

Review

General comments

I would like to thank the authors for the detailed responses they provided. The clarity of the manuscript has improved. I found it easier to understand the link between the model and the experiment.

I also note that a certain number of the suggestions proposed were not followed and I understand that our points of view differ on some of the points raised.

I nevertheless still have several questions and I allow myself to propose a couple of suggestions to further improve 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.

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.

Plastic or 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 beginning 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.

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.

Link between the model and the observations

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

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.

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. ☺

Typos and small suggestions

(I 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...

J: $\Sigma_8^2 = 29.92$, $p < 0.001$). id: $\Sigma_8^2 = 19.44$, $p = 0.013$

451 Life history assay

452 Neither fighter size ($\Sigma_4^2 = 6.58$, $p = 0.160$, mean size: $439.76 \mu\text{m} \pm 2.39 \text{ SE}$, $n = 190$) (Fig. 7A), scambler

453 size ($\Sigma_4^2 = 2.77$, $p = 0.597$, mean size: $401.07 \mu\text{m} \pm 7.33 \text{ SE}$, $n = 14$) (Fig. 7B) or female adult size ($\Sigma_4^2 =$

454 3.75 , $p = 0.440$, mean size: $334.66 \mu\text{m} \pm 1.65 \text{ SE}$, $n = 206$) (Fig. 7C) were affected by the harvesting

455 treatments in the common garden life history assay. Fighter expression was also not significantly

456 affected by harvesting treatment ($\Sigma_4^2 = 0.51$, $p = 0.973$) (Fig. 7D); however, only 14 scambler

*“If scambler 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.

“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

Fig. 4f -> 1f Verify the figure number

*“This plastic (ecological) response, in turn, **will** further fuel the evolution towards developmental”*

Will -> can? See previous comment.

Figure 5 A, you could put the D100 on the right side to better show the gradient of selection pressure.

You could present the variables in the statistical analysis section in the same order as in the result section.