**Supplemental Information for**

**Chromosomal inversions can limit adaptation to new environments**

Molecular Ecology

**Marius Roesti**, **Kimberly J. Gilbert, Kieran Samuk**

**Table of Contents**

Supplementary Methods, pages 2–3

Supplementary Figures S1-S3, pages 4-6

Supplementary References, page 7

**Supplementary Methods**

*Description of simulations*

Our illustrative simulations were performed using SLiM version 3.7.1 (Haller and Messer 2019). The code to produce all the simulation results as well as to create the plots shown in the manuscript is available at <https://github.com/ksamuk/inversion_constraint>. A brief description of the simulations follows.

*General simulation structure*

Each simulation had the following basic structure: Two populations, population 1 and population 2 (hereafter p1 and p2), are initialized in the first generation. Each population is composed of 2500 hermaphroditic individuals, with each individual having a diploid genome composed of three loci: two fitness-affecting loci, and one “inversion” locus. Each of the two fitness-affecting loci has two alleles: the “m1” allele and the “m2” allele, each favored in p1 and p2, respectively. Each allele has a selection coefficient of *s* = 0.05 in its home population, and a coefficient of -*s* (i.e., -0.05) in the alternate population (i.e., symmetrical divergent selection). All alleles have dominance coefficients of *h* = 0.5 (i.e., pure additivity). The inversion locus similarly has two alleles, “m3” and “m4”, corresponding to the genomic inversion rearrangements in p1 or p2, respectively. No de novo mutations occur (i.e., adaptation occurs entirely from standing variation). Throughout the simulation, the two populations exchange migrants at a rate of *m* = 0.01 (symmetrical gene flow).

Because of the simplified genomic architecture, we set baseline recombination rates at a value of 1e-2 (SLiM units) to observe sufficient recombination over the course of the simulations. We modelled recombination suppression by the inversion after the example in the SLiM manual, i.e., if an individual is heterozygous at the inversion locus, recombination is suppressed at all three loci (i.e., the fitness loci and the inversion locus). Otherwise, recombination proceeds at the baseline rate.

Simulations were run for 5000 generations. At each generation, we output the mean relative fitness of each population, scaled against the maximum possible fitness based on the known optimal genotypes. To examine the effect of the inversion, we ran simulations with and without the inversion active. We ran 100 replicates of each simulation, and then processed and plotted the output from SLiM using R version 4.1.2 (R Core Team, 2021) and the tidyverse package (Wickham et al. 2019).

Using this core simulation structure, we simulated three different scenarios:

*(i) “Novel environment” scenario*

In this scenario, the simulation follows the general simulation structure described above, but at *t* = 1000 generations, a third population (hereafter p3) is founded with half of its individuals sourced from p1 and the other half from p2. In p3, at the first fitness locus, the m1 allele has a selection coefficient of *s* = 0.05 and the m2 allele has a selection coefficient of *s* = -0.05. At the second fitness locus, the m2 allele has a selection coefficient of *s* = 0.05 and the m1 allele has a selection coefficient of *s* = -0.05. As such, the optimal genotype in p3 is m1/m1 at the first locus, and m2/m2 at the second locus, i.e., an intermediate between p1 and p2. After the initial founding event, gene flow between all populations continues at the rate of *m* = 0.01.

*(ii) “Environmental change” scenario*

This scenario is similar to the “novel environment scenario” above, but instead of a new population being founded at *t* = 1000, only two populations still exist and instead the selection coefficients in p2 shift to those described for p3 above (i.e., the optimal genotype is intermediate between the original p1 and p2). All other parameters, including migration rate, remain unchanged.

*(iii) “Polygenic” scenario*

To explore the effect of a more complex genetic architecture, we simulated the novel environment scenario with a genome containing 101 loci: 100 fitness-affecting loci and one inversion locus, or with 40 fitness-affecting loci, 60 neutral loci, and one inversion locus. The simulation was otherwise analogous to the simulation with only two fitness-affecting loci, that is, exactly half of all fitness-affecting loci had m1 favored in p3 while the other half of the loci had m2 favored in p3. Loci were arranged such that at all even and uneven loci along the genome, m1 and m2 were adaptive in p3, respectively.

