

Dear Editor,

We thank you and the reviewer for constructive comments. We have addressed all remaining comments with a focus on clarifying the forward model and the limit of such approach.

We hope our reply will address all the small remaining issues. The correction in the text are highlighted in yellow. A few typos (missing commas, extra spaces, grammar) were corrected directly.

We also apologise for the delay,

Best Regards,

Quentin Rougemont, On the behalf of all the authors.

Round #2

revision required

This revision is a great improvement and I agree with the reviewer that it addresses many of the previous concerns. The Dadi reconstruction showing increase in Ne in polymorphic numata is a nice addition.

The simulations are also a welcome addition, but need better description and presentation. In Figure 4 the legend is incomplete - I assume the colours represent the migration parameter but this is not stated. Why not put assortative - disassortative mating on a single panel as this is a continuum.

Reply: Indeed, colours represent migration. As advised, we've merged assortative mating and disassortative mating runs in a single continuum with random mating in the centre, and we have clarified the methods.

It would also be helpful to outline the justification for the modelling of local adaptation with gene flow - is that based on observed spatial heterogeneity in mimetic models in *H. numata*?

Reply: We used a stepping stone model to represent the classic spatial mosaic of selection as observed in the models of *H. numata*, and generally in mimicry variation in *Heliconius* species across their range. We have now added a justification in the SLiM methods line 274 – 318.

I also agree with the reviewer that random mating should be the null hypothesis - there isn't any *Heliconius* where assortative mating has been demonstrated between sympatric morphs as far as I am aware, so not clear that the null hypothesis should be assortative mating.

Reply: Indeed, random mating is a null hypothesis and it is now better presented in the results of Figure 4. Assortative mating between divergently-coloured geographic variants ("races") is reported in *H. melpomene* (Jiggins et al. 2004 J Evol Biol) or in *H. timareta* (Sanchez et al 2015 Evolution), so comparing with assortative mating has a rationale, when trying to give a simulation framework to compare to empirical observations. Both are now clearly visible on figure 4 (see whole method paragraph on Forwards simulations as well (line 274 to 318)). Few polymorphic populations have been studied in *Heliconius*, besides *H. numata*, but partial assortative mating based on colour differences was found in *H. cydno alithea* (Chamberlain et al. 2009 Science). We chose to model a situation with two distinct forms benefitting from mimicry with local co-mimics (one enjoying high protection, the other with intermediate protection), and only vary the mating systems. While disassortative mating may lead to the maintenance of polymorphism, as in *H. numata*, assortative mating tends to eliminate the intermediate form, and in that case simulations are generally run in a situation of local monomorphism, mimicking the classic situation in most other *Heliconius* species.

From the figure, it would seem that disassortative mating does not make much difference to diversity as compared to random mating, it is rather assortative mating that reduces genomic variation? The authors need to clarify the interpretation here.

Reply: This is exactly what we observed, an effect of assortative mating reducing diversity rather than disassortative mating *per se*. We now provide a clearer interpretation line 475 – 478: “*These hypotheses are also supported by our forward simulation which suggests that revealed an effect of the mating system. Yet, most of the differences in genetic diversity was due to assortative mating reducing genetic diversity rather than a diversifying effect of disassortative mating.*”

Nonetheless, the simulations do show that mating patterns at a single locus can significantly affect linked diversity, so provide some support for the overall hypothesis. I also found myself wondering whether unlinked diversity would be similarly influenced - as only a 1MB chromosome was simulated. The authors could perhaps address this in the discussion.

Reply: Indeed, we now discuss this line 496-500: “*Another limit of our model is that we tested the effect on genetic diversity on a 1 Mb segment of a single chromosome, not on unlinked chromosomes. Although testing the effect on unlinked data would be relevant, it would substantially increase the compute time so we choose to focus on this simpler model and leave the question of large scale effect for future investigations.*”

I therefore request the authors to address these concerns and those listed in the review. I do not anticipate that any further review is necessary.

Reply: Thank you.

