# Review History

# RSPB-2019-0961.R0 (Original submission)

## Review form: Reviewer 1

**Recommendation**
Major revision is needed (please make suggestions in comments)

**Scientific importance: Is the manuscript an original and important contribution to its field?**
Good

**General interest: Is the paper of sufficient general interest?**
Excellent

**Quality of the paper: Is the overall quality of the paper suitable?**
Acceptable

**Is the length of the paper justified?**
Yes

**Should the paper be seen by a specialist statistical reviewer?**
No

**Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.**
No

**It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.**

   **Is it accessible?**
   N/A

   **Is it clear?**
   N/A

   **Is it adequate?**
   N/A

**Do you have any ethical concerns with this paper?**
No

**Comments to the Author**
Please see attached review as a pdf file. (See Appendix A)


# Review form: Reviewer 2

**Recommendation**
Accept with minor revision (please list in comments)

**Scientific importance: Is the manuscript an original and important contribution to its field?**
Good

**General interest: Is the paper of sufficient general interest?**
Good

**Quality of the paper: Is the overall quality of the paper suitable?**
Good

**Is the length of the paper justified?**
Yes

**Should the paper be seen by a specialist statistical reviewer?**
No

**Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.**
No

**It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.**

   **Is it accessible?**
   Yes

**Is it clear?**
Yes

**Is it adequate?**
Yes

**Do you have any ethical concerns with this paper?**
No

**Comments to the Author**
I enjoyed reading the manuscript and think it is important to have good theoretical evolutionary frameworks in ecology and evolution and in this case to understand changes in migratory behaviour. The authors propose a realistic model that can be tested and this is a valuable contribution.

General comments:
This is an interesting manuscript aiming to explain the genetic mechanisms underlying changes in migratory behaviour and the transition from migration to residency. The manuscript is well written and the simulation work done is clear but the hypotheses tested are not fully explained in the introduction and it is not clear to me how this paper will enable us to understand which species may be able to adapt to changing environmental conditions by changing their migratory behaviour and becoming partial migrants or fully residents. The model assumes there is a threshold for phenotypic expression of the migratory trait, this fits well with the blackcap data used to validate the model outcome and the crossing of the threshold triggers a change in behaviour through a change in fitness. The mechanisms underlying the model simulations are well explained but it is not clear how this can then be used with real data to make predictions of what species may be close to the threshold or not, it is also not clear how the threshold position may change and whether having this model will enable us to make predictions about threshold values for other species.

Specific comments:
Title: is this a question or a statement? Clarify

Abstract:
Line 11 and 16: There is a strong directional selection due to what factors and how does it operate, these sentences needs to be further explained. The authors argue that it is directional selection, and that residents have higher fitness (this is vague). Fitness may change with varying environmental conditions, how is fitness assessed? This is not fully clear in the abstract.

Introduction:
Line 74-78: The objectives of the manuscript need to be made clearer. Either state the hypotheses or explain the factors that will be explored, the link between the objectives and the methods is poorly done at present.
Line 55-65- explain how selection may operate to change migratory distance.

Methods:
Explain why you add up the allelic values of an individual and the environmental effects to yield the phenotypic value of migratory liability. The rational is not fully explained?
Environmental threshold positions are used to determine if individuals migrate or not, or their migratory propensity, but environmental variation could also affect fitness.
The assumption that migration is more detrimental, and eliminates at least 50% of migrants but does not influence the residents, is not a realistic under environmental change (unless it is assumed that residents have access to resources and are not affected by environmental conditions in the winter). Below certain environmental thresholds all resident individuals may have reduced survival and this could influence the fitness values obtained and the selection pressure acting on

migratory bahaviour. This was not used in the modelling procedure but if the fitness of residents is allowed to vary or decline under certain conditions (and climate extremes will be more likely in the future), would influence the results obtained.

Line 199: Explain why you used the thresholds of 100 and 300 values for environmental variability. Environmental variability is only used to define a migratory propensity threshold, hence it is likely that it will have low impact on the results. Environmental variability may influence fitness but this is not considered. Clarify why.

Results:
Assortative mating can produce a large impact on the frequency of migrancy in the case a discrete function is used, suggesting when assortative mating is present in a population different fitness functions may operate.

Discussion:
The parameters used to test the model were from a blackcap population, the fitness functions tested were contrasted with data from a blackcap population in Germany. How generalizable are these results? How can they help us understand/ predict changes in migratory behaviour for other organisms?

Hope these comments help the authors improve the manuscript.

# Decision letter (RSPB-2019-0961.R0)

03-Jul-2019

Dear Dr Pulido,

I am writing to inform you that your manuscript RSPB-2019-0961 entitled "How migratory populations become resident" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of the Associate Editor and the referees. We are all agreed that this is a very interesting topic and a potentially important manuscript, but the very thorough comments from the referees and AE have identified several issues that require substantial revisions. With this in mind we would be happy to consider a resubmission, provided the comments of the referees and the AE are fully addressed. However please note that this is not a provisional acceptance.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A 'response to referees' document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

With best wishes
Professor Loeske Kruuk
Editor
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member: 1
Comments to Author:
Dear authors,
This MS is a theoretical investigation into the evolution of migration/residency under climate change that should be of general interest. The model extends on previous work and by doing so generates some important new insights (e.g. the shape of fitness function, not the number of genes involved are crucial). The combination of developing a model and testing it with long-term empirical data is appealing as well, and make the MS stand out. Both reviewers and I think it has the potential to be suitable for PRSB. However, reviewer 1, myself and to a lesser extent reviewer 2 point out that the MS still needs quite a lot of work to be up to standard. Specifically reviewer 1 raises concerns about the robustness of the model results. Furthermore, both reviewers and I have substantial comments about the current and future empirical validation of the model, which need much more clarification/substantiation. Finally, reviewer 1 and I emphasizes that the hypothesis being tested/ aims of the MS are poorly explained/developed in the Introduction. As these are all points that go at the heart of the MS a quite substantial amount of works appears to be needed, but I acknowledge that this seems to be feasible and would not expect this to necessarily change the value of the MS.

Best wishes
Martijn van de Pol

Major comments:
1. End of introduction is too vague: "The aim of this model is to identify the most important parameters that affect the evolutionary dynamics of migratory behaviour, to understand and predict adaptive evolutionary changes in migration, and to make recommendations for future studies." Understand and predict what specifically? The MS would greatly benefit by clarifying the aims more specifically. And what empirical test will be done on the model predictions specifically?
2. A rather crucial aspect of the study is not explained well. How were empirical data derived used to test the models? Specifically, L210: "Model fit was compared by plotting the observed migratory activity over time against the loess line of best fit of each simulated scenario and a linear regression of the empirical data." I do not understand this sentence. Furthermore, it reads: "We compared the evolutionary response resulting from simulations with different combinations of parameters with the change in migratory activity documented in German blackcaps between 1988" . How is change in migratory activity measured? Fig. 5 plots empirical liability data, which sounds like something different than activity data. How does one measure liability in the field to make it comparable to how it is defined in the model. At the moment I have no idea what was done here without having to read the references given.
3. When looking at Fig. 5 I see a broad scatter of empirical data points with model predictions imposed. Is this considered a good or bad model fit, and based on what criteria? At the moment the most direct comparison of model and empirical data is hidden in the Suppl (Fig. S6) and the results section does not say much about the empirical data (and model fit) at all. This makes it hard for me to understand at the moment where key conclusions such as those in L315( The

expected rate of evolutionary change in migratory behaviour is in accord with that observed in the European blackcaps [8].) and L325 (In Southern German blackcaps, recent evolutionary changes towards residency [8] were most consistent with a scenario in which residents have the highest fitness and fitness declines exponentially with increasing migratory distance) are based on. This needs to be substantiated/ explained better to the reader.

Minor points
1. L16 populations
2. First paragraph of Introduction is very long, can it be cut into smaller pieces?
3. L49: explain how they are at odds. Any model is a simplification and all arguably all models are wrong, so adding more complexity to models may make them more realistic, but this does not make them necessarily more useful or predictive models. The authors appear to suggest that the additions they make are useful, so one would like to understand the argumentation for this a bit better.
4. L53: "To realistically model migration" As above point, is 'realistic modelling' the aim? Some rewording may be required.
5. L66 "the best data", based on what criteria? I can see that is very comprehensive data and a model system.
6. L70 "This prediction, however, has not been formalised and quantified in a model." Are you suggesting the model predicts in hindsight what was observed in the field? Some rewording maybe needed too. What is meant by a prediction being 'formalised and quantified'?
1. L210: "Model fit was compared by plotting the observed migratory activity over time against the loess line of best fit of each simulated scenario and a linear regression of the empirical data." I do not understand this sentence.
1. L222: "Fitness functions that imposed selection on migratory activity always led to a faster evolution towards residency than their corresponding step functions (Figure 2a,b)." Calrify what are step functions and what are the other ones. As far as I can see in Fig. 1 all fitness functions have a step at trait=0. Please clarify wording.
2. L288-291: where can we see this in the results?

