

Dear Dr Fragata, dear reviewers,

We were happy to see your comments on our manuscript “Landscape connectivity alters the evolution of density-dependent dispersal during pushed range expansions”, and that the general tone from these comments was positive. As is usual, there were some comments and suggestions, and we have done our best to include these in this revised version. In particular:

- The reviewers and recommender felt some of the existing theory on connectivity and evolution during expansions should be better cited and acknowledged. We agree, to a point: most of these references, while very interesting and important, assume the dispersal trait that may evolve is always density-independent, or that its density-dependence cannot evolve. Part of our work in this experiment and series of papers is about what happens when these assumptions are not valid. As such, many of these previous results are not as directly relevant as they may seem. However, we acknowledge that it is still important to spell out explicitly how and why our work contrasts from this existing corpus, rather than leave it implicit; we have thus made a conscious effort to do so in the rewritten Introduction and Discussion (see e.g. answers to comments **C0.1, C0.11, C1.2**)

- While the reviewers acknowledged how we managed to fit many disparate phenotypic assays in one common story, they felt the way we wrote the paper made the aforementioned story hard to follow at times. We have taken this comment to heart, and (i) added an “overview of the experiments” figure to help make clearer which populations were sampled for which traits (ii) streamlined the discussion of some results to focus on the most important ones. We feel, and hope, this make things clearer for both the reviewers and future readers. (see e.g. answers to comments **C0.2, C1.3, C1.4, C2.5 to C2.8**)

- The reviewers and recommender point rightly that, while our results are interesting and intriguing, they are still (especially the most interesting ones) based on a limited number of strains and replicates. We had addressed this in part in the original manuscript, but agree with the reviewers that this was not enough. We have now developed this into a full-fledged limitations paragraph, which importantly is written not as a way to “take down” the present study, but to point out how future studies may pursue these ideas further and better. (see e.g. answers to comments **C0.12, C1.11, C2.1 and C2.2**)

We additionally corrected a small coding error in the Supplementary Material: due to the way the functions we used are written, the pairwise comparisons plots for fecundity were originally made *without* the zero-inflation components, where they should have been included. This is now corrected; our conclusions remain unchanged.

Our detailed point-by-point replies to all comments follow below.

We hope this revision will be well received, and thank you for your time and comments.

Maxime Dahirel, on behalf of all authors.

DETAILED COMMENTS AND REPLIES TO COMMENTS

- we numbered comments that were not numbered by recommender/reviewers (and our corresponding replies), for clarity

- unless specified, lines numbers in our replies refer to the *revised* version.

Replies to notes from the managing board (i.e. from the “When revising your article, we remind you that:” paragraph)

- all data and annotated code were already publicly available (with a DOI) in the original submission; this remains unchanged

- experimental details that were missing according to reviewers have been added (see replies to their comments)

- A “conflict of interest disclosure” paragraph has now been added, including information on the authors that are PCI recommenders

**Comments from the Recommender
by Inês Fragata, 2021-04-23 08:45**

Manuscript: <https://www.biorxiv.org/content/10.1101/2021.03.03.433752v2>

Revisions needed

Dear Authors,

Your work has been assessed by three reviewers and me. Whereas we generally agree that the work is very interesting and clearly has an added value to the study of the eco-evolutionary dynamics of range expansion, there are some changes that would improve the current manuscript and are needed to clarify some aspects of your experimental design.

General comments

I think that the manuscript is very interesting and relevant, but there are several things that need to be improved. Reviewers 1 and 2 provide several comments that will allow you to do this. I am adding below comments of my own, with the intent of also helping on this matter. I think the supplementary material is great!

[We thank the recommender and the reviewers for their comments.](#)

Specific comments

Introduction

Comment 0.1 (C0.1)> As reviewer 1 stated there are some literature, especially theoretical, that can be added to the introduction and later on discussed (reviewer 1 provides most of these examples).

Some additional references that may (or not) be of use also:

- Miller TEX, Angert AL, Brown CD, Lee-Yaw JA, Lewis M, Lutscher F, Marculis NG, Melbourne BA, Shaw AK, Sz?cs M, Tabares O, Usui T, Weiss-Lehman C, Williams JL. Eco-evolutionary dynamics of range expansion. *Ecology*. 2020 Oct;101(10):e03139. doi: 10.1002/ecy.3139. Epub 2020 Sep 2. PMID: 32697876.

- Zaker N, Ketchemen L, Lutscher F. The Effect of Movement Behavior on Population Density in Patchy Landscapes. *Bull Math Biol*. 2019 Dec 23;82(1):1. doi: 10.1007/s11538-019-00680-3. PMID: 31919597.

