The distribution of fitness effects during adaptive walks using a simple genetic network

Nicholas L. V. O'Brien[1,2], **Barbara Holland**[3,4], **Jan Engelstädter**[1,2,‡], **and Daniel Ortiz-Barrientos**[1,2,‡]

[1]School of the Environment, The University of Queensland, St Lucia 4072, QLD, Australia
[2]ARC Centre of Excellence for Plant Success in Nature and Agriculture, The University of Queensland, St Lucia 4072, QLD, Australia
[3]School of Natural Sciences, University of Tasmania, Hobart 7001, Tasmania, Australia
[4]ARC Centre of Excellence for Plant Success in Nature and Agriculture, University of Tasmania, Hobart 7001, Tasmania, Australia
[‡]These authors contributed equally to this work.
Correspondence: n.obrien@uq.edu.au and d.ortizbarrientos@uq.edu.au

January 30, 2024

Abstract

The tempo and mode of adaptation depends on the availability of beneficial alleles. Genetic interactions arising from gene networks can restrict this availability. However, the extent to which networks affect adaptation remains largely unknown. Current models of evolution consider additive genotype-phenotype relationships while often ignoring the contribution of gene interactions to phenotypic variance. In this study, we model a quantitative trait as the product of a simple gene regulatory network, the negative autoregulation motif. Using forward-time genetic simulations, we measure adaptive walks towards a phenotypic optimum in both additive and network models. A key expectation from adaptive walk theory is that the distribution of fitness effects of new beneficial mutations is exponential. We found that both models instead harbored distributions with fewer large-effect beneficial alleles than expected. The network model also had a complex and bimodal distribution of fitness effects among all mutations, with a considerable density at deleterious selection coefficients. This behavior is reminiscent of the cost of complexity, where correlations among traits constrain adaptation. Our results suggest that the interactions emerging from genetic networks can generate complex and multimodal distributions of fitness effects.

***Keywords**: population genetics; quantitative genetics; systems biology; adaptive walk; genotype-phenotype map*

Introduction

A lingering question in the study of adaptation concerns the distribution of effect sizes of adaptive alleles that affect trait and fitness variation in natural populations. Research across various species and traits has yielded mixed results. In some cases, natural selection favors alleles with small phenotypic effects (e.g. Visscher *et al.* 2012; Boyle *et al.* 2017; Kosheleva and Desai 2018), such as the alleles contributing to the evolution of body size in mammals (Kemper *et al.* 2012). Other times, however, alleles with larger effects are preferred, such as those found in flower color changes that affect pollinator preferences (e.g. Dorweiler *et al.* 1993; Bradshaw *et al.* 1995; Bomblies and Peichel 2022). Understanding when allelic effects of different sizes are favored is crucial to comprehend the "adaptive walk" – the metaphorical path a population takes through genetic space as it adapts to its environment (Walsh and Lynch 2018b). Each "step" of the walk represents a beneficial allele fixing in the population (i.e. reaching 100% frequency), moving it closer to phenotypic states that maximize fitness (Maynard Smith 1970; Kauffman and Levin 1987). The distribution of fitness effects (DFE) among beneficial alleles is a key feature of adaptation, as it describes the tempo and mode of the adaptive walk in phenotypic space: how many steps there are, how large each step is, and in which direction each step drives the population.

Adaptive walks and the distribution of fitness effects

The leading theory of adaptive walks was developed by Gillespie and Orr (Gillespie 1983, 1984; Orr 1998). Under this theory, adaptive walks are characterized by a sequence of mutations with diminishing effect sizes. A key prediction of the Gillespie-Orr model is that the distribution of beneficial fitness effects (DBFE) during adaptation is negative exponential (we will refer to this DBFE as exponential from here on) (Orr 1998, 2003a). This shape emerges as populations approach the optimum phenotype: large effect mutations become progressively disadvantageous due to the increased risk of "overshooting" the optimum. Gillespie recognized that extreme-value theory – a

statistical branch focused on sampling from extreme tails of distributions – could be integrated into the genetic theory of adaptation to study the DBFE. To understand this connection, consider the distribution of fitnesses among all genotypes. This distribution is analogous to the distribution of fitness effects (DFE) among all new mutations. Under normal circumstances, the wild-type genotype should have relatively high fitness, and hence it belongs to the right tail of this distribution. The wild-type is a threshold which separates the deleterious genotypes to the left of the wild-type in the distribution from the beneficial right tail. This tail is analogous to the DBFE. Extreme value theory predicts that the DBFE can be described by a family of distributions termed the extreme value distribution (EVD). Gillespie posited that adaptive steps are samples from the EVD (Gillespie 1984). Studying the shape of the EVD highlights the availability of beneficial mutations to a population and which effect sizes are likely to contribute to adaptation.

The EVD can be broadly categorized into one of three "domains of attraction", or shapes: the Gumbel, Weibull, or Fréchet family of distributions (Fig. 1; (Rokyta *et al.* 2008; Joyce *et al.* 2008; Martin and Lenormand 2008)). Gillespie argued that the Gumbel is most likely to represent empirical DFEs as it captures common distributions such as the normal, gamma, and log-normal (Gillespie 1983, 1984). The Gumbel domain describes the DBFE's exponential expectation outlined above (Orr 1998). Mutagenesis studies and spontaneous mutation experiments largely support this theoretical expectation (e.g. Orr 1998; Imhof and Schlötterer 2001; MacLean and Buckling 2009; Kassen and Bataillon 2006; Sanjuán *et al.* 2004; Vale *et al.* 2012; Eyre-Walker and Keightley 2007). However, evidence in favor of alternative domains does exist, suggesting the field is still open to exploration. Adaptive walks with Weibull-distributed DBFEs are characterized by fewer large-effect beneficial mutations compared to Gumbel EVDs (Charras-Garrido and Lezaud 2013, Fig. 1). Rokyta *et al.* (2008) observed a Weibull distribution in the ID11 ssDNA phage, hinting at an upper limit on the size of beneficial fitness effects. Further, stabilizing selection can generate a Weibull-distributed EVD by limiting the size of beneficial mutations as populations approach the optimum (Martin and Lenormand 2008). On the other hand, Fréchet EVDs have more frequent beneficial mutations (Charras-Garrido and Lezaud 2013, Fig. 1). Schenk *et al.* (2012) observed a Fréchet EVD in *Escherichia coli* adapting to antibiotics. Although it is unclear as to which conditions cause Fréchet EVDs to arise, environments which invoke strong selective pressures on populations are associated with their appearance (Neidhart *et al.* 2014; Schenk *et al.* 2012; Bank *et al.* 2014; Foll *et al.* 2014).

There are a large number of factors which can influence the DBFE and these can produce non-Gumbel EVDs. Some of these factors are environmental, such as stabilizing selection driving Weibull EVDs (Martin and Lenormand 2008). However, developmental and selective constraints also contribute to the shape of the DBFE, implicating the structure of the genotype-phenotype-fitness (GPW) map in the DBFE's shape (Connallon and Clark 2015; Eyre-Walker 2010; Salvador-Martínez *et al.* 2018).

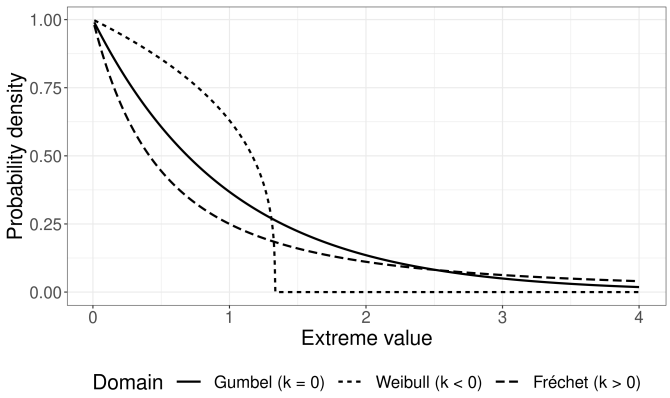


Figure 1 Examples of how the shape parameter, κ, can change the generalized Pareto distribution (GPD). The GPD, an extreme value distribution, characterizes the behavior of extreme values: random samples from the right tail of a continuous distribution. Extreme values are used to represent beneficial mutations during an adaptive walk. Depending on κ, the GPD can belong to one of three domains: Gumbel, Weibull, or Fréchet. When $\kappa = 0$, the extreme values align with the Gumbel distribution, which is the anticipated distribution for beneficial mutations during an adaptive walk (solid line). If $\kappa < 0$, the extreme values take on a Weibull distribution, exhibiting a truncated right tail (dotted line). Decreasing κ shifts the maximum extreme value (the truncation point) towards zero. Conversely, for $\kappa > 0$, the extreme values follow the Fréchet distribution, characterized by an extended right tail (dashed line). The specific κ values used in this plot are 0, -0.75, and 1 for Gumbel, Weibull, and Fréchet, respectively.

Genotype-phenotype maps in adaptation

The genotype-phenotype-fitness (GPW) map is composed of two distinct components: the genotype-phenotype (GP) and the phenotype-fitness (PW) relationships. The GP map describes how developmental and physiological processes translate genotypes into phenotypes, while the PW map describes how ecological and selection regimes create fitness differences between phenotypes. For example, stabilizing selection favors intermediate phenotypes, while directional selection favors extreme phenotypes in one direction (Simpson 1944). The PW map, along with the related genotype-fitness (GW) map, has been extensively studied in the fields of quantitative and population genetics (e.g. Wright 1931, 1932; Thornton 2019; Maynard Smith *et al.* 1985; Lande and Shannon 1996). Because the GP map is notoriously challenging to estimate, many quantitative genetics models assume an additive relationship between genotype and phenotype (Walsh and Lynch 2018a; Lande 1976). These models derive from Fisher's (1918) infinitesimal model, which supposes that continuous trait distributions can be produced by loci under Mendelian segregation as long as those loci are a) many, and b) have small, additive effects on the phenotype (Fisher 1918). This model has become the basis of modern quantitative genetics, providing a simple GP map that requires no information about the underlying genetic systems that describe the developmental and physiological under-

pinnings of traits (Falconer 1996; Walsh and Lynch 2018a).

Overwhelming empirical evidence suggests that non-additive gene interactions (epistasis) are ubiquitous in nature (Álvarez-Castro and Carlborg 2007; Hansen 2013; Ang *et al.* 2023) and play a crucial role in adaptation (Draghi *et al.* 2011; Breen *et al.* 2012; Bendixsen *et al.* 2017; Weinreich *et al.* 2006). Additive models, including the infinitesimal, often fail to capture these gene interactions, and when identified, it is unclear how they represent biological gene action (Huang and Mackay 2016). Hence, it is unknown how functional epistasis might contribute to the developmental constraints that limit evolution (Hansen 2013).

Integrating information about the underlying systems describing trait development enables us to capture genetic interactions affecting fitness. Modeling the molecular networks underpinning trait development and expression can help us understand how gene interactions and regulatory processes shape the survival and reproductive success of organisms, and provide a mechanistic view of how variation arises in natural populations. As we delve deeper into the nature of the genotype-phenotype-fitness landscape, it becomes important to re-evaluate our foundational quantitative genetic models in light of our current understanding of the molecular basis of traits. For instance, how does the distribution of fitness effects change when non-additive GP maps are considered? Does this affect a population's chance to adapt to a changing environment? To address these questions, we consider a simple gene regulatory network motif to motivate our approach and implement a nonlinear GP map into adaptive walk theory.

At the foundation of any trait lies a complex network of interacting genes and regulatory elements that respond to environmental cues by producing specific proteins (Doane and Elemento 2017). Systems biologists employ mathematical networks known as gene regulatory networks (GRNs) to model such systems (Alon 2019). Empirical networks harbor startling complexity: for example, consider the circadian clock network in *Arabidopsis thaliana*. This network is driven by a number of interacting network motifs: small, common subnetworks with particular effects on expression behavior. For instance, a feedforward loop motif in this network generates pulses of expression in *PRR9/7* (Joanito *et al.* 2018). In addition to this behavior, a negative feedback loop (another motif) between *CCA1/LHY* and *PRR5/TOC1* generates a bistable switch (Joanito *et al.* 2018). This switch is toggled at daybreak by the aforementioned *PRR9/7* feedforward loop and again at dusk by yet another circuit (Joanito *et al.* 2018). Given the complexity of many GRNs, it is common to study motifs as separate units to elucidate the reasons for their recruitment into so many systems (Alon 2007; Hallinan and Jackway 2005; Seo *et al.* 2009). This allows systems biologists to probe the general effects of motifs on a broad range of networks. In this study, we focus on the simplest network motif, negative autoregulation (NAR; Fig. 2). This motif consists of two genes: gene X activates gene Z (indicated by the pointed arrow in Fig. 2A), while the presence of Z's product limits its further expression (indicated by the flat arrow in Fig. 2A).

NAR motifs are highly prevalent in biological networks. For example, close to half of all transcription factors in *E. coli* are negatively autoregulated (Stewart *et al.* 2013; Shen-Orr *et al.* 2002). In plants, expression of the oil biosynthesis transcription factor *WRINKLE1* (*WRI1*) is driven via a NAR network that has been evolutionarily conserved from at least the split between monocotyledons and dicotyledons (Snell *et al.* 2019). Snell *et al.* (2019) found that *WRI1* expression was limited by the concentration of WRI1 protein *in planta*. The ubiquity of the NAR motif in nature comes as a result of its self-balancing property: when a gene is overexpressed, the NAR mechanism triggers to reduce that gene's production, leading to a steady state of gene expression (Alon 2019). This reduces variability in gene expression between cells and accelerates transcriptional responses to environmental cues (Rosenfeld *et al.* 2002; Nevozhay *et al.* 2009). Owing to its simplicity and ubiquity, we consider the NAR motif a reasonable toy model for beginning to explore the evolution of complex traits mediated by genetic networks.

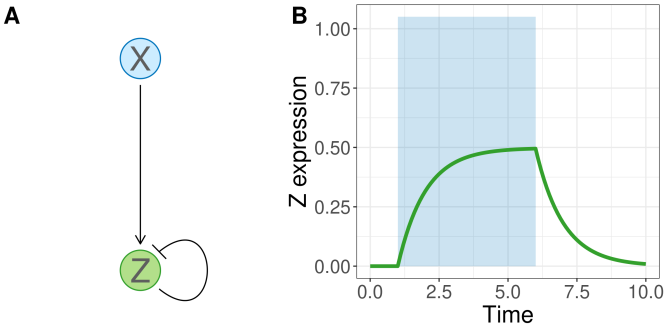


Figure 2 A negative autoregulation (NAR) motif and its expression curve. The NAR motif consists of two genes, X and Z. The expression of X activates Z, and the expression of Z inhibits further Z production (A). This results in a characteristic expression curve (B). Z production begins when X is activated (blue shaded area) and stops when it is inactivated. Gene Z product approaches a steady state concentration and then quickly falls off in the absence of X.

How might we expect networks to affect adaptation? Genetic networks impose functional epistasis, biological interactions between genes, which can create nonlinear GP maps (Hether and Hohenlohe 2014). This results in rugged fitness landscapes where the presence of multiple local optima might obstruct adaptation to a global fitness peak (Romero and Arnold 2009). In other words, populations can get trapped at local optima because the path towards the global peak (which maximizes fitness) is paved by low-fitness intermediate genotypes (Wright 1931). Adaptation should then be limited by the structure of the fitness landscape: the more local peaks there are, the less likely it is for a population to be able to find the global optimum. In turn, the structure of the fitness landscape will depend on the nature of the underlying network. For instance, the NAR network provides a relatively simple fitness landscape. Kozuch *et al.* (2020) investigated the fitness landscape of *E. coli*'s *lexA* NAR network. The authors found that the fitness landscape was ridge-like, with fitness maximized along a pa-

rameter space that balanced *lexA* production with the strength of autoregulation (Kozuch *et al.* 2020). However, other motifs might produce different constraints on adaptation. Recent work by Baier *et al.* (2023) found that a synthetic gene network based on a feedforward motif produced reciprocal sign epistasis, a prerequisite for a rugged fitness landscape. Further, empirical fitness landscapes suggest that while ruggedness is common (although not ubiquitous) (de Visser and Krug 2014), connectivity is also high, making these rugged landscapes searchable (Gavrilets 1997).

In this paper, we use a novel approach to model a quantitative trait as the product of a NAR motif. Previous attempts to reconcile quantitative and population genetics with systems biology have used a variety of approaches. Some have focused on the explicit modeling of network structures (e.g. Wagner 1996; François 2014), whilst others have considered systems as hierarchical structures of developmental parameters that resemble quantitative traits (e.g. Slatkin 1987; Rice 1998). Our approach is the first to our knowledge to combine both approaches. We model a quantitative trait as the expression of the NAR motif via a system of ordinary differential equations (ODE) similarly to François (François 2014). However, instead of modeling the addition/removal of nodes, we instead consider the perturbation of expression dynamics via quantitative changes in the coefficients of the ODE/s, similarly to Slatkin (Slatkin 1987). This approach combines the biological realism of network modeling with the wealth of existing tools for studies of the evolution of quantitative traits. We use our approach to evaluate the shape of the DBFE during adaptive walks, finding stark differences from classical expectations in some cases. We use Wright-Fisher simulations to describe the adaptive walks of populations with either an additive or NAR GP map, and examine which extreme value distribution domain (Gumbel, Weibull, or Fréchet) best captures the behavior of the proposed systems model of adaptation. We then identify if the network imposes constraints on adaptation compared to an additive model. Finally, we discuss the contributions of genetic network architectures to adaptation and propose avenues for further exploration using similar systems models.

Materials and methods

The Model

Modeling the NAR motif

To model the expression patterns of the NAR motif, we first translate its network diagram (Figure 2A) to a system of ordinary differential equations (ODE). ODEs are commonly used to model gene networks due to their balance between efficiency and realism (Schlitt and Brazma 2007; Karlebach and Shamir 2008). The solution to the ODE predicts gene expression over a time period, such as during cell development. The NAR ODE is given by:

$$Z'(t) = \beta_Z \frac{X^{n_{XZ}}}{K_{XZ}^{n_{XZ}} + X^{n_{XZ}}} \frac{K_Z^{n_Z}}{K_Z^{n_Z} + Z^{n_Z}} - \alpha_Z Z \qquad \textbf{(1)}$$

The coefficients in this equation have biological relevance. X and Z represent the cellular concentrations of the two genes. n_{XZ} and n_Z are steepness coefficients (Hill coefficients) reflecting the sensitivity of the system to the presence of X and Z products. Higher values indicate a more rapid activation (n_{XZ}) or repression (n_Z) response as X or Z concentration increases (Alon 2019). α_Z is the rate at which Z is removed from the cell. This might reflect, for example, the activity of ubiquitinase in tagging Z for removal and the 26S proteasome in breaking down the protein (Morimoto *et al.* 2016; Schmidtke *et al.* 2006). β_Z represents the production rate of Z, which is influenced by factors such as transcription factor binding affinity, enhancers, silencers, and *trans*-acting regulatory elements (Halfon 2020). These biological interpretations provide some realism to modeling quantitative traits. It also gives interpretability to the action of mutations that affect these coefficients, which will be explained in the sections following.

X follows a step function and is only expressed within the interval $t \in [t_{start}, t_{stop}]$:

$$X(t) = \begin{cases} 1, & \text{if } t_{start} \leq t \leq t_{stop}, \\ 0, & \text{otherwise} \end{cases} \qquad \textbf{(2)}$$

where t_{start} and t_{stop} are the time points at which X is activated and deactivated. Values for these parameters are given below.

The solution of the ODE is characterized by a nonlinear approach to maximum Z expression, followed by a decline after X expression ceases (Figure 2B). The area underneath the expression curve (Figure 2B) is the total amount of Z produced, which we can take as a quantitative trait value. To create variation in this trait value, the coefficients of Eq. 1 can be varied. We refer to these coefficients as "molecular components". In this study, we focus on modeling mutations in two of the NAR's molecular components: the Z degradation rate, α_Z, and the Z production rate, β_Z. The constant values of the other molecular components (K_Z, K_{XZ}, n_Z, and n_{XZ}) can be found in Supp. Table 1. Mutations in our model have direct effects on either α_Z or β_Z. This leads to a differently-shaped expression curve when the ODE is solved, and hence a (potentially) different trait value.

We adopt a model where loci contribute multiplicatively to the values of molecular components. We refer to these loci as molecular quantitative trait loci (mQTLs). The multiplicative transformation ensures that the molecular components are always positive, which is essential as these values represent rates which cannot be negative. The evolution of these mQTLs can be studied using the Wright-Fisher model, a foundational model in population genetics that describes the stochastic process of allele frequency change in a finite population (Wright 1931; Fisher 1930). In our implementation, individuals in the Wright-Fisher population are diploid.

