# Introduction

We study evolution to understand natural diversity; adaptation via natural selection is the cause of complex forms; natural selection acts on genetic diversity; the amount and direction of diversity limits a population to a certain range of possible phenotypes; particularly the additive variance in traits is important because it is heritable;

Additive variance is heritable; explains polygenic traits

At the heart of the evolutionary sciences is the need to understand the natural world’s diversity. Darwin’s (1863; SOURCE) introduction of natural selection some 140 years ago led to increasingly accurate glimpses into the units of evolution, genes, and their movement through a population in response to selection (SOURCE). However, these movements, particularly in a multivariate trait space, become a challenging realm to predict (SOURCE; Lande 1979, 80 etc.). To navigate this space, it is necessary to reduce the predictors of trait trajectories to their principles: how they affect additive genetic variance, the heritable component of trait variability.

Additive genetic variability, VA, is regarded as the most important predictor of a population’s adaptability (Lynch and Lande 1998; Aguirre *et al.* 2014; Careau *et al.* 2015), and hence it’s trajectory through time towards a phenotypic optimum. Although a multitude of stochastic and deterministic processes also contribute to the population’s total trait variability, VA is heritable, and therefore predictive of a population’s trajectory over micro-evolutionary time. The VA of a population determines the phenotypic space that population can explore. Hence, it is predicted that populations with large amounts of VA are best suited to adapting to novel environments (Barton and Charlesworth 1998). Such an example is X. However this is not always the case, standing genetic variation is characterized by a variety of architectural and population-level constraints such as rates of pleiotropy, selection strength, additive effect size, linkage, and deleterious mutation/background selection (SOURCE). For example, under infinitesimal models, selection has a trivial impact on standing variation (Barton 2017).

including genetic drift, selective pressures, additive effect sizes, between- and within-gene interactions, and heritability (SOURCES).

* Natural diversity, population movements in trait space
* Heritable variation
* Stabilising selection, effect on variation/need for variation vs drift
  + Expected to remove variation, mutation alone can’t explain why in natural populations we see so much variation: why?
* Additive effect sizes, effects on variation
* Background selection, effect on variation
* Population genetics expectations of variation under bkg sel, additive effects

# Methods

Using the forward-genetics modelling package SLiM 3.4 (Haller and Messer 2019), I constructed two models to explore a portion of the multivariate parameter space that explains genetic variability in natural populations. These parameters included genome wide recombination rate, the additive effect size distribution, and the selection strength multiplier, (Table 1). The rate of universal pleiotropy, and the amount of mutational covariance between traits was also varied across models but was not considered for analysis. The relative rate of deleterious mutation compared to trait mutations was also varied across models, but not considered for analysis due to the confounding effects of reduced trait mutation rate with background selection. Among these models, multiple conditions and assumptions were shared.

## Common model elements

Both of my experimental models consisted of a SLiM 3.4 model simulating a Wright-Fisher population of 8000 diploid individuals evolving over 100,000 generations. Populations first were subject to 50,000 generations of burn-in to build standing variation to mutation-drift balance (figure S1). Individuals were characterized by 8 traits, controlled by 100 loci each. Each trait had an identical effect on fitness, forming a ‘mega-trait’ with varying variance-covariance structures depending on pleiotropy rates. Each locus was assumed to have identical length, and each base pair within it mutationally independent, an assumption supported by a study by Thornton (2019), which found that within-locus differences in linkage had no average effect on either genetic variance or the mean trait value, indicating within-locus independence. In addition, the average number of base pairs per locus is highly conserved within eukaryotes (Xu *et al.* 2006), indicating that the assumption of equal gene length is valid. Mutations were modelled as occurring at an arbitrary position within the locus (or its regulatory regions) and is of arbitrary form. Mutations were assumed to be completely additive in effect, with no dominance or epistatic interactions, aside from additive epistasis occurring as a result of the fitness function. All loci were assumed to be on the same chromosome, with genetic distance being determined by the recombination rate parameter, r (Table 1). Both models had a genome-wide germline mutation rate of 8.045x10-6 per locus per generation, based on an average of five groups of eukaryotes (Aston *et al.* 2017).

The effective population size, Ne = 8000, was chosen to compromise between computational performance and the effect of genetic drift on populations under stabilizing selection. This value results in weak genetic drift in comparison with the strength of selection (under strong selection pressures), and appropriate standing genetic variation following burn-in to allow for adaptation (Lynch and Lande 1998).

Mutational effects on trait values were sampled from a normal distribution,

where λ is the additive effect size (Table 1). In the case of pleiotropy, a multivariate normal distribution was used, where n = 8, and

where **Σ** is a covariance matrix with diagonal values equal to λ and non-diagonals pulled from a normal distribution:

where is the parameter value of mutation correlation. **Σ** was ensured to be positive definite by multiplication with its transpose,

Non-trait deleterious mutations had fitness effects sampled from a gamma distribution:

Where and (SLiM Manual). This describes a distribution of fairly weak deleterious mutations on average.

All models were subject to 50,000 generations of burn-in, where mutations accumulate until the population reaches mutation-drift equilibrium. This is tracked as heterozygosity through the simulation, where mutation-drift equilibrium occurs when:

where µ represents the per-locus mutation rate per generation (Kimura and Crow 1964). A population at equilibrium was assumed sufficiently burnt in. Trials indicated that 50,000 generations of burn-in was sufficient for our population size (Figure S1). Deleterious mutation/mutation rate lowered the value of away from expectation in initial burn-in tests, however an alternative equilibrium was reached, satisfying the requirements of burn-in regardless of the parameter (Figure S1). During the simulation run, trait variances, covariances, and trait means were collected every 500 generations to track distances from the optimum and trait variability over time. At the end of the simulation, the allelic effects of segregating mutations in all populations were collected.

## Model specific characteristics

After reaching equilibrium, populations evolved for 100,000 generations of neutral drift or stabilizing selection, depending on the treatment. Neutral drift entailed no change from the properties of the burn-in, whereas stabilizing selection imposed a fitness function on phenotypes, invoking a multivariate optimum a fixed distance from the population mean phenotype post-burn-in. The position of the optimum is defined as:

Where is the vector of phenotype means, is the per-locus, per-generation mutation rate, , is the number of mutational steps to reach the optimum, and is the number of generations of burn-in. For our purposes, and . This distance was quite close to the original phenotypes, meaning most of the simulation investigated the maintenance of variation at a fitness optimum. The fitness of an individual in the population was defined as:

Where s is the selection coefficient, represents the gradient of the selection curve, n is the number of traits, and xn is the phenotype for trait n. To ensure a theoretical minimum and maximum fitness, s was fixed at 0.9, ensuring minimum fitness was , and maximum fitness was 1. This results in individuals at the optimum being at most ten times as fit as those infinitely far from the optimum. The model-specific maximum fitness difference depends on, which adjusts the realized fitness gradient via the curvature of the fitness function.

