Review of Manuscript GENETICS-2024-307195 by Sakamoto *et al.* "The persistence of locally adapted polymorphisms under mutation swamping"

Simon Aeschbacher

20 August 2024

Summary

Sakamoto and co-authors use analytical theory and simulations to investigate the fate of a focal, locally beneficial allele of large effect as a function of the overall mutation rate at an infinitesimal number of small-effect loci in the background. All loci contribute to a quantitative trait under spatially divergent stabilising selection. The authors' undertaking is of high interest because it contributes to our understanding of why and when fundamentally different types of genetic architectures of quantitative traits evolve under divergent selection with gene flow. Methodologically, the authors make a significant contribution to linking population genetic with quantitative genetic models in the context of a trait under locally divergent selection for different optima. This endeavour is challenging because of the different natures of the respective models and the disjoint, if not opposing, sets of assumptions about the respective sets of genetic loci. The authors tackle this challenge by decoupling the two models as much as possible at the genetic level (there is free recombination between the focal locus and the background loci), while keeping them connected through the joint dependence on the equilibrium mean trait of the focal allele frequency and the mean genotypic value of the background regions. This is a clever approach that yields results that increase our understanding of how different evolutionary forces interact to cause and modulate the evolution and maintenance of local adaptation. I am very positive about the manuscript and have no doubt that the overall quality and scope of the analyses and results justify its publication in GENETICS. The paper is well written. I liked how the authors embed the study into the context of the existing literature. However, I have a few major concerns regarding the authors' agenda of promoting "mutation swamping" as a description of what they found, the robustness of the results to the assumption of linkage equilibrium, and the metric the authors use to classify their simulations into those leading to a stable vs. transient architecture. I also had a substantial number of Minor Comments regarding language and I think the authors should describe their simulations in more detail. Overall, I am positive that these minor issues can be fixed fast. The major issues will take some effort to fix and/or address, but I am confident that the manuscript can ultimately be published in GENETICS after a thorough revision.

Major Comments

1. I am concerned about the coining and use of the term "mutation swamping", which the authors justify based on a supposed analogy to "migration swamping". I am afraid this analogy is superficial at least, if not conceptually wrong and thus misleading. I argue against using "mutation swamping" in the manuscript for the following reasons:

