# 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 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, additive genetic variance is heritable, and hence is the component that can be

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

* Importance of genetic variability for adaptation
* Redundancy and genetic pathways to polygenic adaptation
* Additive models of quantitative genetics, Fisher’s geometric model
* Connection between quantitative and population genetics approaches (Geometric model, interaction of traditionally mutation-driven pop gen features (e.g. deleterious mutation, allele frequencies, selective sweeps) with more traits)
* Effects of deleterious mutation on genetic variability, constraining adaptation
* Recombination and linkage in the context of creating largely deleterious haplotypes with non-trait affecting deleterious mutations
* Effects of pleiotropy and the cost of complexity
* Deleterious mutation and pleiotropy as constraints on adaptation – expectations under geometric model
* Theories of adaptation in quantitative genetics – stabilising, disruptive, directional, squashed stabilising
* SLiM as a tool to computationally study these effects over long time scales
* Introduction of aims – to quantify the effects of deleterious mutation and pleiotropy on neutral evolution in an intermediate-sized population & to quantify the effects of del muts and pleio on adaptation to an intermediate optimum
* Novelty: effects of deleterious mutation and recombination on multiple traits – do they behave the same way as with just one or two traits?

# 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 amount of deleterious mutation, the rate of universal pleiotropy, the mutational correlation between trait effects from a single pleiotropic mutation, the additive effect size distribution, and the selection strength multiplier, (Table 1). Among these models, multiple conditions and assumptions are 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 (with an additional 50,000 generations of burn-in (more information in figure S1 – heterozygosity figure from burn-in test). Each individual is characterized by 8 traits, controlled by 100 loci each, unless a pleiotropic treatment is applied which will randomly reduce this by an approximately uniform amount per trait (further detail below). Each locus is assumed to have identical length, and each base pair within it is assumed to be mutationally independent. This behavior seems reasonable, as a study by Thornton (2019) 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 not too far-fetched. The mutation is modelled as occurring at an arbitrary position within the locus (or its regulatory regions) and is of arbitrary form. The effect of the mutation on chromosomal structure (e.g. effects of deletions, insertions etc.) is not explicitly modelled, but is implied via their effect on fitness and/or the trait. Mutations are assumed to be completely additive in effect, with no dominance or epistatic interactions. All loci are assumed to be on the same chromosome, with genetic distance being determined by the recombination rate parameter, r (Table 1). Both models have 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 chosen 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, 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.

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 was sufficient for our population size (FIGURE S1: Plot of heterozygosity). Deleterious mutation (δ) lowered the value of away from expectation in initial burn-in tests, however an alternative equilibrium was reached, satisfying the requirements of burn-in (Figure S1).

## 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 .

The fitness of an individual in the population was defined as:

Where s represents strength of selection, represents the gradient of the selection curve, n is the number of traits, and xn is the phenotype for trait n. For my experiments, s was fixed at s = 0.9, ensuring minimum fitness was 0.1, 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. This value differs depending on, which adjusts the realized fitness gradient.

## 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 192 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) using the runif() function in base R (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.

## Null model analysis

To compare null models generally, I compared mean variances, covariances, population means, and heterozygosities at generation 150,000 using a linear mixed effects model, with the SLiM seed value as a random effect. 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 (Lumley *et al.* 2002). This was verified with diagnostic tools in the R package “jtools” (Long 2020), along with the homoscedasticity of variance.

To compare the complete range of variance and covariance via principal components analysis (PCA), I binned each parameter into eight equidistant factor levels. From the population variances and covariances, I extracted **G** matrices and performed PCA on them, constructing **G** ellipses from the first two traits and principal components for each parameter factor level. Using five-way type-I ANOVAs, I compared the areas of each ellipse, the ratio of the major and minor axes of variation, and the angle of rotation of the ellipse around its center (the two-trait mean). I then performed post-hoc Tukey honest significant differences tests to determine which groups were significantly different. **MANOVA ?? Results were exactly the same from MANOVA vs 3 separate type-I ANOVAs**