## Model Parameterization

Five parameters were shared between models, with a sixth for testing selection (Table 1). These were sampled using a Latin hypercube sampling design, with 1024 parameter combinations testing the null model, and 256 for the selection model. These samples were generated using the R packages ‘DoE.Wrapper’ and ‘LHS’, using the maximin algorithm (Melo *et al.* 2015; R Developmental Core Team 2019). Each model was repeated 100 times, using 100 seed values fed to SLiM. These seeds were randomly sampled from a uniform distribution of the total range of unsigned 32 bit integers (1 to 232 – 1) (R Developmental Core Team 2019). The array of parameter combinations and replicates was processed across 1152 cores on the University of Queensland’s Tinaroo high performance computing (HPC) system, using embedded Nimrod scripts to feed parameter/seed combinations to individual SLiM processes. Of these models, only 512 null and 128 selection model combinations were kept. Models with the top and bottom 25% of deleterious mutation rates were discarded to keep trait mutation rates roughly even.

## Analysis

Despite not all data conforming to normality, no data was transformed owing to the large sample sizes. Previous work into the robustness of regression modelling, t-tests, and F-tests have shown that departures from normality can usually still provide reliable estimates, provided the number of observations is large enough that coefficient estimates are approximately normally distributed due to the central limit theorem (Lumley *et al.* 2002). This was verified with diagnostic tools in the R package “jtools” (Long 2020).

For analysis, each parameter was grouped into three categories: low, medium, and high, with each bin containing a third of the total data. was the exception to this, with a fourth bin, null, describing the neutral models with no value at all. Across all analyses, means of responses were compared at the final generation of the simulation (100,000). Trait variances and covariances were pooled and averaged to form a ‘mega-trait’ average variance and covariance, since traits were functionally identical. Mean trait variances were compared with a multiple regression model including additive effect size, recombination, and selection strength.

In addition, I computed the population mean Euclidean distance from the optimum for each replicate and model:

Where pi and qi are the population mean and optimum value, respectively, for trait *i*. I compared distances with multiple regression, using EHW robust standard errors

To better visualize distances from the optimum, I calculated the probability that parameter combinations could reach the optimum over their 100 replicates. Populations that were ≤ 1 unit away from the optimum at generation 100,000 were considered at the optimum. I constructed a generalized linear model using a Poisson fit with log link to ascertain the effects of additive effect size, recombination rate, and selection strength () on the probability of reaching the optimum.

I also collected the mutational effects of segregating alleles at the end of the simulation. With this, I compared mean distributions of allelic effect sizes according to additive effect size, recombination rate, and selection strength with multivariate multiple regression. Responses included mean allelic effect, variance, and kurtosis of the distribution. I adjusted for heteroskedasticity with EHW robust standard errors. Regressions were calculated across 50 replicates owing to RAM limitations.

# Results

## Diagnostics

To determine the dynamics of the model under selection, I plotted variance and distance to the optimum over time across selection strengths. By generation 100,000, models have not yet reached mutation-selection-drift equilibrium, as variance continues to increase, however trajectories approach stability. Variance decreased initially after introducing the selection regime under low and medium models, before increasing over time due to the prevalence of genetic drift (Figure 1A). Distance to the optimum acted similarly: under low and medium selection, the initial response was a rapid movement towards the optimum. High selection led to no such drop in distance, indicating a lack of ability to adapt given the strong selection regime (Figure 1B). Following the initial decline, all models travelled further away from the optimum before stabilizing to their local ‘best fit’, the closest distance to the optimum where selection strength alone could take them (Figure 1B). The distance from the optimum remains stable after generation 50,000 across all selection strengths. Assuming we are approaching mutation-selection-drift equilibrium, we can describe which existing quantitative genetics models best describe our simulations.

Existing quantitative genetics models predict each locus’s distribution of allelic effects via the strength of selection relative to the mutation rate (Walsh and Lynch 2018). High mutation rates relative to selection strength is an assumption of Gaussian approximation models, and strong selection relative to mutation is an assumption of HOC models. I compared mean trait variance with selection strength to determine the range of CoA models sampled by the parameter space (F6, 12793 = 6686, p < 0.0001, Adjusted R2 = 0.798). Selection strength had weak effects on mean variance, with a 10% increase in selection strength increasing mean variance by 6.584 ± 1.127 units (t12793 = -5.849, p < 0.0001). High levels of unexplained variance within levels of selection indicated that genetic architecture parameters may be more important determining trait variance in populations (Figure 2).

## Patterns of variation with background selection and allelic effect size distributions

To determine the effects of genetic architecture on trait variance, I compared the effects of deleterious mutation rate, and additive effect size on mean trait variance with a linear model (F3, 127996 = 349063.4, p < 0.0001, Adjusted R2 = 0.8911). Recombination rate, pleiotropy rate, and selection strength (along with their pairwise interactions described 0.0051% of variance and were removed from the model. Deleterious mutation rate and allelic effect size contributed strongly to variance (Figure 3). A unit increase in additive effect size resulted in an increase of 110.0 ± 0.1341 units of mean trait variance (t127996 = 820.09, p < 0.0001). A 10% increase in deleterious mutation rate resulted in a decrease of 64.334 ± 0.1327 units (t127996 = -484.7, p < 0.0001). However, these main effects were masked by a strong interaction: the effect of deleterious mutation rate on mean trait variance decreased with increasing additive effect size (β3 = -174.2782 ± 0.4452; t127996 = -391.5, p < 0.0001). Low deleterious mutation rates enabled large increases in mean variance with increasing additive effect size, however this increase was largely constrained under high deleterious mutation rates (Figure 4). These changes to trait variance have strong predictions for adaptation under stabilising selection, specifically in the (Zhang 2012) adherence of trait means to a trait optimum.

## Adherence to a multi-trait optimum with increasing background selection and additive effect size

It is commonly theorised that genetic variability is strongly linked to the adaptability of populations under stabilising selection (Zhang 2012; Walsh and Lynch 2018). To measure this, I calculated Euclidean distances of populations () from the optimum at generation 100,000, and the probability of a given model to reach the optimum, . Deleterious mutation rate and additive effect size, along with their interaction were included in a linear model (F3, 127996 = 122193.3, p < 0.0001, Adjusted R2 = 0.7412; Figure 3). Again, pleiotropy rate, recombination rate, and selection strength (along with their pairwise interactions) explained <1% of variation and were excised from the linear model. Increasing deleterious mutation rate by 10% decreased distance to the optimum by 23.305 ± 0.088 units (t127996= -252.5, p < 0.0001), whilst additive effect size by 1 unit increased the distance to the optimum by 47.874 ± 0.089 units (t127966 = 536.37, p < 0.0001). Again, a significant interaction was observed (Figure 5). Increasing additive effect size under high rates of deleterious mutation resulted in smaller increases in distance to the optimum than under low rates of deleterious mutation (β3 = -36.860 ± 0.296; t127996 = -124.4, p < 0.0001). The probability of reaching the optimum reflected this, with the probability of reaching the optimum increasing with deleterious mutation rate under low effect sizes, but remaining constant at 0 with effect sizes greater than 3 (Figure 6). The probability of reaching the optimum increased linearly under strong selection pressure, however more quadratic trends were visible at lower selection strengths (Figure 6). To understand these patterns, the alleles underpinning trait variation and adaptation needed to be quantified.

