental fig 3.

4. Related to point 3 above. In the Fig 3 legend, the authors state “The number of genes shared between all three stages was greater than expected by chance”. Did the authors make this comparison for any of the individual stage comparisons?

Yes, we now report statistical results for each of the overlaps between individual stages. All the overlaps are significant, see lines [201-202], and supplemental table 7.

5. Fig 4: I think that it is misleading to present and test for significant differences in tau for each developmental stage. Here, tau is a measure of the stage specificity of gene expression and is based on the gene expression in all stages (see comment 6 below). Thus, there is only one tau value calculated for each gene and the tau values for each stage are largely overlapping. I think it would be more appropriate for the authors to simply present and test for significant differences in tau for males and females for each gene category rather than also include the developmental stage.

We agree that it is slightly misleading to have the same tau value three times. However, a gene’s sex-bias category can vary between developmental stages. For instance, a gene can be female biased in hatchlings, but not biased in juvenile and adult stages. The tau value for this gene is then analyzed in two different gene categories. Nevertheless, we added analyses with a single tau per gene category as suggested by the reviewer and we found the same trend. This result is presented in Supplementary figure 4.

6. Related to point 5 above. Lines 474–480, page 24: I think that it would be nice for the reader if the authors included how tau is calculated in this section. I admit that this is a little pedantic,

but tau is actually originally a measure of tissue-specificity for gene expression, which is how it is proposed and used in Yanai et al, which the authors cite. In recent years it has also successfully been applied as a measure of stage specificity.

We clarified this in lines [519-521].

7. Lines 238–241, page 14 and Fig 8: It would be nice if the authors tested for significant differences in the proportion of sex-biased genes in *T. californicum* versus *D. melanogaster* here to provide statistical support for their claims that *T. californicum* shows lower levels of sex bias. It would also be nice if they tested for significant differences in the proportion of sex biased genes between stages within each species to better support their observations of how sex-biased gene expression changed differently over time between the two species.

We now compare the number of sex-biased genes in adults of the two species see lines [268-269] and we as well added the suggested comparisons within the species, lines [263-267] and [131-134]. There is a significant difference in the proportion of sex-bias in all the comparisons. However, we do not want to compare the number of sex-biased genes between the two species for pre-adult stages directly since the stages are likely not homologous (see legend of Figure 8).

8. Page 19, lines 360–364: While it is nice that the authors acknowledge that gonad size may play a role, sex bias varies depending upon tissue and it would be nice if they expanded a bit more about how using whole bodies versus individual tissues may affect their findings.

We have expanded this section as suggested, see lines [400-404].

Minor comments:

1. Lines 46–49, page 3: I agree with the authors and there are a lot of studies focusing on this topic in adult stages in various species. Therefore, it would be nice if the authors provided a few additional examples here.

We now cite Mank 2017 instead, [51-52].

2. Line 60, page 3: I think that the “?” at the end of this line should be a “.”.

Done, [64].

3. Line 80, page 4: I think that “in hemimetabolous insects” should be “in a hemimetabolous insect”.

Done [89].

4. Lines 92–101, page 5: This is a purely stylistic suggestion: I think that it helps with the flow and reader understanding when the final paragraph of the introduction ends with a brief summary of the major results and their take-home message and the authors could consider doing this.

Done, lines [111-117].

5. Line 112, page 6: I think that “over development” should either be “during development” or “over developmental time”.

Done, line [128], as well corrected throughout the manuscript.

6. Figs 2, 6, and 7a: Perhaps it is my computer, but the colors for the different categories within each sex are a bit difficult to distinguish.

We changed the colors to increase the contrast between different sex-bias gene categories in Fig 2, 6 and 7a. We also corrected the threshold our for sex-bias classification, which previously classified genes as strongly sex-biased when  $>2 \log_2FC$  rather than  $>1 \log_2FC$ , lines (192, 193, 194, 286, 287, 288), lines [139-141 and 271-273].

7. Lines 119, 121, 122 (page 6), 243, and 245 (page 14): I find the inclusion of P-values here confusing. In their methods the authors already state the p-value threshold that they used to assess significance. When I initially saw them here, I thought that they were referring to a statistical test.

Removed p-values, lines (140, 142), line 272.

8. Fig 5: Why are the M-biased genes NA for the hatchling stage? Could you please explain this in the figure legend?

None of the four M-biased genes at the hatchling stage had their ortholog identified in the other *Timema* species, which is why the dN/dS value is missing. This was explained in the methods, see lines [536-538]. For clarification, we have now added this explanation in the figure legend, [243-244].

9. Fig 6: I think it would be nice to label each box with the sex in addition to the stage.

Done, Figure 6.

10. Page 21, lines 399–401: Did the authors do any kind of quality control to ensure that the inclusion of DNA in their RNA samples during library prep did not affect their results?

Samples were treated with DNase. We clarified the corresponding section in the methods, see lines [435-436].