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
Please see attached review as a pdf file.

Referee: 2

Comments to the Author(s)
I enjoyed reading the manuscript and think it is important to have good theoretical evolutionary frameworks in ecology and evolution and in this case to understand changes in migratory behaviour. The authors propose a realistic model that can be tested and this is a valuable contribution.

General comments:
This is an interesting manuscript aiming to explain the genetic mechanisms underlying changes in migratory behaviour and the transition from migration to residency. The manuscript is well written and the simulation work done is clear but the hypotheses tested are not fully explained in the introduction and it is not clear to me how this paper will enable us to understand which species may be able to adapt to changing environmental conditions by changing their migratory behaviour and becoming partial migrants or fully residents. The model assumes there is a threshold for phenotypic expression of the migratory trait, this fits well with the blackcap data used to validate the model outcome and the crossing of the threshold triggers a change in

behaviour through a change in fitness. The mechanisms underlying the model simulations are well explained but it is not clear how this can then be used with real data to make predictions of what species may be close to the threshold or not, it is also not clear how the threshold position may change and whether having this model will enable us to make predictions about threshold values for other species.

Specific comments:
Title: is this a question or a statement? Clarify

Abstract:
Line 11 and 16: There is a strong directional selection due to what factors and how does it operate, these sentences needs to be further explained. The authors argue that it is directional selection, and that residents have higher fitness (this is vague). Fitness may change with varying environmental conditions, how is fitness assessed? This is not fully clear in the abstract.

Introduction:
Line 74-78: The objectives of the manuscript need to be made clearer. Either state the hypotheses or explain the factors that will be explored, the link between the objectives and the methods is poorly done at present.
Line 55-65- explain how selection may operate to change migratory distance.

Methods:
Explain why you add up the allelic values of an individual and the environmental effects to yield the phenotypic value of migratory liability. The rational is not fully explained?
Environmental threshold positions are used to determine if individuals migrate or not, or their migratory propensity, but environmental variation could also affect fitness.
The assumption that migration is more detrimental, and eliminates at least 50% of migrants but does not influence the residents, is not a realistic under environmental change (unless it is assumed that residents have access to resources and are not affected by environmental conditions in the winter). Below certain environmental thresholds all resident individuals may have reduced survival and this could influence the fitness values obtained and the selection pressure acting on migratory bahaviour. This was not used in the modelling procedure but if the fitness of residents is allowed to vary or decline under certain conditions (and climate extremes will be more likely in the future), would influence the results obtained.
Line 199: Explain why you used the thresholds of 100 and 300 values for environmental variability. Environmental variability is only used to define a migratory propensity threshold, hence it is likely that it will have low impact on the results. Environmental variability may influence fitness but this is not considered. Clarify why.

Results:
Assortative mating can produce a large impact on the frequency of migrancy in the case a discrete function is used, suggesting when assortative mating is present in a population different fitness functions may operate.

Discussion:
The parameters used to test the model were from a blackcap population, the fitness functions tested were contrasted with data from a blackcap population in Germany. How generalizable are these results? How can they help us understand/ predict changes in migratory behaviour for other organisms?

Hope these comments help the authors improve the manuscript.

# Author's Response to Decision Letter for (RSPB-2019-0961.R0)

See Appendix B.

# RSPB-2019-3011.R0

## Review form: Reviewer 1

**Recommendation**
Accept with minor revision (please list in comments)

**Scientific importance: Is the manuscript an original and important contribution to its field?**
Excellent

**General interest: Is the paper of sufficient general interest?**
Excellent

**Quality of the paper: Is the overall quality of the paper suitable?**
Excellent

**Is the length of the paper justified?**
Yes

**Should the paper be seen by a specialist statistical reviewer?**
No

**Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.**
No

**It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.**

      **Is it accessible?**
      Yes

      **Is it clear?**
      Yes

      **Is it adequate?**
      Yes

**Do you have any ethical concerns with this paper?**
No

**Comments to the Author**
See attached pdf. (See Appendix C)

# Decision letter (RSPB-2019-3011.R0)

24-Jan-2020

Dear Francisco

Thank you for submitting a revised version of this manuscript. Your manuscript has now been peer reviewed and the reviews have been assessed by an Associate Editor. Whilst the reviewer is very happy with the revisions, he/she raises some minor points. The Associate Editor has also provided a very thorough review, identifying several areas where the text requires clarification. The reviewers' comments (not including confidential comments to the Editor) and the comments from the Associate Editor are included at the end of this email for your reference. We would like to invite you to revise your manuscript to address them.

We do not allow multiple rounds of revision so we urge you to make every effort to fully address all of the comments at this stage. If deemed necessary by the Associate Editor, your manuscript will be sent back to one or more of the original reviewers for assessment. If the original reviewers are not available we may invite new reviewers. Please note that we cannot guarantee eventual acceptance of your manuscript at this stage.

To submit your revision please log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions", click on "Create a Revision". Your manuscript number has been appended to denote a revision.

When submitting your revision please upload a file under "Response to Referees" in the "File Upload" section. This should document, point by point, how you have responded to the reviewers' and Editors' comments, and the adjustments you have made to the manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Your main manuscript should be submitted as a text file (doc, txt, rtf or tex), not a PDF. Your figures should be submitted as separate files and not included within the main manuscript file.

When revising your manuscript you should also ensure that it adheres to our editorial policies (https://royalsociety.org/journals/ethics-policies/). You should pay particular attention to the following:

Research ethics:
If your study contains research on humans please ensure that you detail in the methods section whether you obtained ethical approval from your local research ethics committee and gained informed consent to participate from each of the participants.

Use of animals and field studies:
If your study uses animals please include details in the methods section of any approval and licences given to carry out the study and include full details of how animal welfare standards were ensured. Field studies should be conducted in accordance with local legislation; please include details of the appropriate permission and licences that you obtained to carry out the field work.

Data accessibility and data citation:
It is a condition of publication that you make available the data and research materials supporting the results in the article. Datasets should be deposited in an appropriate publicly available repository and details of the associated accession number, link or DOI to the datasets must be included in the Data Accessibility section of the article

(https://royalsociety.org/journals/ethics-policies/data-sharing-mining/). Reference(s) to datasets should also be included in the reference list of the article with DOIs (where available).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should also be fully cited and listed in the references.

If you wish to submit your data to Dryad (http://datadryad.org/) and have not already done so you can submit your data via this link http://datadryad.org/submit?journalID=RSPB&manu=(Document not available), which will take you to your unique entry in the Dryad repository.

If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link.

For more information please see our open data policy http://royalsocietypublishing.org/data-sharing.

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI. Please try to submit all supplementary material as a single file.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

Please submit a copy of your revised paper within three weeks. If we do not hear from you within this time your manuscript will be rejected. If you are unable to meet this deadline please let us know as soon as possible, as we may be able to grant a short extension.

Thank you for submitting your manuscript to Proceedings B; we look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,
Professor Loeske Kruuk
Editor
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
Dear authors,

One of the original reviewers and myself have read the revision. We are both happy with the improvements made, and still think it is an interesting MS that fits the scope of the journal well. However, reviewer 1 has a list of remaining minor comments. Furthermore, I am still quite critical about parts of the writing of the MS in terms of clarity. I think it would be important to resolve these issues, hence my recommendation for major revision to address these long list of points.

Specific comments AE:

Introduction states that : "Yet, all models of the evolution of migration, or the persistence of

partial migration, were based on broad simplifications, not accounting for what we currently know on the genetics of migratory behaviour,i.e. that it is a polygenic threshold trait, controlled by many genes with small effects [6,8,14]. As it is unknown how these simplifying assumptions affect the reliability of models, there is an urgent need for an evolutionary model of migration, which incorporates these findings and is tested against empirical data." However, the results suggest actually that it is not the genetic structure that is crucial, as the number of loci has little influence on predictions and model fit. In fact, it is the shape of the fitness function and modelling liability and activity as a semi-continuous trait that has the strongest effect. The discussion makes the latter point already, but it does not place that into perspective with the remark in the Introduction that the polygenic architecture is crucial. You do make this point later in L354, which was a bit late in the Discussion, given that the genetic architecture argument is quite prominent in the Introduction as a reason why the models should become more 'realistic' (i.o.w. the aim of the study).

