COMMENTS TO THE EDITOR

Comment: I have now received comments from the original referees that commented also on the initial submission. Both referees found the manuscript substantially improved. They both made a number of smaller comments that would be worth considering.

Reply:

Dear Ulrich,

Thank you for the helpful reviews and your own comments. Below we address each of them one-by-one; comments are in black text and our replies in blue.

With best wishes,

Isabel & Jacques

COMMENTS TO REVIEWER 1

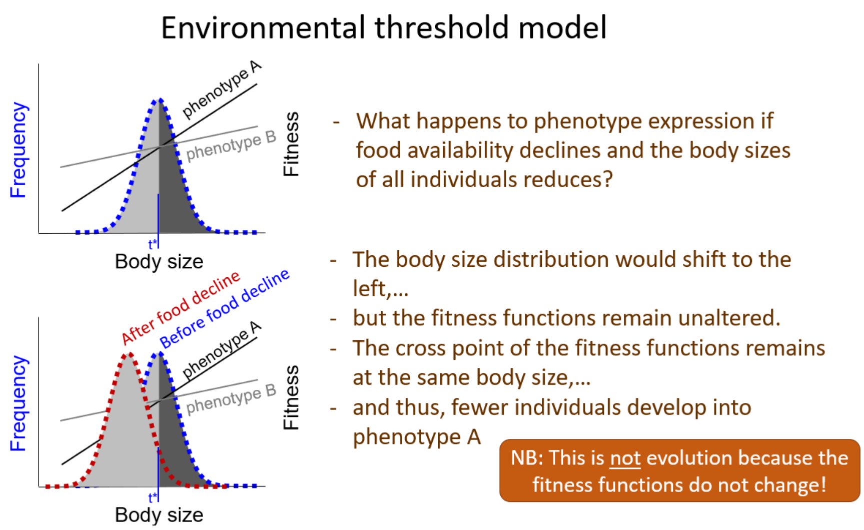
We would like to thank the reviewer for their comments, which have contributed to clarifying key aspects of the manuscript.

Comment: Line 41: It's not obvious what is "population's structure" here. Is it population stage/age structure?
Reply: Yes, this is age or stage structure. We have added this to the text (L38).

Comment: Line 82 onwards: Perhaps, the relation among environmentally cued threshold trait, individual liability and its distribution, and fitness functions could be written in a clearer way.
Reply: The text in its current form is, in our view, the clearest representation of the ideas we are conveying and uses the same explanations as from one of our previous published papers.

Comment: Question: Does the cue of the environment correlate well with body size in bulb mites? If so, can the threshold actually evolve through changes in individual liabilities and the environmental cue, in bulb mites?
Reply: In case of bulb mites, the cue is condition, which is correlated with body size. We mention this in Line 284, referring to Rhebergen et al. (2022), who present these results.

Comment: Question: If threshold can evolve in response to different selection pressures which is implied here, can liability distribution be affected by poor diet which perhaps can affect fighter expression even in a single generation?
Reply: Yes, the liability distribution can also change plastically. For example, in our case, poor diet can reduce overall condition and move the condition distribution, which alters alternative phenotype expression, but there has been no heritable change. We've tried to explain this in the below image for your information:



Comment: Line 102: The current phrasing of the objective makes it appear that the authors want to evaluate how anticipatory and mitigatory plastic responses could affect eco-evolutionary responses and population dynamics in each case. But it rather looks at whether selective juvenile harvesting of different stages could lead to anticipatory or mitigatory plastic responses, and if that has an effect on population dynamics.
Reply: Yes, this is a good point. We have rephrased the aim (Lines 96-98).

Comment: Line 105: “using ET model” perhaps change it to “under the assumptions of the ET model”..
Reply: Done.

Comment: Line 108: Wasn’t it that the fighter males develop slower than the scrambler males?
Reply: Yes, if they are on the same diet. We have changed the text to make clear that the comparison is between good versus bad condition juvenile males (Lines 110).

Comment: Line 110-111: if the mite density does not affect the fighter/scrambler male expression, will the threshold evolve for the populations under selection? Aren’t the individual liability distribution be affected by body size threshold and shouldn’t body size be density-dependent?
Reply: In other mite species within the same genus or family, male morph expression depends on pheromone cues emitted by the population; this is what we refer to here. We don’t refer to the effect of density-dependence through food competition. We have clarified the text that we mean that *R. robini* does not have pheromone-dependent expression of male morphs (Lines 110).