## Distributions of segregating alleles under background selection and growing additive effect size

I compared rare allele frequency (RAF) with increasing additive effect size and deleterious mutation rate. The resulting linear model found significant differences between additive effect size, deleterious mutation rate, pleiotropy rate, stabilizing selection presence/absence and interactions between these four parameters (F15, 63937 = 1174, p < 0.0001, Adjusted R2 = 0.2784).Increasing deleterious mutation rates by 10% decreased RAF by 1.158 ± 0.609, however this difference was marginally insignificant effect (t63937 = -1.903, p = 0.057). Under selection, this decrease became highly significant, decreasing deleterious mutation rate’s effect on RAF by 9.381 ± 0.924 (t63937 = -10.148, p < 0.0001). Increasing additive effect size showed a small increase in the number of rare alleles under no selection (2.288 ± 0.6658 rare alleles per unit increase in additive effect size; t63937 = 3.437, p = 0.0006). Under stabilising selection, this effect was reduced by 5.113 ± 0.894 rare alleles (t63937 = -5.721, p < 0.0001). Under null conditions, the RAF reduction due to increasing deleterious mutation rate decreases by a further 3.020 ± 0.893 for a unit increase in additive effect size (t63937 = -3.388, p = 0.0007). Hence, increasing deleterious mutation rate reduces RAF to a greater extent under higher additive effect sizes. Under selection, this effect was reversed; the RAF reduction due to deleterious mutation is alleviated by 12.437 ± 1.314 units for a unit increase in additive effect size (t63937 = 9.464, p < 0.0001). This implies that deleterious mutation rate reduces RAF to a greater extent under lower additive effect sizes when under stabilising selection. The effects of additive effect size and deleterious mutation rate were highly visible in terms of the distributions of segregating alleles (Figure 7). The total numbers of all mutations decreased with increasing selective pressures (either by the presence of stabilising selection in Figure 7B or increasing deleterious mutation rates).

# Discussion

Surprising result: deleterious mutation reduces distance to optimum under maintenance, particularly under large size effects; still anchored in quant gen theory, even though pop gen predicts the opposite; pop gen – Ne reduced with BS, decreased variation expected, worse selection, more drift etc.

We found that increasing rates of deleterious mutation resulted in populations being more able to maintain their position around the optimum, overcoming some of the difficulties of fending with large-effect alleles that may pull populations away from the optimum (Figures 3, 5). Although it may at first seem counter-intuitive that stronger background selection increases the ability of populations to maintain their position at an optimum (Figure 5, 6), the effect can be explained with existing quantitative genetics expectations. To understand this, we must first explore the population genetics expectations of the effects of background selection on adaptation, and distinguish the difference in expectations between a population approaching the optimum and maintaining its position once it has arrived there.

It is well understood in population genetics that background selection reduces effective population size, reducing the effectiveness of selection and increasing the strength of genetic drift (Charlesworth *et al.* 1997; Houle 1998). As deleterious mutations are removed from the population, close-by linked QTLs are also removed (Charlesworth and Charlesworth 2010). The effect of this is decreased genetic diversity. In population genetics studies this is usually expressed in terms of FST or , whereas in quantitative genetics the analog is additive genetic variance (Falconer 1996; Charlesworth *et al.* 1997). Reductions in VA with increasing background selection were observed in this study, supporting this expectation (Figure 3A, 4). The expected effect of this on adaptation is quite clear when considering the initial approach towards the optimum: in quantitative genetics models, genetic variability is expected to increase the trait space that populations are able to explore, improving their ability to travel towards an optimum (Fisher 1930; Charlesworth and Charlesworth 2010; Aguirre *et al.* 2014). Indeed, these theoretical expectations have been found in natural populations: for example, Pujol and Pannell (2008) showed that populations of annual mercury, *Mercualis annua,* were able to respond to selection for pollen production when standing genetic variation was higher. Similarly, studies into the adaptation of red flour beetle (*Tribolium castaneum*) populations to new niches found high standing variation decreased the likelihood of extinction, and increased rates of niche expansion (Agashe and Bolnick 2010; Agashe *et al.* 2011). However, these expectations do not describe what we found in the current study: the most well-adapted populations consistently have higher rates of deleterious mutation, and hence lower standing genetic variance. The key to this lies in the expectations of the *maintenance* of variation and fitness around an optimum rather than the *approach* towards said optimum. The expectations surrounding this temporal space is considerably less extensive than that of the adaptive walk.

While reduced standing variation is expected to increase the time a population takes to reach an optimum (or perhaps prevent populations from reaching it at all), once a population has reached its optimum or stabilizes around its ‘local optimum’, the closest position it can maintain given the selected traits’ genetic architectures, mutation rates, and the population size – where does the population go?

In theoretical quantitative genetics, much debate is had over which particular models best describe the maintenance of variation in the presence of stabilizing selection over time. Selection is able to retain variation, particularly in large populations where drift is weak, and scenarios where balancing selection creates a non-linear fitness landscape, however the extent of this differs depending on many factors, including selection strength, genetic architectures, epistatic and dominant interactions, and the strength of selection relative to mutation (Walsh and Lynch 2018). Understanding the relative strength of selection to mutation has led to two distinct approximations of expected distributions of allelic effects. When mutation is much stronger than selection, Kimura (**1965a**) and Fleming’s (**1979**) Gaussian approximation holds true, whereas when the opposite occurs, Turelli’s (**1984**) house of cards approximation is more accurate. This distinction between models is arbitrarily granular, mostly for analytical viability. Computational methods allow for a continuous exploration of this space of models.

Figure 3: decrease in var with deleterious mutation is analogous to effects of lower Ne, but on a per locus level rather than genome wide. Hence, gives a proxy of the assumptions of CoA models with N -> Inf

Loss of fitness due to variation around optimum: expected to be 1/4Ne without any background selection (will vary with Ne due to effect on local Ne ) – Lande 1976

# Snippets

Underpinning this model is the continuum of alleles model of allelic effects, suggesting large numbers of alleles at many loci forming a continuous distribution of effect sizes, usually Normal in shape (Lande).

Pleiotropy fundamentally alters the signatures of HCA vs Gaussian approximation in COA so they approach each other - other parameters may as well?

Most effort in understanding stabilizing selection has focused on assuming either a Gaussian (as in this paper) or quadratic fitness function

Pleiotropy also had strong effects, due to contributing more than one trait value per mutation. Increasing pleiotropy rate by 10% increased RAF by 59.366 ± 2.531 alleles under no selection (t63937 = 23.458, p < 0.0001). Increasing deleterious mutation with pleiotropy rate significantly reduced this effect, with a 10% increase in pleiotropy rate and deleterious mutation rate simultaneously leading to a total loss of 22.555 ± 5.994 alleles (t63937 = -17.795, p < 0.0001). Under stabilising selection, a simultaneous 10% increase in pleiotropy rate and deleterious mutation rate led to an increase of 23.553 ± 8.567 alleles (t63937 = 19.193, p < 0.0001).

