

1 Summary

The author investigates the role of many loci in maintaining local adaptation in the face of migrational swamping. The paper illustrates the importance of statistical linkage disequilibrium (or between-population covariance in allelic effects among loci) and recurrent mutation in maintaining local adaptation if selection coefficients at individual contributing loci are so small that these could not be maintained in a single-locus model. The genetic architecture of local adaptation in such scenarios is transient, with individual beneficial mutations constantly establishing via recurrent mutation, and going extinct after some time due to genetic drift and migration. Individual loci show low population divergence, but overall, local adaptation is maintained. This is in stark contrast to the case where individual mutations are strongly selected, survive much longer in the face of gene flow, and hence contribute to much more stable architectures.

2 General issues

1. Overall I think the manuscript could be a useful contribution to the debate on the genetic architecture of local adaptation. However, as currently presented I think the paper is somewhat confusing. I did not finish the paper with a clear sense of what the simulations were telling us. For example, I not get a clear sense of the role of drift and N_e in the model, and whether drift was helping to establish some migration-selection polymorphisms (against initially unfavorable conditions) or if the larger mutational input with larger N helps. The author should be clear to spell out what the predictions of the simulations are, and what are hypotheses for further investigation.
2. In the Abstract, Introduction, and to some extent in the Discussion, a contrast is built up between population-genetic and quantitative-genetic models. It is then claimed that this paper shows that these approaches complement each other. I think this creates a wrong impression; the apparent contrast claimed here is merely due to the fact that the population-genetic models referred to are not appropriate for the question under study. They are almost exclusively one-locus models of migration and selection, without recurrent mutation. Of course, these cannot predict the dynamics of multilocus systems. They do not account for physical nor statistical LD, and because there is no mutation, genetic variance is ultimately lost in the face of drift. If the appropriate population-genetic model were used, no contradiction would be seen. I agree that such a model might be hard to analyse analytically, and it is therefore a good first step to resort to simulations. (It might be a promising endeavour to obtain analytical approximations.) All of the simulations done here are of a explicit population-genetic model after all. It would seem better to alter the story, shifting the focus to a comparison of one- to multilocus models, and to the importance of recurrent mutation in the face of drift and migration. Of course, emphasis should still be on weakly locally beneficial mutations. However, these questions have been addressed in several studies before, and it seems that this paper can only contribute a limited amount of conceptual novelty. As of now, it is even at risk of introducing an unnecessary confusion about incompatibility of population and quantitative-genetic predictions.
3. I did not like the mixing of models with directional selection (population-genetic predictions) and stabilising selection (quantitative-genetic arguments and simulation). Although I believe that when $n_{\text{tot}} \leq n_{\text{opt}}$, selection is essentially directional on all loci, it is a conceptual weakness of this paper. In general, in a model of spatially heterogeneous stabilising selection, migration can increase or decrease local mean fitness, depending on the constellation. With diverging directional selection, gene flow is always maladaptive.