Comment: Q: But isn’t this expected given that largely deutonymphs develop into fighters? Therefore, the selective removal of deutonymphs should result in a lower frequency of fighter males in the population, not evolutionarily because of the mitigating plastic response, but simply due to removal of duetonymphs in each generation that could have metamorphose into fighters. Probably, it will be helpful to plot the data on how the removal of deutonymphs affects the proportion of fighter males in the first 7/14 days of harvesting, so that it becomes clearer that populations have evolved mitigating plastic response as a result of many generations of harvesting, and the reduction in fighter expression is not due to harvesting of deutonymphs (which could have become fighters) every generation.
Reply: Yes, our hypothesis is that the removal of deutonymphs selects against good-condition individuals and thus fighter expression. At first, the harvesting will only result in a demographic (plastic) response as you describe, but over a longer period of time, this harvesting regime can impose selection on fighter expression. This is effectively shown in Fig. 5c as the time period 1 encompasses any initial demographic (plastic) response. If the proportion of fighter males are driven by a demographic response, we would expect a similar proportion of fighter males over the three time periods in treatments where deutonymphs have been removed. To unequivocally show the evolution of fighter expression requires quantifying shifts in the threshold for male morph expression. Unfortunately, logistical constraints prevented us from conducting the time-consuming and labour-intensive assays to measure the threshold for alternative male phenotype expression itself (Lines 326-328).

Comment: Q: what is the difference between the D and F plots of figure 1, it appears that upon facing the stressful condition both the deutonymphs harvesting population (D-100) and juvenile harvesting populations (Jd-100) should react similarly under the ET model predictions? And, if this is the case, what is driving the difference in fighter expression between the juvenile (JD-100) and deutonymph (D-100) harvesting populations in the experiment?
Reply: The difference between D and F is that removing deutonymphs removes good-condition individuals, and thus removes the right-hand part of the condition distribution (compare blue dotted with red dashed on the right-hand side of each distribution). Randomly removing juveniles, means that individuals of all conditions can be removed along the full condition distribution, and not just on the right-hand size of the distribution. We explain this in Lines 197-200.

Comment: Line 217: It is not clear what authors mean by scrambler males not being inbred in the stressful environment?
Reply: We are quoting results presented in Stewart et al. (2019): the stressful environment refers to an environment with limited food availability. We have amended this and provided additional details for clarity (Line 220).

Comment: Line 217: It would also be good if the present findings could be discussed in the light of populations being highly inbred, and if the authors expect different outcomes if populations were outbred.
Reply: Unfortunately, there are no studies yet that present data on male morph expression and its evolution in the field, assuming that these populations are less inbred than lab populations. We refer the reviewer to Stewart et al. (2019), for an in-depth discussion of the role of genetic diversity in the evolution and maintenance of male morphs in *R. robini*.

Comment: Line 236: Will it be possible to provide the reasoning (not just the citations) for why the removal of just a few individuals should induce stress in populations, because the number of deutonymphs are very small, as compared to the total population size. This expected selective response in populations as a result of removing just a few juveniles does not seem okay to me
Reply: As we state in the lines that follow this statement, we anticipate that harvesting only a limited number of individuals will still induce population responses. For example, in previous experiments, we also imposed harvesting regimes within which we removed only a few individuals on a regular basis and found these selections to have significant impacts on the threshold for male adult phenotype expression and population size-structure (Smallegange & Deere 2014; Smallegange & Ens 2018) (Lines 254-256). We would also like to clarify that the harvesting treatment was not aimed at inducing stress per se, but at inducing an evolutionary response in fighter expression.

Comment: Line 244-245: It is not clear to me that what authors mean here by that "populations matched the dynamics of the other treatments"? Had the population sizes become equal in all treatments? Or were there similar number of males/females and juveniles in each treatment?
Reply: We looked at total population size (Fig. 4), which were all very close across the population tubes. We have included this in the text (Lines 259-260).

Comment: Line 248: Is mean total population size just the adult numbers or it is inclusive of all the life-stages?
Reply: This refers to all individuals in the population. We have added this to Line 358.