L54 "migratory activity" please define what to the reader what this key term implies. In l56 you state it has an amount and timing? What does amount refer to? the distance travelled, the energy spent on migration? A behavior? This needs to be very clear from the start, we now have to wait for an explanation till L210 (restlessness, but L320 talks about distance!?) The timing aspect is also not returned to in the remainder of the paper/model I think, so the mentioning of timing may raise expectations that are not delivered on

L63 is the model species for the study -> a model species? There are other model species I assume?

L67 to me the wording of "realistic modelling"" is still vague/confusing and this previous comment has not been resolved. Realistic to me suggest that it should reflect reality as much as possible, i.e. a very complex model that incoprporates as much of what we know about the biological processes involved. However, typically in modelling we only want to include the processes that we think that are important for reliable prediction or understanding. If this is also what the authors also mean (the response letter seems to confirm this, we probably agree what it important to include in models and what not), then I suggest to word more precisely what they mean. The term 'realistic' is ambiguous I think., at least for me as a reader.

L69 is it completely new or does it extend on previous models, if so cite and explain how it extends? In other words, it is unclear if the entire model is new (e.g. it is individual based and previous models were analytical) and in addition you added the genetics. Or that you use and exisiting individual based model and added the genetics. This is linked to L79: please make explicit what part you use from the previous model and what part is new.

L94-99: here I wondered what fitness function was used, as the type of selection is quite crucial. But this is explained later in methods, please make a forward reference to this.

L118-120 I do not understand this sentence, please reword.

L131: 1045 units of what? From which studies were these numbers taken?

L161: can we not see this function in Fig. 1?, if so please refer to Fig. 1 here too. If not, why not in Fig. 1

L175 superscript in equation

L198 100 and 300, units?

L207-213. This text is unclear, but crucial for the paper. Is restlessness a measure of migratory activity, liability or both. I assumed it was a measure of activity, and that liability is an unobservable (latent) variable. However, in Fig. 5 we see empirical measures of liability, suggesting restlessness is a measure of liability. In addition, above lines also suggest that migratory activity = migratory liability in the simulations, so are these concepts identical for non-resident birds? I was getting confused, which may be related to the fact that I am not working on migration, but the text should also be clear for non-experts.

Fig. 5 has redundancy, panel B shows all information from panel A plus more.

L296. Wording: If I look at Fig. 5 the difference between predictions with 10 and 200 loci is marginal. Also there is a crossover in the fits, whether the decline is stronger or weaker with 10 or 200 loci depends on the timescale.

L338 At the moment, the model is parameterized with data from the same population, and predictions are compared to the temporal trends in activity. A more independent validation

would be to show for another population that the model also predict change in migratory activity well, or across a spatial environmental gradient. This also sounds more productive than waiting for the year 2184 to see if your model has good predictive value. In other words, what future studies could bring/test these ideas further? Now, overall the Dsicussion text reads very much like 'the prove is already here in this MS, and this is how it works'. Some critical notes or ideas for further scrutinizing the proposed ideas would make the discussion more balanced.
L:351-353: this idea is not unique to traits such as migration behavior, there are many other reasons why strategies may coexists despite unequal fitness payoffs, A more geneal connection to the wider evolutionary literature would be useful.


Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).
See attached pdf.


# Author's Response to Decision Letter for (RSPB-2019-3011.R0)

See Appendix D.


# Decision letter (RSPB-2019-3011.R1)

24-Feb-2020

Dear Dr Pulido

I am pleased to inform you that your manuscript entitled "How migratory populations become resident" has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know.  Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

Your article has been estimated as being 10 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Paper charges
An e-mail request for payment of any related charges will be sent out after proof stage (within approximately 2-6 weeks). The preferred payment method is by credit card; however, other payment options are available

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,
Professor Loeske Kruuk
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Associate Editor:
Board Member
Comments to Author:
I am happy with all the responses and the revision.

# Appendix A

This is a potentially important paper offering a roadmap to assessing the role of microevolutionary change in migration in response to climate change. In addition, the paper presents an extension to the threshold model of a quantitative trait by Roff that allows for it to be semi-continuous. There are important qualitative consequences this extension, one of which being more efficient selection in the semi-continuous case. The research group of Pulido is in an ideal position to assess the role of microevolutionary change in migration in blackcaps, having dedicated a long career to estimating key parameters necessary to make reliable predictions.

My comments are mostly technical, but I think they need to be addressed for the paper to be considered reasonably robust in its methodology. It would also be helpful for the Discussion to indicate future work in blackcaps that would test the validity of the theory (without having to wait for 6 or more generations) and to assess whether alternate processes may be affecting migratory behaviour.

Major comments from the simulation methods (Supplementary Material)

- Initialization
    - The norm in quantitative genetic simulations is to undergo a longer "burn-in" period after initialization. Three generations is too short. The consequence of too short of a burn-in is that the number of alleles per locus, allele frequencies and corresponding effects are not well-characterized and do not conform to acceptable norms, such as what is expected under mutation-drift equilibrium (e.g., Wright 1931 Genetics 16: 97, Lynch and Hill 1986 Evolution 40: 915). A consequence is that the 25 - 100 generations of selection that follow initialization can reflect, in large part, the unrealistic nature of the initialization process.
        * One solution is to have a burn-in period that is longer, such that the population conforms to what is expected under mutation-drift equilibrium. For a diploid population of size 1000, this is expected to take over 4000 generations, on average, since the expected coalescence time for even a subsample of a population is approximately $4N$ generations (Tajima 1983 Genetics 105:437+). Although reasonable approximations are likely for $<4000$ generations since the distribution of the time to coalescence of an entire sample is approximately exponentially distributed.
        * Alternatively, one could invoke weak stabilizing selection during the burn-in, which would shorten the number of generations that are required to reach mutation-selection-drift equilibrium. For example, you could invoke weak stabilizing selection for a mean liability of 1045 that gives rise to a standard deviation of about 376 at equilibrium, as observed in blackcaps. Noting that the key aspect of this is to get the correct distribution of allele frequencies and corresponding phenotypic effects (not just the genetic variance).

* Note, I do not expect a correct initialization and burn-in to quantitatively (that much) nor qualitatively change the results. It's a matter of doing things in an acceptable manner and being a model for future research of a similar nature.

- Mutation
  - To connect to the quantitative genetic literature, your mutation model is analogous to a "house-of-cards" model, since you draw a new allelic effect with a particular mean and variance (Kingman 1977 Math. Proc. Camb. Phil. Soc. 81: 443+, 1978 J. App. Prob. 15:1+). This is in contrast to a "continuum of alleles" model, in which you add a random effect to the state of an allele (Kimura 1965 Proc. Natl. Acad. Sci 54: 731).
    * Note, from my understanding, new mutations have an average effect of $\mu_p/2n$, which is biased in the direction of migration, even after initialization. You may want to check this assumption, e.g. eq. 11 of the Supplement.
  - I liked how the paper sought to justify a per locus mutation rate (Mutation Rate section of SM, p. 8 of main text), but in the context of quantitative genetics, a consideration of the phenotypic effect of mutations, not just the rate, is needed. The combination of mutation rate and phenotypic effect gives rise to a property called "mutation variance" (ref. Lynch and Walsh 1998 "Genetics and Analysis of Quantitative Traits": Sinauer, p. 328 - 340). Often mutational variance is measured relative to environmental variance, which allows for standardization and comparisons between organisms. A summary of empirical estimates of these mutational heritabilities is provided on p. 338 in Lynch and Walsh (1998). It would add robustness to calculate the mutational heritability for your study to check that it is acceptable compared to published values. Mutational variance and heritability is important because it will affect the response to selection and the replenishment of genetic variance lost by selection.

- Strength of assortative mating
  - I was not able to follow your explanation. Can it be simplified to focus attention on the key principle that results in a pattern of assortative mating?

Major comments from simulation model (Main Text)

- Tested parameters: (i) Strength and form of selection
  - There was not enough support given for the magnitude of the drop in fitness of residents versus migrants (a 50% drop). All of the supporting citations are for birds other than blackcaps and the context and direct correspondence to blackcaps is not clear.

∗ I am wondering if alternatively you could "estimate" the drop in fitness using the data in figure 5a? It looks like the exponential model with a 50% reduction in fitness is a good fit. Having an estimate of parameters in the fitness function that arises from data would be more robust.

