

In this manuscript, Kessinger and Van Cleve use simulations to study the rate at which asexually evolving rapidly adapting populations cross fitness valleys. Their results suggest that the average time to crossing a valley may increase as the variance in fitness in a population increases, and that in more rapidly adapting populations the depth of the valley may play a smaller role in determining the time to cross it.

While I think the work is well motivated, and that the authors have focused on an interesting question with broad appeal to a wide range of biologists and geneticists, I have several technical concerns with the study that I believe should be addressed before the manuscript is published. Primarily, I'm concerned that the simulations have been used to explore a very sparse set of parameters, and I'm not convinced they provide a comprehensive picture. Furthermore, the theoretical motivation for picking these exact parameters is also not obviously justified. In sum, these shortcomings make it difficult to evaluate the conclusions drawn from these results as well as the mechanisms put forward as explanations. The work would be substantially improved if the simulations were expanded to cover a broader range of parameters, so that known limiting forms at the ends of the parameter range can be recovered, and sanity checks are added. I give several suggestions below.

Major comments

1. The authors justify the choice of parameters and regimes to focus on by quoting the results from the analysis of valley crossing in non-drafting populations by Weissman et al. (2009). Yet, as the authors point out at several points, evolutionary dynamics and genealogies are very different in rapidly adapting populations. It is not clear to me that the regimes in rapidly adapting populations are likely to be the same as the regimes in the cases studied by Weissman et al. (2009), or that they occur at the same transition points in terms of the parameters.

For instance, the authors primarily focus on tunneling (where the beneficial mutation arises on the background of a polymorphic deleterious mutation that is not destined to fix), which in the limit studied by Weissman et al. (2009) requires that $\frac{1}{\sqrt{\mu s}} \lesssim N \lesssim \frac{1}{\mu}$. The lower limit of $\frac{1}{\sqrt{\mu s}}$ comes from the calculation of bubble sizes in the limit considered by Weissman et al. (2009). However, the distribution of bubble sizes will be very different in the limit that the authors consider. Therefore it is unclear why this is a relevant scale. Furthermore, it is unclear whether (depending on the population size, fitness effects, mutation rates, etc.) tunneling is likely to be an equally broad regime in rapidly adapting populations. Here, deleterious hitchhikers can have a substantial probability of fixation, which sensitively depends on the mentioned parameters and increases with the size of the population (see e.g. Good and Desai (2014)). Thus, I would expect that the regime of sequential fixation can be much broader in rapidly adapting populations, and that the precise population size at which it gives way to the tunneling regime would be a much more complicated

function of the parameters.

In principle, it should be possible to increase the range of parameters explored by simulation and see whether, e.g. transitions in the quantities measured seem to occur at anticipated transition points, whether the behavior is monotonic (see next point), and at what parameter values known limiting forms are achieved (e.g. in Figure 2, how do the rates compare to the values anticipated in the limits of low σ calculated by Weissman et al. (2009)?).

2. Furthermore, there seems to be substantial variance in the outcomes in drafting populations, so that though the average time to crossing the valley is increased, there are lucky occurrences where the time to crossing the valley can be substantially shorter than in non-drafting populations. This variance appears larger than the difference between the means, so that reporting the average behavior doesn't seem to capture the complete picture. Is this variance in outcomes driven by differences in the fitness of the background that the mutation arises on, so that some valleys are guaranteed to be crossed if the process is initiated in a high-fitness individual? Is the variance reduced when the depth of the valley is reduced? In addition, in Figure 2, though the average appears to asymptote, the fastest time appears non-monotonic. Is this a result of the relatively small number of replicates of the simulations (~ 100), or an phenomenon that arises from the dynamics? Answering these questions would be important to interpreting the results, and would make me more comfortable about the verbal interpretations of the figures.
3. In addition to this, because there are no calculations to guide interpretation, it would be important to add additional simulations to measure these quantities. Specifically, I might be missing something in the design of the simulations, but it's not entirely clear to me that what the authors are observing is true tunneling for all parameters considered. Might it be that the deleterious intermediates are fixing prior to the arising of the second "valley crossing mutation"? It might be a good idea to record the fixation times of both the intermediate and the double mutant to help interpret these results.
4. The way the simulations have been set up is non-standard and makes it hard to use previous results to interpret which regime the populations are in. For instance, the authors' claim is that the population is in the infinitesimal limit, where the only relevant parameter is σ (or, more accurately, $N\sigma$). Yet Good et al. (2014) showed that for moderate $N\sigma$ there is a whole manifold of behaviors depending on the mutation rates and selection coefficients, and that observables do not always agree with asymptotically large $N\sigma$ limits. The fact that fitness effects of "background" mutations in these simulations are adjusted every generation to keep σ constant makes it difficult to understand what the background selection coefficients and overall background mutation rates are. Furthermore, the choice of exponential distribution