- Williams, J.L. and Levine, J.M. (2018), Experimental evidence that density dependence strongly influences plant invasions through fragmented landscapes. *Ecology*, 99: 876-884. <https://doi.org/10.1002/ecy.2156>
- Urquhart, C.A., Williams, J.L. Trait correlations and landscape fragmentation jointly alter expansion speed via evolution at the leading edge in simulated range expansions. *Theor Ecol* (2021). <https://doi.org/10.1007/s12080-021-00503-z>
-

Answer to C0.1: In the Introduction, we've added a new paragraph to point out existing research on connectivity during range expansions (LINES 97-114). In addition to acknowledging (some of) the references above, this serves to point out that most of the literature on the subject assumes density-independent dispersal and/or does not allow for dispersal evolution. This highlights the novelty of our approach, which is based on both allowing evolution and in some cases explicitly testing for the evolution of density-dependence.

Methods

C0.2> You should add a figure that explains better your experimental design (i.e. how many strains and replicates per strain x type of connectivity you have) and also a table explaining which traits were sampled in each generation and which replicates from each population were used. I think this is a really important piece of information that you need to add to the manuscript, it was super confusing to understand what you did in each generation and knowing that you selected a replicate at random was not helpful at all.

Answer to C0.2: We have added a new figure (now Figure 2) which we believe help make sense on the traits that were sampled, in which replicates and when.

C0.3> Another important piece of information that you should is also the size of the landscapes (i.e. the number of patches). This information is available in your initial paper, but I think it should also be here.

Answer to C0.3: To avoid cluttering the Methods, we have added this info in the form of a new panel in Figure 1 (1C) rather than as text.

C0.4> In your analyses, have you considered using the number of the edge patch as a covariate? I am assuming that different landscapes had different sizes. If so, adding the expansion size as covariate (or the number of the edge patch were you took the individuals from) allows to test if the several traits analysed were in some way correlated to how much the population expanded.

Answer to C0.4: We did explore this during preliminary look, but there was simply not enough variation in our samples, especially for the density-dependence experiments (since only 6 landscapes were sampled in total for these). This might have worked if we had sampled all patches in all replicates (vs simply the extremes in a few), but this was too heavy a workload. We therefore fell back on the core-edge contrast. We add to the new limitation paragraph (LINES 506-523) an explicit mention of the fact we may have missed an effect of expansion speed (vs. simply core/front contrast) on traits. We note however that, on average, the two treatments did not differ on speed, despite the reduced connectivity (this was a key point in paper 1, Dahirel et al. 2021 *Oikos*, which can be seen on the new panel Fig. 1C)

C0.5> Another suggestion that I have is to do a multivariate analysis with all traits assuming that, in general, you used the same replicate population across time (and with the BIG caution note that the traits were analysed in different generations). This is because invasive populations often have the invasion syndrome and I wonder if here, by checking each trait individually, you might miss the sign. I know this is more unconventional, so really up to you to see whether this makes sense in your system.

Answer to C0.5: Thank you for the suggestion. The idea of implementing a multivariate mixed model did come up during the initial analysis design phase, as it could indeed be very interesting for precisely the reasons you mention. We ultimately chose to not do it, for two key reasons both tied to our limited number of populations:

- As can now be seen from Figure 2, traits do not overlap perfectly in terms of which populations were sampled. While this does not, in general, preclude the fitting of multivariate models (there are ways to deal with “missing” populations), the low number of populations here makes it difficult (as the % of “missing” populations would be high for any given bivariate correlation)

- In the absence of a resolved phylogeny linking our populations together, we use nested random effects to encode (phylogenetic) proximity while accounting for the “populations nested in landscapes nested in strains” nature of our design. This means that, if we wanted to fit a multivariate model properly, we would need to effectively estimate three covariance matrices (strain-level, replicate-level, population-level). Compare with a model that would have proper phylogenetic random effects, in which there would only be one phylogenetic correlation matrix. Due to both the small sample size in some traits and especially the small number of levels for the “mix/strain” random effect, we do not, in our opinion, have enough data to accurately estimate all three correlation matrices.

If the recommender is really keen on this, we are open to presenting an example analysis in the code, to describe explicitly why we chose not to analyse these data this way (although making this response to reviews publicly available may already fill this purpose). For now however, we choose to keep the analysis as is.

C0.6> Please clarify why do you use two measures of dispersal.

Answer to C0.6: We initially used two movement metrics (effective dispersal and general activity) as while they may be correlated, they reflect different aspects of movement behaviour (willingness to leave a resource patch + ability to survive the transience phase of dispersal for the former, movement propensity in the absence of resource for the latter). While we only need the former for a study of pushed expansions, knowing the latter, and whether responses match between the two responses, may be useful in general, to understand the causes of differences between connectivity contexts, and as a proxy to design future experiments. We rewrote some parts of the Methods to make that a bit clearer (LINES 249-253).