4. I find the terms ‘swamping-prone’ and ‘swamping-resistant’ as defined here problematic. In cases where a population-genetic one-locus migration–selection model (OLM) predicts that $m_{\text{OLM}}^* < m$ for a given mutation, the same mutation may be associated with a critical rate $m_{\text{MLM}}^* > m$ in a model with multiple locally beneficial mutations (MLM), even if unlinked Bürger & Akerman 2011 Aeschbacher & Bürger 2014 . It really depends on what is considered the ‘relevant’ reference model. Similarly, the probability of establishment of a weakly beneficial mutation can switch from zero to positive if there are other beneficial mutations in the background. Establishment in such cases may be of very short duration, and the focus would have to shift to survival times. Weakly beneficial mutations linked to other beneficial alleles can have a positive establishment probability and a short expected survival time at the same time (?). I suspect that this is the case when divergence is maintained with individually ‘swamping-prone’ alleles in this paper. These alleles would not establish in a one-locus model. But in the context of large numbers of cosegregating weakly beneficial mutations, the critical migration rate is increased (or, in other words, the effective migration rate reduced), so that they are maintained for some time. They do go extinct fairly soon, but recurrent mutation maintains variation overall. It might be beyond the scope of this paper to investigate this in detail (I think some analytical work could be done), but I think these aspects should be mentioned. If the definition of ‘swamping-prone’ were extended to include short survival times after establishment (as opposed to being based only on m^* for establishment), it would seem better.
5. It is not until the Discussion that one of the crucial forces driving some of the more interesting results observed here is mentioned: statistical linkage disequilibrium – or, as termed here, between-population covariances in effect sizes among loci. (The other force being recurrent mutation.) I feel that this aspect should be introduced much earlier, perhaps as a hypothesis, with the paper geared at assessing if such LD is relevant. As acknowledged in the Discussion, this brings the current paper very close to Flaxman et al. 2014, and I am not totally convinced that it provides a substantial amount of further insight (beyond using stabilising rather than directional selection). I might be wrong on that, but then the justification should be worked out more clearly.
6. The drawbacks of the quantitative genetics model are only briefly mentioned as a need to specify V_G . The author should be clearer that quantitative genetics models have a number of serious drawbacks, and necessarily reflect a series of approximations to an explicit population genetic model (featuring a large number of loci) that may not hold over time-scales of more than a short number of generations. My reading of Hendry, Day, & Taylor is that no simulations were performed to show that their model held in a reasonable parameter range (perhaps this has been shown in subsequent publications?).
7. A number of terms/variables/models are introduced throughout the manuscript without clear definitions and explanations. E.g. on page 7, what are the definitions of D , V_G , ω , & m . What model leads to equation 1? Please ensure that all models and variables are well defined.
8. The author’s simulations assume that all loci have the same small effect on the trait. It seems likely that if there was a distribution of effect sizes larger effect loci would persist much longer (as per the author’s previous work in this area) and potentially the end result would be that the loci underlying the trait divergence may be relatively stable.

3 Specific comments

3.1 Abbreviations used

Q Question

C Comment

S Suggestion

R Re-formulation or change needed (usually followed by a suggestion)

→ Suggested change/correction

3.2 Title

S: “small-effect alleles” → alleles of small effect

3.3 Abstract

1.26–27 C: I would avoid focussing the story on an apparent contradiction between population- and quantitative genetic models. Rather, shift it to contrasting one-locus vs. multilocus models.

3.4 Introduction

1.53 S: Insert comma after ‘genetic architecture’.

1.55–71 C: It was not clear to me if you are discussing one-locus models only here.

1.62–65 C: If multilocus models are to be included, the mean fitness also depends on the relative amount of time spent in the various genotypes (marginal fitness).

1.65 S: Insert comma after ‘environment’

1.66–67 Q: Does this not also depend on the ratio of migration rates in various directions? Think of the continent–island model, where the locally beneficial allele may only be present on the island and spend no time at all on the continent.

1.66 R: ‘an’ → a

1.68 S: ‘an intermediate, non-locally adapted allele’ → an allele with spatially uniform fitness

1.69 C: Polymorphism will just as well be disfavoured in a stochastic model. **S:** ‘...disfavored, although...’ → ...disfavored. In addition, ...

1.62–71 S: Mention that ‘critical migration rate’ here refers to the process of establishment and does not have an immediate bearing on survival times after establishment (although there is of course a correlation).

1.72–74 C: But this depends on recombination and LD (or is this statement meant to apply to one-locus models only?).

1.75–79 C: I am not sure if this is true. If a deterministic model predicts no invasion, I expect that the corresponding stochastic process is subcritical, i.e. that the probability of establishment is zero. Aeschbacher & Bürger 2014 is incorrectly cited here. They have shown that *linkage to previously existing migration–selection polymorphisms* can increase establishment probabilities and survival times relative to the one-locus case. They also show that the criterion for invasion in the deterministic model is equivalent to the condition for a non-zero probability of establishment under a branching process (contrary to what is claimed here).

1.79–80 R: ‘...the distribution of allele effect sizes underlying a locally adapted trait...’ → ...the distribution of effects of alleles that underly a locally adapted trait...

1.80 R: ‘alleles’ → effects

1.81 R: ‘migration–selection balance’ → migration–selection balance. **S:** Insert “The same degree of local adaptation is reached with fewer alleles of larger effect.”

