Article:
Yamamichi, M., Hairston, N.G., Rees, M. et al. (2019) Rapid evolution with generation overlap: the double-edged effect of dormancy. Theoretical Ecology, 12 (2). pp. 179-195. ISSN: 1874-1738

https://doi.org/10.1007/s12080-019-0414-7

Abstract

In life histories with generation overlap, selection that acts differently on different life-stages can produce reservoirs of genetic variation, for example in long-lived iteroparous adults or long-lived dormant propagules. Such reservoirs provide "migration from the past" to the current population, and depending on the trend of environmental change, they have the potential either to slow adaptive evolution or accelerate it by re-introducing genotypes not affected by recent selection (e.g., through storage effect in a fluctuating environment). That is, the effect of generation overlap is a "double-edged sword," with each edge cutting in a different direction. Here we use sexual (quantitative trait) and asexual (clonal) models to explore the effects of generation overlap on adaptive evolution in a fluctuating environment, either with or without a trend in the mean environment state. Our analyses show that when environmental stochasticity scaled by strength of selection is intermediate and when the trend in mean environment is slow, intermediate values of generation overlap can maximize the rate of response to selection and minimize the adaptation lag between the trait mean and the environmental trend. Otherwise, increased generation overlap results in smaller selection response and larger adaptation lag. In the former case, low generation overlap results in low heritable trait variance, while high generation overlap increases the "migration load" from the past. Therefore, to understand the importance of rapid evolution and eco-evolutionary dynamics in the wild for organisms with overlapping generations, we need to understand the interaction of generation overlap, environmental stochasticity, and strength of selection.

1 Introduction

Life histories that result in generation overlap – that is, iteroparous adults that reproduce at the same time as some of their offspring, or semelparous adults that produce propagules with prolonged dormancy that emerge asynchronously – can result in intriguing ecological and evolutionary dynamics. When natural selection acts differently on adult and immature stages, or differently on dormant and active individuals, subpopulations may possess distinct genotypic distributions. If the strength or direction of selection varies over time, the result is a reservoir of genetic variation in either the long-lived iteroparous adults or the long-lived dormant propagules. That reservoir is a "double-edged sword" with respect to adaptive evolution: it can slow down adaptive evolution, or speed it up by preserving or re-introducing genotypes adapted to current conditions that were not exposed to recent selection.

Here we explore the conditions under which each effect dominates the rate of evolution of a population with generation overlap in a varying environment with a possibly moving adaptive peak. Because our biological questions center on a population's ability to adapt to fluctuating or trending environmental conditions, we measure the speed of trait change in "clock time" rather

than on a per-generation basis. Our exploration was stimulated by and focuses on long-lived dormant propagules with selection acting on active stages, though our analyses and general conclusions should also apply to populations with iteroparous adults when selection acts mostly on juvenile stages. The production of long-lived dormant eggs, cysts, spores, or seeds is a widespread life-history trait in many types of organisms (**?**) including bacteria (**?**), fungi (**?**), algae (**?**), rotifers and inland-water crustaceans (**???**), and a wide range of plant species (**??**). With prolonged dormancy, only a fraction of the propagules produced in the previous growing season hatch (or germinate) in the next season, while the remaining propagules are added to a persistent dormant subpopulation in sediments or soil (i.e., a propagule, egg, seed, or spore bank), where they may remain viable for long periods of time (**??**). This observation, together with the fact that natural selection on these organisms often fluctuates in direction and magnitude (**?**), leads to counter-acting effects on the rate of evolution.

First, dormancy may slow down the speed of evolution for two reasons. Somewhat trivially, increased generation overlap increases generation time, and response to selection is slower when generation time is longer. More substantively, evolution under simple directional selection is slowed because the dormant fraction of the population is not exposed to contemporary selection pressures (**???**). Thus, evolutionary dynamics of the population as a whole can be slow despite strong selection due to introduction of maladapted individuals, depending on the proportion of individuals in dormant versus active stages, the mean time an egg or seed spends in dormancy, and the stages upon which selection is acting (i.e., on the dormant stages or on the active stages). This is one edge of the double-edged sword, and is in essence the effect of migration load slowing down evolution (**??**) where migration is in time rather than across space.

Second, dormancy can enable genotypes to survive disadvantageous times and then to "migrate from the past" and influence the current evolutionary response. When the "migrants from the past" are adapted to the current conditions, for example, in stationary but fluctuating environments, they facilitate contemporary evolution. This can be very important for population persistence under environmental change, including through evolutionary rescue, in which extinction is prevented by rapid adaptive evolution (**????**). This is the other edge of the sword. This "second edge" can be especially effective when temporally fluctuating selection promotes the maintenance of genetic variation (e.g., a temporal storage effect: **???**).

Several previous studies have analyzed the "moving optimum" model for directional evolution when the optimum trait value is continuously moving in one direction, and the population trait mean tracks the trend with an "adaptation lag" (reviewed in **?**). If the optimum trait moves at a constant rate (i.e., a linear increase or decrease), the lag stays constant after a period of transient dynamics. Various factors determine the size of the lag between the optimal and mean trait values, which in turn can determine whether the population goes extinct or is rescued by evolution (**??**). The moving optimum model can incorporate environmental stochasticity in addition to the directional change, but most of the above studies did not consider generation overlap. As

a notable exception, **?** considered evolution of quantitative traits in stage-structured populations, and more recently, **?** examined adaptation lag and evolutionary rescue in structured populations. They found that temporal genetic migration via reproduction by iteroparous long-lived adults slowed the rate of evolution, increasing the adaptation lag. However, they did not allow fluctuating selection and so did not consider the potentially beneficial effects of migration from the past of genotypes adapted to the current environment

To understand how these factors interact and affect the speed of evolution, we use a simple asexual (clonal) population model developed to describe annual plants with dormant seeds (**?**) or zooplankton with dormant eggs (**?**) as well as a sexual (quantitative genetics) version of the model and explore the conditions under which generation overlap caused by prolonged dormancy promotes or inhibits rapid evolution. An important difference between the two models is the potential for evolutionary branching, which creates bimodal trait distributions: it is likely in the asexual model when environmental variance and generation overlap are high (**???**), but unlikely in the sexual model due to continual production of intermediate phenotypes by random mating (**?**). We consider both stationary environments, where conditions in each year vary randomly around an unchanging mean, and an environment with a steady directional trend superimposed on year-to-year random variation, where populations must evolve to track a moving optimum phenotype. We find that when environmentally-driven fluctuations scaled by strength of selection are intermediate, and when the trend in mean environment is slow enough, intermediate values of generation overlap can maximize the selection response and minimize the adaptation lag between the trait mean and the environmental trend. Otherwise, increased generation overlap results in a smaller selection response and larger adaptation lag. This suggests that understanding the interactions between generation overlap, environmental stochasticity, and strength of selection will be important to predict rapid evolution for organisms in the wild.

2 Models

We constructed and studied two evolutionary models: a clonal model which explicitly tracks the frequency of different asexual genotypes for a trait subject to fluctuating selection, and a quantitative trait (or "quantitative genetics") model in which diploid individuals are characterized by their breeding value for the trait, and trait inheritance follows the standard infinitesimal model (**???**). Model simulations were done in R (**?**) (see Supplementary Information for R scripts to replicate all figures in this paper).

2.1 Clonal model

We employed one of the simplest models for evolutionary dynamics with generation overlap (**?**). The population is assumed to consist of a set of haploid lineages (clones), which breed true

for the trait under selection, apart from effect of mutation (described below). Following **?**, the population dynamics (without mutation) are described by

$$X_i\left(t+1\right) = X_i\left(t\right)\left[HY_i\left(t\right)+\gamma\right], \tag{1}$$

where $X_i(t)$ is the abundance of type i individuals in year t, $Y_i(t)$ is their per-capita fecundity, and H (for "hatching") is the fraction of the population that engages in reproduction ($0 < H \leq 1$). The average survivorship across dormant and active individuals is

$$\gamma = Hs_a + \left(1-H\right)s_d, \tag{2}$$

where $1-H$ is the fraction that stays dormant, s_a is survival fraction in the active stage, and s_d is survival fraction in the dormant stage ($0 \leq s_a, s_d \leq 1$). γ is also the amount of generation overlap at steady state, where the mean population size is constant over time. We assume $s_a < s_d$, i.e., there is a survival benefit to being dormant, because if there were no benefit, there would be no reason for dormancy to evolve. Our simulations assume $s_a \ll 1$ so that most generation overlap is through dormancy. However, evolutionary dynamics in our models is determined by the overall value of γ, because we assume that selection affects recruitment and so the trait distribution is identical (apart from sampling variability) in active individuals and those entering dormancy.

Model simulations tracked the frequency of G genotypes with evenly spaced trait values ($z_{i+1} - z_i \equiv \Delta z$). Reproduction included mutation to each neighboring genotype with probability $m \ll 1$. With mutation, the dynamics become

$$\vec{X}(t+1) = \mathbf{U}[H\vec{X}(t) \circ \vec{Y}(t)] + \gamma\vec{X}(t), \tag{3}$$

where $\vec{X} = (X_1, X_2, \cdots, X_G), \vec{Y} = (Y_1, Y_2, \cdots, Y_G)$, "$\circ$" denotes the element-by-element product, and

$$\mathbf{U} = \begin{pmatrix} 1-m & m & 0 & ... & ... & 0 \\ m & 1-2m & m & 0 & ... & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & ... & 0 & m & 1-2m & m \\ 0 & ... & ... & 0 & m & 1-m \end{pmatrix}. \tag{4}$$

For per capita fecundity we assume the "saturating yield" model (**?**),

$$Y_i(t) = \frac{KR(z_i - M_t)}{\sum_j HX_j(t)R(z_j - M_t)}, \tag{5}$$

where z_i is the trait value of type i individuals, M_t is the optimal phenotype in year t, and R is

the relative fecundity under stabilizing selection,

$$R(z_i - M_t) = \exp\left[-\frac{(z_i - M_t)^2}{2\sigma_w^2}\right]. \tag{6}$$

The selection variance σ_w^2 determines the intensity of selection. The mean and variance of the temporally varying optimal phenotype, M_t, are μ_t and σ_M^2, respectively. We consider the "stationary environment" case where μ_t is constant as well as the "moving trait optimum" case where μ_t is changing through time: $\mu_t = \eta(t - t_0)$ where η and t_0 are the speed and onset of the environmental change, respectively (see **?** for review of moving optimum models). We assume that fluctuations in M_t is symmetric around μ_t, without temporal autocorrelation. Note that the total offspring production $\sum_i HX_i(t)Y_i(t)$ equals K each year, regardless of the population state or the value of M_t. Because of this, the total population size $X(t) = \sum_i X_i(t)$ satisfies $X(t) = K + \gamma X(t-1)$ and therefore converges to the stable equilibrium $\bar{X} = K/(1-\gamma)$. Once that equilibrium is reached the population each year consists of new recruits and survivors in constant proportions $(1-\gamma) : \gamma$.

2.2 Quantitative trait model

As an opposite extreme to the clonal model, we also consider a diploid, hermaphroditic (i.e., cosexual) population (e.g., a hermaphroditic plant with a persistent dormant seed bank), in which the focal trait z is a quantitative trait with transmission dynamics described by the infinitesimal model from quantitative genetics (**?**). Various versions of the infinitesimal model are based on a set of nested approximations (**?**). The central one is that the trait is determined by very many loci, with allele effects that are small and additive within and across loci. The "Gaussian descendants" approximation adds that the pre-selection breeding values of offspring from a mating are drawn from a (potentially multivariate) Gaussian distribution with mean equal to the mean parental breeding value and a variance-covariance structure that only depends on the relatedness of the parents, regardless of whether selection, mutation, or migration are or have been occurring. Recent theory has validated the Gaussian descendents approximation, showing that its error decreases to 0 as the number of loci increases when the loci are unlinked (**?**). The "Gaussian population" approximation further states that individuals on a pedigree have breeding values and phenotypes with a multivariate Gaussian distribution whose variance-covariance matrix depends only on the pedigree; this is the most questionable approximation because selection typically produces non-Gaussian trait distributions, but it is often fairly accurate even with strong selection for populations with non-overlapping generations (**?**). Our individual-based simulations, described below, use the Gaussian descendents approximation (which makes it sufficient to track the breeding value of each individual, rather than their genotype), but not the Gaussian population approximation. Some of our model analyses also use the Gaussian population approximation.

We assume that individuals differ in heritable trait z but otherwise the population is unstructured. For simplicity we assume that the trait value equals the breeding value for the trait, i.e., that the "environmental" component of the phenotype is zero. This is not much of a restriction, as a breeding value-dependent fitness function can be defined by integrating the trait-dependent fitness function over the conditional trait distribution given breeding value. When selection and the conditional trait distribution are both Gaussian, the result is Gaussian selection on breeding value (with a wider kernel, implying weaker selection).

