Dear Dr. Calcagno,
This letter accompanies our revised preprint “Direct and transgenerational effects of an experimental heat wave on early life stages in a freshwater snail”. We are grateful for your and the reviewers’ helpful feedback, and we have revised the preprint according to them. Most importantly, we have clarified the conducted statistical analyses, restructured parts of the preprint (e.g. discussion on the potential role of selection in explaining our results), homogenized the figures and combined some of them. Below we explain the changes made in our revision. The text from the reviewers is in Roman font and our responses are in italics. References to the line numbers are for the revised preprint.

Sincerely,
Otto Seppälä

I have read the preprint and had it evaluated by three expert reviewers. All reviewers expressed concerns related to the presentation of the results and some aspects of experimental or statistical methods that could not easily be parsed from the manuscript. Importantly, all reviewers pointed that egg number/size is not shown and very little analyzed, and that it would be important to do so because egg size and egg number might show correlated responses in the context of a size-fertility tradeoff. This may relate to a, possibly adaptive, shift in the egg-laying strategy of females as a function of temperature, and deserves consideration.

> *We agree with the reviewers and the recommender that this is an interesting question that definitely deserves further attention. We, however, do not believe that our data is suitable for testing this. In this experiment, estimates for adult fecundity were not quantified. We used only one egg clutch per parental snail, and we do not think this is enough for estimating snail fecundity. This is because both the rate at which snails lay egg clutches as well as the size of individual clutches show high variation. Therefore, estimating snail fecundity requires quantification of reproductive output over several days.*
> *Furthermore, we could not use egg clutches with low egg numbers in this study (L 191-193). This was because we needed higher level of replication for the analyses on offspring performance that were the main focus of the study. Because of this, the suggested analysis would be confounded if conducted using the available data. We believe this question can be addressed in future experiments.*

There were also several concerns regarding the statistical procedures and interpretations. These should be clarified (possibly, as suggested, with the help of an additional Figure). Similarly, the comparison of effect sizes between direct and indirect effect should be moderated somehow, since not all traits were not equally affected.

> *Responses to all these points are provided below when responding to the feedback from individual reviewers.*

I agree with the reviewers. Personally,I recommend the authors to homogenize the different result figures: whilst several follow the exact same pattern, some are presented as barplots, some as simple dots, some show mean+-SE, some show median/quartiles. This is confusing. I

even think that figures are pretty small and contain one panel only, and that they would benefit from being lumped (at least, Figs 1&2 and 3&4 constitute pairs of companion panels). This way, they might get some extra space for the reporting of the additional results and figures that reviewers have asked for.

> *We have now homogenized the figures so that we use circles to represent means in all figures. The original figure 2 was obviously confusing as commented also by reviewer 3. The aim of this figure was to summarize the patterns in hatching, but seems like we did not reach that goal. Now we present the onset of hatching, median developmental time, and the end of hatching in separate plots of figure 2. We have also combined the original figures 3 and 4.*

Last, the manuscript needs some restructuring, and the reviewers made several suggestions to this end. For instance, some parts are found in the Material & Methods section while they really are Results material (I am thinking for instance of survival and fertility values, provided on page 7 in the Methods, and that one reviewer found missing in the Results). An additional figure highlighting the general experimental setup would certainly help.

> *We have considered the suggested changes in the structure of the manuscript. We agree with the reviewers about some, but not all, of them. Responses to all these suggestions are provided below when responding to the feedback from individual reviewers.*

Considering all these elements, I cannot recommend the preprint as it is, even though it addresses an interesting topic and reports some nice results. If the authors can, and are willing to, address all the points in the present letter and in the review documents, the preprint may certainly become recommendable. In any case, I hope these reviews will be of some help to the authors.

Reviewer 1

Review of “Direct and transgenerational effects of an experimental heat wave on early life stages in a freshwater snail”.

Overall comment:
Investigating the role of transgenerational effects in the context of extreme climatic events such as heat waves is certainly an important and currently understudied topic in ecology. A strength of the study is that it compares the magnitude of transgenerational effects versus the direct effects of temperature on offspring performance which is rarely analysed in the context of thermal effects. I appreciated the general idea of the manuscript and the amount of work invested by the authors to produce these interesting results. The Introduction and Discussion are well written. However, I have some concerns about the link with previous studies (Leicht et al. 2013: Leicht et al. 2017), the statistical analyses, the result interpretation and the possible selection effects of temperature on both the parental and offspring generations. I detail these concerns below.

On lines 150-152, the authors refer to two studies published by the first author in which the effects of temperature on egg numbers were analyzed (Leicht et al. 2013: Leicht et al. 2017).

As these studies used a similar experimental design, it is not clear if the present study involve another set of “new” experiments or if it analyses unpublished data from these previous studies. In my opinion, this is not an issue but it should be clarified. I suggest adding a sentence to clarify this point.

> *This and our earlier experiments cited in the text are all separate studies. We have now clarified this by rewriting the sentence the reviewer refers to. It now reads “We did not measure the number of oviposited eggs in this experiment as the effect of temperature on snail fecundity has been described in detail in earlier studies (Leicht et al. 2013; 2017)” (L 174-176).*

An important point related to this is the link between egg numbers and egg size. We may expect a trade-off between egg size and egg number: the bigger they are, the less abundant they are. By ignoring egg number in the present study, we miss a part of the story. For instance, producing smaller eggs but a larger number of them could be an adaptive response to warming. It would thus be interesting to analyses the relationship between egg number and egg size at the two temperatures.

