Universidad Austral de Chile
Instituto de Ciencias Ambientales y Evolutivas

Professor Wolf Blanckenhorn
Managing Board PCI Evol Biol

Dear Prof. Blanckenhorn,

Please find a revision of our manuscript “Natural selection on plasticity in a highly seasonal environment” which we have revised in accordance to the suggestions provided by all three reviewers.

The reviewers provided very constructive, detailed and clear comments that we respond below. We have carried out new analyses and substantially rearranged some parts of the manuscript in order to make our message clear to the reader. Overall, this resulted in a more succinct form of the ms.

We are grateful for the very helpful suggestions of the reviewers and believe that the consideration of all the comments has greatly improved our ms. Thus we hope that our changes are sufficient to make this manuscript publishable; otherwise we are more than willing to further improve the ms.

Thank you for your attention to this manuscript. If we can provide any further information regarding the submission, then please do not hesitate to enquire.

Sincerely yours,

Juan Diego Gaitan-Espitia
Postdoctoral Fellow
Oceans and Atmosphere Division
CSIRO - Australia

Editor: We obtained 3 external reviews for the paper "Natural selection on plasticity of thermal traits in a highly seasonal environment" by Leonardo Bacigalupe et al. All reviewers saw merit in this study and think that it can eventually be recommended by PCI EvolBio after extensive revisions, which however are necessary and strongly recommended before. The reviews were consistent in the sense that the reviewers identified two aspects that need extensive revision. First, the context of climate change needs to be revised, as only a single population of this frog was studied in its natural habitat, thus experiencing no climate change but natural seasonal fluctuations. Second, the statistical analysis needs extensive revision according to the comments of the reviewers.

R: Many thanks indeed for your positive response. Below we comment on all of the suggestions or requests for clarification. We have tried as much as possible to explain better, trim and re-organize the whole MS. In particular, we have completely removed the climate change context of the problem and reframed the whole ms in terms of selection in seasonal environments. Also, we have made new analyses with the raw data. The overall results are more straightforward than on the previous versions and have strengthen our conclusions. Also we have attempted to streamline the methods and results also whilst accommodating the suggestions made by the Reviewers.

Nadia Aubin-Horth

I think the ms has great potential and presents solid data, however I have concerns that must be addressed in a new version before I can recommend this manuscript.

R: Thank you very much indeed for this positive opinion. Below we comment on all of your suggestions or requests for clarification.

The authors place their work within the framework of predicting the response of populations to global warming. However, they do not test survival in a warm-world scenario (in the lab for example), but rather in current temperatures in the wild. Therefore, they test if plasticity is under selection in the actual thermal regime. Selection could therefore have already acted on plasticity. The authors themselves seem to suggest it did on line 73-74 “However, we still do not know whether that physiologicl plasticity, WHICH RESULTS FROM INHABITING A HIGHLY VARIABLE ENVIRONMENT, is being targeted by natural selection” Their selection measurement is thus not really related to their introduction that aims to quantify the direction and strength of selection on plasticity in a global warming context, since it is not done in conditions that mimic these higher expected temperatures. Why was this type of survival measurement chosen, given the context for the study? If there is selection now, what does it tell us about future response to global warming, if the selection regime changes anyway?

R: Point taken. We have now removed the climate change framework from all the ms (it was something highlighted by all 3 reviewers). We have reframed the story in terms of selection acting in highly seasonal environments that also experiment daily extremes of temperatures.

Second, they present selection gradients but never mention the second half of the prerequisite to a response to selection: heritability of the traits. The interindividual variation in plasticity of each trait must be underlined by genetic variation, otherwise there can be selection on the current generation but there will be no evolutionary response in the next generation. This is not the direct topic of the manuscript, but must be addressed to highlight how these selection measurements are relevant to make predictions about response to climate change.