∗ If there really is a 50% drop in fitness presently then it is not clear how migration seemed to be maintained prior to the recent era. It would be informative to explain/infer processes that would have allowed for the maintenance of migration prior to the recent era. By the logic presented on lines 165 - 171 (p. 8-9, main text), I am wondering how migration could have persisted historically and what brought about the large shift such that there is a 50% decline, presently.

Minor comments

- A stronger argument needs to be made about invoking only genetic change and not plasticity in the case of blackcaps. The statement on lines 32-33 is not adequate, particularly given the position of the research group who author the present paper. There is an extensive literature on plasticity related to migration, at least theoretically (e.g. McNamara, Houston and colleagues). I am not as familiar with empirical work on plasticity and migration, but see Gienapp et al. (2007, Climate Res. 35: 25+). For other mechanisms, see Gill et al. (2014, Proc. Roy. Soc. B 281: 20132161). Furthermore, in Pulido & Berthold (2010, Proc. Natl. Acad Sci USA 107:7341+) there is a reaction norm associated with hatching date and migratory activity, from my understanding. Plasticity could weaken the response to selection.

- line 16, p.2: Insert "a" before "resident"

- lines 28-29, p. 3: It is not clear what is meant by "migratory organisms need to re-adjust their behaviour, for migration to remain adaptive." Please break this down more. It seems like you are thinking that migration is multifaceted, so how can migration remain adaptive, if there is directional selection against it, such that residency is optimal?

- lines 90 - 91, p. 5: If a revision includes plasticity, perhaps have migration activity be a probabilistic.

- lines 93 - 95, p. 5: Awkward reference to Figure S1.

- lines 118 - 119, p. 6: Revise corresponding sentence, awkwardly phrased presently.

- lines 139 - 141, p. 7: Not clear. One sentence is not enough to explain this aspect of the simulation for the Main Text.

- lines 153-154, p. 8: I do not understand what is meant by "allele frequencies reach an asymptote"

- lines 187 - 189, p. 9: Note here that allelic effects were scaled to be smaller with more loci.

# Appendix B

## Associate Editor
Board Member: 1
Comments to Author:

**Major comments:**

*1.     End of introduction is too vague: "The aim of this model is to identify the most important parameters that affect the evolutionary dynamics of migratory behaviour, to understand and predict adaptive evolutionary changes in migration, and to make recommendations for future studies." Understand and predict what specifically? The MS would greatly benefit by clarifying the aims more specifically. And what empirical test will be done on the model predictions specifically?*

This paragraph has been rewritten following your advice and the comments of reviewer 2. We hope that this change makes the specific aims of the study clearer to the reader.

*2.     A rather crucial aspect of the study is not explained well. How were empirical data derived used to test the models? Specifically, L210: "Model fit was compared by plotting the observed migratory activity over time against the loess line of best fit of each simulated scenario and a linear regression of the empirical data." I do not understand this sentence. Furthermore, it reads: "We compared the evolutionary response resulting from simulations with different combinations of parameters with the change in migratory activity documented in German blackcaps between 1988" . How is change in migratory activity measured? Fig. 5 plots empirical liability data, which sounds like something different than activity data. How does one measure liability in the field to make it comparable to how it is defined in the model. At the moment I have no idea what was done here without having to read the references given.*

We have made several changes to clarify this point:

1. We have re-written the statement that generated confusion. (see, "Analysis of simulation results" in "Methods")

2. We have added an explanation of how the migratory activity data were estimated and that they are comparable to simulation data (see, "Analysis of simulation results" in "Methods")

3. To help avoid confusion between migratory activity and migratory liability, the difference is explained in the introduction and in the methods, we changed the figure so that it should now become clear that residents have zero migratory activity, but migratory activity equals migratory liability for migrants

*3.      When looking at Fig. 5 I see a broad scatter of empirical data points with model predictions imposed. Is this considered a good or bad model fit, and based on what criteria? At the moment the most direct comparison of model and empirical data is hidden in the Suppl (Fig. S6) and the results section does not say much about the empirical data (and model fit) at all. This makes it hard for me to understand at the moment where key conclusions such as those in L315( The expected rate of evolutionary change in migratory behaviour is in accord with that observed in the European blackcaps [8].) and L325 (In Southern German blackcaps, recent evolutionary changes towards residency [8] were most consistent with a scenario in which residents have the highest fitness and fitness declines exponentially with increasing migratory distance) are based on. This needs to be substantiated/ explained better to the reader.*

1. To make the visual assessment of model fit clearer and solve the "overplotting" issue with the original figure, we modified figure 5 so that the empirical data are summarised using a boxplot rather than a scatterplot.

2. We modified the plot comparing observed and predicted mean migratory activity in the supplementary materials (see, Fig. S6) by including all fitness functions and genetic architectures to make it easier to compare the accuracy of predictions.

2. Model fit was further assessed quantitatively by calculating the variance explained by cross validation and the mean absolute error (see, explanation in "Methods" and "Results").

3. Changes to the "Prediction of evolutionary response in the blackcap" in "Results" were made so that it now should become clear to the reader what results substantiate these conclusions.

**Minor points**

*4.      L16 populations*

Corrected.

*5.      First paragraph of Introduction is very long, can it be cut into smaller pieces?*

We have split the first paragraph and have shortened it a bit to make it more readable.

*6.      L49: explain how they are at odds. Any model is a simplification and all arguably all models are wrong, so adding more complexity to models may make*

*them more realistic, but this does not make them necessarily more useful or predictive models. The authors appear to suggest that the additions they make are useful, so one would like to understand the argumentation for this a bit better.*

We agree with the reviewer that any model is a simplification, and necessarily has to make simplifying assumptions to be tractable and useful. However, simplifications need to be realistic and their effects need to be tested against empirical facts. Many of the previous models made assumptions that were clearly at odds with what we know on the genetics of migratory behaviour, for instance, assuming a monogenic trait or disregarding that it is a threshold trait. However, they did not test how this affected the dynamics of their model and results. Here we propose using current knowledge on the genetics of migratory behaviour to make a more realistic (but not necessarily more complex) model of migration and to test the effects of making different assumptions (e.g. number of genes, assortative mating, environmental effects) against empirical data. We have rewritten this paragraph and hope that the point we are making is clearer now.

*7.    L53: "To realistically model migration" As above point, is 'realistic modelling' the aim? Some rewording may be required.*

We changed the wording accordingly. See, changes in text and response to previous point.

*8.    L66 "the best data", based on what criteria? I can see that is very comprehensive data and a model system.*

Changed "best data" for "most comprehensive data".

*9.    L70 "This prediction, however, has not been formalised and quantified in a model." Are you suggesting the model predicts in hindsight what was observed in the field? Some rewording maybe needed too. What is meant by a prediction being 'formalised and quantified'?*

Our point is that to make quantitative predictions about the evolution of migration, such as the time it takes for residency to evolve under different forms of selection, and the dynamics of how migratory behaviour changes over time, we a need a mathematical model that takes into account what we currently know about this trait. We have changed this paragraph (and this sentence was deleted), so that there is should be no confusion on its meaning now.

*10.    L210: "Model fit was compared by plotting the observed migratory activity over time against the loess line of best fit of each simulated scenario and a linear regression of the empirical data." I do not understand this sentence.*

See, changes in text. This section of the methods was rewritten to improve clarity and to incorporate the changes in response to major comment 3 (see, above).

*11.    L222: "Fitness functions that imposed selection on migratory activity always led to a faster evolution towards residency than their corresponding step functions (Figure 2a,b)." Calrify what are step functions and what are the other ones. As far as I can see in Fig. 1 all fitness functions have a step at trait=0. Please clarify wording.*

We have made changes throughout the complete manuscript using a consistent terminology to ensure the two types of fitness functions are clear and not confounded. Discrete fitness functions are those where all migrants have the same fitness regardless of the amount of migratory activity. Semi-continuous fitness functions are those where there is a difference in fitness between migrants and residents but there is also a difference in fitness between migrants showing different amounts of migratory activity.

*12.     L288-291: where can we see this in the results?*

See changes in the text (this section was re-written). This result can be seen in figure 5 and this was made clear.

