The cost of selection in selective deaths paid out of reproductive excess

Joseph Matheson[1], Moises Exposito-Alonso[2], Joanna Masel[1]

[1]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA

[2]Departments of Plant Biology & Global Ecology, Carnegie Institution for Science, Stanford University, Stanford, California 94305, USA

Corresponding author: Joanna Masel

ABSTRACT

Haldane's Dilemma refers to the concern that the need for many "selective deaths" to complete a substitution (i.e. selective sweep) creates a speed limit to adaptation. However, discussion of this concern has been marked by confusion over which features of adaptation produce speed limits, what those limits are, and the consequences of violating speed limits. The term "substitution load" has been particularly unhelpful in this regard. Here we distinguish different historical lines of reasoning that lead to speed limits. We identify one line of reasoning, focused on finite reproductive excess, that has not yet been fully addressed. We develop this into a more general theoretical model that can apply to populations with any life history, even those for which a generation or even an individual are not well defined. We then apply this model to a dataset measuring survival of 517 different genotypes of *Arabidopsis thaliana* grown in eight different environmental conditions. These data suggest highly permissive limits to the speed of adaptation in all environmental conditions. This is because a much higher proportion of deaths contribute to adaptation than were anticipated during historical discussions of speed limits.

Keywords: substitutional load, adaptation rate, genetic load, fitness component, biological individual

INTRODUCTION

During an adaptive sweep, new alleles need to be substituted for old alleles across an entire population. This means that all individuals with the old alleles need to die, and individuals with new alleles must produce enough offspring to replenish the population. These requirements put a limit on the speed at which sweeps can happen. If sweeps need to happen rapidly at many sites, this limit could be prohibitive. Haldane (1957) was the first to use this reasoning to propose a rough estimate of the maximum speed at which sweeps could accumulate. This speed limit, later known as Haldane's dilemma (Van Valen 1963) motivated the development of Kimura's neutral theory (Kimura 1968). However, the logic behind Haldane's dilemma and the use of Haldane's specific estimate to motivate neutral theory have been challenged on multiple counts (Ewens 1970; Maynard Smith 1968; Felsenstein 1971; Kern and Hahn 2018). In these discussions, conceptually distinct approaches to quantifying the issue (and criticisms of those quantifications) are often described using identical terms, which apart from being confusing, leaves unresolved the critical question: what is the upper limit on the speed of adaptation and does it matter for natural populations?

Haldane made two somewhat different arguments in his seminal 1957 paper, muddying the waters from the beginning. In the first argument, he defined "selective deaths" as the subset of deaths $s(1-p)N$ that contribute to a change in the allele frequency p. This dramatic term is somewhat misleading; because reduced fecundity is mathematically equivalent to increased mortality, selective "deaths" can result from losing potential offspring, not just literal deaths. Haldane defined the "cost of selection" as the number of selective deaths required for a substitution (i.e. a selective sweep from low allele frequency to fixation). He calculated this cost as the integral of $s(1-p)N$ over the course of a sweep during which the frequency p of the adaptive allele rises from p_0 to close to 1 (Figure 1A). For modest s, most deaths during this time will not be selective deaths with respect to this single locus.

This cost of a sweep is typically expressed as a multiple of an (assumed constant) population size N. In a haploid, the number of selective deaths required for a sweep is equal to $N \times D$, where $D = -\ln(p_0)+\mathrm{O}(s)$, s is the difference in fitness between the two alleles, and p_0 is the initial frequency of the adaptive allele. For appropriately small s and p_0 (Haldane suggests $s < \frac{1}{3}$ and $p_0 = 10^{-4}$), the first term will be much larger than the second, meaning that D is nearly independent of the selection coefficient. For alternative assumptions about ploidy, dominance, and degree of inbreeding, the dependence of D on p_0 is different, but s remains unimportant unless close to 1 (Haldane 1957). Violating the assumption of constant population size does not change the calculation of D, only of N. In a representative case, such as $p_0 = 10^{-4}$ at a diploid autosomal locus with no dominance, $D = 18.4$. Haldane therefore estimated conservatively that 20-30 N selective deaths are likely to be typical for a sweep, after considering extreme cases where D can be significantly smaller (selection coefficients close to 1) or significantly larger (fully recessive beneficial alleles).

For his second argument, Haldane quantified the relationship between the speed of adaptation and the fitness reduction due to adaptation using the following model. Assume there are x loci independently undergoing sweeps, such that the current allele frequency at the ith locus reduces population fitness by d_i relative to its post-sweep value. The fitness of the population is then reduced below that of a hypothetical perfect population by a factor of $\prod_{i=1}^{x}(1-d_i)$, or about $e^{-\sum_{i=1}^{x} d_i}$. Haldane claimed (incorrectly) that this implies $N\sum d_i$ selective deaths per generation, allowing him to convert a given fitness difference into a rate of selective deaths per generation. With $30N$ selective deaths required to complete a selective sweep,

it would take $\frac{30}{\sum d_i}$ generations per sweep. During this time, mean fitness will be reduced relative to the fitness of a population in which all sweeps are complete, arbitrarily designated as 1, down to $e^{-30/n}$, where *n* is the offset time (in generations) between sweeps. This means that the faster sweeps accumulate, the more that population fitness is reduced relative to the optimum.

This second argument of Haldane's is a load argument. Load is a reduction of a population's fitness relative to a reference optimal genotype (Figure 1B). When the reduction in fitness is measured relative to an ideal genotype, it is referred to as lag load (Maynard Smith 1976). The term lag load is inspired by adaptation to a changing environment, where new mutations are required to keep up (Bertram, Gomez, and Masel 2017). (Note that lag load models are distinct from "evolutionary rescue" models (Bell 2017), which start from a scenario in which adaptation has previously failed to catch up and ask whether extinction can be avoided in the immediate aftermath, without asking whether the scenario will promptly recur.) The concept of lag load can be extended to a static environment where innovative new adaptive alleles reveal the possibility of an even better optimal genotype. The interpretation of what it means to lag behind the environment is important when interpreting arguments about speed limits on adaptation. A species adapting in a constant environment has no need to worry about speed limits, but real species must also reckon with the fact that their environment (biotic and abiotic) is rapidly changing in ways that might threaten population persistence in the absence of adaptation.

Haldane argues *a priori* that species could probably only sustain about 10% selective deaths (which he incorrectly equated with 10% lag load) for any serious length of time. This led him to suggest that the maximum speed of adaptation is likely around one sweep every 300 generations, a speed limit later called "Haldane's dilemma" (Van Valen 1963). Here we will revisit Haldane's *a priori* assumption about 10% being the highest tolerable fraction of deaths that are selective, by directly quantifying this fraction using data on the survival and reproduction of different *Arabidopsis thaliana* genotypes.

The fact that each amino acid substitution requires a sweep, along with the observation that amino acid sequences diverge greatly among species — far faster than Haldane's limit — was the original evidence supporting neutral theory (Kimura 1968). Kimura and Ohta (1971) plugged in estimates of the actual rate of substitution in mammalian lineages as n in Haldane's equation $L = e^{-30/n}$, which produced what they considered to be an excessively large lag load. Although still a lag load argument, their argument was subtly different from Haldane's, arguing that a high lag load implies that typical individuals would need to have a biologically implausible fraction of their offspring die (Kimura and Ohta 1971).

Ewens pointed out that Haldane's load argument improperly assumes that selective deaths (the crucial factor for maximum adaptation rate) depend on the lag load, whereas in populations with many loci undergoing sweeps, the load that matters is mean fitness relative to the most fit individual present (Ewens 1970). Haldane's lag load calculation is instead for the reduction in fitness relative to an ideal genotype. In a population with many sweeps occurring at once, the likelihood that even a single individual has the ideal combination of alleles is vanishingly small (Figure 1B). More recent travelling wave theories have rediscovered the importance of this relative type of load, which has come to be known as the "lead" (Desai and Fisher 2007).

One approach to quantify the maximum rate of sweeps, based only on the fitness of individuals actually present (i.e. the lead), is to begin with the variance in fitness (Ewens 1970; Kimura 1969). In the case of many independent sweeps at once, the variance in relative fitness (after normalizing mean population fitness as 1) is approximately s/n, where s is the selection coefficient of an adaptive allele and n is the number of generations between fixation events (Ewens 1970). The fittest genotype likely to be present

can be estimated using the statistics of extreme values. E.g., for a population of size 10^6, the most extreme fitness value likely to appear is around 4.9 standard deviations above the mean (Ewens 1970, see page 135). If we use Haldane's 10% as an estimate of the lead (instead of the lag load as above), then we obtain $4.9\sqrt{s/n} = 0.1$. For $s = 0.01$, n is around 20, much less than Haldane's estimate of 300, and the maximum rate of independent sweeps rises even higher for lower values of s. Ewens (1970) thus concludes that once previous load arguments are appropriately stated in terms of the lead, rather than in terms of lag load, there is no serious speed limit relative to observed rates of amino acid divergence. In other words, many simultaneous sweeps do not imply an implausibly large lead. Reframing load problems to instead be about variance has similarly been applied to deleterious mutation load arguments, e.g. as pertaining to an upper limit on the functional fraction of the human genome (Galeota-Sprung, Sniegowski, and Ewens 2020).