C0.7> In line with reviewer 1, I find the use of the word context very confusing, as it can relate to different things. I think that you should specify the comparisons you are doing and not give a general

nomenclature (both here and in the results).

Answer to C0.7: We initially used a general nomenclature to explicitly reflect the way what we call “experimental contexts” are encoded in the models : as 5 levels of a single categorical variable rather than the interaction of, say, “location” and “experimental connectivity treatment”. We now try to avoid relying on that “context” generalization, and explicitly say what we compare, whenever possible (see e.g. the model descriptions LINES 298-330).

C0.8> In your analyses have you nested experimental replicate in the interaction between experimental landscape and strain? This will be important to guarantee the level at which the intercept is varying.

Answer to C0.8: We are not exactly sure of what is meant here. There are actually no connectivity x strain interactions/ no random slope of connectivity (which is what I assume the recommender means by “interaction between experimental landscape and strain”) in our models, since with only three strains and in some cases only one replicate per strain*connectivity combination, trying to estimate the relevant random effect variances would be next to impossible, even with strong priors.

Even if there were such interactions, including “connectivity: strain: replicate” random interactions is impossible, since each replicate landscape only belongs to one connectivity level. One can only include “strain:replicate” random intercepts. And that is we actually did: the random effect structure is, in practice, using R syntax, (1|strain/landscape/population). This is an approximation of the presumed phylogenetic relationship between the population, as suggested in e.g. Clutton-Brock & Harvey, 1977; Hadfield & Nakagawa, 2010. We very slightly rewrote the sentence of the Methods where we describe the random effects to make their nested structure clearer (LINES 300-303).

Results

C0.9> You should use more precise language in the description of the results.

Answer to C0.9: We have rewritten results in a way that is hopefully more precise. This includes removing mentions of “potential effects” (see also reply to **C1.6**), rephrasing some overly long sentences, and adding effect size values for the key comparisons that are different from 0 (rather than simply suggesting the reader to look at the pairwise comparison supplementary, where these effect sizes are given graphically).

C0.10> To improve plot readability on all figures, I would put a line to mark the median value of the stock populations (like you do in fig 4C) across all other facets in the plot. This way the comparison between stock and the evolved populations would be easier. In addition, it will also make it easier to see how different the two evolved populations are from each other.

Answer to C0.10: Figures have been updated to do this (more exactly to display the mean of the stock posterior across all facets _ except in Fig 4C-5C, where the line marks 0). Note that in facets with low and high density data, we only display the line for the “low-density” stock population, to

avoid cluttering the figure with two sets of lines, which would ultimately hurt rather than improve readability. We note that the purpose of adding such lines (make comparisons with stock populations explicit), is also fulfilled by the Supplementary Material displaying all pairwise comparisons.

Discussion

C0.11> I think the discussion needs to incorporate a bit more some predictions from theoretical studies. I think that it would be great if you add some of predictions to the introduction and then revisit them in the discussion, to give a greater depth to these results.

Answer to C0.11: see also our answer to **C0.1**. We have expanded the second paragraph of the discussion (the one that deals with the negative results) following this comment. We discuss our results in the light of Williams et al empirical and theoretical results (LINES 420-431); in addition, we also draw from both Williams et al 2016 and Ochocki et al 2020's simulations to go a bit beyond the "sometimes range expansion studies find no evolution, and that's it" statement that originally closed that paragraph. We now discuss a bit the idea that trait architecture, and trade-offs in particular may explain why some studies find evolutionary responses while some don't (LINES 428-435).

We did not change much the rest of the Discussion in response to this comment since, as mentioned in our answer to **C0.1**, existing theoretical models mostly deal with either the evolution of density-independent dispersal in response to connectivity, or the effect of non-evolving density-dependent dispersal and connectivity on expansion speed. As such, it does not really bring useful predictions to discuss to the rest of the Discussion (but see the discussion of e.g. Rodrigues & Johnstone, 2014; Travis et al., 1999 LINES 482-485, which already existed in the original, with the caveat that variation in dispersal costs may be different from variation in connectivity). To conclude on this comment, we feel that in the Introduction, adding some explicitly new predictions (vs. nuancing/detailing predictions *we already made* by adding info from these theory papers _ see **C0.1**) would amount to HARKing and should be avoided.

C0.12> As reviewer 2 pointed out, beware the fact that you only analyzed one single replicate for some traits, which limits your ability to generalize your results.