# *Referee 1:*

*This is a potentially important paper offering a roadmap to assessing the role of microevolutionary change in migration in response to climate change. In addition, the paper presents an extension to the threshold model of a quantitative trait by Roff that allows for it to be semi-continuous. There are important qualitative consequences this extension, one of which being more efficient selection in the semi-continuous case. The research group of Pulido is in an ideal position to assess the role of microevolutionary change in migration in blackcaps, having dedicated a long career to estimating key parameters necessary to make reliable predictions. My comments are mostly technical, but I think they need to be addressed for the paper to be considered reasonably robust in its methodology. It would also be helpful for the Discussion to indicate future work in blackcaps that would test the validity of the theory (without having to wait for 6 or more generations) and to assess whether alternate processes may be affecting migratory behaviour.*

We are very thankful for the thorough analysis of the technical aspects of the model and for the very helpful recommendations provided by reviewer 1. To address the issues raised by the reviewer, we have modified the simulation model and repeated all simulations. We believe that these changes have improved the manuscript by addressing some important concerns about mutation and initial conditions, and also by clarifying how the model we present fits into the quantitative genetics literature.

We are currently not conducting experiments with blackcaps so that we cannot run experiments that could provide data on different aspects of our model. However, throughout the discussion (and particularly at the end), we point at areas, for which we have a clear deficit of data, and which need to be collected to refine the model. We think that our model is a first step to understand the determinants and consequences of evolutionary changes in migration, and that future studies need to propose experiments that can test some of the conclusions we draw.

*Major comments from the simulation methods (Supplementary Material)*

• *Initialization*
*– The norm in quantitative genetic simulations is to undergo a longer "burn-in" period after initialization. Three generations is too short. The consequence of too short of a burn-in is that the number of alleles per locus, allele frequencies and corresponding effects are not well- characterized and do not conform to acceptable norms, such as what is expected under mutation-drift equilibrium (e.g., Wright 1931 Genetics 16: 97, Lynch and Hill 1986 Evolution 40: 915). A consequence is that the 25 - 100 generations of selection that follow initialization can reflect, in large part, the unrealistic nature of the initialization process.*

*\* One solution is to have a burn-in period that is longer, such that the population conforms to what is expected under mutation- drift equilibrium. For a diploid population of size 1000, this is expected to take over 4000 generations, on average, since the ex- pected coalescence time for even a subsample of a population is approximately 4N generations (Tajima 1983 Genetics 105:437+). Although reasonable approximations are likely for <4000 generations since the distribution of the time to coalescence of an entire sample is approximately exponentially distributed.*
*\* Alternatively, one could invoke weak stabilizing selection during the burn-in, which would shorten the number of generations that are required to reach mutation- selection-drift equilibrium. For example, you could invoke weak stabilizing selection for a mean liability of 1045 that gives rise to a standard deviation of about 376 at equilibrium, as observed in blackcaps. Noting that the key aspect of this is to get the correct distribution of allele frequencies and corresponding phenotypic effects (not just the genetic variance).*

*\* Note, I do not expect a correct initialization and burn-in to quantitatively (that much) nor qualitatively change the results. It's a matter of doing things in an acceptable manner and being a model for future research of a similar nature.*

We completely agree with the reviewer that by using a short burn-in period, the number of alleles per locus, allele frequencies and corresponding allele effects may not correspond to those of a population at mutation-drift or mutation-selection-drift equilibrium. In our original approach, we assumed that this population was not in equilibrium. However, this may have significantly affected results. Our original approach was only effective at setting the initial genetic variance to match that of the blackcap population. Therefore, our assumption that there are many alleles with normally distributed effects that segregate at each locus may not be hold in natural populations. Given that this has the potential to change how populations respond to selection, we have followed the reviewer's suggestion and included a period of weak stabilizing selection for a total of 15000 generations of burn-in (see, "Initialisation" in "Methods") before introducing selection towards residency.

• *Mutation*
*– To connect to the quantitative genetic literature, your mutation model is analogous to a "house-of-cards" model, since you draw a new alelic effect with a particular mean and variance (Kingman 1977 Math. Proc. Camb. Phil. Soc. 81: 443+, 1978 J.*

*This is in contrast to a "continuum of alleles" model, in which you add a random effect to the state of an allele (Kimura 1965 Proc. Natl. Acad. Sci 54: 731).*

*\* Note, from my understanding, new mutations have an average effect of µp/2n, which is biased in the direction of migration, even after initialization. You may want to check this assumption, e.g. eq. 11 of the Supplement. – I liked how the paper sought to justify a per locus mutation rate (Mutation Rate section of SM, p. 8 of main text), but in the context of quantitative genetics, a consideration of the phenotypic effect of mutations, not just the rate, is needed. The combination of mutation rate and phenotypic effect gives rise to a property called "mutation variance" (ref. Lynch and Walsh 1998 "Genetics and Analysis of Quantitative Traits": Sinauer, p. 328 - 340). Often mutational variance is measured relative to environmental variance, which allows for standardization and comparisons between organisms. A summary of empirical estimates of these mutational heritabilities is provided on p. 338 in Lynch and Walsh (1998). It would add robustness to calculate the mutational heritability for your study to check that it is acceptable compared to published values. Mutational variance and heritability is important because it will affect the response to selection and the replenishment of genetic variance lost by selection.*

**(i) Mutation model:** As the referee rightly points out, our original mutation model was such that the effect of new mutations was biased towards the initial state of the population. Initially the mutations would be equally likely to increase or decrease the contribution towards migratory activity of the affected allele. However, as populations respond to selection and the mean allele effect becomes predominantly negative, mutations would tend to increase migratory activity. Some bias is inherent in the house-of-cards mutation model where mutations have an effect that is independent on the previous allele effect. Thus, even in a house-of-cards mutation model where the average effect of new mutations is sampled from a distribution with a mean of 0 would be biased towards this particular state, with the strength of bias proportional to the distance between this value and the mean allele effect in the population. However, there is no reasonable criterion to determine what state mutations may be biased towards. Based on these concerns, we changed the mutation model to a continuum-of-alleles model where the effect of an allele after mutation corresponds to that allele's effect before mutation plus a random mutational effect of mean zero (see, "Mutation" in "Methods").

**(ii) Mutation rate:** In our original manuscript we estimated the probability of mutations occurring at loci affecting migratory activity in blackcaps based on assumptions about the per-base-pair mutation rate and mutational target size. However, reviewer 1 rightly pointed out that, in quantitative genetics, the relevant parameter that determines the input of genetic variation by mutation is the mutational variance, which is often reported after being scaled by the environmental variance as the mutational heritability. Estimating the mutational heritability for our original choice of parameters gives values that are orders of magnitude below those reported for other quantitative traits ($h_m^2 = 5 \times 10^{-9}$ with 200 loci, and $5 \times 10^{-7}$ with 10). This is not surprising given that similar attempts to arrive at an estimate of the mutational heritability often fall short of the values empirically estimated for other quantitative traits (Walsh and Lynch, 2018). Therefore, in our new simulations, we decided to use the typical value of mutational heritability observed for other quantitative traits. We believe that this more likely to represents the mutational heritability for migratory activity than our original

estimate that was based on assumptions about the mutation rate and the variance in the effect of new mutations for which there is less information.


*• Strength of assortative mating*
*– I was not able to follow your explanation. Can it be simplified to focus attention on the key principle that results in a pattern of assortative mating?*

We changed the wording on the main text to improve clarity. The key principle should now be clear in the main text, while the supplementary material we give all details necessary to reproduce the work.

*Major comments from simulation model (Main Text)*
*• Tested parameters: (i) Strength and form of selection*
*– There was not enough support given for the magnitude of the drop in fitness of residents versus migrants (a 50% drop). All of the supporting citations are for birds other than blackcaps and the context and direct correspondence to blackcaps is not clear.*

*\*I am wondering if alternatively you could "estimate" the drop in fitness using the data in figure 5a? It looks like the exponential model with a 50% reduction in fitness is a good fit. Having an estimate of parameters in the fitness function that arises from data would be more robust.*

Unfortunately, there are no data for blackcaps on the difference in fitness between residents and migrants. We, therefore, cite available data on other migratory bird species and conducted a sensitivity analysis to test how the assumption that the minimum difference in fitness between migrants and residents was 50% affected the results of the simulations. We ran additional simulations assuming this value to be 0 or 25 % with the fitness function that best fit the data (exponential) and with the one that led to the second best fit (logistic). Interestingly, decreasing the minimum fitness difference between migrants and residents had little impact on the change in the proportion of resident or heritability under the exponential fitness function. With the logistic fitness function, however, decreasing this value led to a slower response and would therefore be associated with an even worse fit to the data (Fig. S6). Therefore, assuming a value of 50% is conservative regarding our claim that the exponential fitness function provides a better fit to the data. Moreover, the small effect under the exponential fitness function demonstrates that this parameter is not critical for explaining the observed evolutionary change in migratory activity in blackcaps.