For a more abstract model that applies to any network and any number of traits with any number of mQTLs, refer to S1 Appendix.

Phenotype calculation

To calculate an individual's phenotype from their set of mQTLs, we follow the below algorithm:

1. Take the exponent of the sum of all alleles for α_Z across all loci and both homologous chromosomes (this is analogous to an additive model summing effects across all loci, but with a multiplicative transformation as described above).

2. Repeat step 1 for β_Z.

3. Substitute the α_Z and β_Z values into the ODE, Eq. 1, and solve the ODE between time points 0 and 10 to get an expression curve.

4. Take the area under the expression curve to get the total amount of Z expression (the phenotype).

5. Repeat steps 1-4 for each individual in the population.

This algorithm is described mathematically below.
Let $\mathbf{C}$ be a vector of the molecular components:

$$\mathbf{C} = (\alpha_Z, \beta_Z)$$

An allele at locus i and chromosome j has an effect a^{ij} on one of the molecular components.
Given an individual's alleles, the value for each molecular component is calculated by exponentiation of the sum of allelic effects across all mQTLs and chromosomes. For molecular component k:

$$C_k = \exp\left(\sum_{i=1}^{L_Q}\sum_{j=1}^{2} a_k^{ij}\right), \tag{3}$$

where $C_k \in [0, \infty)$. L_Q represents the number of causal loci along the genome. This treatment of allelic effects assumes no explicit epistasis or dominance deviations, however these might arise as consequences of the nonlinear ODE solution.
After calculating the molecular component values ($\mathbf{C}$), we can determine the amount of Z expression, the phenotype. We achieve this by constructing a system of ODEs (Eq. 1) and solving for the area under the expression curve (i.e. the total Z produced during a time period).
We can express the ODE as an initial value problem where given the "initial" or starting value of the function we can determine the function's behavior for subsequent time points using the ODE:

$$Z'(t) = F(Z(t))$$
$$Z(0) = Z_0$$

where F is the function that defines the differential equation. To calculate Z, we integrate the expression curve:

$$Z = \int_0^{t_{\max}} Z(t)dt \tag{4}$$

where $t_{\max} = 10$.
In Eq. 2, $t_{start} = 1$ and $t_{stop} = 6$. Z activity is evaluated for $0 \leq t \leq 10$, with $t_{start} \leq t \leq t_{stop}$ representing 50% of the total measured time.

Fitness calculation

To map the trait value (P) to fitness (w), we utilize a Gaussian fitness function, which allows us to model directional and stabilizing selection. The Gaussian fitness function, originally described by Lande (Lande 1976), is defined as follows:

$$w(P) = \exp\left(\frac{-\Delta z^2}{2\sigma_w^2}\right) \tag{5}$$

In this function, $\Delta z = P - P_O$, where P_O is the optimal phenotype, and σ_w is the width of the fitness function. Wider functions represent weaker selection.

Evolutionary simulation

To model the evolution of the network over multiple generations, we employ a Wright-Fisher (WF) model (Fisher 1923; Wright 1931). We consider a diploid population of $N = 5000$ individuals with random mating and non-overlapping generations. This satisfies the requirement that $2N\mu \ll 1$ (where $2N$ is the total number of genomes in the population; $2N\mu = 0.0915$), which is necessary to meet the Gillespie-Orr model's assumption of strong selection relative to mutation (see section Model validation for more information on this) (Gillespie 1983; Orr 2010). Offspring are generated by sampling with replacement from parents, with the sampling probability weighted by their fitness. Individuals possess a pair of homologous chromosomes with two mQTLs: one for each molecular component. Random mutations occur at a rate of $\mu = 9.1528 \times 10^{-6}$ per locus per generation. This mutation rate is based on the average mutation rates observed in *A. thaliana* and adjusted to per-locus rates by multiplying it with the average length of a eukaryotic gene (1346bp) (Aston *et al.* 2017; Xu *et al.* 2006). *A. thaliana* was chosen as it is a model plant species with good estimates of mutation rates. The mutational effects of mQTLs on molecular components were drawn from a standard normal distribution and the effect added to the previous allele at that locus such that $a_{\text{new}} \sim \mathcal{N}(a_{\text{old}}, 1)$, where a_{new} is the new allele at a given locus and a_{old} is the previous allelic effect at that locus. After being sampled, the allelic effects were exponentiated as per Eq. 3. We assume free recombination between loci.
We also considered an additive model. In this scenario, genomes also consisted of two causal loci to match the mutational target size of the NAR model. However, the trait value was given by summing allelic effects across QTLs instead of solving an ODE:

$$P_{\text{Additive}} = \sum_{i=1}^{L_Q}\sum_{j=1}^{2} a_k^{ij}, \tag{6}$$

where a_k^{ij} is the allelic effect of locus i on chromosome j.

Computational implementation

To investigate our model's behavior during an adaptive walk, we implemented the model in a custom version of the forward-time, individual-based simulation software SLiM 3.7.1 (Haller and Messer 2019). Our modified SLiM implementation can be accessed openly at `https://github.com/nobrien97/SLiM/releases/tag/AdaptiveWalks2023`. SLiM

scripts are available at https://github.com/nobrien97/NARAdaptiveWalk2023. A flowchart of the SLiM implementation is given in Supp. Fig. 1. The ODE was solved by integrating the Ascent numerical solution library (Berry *et al.* 2021) into SLiM and using a custom C++ function to call it. Ascent solved Eq. 4 to produce trait values for the individuals in the population. Since individuals often shared genotypes (and hence had the same ODE inputs), ODE solutions were cached in memory to reduce redundant solutions and improve performance.

Both models underwent 50,000 generations of burn-in. The burn-in ensured that populations had high fitness just prior to the optimum shift, an assumption of the Gillespie-Orr model (Gillespie 1983, 1984; Orr 1998). During burn-in, populations adapted to a phenotypic optimum at $P_O = 1$. The phenotypic optimum was then instantaneously shifted to $P_O = 2$, and the population was monitored for 10,000 generations of adaptation. We measured the phenotypic means and allelic effects of segregating and fixed mutations every 50 generations. We replicated each model 2,880 times with 32-bit integer seeds sampled from a uniform distribution in R 4.3.1 (R Core Team 2023). Detailed simulation parameters can be found in Table 1. Simulations were executed on the National Computational Infrastructure's Gadi HPC system.

Table 1. Simulation parameters. Table of symbols, names, descriptions, and values for relevant parameters used in the forward-time Wright-Fisher simulations.

Of the simulation parameters, the most important were the fitness function width, σ_w, and the optimum shift amount, O_{shift}. σ_w and O_{shift} are important because the Gillespie-Orr model assumes that the wild-type has relatively high fitness, so that in the total space of genotypes, it exists on the right-tail of the distribution (Orr 1998). Hence, the fitness function and optimum shift needs to be chosen so that at the optimum shift (generation 50,000) fitness is still relatively high. We chose σ_w and O_{shift} so that this assumption is met: at the optimum shift, individuals perfectly adapted to the burn-in optimum suffered a $\sim 5\%$ drop in fitness.

Model validation

The Gillespie-Orr model assumes that there is strong per-locus selection relative to mutation (Gillespie 1983, 1984). This regime is often referred to as the strong selection weak mutation (SSWM) paradigm, as opposed to the weak selection strong mutation regimes found in polygenic models (e.g. the infinitesimal model) (Fisher 1918; Barghi *et al.* 2020).

To ensure we were in the SSWM domain assumed by adaptive walk theory, we measured population heterozygosity and the effects of segregating alleles on trait variation. To measure heterozygosity, H_O, we used the equation

$$H_O^i = N_{het}^i / N \tag{7}$$

where N_{het} is the number of individuals heterozygous at locus i, and N is the total population size. Since we had two genes in our simulations (one for each molecular component, and two for the additive model), we took the average heterozygosity across both. To measure the effects of segregating alleles on the trait, we measured the ratio between the phenotype created by only fixations and the population mean phenotype,

$$r_{fs} = P_{Fixed} / \bar{P} \tag{8}$$

where P_{Fixed} is the phenotype due to only fixed alleles and $\bar{P}$ is the mean population phenotype. P_{Fixed} was calculated by removing any segregating effects from the α_Z and β_Z values and recalculating the phenotype via Eq. 4.

Fitness effect calculation

An important part of adaptive walk theory is the shape of the DFE. To assess this, we needed to measure the fitness effects of mutations that arose during the adaptive walk. We performed single gene knockouts and measured the difference between the "wild-type" and knockout phenotypes and fitness. A flowchart is given in Supp. Fig. 2A. To measure a given mutation's effect on fitness, we first subtracted that mutation's effect on the molecular component (α_Z or β_Z) and recalculated the trait value via Eq. 4. We then calculated the fitness of this knocked-out individual via Eq. 5. To get our final fitness effect, we subtracted the fitness of the individual with the mutation knocked out (w_k) from the fitness of a "wild-type" individual (w):

$$s = w - w_k \tag{9}$$

We built a custom tool to recalculate the phenotype and fitness recalculations, implementing the Ascent C++ library to solve ODEs (Berry *et al.* 2021). This library was also used in the SLiM simulations, keeping the phenotype calculation consistent. The source code for our tool can be found at https://github.com/nobrien97/odeLandscape. The fitness effects were calculated in a similar manner for additive simulations. However, the phenotype was computed as the sum of additive effects instead of using the network ODE solution.

In silico mutant screen

In addition to the DFE of mutations that arose during the adaptive walk, we were also curious about the total DFE across the space of possible mutations (including deleterious mutations that would be lost quickly). We ran an *in silico* mutant screen experiment to determine whether the fitness distribution of possible mutations differed between additive and NAR populations. A flowchart is given in Supp. Fig. 2B. We sampled 1,000 mutations from a standard normal distribution and added them to the molecular component values/quantitative trait values of individuals at each step of the adaptive walk. This was done for both network and additive models. We chose a standard normal distribution as it matched the distribution of new mutations that populations faced in the simulations. For the network models, mutations were applied separately to α_Z and β_Z, while in the additive models, the mutation effect was added to the quantitative trait value. We then recalculated the phenotypic and fitness effects of the sampled mutations using the fitness effect calculation methods described above. The source code for this experiment can be found at https://github.com/nobrien97/NARAdaptiveWalk2023.

Table 1 Simulation parameters

Symbol	Parameter	Description	Value
N	Population size	Number of individuals in the population	5,000
μ	Mutation rate	Per-locus per-generation rate of mutation in the genome	9.1528e-06
τ	Mutational effect variance	Variance in the sampling distribution of mutational effects on molecular components, $a \sim \mathcal{N}(0, \tau)$	1.0
σ_w	Fitness function width	Width of the Gaussian fitness function (5% drop in fitness for individuals one unit away from the optimum).	$\sqrt{10}$
O_{shift}	Optimum shift amount	Magnitude of shift in the phenotypic optimum after the burn-in phase	1
$n_{burn-in}$	Number of burn-in generations	Number of generations of burn-in prior to shifting the optimum.	50,000
n_{gen}	Number of adaptation generations	Number of generations the model was run for following the optimum shift	10,000

Expected waiting times to beneficial mutation

To tie the mutant screen findings to a quantitative measure of adaptation, we used the mutant screen fitness effect distribution to estimate the expected waiting time to a beneficial mutation between models. We first calculated waiting times, t_{wait} for a particular simulation replicate at a given adaptive step using the equation

$$t_{\text{wait}} = \frac{1}{4N\mu p_{s>0}} \tag{10}$$

where N is the population size, μ is the per-locus, per-generation mutation rate, and $p_{s>0}$ is the probability that a new mutation is beneficial. N is multiplied by 4 because the population is diploid, and because there were two loci contributing to the trait. $p_{s>0}$ is given by

$$p_{s>0} \approx n^{ISMS}_{s>0} / n^{ISMS}$$

where n^{ISMS} is the total number of mutations generated in the *in silico* mutation screen experiment and $n^{ISMS}_{s>0}$ is the number of mutation screen alleles with beneficial effects on fitness.

Waiting times were calculated for each adaptive step and model. Comparisons between models were done using a bootstrap analysis. We sampled 100,000 different NAR-Additive model pairs across all adaptive steps, calculating the difference between their waiting times. We then calculated the means and 95% confidence intervals (CIs) for each model's waiting times and the difference between models (Δt_{wait}). The source code for this analysis can be found at `https://github.com/nobrien97/NARAdaptiveWalk2023`. A flowchart describing this methodology is shown in Supp. Fig. 3.

We also measured the difference in $p_{s>0}$ between models and across adaptive steps using a linear model:

$$\begin{aligned} p_{s>0} = &\alpha + \beta_1(\text{Model}_{\text{NAR}}) + \beta_2(\text{Adaptive step}_1) + \\ &\beta_3(\text{Adaptive step}_2) + \beta_4(\text{Adaptive step}_{\geq 3}) + \\ &\beta_5(\text{Model}_{\text{NAR}} \times \text{Adaptive step}_1) + \\ &\beta_6(\text{Model}_{\text{NAR}} \times \text{Adaptive step}_2) + \\ &\beta_7(\text{Model}_{\text{NAR}} \times \text{Adaptive step}_{\geq 3}) + \epsilon \end{aligned} \tag{11}$$

Where α is the intercept, β_i is the slope coefficient for dependent variable i, and ϵ is the residual error. To estimate marginal means and contrasts, we used the R package `emmeans` 1.8.5 in R 4.3.1 (Lenth 2023; R Core Team 2023).

Fitness landscapes

With the mutant DFE evaluated, we turned to how the underlying network might drive the complex distribution that we discovered. This involved constructing a fitness landscape to show how fitness changed with different combinations of α_Z and β_Z, introducing epistasis to the system. We generated 160,000 combinations of α_Z and β_Z, sampling α_Z and β_Z values between 0 and 3. We then calculated the phenotype and fitness for each combination using the ODE landscaper tool we used to calculate fitness effects. We plotted the resulting fitnesses against α_Z and β_Z using `ggplot2` 3.4.2 (Wickham 2016) in R 4.3.1 (R Core Team 2023). A similar method was used to estimate the fitness landscape of the ratio β_Z/α_Z as α_Z increased. For this analysis, we generated another 160,000 combinations of α_Z and β_Z/α_Z ratios. Both α_Z and β_Z/α_Z values were sampled between 0 and 3.

DFE analysis

The final part of our analysis involved measuring the shape of the distribution of beneficial fitness effects (DBFE), which should be Gumbel-distributed under Gillespie-Orr predictions (Orr 2003a; Gillespie 1984). We used a method developed

by Beisel *et al.* (2007) to fit a generalized Pareto distribution (GPD) to the DBFE obtained from the mutant screen. A flowchart describing this approach is provided (Supp. Fig. 4). The GPD characterizes a distribution of extreme events that exceed a threshold. The threshold is the high-fitness wild-type and the extreme events are beneficial mutations. The shape parameter of the GPD, κ, (shape) indicates which domain of attraction the underlying distribution belongs to. Beisel et al.'s (2007) method uses a likelihood ratio test to evaluate if $\kappa = 0$, corresponding to the Gumbel case. If $\kappa > 0$, the DFE belongs to the Fréchet domain, and if $\kappa < 0$, it belongs to the Weibull domain (Fisher and Tippett 1928; Beisel *et al.* 2007).

We used two different sampling approaches to fit our data to GPDs (Supp. Fig. 4). In the first method, we pooled mutant screen alleles across all replicates to create a pooled DBFE (pDBFE) for additive and network models. We sampled 1,000 mutations from the pDBFE and fit the GPD to the resulting sampling distribution. We repeated this process 10,000 times, yielding a distribution of estimates of κ. The pooled approach shows the average shape of the DBFE across many adaptive walks. In the second method, we sampled up to 100 mutations from each simulation's mutant screen DBFE and fit the GPD to that sample. This was done per adaptive step. Simulation-step pairs with fewer than 10 beneficial alleles were discarded as they could not be reliably used to estimate the GPD shape (Beisel *et al.* 2007). We obtained a distribution of κ estimates, one for each simulation-step pair. This approach shows the variability in the shape of the DFE during different adaptive walks. In both methods, p-values from the likelihood ratio tests were combined across replicates using Fisher's method (Fisher 1936). Both approaches fit the GPD using the R package `GenSA` 1.1.8 (Xiang *et al.* 2013) in R 4.3.1 (R Core Team 2023), with code modified from Lebeuf-Taylor *et al.* (2019).

Results

Assumption validation

We first checked if we were in the strong selection weak mutation (SSWM) regime assumed by the Gillespie-Orr model. Under SSWM, each beneficial mutation should fix before another arises. Hence, heterozygosity at QTLs/mQTLs should be close to zero most of the time. We found observed heterozygosity (H_O) at the QTLs/mQTLs was low on average, peaking just after the optimum shift at $11.39\% \pm 0.423\%$ (95% CI) and $9.637\% \pm 0.400\%$ (95% CI) in additive and NAR models, respectively (Supp. Fig. 8). The grand mean H_O across all time points was $8.368\% \pm 0.028\%$ (95% CI) and $7.841\% \pm 0.027\%$ (95% CI) for additive and NAR models. Although H_O was not zero, the majority of variation arising was not beneficial (Fig. 5B), meaning that the chance that two beneficial mutations could co-segregate was lower than H_O suggests.

To further test the SSWM assumption, we also measured the ratio between phenotypes generated by only fixations (i.e. the phenotype in absence of all segregating variants) and the mean population phenotype (Eq. 8. This ratio describes how much trait variation was contributed by segregating variants compared to fixations. When $r_{FS} = 1$, segregating variants contribute no trait variation. When $r_{FS} > 1$, segregating variants decrease the trait value on average compared to the phenotype due to fixations, and vice versa when $r_{FS} < 1$. We found that most populations had no variation contributed by segregating variants (Supp. Fig. 9). The average r_{FS} was 1.019 ± 0.007 (95% CI) and 1.024 ± 0.011 (95% CI) for additive and NAR populations respectively. From these figures, it seems that the populations are in the SSWM domain, enabling us to explore the extreme value theory expectations set out by Gillespie and Orr (Gillespie 1983, 1984; Orr 1998).

NAR and additive populations differ in their ability to adapt

We first examined the phenotypic response to selection during adaptive walks in populations with either an additive or a network (NAR motif) GP map. We defined an adapted population as one with a mean phenotype within 10% of the optimum ($1.9 < \bar{P} < 2.1$). We chose 10% as the threshold as it only decreased fitness by 0.04% relative to a 5% threshold, but increased the sample size of adapted populations by 42% (2636 adapted populations vs. 1529). By the end of the simulation, 50.5% of the additive simulations and 41% of the network simulations had adapted. For the remainder of the results, we consider only the adapted populations.

Among the adapted populations, there were differences in the rate of trait evolution between the models (Fig. 3). Network model populations took longer to reach the optimum on average (Fig. 3A), although the changes in phenotype across adaptive steps were similar between the models (Fig. 3B). Note that this difference cannot be solely attributed to the NAR ODE, as the multiplicative scaling of the allelic effects also contributes to differences between the models. See the discussion for more on this distinction. There was no difference between the models in the number of adaptive steps taken. Of the populations that adapted, 66.3% did so in a single step, 28% in two steps, and the remaining 5.7% adapted in three or more steps. On average, the number of steps taken was 1.4 ± 0.026 (95% CI). This agrees with results from Orr (Orr 2003b), who found a mean walk of $\approx$ 1.72 steps, assuming that the genotype was comprised of many loci and that the starting genotype of the walk was randomly sampled.

Among populations that had not yet adapted after one step (the left tail in Fig. 3), the median step in a network model occurred 900 (95% CI [150, 1600]) generations later than in the additive models (Supp. Fig. 5). To investigate the underlying causes of the differences in adaptation success between models, we examined the DBFE among fixations during the adaptive walk.