Ewens' approach continues the traditional emphasis of evolutionary genetics on the relationship between variance in fitness and the rate at which populations adapt (Fisher 1930; Crow 1958; Ewens 1979). Modern traveling wave theory instead derives the lead directly from s, N, and the beneficial mutation rate U, and obtains the fitness variance from that (Desai and Fisher 2007), rather than relying on our ability to directly measure fitness and its variance as an input to the calculation. However, although Ewens' variance-based approach puts to rest arguments attempting to convert lag load into a speed limit (i.e. Haldane's second argument), it doesn't address Haldane's first line of argument as discussed above: the cost of natural selection in terms of selective deaths.

Maynard Smith made a quite different argument against speed limits, claiming that the reason Haldane's dilemma is not a problem is pervasive synergistic epistasis. Synergistic epistasis can increase differences in fitness above those expected from differences in the numbers of beneficial mutations, thereby making each selective death more likely to count towards multiple sweeps at once (Maynard Smith 1976). A persistent source of confusion has been that in his model of truncation selection, Maynard Smith also made the shift from Haldane's absolute fitness to a more standard population genetic relative fitness, and hence from lag load to lead (Maynard Smith 1976). The fact that Haldane's dilemma did not arise in this model might therefore be due to reasons put forth by Ewens, rather than due to epistasis.

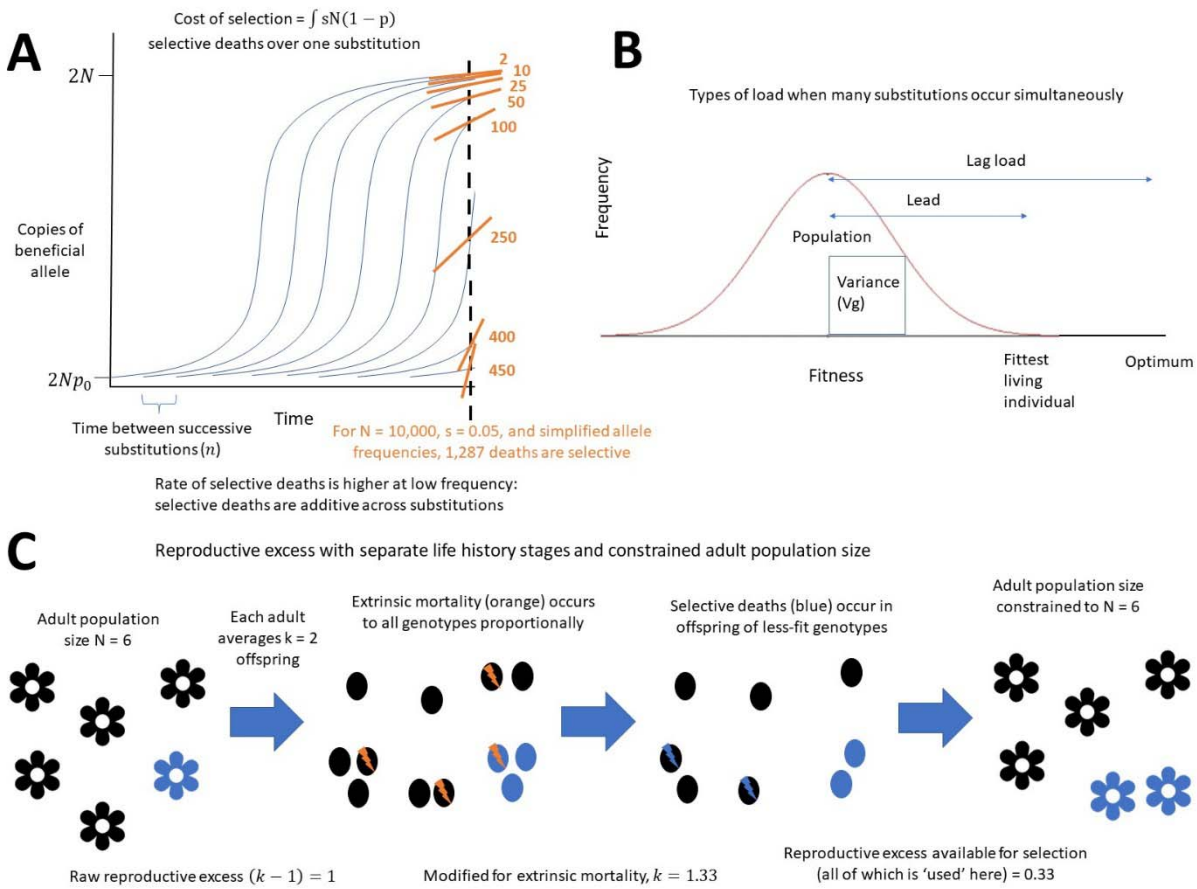

Figure 1. Three different types of arguments have been used to argue for limits to the speed of adaptation. A) The cost of selection is the number of selective deaths that must occur over time to complete a single sweep (each sweep shown as a logistic curve). The cost of selection at any given moment in time is the sum of the costs for each current sweep, illustrated as the slopes of the orange lines. B) Load arguments calculate the reduction in mean fitness of a population from what it could be. C) Finite reproductive excess imposes an upper limit on how many selective deaths per generation a population can sustain, which sets an upper bound to how fast substitutions can occur.

A further source of confusion is that different papers use the term "substitutional load", which we avoid here, to mean very different things. 'Substitutional load' has been used to refer to what we here call the lag load (Kimura and Ohta 1971), the cost of selection (Kimura 1968), the lead (Maynard Smith 1976), the number of offspring that the most fit genotype must produce (Ewens 1979), the sum of lag load across all generations involved in a substitution (Nei 1971; Kimura 1960), and even more broadly to refer to variance rather than load-based arguments when made in the context of similar questions (Ewens 1970). This confusion in terminology has obscured the consequences of formulating Haldane's dilemma in different ways.

A key overlooked piece in this debate is that selective deaths can occur at any life history stage. When their "cost" is paid out of a reproductive excess that would in any case subsequently be pruned during density regulation, its payment need not cause a drop in adult population size. However, during hypothetical excessively rapid adaptation, the cost might exceed the available reproductive excess. The

development of Haldane's selective deaths argument to accommodate these considerations therefore requires explicit consideration of reproductive excess (Felsenstein 1971; Nei 1971). The Wright-Fisher and related relative fitness models, which dominate population genetics, implicitly assume inexhaustible reproductive excess (Bertram and Masel 2019). However, real populations have a finite reproductive excess, leading to exponential population decline when exhausted.

The speed limit implied by the exhaustion of reproductive excess has been calculated for independently evolving sites in a haploid population with adult population size N (Felsenstein 1971; Nei 1971). Each adult has fecundity k, i.e. produces k offspring prior to juvenile deaths (Figure 1C, first arrow). In their deterministic models, k is exact for all adults, but the theory readily generalizes to interpreting k as an expectation. The raw reproductive excess is defined as $(k-1)N$, with $k > 1$. Some of this reproductive excess is lost to non-selective mortality; in these models, the non-selective mortality rate is not set extrinsically but derived from population size regulation occurring subsequent to selective mortality (Figure 1C, far right). Extrinsic mortality occurring prior to selective mortality, and hence at a fixed rate, can be folded into a lower value of k; the number of selective deaths then available for adaptation is then still denoted $(k-1)N$, although this no longer represents raw reproductive excess (Figure 1C, second arrow). Haldane's estimate of $k = 1.1$ (resulting in a maximum of 10% selective deaths), which Nei (1971) and Felsenstein (1971) retain, presumably already includes an expectation that extrinsic mortality substantially reduces the fecundity available to be 'used' for selective deaths.

The population then undergoes sweeps, all with the same selection coefficient s, which Nei (1971) and Felsenstein (1971) assume applies to survival rather than fecundity. Each begins at the same frequency p_0 (e.g. $1/N$ for *de novo* mutations in a haploid population). Assuming that each sweep is independent implies that each sweep will follow the same trajectory with a mean delay of n generations between sweeps (Figure 1A), in a similar manner to the more modern approach of Desai and Fisher (2007). For any given point in time (e.g. the dashed vertical line in Figure 1A), the cost of selection can be calculated for each locus individually (slopes of orange lines in Figure 1A); Haldane's integral takes the expectation of this sum. Comparing this cost to the reproductive excess of the population produces the result that the minimum spacing n that can be sustained without reproductive excess dropping below 0 (i.e. k dropping below 1) is $-{\ln(p_0)}/{\ln(k)}$. For Haldane's estimates of $p_0 = 10^{-4}$ and $k = 1.1$, this yields $n = 97$ generations between selective sweeps. This can be compared to Haldane's original spacing of $-{\ln(p_0)}/{\ln\left(\frac{W_{max}}{\bar{W}}\right)} = 92$ for a denominator described by Haldane as a selection intensity $= \ln\left(\frac{W_{max}}{\bar{W}}\right) = 0.1$.

This limit is much slower than the speeds predicted by variance- or lead-based arguments, but is similar in magnitude to Haldane's original result. Importantly, this speed limit calculation is not subject to the same criticisms as Haldane's original argument. Where Haldane compared the mean fitness of the population to the mean fitness of a hypothetical perfect population, Nei's (1971) and Felsenstein's (1971) approach compares the available reproductive excess to the reproductive excess required to effect changes in allele frequencies. Even if no individual exists who possesses the beneficial allele at every segregating site, each sweep still needs to proceed on its trajectory, in the process generating a corresponding cost of selection. It is the finite nature of reproductive excess by all members of the population (shared in Nei (1971) and Felsenstein (1971)) that directly produces this limit on the rate of adaptation, past which the population size will exponentially decrease.

