Estimating the additive genetic variance for relative fitness from changes in allele frequency

Manas Geeta Arun*[1], Aidan Angus-Henry[1,2], Darren J. Obbard[1], and Jarrod D. Hadfield[1]

[1]Institute of Ecology and Evolution, The University of Edinburgh, Ashworth Laboratories Charlotte Auerbach Road, Edinburgh, EH9 3FL, United Kingdom.
[2]Charité - Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany.

Running head

Estimating heritable variance in fitness

Keywords

Fisher's Fundamental Theorem of Natural Selection, rate of adaptation, quantitative genetics, evolve and resequence.

*Corresponding author: manas.geetaarun@ed.ac.uk

Abstract

The rate of adaptation is equal to the additive genetic variance for relative fitness (V_A) in the population. Estimating V_A typically involves measuring fitness proxies on a large number of individuals from a known pedigree. Such data are hard to collect and the results are often sensitive to the definition of fitness used. Here, we present a new method for estimating V_A that does not involve making measurements of fitness proxies on individuals, but instead tracks changes in the genetic composition of the population. First, we show that V_A can readily be expressed as a function of the genome-wide diversity/linkage disequilibrium matrix and genome-wide expected change in allele frequency due to selection. We then show how independent experimental replicates can be used to infer the expected change in allele frequency due to selection and then estimate V_A via a linear mixed model. Finally, using individual-based simulations, we demonstrate that our approach yields precise and unbiased estimates over a range of biologically plausible levels of V_A.

Introduction

Despite its simplicity, the fundamental theorem of natural selection (FTNS) (Fisher, 1930, 1958) is arguably one of the most central results in evolutionary biology, providing a concise mathematical statement of how quickly a population is expected to adapt. It describes the per generation gain in the mean fitness of a population as a result of natural selection, assuming the 'environment' (including variables intrinsic to the population, such as density or allele frequencies) are held constant (Ewens, 1989; Frank and Slatkin, 1992). The crucial insight the FTNS provides is that the rate of this increase in mean fitness is exactly equal to the additive genetic variance for relative fitness (V_A) in the population (Burt, 1995; Grafen, 2015). Consequently, estimating V_A is one of the '*holy grails*' of evolutionary genetics (Walsh, 2022), despite the FTNS being criticised for its inability to predict the actual gain in mean fitness (Price, 1972; Ewens, 1989, 2024).

A number of attempts have been made to measure V_A in wild populations. Typically, these have involved long term studies on natural populations in which the lifetime reproductive success of a large number of individuals has been measured. Combining these fitness data with information on the relatedness among individuals in a (generalised) linear mixed model approach yields estimates of V_A (Kruuk, 2004). This is far from straightforward in natural populations, as it can be notoriously difficult to tease apart additive genetic effects from common environmental effects, such as parental effects (Kruuk and Hadfield, 2007; Shaw and Shaw, 2014). In addition,

wild study systems are rarely closed, meaning emigration can be misinterpreted as mortality and offspring sired outside the study area can be overlooked. Furthermore, many studies on wild populations lack genetic pedigrees, and as a consequence may miss substantial fitness variation acquired through undetected polygamy (Vedder *et al.*, 2011; Charmantier and Sheldon, 2006). In addition to these biases, estimates from wild populations also tend to come with considerable uncertainty. Burt (1995) reviewed studies estimating V_A in 3 species of plants and 3 species of animals, and found that most estimates of V_A were not significantly different from 0. They argued that the upper bound for estimates of V_A could be as high as 0.3, but most likely less than 0.1. Consistent with this, and using a larger data set, Hendry *et al.* (2018) reported that estimates of V_A varied between 0 and 0.85, with the vast majority of estimates (73%) being less than 0.2. Overall, the mean $\widehat{V_A}$ across studies was 0.08. In a recent meta-analysis, Bonnet *et al.* (2022) applied Bayesian quantitative genetic methods to data obtained from 19 long term studies on wild vertebrate populations. They reported a meta-analytic mean V_A of 0.185 across studies, considerably higher than those obtained by Burt (1995) and Hendry *et al.* (2018). Surprisingly, estimates of V_A in populations of Spotted Hyenas (*Crocuta crocuta*), as well as two of the three populations of Blue Tits (*Cyanistes caeruleus*) were higher than 0.4. This is a remarkable finding, since it suggests that growth rates in these populations should increase nearly 1.5 fold every generation due to selection, provided the environment remains constant. All three meta-analyses investigating V_A (Burt, 1995; Hendry *et al.*, 2018; Bonnet *et al.*, 2022) have detected substantial variability between study systems in their estimates of V_A. Although most of this variability will be sampling error, the best estimate suggests the variability in actual V_A across populations is also large with a standard deviation of 0.11, although there is substantial uncertainty about its exact value (95% credible intervals: 0.01–0.26) (Bonnet *et al.*, 2022).

Measuring V_A in the laboratory is considerably more straightforward, and involves either quantitative genetic breeding designs such as the full-sib half-sib design (Falconer and Mackay, 1996; Lynch and Walsh, 1998), or experimental techniques such as hemiclonal analysis (Abbott and Morrow, 2011). While the controlled environment of the laboratory can, to a large extent, help overcome some of the challenges faced by field studies, laboratory environments often lack important features that are likely to generate fitness variation, such as parasites, predators and competitors. Therefore, it is not entirely clear if laboratory estimates of V_A are particularly relevant. Furthermore, many laboratory studies standardise their fitness measurements in a way that makes it hard to infer estimates of V_A (eg. Ruzicka *et al.* (2019)) or work with absolute fitness but then fail to report V_A (eg. Singh *et al.* (2023)). However, these studies do suggest that genetic variance for fitness is likely highly dependent

on the specific environment in which fitness is assayed (eg. Punzalan *et al.* (2014)). In one of the few laboratory studies that estimated the genetic variance for *relative* fitness, Martinossi-Allibert *et al.* (2018) reported estimates from isofemale lines of the bean beetle, *Acanthoscelides obtectus*, that were highly sensitive to evolutionary history, assay environment, and sex, being higher in males (0.13-0.42) than in females (0.013-0.056).

A common difficulty for current field and laboratory approaches is that, while Darwinian fitness is a deceptively intuitive concept, there is little consensus on its precise definition. In fact, it has been argued that the appropriate definition of fitness can vary depending on the context (Hendry *et al.*, 2018). In the absence of a universal definition, empiricists can only measure proxies of fitness, and it is reasonable to assume that estimates of V_A are likely to be sensitive to the fitness proxy used. A useful illustration of this point is provided by two studies that estimated V_A in a wild population of Red Deer (*Cervus elaphus*), using largely overlapping datasets, but markedly different definitions of fitness (Kruuk *et al.*, 2000; Foerster *et al.*, 2007). Kruuk *et al.* (2000) defined fitness as the total number of progeny produced by an individual in its lifetime and estimated V_A to be 0.1, whereas Foerster *et al.* (2007) employed an alternative definition of fitness that measured an individual's contribution to population growth (Coulson *et al.*, 2006) and obtained the appreciably higher estimate of 0.64.

Some of the definitional difficulties of measuring V_A can be overcome by measuring V_A as the rate of adaptation, rather than comparing the fitness proxies of relatives. One of the earliest attempts to explore this idea was made by Fowler *et al.* (1997) (also see Gardner *et al.* (2005)) in laboratory populations of *Drosophila melanogaster*. Using balancer chromosomes that allow recombination to be suppressed on the third chromosome, Fowler *et al.* (1997) could track the frequency trajectories of newly introduced wild-type third chromosomes over the course of 43 weeks. By modeling fitness using genotype frequency data for a number of different wild-type third chromosomes (Barton and Partridge, 2000), they demonstrated the presence of substantial V_A on the third chromosome. More recently, in a landmark study, Buffalo and Coop (2019) developed a method to estimate the amount of genome-wide allele frequency change that can be attributed to selection. The linchpin of their theory is the idea that linked selection induces across-generation covariances in allele frequency change at neutral loci, since associations between these neutral loci and their respective non-neutral backgrounds persist across generations. This new theoretical framework has the potential to pave the way for a powerful empirical tool to detect genomic signatures of linked selection (Buffalo and Coop, 2020; Simon and Coop, 2024). Of particular relevance here, Buffalo and Coop (2019) also show that

their method can be used to obtain estimates of V_A, albeit under some potentially restrictive assumptions.

In this study, we present an alternative theoretical framework that relates V_A to genome-wide changes in allele-frequency. Using mathematical identities only, we show how V_A can be obtained from an initial linkage disequilibrium (LD) matrix and expected changes in allele frequency due to selection, without making any assumptions about patterns of gene action or the relationships between genotype fitnesses and genotype frequencies. Our approach, like that of Buffalo and Coop (2019), relies on temporal genomic data and does not necessitate measuring fitness in individuals. However, in contrast to the 'bottom-up' population genetic approach of Buffalo and Coop (2019), we use a 'top-down' quantitative genetic approach which simplifies and generalises some aspects of the problem.

Our aim here is three-fold. First, we derive our central theoretical result from first principles. Second, we develop the statistical machinery required to apply our result to real biological data, and validate it with individual based simulations. Third, we make a detailed comparison of our method with that of Buffalo and Coop (2019).

Materials and Method

Outline of the theory

Table 3 summarises the notation used in the following text. We consider a population consisting of N diploid individuals. We assume that there are n_L segregating biallelic loci in the population. Let $c_{k,i}$ and α_i represent the proportion of copies of an arbitrarily chosen reference allele at locus i in individual k and Fisher's average effect (Fisher, 1941) for fitness at locus i, respectively. The widely accepted mathematical definition of the α's are the regression coefficients obtained from a multiple regression of the c's on relative fitness, w (Fisher, 1941; Lee and Chow, 2013). The vector $\boldsymbol{\alpha}$ can be expressed as

$$\boldsymbol{\alpha} \quad = \mathbf{L}^{-1} Cov(\mathbf{c}, w) \tag{1}$$

where $\mathbf{c}$ is the vector of predictors representing the c's at all loci for an individual and $\mathbf{L}$ is a symmetric $n_L \times n_L$ matrix whose i^{th} (diagonal) element is the variance in the c's at locus i computed over individuals, while the ij^{th} (off-diagonal) element describes the covariance between the c's at locus i and locus j computed over individuals.

The breeding value for the relative fitness of individual k is then

$$u_k = \mathbf{c}_k^\top \boldsymbol{\alpha}$$

and the additive genetic variance for relative fitness is the variance of this quantity across individuals:

$$\begin{aligned} V_A &= Var(\mathbf{c}^\top \boldsymbol{\alpha}) \\ &= \boldsymbol{\alpha}^\top Var(\mathbf{c}, \mathbf{c}^\top)\boldsymbol{\alpha} \\ &= \boldsymbol{\alpha}^\top \mathbf{L}\boldsymbol{\alpha} \end{aligned} \tag{2}$$

This follows from the fact that at any given point in time, α's are constant across individuals. In the absence of mutation and meiotic drive, the allele frequency in parents is transmitted to offspring without bias, such that the vector of expected change in allele frequencies due to selection can be expressed as Robertson's covariance (Robertson, 1966; Price, 1970; Queller, 2017):

$$\begin{aligned} E[\Delta\mathbf{p}] &= E[\Delta\bar{\mathbf{c}}] \\ &= Cov(\mathbf{c}, w) \end{aligned} \tag{3}$$

Substituting Equation 1 into Equation 3 gives $E[\Delta\mathbf{p}] = \mathbf{L}\boldsymbol{\alpha}$, which is the multivariate analogue of Equation 10 in Kirkpatrick *et al.* (2002). Combining this with Equation 2 yields,

$$\begin{aligned} V_A &= \left(\mathbf{L}^{-1}E[\Delta\mathbf{p}]\right)^\top \mathbf{L}\left(\mathbf{L}^{-1}E[\Delta\mathbf{p}]\right) \\ &= E[\Delta\mathbf{p}]^\top \mathbf{L}^{-1}E[\Delta\mathbf{p}] \end{aligned} \tag{4}$$

Equation 4 is a general result and involves no assumptions about the patterns of dominance or epistasis for fitness, or about patterns of mating. An intuitive explanation of why V_A can be calculated this way is to note Fisher's Fundamental Theorem states that V_A is equal to the (partial) increase in mean fitness caused by evolutionary change through natural selection. Equation 2 can be expressed as a sum of $\alpha_i E[\Delta p_i]$ over all loci, i:

$$V_A = \boldsymbol{\alpha}^\top E[\Delta\mathbf{p}] \tag{5}$$

In this, α represents the proportional change in the population mean fitness that is caused by a unit change in allele frequency at a locus (Fisher, 1941; Kojima, 1959; Lee and Chow, 2013). Therefore, multiplying α by the actual change caused by natural selection, $E[\Delta p]$, we get the proportional change in mean fitness caused by evolutionary change by natural selection at that locus. If we then add these changes at every locus in the genome, we obtain the total proportional change in mean fitness due to evolutionary change by natural selection, and hence V_A.

Extending our approach to practical situations

Our theoretical approach assumes $\boldsymbol{\alpha}$ or $E[\Delta\mathbf{p}]$ are known. In reality, neither can be directly observed and must be inferred from data on observed allele frequency change, $\Delta\mathbf{p}$. Since $\Delta\mathbf{p}$ will vary around $E[\Delta\mathbf{p}]$ due to genetic drift, $E[\Delta\mathbf{p}]$ must be inferred using replicate observations of $\Delta\mathbf{p}$. Since time cannot be replayed, we infer $E[\Delta\mathbf{p}]$ through experimental replicates (see Buffalo and Coop (2020) also). Our theoretical model also assumes $\mathbf{L}$ is known, but since it is hard to generate experimental replicates without at least one round of reproduction, we condition on $\mathbf{L}_0$ ($\mathbf{L}$ in a generation prior to the first generation over which allele frequency change is measured). In what follows we will refer to the population at time zero (generation 0) as the 'base population'. Our aim is then to estimate an additive genetic variance for fitness in the base population, $V_A(0) = \bar{\boldsymbol{\alpha}}^\top \mathbf{L}_0 \bar{\boldsymbol{\alpha}}$, where $\bar{\boldsymbol{\alpha}}$ is the vector of mean average effects across time and replicates.

We also allow allele frequency changes to be measured over multiple generations, rather than a single generation. Thus, $\Delta\mathbf{p}_m = \Delta\mathbf{p}_{t_m,m} + \Delta\mathbf{p}_{t_m+1,m} \ldots \Delta\mathbf{p}_{\tau_m-1,m}$ is the observed change in allele frequency from generation t_m to τ_m in replicate m, with $\Delta\mathbf{p}_{t,m}$ being the change from time t to $t+1$. Note that t_m (i.e. the generation in which we start recording allele frequency changes) can be any positive integer. If replicate populations are derived from the base population with exactly one round of reproduction, t_m would be 1. τ_m can be any integer greater than t_m. We first note that the total change in allele frequency in replicate m between times t_m and τ_m is

$$\Delta\mathbf{p}_m = \sum_{t=t_m}^{\tau_m-1} \left(\mathbf{L}_{t,m}\boldsymbol{\alpha}_{t,m} + \underset{D}{\Delta}\mathbf{p}_{t,m} \right) \tag{6}$$

where $\mathbf{L}_{t,m}\boldsymbol{\alpha}_{t,m} = E[\Delta\mathbf{p}_{t,m}]$ is the expected change in allele frequency due to selection in replicate m between generation t and $t+1$ and $\underset{D}{\Delta}\mathbf{p}_{t,m}$ is the change due to drift. It is important to note that $\mathbf{L}_{t,m}\boldsymbol{\alpha}_{t,m}$ captures responses to both direct and indirect selection. $\mathbf{L}$ will vary over time and we decompose $\mathbf{L}$ at a particular time into a part that can be predicted by $\mathbf{L}_0$ and the action of drift and recombination, and a part that cannot be predicted. To do this we decompose $\mathbf{L}$ into $\mathbf{L}'$ and $\mathbf{L}''$, following the notation of Buffalo and Coop (2019). $\mathbf{L}'$ represents the (co)variances in the c's computed over the haploid genomes of all $2N$ gametes that constitute the base population. Thus, the diagonal elements of $\mathbf{L}'$ are proportional to gametic (gene) diversity and the off-diagonals are proportional to gametic-phase disequilibrium. On the other hand, $\mathbf{L}''$ represents the (co)variances that arise due to alleles within the different gametic contributions of a genotype, and thus the diagonal elements are

proportional to the additive coefficients of Hardy-Weinberg disequilibrium (Bulmer *et al.* (1980) and Chapter 12 in Weir *et al.* (1989)), and the off-diagonal elements are proportional to the nongametic-phase disequilibrium. Given this decomposition (See Appendix S1 for details):

$$\mathbf{L}_{t,m} = \mathbf{N}_{t,m} \circ \tilde{\mathbf{L}}_0 + \Delta\mathbf{L}^{'}_{t,m} + \mathbf{L}^{''}_{t,m} \tag{7}$$

where $\circ$ is the Hadamard product, $\tilde{\mathbf{L}}_0$ is the weighted sum of $\mathbf{L}^{'}_0$ and $\mathbf{L}^{''}_0$, with the ij^{th} element of $\mathbf{L}^{''}_0$ weighted by $r_{i,j}/(1-r_{i,j})$ and $\mathbf{N}_{t,m}$ is a matrix with the ij^{th} element being $(1-r_{i,j})^t \prod_{k=0}^{t-1}(1-\frac{1}{2N_{e_{k,m}}})$. $r_{i,j}$ is the recombination rate between the two loci and $N_{e_{k,m}}$ is the effective population size in generation k in replicate m. $\Delta\mathbf{L}^{'}_{t,m}$ is a stochastic term with zero expectation that represents the accumulated change in $\mathbf{L}^{'}$ between generations 0 and t in replicate m that cannot be predicted. $\mathbf{L}^{''}_{t,m}$ is the matrix of nongametic-phase disequilbria that arises in generation t in replicate m. In the following, it will be useful to designate $\boldsymbol{\mathcal{L}}_{t,m} = \mathbf{N}_{t,m} \circ \tilde{\mathbf{L}}_0$ as the expected $\mathbf{L}_{t,m}$, for $t > 0$, conditional on $\mathbf{L}^{'}_0$ and $\mathbf{L}^{''}_0$. Also, we represent the deviation of $\mathbf{L}_{t,m}$ from this expectation as $\Delta\mathbf{L}_{t,m} = \Delta\mathbf{L}^{'}_{t,m} + \mathbf{L}^{''}_{t,m}$. It will also be useful to denote $\boldsymbol{\alpha}_{t,m} = \bar{\boldsymbol{\alpha}} + \Delta\boldsymbol{\alpha}_{t,m}$ where $\Delta\boldsymbol{\alpha}_{t,m}$ is the deviation of the average effects (in generation t in replicate m) from the mean of the average effects across time and replicates. We can then decompose the change due to selection into two terms:

$$\begin{aligned} \Delta\mathbf{p}_m = & \sum_{t=t_m}^{\tau_m-1}\left(\boldsymbol{\mathcal{L}}_{t,m}\bar{\boldsymbol{\alpha}} + \Delta\mathbf{L}_{t,m}\bar{\boldsymbol{\alpha}} + \boldsymbol{\mathcal{L}}_{t,m}\Delta\boldsymbol{\alpha}_{t,m} + \Delta\mathbf{L}_{t,m}\Delta\boldsymbol{\alpha}_{t,m} + \underset{D}{\Delta}\mathbf{p}_{t,m}\right) \\ = & \sum_{t=t_m}^{\tau_m-1}\left(\boldsymbol{\mathcal{L}}_{t,m}\bar{\boldsymbol{\alpha}} + \underset{U}{\Delta}\mathbf{p}_{t,m} + \underset{D}{\Delta}\mathbf{p}_{t,m}\right) \end{aligned} \tag{8}$$

where the predictable change due to selection is $\boldsymbol{\mathcal{L}}_{t,m}\bar{\boldsymbol{\alpha}}$ and is similar to that derived for a single locus in Santiago and Caballero (1998), and the unpredictable change due to selection caused by stochastic changes in $\mathbf{L}$ and/or $\boldsymbol{\alpha}$ is $\underset{U}{\Delta}\mathbf{p}_{t,m} = \Delta\mathbf{L}_{t,m}\bar{\boldsymbol{\alpha}} + \boldsymbol{\mathcal{L}}_{t,m}\Delta\boldsymbol{\alpha}_{t,m} + \Delta\mathbf{L}_{t,m}\Delta\boldsymbol{\alpha}_{t,m}$. Note that $\underset{D}{\Delta}\mathbf{p}$ and $\underset{U}{\Delta}\mathbf{p}$ have expectation zero and that the $\underset{D}{\Delta}\mathbf{p}$ terms are independent across replicates, generations, and generations within replicates (Buffalo and Coop, 2019). The $\underset{U}{\Delta}\mathbf{p}$ are never independent across generations within replicates since the $\Delta\mathbf{L}^{'}$ are sums of changes over all previous generations after generation 0. They may, however, be independent across replicates if each replicate is initiated from individuals independently generated from Generation 0 individuals (e.g. $t_m = 1$ for all m). If replicates are derived from a single population that was derived from generation 0 $t_m - 1$ generations before, then the $\Delta\mathbf{L}$ will be dependent due to the shared changes in $\mathbf{L}$ from generation 0 to $t_m - 1$. However, even

when replicates are independently generated from generation 0, there may be cases where the $\underset{U}{\Delta}\mathbf{p}$ are non-independent between replicates, because of some structure to either the $\Delta\mathbf{L}'$'s or the $\Delta\boldsymbol{\alpha}$'s. The $\Delta\boldsymbol{\alpha}$'s may be non-independent across replicates, if a selection regime is unique to a time point but experienced by all replicates. The $\Delta\mathbf{L}$'s may also be non-independent if selection induces correlated changes in $\mathbf{L}$ across replicates, although here we follow the assumption of the infinitesimal model where such changes are negligible (Barton *et al.*, 2017), at least for the diagonal elements (Bulmer, 1971). Making t_m as small as possible so that changes in $\mathbf{L}$ from $\mathbf{L}_0$ are minimised (i.e. the replicates are initiated using the offspring of the base population such that $t_m = 1$) and calculating change in allele frequency over a single generation (i.e. $\tau_m - t_m = 1$) will result in the least bias since any change in $\mathbf{L}$ due to selection should be minimised and the approximations that follow should hold well. However, although the approximations will hold better when $\tau_m - t_m = 1$, increasing $\tau_m - t_m$ (i.e. calculating change in allele frequency over multiple generations) will increase power since the changes in allele frequency due to selection will be larger.

Assuming the $\Delta\mathbf{L}$'s and the $\Delta\boldsymbol{\alpha}$'s are independent across replicates we can then derive the mean and covariance structure of allele frequency change across replicates in terms of the mean ($\boldsymbol{\mu}_{\bar{\alpha}}$) and covariance structure ($\mathbf{V}_{\bar{\alpha}}$) of the mean average effects (Appendix S2). Note that here we treat the average effects as random variables rather than fixed quantities (See Appendix S3 and Gianola *et al.* (2009) for a discussion on what this implies):

$$E\left[\Delta\mathbf{p}_m\right] = \boldsymbol{\mathcal{L}}_m\boldsymbol{\mu}_{\bar{\alpha}} \tag{9}$$

where $\boldsymbol{\mathcal{L}}_m = \sum_{t=t_m}^{\tau_m-1}\boldsymbol{\mathcal{L}}_{t,m}$.

The conditional among-replicate covariance is also straightforward

$$COV(\Delta\mathbf{p}_m, \Delta\mathbf{p}_n^\top) = \boldsymbol{\mathcal{L}}_m\mathbf{V}_{\bar{\alpha}}\boldsymbol{\mathcal{L}}_n \tag{10}$$

where m and n are a pair of replicates. The within-replicate (co)variances are more challenging to derive, as not only do they include the predictable response to selection, but also the cumulative effects of drift and the unpredictable response to selection. There is no easy form for the within-replicate (co)variances in the allele frequency change caused by the total impact of the unpredictable response to selection. Here, we simply write the covariance due to the cumulative unpredictable response to selection as $\boldsymbol{\mathcal{U}}_m$, which gives

$$VAR\left(\Delta\mathbf{p}_m\right) = \boldsymbol{\mathcal{L}}_m\mathbf{V}_{\bar{\alpha}}\boldsymbol{\mathcal{L}}_m + \boldsymbol{\mathcal{D}}_m + \boldsymbol{\mathcal{U}}_m \tag{11}$$

where $\boldsymbol{\mathcal{D}}_m$ is the covariance due to the cumulative action of drift and is equal to $\tilde{\mathbf{L}}_0 \circ \mathbf{M}^{(m)}$ where $\mathbf{M}^{(m)} = \sum_{t=t_m}^{\tau_m - 1} \mathbf{M}_{t,m} \circ \mathbf{N}_{t,m}$ and the ij^{th} element of $\mathbf{M}_{t,m}$ is $(1 - r_{i,j})/N_{E_{t,m}}$ (Appendix S2). Note that since the predictable response to selection includes both direct and indirect selection, the relevant effective population size for the drift covariance in allele frequency does not include the effects of linked selection and for this reason we denote it as N_E rather than N_e. Using this notation, $N_E = 4N/(2 + V_o)$ where N is the census population size and V_o is the variance in offspring number in the absence of genetic fitness variation (Wright, 1938). Although the dynamics of $\mathbf{L}$ (Equation 7, also see Appendix S1) were derived under drift and recombination in the absence of selection, and hence $N_e = N_E$, the dynamical equations for $\mathbf{L}$ will better approximate reality when an effective population size that incorporates all excess variation in fitness is used. Consequently, we use N_e and N_E to distinguish the effective population sizes that are relevant for stochastic changes in $\mathbf{L}$ and allele frequency, respectively. It should also be noted that if $VAR(\Delta\boldsymbol{\alpha})$, and hence $\boldsymbol{\mathcal{U}}$, is non-zero then the additive genetic variance exhibited in replicate m at generation t will in general exceed $V_A(0)$. When $VAR(\Delta\boldsymbol{\alpha})$ is non-zero, we should more generally think of $V_A(0)$ not as the additive genetic variance for fitness in the base population, but the additive genetic covariance in fitness between replicates.