We implemented the model as an individual-based simulation with a finite number of individuals. At the start of each discrete time step, each individual is assigned to be reproductively active in that time step with probability H, and otherwise stays dormant. K new recruits are produced at each time step. The father (i.e., male zygote contributor) and mother (female zygote contributor) of each recruit are chosen at random, with replacement, from the reproductively active individuals, with each individual's probability of being chosen (for either role) proportional to its trait-dependent fitness $R(z_i - M_t)$. As parents are unrelated, the Gaussian descendents approximation specifies that each offspring's breeding value is drawn from a Gaussian distribution with mean equal to the average parent breeding value, and variance V_0 (the "segregation variance") that is constant across offspring and across time, despite the action of selection.

Dormant and active individuals survive to the next time-step with probabilities s_d and s_a, respectively. Survivors and new recruits are then combined to form the starting population for the next time step. In analyses below where we compare outcomes across different H values, K varies with H so that the expected total population $\bar{X} = K/(1-\gamma)$ is constant and the degree of demographic stochasticity does not vary across different values of H.

2.3 Scaling

Without loss of generality, the origin for the trait z is chosen so that $\mathbb{E}[M_t] = 0$ either for all time (a stationary environment) or prior to the onset of a directional trend. We could also make $\sigma_M^2 = 1$, or $\sigma_w^2 = 1$, or (in the quantitative trait model) $V_0 = 1$, by scaling z relative to σ_M, σ_w, or $\sqrt{V_0}$. We choose not to do this, because it is simpler (for example) to ask what happens as $\sigma_w^2 \to \infty$ (weak selection) instead of asking what happens as σ_M^2 and V_0 converge to 0 such that σ_M^2/V_0 remains constant, and easier to interpret the results.

3 Analysis and Results: Quantitative Trait Model

3.1 Stationary Environment

We consider first the situation where the mean environment $\mathbb{E}(M_t) \equiv 0$, and ask what effect dormancy has on the population's heritable trait variation and short-term responses to fluctuat-

ing selection. Let $\bar{z}(t)$ and $\sigma_z^2(t)$ denote the trait mean and variance in year t, for the combined population of survivors and new recruits. As in the clonal model, apart from small fluctuations due to finite population size, the total population size converges to $\bar{X} = K/(1-\gamma)$ consisting (before selection) of survivors and new recruits in proportions $\gamma : (1-\gamma)$, and we assume that this has already occurred.

Assuming a Gaussian trait distribution before selection, under Gaussian selection the trait mean in selected parents (and therefore in their offspring), $\bar{z}_r$, is (**?**, ch. 9)

$$\bar{z}_r(t+1) = \bar{z}(t) + \frac{\sigma_z^2(t)}{\sigma_w^2 + \sigma_z^2(t)}[M_t - \bar{z}(t)]. \tag{7}$$

Then, combining new recruits with survivors (whose trait mean is still $\bar{z}(t)$) and defining

$$a(t) = \frac{\sigma_z^2(t)}{\sigma_w^2 + \sigma_z^2(t)}, \tag{8}$$

the trait mean in the next year is

$$\bar{z}(t+1) = \gamma\bar{z}(t) + (1-\gamma)\bar{z}_r(t+1) = \bar{z}(t) + (1-\gamma)\, a(t)\, [M_t - \bar{z}(t)]\,. \tag{9}$$

This shows one edge of the sword: all else being equal, with higher γ the response to selection

$$\mathcal{R} = \bar{z}(t+1) - \bar{z}(t) = (1-\gamma)\, a(t)\, [M_t - \bar{z}(t)] \tag{10}$$

becomes smaller in magnitude.

The other edge is how generation overlap affects the trait variance. Assuming the set of potential parents has a Gaussian trait distribution with variance $\sigma_z^2(t)$, under Gaussian selection the trait variance in selected parents (i.e., the variance of the parental trait distribution weighted by relative fitness) is (**?**)

$$\frac{\sigma_w^2\sigma_z^2(t)}{\sigma_w^2 + \sigma_z^2(t)} = \sigma_z^2(t)[1 - a(t)]. \tag{11}$$

Note that this does not depend on M_t, because the curvature of log-fitness is constant under Gaussian selection. Offspring traits in the infinitesimal model are the average of their parents' traits, plus an independent effect of segregation variance, so the trait variance in new recruits is

$$\sigma_r^2(t+1) = \frac{1 - a(t)}{2}\sigma_z^2(t) + V_0. \tag{12}$$

To compute $\sigma_z^2(t+1)$ we use the "law of total variance" (if X and Y are random variables, then $\mathrm{Var}(X) = \mathbb{E}\left[\mathrm{Var}(X|Y)\right] + \mathrm{Var}\left[\mathbb{E}(X|Y)\right]$), with $X = z$ and $Y = 0$ or 1 to indicate if the individual is

a recruit or survivor. Then we have

$$\begin{aligned}\sigma_z^2(t+1) &= \gamma\sigma_z^2(t) + (1-\gamma)\sigma_r^2(t+1) + \gamma(1-\gamma)[\bar{z}_r\,(t+1) - \bar{z}\,(t)]^2 \\ &= \gamma\sigma_z^2(t) + (1-\gamma)\left[\frac{1-a(t)}{2}\sigma_z^2(t) + V_0\right] + \gamma(1-\gamma)a(t)^2\,[M_t - \bar{z}(t)]^2\end{aligned} \tag{13}$$

Eqn. (**??**) is nonlinear and depends on the random variable M_t so we cannot analyze it exactly, but some simple approximations are informative. First, we can take expectations of both sides conditional on the current value of $\sigma_z(t)$. Assume (for the rest of this section) that fluctuations in M_t are independent over time and symmetric. The model is then symmetric around 0 hence $\mathbb{E}[\bar{z}(t)] = 0$ and $\mathbb{E}\left[(M_t - \bar{z})^2\right] = \sigma_M^2 + \mathrm{Var}(\bar{z})$. Making the approximation that $a(t)$ remains constant at its mean value $\bar{a}$ in eqn. (**??**) and taking the variance of both sides (using the fact that M_t is independent of the trait distribution), the stationary value of $\mathrm{Var}(\bar{z})$ is found to be $\sigma_M^2(1-\gamma)\bar{a}/[2-(1-\gamma)\bar{a}]$ and therefore

$$\mathbb{E}\left[(M_t - \bar{z})^2\right] = \sigma_M^2 + \mathrm{Var}(\bar{z}) \;\approx\; \frac{2\sigma_M^2}{2-(1-\gamma)\bar{a}}. \tag{14}$$

Substituting this into (**??**) we have (approximately)

$$\sigma_z^2\,(t+1) = \gamma\sigma_z^2(t) + (1-\gamma)\left[\frac{1-\bar{a}}{2}\sigma_z^2(t) + V_0\right] + \frac{2\gamma(1-\gamma)\bar{a}^2\sigma_M^2}{2-(1-\gamma)\bar{a}} \tag{15}$$

To approximate the long-term behavior we drop the time-dependence in (**??**) and solve for σ_z^2, getting

$$(1+\bar{a})\bar{\sigma}_z^2 = 2V_0 + \frac{4\gamma\bar{a}^2\sigma_M^2}{2-(1-\gamma)\bar{a}}. \tag{16}$$

With the further (rough) approximation $\bar{a} = \bar{\sigma}_z^2/(\bar{\sigma}_z^2 + \sigma_w^2)$, (**??**) is an approximate implicit equation for $\bar{\sigma}_z^2$.

Because $\bar{a}$ depends nonlinearly on $\bar{\sigma}_z^2$, in general eqn. (**??**) can only be evaluated numerically. However, two limiting cases are tractable. In the limit of very weak selection ($\sigma_w \to \infty$), $a(t) \to 0$ and the solution to (**??**) converges to the correct value for a neutral trait, $\sigma_z^2 = 2V_0$. In simulations (below) we observe that (**??**) loses accuracy as selection becomes stronger, but surprisingly it regains accuracy under very strong selection ($\sigma_w^2 \to 0$). In that limit $a(t) \to 1$, so the solution of (**??**) is

$$\bar{\sigma}_z^2 = V_0 + \left(\frac{2\gamma}{1+\gamma}\right)\sigma_M^2. \tag{17}$$

In the strong selection limit, all selected parents in year t have trait values very near M_t. This simplification allows an exact analysis (see Appendix **??**) which confirms that (**??**) is the long-term average trait variance. So under very strong selection, the steady-state average trait variance is a monotonically increasing function of generation overlap γ. Genetic variation in this case is composed of variation generated in the current year among offspring as a result of segregation,

V_0, and "migration from the past." This second term depends on the amount of generation overlap, γ and the variability in the environment, σ_M^2. Without variation in the environment, migration from the past does not introduce new genotypes and so cannot increase σ_z^2.

Numerical solutions of (**??**) with randomly generated values of σ_w and V_0 (scaling z so that $\sigma_M = 1$) confirm that the solution is an increasing function of γ: more generation overlap is predicted to result in higher steady-state genetic variance, as we expect (R script `IterateApproximateVz.R`).

Compared with simulations, (**??**) is most accurate when selection is weak (σ_w is large) and loses accuracy as selection becomes stronger (Fig. **??**A); as noted above, accuracy is regained for extremely strong selection. It is also accurate for moderately strong selection when generation overlap is very low (see Fig. **??**). This pattern is related to the fact that several steps in deriving (**??**) assume a Gaussian trait distribution at the start of each time-step. Without generation overlap, trait distributions generally remain close to Gaussian even under strong selection (**?**), so (**??**) is still accurate. But with substantial generation overlap, severely non-Gaussian distributions can result from large differences between survivor and recruit trait means. The population distribution can even be bimodal after years when $M_t - \bar{z}(t)$ is large relative to σ_w.

Eqn. (**??**) typically over-estimates the average steady-state trait variance in simulations with high generation overlap (e.g., Fig. **??**A). This occurs because (**??**) over-estimates the selection response conditional on the trait variance (e.g., Fig. **??**B), which results in an over-estimate of the trait variance generated by recruit-survivor differences (the final term in (**??**)). To see why this happens, suppose that the trait distribution at the start of the time step can be approximated as a mixture of two Gaussians (new recruits and survivors), both on the same side of M_t. Each component of the mixture will respond to selection with its own trait variance, which is smaller than the population's overall trait variance. Hence the selection responses of both components (and so the overall response) are smaller than in the Gaussian approximations leading to (**??**).

We can now ask how generation overlap affects the magnitude of response to selection. In a stationary environment the selection response $\mathcal{R}$ has mean 0, so its standard deviation $\sigma_{\mathcal{R}}$ is the RMS (root mean square) response magnitude. From eqns. (**??**) and (**??**) we have

$$\sigma_{\mathcal{R}} = (1-\gamma)\bar{a}\sqrt{\mathbb{E}\left\{[M_t - \bar{z}(t)]^2\right\}} = (1-\gamma)\bar{a}\sigma_M\sqrt{\frac{2}{2-(1-\gamma)\bar{a}}}. \tag{18}$$

The RMS response $\sigma_{\mathcal{R}}$ falls to 0 as $\gamma \to 1$ (with little recruitment there is little response). In the strong selection limit ($\sigma_w^2 \to 0, a(t) \to 1$), eqn. (**??**) implies that $\sigma_{\mathcal{R}}$ is proportional to $(1-\gamma)/\sqrt{1+\gamma}$ so the selection response is a steadily decreasing function of generation overlap. In this limit, the response to selection has become independent of σ_z^2 because all selected parents and therefore all recruits have trait values near M_t.

In the weak-selection limit ($\sigma_w^2 \to \infty, a(t) \to 0$), the stationary trait variance is $2V_0$ indepen-

dent of γ (e.g., Fig. **??**A, blue and magenta curves). Again the response to selection is always slowed by greater generation overlap, though for a different reason – "migration from the past" in this case does not introduce genotypes different from those already present.

These two limiting results suggest that generation overlap will always slow the response to selection, despite its potentially positive effect on trait variance. Surprisingly, that is not true: under moderately strong selection, the selection response can be maximized at intermediate generation overlap (e.g., Fig. **??**B).

To understand when this occurs, because (**??**) is accurate when γ is small we can use it to ask when $\sigma_{\mathcal{R}}$ is an increasing function of γ for $\gamma \approx 0$; this will imply that the RMS response to selection is largest at intermediate generation overlap. $\sigma_{\mathcal{R}}$ is proportional to $(1-\gamma)\bar{a}\sqrt{\sigma_M^2 + \mathrm{Var}(\bar{z})}$ (eqn. **??**) where $\bar{a}$ and $\mathrm{Var}(\bar{z})$ depend on γ. Using the approximations above for $\bar{\sigma}_z^2$ and $\mathrm{Var}(\bar{z})$, some straightforward calculus (done using Maxima in `QG-IBM-noTrend.max`, see Appendix **??**) shows that the condition for this product to be increasing at $\gamma = 0$ can be expressed in the form $\sigma_M/\sigma_w > \sqrt{F(\sigma_w^2/V_0)}$, where F is given by eqn. (**??**). Over a wide range in selection strength ($0.6\sqrt{V_0} < \sigma_w < 7\sqrt{V_0}$, which goes from weak to moderately strong selection across the trait scale of segregation variance $2\sqrt{V_0}$), $\sqrt{F}$ does not vary greatly ($2 < \sqrt{F} < 4$), so the response to selection is maximized at intermediate γ if $\sigma_M > (2 \text{ to } 4) \times \sigma_w$. Such parameters imply that in a typical year, any individuals with traits near M_t will have much higher fitness than those with traits near the long-term trait mean $\bar{z} = 0$.