The distribution of beneficial fitness effects among fixations is similar between NAR and additive populations

Fixations had similar effect sizes in both models and became smaller over the course of the adaptive walk (Fig. 4). We found that the mode of the distribution of beneficial fixations was offset from 0 (Supp. Fig. 6), supporting a gamma distribution of fixations rather than an exponential distribution. We fit gamma distributions to the DFE among fixations in both the additive and network models using `fitdistrplus` 1.1.8 in R 4.3.1

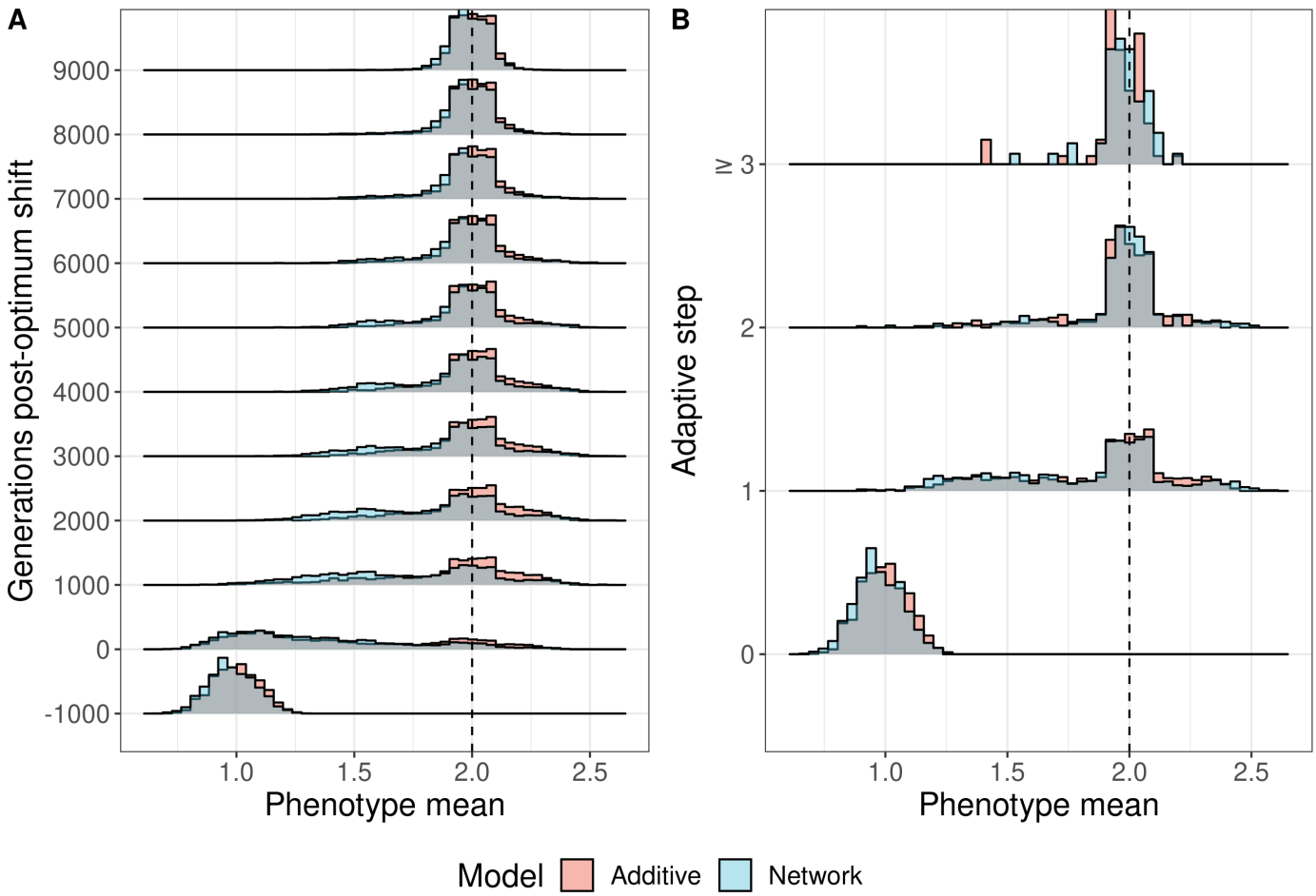


Figure 3 Phenotypic evolution over 10,000 generations of adaptation across 2,880 replicates of additive and network models. (A) The distribution of mean phenotypes among replicate populations during the adaptive period. The -1000 timepoint represents 1,000 generations before the optimum shifted. (B) The mean phenotype at each adaptive step (each fixation along the adaptive walk). The last step is pooled across all steps greater than or equal to step three in the walk. The dotted line in A and B represents the post-shift optimum. The grey area shows the overlap between additive and network distributions.

(R Core Team 2023; Delignette-Muller and Dutang 2015). The deleterious fixations shown in Fig. 4 were excluded from the dataset when fitting the gamma distribution. These deleterious fixations were neutral or adaptive prior to the shift in the optimum and typically reached high frequencies before the adaptive walk commenced (Supp. Fig. 7). Therefore, these fixations can be attributed to the action of genetic drift.

We compared the shape and rate parameters of the gamma fits between the models. The network model populations exhibited slightly fewer small effect fixations and had a shorter tail compared to the additive models (NAR: Shape $= 2.032 \pm 0.147$ 95% CI, Rate $= 44.392 \pm 3.651$ 95% CI; Additive: Shape $= 1.722 \pm 0.108$ 95% CI, Rate $= 37.588 \pm 2.730$ 95% CI). However, the DFE among beneficial fixations did not differ strongly between the models. To investigate this further, we needed to determine the availability of beneficial mutations to the populations. We did this by conducting an *in silico* mutant screen experiment, examining the DFE among new mutations at each adaptive step.

The distribution of fitness effects among possible mutants is complex and deleterious in NAR populations

The *in silico* mutation screen experiment revealed that the DFE among new mutations differed dramatically between models (Fig. 5A). Both were strongly negatively skewed; however, mutations in NAR populations were overall more deleterious than those in additive models, with a longer and more pronounced tail of strongly deleterious mutations. In addition, the DFE in NAR populations was bimodal (Fig. 5A). The first mode was at $s \approx 0$, similar to the additive DFE mode (NAR: $s = -0.0003$; Additive: $s = -0.0004$). The second mode was at $s = -0.113$. Among beneficial mutations, the distributions were more similar.

Data pooling affects generalized Pareto distribution fitting results

We fit a generalized Pareto distribution (GPD) to the DBFEs generated by the mutant screen experiment using methods de-

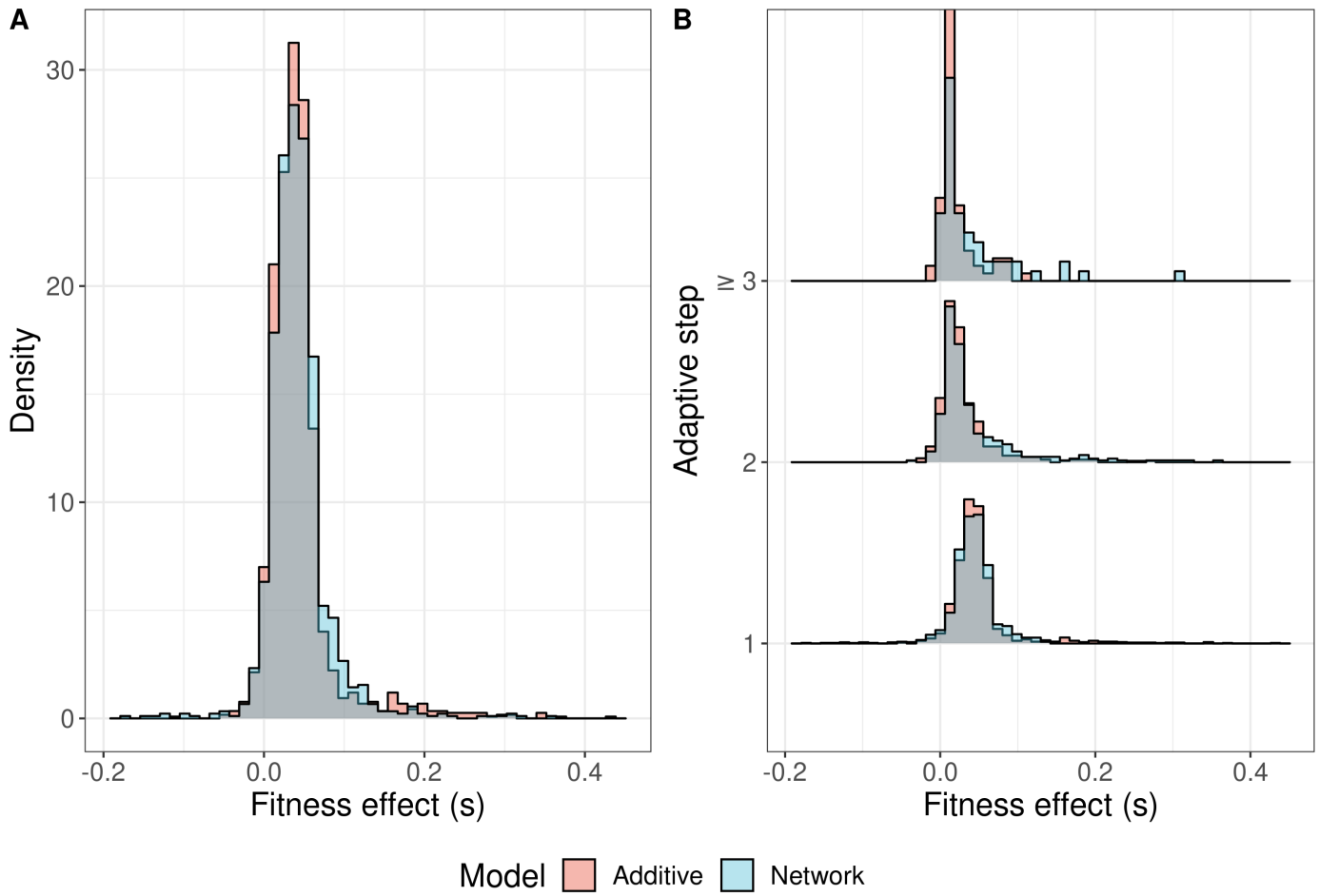


Figure 4 Fitness effects among fixations during adaptive walks to a phenotypic optimum across 2,880 replicates of additive and network models. (A) The overall distribution of all fixations across the entire walk. (B) The distribution of fixations at each adaptive step. The grey area shows overlap between the additive and network distributions of effects.

veloped by Beisel *et al.* (2007). This was done separately for additive and network models. The shape parameter of the GPD (κ) describes what domain the DFE belongs to: $\kappa = 0$ specifies a Gumbel domain, $\kappa < 0$ specifies a Weibull domain, and $\kappa > 0$ specifies a Fréchet domain. We fit this in two ways: first, by pooling all replicates and sampling from this pooled distribution to estimate the GPD parameters; and second, by sampling each replicate at each of its adaptive steps separately and estimating the GPD for each estimate (for more information, see the Methods section). Under the pooled approach, we found evidence for a minor deviation from an exponential distribution towards the Weibull domain in both the additive and network models (Network: $\hat{\kappa} = -0.130 \pm 0.0007$; Additive: $\hat{\kappa} = -0.118 \pm 0.0005$). However, this deviation is small enough that the adaptive walk should behave as if its DFE belonged to the Gumbel domain (Joyce *et al.* 2008).

Under the per-simulation method, we saw a different result. The DBFE in both NAR and additive populations were strongly Weibull distributed. In the NAR model, κ tended to decrease over the adaptive walk. However, the additive model κ estimate was relatively stable at around $\kappa \approx -2$, slightly increasing over the adaptive walk (Supp. Table 2; Supp. Fig. 11). In addition, these non-pooled fits were considerably more variable than under the pooled approach (Supp. Fig. 10), particularly at early adaptive steps. Under the per-simulation/adaptive-step approach, across all adaptive steps, at least 99.8% of additive replicates and at least 99.7% of NAR replicates had DFEs of beneficial mutations that fit the Weibull domain. Under the pooling method, all replicates in both models were approximately Gumbel-distributed (i.e. $\kappa \approx 0$).

Beneficial mutations are less common in NAR populations than additive populations

Despite sharing similar DFEs, beneficial mutations were on average less common in network models than in additive models (Linear model, Eq. 11; $F_{7,5171} = 1239, R^2 = 0.626$, Fig. 5B). The difference between models diminished with adaptive steps (as populations neared the optimum). Immediately after the optimum shift, mutations in NAR populations were $9.504 \pm 0.639\%$ (95% CI) less likely to be beneficial. At the first adaptive step, this difference decreased to $4.648 \pm 0.639\%$ (95% CI). At the second adaptive step, the NAR models were

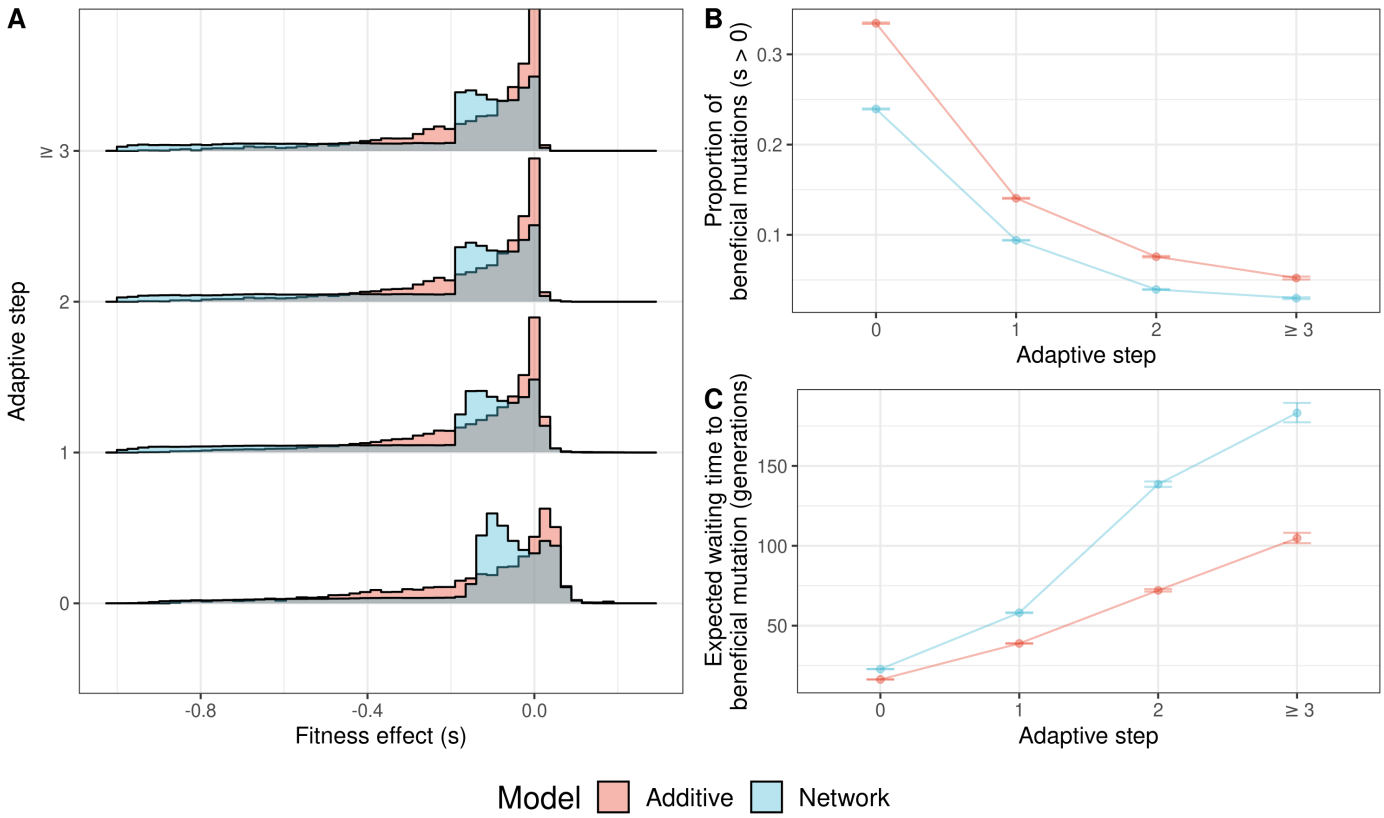


Figure 5 Fitness effects among possible mutations and the consequences for adaptation in 2,880 replicates of additive and network models. (A) The distribution of fitness effects among 1,000 randomly sampled mutations in additive and network models. The grey area represents overlap between additive and network models. (B) The proportion of mutations that are beneficial over the adaptive walk. (C) The change in beneficial mutation probability corresponds to the expected waiting time for a beneficial mutation.

$3.630 \pm 1.107\%$ (95% CI) less likely to produce a beneficial mutation. At further adaptive steps, there was no difference between models. We measured the effect of the lower beneficial mutation rate on adaptation by estimating the waiting time for a new beneficial mutation to arise. We found that the mean expected waiting time for a beneficial mutation in NAR populations was 98.555 ± 5.816 (95% CI) generations longer than in additive populations (Additive mean waiting time: 136.382 ± 2.131 (95% CI) generations; NAR mean waiting time: 234.937 ± 5.414 generations; Fig. 5C). Given that the DFE differed between models, we wanted to explore the causes underpinning this. To do so, we investigated the fitness landscape of α_Z and β_Z.

The NAR fitness landscape is ridge-like across molecular component space

We found a pronounced fitness ridge on the NAR fitness landscape, surrounded by valleys of low fitness (Fig. 6A). This ridge lay diagonally across α_Z and β_Z space, suggesting that fitness depended on the ratio of the molecular components. This relates to the underlying ODE: β_Z/α_Z is the level of steady state Z expression under simple gene regulation (i.e. without any negative feedback or induction) (Alon 2019). Plotting this ratio over α_Z confirmed this: β_Z/α_Z was the arbiter of fitness as long as $\alpha_Z \gtrapprox 0.5$ (Fig. 6B). We investigated how the β_Z/α_Z ratio contributed to the shape of the GP map, finding a non-linear relationship between β_Z/α_Z and phenotype (Fig. 6C). Around the optimum, increasing β_Z/α_Z quickly increases the trait value, however with larger β_Z/α_Z increases, the trait value changes less (Fig. 6C). We approximated the optimum β_Z/α_Z ratio ($\theta_{\beta_Z/\alpha_Z}$), by substituting our fixed molecular components ($K_{XZ} = 1, K_Z = 1, n_Z = n_{XZ} = 8$) into Eq. 1. We considered the case where $X = 1$, simplifying to

$$Z'(t) = -Z\alpha_Z + \frac{\beta_Z}{2 + 2Z^8}$$

The equilibrium fulfills the equation

$$\frac{\beta_Z}{2\alpha_Z} = Z(1 + Z^8)$$

If we assume a model of simple gene regulation,

$$\theta = \frac{\beta_Z}{\alpha_Z}$$

$$\frac{\theta}{2} = Z(1 + Z^8)$$

where θ is the steady state of gene expression under simple gene regulation (Alon 2019). For small values of Z (where $Z < 1$),