*\*If there really is a 50% drop in fitness presently then it is not clear how migration seemed to be maintained prior to the recent era. It would be informative to explain/infer processes that would have allowed for the maintenance of migration prior to the recent era. By the logic presented on lines 165 - 171 (p. 8-9, main text), I am wondering how migration could have persisted historically and what brought about the large shift such that there is a 50% decline, presently.*

The explanation of the new initialisation process in "Methods" clarifies this. It is thought that migration was advantageous before recent climate change has created environmental conditions that select for residency.

*Minor comments*

*• A stronger argument needs to be made about invoking only genetic change and not plasticity in the case of blackcaps. The statement on lines 32-33 is not adequate, particularly given the position of the research group who author the present paper. There is an extensive literature on plasticity related to migration, at least theoretically (e.g. McNamara, Houston and colleagues). I am not as familiar with empirical work on plasticity and migration, but see Gienapp et al. (2007, Climate Res. 35: 25+). For other mechanisms, see Gill et al. (2014, Proc. Roy. Soc. B 281: 20132161). Furthermore, in Pulido & Berthold (2010, Proc. Natl. Acad Sci USA 107:7341+) there is a reaction norm associated with hatching date and migratory activity, from my understanding. Plasticity could weaken the response to selection.*

This is a good point. Indeed, there is some evidence that phenotypic change in migratory behaviour is not only due to evolutionary change but also a consequence of adaptive phenotypic plasticity, and that this could modify evolutionary response. However, in the present model, we chose not to consider phenotypic plasticity for two reasons:

(1) The data on evolutionary change in migratory activity in the blackcap were obtained in the laboratory, i.e. under controlled conditions, which were chosen to minimize environmental variation. For comparing our predictions with the empirical data obtained in the blackcap, we have to assume that the contribution of phenotypic plasticity to phenotypic change is zero.

(2) We have no good data in general, and specifically for the blackcap on how phenotypic plasticity will affect the rate and direction of evolutionary response. The reaction norm of onset of migration with hatching date could be a good starting point to include in our simulation model. However, it is not straightforward since the effect of hatching date on the amount of migratory activity (i.e. on liability) is week and indirect. It is mainly due to the effect of hatching date on the timing of migration and the genetic correlation between timing and amount of migration. Moreover, we have little reliable data on the changes in the timing of breeding in "our" southern German blackcap population.

Since we think this is important aspect to consider in future models, and also to consider in the interpretation of results, we have included this in the discussion.


• line 16, p.2: Insert "a" before "resident"

We corrected this to "resident populations" as suggested by the Associate Editor.


• lines 28-29, p. 3: It is not clear what is meant by "migratory organ- isms need to re-adjust their behaviour, for migration to remain adaptive." Please break this down more. It seems like you are thinking that migration is multifaceted, so how can migration remain adaptive, if there is directional selection against it, such that residency is optimal?

Yes, we believe that an environmental change not only changes between conditions that favour residency and migration, but that changes may be more subtle and

gradual. Climate change may thus first favour individuals that advance spring migration and delay autumn migration a little. At the same time a northward shift of the wintering area may be favoured. As climate change advances, the optimum of the timing of migration and migration distance may shift further. I have reworded this sentence to make this clearer.

• lines 90 - 91, p. 5: If a revision includes plasticity, perhaps have migration activity be a probabilistic.

Including phenotypic plasticity is outside the scopes of the manuscript. In this study, we focussed on evolutionary change because this is the mechanism that had been shown to determine the observed phenotypic change in blackcap migratory behaviour. Extending the model to include phenotypic plasticity would be very interesting but would be difficult to test and validate due to a lack of empirical data (see above)

• lines 93 - 95, p. 5: Awkward reference to Figure S1.

Corrected

• lines 118 - 119, p. 6: Revise corresponding sentence, awkwardly phrased presently.

We changed the wording to improve clarity.

• lines 139 - 141, p. 7: Not clear. One sentence is not enough to explain this aspect of the simulation for the Main Text.

We revised the wording and added a reference, where further discussion and explanation of the model can be found.

• lines 153-154, p. 8: I do not understand what is meant by "allele frequencies reach an asymptote"

This section was re-written.

• lines 187 - 189, p. 9: Note here that allelic effects were scaled to be smaller with more loci.

This section was re-written.

# Referee 2

*General comments:*
*This is an interesting manuscript aiming to explain the genetic mechanisms underlying changes in migratory behaviour and the transition from migration to residency. The manuscript is well written and the simulation work done is clear but the hypotheses tested are not fully explained in the introduction and it is not clear to me how this paper will enable us to understand which species may be able to adapt to changing environmental conditions by changing their migratory behaviour and becoming partial migrants or fully residents. The model assumes there is a threshold for phenotypic expression of the migratory trait, this fits well with the blackcap data used to validate the model outcome and the crossing of the threshold triggers a change in behaviour through a change in fitness. The mechanisms underlying the model simulations are well explained but it is not clear how this can then be used with real data to make predictions of what species may be close to the threshold or not, it is also not clear how the threshold position may change and whether having this model will enable us to make predictions about threshold values for other species.*

We are pleased that you find the manuscript interesting and well written. Unfortunately, our study cannot tell, which migratory species specifically will adapt to climate change and which will become partially migratory or resident, since we do not have the data we would need (i.e. genetic and phenotypic variance, fitness functions, etc.) to make such predictions. However, if these data are collected – and we identify the data are essential (and which are not) to make these predictions – the model can make clear prediction on the short-term and long-term evolution of migratory behaviour. We believe that by identifying which data are crucial (e.g. fitness functions, genetic variation) and which are not (e.g. number of genes, assortative mating) much less data need to be collected to make good predictions.

*Specific comments:*
*Title: is this a question or a statement? Clarify*
It is a statement. Our paper describes how migratory populations become resident.

*Abstract:*
*Line 11 and 16: There is a strong directional selection due to what factors and how does it operate, these sentences needs to be further explained. The authors argue that it is directional selection, and that residents have higher fitness (this is vague). Fitness may change with varying environmental conditions, how is fitness assessed? This is not fully clear in the abstract.*

Our model is a genetic model and what is important for evolutionary trajectories is the strength of selection and the form of the fitness function. In our paper, we identify the kind of fitness differences that would need to exist between individuals with different migratory behaviour to explain the evolutionary response that was observed. Determining the specific ecological factors that generate this selective pressure is beyond the scope of this work.

*Introduction:*
*Line 74-78: The objectives of the manuscript need to be made clearer. Either state the hypotheses or explain the factors that will be explored, the link between the objectives and the methods is poorly done at present.*

This paragraph has been rewritten following the advice of the reviewer and the associate editor (see above)


*Line 55-65- explain how selection may operate to change migratory distance.*

This paragraph is on the threshold model of migration and how it needs to be modified to realistically model the evolution of migration. We mention the advantages of migrating shorter migration distances in different parts of the manuscript: in the introduction (lines 37-39 of the original manuscript) in "Materials and Methods" (lines 167-171 of original manuscript) and, in more detail, in the discussion (lines 316-324 of the original manuscript).


*Methods:*
*Explain why you add up the allelic values of an individual and the environmental effects to yield the phenotypic value of migratory liability. The rational is not fully explained?*

This is the standard approach in quantitative genetics. The genetic value is the sum of the allelic values and the phenotypic value is the sum of the genetic value and an environmental effect. Since our model follows directly from this principle, we think it is better to prioritise explaining those parts of the model which are novel and require more explanation.


*Environmental threshold positions are used to determine if individuals migrate or not, or their migratory propensity, but environmental variation could also affect fitness. The assumption that migration is more detrimental, and eliminates at least 50% of migrants but does not influence the residents, is not a realistic under environmental change (unless it is assumed that residents have access to resources and are not affected by environmental conditions in the winter). Below certain environmental thresholds all resident individuals may have reduced survival and this could influence the fitness values obtained and the selection pressure acting on migratory bahaviour. This was not used in the modelling procedure but if the fitness of residents is allowed to vary or decline under certain conditions (and climate extremes will be more likely in the future), would influence the results obtained.*

- We conducted a sensitivity analysis to explore the effect of the minimum difference in fitness between migrants and residents (see, response to referee 1).

- Note that a 50% difference in fitness between migrants and residents implies that residents are twice as likely to survive and reproduce. We do not assume that all residents survive, and half the migrants die due to environmental change. The fitness functions are a feature of the model that represents the nature of selection acting on this population without explicitly modelling the ecological and demographic processes that affect the fitness of migrants and residents. We agree that it is not straightforward to predict how environmental change will affect the fitness of individuals with different migratory behaviours and we point this out in the last paragraph of the discussion.