The overall qualitative pattern, then, is that average selection response is largest at intermediate levels of generation overlap when there is sufficiently high environmental variability and sufficiently strong (but not too strong) selection. When selection is very weak or very strong, or environmental variability is low, generation overlap always slows the response to selection. A second general pattern is that moment equations that assume a Gaussian trait distribution become inaccurate when there is nontrivial generation overlap, and generally over-estimate the response to selection.

3.2 Moving trait optimum

We now consider a population experiencing a directional trend in the optimal phenotype, and ask how generation overlap affects the population's ability to track the trend. Specifically, we assume that $\mu_t = \mathbb{E}(M_t)$ equals 0 up to time t_0, and thereafter it increases at a constant rate, $\mu_t = \eta(t - t_0)$, as in **?**. Following a transient period, trait evolution "reaches a steady state where the mean phenotype lags a constant distance behind the changing optimum phenotype" (**?**, p. 3; we observe the same in our simulations). **?** analyzed a model with discrete, nonoverlapping generations (and phenotypic plasticity). Here we ask how the lag L between the environmental trend and the trait mean ($L = \mu_t - \bar{z}(t)$) depends on the generation overlap γ.

At steady state the change in trait mean is exactly keeping pace with the environment trend:

$\mathbb{E}\left[\bar{z}(t+1) - \bar{z}(t)\right] = \eta$. Hence from eqn. (**??**), assuming a Gaussian trait distribution the average lag $\bar{L}$ satisfies

$$\begin{aligned} \eta &= (1-\gamma)\mathbb{E}\left\{a(t)[M_t - \bar{z}(t)]\right\} \approx (1-\gamma)\mathbb{E}\left[a(t)\right]\mathbb{E}\left[M_t - \bar{z}(t)\right] \\ &= (1-\gamma)\mathbb{E}\left[a(t)\right]\left[\mu_t - \bar{z}(t)\right] = (1-\gamma)\bar{a}\bar{L} \end{aligned} \tag{19}$$

the "$\approx$" being that correlation between $a(t)$ and $(M_t - \bar{z}(t))$ is ignored. Thus

$$\bar{L} \approx \frac{\eta}{(1-\gamma)\bar{a}}. \tag{20}$$

?, eqn. 8a previously derived this result for the case $\gamma = 0$ (no generation overlap). So as was true for selection response $\mathcal{R}$ in a stationary environment, the relationship between $\bar{L}$ and γ depends on how strongly γ affects $\bar{a}$ through its effect on $\bar{\sigma}_z^2$.

The strong and weak selection limits are again both tractable. In the strong selection limit ($\sigma_w^2 \to 0$), $\bar{a} \to 1$ so eqn. (**??**) predicts that $\bar{L} \to \eta/(1-\gamma)$. An exact analysis of the strong selection limit (see Appendix **??**) confirms that $\bar{L} = \eta/(1-\gamma)$ and therefore higher generation overlap always increases the adaptation lag. In the weak-selection limit, $\bar{a} \to 0$ as $\sigma_w^2 \to \infty$, and the extension of (**??**) to the case of a moving trait optimum (see eqn. (**??**) in Appendix **??**) reduces to

$$\bar{\sigma}_z^2 = 2V_0 + \frac{2\eta^2\gamma}{(1-\gamma)^2}. \tag{21}$$

Using (**??**) to approximate $\bar{a}$ in (**??**) gives an approximate adaptation lag. For $V_0 > \eta^2$ (which is required for the population to persist by adapting to the trend), this approximation predicts that the adaptation lag is an initially increasing function of γ, but drops as $\gamma \to 1$ because $\bar{\sigma}_z^2$ diverges to $+\infty$ (which cannot really be true).

Figures **??**A-D show that simulations align well with these asymptotic results, except for weak selection with very high generation overlap. Because the increase of $\bar{\sigma}_z^2$ as $\gamma \to 1$ is over-predicted (by eqn. (**??**)), the adaptation lag is under-predicted. The inaccuracy of eqn. (**??**) is again a result of non-Gaussian trait distributions. The final term in (**??**) includes the contribution to $\sigma_z^2(t)$ from the trait difference between recruits and survivors, which (as discussed above) is smaller than predicted when the trait distribution is non-Gaussian. As a result, simulations consistently do not exhibit the predicted decrease in adaptation lag when γ is very near 1. As was the case for selection response $\mathcal{R}$, with very weak or very strong selection the restraining effect of generation overlap is always dominant.

And again, at intermediate selection the situation can be different. Consider first a situation where η is very small. The population then tracks the trend closely at all times, and the standing variation $\sigma_z^2(t)$ and trait distribution (re-centered at μ_t) are close to what they would be in a stationary environment. The denominator on the right-hand side of (**??**) is the inverse of the coefficient determining the selection response $\mathcal{R}$ in a stationary environment (eqn. **??**). The

circumstances where $\mathcal{R}$ is maximized at intermediate γ will then be circumstances where $\bar{L}$ is mimimized at intermediate γ for small η, discussed above. Fig. **??**E,F shows an example, a slow trend ($\eta = 0.01$) starting at $t = 1000$ with the same parameters as the dotted (red) curve **??**B: the minimal lag (at $\gamma \approx 0.4$) aligns with the largest short-term responses in a stationary environment. However, the approximation (**??**) is poor, except at small γ, because selection is strong enough that the trait distribution is often bimodal. Each year's selection response is then smaller than predicted (as discussed above) so the lag is larger than predicted.

The transient period following the start of a trend will initially reflect the population state in the previously stationary environment. Fig. **??**A is an expanded view of the initial response to a fairly rapid trend ($\eta = 0.05$/yr) starting in year 500 for the same parameters as the dotted (red) curve in Fig. **??**B. Each curve in panel A) is the average adaptation lag across 500 replicate simulations (γ values shown in the figure legend). Initially the greatest lags are at low overlap, with smaller lags at intermediate and high generation overlap where the trait variation was highest before the trend. But after 100 years a different long-term pattern has emerged, in which higher generation overlap implies greater lag.

Fig. **??**B illustrates another challenge to a general analysis of adaptation lag under intermediate selection: approximating how γ affects $\bar{\sigma}_z^2$. The extension of (**??**) for $\eta > 0$ (Appendix **??**) and the resulting approximate $\bar{\sigma}_z^2$ (eqns. **??** and **??**) predict that trait variance is larger when $\eta > 0$, and our simulations (such as the ones plotted in Fig. **??**A, B) confirm this for weak selection. But with stronger selection, $\sigma_z^2(t)$ often actually decreases substantially after the onset of a directional trend (Fig. **??**B).

The erosion of trait variance at high γ is again related to non-Gaussian trait distributions. When $\eta > 0$ the population trait mean gradually moves upward, leaving behind a tail of individuals who have been dormant for a while and are maladapted when they become active. The trait distribution therefore develops negative skew when there is high generation overlap (Fig. **??**C). The distribution also develops a very fat tail (Fig. **??**D; a Gaussian has zero kurtosis). These departures from Gaussian are enormously larger than what can occur under constant directional selection on a quantitative trait (standardized skew<0.05 and kurtosis<0.01, **?**). Large negative skew combined with positive directional selection can produce a negative value of the quadratic selection differential C, leading to a decrease in trait variance (see **?**, eqn. 23 and Appendix **??**). In terms of our approximations, the effect of negative skew is that the trait variance in selected parents is lower than the value (**??**) that would occur with a Gaussian trait distribution (see Fig. **??**) because the wide left tail is chopped off by selection.

The reduced trait variance flattens the positive relationship between γ and $\bar{\sigma}_z^2$, so that the relationship between γ and adaptation lag is dominated by the $(1-\gamma)$ factor in (**??**). Individual-based simulations with many randomly-generated parameter sets (described in Appendix **??**) confirm the generality of the pattern in Fig. **??**: except when the trend is slow enough to track closely, the adaptation lag is an increasing function of γ, all else being equal.

3.3 A triple-edged sword? Evolutionary modification of segregation variance

Heritable variation among offspring from a mating, measured by the segregation variance V_0, injects new heritable variation into the population each year. In the infinitesimal model the value of V_0 results from mutation-drift balance (**?**, Chapter 16). New mutations generate heritable variation; drift (in this context) is new allelic identity by descent resulting from random mating in a finite population, which reduces heritable variation. Thus V_0 is an "exogenous" parameter in the infinitesimal model, affecting evolution but not affected by it. However, the value of V_0 is actually a potential target of selection through evolution of the mutation rate, providing yet another mechanism whereby generation overlap and fluctuating selection can promote heritable variation (**?**).

To explore this possibility in a simple model, we assume that segregation variance itself is a quantitative trait (specifically, offspring inherit the average of their two parents' V_0 values, plus lognormal "segregation variance" with standard deviation 0.05 on log scale). This model is mechanistically indefensible. However, as an add-on to a model for flowering strategy evolution it exhibited the same long-term behavior as a mechanistic model where the mutation rate at loci affecting z is evolvable (**?**, Appendix H), probably because the averaging of parental V_0 values approximates the statistical outcome of matings between individuals with different mutation rate genotypes.

Given enough time, evolution of V_0 can substantially flatten the relationship between generation overlap and adaptation lag (Fig. **??**A,B). With higher generation overlap and higher initial lag, there is increasing selection for larger V_0 (Fig. **??**C) which allows the population to better track the environmental trend (Fig. **??**D).

As in **?** the time scale of V_0 evolution is very long, much longer than a steady environment trend could really continue. In this simulation the slow evolution is in part due to weak selection ($\sigma_w = 2$), but even with strong selection V_0 adapts very slowly except when generation overlap is low and adaptation lag is small (Fig. **??**). At higher generation overlap the adaptation lag is larger, but nontrivial changes in V_0 require thousands of generations (or even more when the environment is stationary, Fig. **??**). Moreover, these simulations assume that changes in V_0 are cost-free, whereas an increased mutation rate would probably result in more uniformly deleterious mutations. Because of all these caveats, we view this section mainly as a cautionary reminder that V_0 (and mutation rate in our clonal model) are not really "exogenous" parameters. They are likely to be products of natural selection, with the benefits of elevated mutation rate depending on the amounts of environmental variability and generation overlap.

4 Clonal Model

In a stationary environment, the clonal model's trait distribution is strongly affected by the evolutionary branching that occurs when $\gamma\sigma_M^2$ exceeds σ_w^2 (**???**). When generation overlap γ and/or environmental variance σ_M^2 are low, there is a monomorphic ESS (evolutionarily stable strategy) trait value, equal to the mean trait optimum. In our simulations, mutations blur the ESS into a trait distribution with positive (but small) variance which can respond to selection. As γ and σ_M^2 increase the ESS branches into an ESC (evolutionarily stable combination; **??**) consisting of two widely separated phenotypes (and possibly more as γ and σ_M^2 become larger), symmetric about the mean trait optimum (because selection and the distribution of M_t are both symmetric, by assumption). These two situations need to be analyzed separately. Because the evolutionary branching is driven by σ_M^2 and γ, we study the clonal model using these (rather than σ_w and γ) as our "control parameters".

4.1 Analysis: one-phenotype trait distribution

When the trait distribution has small variance, the response to selection in model (**??**) can be approximated as in the Appendix of **?**,

$$\mathcal{R} = \bar{z}(t+1) - \bar{z}(t) = \sigma_z^2(t)\frac{1}{\overline{W}_t}\left.\frac{\partial W\left(z_i,\bar{z}(t),t\right)}{\partial z_i}\right|_{z_i=\bar{z}(t)} \tag{22}$$

where $W\left(z_i,\bar{z},t\right)$ is the fitness of a trait-value z_i individual in a population with trait mean $\bar{z}$ in year t, and $\overline{W}_t$ is population mean fitness in year t. This is eqn. (A6) in **?** except that we use $\overline{W}_t$ instead of approximating it by $W(\bar{z}(t),\bar{z}(t),t)$.