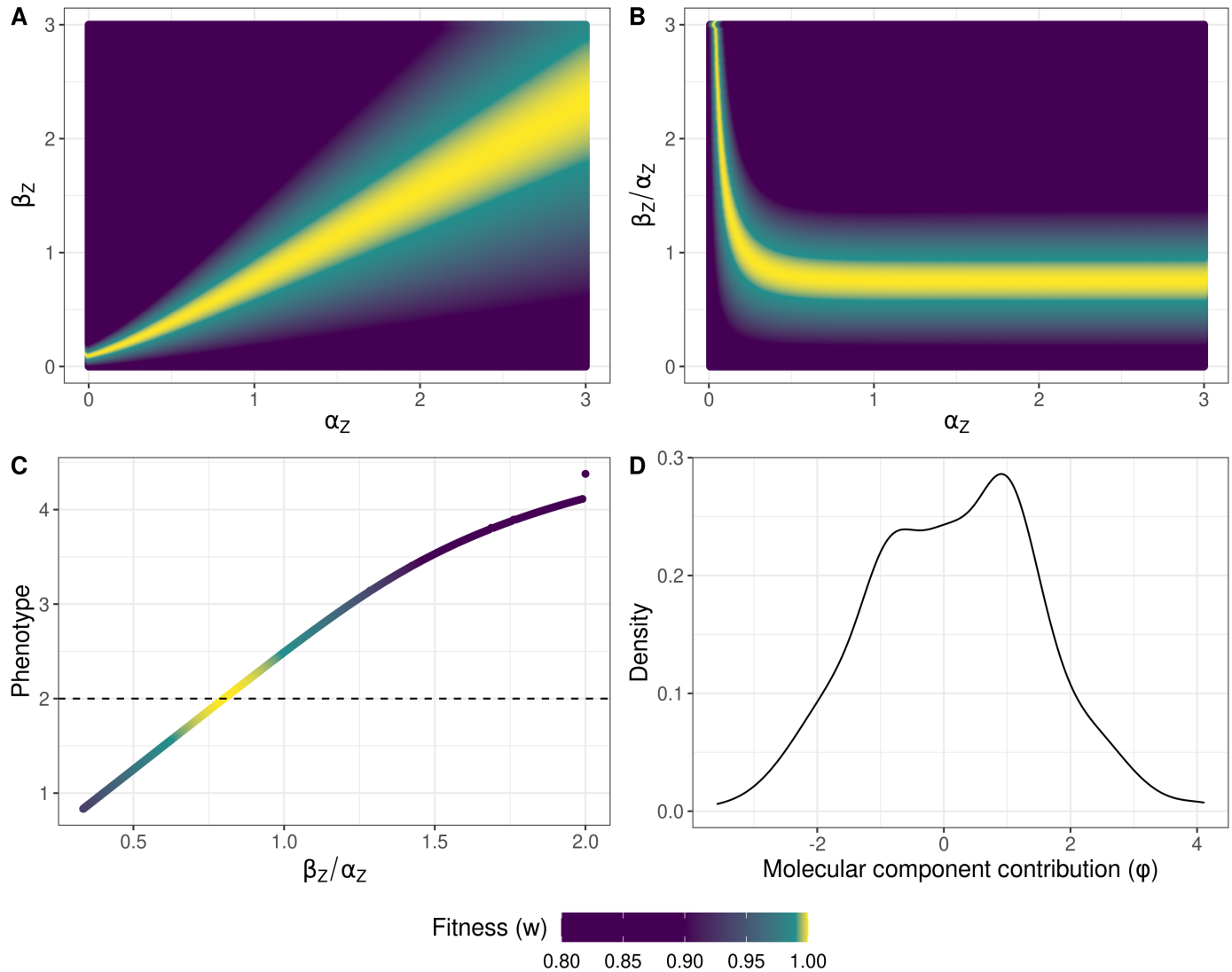


Figure 6 The fitness and phenotype landscapes of the network model. (A) The fitness landscape with regards to the molecular components of the network, α_Z and β_Z. The landscape consists of a high-fitness ridge (yellow) surrounded by low-fitness valleys (purple). (B) The ratio of β_Z/α_Z, the steady state of the network, is what matters for adaptation. As α_Z increases relative to β_Z/α_Z, fitness becomes constant over α_Z. (C) Fitness is a nonlinear function of the network steady state, β_Z/α_Z, although largely linear around the phenotypic optimum. An optimum ratio at $\theta_{\beta_Z/\alpha_Z} = 0.8$ maximizes fitness. The dashed black line shows the phenotypic optimum. (D) The contributions of mutations in each molecular component to the adaptive walk. This is taken for populations with at least 2 steps in the walk. ϕ shows the difference in sums of absolute allelic effects among fixations in α_Z and β_Z across an adaptive walk. Values greater than 0 represent larger changes in α_Z than β_Z, and vice versa for values smaller than 0. At 0, both α_Z and β_Z changed by the same amount during the adaptive walk.

this is well approximated by

$$Z = \frac{\theta}{2}$$

Using this, we approximate the solution of the ODE by the equilibrium whenever $X = 1$ and by 0 when $X = 0$. Hence, the solution can be given by

$$M = \int_{t_{start}}^{t_{stop}} \frac{\theta}{2} dt = \int_1^6 \frac{\theta}{2} dt = \frac{5\theta}{2}$$

When the optimum is at $P = M = 2$, the optimum steady state is

$$\begin{aligned} \theta_{\beta_Z/\alpha_Z} &= \frac{2M}{5} \\ &= 0.8 \end{aligned}$$

This assumes that the increase in Z after X is activated is rapid, and likewise for the decrease in Z when X is deactivated (i.e. in Fig. 2B, the green curve should be close to rectangular around X activation). This behavior is driven by n_Z and n_{XZ}: for our

chosen values, the above approximation predicts the observed $\theta_{\beta_Z/\alpha_Z}$ in Fig. 6B.

To explore how the nonlinearity of the GP map might have influenced how populations navigated molecular component space, we examined the ratio of absolute additive effect sizes on molecular components among populations that adapted by two steps or more, denoted by ϕ (Fig. 6D). Mathematically,

$$\phi = |2a_{\alpha_Z}| - |2a_{\beta_Z}|,$$

where a is the allelic effect on α_Z or β_Z without the multiplicative transformation (i.e. the value composed by Eq. 6 instead of Eq. 3. Allelic effects were multiplied by two since we simulated a diploid population. When $\phi > 0$, fixations have larger effects on the α_Z axis than the β_Z axis, and vice versa for when $\phi < 0$. When $\phi = 0$, there are equal contributions from mutations affecting α_Z and β_Z. On average, $\phi = 0.164 \pm 0.150$, suggesting that α_Z mutations contributed most to adaptation, albeit by a small amount. Of the populations considered for this analysis, 29.5% adapted with only α_Z mutations, 25% with only β_Z mutations, and the remaining 45.5% used a combination of both to reach the optimum.

Discussion

Overview

In this study, we explored how a simple gene network might influence trait evolution during an adaptive walk scenario. Models of adaptation have traditionally relied on direct genotype-phenotype relationships (Falconer 1996; Kopp and Hermisson 2009; Matuszewski *et al.* 2015), overlooking the intricate gene interaction networks that drive trait expression at different stages of development (but see (Rice 1998; Slatkin 1987; Wagner 1996)). Understanding how genetic networks influence adaptation is an under-explored area of research, particularly in the study of DFEs.

We observed that both network and additive models exhibit similarities in key aspects of their adaptive walks, specifically the number of adaptive steps and the distribution of fitness effects (DFE) among both fixations and beneficial mutations (DBFE) (Fig. 3, 4). However, the characterization of the DBFE was sensitive to the estimation method employed. Utilizing a pooling approach under the assumptions of the Central Limit Theorem leads to a homogenization of the shape parameter κ, nudging the distribution towards the Gumbel domain, and thus suggesting exponential spacing between adjacently-ranked beneficial alleles (Joyce *et al.* 2008). The averaging effect of pooling samples can mask underlying Fréchet or Weibull behaviors in κ, effectively reducing our ability to discriminate between different extreme value distributions (EVDs).

In contrast, using a non-pooled approach preserves the contributions of each replicate's unique combination of mutations to the shape parameter, improving the detection of non-Gumbel behavior. However, this is not always possible in experimental studies where the number of beneficial mutations obtained per round of adaptation is small. Under these conditions, pooling between replicates is required to reduce the risk of incorrectly rejecting a Gumbel distribution (Joyce *et al.* 2008; Beisel *et al.* 2007). We suggest that researchers should be cautious of the homogenization effect of the pooling approach when estimating the DBFE. Given these considerations, we next discuss the variation in estimates of the shape parameter κ when using pooled and non-pooled approaches.

When we fit GPDs to each replicate with their own set of mutations, almost all the fitted GPDs belonged to the Weibull domain (Supp. Fig. 10). In addition, κ was quite variable between simulations, models, and adaptive steps. κ affects the GPD by modulating an upper bound on the effect size of beneficial mutations. With decreasing κ, this upper bound approaches zero. We observed that κ tended to decrease over longer walks under the network model. In the additive model, κ was relatively stable, however it slightly increased (particularly when comparing κ before the optimum shift and at adaptive steps ≥ 3 (Supp. Fig. 11; Supp. Table 2). Regardless of this increase, both models were largely Weibull-distributed (over 99% of simulations/adaptive steps belonged to the Weibull domain), which rejects Gillespie's Gumbel expectation.

Weibull-distributed DFEs of beneficial effects have been observed in a number of empirical studies in viral and bacterial populations (e.g. Rokyta *et al.* 2008; MacLean and Buckling 2009; Foll *et al.* 2014). Furthermore, a Weibull distribution of effects is predicted when populations are close to an optimum (Martin and Lenormand 2008). Hence, our results match theoretical expectations. Evidence for limits on adaptation seems common and both models seem to be affected by such limitations. While both models were starved of large-effect beneficial mutations relative to the exponential expectation, this effect was more limiting for network models than additive as the walk went on. This was further highlighted by a major difference between network and additive responses to selection: the rate of trait evolution.

Trait adaptation and network evolution

In our simulations, we found that network simulations approached the optimum more slowly than additive models on average (Fig. 3A). This difference in rate might not have only been due to the NAR structure, but also because of the difference in allelic effect scaling between additive and network models, as shown in Eq. 3 and Eq. 6. This scaling difference can influence rates of adaptation, mandating caution in attributing observed differences solely to the network's complexity. Hence, we do not make conclusions on how much the observed differences are due to the NAR network's structure compared to the difference in allelic effect scaling. Nonetheless, there was a difference in the rate of adaptation between the models and the NAR DFE does not completely reflect the expectation under a multiplicative model, so the distribution of effects is at least partly mediated by the NAR structure. This will be the subject of a future study.

We explored the distribution of new mutations and discovered a complex DFE in the network models, contrasting with the simple distributions observed in additive models (Fig. 5). The additive model distribution largely met Gillespie's expectations: the DFE was comprised of mostly neutral mutations with a small, exponential tail of beneficial mutations, and a longer tail of

deleterious mutations (Fig. 5, (Gillespie 1991) pg. 267, Fig. 6.5).

In the network model, we found a bimodal distribution of fitness effects. The bimodality is more pronounced than expected compared to a multiplicative model, suggesting that the NAR contributes to this shape. This observation aligns with previous studies that have reported complex and multimodal distributions of deleterious fitness effects (e.g. Kousathanas and Keightley 2013; Zeyl and DeVisser 2001), potentially attributed to different classes of mutations with distinct DFEs (Eyre-Walker and Keightley 2007). Mutations in gene regulatory regions can have widely varying effects on fitness (e.g. Halligan *et al.* 2013; Williamson *et al.* 2014; Ohta 2002), suggesting that network structure can strongly affect the fitness distribution of mutants on a given molecular component (Erwin and Davidson 2009). The ridge-like shape of the α_Z/β_Z fitness landscape reflects this, showing that dependencies between the molecular components drive the fitness of organisms in the NAR model and that both the phenotypic and fitness effect of a mutation depends on its genetic background (Fig. 6).

Unlike the additive model, the fitness landscape between the molecular components of the NAR revealed non-interchangeable per-locus effects (Fig. 6A,C). The ridge-like shape of the landscape corroborates findings by Kozuch *et al.* (2020), who investigated a similar NAR motif in the *E. coli lexA* transcription factor. NAR populations also faced lower beneficial mutation rates and a slight bias favoring α_Z mutations over β_Z, which might have been caused by the ridge-like landscape (Fig. 5, 6D). The bias suggests α_Z mutations should be less deleterious on average than β_Z mutations. Supporting this, α_Z mutations were slightly less likely to be strongly deleterious, instead falling between the distribution's modes (Supp. Fig. 13). However, both mutation types remained largely deleterious, and had similar DBFEs. A potential explanation for the bias towards α_Z mutations involves the optimal β_Z/α_Z ratio of 0.8 post-shift, versus 0.4 at burn-in. To reach this optimum from burn-in, populations could increase β_Z or decrease α_Z. However, the required α_Z decrease exceeded the β_Z increase. In essence, larger α_Z mutations were needed to reach the new optimum. Crucially, α_Z and β_Z mutations did not affect the phenotype equally.

Complexity and the cost of adaptation

In our study, the NAR model demonstrated greater complexity compared to traditional additive models. We discovered a nonlinear relationship between β_Z/α_Z and phenotype, embodied by a nonlinear genotype-phenotype (GP) map (Fig. 6C). Such maps are known to impose adaptive constraints (Wade *et al.* 2001; Wagner 2011), reflected in the multimodal shape of the deleterious DFE for NAR populations (Fig. 5A). The NAR model's increased complexity through multiple molecular components and dynamic biochemical interactions renders it more susceptible to deleterious mutations. This vulnerability could amplify Hill-Robertson interference, with beneficial alleles overshadowed by deleterious backgrounds (Hill and Robertson 1966). For instance, consider the persistence of epistasis over time. Gene interactions can lead to situations where genetic combinations have greater effects on fitness than the sum of their parts (positive epistasis), or the combinations can reduce fitness compared to the sum expectation (negative epistasis). Drift and mutation generate positive and negative epistasis in similar proportions (Ortiz-Barrientos *et al.* 2016). However, when recombination is rare, selection will quickly fix synergistic combinations, leading to an excess of deleterious gene combinations which persist through drift and limit the efficiency of selection (Hill and Robertson 1966). When a gene network introduces further epistasis on traits, this might compound the imbalance between negative and positive epistasis. Thus, network complexity may trade off with adaptive potential.

Trade-offs between complexity and adaptability echo the cost of complexity described by Orr (Orr 2000). The cost of complexity postulates that an increasing number of traits under selection leads to an increasingly deleterious mutation space, impeding adaptation. This cost could also apply to network models of traits, where molecular components under selection limit adaptation. This notion finds support in studies that have explored network topology in different species. For instance, Costanzo *et al.* (2010) found that increasing the degree of genetic interaction in *Saccharomyces cerevisiae* genes was negatively correlated with single mutant fitness, suggesting an adaptive cost to maintaining a highly connected gene network. This cost might lead to the long-term stability of gene networks or complex traits: comparisons between *S. cerevisiae* and *Schizosaccharomyces pombe* genetic interaction networks revealed remarkable similarity in network structure (Koch *et al.* 2012). Another study by Barua and Mikheyev (Barua and Mikheyev 2021) found features of housekeeping gene networks involved in reptile venom production were conserved between amniote clades, with those genes contributing to saliva production in mammals (Barua and Mikheyev 2021). The cost of complexity suggests some rather strong limitations on the evolution of complex genetic networks: in Orr's (Orr 2000) model, the rate of adaptation declines according to the inverse square root of the number of traits, $1/\sqrt{n}$. Empirical evidence suggests that biological networks can be extremely complicated, connecting hundreds of nodes. Adaptation involving such a highly connected network would be extremely slow if Orr's (Orr 2000) theory is correct. However, aspects of network structure, including local connectivity and modularity, might alleviate such a cost of molecular complexity.

One empirical example of network structure modulating the rate of evolution is the biofilm production network in *Candida* species (Mancera *et al.* 2021). In this network, seven master regulators control the expression of about one-sixth of the *C. albicans* genome and are required to drive biofilm production (Mancera *et al.* 2021). Genes that are connected to one or more of these master regulators are much less connected than the master regulators themselves and more "free" to evolve without affecting the expression of hundreds of other genes. Mancera *et al.* (2021) found that the master regulators evolve slowly compared to the target genes, possibly due to the structure of the network. Since the master regulators affect the expression of hundreds of genes, the space of beneficial mutations at master regulator loci might scale according to the cost of complex-

ity. On the other hand, the authors found that target genes are considerably more divergent between *Candida* species and populations, perhaps due to target genes harboring fewer interactions within the network than master regulators (Mancera *et al.* 2021).

The NAR motif serves as a simplified regulatory network, effectively sidestepping the full cost of complexity outlined by Orr (Orr 2000). Despite this, it seems to impose greater adaptive constraints than the additive model, a point underlined by the density of strongly deleterious alleles in the DFE of new mutations (Fig. 5). Fixing multiple NAR molecular components (e.g. K_Z, K_{XZ}, n_Z, n_{XZ}) renders the model's behavior more predictable. For sufficiently large α_Z, individual fitness correlates solely with the ratio β_Z/α_Z (Fig. 6B), reflecting the steady-state cellular concentration under simple regulation (Alon 2019). This outcome may be attributed to our parameter choices; specifically, the rapid onset of Z activation/suppression due to $n_Z = n_{XZ} = 8$. Varying these Hill coefficients or treating them as evolvable components in future work could offer deeper insights into how more gradual responses to X activation/deactivation and Z production might modulate total Z expression (Alon 2019). Moreover, evolving parameters like K_Z and K_{XZ} could elucidate the full NAR fitness landscape. In essence, our simplified NAR model balances complexity and tractability, avoiding certain complexities while retaining meaningful constraints.

Should K_Z be allowed to evolve, the fitness landscape would likely exhibit increased ruggedness. In this context, the optimal β_Z/α_Z ratio would be contingent upon the K_Z value. A similar effect is anticipated for K_{XZ}, the regulator of Z's production rate in response to X. Consequently, both steady-state and fitness would depend on K_Z and K_{XZ}. Introducing more evolvable molecular components is expected to amplify the landscape's ruggedness due to interdependencies that influence fitness, exacerbating the previously noted cost of complexity. To substantiate these projections, future research should explore more intricate networks, as exemplified by empirical studies such as Bertheloot *et al.* (2020). In addition, exploring the behavior of other network motifs will elucidate the generality of our findings.

There are other motifs that are common in nature that could be reasonable choices for a toy model of network-mediated traits. Other options include feedback loops, feedforward loops, single input modules, and cascades. Each of these motifs have different behaviors that could lead to different evolutionary dynamics. For instance, feedback loops can create oscillatory expression patterns (Pigolotti *et al.* 2007). Autoregulation is a simple example of a feedback loop, however multi-locus feedback loops also exist. The p53–Mdm2 feedback loop is an example of a two-gene feedback loop in vertebrates. p53 is an important tumor suppressor transcription factor. Its expression is kept at an equilibrium cell concentration by the presence of Mdm2 (Wu *et al.* 1993). p53 is a positive regulator of Mdm2, whilst Mdm2 inhibits p53 expression. When this balance is interrupted, increased Mdm2 production results in cell proliferation, whilst increased p53 production leads to cessation of growth (Wu *et al.* 1993). Another common motif, the feedforward loop, has eight different configurations, each with different behaviors (Alon 2019). Some of these configurations are more common in transcription networks than others, suggesting a selective pressure favoring these ubiquitous forms (Mangan and Alon 2003).

Ramifications of strong selection weak mutation (SSWM) and alternative approaches

This study focused on adaptation by employing a Gillespie-Orr model, which operates under the assumption of SSWM, characterized by $Ns \gg 1$ and $N\mu \ll 1$. We ensured adherence to the SSWM assumption by choosing an appropriate population size and mutation rate (see Table 1). In our simulations, $N\mu = 0.092$. However, Ns varies due to selection coefficients being dependent on individuals' genetic backgrounds and phenotypes. Nonetheless, $Ns > 1$ holds at $s > 0.0002$ for this model, so most beneficial mutations should meet the $Ns \gg 1$ criterion. In the instance they do not, these mutations are driven by drift (Lynch 2010). This also applies to slightly deleterious mutations. $N|s| = 1$ represents a "drift barrier", where the fitness effects of alleles become effectively too small to drive allele frequency changes, leading to drift-dominated dynamics. The drift barrier explains our finding that the DFE among fixations followed a gamma distribution as opposed to an exponential in both network and additive models (6). As populations approach the optimum, selection's efficiency is weakened by the drift barrier (Lynch 2010). At the barrier, the probability of fixation (π) is no longer predicated by $\pi \approx 2s$ when the population is close to the optimum. Instead, $\pi = 1/2N$ becomes more applicable (Charlesworth and Charlesworth 2010). The switch from selection-dominated to drift-dominated dynamics occurs when $N|s| = 1$ (Koonin 2016). In our simulations, this means that when $|s| < 0.0002$, drift is expected to dominate. Owing to the relatively large population size in our simulations ($N = 5000$), drift is unlikely to drive allele fixations early in the adaptive walk (i.e. when most beneficial alleles have $s > 0.0002$), however later steps might be more affected as the possible fitness improvement of any given allele faces diminishing returns.

Another assumption of the SSWM model is that only one mutation segregates at any given time point, i.e. before a new beneficial mutation arises, the previous beneficial mutation must have fixed. This was met most of the time (see Supp. Fig. 8, 9), however, some simulations revealed instances where segregating variation contributed to adaptation, including the presence of alleles under balancing selection (Supp. Fig. 12). Future research should therefore explore the model using a polygenic approach, where many loci contribute small amounts to the molecular components underlying complex traits. Adaptation would then occur by small frequency shifts at these loci (Walsh and Lynch 2018c). The genetic architecture of polygenic traits, that is, the number of loci, frequencies, and the distributions of their additive and non-additive effects on phenotypes, can greatly affect the response to selection (Barghi *et al.* 2020). There are also expectations for how genetic networks might determine the genetic architectures of polygenic traits: large-effect loci are likely linked to *cis*-regulatory features and coding region

mutations, whilst small-effect loci are more likely to represent *trans*-regulatory elements (Boyle *et al.* 2017; Liu *et al.* 2019). However, the non-additive effects of genetic networks on phenotypes remain understudied. These interactions have ramifications for theories of the evolution of recombination (Ritz *et al.* 2017), and the persistence of linkage disequilibrium in natural populations (Ortiz-Barrientos *et al.* 2016).

Opportunities and challenges in modeling complex genetic networks

Our study serves as an introductory venture into the evolutionary modeling of genetic networks, elucidating how they guide evolutionary trajectories. The strong selection and weak mutation assumptions grant computational tractability but sideline rich polygenic dynamics (Barghi *et al.* 2020; Thornton 2019). The NAR motif offers initial insights but lacks empirical complexity, making future work with more complex architectures important. Moving forward, our focus will shift towards a spectrum of genetic architectures to better capture non-additive effects. This will include investigating other network motifs, including several feedforward loops. By doing so, we will enrich our understanding of how topology influences evolutionary pathways (Siegal *et al.* 2007). This transition will require further statistical and computational innovations to maintain feasibility. Collaborations with molecular biologists will enrich our models with further mechanistic details. Rigorous validation against empirical data and existing theories will be a cornerstone in constructing more realistic models of adaptation (e.g. Bertheloot *et al.* 2020).

As well as this simple approach, we also constructed a general form of our network model in S1 Appendix. In the full model, we envision a phenotype as a combination of "molecular traits": intermediate traits like methylation or gene expression that represent cellular or developmental processes (Claringbould *et al.* 2017). Our NAR model represents the simplest case, where there is one molecular trait, Z expression, which we treat as the phenotype. In more complex implementations, phenotypes can emerge from a composite of molecular traits. Each molecular trait has its own differential equation and its own set of molecular components driving gene expression, creating a hierarchical model of gene expression driving quantitative trait variation.

As we delve into more complex systems, computational hurdles become prominent. The complexity of the ODE system significantly impacts *in silico* performance (Agocs 2020; Städter *et al.* 2021). We are considering optimization strategies such as smarter caching of common genotypes, data-oriented design, and adaptive step-size algorithms. Pre-computed ODE solutions stored on disk are also under exploration as a means to tackle the computational load. In summary, our study, while preliminary, opens up avenues for understanding the complex interplay between genetic networks and evolutionary dynamics, leaving us optimistic about the future of this research area.

Conclusion

Our network approach contributes methodologically to the field and holds broad applicability to many open questions in evolutionary biology, from quantitative genetics to the mechanisms underpinning adaptability. It offers a robust framework to investigate issues such as the maintenance of genetic variation, the evolution of recombination, and the dynamics of adaptation from both standing variation and *de novo* mutations. We have introduced how networks can generate epistasis, with implications for the evolution of recombination, maintenance of linkage, and the trade-offs between network complexity and adaptability. Our study here shows that networks can create complex distributions of fitness effects that differ from additive expectations. While promising, it is essential to acknowledge the model's limitations, particularly concerning its computational demands and simplifications, which future work should aim to address. Empirical validation through evolutionary experiments will be crucial in assessing our model's real-world applicability. Expanding into adaptation by standing variation will further illuminate the role of genetic interactions during polygenic adaptation. As detailed systems models of complex traits become increasingly available, our framework can be employed to simulate the evolution of traits based on empirically-described networks. The utility of our model also extends beyond evolutionary biology, offering valuable insights for interdisciplinary collaborations that aim to integrate molecular biology, computational science, and statistical genetics. Future work should focus on leveraging this interdisciplinary potential and on exploring more complex genetic networks to better understand how gene interactions constrain evolution.

Author contributions

DO conceived the original idea. NO, JE and DO developed together the new theoretical approach, constructed the models and designed simulation experiments. NO implemented and ran the simulations. BH and JE provided mathematical guidance. NO wrote the paper with assistance from all authors.

Data availability

The data used in this study is available at `https://doi.org/10.48610/f3850b0`. Instructions and scripts for repeating the analysis and for running the SLiM simulations are available at `https://github.com/nobrien97/NARAdaptiveWalk2023`. SLiM modifications are available at `https://github.com/nobrien97/SLiM/releases/tag/AdaptiveWalks2023`: note that this modification requires the Ascent ODE solution library installed `https://github.com/AnyarInc/Ascent/releases/tag/v0.7.0`.

Acknowledgments

We would like to acknowledge members of the Ortiz-Barrientos lab for their feedback on earlier versions of this manuscript. This research was conducted using the Gadi HPC system maintained by the National Computational Infrastructure (NCI), which is supported by the Australian Government.

Funding

This work was supported by an Australian Research Council grant awarded to DO (FT200100169) and the Australian Research Council Centre of Excellence for Plant Success in Nature and Agriculture (CE200100015).

Conflicts of interest

The authors declare no conflicts of interest.

Literature cited

Agocs FJ. 2020. (py)oscode: Fast solutions of oscillatory ODEs. Journal of Open Source Software. 5:2830.

Alon U. 2007. Network motifs: Theory and experimental approaches. Nat Rev Genet. 8:450–461.

Alon U. 2019. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press. Boca Raton, FL, USA. second edition.

Álvarez-Castro JM, Carlborg Ö. 2007. A unified model for functional and statistical epistasis and its application in quantitative trait loci analysis. Genetics. 176:1151–1167.

Ang RML, Chen SAA, Kern AF, Xie Y, Fraser HB. 2023. Widespread epistasis among beneficial genetic variants revealed by high-throughput genome editing. Cell Genomics. 3:100260.

Aston E, Channon A, Belavkin RV, Gifford DR, Krašovec R, Knight CG. 2017. Critical mutation rate has an exponential dependence on population size for eukaryotic-length genomes with crossover. Sci Rep. 7:15519.

Baier F, Gauye F, Perez-Carrasco R, Payne JL, Schaerli Y. 2023. Environment-dependent epistasis increases phenotypic diversity in gene regulatory networks. Science Advances. 9:eadf1773.

Bank C, Hietpas RT, Wong A, Bolon DN, Jensen JD. 2014. A Bayesian MCMC Approach to Assess the Complete Distribution of Fitness Effects of New Mutations: Uncovering the Potential for Adaptive Walks in Challenging Environments. Genetics. 196:841–852.

Barghi N, Hermisson J, Schlötterer C. 2020. Polygenic adaptation: A unifying framework to understand positive selection. Nat Rev Genet. 21:769–781.

Barua A, Mikheyev AS. 2021. An ancient, conserved gene regulatory network led to the rise of oral venom systems. PNAS. 118:e2021311118.

Beisel CJ, Rokyta DR, Wichman HA, Joyce P. 2007. Testing the extreme value domain of attraction for distributions of beneficial fitness effects. Genetics. 176:2441–2449.

Bendixsen DP, Østman B, Hayden EJ. 2017. Negative epistasis in experimental RNA fitness landscapes. J Mol Evol. 85:159–168.

Berry S, Walcott M, Avalos CG, Du R. 2021. Ascent. AnyarInc.

Bertheloot J, Barbier F, Boudon F, Perez-Garcia MD, Péron T, Citerne S, Dun E, Beveridge C, Godin C, Sakr S. 2020. Sugar availability suppresses the auxin-induced strigolactone pathway to promote bud outgrowth. New Phytologist. 225:866–879.

Bomblies K, Peichel CL. 2022. Genetics of adaptation. PNAS. 119:e2122152119.

Boyle EA, Li YI, Pritchard JK. 2017. An expanded view of complex traits: From polygenic to omnigenic. Cell. 169:1177–1186.

Bradshaw HD, Wilbert SM, Otto KG, Schemske DW. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). Nature. 376:762–765.

Breen MS, Kemena C, Vlasov PK, Notredame C, Kondrashov FA. 2012. Epistasis as the primary factor in molecular evolution. Nature. 490:535–538.

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

Charras-Garrido M, Lezaud P. 2013. Extreme value analysis: An introduction. Journal de la SFdS. 154:66–97.

Claringbould A, de Klein N, Franke L. 2017. The genetic architecture of molecular traits. Curr. Opin. Syst. Biol. 1:25–31.

Connallon T, Clark AG. 2015. The distribution of fitness effects in an uncertain world. Evolution. 69:1610–1618.

Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S *et al*. 2010. The genetic landscape of a cell. Science. 327:425–431.

de Visser JAG, Krug J. 2014. Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet. 15:480–490.

Delignette-Muller ML, Dutang C. 2015. fitdistrplus: An R package for fitting distributions. J. Stat. Softw. 64:1–34.

Doane AS, Elemento O. 2017. Regulatory elements in molecular networks. WIREs Systems Biology and Medicine. 9:e1374.

Dorweiler J, Stec A, Kermicle J, Doebley J. 1993. *Teosinte Glume Architecture 1*: A genetic locus controlling a key step in maize evolution. Science. 262:233–235.

Draghi JA, Parsons TL, Plotkin JB. 2011. Epistasis increases the rate of conditionally neutral substitution in an adapting population. Genetics. 187:1139–1152.

Erwin DH, Davidson EH. 2009. The evolution of hierarchical gene regulatory networks. Nat Rev Genet. 10:141–148.

Eyre-Walker A. 2010. Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. Proceedings of the National Academy of Sciences. 107:1752–1756.

Eyre-Walker A, Keightley PD. 2007. The distribution of fitness effects of new mutations. Nat Rev Genet. 8:610–618.

Falconer DSM. 1996. *Introduction to Quantitative Genetics*. Pearson Education Limited. Longmans Green, Harlow, Essex, UK. fourth edition.

Fisher RA. 1918. The correlation between relatives on the supposition of Mendelian inheritance. Trans. R. Soc. Edinburgh. 52:399–433.

Fisher RA. 1923. XXI.—On the dominance ratio. Proc. R. Soc. Edinb. 42:321–341.

Fisher RA. 1930. *The Genetical Theory of Natural Selection*. The Clarendon press. Oxford, UK.

Fisher RA. 1936. Statistical methods for research workers. Statistical methods for research workers.. .

Fisher RA, Tippett LHC. 1928. Limiting forms of the frequency

distribution of the largest or smallest member of a sample. Math. Proc. Camb. Philos. Soc. 24:180–190.

Foll M, Poh YP, Renzette N, Ferrer-Admetlla A, Bank C, Shim H, Malaspinas AS, Ewing G, Liu P, Wegmann D *et al*. 2014. Influenza virus drug resistance: A time-sampled population genetics perspective. PLOS Genetics. 10:e1004185.

François P. 2014. Evolving phenotypic networks in silico. Semin. Cell Dev. Biol. 35:90–97.

Gavrilets S. 1997. Evolution and speciation on holey adaptive landscapes. Trends in Ecology & Evolution. 12:307–312.

Gillespie JH. 1983. A simple stochastic gene substitution model. Theor. Popul. Biol. 23:202–215.

Gillespie JH. 1984. Molecular evolution over the mutational landscape. Evolution. 38:1116–1129.

Gillespie JH. 1991. *The Causes of Molecular Evolution*. University Press.

Halfon MS. 2020. Silencers, enhancers, and the multifunctional regulatory genome. Trends in Genetics. 36:149–151.

Haller BC, Messer PW. 2019. SLiM 3: Forward genetic simulations beyond the Wright-Fisher model. Mol Biol Evol. 36:632–637.

Halligan DL, Kousathanas A, Ness RW, Harr B, Eöry L, Keane TM, Adams DJ, Keightley PD. 2013. Contributions of protein-coding and regulatory change to adaptive molecular evolution in murid rodents. PLOS Genetics. 9:e1003995.

Hallinan J, Jackway P. 2005. Network Motifs, Feedback Loops and the Dynamics of Genetic Regulatory Networks. In: *2005 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology*. pp. 1–7.

Hansen TF. 2013. Why epistasis is important for selection and adaptation. Evolution. 67:3501–3511.

Hether TD, Hohenlohe PA. 2014. Genetic Regulatory Network Motifs Constrain Adaptation Through Curvature in the Landscape of Mutational (co)variance. Evolution. 68:950–964.

Hill WG, Robertson A. 1966. Effect of linkage on limits to artificial selection. Genet Res. 8:269–294.

Huang W, Mackay TFC. 2016. The genetic architecture of quantitative traits cannot be inferred from variance component analysis. PLOS Genetics. 12:e1006421.

Imhof M, Schlötterer C. 2001. Fitness effects of advantageous mutations in evolving *Escherichia coli* populations. PNAS. 98:1113–1117.

Joanito I, Chu JW, Wu SH, Hsu CP. 2018. An incoherent feed-forward loop switches the Arabidopsis clock rapidly between two hysteretic states. Sci Rep. 8:13944.

Joyce P, Rokyta DR, Beisel CJ, Orr HA. 2008. A general extreme value theory model for the adaptation of DNA sequences under strong selection and weak mutation. Genetics. 180:1627–1643.

Karlebach G, Shamir R. 2008. Modelling and analysis of gene regulatory networks. Nat Rev Mol Cell Biol. 9:770–780.

Kassen R, Bataillon T. 2006. Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. Nat Genet. 38:484–488.

Kauffman S, Levin S. 1987. Towards a general theory of adaptive walks on rugged landscapes. J. Theor. Biol. 128:11–45.

Kemper KE, Visscher PM, Goddard ME. 2012. Genetic architecture of body size in mammals. Genome Biol. 13:244.

Koch EN, Costanzo M, Bellay J, Deshpande R, Chatfield-Reed K, Chua G, D'Urso G, Andrews BJ, Boone C, Myers CL. 2012. Conserved rules govern genetic interaction degree across species. Genome Biology. 13:R57.

Koonin EV. 2016. Splendor and misery of adaptation, or the importance of neutral null for understanding evolution. BMC Biology. 14:114.

Kopp M, Hermisson J. 2009. The genetic basis of phenotypic adaptation II: The distribution of adaptive substitutions in the moving optimum model. Genetics. 183:1453–1476.

Kosheleva K, Desai MM. 2018. Recombination alters the dynamics of adaptation on standing variation in laboratory yeast populations. Mol Biol Evol. 35:180–201.

Kousathanas A, Keightley PD. 2013. A comparison of models to infer the distribution of fitness effects of new mutations. Genetics. 193:1197–1208.

Kozuch BC, Shaffer MG, Culyba MJ. 2020. The Parameter-Fitness Landscape of lexA Autoregulation in Escherichia coli. mSphere. 5:e00718–20.

Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution. 30:314–334.

Lande R, Shannon S. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. Evolution. 50:434–437.

Lebeuf-Taylor E, McCloskey N, Bailey SF, Hinz A, Kassen R. 2019. The distribution of fitness effects among synonymous mutations in a gene under directional selection. eLife. 8:e45952.

Lenth RV. 2023. *Emmeans: Estimated Marginal Means, Aka Least-Squares Means*.

Liu X, Li YI, Pritchard JK. 2019. Trans effects on gene expression can drive omnigenic inheritance. Cell. 177:1022–1034.e6.

Lynch M. 2010. Evolution of the mutation rate. Trends in Genetics. 26:345–352.

MacLean RC, Buckling A. 2009. The distribution of fitness effects of beneficial mutations in *Pseudomonas aeruginosa*. PLOS Genetics. 5:e1000406.

Mancera E, Nocedal I, Hammel S, Gulati M, Mitchell KF, Andes DR, Nobile CJ, Butler G, Johnson AD. 2021. Evolution of the complex transcription network controlling biofilm formation in *Candida* species. eLife. 10:e64682.

Mangan S, Alon U. 2003. Structure and function of the feed-forward loop network motif. Proceedings of the National Academy of Sciences. 100:11980–11985.

Martin G, Lenormand T. 2008. The distribution of beneficial and fixed mutation fitness effects close to an optimum. Genetics. 179:907–16.

Matuszewski S, Hermisson J, Kopp M. 2015. Catch me if you can: Adaptation from standing genetic variation to a moving phenotypic optimum. Genetics. 200:1255–1274.

Maynard Smith J. 1970. Natural selection and the concept of a protein space. Nature. 225:563–564.

Maynard Smith J, Burian R, Kauffman S, Alberch P, Campbell J, Goodwin B, Lande R, Raup D, Wolpert L. 1985. Developmental constraints and evolution: A perspective from The

Mountain Lake Conference on Development and Evolution. Q. Rev. Biol.. 60:265–287.

Morimoto D, Walinda E, Fukada H, Sugase K, Shirakawa M. 2016. Ubiquitylation directly induces fold destabilization of proteins. Sci Rep. 6:39453.

Neidhart J, Szendro IG, Krug J. 2014. Adaptation in Tunably Rugged Fitness Landscapes: The Rough Mount Fuji Model. Genetics. 198:699–721.

Nevozhay D, Adams RM, Murphy KF, Josić K, Balázsi G. 2009. Negative autoregulation linearizes the dose–response and suppresses the heterogeneity of gene expression. PNAS. 106:5123–5128.

Ohta T. 2002. Near-neutrality in evolution of genes and gene regulation. PNAS. 99:16134–16137.

Orr HA. 1998. The population genetics of adaptation: The distribution of factors fixed during adaptive evolution. Evolution. 52:935–949.

Orr HA. 2000. Adaptation and the cost of complexity. Evolution. 54:13–20.

Orr HA. 2003a. The distribution of fitness effects among beneficial mutations. Genetics. 163:1519–1526.

Orr HA. 2003b. A minimum on the mean number of steps taken in adaptive walks. J. Theor. Biol.. 220:241–247.

Orr HA. 2010. The population genetics of beneficial mutations. Philos. Trans. R. Soc. Lond., B, Biol. Sci.. 365:1195–1201.

Ortiz-Barrientos D, Engelstädter J, Rieseberg LH. 2016. Recombination rate evolution and the origin of species. TREE. 31:226–236.

Pigolotti S, Krishna S, Jensen MH. 2007. Oscillation patterns in negative feedback loops. Proceedings of the National Academy of Sciences. 104:6533–6537.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Rice SH. 1998. The evolution of canalization and the breaking of Von Baer's laws: Modeling the evolution of development with epistasis. Evolution. 52:647–656.

Ritz KR, Noor MAF, Singh ND. 2017. Variation in recombination rate: Adaptive or not? Trends in Genetics. 33:364–374.

Rokyta DR, Beisel CJ, Joyce P, Ferris MT, Burch CL, Wichman HA. 2008. Beneficial fitness effects are not exponential for two viruses. J Mol Evol. 67:368–376.

Romero PA, Arnold FH. 2009. Exploring protein fitness landscapes by directed evolution. Nat Rev Mol Cell Biol. 10:866–876.

Rosenfeld N, Elowitz MB, Alon U. 2002. Negative autoregulation speeds the response times of transcription networks. J Mol Biol. 323:785–793.

Salvador-Martínez I, Coronado-Zamora M, Castellano D, Barbadilla A, Salazar-Ciudad I. 2018. Mapping Selection within Drosophila melanogaster Embryo's Anatomy. Molecular Biology and Evolution. 35:66–79.

Sanjuán R, Moya A, Elena SF. 2004. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. PNAS. 101:8396–8401.

Schenk MF, Szendro IG, Krug J, de Visser JAGM. 2012. Quantifying the adaptive potential of an antibiotic resistance enzyme. PLOS Genetics. 8:e1002783.

Schlitt T, Brazma A. 2007. Current approaches to gene regulatory network modelling. BMC Bioinformatics. 8:S9.

Schmidtke G, Kalveram B, Weber E, Bochtler P, Lukasiak S, Hipp MS, Groettrup M. 2006. The UBA domains of NUB1L are required for binding but not for accelerated degradation of the ubiquitin-like modifier FAT10 *. JBC. 281:20045–20054.

Seo CH, Kim JR, Kim MS, Cho KH. 2009. Hub genes with positive feedbacks function as master switches in developmental gene regulatory networks. Bioinformatics. 25:1898–1904.

Shen-Orr SS, Milo R, Mangan S, Alon U. 2002. Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet. 31:64–68.

Siegal ML, Promislow DEL, Bergman A. 2007. Functional and evolutionary inference in gene networks: Does topology matter? Genetica. 129:83–103.

Simpson GG. 1944. *Tempo and Mode in Evolution.*. Columbia University Press, N.Y.

Slatkin M. 1987. Quantitative genetics of heterochrony. Evolution. 41:799–811.

Snell P, Grimberg Å, Carlsson AS, Hofvander P. 2019. WRINKLED1 Is Subject to Evolutionary Conserved Negative Autoregulation. Frontiers in Plant Science. 10.

Städter P, Schälte Y, Schmiester L, Hasenauer J, Stapor PL. 2021. Benchmarking of numerical integration methods for ODE models of biological systems. Sci Rep. 11:2696.

Stewart AJ, Seymour RM, Pomiankowski A, Reuter M. 2013. Under-Dominance Constrains the Evolution of Negative Autoregulation in Diploids. PLOS Computational Biology. 9:e1002992.

Thornton KR. 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.

Vale PF, Choisy M, Froissart R, Sanjuán R, Gandon S. 2012. The distribution of mutational fitness effects of phage Φx174 on different hosts. Evolution. 66:3495–3507.

Visscher PM, Brown MA, McCarthy MI, Yang J. 2012. Five years of GWAS discovery. The American Journal of Human Genetics. 90:7–24.

Wade MJ, Winther RG, Agrawal AF, Goodnight CJ. 2001. Alternative definitions of epistasis: Dependence and interaction. TREE. 16:498–504.

Wagner A. 1996. Does evolutionary plasticity evolve? Evolution. 50:1008–1023.

Wagner A. 2011. Genotype networks shed light on evolutionary constraints. TREE. 26:577–584.

Walsh B, Lynch M. 2018a. *Evolution and Selection of Quantitative Traits*. Oxford University Press. New York, NY, USA.

Walsh B, Lynch M. 2018b. Long-term response: 3. Adaptive walks, In: *Evolution and Selection of Quantitative Traits*, Walsh B, Lynch M, editors, Oxford University Press. pp. 991–1013.

Walsh B, Lynch M. 2018c. Maintenance of quantitative genetic variation, In: *Evolution and Selection of Quantitative Traits*, Walsh B, Lynch M, editors, Oxford University Press. pp. 1016–1078.

Weinreich DM, Delaney NF, DePristo MA, Hartl DL. 2006. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science. 312:111–114.

Wickham H. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Use R!. Springer-Verlag New York.

Williamson RJ, Josephs EB, Platts AE, Hazzouri KM, Haudry A, Blanchette M, Wright SI. 2014. Evidence for widespread positive and negative selection in coding and conserved noncoding regions of *Capsella grandiflora*. PLOS Genetics. 10:e1004622.

Wright S. 1931. Evolution in Mendelian populations. Genetics. 16:97–159.

Wright S. 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the XI International Congress of Genetics. 8:209–222.

Wu X, Bayle JH, Olson D, Levine AJ. 1993. The p53-mdm-2 autoregulatory feedback loop. Genes Dev.. 7:1126–1132.

Xiang Y, Gubian S, Suomela B, Hoeng J. 2013. Generalized simulated annealing for efficient global optimization: The GenSA package for R. The R Journal. 5/1.

Xu L, Chen H, Hu X, Zhang R, Zhang Z, Luo ZW. 2006. Average gene length is highly conserved in prokaryotes and eukaryotes and diverges only between the two kingdoms. Mol Biol Evol. 23:1107–8.

Zeyl C, DeVisser JAGM. 2001. Estimates of the rate and distribution of fitness effects of spontaneous mutation in *Saccharomyces cerevisiae*. Genetics. 157:53–61.

Supplementary material

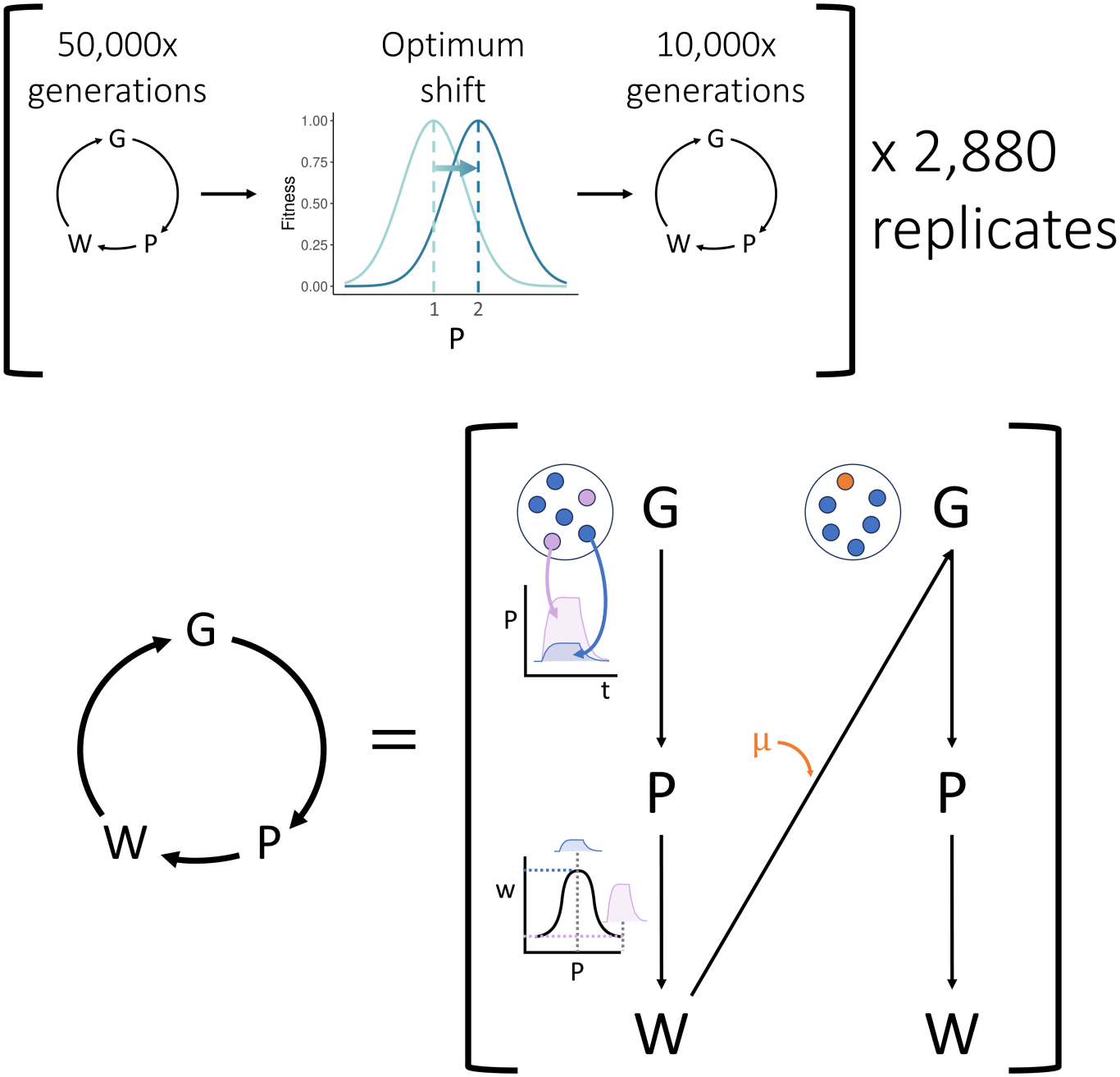
S1 Appendix. A hierarchical molecular network model for quantitative traits. An extension of the negative autoregulation model to a general form to model quantitative traits under the control of an arbitrary genetic network.

Supplementary Table 1 Model symbols and parameters.

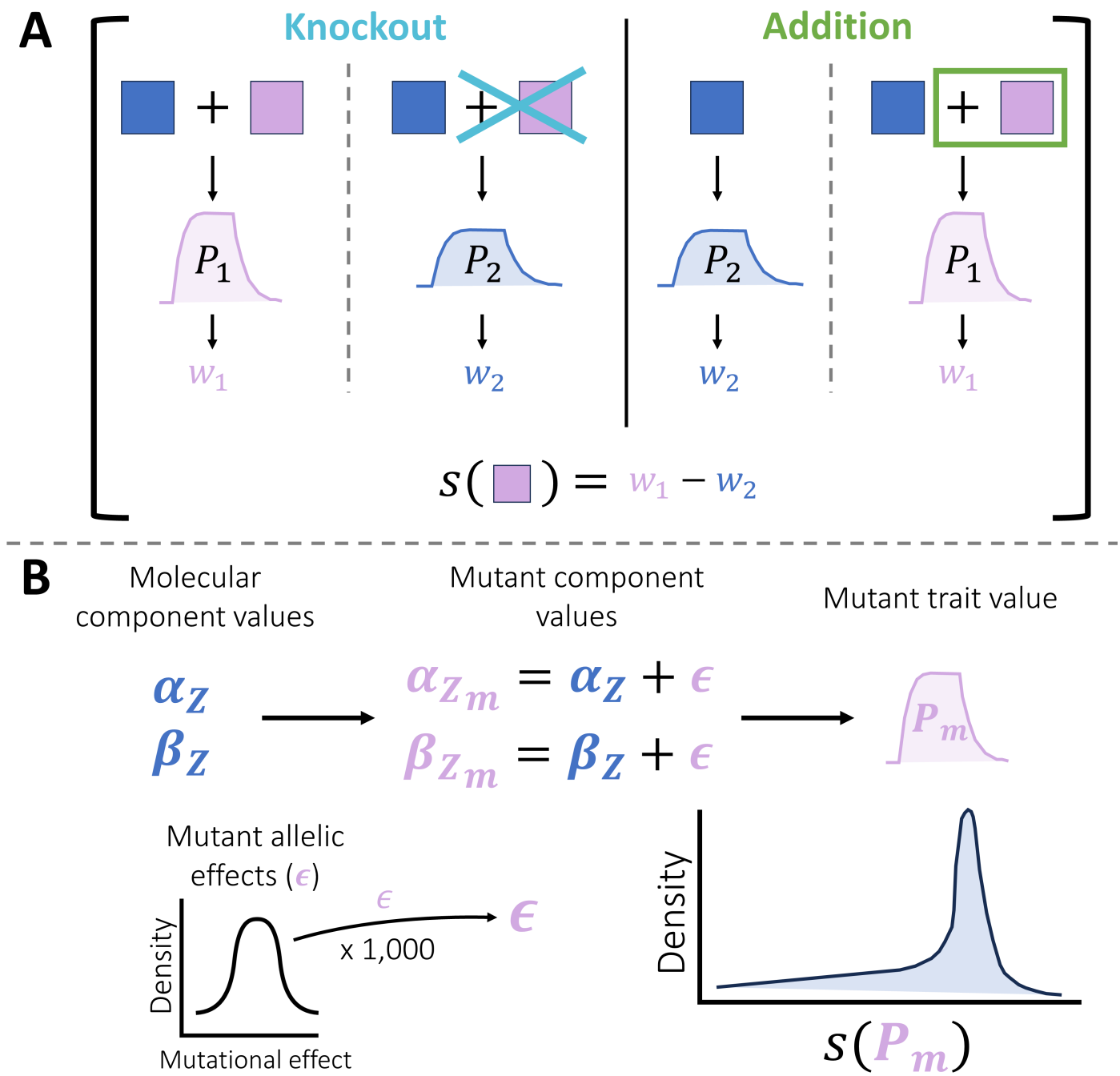
Symbol	Parameter	Description
Molecular components		
α_Z	Z removal rate	The rate at which Z product is removed from the cell.
β_Z	Z production rate	The rate at which Z is produced.
K_Z	Repression coefficient	The Z product concentration at which further Z expression is reduced by half. Fixed at $K_Z = 1$.
K_{XZ}	Activation coefficient	The X product concentration at which Z expression is half the maximum. Fixed at $K_{XZ} = 1$.
n_Z	Hill repression coefficient	The steepness coefficient of the Z repression curve. Fixed at $n_Z = 8$.
n_{XZ}	Hill activation coefficient	The steepness coefficient of the Z activation curve. Fixed at $n_{XZ} = 8$.
Integration parameters		
t_{start}	X activation time	The left boundary where $X = 1$ when solving the NAR ODE. Before this point, $X = 0$.
t_{stop}	X deactivation time	The right boundary where $X = 1$ when solving the NAR ODE. After this point, $X = 0$.
t_{max}	Development time	The amount of time for which $Z(t)$ is solved for. Models the developmental time of the trait in an individual.
Model abbreviations/definitions		
$mQTL$	Molecular quantitative trait locus	A QTL for a molecular component.
$\mathbf{C}$	Molecular components	A vector of ODE coefficients contributing to Z expression. Includes α_Z and β_Z.
w	Relative fitness	Relative fitness of an individual.
$\mathbf{P_O}$	Phenotypic optimum	The phenotype where $w = 1$.
σ_w	Fitness function width	The width of the Gaussian fitness function. A measure of selection strength.
a	Allelic effect	The allelic effect of an mQTL on a molecular component. Note this is different from s, which is the effect of an allele on fitness. It also is not a measure of the allelic effect on the phenotype due to the ODE intermediate between a and the phenotype.
s	Selection coefficient	The effect of a mutation on fitness.

Supplementary Table 2 Mean generalized Pareto distribution (GPD) parameters fit to mutant screen distributions of fitness effects over adaptive walks in additive and network populations. Brackets indicate 95% confidence intervals. Parameters were estimated by fitting a GPD to a random sample of mutations from the mutant screen experiments conducted on each replicate at each adaptive step. κ is the shape parameter of the GPD. $\bar{\kappa}$ is the mean κ value across n replicate simulations. The log-likelihood ratio tests the null hypothesis that a sample of beneficial mutations belongs to an exponential distribution (which meets $\kappa = 0$). P-values across the n replicates were combined using Fisher's method.

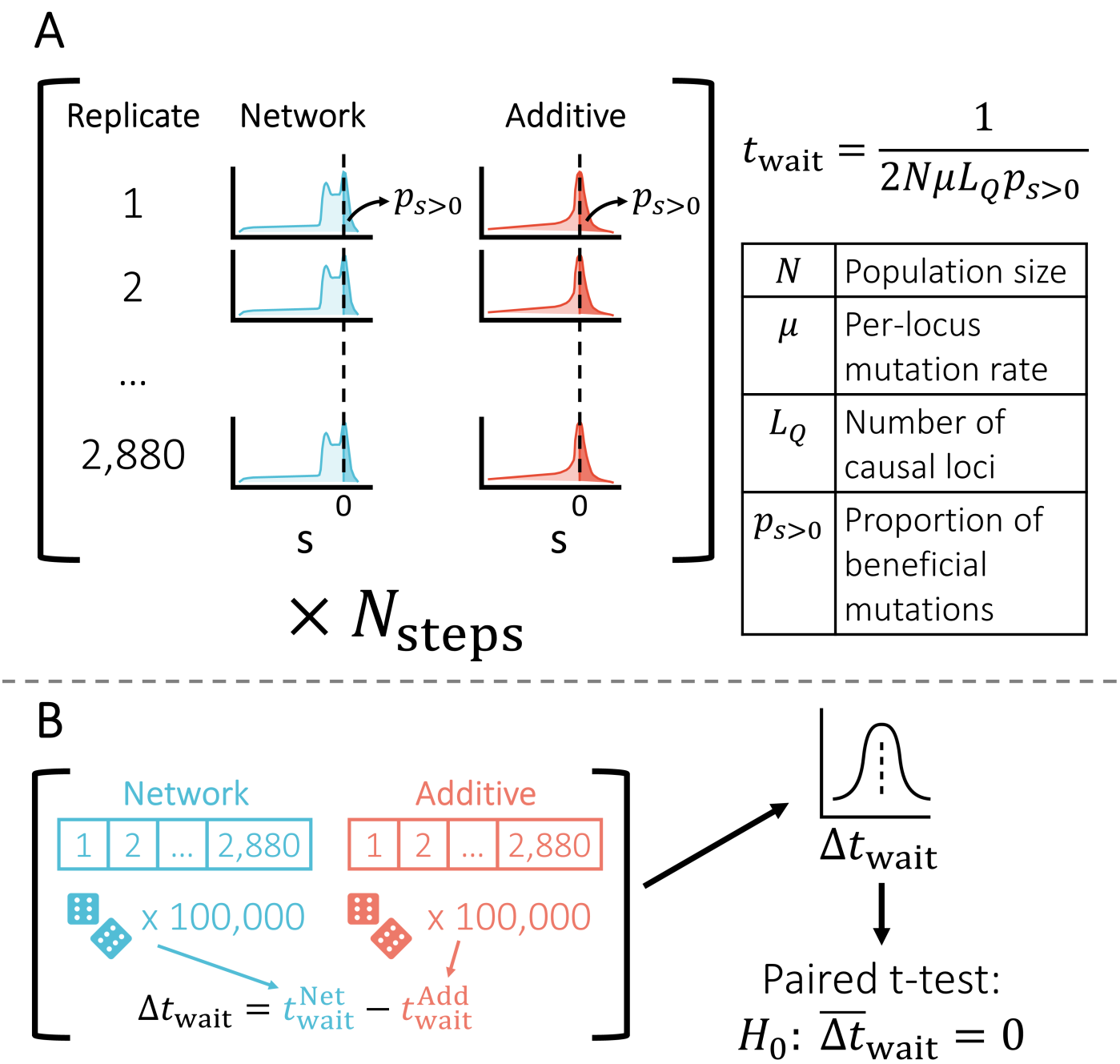
Additive model				
Adaptive Step	n	$\bar{\kappa}$	**Log-likelihood ratio**	**p-value**
Before optimum shift	1213	$-2.12(\pm 0.04)$	$0.73(\pm 0.003)$	< 0.001
1	1110	$-2.10(\pm 0.04)$	$0.83(\pm 0.004)$	< 0.001
2	333	$-2.02(\pm 0.08)$	$0.87(\pm 0.006)$	< 0.001
≥ 3	59	$-1.85(\pm 0.11)$	$0.89(\pm 0.013)$	< 0.001
Network model				
Adaptive Step	n	$\bar{\kappa}$	**Log-likelihood ratio**	**p-value**
Before optimum shift	950	$-1.88(\pm 0.03)$	$0.77(\pm 0.002)$	< 0.001
1	912	$-2.38(\pm 0.25)$	$0.83(\pm 0.004)$	< 0.001
2	270	$-4.34(\pm 1.44)$	$0.86(\pm 0.006)$	< 0.001
≥ 3	48	$-3.81(\pm 2.00)$	$0.88(\pm 0.012)$	< 0.001



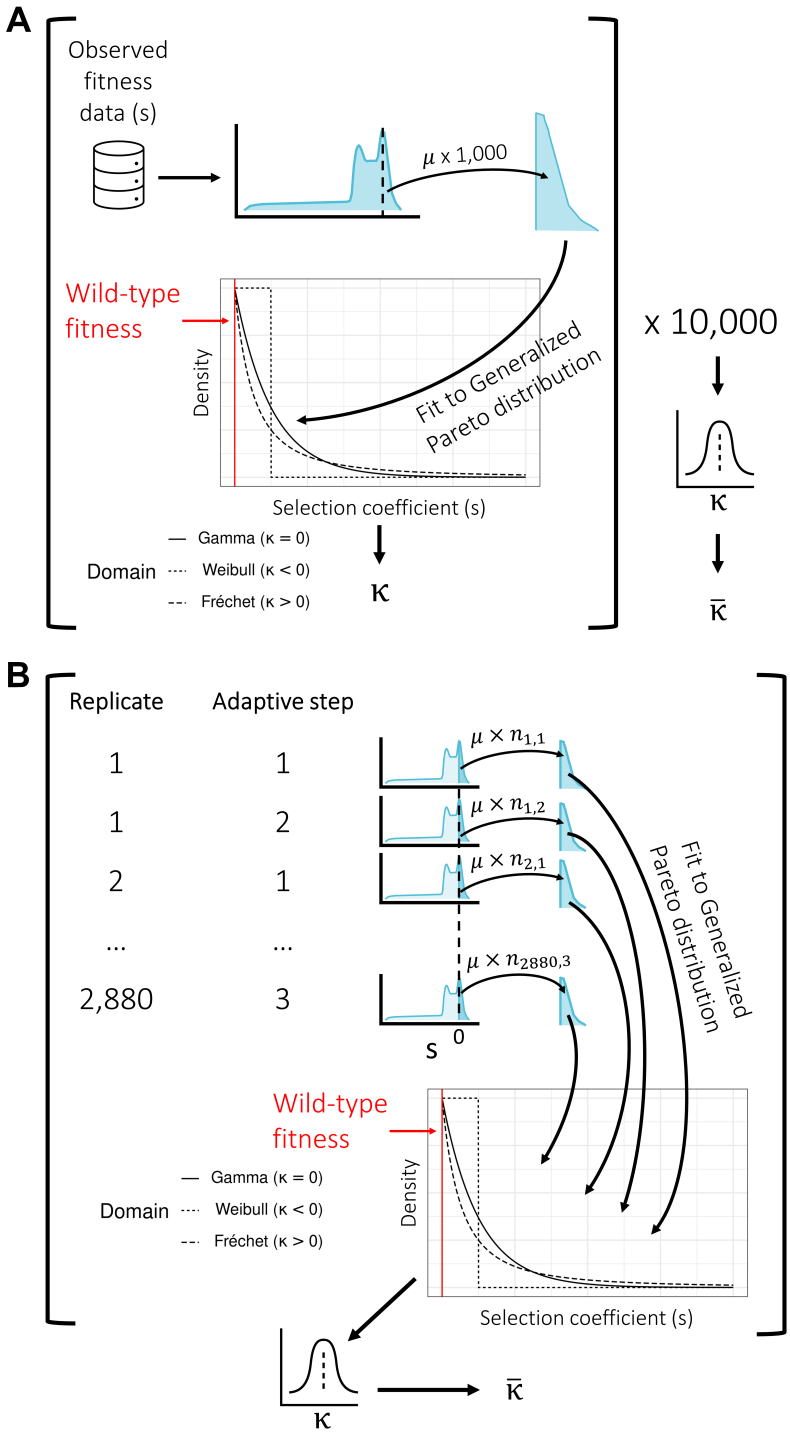
Supplementary Figure 1 Flowchart of SLiM simulation procedure. SLiM simulates a Wright-Fisher process to model evolutionary change through a genotype-phenotype-fitness (GPW) map. Our simulation began with 50,000 generations of burn-in to ensure populations were adapted to the environment. We then shifted the phenotypic optimum and adaptation was tracked for a further 10,000 generations. This process was repeated 2,880 times per model for replication purposes (i.e. there were 2,880 replicates of network and additive adaptive walks for a total 5,760 simulations). The GPW map consisted of several stages: first, the genotype was translated to phenotype via summing genetic effects at QTLs (additive models; not shown) or by solving a system of ordinary differential equations (network models, P/t figure shows solutions to the differential equation for the blue and purple genotypes). Phenotype was then translated to fitness by a stabilizing selection fitness function (shown by the w/P figure – the purple phenotype has lower fitness than the blue phenotype). The fitness value then influenced the chance that an individual was sampled as a parent for the next generation. After parents were chosen, random mutations (μ) could occur to introduce further genotypic and phenotypic variation in the next generation (shown in orange).



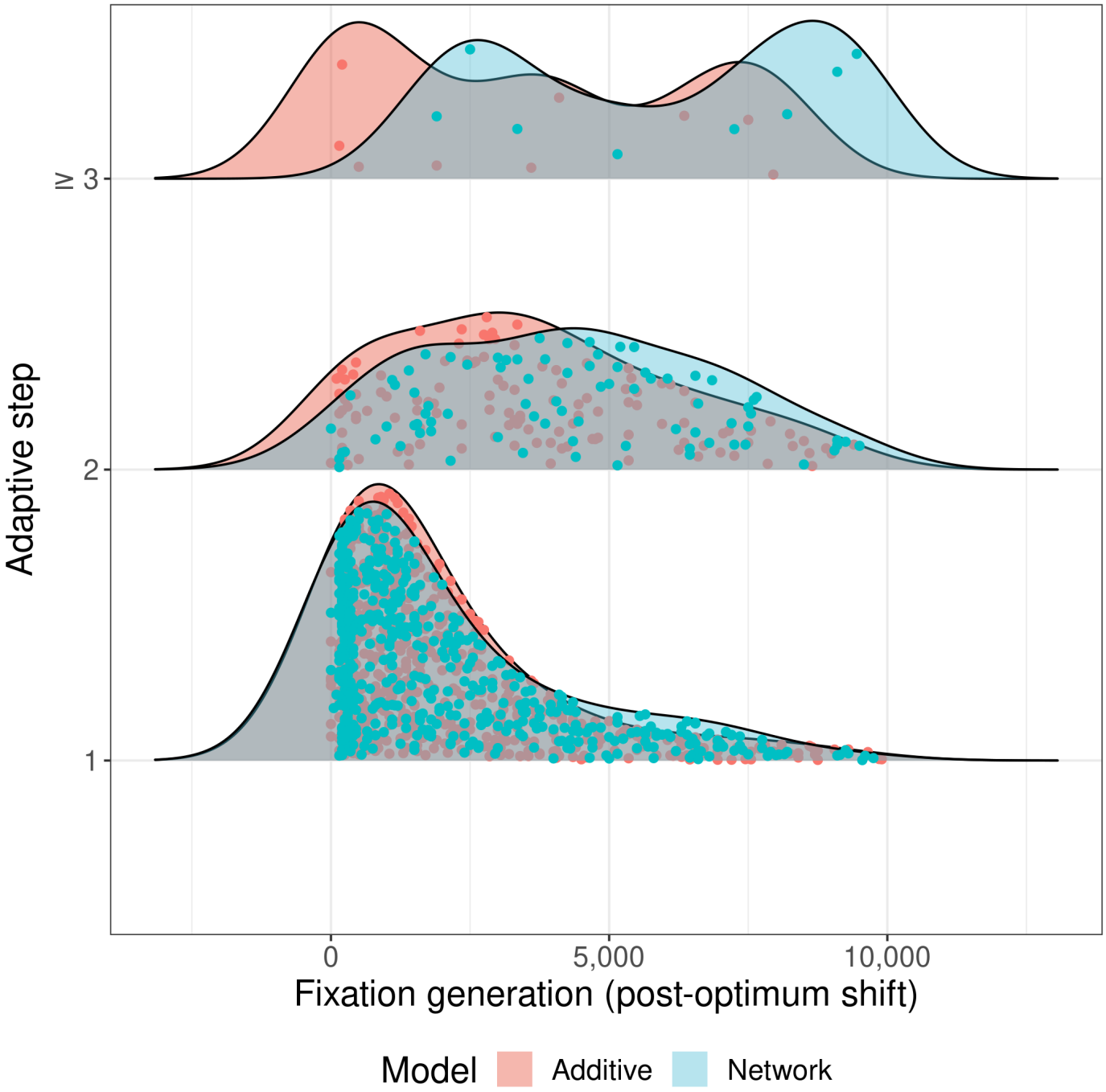
Supplementary Figure 2 Flowchart describing how selection coefficients and the distribution of fitness effects (DFE) among new mutations were estimated. (A) Consider two alleles contributing to the phenotype (blue and purple boxes) at either the same or different loci. Individuals with both alleles have some phenotype (P_1), which gives rise to a fitness (w_1). By removing the purple allele and recalculating the phenotype, we achieve a different phenotype (P_2) and fitness (w_2). The difference between w_1 and w_2 represents the selection coefficient (s) of the purple allele. This difference can be measured either through addition – adding the purple allele to the genotype, or by knockout (removing the purple allele from the genotype). (B) To estimate the distribution of fitness effects (DFE) among new mutations, we conducted a mutation screen experiment. We generated mutants by taking 1,000 samples (ϵ) from a standard normal distribution and adding those to the molecular component values from the SLiM simulations (α_Z and β_Z). Mutant phenotypes ($\mathbf{P_m}$) were calculated by inputting the mutant component values into an ordinary differential equation and solving it. The 1,000 samples were independently added to each molecular component to measure the DFE of both components. We then calculated the selection coefficients of ϵ by the addition method in (A). P_m represents P_1 in (A), whilst the phenotypes without the added ϵ represents P_2. We then plotted the joint distribution of s across all sampled ϵ.



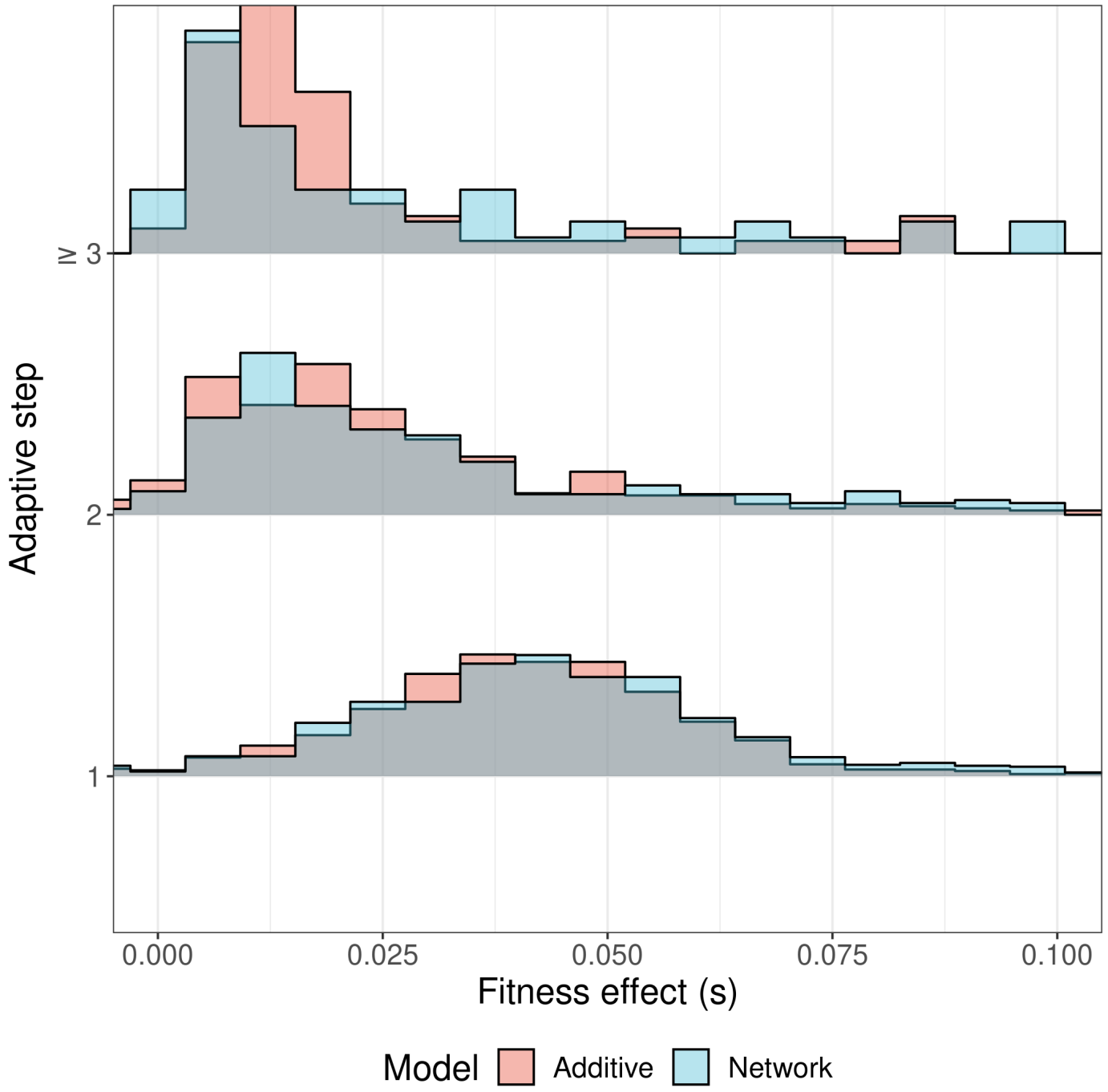
Supplementary Figure 3 Flowchart describing how waiting times to new beneficial mutations were calculated. (A) For each of the 2,880 replicates, we took the distribution of s calculated during the mutation screen experiments and extracted the proportion of new mutations with $s > 0$, $p_{s>0}$. This is the area under the curve shown in dark blue/red (to the right of the dashed line). This was repeated for each adaptive step (N_{steps} within a replicate (i.e. replicates with more than one adaptive step had $p_{s>0}$ calculated for each adaptive step). The waiting time to a new beneficial mutation for a given replicate at a given step was calculated as $t_{\text{wait}} = 1/(2N\mu L_Q p_{s>0})$, where N is the population size (multiplied by 2 because we simulate a diploid organism), μ is the per-locus, per-generation mutation rate, and L_Q is the number of causal loci. (B) To calculate the difference between models in the waiting time to a new beneficial mutation, we ran a bootstrap analysis. We sampled 100,000 random pairs of additive and network models calculating t_{wait} for each. We calculated the difference between their waiting times, Δt_{wait}. The random sampling gave a distribution of differences between network and additive waiting times. We used a paired t-test to determine if the mean difference in waiting times between models, $\bar{\Delta} t_{\text{wait}}$ was not zero (and hence there was a difference between models).



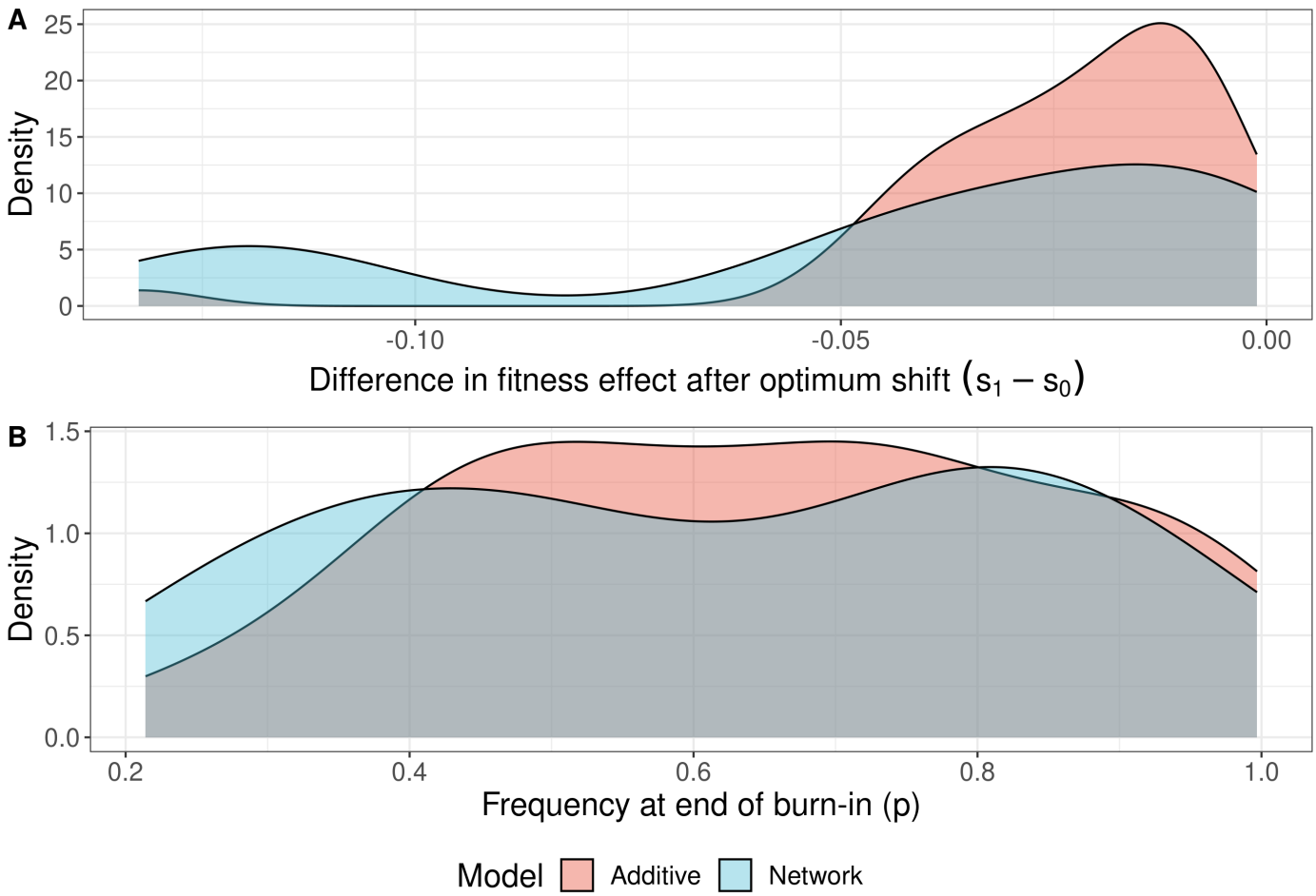
Supplementary Figure 4 Flowchart describing how Beisel *et al.*'s (2007) method was used to estimate the shape of the distribution of fitness effects (DFE) among beneficial mutations via two approaches. (A) Using the pooled distribution of s across all simulations and adaptive steps that we calculated during the mutation screen experiments, we sampled 1,000 beneficial mutations (where $s > 0$) to generate a distribution of beneficial mutations. We then fit this to a generalized Pareto distribution (GPD) using the method outlined by Beisel *et al.* (2007). From the fit, we extracted the shape parameter of the fit, κ. This process was repeated 10,000 times to generate a distribution of κ. We treated the mean κ, $\bar{\kappa}$ as the estimate for the shape parameter. (B) We sampled $n_{x,y}$ mutations from each simulation's s distribution at each adaptive step. The s distributions were generated during the mutation screen experiments. $n_{x,y}$ was $\min(n^{x,y}_{s>0}, 100)$, where $n^{x,y}_{s>0}$ was the number of mutant screen alleles with $s > 0$ in replicate x and adaptive step y. The sampled distributions were fit to a GPD as per (A), generating a distribution of κ from each replicate and adaptive step. The mean was again treated as the estimate for the shape parameter.



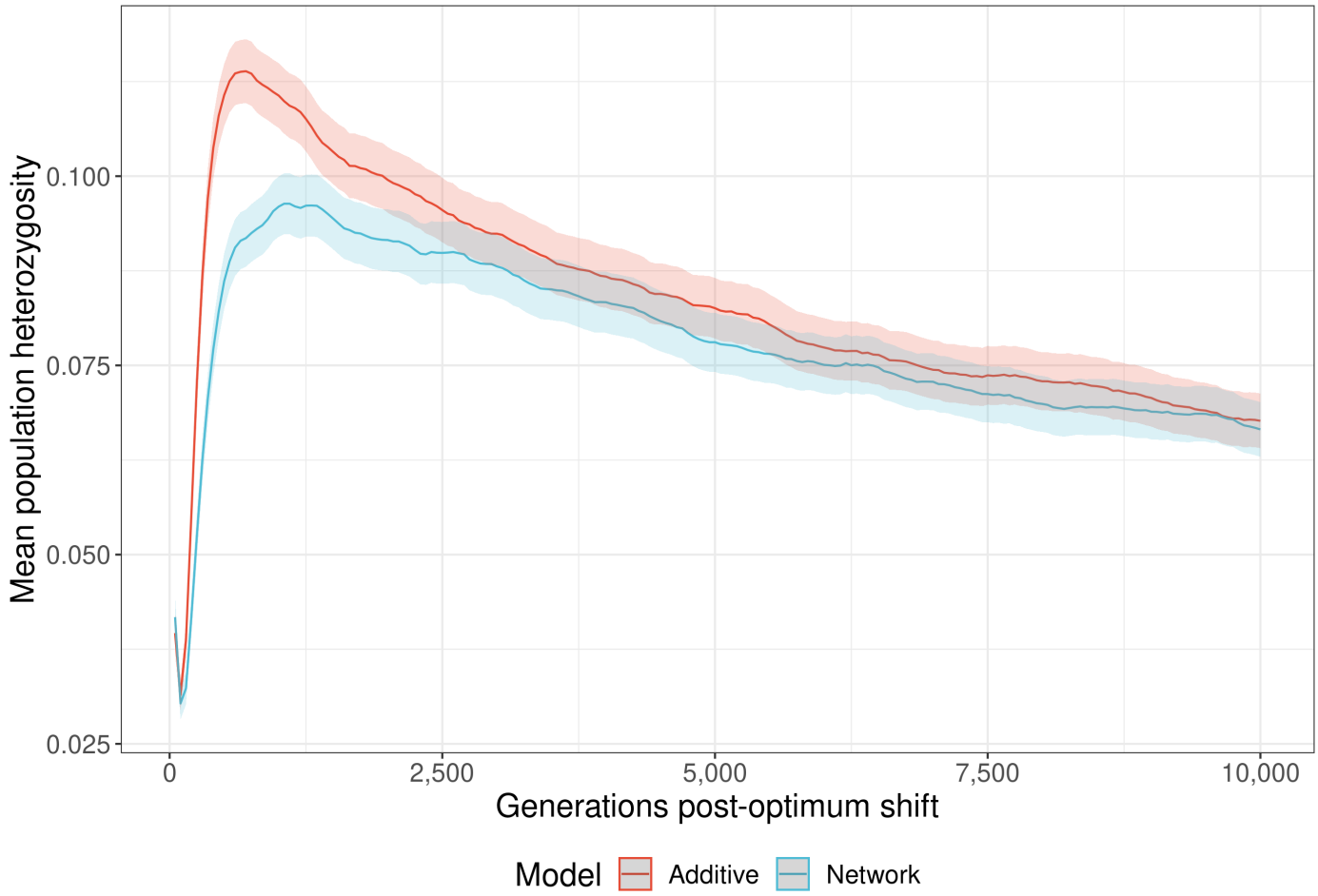
Supplementary Figure 5 Distribution of generations at which mutations fixed in adapting additive (red) and network (blue) populations at each adaptive step. There was no significant difference between populations at adaptive step 1. Generations are given relative to the optimum shift so that Generation 0 is when the optimum shifted.