Answer to C0.12: We agree (with the caveat that for the relevant traits we actually sampled one single replicate *per strain*, not one single replicate _ we rewrote Methods to make that clearer). We make the need for caution even more explicit in the discussion, by expanding the new "limitations" paragraph (LINES 506-523, see also replies to **C2.1** and **C2.2**)

C0.13> L374: 376 – The sentence starting by "This is despite..." is a bit confusing, i am not sure if it is a comma missing or the structure of the sentence that is not correct.

Answer to C0.13: We deleted this entire sentence.

C0.14> L439:441 - The sentence starting by "By contrast..." is also a bit weirdly phrased

Answer to C0.14: This sentence has been rewritten as (LINE 475): “Indeed, other empirical studies show shifts to negative density-dependent dispersal can happen at the edge of range expansions”

Reviews

Reviewer 1 (Reviewed by anonymous reviewer, 2021-04-06 16:31)

Dahirel and colleagues present the results of an experimental study to examine the effects of connectivity and density dependence on trait selection during range expansion. The purpose of the study is well motivated and clear. Based on a previous experiment and theory for pushed vs pulled expansions, we would expect that the high connectivity (reference) landscapes would have higher fitness at the leading edge than the core, and/or negative or non-density dependent dispersal. The reduced connectivity landscapes should have lower fitness at the leading edge and/or positive density dependent dispersal. Although the results were somewhat complicated, the figures were clear. The authors did find that the density dependence of dispersal varied between treatments in the manner we would expect given high connectivity landscapes induced pulled waves, and reduced connectivity resulted in pushed waves. The manuscript would benefit from being more strongly grounded with theory – both clarifying in the introduction and stronger links in the discussion. Below I provide more specific feedback to improve the manuscript.

C1.1> 1. L67-71 I think this is a bit misleading, “spatial sorting” is typically used to refer specifically to traits that enhance dispersal, and there are a wider range of traits potentially under selection during spread that are not described by this term (the authors describe some of these in lines 74-75). Related, I disagree with the use of ‘spatial selection’ at L83 to refer to all of these processes – if the authors wish to equate spatial and natural selection (as it’s arguably still debated in the literature), they will need to explicitly define this. Otherwise, I think it would be more suitable to use natural selection for traits like fecundity.

Answer to C1.1: We agree with the reviewer that we have been careless with vocabulary here, especially since we did not define terms. Since we only used these terms a handful of times, and the rest of the text shows we don’t need them to make our argument, we simply decided to remove them from the text (we keep the reference to “spatial sorting” in key words, however, for referencing purposes)

C1.2> 2. I was surprised the authors did not incorporate any of the literature that explicitly considers landscape connectivity in shaping range expansions. There are a number of previous studies that look at both density dependence during spread in fragmented landscapes, as well as how landscape connectivity alters evolutionary trajectories of leading edge populations. I’ve provided some suggestions below for papers that should be highly relevant, but I would suggest the authors incorporate more of the previous literature on this topic to better motivate their work. (e.g. Dowdall et al. 2018, J. of Mathematical Biology; Lutscher & Musgrave 2017, Ecology; Gralka & Hallatschek 2019, eLife; Williams et al. 2016, American Naturalist & Science)

Answer to C1.2: We thank the Reviewer for these suggestions. As mentioned in our replies to the Recommender (**C0.1, C0.11**), we’ve updated the Introduction and Discussion to include some of these references. (see e.g. LINES 97-118; 420-435).

C1.3> 3. In some places, the methods need clarifying to help the reader follow the complex story. Adding a summary of the traits studied and from which generation would help – this could be in the text (~L 155 – 168) and/or in a table. It wasn't clear why traits were tested at different generations, and not the latest one (12th generation, I believe?). I'm not sure it really matters for the results but I found it confusing. Please also add an explanation of why two different measures of dispersal were used, how are they related to each other, and what the expectations were for each treatment. (I'd also recommend putting these two sections next to each other in the methods).

Answer to C1.3: See answers to Recommender comments **C0.2** and **C0.6**. We believe the new figure (Figure 2) addresses the need for clarity about the experiment "story" (and does it better than a Table), and answer **C0.6** is about the differences between dispersal and activity. We've placed the dispersal and activity sections closer together in the Methods too.

C1.4> 4. For the measurements of wasp size, where 8 core-edge populations were selected at random, how did you choose from across the isolines? Maybe I missed something, but I thought the original design had 8 replicates of each of the three types.

Answer to C1.4: Yes, the original design had 8 replicates per strain, but we only selected some of these, since measuring individual microwasps is extremely time-consuming. We aimed to sample each strain as equally as possible (strains had 3/8,3/8,2/8 replicates sampled), but within each strain*treatment combination, decided which replicate to sample at random. We added a new Figure (Figure 2), and generally rewrote these parts of the Methods to make it clearer (LINES 192-194, 238-242, 281-284; see also our replies to comments **C2.5** to **C2.8** by reviewer #2).