With the saturating yield function for recruitment, $\overline{W}_t \equiv 1$ once the total population size has converged to $\bar{X}$. In model (**??**) we have

$$W\left(z_i,\bar{z}(t),t\right) = \frac{X_i(t+1)}{X_i(t)} = HY\left(z_i,\bar{z}(t),t\right) + \gamma. \tag{23}$$

The per capita fecundity Y (eqn. **??**) depends on the entire trait distribution, but as a first approximation, we can compute the denominator if all individuals have the mean trait value. This gives (for total population at $\bar{X}$)

$$Y\left(z_i,\bar{z}(t),t\right) \approx \frac{KR\left(z_i - M_t\right)}{H\bar{X}R\left(\bar{z}(t) - M_t\right)} = \left(\frac{1-\gamma}{H}\right)\frac{R\left(z_i - M_t\right)}{R\left(\bar{z}(t) - M_t\right)}. \tag{24}$$

Substituting (**??**) and (**??**) into (**??**), we have

$$\mathcal{R} \approx \sigma_z^2(t)\left(1-\gamma\right)\left.\frac{R'\left(z_i - M_t\right)}{R\left(\bar{z}(t) - M_t\right)}\right|_{z_i=\bar{z}(t)} = \frac{\sigma_z^2(t)\left(1-\gamma\right)}{\sigma_w^2}\left[M_t - \bar{z}(t)\right]. \tag{25}$$

This is in effect a small variance (equivalently, weak selection) approximation to eqn. (**??**), dropping $\sigma_z^2(t)$ from the denominator of $a(t)$. As in the quantitative trait model, we see the "double-edge": higher generation overlap decreases $(1-\gamma)$ but often increases $\sigma_z^2(t)$, so it can be the case that intermediate values of H and γ maximize the response rate.

The variance dynamics in a stationary environment ($\mathbb{E}(M_t) \equiv 0$) can be approximated as in the quantitative trait model, giving

$$\sigma_z^2(t+1) \approx \gamma\sigma_z^2(t) + (1-\gamma)\left[\sigma_z^2(t)\left(1 - \frac{\sigma_z^2(t)}{\sigma_w^2}\right) + V_0\right] + \gamma(1-\gamma)\frac{\sigma_z^4(t)}{\sigma_w^4}\left[M_t - \bar{z}(t)\right]^2 \tag{26}$$

where V_0 is the offspring trait variance due to mutation (mutation rate $2m$ times the squared difference between adjacent genotypes: $V_0 = 2m(\Delta z)^2$). The differences from (**??**) are due to the small-trait-variance approximation of $a(t)$, and uniparental rather than biparental inheritance. The resulting approximate condition for the steady-state trait variance, analogous to (**??**), is

$$(1+\bar{b})\bar{\sigma}_z^2 = V_0 + \frac{2\gamma\bar{b}^2\sigma_M^2}{2-(1-\gamma)\bar{b}} \tag{27}$$

where $b(t) = \sigma_z^2(t)/\sigma_w^2$. With a trending optimum, as in the quantitative trait model the steady-state adaptation lag is characterized by $\mathbb{E}(\mathcal{R}) = \eta$, giving in this case

$$\bar{L} \approx \frac{\eta}{(1-\gamma)\bar{b}}. \tag{28}$$

analogous to (**??**).

In the ESS case we can alternatively use Gaussian approximations for the clonal model, as we did for the sexual quantitative trait model. Doing so produces exactly the same results as above, except that $b(t)$ is everywhere replaced by $a(t) = \sigma_z^2(t)/(\sigma_w^2 + \sigma_z^2(t))$.

4.2 Analysis: two-phenotype trait distribution

Now we consider the situation where the trait distribution is approximately bimodal at the two-phenotype ESC $z = \pm z_s$ that arises when γ and σ_M^2 are high enough to maintain trait variation through the storage effect.

We first need to approximate the steady-state value of $\sigma_z^2(t)$. From **?** we know that the z_s is largely independent of γ, unless γ is close to the threshold for evolutionary branching – it only depends on σ_M^2 and σ_w^2. In the following, we assume that z_s is independent of γ. $\sigma_z^2(t)$ is then determined by the fraction of type z_s year t, and $\bar{\sigma}_z^2$ is determined by the distribution of that fraction.

At the two-type ESC, a good approximation is that all reproduction in any one year is done by whichever type is closest to M_t (**?**) – which is a kind of strong selection limit. So if n_1 is the

number of type z_s in the population, we have

$$n_1(t+1) = KB(t) + \gamma n_1(t), \tag{29}$$

where $B(t)$ is 0 or 1 with equal probability. Taking the variance of both sides (recall that $B(t)$ and $n_1(t)$ are independent) and solving, the stationary variance of n_1 is $0.25K^2/(1-\gamma^2)$. The total population is $K/(1-\gamma)$, so defining p to be the fraction of type z_s in the population we have

$$\mathrm{Var}(p) = \frac{0.25K^2}{1-\gamma^2}\frac{(1-\gamma)^2}{K^2} = 0.25\left(\frac{1-\gamma}{1+\gamma}\right). \tag{30}$$

Hence increasing generation overlap gives lower variance in the population composition, as expected.

Conditional on p the trait variance is $4z_s^2p(1-p)$ so the expected trait variance is $4z_s^2\mathbb{E}[p(1-p)]$. Using (**??**) and $\mathbb{E}(p) = 0.5$ by symmetry, we find $\mathbb{E}[p(1-p)] = 0.5\gamma/(1+\gamma)$, so

$$\bar{\sigma}_z^2 \approx \frac{2z_s^2\gamma}{1+\gamma}. \tag{31}$$

The selection response implied by (**??**) is

$$\mathcal{R} = \bar{z}(t+1) - \bar{z}(t) = 2z_s(1-\gamma)[B(t) - p(t)]. \tag{32}$$

We again use σ_R, the standard deviation of $\mathcal{R}$, as a measure of the average magnitude of response. As B and p are independent, combining (**??**) with (**??**) we get

$$\sigma_{\mathcal{R}} = \frac{\sqrt{2}(1-\gamma)z_s}{\sqrt{1+\gamma}} \tag{33}$$

which is monotonically decreasing in γ. As in the quantitative trait model, a strong selection limit removes one "edge of the sword", because all selected parents have a single value determined by M_t regardless of the overall trait distribution. The slowing effect of generation overlap is then dominant.

In a nonstationary environment, if the mean trait optimum moves slowly enough for the population to track it closely, the trait distribution (in simulations) remains bimodal. If the standing genetic variation allows the two trait modes to move in parallel with the mean trait optimum, the trait difference between the two modes, and the fluctuations in their relative frequencies, are (approximately) the same as they are in a stationary environment. The trait variance is then still approximated by (**??**), and substituting this into (**??**) the adaptation lag is approximated as

$$\bar{L} \approx \left(\frac{\sigma_w^2\eta}{2z_s}\right)\frac{1+\gamma}{\gamma(1-\gamma)}. \tag{34}$$

Because the two-type ESC does not have small trait variance as assumed in (**??**), eqn. (**??**) can only be justified as a weak-selection approximation. If we assume that z_s is independent of γ, (**??**) predicts that the adaptation lag is minimized at $\gamma = \sqrt{2} - 1 \approx 0.41$.

4.3 Simulation Results

Model simulations confirm that generation overlap promotes the maintenance of genetic variation in stationary environments (**?**). However, unlike the quantitative trait model, the clonal model results in a multimodal trait distribution when generation overlap, γ, and environmental stochasticity, σ_M^2, are high (**??**). Hence, the approximation based on a two-phenotype trait distribution (eqn. (**??**)) works when environmental stochasticity is intermediate ($\sigma_M = 10$) and generation overlap is high (Fig. **??**A). When environmental stochasticity is small ($\sigma_M = 5$), trait distributions are unimodal, and thus the approximation based on small-variance (eqn. (**??**)) fits as long as generation overlap is not very high. When environmental stochasticity is high ($\sigma_M = 15$), the trait distribution can become trimodal, but the approximation based on the bimodal trait distribution (eqn. (**??**)) only slightly underestimates the trait variance (Fig. **??**A).

Because of the positive relationship between generation overlap and genetic variation (Fig. **??**A), the response to selection can be maximized at intermediate values of generation overlap when $\sigma_M = 10$ (Fig. **??**B). The approximation based on the two-phenotype trait distribution (eqn. **??**) works well due to the bimodal trait distribution (Fig. **??**B). Please note that z_s is positively correlated with γ here because evolutionary branching occurs when γ is small with $\sigma_M = 10$. When environmental stochasticity is small ($\sigma_M = 5$), on the other hand, generation overlap decreases the response to selection due to small genetic variance, and the approximation based on small-variance (eqn. (**??**)) works very well (results not shown).

As with the quantitative trait model, when the speed of environmental change (η) is slow and when environmental stochasticity is intermediate ($\sigma_M = 7$), intermediate generation overlap can minimize adaptation lag (Fig. **??**E, F). On the other hand, generation overlap increases adaptation lag when stochasticity is small ($\sigma_M = 5$, Fig. **??**A, B) or large ($\sigma_M = 13$, Fig. **??**C, D). When environmental stochasticity is not large ($\sigma_M = 5, 7$ in Fig. **??**B, F), the small-variance approximation (eqn. (**??**)) works well because the trait distribution is unimodal. However, adaptation lag is minimized at intermediate generation overlap near $\gamma = 0.41$ (Fig. **??**E, F), as predicted by the two-trait ESC approximation (**??**) when z_s is constant. The small variance approximation does not work well when environmental stochasticity is higher ($\sigma_M = 10$) and generation overlap is high, because the trait distribution remains bimodal (Fig. **??**C, D).

5 Discussion

Previous studies that did not consider environmental stochasticity in fitness components have predicted that greater generation overlap invariably slows down the speed of evolution (e.g., **?**) and reduces the chances for population persistence or evolutionary rescue in the face of rapid environmental change (**?**). The picture that emerges from our models and analyses is less simple, because of the potentially opposing effects of generation overlap on two components of response to selection: the standing level of trait variation, and the fraction of the population exposed to selection on traits affecting reproductive success. As a result we find complicated effects of generation overlap (e.g., long-term dormancy) on evolutionary responses. In some circumstances, intermediate generation overlap is best for rapid initial response to directional environmental change (Figs. **??**, **??**) as well as for long-term responses (Figs. **??**, **??**): this is because smaller generation overlap reduces the amount of genetic variation, whereas larger generation overlap results in a population with few active individuals. Although the short-term response is transient, it may nonetheless be important (**?**), especially for adaptation that prevents extinction (i.e., evolutionary rescue; **????**).

Our overall conclusions can be summarized conveniently in terms of σ_M/σ_w, which is the environmental stochasticity parameter when the trait is scaled by σ_w. For both the quantitative trait and clonal models, there is only one "edge of the sword" when σ_M/σ_w is very small (low variance, weak selection) or very large (large variance, strong selection) because γ has little or no effect on the trait variance and the slowing effect of generation overlap dominates ($\sigma_M/\sigma_w = 0.25$ and 1 in Fig. **??**A, B and Fig. **??**A, B, respectively, and $\sigma_M/\sigma_w = 10$ and 2.6 in Fig. **??**C, D and Fig. **??**C, D, respectively). At intermediate values (intermediate environmental stochasticity, intermediate selection), trait variance becomes sensitive to γ. This is clearest in the clonal model which goes from ESS to ESC, but is true in both models. Both edges are then present, and intermediate γ can maximize the selection response ($2 \leq \sigma_M/\sigma_w < 3.34$ in Fig. **??**, $\sigma_M/\sigma_w = 2$ in Fig. **??**B) and minimize the adaptation lag ($\sigma_M/\sigma_w \approx 3.33$ in Fig. **??**E, F, $\sigma_M/\sigma_w = 1.4$ in Fig. **??**E, F).

Theoretical models for rapid evolution have typically assumed constant additive genetic variance in evolving traits (e.g., **??**). However, genetic variance itself can change through time in response to environmental change (**?**). Furthermore, studies of rapid evolution have shown the potential role of temporally fluctuating selection in maintaining genetic variance (**???**), whereas classic population genetic models, without generation overlap, predict that temporally fluctuating selection is not very efficient at maintaining genetic variation (**??**); the only robust mechanism is heterozygote advantage, in models where the optimal or ESS phenotype can only be realized by a heterozygote. When fitness has a stochastic component and generations overlap, the temporal storage effect may play an important role in the maintenance of standing genetic variation (**?**) and thus promote rapid evolution. Our results here illustrate this fundamental relationship.

In both of our models, trait distributions can be far from Gaussian, and this has a significant impact on evolutionary dynamics. In the clonal model, evolutionary branching produces strongly bimodal distributions when environmental variance and generation overlap are high (**???**), and long-term adaptation lag depends on whether or not the bimodality persists. The quantitative trait model can exhibit transient bimodality, high kurtosis, and large negative skew (Fig. **??**), even though each mating produces a Gaussian distribution of offspring trait values; none of these occur under constant directional selection with nonoverlapping generations (**?**). Gaussian-based approximations are convenient and powerful, for example their use allowed **?** to give exact conditions under which adaptation and plasticity allowed population persistence in a deterministically trending environment. But other approaches will be needed for populations with generation overlap and fluctuating selection (perhaps adding age as a second individual-level state variable and allowing a different trait mean and variance for each age; **?**) and they will probably be far less analytically tractable.