To describe the total structure of differences between **G** matrices across all eight traits, I used relative PCA, comparing differences in variance-covariance structure both within and between levels of each predictor variable. This was done using the ‘vcvComp’ package for R (Le Maître and Mitteroecker 2019). Relative principal components analysis produces the generalized variances between two tested models, which is the product of all relative eigenvalues (Le Maître and Mitteroecker 2019). This is equivalent to the ratio of the determinants of the two covariance matrices (or models) being compared. The log generalized variance is a useful metric for comparing the magnitude of variation across all traits between the two groups (Le Maître and Mitteroecker 2019). I sampled 128 models of the total 1024 for 812,800 pairwise comparisons between models (8128 comparisons replicated 100 times), computing a relative PCA for each comparison. To reduce the noise in the data, I retained only relative PCAs between models with minimal difference in each parameter (i.e. they belonged to the same bin), and those with the maximum difference (models were in the most extreme bins, with the highest and lowest values for their parameter ranges). These two ranges defined two groups to perform statistical analysis on.

I compared mean log generalized variances for each parameter with unpaired t-tests, adjusting for multiple comparisons with Bonferroni correction. I then compared the distributions of log generalized variances between groups using Kolmogorov-Smirnov tests, and the variance between groups with a five-way type I ANOVA comparing only first-order interactions between predictor variables. To assess which interactions were significant, I performed a post-hoc Tukey honest significant differences test.

## Selection model analysis

I repeated the above analyses on the selection model, this time including the sixth predictor, , in statistical models. In addition, I computed the population mean distance from the optimum for each replicate and model, comparing these distances with another type I ANOVA.

To compare null models to selection models, eigentensor analysis was used to compare groups of **G** matrices with similar parameter values.

# Results

## Null model

I first analysed the effects of background selection, recombination rate, pleiotropic covariance, rate of pleiotropy, and additive effect size (along with their first-order interactions) on heterozygosity with a linear mixed effects model. This model explained 81.7% of total variation. I found that increasing background selection reduced heterozygosity, (β = -0.137, t p < 0.0001), recombination rate increased it (β = 155.5, p < 0.0001), and recombination alleviated some of the effects of background selection on heterozygosity via interaction (β = 122.9, p < 0.0001). Although the residuals of groups strayed from normality, variance was homoscedastic and the number of observations (102400) led to normality of coefficient estimates through the central limit theorem, providing some robustness from the effects of non-normality (Lumley *et al.* 2002).

I then analysed the effects of the parameters on variance and covariance, again using a linear mixed effects model, and choosing to focus on the variance of trait 1 and covariance between traits 1 and 2 as a proxy for all of the variance and covariance terms. I did a more complete analysis of the total variance-covariance structure with PCA and relative PCA later on. **Nonsense linear regression results – by adding other factors, delmu increases variation (?), despite the actual data not fitting that trend at all, but by excluding the other factors delmu decreases variation?**

To compare the effects of parameters on variance-covariance structure, I computed **G** ellipses of traits 1 and 2, comparing ellipse area major-minor axis ratios, and angles of rotation around the center with a MANOVA.

To compare the effects of the parameters on total variance-covariance structure, I used relative PCA, comparing pairs of models in the same bin against those the furthest apart (i.e. relative PCA between two models with very similar values for a given parameter, compared against relative PCA between two models with maximum difference in values for a given parameter). Mean values of log generalized variance were compared between these two groups for each parameter (Figure 1). Large discrepancies in deleterious mutation between pairs resulted in higher log generalized variance (t17626 = -23.433, p < 0.0001), with a mean difference of 2.539 (**log generalized variance – not sure how to interpret what this number actually means in a quantitative sense**). The same was true for rate of pleiotropy (t24740 = -5.709, p < 0.0001; = 0.61), pleiotropic covariance (t14700 = -7.017, p < 0.0001; = 0.925), and additive effect size (t17295 = -19.961, p < 0.0001; = 2.356). Recombination rate differed from the other results, with log generalized variance decreasing with increasing differences between paired models (t7275.4 = 9.405, p < 0.0001; = -1.51). As well as differences in means, distributions of log generalized variances also differed significantly.