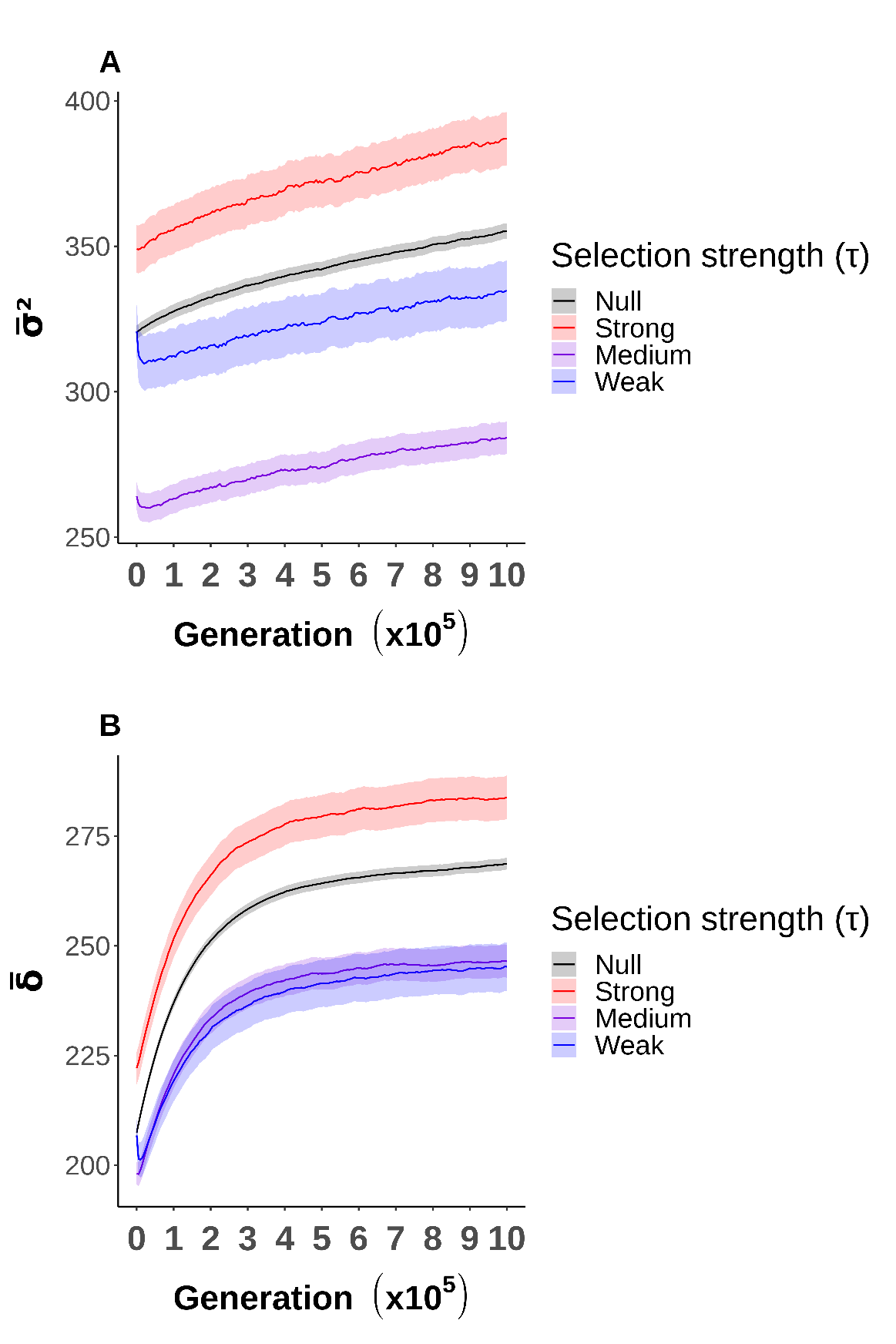


Figure 1 – Mean trait variance (A) and Euclidean distance from the optimum (B) over 100,000 generations of stabilizing selection of different strengths (). 128 total models were sampled across the spectrum of selection strengths () with an additional 512 models sampling the null space of parameters ().

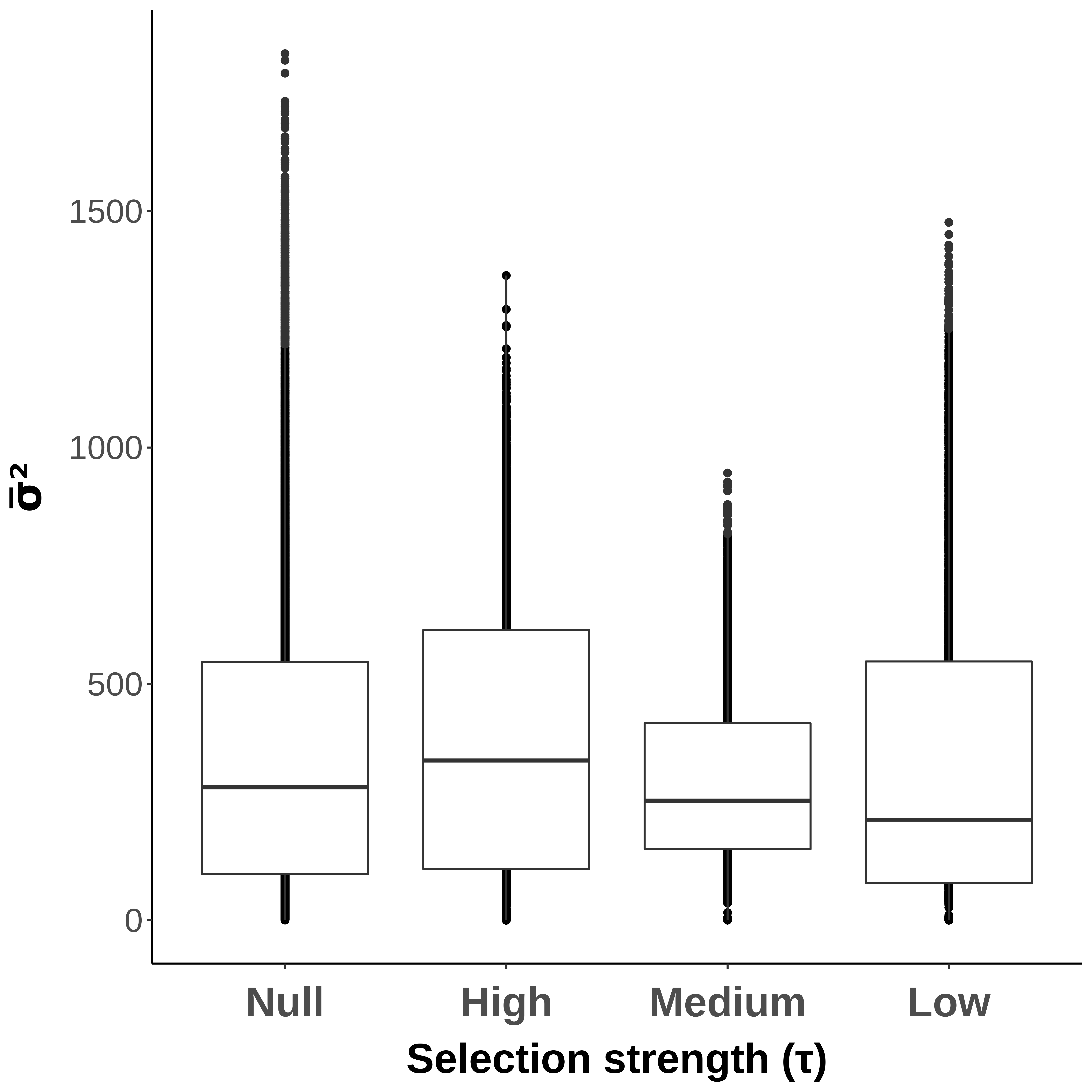


Figure 2: Mean trait variance with increasing selection strength explains very little shift in variance patterns across models. Boxplots represent variation within and between 512 null parameter combinations and 128 selection treatment parameter combinations.