Inference outline

By applying sum of squares theory (Searle, 2006, page 355) to Equation 3 we can obtain the expected $V_A(0)$ after averaging over the distribution of average effects:

$$\begin{aligned} E[V_A(0)] &= E[\bar{\boldsymbol{\alpha}}^\top \mathbf{L}_0 \bar{\boldsymbol{\alpha}}] \\ &= Tr(\mathbf{L}_0 \mathbf{V}_{\bar{\alpha}}) + \boldsymbol{\mu}_{\bar{\alpha}}^\top \mathbf{L}_0 \boldsymbol{\mu}_{\bar{\alpha}} \end{aligned} \tag{12}$$

where we aim to estimate $\boldsymbol{\mu}_{\bar{\alpha}}$ and $\mathbf{V}_{\bar{\alpha}}$ through Equations 9 and 10 using multiple evolutionary replicates starting from a common base population.

Rather than working with the allele frequency changes directly, we project them on to a new (reduced) basis and denote this new vector of changes as $\Delta\overrightarrow{\mathbf{p}} = \mathbf{P}\Delta\mathbf{p}$ where $\mathbf{P}$ is some projection matrix. We chose a projection that collapses allele frequency changes into the non-null subspace of $\mathbf{L}_0$, since $V_A(0)$ only depends on this subspace (de Los Campos *et al.*, 2015). To do this, let $\mathbf{U_L}$ be the eigenvectors of $\mathbf{L}_0$ with non-zero eigenvalues and then the drift covariance in the reduced subspace is $\mathbf{U}_\mathbf{L}^\top(\tilde{\mathbf{L}}_0 \circ \mathbf{M}^{(m)})\mathbf{U_L}$. If we have $\mathbf{U}_2$ and $\mathbf{D}_2$ as the eigenvectors and a diagonal matrix of square-rooted eigenvalues of this matrix, then the projection matrix $\mathbf{P} = \mathbf{D}_2^{-1}\mathbf{U}_2^\top\mathbf{U}_\mathbf{L}^\top$ results in projected allele frequency changes that are identically and independently distributed under drift in the reduced subspace only (see Appendix S4).

The mean and covariance of the projected allele frequency changes due to predictable selection are:

$$E[\Delta \overrightarrow{\mathbf{p}}_m] = \mathbf{P}_m \boldsymbol{\mathcal{L}}_m \boldsymbol{\mu}_{\bar{\alpha}} \tag{13}$$

and

$$COV(\Delta \overrightarrow{\mathbf{p}}_m, \Delta \overrightarrow{\mathbf{p}}_n) = \mathbf{P}_m \boldsymbol{\mathcal{L}}_m \mathbf{V}_{\bar{\alpha}} \boldsymbol{\mathcal{L}}_n \mathbf{P}_n^\top \tag{14}$$

In Appendix S3 we determine permissible models for $\boldsymbol{\mu}_{\bar{\alpha}}^\top$ and $\mathbf{V}_{\bar{\alpha}}$. Of those, we identify the sensible biological model for the mean average effect:

$$\boldsymbol{\mu}_{\bar{\alpha}} = \beta_{\bar{\alpha}}^{(0)} + \beta_{\bar{\alpha}}^{(1)}(\mathbf{p}_0 - \mathbf{q}_0) \tag{15}$$

where $\mathbf{p}_0$ is the frequency of the reference alleles at each locus in the base population and $\mathbf{q}_0 = 1 - \mathbf{p}_0$. For the variance of the average effects, we identify the model

$$\mathbf{V}_{\bar{\alpha}} = \sigma_{\bar{\alpha}}^2 \mathbf{L}_0^{p_{\bar{\alpha}}} \tag{16}$$

although in our analysis of simulated data we set the off-diagonal elements of $\mathbf{L}_0$ to zero in the above equation, such that the variance in average effects is simply a power function of the genetic diversities (Zeng *et al.*, 2018). With $p_{\bar{\alpha}}$ known, the model is a linear mixed model with covariance structure due to the predictable response to selection being proportional (by a factor $\sigma_{\bar{\alpha}}^2$ that is to be estimated) to

$$\mathbf{V}_{m,n} \propto \mathbf{P} \boldsymbol{\mathcal{L}}_m \mathbf{L}_0^{p_{\bar{\alpha}}} \boldsymbol{\mathcal{L}}_n \mathbf{P}^\top \tag{17}$$

between replicates m and n. When $\boldsymbol{\mathcal{L}}^{(m)} = \boldsymbol{\mathcal{L}}^{(n)}$ for all n and m, then this can more easily be fitted by incorporating locus as a random effect with the above covariance structure. The fixed-effect covariate (shown for replicate m)

$$\mathbf{P} \boldsymbol{\mathcal{L}}^{(m)} (\mathbf{p}_0 - \mathbf{q}_0) \tag{18}$$

can also be fitted, with associated coefficient $\beta_{\bar{\alpha}}^{(1)}$ to be estimated in addition to the intercept, $\beta_{\bar{\alpha}}^{(0)}$.

We estimate the parameters of the model by treating it as a separable linear mixed model problem (Richards, 1961). Conditional on a value of $p_{\bar{\alpha}}$ the model is linear mixed model and the conditional (restricted) maximum likelihood can be obtained from asreml (Butler *et al.*, 2023). In order to maximise the (unconditional) likelihood we use the R function 'optim()' to find the value of $p_{\bar{\alpha}}$ that results in the highest conditional likelihood. Note that the residual variance should equal one if there is no unpredictable response to selection.

Although linear (mixed) models generate unbiased estimates of $\boldsymbol{\mu}_{\bar{\alpha}}$, the quadratic form $\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}^{\top}\mathbf{L}_0\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}$ will be upwardly biased by sampling error in $\boldsymbol{\beta}_{\bar{\alpha}}$. To correct for this bias, we use the inverse Hessian (conditional on the best estimate of $p_{\bar{\alpha}}$) to get an approximate expression for the sampling (co)variance matrix ($\mathbf{S}_{\bar{\alpha}}$) of the two parameters, and use this in order to get an improved estimate of the quadratic form as follows (see Appendix S5):

$$\widehat{\boldsymbol{\mu}_{\bar{\alpha}}^{\top}\mathbf{L}_0\boldsymbol{\mu}_{\bar{\alpha}}} = \widehat{\boldsymbol{\mu}_{\bar{\alpha}}}^{\top}\mathbf{L}_0\widehat{\boldsymbol{\mu}_{\bar{\alpha}}} - Tr\left(\mathbf{X}^{\top}\mathbf{L}_0\mathbf{X}\mathbf{S}_{\bar{\alpha}}\right) \tag{19}$$

where $\mathbf{X}$ is the fixed effect design matrix.

Comparison with Buffalo and Coop (2019)

To connect our work with Buffalo and Coop (2019) (henceforth 'B&C') it will be useful to express the covariance matrix $\mathbf{L}$ in terms of a diagonal matrix of standard deviations, $\mathbf{B}$, (half the square-root of the genetic diversities under random mating) and the correlation matrix, $\mathbf{R}$, such that $\mathbf{L} = \mathbf{BRB}$. We can then split the response to selection in generation t into two parts:

$$\begin{aligned} \mathbf{L}_t\boldsymbol{\alpha}_t = & \ \mathbf{B}_t\mathbf{B}_t\boldsymbol{\alpha}_t & + & \ \mathbf{B}_t(\mathbf{R}_t - \mathbf{I})\mathbf{B}_t\boldsymbol{\alpha}_t \\ = & \ \underset{S}{\Delta}\mathbf{p}_t & + & \ \underset{L}{\Delta}\mathbf{p}_t \end{aligned} \tag{20}$$

where the first term is due to direct selection at the loci, and the second term is due to linkage-disequilibria with other selected loci. It will also be useful to distinguish the additive *genetic* variance ($V_A(t)$) from the additive *genic* variance ($V_a(t)$) in generation t, i.e. :

$$V_A(t) = \boldsymbol{\alpha}_t\mathbf{L}_t\boldsymbol{\alpha}_t^{\top} \tag{21}$$

versus

$$V_a(t) = \boldsymbol{\alpha}_t\mathbf{B}_t\mathbf{B}_t\boldsymbol{\alpha}_t^{\top} \tag{22}$$

with the distinction being that the additive *genetic* variance ($V_A(t)$) captures the contributions of both genetic diversities at individual loci and the linkage disequilibria between pairs of loci, while the additive *genic* variance ($V_a(t)$) captures the contributions of genetic diversities only.

We can also think about the additive genetic/genic covariance in fitness between time points t and τ for a population with genetic structure equal to that in generation t:

$$C_A(\tau \to t) = \boldsymbol{\alpha}_t\mathbf{L}_t\boldsymbol{\alpha}_{\tau}^{\top} \tag{23}$$

and

$$C_a(\tau \to t) = \boldsymbol{\alpha}_t \mathbf{B}_t \mathbf{B}_t \boldsymbol{\alpha}_\tau^\top \tag{24}$$

These will differ from $V_A(t)$ or $V_a(t)$ when the average effects at generation τ are different from those at generation t. This could happen when the selective environment changes between generations t and τ, meaning parameters such as s and h change. The average effects will also change if the genetic composition of the population changes between generations t and τ (i.e. $\mathbf{L}_t \neq \mathbf{L}_\tau$) and there is non-additivity (e.g. $h \neq 0.5$) since the average effects are the coefficients of a multiple regression of the c's on relative fitness (Equation 1).

In the theory section above, we assumed $\mathbf{L}_t$ and $\boldsymbol{\alpha}_t$ were known and so the change due to both direct and linked selection are fixed quantities, as is $V_A(t)$. In contrast, B&C condition on $\mathbf{B}_t$ and $\boldsymbol{\alpha}_t$, and treat $\mathbf{R}_t$ as a random variable. Consequently, in B&C, the change due to direct selection and $V_a(t)$ are fixed, as in our approach, since genetic diversities and average effects are known. However, the change due to indirect selection and $V_A(t)$ is random since the linkage-disequilibrium is unknown. In our inference section we acknowledge that $\boldsymbol{\alpha}$ cannot be directly observed, but develop a model for the distribution of α's, where estimates of $V_A(t)$ can be made after marginalising $\boldsymbol{\alpha}$. Similarly, in the approach of B&C the $\boldsymbol{\alpha}$'s and the elements of $\mathbf{B}$ for selected sites are also not directly observed, but they make a number of strong assumptions that effectively allow the joint distribution of the average effects and selected site diversities to be marginalised.

In Appendix S6 we work through the derivation of B&C using our own notation and retaining a full multi-locus treatment. For easy comparison with the present work, here we summarise both the explicit and implicit assumptions underlying the general approach of B&C:

- A) Allele frequency changes are only measured over a single generation.
- B) There is no direct selection on the loci for which allele frequency change is measured.
- C) The reference allele at a neutral locus is chosen arbitrarily.
- D) The signed linkage-disequilibrium between selected sites is zero, which precludes processes such as Hill-Robertson interference. Under this assumption $V_A = V_a$.

- E) Changes in $\mathbf{R}$ are due to recombination alone - selection and drift are absent.
- F) Nongametic-phase linkage disequilibrium is absent.
- G) There is no relationship between the additive genic variation at a selected site and its LD with neutral sites, as measured through $R_{i_t,j_t}R_{i_\tau,j_\tau}$ where i is a neutral locus and j a selected locus. Note that since genetic diversity determines the additive genic variation and LD (even measured as a correlation) is constrained by the genetic diversities at the two loci, this is unlikely to be met (Sved and Hill, 2018).
- H) The average effects are constant, i.e. $\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_\tau$ if estimates are to be interpreted as the additive genic variance (or additive genetic variance, if Assumption D is met) in generation τ. If they are not constant, it is the additive genic/genetic covariance between generations t and τ that is measured.
- I) The initial expected LD-structure between selected and neutral loci, $E[R^2_{i_t,j_t}]$, is approximated from an expression for $E[L^2_{i_t,j_t}]$, which is itself derived under mutation-recombination-drift balance (i.e. no selection). This assumption is partly relaxed in the appendix by assuming the expected LD-structure between selected and neutral loci is equal to the observed LD-structure between all loci, not distinguishing between selected and neutral loci (Assumption I-b). However, in many cases the genetic diversity at selected loci will be less than that found at (or assumed for) neutral loci and so $E[R^2_{i_t,j_t}]$ will likely be smaller than assumed (Sved and Hill, 2018).
- J) The ratio of genetic diversity in generation τ to genetic diversity in generation t ($\phi_{t,\tau}$), is constant across all selected loci. Under this assumption, dividing the estimate of $V_a(\tau)$ by $\phi_{t,\tau}$ gives an estimate of $V_a(t)$.
- K) The recombination rate between two sites g base pairs apart is given by $(1 - e^{-2g\bar{r}})/2$ (i.e. Haldane's (1919) mapping function) where $\bar{r}$ is the average crossover rate per site per generation.
- L) Neutral and selected loci are distributed uniformly and independently across the genome such that the distance between them has a triangular distribution.
- M) Since the genetic diversity at selected sites is not measured, it is assumed that $\phi_{t,\tau}$ is equal to the average ratio of genetic diversity in generation τ to genetic diversity in generation t across all neutral loci.

When applying the theory to data, the method of B&C requires two additional assumptions:

- Ib) The average LD between selected and neutral sites is equal to the average LD observed between all sites: see Assumption I) above.
- N) The (co)variance in allele frequency changes divided through by the average genetic diversity is a good approximation for the (co)variance of weighted allele frequency changes where the weights are the inverse of the square root of the genetic diversities.
- O) If allele frequencies are estimated from a sample, then the estimation error is binomially distributed.

Our approach relaxes most of these assumptions, but not all. Assumption E is only partly relaxed - the change in $\mathbf{R}$ does not assume a lack of drift but it does assume a lack of selection. We also make Assumption H if we choose to interpret $V_A(0)$ as an additive genetic variance rather than an additive genetic covariance. Assumptions E & H result in the unpredictable response to selection being zero. While we do not make this assumption, we do assume that the unpredictable responses to selection are independent across replicates. However, when applying the theory to data we assume that the within-replicate (residual) (co)variances are driven by drift only (rather than an unpredictable response to selection) and that the drift (co)variances are equal to their expected values (conditional on $\mathbf{L}_0$). In reality, the within-replicate covariances are likely larger, due to the unpredictable response to selection and evolutionary variance in the drift (co)variances. These assumptions might not be too severe if these two additional processes result in within-replicate (co)variances that are proportional to those under pure drift. Under this scenario, the model should perform well, although the residual variance may be higher than expected. In addition, in the absence of a recombination map, Assumption K may also be applied in our method, and if $\mathbf{L}_0$ cannot be partitioned into gametic-phase and nongametic-phase contributions, assumption F might also be made by substituting $\mathbf{L}_0$ for $\mathbf{L}_0^{'}$ in $\tilde{\mathbf{L}}_0 \circ \mathbf{N}$ and $\tilde{\mathbf{L}}_0 \circ \mathbf{M}$. By relaxing some of the more extreme assumptions of B&C we are instead forced into making assumptions about the mean and covariance structure of the average effects of projected loci when making inferences.

Multilocus simulations

To validate our method and inference approach, we performed multilocus simulations using msprime (version 1.2.0) (Kelleher *et al.*, 2016) and SLiM (version

4.2.2) (Haller and Messer, 2023). The code for the simulations and the downstream analyses performed in R (version 4.3.3) is available on GitHub (https://github.com/manas-ga/Va_simulations) and includes the R library, `Vw`. We first describe the structure of the simulations in general terms. Towards the end of this section (see 'Varying simulation parameters') we discuss the specifics of our parameter choices in the context of the scaling used in the simulations.

We simulated a single, contiguous 1 million base-pair long genomic region. Our simulations had two distinct phases: the first phase simulated the history of an ancestral population and the second phase simulated a typical evolve and resequence experiment with independent replicate experimental populations derived from a base population drawn from the ancestral population.

Model for fitness

We used an additive model for log absolute fitness, $log(W)$, as in Buffalo and Coop (2019). The breeding value of log fitness for individual k is equal to $u_k = \sum_{i=1}^{n_L} c_{k,i}\eta_i$, where, as before, $c_{k,i}$ represents the proportion of copies of the reference allele in individual k at locus i, and η_i is the (average) effect of the allele on log fitness. The average effects for (relative) fitness are well approximated by the η's when they are small in magnitude, but higher-order approximations are required when the η's are large (see Appendix S6.1). To obtain the log fitness of each individual we added a noise term drawn from a standard normal distribution (mean = 0, variance = 1) to the breeding value. The fitness of each individual is then obtained by exponentiating the individual's log fitness.

We sampled the η's from a distribution comprising a weighted mixture of three distributions: (1) a point mass at $\eta = 0$ representing neutral mutations, (2) a reflected gamma distribution with shape = 0.3 and scale = η_{scale} (typically 0.033, unless specified otherwise) representing deleterious mutations, and (3) a gamma distribution with shape = 0.3 and scale = η_{scale} (typically 0.033, unless specified otherwise) representing beneficial mutations. Thus, for both deleterious and beneficial mutations, the mean absolute η (scale $\times$ shape) was, typically, 0.01. The ratio of the frequency of beneficial to deleterious mutations, and η_{scale} varied among simulations. Below, we refer to the distribution of non-neutral η's (i.e. omitting mixture component (1) described above) as the 'distribution of fitness effects' (DFE), but note that this is in fact the distribution of non-zero effects on log fitness.

Phase 1 ('history phase'): Simulating the history of an ancestral population

We first used a neutral coalescent simulation implemented in msprime (Kelleher *et al.*, 2016) to construct genealogies for 2,500 diploid genomes (i.e. $N_e = 2{,}500$). To initialise a (non-equilibrium) set of selected loci, we then simulated mutations at a rate μ_{msp}, using the pyslim package to attach fitness effects (η) drawn randomly from the non-neutral part of the distribution described above. Since derived mutations are rare, and the DFE is predominantly deleterious, this will generate a positive relationship between the reference allele frequency and the fitness effects η's (and therefore the α's, i.e. the average effects for relative fitness), such that $\beta^{(1)}_{\bar{\alpha}} > 0$ but there should be no relationship between genetic diversities and the η's (i.e. $p_{\bar{\alpha}} = 0$). We implemented the fitness model described above by adapting SLiM's (Haller and Messer, 2023) default recipe for polygenic selection in the Wright-Fisher mode. To reach mutation-selection-drift balance, we then let this population of 2,500 individuals evolve forward in time with selection for 25,000 generations (see below for an explanation on the consequence of setting $N_e = 2500$ on other parameter choices). Non-neutral mutations drawn from the same DFE were allowed to occur in this period at a rate given by μ_{SLiM}. At generation 25,000, as the alleles reach mutation-selection-drift balance, we expect $\beta^{(1)}_{\bar{\alpha}}$ to have become more positive and $p_{\bar{\alpha}}$ to have become more negative, better reflecting a real population undergoing selection.

In generation 25,000, we sampled N_0 diploid individuals (typically 1000, unless specified otherwise) from this population, which then go on to become the base population in the next phase of the simulation. At this stage, we generated complete genomes for the N_0 individuals in the base population by using pyslim to add neutral mutations to the tree sequence recorded so far. To obtain our target number of loci, n_L (typically 65,000, unless specified otherwise), we set the neutral mutation rate to be $(n_L - n_{L_S})/g$ where n_{L_S} is the number of non-neutral segregating sites already present and g is the total branch length of the recorded tree sequence. We recorded the phased genotype of each parent at each locus, allowing us to construct $\mathbf{L}_0$ and $\mathbf{L}_0^{'}$.

Phase 2 ('experiment phase'): Simulating an evolve and resequence experiment

In the second phase of our simulations, again implemented in SLiM, we first allowed the base population to undergo one round of reproduction without selection to establish replicate experimental populations (typically 10 replicates, unless specified otherwise). Next, we allowed each of these populations to evolve forward in time

with selection as in the history phase. We allowed the number of generations in the experiment to be either 1, 3, or 5. Since our goal was to estimate V_A in the base population, we restricted our analyses only to the set of loci segregating in the base population. Any new mutations in subsequent generations would, in all likelihood, occur at loci outside this set. Therefore, we did not simulate new mutations during the experiment phase. For each of the independent replicate populations, we recorded the genome-wide vector of allele frequencies in each generation of the experiment.

Varying the true levels of V_A

We varied the true (i.e. simulated) levels of V_A in the base population from *ca.* 0.01 to 0.1 by changing the rate of non-neutral mutations in the history phase, implemented in msprime (μ_{msp}) and SLiM (μ_{SLiM}). In simulations with $\eta_{scale} = 0.033$ and no beneficial mutations, we varied μ_{msp} between 5.56×10^{-9} and 5.56×10^{-8} and μ_{SLiM} between 5.56×10^{-7} and 5.56×10^{-6} (for other scenarios see Table 1). We set μ_{msp} to be an order of magnitude smaller than μ_{SLiM} in most simulations because otherwise all individuals would have had a fitness of zero at the end of the coalescent part of the history phase. This extreme mutation load arises because deleterious alleles, having previously evolved under neutrality, would segregate at fairly high frequencies. However, the choice of μ_{msp} is unlikely to significantly affect the composition of the base population in most simulations, since the non-neutral genetic diversity of the base population of the experiment phase was primarily determined by the drift-recombination-mutation-selection equilibrium reached over the course of the 25,000 generation long forward simulation, and therefore primarily dependent on μ_{SLiM}.

Simplified simulations

In addition to the full simulations described above, we performed a set of simplified, proof-of-principle simulations to test the logic of our method. These simulations were different from the full simulations in three ways:

(1) The history phase was highly abbreviated – the forward simulation implemented in SLiM lasted only a single generation without any selection. This meant that the ancestral population – from which replicate experimental population were founded – had evolved entirely under neutrality.

(2) Rather than attach η's to the derived alleles generated with msprime, we attached the η's randomly with respect to either of the two alleles segregating at the locus. Amendments (1) and (2) generate a scenario where there is no relationship between allele frequency and η: both $p_{\bar{\alpha}}$ and $\beta^{(1)}_{\bar{\alpha}}$ are expected to be zero.

η_{scale}	$\mu_{ben} : \mu_{del}$	μ_{msp}	μ_{SLiM}	n_{L_S}
0.033	0	5.56×10^{-8} – 5.56×10^{-7}	5.56×10^{-7} – 5.56×10^{-6}	10058 – 63755
0.033	0.0002	5.56×10^{-8} – 5.56×10^{-7}	5.56×10^{-7} – 5.56×10^{-6}	9742 – 63797
0.033	0.02	2.0×10^{-8} – 2.0×10^{-7}	2.0×10^{-7} – 2.0×10^{-6}	3127 – 29999
0.045	0	3.6×10^{-8} – 3.6×10^{-7}	3.6×10^{-7} – 3.6×10^{-6}	6565 – 42097
0.1	0	1.8×10^{-8} – 1.6×10^{-7}	1.8×10^{-7} – 1.6×10^{-6}	2998 – 18011

Table 1: The ranges for the rates for non-neutral mutations used in the history phase of the full simulations implemented in msprime (μ_{msp}) (the coalescent simulation) and SLiM (μ_{SLiM}) (the forward simulation of the history), along with the ranges for the resulting number of non-neutral sites segregating in the population at the end of the history phase (n_{L_S}) in simulations with different values of the scale (η_{scale}) of the gamma distribution from which η's for non-neutral mutations were sampled, and the ratio of the rate of beneficial mutations to the rate of deleterious mutations in the history phase ($\mu_{ben} : \mu_{del}$). Note that the ranges for the number of selected loci, n_{L_S}, are only shown for the simulations where the map length in the history phase was 0.5 morgans.

(3) Given that the ancestral population had evolved without selection, these simplified simulations had considerably higher genetic diversities compared to the full simulations. This meant that fewer segregating non-neutral sites (n_{L_S}) were required to achieve true levels of V_A between 0.01 and 0.1. Therefore, in these simulations, we set n_L to be 3000, much lower than the 65,000 used for the full simulations. We achieved this by varying μ_{msp} between 3×10^{-9} and 2.35×10^{-8}. μ_{SLiM} was set to 0 in these simulations. This resulted in n_{L_S} at the end of the history phase being between 160 and 1902. We then added neutral mutations using a suitable rate as described above such that n_L was expected to be 3000.