*Line 199: Explain why you used the thresholds of 100 and 300 values for environmental variability. Environmental variability is only used to define a migratory propensity threshold, hence it is likely that it will have low impact on the results. Environmental variability may influence fitness but this is not considered. Clarify why.*

- We have added a justification for the choice of the values for the threshold variance in the text. There are no empirical estimates for this parameter so we have explored a reasonable range of values.

- We don't think it is trivial to claim that increased threshold variance will have no effect on the results. In fact, if we assume a discrete fitness function as done in all previous evolutionary models of migration, increased threshold variance greatly speeds up the evolutionary response (see, Fig.3). In our simulations, introducing threshold variance affects fitness indirectly by influencing individual phenotypes. We did not model environmental conditions or the ecological processes that may lead to these effects since it is not the focus of the manuscript.

*Results:*
*Assortative mating can produce a large impact on the frequency of migrancy in the case a discrete function is used, suggesting when assortative mating is present in a population different fitness functions may operate.*

We discuss the situations where assortative mating is expected to influence the evolutionary response (see, lines 384 -391 of original manuscript) based on this result. This does not mean that if there is assortative mating there is a different fitness regime. It means that if there is assortative mating it will only modify selection response and the rate of evolutionary change under some (discrete) fitness functions, but not under the exponential fitness function, which we found to fit the blackcap data best

*Discussion:*
*The parameters used to test the model were from a blackcap population, the fitness functions tested were contrasted with data from a blackcap population in Germany. How generalizable are these results? How can they help us understand/ predict changes in migratory behaviour for other organisms?*

Our model is very flexible and can be used in other organisms and for simulating evolution of other semi-continuous traits. The general conclusion we draw on the effects of different factors, like the number of genes or fitness functions, hold for any species or semi-continuous trait. However, the exact fitness function or the speed of selection will depend on the data on phenotypic and genetic variation and change over time for each species and population. Yet if these data are available, the rate and direction of evolutionary change of a trait and its variance can be readily predicted using our model.

# Appendix C

The authors did an excellent job revising the paper, such that the methods are inline with standard quantitative genetic approaches. These revisions aid the robustness of the paper, the ability to compare with existing literature and allow for a useful baseline for future research. My comments at this point are very minor.

- line 9 & elsewhere: Suggest replacing "We show . . ." with "Our model indicates . . ." All of the results arise from a model and my opinion is that the statement "We show . . . is the compatible with . . ." is too strong of a statement. (Note, this recommendation is based on me making general and strong statements in some of my papers, without the important qualifier that the results arise from a particular model, with particular assumptions, with some statements having come back to be shown to be too general and too strong.)

- lines 23-25: Is there a citation for this statement?

- lines 31: Is the state of the field such that we can say "most" migratory species show little flexiblity? With the use of geolocators and stable isotopes, I think we are better appreciating that an individual's migratory behaviour can respond to environmental conditions and their state (such as health) [Hegemann et al. 2015, Am. Nat.]. Sometimes this leads to a lack of a complete migration within a year.

- line 69: Is the use of the word "new" needed?

- line 80: Suggest: "Individuals are diploid at each of $n$ loci."

- line 82: Suggest: "The sum of . . . , such that alleles are co-dominant."

- line 87: "0" -> "no"

- line 99: Add a "see below" otherwise a reader may think the explanation of selection is incomplete.

- line 113-114: Note that in some cases mutational bias is appropriate. For example, a phenotype may be near its extreme of possible values, such that mutation is biased in the opposite direction. Instead of indicating an "advantage" merely state the difference.

line 120: I think "above" should be "below".

line 125: "environmental effects" -> "variation in the environment"; (To not confuse the reader with a genetic "environmental effect".

lines 175, 177: subscript the "m"

lines 226, 228: It would have been clarifying to add the word "differential" before "selection".

line 230: "some" -> "a few"

line 265: "neither" -> "not"

line 346, 396: "will" -> "is expected to"

Supplement:

- p.2, 1st paragraph: Point reader to $\mu_p$ and $\sigma_\alpha$ definitions and values table to get them started about where to refer to definitions and values.

# Appendix D

## Response to reviewer's comments

We would like to thank again the referee and the associate editor for their effort in revising our new version of our manuscript and for their helpful comments and suggestions.

In the following, we explain point-by-point how we have addressed your comments. As in the previous response letter, we have put your comments in a cursive font and our answers in blue. We hope that this makes it easier to follow. In the uploaded "track-changes" version of the manuscript you will easily find the changes made.

_Specific comments AE:_

_Introduction states that : "Yet, all models of the evolution of migration, or the persistence of partial migration, were based on broad simplifications, not accounting for what we currently know on the genetics of migratory behaviour,i.e. that it is a polygenic threshold trait, controlled by many genes with small effects [6,8,14]. As it is unknown how these simplifying assumptions affect the reliability of models, there is an urgent need for an evolutionary model of migration, which incorporates these findings and is tested against empirical data." However, the results suggest actually that it is not the genetic structure that is crucial, as the number of loci has little influence on predictions and model fit. In fact, it is the shape of the fitness function and modelling liability and activity as a semi-continuous trait that has the strongest effect. The discussion makes the latter point already, but it does not place that into perspective with the remark in the Introduction that the polygenic architecture is crucial. You do make this point later in L354, which was a bit late in the Discussion, given that the genetic architecture argument is quite prominent in the Introduction as a reason why the models should become more 'realistic' (i.o.w. the aim of the study)._

The first aim of our study was "to assess the importance of the parameters that are likely to affect evolutionary change in migratory behaviour". To fulfil this aim, we explored the effect of the number of genes on evolutionary change and on the change in genetic variation. One of the conclusions was that the number of genes was not important for the rate of change, but it was for the amount of genetic variation preserved if selection was strong. We think that these conclusions are based on realistic assumptions about the number of genes controlling migration, as values were derived from empirical studies on migration or other complex traits. All these studies indicate (as we discuss in the text) that migratory behaviour is controlled by, at least, 4-18 loci, but most likely the number of genes involved is much higher. Clearly, migratory behaviour is not controlled by one gene with two alleles, as was assumed in some models on the evolution of migration. Unfortunately, we do not know, how the dynamics of evolutionary change would change if we assumed migration to be monogenic, but we expect results to differ clearly. Thus, we think that for modelling the evolution of migration it is important to take into account that it is expressed as a semi-continuous threshold trait and that it is polygenic, but it is not crucial to know the exact number, as long as the number of genes is in a realistic range (>4, but the exact number would needed to be explored).

We have tried to make this point clearer in the discussion (see, lines 300-302 and 343-355).

*L54 "migratory activity" please define what to the reader what this key term implies. In l56 you state it has an amount and timing? What does amount refer to? the distance travelled, the energy spent on migration? A behavior? This needs to be very clear from the start, we now have to wait for an explanation till L210 (restlessness, but L320 talks about distance!?) The timing aspect is also not returned to in the remainder of the paper/model I think, so the mentioning of timing may raise expectations that are not delivered on*

We have re-written this section, and we hope that this change makes it easier to understand what migratory activity is in this context.

*L63 is the model species for the study -> a model species? There are other model species I assume?*

Although the European blackcap is the most important model, particularly for the control and evolution of passerine migration, there are, of course other model species. We have changed the sentence, to avoid misunderstandings.

*L67 to me the wording of "realistic modelling"" is still vague/confusing and this previous comment has not been resolved. Realistic to me suggest that it should reflect reality as much as possible, i.e. a very complex model that incoprporates as much of what we know about the biological processes involved. However, typically in modelling we only want to include the processes that we think that are important for reliable prediction or understanding. If this is also what the authors also mean (the response letter seems to confirm this, we probably agree what it important to include in models and what not), then I suggest to word more precisely what they mean. The term 'realistic' is ambiguous I think., at least for me as a reader.*

We have removed the word "realistic" and changed the sentence to emphasize the ability to validate the model by comparison its predictions with empirical data.

*L69 is it completely new or does it extend on previous models, if so cite and explain how it extends? In other words, it is unclear if the entire model is new (e.g. it is individual based and previous models were analytical) and in addition you added the genetics. Or that you use and existing individual based model and added the genetics. This is linked to L79: please make explicit what part you use from the previous model and what part is new.*

Removed the word "new" from line 69.