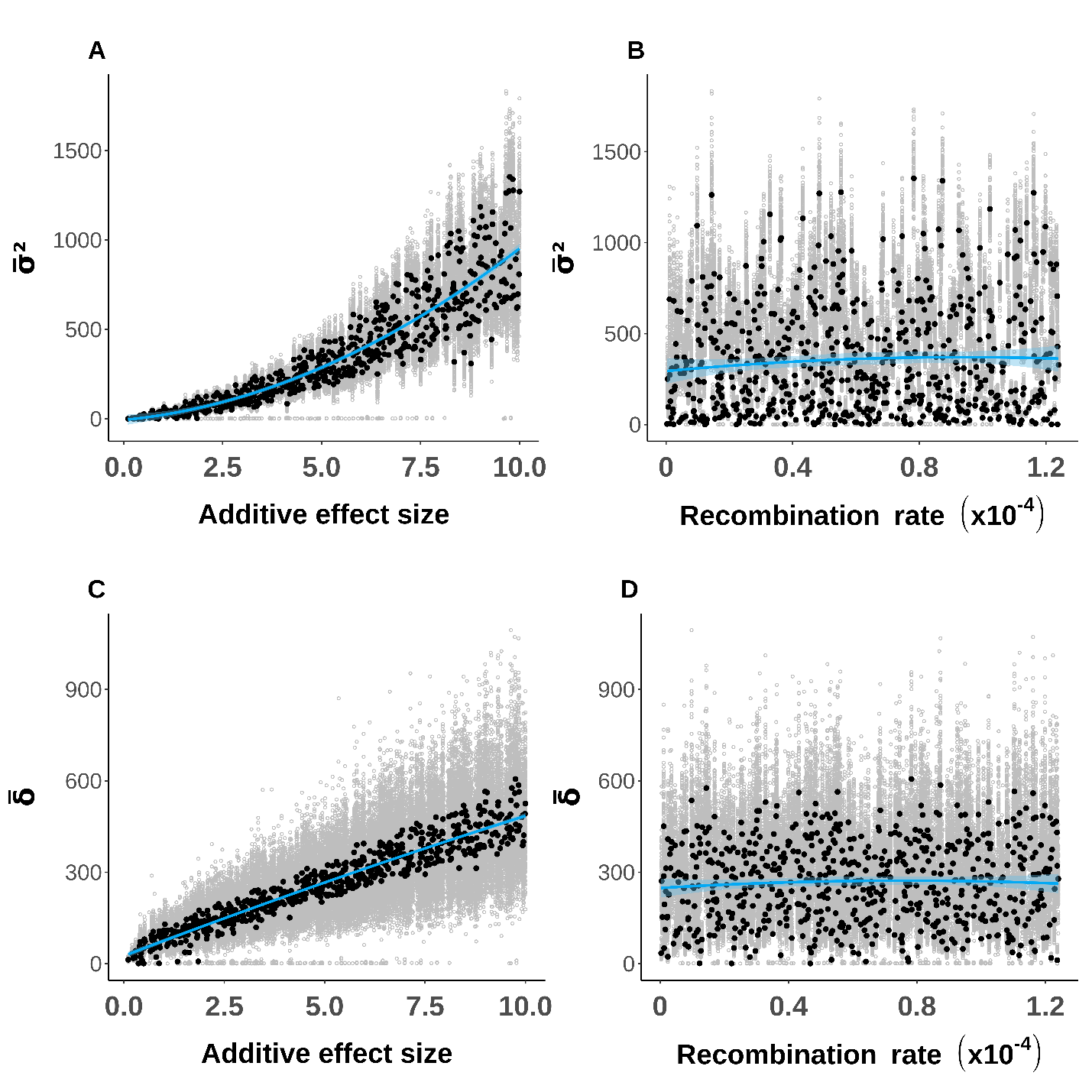


Figure 3: Mean trait variances () and Euclidean distances from the optimum () with increasing additive effect size (A and C) and per-locus recombination rate (B and D). Black dots indicate mean distances/trait variances of 100 replicates, grey dots indicate distances/variances of individual populations.

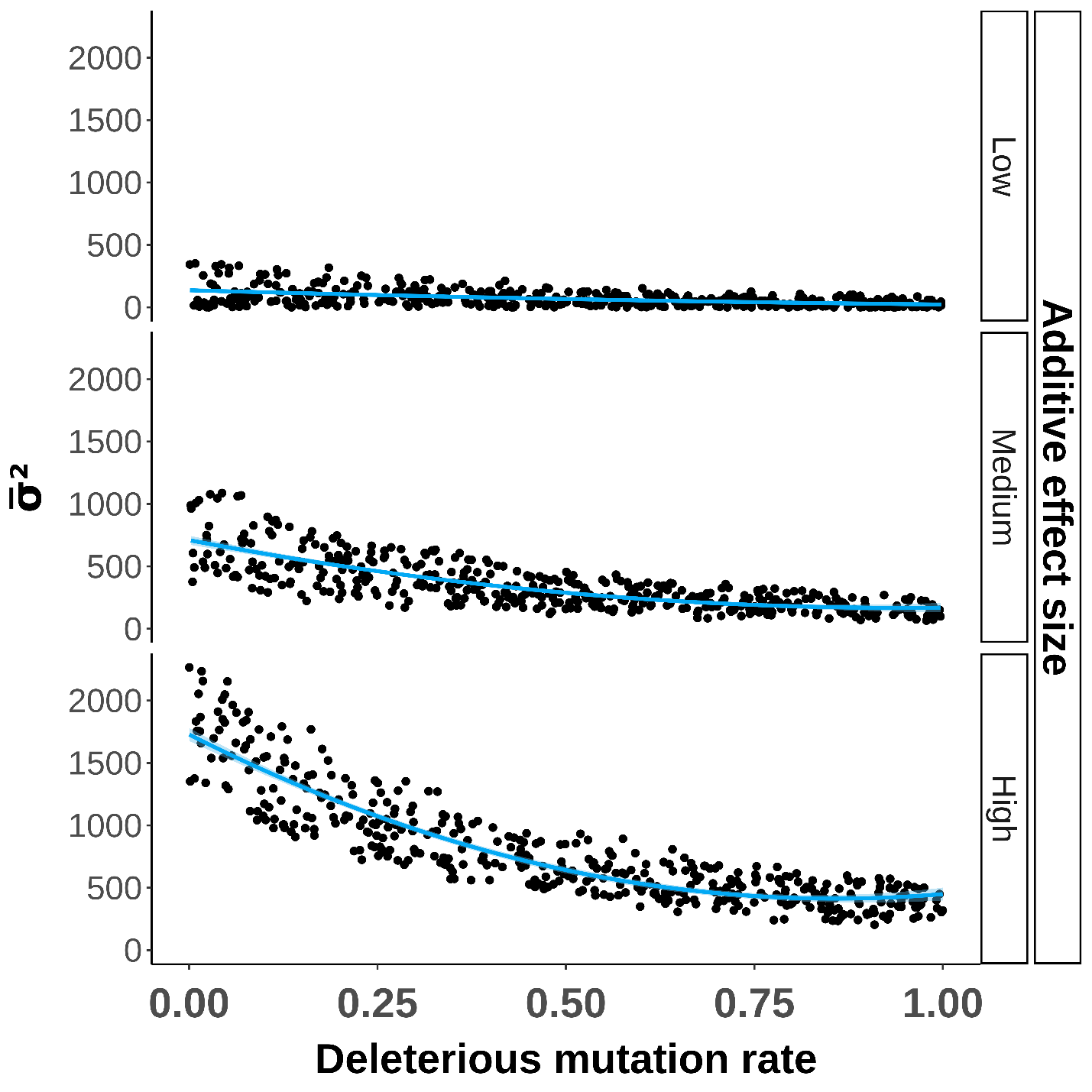


Figure 4: Mean trait variance with increasing deleterious mutation rate and additive effect size. Taken across null and selection models, for a total of 1280 means of n = 100 replicates.