Comment: Line 374-375: Is it possible to plot the number of fighters and scramblers numbers in each treatment rather than the proportions? Because it is the main result of the paper and the reader could get an idea what are the actual differences among the treatments, the proportions are not fully informative.
Reply: Because our statistical analysis concerns fighter expression, i.e. the percentage of males that are fighters, we plot the results as percentages to match our statistical outcome. We have added fighter and scrambler numbers to the Appendix (Lines 400-401).

Comment: Line 389 onwards: It is interesting that fighter expression is declining as the experiment is progressing. Is it because the situation in tubes might be becoming worse overtime by lower per-capita availability of food and waste accumulation?

Reply: Yes, this could indeed be the case. We postulated that "This 'delayed' response could be due to that fact that our proportional harvesting treatments became more severe as population size increased over the course of the experiment.". We have added your suggestion as well there (Lines 497-502).

Comment: Line 403 onwards: Why is the increase in deutonymph numbers later in the experiment in D-100 counterintuitive, wouldn't we expect that given D-100 treatment had the highest population size (Figure 4& 5A) it will have lowest per-capita availability of resources, and in such situations juveniles would become deutonymphs. The main question over here is: why removing all the deutonymphs increases the population size over time? If this is due to duetonymphs freeing up the resources for other individulas, we would expect a similar response in JD-100, but it doesn't have a high population size as D-100.
Reply: The response is counterintuitive because you would expect that if you harvest a particular life stage, the relative occurrence of that life stage would decline. We do list potential explanations for this phenomenon but you have raised a good point which we have added as well (second paragraph of the Discussion). We, however, cannot determine whether the removal of deutonymphs caused the increase in population size.

Comment: Line 434: Figure 7 caption: The current caption does not describe all the panels; also there are errors in naming.
Reply: Thanks, corrected.

Comment: Figure 7D: Y-axis should be fighter expression proportion.
Reply: Corrected.

Comment: Line 484-487: This statement is unnecessary given there isn't any difference both statistically and visually.
Reply: We are not sure what you refer to here because these lines are the end and beginning of two different paragraphs:

> Mojica and Kelly 2010; Lürig & Matthews 2021; Rhebergen 2022), and how such trait dynamics influence, and are influenced by, population dynamics (Smallegange 2022).
>
> One unexpected finding in our experiment was that the number of deutonymphs was highest when we harvested all of them, each week. Further, despite being harvested, deutonymph

Comment: Line 491: Is the reduction in fighter expression in D-100 treatment largely due to the harvesting of all the deutonymphs (that could become fighter males) rather than evolution of mitigating plasticity, because the decline in the fighter expression seems to be really low (looking at the magnitude of proportions). Again it will be helpful to look at the real numbers of fighters and scrambler males in each of these treatments and how it declined over different time periods.
Reply: Please see our reply to your previous comment above on this subject. Our life history assay at the end of the population experiment did not have sufficient statistical power to tell if these shifts in fighter expression were (in part) evolutionary.

REVIEWER 2

We would like to thank the reviewer for their comments, which have contributed to clarifying key aspects of the manuscript.

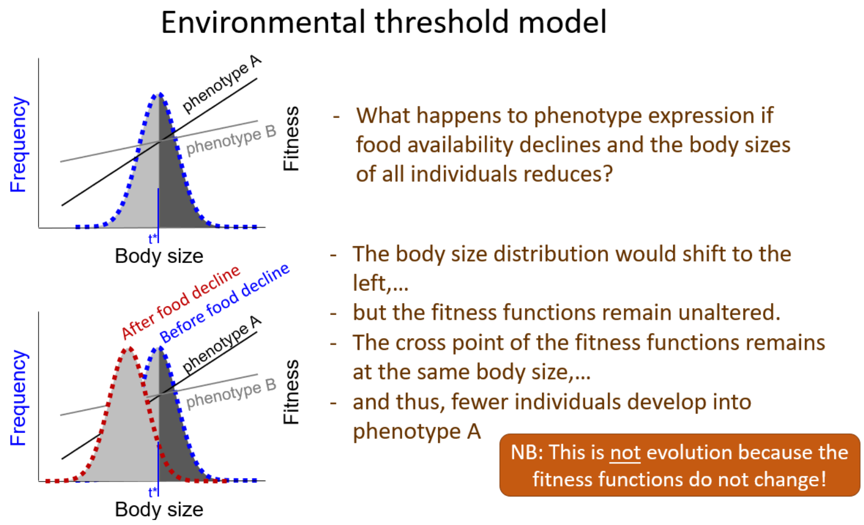
Harvesting treatment. In the experimental procedure, I think that it would be useful to explain in more detail how the % of harvesting in the JD100 and JD50 have been computed. For example, was this proportion averaged across the different replicates each week and applied similarly to different controls? I think it would be useful to give in a few words an order of magnitude on what this harvesting represent in%, in number of individuals and to what extent this harvesting varied during the course of the experiment.
Reply: Thank you for the suggestion. We have now added additional details (Lines 234-243) to provide more clarity on how harvesting was done in treatments where juveniles other than deutonymphs were removed. The numbers harvested changed during the course of the experiment as the population sizes changed which we have now provided as a figure in the appendix (Fig. S6).