Varying simulation parameters

Our aim was to investigate the sensitivity of our method to the following parameters: (1) map length in the history phase (0.5 morgans, 5 morgans, 50 morgans, 250 morgans), (2) map length in the experiment phase (0.01 morgans, 0.2 morgans, 2 morgans), (3) number of replicate populations in the experiment phase (3, 5, 10), (4) the population size of each of the replicate populations (100, 500, 1000), (5) number of generations over which allele frequency changes were recorded in the experiment phase (1, 3, 5), (6) the ratio of the rates of beneficial to deleterious mutations (0%, 0.02 %, 2%), and (7) the scale of the DFE, η_{scale} (0.033, 0.045, 0.100).

Our final choice of parameter ranges was limited by computational resources. Namely, the requirement for large $n_L \times n_L$ matrices (e.g. $\mathbf{L}_0$, $\mathbf{L}_0^{'}$, $\mathbf{M}$, and $\mathbf{N}$, etc.) imposed an upper bound of around 70,000 on n_L to permit the analysis of simulations in parallel. For full simulations, the number of non-neutral segregating sites ($n_{L_\mathcal{S}}$) was typically smaller than 65,000 (see Table 1). Therefore, in general, we set n_L to 65,000 in the full simulations. However, as the map length in the history phase increased, $n_{L_\mathcal{S}}$ also increased, presumably as a consequence of the strength of background selection being weakened by greater recombination. For the simulations where η_{scale} was 0.033, accommodating all levels of map length in the history phase (up to 250 morgans), would have required us to increase n_L to more than 100000, which was beyond the maximum n_L we could handle. Therefore, for $\eta_{scale} = 0.033$, we restricted ourselves to only those simulations where $n_{L_\mathcal{S}}$ was less than or equal to 65000. When the map length in the history phase was 0.5 morgans, we could analyse simulations in which the true V_A spanned the whole of the target range (0.01 to 0.1). However, for higher map lengths, the analyses we present here do not include simulations where the true V_A was close to the higher end of the target range (see Figure 4B).

Our choices for recombination rates (or map lengths) were dictated by the fact that our simulations were scaled in two ways relative to, say, the global *Drosophila melanogaster* population: (1) in the history phase of our simulations, we simulate a population of 2,500 individuals, which is 532 times smaller than 1.33×10^6, the estimate of N_e in *D. melanogaster* (Campos *et al.*, 2013; Campos and Charlesworth, 2019), and (2) we assume that all the heritable variation in fitness is limited to a genomic region that is one million base-pairs long - nearly 200 times smaller than the *D. melanogaster* genome. Therefore, the effective crossover rate in *D. melanogaster* would need to be scaled appropriately. We assumed the true effective crossover rate in *D. melanogaster* to be 1×10^{-8} per generation per site, which – to control for the absence of crossovers in males – is half the standard estimate of female recombination rate in European *D. melanogaster* (Wang *et al.*, 2023) (also see Comeron *et al.* (2012)). To account for (1), the recombination rate would need to be scaled up by a factor of 532 such that the product of N_e and the recombination rate remains conserved (similar to Table 1 in Campos and Charlesworth (2019)). The map length of the simulated region would then be 5.32 morgans. However, our study focuses on the *total* additive genetic variation for a trait, and therefore, we also need to account for (2). Specifically, since all the sites that code for fitness in our simulations are concentrated in a region that is about 50 times smaller than a typical *D. melanogaster* chromosome, we would need to further scale up the recombination rate 50-fold such that the average number of crossovers between a randomly chosen pair of sites is

comparable to a typical *D. melanogaster* chromosome. Under this scaling the map length of the simulated region would then be 266 morgans.

It is important to note that our method relies on linkage disequilibria D_{ij} at the end of the history phase, as well as the crossover rate r_{ij} between pairs of loci during the experiment phase. Accounting only for (1), and using a map length of 5.32 morgans, could in principle lead to reasonable D_{ij}'s. However, the r_{ij}'s will be considerably lower. Accounting for both (1) and (2), however, would result in reasonable r_{ij}'s but much lower D_{ij}'s. Therefore, in our simulations, we try to span this entire range by running simulations with the map length in the history phase chosen to be either 0.5, 5, 50, or 250 morgans. We use a similar logic for choosing recombination rates in the experiment phase. However, our goal in the experiment phase was to simulate an actual population with census size N_0 individuals, rather than mimicking the effects of evolution at a global scale with a realistic N_e. Therefore, in the experiment phase we do not account for (1). Instead we vary the map length in the experiment phase over a broad range (0.01 to 2 morgans) that spans both extremes – having a reasonable crossover rate per site per generation (1×10^{-8}) at one end of the simulated parameter range - which translates to a map length of 0.01 morgans - and having a sensible average r_{ij} between pairs of site at the other end of the range, by using a map length comparable to that in *D. melanogaster*.

Rather than vary all parameters in a fully factorial design, we selected a reference parameter set and explored the sensitivity of the method by changing each parameter in turn. The reference parameter set was (1) a map length of 0.5 morgans in the history phase (2) a map length of 2 morgans in the experiment phase, (3) 10 replicate populations in the experiment phase, (4) a population size of 1000 in each of the replicate populations in the experiment phase, (5) 3 generations over which allele frequency changes were recorded in the experiment phase, (6) no beneficial mutations in the history phase, and (7) 0.033 for the scale of the DFE, η_{scale}. The reference parameter set was chosen to best reflect evolve-and-resequence experiments using *D. melanogaster*.

For logistical reasons, in the simulations in which we varied the map length in the history phase, we could not span the entire target range for V_A (0.01 to 0.1) when η_{scale} was set to its default level of 0.033 (see above). Therefore, we ran an additional set of full simulations in which we varied the map length in the history phase while setting η_{scale} to 0.045. In these simulations, even at a map length of 250 morgans, n_{L_S} seldom exceeded 65,000, allowing us to set n_L to 67,500 for all map lengths.

Comparison with B&C

To compare the precision and the accuracy of our approach to that of B&C, we performed additional simplified and full simulations. Although originally designed for the covariance in allele frequency change between multiple time points, B&C's method can be readily adapted to allele frequency changes recorded over a single generation in multiple independent evolutionary replicates (see Equation S49 and Buffalo and Coop (2020)). We set simulation parameters to their reference values (see above) with a few modifications. First, to make comparisons of biases clearer, we reduced noise by employing 50 replicate populations, and second, allele frequency change in the experiment phase was recorded over a single generation. We also ran these simulations at four different levels of map length in the history phase: 0.5 morgans, 5 morgans, 50 morgans, and 100 morgans.

To investigate the consequences of Assumptions B, Ib, and N, we implemented B&C's method using six different approaches. For clarity, we label the six approaches using a code (see Table 2) that indicates with superscripts whether an assumption is required (+) or not required (0) for an approach to work. To test Assumption B, we recorded allele frequency changes at either all segregating sites (B^+) or neutral sites only (B^0). To test Assumption Ib, we used either the average LD between all sites (Ib^+) or the average LD between selected sites with neutral sites (Ib^0). And finally, we either divided the (co)variance in allele frequency change throughout by twice the average genetic diversity (N^+; analogous to Equation 16 in B&C) or used the (co)variance of weighted allele frequency change, where the weights are the inverse of the square root of twice the genetic diversity (N^0).

Results

Simplified simulations

We begin by discussing results from the simplified simulations in which we expect there to be no relationship between allele frequencies and α's. Under our reference parameter set, our method provided precise and unbiased estimates of V_A throughout the simulated range (0.01– 0.1) (Figure 1A). Furthermore, the estimates of $p_{\bar{\alpha}}$, $\beta_{\bar{\alpha}}^{(0)}$, and $\beta_{\bar{\alpha}}^{(1)}$ were centred around 0 as expected (Figure 1B-D). We next investigated how this relationship between the true and estimated levels of V_A was affected by changing (1) map length of the genomic region being simulated, (2) population size of each replicate population in the evolve and resequence experiment, (3) number of replicate populations, and (4) number of generations over which allele frequency

Approach	$\Delta\mathbf{p}$	Average LD between	$Cov(\Delta\mathbf{p}_m, \Delta\mathbf{p}_n)$
$B^+Ib^+N^+$	all sites	all sites	divided throughout by $(\bar{p}_m\bar{q}_m + \bar{p}_n\bar{q}_n)/2$
$B^0Ib^+N^+$	neutral sites	all sites	divided throughout by $(\bar{p}_m\bar{q}_m + \bar{p}_n\bar{q}_n)/2$
$B^0Ib^0N^+$	neutral sites	selected & neutral sites	divided throughout by $(\bar{p}_m\bar{q}_m + \bar{p}_n\bar{q}_n)/2$
$B^+Ib^+N^0$	all sites	all sites	replaced by $Cov(\Delta\overrightarrow{\mathbf{p}_m}, \Delta\overrightarrow{\mathbf{p}_n})$
$B^0Ib^+N^0$	neutral sites	all sites	replaced by $Cov(\Delta\overrightarrow{\mathbf{p}_m}, \Delta\overrightarrow{\mathbf{p}_n})$
$B^0Ib^0N^0$	neutral sites	selected & neutral sites	replaced by $Cov(\Delta\overrightarrow{\mathbf{p}_m}, \Delta\overrightarrow{\mathbf{p}_n})$

Table 2: We applied B&C's method to our simulations using six different approaches. The requirement of Assumptions B, Ib, and N is indicated using superscripts ("+" when required and "0" when not required). We used allele frequency change ($\Delta\mathbf{p}$) at either all segregating sites (Assumption B required) or at neutral segregating sites only (Assumption B not required). We computed average LD using either the LD between all segregating sites (Assumption Ib required) or the LD between selected and neutral sites (Assumption Ib not required). We used either the (co)variance in $\Delta\mathbf{p}$ divided throughout by twice the average genetic diversity (Assumption N required) or the (co)variance in weighted $\Delta\mathbf{p}$ (i.e. $\Delta\overrightarrow{\mathbf{p}}$) where the weights are the square roots of twice the genetic diversity at each site (Assumption N not required).

changes were recorded in the experiment. Our estimates of V_A remained unbiased as the map length in the experiment became smaller, although estimates became noisier, particularly at higher values of V_A (Figure 2A). As expected, estimates also became noisier as the number of individuals, replicate populations, or generations became smaller (Figure 2B-D). However, estimates appeared upwardly biased when the number of replicate populations or generations was small (Figure 2C-D).

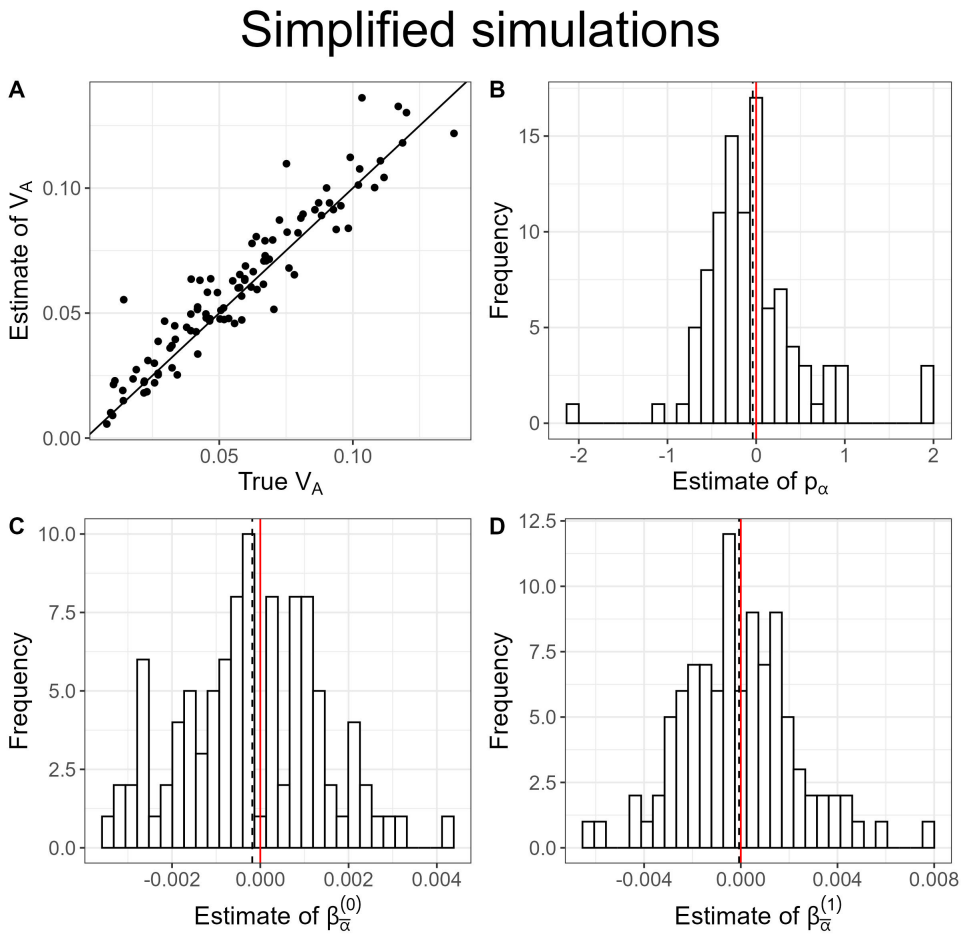

Figure 1: Results of simplified simulations (map length in the history phase = 0.5 morgans, map length in the experiment phase = 2 morgans, number of replicate populations = 10, population size = 1000, number of generations = 3, $\eta_{scale} = 0.033$). (A) A scatter plot of estimates of V_A vs true values of V_A. The solid black line indicates the 1:1 line. (B)-(D) Histograms of the estimates of $p_{\bar{\alpha}}$, $\beta_{\bar{\alpha}}^{(0)}$, and $\beta_{\bar{\alpha}}^{(1)}$, respectively. The vertical red line indicates 0.

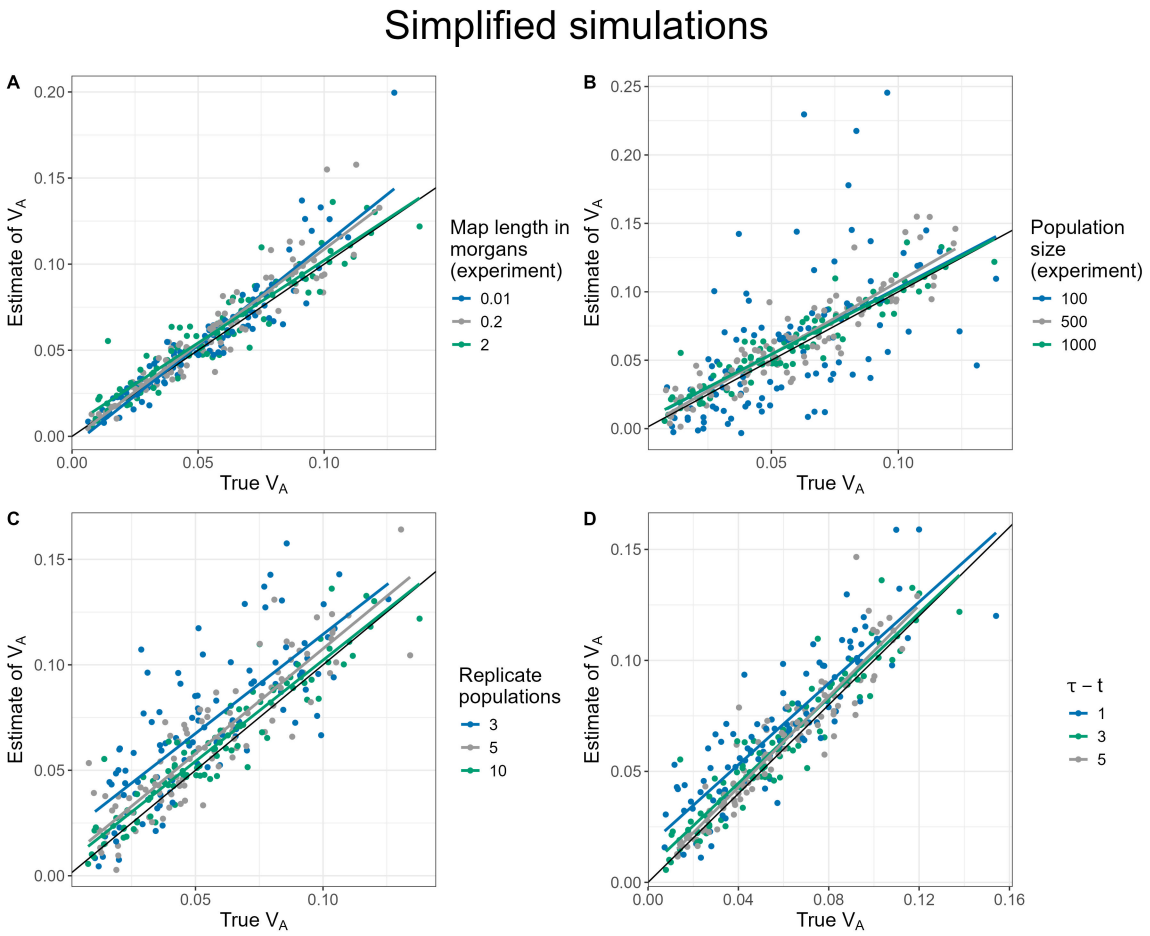

Figure 2: Scatter plots of estimates of V_A vs true values of V_A for simplified simulations at different levels of (A) map length in the experiment, (B) population size of each replicate population in the evolve and resequence experiment, (C) number of replicate populations, and (D) number of generations for over which allele frequency changes were recorded in the experiment. In each case, other than the parameter to be varied, the other parameters were fixed at their default values ($\eta_{scale} = 0.033$, map length in the history phase = 0.5 morgans, map length in the experiment phase = 2 morgans, number of replicate populations = 10, population size = 1000, number of generations = 3). The solid black line indicates the 1:1 line. The coloured lines represent regression lines for estimates of V_A vs true values of V_A.

Full simulations

In our full simulations we let the ancestral population evolve forward in time with selection for 25000 generations before simulating the experiment. In these simulations, we expect the population to be at drift-recombination-mutation-selection equilibrium, such that p_α is negative and $\beta_{\tilde{\alpha}}^{(1)}$ is positive. Results from our standard set (Figure 3) suggest that not only does our method provide precise and unbiased estimates of V_A, but it does so by correctly estimating the signs of p_α and $\beta_{\tilde{\alpha}}^{(1)}$ (Figure 3B,D). As before, estimates of V_A became slightly noisier at shorter map lengths during the experiment (Figure 4A). The quality of our estimates of V_A was not significantly affected by adding beneficial mutations during the history phase (Figure 4C). On the other hand, our estimates were downwardly biased at larger map lengths in the history phase (Figure 4B, also see Figure S1) and when non-neutral loci had on average larger fitness effects (i.e. when η_{scale} was larger) (Figure 4D). The bias at larger η_{scale} was likely driven by the loss of additive genic variance at large-effect loci during the experiment (Figure S2A-D).

Finally, we re-analysed our standard set of full simulations in two different ways (Figure 4E-F). First, to accommodate cases in which phased genomes, and therefore $\mathbf{L}_0^{'}$, are unavailable, we assumed that $\mathbf{L}_0^{''}$ was 0 and set $\mathbf{L}_0^{'} = \mathbf{L_0}$. In other words, we assumed that $\tilde{\mathbf{L}}_0 = \mathbf{L}_0$. Our analyses suggest that this assumption leads to a slight downward bias in our estimates of V_A (Figure 4E). Second, to account for instances where N_E in the experiment phase is unknown, we replaced our estimate of N_E in the experiment phase by the number of individuals. Our analyses suggest that this does not affect our estimates of V_A in any way (Figure 4F), with the faster than expected drift being absorbed by an increased residual variance (Figure S3): estimates of N_E can be obtained by dividing the assumed value of N_E by the estimated residual variance.

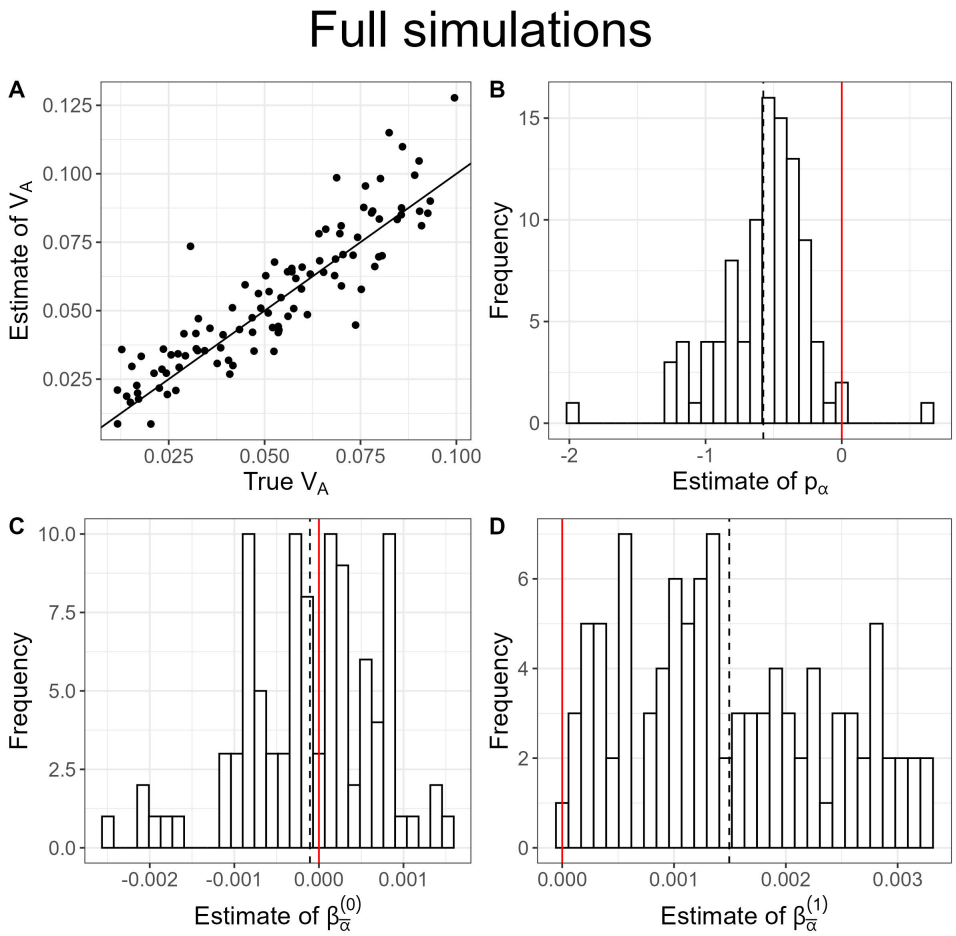

Figure 3: Results of full simulations with a burn-in phase of 25000 generations (map length in the history phase = 0.5 morgans, map length in the experiment phase = 2 morgans, number of replicate populations = 10, population size = 1000, number of generations = 3, η_{scale} = 0.033). (A) A scatter plot of estimates of V_A vs true values of V_A. The solid black line indicates the 1:1 line. (B)-(D) Histograms of the estimates of $p_{\bar{\alpha}}$, $\beta_{\bar{\alpha}}^{(0)}$, and $\beta_{\bar{\alpha}}^{(1)}$, respectively. The vertical red line indicates 0.

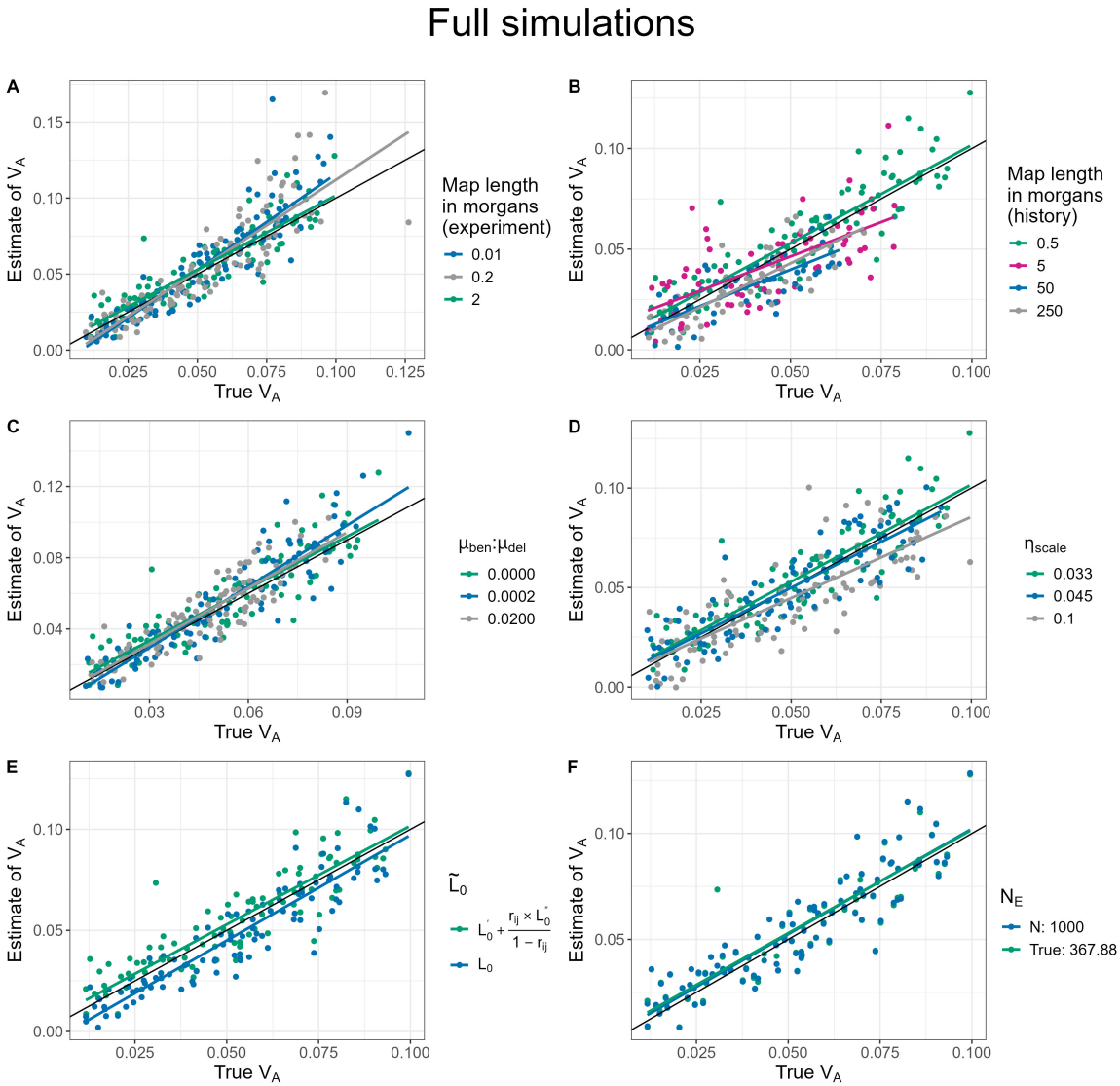

Figure 4: Scatter plots of estimates of V_A vs true values of V_A for full simulations with a burn-in phase of 25000 generations at different values of (A) map length in the experiment phase, (B) map length during the history phase, (C) the ratio of rates of beneficial and deleterious mutations in the history phase, and (D) η_{scale} for non-neutral mutations. In each case, other than the parameter to be varied, the other parameters were fixed at their default values (map length in the history phase = 0.5 morgans, map length in the experiment phase = 2 morgans, number of replicate populations = 10, population size = 1000, number of generations = 3), η_{scale} for non-neutral mutations = 0.033. The solid black line indicates the 1:1 line. The coloured lines represent regression lines for estimates of V_A vs true values of V_A. The effect of analysing the standard set of simulations (Figure 3) (E) without partitioning $\tilde{\mathbf{L}}_0$ into its gametic ($\mathbf{L}_0'$) and non-gametic phase ($\mathbf{L}_0''$) components, and (F) using $N_E = N = 1000$. Note that in (F), a large number of green points are eclipsed by the blue points.

Comparison with B&C

In both simplified and full simulations, all implementations of B&C's method resulted in a strong upward bias at low recombination rates in the history phase (Figures 5 and6). This was likely a consequence of Assumption G being violated; the correlation between the contribution of a selected site j to V_a and the persistent association of this site with neutral alleles ($\sum_{\mathcal{N}_i}(R_{i_m j_m} R_{i_n j_n})$) was considerably higher than 0 when the map length in the history phase was low (panel H of Figures 5 and 6). In contrast, the estimates of V_A provided by our method were unbiased and had far greater precision compared to any of the six implementations of B&C's method in the simplified simulations (Figure 5G), although there was a small upward bias at low recombination rates in the full simulations (Figure 6G). Due to the difficulty of distinguishing selected from unselected sites, B&C suggested that the average LD between all sites could be used instead of the LD between selected and unselected sites (Assumption Ib) and indeed this assumption seems to be well justified (panels D & F in Figures 5 and 6 are very similar). However, an inability to distinguish selected from unselected sites would likely mean that B&C's method is applied to allele frequency change data from *all* segregating sites, contravening Assumption B. This results in considerable overestimation of V_A in simplified simulations (Figure 5A-B) but not in full simulations (Figure 6A-B). Under all scenarios the bias in the method of B&C can be reduced (but not eliminated) by relaxing Assumption N (compare panels B, D, and F in Figures 5-6 to panels A, C and E). Under assumption N, it is assumed that the (co)variance in allele frequency change weighted by the square root of twice the genetic diversities can be approximated by the (co)variances in allele frequency change divided through by twice the average diversities. However, this approximation is not required.

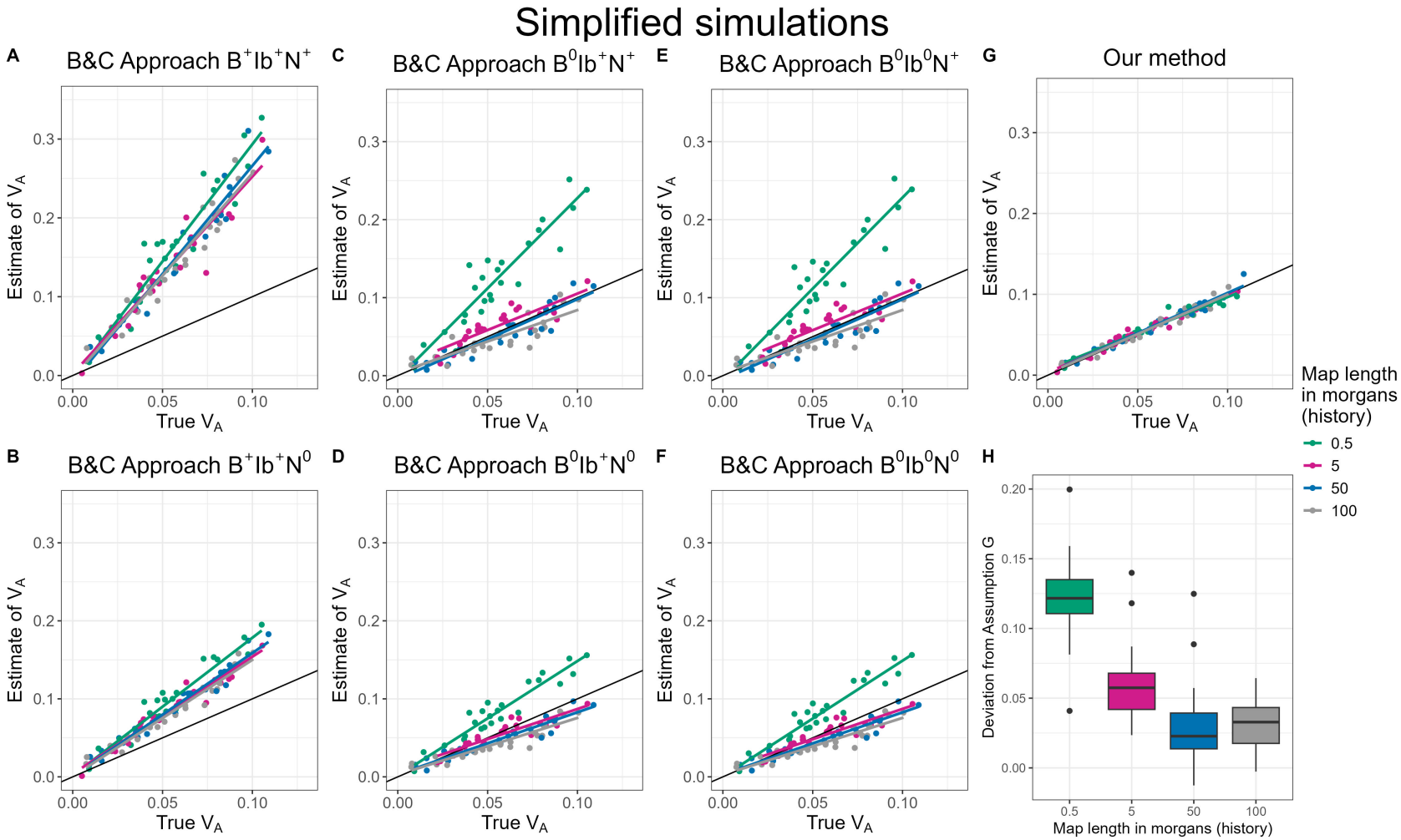

Figure 5: Estimates of V_A obtained using our method (G) and various applications of the B&C Approach (A–F) from simplified simulations with varying map lengths in the history phase: 0.5 morgans (green), 5 morgans (magenta), 50 morgans (blue), and 100 morgans (grey). The solid black line indicates the 1:1 line. Each application of B&C is defined in terms of which three assumptions (B, Ib, N) are required or not (designated with a + or 0 superscript, respectively). Assumption B is that allele frequencies are tracked at unselected loci only, Assumption Ib is that the average LD between all sites is equal to the average LD between selected and unselected sites, and Assumption N is that the covariance between projected allele frequency changes is equal to the covariance between allele frequencies changes scaled by the average projection (see Table 2 for details). Panel H is the correlation between the contribution of a selected sites to V_a and their persistent associations with neutral alleles (i.e. deviations from Assumption G in B&C)

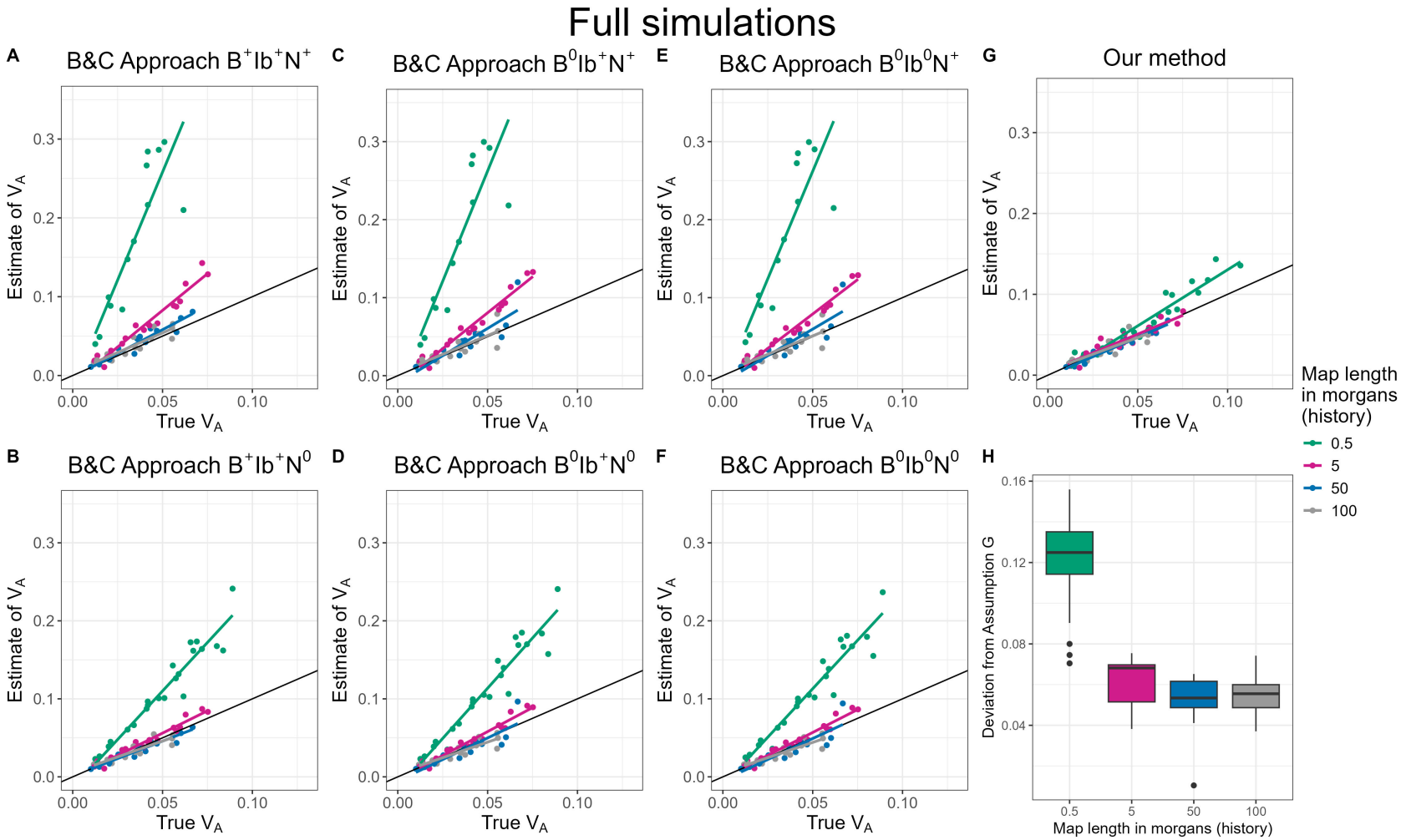

Figure 6: Estimates of V_A obtained using our method (G) and various applications of the B&C Approach (A–F) from full simulations with varying map lengths in the history phase: 0.5 morgans (green), 5 morgans (magenta), 50 morgans (blue), and 100 morgans (grey). The solid black line indicates the 1:1 line. Each application of B&C is defined in terms of which three assumptions (B, Ib, N) are required or not (designated with a + or 0 superscript, respectively). Assumption B is that allele frequencies are tracked at unselected loci only, Assumption Ib is that the average LD between all sites is equal to the average LD between selected and unselected sites, and Assumption N is that the covariance between projected allele frequency changes is equal to the covariance between allele frequencies changes scaled by the average projection (see Table 2 for details). Panel H is the correlation between the contribution of a selected sites to V_a and their persistent associations with neutral alleles (i.e. deviations from Assumption G in B&C). For clarity of presentation, in (A), (C), and (E), we have restricted the range of the y axis between 0 and 0.35. As a consequence, points having an estimate of V_A above 0.35 have been excluded from the plot for 0.5 morgans (green).

Discussion

In this paper we estimate the additive genetic variance for relative fitness (V_A) directly from the change in the genetic composition of a population caused by selection. Assuming only the absence of meiotic drive, we show that V_A can be conveniently expressed as a function of the genome-wide genetic diversity matrix $\mathbf{L}$, and the vector of genome-wide expected allele frequency change due to selection $E[\Delta\mathbf{p}]$. In our inference approach, we describe how a linear mixed model can be employed to estimate $E[\Delta\mathbf{p}]$ via independent evolutionary replicates derived from the same base population with a known $\mathbf{L}$ – a common feature of evolve-and-resequence studies. Unlike alternative methods (Buffalo and Coop, 2019), this allows us to obtain estimates of V_A that are largely robust to the underlying genetic properties of the population and the genetic architecture of fitness. Moreover, the underlying modeling framework not only allows V_A to be estimated, but allows inferences about the relationship between effect sizes and allele frequency.

Although $E[\Delta p]$'s at individual loci cannot be usefully estimated, for our purposes it is sufficient to estimate their distribution as parameterised through the mean vector $\boldsymbol{\mu}_{\bar{\boldsymbol{\alpha}}}$ and the (co)variance matrix $\boldsymbol{V}_{\bar{\boldsymbol{\alpha}}}$ for the distribution of the average effects for fitness, the α's. Our inference approach uses a low-dimensional (4-parameter), but biologically sensible, model for the means and (co)variances. It is superficially surprising that such a simple model for an n_L (the number of loci) dimensional distribution can produce accurate results. However, in a typical dataset one expects the number of individuals, N, to be far smaller than the number of loci, n_L. Therefore, the non-null subspace of $\mathbf{L}$ has only N dimensions, and we can work with allele frequency changes projected into this reduced space. This provides a route to understanding how the distribution of α's can be estimated from selected changes in allele frequency that must be negligible compared to the impact of drift: the projection defines '*chunks*' of genome, and it is the frequency changes in these chunks, rather than individual alleles, that are tracked. Since the aggregate fitness effects of alleles across chunks will be more substantial, they can be more easily detected. Moreover, since these aggregate effects may involve a large number of loci, they will, thanks to the central limit theorem, tend to normality and, conditional on the projection, converge in distribution. Since the projection is defined by $\mathbf{L}$, a model of the α's that conditions on aspects of $\mathbf{L}$ ($p-q$ and pq in our case) is expected to be sufficient.

The results of our simulations demonstrate that our approach provides usefully precise and consistent estimates of V_A over a wide range of parameter combinations and experimental designs. In the simplified simulations, in which the relationship between allele frequencies and α's was expected to be absent, the model requires

only a single parameter: the variance in average effects, $\sigma^2_{\bar{\alpha}}$. If such a model is assumed, this is a relatively straightforward inference problem and the method of Buffalo and Coop (2019) can give accurate estimates if selected and neutral sites can be distinguished, and certain patterns of recombination and linkage-disequilibrium hold. Although our approach does not assume such a model, it can accurately infer V_A and also the lack of relationship between allele frequencies and α's, with the three parameters that determine the relationship between allele frequency and α all centered on their null expectations (Figure 1). However, in reality, independence between allele frequency and α seems implausible at mutation-selection-drift balance, and the elevated contribution of high-diversity loci to V_A under the simple, but unrealistic, scenario may result in greater power to detect V_A.

At mutation-selection-drift balance a negative relationship is expected between α's and allele frequencies (Charlesworth and Charlesworth, 2010). Although it could be argued that the strength of this relationship will be reduced in experimental evolution studies, which generally expose populations to novel environmental stressors (such as a laboratory environment, pathogens (Basu *et al.*, 2024), extreme population densities (Joshi and Mueller, 1996) and temperatures (Singh *et al.*, 2015; Hsu *et al.*, 2024), desiccation (Gibbs *et al.*, 1997), malnutrition (Kawecki *et al.*, 2021), or toxic substances (Godinho *et al.*, 2024; Xiao *et al.*, 2019)) it seems unlikely that no relationship would persist. Even under the more realistic scenario of our full simulations, in which the base population had evolved under selection for 25,000 generations, our approach provided reliable estimates of V_A that were remarkably robust to the details of the distribution of fitness effects, such as the relative frequency of beneficial mutations, or properties of the population such as the recombination rate. Furthermore, it did so by correctly inferring a negative relationship between fitness effects of alleles and genetic diversity, and a positive relationship between fitness effects and $p - q$, as expected under mutation-selection balance.

Although our method generally performs well, some biases were observed. In terms of experimental design, small upward biases are evident when power is low - either because there are few replicates and/or allele frequency change is only calculated over a single generation. In terms of genetic architecture, our approach marginally underestimates V_A when the fitness effects of new non-neutral mutations are large. This downward bias seems to arise because there is a large contribution of rare highly deleterious variants to V_A and these get lost during the experimental phase. Nevertheless, even when 25% of V_A was lost during the course of the experiment the impact on estimates was rather minor. Similarly, modest downward biases were observed when the recombination rate in the history phase was increased, but the source of this bias has been harder to diagnose. Increasing the recombination

rate resulted in an increase in the number of segregating selected sites and their genetic diversity, and a steeper relationship between α and genetic diversity, consistent with a reduction in the effects of background selection (Charlesworth *et al.*, 1993) on weakly selected sites (Stephan *et al.*, 1999). While it is not clear why this causes downward bias in the estimates, it is also unclear whether real populations would ever have such extreme genetic architecture where the bulk of the additive genetic variance for fitness is caused by highly deleterious variants segregating at very low frequencies. In order to make the forward simulations manageable we were working with considerably smaller population and genome sizes than are typical of real populations. It is not clear, however, whether the standard rescaling of mutations rates, recombination rates and selection coefficients to accommodate this downsizing results in genetic architectures that would be typical of larger populations and larger genomes (Dabi and Schrider, 2025).

While our method performs well when applied to simulated data, application to real-world data would involve overcoming a number of challenges. First, we require genome-wide allele frequency change data from multiple independent evolutionary replicates – although this should be readily available for most evolve-and-resequence experiments. Second, we require the genetic diversity matrix $\mathbf{L}$ in the base population from which the replicates are derived. This is not always the case for evolve-and-resequence studies, in which base populations are often split into replicate baseline populations long before the experiment, or when newer selection regimes are derived mid-experiment (Burke *et al.*, 2010; Singh *et al.*, 2015; Gupta *et al.*, 2016; Robinson *et al.*, 2023). Furthermore, our method requires that sufficient individuals from the base population are individually sequenced to estimate $\mathbf{L}$, or – even better – its gametic phase ($\mathbf{L}^{'}$) and non-gametic phase ($\mathbf{L}^{''}$) components, although the required phasing should become more readily available with long-read sequencing. Third, to predict how $\mathbf{L}$ evolves, we require an estimate of the recombination probability between all pairs of segregating sites, such as a recombination map for the population. In reality, recombination maps are likely to have been derived from other populations, in which recombination patterns may differ (Johnston, 2024). Although recombination maps can be approximated, for example by using Haldane's mapping function, it is not clear how sensitive the method is to errors in the recombination map. Fourth, our method requires the mean number of generations over which allele frequency changes are calculated, which may be hard to infer with overlapping generations. As a first approximation, $V_A \approx E[\Delta(\mathbf{p}^\top)]\mathbf{L}^{-1}E[\Delta(\mathbf{p})]/(\tau - t)^2$ where $\tau - t$ is the number of generations over which allele frequency is tracked. Consequently, estimates might be out by a factor $(\tau - t)^2/(\widehat{\tau - t})^2$. Fifth, our method uses effective population size to predict how $\mathbf{L}$ evolves (N_e) and to derive expressions for the drift (co)variance

(N_E). Although reliable estimates of both effective population sizes may be hard to obtain, this is unlikely to be a major issue because: (i) $\mathbf{L}^{'}$ decays with a rate roughly proportional to $1 - 1/N_e$ (likely making it insensitive to errors in estimating N_e) and (ii) our simulations suggest that using the wrong N_E does not adversely affect our estimates of V_A because the mispecification is absorbed by the residual variance of the model. Sixth, our simulations minimise the unpredictable response to selection because our fitness model is close to being additive and so the average effects remain relatively constant as allele frequencies change. However, with greater non-additivity and/or selection coefficients that vary in time or across replicates, the unpredictable response to selection will be greater. Indeed in outdoor mesocosms of *D. melanogaster*, Bitter *et al.* (2024) report that allele frequency changes can switch signs over time-points separated by a matter of weeks in spite of exhibiting highly concordant evolution between replicates. It is not clear to what degree this will affect inferences. Finally, selection can often act in different ways in different contexts such as space (Whitlock, 2015; Delph, 2018), time, and between the two sexes (Schenkel *et al.*, 2018). Our approach captures the effects of selection averaged over all these different contexts. Specifically, if loci have different fitness effects in males and females, we effectively estimate $(V_{A,f} + V_{A,m} + 2COV_{A,mf})/4$, where $V_{A,f}$ and $V_{A,m}$ are the additive genetic variances for relative fitness in females and males respectively, and $COV_{A,mf}$ is the intersexual additive genetic covariance for relative fitness. For these reasons, when applied to replicate populations, our estimates of V_A are perhaps best thought of as the additive genetic *covariance* in fitness between replicates. Although rarely made explicit, estimates from wild systems should be interpreted in the same way: the additive genetic covariance between the environments in which relatives live (Vehviläinen *et al.*, 2008).

Analysing data from wild populations is likely to entail further challenges. For example, in natural populations, allele frequency change from immigration may be consequential, and without some modification is likely to be mistaken for allele frequency change caused by natural selection (Simon and Coop, 2024). Furthermore, replicate populations are unlikely to be available for natural systems – with some exceptions such as Trinidadian guppies (Reznick and Bryga, 1996) – although with some modifications our method could be adapted to use allele frequency change data from multiple time points. However, a lack of individual-level sequences in the base population would mean that the estimates of $\mathbf{L}$ will likely be considerably noisier in natural populations.

While we accept that the data requirements for our method are steep, and the list of caveats appears long, we do think that information about V_A can be successfully leveraged from current evolve and resequence studies. Going forward, we hope this

work will inform future evolve and resequence study design, and that estimates of V_A from a wide range of organisms and environments becomes available. Partitioning V_A into genomic features is an obvious next step, and in the future we hope to go beyond simply knowing the magnitude of V_A and start to understand its underlying causes.

Data Availability Statement

This study does not use any data. The code used for simulations and analyses is available in the following GitHub repository: https://github.com/manas-ga/Va_simulations.

Acknowledgments

The authors would like to thank Bill Hill, Brian Charlesworth, Peter Keightley, Konrad Lohse, Bruce Walsh, Ben Longdon and Vince Buffalo for their insightful comments, and The Argyle for hosting us at the early stages of this project. Computer simulations described here were performed on the AC3 computing cluster based at Ashworth Laboratories, the University of Edinburgh.

Funding

This work was funded by a Natural Environment Research Council (NERC) grant (reference: NE/W001330/1). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission.

Conflict of Interest

The authors have no conflicts of interest to declare.

References

Abbott, J. K., and E. H. Morrow, 2011 Obtaining snapshots of genetic variation using hemiclonal analysis. Trends in Ecology & Evolution **26**: 359–368.

Barton, N. H., A. M. Etheridge, and A. Véber, 2017 The infinitesimal model: Definition, derivation, and implications. Theoretical population biology **118**: 50–73.

Barton, N. H., and L. Partridge, 2000 Measuring fitness by means of balancer chromosomes. Genetics Research **75**: 297–313.