C1.5> 5. Throughout the methods and results, I found the use of the word 'context' hard to keep track of (in particular the contrasts between core and edge and the two connectivity treatments). Upon revisiting this section of the methods, I see now that context is used more generically to mean both core vs. edge and reference vs. reduced connectivity. Since specific hypotheses are laid out for each, to me it does not make sense to combine them into one word. To me it would be clearer to give each more precise names, or at least occasionally redefining context.

Answer to C1.5: (see also our reply to recommender comment **C0.7**). Our use of a single "context" word stemmed from the fact this is how the "categories" are encoded in-model. As mentioned in reply to Recommender comment, we agree that this can be confusing; we have tried to shift to explicitly say "location" or "connectivity level" whenever a sentence refers only to one of those.

C1.6> 6. I struggled with some of the language in the results section around whether or not two groups were different from each other. What does it mean by 'some indication' or the 'only potential effect'? Is there an alternative way to make it clearer that the groups are not significantly different (using the appropriate Bayesian language here, of which I am not an expert), but the mean is higher in one treatment than another?

Answer to C1.6: For simplicity, we removed direct mentions to these "close to significant" effects (as they would be referenced in some frequentist result sections) from the text, since they have no influence on the main message of the paper, and the interested reader can look at the shape of the posteriors themselves, including the shape of the pairwise differences from the supplementary figures.

C1.7> 7. L 381-384 This was interesting to me and I wish the authors gave more information in the

main text. From Figure S.2.1: wasps were smaller after the experiments than stock populations, and edge populations in reference landscapes were smaller than core populations as well as those from the reduced connectivity landscape. But there were no differences between edge and core in the reduced connectivity landscapes? Earlier, the authors state that body size is linked to fitness, which I assumed to mean that larger individuals reproduced more. It wasn't immediately clear why they expected the differences in body size they reported to be most important in the context of dispersal-competition trade-offs. I wondered if the authors could provide some further insight on how body size relates to competition (if at all) and perhaps whether this could be indicative of selection in the reference landscapes? Further in L385-400, can you relate any of these trait changes to the changes in speed from the first experiment?

Answer to C1.7: As a consequence of the changes made in response to the previous comment, the sentences mentioned in this comment have been deleted. In any case, we believe that we don't have the data to make it as major a part of the discussion as the reviewer seems to want (we would be happy to discuss it in detail if we had data on the links between dispersal, fecundity and size, but we don't _ see also Reply to comment **C0.5** by recommender).

C1.8> 8. In a few places in the discussion, the complexity of the results gets washed over (e.g. L377-379, 404-406) – while I appreciate trying to simplify, the arguments would be more powerful if the nuance were included.

Answer to C1.8: We have tried to put back some nuance wherever it is needed. In the specific examples mentioned by the reviewer, we hedge our results by saying we found “indications of” or that “while our data are limited (see below),”(e.g. LINE 441). The “see below” refers to a more detailed limitations paragraph that is now present just before the conclusion (LINES 506-523)

C1.9> 9. L410-424 – can you provide an alternative hypothesis? What if the egg number-egg size tradeoff was not supported with further work?

Answer to C1.9: An alternative hypothesis would be that core wasps evolve a higher propensity to superparasitise, as they experience high densities more frequently than especially edge wasps. Since wasps emerging from superparasitised eggs tend to be smaller (and so probably less fecund, Durocher-Granger et al., 2011), this may explain our results. We have rephrased this paragraph to introduce both hypotheses (LINES 455-460).

C1.10> 10. L445 – In this paragraph, I think more can be said from theory in the metapopulation context that makes predictions about why dispersal might be lower when patches are more dispersed.

Answer to C1.10: We see the point the reviewer is making, but we are not sure it is relevant here. As mentioned elsewhere in these replies, ultimately what we are interested in *this part of the Discussion* is the density-dependence of dispersal, and whether the experimental conditions select for higher or lower dispersal overall besides that is not directly relevant here. We already discuss relevant references and results on how dispersal costs/connectivity may shape the evolution of density-dependence in a metapopulation context elsewhere (LINES 482-485, Justin M. J. Travis, Murrell, and Dytham 1999; António M. M. Rodrigues and Johnstone 2014; Govindan et al. 2015).

We are of course open to expanding this in a next revision if the reviewer feels strongly it is needed, but for now we chose to let this paragraph as is, given the current length of the manuscript.

C1.11> 11. Overall, the discussion would benefit from a paragraph describing the limitations of this experiment to help put it into a larger context. There is just a little on L471-473.