Analysis. Could you explain why you did not use a Poisson model with an offset term to analyze the deutonymph expression as for fighter expression? In both cases you did count the number of individuals and it is not clear why you did not use the same type of analysis.: .
Reply: The difference between these analyses is that fighter expression is quantified as the proportion of fighters of the total number of adult males, which is why we offset number of fighters against the total number of males. We did not analyse total number of deutonymphs or total population size as proportions of something. Because these are discrete numbers with a minimum of zero, we used the Poisson distribution.

Plastic of Evolutionary response: Regarding the genetic or plastic basis of the observed phenotypic responses, the authors explain that evolutionary response is possible on the timescale of the experiment with this biological model. But I guess that such rapid evolution requires to have a sufficiently large initial genetic variability within the different populations. According to what I read, isn't it possible that the level genetic diversity at the begging of the experiment was too low to allow an evolutionary process to take place (inbred lines)? This lack of initial diversity could also explain why no evolutionary responses have finally been observed in the common garden experiment. I think that this should be discussed. I also think that the way the authors interpret the absence of genetic difference among their groups of individuals in the discussion could be rephrased.
Reply: We have used mites from the same populations in other experimental evolution experiments and have found an evolutionary response on similar time scales (Smallegange & Deere 2014). This difference between this and the previous study is how we conducted the selective harvesting (directly harvesting fighters or scramblers, versus indirectly via juveniles). We now mention that future studies might want to consider using outbred populations (Lines 592-597).

For instance, in the discussion the authors write "*Our life history assay at the end of the population experiment did not have sufficient statistical power to tell if these shifts in fighter expression were (in part) evolutionary.*" -> This sentence could be rephrased as something like "Our life history assay at the end of the population experiment failed at showing any sign of genetic differentiation between our treatments… which could be due to 1) no evolution…, 2) insufficient initial genetic diversity, 3) to little statistical power… blabla." Regarding the first possibility (no evolution), I do not fully agree when the authors explain that the change is the selection pressure due to the harvesting treatment is necessarily going to produce an evolutionary response. Given that if you have an adaptive plastic response, the plasticity will in return modify the selection pressure due to the harvesting treatments and this could limit the evolutionary response. The fact that the observed phenotypic plasticity could well reduce the selection pressure should I think be mentioned and discussed.

Reply: We have included your text suggestion (Lines 498-503). Regarding the role of plasticity: this operates through the condition distribution. For example, in our case, poor diet can reduce overall condition and move the condition distribution, which alters alternative phenotype expression, but there has been no heritable change. We've tried to explain this in the below image for your information:



However, such plasticity would not directly impact selection per se, unless the shift in the condition distribution directly affects how many deutonymphs occur, and how strong the selection pressure is through our selective harvesting of deutonymphs. We have added the latter to the Discussion (Lines 501-503).

Link between the model and the observations:

systems that produce scramblers in response to the juvenile harvesting selection pressure. What is important to note is that the ET model states that any change in the threshold for alternative male phenotype expression will affect the proportion of individuals developing either phenotype because it is expected to track the intersection of the alternative phenotype fitness functions. Therefore, we can deduce evolutionary shifts in the threshold from evolutionary changes in fighter expression (the proportion of adult males that are fighters).