I used two-sample Kolmogorov-Smirnov tests to assess differences between the distributions of log generalized variance between similar- and distinct-parameter pairs (Figure 2). All distributions were significantly different (Table 2). Comparisons between models with large differences in deleterious mutation frequency showed considerable increases in frequencies of log generalized variances around zero, with bimodal peaks on either side of zero (Figure 2). Comparisons between models with large differences in either pleiotropy rates or additive effect size also led to multimodal distributions, with a bottleneck effect appearing on either side of the center, with compression of the tails (Figure 2). Large differences in recombination similarly compressed the range of variation in log generalized variance between models, with a strong bottleneck appearing below zero log generalized variance. Large differences in pleiotropic covariance led to more subtle patterns, where small bottlenecks and tail compression were visible, but to a much lesser extent than with the other parameters.

, rate of pleiotropy (β = 0.012, p < 0.0001), and additive effect size (β = 0.00022, p = 0.0023) on heterozygosity, as well as significant interactions. The most biologically meaningful of these interactions included interactions between deleterious mutation and recombination rate (β = 122.9, p < 0.0001).

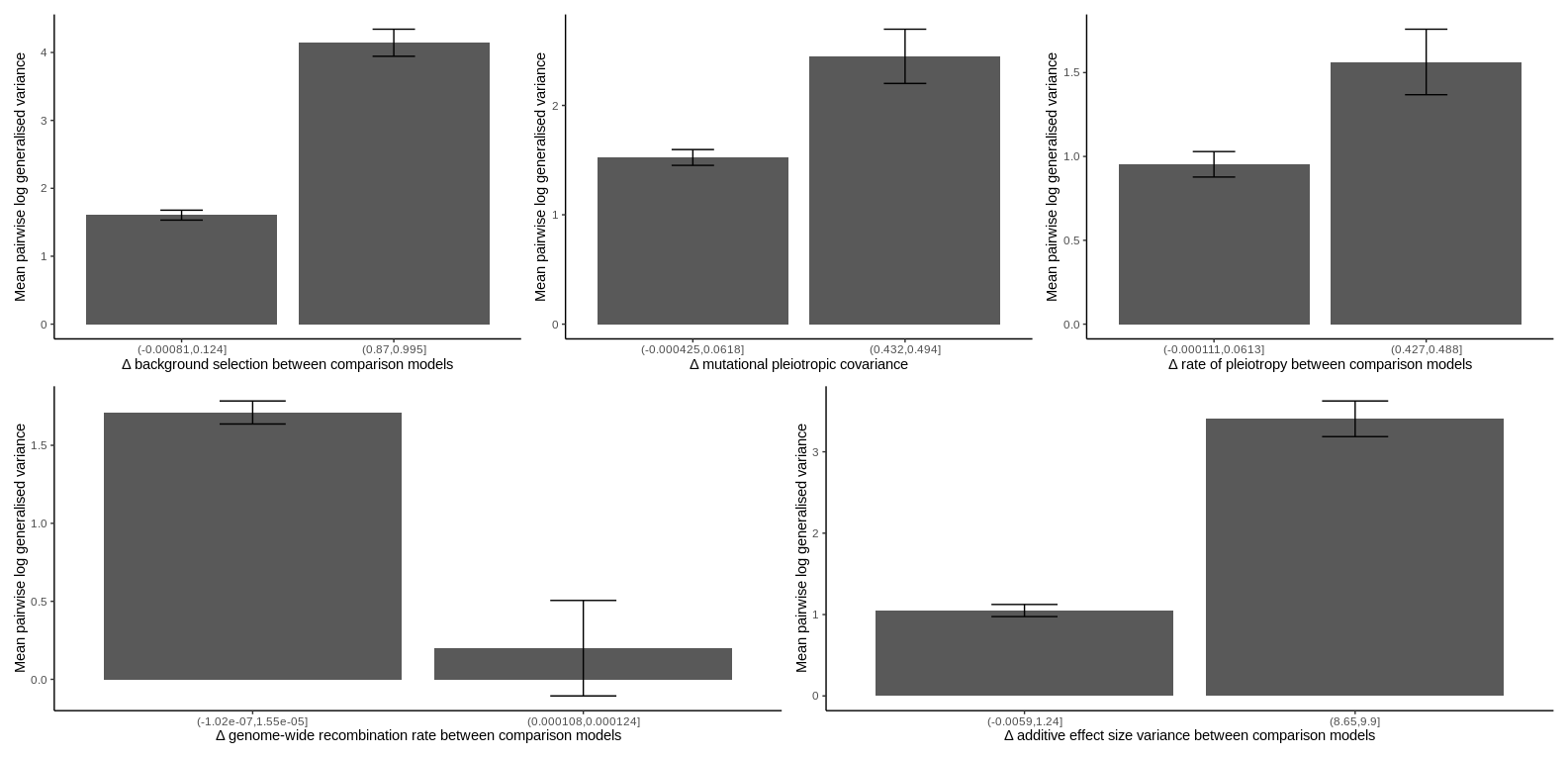
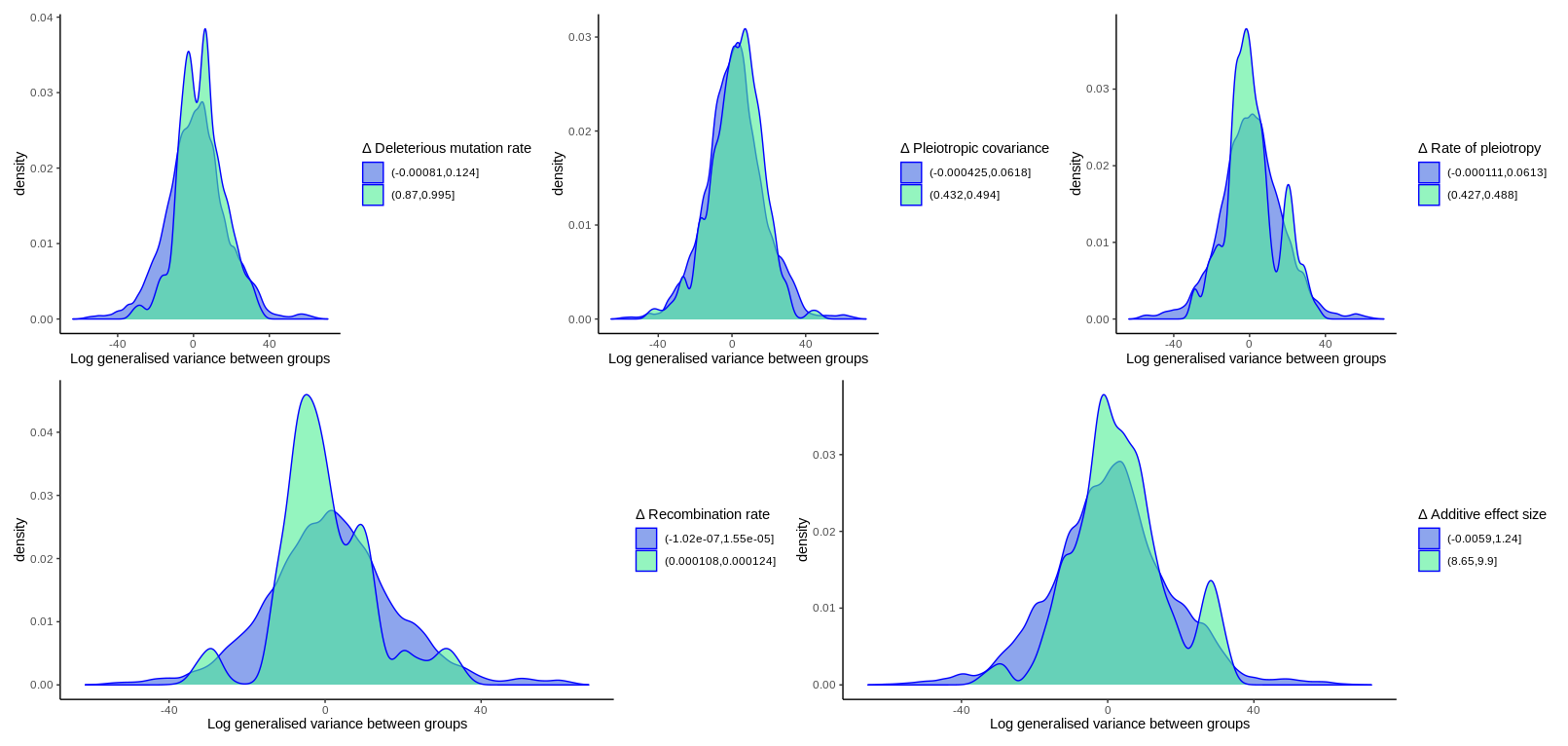
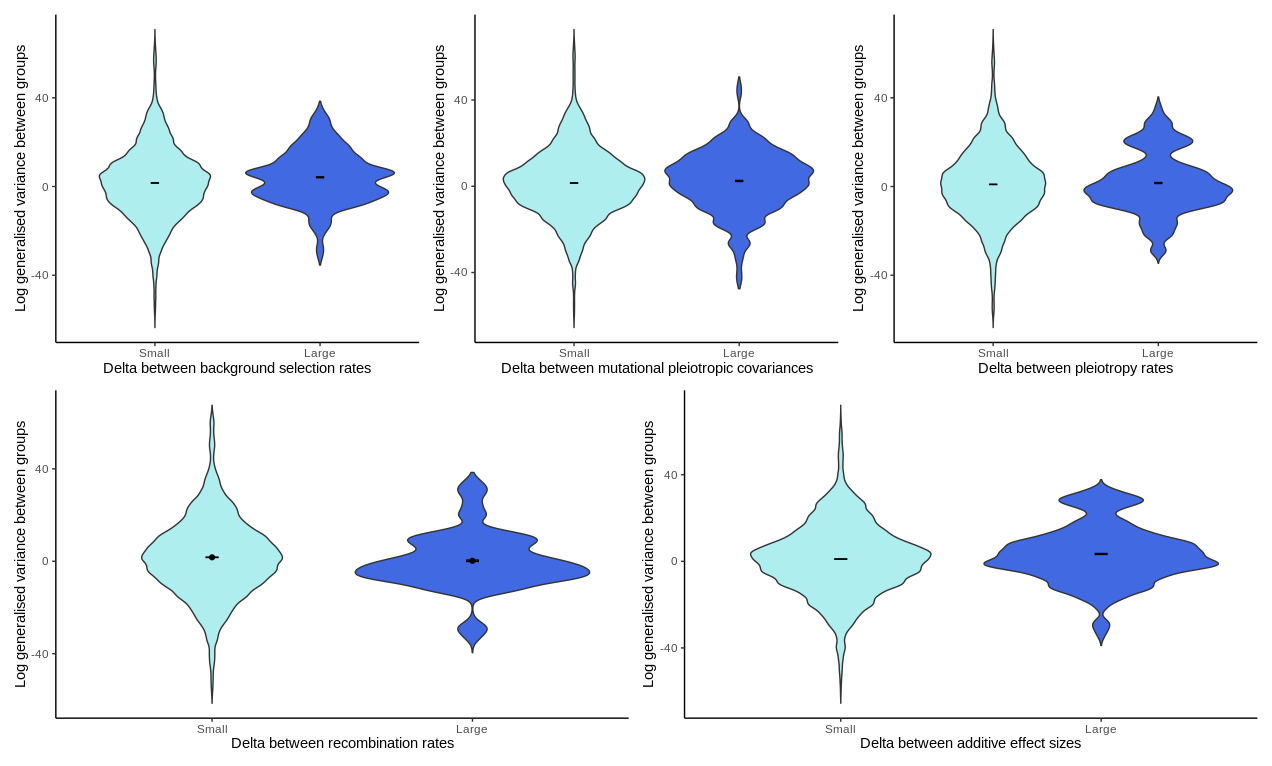
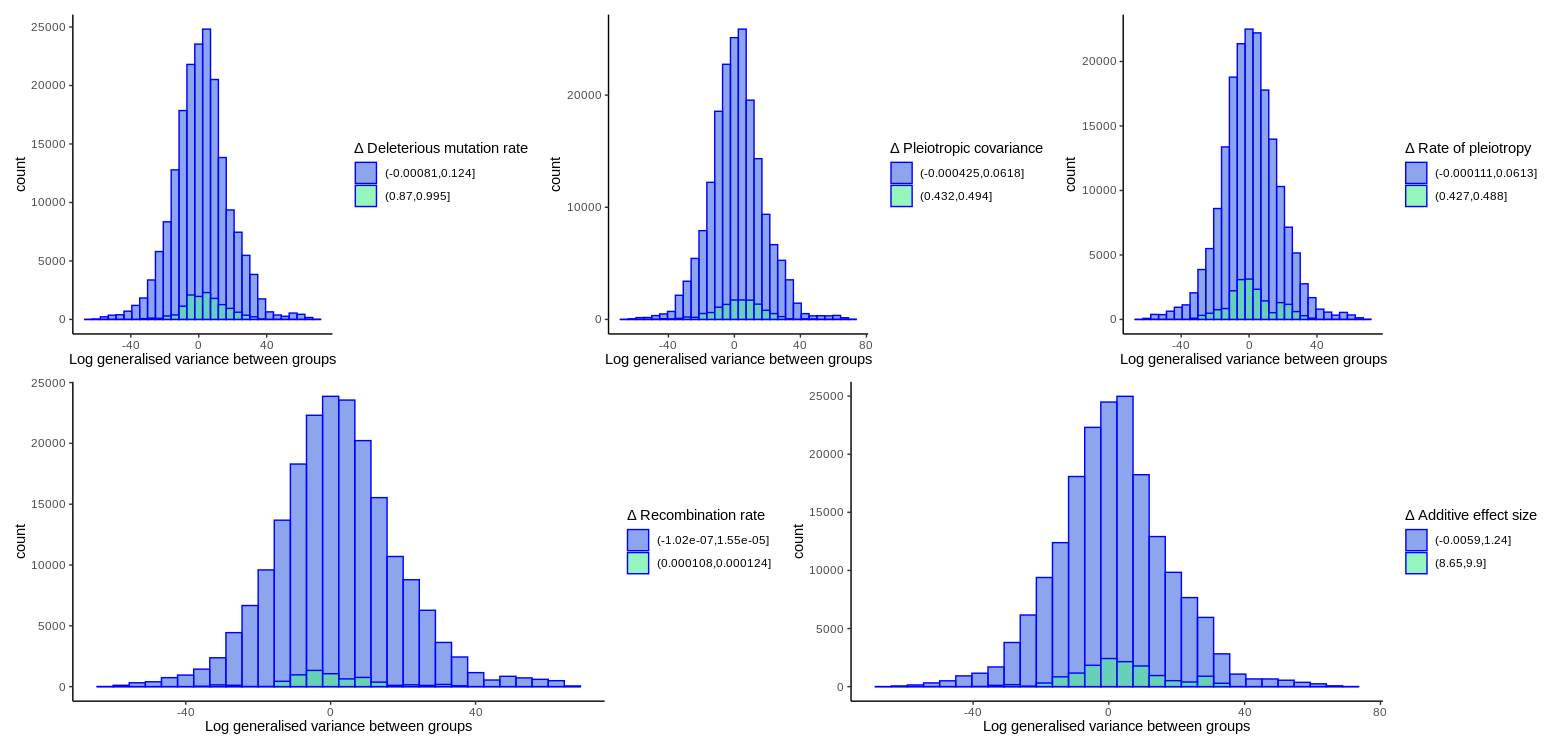


Figure 1 – Mean differences in log generalized variance between relative principal components analysis outcomes comparing similar models and distinctive models by deleterious mutation rate/background selection prevalence, mutational pleiotropic correlation, rate of pleiotropy, recombination rate, and additive effect size. All pairwise comparisons were significant.





One of these for Figure 2

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. |  |

Table 2: Test results for Kolmogorov-Smirnov tests between distributions of log generalised variance, the output of relative PCA comparing similar and distinct pairs of parameter models. All tests were two-sided.

|  |  |  |
| --- | --- | --- |
| Model parameter | D statistic | p-value |
| Deleterious mutation | 0.13168 | <0.0001 |
| Recombination rate | 0.11125 | <0.0001 |
| Pleiotropic correlation | 0.061734 | <0.0001 |
| Pleiotropy rate | 0.073615 | <0.0001 |
| Additive effect size | 0.1044 | <0.0001 |

# References

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.

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.

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.

Kimura, M., and J. F. Crow, 1964 The Number of Alleles That Can Be Maintained in a Finite Population. Genetics 49**:** 725-738.

Le Maître, A., and P. Mitteroecker, 2019 Multivariate comparison of variance in R. Methods in Ecology and Evolution 10**:** 1380-1392.

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.

R Developmental Core Team, 2019 R: A language and environment for statistical computing, pp. 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.

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. Mol Biol Evol 23**:** 1107-1108.