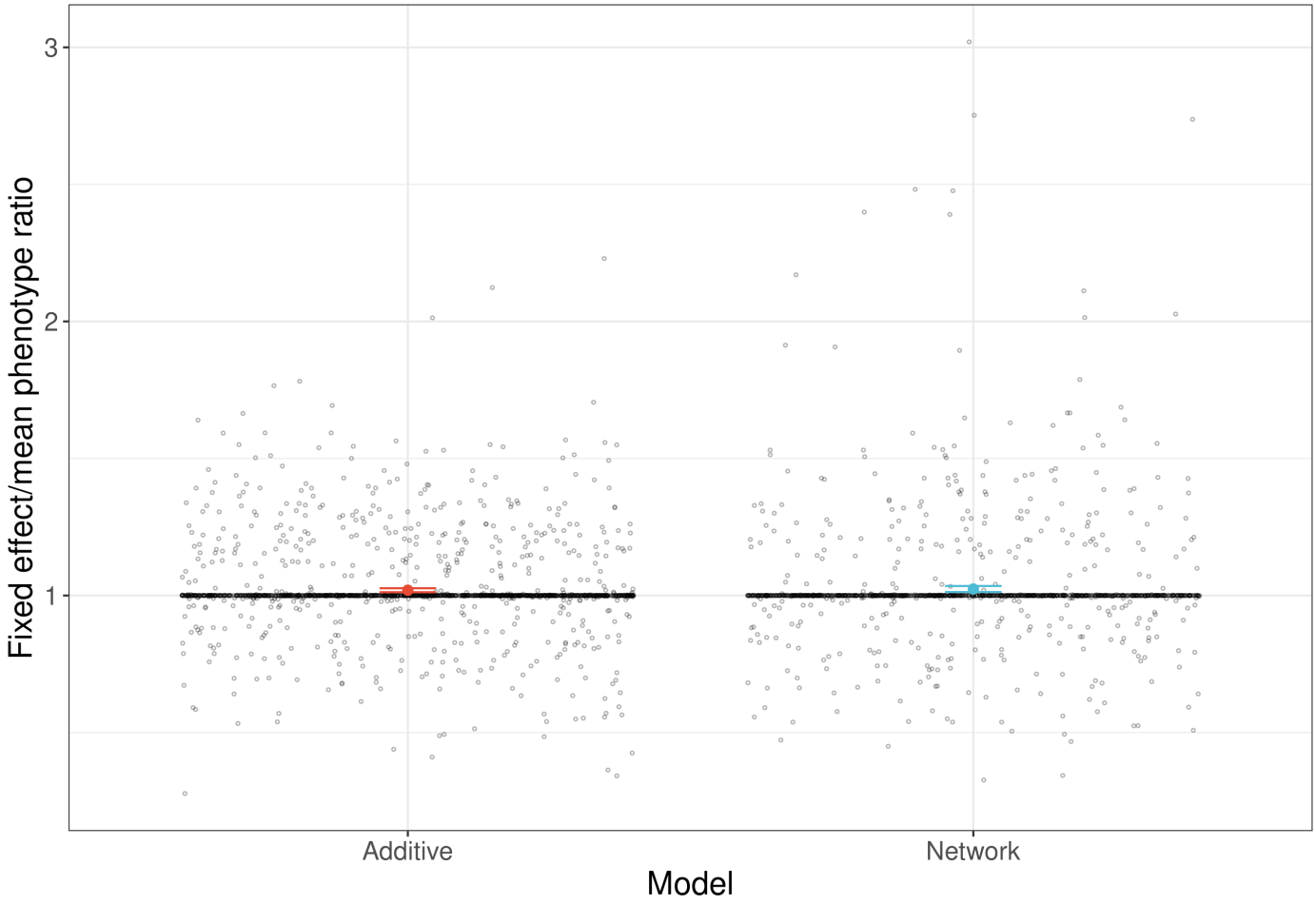
Supplementary Figure 6 Zoomed view of Fig 4B. The mode of the distribution is clearly greater than 0, possibly due to small effect alleles being more susceptible to loss by drift.



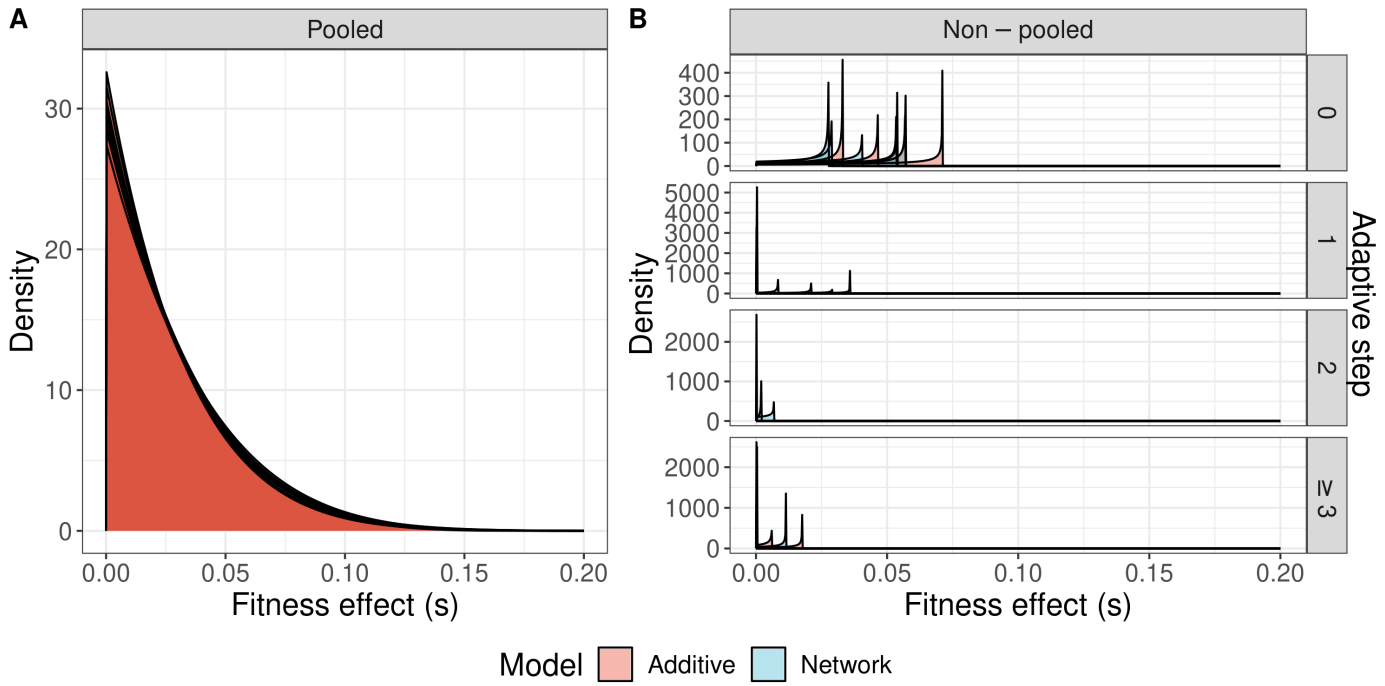
Supplementary Figure 7 Behavior of deleterious fixations in additive (red) and network (blue) models. (A) The selection coefficient, s, decreases after the optimum shift at generation 50,000. (B) The allele frequency, p, of the fixations at the optimum shift.



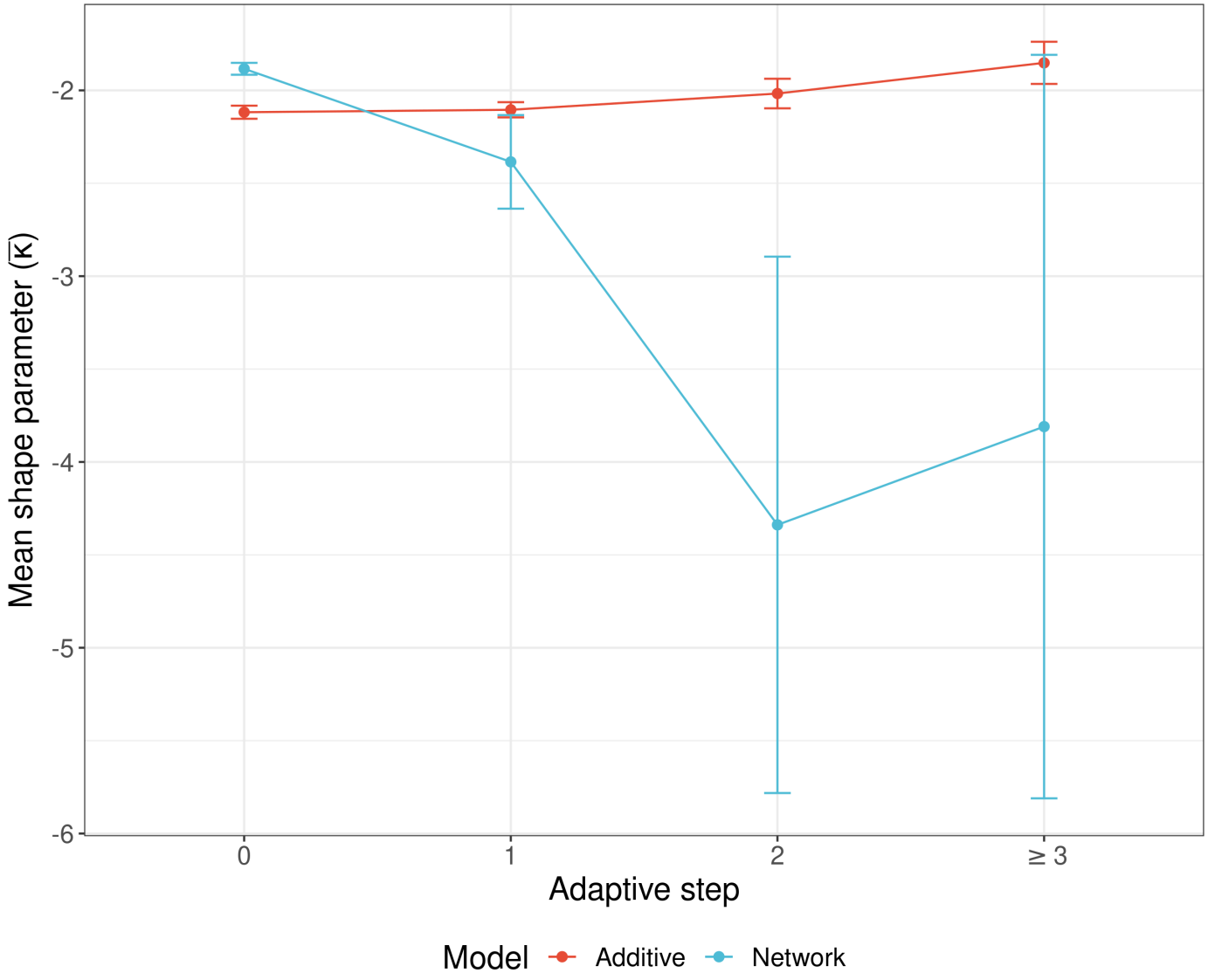
Supplementary Figure 8 Mean observed heterozygosity in additive (red) and network (blue) populations over the adaptation period. The mean is taken over 2,880 replicates per group. Heterozygosity was measured at the two QTLs (mQTLs in NAR models) in both models.



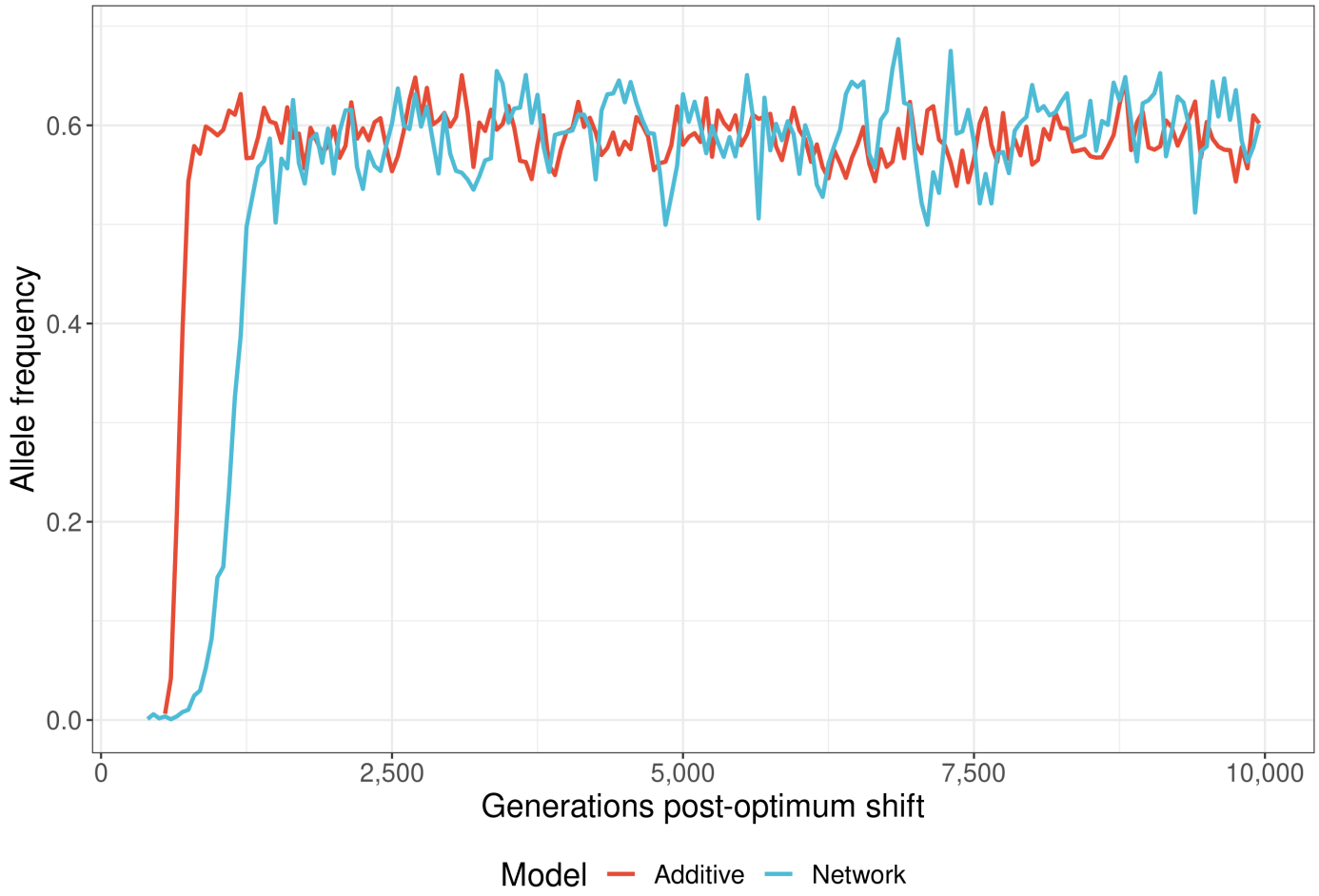
Supplementary Figure 9 Ratio of fixed effect phenotypes to mean population phenotype in additive and network models. Values > 1 indicate that segregating variation decreases the mean population phenotype, and vice-versa for values < 1. Values $= 1$ indicate no segregating variation contributing to trait variance in the population.



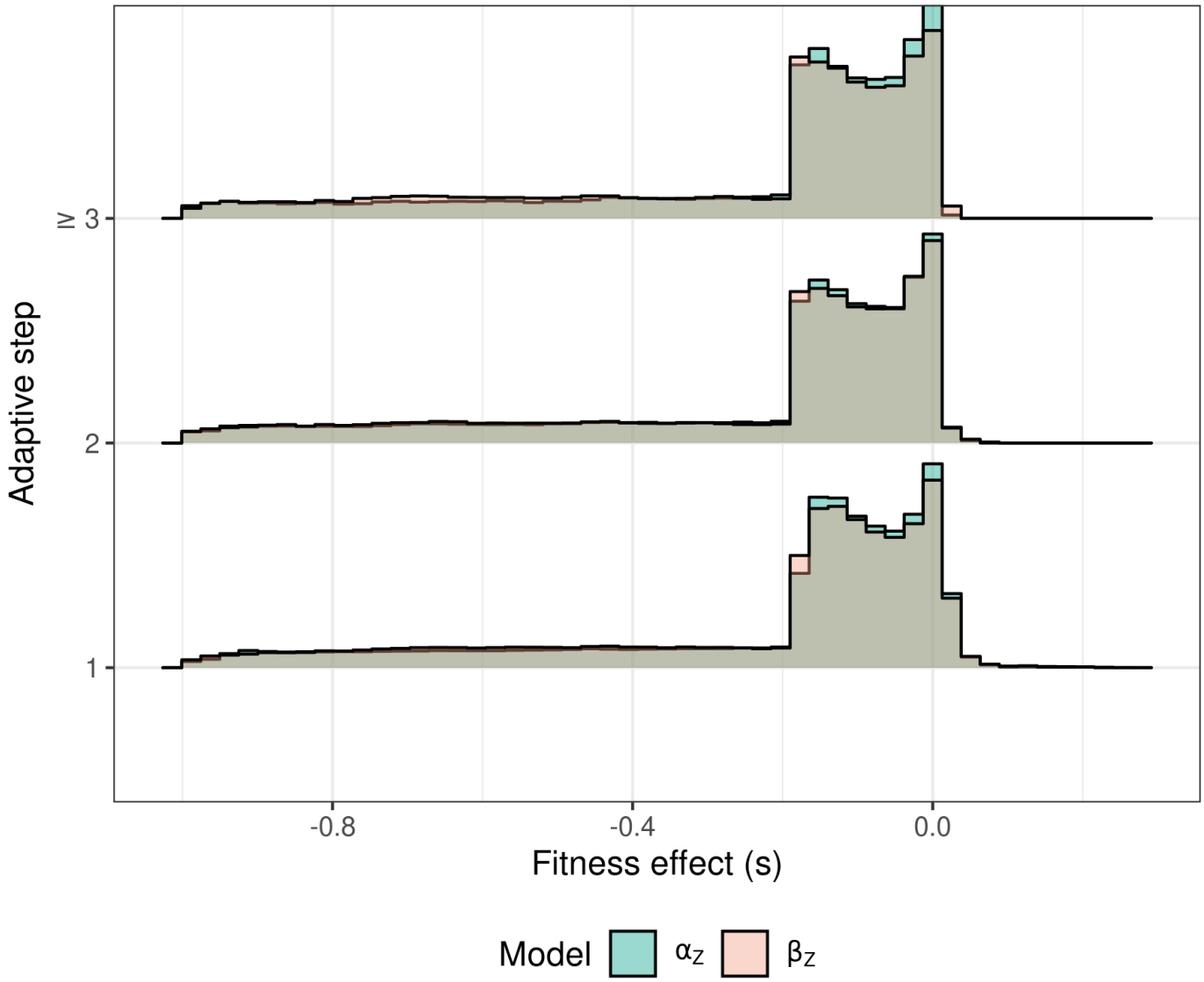
Supplementary Figure 10 Comparison of sampled generalized Pareto distribution (GPD) fits to the distributions of fitness effects (DFE) among beneficial mutations fit onto mutations sampled by two methods. Mutations were created by a mutant screen experiment where 1,000 mutations were randomly generated and their fitness effects measured relative to the phenotype of a population at a given adaptive step. The pooled method (A) pooled together the mutations from all mutant screens across simulation replicates and adaptive steps and used bootstrapping to repeatedly sample 1,000 alleles from the pooled distribution to fit a GPD onto. The non-pooled method (B) fit the GPD onto a sample of 100 mutant screen alleles from each replicate simulation at each of its adaptive steps. Each figure shows the resulting GPD fits of 5 randomly sampled replicates per model of these two methods (i.e. either 5 simulations in the non-pooled case or 5 repeated samples of 1,000 alleles for the pooled case). Note that fewer than 5 samples existed for adaptive steps > 1 in (B), owing to the rarity of walks that long.



Supplementary Figure 11 Comparison of mean shape parameters, κ, from generalized Pareto distribution (GPD) fits to beneficial alleles randomly sampled from each simulation's mutant screen experiment between models and at each adaptive step. κ describes the shape of the GPD, with negative values (seen here) indicating a Weibull domain of attraction. As κ decreases, the maximum size of a beneficial mutation decreases. Error bars are 95% confidence intervals. Sample sizes for each group are given in Supp. Table 2.



Supplementary Figure 12 Two examples of alleles under balancing selection during adaptation in both additive (red) and network (blue) models. The additive mutation had a phenotypic effect $\alpha = -0.618$ while the NAR allele had a phenotypic effect $\alpha = -0.689$



Supplementary Figure 13 The distribution of fitness effects among new mutations in network models for both α_Z and β_Z molecular components. Compared to β_Z mutations, α_Z mutants were less likely to have strongly deleterious effects and more likely to have slightly deleterious effects existing on one of the two modes of the distribution. The shape of the distribution was similar between the molecular components.