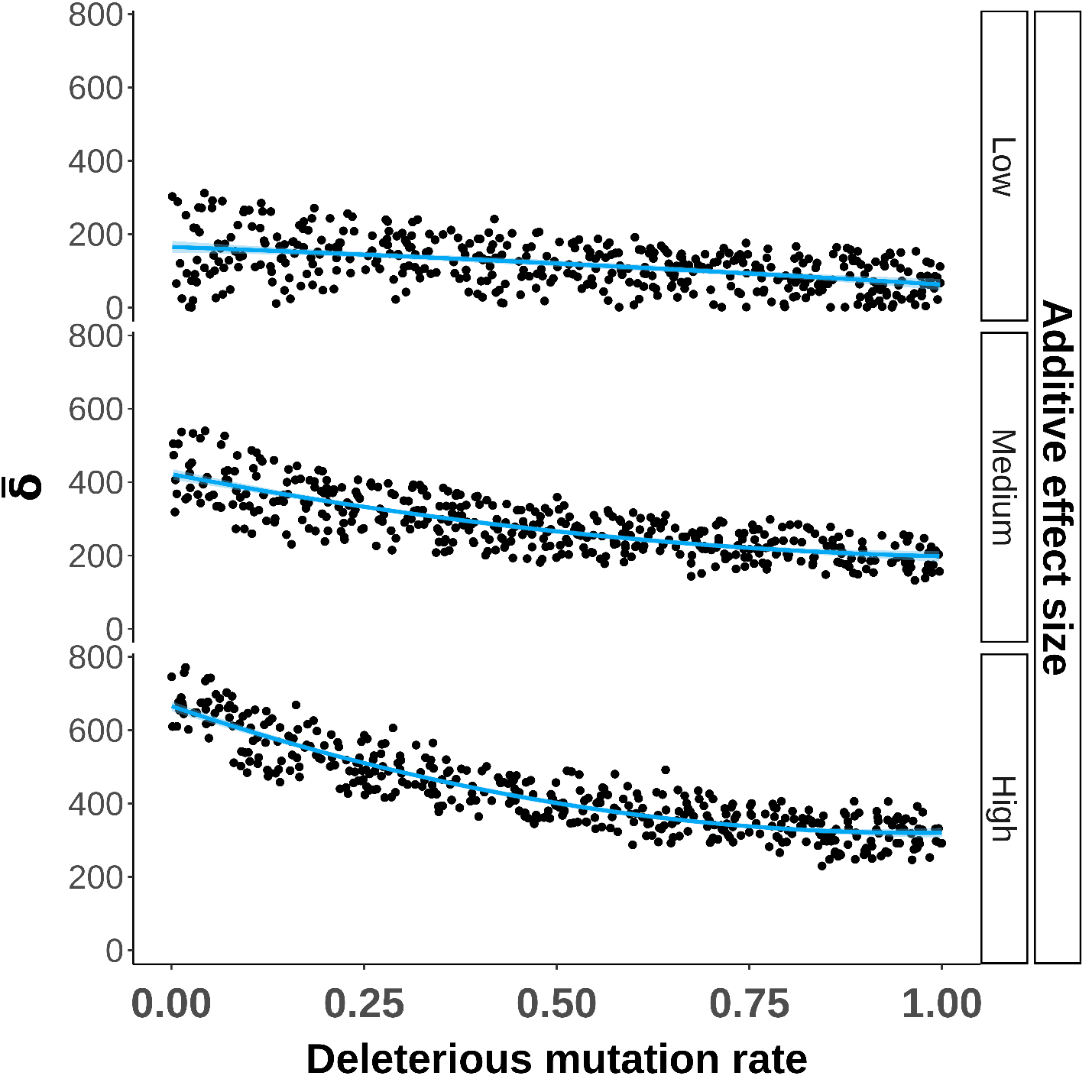


Figure 5: Mean Euclidean distance from the optimum with increasing deleterious mutation rate and additive effect size. Lower values indicate better ability of the population to maintain position at the optimum. Taken across null and selection models, for a total of 1280 means of n = 100 replicates.