Answer to C1.11: We now expand the sentence that was originally L471-473 into a full paragraph about limitations (LINES 506-523). Besides the limited number of replicates, we highlight the fact that we “only” used three strains (so larger comparative studies may be needed to see if our results are strain-specific), and the fact that although these strains derived from crossing isolines, so are genetically diverse, their source genetic material still comes from a few isolines, meaning standing variation may be limited. All these may explain why we saw no response for the other traits.

C1.12> 12. The supplementary material sections include several sections – it would be clearer in the main text if the authors referred to specific figures rather than referring the reader to the entire section.

Answer to C1.12: We now mention the specific supplementary figure referenced (whenever relevant)

A few minor comments:

C1.13> L24. The reader does not yet know what ‘treatments’ refers to – please be more precise here.

Answer to C1.13: Indeed. We replaced it by “between reference and low connectivity replicates” (LINES 22-23)

C1.14> L92. The use of the word surprising in this case seemed a bit strong. To me, the lack of studies is not surprising – this is a new field and these kinds of experiments are very difficult to do. Instead, what about something like, “the next step’ or ‘the way forward’ is to...

Answer to C1.14: Indeed. We rewrote the sentence, which now reads “Exploring this is in our opinion the next step in pushed expansion studies, given the distinction between pushed and pulled expansions rests, at its core, on traits (dispersal and fecundity) we now know can be under selection during range expansions.” (LINES 86-89).

C1.15> L 98-88. What does ‘appears to move edge populations away from’ mean?

Answer to C1.15: We rewrote the entire sentence for clarity. It now reads: “(…), evolution at range edges may lead pushed expansions to become pulled (Erm & Phillips, 2020), as an initial positive density-dependence in growth or dispersal is lost during expansion (Erm & Phillips, 2020; Fronhofer et al., 2017; Travis et al., 2009; Weiss-Lehman et al., 2017; but see Mishra et al., 2020)” (LINES 92-96)

C1.16> L203. What is a tracklet?

Answer to C1.16: We use it to mean “any continuously tracked bout of movement longer than 2

seconds”, although the word itself has been used elsewhere to refer to any small (part of a) track. We rewrote the sentence (actually split it in two) to make it clearer that the definition in parentheses referred to this word (LINES 225-227).

C1.17> L391. What does “but in the other in the second” mean?

Answer to C1.17: We rewrote that sentence which was admittedly unclear. It now reads: “However, our experiments are here inconsistent: low-density dispersal was higher in edge vs. core patches only in “reduced connectivity” landscapes in one dispersal experiment, and only in “reference” landscapes in the other dispersal experiment (Fig. 5, Supplementary Figures S.2.3 and S.2.4). There was also no clear divergence from the starting stock populations themselves (Fig. 5, Supplementary Figures S.2.3 and S.2.4). (LINES 416-420) (this also answers **C.3.4**)

C1.18> L410. Appearance would be more appropriate than apparition here.

Answer to C1.18: We corrected accordingly, LINE 447.

Reviewer 2 (reviewed by anonymous reviewer, 2021-03-31 15:28)

the authors present new results from a previously published experiment on range expansions in microcosms using a parasitic wasp to examine the effects of connectivity. In their previous work, the authors showed that the different connectivity treatments resulted in different expansion dynamics, as quantified by the degree to which the expansions were pushed vs. pulled. In the new data and analyses, the authors attempt to link these observations to patterns in trait evolution by measuring a suite of traits in wasps collected at various points throughout the experiment. I found this manuscript well-written and very intriguing. The question put forward is highly relevant and important to current research on eco-evolutionary dynamics in range expansions. For the most part the authors describe their experiments and statistical analyses well (see a few minor comments below) and the analyses are appropriate and well justified.

We thank the reviewer for these comments.

C2.1> My only concern with this manuscript is one that the authors appear to be well aware of, which is the difficulty of drawing meaningful inference from the low sample sizes used in the experiments quantifying the density dependence of fecundity and dispersal. My concern regarding this arises for two reasons. First and most obvious, it is extremely difficult to draw meaningful interpretations on trait evolution from a single replicate per connectivity X strain treatment. I realize multiple wasps were taken from core and edge locations of these landscapes, but they only represent a single replicate of the evolutionary processes underpinning these range expansions. The authors are aware of this limitation and mention it several times (with a particularly nice discussion of it on lines 389-398). However, there are also several instances where the authors take these results and interpret them more broadly in the context of the behavior of edge or core populations broadly (e.g. lines 376-379, 384-389, and 429-431), which is a difficult assertion to justify given the small sample size.

