Review of the manuscript “Negative impact of mild arid conditions on a rodent revealed using a physiological approach in natura”

This paper reports interesting findings on physiological responses to increased aridity in natural rodent populations inhabiting a semi-arid region of South Africa. The authors compared the habitat and the physiology of several populations of two mouse species at the onset and the end of the dry season. The comparison is based on several measures of environmental (i.e. vegetation composition and greenness) and individual (i.e. body condition and blood metabolites related to nutrition and liver function) quality. The authors found that mice responded to the seasonal variation of dry conditions, as the levels of half of the blood parameters shifted as one would have expected under dietary restriction. Although both species responded to increasing dryness and did not differ in body condition, the authors found some interspecific differences for some markers, which they attributed to the fact that one species is supposedly better adapted to dry environments than the other is.

Such studies are of critical relevance in the actual context of global change, in particular climate warming and disrupted precipitation patterns, which promote drought occurrence. Understanding how, and to which extent, natural populations will cope with these changes is essential and has become a central question in ecology. Since physiology mediates the relationship between life history and the environment, studying wildlife physiological responses is a promising way to do so.

The manuscript is well written and the rationale behind the study is comprehensive. While the design is sound and the authors considered various complementary and pertinent variables, part of the methods, and in particular statistical analyses, may be flawed and deserve clarification. Most of my concerns and critics refer to this section.

Evaluation of the various components of the article

Title and abstract

The title does reflect the main content of the article, but it could nevertheless be improved. First, referring to “a rodent” reduces the significance of the study, since the authors actually used a comparative approach to study several populations of two rodent species. So I would rather suggest referring to “natural rodent populations” or to an equivalent wording. Second, the term “physiological approach” is vague and little informative. Likewise, I would suggest being more precise on what has been done in the study, by referring to “markers of physiological condition”, for instance (or something else equivalent but more detailed than the actual formulation).

Overall, the abstract summarizes properly the objectives and the main findings of the study, except for body condition. However, the authors did not state precisely which physiological metabolites were investigated, and neither did they mention body condition nor habitat characteristics (i.e. bullet points #2 and #3). Thus, I would suggest replacing “blood concentrations of physiologically relevant metabolites” by “blood concentrations of markers of kidney and liver function” (as stated later in the text), and I would at least mention that body condition was also measured. Finally, the idea that “it is crucial to identify the potential for species responses” (i.e. bullet point #1) could be even more explicit, by stating for instance that it is crucial to “investigate/study/define the extent and limits of species responsiveness”.

Introduction

The introduction clearly explains the motivation for the study, and both the research question and the associated predictions are well presented. The introduction also builds on relevant research in the field. I only have a few minor comments.

Since "Adaptive evolution" is defined, I would also provide a short definition and an associated reference for "phenotypic plasticity" (eventually within brackets).

At line 72, please provide a definition of homeostasis. This is a central concept in physiology, which will be known by the community of ecophysiologists. Yet, a definition (or a reminder) might be useful for ecologists and biologists who specialized in other fields.

At line 101, it would be more appropriate to refer to "physiological responses" (i.e. plural) as these could involve different processes.

At line 107, reword as "generate additional selective pressures to those already experienced".

At lines 112 and 115 : add "a third species of the same genus" after the first occurrence of "*R. pumilio*", and replace the second occurrence of "*R. pumilio*" by "individuals" since the pronoun "they" is used later in the sentence.

At line 126, please specify there which species you are referring to, since that last one to be mentioned was *R. pumilio*, but that species was not included in the study. Then in the next sentence (line 128), simply mention "parapatric populations of the two species".

At lines 129 and 133, add "that" after "we expected".

Materials and methods

This study relies on the combination of several environmental and individual measures to investigate the physiological consequences of seasonal variation in dry conditions, which is a strength of the study. Consequently, the description of the materials and the methods is long, and should be as clear as possible to help the reader follow the experimental plan. In particular, the way this section is organized could be optimized, as I would have expected some information to be given earlier than in the last subsection named "Data preparation and analysis". Typically, I first thought that information regarding body condition assessment and age class determination were missing because they were not introduced before the data analysis subsection. By contrast, specific paragraphs are dedicated to the description of the measurement of habitat characteristics and physiological markers (i.e. blood metabolite concentrations), before presenting how these data were then analyzed. I thus recommend that this section be rearranged and homogenized to separate variables description from data analysis (i.e. statistics).