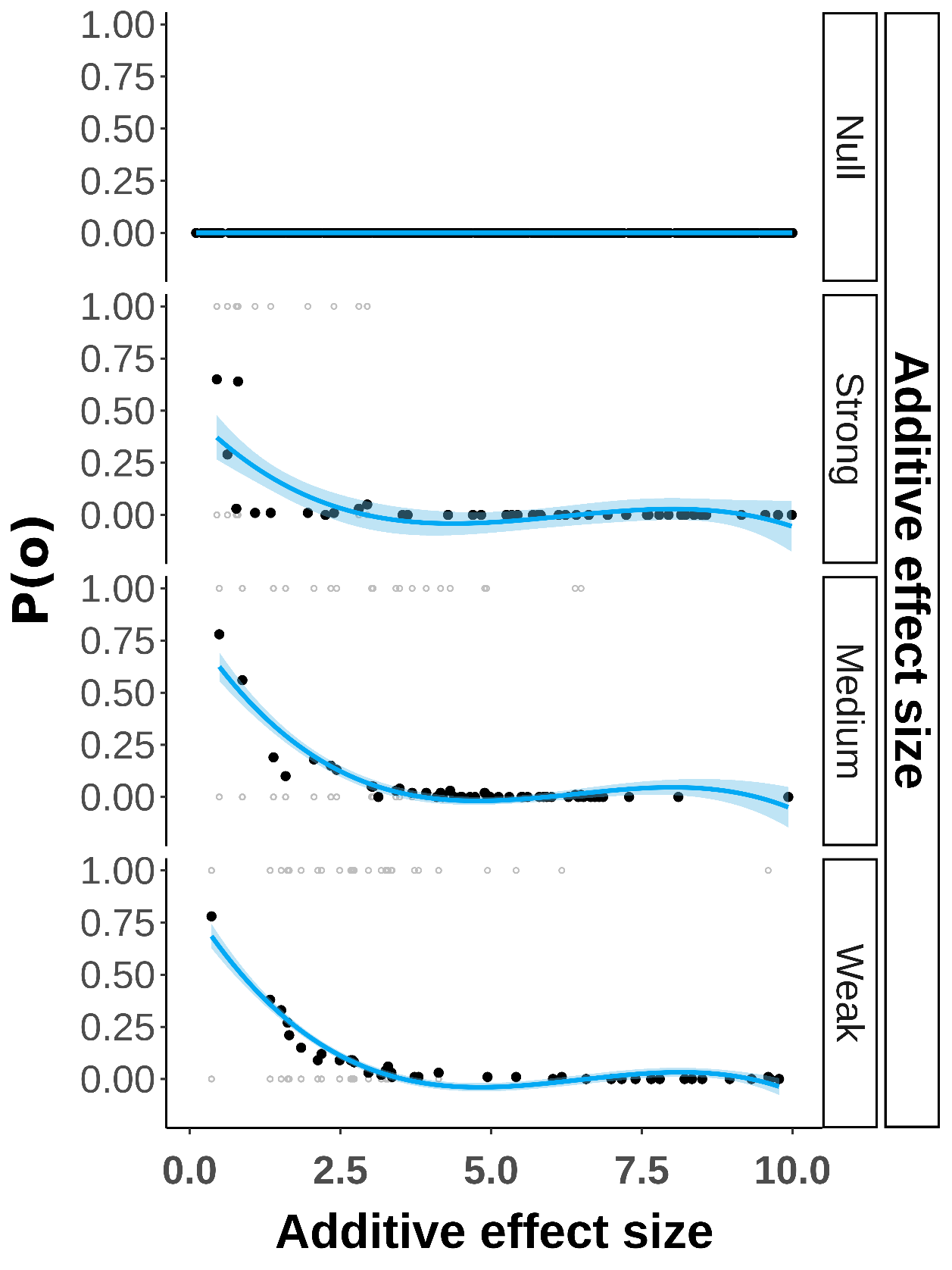


Figure 6: Probability of being at the optimum (+1 unit of tolerance) at generation 100,000 with increasing additive effect size and selection strength (. Black dots represent means of 100 replicates. Grey dots represent individual populations at (P(o) = 1) and away from (P(o) = 0) the optimum.

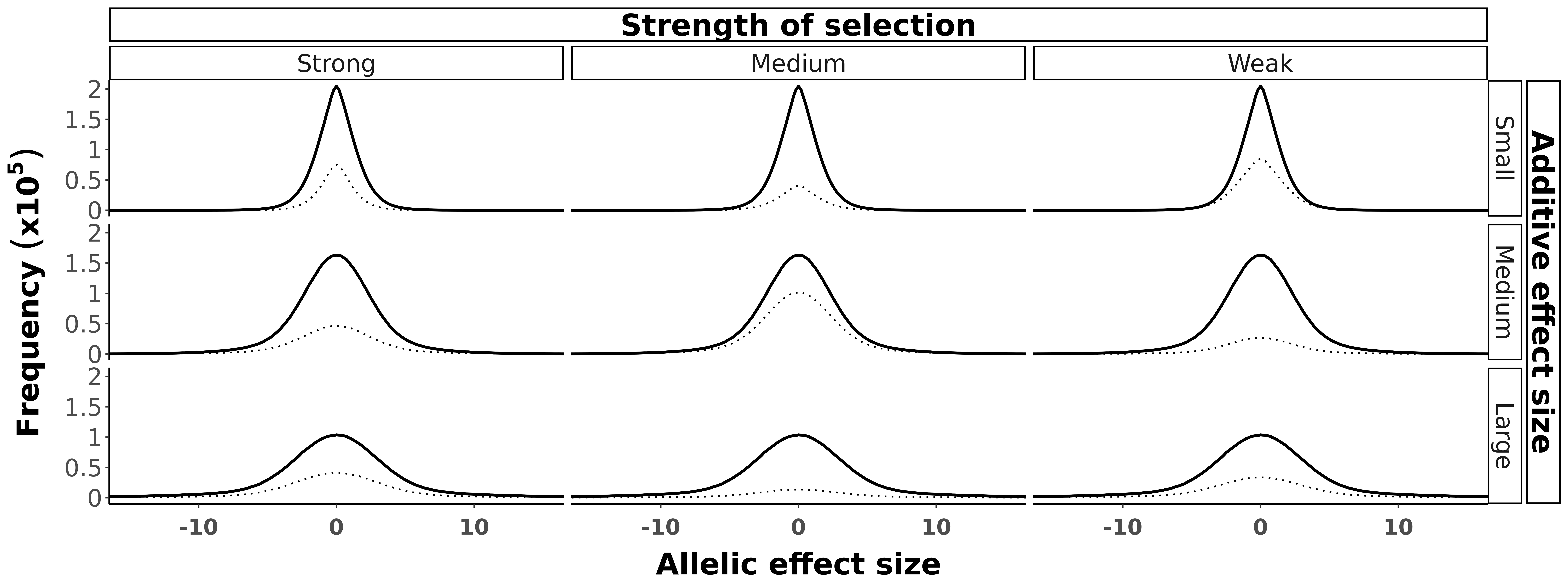


Figure 7: Frequency distribution of mutational effect sizes at generation 100,000 under no selection (solid line) and some strength of stabilizing selection (dotted line), with additive effect size. Both figures represent total distributions of 100 replicates of 128 models. 128 of the 512 null models were randomly sampled to calculate the totals of null models.