Answer to C2.1: As mentioned elsewhere in these replies, we have now expanded our discussion of the sample size related limitations to a fuller paragraph (LINES 506-523). In addition, we use hedge

words to slightly nuance the statements flagged by the reviewer (LINES 405-408, 410-414, 465-467), foreshadowing the limitations paragraph coming later on.

C2.2> My second concern regarding the small sample sizes is that evolutionary dynamics in range expansions are known to be strongly affected by neutral, stochastic forces (i.e. gene surfing). This means that alleles can rise to high frequencies (or be lost) in edge populations purely by chance in range expansions, making it even more difficult to draw inferences on trait evolution. Gene surfing increases the importance of replication in these studies as it makes distinguishing evolutionary changes due to adaptation vs. neutral dynamics very difficult in a single range expansion. This is not to say that the authors should not report on their results, as they do provide intriguing directions for future research. However, I would suggest the authors use more careful language throughout their manuscript as this experiment's ability to make broad inferences on trait evolution in edge populations at varying connectivity levels is a bit low.

Answer to C2.2: The "limitations" statement we now include now explicitly mention this stochasticity "problem"(LINES 506-511).

Below, I have included a few minor comments with line numbers.

C2.3> Line 83: The authors switch to "spatial selection" here from "spatial sorting." This is in line with existing definitions of spatial selection vs. spatial sorting (e.g. Perkins et al. 2013 which the authors cite), but the authors should perhaps include an explicit definition if they are to use this term so that casual readers do not conflate the two.

Answer to C2.3: We've removed these mentions to spatial selection and spatial sorting, since they are not needed (see also answer to C1.1)

C2.4> Line 101-105: This is the first time the authors introduce the idea of ecological context affecting trait evolution during range expansions. As ecological context could refer to a wide variety of things (e.g. biotic interactions), I suggest the authors define what they mean by ecological context and/or use more precise language here.

Answer to C2.4: In line with other comments by the Recommender and Reviewer 1, we've recreated an entire paragraph here to better introduce the existing connectivity context (LINES 97-118). As a result, the sentence mentioned by the reviewer (now at the end of this paragraph) now refers to connectivity specifically.

C2.5> Line 169-172: Were the 8 landscapes truly chosen randomly (i.e. was there no attempt to balance among strains or connectivity levels)?In general, I have this same question for most of the traits measured.

Answer to C2.5: *Generally speaking*, landscapes were chosen randomly *within each strain x connectivity combination*. The new Figure 2 make this much clearer in our opinion, by making the distribution of sampled replicates between strains and connectivity levels explicit. We modified the text throughout the Methods to refer to it.

C2.6> Line 219-222: The authors should state whether it was an edge or core population that was excluded here and why (presumably low numbers).

Answer to C2.6: It was indeed an edge population, and it was indeed due to low numbers. This is now added (LINES 194-195)

C2.7> Line 227-228: How did the 47 populations divide into 99 groups? Also, which four populations were excluded and why? I realized this is asking for a lot of details, but given the low replication of this experiment (4 replicates per strain x connectivity combination) these numbers of excluded populations could make a difference depending on how they are distributed.

Answer to C2.7: This comment contains two questions:

- "how did the 47 populations divide into 99 groups?". It is because these populations are actually 44 populations + 3 stock populations, and we used 4 replicates per stock populations versus 2 for the other ones. This makes $44*2+3*4=100$, but one of the non-stock replicates was lost (the reason was not recorded, but no parasitoid development during the trial _ so nothing can be recorded _ is likely)

- which populations were excluded (since we sampled 44 out of the 48 that were available)? One population per location*connectivity treatment was excluded (see the new Figure 2). The strain for which one population was excluded was chosen at random for each combination.

We have tried to put as many of these details as possible in the revised Methods without making it too unwieldy, with the help of the new Figure 2 (LINES 238-242). Some details may still be missing; however, we believe the fact that this present reply + the totality of the data will be/are publicly available should make up for it. We are open to revising this part of the Methods further if not.

C2.8> Line 252-255: I'm again having trouble figuring out how these numbers add up to 72.

Answer to C2.8: Stock populations: 4 replicates per density (low-high, 2) per strain (3)=24

Expanding populations: 2 replicates per density (2) per strain (3) per location (core-edge, 2) per connectivity (reference-reduced, 2)=48

$48+24=72$

We have tried to rewrite the methods to make that a bit clearer, especially the fact that we tested more replicates for stock populations (LINES 281-284)

C2.9> Line 286-291: Is the choice to not break out results for p and lambda separately driven by a parameter identifiability issue or simply because it does not add any new information? I am curious as it seems possible that experimental factors could impact these two variables differently, so it would be interesting to know if the authors are able to show how they each respond to connectivity, location, strain, etc.