R: Thanks for highlighting this. You are completely correct and we really appreciate the suggestion about reframing the ms in other terms than climate change. That said, we decided to stick with just

Campus Isla Teja · Valdivia · Chile
Casilla 567 · Fono: 56 63 221344 · Fax: 56 63 221344 · email: ecologia@uach.cl · www.uach.cl

the selection analyses and its interpretation, leaving out any potential speculation about the evolutionary response to selection in order to keep the ms as short as possible.

Because of these two points, I think the manuscript introduction must be rewritten so that the core of the work is presented in an appropriate context. I agree that we need to know more about selection on plasticity itself, and that acclimation to temperature is an ecologically-relevant trait in a frog species. In my opinion, this important and novel dataset could be presented without relying on climate change as the background context for doing this study, but a large amount of work on the text must then be done.
R: Completely agree. As we mentioned, we have now removed the climate change framework from all the ms and have reframed the text in terms of selection acting in highly seasonal environments that also experiment daily extremes of temperatures. In our opinion, this has made the ms definitively much more strong and integrated.

My other main point is that the discussion is on the absolute values of the traits that were associated with survival (line 214-2017). However, these results are not clear at all from the results section (where can we see this result?), and since this question was not presented in the introduction, the reader is left wondering what is the link with plasticity (the main question at the beginning of the ms). If these results become more central in a revised version, I would suggest using a 3-D graph showing survival and an associated physiological trait with the two acclimation temperature on the third axis, so we can see the different fitness peaks in different "trait combinations"
R: Many thanks for highlighting this. We have completely re-analyzed the data using the raw data instead of the plasticity as a difference of means as suggested by another reviewer. This clearly resulted in a more clear presentation and interpretation of the Results.

Minor comments:
*Predictions presented in the introduction (line 77) should be represented graphically.
R: This is a great idea. Many thanks! We have included a new figure (figure 1) with a graphical representation of the main predictions.

*Some terms such as Q10 etc should be defined for non physiology specialists if a broad readership is targeted. Similarly, terms like norm of reaction should be defined.
R: Done

*Please include how many frogs were tested first at 10 C and then at 20C and vice-versa. Was it about equal?
R: Yes, it was half and half. We made it clear on this new version.

*Why not include mass as a covariate of some sort (non-specialist question here) instead of using it in the MCR model.
R: We included body mass as a covariate in those models where it was needed (i.e. models including traits correlated with body mass).

*The calculation made on line 200-202 need to be explained in more details.
R: The line "Survival in relation to each covariate was obtained as the model averaged value across all candidate models (Table 1), weighted by individual model probability" simply means that the coefficients relating each independent variable to survival were obtained by model averaging. For example, assuming the basic linear regression model y = a + bx. If model one has an Akaike weight of 0.75

and b = 0.3, and model two has an Akaike weight of 0.25 and b = 0.23, then the model averaged value of b is calculated as: 0.3*0.75 + 0.23*0.25 = 0.283 (i.e. it is "closer" to 0.3 as this model has a higher probability).

*Figures: figure 2 need finer lines, maybe color, so the reader can follow individual norms of reaction.
R: We appreciate the suggestion but this is rather difficult to implement it. The lines are very thin although they appear not to be because many individuals had the same value for some of the traits (e.g. CTMax).

Figure 3: need to re-order the panels so they match the 3 predictions (and the new figure showing the predictions graphically)
R: We have now a new figure (Figure 4)

##########################
Wolf Blanckenhorn
This is definitely an interesting study because selection on plasticity in thermal traits, as opposed to selection on the thermal traits themselves, is rarely reported in animal species. The manuscript is overall clear although several small (language and style) edits are still needed to streamline the presentation.
R: Many thanks indeed for your positive response. We have modified the ms accordingly to your suggestions, whilst accommodating the suggestions made by the Reviewers.