*Basic exploration of parameter space*

While not our primary goal, we explored the robustness of our results by varying the strength of selection (*s*) and migration rate (*m*) at 0.01, 0.05, and 0.1 in both the “novel environment” scenario (Figure S1) and the “environmental change” scenario (Figure S2). All simulations produced qualitatively similar results, i.e., the inversion limits adaptation to changing ecological conditions.

**Supplemental Figures**

**Figure S1**

**Graphical user interface, diagram

Description automatically generated with medium confidence**

**Figure S1. Simulation results for the three populations in the “novel environment” scenario across nine parameter combinations.** Each panel depicts the change in scaled mean fitness of a simulated population (colors) over time in the presence or absence of a chromosomal inversion (solid and dashed lines, respectively) under a single parameter combination. Each column depicts simulations performed at different total strengths of selection (the sum of the magnitude of all selection coefficients across all loci in any of the given environments). Each row depicts simulations performed at different migration rates. Note that under each parameter combination, the presence of an inversion limits adaptation of population 3 into the novel habitat (i.e., the dashed green line is always above the solid green line).

There are two further observations that may be worth explaining. First, in both populations 1 and 2, fitness is generally reduced less in the presence of the inversion as compared to when it is absent. This is because when the inversion is present, population 3 cannot reach its optimal haplotype (with locally favorable alleles at both fitness-affecting loci) and thus will contain a mix of the alternative inversion arrangements adaptive in either population 1 or 2. As a consequence, gene flow from population 3 into population 1 and 2 is less maladaptive, on average, in the presence of the inversion. Second, in the absence of the inversion, fitness is generally higher in population 3 than in population 1 and 2. This is because without the inversion, locally beneficial haplotypes can form for all three populations. Overall gene flow into population 3 is thus less maladaptive than overall gene flow into population 1 and 2. In other words, total migration load into population 3 is lower than into population 1 and 2. For further details on the simulations, see Supplementary Methods and the caption of Figure 1b.

**Figure S2**

**A picture containing diagram

Description automatically generated**

**Figure S2. Simulation results for the two populations in the “environmental change” scenario across nine parameter combinations.** In this scenario, at generation 1000, the optimal genotype in population 2 shifts to the optimal genotype from population 3 in the “novel environment” scenario described in the text. Each panel depicts the change in scaled mean fitness of a simulated population (colors) over time in the presence or absence of a chromosomal inversion (solid and dashed lines respectively) under a single parameter combination. Each column depicts simulations performed at different total strengths of selection (the sum of the magnitude of all selection coefficients across all loci in any of the given environments). Each row depicts simulations performed at different migration rates. Note that under each parameter combination, the presence of an inversion limits adaptation of population 2 into the changed habitat (i.e., the dashed blue line is always above the solid blue line after generation 1000).

It is also interesting to note that population 1 has an increased fitness when the inversion is present. The reason for this is that population 2 contributes less maladaptive gene flow in return to population 1 because population 2 is unable to attain its optimal inversion haplotype and has more population 1-adaptive alleles present as compared to without the inversion. For further details on the simulations, see Supplementary Methods.

**Figure S3**

Diagram

Description automatically generated with medium confidence

**Figure S3. Simulation results for the “polygenic” scenario (100 loci) across nine parameter combinations.** This scenario is identical to the “novel environment” scenario described in the text, except fitness is determined by 100 loci instead of two. Each panel depicts the change in scaled mean fitness of a simulated population (colors) over time in the presence or absence of a chromosomal inversion (solid and dashed lines respectively) under a single parameter combination. Each column depicts simulations performed at different total strengths of selection (the sum of the magnitude of all selection coefficients across all loci in any of the given environments). Each row depicts simulations performed at different migration rates. Note that under all parameter combinations except when the total strength of selection is very weak compared to migration (left panels), the presence of an inversion limits adaptation of population 3 into the novel habitat (i.e., the dashed green line is above the solid green line).

We further note that identical simulations but with 40 instead of 100 loci under selection produced highly similar results and are thus not shown.

**Supplementary References**

Haller, B. C., & Messer, P. W. (2019). SLiM 3: Forward genetic simulations beyond the Wright–Fisher model. *Molecular Biology and Evolution, 36*(3), 632-637. doi:10.1093/molbev/msy228

R Development Core Team (2021). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43). doi:10.21105/joss.016