Overall, the experimental plan is consistent with the questions. The methods are described in sufficient detail for most parts, but not all. Specifically, I believe that breeding status assessment and age class assignment deserve clarification. Also, I wonder whether the method for body condition calculation is appropriate. Most importantly, data analysis relies on various tests and models, and I found it hard to follow the analytical approach. Actually, some of the chosen methods do not convince me, and I don't understand the model selection procedure for physiology data (although I briefly checked the statistical scripts, I did not evaluate them in detail). My point-by-point comments are detailed below.

At lines 165-166, it is said that breeding status was assessed (remove the first "s" from "sassessed") based on external morphological metrics, but these metrics are not specified. Please explain clearly how breeding status was determined.

At line 249, the authors stated that trapped mice were assigned to age categories based on their "general appearance, size and/or breeding status". First, while I agree that both size and breeding status can be objective criteria, I wonder how general appearance could be too without further information. Please provide detail on how mice general appearance, and which particular elements of it, were included in the assignment of age classes. Second, I struggle to understand how many/which categories were considered in the end in the analyses, since it is first said that mice were assigned to either juveniles, subadults or adults (i.e. 3 categories, line 248), but it is then written that 4 length/age classes were considered (lines 255-256). Please clarify.

At line 237, the authors explain that body condition was assessed through the ratio of log(mass) to log(length). Assessing body condition is particularly relevant in the context of this study, and common practice in ecology. Several methods have been used to calculate it, and despite no real consensus, most of them have been criticized (see Peig and Green 2010 *Functional Ecology* for a review). I am under the impression that simply computing the ratio of mass to length is outdated, as this method was not even considered in the above-cited paper. According to the authors of that review, the dominant method consists in calculating the residuals from an OLS regression of mass against length. Yet, these authors also proposed an alternative method, the scaled mass index, and argued that their index is a better indicator of energy reserves than OLS residuals (Peig and Green 2009 *Oikos*). Therefore, I strongly recommend that the authors of the present paper re-evaluate mice body condition, ideally using the scaled mass index, and check whether their conclusions hold using another estimate of body condition.

At lines 218-223, the authors stated that they performed an ACP on the 7 variables of vegetation composition. Both the method of measuring these variables and the way to analyze them are puzzling me.

First, it is said that these variables were measured within the $4m^2$ and within the $100m^2$ quadrats for each successful trap, and earlier in the section (i.e. line 184), it was specified that both small and large quadrats were centered on the trap position. So from my understanding, small and large quadrats overlap, in the sense that small quadrats were positioned within the larger ones, meaning that all items that were considered in the small ones were also considered in the larger ones. How is that not redundant? How did the authors take into account the fact that they measured the same grass/bushes/plants twice, but included both measurements in their analyses later on? What was the interest of considering both small and larger quadrats? Actually, the first 2 PCs are displayed in the supplementary figures 4 and 5, and as I would have expected, the position of the different habitat structure variables are very similar in $4m^2$ and $10m^2$ quadrats (only "dry bush" and "no cover" items slightly differ but still remain in the same area on the graph). Then, the authors mentioned that they considered 7 variables in the following PCA (lines 218-219), corresponding to the 7 vegetation variables previously listed (lines 182-183). Thus, it seems that they ran two separate analyses, one for small quadrats and one for large quadrats (it would seem so from the results section, the supplementary information and the R scripts), but again, it appears highly redundant to me.

Second, it is written than 80% of the variance was explained by the five first principal components (PCs), and that these 5 PCs were then used as response variables of a PERMANOVA. While most of the variance was indeed captured by the retained PCs (which is desirable when performing a PCA), retaining 5 PCs when the analysis initially included only 7 variables appears pointless and a failure to summarize the original dataset. Multivariate analyses such as PCA are typically used to deal with datasets containing multiple quantitative variables, aiming at summarizing the information by reducing the number of variables (generally to 2 or 3 PCs are being considered, Greenacre et al. 2022 *Nature Reviews Methods Primers*). In the present case, having 5 composite variables (or PCs) rather than 7 raw variables doesn't seem effective nor relevant to me, even more since I suspect that variables of vegetation composition might be redundant. Finally, I wonder why the authors did not consider including the NDVI into the PCA, since it is an index of vegetation greenness, which contributed to characterize vegetation and thus habitat quality.