Explicit consideration of reproductive excess means abandoning population genetic simplifications like the Wright-Fisher model that consider only one life history stage with adult population size N, with an implied infinite number of juveniles (Bertram and Masel 2019). We need at least two life history stages: adults and juveniles, where the latter shows reproductive excess relative to the former. Haldane's first argument interpreted the integral of selective deaths over the course of a sweep as relative in magnitude to N. It is now clear that it must be interpreted relative to a denominator describing the population size at the appropriate life history stage (Kimura and Crow 1969). I.e., a population that pays a cost through juvenile selective deaths must do so in proportion to the total number of juveniles, not in proportion to the smaller population size of adults.

The emphasis on life history transitions rather than generations is a strength rather than a weakness of the selective deaths scheme. One of the many flaws of the concept of "fitness" (Van Valen 1989) is the difficulty of defining a "generation" for many species, especially colonial species for which an "individual" is not well defined (Wilson and Barker 2021). Consider for example the budding yeast *Saccharomyces cerevisiae.* Is each mitotic division a generation? Or each life cycle spanning from meiosis to meiosis, with variable number of mitoses in between? Or the span between outcrossing events, with variable occurrences of selfing as well as mitoses in between? Or is it best defined ecologically with respect to dispersal between resources that allow growth? Problems defining a generation arise for a broad range of species (albeit not humans), but are resolved when population genetics is viewed through the lens of costs of selection. The "generation" that matters in this view is literally the generation of reproductive excess, in contrast to other life history transitions that involve survival.

Felsenstein's (1971) and Nei's (1971) formulations of Haldane's dilemma ask how much reproductive excess is available for selective deaths, defined as $(k-1)N$ after controlling for extrinsic mortality, where N is the population size prior to the generation of reproductive excess. But the value of k is an effective value that can conceal much, the "-1" assumes that perfect density regulation demands no excess individuals above N, and the N refers always to the same adult life history stage. This "effective" reproductive excess parameter is better conceived of as a measure of the proportion of deaths that are selective than as a true reproductive excess, and indeed they used Haldane's estimate for 10% deaths being selective to set $k = 1.1$. While promising, their approach is not yet a general theory applicable to the full diversity of species.

Here we reformulate and generalize these ideas to match them with two empirically accessible questions. First, how much reproductive excess does a genotype or population produce, beyond what is needed to avoid decline? Second, what fraction of deaths are selective (and how does this compare to the 10% guess made by Haldane)? Posing questions in this form allows us to make the first empirical estimates with which to ground Haldane's approach. We use data from Exposito-Alonso et al. (2019), who counted or estimated every plant grown and seed produced of *Arabidopsis thaliana* cultivars from 517 different genotypes in one season, under 8 distinct environmental conditions. These data are not representative of natural conditions, but they suffice to illustrate how such an analysis can be done. Ours is the first direct application of Haldane's selective death arguments to empirical data, representing proof of principle.

METHODS

Environmental conditions

We re-analyze the data of Exposito-Alonso et al. (2019), who used a 2 × 2 × 2 design for environmental conditions, with the three treatments being climate, water availability, and adult density. We treat these as eight separate populations throughout the analysis that follows (each has its own tab in Supplementary

File 1). Within each population, seeds from 517 fully homozygous plant genotypes (taken from a parental generation grown under controlled conditions to control for parental effects) were grown in pots that included only plants of that genotype (first column in Supplementary File 1). The number of replicate pots per genotype per population was occasionally as few as one due to experimental losses, but mostly ranged between five to seven replicates (fifth column in Supplementary File 1). Our interest is in differences among genotypes, not among replicates. For this reason, we calculate key quantities of interest for each genotype-environment combination by averaging across all replicates. In particular, selective deaths are based on the difference between each suboptimal genotype and the highest performing genotype in that environmental condition.

For climate, plants were grown in outdoor field stations in either Tübingen, Germany (near the center of the species range of *A. thaliana*) or in Madrid, Spain (at the southern edge of the range). Plants were all artificially watered. The high-water treatment matched soil moisture levels near the station in Germany, and the low-water treatment matched soil moisture levels near the station in Spain. To generate high adult density, thirty seeds of the same genotype were planted per pot. For low density, several seeds (~10) were planted per pot, enough to ensure that at least one seed would germinate, but few enough that the seeds were unlikely to inhibit each other pre-germination. To avoid any competition between adult plants in this treatment, only one plant, chosen at random, was retained after germination, and the rest were plugged out and discarded. We refer to each treatment with a three-letter abbreviation: M or T for Madrid or Tübingen, L or H for low or high water, and I or P for a single individual plant or a population of thirty plants per pot. For example, the treatment with thirty seeds per pot grown in Madrid with high water is abbreviated as MHP.

Selective deaths

The experiment can be mapped reasonably easily onto theoretical treatments. In each environmental treatment, a starting population of seeds grows into adult plants, experiencing both selective deaths and non-selective deaths as they proceed from seeds to seedlings to adults. Plants which survive to become adults then produce k seeds on average, some of which would normally constitute the next generation, although the experiment concludes at the end of season. The experiment does not capture the life history stage of seed dispersal to fertile ground, to complete the life cycle that began with seeds planted in a pot.

Juvenile deaths must be treated differently for the low- and high-density treatments. In the low-density treatment, where exactly one seedling is retained after germination, we do not have access to data on selective seed deaths, and so consider only seedling selective deaths. In the high-density treatment, any of the thirty seeds that fail to survive to the end of the experiment are counted as deaths, whether due to seed death before germination or to subsequent seedling death; our selective death calculations do not differentiate between these two life history transitions. This means that across the two life history transitions at which juvenile plants can die (as planted seeds before germination and as seedlings), only one set of juvenile deaths is recorded in each density treatment, but they are not comparable. They are combined seed and seedling deaths in the high-density case, and seedling deaths alone in the low-density case. Histograms are shown in Supplementary Figures 1-2.

In each treatment, we score the observed juvenile death rate of the highest performing genotype as the baseline extrinsic mortality for all genotypes (i.e. as non-selective deaths). Conceptually (ignoring a correction for extreme value bias that is treated below), for each life history transition in each environmental condition we have:

$$\text{Selective deaths in the population} = \sum_i n_i(d_i - d_{best})$$

where n_i is the starting population of genotype i at that life history transition, d_i is the genotype's average death rate during that life history transition, and d_{best} is the average death rate of the genotype with the lowest death rate for that life history transition in that environment.

Test for genetic variance in fecundity

Selective "deaths" can also be defined for unrealized fecundity. We did not analyze this here, because of lack of evidence for significant genetic differences in fecundity. In support of this, we performed an ANOVA test on fecundity in each environmental condition. We only have information on fecundity as an aggregate per replicate pot (rather than per individual plant in the high-density condition), so we compare among-genotype variance to among-replicate variance. Note, however, that we remove replicates that had no adults surviving to reproductive maturity, as well as genotypes with only a single replicate pot with surviving adults (and therefore no way of estimating among-replicate variance). All surviving adults produced at least some seeds. We Box-Cox transformed the data for each pot with surviving adults in each environmental condition (see Supplementary Figure 3 for post-transform histograms) before performing the ANOVA.

Proportion of juvenile deaths selective

Because *Arabidopsis* is an annual plant, all juveniles will die by the end of the season, whether as selective deaths during the experiment, non-selective deaths during the experiment, or non-selective deaths after the end of the experiment. From this we obtain, for each environmental condition:

$$\text{Fraction of juvenile deaths selective} = \frac{Selective\ deaths\ of\ juveniles}{Starting\ population}$$

In the high-density populations, the starting population is 30 seeds per pot. In the low-density populations, the starting population is 1 seedling per pot.

Correction for extreme value bias

The estimated best genotype is subject to extreme value bias, leading to overestimation of the number of selective deaths. I.e., the best genotype observed is likely not only to be a superior genotype, but also to have outperformed its own expected death rate by chance. The more uncertainty in estimated genotypic survival rates, relative to true genetic variance, the worse the extreme value bias problem. Here we attempt a rough estimate of the magnitude of extreme value bias using the observed noise among replicates and among genotypes. We then subtract a conservative estimate of bias from our best observed genotype, and redo our calculations of selective deaths.

We perform 10,000 simulations per environmental treatment. In each simulation, we assume that the observed 517 genotypic death rates are the 'true' values for each genotype (thus slightly overestimating genetic variance within the population). We then resample the number of surviving seeds 'observed' for each replicate of that genotype using a binomial distribution and calculate the resulting observed death rate.

In each simulated dataset, we took the best observed genotype and recorded the difference between its observed value and its actual genotypic value, then averaged these differences across the 10,000

simulations to obtain estimated bias. We then adjusted our estimate of the number of selective deaths to incorporate this estimated bias:

$$\text{Total adjusted selective deaths} = \sum_{i=1}^{517} n_i(d_i - (d_{best} + bias)).$$

Adjustments to selective deaths are shown in Table 1. Across all environmental treatments, adjusting for extreme value bias leads to negligible change in the estimate of the proportion of deaths which are selective.