> *See the response to the recommender’s first comment above.*

The statistical analyses section is very vague and should be improved to better understand which statistical analyses were computed. For instance:
Lines 195-198: if the ANOVA has a random factor, then it is a Mixed Model ANOVA.

> *Corrected accordingly throughout the manuscript.*

Lines 199-205: Why family was not included as a random factor in this analysis? Specify that the interaction between maternal temperature treatment and offspring temperature treatment was included in the model.

> *In this analysis, egg clutch was the unit of observation (L 234-236). One egg clutch per maternal snail was used in the experiment. Family (i.e. egg clutch ID) could be used as a factor only in those analyses where an individual offspring was the unit of observation. The use of the term “family” may be confusing here. Therefore, we have replaced it with the term “egg clutch” throughout the manuscript. The included interaction term is specified in the text now (L 237).*

Lines 209-217: I have the same questions about the random family effect and the interaction for the MANOVA analysis.

> *Egg clutch was the unit of observation also here (L 244-248). The included interaction term is specified in the text now (L 250).*

Lines 218-227: a GLM with a random effect is a GLMM. Same comment as above for the interaction term.

> *Terminology is corrected accordingly throughout the manuscript. The included interaction term is specified now (L 259-260).*

Analyses on offspring survival: I was surprise to see that the interaction between maternal temperature and offspring temperature was not significant. Figure 3 suggests that it is significant. In table 2 (GLM for offspring survival), it is not clear if family is a random factor or a fixed one. If it is fixed (as in a GLM) then this might explain why the interaction term appears non-significant.

> *The referee is correct that in figure 3, the effect of maintenance temperature on snail survival shows tendency to depend on maternal temperature. Considering the strength of this effect and the variation within treatment combinations this effect is not statistically significant. Egg clutch (i.e. family) is treated as a random factor, which is now clarified in the table legend.*

I also noticed that, in the control group (parents and offspring maintained at 15°C), survival is lower at the offspring generation compared to the parent generation: approximately 80 % vs 94.4%. Why is it the case? Maybe it is because I am comparing survival of juveniles versus adults?

> *We are not aware how much adult and juvenile mortality differ in our study species, and we have never conducted an experiment to test that. Snail survival may be age dependent, which could explain the observed difference at least partly. In this study, however, it is important to note that the survival rates the reviewer refers to are based on time periods with different lengths. Adults were followed for one week whereas offspring were followed for five weeks. The observed higher mortality in juveniles may thus be because of a longer observation period.*

Shell length: Table 3 suggests that family was included as fixed effect in the model. It should be random unless the authors are interested in discussing family effects.

> *We have clarified the table legend to avoid confusion. Furthermore, the table indicates which terms of the model were used as error terms in significance testing. Also that indicates that egg clutch (i.e. family) was used as a random factor.*

The increase in shell length in response to warm temperature (direct effect on offspring) is spectacular (97.4% increase in size). It is also opposite to the temperature size rule (i.e. smaller individual at warmer temperature). I wonder how much of this result is linked to plastic versus selection effects. The effect of family is very strong (Table 3). Is this because there was a selection for families with large snails? It would be interesting to look at the size of snails that died both at the parent and offspring generations to determine if body size is important for survival at warm temperature (if yes, we expect that the dead snails were smaller). If there is a selection for larger size, this might explains some of your results. This should be clarified, especially because the low survival of adults at warm temperature (66.1 % survival) suggests possible strong directional selection that could strongly contribute to the results.

> *We do not believe we can link our data to the temperature size rule the reviewer refers to. This is because we have not measured the final size snails can reach when adults. Our measurements from juvenile snails are more likely to reflect differences in growth rate owing to variation in metabolic*

activity. Thus, it could be that at high temperature snails simply reach their final adult size earlier. Whether or not final size depends on temperature should be addressed in future studies.

Examining the potential role of selection in determining our results is a great idea. Unfortunately, we do not have any measurements collected from the parental snails that could be used to evaluate if they affected snail survival or probability to reproduce during the egg laying period. Similarly, we measured the size of only those juvenile snails that survived to the end of the experiment. Measuring juvenile growth rate earlier during the study would have made it possible to examine how variation in growth affects snail survival. Although we were not able to test the role of selection in our study we have discussed this more thoroughly and moved this part to the discussion section (L 367-378).

General conclusion: the conclusion that transgenerational effects and direct effects of temperature are equally strong should be lowered as it is not general but depends on the traits. For instance the direct effect of temperature on shell length is much stronger than the maternal effect (Fig. 4). The conclusion should thus be that transgenerational effects and direct effects of temperature **can be** equally strong **depending on the trait considered**.

We have now specified throughout the text that whether or not transgenerational effects are equally strong to the direct effect of temperature depends on the examined trait (e.g. L 44-45, 127-128, 365-366, 445-446).

Overall, I think that it is an interesting study that deserves to be published after clarifications on the statistical analyses and on the relative importance of the potential selection effect. Investigating the relationship between egg size and egg number would also be an interesting addition to this study.