At lines 240-241, the authors wrote that they performed an ANOVA to test whether several variables (e.g. breeding status, sex, habitat quality) influence body condition, including site as a random factor. It might just be a semantic issue, but what they are describing seems to be a linear mixed effects model to me, leaving me to wonder why they did not say so? Actually, checking the R script, I saw that the "lmer" function of the package "lme4" was used for the analysis, so I really think that it would be more informative to indicate it. I am under the impression that the term "ANOVA" is being used in its general sense throughout the text, while a more precise terminology could be used to describe the models that were computed. Also, the authors included the three-way interaction between session, species and sex (i.e. session*species*sex, line 241 and in R scripts), without explaining the reason for the consideration of such a complex interaction. Three-way interactions are particularly difficult to interpret, and should not be included unless one has solid hypotheses behind them and can express associated predictions. I would recommend removing the triple interaction from the model, eventually including the two-way interactions instead (i.e. session*species + session*sex + species*sex), unless providing a robust justification for it.

At lines 260-262, the authors explained that they performed another PCA to identify outliers in the data set, what led them to remove 12 data points from further analyses. I never heard of the use of PCA to identify outliers (simply plotting the distribution of a given variable through a boxplot might have worked just fine), and I don't understand how they proceeded exactly without further explanation. Typically, what was the cut-off criterion? In fact, I am concerned by the removal of data points. Again, one must have a solid argument to consider a given data point as an outlier. For instance, it can be the case if its value is far out of the documented physiological range of the variable for the studied species. While I am well aware that reference values are often lacking for wildlife, it remains that data points removal must be properly justified. The authors mentioned that some samples were heavily hemolyzed, what can effectively be an issue, but I don't understand why values from smaller blood volumes would end up as outliers, and I still don't see what the PCA brings to the matter. If the Vetscan was unable to compute reliable values from hemolyzed samples or from low volume samples, simply state it and discard the values for that specific reason.

At line 272, it is said that a multivariate analysis was performed, please indicate here which analysis (I assume it was a PERMANOVA since this analysis is mentioned later at line 276, but it would be clearer to precise it right there). I am not familiar with PERMANOVAs, so I cannot really comment on it. However, I am confused with the following procedures of model selection: the authors first state that they performed a backwards stepwise model selection to retain the most parsimonious model (lines 277-278), and then that the best-fitting model was determined by comparing all previous models using the AICc. So it seems that they applied two different selection procedures on the same full/saturated model. While both approaches are valid independently of one another (although their respective use depends on the approach, e.g. exploration *versus* inference), and both aim at defining the final model with the fewest predictors, combining them seems irrelevant. Following the backwards stepwise model selection, the authors were already supposed to get a reduced model that best explains the data.

Then, I do not get either why they performed 12 ANOVA tests using the parameters of the most parsimonious PERMANOVA model (lines 282-284). I would have thought that performing a PERMANOVA including all markers or computing separate models for the 12 markers would be two alternative approaches and that they should chose one or the other. Again, they seemed to have proceeded the same way than before for model selection. Overall, this subsection on physiology needs simplification and clarification.

Results

Results are properly described and their interpretation makes sense based on the output of the current analyses, although some of these analyses are questionable, as stated previously. I have checked the raw data, but I have not rerun the statistical analyses.

I did not detect obvious manipulation of data apart from the stated removal of 12 outliers (lines 261-262), which I already commented on earlier. Overall, the statistical results seem to support the conclusion, although the significance is not always very strong, especially for analyses on body and physiological condition where p-values are often between 0.01 and 0.05. It remains nonetheless that the authors found significant differences between sessions and/or species.

Discussion

The authors discussed all the hypotheses they had initially formulated, based on their own results and published research in the field (on both mice and other species). They thoroughly commented on the observed variations of body and physiological condition, as well as on the differences between species. For each item, they proposed one or several alternative explanations, and their interpretation is in line with their observations and previous findings in the literature. They also commented on the absence of variation for some markers (e.g. sodium, blood urea nitrogen), and considered results that went against their predictions (e.g. body condition, TBIL). Finally, they emphasized a few limitations of the study and suggested a few avenues for future studies, which is nice. I have no major concern on this section, although I think that the structuration could be improved by reorganizing some paragraphs.

Line 362. Please specify here that the sister species were rodent (or at least mammal) species, to clarify the context.

Line 365. The authors state here that the study took place in a relatively wet year, indicating “La Nina” within brackets. Unless I’m mistaken, they did not mention the role of the “El Nino Southern

Oscillation phenomenon" in driving South Africa climate in the introduction, and they describe it only after talking about "La Nina" (which should have been the opposite to promote non-familiar readers comprehension). It is an important feature for the study, so I suggest introducing this phenomenon from the introduction.

From lines 375 to 379, the authors sum up their main results. Such a concise summary paragraph is useful at the beginning of the discussion, but comes a bit late here (since it is placed after another longer paragraph on climate and habitat quality). I believe it would be more relevant to move up this paragraph and to combine it with the first one, which sums up the main questions of the study.

From lines 380 to 394, this paragraph on body condition also deals with seasonal variation of physiological condition to some extent, so I see no reason not to include it the sub-section named "seasonal variation in physiology" (which should then be renamed "seasonal variation of body and physiological condition", for instance). Otherwise, as it is, it is a bit weird to have most of the discussion shared between 2 sub-sections (i.e. "seasonal variation in physiology" and "interspecific differences") but a few paragraphs out of any sub-section. An alternative would consist in creating another subsection called "(seasonal) variation in body condition", which would include the current third and fourth paragraphs.

In line with my previous comment, the last paragraph of the sub-section on seasonal variation in physiology (lines 484-488) deals with body condition, which has already been discussed earlier. For more consistency, and to avoid going back and forth, I suggest grouping all information on body condition within a single sub-section, or at least moving this paragraph closer to the other ones also dealing with body condition.

At lines 522-524, it is written that overall differences between species can be accounted for by lower blood concentrations in *R bechuance*, and on average higher blood concentrations in key metabolites in R. d. *dilecticus*. However, the concentrations involved in the first part of the sentence are not specified, and the key metabolites are not listed either in the second part. Thus, I find it hard to keep track. Please specify which concentrations you are referring to.

Line 526. I am under the impression that "rather than" would better fits than "although" in this sentence, unless I misunderstood it.

Line 547. Replace "is" by "being" in "starch is the best substrate for […]".

From line 552 to 566. This last paragraph is a bit out of topic within the sub-section on interspecific differences, as it has a broader scope. Its ideas are more in line with those presented in the conclusion section, part of which does not fall within the scope of a conclusion *per se*. Thus I wonder whether it would be preferable to merge the text of the last 2 paragraphs (or at least most of it) into a single subsection named "perspectives" (or "perspectives and conclusions" if the entirety of the text was to be merged).

<u>Tables and figures</u>

The manuscript includes a lot of tables and figures, of which some are presented as supplementary material. I believe a few could be merged to reduce their total number.

Table 2. It is said in the legend of this table that values in the column "Aridity Index" correspond to the average index ± standard errors. Yet, in the table, the symbol "±" does not appear between the 2 values, only a "-". Please correct it.

Tables 3, 4 and 5. Titles and captions are not detailed or specific enough to understand these tables without reading the manuscript. More precision is required about the analysis/model that was performed in each case. In particular, tables 3 and 4 have the same title except for the name of the response variable within brackets, and the term “physiological response” in table 5’s title remains vague and uninformative.

Figure 1. I personally do not see the difference between translucent and full color dots on the map.

Figures 3, 4, 5 and 6. Overall, labels on y and x axes are too small, especially compared to the heading of each panel.

Supplementary table 4. Please provide the significance of the different acronyms in the caption (e.g. ALB, ALP, ALT, etc.).

Supplementary methods. In the section about the Aridity Index formula, the reference for “PET by Thornthwaite method” is missing (i.e. [REF]).

Supplementary table 2. Is this table really useful? Mean value for trapped surface ± standard deviation (and min-max values) could have simply been indicated in the main text, in the sub-section on sampling periods and sites. This would allow the reader to get the main information without having to consult the supplementary material.

Supplementary figures 2, 3, 4 and 5. It would be nice to plot on the same graph both the habitat structure variables and the traps position (i.e. biplots, see Fig 3 in Greenacre et al. 2022 *Nature Reviews Methods Primers* for an example), so that the reader could easily see how sites and traps differ in vegetation structure. This would consist here in combining figures 2 and 4 on one hand, and figures 3 and 5 in the other hand. Not only this would contribute reducing the number of figures, it would also improves the visualization of inter-site differences.