When our models include rapid directional environmental change (superimposed on random year-to-year variation), migration from the past always acts as a kind of migration load. However, temporal migration may promote evolution when directional environmental change is slow enough. Thus, dormancy can be regarded as a bet-hedging strategy in a stationary but fluctuating environment, as genetic variants that have been selected in the past may be adaptive again in the future. On the other hand, individuals emerging from dormancy are most likely maladapted when the environmental trend is strong relative to the fluctuation intensity (σ_M^2). For example, long-lived dormant eggs of freshwater zooplankton can introduce generation overlap and have been shown to affect their rate of evolution in a changing environment. In small water-bodies, the copepod *Onychodiaptomus sanguineus* has life histories adapted to avoid the springtime onset of fish predation, which varies among years in timing and intensity, causing the optimal life history to fluctuate. The dormant eggs produced each year reflect that fluctuating selection, and store trait variation in a sediment egg bank (**?**). Each year, copepod trait change reflects both the contribution of variance from eggs hatching from the sediments and the direction and intensity of selection from fish predation (**?**), illustrating the "good edge" of generation overlap. However, in one instance, when a severe drought removed fish entirely from a pond, natural selection on optimal life history was intense, and the trait distribution emerging from the egg bank strongly retarded selection response (**?**) revealing the "bad edge" of generation overlap.

Recent theoretical papers have studied evolution of the storage effect, but mainly focused on the long-term outcome of adaptive evolution (**???**). Our study is instead about how the storage effect can affect contemporary microevolutionary dynamics. Thus it will be interesting in future studies to examine eco-evolutionary dynamics when the storage effect is present. That is, how natural selection affects species coexistence, and then species coexistence affects evolutionary dynamics. Future studies may be able to address more complicated dynamics by using more thorough numerical simulations and analytical methods for structured populations (**?**). While most microcosm experiments do not have dormant stages, it will be interesting to construct an

experimental system with dormant stages to examine the roles of dormancy on eco-evolutionary dynamics to mimic processes in the wild (**??**).

Acknowledgements

We thank two anonymous referees for detailed and helpful comments on the original manuscript. M.Y. was supported by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (KAKENHI) 15H02642, 16K18618, 16H04846, and 18H02509 and by Hakubi Center for Advanced Research and John Mung Program of Kyoto University. N.G.H. was supported by USDA National Institute of Food and Agriculture, Hatch Grant 1007884, and by Eawag, the Swiss Federal Institute of Aquatic Science and Technology, while this paper was being prepared for publication. M.R. was supported by NERC-NE/K014048/1. S.P.E. was supported by US National Science Foundation grant DEB 1353039.

Literature Cited

Abrams, P. A., C. M. Tucker, and B. Gilbert. 2013. Evolution of the storage effect. Evolution 67:315–327.

Alexander, H. K., G. Martin, O. Y. Martin, and S. Bonhoeffer. 2014. Evolutionary rescue: linking theory for conservation and medicine. Evolutionary Applications 7:1161–1179.

Barfield, M., R. D. Holt, and R. Gomulkiewicz. 2011. Evolution in stage-structured populations. The American Naturalist 177:397–409.

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

Baumgartner, M. F., and A. M. Tarrant. 2017. The physiology and ecology of diapause in marine copepods. Annual Review of Marine Science 9:387–411.

Bell, G. 2010. Fluctuating selection: the perpetual renewal of adaptation in variable environments. Philosophical Transactions of the Royal Society B: Biological Sciences 365:87–97.

———. 2017. Evolutionary rescue. Annual Review of Ecology, Evolution, and Systematics 48:605–627.

Bolnick, D. I., and P. Nosil. 2007. Natural selection in populations subject to a migration load. Evolution 61:2229–2243.

Brendonck, L., and L. De Meester. 2003. Egg banks in freshwater zooplankton: evolutionary and ecological archives in the sediment. Hydrobiologia 491:65–84.

Bridle, J. R., J. Polechová, M. Kawata, and R. K. Butlin. 2010. Why is adaptation prevented at ecological margins? New insights from individual-based simulations. Ecology Letters 13:485–494.

Brown, J. S., and D. L. Venable. 1986. Evolutionary ecology of seed-bank annuals in temporally varying environments. The American Naturalist 127:31–47.

Bruns, T. D., K. G. Peay, P. J. Boynton, L. C. Grubisha, N. A. Hynson, N. H. Nguyen, and N. P. Rosenstock. 2009. Inoculum potential of *Rhizopogon* spores increases with time over the first 4 yr of a 99-yr spore burial experiment. New Phytologist 181:463–470.

Bulmer, M. G. 1980. The Mathematical Theory of Quantitative Genetics. The Clarendon Press, Oxford, UK.

Bürger, R., and M. Lynch. 1995. Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49:151–163.

Chevin, L.-M., R. Lande, and G. M. Mace. 2010. Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. PLOS Biology 8:e1000357.

Cohen, D. 1966. Optimizing reproduction in a randomly varying environment. Journal of Theoretical Biology 12:119–129.

Cohen, D., and S. A. Levin. 1991. Dispersal in patchy environments: the effects of temporal and spatial structure. Theoretical Population Biology 39:63–99.

Dempster, E. R. 1955. Maintenance of genetic heterogeneity. Pages 25–32 *in* Cold Spring Harbor Symposia on Quantitative Biology. Vol. 20. Cold Spring Harbor Laboratory Press.

Doebeli, M., and U. Dieckmann. 2000. Evolutionary branching and sympatric speciation caused by different types of ecological interactions. The American Naturalist 156:S77–S101.

Ellegaard, M., and S. Ribeiro. 2017. The long-term persistence of phytoplankton resting stages in aquatic 'seed banks'. Biological Reviews .

Ellner, S., and N. G. Hairston, Jr. 1994. Role of overlapping generations in maintaining genetic variation in a fluctuating environment. The American Naturalist 143:403–417.

Ellner, S., and A. Sasaki. 1996. Patterns of genetic polymorphism maintained by fluctuating selection with overlapping generations. Theoretical Population Biology 50:31–65.

Ellner, S. P., N. G. Hairston, Jr., C. M. Kearns, and D. Babaï. 1999. The roles of fluctuating selection and long-term diapause in microevolution of diapause timing in a freshwater copepod. Evolution 53:111–122.

Evans, M. E. K., and J. J. Dennehy. 2005. Germ banking: bet-hedging and variable release from egg and seed dormancy. The Quarterly Review of Biology 80:431–451.

Gomulkiewicz, R., and R. D. Holt. 1995. When does evolution by natural selection prevent extinction? Evolution 49:201–207.

Gyllström, M., and L.-A. Hansson. 2004. Dormancy in freshwater zooplankton: induction, termination and the importance of benthic-pelagic coupling. Aquatic Sciences 66:274–295.

Hairston, N. G., Jr., and B. T. De Stasio, Jr. 1988. Rate of evolution slowed by a dormant propagule pool. Nature 336:239–242.

Hairston, N. G., Jr., S. Ellner, and C. M. Kearns. 1996*a*. Overlapping generations: the storage effect and the maintenance of biotic diversity, pages 109–145. University of Chicago Press, Chicago, IL.

Hairston, N. G., Jr., C. M. Kearns, and S. P. Ellner. 1996*b*. Phenotypic variation in a zooplankton egg bank. Ecology 77:2382–2392.

Haldane, J. B. S., and S. D. Jayakar. 1963. Polymorphism due to selection of varying direction. Journal of Genetics 58:237–242.

Hastings, A. 2004. Transients: the key to long-term ecological understanding? Trends in Ecology & Evolution 19:39–45.

Hedrick, P. W. 1995. Genetic polymorphism in a temporally varying environment: effects of delayed germination or diapause. Heredity 75:164–170.

Iwasa, Y., A. Pomiankowski, and S. Nee. 1991. The evolution of costly mate preferences ii. the 'handicap' principle. Evolution 45:1431–1442.

Kinnison, M. T., and N. G. Hairston, Jr. 2007. Eco-evolutionary conservation biology: contemporary evolution and the dynamics of persistence. Functional Ecology 21:444–454.

Kopp, M., and S. Matuszewski. 2014. Rapid evolution of quantitative traits: theoretical perspectives. Evolutionary Applications 7:169–191.

Lennon, J. T., and S. E. Jones. 2011. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nature Reviews Microbiology 9:119–130.

Levin, S. A., D. Cohen, and A. Hastings. 1984. Dispersal strategies in patchy environments. Theoretical Population Biology 26:165–191.

Ludwig, D., and S. A. Levin. 1991. Evolutionary stability of plant communities and the maintenance of multiple dispersal types. Theoretical Population Biology 40:285–307.

Lynch, M., and R. Lande. 1993. Evolution and extinction in response to environ mental change, pages 234–250. Sinauer Associates, Inc., Sunderland, MA.

Messer, P. W., S. P. Ellner, and N. G. Hairston, Jr. 2016. Can population genetics adapt to rapid evolution? Trends in Genetics 32:408–418.

Miller, E. T., and C. A. Klausmeier. 2017. Evolutionary stability of coexistence due to the storage effect in a two-season model. Theoretical Ecology 10:91–103.

Miner, B. E., L. De Meester, M. E. Pfrender, W. Lampert, and N. G. Hairston, Jr. 2012. Linking genes to communities and ecosystems: *Daphnia* as an ecogenomic model. Proceedings of the Royal Society B: Biological Sciences 279:1873–1882.

Nuismer, S. L., M. Doebeli, and D. Browning. 2005. The coevolutionary dynamics of antagonistic interactions mediated by quantitative traits with evolving variances. Evolution 59:2073–2082.

Orive, M. E., M. Barfield, C. Fernandez, and R. D. Holt. 2017. Effects of clonal reproduction on evolutionary lag and evolutionary rescue. The American Naturalist 190:469–490.

Pantel, J. H., C. Duvivier, and L. De Meester. 2015. Rapid local adaptation mediates zooplankton community assembly in experimental mesocosms. Ecology Letters 18:992–1000.

R Core Team. 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rees, M. 1996. Evolutionary ecology of seed dormancy and seed size. Philosophical Transactions of the Royal Society B: Biological Sciences 351:1299–1308.

Rees, M., and S. P. Ellner. in press. Why so variable: can genetic variance in flowering thresholds be maintained by fluctuating selection? The American Naturalist .

Sasaki, A., and S. Ellner. 1997. Quantitative genetic variance maintained by fluctuating selection with overlapping generations: Variance components and covariances. Evolution 51:682–696.

Snyder, R. E., and P. B. Adler. 2011. Coexistence and coevolution in fluctuating environments: can the storage effect evolve? The American Naturalist 178:E76–E84.

Templeton, A. R., and D. A. Levin. 1979. Evolutionary consequences of seed pools. The American Naturalist 114:232–249.

Turelli, M. 2017. Fisher's infinitesimal model: A story for the ages. Theoretical Population Biology 118:46–49.

Turelli, M., and N. H. Barton. 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138:913–941.

Turelli, M., D. W. Schemske, and P. Bierzychudek. 2001. Stable two-allele polymorphisms maintained by fluctuating fitnesses and seed banks: protecting the blues in *Linanthus parryae*. Evolution 55:1283–1298.

Walsh, B., and M. Lynch. 2018. Evolution and Selection of Quantitative Traits. Oxford University Press, Oxford, UK.