1. I agree that the replacement of what the authors call a "stable divergent [genetic] architecture" by a "transient architecture" at high background mutation rates is reminiscent of what Yeaman and Whitlock (2011) observed at intermediate migration rates under divergent selection for different optima. However, as far as I can see, "swamp"-ing by high levels of gene flow was used in Yeaman and Whitlock (2011) to describe the *loss of divergence at the weakly selected alleles*, while *divergence is maintained at one or a few strong loci*. In contrast, the authors of the current manuscript are concerned with the *loss* of divergence at the *focal* locus in favour of what they call a "transient" architecture under which divergence is maintained by ephemeral weak-effect alleles at many background loci: "However, as the mutation rate increases, divergence at large effect loci becomes unstable, and the architecture transitions to a transient one driven by fleeting divergence at many small effect loci dispersed across the genome. We termed this collapse of the stable divergent architecture 'mutation swamping', with an analogy to the classic migration swamping (Haldane 1930; Wright 1931; Felsenstein 1976; Lenormand 2002)." Thus, the analogy between migration swamping *sensu* Yeaman and Whitlock (2011) and "mutation swamping" as used in the current manuscript remains fuzzy at the level of an evolutionary pattern.
2. Related to the previous argument, but more fundamentally, the term "mutation swamping" in analogy to migration swamping is unjustified in the current context at the genetic and mechanistic level. In the classic literature also cited by the authors (Haldane, 1930; Wright, 1931; Felsenstein, 1976), migration swamping describes the removal of a locally beneficial allele at a focal locus by an immigrant allele at the same locus coming in through gene flow via migration. This is both a *directed* process – migration is directional and running counter to selection on average unless in special configuration – as well as a *direct* process in two ways: it is direct in the sense that migration is the force immediately acting on and causing the loss of the locally beneficial allele, and direct in the sense that the process acts on the focal locus. In contrast, what the authors term "mutation swamping" neither is the result of a directed process nor does it act directly on the focal locus where divergence is lost. Instead, mutation as modelled in the quantitative framework that the authors use for the background loci is not directed: effect sizes follow a truncated Gaussian distribution. Moreover, the mutational process that leads to what the authors term "mutation swamping" (i.e., the loss of divergence at the focal locus) does not act directly on the focal locus – there *is* no mutation at this locus. From the perspective of the focal locus, mutation at the background loci merely has a *modulating* effect on the joint action of selection and migration at the focal locus. In my opinion, there is just too much conceptual incongruence for there to be any justification in using the term "mutation swamping". In this context, the authors also imprecisely cite Höllinger et al. (2019, 2023) on lines 16 to 19: Höllinger et al. did not show that mutation rate *drives* selective sweeps nor small allele-frequency shifts at many loci. These authors only showed that mutation rate *modulates* the way by which polygenic adaptation occurs. The *driving* force is and remains selection, both in the sweep-like and in the shift-like scenarios of oligo- or polygenetic adaptation. In my view, this imprecise citation reflects the conflation of direct and indirect effects of mutation that the authors seem to be willing to accept in using "mutation swamping" in the given context. I think this conflation is misleading and must be avoided.
3. My concern was reinforced when I read lines 43 to 53 of the current manuscript: "The range of migration rates within which mutation swamping occurs shifts toward lower values as the effect size at the focal locus decreases (Figure 4D). This pattern suggests that *mutation swamping will also re-shape the architecture of adaptation through homogenizing small*

effect alleles while preserving divergence at large effect loci. This effect is analogous to the evolution of a concentrated architecture by migration swamping [emphasis added] (Yeaman and Whitlock 2011; Yeaman 2015; Rafajlović et al. 2016) [...]". The authors introduce and repeatedly refer to "mutation swamping" as the *loss* of divergence at the focal locus in favour of a transient architecture involving many background alleles. Yet, here the authors seem to imply that "mutation swamping" could lead to exactly the opposite, i.e., an architecture with a large-effect locus at which there is divergence, and no divergence at the background loci. This passage thus read to me as if it inverted the authors' own definition of "mutation swamping" (see above), and so would also break with the analogy to what happens at the focal locus under migration swamping under classic theory (e.g., the works by Haldane, 1930; Wright, 1931; Felsenstein, 1976 cited earlier by the authors). I thus felt the argument became not only inconsistent, but also confusing.