Environmental condition	Among-genotype variance	Observed maximum survival rate	Estimated bias	Unadjusted proportion of deaths selective	Adjusted proportion of deaths selective
MLI	0.0618	1	0	0.664	0.664
MHI	0.0144	1	0	0.0847	0.0847
TLI	0.0337	1	0	0.564	0.564
THI	0.0139	1	0	0.109	0.109
MLP	0.0113	1	0	0.96	0.96
MHP	0.0433	1	0	0.564	0.564
TLP	0.023	0.947	0.00418	0.633	0.629
THP	0.0195	0.871	0.0195	0.391	0.372

Table 1: Adjusting for extreme value bias has little effect on the estimated proportion of selective deaths.

In six out of eight environmental conditions, at least one genotype had all plants survive until the end of the experiment. This genotype then has an observed death rate of zero and an observed between-replicate variance of zero, and thus our approach will not find any bias. Determining bias in this case would require attempting to fit and then resample from a true distribution of genotypic values that includes death rates near but not equal to zero. We did not pursue this avenue, because resampling genotypes would add a new source of variance, and we need only correct for the extreme value bias pertaining to the genotypes actually studied. Given the small degree of bias observed in the two environmental conditions in which no genotype experienced perfect survival, we consider our approach sufficient to conclude that extreme value bias has little quantitative effect on our results.

Pairwise genotype comparisons

For every possible pair of genotypes, we repeat the analysis above to estimate selective deaths and the proportion of deaths which are selective, using the better genotype of the pair as the 'best' genotype in the calculation of selective deaths. With only two genotypes, we do not adjust for extreme value bias. Using whole-genome information, we calculated the total number of SNP differences between each pair (Hamming distance, number of allele differences out of 1,353,386 biallelic SNPs) using PLINK v1.9.

RESULTS

Theory

Previous theoretical treatments by Nei (1971) and Felsenstein (1971) assume that all genotypes have the same fecundity (k), i.e. that there is no selection on fecundity, only one life history transition representing survival. They also assume that extrinsic mortality has a density-dependent component such that the combination of selective and non-selective mortality is perfectly balanced with fecundity. These are obviously not realistic assumptions. However, the theory can be extended in a variety of ways.

Selective deaths during generative life history transitions

In the Methods, we formalize selective deaths during survival life history transitions. We can also describe differences in fecundity as selective 'deaths'. This is because mathematically, foregone fecundity is equivalent to deaths that take place immediately after fecundity, and can be treated as:

$$\text{Selective } '\text{deaths}' \text{ during differential fecundity} = N_i(b_{best} - b_i)$$

where N_i is the number of reproductive mature adults, b_i is the fecundity of genotype i, and b_{best} is the fecundity of the genotype with the highest fecundity in that environment.

Reproductive excess within a fixed life cycle

Next we generalize from just one life history transition experiencing selection to multiple that occur in a consistent order. We consider a life history transition j that starts with population size N_j and ends at population size $N_{j+1} = k_j N_j$:

$$\text{Reproductive excess after transition } j = k_j N_j - N_{min,j+1}$$

where $N_{min,j+1}$ is the minimum population size at the end of transition j that is required to result in a population size of N_j at the beginning of transition j in the next life history cycle. Note that $k_j > 1$ indicates fecundity while $k_j \leq 1$ indicates survival.

To produce selective deaths, $k_{i,j}$ must depend on genotype i. To capture density regulation, $k_{i,j}$ for at least some values of j must depend on population size N_j. The values $k_{i,j}$ can also be functions of the genotype frequencies and/or an absolute measure of time. Two life history transitions (survival and fecundity) is the minimum, but each of these can be broken up into multiple transitions. For example, survival ($k > 1$) can be broken into components representing survival at different ages, or a selective component depending only genotype vs. a density-dependent extrinsic mortality component depending only on N_j vs. an extrinsic mortality component occurring at a constant rate.

Reproductive excess can be calculated either with respect to the best genotype present (i.e. the one most likely to become fixed), or with respect to the population mean, by using different values of k_j and $N_{min,j+1}$. Reproductive excess with respect to the population mean is needed to avoid population decline in the next generation, while reproductive excess with respect to the best genotype describes the ability to avoid population decline that would continue even after the best genotype has swept to fixation. The best choice depends on the particulars of the population in question. For example, studying balancing selection calls for the population mean, while studying evolutionary rescue calls for the best genotype.

To calculate reproductive excess with respect to the population mean, $k_j = \sum k_{i,j} f_i$ where f_i is the frequency of genotype i at the beginning of the transition. To calculate $N_{min,j+1}$ we solve for it in:

$$N_j = N_{min,j+1} \sum_i f_i \prod_{x \neq j} k_{i,x}.$$

With respect to the best genotype, we use $k_{best,j}$ and instead solve for $N_{min,j+1}$ in:

$$N_j = N_{min,j+1} \prod_{x \neq j} k_{best,x}.$$

Nei and Felsenstein as a special case

Nei (1971) and Felsenstein (1971) treated reproductive excess in the conceptually simple case of only two alternating life history transitions: births and deaths. Deaths included *only* selective deaths, while the "effective" fecundity transition was non-selective. They handled non-selective deaths by collapsing them into the value of $k_{fecundity_effective}$, either before selection on survival (in which case non-selective deaths reduce $k_{fecundity_effective}$ in the current generation), or after selection on survival (in such a way as to exactly balance out any available deaths that were "unused" by selection in the current generation, by reducing $k_{fecundity_effective}$ in the subsequent generation). The product $k_{fecundity_effective} \times k_{selective_mortality}$ was thus constrained to not exceed 1, via a fudge factor in the former. They solved for equality to 1 in order to calculate the maximum amount of selective deaths. When this equality is satisfied, $k_{fecundity_effective}$ can be interpreted as the product of actual fecundity and non-selective survival. Fig. 1C interprets this scheme in a temporal manner, proceeding first through non-selective fecundity to produce raw reproductive excess, then the non-density-dependent component of extrinsic mortality, then selective deaths, and finally density-dependent extrinsic mortality to cap the population size at *N* adults. They score reproductive excess as subject to the first but not the second form of reduction down to "effective" fecundity, minus the *N* individuals needed to replace the population.

Reproductive excess beyond a fixed life cycle

Not all organisms proceed through the exact same sequence of life history transitions every time, e.g. with budding yeast experiencing a variable number of mitoses in between each meiosis, and a variable number of selfing events between each outcrossing. In this case we cannot take the product of an exact series of transitions. Instead, we privilege the life history transition that produces the most severe bottleneck, assuming that the population will spring back to vibrancy after. We define a minimum number of individuals N_{bot} who need to make it through to the other side of the bottleneck, and define

$$\text{Reproductive excess at transition } j = k_j N_j - \text{minimum needed to ensure } N_{bot} \text{ after bottleneck step}$$

We now need to take the expectation over all possible series of life history transitions, and solve for $N_{min,j+1}$ in

$$N_{bot} = \mathrm{E}\left(N_{min,j+1} \prod_{x}^{\substack{life\ history\ stages \\ between\ j\ and\ bot}} k_{best,x} \right).$$

The precise value of N_{bot} will be informed by the ecology of the species in question. It may be small, such as when just a modest number of new hosts, each colonized by just one infectious microorganism, is sufficient to ensure the population's future. The most appropriate value of N_{bot} is the smallest population size that reliably escapes extinction.

Comparison to fitness

It is instructive to calculate the required reproductive excess and the proportion of deaths selective in the asexual relative fitness model treated by Desai and Fisher (2007), whose Eq. 39 solves for the lead qs as a function of s, U, and N. Here qs is defined in Malthusian fitness terms, but we approximate it here in terms of per-generation fitness. If the mean genotype has absolute fitness ~ 1 given a constant population size, then the best genotype present (the nose) has absolute fitness $\sim 1 + qs$. The lead qs can thus be conceived of as the minimum reproductive excess, with respect to the optimal genotype, that is required in order to avoid limits to adaptation. This aligns closely with the parameter $k - 1$ of Nei (1971) and Felsenstein (1971). The actual reproductive excess in this model is infinite, as for all models that assume a constant population size and treat only relative fitness.

We next calculate the proportion of deaths that are selective. The entire reproductive excess of the best genotype present, $RE \geq 1 + qs$, represents non-selective deaths (or foregone fecundity) among its offspring. Other genotypes all experience the same rate of non-selective deaths. The strength of selection tells us that the per-capita odds that the next death hits a specific average individual rather than a specific nose individual are $1 : 1/(1 + qs)$. An average parent therefore experiences a total of $RE(1 + qs)$ offspring deaths during the time in which it experiences the RE non-selective offspring deaths that represent a generation, making the proportion of deaths that are selective equal to $1 - 1/(1 + qs)$.

For a substantial range of parameters, especially large s, U, and N, both the minimum required reproductive excess, and the proportion of juvenile deaths that are selective, exceed the previously assumed value of 0.1 (Figure 2). This application to the model of Desai and Fisher (2007) helps clarify the distinction between these two related properties, which were confounded into a single value of k by Nei (1971) and Felsenstein (1971). In this particular model, with its explicit adults and implicit juveniles under selection, they are both functions of the lead qs, but this need not continue to be so simple when more complex life histories are considered.