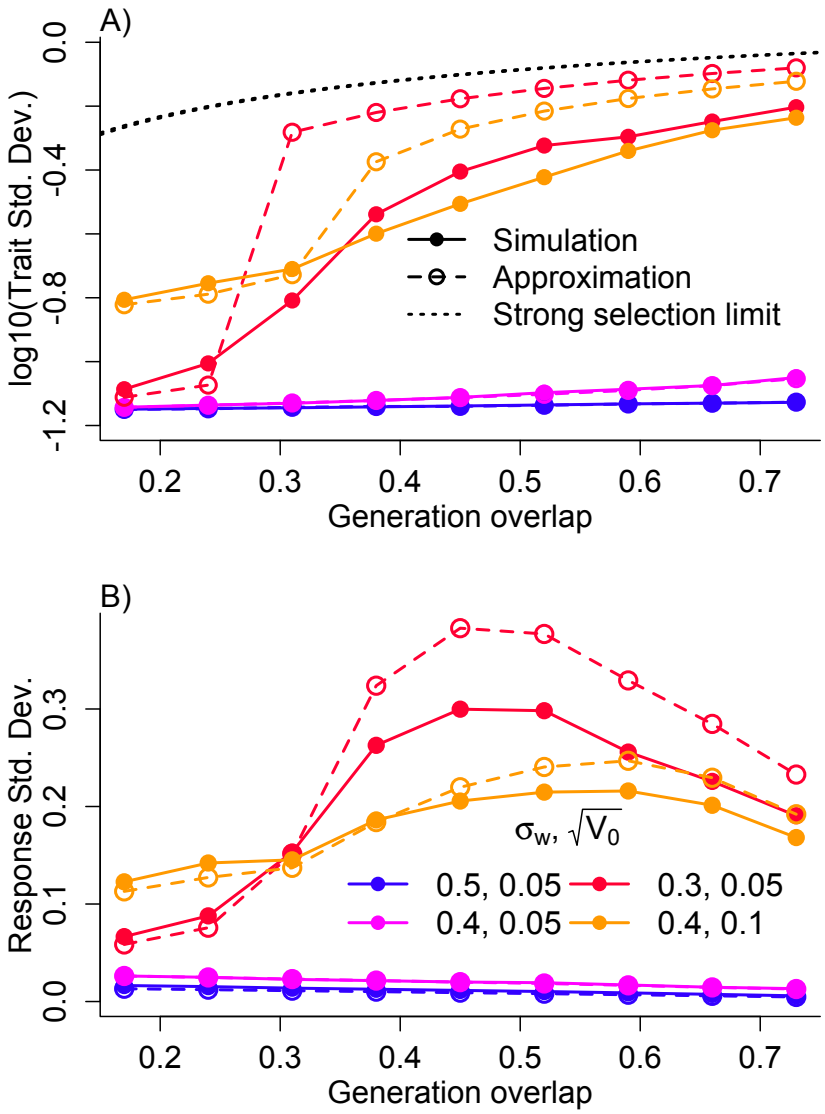

Figure 1: Effect of generation overlap γ on trait variance $\bar{\sigma}_z^2$ and selection response $\mathcal{R}$ in the quantitative trait model. A) Steady-state average trait variance. B) $\sigma_{\mathcal{R}}$, the standard deviation of $\mathcal{R}$. Solid curves (with closed symbols) are simulation results for 10000 years of a population of expected size 10000; plotted points are the within-year trait variance (in A) or standard deviation of the selection response (in B) averaged over the last 8000 years. In panel A, dashed curves (with open circles) are numerical solutions of the moment approximation equation (**??**). Dotted black curves (almost completely overlapping) are the strong selection limits for the two values of V_0 used in the simulations. In panel B the dashed curves are (**??**) computed using mean σ_z^2 from the simulations rather than the approximate σ_z^2 in calculating $\bar{a}$. Note that there are dashed lines nearly hidden by the blue and magenta solid lines. Parameter values were $s_d = 0.8, s_a = 0.1, \sigma_M = 1$ for all simulations. The numerical values in the panel B legend are σ_w and $\sqrt{V_0}$; these apply to panel A as well. In panel A, the curves for $\sqrt{V_0} = 0.05$ converge at low generation overlap because the solution to (**??**) at $\gamma = 0$ is $\sigma_z^2 = 2V_0(1 + O(V_0/\sigma_w^2))$. Figure produced by script `QG-IBM-noTrend.R`.

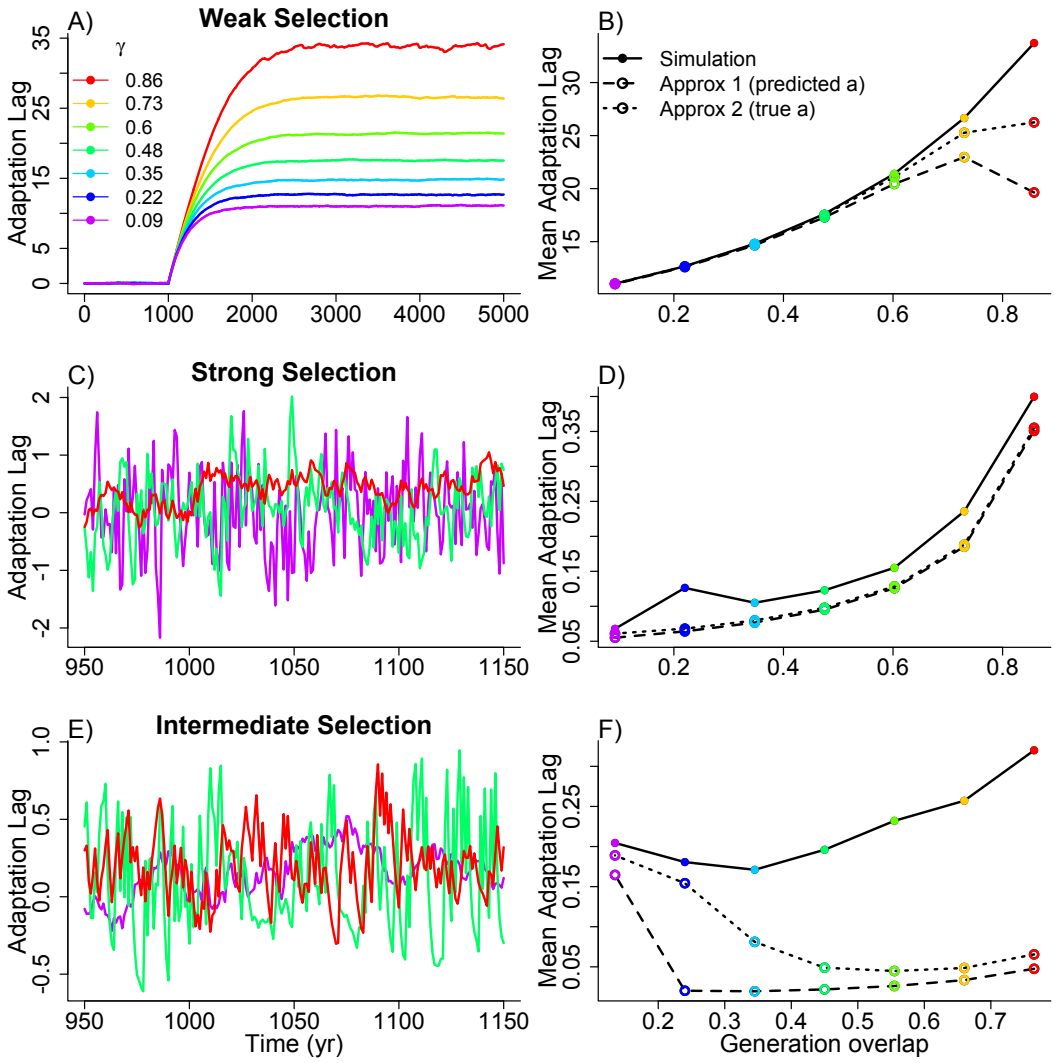

Figure 2: Effect of generation overlap on adaptation lag in the quantitative trait model with a moving expected trait optimum. A), B): $\eta = 0.05$ with weak selection, $\sigma_w = 4$. C), D): $\eta = 0.05$ with strong selection, $\sigma_w = 0.1$. E), F): $\eta = 0.01$ with intermediate selection, $\sigma_w = 0.3$. The trend begins in year 1000. Panel A) legend shows the generation overlap γ (same order, top to bottom, as the curves in panel A). "Adaptation lag" is the difference between expected trait optimum μ_t and trait mean $\bar{z}(t)$. A), C), and E) show adaptation lag over portions of 12000 year individual-based simulations with different generation overlap; note the difference in scale resulting from the differing strength of selection. B), D), and F) show the average adaptation lag over the last 6000 years of the simulations (solid curve and symbols) and the theoretical predictions (dotted and dashed curves). The theoretical predictions are eqn. (**??**) with either the true $\bar{a}$ over the second half of the simulations (dotted curves) or the approximate predictions $\bar{a} = 2V_0/(2V_0 + \sigma_w^2)$ (very weak selection) in B), $\bar{a} = 1$ (very strong selection) in D), and $\bar{a}$ resulting from the solution to (**??**) in F) (dashed curves). Other parameters were $\sigma_M = 1$, and $s_d = 0.9, s_a = 0.05, \sqrt{V_0} = 0.2$ for A)-D), $s_d = 0.8, s_a = 0.1, \sqrt{V_0} = 0.05$ for E) and F). Figure produced by `QG-IBM-Trend.R`.

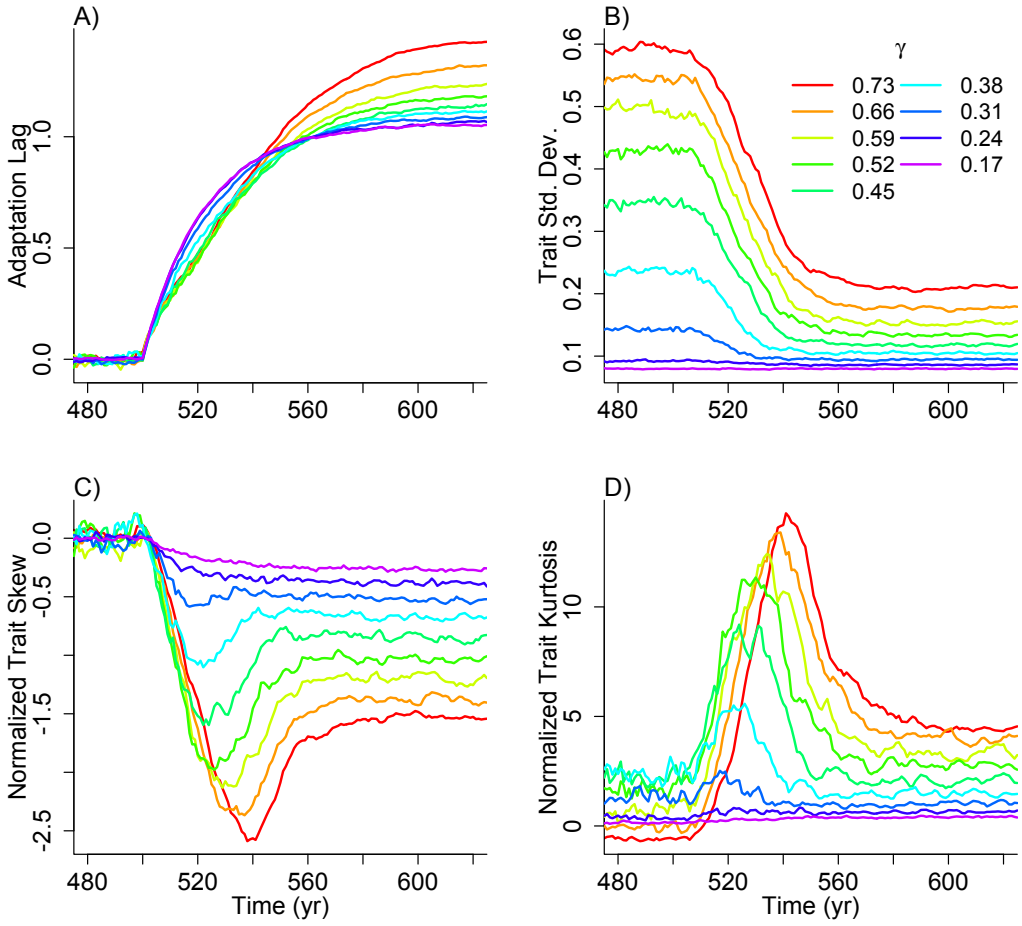

Figure 3: Effect of generation overlap on the transient dynamics of the quantitative trait model at the start of an environmental trend. A) Mean adaptation lag. B), C), D) Mean among-individual standard deviation, normalized skew, and kurtosis of the trait. Normalized skew = $m_3/m_2^{3/2}$, normalized kurtosis = $m_4/m_2^2 - 3$ where m_n is the n^{th} central moment. The trend begins in year 500. Each curve is the average across 500 independent simulations with the same H value. Legend in panel B) gives the generation overlap γ. Parameter values for all simulations are: $\sigma_w = 0.3, \sigma_M = 1, s_d = 0.8, s_a = 0.1, \eta = 0.05$, and $\sqrt{V_0} = 0.05$. Figure produced by script `QG-IBM-Trend-Initial.R`.