Basu, A., K. Tekade, A. Singh, P. N. Das, and N. G. Prasad, 2024 Experimental evolution for improved postinfection survival selects for increased disease resistance in *Drosophila melanogaster*. Evolution **78**: 1831–1843.

Bitter, M., S. Berardi, H. Oken, A. Huynh, E. Lappo *et al.*, 2024 Continuously fluctuating selection reveals fine granularity of adaptation. Nature **634**: 389–396.

Bonnet, T., M. B. Morrissey, P. de Villemereuil, S. C. Alberts, P. Arcese *et al.*, 2022 Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. Science **376**: 1012–1016.

Buffalo, V., and G. Coop, 2019 The linked selection signature of rapid adaptation in temporal genomic data. Genetics **213**: 1007–1045.

Buffalo, V., and G. Coop, 2020 Estimating the genome-wide contribution of selection to temporal allele frequency change. Proceedings of the National Academy of Sciences **117**: 20672–20680.

Bulmer, M., 1971 The effect of selection on genetic variability. The American Naturalist **105**: 201–211.

Bulmer, M. G. *et al.*, 1980 *The mathematical theory of quantitative genetics.*. Clarendon Press.

Burke, M. K., J. P. Dunham, P. Shahrestani, K. R. Thornton, M. R. Rose *et al.*, 2010 Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature **467**: 587–590.

Burt, A., 1995 The evolution of fitness. Evolution **49**: 1–8.

Butler, D., B. R. Cullis, A. R. Gilmour, B. J. Gogel, and R. Thompson, 2023 *ASReml-R Reference Manual, Version 4.2*. Queensland DPI, Brisbane, Australia.

Campos, J. L., and B. Charlesworth, 2019 The effects on neutral variability of recurrent selective sweeps and background selection. Genetics **212**: 287–303.

Campos, J. L., K. Zeng, D. J. Parker, B. Charlesworth, and P. R. Haddrill, 2013 Codon usage bias and effective population sizes on the X chromosome versus the autosomes in *Drosophila melanogaster*. Molecular biology and evolution **30**: 811–823.

Charlesworth, B., and D. Charlesworth, 2010 *Elements of evolutionary genetics*, volume 42. Springer.

Charlesworth, B., M. Morgan, and D. Charlesworth, 1993 The effect of deleterious mutations on neutral molecular variation. Genetics **134**: 1289–1303.

Charmantier, A., and B. C. Sheldon, 2006 Testing genetic models of mate choice evolution in the wild. Trends in Ecology & Evolution **21**: 417–419.

Comeron, J. M., R. Ratnappan, and S. Bailin, 2012 The many landscapes of recombination in *Drosophila melanogaster*. PLoS Genetics **8**: e1002905.

Coulson, T., T. Benton, P. Lundberg, S. Dall, B. Kendall *et al.*, 2006 Estimating individual contributions to population growth: evolutionary fitness in ecological time. Proceedings of the Royal Society B: Biological Sciences **273**: 547–555.

Dabi, A., and D. R. Schrider, 2025 Population size rescaling significantly biases outcomes of forward-in-time population genetic simulations. Genetics **229**: iyae180.

de Los Campos, G., D. Sorensen, and D. Gianola, 2015 Genomic heritability: what is it? PLoS Genetics **11**: e1005048.

Delph, L. F., 2018 The study of local adaptation: a thriving field of research. Journal of Heredity **109**: 1–2.

Ewens, W. J., 1989 An interpretation and proof of the fundamental theorem of natural selection. Theoretical population biology **36**: 167–180.

Ewens, W. J., 2024 The fundamental theorem of natural selection: the end of a story. Evolution **78**: 803–808.

Falconer, D., and T. Mackay, 1996 *Introduction to Quantitative Genetics, 4th edn*. Addison Wesley Longman, Harlow.

Fisher, R. A., 1930 *The genetical theory of natural selection*. Oxford University Press.

Fisher, R. A., 1941 Average excess and average effect of a gene substitution. Annals of Eugenics **11**: 53–63.

Fisher, R. A., 1958 *The genetical theory of natural selection, 2nd revised edn*. Dover Publications, New York.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock *et al.*, 2007 Sexually antagonistic genetic variation for fitness in red deer. Nature **447**: 1107–1110.

Fowler, K., C. Semple, N. H. Barton, and L. Partridge, 1997 Genetic variation for total fitness in *Drosophila melanogaster*. Proceedings of the Royal Society of London. Series B: Biological Sciences **264**: 191–199.

Frank, S. A., and M. Slatkin, 1992 Fisher's fundamental theorem of natural selection. Trends in Ecology & Evolution **7**: 92–95.

Gardner, M. P., K. Fowler, N. H. Barton, and L. Partridge, 2005 Genetic variation for total fitness in *Drosophila melanogaster*: complex yet replicable patterns. Genetics **169**: 1553–1571.

Gianola, D., G. de Los Campos, W. G. Hill, E. Manfredi, and R. Fernando, 2009 Additive genetic variability and the Bayesian alphabet. Genetics **183**: 347–363.

Gibbs, A. G., A. K. Chippindale, and M. R. Rose, 1997 Physiological mechanisms of evolved desiccation resistance in *Drosophila melanogaster*. Journal of Experimental Biology **200**: 1821–1832.

Godinho, D. P., I. Fragata, A. Majer, L. R. Rodrigues, and S. Magalhães, 2024 Limits to the adaptation of herbivorous spider mites to metal accumulation in homogeneous and heterogeneous environments. Journal of Evolutionary Biology : voae003.

Grafen, A., 2015 Biological fitness and the fundamental theorem of natural selection. The American Naturalist **186**: 1–14.

Gupta, V., S. Venkatesan, M. Chatterjee, Z. A. Syed, V. Nivsarkar *et al.*, 2016 No apparent cost of evolved immune response in *Drosophila melanogaster*. Evolution **70**: 934–943.

Haldane, J. B. S., 1919 The combination of linkage values, and the calculation of distances between the loci of linked factors. Journal of Genetics **8**: 299–309.

Haller, B. C., and P. W. Messer, 2023 SLiM 4: multispecies eco-evolutionary modeling. The American Naturalist **201**: E127–E139.

Hendry, A. P., D. J. Schoen, M. E. Wolak, and J. M. Reid, 2018 The contemporary evolution of fitness. Annual Review of Ecology, Evolution, and Systematics **49**: 457–476.

Hsu, S.-K., W.-Y. Lai, J. Novak, F. Lehner, A. M. Jakšić *et al.*, 2024 Reproductive isolation arises during laboratory adaptation to a novel hot environment. Genome Biology **25**: 141.

Johnston, S. E., 2024 Understanding the genetic basis of variation in meiotic recombination: past, present, and future. Molecular Biology and Evolution **41**: msae112.

Joshi, A., and L. D. Mueller, 1996 Density-dependent natural selection in *Drosophila*: trade-offs between larval food acquisition and utilization. Evolutionary Ecology **10**: 463–474.

Kawecki, T. J., B. Erkosar, C. Dupuis, B. Hollis, R. C. Stillwell *et al.*, 2021 The genomic architecture of adaptation to larval malnutrition points to a trade-off with adult starvation resistance in *Drosophila*. Molecular Biology and Evolution **38**: 2732–2749.

Kelleher, J., A. M. Etheridge, and G. McVean, 2016 Efficient coalescent simulation and genealogical analysis for large sample sizes. PLoS computational biology **12**: e1004842.

Kirkpatrick, M., T. Johnson, and N. Barton, 2002 General models of multilocus evolution. Genetics **161**: 1727–1750.

Kojima, K., 1959 Role of epistasis and overdominance in stability of equilibria with selection. Proceedings of the National Academy of Sciences **45**: 984–989.

Kruuk, L. E., 2004 Estimating genetic parameters in natural populations using the 'animal model'. Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **359**: 873–890.

Kruuk, L. E., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, S. Brotherstone *et al.*, 2000 Heritability of fitness in a wild mammal population. Proceedings of the National Academy of Sciences **97**: 698–703.

Kruuk, L. E., and J. D. Hadfield, 2007 How to separate genetic and environmental causes of similarity between relatives. Journal of evolutionary biology **20**: 1890–1903.

Lee, J. J., and C. C. Chow, 2013 The causal meaning of Fisher's average effect. Genetics research **95**: 89–109.

Lynch, M., and B. Walsh, 1998 *Genetics and analysis of quantitative traits*. Sinauer Associates, Inc., Sunderland.

Martinossi-Allibert, I., U. Savković, M. Đorđević, G. Arnqvist, B. Stojković *et al.*, 2018 The consequences of sexual selection in well-adapted and maladapted populations of bean beetles. Evolution **72**: 518–530.

Price, G. R., 1970 Selection and covariance. Nature **227**: 520–521.

Price, G. R., 1972 Fisher's 'fundamental theorem' made clear. Annals of human genetics **36**: 129–140.

Punzalan, D., M. Delcourt, and H. Rundle, 2014 Comparing the intersex genetic correlation for fitness across novel environments in the fruit fly, *Drosophila serrata*. Heredity **112**: 143–148.

Queller, D. C., 2017 Fundamental theorems of evolution. The American Naturalist **189**: 345–353.

Reznick, D. N., and H. A. Bryga, 1996 Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). V. genetic basis of parallelism in life histories. The American Naturalist **147**: 339–359.

Richards, F. S., 1961 A method of maximum-likelihood estimation. Journal of the Royal Statistical Society Series B: Statistical Methodology **23**: 469–475.

Robertson, A., 1966 A mathematical model of the culling process in dairy cattle. Animal Science **8**: 95–108.

Robinson, C. E., H. Thyagarajan, and A. K. Chippindale, 2023 Evolution of reproductive isolation in a long-term evolution experiment with *Drosophila melanogaster*: 30 years of divergent life-history selection. Evolution **77**: 1756–1768.

Ruzicka, F., M. S. Hill, T. M. Pennell, I. Flis, F. C. Ingleby *et al.*, 2019 Genome-wide sexually antagonistic variants reveal long-standing constraints on sexual dimorphism in fruit flies. PLoS biology **17**: e3000244.

Santiago, E., and A. Caballero, 1998 Effective size and polymorphism of linked neutral loci in populations under directional selection. Genetics **149**: 2105–2117.

Schenkel, M. A., I. Pen, L. W. Beukeboom, and J.-C. Billeter, 2018 Making sense of intralocus and interlocus sexual conflict. Ecology and Evolution **8**: 13035–13050.

Searle, S. R., 2006 *Matrix algebra useful for statistics*. John Wiley & Sons.

Shaw, R., and F. Shaw, 2014 Quantitative genetic study of the adaptive process. Heredity **112**: 13–20.

Simon, A., and G. Coop, 2024 The contribution of gene flow, selection, and genetic drift to five thousand years of human allele frequency change. Proceedings of the National Academy of Sciences **121**: e2312377121.

Singh, A., A. Hasan, and A. F. Agrawal, 2023 An investigation of the sex-specific genetic architecture of fitness in *Drosophila melanogaster*. Evolution **77**: 2015–2028.

Singh, K., E. Kochar, and N. Prasad, 2015 Egg viability, mating frequency and male mating ability evolve in populations of *Drosophila melanogaster* selected for resistance to cold shock. PloS one **10**: e0129992.

Stephan, W., B. Charlesworth, and G. McVean, 1999 The effect of background selection at a single locus on weakly selected, partially linked variants. Genetics Research **73**: 133–146.

Sved, J. A., and W. G. Hill, 2018 One hundred years of linkage disequilibrium. Genetics **209**: 629–636.

Vedder, O., J. Komdeur, M. van der Velde, E. Schut, and M. J. Magrath, 2011 Polygyny and extra-pair paternity enhance the opportunity for sexual selection in blue tits. Behavioral Ecology and Sociobiology **65**: 741–752.

Vehviläinen, H., A. Kause, C. Quinton, H. Koskinen, and T. Paananen, 2008 Survival of the currently fittest: genetics of rainbow trout survival across time and space. Genetics **180**: 507–516.

Walsh, B., 2022 How full is the evolutionary fuel tank? Science **376**: 920–921.

Wang, Y., P. McNeil, R. Abdulazeez, M. Pascual, S. E. Johnston *et al.*, 2023 Variation in mutation, recombination, and transposition rates in *Drosophila melanogaster* and *Drosophila simulans*. Genome Research **33**: 587–598.

Weir, B. S., C. C. Cockerham, and M. Feldman, 1989 Complete characterization of disequilibrium at two loci. Mathematical evolutionary theory **86**: 110.

Whitlock, M. C., 2015 Modern approaches to local adaptation. The American Naturalist **186**: S1–S4.

Wright, S., 1938 Size of population and breeding structure in relation to evolution. Science **87**: 340–431.

Xiao, C., N. Bayat Fard, K. Brzezinski, R. M. Robertson, and A. K. Chippindale, 2019 Experimental evolution of response to anoxia in *Drosophila melanogaster*: recovery of locomotion following CO_2 or N_2 exposure. Journal of Experimental Biology **222**: jeb199521.

Zeng, J., R. De Vlaming, Y. Wu, M. R. Robinson, L. R. Lloyd-Jones *et al.*, 2018 Signatures of negative selection in the genetic architecture of human complex traits. Nature genetics **50**: 746–753.

Symbol	Description
w	Relative fitness
W	Absolute fitness
V_A	Additive genetic variance for relative fitness
V_a	Additive genic variance for relative fitness
t_m	Time at which allele frequencies are first measured in replicate m.
τ_m	Time at which allele frequencies are finally measured in replicate m.
$C_A(t \to \tau)$	Additive genetic covariance for relative fitness between generation t and τ for a population with genetic structure equal to that in generation t.
$C_a(t \to \tau)$	Additive genic covariance for relative fitness between generation t and τ for a population with genetic structure equal to that in generation t.
$\boldsymbol{\alpha}_{t,m}$	Vector of average effects for relative fitness at time t in replicate m.
$\bar{\boldsymbol{\alpha}}$	Vector of mean average effects for relative fitness.
$\Delta\alpha_{t,m}$	Vector of deviations of the average effects for relative fitness at time t in replicate m from the global mean.
$c_{k,i}$	Number of reference alleles at locus i for individual k divided by 2.
n_L	Number of loci.
$N_{t,m}$	Census population size at time t in replicate m
$N_{e_{t,m}}$	Variance effective population size at time t in replicate m
$N_{E_{t,m}}$	Variance effective population size at time t in replicate m ignoring the impact of linked-selection
$\boldsymbol{\eta}_{t,m}$	Linear model coefficients for the c's on $log(W)$
$\mathbf{p}_{t,m}$	Vector of reference allele frequencies at time t in replicate m.
$\mathbf{q}_{t,m}$	Vector of alternate allele frequencies at time t in replicate m.
$\Delta\mathbf{p}_{t,m}$	Vector of reference allele frequency changes between time t and $t+1$ in replicate m.
$\Delta\mathbf{p}_m$	Vector of reference allele frequency changes between time t_m and τ_m in replicate m.
$\mathbf{L}_{t,m}$	Covariance matrix of c's at time t in replicate m.
$\mathbf{L}'_{t,m}$	Covariance matrix of c's at time t in replicate m due to being on the same gametic contribution.
$\mathbf{L}''_{t,m}$	Covariance matrix of c's at time t in replicate m due to being on different gametic contributions.
$\Delta\mathbf{L}'_{t,m}$	The stochastic change in $\mathbf{L}'$ between time zero and t in replicate m.
z_{i_t,j_t}	$(1 - r_{i_t,j_t})(1 - \frac{1}{2N_{e_t}})$
ζ_{i_t,j_t}	$r_{i_t,j_t}(1 - \frac{1}{2N_{e_t}})$

$\tilde{\mathbf{L}}_0$	Weighted sum of $\mathbf{L}_0^{'}$ and $\mathbf{L}_0^{''}$ with weights for element ij being 1 and $z_{i_0,j_0}/\zeta_{i_0,j_0}$ respectively.
$r_{i,j}$	Recombination rate between loci i and j.
$\mathbf{R}_+$	Matrix of recombination probabilities.
$\mathbf{R}_-$	Matrix of non-recombination probabilities.
$\mathbf{N}_{t,m}$	Matrix of weights for $\tilde{\mathbf{L}}_0$ that gives the expected $\mathbf{L}$ at time $t > 0$ in replicate m.
$\mathbf{N}_m$	Matrix of weights for $\tilde{\mathbf{L}}_0$ that gives the sum of the expected $\mathbf{L}$ from time $t_m > 0$ to τ_m in replicate m.
$\boldsymbol{\mathcal{L}}_m$	A matrix that gives, when post-multiplied by $\boldsymbol{\alpha}$, the predictable change in allele frequency due to selection between generation t_m and τ_m in replicate m.
$\mathbf{M}_{t,m}$	Matrix of weights for $\tilde{\mathbf{L}}_0$ that gives the covariance in allele frequency changes due to drift between time $t > 0$ and $t + 1$ in replicate m.
$\mathbf{M}_m$	Matrix of weights for $\tilde{\mathbf{L}}_0$ that gives the covariance in allele frequency changes due to drift between time $t_m > 0$ and τ_m in replicate m.
$\boldsymbol{\mathcal{D}}_m$	Matrix that gives the covariance in allele frequency changes between generation t_m and τ_m in replicate m due to drift.
$\boldsymbol{\mathcal{U}}_m$	Matrix that gives the covariance in allele frequency changes between generation t_m and τ_m in replicate m due to the unpredictable response to selection.
$\boldsymbol{\mu}_{\bar{\alpha}}$	Vector of expected values for the mean average effects.
$\beta_{\bar{\alpha}}^{(0)}$	Intercept of the regression of the mean average effects on some covariate.
$\beta_{\bar{\alpha}}^{(1)}$	Slope of the regression of the mean average effects on some covariate.
$\mathbf{X}$	Design matrix for the regression of the mean average effects on some covariate.
$\mathbf{S}_{\bar{\alpha}}$	Sampling covariance matrix for the parameters of the regression of the mean average effects on some covariate.
$\mathbf{V}_{\bar{\alpha}}$	Covariance matrix for the mean average effects.
$p_{\bar{\alpha}}$	Parameter that takes $\mathbf{L}_0$ to some power
$\sigma_{\bar{\alpha}}^2$	Proportionality constant that relates $\mathbf{V}_{\bar{\alpha}}$ to $\mathbf{L}_0^{p_{\bar{\alpha}}}$.
$\mathbf{P}$	Projection matrix for allele frequencies.
$\mathbf{U}_{\mathbf{L}}$	Eigenvectors of $\mathbf{L}_0$.
$\mathbf{U}_2$	Eigenvectors of $\boldsymbol{\mathcal{D}}$
$\mathbf{D}_2$	Diagonal matrix of square-rooted eigenvalues of $\boldsymbol{\mathcal{D}}$
$\mathbf{B}$	Diagonal matrix of standard deviations for the c's
$\mathbf{R}$	Correlation matrix of the c's.

R_{j_t,i_t}	Correlation in allele count between locus i and j at time t (Note the use of the uppercase to distinguish from the recombination rate r).
$\mathbf{W}_{t\tau}$	A matrix with the ij^{th} element equal to $R_{j_t,i_t}R_{k_\tau,i_\tau}(b_{i_\tau}/b_{i_t})$
$\mathbf{H}_{t\tau}$	$\mathbf{W}_{t\tau}$ but with the off-diagonals set to zero.
$\mathbf{F}_{t\tau}$	$\mathbf{W}_{t\tau}$ but with the diagonals set to zero.
$\mathcal{S}$	Used as a subscript to indicate the set of selected loci.
$\mathcal{N}$	Used as a subscript to indicate the set of neutral loci.
$\phi_{t,\tau}$	The ratio of genetic diversity in generation τ to genetic diversity in generation t assumed constant across all selected loci.

Table 3: Notation

Supplementary data

S1 Appendix 1: The dynamics of L under drift and recombination.

The matrix $\mathbf{L}$ is a covariance matrix whose elements are proportional to the genotypic linkage-disequilbria (off-diagonals) or variance in genotypic allele frequencies (diagonals) at the start of a generation, before selection has acted. We can decompose $\mathbf{L}$ into $\mathbf{L}^{'}$ and $\mathbf{L}^{''}$ following the notation of Buffalo and Coop (2019) where $\mathbf{L}^{'}$ represents the (co)variances that arise due to alleles in the same gamete and $\mathbf{L}^{''}$ represents the (co)variances that arise due to alleles in the different gametic contributions of a genotype. Note that in Santiago and Caballero (1998) the primes have a subtly different meaning after the initial generation, as the double prime in following generations designates gametic phase disequilibrium due to recombination and nongametic phase disequilibrium in the initial generation. The elements of $\mathbf{L}^{'}$ and $\mathbf{L}^{''}$ have direct correspondences with genetic diversities and additive measures of disequilibria. Under the notation of Weir *et al.* (1989) (see also Bulmer *et al.* (1980), Chapter 12), the diagonal elements of $\mathbf{L}^{'}$ are half the gametic genetic diversities ($\pi_i = p_i(1-p_i) = 2L^{'}_{i,i}$), and the off-diagonals are half the gametic-phase disequilibria ($D_{i,j} = 2L^{'}_{i,j}$). Note that Weir *et al.* (1989) uses π to denote $p_i(1-p_i)$ rather than the more usual (and less natural) $2p_i(1-p_i)$. The diagonal elements of $\mathbf{L}^{''}$ are half the additive coefficients of Hardy Weinberg disequilibria ($D_i = 2L^{''}_{i,i}$), and the off-diagonals are half the nongametic-phase disequilibria ($D_{i/j} = 2L^{''}_{i,j}$).

To see these correspondences, imagine two bi-allelic loci, i and j, with reference/alternate alleles A/a and B/b, respectively. There are four possible gametic haplotypes, and we can denote the frequency of haplotype mn as p_{mn} in the gametes and the frequency of allele m as p_m (m and n are indexing variables for allelic states at locus A/a and locus B/b, respectively). The proportion of copies of the reference allele at locus i in a randomly chosen individual – i.e. c_i – can be decomposed into into the sum of maternal and paternal contribution: $c_i = m_i + f_i$ where m_i (or f_i) takes the value 1/2 if the mother (or the father) contributed a reference allele and 0 if not. Then,

$$\begin{aligned} L_{i,j} = & \; COV(c_i, c_j) \\ = & \; COV(m_i + f_i, m_j + f_j) \\ = & \; COV(m_i, m_j) + COV(f_i, f_j) + COV(m_i, f_j) + COV(f_i, m_j) \end{aligned} \tag{S1}$$

Assuming haplotype frequencies are identical in male and female gametes we get

$$\begin{aligned} L_{i,j} = & (p_{AB} - p_A p_B)/4 + (p_{AB} - p_A p_B)/4 + (p_{A/B} - p_A p_B)/4 + (p_{A/B} - p_A p_B)/4 \\ = & (p_{AB} - p_A p_B)/2 + (p_{A/B} - p_A p_B)/2 \\ = & D_{i,j}/2 + D_{i/j}/2 \end{aligned} \tag{S2}$$

where $p_{A/B}$ is the frequency of zygotes that have an A from their mother and a B from their father, or vice versa. When $i = j$,

$$\begin{aligned} L_{i,i} = & COV(c_i, c_i) \\ = & (p_A^2 - p_A)/2 + (p_{A/A} - p_A^2)/2 \\ = & \pi_i/2 + D_i/2 \end{aligned} \tag{S3}$$

where $p_{A/A}$ is the frequency of A homozygotes. The term D_i is an alternative, additive, measure of deviation from Hardy-Weinberg Equilibrium than the more commonly used inbreeding coefficient, F.

To derive expressions for the dynamics of $\mathbf{L}^{'}$ and $\mathbf{L}^{''}$, note that $\mathbf{L}^{''}$ is generated anew each generation under random mating and in a finite population has zero expectation such that

$$L^{''}_{i_{t+1},j_{t+1}} = e^{''}_{i_{t+1},j_{t+1}} \tag{S4}$$

where $e^{''}_{i_{t+1},j_{t+1}}$ is a stochastic term with mean zero and variance given in Weir (1996). The elements of $\mathbf{L}^{'}$ under recombination and drift are (Hill and Robertson, 1968; Santiago and Caballero, 1998)

$$\begin{aligned} L^{'}_{i_{t+1},j_{t+1}} = & \left(1 - \tfrac{1}{2N_{e_t}}\right)\left((1 - r_{i_t,j_t})L^{'}_{i_t,j_t} + r_{i_t,j_t}L^{''}_{i_t,j_t}\right) + e^{'}_{i_{t+1},j_{t+1}} \\ = & z_{i_t,j_t}L^{'}_{i_t,j_t} + \zeta_{i_t,j_t}L^{''}_{i_t,j_t} + e^{'}_{i_{t+1},j_{t+1}} \end{aligned} \tag{S5}$$

where $z_{i_t,j_t} = (1 - r_{i_t,j_t})(1 - \frac{1}{2N_{e_t}})$ and $\zeta_{i_t,j_t} = r_{i_t,j_t}(1 - \frac{1}{2N_{e_t}})$, with r_{i_t,j_t} being the recombination rate between locus i and j in generation t, and N_{e_t} the effective population size in generation t. The stochastic terms, $e^{'}_{i_{t+1},j_{t+1}}$, have zero mean and are uncorrelated over time, with variances given by Ohta and Kimura (1969, Eq 25, although those variances must be divided by 4 here since we are working with the frequency of alleles in individuals rather than gametes). The covariances between, for example, $e^{'}_{i_{t+1},j_{t+1}}$ and $e^{'}_{k_{t+1},l_{t+1}}$ are not given in Ohta and Kimura (1969), and may be unknown. As a consequence,

$$
\begin{aligned}
L'_{i_1,j_1} &= z_{i_0,j_0}L'_{i_0,j_0} + \zeta_{i_0,j_0}L''_{i_0,j_0} + e'_{i_1,j_1} \\
L'_{i_2,j_2} &= z_{i_1,j_1}(z_{i_0,j_0}L'_{i_0,j_0} + \zeta_{i_0,j_0}L''_{i_0,j_0} + e'_{i_1,j_1}) + \zeta_{i_1,j_1}L''_{i_1,j_1} + e'_{i_2,j_2} \\
L'_{i_3,j_3} &= z_{i_2,j_2}(z_{i_1,j_1}(z_{i_0,j_0}L'_{i_0,j_0} + \zeta_{i_0,j_0}L''_{i_0,j_0} + e'_{i_1,j_1}) + \zeta_{i_1,j_1}L''_{i_1,j_1} + e'_{i_2,j_2}) \\
&\quad +\zeta_{i_2,j_2}L''_{i_2,j_2} + e'_{i_3,j_3} \\
L'_{i_3,j_3} &= z_{i_2,j_2}z_{i_1,j_1}z_{i_0,j_0}L'_{i_0,j_0} + z_{i_2,j_2}z_{i_1,j_1}\zeta_{i_0,j_0}L''_{i_0,j_0} + z_{i_2,j_2}z_{i_1,j_1}e'_{i_1,j_1} \\
&\quad +z_{i_2,j_2}\zeta_{i_1,j_1}L''_{i_1,j_1} + z_{i_2,j_2}e'_{i_2,j_2} + \zeta_{i_2,j_2}L''_{i_2,j_2} + e'_{i_3,j_3} \\
\cdots &= \cdots \\
L'_{i_t,j_t} &= L'_{i_0,j_0}\textstyle\prod_{k=0}^{t-1} z_{i_k,j_k} + \sum_{k=0}^{t-1} \zeta_{i_k,j_k}L''_{i_k,j_k}\prod_{u=k+1}^{t-1} z_{i_u,j_u} \\
&\quad + \textstyle\sum_{k=1}^{t-1} e'_{i_k,j_k}\prod_{u=k}^{t-1} z_{i_u,j_u} + e'_{i_t,j_t} \\
L'_{i_t,j_t} &= L'_{i_0,j_0}\textstyle\prod_{k=0}^{t-1} z_{i_k,j_k} + \zeta_{i_0,j_0}L''_{i_0,j_0}\prod_{k=1}^{t-1} z_{i_k,j_k} \\
&\quad + \textstyle\sum_{k=1}^{t-1} \zeta_{i_k,j_k}L''_{i_k,j_k}\prod_{u=k+1}^{t-1} z_{i_u,j_u} + \sum_{k=1}^{t-1} e'_{i_k,j_k}\prod_{u=k}^{t-1} z_{i_u,j_u} + e'_{i_t,j_t} \\
L'_{i_t,j_t} &= L'_{i_0,j_0}\textstyle\prod_{k=0}^{t-1} z_{i_k,j_k} + \frac{\zeta_{i_0,j_0}}{z_{i_0,j_0}}L''_{i_0,j_0}\prod_{k=0}^{t-1} z_{i_k,j_k} + \sum_{k=1}^{t-1} \frac{\zeta_{i_k,j_k}}{z_{i_k,j_k}}L''_{i_k,j_k}\prod_{u=k}^{t-1} z_{i_u,j_u} \\
&\quad + \textstyle\sum_{k=1}^{t-1} e'_{i_k,j_k}\prod_{u=k}^{t-1} z_{i_u,j_u} + e'_{i_t,j_t} \\
L'_{i_t,j_t} &= \left(L'_{i_0,j_0} + \frac{\zeta_{i_0,j_0}}{z_{i_0,j_0}}L''_{i_0,j_0}\right)\textstyle\prod_{k=0}^{t-1} z_{i_k,j_k} \\
&\quad + \textstyle\sum_{k=1}^{t-1} \left(e'_{i_k,j_k} + \frac{\zeta_{i_k,j_k}}{z_{i_k,j_k}}L''_{i_k,j_k}\right)\prod_{u=k}^{t-1} z_{i_u,j_u} + e'_{i_t,j_t}
\end{aligned}
\tag{S6}
$$

Having the matrix $\mathbf{N}_t$ with the ij^{th} element being the relevant $\prod_{k=0}^{t-1} z_k$, we have for $t > 0$

$$
\mathbf{L}_t = \mathbf{N}_t \circ \tilde{\mathbf{L}}_0 + \Delta\mathbf{L}'_t + \mathbf{L}''_t \tag{S7}
$$

where $\circ$ is the Hadamard product. The first term is the expected $\mathbf{L}$ (and $\mathbf{L}'$) in generation t, conditional on the genotypic composition of the population in generation 0, where $\tilde{\mathbf{L}}_0$ is the sum of $\mathbf{L}'_0$ and $\mathbf{L}''_0$ with the elements of the latter weighted by $\zeta_{i_0,j_0}/z_{i_0,j_0} = r_{i_0,j_0}/(1 - r_{i_0,j_0})$. $\Delta\mathbf{L}'_t$ is a matrix with elements equal to the sum of the stochastic terms involving e in Equation S6 and represents stochastic changes

in $\mathbf{L}^{'}$ from generation 0 to generation t. $\mathbf{L}_t^{''}$ are the new nongametic-phase disequilibria that arise in generation t. Note that if replicate populations are initiated from the offspring of Generation 0, then the stochastic terms, $\Delta\mathbf{L}_t^{'}$ and $\mathbf{L}_t^{''}$, will be unique in each replicate although the deterministic part is shared. If recombination rates are constant in time $r_{i_t,j_t} = r_{i,j}$ then $\zeta_{i_0,j_0}/z_{i_0,j_0} = r_{i,j}/(1-r_{i,j})$ and $\prod_{k=0}^{t-1} z_{i_k,j_k} = (1-r_{i,j})^t \prod_{k=0}^{t-1}(1-\frac{1}{2N_{e_k}})$. When population sizes are also constant then $\prod_{k=0}^{t-1} z_{i_k,j_k} = (1-r_{i,j})^t(1-\frac{1}{2N_e})^t$.

S2 Appendix 2: Derivation for the mean, within-replicate (co)variances, and between replicate (co)variances of allele frequency changes.

Both $\underset{D}{\Delta}\mathbf{p}$ and $\underset{U}{\Delta}\mathbf{p}$ have expectation zero such that the conditional mean is

$$\begin{aligned} E\left[\Delta\mathbf{p}_m\right] = & \textstyle\sum_{t=t_m}^{\tau_m-1} E\left[(\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0)\bar{\boldsymbol{\alpha}}\right] \\ = & \left(\textstyle\sum_{t=t_m}^{\tau_m-1}(\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0)\right) E\left[\bar{\boldsymbol{\alpha}}\right] \\ = & \left(\tilde{\mathbf{L}}_0\circ\textstyle\sum_{t=t_m}^{\tau_m-1}\mathbf{N}_{t,m}\right) E\left[\bar{\boldsymbol{\alpha}}\right] \\ = & \left(\tilde{\mathbf{L}}_0\circ\mathbf{N}^{(m)}\right) E\left[\bar{\boldsymbol{\alpha}}\right] \end{aligned} \tag{S8}$$

where $\mathbf{N}^{(m)} = \sum_{t=t_m}^{\tau_m-1}\mathbf{N}_{t,m}$ which is null if $t_m = 0$ and $\tau_m = 1$. Under the same assumptions, the conditional between-replicate covariance has a similar form

$$COV(\Delta\mathbf{p}_m, \Delta\mathbf{p}_n^\top) = \left(\tilde{\mathbf{L}}_0\circ\mathbf{N}^{(m)}\right) VAR(\bar{\boldsymbol{\alpha}}) \left(\tilde{\mathbf{L}}_0\circ\mathbf{N}^{(n)}\right) \tag{S9}$$

where m and n are a pair of replicates. The within-replicate (co)variances are more challenging to derive as they not only include the predictable response to selection, but also the effects of drift and the unpredictable response to selection. In what follows, we assume $E\left[\Delta\boldsymbol{\alpha}\right] = 0$ and $COV(\Delta\mathbf{L}^{'} + \mathbf{L}^{''}, \bar{\boldsymbol{\alpha}}) = 0$ such that the predictable and unpredictable response to selection are independent. In addition, we will assume $COV(\Delta\mathbf{L}^{'} + \mathbf{L}^{''}, \Delta\boldsymbol{\alpha}) = 0$ and that all $\Delta\boldsymbol{\alpha}$ are independent of each other. Since $\Delta\mathbf{L}_t^{'}$ is the sum of stochastic changes from generation 1 to t, $COV(\Delta\mathbf{L}_{t_1,m}^{'}, \Delta\mathbf{L}_{t_2,m}^{'}) = VAR(\Delta\mathbf{L}_{t_1,m}^{'})$ when $t_1 \leq t_2$. Note $COV(\mathbf{L}_{t_1,m}^{''}, \mathbf{L}_{t_2,m}^{''}) = 0$. These assumptions allow the simplification:

$$
\begin{aligned}
VAR(\Delta\mathbf{p}_m) = &\quad VAR\left(\sum_{t=t_m}^{\tau_m-1}\left((\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0)\bar{\boldsymbol{\alpha}} + \underset{U}{\Delta}\mathbf{p}_{t,m} + \underset{D}{\Delta}\mathbf{p}_{t,m}\right)\right)\\
VAR(\Delta\mathbf{p}_m) = &\quad \sum_{t_1=t_m}^{\tau_m-1}\sum_{t_2=t_m}^{\tau_m-1} COV\left(((\mathbf{N}_{t_1}\circ\tilde{\mathbf{L}}_0)\bar{\boldsymbol{\alpha}} + \underset{U}{\Delta}\mathbf{p}_{t_1,m} + \underset{D}{\Delta}\mathbf{p}_{t_1,m}, ((\mathbf{N}_{t_2}\circ\tilde{\mathbf{L}}_0))\bar{\boldsymbol{\alpha}} + \underset{U}{\Delta}\mathbf{p}_{t_2,m} + \underset{D}{\Delta}\mathbf{p}_{t_2,m}\right)\\
VAR(\Delta\mathbf{p}_m) = &\quad \sum_{t=t_m}^{\tau_m-1} VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right) + \sum_{t_1=t_m}^{\tau_m-1}\sum_{t_2=t_m}^{\tau_m-1} COV\left(((\mathbf{N}_{t_1}\circ\tilde{\mathbf{L}}_0))\bar{\boldsymbol{\alpha}} + \underset{U}{\Delta}\mathbf{p}_{t_1,m}, ((\mathbf{N}_{t_2}\circ\tilde{\mathbf{L}}_0))\bar{\boldsymbol{\alpha}} + \underset{U}{\Delta}\mathbf{p}_{t_2,m}\right)\\
VAR(\Delta\mathbf{p}_m) = &\quad \sum_{t=t_m}^{\tau_m-1} VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right) + \left(\sum_{t_1=t_m}^{\tau_m-1}((\mathbf{N}_{t_1}\circ\tilde{\mathbf{L}}_0))\right) VAR(\bar{\boldsymbol{\alpha}}) \left(\sum_{t_2=t_m}^{\tau_m-1}((\mathbf{N}_{t_2}\circ\tilde{\mathbf{L}}_0))\right)\\
&\quad + \sum_{t_1=t_m}^{\tau_m-1}\sum_{t_2=t_m}^{\tau_m-1} COV\left(\underset{U}{\Delta}\mathbf{p}_{t_1,m}, \underset{U}{\Delta}\mathbf{p}_{t_2,m}\right)\\
VAR(\Delta\mathbf{p}_m) = &\quad \sum_{t=t_m}^{\tau_m-1} VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right) + \left(\sum_{t_1=t_m}^{\tau_m-1}((\mathbf{N}_{t_1}\circ\tilde{\mathbf{L}}_0))\right) VAR(\bar{\boldsymbol{\alpha}}) \left(\sum_{t_2=t_m}^{\tau_m-1}((\mathbf{N}_{t_2}\circ\tilde{\mathbf{L}}_0))\right)\\
&\quad + \sum_{t=t_m}^{\tau_m-1}(2(\tau^m - t)+1) VAR\left(\underset{U}{\Delta}\mathbf{p}_{t,m}\right)\\
VAR(\Delta\mathbf{p}_m) = &\quad \sum_{t=t_m}^{\tau_m-1} VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right) + \left(\tilde{\mathbf{L}}_0\circ\mathbf{N}^{(m)}\right) VAR(\bar{\boldsymbol{\alpha}}) \left(\tilde{\mathbf{L}}_0\circ\mathbf{N}^{(m)}\right)\\
&\quad + \sum_{t=t_m}^{\tau_m-1}(2(\tau^m - t)+1) VAR\left(\underset{U}{\Delta}\mathbf{p}_{t,m}\right)
\end{aligned}
\tag{S10}
$$

which requires an expressions for $VAR\left(\underset{U}{\Delta}\mathbf{p}_{t,m}\right)$ and $VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right)$. The variance due to unpredictable selection is:

$$
\begin{aligned}
VAR\left(\underset{U}{\Delta}\mathbf{p}_{t,m}\right) &= VAR\left((\Delta\mathbf{L}'_{t,m} + \mathbf{L}''_{t,m})\bar{\boldsymbol{\alpha}} + ((\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0))\Delta\boldsymbol{\alpha}_{t,m} + (\Delta\mathbf{L}'_{t,m} + \mathbf{L}''_{t,m})\Delta\boldsymbol{\alpha}_{t,m}\right)\\
&= VAR\left((\Delta\mathbf{L}'_{t,m} + \mathbf{L}''_{t,m})\bar{\boldsymbol{\alpha}}\right) + VAR\left(((\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0)\Delta\boldsymbol{\alpha}_{t,m}\right) + VAR\left((\Delta\mathbf{L}'_{t,m} + \mathbf{L}''_{t,m})\Delta\boldsymbol{\alpha}_{t,m}\right)\\
&= VAR\left(\Delta\mathbf{L}'_{t,m} + \mathbf{L}''_{t,m}\right) VAR\left(\bar{\boldsymbol{\alpha}}\right)\\
&\quad + ((\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0) VAR\left(\Delta\boldsymbol{\alpha}_{t,m}\right) ((\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0) VAR\left(\Delta\mathbf{L}'_{t,m} + \mathbf{L}''_{t,m}\right) VAR\left(\Delta\boldsymbol{\alpha}_{t,m}\right)
\end{aligned}
\tag{S11}
$$

and cannot be simplified.

For the drift covariances, note that under random mating, haplotypes are drawn from a multinomial with $2N$ trials. Using AB, Ab, AB and aB to denote the *number* of each haplotypes in the gamete pool, the number of mn haplotypes has variance $2Np_{mn}(1-p_{mn})$ and the covariance in the numbers of mn and op haplotypes is $-2Np_{mn}p_{op}$ where p_{mn} is the frequency of the mn halpotype in the gamete pool (i.e the parental haplotype frequencies modified by recombination). The covariance in the number of A and B alleles sampled, is therefore

$$
\begin{aligned}
COV(AB+Ab, AB+aB) &= COV(AB,AB) + COV(AB,aB)\\
&\quad + COV(Ab,AB) + COV(Ab,aB)\\
&= 2N\left[p_{AB}(1-p_{AB}) - p_{AB}p_{aB} - p_{Ab}p_{AB} - p_{Ab}p_{aB}\right]\\
&= 2N\left[p_{AB}(1-p_{AB}-p_{aB}) - p_{Ab}(p_{AB}+p_{aB})\right]\\
&= 2N\left[p_{AB}(1-p_B) - p_{Ab}p_B\right]\\
&= 2N\left[p_{AB} - (p_{Ab}+p_{AB})p_B\right]\\
&= 2N\left[p_{AB} - p_A p_B\right]\\
&= 4N\bar{L}'_{i,j}
\end{aligned}
\tag{S12}
$$

where $\bar{L}^{'}_{i,j}$ is the gametic-phase linkage disequilibria that would be achieved in an infinite population. We can divide through by $(1/2N)^2$ to obtain the drift covariance in frequency (rather than counts) as $\bar{L}^{'}_{i,j}/N$ (i.e the drift (co)variances in a allele frequency from generation t to $t+1$ are proportional to the gametic-phase disequilibria in generation $t+1$ that would be achieved in an infinite population conditional on the genotypic composition of the population in generation t). This recovers the well known result for the drift variance in allele frequency: $\bar{L}^{'}_{i,i}/N = p_i(1-p_i)/2N$. We can replace the census population size, N, with the effective population size N_E, since this approximates the sampling of genotypes in non-idealised populations well (Ethier and Nagylaki, 1980). However, note that N_E differs from N_e in that it does include the impact of linked selection since this is conditioned on in the expectation, $E[\mathbf{p}]$. In matrix terms (since $\bar{L}^{'}_{i_{t+1},j_{t+1}} = L^{'}_{i_t,j_t} r_{i,j} + L^{''}_{i_t,j}(1-r_{i,j})$):

$$\begin{aligned} VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right) &= \tfrac{1}{N_{E_{t,m}}}\bar{\mathbf{L}}^{'}_{t+1,m} \\ VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right) &= \tfrac{1}{N_{E_{t,m}}}\left[\mathbf{R}_{-}\circ\mathbf{L}^{'}_{t,m} + \mathbf{R}_{+}\circ\mathbf{L}^{''}_{t,m}\right] \\ VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right) &= \tfrac{1}{N_{E_{t,m}}}\left[\left(\mathbf{R}_{-}\circ\left(\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0 + \Delta\mathbf{L}^{'}_{t,m}\right) + \mathbf{R}_{+}\circ\mathbf{L}^{''}_{t,m}\right)\right] \end{aligned} \tag{S13}$$

where $\mathbf{R}_{+}$ and $\mathbf{R}_{-}$ are matrices with the ij_{th} being $r_{i,j}$ and $1-r_{i,j}$ respectively. The expected drift terms (conditional on $\mathbf{L}_0$) are therefore:

$$E\left[VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right)\right] = \tfrac{1}{N_{E_{t,m}}}\left(\mathbf{R}_{-}\circ\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0\right) \tag{S14}$$

We define a new matrix $\mathbf{M}_{t,m}$ with the ij^{th} element being $(1-r_{i,j})/N_{E_{t,m}}$ ($\mathbf{M}_{t,m} = \frac{1}{N_{E_{t,m}}}\mathbf{R}_{-}$) to give

$$E\left[VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right)\right] = \mathbf{M}_{t,m}\circ\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0 \tag{S15}$$

Since the drift terms are independent, this gives:

$$\begin{aligned} E\left[VAR\left(\underset{D}{\Delta}\mathbf{p}_{m}\right)\right] &= \textstyle\sum_{t=t_m}^{\tau_m-1} E\left[VAR\left(\underset{D}{\Delta}\mathbf{p}_{t,m}\right)\right] \\ &= \textstyle\sum_{t=t_m}^{\tau_m-1} \tfrac{1}{N_{E_{t,m}}}\left(\mathbf{R}_{-}\circ\mathbf{N}_{t,m}\circ\tilde{\mathbf{L}}_0\right) \\ &= \tilde{\mathbf{L}}_0\circ\textstyle\sum_{t=t_m}^{\tau_m-1}\mathbf{M}_{t,m}\circ\mathbf{N}_{t,m} \\ &= \tilde{\mathbf{L}}_0\circ\mathbf{M}^{(m)} \end{aligned} \tag{S16}$$

where $\mathbf{M}^{(m)} = \sum_{t=t_m}^{\tau_m-1}\mathbf{M}_{t,m}\circ\mathbf{N}_{t,m}$ and is null when $t_m = 0$ and $\tau_m = 1$. Note the ij^{th} element of $\mathbf{M}_{t,m}\circ\mathbf{N}_{t,m} = (1-r_{i,j})^{t+1}\frac{1}{N_{E_t}}\prod_{k=0}^{t-1}(1-\frac{1}{2N_{e_{k,m}}})$ which reduces

to $(1-r_{i,j})^{t+1}\frac{1}{N_{e_m}}(1-\frac{1}{2N_{e_m}})^t$ with constant population size and $N_e = N_E$. It is not clear how large the evolutionary variance in $VAR\left(\underset{D}{\Delta}\mathbf{p}_m\right)$ is, and whether this would need to be accommodated.

S3 Appendix 3: Treating $\boldsymbol{\alpha}$ as random

Before we consider $\boldsymbol{\alpha}$ as random rather than fixed, it is important to understand that quantities such as V_A and genomic best linear unbiased predictors (gBLUP) are insensitive to which allele at a locus we consider the reference allele and which allele we consider the alternate allele. To make this explicit, consider the diagonal 'assignment' matrix $\mathbf{A}$ for which the diagonal elements are either 1 (the fittest allele is the reference allele) or -1 (the fittest allele is the alternate allele). Under a particular assignment, $\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\alpha}_+$ and $\mathbf{L} = \mathbf{A}\mathbf{L}_+\mathbf{A}$, where the subscript + indicates the quantity had the fitter of the two alleles been the reference allele at all loci. If we consider V_A conditional on a particular assignment we get (since $\mathbf{A}\mathbf{A} = \mathbf{I}$):

$$\begin{aligned} V_A &= \boldsymbol{\alpha}^\top \mathbf{L}\boldsymbol{\alpha} \\ &= (\mathbf{A}\boldsymbol{\alpha}_+)^\top \mathbf{A}\mathbf{L}_+\mathbf{A}\mathbf{A}\boldsymbol{\alpha}_+ \\ &= \boldsymbol{\alpha}_+^\top \mathbf{A}\mathbf{A}\mathbf{L}_+\mathbf{A}\mathbf{A}\boldsymbol{\alpha}_+ \\ &= \boldsymbol{\alpha}_+^\top \mathbf{L}_+\boldsymbol{\alpha}_+ \end{aligned} \tag{S17}$$

showing we get the same value of V_A irrespective of the assignment we choose. In contrast, quantities such as $E[\mathbf{L}\boldsymbol{\alpha}]$ are sensitive to the assignment, since they undergo a sign reversal under a different choice:

$$\begin{aligned} \mathbf{L}\boldsymbol{\alpha} &= \mathbf{A}\mathbf{L}_+\mathbf{A}\mathbf{A}\boldsymbol{\alpha}_+ \\ &= \mathbf{A}\mathbf{L}_+\boldsymbol{\alpha}_+ \end{aligned} \tag{S18}$$

It is tempting to use the argument that $E[\mathbf{L}\boldsymbol{\alpha}] = \mathbf{0}$ when the reference allele is chosen at random, since α is equally likely to be positive as negative. The logic behind this argument can be expressed mathematically as $E[\boldsymbol{\alpha}] = E[\mathbf{A}]E[\boldsymbol{\alpha}_+]$ since the reference allele is chosen at random and so $\mathbf{A}$ must be independent of $\boldsymbol{\alpha}_+$. Under this same assumption $E[\mathbf{A}] = \mathbf{0}$, since any diagonal element has an equal chance of being -1 or 1, such that $E[\mathbf{L}\boldsymbol{\alpha}] = 0$. However, this logic is incorrect. The argument envisages $\mathbf{A}$ as random, yet for any particular analysis $\mathbf{A}$ is no longer a random variable but fixed - a choice has been made as to which allele is the reference allele - even if there remains epistemic uncertainty as to whether the reference allele is the fitter of the two alleles.

In our inference section we show that

$$\begin{aligned} E_{\mathbf{L}_0}(V_A) &= E_{\mathbf{L}_0}[\bar{\boldsymbol{\alpha}}^\top \mathbf{L}_0 \bar{\boldsymbol{\alpha}}] \\ &= Tr(\mathbf{L}_0 \mathbf{V}_{\bar{\alpha}}) + \boldsymbol{\mu}_{\bar{\alpha}}^\top \mathbf{L}_0 \boldsymbol{\mu}_{\bar{\alpha}} \end{aligned} \tag{S19}$$

where $\boldsymbol{\mu}_{\bar{\alpha}}$ and $\mathbf{V}_{\bar{\alpha}}$ are the mean vector and covariance matrix of the expected average effects. When the reference allele is chosen arbitrarily, any sensible distribution for the $\bar{\alpha}$'s must induce the same distribution on the $\bar{\alpha}_+$'s regardless of the assignment. If $\boldsymbol{\mu}_{\bar{\alpha}_+}$ and $\mathbf{V}_{\bar{\alpha}_+}$ are the means and (co)variances of the expected average effects had all reference alleles been the fitter allele, then the distribution for a particular assignment becomes $\boldsymbol{\mu}_{\bar{\alpha}} = \mathbf{A}\boldsymbol{\mu}_{\bar{\alpha}_+}$ and $\mathbf{V}_{\bar{\alpha}} = \mathbf{A}\mathbf{V}_{\bar{\alpha}_+}\mathbf{A}$. Given $\mathbf{A}^{-1} = \mathbf{A}$ this implies $\boldsymbol{\mu}_{\bar{\alpha}_+} = \mathbf{A}\boldsymbol{\mu}_{\bar{\alpha}}$ and $\mathbf{V}_{\bar{\alpha}_+} = \mathbf{A}\mathbf{V}_{\bar{\alpha}}\mathbf{A}$. For, $\boldsymbol{\mu}_{\alpha}$ this implies that suitable models should be (weighted) sums of differences between invariant properties of the alleles such that the difference reverses sign when the reference and alternate allele are switched. This might be their (log) frequency, such that the model is $\beta(p-q)$, with β a parameter, or it might be the difference in derived vs ancestral coded as 1 vs -1, such that the model is 2β or -2β depending on whether the reference allele is derived or ancestral, respectively.