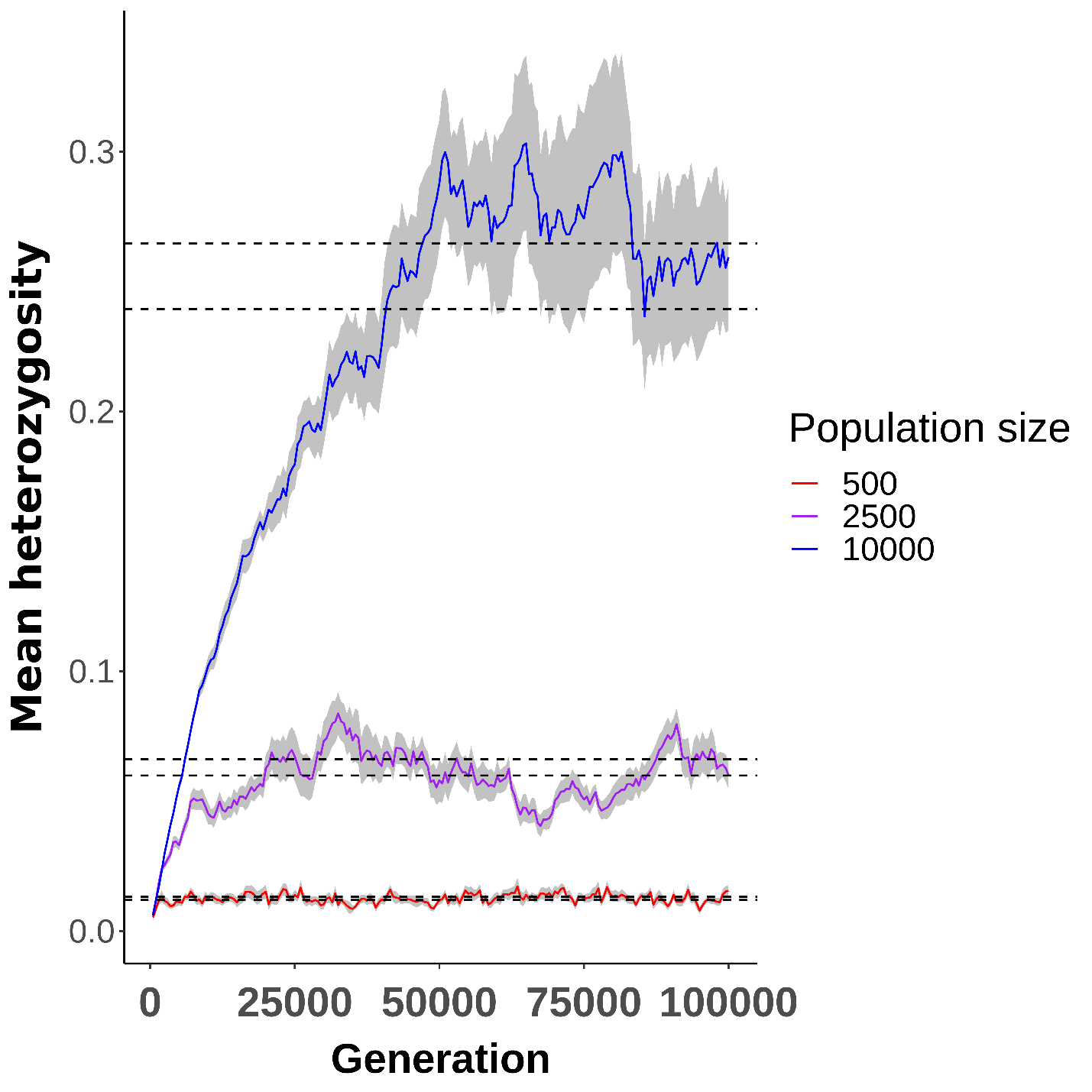


Figure S1: Mean population heterozygosity over time. Lines represent mean trajectories of 20 replicates, with ribbons representing standard errors. Dotted lines represent expected heterozygosities ± 5%, given by .

Table 1: Model parameters for both null and stabilizing selection models. The range of values is based on literature, but values are adjusted to be practical for the time of the experiment.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | Symbol | Range | Description | Source(s) |
| Genome wide recombination rate | r | 0 to 1.241x10-4 per locus | The singular recombination rate used across the entire simulated genome. | Stapley et al. 2017 |
| Background selection rate | δ | 0 to 1 | The number of non-trait, deleterious mutations that occur relative to trait mutations. |  |
| Rate of universal pleiotropy | ϖ | 0 to 0.5 | The proportion of trait mutations that affect all traits rather than a single trait. While 100 loci control a trait independently by default, this may be changed by this parameter. However ratios of loci affecting each trait will remain constant, especially across multiple replicates. | Chesmore et al. 2017; |
| Mutational pleiotropic correlation | m | 0 to 0.5 | The mutational correlation between additive effects of pleiotropic mutations determines the similarity of trait effects between traits for the same pleiotropic mutation. |  |
| Additive effect size | λ | 0.1 to 10 | Additive effect size controls the variance of trait effect size around mean 0, so that N(0, λ). | Albert et al. 2008; |
| Selection strength (selection model only) |  | 10 to 10000 | The parameter that controls the curve of the fitness function (eq. 3), with higher values resulting in a smaller difference in fitness between trait-differing individuals. |  |

# References

Agashe, D., and D. I. Bolnick, 2010 Intraspecific genetic variation and competition interact to influence niche expansion. Proc Biol Sci 277**:** 2915-2924.

Agashe, D., J. J. Falk and D. I. Bolnick, 2011 Effects of founding genetic variation on adaptation to a novel resource. Evolution 65**:** 2481-2491.

Aguirre, J. D., E. Hine, K. McGuigan and M. W. Blows, 2014 Comparing G: Multivariate analysis of genetic variation in multiple populations. Heredity 112**:** 21-29.

Aston, E., A. Channon, R. V. Belavkin, D. R. Gifford, R. Krasovec *et al.*, 2017 Critical mutation rate has an exponential dependence on population size for eukaryotic-length genomes with crossover. Sci Rep 7**:** 15519.

Barton, N. H., 2017 How does epistasis influence the response to selection? Heredity (Edinb) 118**:** 96-109.

Barton, N. H., and B. Charlesworth, 1998 Why sex and recombination? Science 281**:** 1986-1990.

Careau, V., M. E. Wolak, P. A. Carter and T. Garland, Jr., 2015 Evolution of the additive genetic variance-covariance matrix under continuous directional selection on a complex behavioural phenotype. Proc Biol Sci 282.

Charlesworth, B., and D. Charlesworth, 2010 *Elements of evolutionary genetics*. Roberts and Company, Greenwoord Village, Colorado, USA.

Charlesworth, B., M. Nordborg and D. Charlesworth, 1997 The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genetical Research 70**:** 155-174.

Eicker, F., 1967 Limit theorems for regressions with unequal and dependent errors, pp. 59-82 in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. University of California Press, Berkeley, Calif.

Falconer, D. S. M., T. F. C., 1996 *Introduction to quantitative genetics*. Pearson Education Limited, Longmans Green, Harlow, Essex, UK.

Fisher, R. A., 1930 *The genetical theory of natural selection*. The Clarendon press, Oxford, UK.

Haller, B. C., and P. W. Messer, 2019 Slim 3: Forward genetic simulations beyond the wright-fisher model. Molecular Biology and Evolution 36**:** 632-637.

Houle, D., 1998 How should we explain variation in the genetic variance of traits? Genetica 102-103**:** 241-253.

Huber, P. J., 1967 The behavior of maximum likelihood estimates under nonstandard conditions, pp. 221-233 in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*. University of California Press, Berkeley, Calif.

Kimura, M., and J. F. Crow, 1964 The number of alleles that can be maintained in a finite population. Genetics 49**:** 725-738.

Long, J. A., 2020 Jtools: Analysis and presentation of social scientific data

Lumley, T., P. Diehr, S. Emerson and L. Chen, 2002 The importance of the normality assumption in large public health data sets. Annu Rev Public Health 23**:** 151-169.

Lynch, M., and R. Lande, 1998 The critical effective size for a genetically secure population. Animal Conservation 1**:** 70-72.

Melo, D., G. Garcia, A. Hubbe, A. P. Assis and G. Marroig, 2015 EvolQG - an R package for evolutionary quantitative genetics. F1000Research 4**:** 925.

Pujol, B., and J. R. Pannell, 2008 Reduced responses to selection after species range expansion. Science 321**:** 96.

R Developmental Core Team, 2019 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Thornton, K. R., 2019 Polygenic adaptation to an environmental shift: Temporal dynamics of variation under gaussian stabilizing selection and additive effects on a single trait. Genetics 213**:** 1513-1530.

Walsh, B., and M. Lynch, 2018 *Evolution and selection of quantitative traits*. Oxford University Press, New York, NY.

White, H., 1980 A heteroskedasticity-consistent covariance-matrix estimator and a direct test for heteroskedasticity. Econometrica 48**:** 817-838.

Xu, L., H. Chen, X. Hu, R. Zhang, Z. Zhang *et al.*, 2006 Average gene length is highly conserved in prokaryotes and eukaryotes and diverges only between the two kingdoms. Molecular Biology and Evolution 23**:** 1107-1108.

Zhang, X. S., 2012 Fisher's geometrical model of fitness landscape and variance in fitness within a changing environment. Evolution 66**:** 2350-2368.