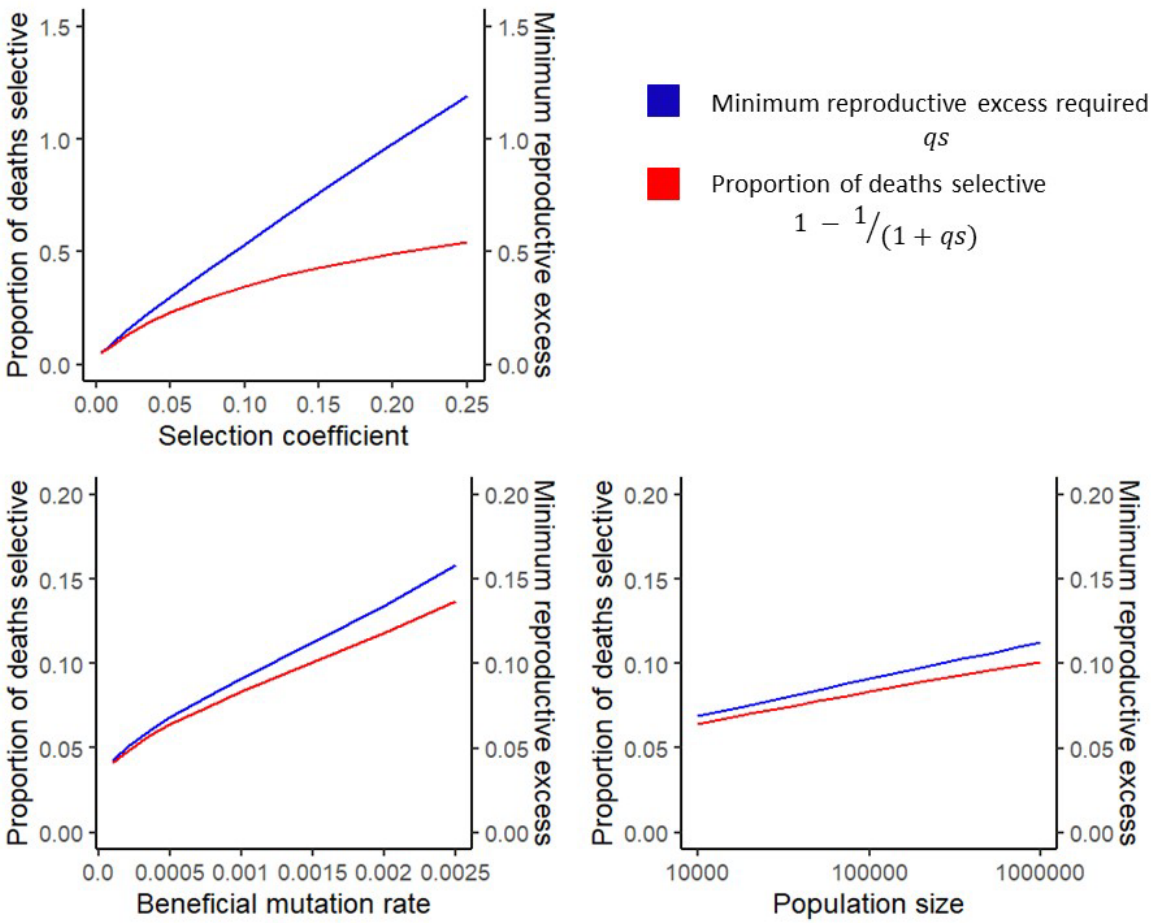

Figure 2. Minimum reproductive excess and the corresponding proportion of deaths that are selective in the model of Desai and Fisher (2007), as a function of the model's parameters U, s, and N. Parameter ranges are truncated to avoid the regime $\frac{s}{U} < 3$, for which the assumptions of the model break down. The minimum required reproductive excess qs and the corresponding proportion of selective deaths $1 - 1/(1+qs)$ were calculated by numerically solving Equation 39 for q in Desai and Fisher (2007).

Values of k in our framework are equivalent to fitness components, with respect to absolute rather than relative fitness. Haldane obtained selective deaths from $sN(1-p)$ over a time step of a complete generation, where s is the selection coefficient with respect to relative fitness. Applied to the asexual model of Desai and Fisher (2007) as a one locus model with many alleles instead of just two, this yields qs, i.e. the minimum reproductive excess, not the proportion of deaths that are selective. While beyond the scope of this work, we presume that the more pertinent quantity for affecting change in allele frequencies is the proportion of deaths that is selective, given its close relationship with fitness components k, and given that actual reproductive excess might greatly exceed the minimum required reproductive excess. We have shown how selective deaths can be derived directly from the underlying population dynamic model, without requiring either generation or relative fitness to be defined first.

Antagonistic pleiotropy is treated quite differently in a selective deaths framework than for fitness components. Per-generation fitness is the product of fitness components, such that when a genotype that

benefits fitness in one life history transition bears an antagonistically pleiotropic cost at another, the costs and benefits at least partially cancel out. In contrast, the cost of selective deaths accrues across life history transitions, with no cancelling out. Seasonal fluctuations can alter the trade-off balance, leading to large genotypic changes from fluctuating selection (Kelly 2022; Machado et al. 2021; Rudman et al. 2022). The cost of selection applies not only to the observed genotypic fluctuations in such cases, but also to the unobserved, cancelled out effects between life history transitions.

In the simple case of just births at rate b and deaths at rate d, classic population genetic per-generation fitness corresponds to the effective reproduction number ${}^{b}/_{d}$, while the Malthusian parameter gives an alternative formulation of fitness as $b - d$. A Malthusian approach is sometimes preferred to avoid the complications of age- or stage-structured populations. Our approach extends an effective reproduction number framing to these more complex scenarios, while avoiding dependence on the definition of one "generation". While the Malthusian approach is sufficient for many purposes, something closer to an effective reproduction number approach is required to capture the finite nature of reproductive excess and the corresponding limits to selective deaths and hence adaptation. We note that often the question being asked is simply what will invade, in which case either approach can be used (Lehmann et al. 2016; Metz, Geritz, and Nisbet 1992; Roff 2008).

Application of theory to Arabidopsis *data*

The experiment analyzed here has a fixed life cycle of four life history transitions (adults producing seeds, seeds successfully dispersing to suitable habitat, seeds surviving to be seedlings, seedlings surviving to be adults). Matching this, we define reproductive excess (RE) four different ways, as illustrated in Figure 3. Given the presence of many poorly adapted genotypes in the experiment, we perform each calculation with respect to the best genotype (denoted by the prime symbol $'$), yielding:

$$\mathrm{RE(seed\ survival)} = k'_{seed_survival} N_{seeds} - \frac{N_{seeds}}{k'_{seedling_survival} k'_{fecundity} k'_{dispersal}}$$

$$\mathrm{RE(seedling\ survival)} = k'_{seedling_survival} N_{seedlings} - \frac{N_{seedlings}}{k'_{fecundity} k'_{dispersal} k'_{seed_survival}}$$

$$\mathrm{RE(fecundity)} = k'_{fecundity} N_{adults} - \frac{N_{adults}}{k'_{dispersal} k'_{seed_survival} k'_{seedling_survival}}$$

$$\mathrm{RE(dispersal)} = k'_{dispersal} N_{seeds_produced} - \frac{N_{seeds_produced}}{k'_{seed_survival} k'_{seedling_survival} k'_{fecundity}}$$

In the high density environmental conditions, we use a single survival transition to cover both seeds and seedlings. Under low density conditions, where one seedling was chosen at random from the product of 10 planted seeds, we use $k_{seed_survival} = 0.1$ for all genotypes. We treat all genotypes as having the same $k_{fecundity}$ (due to lack of evidence for genetic variation – see Experimental Results below). Our experiment provides no information about values of $k_{dispersal}$, so we show calculations in a range from 0.01 to 0.1. We apply the same value to all genotypes, and because we do not have data on the number of seeds after dispersal, we do not calculate a reproductive excess for this transition. This means that we

calculate only two reproductive excesses for each environmental condition, one for survival and one for fecundity, although the calculations are different between low- and high-density conditions. These reproductive excesses are given by:

$$\text{RE(fecundity, low density)} = k_{fecundity} N_{adults} - \frac{N_{adults}}{0.1 k_{dispersal} k'_{seedling_{survival}}} \quad (1)$$

$$\text{RE(survival, low density)} = k'_{seedling_survival} N_{seedlings} - \frac{N_{seedlings}}{0.1 k_{dispersal} k_{fecundity}} \quad (2)$$

$$\text{RE(fecundity, high density)} = k_{fecundity} N_{adults} - \frac{N_{adults}}{k_{dispersal} k'_{survival}} \quad (3)$$

$$\text{RE(survival, high density)} = k'_{survival} N_{seeds} - \frac{N_{seeds}}{k_{dispersal} k_{fecundity}} \quad (4)$$