for the selection coefficients of the genetic background may put the population in a regime where the infinitesimal limit may not always apply (see Fisher (2013)). Finally, the equilibration time of $N/10$ generations seems potentially short (one would hope for minimally a timescale of coalescence, and it's not possible to interpret whether this is satisfied here).

Minor comments

1. Page 2, column 2: “Alleles whose dynamics are primarily governed by genetic drift can be effectively modeled by a diffusion approximation” and “drafting populations experience quite large jumps in allele frequencies that cannot be encapsulated by a diffusion approximation”: I’m not sure this is correct: allele frequency trajectories can in fact be modeled by a diffusion equation, albeit a more complicated one (e.g. the reference cited, Neher and Shraiman (2011a), uses such an approach). I’m also not sure what is meant by ‘jumps’ here – allele frequency trajectories tend to be no less smooth in drafting populations than in drifting populations.
2. Page 3, column 1: “small bubbles are more common, but large bubbles are very rare”: I don’t think the first half of this statement is correct. If mutations arise equally frequently in both cases, small bubbles must be equally common.
3. Page 3, column 2: “In previous approaches, these stochastic effects were smoothed out due to the presence of recombination”: Perhaps I don’t understand what is meant here, but this strikes me as inaccurate. The effects of drift are in the nose of the distribution are of critical importance in the case considered by Neher and Shraiman (2011a): these authors do take them into account to obtain the establishment probabilities and allele frequency trajectories (e.g. Eq. 4 in that work).
4. Page 3, column 2: “This is consistent with the fact that *all* forms of adaptation are slowed down due to clonal interference and genetic background effects”: Does it mean to imply that the rate at which valleys are crossed is reduced by the same amount by the rate at which beneficial mutations of effect s are likely to fix compared to $NU_b s$? This is a quantitative claim, and I don’t think it’s been shown. If I’m missing it, then perhaps a reference to a figure would be appropriate.
5. Page 4, column 1: “By Fisher’s “fundamental theorem”, the rate v at which the mean fitness advances is set to σ^2 ”: this implies that $v = \sigma^2$, ignoring the $\mu\langle s \rangle$ term [in $v = \sigma^2 + \mu\langle s \rangle$] that arises in the full moment expansion, e.g. see Neher and Shraiman (2011b) for a recent review. This term is sometimes subdominant, but not in all regimes (and it’s not obvious it’s subdominant in the simulations carried out by the authors).

6. Page 4, column 2: The analysis of timescales seems inaccurate: the estimate that T_0 is of order $\frac{1}{s} \log(Ns)$ is not appropriate for drafting populations.

References

- FISHER, D. S. 2013. Asexual evolution waves: fluctuations and universality. *Journal of Statistical Mechanics* .
- GOOD, B. H. AND DESAI, M. M. 2014. Deleterious passengers in adapting populations. *Genetics* 198:1183–1208.
- GOOD, B. H., WALCZAK, A. M., NEHER, R. A., AND DESAI, M. M. 2014. Genetic diversity in the interference selection limit. *PLoS genetics* 10:e1004222.
- NEHER, R. A. AND SHRAIMAN, B. I. 2011a. Genetic draft and quasi-neutrality in large facultatively sexual populations. *Genetics* 188:975–996.
- NEHER, R. A. AND SHRAIMAN, B. I. 2011b. Statistical genetics and evolution of quantitative traits. *Reviews of Modern Physics* 83:1283–1300.
- WEISSMAN, D. B., DESAI, M. M., FISHER, D. S., AND FELDMAN, M. W. 2009. The rate at which asexual populations cross fitness valleys. *Theoretical Population Biology* 75:286–300.