Below is the new Figure 4 and its legend as it will be referred to several time in the review:

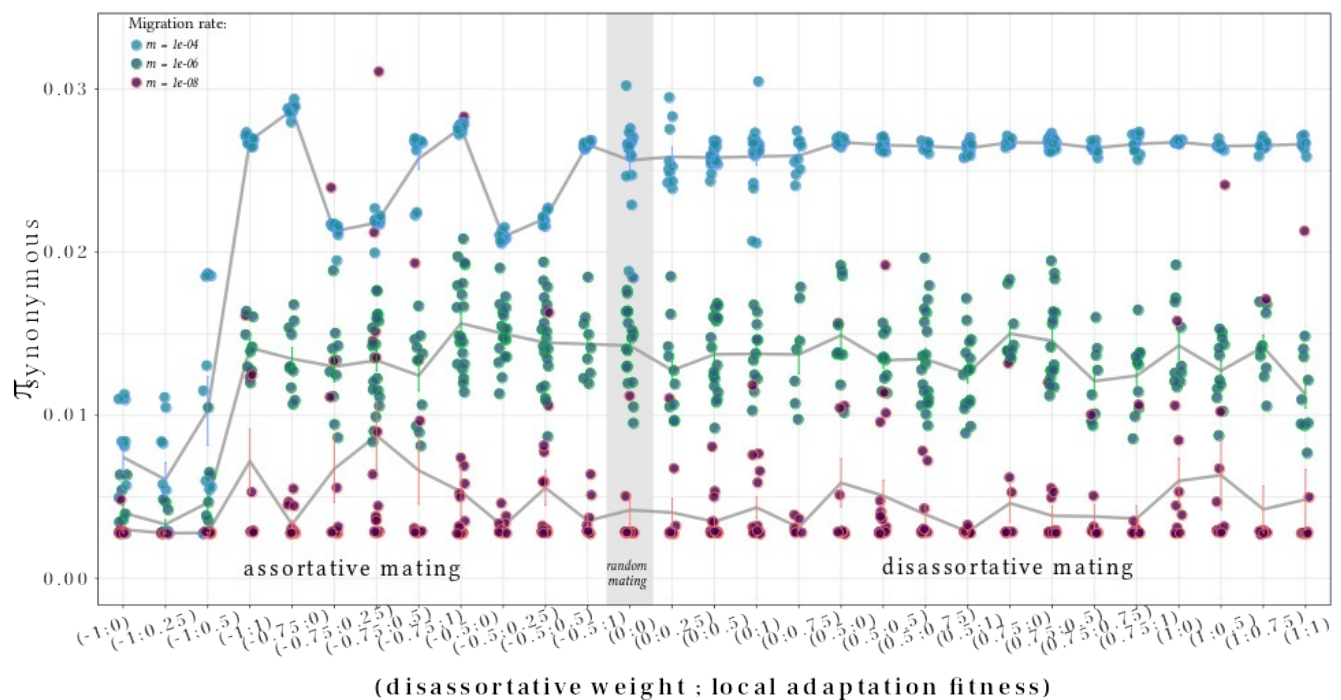


Figure 4 | Weak but significant differences in synonymous nucleotide diversity (π_s) emerged at a genome-wide scale under divergent selection and mating regime. Results from forward simulations of 10 replicate per parameter combination. A set of 10 demes with varying levels of migration, local adaptation fitness and different mating strategy (from assortative mating to disassortative mating) are presented. Shown are levels of synonymous diversity obtained under each combination of parameters. Parameters on the left part of the brackets display the disassortative mating weight (from -1 to 1). Parameter on the right side of the brackets displays the fitness value for local adaptation. A left value of -1 in the bracket means complete assortative mating and 0 means random mating. 1 = complete disassortative mating. A right value of 0 in the bracket means a fitness of 0 for non locally adapted individuals in a deme. A value of 0.5 means a reduced fitness of 0.5 relative to the maximum value. A value of 1 means no loss of fitness.

Reviews

Reviewed by Christelle Fraïsse, 21 Dec 2022 17:22

I am delighted with this new version and the comprehensive reply letter. The author made substantial efforts to address my and other reviewers' concerns, and they performed additional analyses:

- 1) perform two demographic analyses to estimate the timing of N_e changes explicitly.
- 2) perform extensive individual-based forward simulations to confront their observations with predictions under different models.