Answer to C2.9: There were three main reasons:

It was partly because the use of zero-inflated models was not decided from the start, but because the data did not fit the assumptions of a classical Poisson/negative binomial model. Since we had no intention to separate zeroes from non-zeroes if we had used these methods, we decided not to

here for consistency.

In addition, here we are studying phenotypic traits insofar they can inform us on range expansion. What is relevant for this is the mean fecundity; whether a change in mean fecundity is due to a change in the *individual* probability of reproductive failure or a change in the number of offspring produced by the “non-zero” wasps is not of interest.

Finally, rather than statistical identifiability, there is an issue of biological identifiability. We do not have a conceptual model allowing us to associate the probability of excess zeroes to a specific process in *Trichogramma* life cycle, or even a set of processes (failure to mate reducing egg-laying, early death of the egg laying adult, self-superparasitism leading to larvae deaths...).

As such, while it may be interesting, in general, to split p and λ , we felt it was not useful to do it specifically here. We are open to revisit this choice if the reviewer still felt it was needed after this reply. We note that any interested reader can relatively easily get info on p and λ by themselves, using the code and data provided.

Reviewer 3 (Reviewed by anonymous reviewer, 2021-03-22 23:15)

I just completed the review of the manuscript with the title “Landscape connectivity alters the evolution of density-dependent dispersal during pushed range expansions”. It is well-written and describes adequate and well thought off experimental procedures to answer an interesting research question about species range expansion in the context of environmental change.

C3.1> I have an overall question concerning the main conclusions of the preprint: could it be that, as a species expands its range, the lower densities at the expanding range edge cause an initially pushed expansion to move towards a pulled one?

Answer to C3.1: Indeed, we believe that it is, in the end, the core of both our and Erm and Phillips (2020)’s results: low densities at the expanding range edge mean that anything that disperses or grows worse at low densities will be outperformed/outrun, leading to an accumulation of individuals that disperse/grow well at low densities. Which can cause an initially pushed expansion to become pulled, unless the selected individuals have higher growth/dispersal across the board, rather than just at low densities. While the argument was implicit throughout the manuscript before, we have now rewritten the Conclusion to make it explicit (LINES 530-533).

Introduction

The introduction provides an adequate framework for the study. It is well structured, explaining the research context and knowledge gaps and it is based on relevant research. The motivation of the study, as well as the objectives, are clear and relevant in the current context of environmental change.

Materials and methods

Data and code were made available to the reader. The description of the methods is detailed enough to allow replication and the statistical analysis are adequate.

C3.2> I think that the methods would benefit from a new figure, describing schematically the experimental design. As such, it is my recommendation that the authors consider creating a figure were the experimental steps are schematically represented.

Answer to C3.2: We have added such a summary figure (now Figure 2), in response to comments by Recommender.

Results

C3.3> Paragraph around line 355 – I understand what the authors meant. However, when they write “...leading to negative density-dependent dispersal...” that is not always absolutely right, from what I can understand. In the wasps from the “core” in the “reference” landscape and the wasps from the “edge” in the “reduced connectivity” landscape, there is a positive density-dependence. The authors should clarify this paragraph.

Answer to C3.3: Yes, the whole sentence about negative density-dependent dispersal actually refers specifically to the “edge, reference” and “core, reduced” cases, as mentioned at the end of the sentence. Which means both the reviewer’s assessment and the sentence are correct. However, the reviewer’s comment shows that the sentence is not clear about that. We rewrote it and expanded it to hopefully make it better. It now reads:

“After experimental evolution, this pattern was reversed, leading to negative density-dependent dispersal, in two cases: in wasps coming from edge populations of “reference” landscapes ($\log(\text{odds ratio}) = -1.57 [-2.23; -0.90]$) and in wasps from core populations of landscapes with reduced connectivity ($\log(\text{odds ratio}) = -0.60 [-1.06; -0.15]$)(Fig. 5C). Dispersal remained positive density-dependent in the other two connectivity \times location treatments (Fig. 5C). (LINES 357-362)

References

References are adequate.

Discussion

C3.4> Lines 390-391 – Please clarify this sentence.

Answer to C3.4: We rewrote this which should now be clearer. This sentence, and the next, now read:

“However, our experiments are here inconsistent: low-density dispersal was higher in edge vs. core patches only in “reduced connectivity” landscapes in one dispersal experiment, and only in “reference” landscapes in the other dispersal experiment (Fig. 5, Supplementary Figures S.2.3 and S.2.4). There was also no clear divergence from the starting stock populations themselves (Fig. 5, Supplementary Figures S.2.3 and S.2.4).” (LINES 416-420). This also answers **C1.17**