2. For the sake of analytical tractability, the authors assume free recombination between the focal locus and the background loci. I think this assumption is justified as far as the analytical work goes for the current manuscript. However, the assumption could have non-negligible consequences on the generality of the results. The work by Yeaman and Whitlock (2011), follow-up papers by Yeaman and co-authors, as well as work by others collectively have shown that physical linkage between a major locally divergent locus and nearby minor-effect loci has a qualitative effect on the dynamics and outcome of evolution at the genetic level. Such linkage therefore also affects the extent to which alternative genetic architectures of polygenic traits under spatially divergent selection evolve and are maintained. In fact, the authors report that in some of their simulations, "a large effect locus arises in the background region by tight linkage of small alleles and takes over the role of the cluster" (p. 6, l. 4–6). Although it remains unclear what proportion of simulations resulted in this outcome, it is clear that linkage is of qualitative importance. Given that the focal locus starts out from perfect adaptation in the simulations, whatever the proportion of simulations of interest was, it would be a lower estimate of the propensity of clustering to evolve around the focal locus. Such clustering is expected to increase the persistence of the concentrated architecture around the focal locus, and thus shift the odds in favour of the concentrated architecture. For this reason, I think the authors should aim to report simulation results that include some degree of linkage between the focal and the background loci. If this is too cumbersome, I think the authors should at least more thoroughly discuss the limitations and implications of assuming free recombination between the focal and the background loci. I appreciate that the authors discuss the assumption of linkage equilibrium among the background loci in the Discussion (p. 9, l. 24–51). Here, I refer to linkage between the focal and the background loci.
3. My next concern relates to those simulations in which a large-effect locus arose in the background through tight linkage (p. 6, l. 4–6; see previous Major Comment). In these simulations, the initial divergence at the focal locus disappeared and was replaced by a concentrated architecture in the background. The purpose of the simulations was to validate the authors' analytical prediction of a threshold mutation rate beyond which the point-concentrated genetic architecture at the focal locus is replaced by a transient architecture constituted by many weak-effect background loci. The analytical theory assumed free recombination among the background loci, while the simulations assumed weakly linked background loci (Poisson-distributed number of recombination events across the background region with mean 10; p. 3, l. 7–11). The authors seem to first have explored the derived allele frequency at the focal locus as an indicator to distinguish between simulations in which the initial concentrated architecture remained as opposed to those in which it was replaced by a transient architecture. But then the authors argue that because of the simulations that resulted in a concentrated architecture in the background, this natural indicator was not "robust"

enough (p. 6, l.6–7). Instead, the authors developed an indicator T^* that essentially assigns these simulations to the category of stable concentrated architectures. However, this bookkeeping is conceptually wrong as it inflates the proportion of stable concentrated architectures. The theory predicts when divergence at the focal locus collapses, it has no bearing on the evolution of clustered architectures in the background because it assumes free recombination. The authors do not report the proportion of simulations that resulted in a concentrated architecture in the background region (I think they actually should), and so I cannot judge how severe the inflation is. In any case, however, if the authors want to treat the simulations of concern separately, they should *exclude* them from the analysis, not assign them to the 'wrong' category. This inconsistency needs to be fixed and the Results text and figures updated accordingly.

Minor Comments

C: comment; **Q:** question; **S:** suggestion; **R:** request.

Title

- **C:** The term persistence sounds confusing (contradictory) in relation to the high-mutation-rate regime that 'mutation swamping' implies. Essentially, the paper shows that locally adapted polymorphisms tend *not* to persist under high mutation rates. ***S:**: Replace by 'fate'.
- **C/R:** As explained above (Major Comment 1), I find the term "mutation swamping" misleading und not justified. I request its omission from the paper and the title. ***S:**: Replace by "high mutation rates". In combination with my previous comment, I suggest the title be modified to "The fate of locally adapted polymorphisms under high mutation rates". **R:** Please also adjust the running title "Mutation swamping in local adaptation" so that it does not contain "mutation swamping".

Abstract

- [l.2–3] **C:** The two descriptors "*pronounced divergence* [at a few loci]" and "*small allele frequency shifts* [at many loci]" do in my view not not describe two genetic architectures, but a mechanistic/procedural aspect not typically considered to be part of the genetic architecture, but in this case causally linked to the coming-into-place of architectures. **S:** Insert "and evolutionary histories" after "architectures" to solve the problem.
- [l.3] **R:** Insert "the" before "migration rate".
- [l.4] **C/S:** I find "stable contributions" a fuzzy term and suggest "lasting" instead.
- [l.6] **R:** As explained above (Major Comment 1), I request that the authors omit the term "mutation swamping" because it is an inappropriate analogy to "migration swamping" in the context of this paper. **S:** Replace "... 'mutation swamping' phenomenon depends upon ..." by "... transition in genetic architecture depends on ...".
- [l.7] **S:** Replace "determining mutation effects" by "describing the mutation process". **R:** Please use present tense for "develop" to clarify that the theory is developed in your paper.
- [l.7–9] **C/S:** I find this sentence too long and complicated. I suggest to split after "occurs" and to then start with "We find that ...".
- [l.8] **R:** Replace "swamping" by "the transition". Insert "a" before "wide range" and "the" before "parameter space".
- [l.9] **R:** Replace "mutation swamping threshold" by "threshold mutation rate".
- [l.10] **S:** Omit "and" before "weakly"; replace "but is only minimally affected by" by "and marginally on the population size" to streamline the phrasing. **R:** Insert "the" before "parameter

space".