I also liked very much that the authors were more cautious in their interpretation. They made the necessary changes in the abstract, results and discussion.

I still have a few suggestions below to handle before publication; they are minor. Therefore, I fully support the release of their great study in PCI.

Note that line numbers refer to the "tracked changes" version.

Forward simulations:

- Material and methods:

- L288 and following: Justify in biological grounds the setting of your simulations (stepping stone model, number of morphs = 5, 2 or 3 morphs have a fitness advantage per population). Is this setting in agreement with what is known from the system? How are the morphs initially distributed across the demes?

Reply: Please see the new methods (line 274 – 318) where we attempt to provide biological justification. Stepping stone is chosen to reflect isolation by distance, we choose a number of morphs to reflect the number of morphs known to exist in a single locality (between 2 to 7). Initially, in each demes 3 morphs have a fitness advantage. The morphs were distributed randomly.

- L309 – 327: your parametrization of the strength of assortative/disassortative mating is confusing. I would rephrase it in terms of the probability of mating with a different morph, where a value of 0 = assortative mating, of 1 = disassortative mating and 0.5 = random mating, as you did in Figure 4. It will be more intuitive to interpret.

Reply: This is what we have done now (see changes in methods line 247 – 318).

- Moreover, I need clarification on why you tested only values from 0 to 0.5 (in your parametrization) in both the dis-assortative and assortative mating. A value of 1 is omitted, while this corresponds to a "random mating model" (as far as I understand). So in the text, you say that you do not test for "random mating" (which is the null model against which the disassortative model should be tested against). But then, in Figure 4 it is indicated that you tested a case "equivalent to random mating". Sorry for the confusion. Could you clarify please?

Reply: We have added the results for 1 (random mating that were not presented clearly). The method should be clear now and Figure 4 is also updated.

- Results (from L396 to 404):

- I am not sure to understand why you highlighted the difference between the case of assortative mating vs disassortative mating (average π_S of 0.011 vs 0.0145). Is assortative mating the common reproduction mode when the supergene is absent in these butterflies? In addition to this comparison, could you please briefly describe the difference with the simulations under random mating?

Reply: Yes, the idea was to contrast disassortative mating to the common case with assortative mating. See our clarifications in the methods.

- You could put the weak difference obtained in the simulations in perspective with the observed differences (<0.01 for the Brazilian *H. numata* vs >0.025 for the Amazonian *H. numata*). This hints at the minor contribution of assortative mating to the neutral genetic diversity compared to other factors,

such as the migration rate between demes or founder events. This can be highlighted, for example, along lines 490 in the Discussion.

Reply: Excellent, we have added this now line 480: “Overall, the strength of change in genetic diversity due to the mating system is weak compared to migration. Indeed, our simulation results indicated that migration between demes has the stronger effect on synonymous diversity. In addition, the magnitude of difference in the simulation is weak in comparison to the observed differences of genetic diversity between *H. numata* from Brazil and *H. numata* from Peru. Whether these differences between our simulations and those empirical observations can be directly compared is not straightforward however given the simplified assumption required by our modelling approaches. In all cases, our demographic simulations with $\partial a \partial i$, do suggest a strong effect of founder events/bottleneck in reducing the genetic diversity of the Brazilian population.

- Figure S5: It is unclear which simulations were pooled under the label “assortative mating” and under “disassortative mating”. All parameter combinations? It would be clearer to color the points according to the strength of the assortative (or disassortative) mating.

Reply: All Simulations were pooled by mating type to test an average difference. However, this graph does not help a lot since it would be better to decompose the result by migration rate and strength of local adaptation as in the whole figure (Figure 4). Therefore we removed this figure to reduce the reader's confusion.

Minor suggestions to handle before publication:

- L188: “(I)” → “(i)”. **Reply:** corrected.
- L227: typo in “at which LD decay[s]”. **Reply:** corrected.
- L235: please, add “i.e.,” before “migration rate times the duration of the migration band”.

Reply: Done.