However, I had difficulties following the conclusions of the authors, as the data analysis in my opinion is convoluted and suboptimal, in the extreme even flawed. I suggest reanalysis in a more standard way as outlined below, which should lead to clearer results and make this manuscript more understandable.
R: Thanks for highlighting this. You are completely correct and we really appreciate the suggestion about reanalysing the data using raw values instead of a difference in means. Indeed, the results are much easier to interpret in this new version.

Major comments: 1) The authors investigate one apparently marginal, or in some way special ("northermost") population of this species (L93). That's fine, and the desert habitat seems to be indeed practical for monitoring survival in a thermally extreme environment because it is nicely contained. However, only ONE population is assessed here, so any conclusions about the adaptive fit of this species to this special habitat remain unreplicated and hence of limited generality. What remains is a study of thermal plasticity and selection in a single population.
R: The reviewer is correct. In this context, we have substantially toned down our conclusions and reframed the whole ms in terms of selection in seasonal environments instead of discussing about climate change implications.

2) Most crucially, I believe the analysis is suboptimal if not flawed, hampering clear interpretation and discussion of results.
R: We really appreciate your suggestion of using the raw data instead of the plasticity as a difference of means. In this new version we further reduced the number of candidate models (from 27 to 13) to minimize the likelihood of spurious results given our relatively small simple size (N = 83) and number of recaptures (N = 3). Thus, we tested only for a model with body mass and models with directional selection for each trait separately and also for correlational selection (interaction of trait combinations) among the same trait at both acclimation temperatures, which indicates plasticity (as you suggested). Body mass was included as a covariate in the case of just two variables as it was not asso-

ciated with any other trait. The overall results are more straightforward than on the previous versions and have strengthen our conclusions.

c) I do not see the added value of the AIC analysis in Table 1. In particular, it does not seem to help the authors in interpreting their results, as it is unclear whether any effects are significant or not. The logic of AIC is that the model(s) with the lowest AICc should be preferred (if delta AIC is within ca. 2). Here this means, if Table 1 is correct, that most of the terms do not add significantly to the model.
R: This is partially correct. The delta AIC around 2 should not be taken as a sort of cutoff point. As Anderson suggested (Model Based inference in the Life Sciences: A primer on evidence), if a model has a e.g. delta_AIC = 4 it has an evidence ratio (Akaike weight best model / Akaike weight reference model) of 7.4. This means that the best model has 7.4 times the weight of evidence relative to model we are evaluating, which suggest that that reference model still has some information to provide. Usually models with a delta AICc somewhere in the 8–14 range would be judged by most objective people as having little plausibility. In our case (Table 1) the last 2 models are the only ones that are very unlikely given our data. That means that the remaining models do contribute "something". In any case, in spite of this contribution (inferred from the corresponding Akaike weight) we realized that the rate of change of survival with changes in the corresponding covariate is very slight and thus, we decided to explicitly minimizing in the discussion the impact of that trait on survival.

d) Related to (c) above, it is unclear if the linear selection gradients depicted in Fig. 3 are significant or not. I would suggest using standard uni- and multivariate selection analyses for both the traits xi and plasticity (although xi(10)xi(20) would show up as correlational selection in such an analysis), adding linear and quadratic terms (as for term 27 in Table 1) plus the above correlational term. Terms can be removed if not significant. This would make the analysis much more palatable and clear to the average reader.
R: We kindly disagree. A "non-significant" term does not necessarily mean that there is no contribution in the term itself (e.g. it might just be a matter of statistical power). Given our small sample size, few recapture events and the many traits we measured, we wanted to obtain as much information as possible. That is to evaluate the strength of the evidence for some competing hypothesis given by our data set.

L228: The argument that frogs died due to thermal effects rather than predation is plausible, but in the end cannot be proven. Why not add actual field temperature parameters in the model to obtain indirect evidence that indeed temperature is the culprit?!
R: This a good point, but we do not have environmental data as the one reported in Ruiz-Aravena et al. 2014 for all the recapture periods.