Minor comments:
I was surprise that some results and element of discussion were already included within the experimental design section (lines 144-158). I think this information is important and the results should go to the Result section. The discussion on potential selection effects should go to the Discussion.

This is a difficult matter. We had two reasons for giving this information already when describing the experimental design: (1) provided "results" are not direct results for our study question, and (2) colleagues who commented the manuscript before its submission raised the question about the suitability of the design already in the methods section. Earlier, we had included this information into the results and discussion sections as suggested by the reviewers now, but spreading the information into different sections did not satisfy our earlier readers who wanted to get answers immediately when they came up with this question. Because of the conflicting feedback from different readers, it seems like we cannot reach a solution that satisfies everyone. However, because all reviewers as well as the recommender suggested revision we have done that hoping that the majority of our future readers would prefer this order. First, we give exact numbers of snail that survived and laid eggs in the methods section (suggested by reviewer 2) (L 173-174), we give proportions and results from statistical tests in the results section

(suggested by all reviewers) (L 269-277), and discus the potential role of selection in the experiment in the discussion section (suggested by everyone) (L 367-378).

Line 57: replace “cannot” by “can not”

Corrected accordingly (L 80, also L 363).

Line 101, 103, 119, 120: “2015” can be removed as the data are unpublished.

Corrected accordingly (L 136, 137, 165, 167).

Line 125: “large numbers” How many?

We maintained the stock population in the approximate size of 400 individuals. This is now clarified in the text (L 142).

Line 112: define “high temperature”. e.g. >20°C.

We refer to temperatures equal or higher than 25°C. This is now clarified in the text (L 117-118).

Lines 159-160: It would be useful to clarify that temperatures were kept at 15 and 25°C even after the initial exposure to temperature treatments. At the first sight, one may think that temperature treatments stopped here after seven days.

We have clarified this. The sentence now reads “After the initial seven-day exposure to temperature treatments, we removed all egg clutches oviposited by the snails from the cups and continued maintaining the snails under the same experimental conditions. During the following ten days, …” (L 184-186).

Lines 182-185: specifying that the design was full factorial (2 maternal temperatures x 2 offspring temperatures) would ease the understanding of the experimental design.

This is clarified in the text now (L 207-209). We have also added a new figure (Fig. 1) that shows the experimental design.

Lines 234-235: specify that the interaction between maternal temperature and offspring temperature was not significant.

Corrected accordingly (L 281-282).

Reviewer 2

Leicht & Seppälä subjected two generations of a freshwater snail to above optimal, but realistic, short term changes in temperature to determine the effect of “heat waves” on offspring. Using a full-factorial experimental design, they found evidence of transgenerational effects on offspring investment and performance. Importantly, they found

that some maternal effects occurred at the same magnitude as the direct effects of temperature, indicating that they should not be overlooked.

While I feel that this paper has the potential to be a valuable contribution to both the fields of climate change and maternal effects research, there were a few aspects that need improvement and clarification. That being said, they are all (most likely) minor issues and should not thwart the recommendation of this paper.

Introduction

I may be interpreting the results from Vaughn 1953 incorrectly, as I only gave it a cursory read, but it appears that the analysis in that paper was done on juvenile snails, not adults. Does it make sense to consider their research applicable to adult stages as well?

> *The reviewer is correct that the study by Vaughn (1953) focuses on juvenile snails. Because also our study focused on juvenile snails, using the information from that paper when choosing the experimental temperatures is highly relevant. However, we did not choose the experimental temperatures solely based on the results by Vaughn. We had also other lines of evidence to support our choices. These were (1) our own results considering thermal performance of adult snails that originated from the same study population as the individuals used in the current study, and (2) measurements of water temperature in ponds in our region. All this information was used when deciding about the treatments used in the experiment. We have now clarified this in the methods section (L 162-167).*

While it may not be necessary to provide in full detail all the hypotheses and predictions regarding the effect of temperature on offspring investment and performance, some more background on what was been previously found for these snails might provide the reader with some context for what might be expected.

> *We now provide information about the effects of high temperature on adult snails to prepare the reader for its potential transgenerational effects (L 117-121). See also the response to the comment by reviewer 3 to the structure of the introduction.*

Methods

A map of the region that the stock pond is representative of, highlighting the location of the stock pond, would be useful in determining the scope of the research.

> *We agree with the reviewer that maps can be highly useful in ecological research. This is especially the case when several populations that originate, for example, from different regions/environments are compared. Such a map may not only present the locations of the used populations, but also summarize the key factors in the environment that differ between them. Our study used a laboratory stock population that originated from one location. Thus, we do not see that including a map into the manuscript would be highly beneficial. It is also unclear to us what kind of information such a map should include about the environment. We, however, see that we did not provide enough information that would allow readers to easily locate the pond. We have now provided more detailed coordinates of the pond (L 134) so it cannot*

be mixed with any other pond. Provided coordinates also allow readers to evaluate any aspects of the environment if they so desire.

L128, the number of snails haphazardly collected (113) does not line up with the number of snails that were assigned to the temperature treatments (56 + 60). Also, why subject 56 snails to one treatment and 60 to another? Were there mortalities or other issues that prevented you from having equal numbers in each treatment?

The latter numbers were incorrect and are now corrected (L 160-162). We exposed more parental snails to high temperature compared with the control temperature because we expected higher mortality in those snails. This is now clarified in the text (L 167-168).

By placing the snails from the maternal generation in perforated cups, hence the same water circulating between all the cups, is there the possibility that hormonal or other chemical cues could be affecting the study? I noticed that this is the only treatment with the use of perforated cups.

We are not aware of such hormonal effects but can also not exclude their potential importance. The reason we chose to maintain parental snails in a water bath was to provide maximal water quality for them. Adult L. stagnalis snails consume lots of food and thus release high amounts of nutrients into the water. If adult snails are maintained in jars with low volume, high nutrient levels lead to massive growth of microorganisms that activate snail immune function (Seppälä & Leicht 2013, Journal of Experimental Biology). This could possibly affect snails' reproductive strategies as well as the quality of produced offspring. Such effects could be highly problematic in the current study because water quality deteriorates more quickly at high temperature. The use of water baths with biological filtration ensured high water quality for all snails. This justification is now given in the text (L 148-151).

You never mention the possibility that some snails may not have eggs, you allude to it by saying they can store eggs so they don't need a mate in the cup, but is there a chance that they haven't stored eggs? It is unclear to me if it isn't mentioned because it isn't possible, or because it wasn't considered.

The sentence pointed out by the reviewer does not refer to storing eggs from previous matings. The sentence aims to explain how snails can fertilize the eggs they oviposit under the used experimental conditions. Lymnaea stagnalis snails can reproduce through both outcrossing and self-fertilization. First, when mating, snails receive sperm from other individuals that they store. During the next few months, snails can use this sperm to fertilize eggs that they oviposit. This leads to production of outcrossed offspring. Second, if snails have not mated or if they choose not to use the stored allosperm they can self-fertilize their eggs. Therefore, maintaining snails individually during experiments like the one described in this manuscript does not prevent them from ovipositing eggs. We have tried to clarify this in the text. The sentence now reads "Because L. stagnalis snails can reproduce through self-fertilization as well as through outcrossing using allosperm they have stored from previous matings (Cain 1956; Nakadera et al. 2014), experimental snails

did not need a mating partner to oviposit eggs under the used conditions" (L 153-156).

Line 150 – effect on reproduction-do you mean fecundity? There could possibly be a trade-off between egg size and number that should at least be mentioned.

The term "reproduction" has been replaced with "fecundity" (L 175). Please see the response to recommender's comment on trade-off between egg size and egg number above.

L152-158-This justification could be condensed to come across less vague. It may also be better suited for the Discussion.

This justification is revised and moved to the discussion section. See more detailed responses to similar comments by reviewer 1.

L-176 Could this be included as a variable in the analysis?

We have tried to use this variable as a covariate in the analyses. It, however, did not work out. First, the variable is not properly continuous as it has only 10 levels. Second, variation is not similar between the treatment groups, which makes evaluating an important assumption of equality of regression slopes impossible. Therefore, we have only mentioned this factor as a potential additional source for unexplained variance in the study (L 202-203).

After the number of snails undergoing the maternal treatments is provided, the rest of the data is provided as proportions (or no information at all). It would be beneficial to provide the actual numbers (i.e. # of eggs/clutches, # of juvenile snails) as well. It can be deceiving when just proportions are provided. For example, L166- mean clutch size? L182 – how many clutches to each treatment? L189-This resulted in how many hatchlings at each temperature?

We now provide actual numbers for adult snails, egg clutches and hatchlings in the methods section instead of proportions (L 160-162, 173-174, 211-213, 216-218). Additionally, we have included a figure that describes the experimental design (Fig. 1). Sample sizes are provided also in that figure.

L203 – two clutches were not included, without knowing how many clutches there were to begin with, it is difficult to determine the impact of this exclusion.

Information considering the treatment combination from which these clutches were removed is provided now (238-240).

There were a few transformations performed on the data (e.g. L195, L210, L212, L224) without reasons provided.

Transformations were used to homogenize error variance. This information is now given after each transformation is reported (L 229-230, 245, 248, 261-262).

L-227 Were the excluded snails equally distributed between treatments?

The excluded snails were not equally distributed across the treatments. They were excluded because of human errors (L 264) during the maintenance and measurements. Thus, they were distributed randomly across the treatments.

Results

A graphical illustration of the results would be beneficial for understanding the effect of the different treatments; the factorial design of the experiment would make this quite "easy" to do.

Unfortunately we do not understand which graph the reviewer refers to and what kind of changes are suggested. We have made several changes in the figures to make them easier to follow. Those changes are explained in our response to the recommender's comment on the figures. Additionally, we have included a figure that describes the experimental design as suggested by the reviewer 3.

Discussion

While the magnitude of the transgenerational and direct effects of temperature were similar for hatching success and survival, there was a large difference for the other traits, which should be mentioned.

We have specified the text throughout the manuscript to clarify that this was seen in some but not all traits that were examined (see the response to a similar comment above).

L325 - *maturation

Corrected accordingly (L 392).

Reviewer 3

Review of the article "Direct and transgenerational effects of an experimental heat wave on early life stages in a freshwater snail" by Leicht & Seppälä for PCI Ecology.

General comments: The article demonstrates that both direct and maternal effects of temperature are involved in determining traits in a freshwater snail. Authors used a proper factorial design by switching two different temperatures to test such effects. Although they remained overall weak, maternal effects were identified on hatching success, onset of hatching, survival rate and size of the offspring. The abstract would benefit from adding more details on the experimental design and results so readers could better perceive what was done in the study. I believe the abstract is not clear enough as it stands for now. Overall, the structure of the manuscript is a bit confusing. Some information is not provided in the right section, to my opinion (see comments below). It is often difficult to identify parts refereeing to experiments on direct effect, maternal effects, and offspring effects. A summary figure of the experimental design in the material and methods would greatly help! I have spotted several grammar and syntax mistakes throughout the text, but as I am not myself a native English speaker, I would just suggest the authors to revise carefully the use of English in the manuscript.

We have added a figure that describes the experimental design and also specifies the sample sizes used in different steps of the study (Figure 1). Responses to other points are provided below when replying to the reviewer's specific comments.

Specific comments:
L28: I would replace "completely" by "often" as it is not true to affirm that transgenerational plasticity has been completely neglected in the context of climate change.

The sentence is reworded to increase clarity (L 26-28).

L28-31: This sentence is not clear, please rephrase. Do you already know that high temperatures reduce adult performance, or is it something that you will test? Also please write it at the past tense.

This sentence is split into two sentences to increase clarity (L 28-33). The effects of high temperature on adult snails are not the focus of this experiment, but known based on earlier studies. This is now clarified in the text (L 29). We asked a colleague who is a native English speaker about the tense. She suggested not to change it.

L30: "which traits are affected" in the offspring, the maternal generation or both? Precise what kind of traits you are measuring (morphological, physiological ...).

In the sentence the reviewer refers to we now specify that we focus on eggs and hatched juveniles (L 31). We do not believe we can specify the examined traits in this sentence. If that was done, we would lose the benefits gained by splitting the sentence into two separate ones. Based on the previous comment, we think we should keep the sentences short. The traits are described in detail in the next sentence.

L31: "with direct effects of high temperature" Here you are talking about the offspring, right?

This is specified in the text now (L 33).

L37: "Direct effects of high temperature on offspring" from both maternal origins?

We refer to the direct effects of temperature on different traits. This is now clarified in the text (L 38-40).

L38-39: I am not sure it is worth insisting on this similarity in magnitude as the relationship between direct and maternal effect is in one case negative (hatching rate) and in the other case positive (survival).

We do not see why the magnitudes of direct and transgenerational effects could not be compared if the directions of those effects are not consistent among traits. We think that maternal effects are highly important to understand in both cases. If the direction of direct and transgenerational effects is the same the overall effect would be doubled (assuming the magnitudes are the same). If the directions were opposite, the effects would

override each other. Our main point is that transgenerational effects are highly important to consider because they can be strong. Neglecting them when evaluating overall effects of environmental change can lead to false conclusions.

L39-41: I would reverse the order of this sentence so the focus would be made on the importance of transgenerational effects. It could read “This indicates that heat waves cannot only impact natural populations through direct effects of temperature, but that such effects can be equally strong to maternal effects”, or something similar.

We have modified the sentence to increase clarity. It now reads “This indicates that heat waves can affect natural populations through transgenerational effects, and that the magnitude of such effects can be equally strong to the direct effects of temperature, although this depends on the trait considered” (L 42-45).

L42: Replace climate change by climate warming as you focus on temperature only

Corrected accordingly (L 46).

L59-61: This is what evolutionists do. It would require specifying that you are studying transgenerational plasticity (and not long-term evolution).

To avoid giving an impression that we refer to evolutionary change the sentence has been reworded as “Hence, for understanding the effects of climate change on natural populations, studies examining such transgenerational effects are needed” (L 82-84).

L73-74: This is only true if the offspring environment is predictable by the mother environment. Otherwise, bet-hedging strategies could appear, or the “predictive” maternal effect would not be fully adaptive.

We fully agree with the reviewer. That is why we had written “Furthermore, maternal effects <u>can be</u> adaptations to prepare offspring for the future conditions they <u>are about to encounter</u>”. We prefer not to expand the introduction to bet-hedging strategies because the experiment was not designed to test them.

L86: It would be interesting to have information (if available) on the relationship between temperature and egg size in invertebrates and/or in the study species. What are the known factors determining egg size and embryo development in this species?

See the response to the recommenders first comment above.

L93: Please give the authority and (Order: Family) for the species the first time you mention it.

Information considering the authority and taxonomic group are given now (L 116-117).

L91-103: Most of the information here should be moved to the material and methods section. Instead, hypotheses are missing at the end of the introduction and should be clearly stated. What do the authors expect to see on direct and maternal effects and on which traits? Why?

The way how the end of an introduction is written varies a lot among articles, and there is no general agreement how it should be done. Some writing guides recommend stating the main study question and the approach (Davis 2005, Katz 2006) whereas some highlight the importance of a summary of the results (Montgomery 2003, Day and Gastel 2006, Heard 2016). Our approach is to present (1) the main study question, (2) the approach and (3) summary of the results. With this our aim is to very briefly give all the key elements that then hopefully motivate the reader to continue to other parts of the article. The reviewer is correct that the information we have given at the end of the introduction overlaps with the information that is provided in the abstract. We, however, see that it is beneficial to remind the reader about these aspects in the introduction after the "big picture" has been explained in the previous paragraphs.

Some journals have requirements for the structure of the end of the introduction. Because such requirements do not exist for preprints we feel that we can keep the structure that was explained above. However, we have moved the information considering the chosen experimental temperatures to the methods section (L 162-167) as we agree that those details are not necessary in this part. Additionally, we now provide information about the effects of high temperature on adult snails to prepare the reader for its potential transgenerational effects (L 117-121).

Montgomery 2003. The Chicago guide to communicating science.
Davis 2005. Scientific papers and presentations.
Day and Gastel 2006. How to write and publish a scientific paper.
Katz 2006. From research to manuscript: a guide to scientific writing.
Heard 2016. The scientist's guide to writing.

L99: Is there any more recent reference than Vaughn (1953) describing the thermal optimum of this species? In 65 years, it is highly probable that selection would have acted on thermal optima of Lymnaea populations. If nothing is known about current thermal optimum in this species, this is a point that should be discussed in the manuscript. To the same extent, artificial selection could happen in the laboratory, as snails were maintained at constant 15°C for 2 years before the study. It was shown in some (insect) species that thermal plasticity can be highly reduced when maintained over a long time or over several generations at constant temperatures (i.e., it has a narrowing effect on the thermal optima curve).

We are not aware of a more recent paper that could replace the reference to Vaughn (1953). The reviewer is correct that results from that paper may not represent thermal optimum in our snails. This is not only because the study by Vaughn was conducted a long ago, but also because it used another snail population. The experimental temperatures used in our study were, however, not solely based on the results by Vaughn. We had also other lines of evidence to support our choices. These were our own results considering thermal performance of adult snails in the same population the individuals used in this study originated from, and measurements of water temperature in ponds in

our region. All this information was used when deciding about the treatments used in this experiment. We have now modified the text to clarify this (L 162-167).

It is inevitable that selection operates also under laboratory conditions. Our stock population was, however, not maintained over a long time period before the study. We used F_4 generation (L 133). We cannot evaluate how quickly plasticity may be reduced under our culturing conditions. However, if such reduction happened in our lab population that would not create differences that are artefacts, but instead make the conducted study more conservative.

L99: "reduce life-history" is not very informative. Which traits were affected? Also it should read "reduce the value of life-history traits" or "of life histories".

We aimed to refer to both life history traits and immune defence traits. To avoid confusion we have now changed the order (L 162-163).

L103-107: Please remove this part from the introduction. It is a summary of the results and is already mentioned in the abstract. It can be moved to the beginning of the discussion, if needed, to briefly summarize your findings.

See a response to a comment considering the structure of the end of the introduction above.

L111-115: This part should actually be in introduction.

Corrected accordingly (L 117-121).

L152-158: Again, I find it a bit awkward to discuss results and potential experimental bias before exposing the results per se. I would move this part to the discussion section. It has to be discussed in regards to results from the offspring generation. Figure 1, 2, 3 & 4: Please display on figures results of statistical analyses so we can see significant differences among treatments without referring to the text or to the tables.

Potential experimental bias is explained in the discussion section now (L 367-378). See detailed response to a similar comment by reviewer 1 above.

Whether or not statistical significance is presented in figures is a matter of taste. Some stats books even recommend dropping significance testing completely (Hector 2015, The new statistics with R). We also want to encourage readers to focus more on effect sizes and variation within treatment groups rather than p-values. Therefore, we have presented figures and statistical tests separately.

L218 and Figure 3: If daily survival data is available, it would be better to analyze and represent this data using Cox-regression models and survival curves. Using a GLM will only compare mean survival rates among treatments but cannot capture any time effect. Using a Cox model should not alter the conclusions. Survival % at 5 weeks can still be given in the text.

Daily survival data was not collected. This was because it is not relevant for juvenile snails at which point before reaching maturity they die. In case of mortality, they do not gain any fitness. The analysis suggested by the reviewer would be very important if adult snails were the focus of the study.

L221: The "family" effect in the model represents a mother ID effect, correct? Why did you chose to nest this effect within the interaction effect?

In our study species, eggs are laid in clutches that cannot be split to expose individual eggs from the same clutch to different experimental treatments. Therefore, each egg clutch (i.e. family) could be allocated into one maternal temperature by offspring temperature combination.

L235-236: Please provide precise data on how much hatching success was increased/decreased by increasing temperatures?

This information is now provided (L 282-283).

L236: What about the non-significance of the interaction term (M x O)? What does it mean biologically speaking?

It means that the effects of the examined factors affected independently of each other. This is now explained in the text (L 281-282).

Same remark at L239-240.

Corrected as above (L 284-285).

L241: What about differences in median and end of hatching between 25 and 15°C? Are there no significant differences? Please precise.

We were obviously not clear here. Our aim was to present the reasons leading to the results observed in MANOVA, which was the main analysis. We have now clarified this by writing "The effects of temperature on developmental time were first because offspring started to hatch 12.3% earlier when mothers had been exposed to 25°C (ANOVA: $F_{1,53} = 15.806$, $p < 0.001$; Fig. 3a). Second, the onset, median, and the end of hatching were earlier when offspring were maintained at 25°C (first day of hatching: 55.1% reduction; ANOVA: $F_{1,53} = 571.961$, $p < 0.001$; median developmental time: 43.8% reduction; ANOVA: $F_{1,53} = 189.817$, $p < 0.001$; last day of hatching: 43.8% reduction; ANOVA: $F_{1,53} = 62.002$, $p < 0.001$; Fig. 3)" (L 288-293). We have chosen to present test statistics only for those traits and factors that were statistically significant. That is to increase readability of this section.

L241-243: How much earlier? Please provide data in days for onset, median and end of hatching in the main text for as it is not precisely displayed in Fig. 2.

We provide these data in proportions now (L 288-293). Providing the magnitude of change in days would require reporting also a reference point

(e.g. number of days in 15°C treatment). That would be repetitive as the information is already provided in figures.

Figure 2: This figure has a standard display for representing mean±SE data and it could be confusing at first sight. It actually represents onset, median and end of the hatching period. Although the authors' choice makes sense when reading the figure caption, I wonder if a clearer way to display this data could be imagined. Maybe just adding text in the graph, or dashed lines to show onset and end of the hatching period would help. It would also help the reader seeing differences among treatments more clearly.

See the response to the recommenders comment on figures above.

L267-268: By how long survival was reduced in days and in percent? Did they die faster over the five weeks of experiment? A survival curve would allow showing this information better than barplots.

The magnitude of the reduction is given in the text now (L 315). Data about variation in survival over the course of the experiment was not collected. See a response to a similar comment above.

L269-270: Please move this sentence after Fig. 3 as it presents another type of results.

The sentence the reviewer refers to is the last sentence of the results section. We are not willing to move it into a separate paragraph because that would lead to a paragraph with only one sentence. In our opinion one sentence paragraphs should be avoided.

L267-270: What about the M x O effect and the family effect? Please add a sentence for the biological significance of these factors according to the presented results. It is important because it is the part that allows saying that the magnitude between direct and maternal effects is similar. Tables 1, 2 & 3: add "interaction effect" in the table legend.

The fact that the effect of high maternal temperature on offspring survival was independent of the temperature offspring were exposed to is now explained in the text (L 315-316). Family is included into the statistical models only because of the structure of the data; individuals within a family originated from the same egg clutch. This experiment did not aim to examine family-level variation in any of the examined traits, and in our opinion the data is not suitable for that. A study focusing on family-level variation should use individuals from a higher number of egg clutches per mother. Thus, our data does not allow separating variation among families and variation among egg clutches within a family. Therefore, we do not provide any biological interpretation for "family effects". Instead, to avoid confusion we have replaced the term "family" with the term "egg clutch" throughout the manuscript. Interaction terms are now added into the legends of all tables.

L274 and 288: Please briefly precise what is the "family" effect in the legend so readers don't have to refer to the material and methods.

See the response to the previous comment.

L305: “largely negative”. I think it is worth to precise here which traits were negatively affected by direct effects of high temperature.

We have specified both the negative and positive effects in the revised version of the manuscript (L 353-355).

L306-307: “early life stages”: what traits do you consider to be beneficially affected by high maternal temperature? “later stages”: same remark, please precise which “late” traits you found to be negatively affected.

Corrected accordingly (L 357-359).

L309: What is the rationale of using these references here? Pettay et al. is on humans, Heath et al. is a case study on salmons, and Mousseau & Dingle is a review focusing on insects. I suggest removing references from this part of the text and adding specific examples later in the discussion if and when relevant.

These references are chosen because they examine the role of maternal effects on traits expressed at different stages of organisms’ life histories. We do not know such studies on L. stagnalis or any other closely related species. Citations are removed and given later in the section “general conclusions”.

L309-311: Although the direction effect was reversed for hatching rate but not for survival levels, correct?

This is correct and clarified in the text now (L 360-363).

L316: Eggs were significantly smaller at 25°C, but only by 0.20 mm² on average. Is this difference biologically meaningful? What does it imply in terms of fitness?

Eggs produced at 25°C were 14.6% smaller compared with eggs produced at 15°C. We, however, do not know if and how much the reduced size contributed to hatching success of eggs or other traits that are important for snail fitness in later life stages. Our data also does not allow calculating a selection gradient for egg size. Therefore, we have only reported this effect and highlighted that the effect was different compared with hatching success (L 381-384).

L317: replace “benign” by “optimal”?

Based on the available information on the performance of L. stagnalis snails across different temperatures (Vaughn 1953) 15°C is not the exact thermal optimum. This is why we use the term “benign temperature”.

L318: Hatching success is affected, but is it really biologically important (a matter of 9% maximum)?

We see hatching success being highly important for fitness. Those snails that do not hatch have zero fitness. Therefore, an increase of 9% in a chance to hatch can make a big difference for an individual.

L323-328: Would faster development also be associated with lower risks of predator attack at the egg stage? Exposure to parasites? Would snails access to reproduction faster? Also please try to refer to literature on aquatic invertebrate systems, as references on vertebrates and homeothermic organisms may not be relevant for pond snails.

Fast development of eggs can very well reduce the risk of being exposed to natural enemies. We have included this option as a potential benefit into the discussion (L 389-390).