1.82–83 S: ‘Thus, we can test for the importance...in comparisons...’ → One way of assessing the importance...is to compare...

1.85–87 C: Again, it was not clear to me if the focus was exclusively on one-locus models. In general, the critical migration rate depends on the recombination rate between migration–selection polymorphisms.

- 1.87–90 C:** I found the definition vague (see above). What is the ‘relevant’ model? A one-locus or a multilocus model? Also, is swamping defined w.r.t. to establishment, or also to survival? Is the ‘critical threshold’ the critical migration rate, or can it be another threshold?
- 1.92 R:** Delete ‘with’.
- 1.94–95 C:** Given the current definition of swamping-prone, I am confused by the claim that SPAs may have non-zero probabilities of establishment. See comment to 1.75–79 above. Does the reference to Yeaman & Whitlock 2011 only refer to the second half of the sentence?
- 1.103–106 C:** Given that it is not clear that the previous section was apparently dealing with one-locus models only, this section is a bit of a surprise.
- 1.106 C:** If you want to cite Aeschbacher & Bürger 2014 (see comment to 1.75–79), it would seem more appropriate here.
- 1.132–135 C:** I had a problem with this sentence, given that it is not clear whether ‘swamping-prone’ was defined exclusively w.r.t. a one-locus model.
- 1.141 S:** ‘behavior’ → predictions
- 1.156–163 C:** Both V_G and V_P change in reality, but are assumed constant here. **C:** As mentioned above, I find the comparison between population and quantitative-genetic models problematic, because of the difference between directional and stabilising selection. This is discussed later, but might be worth mentioning here.
- 1.166–176 C:** This sentence is too long.
- 1.168 R:** ‘... where all phenotypes = 0, ...’ → where all individuals have a phenotypic value of 0.
- 1.188–192 C:** The difference is due to the fact that V_G is a parameter in the quantitative-genetic model.
- 1.195–197 C:** Because the contradiction between ‘population-’ and ‘quantitative-genetic’ models used here is not a real one, I do not like the way the story is developed (see major issues above). What the paper really is about is the effect of (statistical) LD on establishment of weakly locally beneficial mutations, and the role of recurrent mutation. Both can be phrased in a population-genetic context.
- 1.204 C:** Again, ‘population-genetic models’ is implicitly set equal to one-locus migration–selection models. I think this is a short-cut, because population-genetic models can be with or without linkage, with or without genetic drift, with or without mutations. These choices affect the meaning of the sentence (see comments above).

3.5 Simulation model

- 1.216 C:** Please check the typesetting of the γ in the exponent.
- 1.223 R:** ‘mutations are all diallelic’ → loci are all diallelic.
- 1.228–233 C:** Here, it reads as if statistical LD is of minor importance, but I think it is not, and later in the paper this is also acknowledged. For the same reason, I think that the prediction based on Eq. (2) might actually not be accurate. Again, the correctness of the statement in 1.233 depends on how exactly ‘swamping’ is defined.
- 1.234–236 C:** See comments above about mixing stabilising and directional selection in the argumentation.

3.6 Results

- 1.266–272 C:** My suspicion is that a combination of the reduction in the effective migration rate due to unlinked other locally beneficial alleles, and recurrent mutation, are crucial to explaining the result. The reduced m_e allows weakly beneficial mutations to become established, and recurrent

mutation compensates for their short survival time.