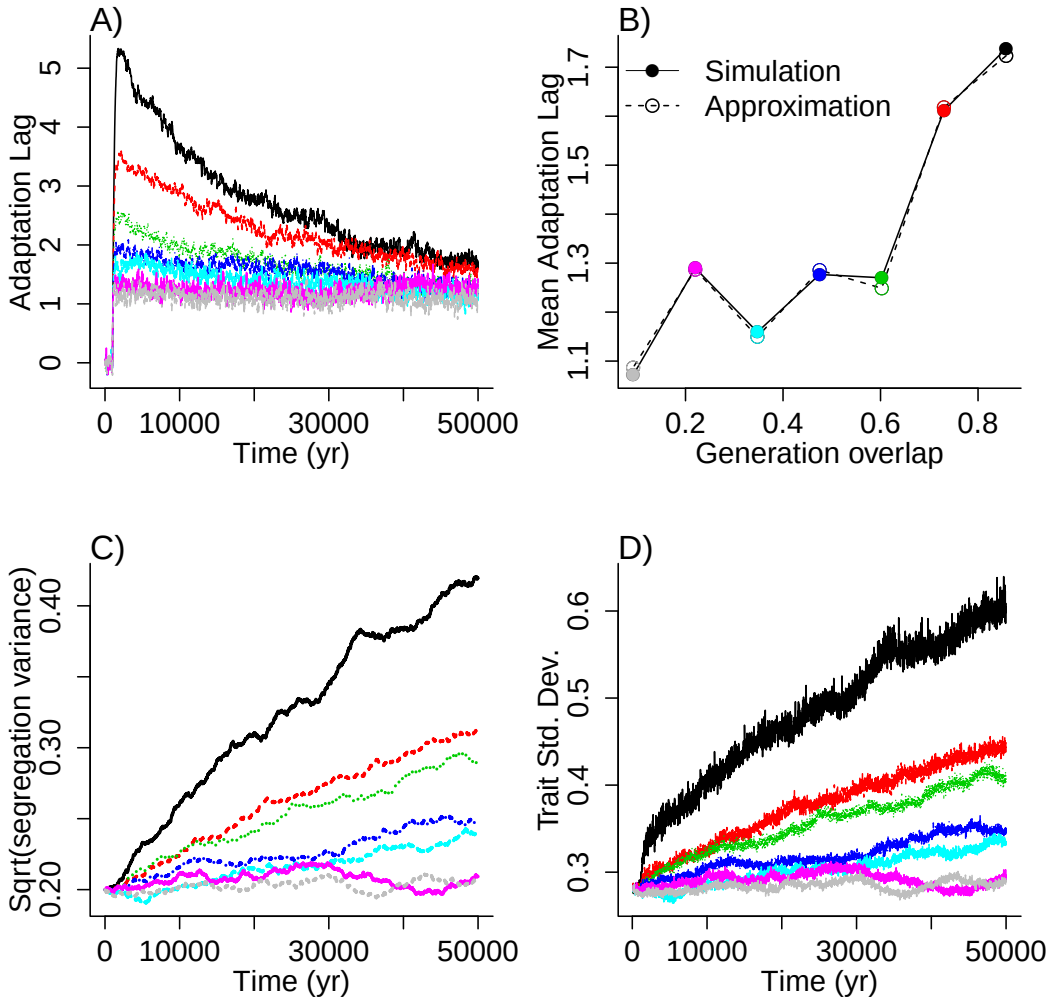

Figure 4: Effect of generation overlap on the adaptation lag in the quantitative trait model with a moving trait optimum, when segregation variance V_0 can evolve. Parameter values $\sigma_w = 2, \sigma_M = 1, s_d = 0.9, s_a = 0.05, \eta = 0.02$, and $\sqrt{V_0} = 0.2$ for all individuals at $t = 0$. Mean trait optimum is 0 until $t = 1000$, then increases by η per year. A) Adaptation lag over time. B) Mean adaptation lag in the simulations (averaged over the last 6000 years; solid curve and symbols) and the theoretical prediction, eqn. (**??**) with $\bar{a}$ computed from the simulations over the same time period. C) $\sqrt{V_0}$ over time. D) Population standard deviation of z over time. Colors indicate different values of generation overlap γ as in panel B). Figure produced by script `QG-IBM-TrendV0.R`.

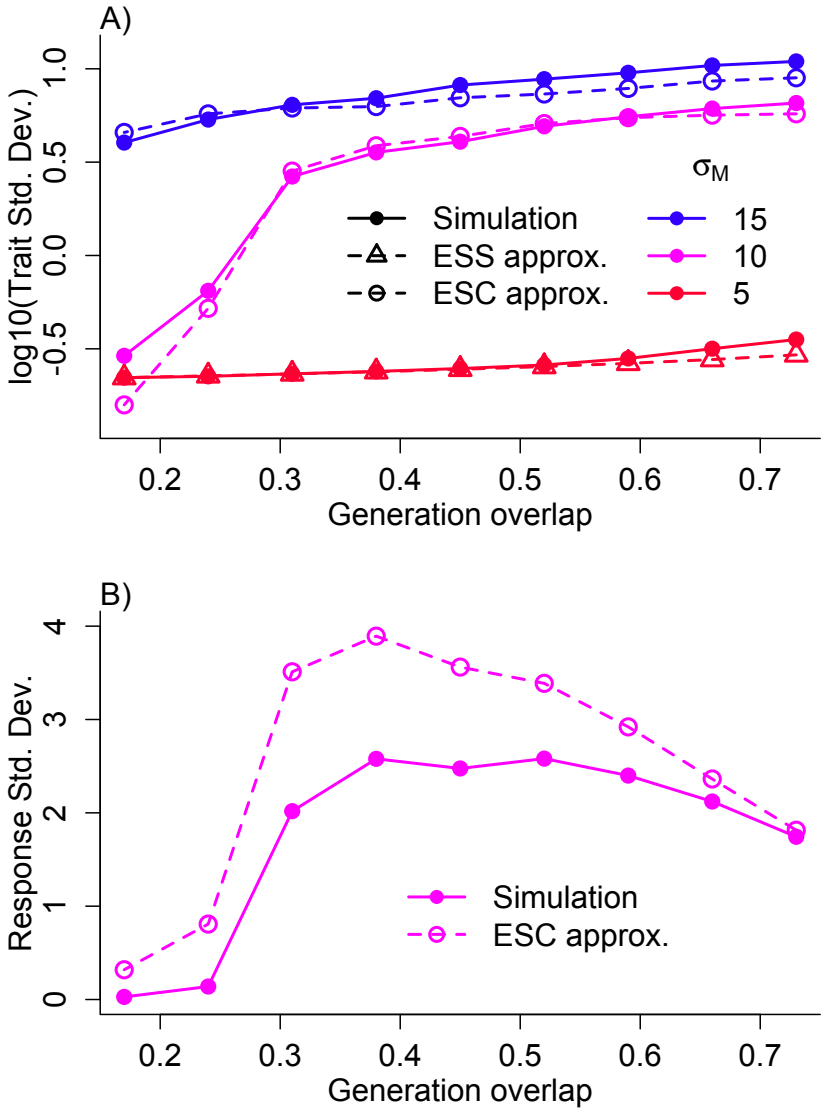

Figure 5: Effect of generation overlap γ on trait variance $\bar{\sigma}_z^2$ and selection response $\mathcal{R}$ in the clonal model in a stationary environment. A) Steady-state average trait variation. B) $\sigma_{\mathcal{R}}$, the standard deviation of the response to selection (change in mean trait between one time and the next) when $\sigma_M = 10$. Solid curves (with closed symbols) are simulation results for 5000 years of a population; plotted points are the within-year trait variance averaged over the last 3000 years. Dashed curves with open triangles are the small-variance approximation (**??**) for a one-trait ESS, dashed curves with open circles are the two-phenotype ESC approximation (**??**). Here values of $\bar{z}_s$ and $\bar{\sigma}_z^2$ were calculated from the simulations, and this may be why the fit is better than that of the quantitative trait model (Fig. **??**). Parameter values were $s_d = 0.8, s_a = 0.1, \sigma_w = 5, K = 1, m = 10^{-3}$, and $\Delta z = 0.2$ for all simulations. The numerical values in the panel A legend are σ_M. In panel B), the dashed curve with open circles is the two-phenotype approximation (**??**). Figure produced by script `HG-DE-noTrend.R` .

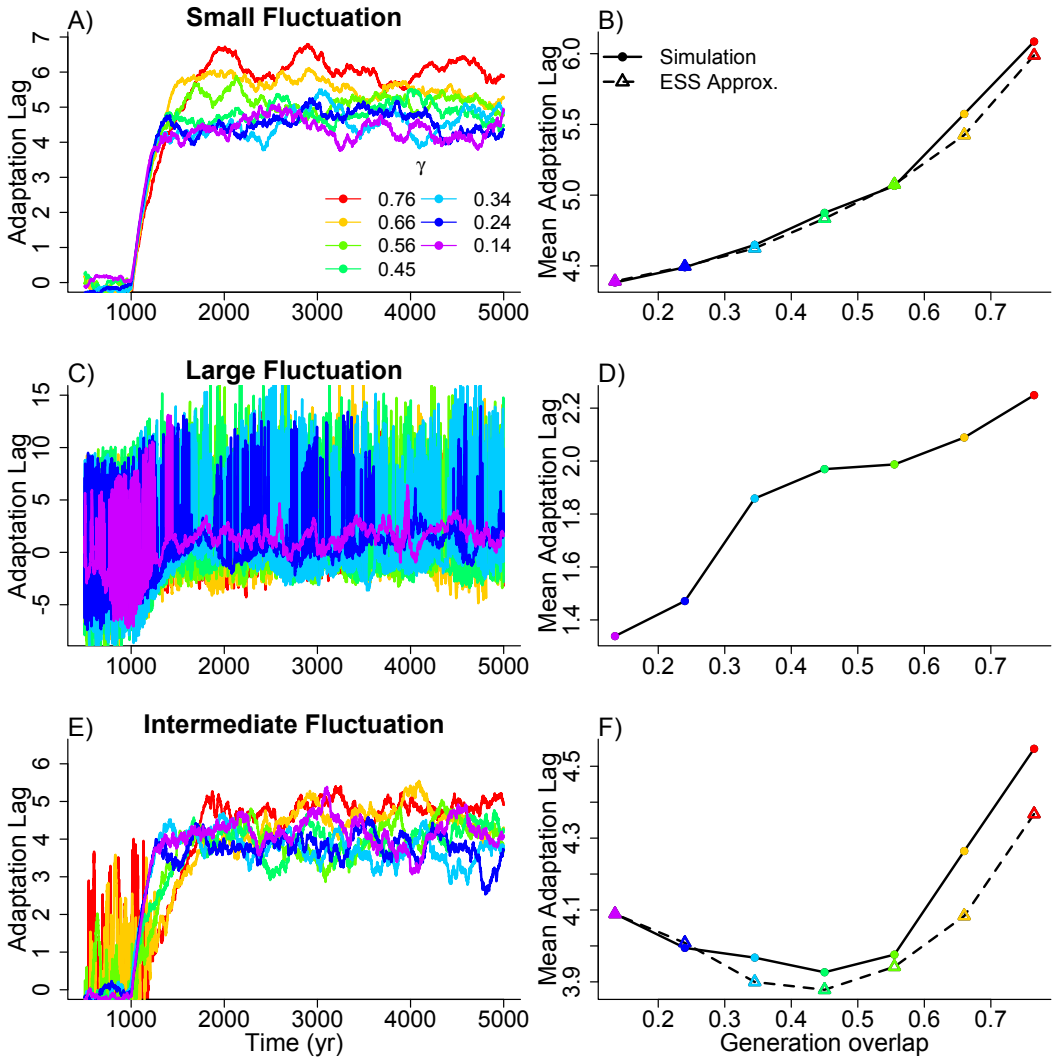

Figure 6: Effect of generation overlap on the adaptation lag in the clonal model with a moving expected trait optimum. A), B): small fluctuation, $\sigma_M = 5$. C), D): large fluctuation, $\sigma_M = 13$. E), F): intermediate fluctuation, $\sigma_M = 7$. The trend begins in year 1000. Panel A) legend shows the generation overlap γ. "Adaptation lag" is the difference between expected trait optimum μ_t and trait mean $\bar{z}(t)$. A), C), and E) show adaptation lag over portions of 22000 year simulations with different generation overlap; note the difference in scale resulting from the differing intensity of fluctuation. B), D), and F) show the adaptation lag in the simulations (averaged over the last 16000 years; solid curve and symbols) and the small-variance approximation, eqn. (**??**), where the average $\sigma_z^2(t)$ value was calculated over the last 16000 years of the simulations. Here $\eta = 0.02$, and other parameter values are the same as Fig. **??**. Figure produced by script `HG-DE-Trend.R`.

Online Supplementary Information

A.1 Supplementary Figures

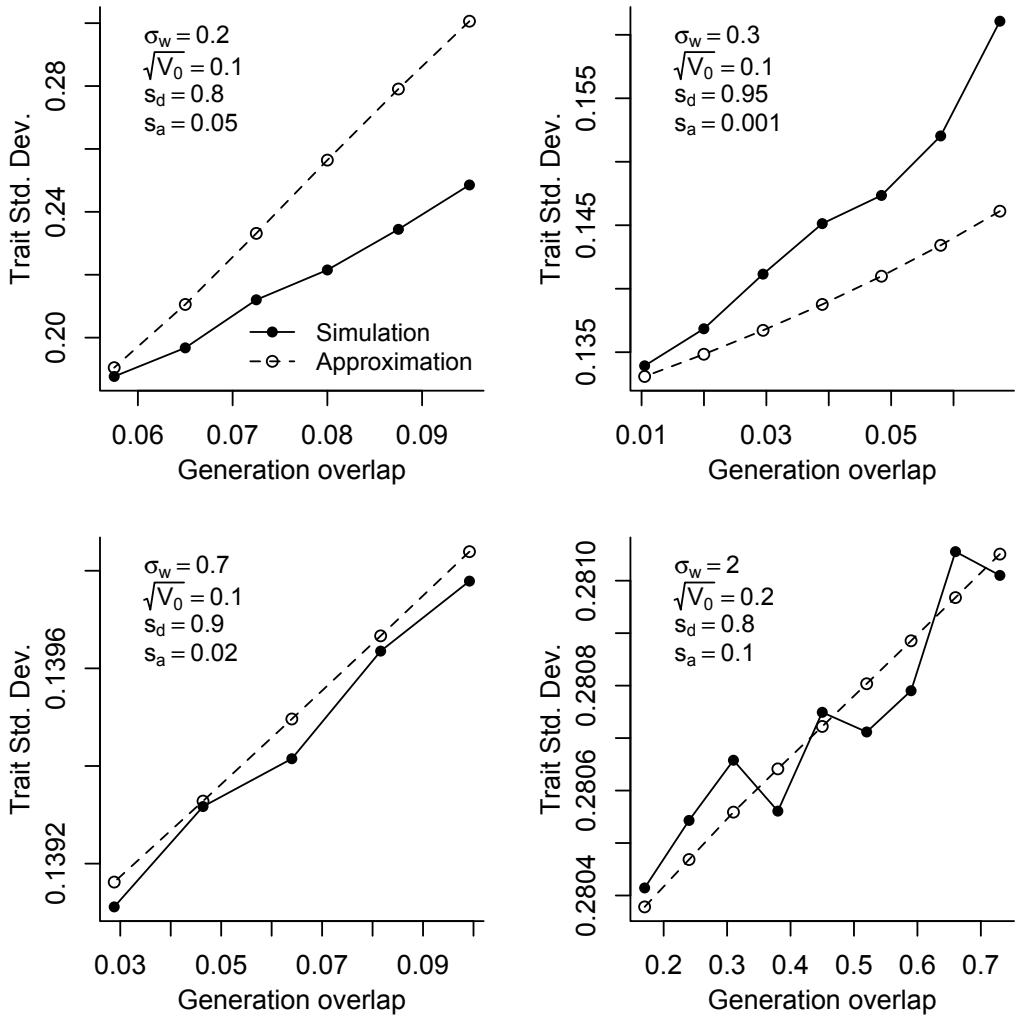

Figure A-1: Effect of generation overlap on steady-state average trait variation in the quantitative trait model, when generation overlap γ is low in a stationary environment. Each panel compares simulation results (as in Fig. **??**) with the moment approximation (**??**) with $H = 0.91, 0.93, \cdots, 0.99$. The values of $\sigma_w, \sqrt{V_0}, s_d, s_a$ are listed in each panel (in that order, from top to bottom). Figure produced by script `QG-IBM-noTrend.R`.