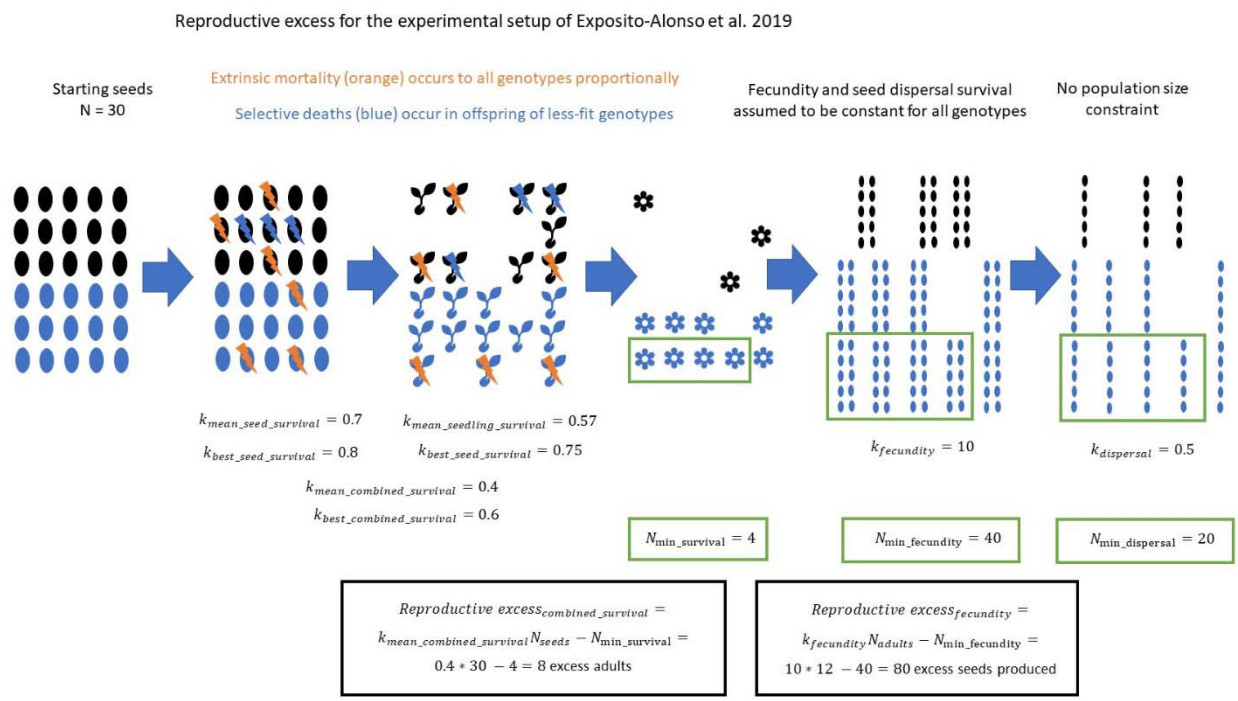

Figure 3. Worked example of reproductive excess for the life history transitions of *Arabidopsis* in the experimental setup from Exposito-Alonso *et al* (2019). We found no evidence for between-genotype differences in fecundity and the experimental setup provides no information about seed dispersal, so we show no selective deaths during these transitions. Specific values of k are chosen for illustrative purposes.

Experimental

Our empirical findings are restricted to selective deaths relating to juvenile survival. We did not analyze selective 'deaths' attributable to differences in fecundity, because an ANOVA on fecundity within and among genotypes showed no statistical support for any difference in mean fecundity among genotypes in six of the eight experimental conditions. Even in the two conditions with statistical significance, among-genotypic variance was three times smaller than among-replicate variance. A more sensitive MCMCglmm model with Poisson errors and controlling for replicate found non-zero heritability in 6/8 environments, but still below 10%, which we consider low enough to neglect.

The proportion of *A. thaliana* deaths that are selective substantially exceeds Haldane's 10% estimate in six out of eight experimental conditions, and is close to it in the other two (Figure 4 y-axis, Supplementary Table 1). In Madrid with low water and high density, as many as 95% of deaths are selective.

A priori, we expect high water and low density to be more benign, which might increase reproductive excess and/or reduce the proportion of deaths selective. While we cannot compare seed deaths at low density to seed plus seedling deaths at high density, these predictions are confirmed for high vs. low water: strongly in the case of selective deaths and excess seeds, and weakly with respect to excess seedlings (Figure 4, Supplementary Table 1). Estimated reproductive excess is fairly insensitive to the 10-fold range we consider for the proportion of seeds that successfully disperse to suitable habitats (vertical line length in Figure 4).

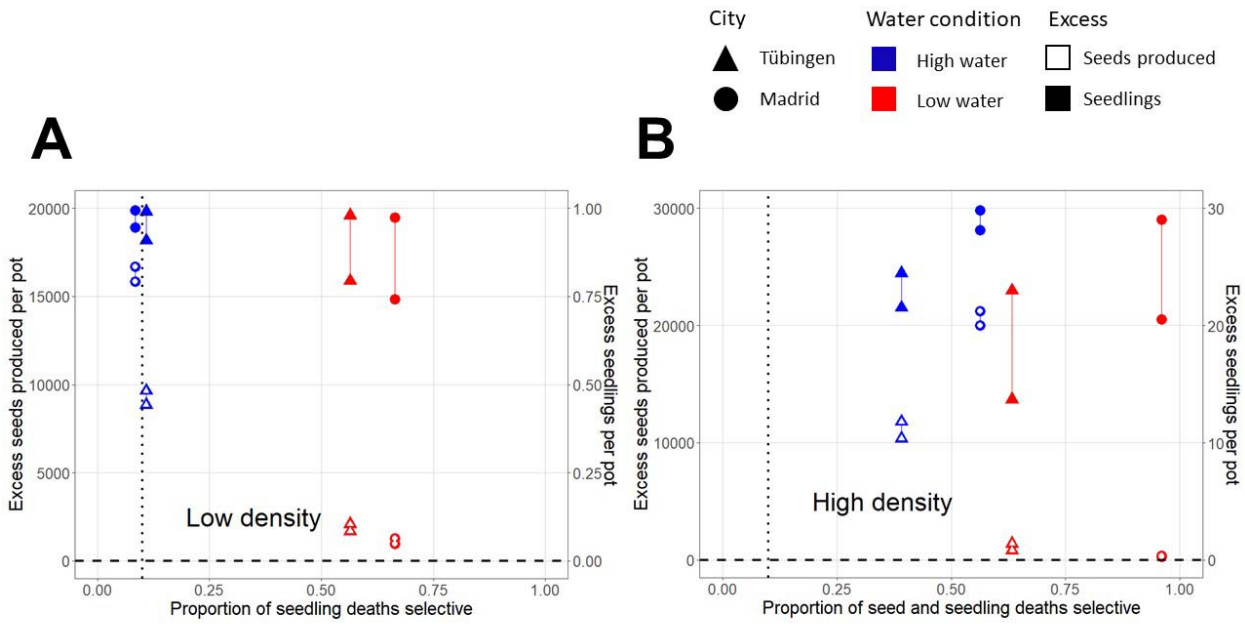

Figure 4. <u>The proportion of deaths that are selective generally exceeds Haldane's 10% estimate (dotted vertical line), with ample reproductive excess (above zero, shown as dashed horizontal line) especially under high-water conditions.</u> Selective deaths shown on the x-axis apply either to seeds plus seedlings (A) or to seedlings alone (B). Reproductive excess in seeds produced vs. seedlings surviving is shown with open vs. closed symbols, corresponding to values shown on the left and right y-axes, respectively. Reproductive excess is with respect to what the genotype with the highest survival produced above what would be required to replace the starting population of a pot. Note that reproductive excess of seedlings cannot exceed one per pot under low density conditions, and 30 per pot under high density conditions. Reproductive excess of seeds produced was calculated using equation 1 for low-density conditions and equation 3 for high-density conditions. Reproductive excess of surviving seedlings was calculated using equations 2 and 4. All reproductive excesses are shown as a vertical range, with the lower bound calculated using $k_{dispersal} = 0.01$ and the upper bound calculated using $k_{dispersal} = 0.1$. The proportions of deaths that are selective are adjusted for extreme value bias as shown in Table 1. Values can be found in Supplementary Table 1.

We expected *a priori* that harsher environmental conditions would have higher extrinsic mortality. However, this wasn't the case. We can estimate extrinsic mortality as the death rate of the best genotype, after correcting for extreme value bias (see columns 3 and 4 of Table 1). Most environmental conditions saw a highest-fitness genotype with perfect survival, or close to it. We saw the most extrinsic mortality in the THP environmental condition, which is not one of the harsher conditions. Nei (1971) and Felsenstein implicitly assume high extrinsic mortality via their choice of value for reproductive excess; we explicitly account for extrinsic mortality during mortality transitions. Extrinsic mortality might of course be much higher in natural conditions, lowering the proportion of deaths below the high values observed here.

The artificially high genetic diversity in our experiment might also act to inflate the proportion of deaths that are selective. If this were the case, then we expect that competition between more similar genotypes should lead to a smaller estimate for this proportion. We tested this prediction by repeating our analysis on every pair of genotypes, as though they were the only two genotypes in the experiment, and looking for a correlation between genetic distance and proportion of deaths selective. Note that we use genetic distance, much of it presumably neutral, as a proxy for genetic differences related to adaptation. Although some statistically significant correlations were observed in some environmental conditions, the direction of correlation was evenly split between negative and positive (as seen in Table 2), and the highest R^2 value observed in any environmental condition was $0.096^2 = 0.0092$, for seedling deaths in the MLI environment, which we deem biologically insignificant. This is reassuring with respect to the artificially high genetic diversity in our experiment.