- L276: please, add after “controlled by parameters s_1 and s_2 ”, the new time parameters (T_{p1} and T_{p2}), which indicate the time of exponential change. **Reply:** Done.
- L285: “and only models with the lowest” → do you mean the “run” with lowest... ?

Reply: indeed, this is the run. We have corrected this now.

- L302: typo in “with a given [set of] deleterious recessive mutation[s], generating overdominance at [these] loci”. **Reply:** corrected.
- L306:
 - remove “reduction” in “fitness reduction varied”.
 - “fitness of zero for migrants in a deme” → “null fitness for non-locally adapted individuals”.
 - remove the “s” in deme[]. **Reply:** Corrected.
- L311:
 - add after “complete disassortative mating”, a short explanation of what it means, such as “where a given individual mates only with different morphs”.
 - “no mating weight” is unclear; please, consider replacing with “random mating”. **Reply:** Corrected.
- L325: please, remove “no disassortative mating but”. **Reply:** removed.
- L326: please, remove “(or no assortative mating)”. **Reply:** removed
- L349: typo in “other Amazonian population[s]”. **Reply:** Corrected.

● L390: you could also mention the inconsistency of the point estimate of N_e for the peruvian *H. numata* population in the two inferences (with *H. pardalinus* or with brazilian *H. numata*). However, they both show a large N_e for the peruvian *H. numata*.

Reply: Indeed. This was not clear and is now written as follows: *“There was a large variance in estimates of effective population size in H. numata from Peru in the two inferences (with H. pardalinus or with Brazilian H. numata). Yet, in these two inferences, the effective sizes of the Peruvian H. numata were large (greater than 11 million).”*

● L442: I would replace “coincide” with “coincide in time”. **Reply:** Done.

● L479: I would replace “translate into an effect on effective migration genomewide” with “translate into a reduction of the effective migration genomewide”. **Reply:** indeed. This is done.

● L489: typo in “results also suggest[]”. **Reply:** Corrected.

● L722: typo in “a[d]mixture”. **Reply:** Corrected.

● L730: a bracket is missing between “numata” and “populations”. **Reply:** Corrected.

● Figure 4: - title: “and mating region” should be removed or rephrased, because it is confusing as it states. **Reply:** it is replace by mating regime

- I am not sure to understand what the left number in brackets means. Is 0.5 random mating? Because you state it is neither “assortative mating”, nor “assortative mating”.

- In the same line, I do not understand the difference between the set of parameters $n^{°7-9}$ in panel A vs $n^{°1-3}$ in panel B.

- Why the right number in brackets is increasing in panel A, while it is decreasing in panel B?

Reply: See the change now.

- Could you please indicate the meaning of the different colors? **Reply:** These were migration rates. This is in the legend now.

- L761: “10 populations” → do you mean 10 replicates? **Reply:** 10 demes and 10 replicate.

- L763: please, add “strength of” before “local adaptation”. **Reply:** Done.

- L767: typo in “in the brack[et] mean[s]”. **Reply:** Corrected.

- L768: typo in “in a deme[]”. **Reply:** Corrected.

Below, line numbers from the biorxiv version.

● Table 1:

- L715: typo in “Biological parameter assum[]ed”. **Reply:** corrected.

- L716: I would remove “, and descending populations 1 and 2”. **Reply:** indeed...

- L725: “for” → “four”. **Reply:** corrected

● Figure S1: typo in “*H. numata* individual[s]”, “microsatel[l]ites”, “20 STRUCTURE run[s]”. **Reply:** corrected

● Figure S2: label the x axis of the left bottom residual plot. **Reply:** done

● Table S7: a note beside the table could indicate that the optimization of the logL was sometimes unsuccessful as some nested models have a higher likelihood than their general version (for example, SI2NG is much less likely than SI2N, while it should be at least as likely because it is the same model but with an extra parameter).

Reply: Indeed, We have added this: *“The optimization of the Log Likelihood was sometimes unsuccessful. Indeed some nested models have a higher likelihood than their less parameterized version.”*

More specifically, the SI2NG which contains an extra parameter should be at least as likely as the SI2N but received a lower likelihood in the H. numata Peru versus H. numata from Brazil comparison”