We appreciate the concern for making the difference between the model we are presenting and that of reference [22] clear within the manuscript. The model we present resembles Roff's model [22] conceptually, because it served as the starting point for the development of our model. However, our model has little else in common with Roff's model other than being an individual-based simulation model of threshold-trait evolution. All key simulation processes (initialisation, selection, reproduction, mutation, etc..) had to be modified extensively or created anew, to achieve the aims of this work.

Therefore, we modified this part of the text to make it clear that the previous model was also simulation-based, but that most components of the model are new. We hope this makes the relevance our model clearer without pointing out each single differences between the 2

models (which would be overwhelming and confusing for the reader that at this point still doesn't know how the model works).

*L94-99: here I wondered what fitness function was used, as the type of selection is quite crucial. But this is explained later in methods, please make a forward reference to this.*

We introduced a forward reference to the section in which we describe the functions.

*L118-120 I do not understand this sentence, please reword.*

The sentence was ambiguous. We reworded it to improve clarity.

*L131: 1045 units of what? From which studies were these numbers taken?*

We introduced units and added a comment directing the reader to reference [8], for more details.

*L161: can we not see this function in Fig. 1?, if so please refer to Fig. 1 here too. If not, why not in Fig. 1*

We added the reference to figure 1.

*L175 superscript in equation*

Was corrected.

*L198 100 and 300, units?*

We added the units to text.

*L207-213. This text is unclear, but crucial for the paper. Is restlessness a measure of migratory activity, liability or both. I assumed it was a measure of activity, and that liability is an unobservable (latent) variable. However, in Fig. 5 we see empirical measures of liability, suggesting restlessness is a measure of liability. In addition, above lines also suggest that migratory activity = migratory liability in the simulations, so are these concepts identical for non-resident birds? I was getting confused, which may be related to the fact that I am not working on migration, but the text should also be clear for non-experts.*

You are right, this may be confusing, and needed to be made clearer. Working on this subject for long time, sometimes we lose perspective on what can be misunderstood…

Liability is the latent variable underlying the expression of migratory behaviour. All evidence we have indicates that this variable is closely correlated to migratory activity, at least in migratory individuals. During the migratory seasons, migratory birds in captivity show an activity at night, which is called migratory restlessness (or zugunruhe) and which can be quantified in the lab. This migratory activity is correlated with migratory activity (and migration distance) in the wild. Thus, we use the amount of migratory restlessness as a proxy for migratory activity and migration liability in migrants. Moreover, we assume that a higher liability results in higher migratory activity in the wild, which corresponds to longer migration distance.

We have tried to explain this more explicitly and have changed the text to make sure that we use the terms consistently. See also, changes made in response to the comment about L54.

*Fig. 5 has redundancy, panel B shows all information from panel A plus more.*

Indeed, there was a redundancy here. As a consequence, we have removed Figure 5A and made some change in figure 5B, to make it more informative.

*L296. Wording: If I look at Fig. 5 the difference between predictions with 10 and 200 loci is marginal. Also there is a crossover in the fits, whether the decline is stronger or weaker with 10 or 200 loci depends on the timescale.*

We have changed the wording to reflect this.

*L338 At the moment, the model is parameterized with data from the same population, and predictions are compared to the temporal trends in activity. A more independent validation would be to show for another population that the model also predict change in migratory activity well, or across a spatial environmental gradient. This also sounds more productive than waiting for the year 2184 to see if your model has good predictive value. In other words, what future studies could bring/test these ideas further? Now, overall the Dsicussion text reads very much like 'the prove is already here in this MS, and this is how it works'. Some critical notes or ideas for further scrutinizing the proposed ideas would make the discussion more balanced.*

Yes, the model is currently fitted to the only data available on the evolution of migratory behaviour. Clearly, there is a need for more data to test this model and to see whether it predicts genetic variation and evolutionary changes in migratory activity, or other traits, in other species. Unfortunately, we cannot make useful recommendations on how to test the model, since data are simply not available (at least, as far as we know) or are very difficult to obtain. We have, however, proposed some ideas on what to do with the model and which data to collect (see, lines 218-326, 388-402, 411-413).

There are some critical notes at the end of our manuscript, on what may limit the value of the current model (e.g. fluctuating selection, phenotypic plasticity, genetic correlations). We have rewritten parts of the discussion to make this clearer (see lines 402-4010).

*L:351-353: this idea is not unique to traits such as migration behavior, there are many other reasons why strategies may coexists despite unequal fitness payoffs, A more geneal connection to the wider evolutionary literature would be useful.*

Yes, we acknowledge in lines 388-392 that the semi-continuous threshold model is probably applicable to many other traits, so-far considered as dichotomous traits, like dimorphisms, alternative mating strategies or complex diseases. We are aware of the vast literature on alternative strategies, where genetic models have been developed for testing, among other hypotheses, whether equal fitness needs to be assumed for alternative strategies to be maintained (see, for instance, review by Tomkins & Hazel 2007, doi: 10.1016/j.tree.2007.09.002 or Buzatto et al. 2014, doi: 10.1098/rstb.2013.0625 and Buzatto et al. 2015; doi: 10.1098/rspb.2015.2075 ). We think, however, that making inferences about the evolution of other dichotomous traits is too speculative here, and off the scope of this manuscript. We believe, however, that our conclusions are important for modelling and

understanding the prevalence of partial migration. In the text we certainly did not mean to say that the genetics and expression of the trait is the only explanation for the prevalence of partial migration.

We have made a few changes in this paragraph and added a reference to make this clearer (see, lines 334-342).

*Reviewer 1*

*The authors did an excellent job revising the paper, such that the methods are inline with standard quantitative genetic approaches. These revisions aid the robustness of the paper, the ability to compare with existing literature and allow for a useful baseline for future research. My comments at this point are very minor.*

*- line 9 & elsewhere: Suggest replacing "We show . . ." with "Our model indicates . . ." All of the results arise from a model and my opinion is that the statement "We show . . . is the compatible with . . ." is too strong of a statement. (Note, this recommendation is based on me making general and strong statements in some of my papers, without the important qualifier that the results arise from a particular model, with particular assumptions, with some statements having come back to be shown to be too general and too strong.)*

We have changed the wording throughout the text following the recommendation of the reviewer.

*- lines 23-25: Is there a citation for this statement?*

References 1 and 2 support the first 3 sentences of the introduction and so are given at the end of the third sentence.

*- lines 31: Is the state of the field such that we can say "most" migratory species show little flexiblity? With the use of geolocators and stable isotopes, I think we are better appreciating that an individual's migratory behaviour can respond to environmental conditions and their state (such as health) [Hegemann et al. 2015, Am. Nat.]. Sometimes this leads to a lack of a complete migration within*

*a year.*

You are absolutely right on this point, as more and more within- and among-individual variation is found in studies using remote-sensing or other tracking techniques. However, these studies show that there is individual variation in migration, but it is not clear to what extent this variation is adaptive and/or if this variation can be transmitted horizontally (by cultural transmission, i.e. learning or imitation) or vertically (by genetic or epigenetic transmission). In our view, most adaptive changes of migratory behaviour in migratory bird populations (with the exception of species with cultural transmission of migration) will require evolutionary change. We have changed the sentence (lines 40-43) to make clear that we are referring here to the <u>"flexibility" of populations to adjust migration to changing conditions over time</u>.

*- line 69: Is the use of the word "new" needed?*

We removed the word "new".

*- line 80: Suggest: "Individuals are diploid at each of n loci."*

We modified the text as suggested.

*- line 82: Suggest: "The sum of . . . , such that alleles are co-dominant."*

We modified the text as suggested.

*- line 87: "0" -> "no"*

Modified as suggested.

*- line 99: Add a "see below" otherwise a reader may think the explanation of selection is incomplete.*

We introduced a forward reference.

*- line 113-114: Note that in some cases mutational bias is appropriate. For example, a phenotype may be near its extreme of possible values, such that mutation is biased in the opposite direction. Instead of indicating an "advantage" merely state the difference.*

We changed the sentence, accordingly.

*line 120: I think "above" should be "below".*

Corrected to "below".

*line 125: "environmental effects" -> "variation in the environment"; (To not confuse the reader with a genetic "environmental effect".*
Modified as suggested.

*lines 175, 177: subscript the "m"*
Corrected.

*lines 226, 228: It would have been clarifying to add the word "differential" before "selection".*
Modified as suggested.

*line 230: "some" -> "a few"*
Modified as suggested.

*line 265: "neither" -> "not"*
Modified as suggested.

*line 346, 396: "will" -> "is expected to"*
Modified as suggested.

*Supplement:*
*- p.2, 1st paragraph: Point reader to μp and definitions and values table to get them started about where to refer to definitions and values.*

We added a reference to table S1

-