If $\mathbf{V}_{\bar{\alpha}}$ is assumed to be diagonal, all models are permissible since the square removes any sign. Since the multiplication of diagonal matrices is not affected by order we can see this directly:

$$\begin{aligned} \mathbf{V}_{\bar{\alpha}_+} &= \mathbf{A}\mathbf{V}_{\bar{\alpha}}\mathbf{A} \\ \mathbf{V}_{\bar{\alpha}_+} &= \mathbf{A}\mathbf{A}\mathbf{V}_{\bar{\alpha}} \\ \mathbf{V}_{\bar{\alpha}_+} &= \mathbf{V}_{\bar{\alpha}} \end{aligned} \tag{S20}$$

However, for non-diagonal matrices a suitable distribution must result in a sign reversal of all covariances at a locus when the reference and alternate alleles are switched. The most obvious way to achieve this is to allow $\mathbf{V}_{\bar{\alpha}}$ to be proportional to $\mathbf{L}^p$ since

$$\begin{aligned} \mathbf{L}^p &= (\mathbf{A}\mathbf{L}_+\mathbf{A})^p \\ &= \mathbf{A}\mathbf{L}_+\mathbf{A}...\mathbf{A}\mathbf{L}_+\mathbf{A} \\ &= \mathbf{A}\mathbf{L}_+^p\mathbf{A} \end{aligned} \tag{S21}$$

where $\mathbf{L}_+$ is the linkage-disequilibrium matrix had the fittest allele been the reference allele at all loci. Under this assumption

$$\begin{aligned} \mathbf{V}_{\bar{\alpha}_+} &= \mathbf{A}\mathbf{V}_{\bar{\alpha}}\mathbf{A} \\ \mathbf{V}_{\bar{\alpha}_+} &\propto \mathbf{A}\mathbf{L}^p\mathbf{A} \\ \mathbf{V}_{\bar{\alpha}_+} &\propto \mathbf{A}\mathbf{A}\mathbf{L}_+^p\mathbf{A}\mathbf{A} \\ \mathbf{V}_{\bar{\alpha}_+} &\propto \mathbf{L}_+^p \end{aligned} \tag{S22}$$

A similar model was proposed by Zeng *et al.* (2018), although there, $\mathbf{L}$ was treated as diagonal such that the variance of the average effects are assumed to be proportional to the genetic diversities to some power.

S4 Appendix 4: Projection matrices

To show that our chosen projection ($\mathbf{P} = \mathbf{D}_2^{-1}\mathbf{U}_2^\top\mathbf{U}_\mathbf{L}^\top$) results in identical and independently distributed residuals, we need to show that the drift covariance matrix is an identity matrix under this projection. First, we note that both $\mathbf{U}_\mathbf{L}$ and $\mathbf{U}_2$ are (semi-)unitary such that $\mathbf{U}\mathbf{U}^\top = \mathbf{U}^\top\mathbf{U} = \mathbf{I}$. Next we write down the eigendecomposition of $\mathbf{U}_\mathbf{L}^\top(\tilde{\mathbf{L}}_0 \circ \mathbf{M}^{(m)})\mathbf{U}_\mathbf{L} = \mathbf{U}_2\mathbf{D}_2\mathbf{D}_2\mathbf{U}_2^\top$, and use the unitary property of $\mathbf{U}_\mathbf{L}$ to note that $\tilde{\mathbf{L}}_0 \circ \mathbf{M}^{(m)}$ can be expressed as $\mathbf{U}_\mathbf{L}\mathbf{U}_2\mathbf{D}_2\mathbf{D}_2\mathbf{U}_2^\top\mathbf{U}_\mathbf{L}^\top$. The drift covariance matrix under the projection is then:

$$\begin{aligned} VAR(\underset{D}{\Delta}\overrightarrow{\mathbf{p}}_m) &= \mathbf{P}VAR(\underset{D}{\Delta}\mathbf{p}_m)\mathbf{P}^\top \\ &= \mathbf{P}(\tilde{\mathbf{L}}_0 \circ \mathbf{M}^{(m)})\mathbf{P}^\top \\ &= \mathbf{P}\mathbf{U}_\mathbf{L}\mathbf{U}_2\mathbf{D}_2\mathbf{D}_2\mathbf{U}_2^\top\mathbf{U}_\mathbf{L}^\top\mathbf{P}^\top \\ &= \mathbf{D}_2^{-1}\mathbf{U}_2^\top\mathbf{U}_\mathbf{L}^\top\mathbf{U}_\mathbf{L}\mathbf{U}_2\mathbf{D}_2\mathbf{D}_2\mathbf{U}_2^\top\mathbf{U}_\mathbf{L}^\top\mathbf{U}_\mathbf{L}\mathbf{U}_2\mathbf{D}_2^{-1} \\ &= \mathbf{I} \end{aligned} \tag{S23}$$

S5 Appendix 5: Bias correction for $\boldsymbol{\mu}_{\bar{\alpha}}^\top\mathbf{L}_0\boldsymbol{\mu}_{\bar{\alpha}}$

Using $\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}^\top\mathbf{L}_0\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}$ as an estimator of $\boldsymbol{\mu}_{\bar{\alpha}}^\top\mathbf{L}_0\boldsymbol{\mu}_{\bar{\alpha}}$ results in upward bias. To understand this, and correct for it, have

$$\widehat{\boldsymbol{\beta}_{\bar{\alpha}}} = \boldsymbol{\beta}_{\bar{\alpha}} + \mathbf{m}_{\bar{\alpha}} \tag{S24}$$

where $\mathbf{m}_{\bar{\alpha}}$ is the vector of deviations of the estimates from their true value. The expected estimate of $\boldsymbol{\mu}_{\bar{\alpha}}^\top\mathbf{L}_0\boldsymbol{\mu}_{\bar{\alpha}}$ is then

$$\begin{aligned} E[\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}^\top\mathbf{L}_0\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}] = &\ \boldsymbol{\beta}_{\bar{\alpha}}^\top\mathbf{X}^\top\mathbf{L}_0\mathbf{X}\boldsymbol{\beta}_{\bar{\alpha}} + E[\mathbf{m}_{\bar{\alpha}}^\top\mathbf{X}^\top\mathbf{L}_0\mathbf{X}\mathbf{m}_{\bar{\alpha}}] \\ = &\ \boldsymbol{\mu}_{\bar{\alpha}}^\top\mathbf{L}_0\boldsymbol{\mu}_{\bar{\alpha}} + E[\mathbf{m}_{\bar{\alpha}}^\top\mathbf{X}^\top\mathbf{L}_0\mathbf{X}\mathbf{m}_{\bar{\alpha}}] \end{aligned} \tag{S25}$$

where $\mathbf{X}$ is the design matrix with the first column all ones, and the second $\mathbf{p}_0 - \mathbf{q}_0$. Here, the expectation is taken over the distribution of estimates, and it is assumed the estimates are unbiased (since then, $E[\mathbf{m}_{\bar{\alpha}}] = \mathbf{0}$, and so $E[\mathbf{m}_{\bar{\alpha}}\boldsymbol{\beta}_{\bar{\alpha}}^\top] = \mathbf{0}$). This same assumption implies

$$E[\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}^\top\mathbf{L}_0\widehat{\boldsymbol{\mu}_{\bar{\alpha}}}] = \ \boldsymbol{\mu}_{\bar{\alpha}}^\top\mathbf{L}_0\boldsymbol{\mu}_{\bar{\alpha}} + Tr\left(\mathbf{X}^\top\mathbf{L}_0\mathbf{X}E[\mathbf{m}_{\bar{\alpha}}^\top\mathbf{m}_{\bar{\alpha}}]\right) \tag{S26}$$

Since $E[\mathbf{m}_{\bar{\alpha}}^\top \mathbf{m}_{\bar{\alpha}}]$ is a matrix of sampling (co)variances for the parameters, to get an improved estimate we can use the inverse Hessian to get an approximate $\mathbf{S}_{\bar{\alpha}}$:

$$\widehat{\boldsymbol{\mu}_{\bar{\alpha}}^\top \mathbf{L}_0 \boldsymbol{\mu}_{\bar{\alpha}}} = \widehat{\boldsymbol{\mu}_{\bar{\alpha}}}^\top \mathbf{L}_0 \widehat{\boldsymbol{\mu}_{\bar{\alpha}}} - Tr\left(\mathbf{X}^\top \mathbf{L}_0 \mathbf{X} \mathbf{S}_{\bar{\alpha}}\right) \tag{S27}$$

In addition, $Tr(\mathbf{L}_0 \widehat{\mathbf{V}_{\bar{\alpha}}})$ is most likely an upwardly biased estimator of $Tr(\mathbf{L}_0 \mathbf{V}_{\bar{\alpha}})$. However, we found no easy way to determine or correct for the degree of bias, and the simulation results suggest that the bias is likely to be small, at least when the number of replicates or generations is moderately large.

S6 Appendix 6: Comparison with the method of Buffalo and Coop (2019)

To make clearer the distinction between the approach of B&C and our approach, here we explicitly express the expectations and covariances appearing in B&C as conditional on $\mathbf{B}$ and $\boldsymbol{\alpha}$. In the sections dealing with our theory and inference, the conditioning (on $\mathbf{L}_0$) is left implicit. B&C work with the quantity

$$COV(\Delta \mathbf{p}_t, \Delta \mathbf{p}_\tau^\top | \mathbf{B}, \boldsymbol{\alpha}) \tag{S28}$$

where τ is some generation after t (B&C actually only work with the diagonal elements of this matrix, but we retain the full multi-locus model here for generality). Importantly, when B&C estimate the additive genetic variance in fitness they assume that both $\Delta \mathbf{p}_t$ and $\Delta \mathbf{p}_\tau$ both represent allele frequency change over a single generation (Assumption A). As the change due to drift will be independent in different generations, we can rewrite B&C's covariance:

$$\begin{aligned} COV(\Delta \mathbf{p}_t, \Delta \mathbf{p}_\tau^\top | \mathbf{B}, \boldsymbol{\alpha}) = & \; COV(\mathbf{L}_t \boldsymbol{\alpha}_t, \boldsymbol{\alpha}_\tau^\top \mathbf{L}_\tau | \mathbf{B}, \boldsymbol{\alpha}) \\ = & \; E\left[\mathbf{L}_t \boldsymbol{\alpha}_t \boldsymbol{\alpha}_\tau^\top \mathbf{L}_\tau | \mathbf{B}, \boldsymbol{\alpha}\right] - E\left[\mathbf{L}_t | \mathbf{B}, \boldsymbol{\alpha}\right] \boldsymbol{\alpha}_t \boldsymbol{\alpha}_\tau^\top E\left[\mathbf{L}_\tau | \mathbf{B}, \boldsymbol{\alpha}\right] \end{aligned} \tag{S29}$$

Since the diagonal elements of $\mathbf{L}$ have to be positive, a sufficient, but not necessary, condition for the final term to be zero is that there is no direct selection on the loci (i.e. $\boldsymbol{\alpha} = \mathbf{0}$). It is not a necessary condition because the change caused by direct selection at all loci could be exactly balanced by the change caused by indirect selection at other loci, although we ignore this unlikely scenario. The assumption that $\boldsymbol{\alpha} = \mathbf{0}$ is achieved in B&C by assuming that sites can be partitioned into neutral and selected sites and that allele frequency change is only tracked at the neutral sites

(Assumption B). To understand the consequences of this assumption we consider all loci are being followed, both a selected set ($\mathcal{S}$) and a neutral set ($\mathcal{N}$). Consequently,

$$\boldsymbol{\alpha} = \begin{bmatrix} \mathbf{0} \\ \boldsymbol{\alpha}_{\mathcal{S}} \end{bmatrix} \tag{S30}$$

and so

$$\boldsymbol{\alpha}\boldsymbol{\alpha}^{\top} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\alpha}_{\mathcal{S}}\boldsymbol{\alpha}_{\mathcal{S}}^{\top} \end{bmatrix} \tag{S31}$$

We can also partition $\mathbf{L}$

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_{\mathcal{N}} & \mathbf{L}_{\mathcal{N},\mathcal{S}} \\ \mathbf{L}_{\mathcal{S},\mathcal{N}} & \mathbf{L}_{\mathcal{S}} \end{bmatrix} \tag{S32}$$

and writing $\mathbf{L}_{\mathcal{N},\mathcal{S}} = \mathbf{B}_{\mathcal{N}}\mathbf{R}_{\mathcal{N},\mathcal{S}}\mathbf{B}_{\mathcal{S}}$, Equation S29 for neutral sites becomes

$$\begin{aligned} COV(\Delta\mathbf{p}_{\mathcal{N}_t}, \Delta\mathbf{p}_{\mathcal{N}_\tau}^{\top}|\mathbf{B},\boldsymbol{\alpha}) = \; & \mathbf{B}_{\mathcal{N}_t} E\left[\mathbf{R}_{\mathcal{N}_t,\mathcal{S}_t}\mathbf{B}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_\tau}^{\top}\mathbf{B}_{\mathcal{S}_\tau}\mathbf{R}_{\mathcal{S}_\tau,\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{N}_\tau} \\ & +\mathbf{B}_{\mathcal{N}_t} E\left[\mathbf{R}_{\mathcal{N}_t,\mathcal{S}_t}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_\tau}^{\top}\mathbf{B}_{\mathcal{S}_\tau} E\left[\mathbf{R}_{\mathcal{S}_\tau,\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{N}_\tau} \end{aligned} \tag{S33}$$

Under the conditioning of B&C, $\mathbf{R}_{\mathcal{S},\mathcal{N}_\tau}$ is a random variable. If we assume that the linkage-disequilibrium between the neutral alleles and the selected alleles has arbitrary sign, then $E[\mathbf{R}_{\mathcal{S},\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha}] = \mathbf{0}$, which will be met if the reference allele is chosen arbitrarily (e.g. not based on minor allele frequency (Good, 2022)). Under this assumption (Assumption C) the final term in Equation S33 disappears to give:

$$COV(\Delta\mathbf{p}_{\mathcal{N}_t}, \Delta\mathbf{p}_{\mathcal{N}_\tau}^{\top}|\mathbf{B},\boldsymbol{\alpha}) = \; \mathbf{B}_{\mathcal{N}_t} E\left[\mathbf{R}_{\mathcal{N}_t,\mathcal{S}_t}\mathbf{B}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_\tau}^{\top}\mathbf{B}_{\mathcal{S}_\tau}\mathbf{R}_{\mathcal{S}_\tau,\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{N}_\tau} \tag{S34}$$

As in our inference section, the vector of allele frequency changes could be transformed using matrix $\mathbf{P}$ ($\Delta\overrightarrow{\mathbf{p}} = \mathbf{P}\Delta\mathbf{p}$) and B&C use the projection $\mathbf{P} = \mathbf{B}_{\mathcal{N}_t}^{-1}$ (they actually multiply this by $\sqrt{2}$ - see Equation S44) which results in

$$\begin{aligned} COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t}, \Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_\tau}^{\top}|\mathbf{B},\boldsymbol{\alpha}) = \; & E\left[\mathbf{R}_{\mathcal{N}_t,\mathcal{S}_t}\mathbf{B}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_\tau}^{\top}\mathbf{B}_{\mathcal{S}_\tau}\mathbf{R}_{\mathcal{S}_\tau,\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{N}_\tau}\mathbf{B}_{\mathcal{N}_t}^{-1} \\ = \; & \mathbf{B}_{\mathcal{N}_t}^{-1} E\left[\mathbf{R}_{\mathcal{N}_t,\mathcal{S}_t}\mathbf{B}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_t}\boldsymbol{\alpha}_{\mathcal{S}_\tau}^{\top}\mathbf{B}_{\mathcal{S}_\tau}\mathbf{R}_{\mathcal{S}_\tau,\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{N}_\tau} \end{aligned} \tag{S35}$$

Further derivation in B&C considers the expected value of a diagonal element of this matrix: the average (over neutral loci) covariance in projected allele frequency change. However, it is perhaps easier to note that the trace of this matrix is equal

to this average multiplied by the number of neutral loci, $n_{L_{\mathcal{N}}}$ (see below). Since the trace of an outer product is equal to the inner product we get:

$$\begin{aligned} Tr\left(COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t}, \Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_\tau}^\top|\mathbf{B},\boldsymbol{\alpha})\right) &= E\left[\boldsymbol{\alpha}_{\mathcal{S}_t}^\top \mathbf{B}_{\mathcal{S}_t}\mathbf{R}_{\mathcal{S}_t,\mathcal{N}_t}\mathbf{B}_{\mathcal{N}_t}^{-1}\mathbf{B}_{\mathcal{N}_\tau}\mathbf{R}_{\mathcal{N}_\tau,\mathcal{S}_\tau}\mathbf{B}_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau}|\mathbf{B},\boldsymbol{\alpha}\right] \\ &= \boldsymbol{\alpha}_{\mathcal{S}_\tau}^\top \mathbf{B}_{\mathcal{S}_t} E\left[\mathbf{R}_{\mathcal{S}_t,\mathcal{N}_t}\mathbf{B}_{\mathcal{N}_t}^{-1}\mathbf{B}_{\mathcal{N}_\tau}\mathbf{R}_{\mathcal{N}_\tau,\mathcal{S}_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau} \end{aligned} \tag{S36}$$

The diagonal element j of $\mathbf{W}_{t\tau} = \mathbf{R}_{\mathcal{S}_t,\mathcal{N}_t}\mathbf{B}_{\mathcal{N}_t}^{-1}\mathbf{B}_{\mathcal{N}_\tau}\mathbf{R}_{\mathcal{N}_\tau,\mathcal{S}_\tau}$ is equal to the sum of selected locus j's $R_{j_t,i_t}R_{j_\tau,i_\tau}(b_{i_\tau}/b_{i_t})$ across all neutral loci i. The jk^{th} off-diagonal element is the sum of $R_{j_t,i_t}R_{k_\tau,i_\tau}(b_{i_\tau}/b_{i_t})$ for selected loci j and k across all neutral loci i. If we write $\mathbf{W}_{t\tau} = \mathbf{H}_{t\tau} + (\mathbf{W}_{t\tau} - \mathbf{H}_{t\tau})$ where $\mathbf{H}_{t\tau}$ and $\mathbf{W}_{t\tau} - \mathbf{H}_{t\tau}$ are zero but for the diagonal and off-diagonal elements respectively, then Equation S36 becomes

$$\begin{aligned} Tr\left(COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t}, \Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_\tau}^\top|\mathbf{B},\boldsymbol{\alpha})\right) &= \boldsymbol{\alpha}_{\mathcal{S}_t}^\top \mathbf{B}_{\mathcal{S}_t} E\left[\mathbf{H}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau} \\ &\quad + \boldsymbol{\alpha}_{\mathcal{S}_t}^\top \mathbf{B}_{\mathcal{S}_t} E\left[\mathbf{W}_{t\tau} - \mathbf{H}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau} \end{aligned} \tag{S37}$$

If we focus on a system with two selected loci, j and k, then the final term in Equation S37 is equal to:

$$\sum_j\sum_k\left(\alpha_{j_t}\alpha_{k_\tau}b_{j_t,j_t}b_{k_\tau,k_\tau}\sum_i\left(E\left[R_{i_t,j_t}R_{i_\tau,k_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]b_{i_\tau}/b_{i_t}\right) + \alpha_{j_\tau}\alpha_{k_t}b_{j_\tau,j_\tau}b_{k_t,k_t}\sum_i\left(E\left[R_{i_\tau,j_\tau}R_{i_t,k}|\mathbf{B},\boldsymbol{\alpha}\right]b_{i_\tau}/b_{i_t}\right)\right). \tag{S38}$$

Under Hill-Robertson interference, if α_j and α_k have the same sign we expect them to be in negative LD with each other, and as a consequence have opposing patterns of LD with the neutral loci (i.e. if $\alpha_j\alpha_k > 0$ then we expect $E[R_{i,j}R_{i,k}] < 0$ and vice versa). Although the terms of these products are evaluated at different generations in Equation S38 (t and τ) we expect the terms to share sign in the same way, generating a negative expectation for the second term in Equation S37. However, assuming an absence of Hill-Robertson interference, or signed linkage-disequilibrium more generally (Assumption D), then the final term in Equation S37 can be dropped:

$$Tr\left(COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t}, \Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_\tau}^\top)|\mathbf{B},\boldsymbol{\alpha}\right) = \boldsymbol{\alpha}_{\mathcal{S}_t}^\top \mathbf{B}_{\mathcal{S}_t} E\left[\mathbf{H}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau} \tag{S39}$$

Based on a deterministic model for changes in linkage-disequilibrium (Assumption E) and assuming nongametic-phase linkage-disequilibrium is absent (Assumption F), Equations 40-44 in B&C derive an expression from which $E\left[\mathbf{H}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]$ can be computed. Under the assumption that drift (or selection) does not alter the dynamics of $\mathbf{R}$ then (Equation 42 in B&C):

$$L_{j_\tau,i_\tau} = L_{j_t,i_t}\frac{b_{j_\tau}^2}{b_{j_t}^2}(1 - r(g_{j,i}))^{\tau-t} \tag{S40}$$

which implies

$$R_{j_\tau,i_\tau} = \quad R_{j_t,i_t}\frac{b_{j_\tau}b_{i_t}}{b_{j_t}b_{i_\tau}}(1-r(g_{j,i}))^{\tau-t} \tag{S41}$$

and so

$$\begin{aligned} E\left[H_{j_t,j_\tau}|\mathbf{B},\boldsymbol{\alpha}\right] &= \textstyle\sum_i E\left[R_{j_t,i_t}R_{j_\tau,i_\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\frac{b_{i_\tau}}{b_{i_t}} \\ &= \frac{b_{j_\tau}}{b_{j_t}}\textstyle\sum_i E\left[R^2_{j_t,i_t}|\mathbf{B},\boldsymbol{\alpha}\right](1-r(g_{j,i}))^{\tau-t} \end{aligned} \tag{S42}$$

where $r(g_{j,i})$ is the recombination rate as a function of the distance $g_{i,j}$ between the two loci. Writing $\mathbf{F}_{t\tau}$ a diagonal matrix with the jth element equal to $\sum_i E\left[R^2_{j_t,i_t}|\mathbf{B},\boldsymbol{\alpha}\right](1-r(g_{j,i}))^{\tau-t}$, then $E\left[\mathbf{H}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right] = E\left[\mathbf{F}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_\tau}\mathbf{B}^{-1}_{\mathcal{S}_t}$ and so

$$\begin{aligned} Tr\left(COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t},\Delta\overrightarrow{\mathbf{p}}^\top_{\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha})\right) &= \boldsymbol{\alpha}^\top_{\mathcal{S}_t}\mathbf{B}_{\mathcal{S}_t}E\left[\mathbf{F}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_\tau}\mathbf{B}^{-1}_{\mathcal{S}_t}\mathbf{B}_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau} \\ &= \boldsymbol{\alpha}^\top_{\mathcal{S}_t}E\left[\mathbf{F}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}_{\mathcal{S}_\tau}\mathbf{B}_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau} \\ &= \boldsymbol{\alpha}^\top_{\mathcal{S}_t}E\left[\mathbf{F}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]\mathbf{B}^2_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau} \end{aligned} \tag{S43}$$

where (under random mating) $\mathbf{B}^2_{\mathcal{S}_\tau}$ is a diagonal matrix with elements proportional to the genetic diversities in generation τ. However, this is still hard to evaluate since $E\left[\mathbf{F}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]$ will vary over loci in a way that may depend on $\mathbf{B}$ and hence $\mathbf{B}^2_{\mathcal{S}_t}$. The method of B&C assumes that the sample correlation between the diagonal elements of $E\left[\mathbf{F}_{t\tau}|\mathbf{B},\boldsymbol{\alpha}\right]$ and the elements $\alpha_{j_t}\alpha_{j_\tau}b^2_{j_\tau}$ is zero (Assumption G). Under this assumption:

$$\begin{aligned} Tr\left(COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t},\Delta\overrightarrow{\mathbf{p}}^\top_{\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha})\right) &= \tfrac{1}{n_{L_\mathcal{S}}}\boldsymbol{\alpha}^\top_{\mathcal{S}_t}\mathbf{B}^2_{\mathcal{S}_\tau}\boldsymbol{\alpha}_{\mathcal{S}_\tau}E\left[Tr(\mathbf{F}_{t\tau})|\mathbf{B},\boldsymbol{\alpha}\right] \\ &= \tfrac{1}{n_{L_\mathcal{S}}}C_a(t\rightarrow\tau)E\left[Tr(\mathbf{F}_{t\tau})|\mathbf{B},\boldsymbol{\alpha}\right] \end{aligned} \tag{S44}$$

where $n_{L_\mathcal{S}}$ is the number of selected loci (note that in B&C the leading term is $\frac{1}{2n_{L_\mathcal{S}}}$ not $\frac{1}{n_{L_\mathcal{S}}}$ and this is because our projection matrices are only proportional by a factor $\sqrt{2}$ - see Equation S35). The method of B&C further assumes (Assumption H) that the average effects are constant in time such that $\boldsymbol{\alpha}_t = \boldsymbol{\alpha}_\tau$ then this reduces to

$$Tr\left(COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t},\Delta\overrightarrow{\mathbf{p}}^\top_{\mathcal{N}_\tau}|\mathbf{B},\boldsymbol{\alpha})\right) \quad = \tfrac{V_a(\tau)}{n_{L_\mathcal{S}}}E\left[Tr(\mathbf{F}_{t\tau})|\mathbf{B},\boldsymbol{\alpha}\right] \tag{S45}$$

Under Assumption H, Assumption G implies that the additive genic variance contributed by a selected locus is independent of its associations with neutral loci as measured by $R^2_{i_t,j_t}$. However, allele frequencies at selected loci enter the expression

for the additive genic variance, and they also dictate the range of R_{i_t,j_t} and therefore the magnitude of $R^2_{i_t,j_t}$: when selected allele frequencies are small compared to frequencies at neutral alleles, R_{i_t,j_t} cannot cover the full range of -1 to 1 (Sved and Hill, 2018).