- [l.11] **R:** Replace "mutation swamping" by "the transition to a transient architecture".

Introduction

Page 1

- [l.46] **C/S:** The term "linked selection" has become widely used, but this does not change my opinion that it is imprecise. I suggest to replace it by "selection at linked sites". #### Page 2
- [l.2, l.6, and other places] **C/S:** To increase clarity, I suggest to replace "small effect" and "large effect" by "small-effect" and "large-effect" here and throughout the paper whenever these pairs of words occur as compound adjectives.
- [l.5–7] **R:** Please back this sentence with (a) reference(s) or rephrase it to be more tentative.
- [l.9] **S:** Omit "rate" after "mutation" or insert "the" before "mutation rate".
- [l.17] **S:** Replace "maintain divergence" by "persist".
- [l.18–20] **C:** I missed an indication after "causal alleles" of the number of loci involved – more than in the case of the constant architecture, right? **S:** Insert "at many loci" after "causal alleles". **C/S:** Replace "balance" by "under the joint action" because "balance" is problematic if understood as applying to the level of individual loci.
- [l.26–31] **S:** Replace "transition into" by "become". **C/S:** I find "as various parameters change" a bit bold given the moderate number of parameters explored. I suggest to write "... threshold change as a function of the effective population size, selection strength, and the effect size of alleles under selection, and (iii) why ...". To me, this phrasing is more accurate and thus gives a better idea of the scope of the study. I suggest to replace the commas separating the three objectives by semi colons, just to make things more clear.
- [l.48] **S:** Replace "... consider their interaction, we ..." by "... consider the interaction of these two types of loci, we ..." to avoid ambiguity.
- [l.50–54] **R:** Correct "combines" to "combined" and replace "one locus two allele model" by "one-locus two-allele model" to increase clarity.
- [l.64] **C:** "Show" would do instead of "demonstrate", I think.
- [l.69–72] **C/R:** As explained in Major Comment 1, the similarity of the process studied here is to "migration swamping" is too superficial to justify an analogy and I therefore strongly argue against coining the term "mutation swamping" in this context. I request that this sentence be omitted.
- [l.72–76] **S:** For consistency, I suggest the authors stick to present tense, and so replace "compared", "confirmed", and "investigated" by "compare", "confirm", and "investigate", respectively.

Model

Page 2

- [l.80] **S:** Insert a hyphen between "discrete" and "generation" (compound adjective).
- [l.85 and other places] **C/S:** The colons before display equations are unnecessary and, as far as I can see, not standard in GENETICS. I suggest these colons be removed everywhere except where language makes them necessary. As far as I can judge, the colons would thus need to be removed before the following equations: equation (1), the equation showing the truncated normal distribution on p. 3, equations (2) to (5), equations (7) to (10), equation (A1), equations (A4) to (A6), equation (B1), equation (B3), and equations (C1) to (C4).

- [l.94–96] **C:** It is currently unclear if "based on the probability weighted by fitness" only refers to the parents chosen from population 2 or also the immigrants from population 1. **R:** Please rephrase to clarify.
- [l.102] **R:** Replace "Phenotype is determined . . . " by "The phenotype is determined . . . ".