Seedling deaths	Spearman's rho	p-value
MLI	0.096	2.2e-16
MHI	0.0004	0.88
TLI	-0.0024	0.39
THI	-0.0056	0.042
Combined seed and seedling deaths	Spearman's rho	p-value
MLP	0.055	2.2e-16
MHP	0.011	4.9e-5
TLP	-0.0081	0.0031
THP	-0.0085	0.002

Table 2: Genetic distance between a pair of genotypes does not consistently correlate with proportion of deaths selective. Visualization of each of the eight relationships is available in Supplementary Figures 4-5.

DISCUSSION

The prevailing consensus is that Haldane's Dilemma poses no real limitation to the speed of adaptation, despite persistent confusion as to the reason. Here we began by clarifying the primary issue with load-based arguments: correcting Haldane's use of an absolute rather than relative standard for fitness enables far more rapid adaptation. However, the historical consensus that adaptation is not significantly limited largely overlooks a different and crucial type of limitation pointed out by Nei (1971) and Felsenstein (1971), one that depends not on relative load but on the finite nature of reproductive excess. If we accepted Haldane's 10% guess as an estimate for the reproductive excess that is available for selective deaths to take place among, or as an estimate of the proportion of deaths that are selective, then the rate of adaptation would be significantly limited. We clarified and extended theoretical arguments regarding selective deaths drawn from finite reproductive excess, along with a simple proof-of-concept application to an experimental dataset. Our extension applies flexibly to different life histories, including those for which a "generation" is poorly defined, e.g. colonial organisms.

Our experimental results suggest a possible resolution to Haldane's concern that the rate of adaptation might be substantively limited. Our illustration of the model of Desai and Fisher (2007) confirms that the cost of selection could indeed be onerous if only 10% of deaths were available for selection. But the smallest proportion of selective deaths we observed across 8 environmental conditions was 8.5%, while in the most adverse environmental conditions, 95% of deaths were selective. Relaxing this auxiliary assumption about a critical parameter value resolves Haldane's concerns.

Our proof-of-concept data measures selective deaths under artificial conditions. An obvious concern with our set-up is that with genotypes representing the worldwide diversity of *A. thaliana*, exaggerated differences between the best-adapted and worst-adapted genotypes would inflate estimates of the proportion of deaths that were selective. These concerns are at least partially mitigated by our unexpected finding that the genetic distance between genotypes is not consistently related to the proportion of deaths which would be selective in a competition between genotypes. However, two similar genotypes in our experiment represent more genetic distance than might be present within a typical natural population. Future work under more natural conditions (e.g. with higher extrinsic mortality) and in different species (e.g. less fecund) remains necessary to reach the conclusion that the proportion of deaths that are selective is typically high.

Empirical demonstration of the concepts, even with serious caveats about the generalizability of the empirical example, makes the concepts more concrete. This is especially important because disparate usage of the term 'substitutional load' in the literature, as well as the variety of underlying lines of reasoning involved, has made this topic unnecessarily opaque. One aspect of our current work is simply to clarify the variety of lines of reasoning that produce limits on the rate of adaptation. Our more specific theoretical and empirical analyses then develop a line of reasoning about reproductive excess and selective deaths that was not previously resolved.

This approach, building on Haldane (1957), Nei (1971), and Felsenstein (1971), is not identical to standard modes of reasoning in genetics. In particular, quantitative genetics approaches follow Fisher (1930) to focus on variances — sums of differences squared — while selective deaths and reproductive excess are, like load, both differences, with no square operation. We hesitate to call this aspect of our approach 'novel', because it is clearly quite old, but it nevertheless comes across as novel with respect to aspects of contemporary evolutionary genetics.

Interestingly, the concept of relative load was later reinvented as the "lead", as part of calculations that derived the actual speed of adaptation v (rather than limits to it) from the beneficial mutation rate U, the population size N, and the selection coefficient s (Desai and Fisher 2007). One reason this solution was

not available to Haldane was that population genetics had not yet begun to treat origination processes (McCandlish and Stoltzfus 2014). Instead of treating a steady input of beneficial new mutations, Haldane considered a scenario in which environmental change activates beneficial variants within standing genetic variation. Indeed, a variant's initial frequency $-\ln(p_0)$ is the primary factor in determining the cost (and inversely, maximum speed) of a sweep. However, in many species, especially asexuals (or selfers like *Arabidopsis*), steady activation of standing genetic variation is unlikely, and it is more appropriate to treat *de novo* mutations each appearing at initial frequency $1/N$ or $1/2N$. For this case, a lead-based approach was used to derive the rate of beneficial sweeps in asexuals as $\frac{2s\ \ln[Ns]}{\ln^2[s/U]}$ for the simple case of constant s (Desai and Fisher 2007), for parameter ranges in which the previously derived rate UNs does not apply. Here we relaxed the assumption of infinite reproductive excess made by this relative fitness model, and calculated the more minimal assumption regarding required reproductive excess. We also reveal the model's implied fraction of deaths (or foregone fecundity) that are selective. Both quantities are functions of the lead.

Another approach, starting with Kimura (1961), is to use the framework of information theory to set bounds on the speed of adaptation. Natural selection increases the information stored in genomes (Adami 2012). Kimura calculates the amount of information acquired per sweep in terms of p_0 and then relates this to the cost of selection using Haldane's equation that $D = -\ln(p_0)$ (Kimura 1961). More recent approaches treat the bounds placed on the accumulation of information in much more detail, while treating either the Malthusian parameter (McGee et al. 2022) or classic discrete time relative fitness (Hledík et al. 2022). Both these approaches define an information "cost", but this is not equal to our cost in terms of selective deaths.

The historical significance of Haldane's arguments about limitations to adaptation is that they were convincingly used to support neutral theory. This was framed as a dilemma because data on the rates of amino acid divergence among species seemed to exceed Haldane's speed limit. The development of neutral theory resolved this apparent dilemma by suggesting that most amino acid substitutions are neutral and do not count against the speed limit. However, this historical argument is now on troubled ground, because recent literature argues that the fraction of substitutions explained by adaptation can be high (Galtier 2016; Sella et al. 2009; Uricchio, Petrov, and Enard 2019), and that on shorter timescales, as much as 37% of allele frequency changes are attributable to adaptation (Buffalo and Coop 2020). For example, recent experiments have shown consistently rapid, pervasive seasonal adaptation seasons in *Drosophila* (Machado et al. 2021; Bertram 2021; Kelly 2022). There are other resolutions, e.g. these estimates include substitutions of neutral alleles via hitchhiking. It is nevertheless curious that the empirical collapse of historical arguments for neutral theory has not led to a re-evaluation of related arguments by Haldane. Here we revise Haldane's arguments for the modern era, finding that Haldane's revised arguments are compatible with empirical evidence for abundant adaptation, while still posing upper limits that might matter in some contexts.

We note, however, that Haldane (1957) used a one-locus model – linkage disequilibrium will typically make the conversion of selective deaths to adaptation less efficient than implied by the assumption of independence used both in his calculations, and in the subsequent calculations of Nei (1971) and Felsenstein (1971). Here we also considered the other extreme, in the form of the asexual model of Desai and Fisher (2007). The focus of the current work is to calculate the required reproductive excess and the proportion of deaths that are selective; more work is required to quantify how this is converted into adaptation in a broader range of models of linkage disequilibrium and epistasis, as well as life histories and antagonistic pleiotropy.

Excitingly, unlike most approaches in evolutionary genetics, the approach we describe does not require the quantity 'fitness', which is deceptively difficult to define in a manner that can be generalized to all circumstances (Ariew and Lewontin 2004; Bertram and Masel 2019; Doebeli, Ispolatov, and Simon 2017; Van Valen 1989). Indeed, while "lineage fitness" solves a number of problems (Akçay and Van Cleve 2016; Graves and Weinreich 2017; King and Masel 2007), it does so at the cost of defining the fittest genotype to be that which eventually prevails, sacrificing much of the quantitative benefit of 'fitness'. In particular, standard quantitative definitions of either relative or absolute 'fitness' require a clear definition of a 'generation' over which change is assessed, which in turn requires a clear definition of an 'individual' whose life cycle a generation captures (Wilson and Barker 2021). Our generalized selective deaths approach is derived from selection, but does not require 'fitness' to be defined. Rather, we measure selective deaths from pairwise differences in fecundity and survival between each genotype and the best genotype present, during each life history transition. Reproductive excess corresponding to that life history transition indicates a potential for more stringent selection without a decline in population size. This approach applies at each life history transition and can therefore be generalized to species with complex life histories, where it becomes difficult to define a 'generation' and therefore fitness.

DATA AVAILABILITY

All analysis was performed in R, and our code is available on GitHub (www.github.com/josephmatheson/selective_deaths)

ACKNOWLEDGEMENTS

This work was supported by the John Templeton Foundation [62028] and [62220]. We thank Taylor Kessinger for sparking the last author's original interest in this topic many years ago by presenting Haldane's 1957 paper in journal club, and Jason Bertram and Ryan Gutenkunst for helpful discussions.