B&C approximate $\mathbf{R}_t$, and therefore $\mathbf{F}_{t\tau}$ under the assumption of mutation-drift-recombination equilibrium (Assumption I). Ohta and Kimura (1971) derived the expectation of $L_{j_t,i_t}L_{j_t,i_t} = L^2_{j_t,i_t}$ under this assumption, although the expectation of $R^2_{j_t,i_t}$ can only be approximated as the expectation of $L^2_{j_t,i_t}$ divided by the expectation of the genetic diversities at the two loci, and is only accurate when the minor allele frequencies are greater than 10% (McVean, 2002) – and can be out by orders of magnitude when allele frequencies are extreme (Song and Song, 2007), as can be expected at loci under selection. Moreover, Ohta and Kimura (1971) derived the expectation $E[L^2_{j_t,i_t}]$ yet the B&C approach actually requires $E\left[L^2_{j_t,i_t}|\mathbf{B},\boldsymbol{\alpha}\right]$ which is considerably more challenging to compute (Good, 2022). In the Appendix B&C also relax, to some extent, Assumption I, where $Tr(\mathbf{F}_{t\tau})$ is calculated empirically using all loci, neutral and selected (Equation 55). In both cases - using the neutral expectation (Assumption I) or empirical LD (Assumption I-b) - the expected LD between selected and neutral loci will be overestimated since in reality selected alleles will be rarer than neutral alleles and so their LD, even measured as a correlation, will be reduced compared to that between neutral loci (Sved and Hill, 2018).

Equation S45 allows $V_a(\tau)$ to be estimated, but B&C aim to estimate $V_a(t)$. Under Assumption H, and assuming that the proportional change in genetic diversity at selected loci between generation t and τ is constant across loci (Assumption J: $\mathbf{B}^2_{\mathcal{S}_t}\mathbf{B}^{-2}_{\mathcal{S}_\tau} = \phi_{t,\tau}\mathbf{I}$), $V_a(t) = \phi_{t,\tau}V_a(\tau)$.

The above derivation assumes that the sites can be partitioned into selected and neutral sites and that the map positions of all sites are known. In practice, this will often be infeasible and so a number of additional assumptions are required. In the absence of map positions, Haldane's (1919) mapping function is assumed for $r(g)$ (Assumption K) but, since the selected loci and their physical position, g, are assumed unknown, a model is also required for g. B&C assume that selected and neutral loci are distributed uniformly and independently such that $g_{j,i}$ has a triangular distribution (Assumption L). Also, since the selected sites are unobserved, $\phi_{t,\tau}$ cannot be computed and so it is assumed that that $\phi_{t,\tau}$ is equal to the ratio of genetic diversity at generation τ to genetic diversity at generation t across all neutral loci (Assumption M).

In addition to the assumptions/approximations made when developing the the-

ory, additional assumptions/approximations are made when making inferences from data. Rather than taking the average (over loci) covariance (over evolutionary replicates) the covariance over loci is taken. However, from the law of total covariance this is expected to yield the correct result under Assumptions B and C. In addition, rather than calculating the (co)variances in projected allele frequency change, it is assumed (Assumption N) that the (co)variances in actual allele frequency divided through by the average projection is a good approximation: $COV_L(\Delta p_t, \Delta p_\tau)/E[b_t^2]$ is a good approximation of $COV_L(\Delta p_t/b_t, \Delta p_\tau/b_t)$ where COV_L designates a covariance over loci (Equation 16 B&C). The two quantities are only expected to agree under restrictive conditions (Bohrnstedt and Goldberger, 1969). However, it is not necessary to make Assumption N and relaxing it can reduce the bias in the estimator considerably. Finally, if there is measurement error in allele frequencies and the same allele frequency measurements are used to calculate change over adjacent intervals then this will generate downward bias in the estimated covariances. This is corrected for in B&C by assuming the sampling error in allele frequencies is binomial around the true value (Assumption O), although in practice non-binomial causes of overdispersion are likely. While we have tested our method assuming allele frequencies are known without error, any error in the allele frequencies will look like drift if allele frequencies are measured independently in each time/replicate. Consequently, the sampling noise will be incorporated into the estimate of the residual variance.

To connect Equation S45 with the derivation in B&C more clearly, we can write

$$Tr\left(COV(\Delta\overrightarrow{\mathbf{p}}_{\mathcal{N}_t}, \Delta\overrightarrow{\mathbf{p}}^{\top}_{\mathcal{N}_\tau}|\mathbf{B}, \boldsymbol{\alpha})\right) = n_{L_\mathcal{N}}\overline{COV(\Delta\overrightarrow{p}_{\mathcal{N}_t}, \Delta\overrightarrow{p}^{\top}_{\mathcal{N}_\tau}|\mathbf{B}, \boldsymbol{\alpha})} \tag{S46}$$

where the overline denotes the sample mean for the $n_{L_\mathcal{N}}$ neutral loci. Similarly,

$$\begin{aligned} Tr(\mathbf{F}_{t\tau}) = &\ \textstyle\sum_j \sum_i E\left[R^2_{j_t,i_t}|\mathbf{B}, \boldsymbol{\alpha}\right](1 - r(g_{j,i}))^{\tau-t} \\ = &\ n_{L_\mathcal{N}}\overline{\textstyle\sum_j E\left[R^2_{j_t,\mathcal{N}_t}|\mathbf{B}, \boldsymbol{\alpha}\right](1 - r(g_{j,\mathcal{N}}))^{\tau-t}} \end{aligned} \tag{S47}$$

Equation S45 can then be written as

$$\overline{COV(\Delta\overrightarrow{p}_{\mathcal{N}_t}, \Delta\overrightarrow{p}^{\top}_{\mathcal{N}_\tau}|\mathbf{B}, \boldsymbol{\alpha})} = \tfrac{V_a(\tau)}{n_{L_\mathcal{S}}}\overline{\textstyle\sum_j E\left[R^2_{j_t,\mathcal{N}_t}|\mathbf{B}, \boldsymbol{\alpha}\right](1 - r(g_{j,\mathcal{N}}))^{\tau-t}} \tag{S48}$$

which is Equation 8 in B&C averaged over neutral loci and can be further simplified to

$$\overline{COV(\Delta\overrightarrow{p}_{\mathcal{N}_t}, \Delta\overrightarrow{p}^{\top}_{\mathcal{N}_\tau}|\mathbf{B}, \boldsymbol{\alpha})} = V_a(\tau)\overline{E\left[R^2_{\mathcal{S}_t,\mathcal{N}_t}|\mathbf{B}, \boldsymbol{\alpha}\right](1 - r(g_{\mathcal{S},\mathcal{N}}))^{\tau-t}} \tag{S49}$$

where the overline indicates the sample average of all the $n_{L_\mathcal{S}} n_{L_\mathcal{N}}$ pairwise comparisons between selected loci and neutral loci.

Extending this theory to allele frequency change measured in independent replicates, rather than at different time-points in a single population is straightforward (Buffalo and Coop, 2020). If allele frequency change is measured in two populations, m and n, both of which have been derived independently from the base population t generation in the past, Equation S49 can be expressed as:

$$\overline{COV(\Delta\overrightarrow{p}_{\mathcal{N}_m}, \Delta\overrightarrow{p}^{\top}_{\mathcal{N}_n}|\mathbf{B}, \boldsymbol{\alpha})} \quad = V_a(m,n)\overline{E\left[R^2_{\mathcal{S}_0,\mathcal{N}_0}|\mathbf{B}, \boldsymbol{\alpha}\right](1 - r(g_{\mathcal{S},\mathcal{N}}))^{2t}} \tag{S50}$$

assuming the average effects have stayed constant.

S6.1 Appendix 7: N_E and average effects under a log-linear model

In the simulations we simulate for individual k

$$y_k = \mathbf{c}_k^{\top}\boldsymbol{\eta} + e_k$$

where the environmental deviations, e, have zero mean and standard deviation σ_e. We then draw $2N_{t+1}$ parents with replacement from a multinomial with the unnormalised probability of k being a parent of an offspring equal to $P_k = exp(y_k)$ and the normalised probability being $p_k = P_k / \sum_j^{N_t} P_j$. Absolute fitness is therefore $2N_{t+1}p_k = 2\frac{N_{t+1}}{N_t E[P]}P_k$. Having $m = 2\frac{N_{t+1}}{N_t E[P]}$, then absolute fitness is $P_k m = exp(y_k)m$. Following Kojima (1959) we can define Fisher's average effect for relative fitness in terms of the partial derivative of each $E[c] = p_0$ with respect to mean fitness and then rescale by mean fitness:

$$\begin{aligned} \alpha_i = & \frac{1}{\bar{W}}\frac{\partial \bar{W}}{\partial p_{0_i}} \\ = & \frac{1}{E[exp(y)m]}\frac{\partial E[exp(y)m]}{\partial p_{0_i}} \\ = & \frac{1}{E[exp(y)]}\frac{\partial E[exp(y)]}{\partial p_{0_i}} \\ = & \frac{\partial log(E[exp(y)])}{\partial p_{0_i}} \end{aligned} \tag{S51}$$

Assuming that y to be normally distributed with mean μ_y and variance V_y then allows us to use the expression for the expectation of a log-normal distribution to write:

$$\begin{aligned}\alpha_i &= \frac{\partial log(exp(\mu_y+V_y/2))}{\partial p_i} \\ &= \frac{\partial(\mu_y+V_y/2)}{\partial p_{0_i}}\end{aligned} \tag{S52}$$

It is worth highlighting that, so far, we have made no assumptions about the distribution of the contributions ($c_i\eta_i$) made by individual loci to y. Next, we split up μ_y and V_y into the contributions of individual loci.

$$\begin{aligned}\mu_y &= \textstyle\sum_{i=1}^{n_L} E[c_i\eta_i] \\ &= \textstyle\sum_{i=1}^{n_L} p_{0_i}\eta_i\end{aligned} \tag{S53}$$

$$\begin{aligned}V_y &= \eta^\top \mathbf{L}_0 \eta + \sigma_e^2 \\ &= \textstyle\sum_{i=1}^{n_L} l_{0_{ii}}\eta_i^2 + 2\sum_{j\neq i}^{n_L} l_{0_{ij}}\eta_i\eta_j + \sigma_e^2\end{aligned} \tag{S54}$$

Assuming all additive genetic variance is genic, V_y simplifies to

$$V_y = \textstyle\sum_{i=1}^{n_L} l_{0_{ii}}\eta_i^2 + \sigma_e^2 \tag{S55}$$

and substituting these expressions for μ_y and V_y yields:

$$\alpha_i = \frac{\partial\left(\sum_{i=1}^{n_L}(p_{0_i}\eta_i + l_{0_{ii}}\eta_i^2/2)+\sigma_e^2\right)}{\partial p_{0_i}} \tag{S56}$$

Under Hardy-Weinberg Equilibrium $l_{0_{ii}} = p_i(1-p_i)/2$ and this simplifies to

$$\begin{aligned}\alpha_i &= \frac{\partial\left(\sum_{i=1}^{n_L}(p_{0_i}\eta_i + p_{0_i}(1-p_{0_i})\eta_i^2/2)+\sigma_e^2\right)}{\partial p_{0_i}} \\ &= \eta_i + \tfrac{1}{4}(1-2p_{0_i})\eta_i^2 \\ &= \eta_i - \tfrac{1}{4}(p_{0_i} - q_{0_i})\eta_i^2\end{aligned} \tag{S57}$$

Note $\frac{1}{4}\eta^2$ is the quantitative genetic dominance deviation, often denoted d.

In order to calculate N_E under this multinomial log-linear model we need to know the variance in offspring number V_o in the presence of environmental variance in fitness only. Then $N_E = 4N/(2+V_o)$ where N is the census population size (Wright, 1938). The expected number of offspring for parent k is $2N_{t+1}p_k$, as given above, and the variance in the number offspring is $2N_{t+1}p_k(1-p_k)$. From the the law of total variance $V_o = 4N_{t+1}^2 Var(p) + 2N_{t+1}E[p(1-p)]$. Since $E[p(1-p)] = E[p] - E[p^2]$ and $Var(p) = E[p^2] - E[p]^2$ then $E[p(1-p)] = E[p] - Var(p) + E[p]^2$. Since $E[p] = 1/N_t$ by definition,

$$
\begin{aligned}
V_o = & \; 4N_{t+1}^2 Var(p) + 2N_{t+1}E[p(1-p)] \\
= & \; 4N_{t+1}^2 Var(p) + \tfrac{2N_{t+1}}{N_t} - 2N_{t+1}Var(p) + \tfrac{2N_{t+1}}{N_t^2} \\
= & \; (4N_{t+1}^2 - 2N_{t+1})Var(p) + \tfrac{2N_{t+1}}{N_t} + \tfrac{2N_{t+1}}{N_t^2}
\end{aligned} \tag{S58}
$$

From the properties of the log-normal (with zero mean) we know $E[P] = exp(\sigma_e^2/2)$ and $Var(P) = (exp(\sigma_e^2) - 1)exp(\sigma_e^2)$. Since $p = P/(N_t E[P])$

$$
\begin{aligned}
Var(p) = & \; Var(P)/(N_t E[P])^2 \\
= & \; (exp(\sigma_e^2) - 1)exp(\sigma_e^2)/(N_t exp((\sigma_e^2)/2))^2 \\
= & \; (exp(\sigma_e^2) - 1)exp(\sigma_e^2)/(N_t^2 exp(\sigma_e^2)) \\
= & \; (exp(\sigma_e^2) - 1)/N_t^2
\end{aligned} \tag{S59}
$$

If N_{t+1} and N_t are large then Equation S58 simplifies to

$$
V_o = 4N_{t+1}^2 Var(p) + 2N_{t+1}/N_t \tag{S60}
$$

such that

$$
V_o = (4N_{t+1}^2/N_t^2)(exp(\sigma_e^2) - 1) + 2N_{t+1}/N_t \tag{S61}
$$

If the population size is constant then this simplifies to $V_o = 4exp(\sigma_e^2) - 2$ and the variance effective population size, N_{E_t}, is $4N_t/(2 + V_o) = N_t/exp(\sigma_e^2)$.

S7 Supplementary Figures

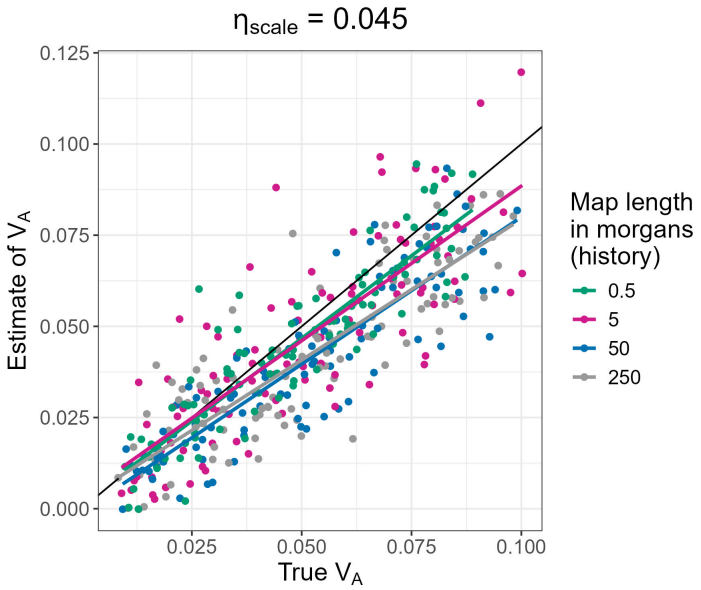

Figure S1: Results of full simulations with a burn-in phase of 25000 generations (map length in the history phase = 2 morgans, number of replicate populations = 10, population size = 1000, number of generations = 3, η_{scale} = 0.045) at different levels of the map length in the history phase (0.5 morgans: pink, 5 morgans: blue, and 50 morgans: grey). The coloured solid lines represent regression lines for estimates of V_A vs true values of V_A.

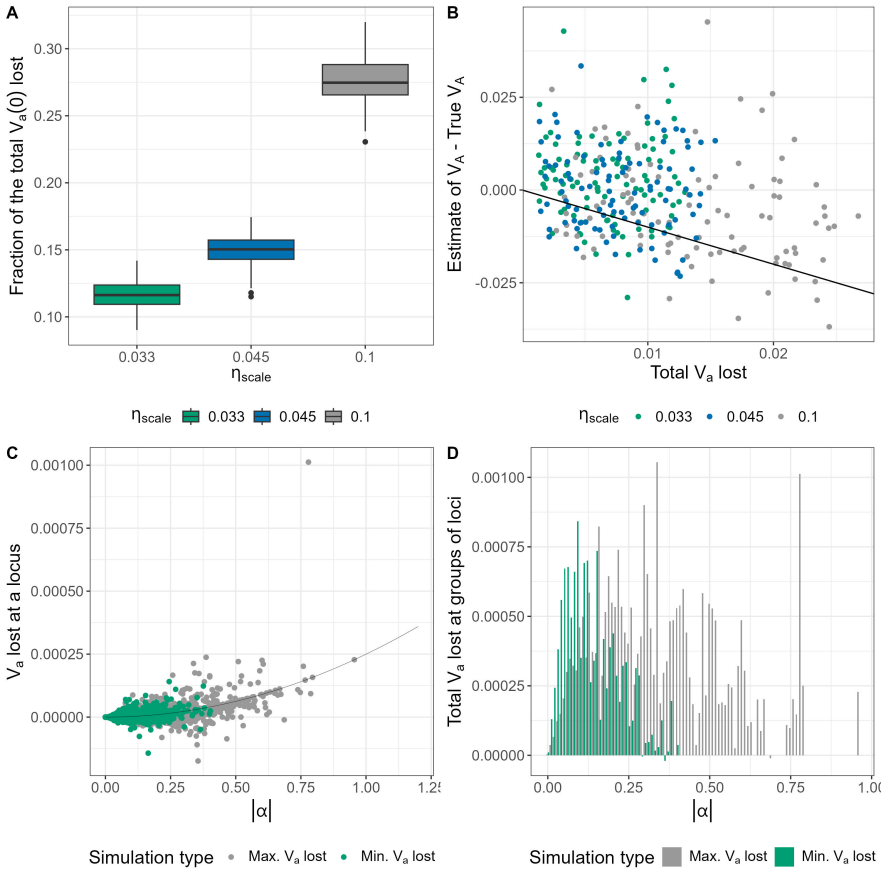

Figure S2: (A) Fraction of the total additive genic variance ($V_a(0)$) lost during the experiment (averaged over all the replicate populations) plotted as a function of η_{scale}. Each boxplot represents 100 independent simulations shown (for the corresponding level of η_{scale}) in Figure 4D. (B) The error on our estimates of V_A plotted versus the total additive genic variance (V_a) lost during the experiment (averaged over all the replicate populations) for simulations where the non-neutral η_{scale} was either 0.033 (green), 0.045 (blue), or 0.1 (grey). The solid line has an intercept of 0 and a slope of -1. (C) The contribution of each non-neutral locus (averaged over all the replicate populations) to V_a lost during the experiment plotted versus $|\alpha|$ for that locus for a simulation with $\eta_{scale} = 0.1$ that had the maximum V_a loss (grey) and a simulation with $\eta_{scale} = 0.033$ that had a comparable true initial V_A but the minimum V_a loss (green) in the experiment. The solid curve represents the additive genic variance lost when a singleton goes extinct. (D) The total V_a lost at groups of loci classified by dividing $|\alpha|$ into intervals of 0.01 for the simulation with the maximum V_a loss (grey) and the simulation with the minimum V_a loss (green) in the experiment. In the simulation with the maximum V_A loss – in which the mean $|\alpha|$ for non-neutral loci was 0.0410 – more than 50% of the total loss in V_a was driven by just 0.96% loci whose mod α was greater than 0.3.

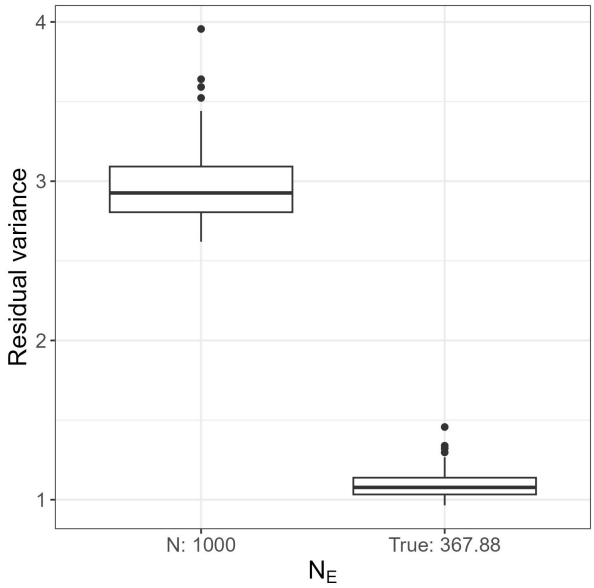

Figure S3: Residual variance in the models analysing the results of full simulations shown in Figure 3 using either the number of individuals (1000) instead of the N_E or the true N_E.

References

Bohrnstedt, G. W., and A. S. Goldberger, 1969 On exact covariance of products of random variables. Journal of the American Statistical Association **64**: 1439–1442.

Buffalo, V., and G. Coop, 2019 The linked selection signature of rapid adaptation in temporal genomic data. Genetics **213**: 1007–1045.

Buffalo, V., and G. Coop, 2020 Estimating the genome-wide contribution of selection to temporal allele frequency change. Proceedings of the National Academy of Sciences **117**: 20672–20680.

Bulmer, M. G. *et al.*, 1980 *The mathematical theory of quantitative genetics.*. Clarendon Press.

Ethier, S. N., and T. Nagylaki, 1980 Diffusion approximations of Markov chains with two time scales and applications to population genetics. Advances in Applied Probability **12**: 14–49.

Good, B. H., 2022 Linkage disequilibrium between rare mutations. Genetics **220**: iyac004.

Haldane, J. B. S., 1919 The combination of linkage values, and the calculation of distances between the loci of linked factors. Journal of Genetics **8**: 299–309.

Hill, W., and A. Robertson, 1968 Linkage disequilibrium in finite populations. Theoretical and applied genetics **38**: 226–231.

Kojima, K., 1959 Role of epistasis and overdominance in stability of equilibria with selection. Proceedings of the National Academy of Sciences **45**: 984–989.

McVean, G. A., 2002 A genealogical interpretation of linkage disequilibrium. Genetics **162**: 987–991.

Ohta, T., and M. Kimura, 1969 Linkage disequilibrium due to random genetic drift. Genetics Research **13**: 47–55.

Ohta, T., and M. Kimura, 1971 Linkage disequilibrium between two segregating nucleotide sites under the steady flux of mutations in a finite population. Genetics **68**: 571.

Santiago, E., and A. Caballero, 1998 Effective size and polymorphism of linked neutral loci in populations under directional selection. Genetics **149**: 2105–2117.

Song, Y. S., and J. S. Song, 2007 Analytic computation of the expectation of the linkage disequilibrium coefficient r^2. Theoretical population biology **71**: 49–60.

Sved, J. A., and W. G. Hill, 2018 One hundred years of linkage disequilibrium. Genetics **209**: 629–636.

Weir, B. S., 1996 *Genetic data analysis II: methods for discrete population genetic data*. Sunderland.

Weir, B. S., C. C. Cockerham, and M. Feldman, 1989 Complete characterization of disequilibrium at two loci. Mathematical evolutionary theory **86**: 110.

Wright, S., 1938 Size of population and breeding structure in relation to evolution. Science **87**: 340–431.

Zeng, J., R. De Vlaming, Y. Wu, M. R. Robinson, L. R. Lloyd-Jones *et al.*, 2018 Signatures of negative selection in the genetic architecture of human complex traits. Nature genetics **50**: 746–753.