We are not aware of studies directly testing these effects in any closely related species. Therefore, we have chosen to cite studies that provide data from other species. In general, we have aimed to choose the cited references so that they are conceptually solid although focus on distantly related species.

L329-330: Is there any evidence of increased metabolic rate within eggs at high temperature in the literature?

To our knowledge, this is not known.

L331: Thus, could the effect of temperature on egg size simply be a plastic response to temperature constraints and not be adaptive? Is it a "maternal decision" or a response to temperature constraints? Are mothers able to lay different quality/type of eggs depending on the conditions they encounter (including temperature)? This is why we need information about determination of egg size and development in this species in the introduction of the paper. Also, egg size is often a good indicator of egg quality because it correlates well with energetic reserves. Here you show that you also have to consider potential trade-offs involving response to high temperature. I think there is a bit more to discuss about the advantages of developing in a small egg at high temperatures (resistance to heat shocks? More parsimonious energy consumption? ...).

We do not have any knowledge on that if and how strongly egg size is determined by decisions made by maternal snails and/or environmental constraints in L. stagnalis. We also do not have any other information considering the dependence of egg quality and environmental conditions than what is presented in this preprint (effect of temperature on hatching success).

Our aim was not to present reduced egg size under high temperature as a definite adaptation. We only wanted to mention reduced size potentially being beneficial when oxygen concentration is reduced. To avoid confusion we have reworded this part of the discussion and included a sentence about typically higher energy/nutrient level of large eggs (L 395-399).

L334: What do you mean by "resource level"?

This term was not necessary and is now removed to avoid confusion (L 402-403).

L337: This second hypothesis would require females to have the choice in laying high quality versus low quality eggs depending on the temperature or mortality risks. Can they?

> *It is not known whether L. stagnalis can choose to lay eggs with certain quality. The hypotheses presented here are only potential explanations for the result. We cannot test their actual relevance.*

L351: Is egg size correlated with offspring size? It would be expected. If so, reduced offspring size at high temperature could be explained by reduced egg size.

> *In this study, random eggs as well as random hatchlings were measured for size. Individuals were selected for these measurements independently of each other. Thus, we cannot analyze the potential relationship between egg size and hatchling size [we do not know which (if any) individuals were measured at both steps]. This would be worth to analyze in future studies. Before that, however, we need to find a way to identify which snail hatched from which egg, which is tricky because eggs are laid in clutches and cannot be separated form each other without harming them.*

L369: Higher temperature usually fastens metamorphosis rate (or organ development rate) but not growth rate, which leads to smaller adults in arthropods, or smaller individuals hatching from eggs. Does this temperature-size rule (see Atkinson, 1994) also apply to aquatic snails? Concerning survival rates, increasing metabolic rates and faster use of energetic reserves could also explain high temperature effect.

> *See the response to a similar comment about temperature-size rule above. Faster depletion of energetic resources is a highly valid point and is included in the text now (L 439-440).*

L381-384: This paper also show unexpected similarities in the magnitude of direct and transgenerational effects, for example hatching success increases at high temperature but decreases when mothers were exposed to high temperatures. Therefore, and as mentioned for the abstract of the paper, I would be more parsimonious in insisting on this “equally strong” effect.

> *See the response to a similar comment above.*

L385: In this paragraph, the authors should temper a bit their claim. Indeed, it is unlikely that climate warming change pond and lake temperatures by +10°C in a short period of time (one generation). Thus, transgenerational plasticity is not the only factor to take into account, but also genetic evolution over several generations. Could it be a mother/offspring conflict on fitness in the context of climate change? If mothers encounter conditions that will not be those that offspring will develop on, the maternal response could not be adaptive.

> *Ambient temperature can change very quickly in freshwater habitats. In our study system (ponds) temperature can show large and rapid changes few times during the life span of an individual snail depending on the occurrence of heat waves. Such extreme effects are becoming more frequent because of climate change. To avoid confusion we have reworded this part so that we speak specifically about heat waves rather than generally about climate*

change (L 466-468). Our data does not allow concluding anything about long-term evolutionary changes in snail populations.

Reviewer's point that offspring may not encounter the same environmental conditions as their parents is also our key point. This is presented few lines later after explaining the lack of interactive effects between maternal and offspring temperature in our study (L 473-475).

L397-398: "none of the observed direct effects of temperature depended on the maternal environment". It is not clear what this sentence refers to. No interaction effect? Please detail a bit more, as it is interesting.

Reviewer's interpretation is correct. This is clarified in the text now (L 471-473).

L400: Could it be that water environments are way more buffered than terrestrial environments, as you mention? Thus, maternal effects are unlikely to evolve if maternal and offspring environments have a high probability to be similar. It is also possible that other environmental factors fluctuate more than temperature does in such environments. Maternal effects could thus be much stronger when looking at resource availability, pH, ...

Terrestrial environments are likely to be less buffered than freshwater environments. However, freshwater environments are likely to be less buffered than marine environments that contain a much higher volume of water. Our discussion compares only freshwater and marine environments because the conducted experiments focus on species living in those habitats. Contrary to the reviewer's suggestion adaptive maternal effects are most likely to evolve in systems where environmental conditions experienced by parents provide reliable cues to predict environmental conditions offspring are exposed to. This is what we state at the end of the discussion (L 473-475).