- 1.287–290 C:** At the early stage, when divergence is building up, selection against migrants is important. These carry several maladapted alleles, and even though these are unlinked, this reduces the effective migration rate. Later, when stabilising selection kicks in as some individuals overshoot the phenotypic optimum, migration has a mixed effect. I am not sure if a conclusion can be made about whether stabilising selection is not necessary for the long-term maintenance of polymorphism just based on the fact that it is not needed for the initiation of divergence.
- 1.305–311 C:** When reading this the first time and looking at Figure 3, I was wondering if established alleles might also disappear because they are hit by recurrent mutation. However, the time-scales of survival and mutation seem too different for this to be the case (the waiting time for a mutation at a given site is longer than the period over which an allele with large d exists in Figure 3C). It might be worth mentioning this, though.
- 1.316–324 C:** I found this very important. I was wondering though if drift was the only agent here, or whether redundancy and overshooting of the optimum also cause migration to contribute more to these dynamics. When alleles contribute to overshooting of the maximum, they are selected against, and the alternative allele receives a benefit when brought in by migration. I think this is reflected in the finding that if alleles are resistant to swamping and $n_{\text{tot}} > n_{\text{opt}}$, migration has a strong impact on V_G (Figure 1D and 1.331–339).
- 1.324–325 Q:** Given large, but still finite population size, would you not still expect to see a turn-over, just on a much longer time scale?
- 1.339 Q:** Would it be worth mentioning that the reverse is not true, i.e. that large $V_G \nRightarrow$ large D ?
- 1.342 S:** Insert ‘at individual loci’ after ‘(F_{ST})’.
- 1.343 S:** As θ_B is not strictly defined here, I would add a reference to Le Corre & Kremer 2003.
- 1.359–360 Q:** Do these theoretical models assume that μ is much smaller than the values you use?

3.7 Discussion

- 1.376 S:** Insert ‘and/or the polygenic’ before ‘level’.
- 1.382–385 C:** The maintenance of allelic diversity still depends on recurring mutation, as populations are still of finite size and variation can be lost by drift. Yet, the time-scale of loss and turn-over of genetic architecture is much longer. Aeschbacher & Bürger 2014 are mis-cited here. They approximated the survival time of beneficial mutations linked to existing migration–selection polymorphisms, showing that these become large, but they are still finite. There was no recurrent mutation in their model. Hence, they cannot have shown that maintenance of adaptive divergence does not depend on mutation. On the contrary, large, but still finite, survival times suggest the opposite.
- 1.395–396 S:** ‘in terms of the susceptibility of alleles to swamping and the factors...’ \rightarrow in terms of the α s or, ultimately, the distribution of fitness effects, and other factors...
- 1.399–401 C:** Surely, this depends on the distribution of fitness effects (DFE). Fixing the α s is artificial. Without knowledge about the DFE, the statement seems speculative. [I see you discuss this later.]
- 1.435–456 C:** This paragraph is important, as it lifts the apparent magic that was hovering above the story. Yet, it has been known before that statistical LD can drastically alter the dynamics of adaptation. Therefore, I think it should be mentioned earlier as a hypothesis to be tested.
- 1.458–461 Q:** Are you certain that a reduction in m_e was not decisive in this study, or is this a speculation? I suspect that a reduction in m_e is part of the explanation why divergence can emerge in the first place from alleles of weak effect. Long-term maintenance then also depends on recurrent mutation. Also, the extent to which a reduction in m_e is relevant depends on the

magnitude of ϕ and m . For instance, if ϕ is small, then a small absolute reduction in m_e can be substantial relative to ϕ . It can make the difference between a zero and a positive probability of establishment.

1.457–475 C: Given the similarity of these two studies, I was wondering about the extent of novelty contributed by the current paper (see major issues above).

1.495 R: ‘isn’t’ \rightarrow is not.

1.550–552 Q: Would such covariances not be very hard to measure?

4 References

general Page ranges should be given using an em-dash (–) instead of an en-dash (-).

general Be consistent with the separator for volume number and page range. Currently, both ‘:’ and ‘,’ are used.

1.606 C: Volume number missing.

1.658 R: “2003” \rightarrow 2004.

5 Figure Legends

1.760–762 C: Divergence in trait means seems to be missing in panel A.

1.772 Q: What is m here?

1.778 R: Substitute B for W as a subscript in the second θ_W .

1.785 R: Replace en-dashes (-) by em-dashes (–) in ‘migration-selection-drift balance’.

1.787–788 C: Because of this, I would not show this figure at all. It is overly simple and hence at risk of being misleading.

6 Figures

1C and 1D C: Dotted lines are indistinguishable from dashed lines. Please mention this in the caption.

2A C: Red line is missing in legend for $n_{\text{tot}} = 5000$.

4A C: It would be interesting to see a version of this plot with a much extended x axis. I suspect that there is a turn-over of architecture that can only be seen on a longer time scale. Show at least the last 10^4 generations. A similar comment applies to Figures 4C and 4D, although the time scale might be unreasonably long for simulations.