Are you sure that you can "*deduce*"? I would rather say "interpret" or something like "Therefore, changes in fighter expression can be interpreted as the result of an evolutionary (or plastic) shift in the threshold." Using the word "*deduce*" supposes that you exclude any other possible interpretation. Changes in fighter expression could also result from other causes like changes in the mean growth trajectories due to changes in the strength of competitive interaction for instance. It could also be a plastic rather than an evolutionary response, what your results suggest anyway.

Reply: We have replaced deduced with interpret (Line 320).

Discussion: When you recall the predictions at the beginning of the discussion, you do not recall that in the mitigating hypothesis, if I am right, you also expect to observe bigger fighters (according to the threshold shift), when you empirically observe a decrease in fighter size during the course of the experiment. I think that you should discuss not only the results that support the mitigating hypothesis but also the discrepancies. A section in the discussion on the limits of the experiment would be useful to help the reader identify what could also explain some of the results and also to focus its attention on what could be nice to do in future work (identify the thresholds, follow individual growth trajectories, etc.). I understand that you removed the section in the discussion on the olfactory cues because it was speculative but I personally regret it. ☺

Reply: The effect of a shift in the threshold is that the proportion of fighters versus scramblers changes; it does not have any predictions on body size. In fact, the assumption of the ET model is that the body size distribution remains in the same position relative to the fitness functions, see e.g. Fig. 4c. It is very difficult to predict the effect of selective harvesting on body size, and thus we have not included these in our manuscript. We predict, e.g. in case of mitigating developmental plasticity, that the selective removal of large individuals with high resource budgets (deutonymphs) selects against males that are likely to develop into a fighter. This then changes their fitness relative to that of juvenile males developing into scramblers, which is displayed in Fig. 4d.

We have added some suggestions on how to improve our experiment at the end of the Discussion (Lines 592-597). Finally, since you emphasise that you appreciated the olfactory cue section, we have again included it in the Discussion.

Typos and small suggestions: worked on the track-change file so I do not refer to line numbers.)

242 verify the chisq symbols.

They are correct

Are you sure? In my file this is not the case…

d: Σ_8^2 = 29.92, p<0.001). d: Σ_8^2 = 19.44, p=0.013)

Life history assay

Neither fighter size (Σ_4^2 = 6.58, p=0.160, mean size: 439.76 µm ± 2.39 SE, n = 190) (Fig. 7A), scrambler size (Σ_4^2 = 2.77, p=0.597, mean size: 401.07 µm ± 7.33 SE, n = 14) (Fig. 7B) or female adult size (Σ_4^2 = 3.75, p=0.440, mean size: 334.66 µm ± 1.65 SE, n = 206) (Fig. 7C) were affected by the harvesting treatments in the common garden life history assay. Fighter expression was also not significantly affected by harvesting treatment (Σ_4^2 = 0.51, p=0.973).) (Fig. 7D); however, only 14 scramblers

Reply: Apologies! We have replaced the capital sigma with chi now everywhere.

"*If scrambler fitness increases relative to that of fighters, the threshold for fighter expression will evolve to **decrease**, both in response to deutonymph harvesting*"

-> Don't you mean **increase** rather than **decrease**? The threshold increases on the ET model figure.

Reply: Yes, you are correct, thanks for pointing out this sloppy mistake. We have corrected it.

"*Specifically, under juvenile harvesting (**J-100** and J-D50), fighter expression is predicted to remain unaltered if it is anticipatory **(Fig. 4e)**"*
J-100-> J-D100
-> Fig. 1e
Reply: Thank you, this has been corrected.

Fig. 4f -> 1f Verify the figure number
Reply: Thank you, this has been corrected.

"*This plastic (ecological) response, in turn, **will** further fuel the evolution towards developmental*"
Will -> can? See previous comment.
Reply: Corrected.

Figure 5 A, you could put the D100 on the right side to better show the gradient of selection pressure.
Reply: Done, that is a good suggestion.

You could present the variables in the statistical analysis section in the same order as in the result section.
Reply: Because different response variables are analysed with statistical models that have the same structure, and are therefore presented together in the statistical analysis section, we refrained from following this suggestion as it would increase the amount and repetition of text.