Further, why not separate analyses for breeding and non-breeding season to correctly identify which frogs died when?
R: Unfortunately we have not enough recaptures for such analyses. The problem is that sometimes we recaptured a frog at the third recapture, which means that if we analyze only the first or second time it would appear as "dead".

######################

Dries Bonte

The work of Bacigalupe and colleagues addresses an important topic in evolutionary biology, i.e. the evolvability of thermal reaction norms. The authors study different component of thermal biology (ranging from acclimation strategies to plasticity in thermal preferences and thermal sensitivity of metabolism in two populations of a frog. The species had been demonstrated to fulfil the conditions by which thermal reaction norms can be expected to be subject to selection. The paper is in general well written and all experiments and analyses seem to be well executed, and repeatable.

R: Many thanks indeed for your positive response. Below we comment on all of the suggestions or requests for clarification. We have tried as much as possible to explain better, trim and re-organize the whole MS. In particular, we have completely removed the climate change context of the problem and reframed the whole ms in terms of selection in seasonal environments. Also, we have made new analyses with the raw data. Also we have attempted to streamline the methods and results also whilst accommodating the suggestions made by the Reviewers.

The authors use model inference approaches of estimated survival rates in function of different plasticity measures to detect signals of stabilising or directional selection. By means of model averaging, the authors do not detect strong and univocal signatures of selection, and instead survival to be primary size dependent. Of the remaining models, 25% of the AIC weights suggest directional selection in one of the measures traits. The authors use this evidence to demonstrate that thermal acclimation is subject to selection. This interpretation of course highly depends on the set of contrasted models, and the interpretation is subject to debate as especially the estimates on the directional component (Fig 3) show large variation, what frequentist statisticians would interpret as non-significant. I do not think that this absence of 'significant' results does jeopardise the importance of the study, but it may overall need to be toned down.

R: The reviewer is correct, as all results are dependent on the set of contrasting models. In this new version we further reduced the number of candidate models (from 27 to 13) to minimize the likelihood of spurious results given our relatively small simple size (N = 83) and number of recaptures (N = 3). Furthermore, we analyzed the raw data instead of the plasticity as a difference of means. Thus, we tested only for a model with body mass and models with directional selection for each trait separately and also for correlational selection (interaction of trait combinations) among the same trait at both acclimation temperatures, which indicates plasticity. Body mass was included as a covariate in the case of just two variables as it was not associated with any other trait. The overall results are more straightforward than on the previous versions and have strengthen our conclusions.

In the abstract you write that the results suggest a complex fitness landscape. One may equally argue it is very simple as none of the directions are really convincing- and none of the complex models showed strong support. I am also wondering to which degree the different thermal components are correlated at the individual level. Such phenotypic correlations may be important for further interpretation. Now, only correlates with body size are reported This is my main point of concern for this study.

R: We have calculated phenotypic correlations among the traits and have discussed some of them in this new version.

I here-under list additional issues that the authors may want to consider:

R: As the ms changed substantially, we respond to the queries that relate directly with this new version

line 72: Is anything known on whether the thermal traits themselves are under selection (compared to plasticity). This seems like a first logical step in the research. If yes, please add a couple of sentences on this as this may be important to further understand the relevance of your findings on acclimation.
R: All the data was re-analyzed using the raw data and the plasticity as correlational selection.

Methods line 163. Explain what you mean with the survival rates extracted as individual covariates?
R: Changed to: Survival in relation to each covariate was obtained as the model averaged value across all candidate models (Table 1), weighted by individual model probability. This means that the coefficients relating each independent variable to survival were obtained by model averaging.

Methods line 170: the selection of alternative models is highly responsible for this interpretation. You need to explain clearly why you always had body size as covariate from the beginning.
R: This a very good point as we realize we did not explained correctly in the previous version. Body mass was included as a covariate in the case of just two variables as it was not associated with any other trait. That means that only 4 of the 13 evaluated models has body mass included.