AUTHOR CONTRIBUTIONS

Conception of the paper by Joanna Masel and Moises Exposito-Alonso. Design and interpretation by Joseph Matheson and Joanna Masel. Data and details of experimental methodology provided by Moises Exposito-Alonso. Data analysis mostly by Joseph Matheson, with the MCMCglmm model analyzed by Moises Exposito-Alonso. Manuscript drafted by Joseph Matheson and substantively revised by all authors.

CONFLICT OF INTEREST

The authors declare no conflicting interests.

BIBLIOGRAPHY

Adami, Christoph. 2012. "The Use of Information Theory in Evolutionary Biology." *Annals of the New York Academy of Sciences* 1256(1):49–65.

Akçay, Erol, and Jeremy Van Cleve. 2016. "There Is No Fitness but Fitness, and the Lineage Is Its Bearer." *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1687).

Ariew, André, and R. C. Lewontin. 2004. "The Confusions of Fitness." *British Journal for the Philosophy of Science* 55(2):347–63.

Bell, Graham. 2017. "Evolutionary Rescue." *Annual Review of Ecology, Evolution, and Systematics*

48:605–27.

Bertram, J., K. Gomez, and J. Masel. 2017. “Predicting Patterns of Long-Term Adaptation and Extinction with Population Genetics.” *Evolution*.

Bertram, Jason. 2021. “Allele Frequency Divergence Reveals Ubiquitous Influence of Positive Selection in Drosophila.” *PLOS Genetics* 17(9):e1009833.

Bertram, Jason, and Joanna Masel. 2019. “Density-Dependent Selection and the Limits of Relative Fitness.” *Theoretical Population Biology* 129:81–92.

Buffalo, Vince, and Graham Coop. 2020. “Estimating the Genome-Wide Contribution of Selection to Temporal Allele Frequency Change.” *Proceedings of the National Academy of Sciences of the United States of America* 117(34):20672–80.

Crow, James F. 1958. “Some Possibilities for Measuring Selection Intensities in Man.” *Human Biology* 30(1):1–13.

Desai, Michael M., and Daniel S. Fisher. 2007. “Beneficial Mutation-Selection Balance and the Effect of Linkage on Positive Selection.” *Genetics* 176(3):1759–98.

Doebeli, Michael, Yaroslav Ispolatov, and Burt Simon. 2017. “Towards a Mechanistic Foundation of Evolutionary Theory.” *ELife* 6:1–17.

Ewens, W. J. 1970. “Remarks on the Substitutional Load.” *Theoretical Population Biology* 139:129–39.

Ewens, Warren. 2004. *Mathematical Population Genetics*. 2nd ed. New York: Springer US.

Exposito-alonso, Moises, Oliver Bossdorf, Rasmus Nielsen, and Detlef Weigel. 2019. “Natural Selection on the Arabidopsis Thaliana Genome in Present and Future Climates.” *Nature* 573(7772):126–29.

Felsenstein, Joseph. 1971. “On the Biological Significance of the Cost of Gene Substitution.” *American Naturalist* 105(941):1–11.

Fisher, Ronald. 1930. *The Genetical Theory of Natural Selection*. Oxford: Clarendon Press.

Galeota-Sprung, Benjamin, Paul Sniegowski, and Warren Ewens. 2020. “Mutational Load and the Functional Fraction of the Human Genome.” *Genome Biology and Evolution* 12(4):273–81.

Galtier, Nicolas. 2016. “Adaptive Protein Evolution in Animals and the Effective Population Size Hypothesis.” *PLoS Genetics* 12(1):1–23.

Graves, Christopher J., and Daniel M. Weinreich. 2017. “Variability in Fitness Effects Can Preclude Selection of the Fittest.” *Annual Review of Ecology, Evolution, and Systematics* 48:399–417.

Haldane, JBS. 1957. “The Cost of Natural Selection.” *Genetics* 55(3):511–24.

Hledík, Michal, Nick Barton, and Gašper Tkačik. 2022. “Accumulation and Maintenance of Information in Evolution.” *Proceedings of the National Academy of Sciences of the United States of America* 119(36):1–10.

Kelly, John. 2022. “The Genomic Scale of Fluctuating Selection in a Natural Plant Population.” *Evolution Letters* 506–21.

Kern, Andrew D., and Matthew W. Hahn. 2018. “The Neutral Theory in Light of Natural Selection.” *Molecular Biology and Evolution* 35(6):1366–71.

Kimura, Motoo. 1960. “Optimum Mutation Rate and Degree of Dominance as Determined by the

Principle of Minimum Genetic Load." *Genetics* 57(1):21–34.

Kimura, Motoo. 1961. "Natural Selection as the Process of Accumulating Genetic Information in Adaptive Evolution." *Genetical Research* 2(1):127–40.

Kimura, Motoo. 1968. "Evolutionary Rate at the Molecular Level." *Nature* 217:624–26.

Kimura, Motoo. 1969. "The Number of Heterozygous Nucleotide Sites Maintained in a Finite Population Due to Steady Flux of Mutations." *Genetics* 61(4):893–903.

Kimura, Motoo, and James F. Crow. 1969. "Natural Selection and Gene Substitution." *Genetical Research* 13(2):127–41.

Kimura, Motoo, and Tomoko Ohta. 1971. "On the Rate of Molecular Evolution." *Journal of Molecular Evolution* 1(1):1–17.

King, Oliver D., and Joanna Masel. 2007. "The Evolution of Bet-Hedging Adaptations to Rare Scenarios." *Theoretical Population Biology* 72(4):560–75.

Lehmann, Laurent, Charles Mullon, Erol Akçay, and Jeremy Van Cleve. 2016. "Invasion Fitness, Inclusive Fitness, and Reproductive Numbers in Heterogeneous Populations." *Evolution; International Journal of Organic Evolution* 70(8):1689–1702.

Machado, Heather E., Alan O. Bergland, Ryan Taylor, Susanne Tilk, Emily Behrman, Kelly Dyer, Daniel K. Fabian, Thomas Flatt, Josefa González, Talia L. Karasov, Bernard Kim, Iryna Kozeretska, Brian P. Lazzaro, Thomas J. S. Merritt, John E. Pool, Katherine O'brien, Subhash Rajpurohit, Paula R. Roy, Stephen W. Schaeffer, Svitlana Serga, Paul Schmidt, and Dmitri A. Petrov. 2021. "Broad Geographic Sampling Reveals the Shared Basis and Environmental Correlates of Seasonal Adaptation in Drosophila." *ELife* 10.

Maynard Smith, J. 1968. "'Haldane's Dilemma' and the Rate of Evolution." *Nature* 219(5159):1114–16.

Maynard Smith, J. 1976. "What Determines the Rate of Evolution?" *American Naturalist* 110(973):331–38.

McCandlish, David M., and Arlin Stoltzfus. 2014. "Modeling Evolution Using the Probability of Fixation: History and Implications." *Quarterly Review of Biology* 89(3):225–52.

McGee, Ryan Seamus, Olivia Kosterlitz, Artem Kaznatcheev, Benjamin Kerr, and Carl T. Bergstrom. 2022. "The Cost of Information Acquisition by Natural Selection." *Biorxiv* 2022.07.02.

Metz, J. A. J., S. A. H. Geritz, and R. M. Nisbet. 1992. "How Should We Define 'Fitness' for General Ecological Scenarios? The Simplest Textbook Model for Selection in Large Populations Relies N(t Fl) = B(E(t)) N(t) (1) with B(E(t)) a Matrix Whose El-Ements Depend on E(t), Implying N(t) = B(E(I-I)) B(E(t-2)). .." *Trends in Ecology & Evolution* 7(6):198–202.

Nei, Masatoshi. 1971. "Fertility Excess Necessary for Gene Substitution in Regulated Populations." *Genetics* 68:169–84.

Remine, Walter. 2006. "More Precise Calculations of the Cost of Substitution." *Creation Research Science Quarterly* 43(September).

Remine, Walter J. 2005. "Cost Theory and the Cost of Substitution — a Clarification." *Journal of Creation* 19(1):113–25.

Roff, Derek A. 2008. "Defining Fitness in Evolutionary Models." *Journal of Genetics* 87(4):339–48.

Rudman, Seth M., Sharon I. Greenblum, Subhash Rajpurohit, Nicolas J. Betancourt, Jinjoo Hanna, Susanne Tilk, Tuya Yokoyama, Dmitri A. Petrov, and Paul Schmidt. 2022. "Direct Observation of Adapting Tracking on Ecological Time Scales in Drosophila." *Science* 375(6586).

Sella, Guy, Dmitri A. Petrov, Molly Przeworski, and Peter Andolfatto. 2009. "Pervasive Natural Selection in the Drosophila Genome?" *PLoS Genetics* 5(6):e1000495.

Uricchio, Lawrence H., Dmitri A. Petrov, and David Enard. 2019. "Exploiting Selection at Linked Sites to Infer the Rate and Strength of Adaptation." *Nature Ecology & Evolution* 3(6):977–84.

Van Valen, Leigh. 1963. "Haldane's Dilemma, Evolutionary Rates, and Heterosis." *American Naturalist* 97(894):185–90.

Van Valen, Leigh. 1989. "Three Paradigms of Evolution." *Evolutionary Theory* 9(2):1–17.

Wilson, Robert A., and Matthew J. Barker. 2021. "Biological Individuals." in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta.