# Introduction

* 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

## 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). Each individual is characterized by 8 traits, which evolve under a variety of conditions. Each trait is controlled by 100 loci each, unless a pleiotropic treatment is applied which will randomly reduce this by an approximately uniform amount per trait. 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. All loci are assumed to be on the same chromosome, with genetic distance being determined by the recombination rate parameter, r. Both models have a fixed 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 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. The mutation is modelled as occurring at an arbitrary position within the locus (or its regulatory regions), and is of arbitrary form.

An effective population size of 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 are sampled from a normal distribution, N(0, λ), where λ is the model parameter additive effect size (Table 1). In the case of pleiotropy, a multivariate normal distribution is used, where n = 8, and N(0, **Σ**), where **Σ** is a covariance matrix with diagonal values equal to λ and non-diagonals pulled from a normal distribution N(mλ, 0.2mλ), where m is the parameter value of mutation correlation. **Σ** is 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 (SOURCE – Kimura 1968). 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). Following this, 100,000 generations of neutral drift or stabilizing selection follow, depending on the treatment. Neutral drift entails no change from the properties of the burn-in, whereas stabilizing selection incurs 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 is 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 ten times as fit as those infinitely far from the optimum.

## 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 models built for testing the null model (Aim 1), and 128 for the selection model (Aim 2). 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 32 bit integers using the runif() function in base R (R Developmental Core Team 2019). The array of parameter combinations and replicates was processed across 960 cores on the University of Queensland’s Tinaroo HPC system, using embedded Nimrod scripts to feed combinations of parameters and seed values to a SLiM process.

Genome wide recombination rate refers to the singular recombination rate used across the entire simulated genome.

The ratio of deleterious mutations refers to the number of non-trait mutations that occur relative to trait mutations, with a purely deleterious effect on fitness. This compounds with fitness values from phenotypes in the selection model.

The rate of universal pleiotropy refers to 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 it is unbiased between traits so ratios of loci affecting a trait will remain constant, especially across multiple replicates.

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 controls the variance of trait effect size around mean 0, so that N(0, λ).

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 | Source(s) |
| Genome wide recombination rate | r | 0 to 1.241x10-4 per locus | Stapley et al. 2017 |
| Ratio of deleterious mutations to all other mutations | δ | 0 to 1 |  |
| Rate of universal pleiotropy | ϖ | 0 to 0.5 | Chesmore et al. 2017; |
| Mutational correlation between additive effects of pleiotropic mutations | m | 0 to 0.5 |  |
| Additive effect size | λ | 0.1 to 10 | Albert et al. 2008; |
| Selection strength (selection model only) | τ | 10 to 100 ? |  |

# References

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.

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. 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.