Page 3

- [l.2] **R:** Please briefly motivate the constant C. I suppose it corrects for the probability mass outside the truncation limits.
- [l.9–11] **R:** This clause needs to be rephrased to be more precise. Brazier and Glémin (2022) report a range of 5 to 6 *per chromosome*, if I am not mistaken. On the other hand, Kong et al. (2002) report an average recombination rate of approx. 1 cM/Mb, which, for the human genome, amounts to a genetic (map) length of 3000 cM, which corresponds to 30 Morgans or recombination events *per genome*. Surely, 10 recombination events in the background region are in the ballpark of 5 to 6 and 30 as numbers, but the question is what proportion of a chromosome or genome the 1000 simulated background loci would typically represent. Or, in other words, how many background loci a typical chromosome or genome would contain. The authors need to expand here to justify their choice. See also my Major Comment 3.
- [l.22] **Q:** Should "registered" read "recorded"?

Results

Page 3

- [l.28–29] **S:** Replace "patterns" by "types".
- [l.54–56] **C:** The definition of a concentrated architecture given here is inconsistent with what the authors later seem to also assign to the category of a concentrated architecture: a group of tightly linked loci with persistently divergent allele frequencies in the background (e.g., l. 8–32; see also my Major Comment 3). My understanding of a concentrated architecture also includes this latter case (see also Fig. 2B). I understand that the assumption of free recombination between the focal locus and the background loci might have contributed to the apparently overly restrictive definition of a concentrated architecture here. **R:** I think the authors should use a more generic definition of "concentrated architecture", and then state what constraints apply to the ability of their theory and simulations to model and capture a concentrated architecture.

Page 4

- [l.11] **R:** Insert "the" before "mean genotypic value".
- [l.16] **S:** Replace "works efficiently and the linkage effect can be ignored" by "is high and linkage among the background loci can be ignored".
- [l.23] **S:** Replace "required" by "necessary" to conform with what I think is the conventional phrasing.
- [l.28–29] **C:** I found this sentence a bit unfortunate, because the equation itself does not explain why the concentrated architecture does not evolve at high mutation rate. However, I agree that it can be used to explain why the concentrated architecture at the focal locus may not be maintained. **S:** Replace "Equation 4 explains why the concentrated architecture does not evolve at high mutation rate. This equation shows . . . " by "Based on Equation 4, one can explain why the concentrated architecture at the focal locus is not maintained at high mutation rate. The equation shows . . . ".

- [l.38] **C/S:** The pronoun "its" dangles. To fix this, I suggest to replace "... of its divergence" by "... of divergence at the focal locus".

Page 5

- [l.2] **S:** I suggest to omit "disruptive selection arising from" and just keep "stabilizing selection" to avoid confusion.
- [l.5–6] **C/S:** The pronoun "its" dangles. I suggest to write "..., divergence at the focal locus remains table" instead of "..., its divergence is stably maintained".
- [l.7] **C/S:** I doubt that *no* stable equilibrium exists if m is higher than the critical value given, but I suppose no stable *internal* (or: polymorphic) equilibrium exists. Please rephrase to fix.
- [l.13] **S:** Insert "at the background loci" after "mutation rate" to increase clarity. **S:** Replace "at" by "in".
- [l.14–15] **C/S:** The phrase "and the phenotypic value becomes closer to its optimum" seems to be a somewhat inaccurate shortcut that mislead to the conclusion that mutation directly moves the trait mean to the optimum. However, the immediate consequence of mutation only is that selection becomes more effective. It is only selection, not mutation, that leads to adaptation at the background loci. I suggest that the authors rephrase this to be more accurate.
- [l.16–17] **S:** Replace "... the threshold of the mutation rate through ..." by "... the threshold mutation rate at which the focal locus becomes unstable by ..." to increase clarity.
- [l.20] **S:** I suggest to remind the reader of what $\bar{z}_2$ is inserting ", the mean genotypic value at the background loci in population 2" at the end of the sentence.
- [l.29] **S:** Replace "may be" by "is".
- [l.31–33] **C/S:** In my understanding, disruptive selection acts at each site, not at each allele. A given allele is always under directional selection, but the direction depends on its frequency (underdominance). This is why I think the term disruptive selection should be applied to the site, not the allele. I would thus write "acts at each background locus" instead of "works on each mutant allele". **S:** In line with the previous comment and to increase clarity, I would extend "preventing its increase" to "preventing the increase in frequency of novel mutations as they are rare".
- [l.41–43] **C/R:** As far as I can judge, the statement " " is only true if you make a restrictive assumption about γ (or, in the particular case, σ) with respect to m. I have not seen this assumption being made explicit in this manuscript. Please state this assumption at some point.
- [l.83] **S:** I suggest that the authors insert a reference to the simulation code at the end of this sentence. Currently, the code is linked only from section "Data Availability".

Page 6

- [l.3–6] **Q:** Do the simulation runs in which a concentrated architecture arises in the background region occur above or/and below the mutation threshold? And what proportion of simulations resulted in such background clusters?
- [l.8–32] **C:** As explained in Major Comment 3, the indicator introduced here is chosen so that it assigns simulations under which a concentrated architecture evolves in the background region to the category of the architecture concentrated at the focal locus. This is conceptually incorrect. It would be better to attribute these cases to a violation of the assumption of free recombination among the background loci, and to thus exclude these simulations from the bookkeeping. Currently, the bookkeeping inflates the proportion of runs that maintain the concentrated architecture at the focal locus.

- [l.19] **S:** Replace "..., which is registered ..." by "..., which we recorded ...".
- [l.83–87] **R:** These two sentences are partially redundant; please merge to a single sentence and streamline.

Page 7

- [l.2–3] **R:** Replace "... are the deterministic ones, ..." by "... are deterministic, ...".
- [l.5] **R:** Replace "variance" by "standard deviation".
- [l.36] **S:** Replace "Of note, ..." by "However, ...".

Discussion

Page 7

- [l.65] **C/S:** I stumbled over the phrase "phenotypic divergence underlying local adaptation" – it just felt a bit awkward in terms of cause and consequence. How about "locally adaptive phenotypic divergence" instead?

Page 8

- [Figure 4] **R:** Replace "variance" by "standard deviation" in the title of panel B.
- [l.2–4] **C/S:** I found this statement imprecise. Please modify to emphasise the *conditions* under which the architectures evolve and are maintained. For a type of wording I encourage, see e.g. Höllinger et al. (2019, Introduction).
- [l.4] **R:** Insert "the" before "mutation".
- [l.8] **R:** Delete "the" before "gene flow".
- [l.12–19] **C/R:** This part must be rephrased according to my Major Comment 1. "Mutation swamping" should be avoided.
- [l.21] **R:** Replace "mutation swamping" by, e.g., "the transition to a transient architecture".
- [l.27] **R:** Replace "mutation swamping" by, e.g., "this phenomenon".
- [l.30] **R:** Replace "mutation swamping" by "transient architectures". **S:** Replace "can occur" by "might underlie".
- [l.35] **S:** Insert "for the transition to a transient architecture" after "threshold mutation rate" to increase clarity.
- [l.37] **C:** I would prefer "standard deviation" instead of "variance" for consistency.
- [l.42] **R:** Replace "... traits would be subject to mutational swamping ..." by "... traits with a transient genetic architecture would be more likely to be found ..." or similar.
- [l.44] **R:** Replace "mutation swamping" by "the transition".
- [l.46] **C/R:** As explained in detail in my Major Comment 1, I struggled making sense of this part. I ask that the authors rephrase this part to clarify, and that they omit the term "mutation swamping".

Page 9

- [l.5] **R:** Replace "weak linkage effects" by "no physical linkage between the focal and the background loci", as this is what the authors said they assumed.
- [l.6–7] **S:** Replace "Of note, ..." by "Importantly, ...".
- [l.29] **S:** Replace "non-small" by "non-negligible". Replace "sustains" by "may be sustained".
- [l.31] **S:** Insert "statistical" before "linkage".
- [l.44] **S:** Replace "reducing" by "which would reduce" to avoid a dangling participle.

- [l.48] **S:** Insert "our" before "simulations" and ", which do include moderte linkage among background loci" after "simulations".
- [l48–51] **Q:** The authors state that the Bulmer effect might affect the evolutionary dynamics under stronger selection or tighter linkage. I wonder if the authors explored this scenario in their simulations, and what the outcome was.
- [l.50–51] **R:** Replace "... dynamic, deviating it from theoretical expectation" by, e.g., "... dynamic, and cause over analytical theory to fail" to fix an awkward phrasing.
- [l.51–52] **C:** I think that somewhere here the authors should also discuss the potential implications of their choice of indicator T^* to delineate concentrated vs. transient architectures. See my Major Comment 3 regarding the bias introduced by an inflation of the proportion of concentrated architectures.
- [l.52] **S:** Insert "stable vs. transient" before "genetic". Replace "architecture" by "architectures".
- [l.61] **R:** Insert "a" before "locally adaptive trait". **S:** Insert ", putatively polygenic," after "adaptive".
- [l.67] **R:** Delete "mutation swamping" and insert "we determined" after "threshold".
- [l.73–74] **R:** Replace "mean the action of mutation swamping" by, e.g., "imply a transient architecture".
- [l.74–77] **R:** Insert "a" before "transient architecture" and before "diffuse architecture".
- [l.91] **R:** Replace "that would permit mutation swamping" by, e.g., "under which we observed the evolution to a transient architecture under a high enough mutation rate".
- [l.93] **R:** Delete "swamping".
- [l.95–96] **C/S:** I found "one might think that it should hardly be the case" colloquial and would thus suggest to rephrase to "a small mutational target in this case would seem surprising".
- [l.99–100] **R:** Insert "a" before "large mutational target".
- [l.103] **S:** Replace "is" by "are" and "effect" by "effects".
- [l.104] **S:** Replace "is" by "are".
- [l.114] **C/S:** I would appreciate some more future perspectives and a concluding paragraph that summarises the main achievement of the study and its relevance. I suggest the authors expand on these points.

Appendices

Page 13

- [l.7–8] **S:** Delete "described as" because this seems unnecessarily wordy.
- [l.9] **S:** Replace "Noting that" by "Because".
- [l.15] **S:** Replace "which is arranged into" by "which simplifies to" or "which can be rearranged to".
- [l.17] **S:** Replace ". After some calculation, the condition is rearranged into:" by ", which can be rearranged to".
- [l.18] **S:** Replace "derived" by "obtained".
- [l.34] **S:** Replace "values" by "trait mean and the frequency of the mutant allele at the focal locus in population 2".
- [l.35] **S:** Replace "is" by "are". See the previous comment.

Page 14

-

Page 15

- [l.2 and other places] **R:** Replace "variation" by "variance". This comment also applies to l.7/8, l.13, l.14, and l.40, as well as in the caption of Figure C1.
- [l.8] **S:** Insert index "i" after "population". Replace "its result" by "it".
- [l.10–13] **C:** I think this additional introductory part becomes oblivious (and it is partially redundant) given the preceding paragraph. I would omit this part.
- [l.15–16] **S:** Delete "it is essentially one population model and". Insert "for a single population" after "theory".
- [l.17] **Q:** Could the authors please verify that this reference to Bürger et al. (1989) is the reference intended by them? I did not find it particularly helpful with respect to backing Equation C1.
- [l.23] **S:** Insert "for" after "account".
- [l.25] **S:** Replace "are finally" by "go".
-

Supplementary Tables and Figures

- Table S1: **S:** Insert "forward" before "migration rate" in row 1. **R:** Fix the indices in row 3, column "default": it should read "$\theta_1 = 0, \theta_2 = 1$".