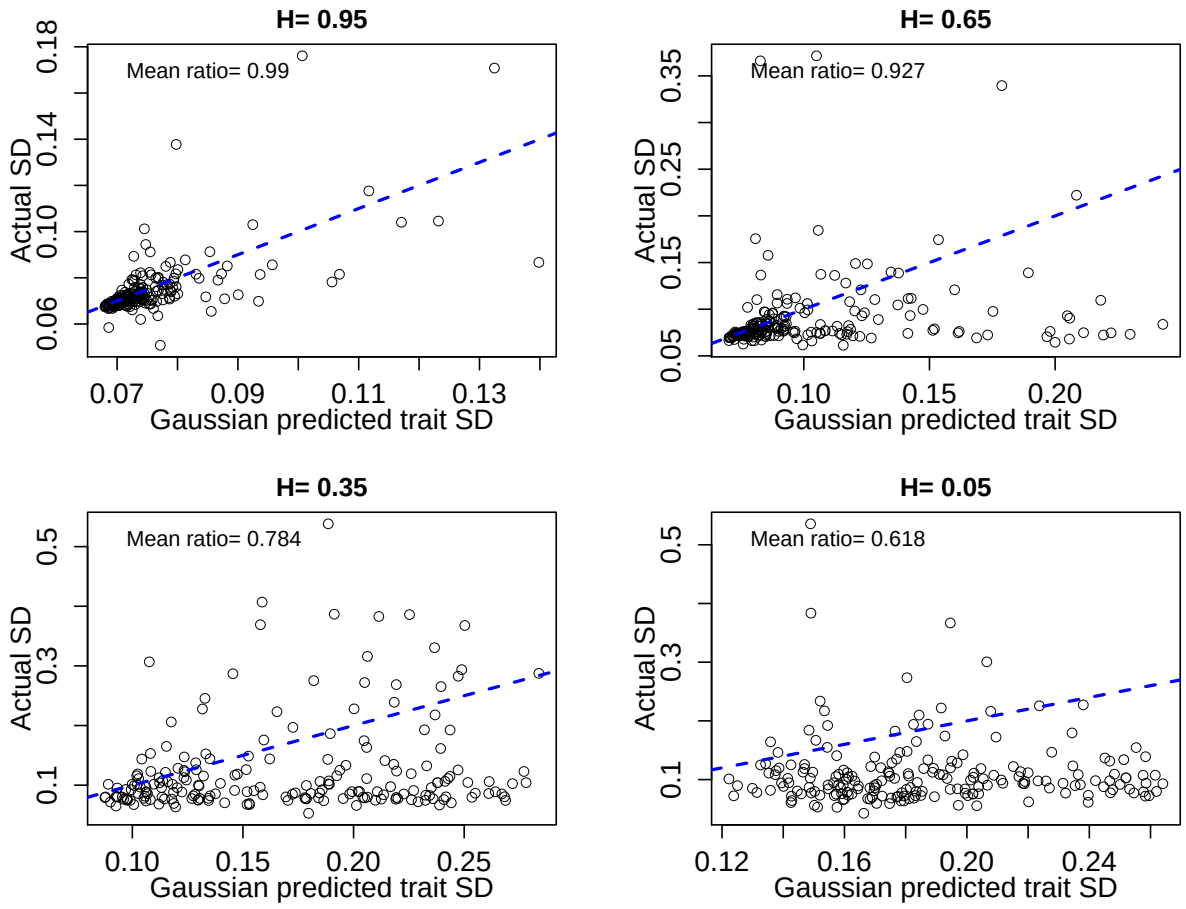

Figure A-2: Relationship between observed trait variance in selected parents, and the predicted trait variance assuming a Gaussian trait distribution, in the quantitative trait model with an environmental trend starting in year 500. Each panel shows the observed and predicted trait standard deviations in years 550 to 750; dashed blue line is the 1:1 line and "Mean ratio" is the average ratio between predicted and observed standard deviations in years 550 to 750. When $H \approx 1$ and $\gamma \approx 0$ the Gaussian approximation is unbiased, but with smaller H/higher γ the variance in selected parents is lower on average than the Gaussian approximation. Parameter values $\sigma_M = 1, \sigma_w = 0.3, \eta = 0.03, s_d = 0.8, s_a = 0.1, \sqrt{V_0} = 0.05$. Figure produced by script `QG-IBM-Test-C-Approx.R`.

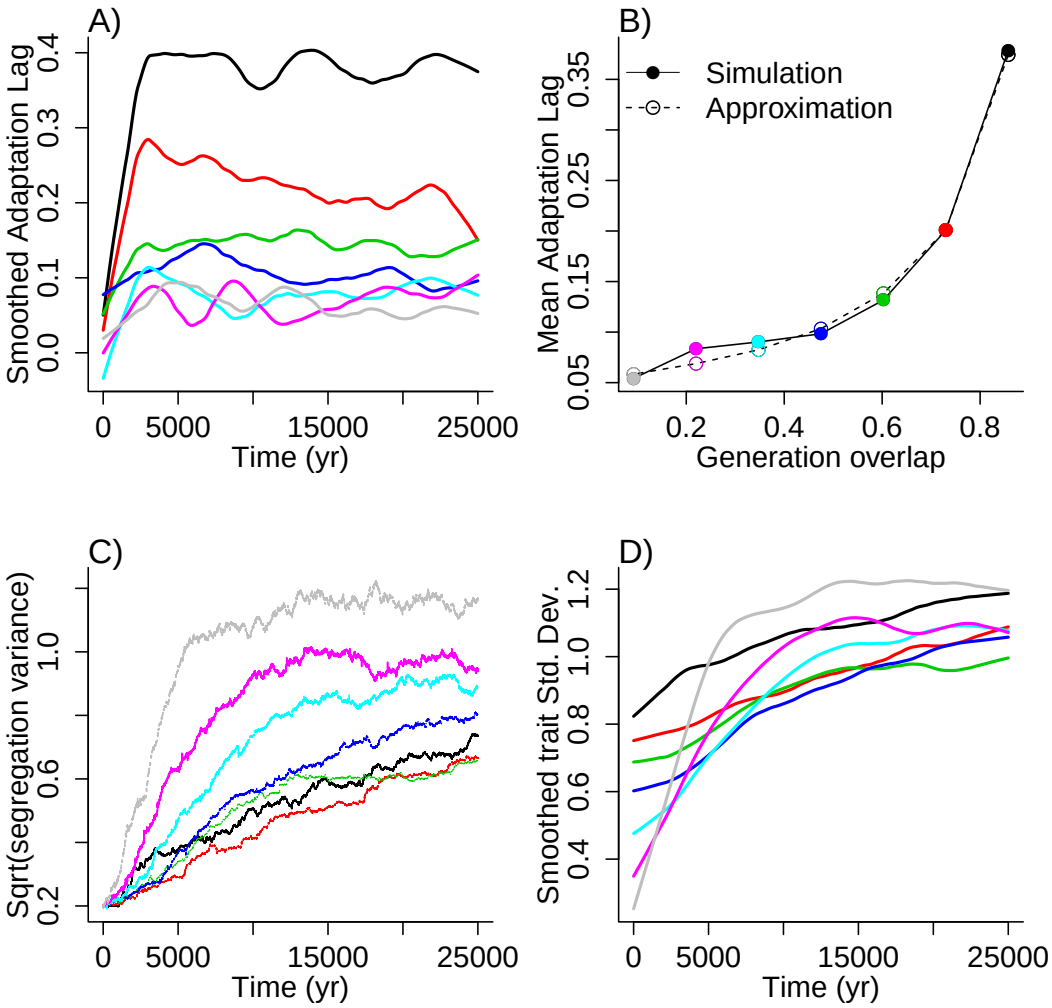

Figure A-3: As in Figure **??** with stronger selection and a faster environmental trend: $\sigma_w = 0.3, \eta = 0.05$. In panels A) and D) the yearly values have been smoothed to show the trends clearly using `loess.smooth` with degree = 1, span = 0.2. Figure produced by script `QG-IBM-TrendV0.R`.

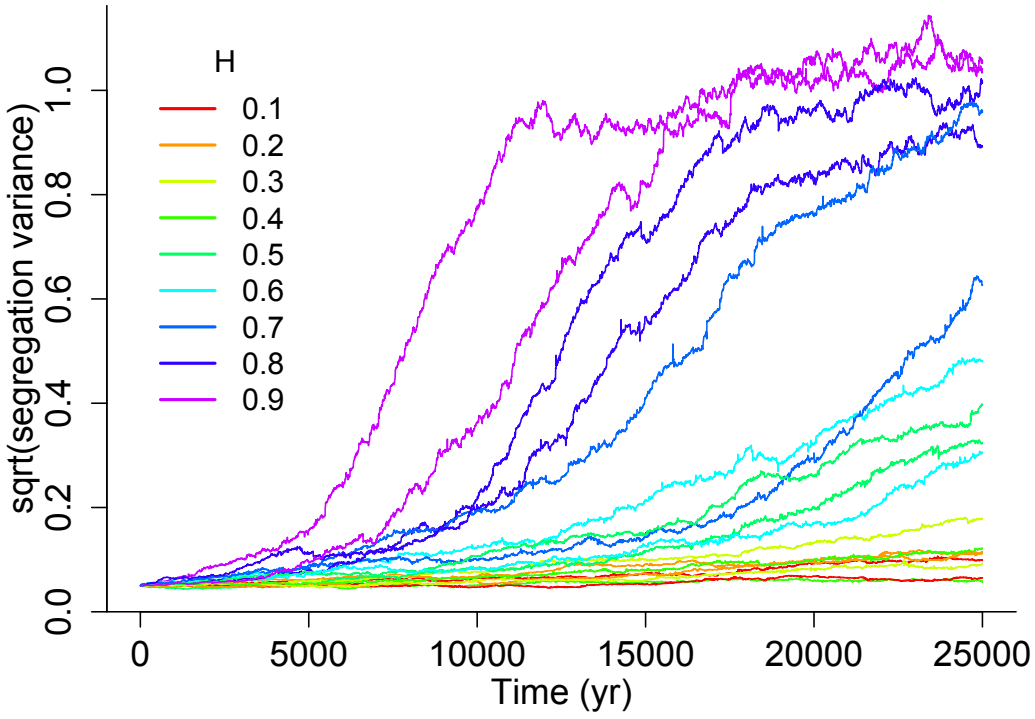

Figure A-4: Evolution of V_0 in a stationary environment starting from $\sqrt{V_0} = 0.05$. Two independent simulations are plotted for each value of H (indicated by the color of the curve). Other parameters for all simulations are: $\sigma_W = 0.3, \sigma_M = 1, s_d = 0.8, s_a = 0.1$. Figure produced by script `QG-IBM-noTrend.R`.

A.2 Trait variance in the strong selection limit of the quantitative trait model, stationary environment

In this Appendix we derive the assertion in the main text that eqn. (**??**) is the stationary trait variance in the strong selection limit of the quantitative trait model with a stationary environment.

The population at any time consists of current new recruits plus the survivors from all past cohorts of new recruits, in proportions $p_j = (1-\gamma)\gamma^j, j = 0, 1, \cdots$ where j is the number of years in the past. In the strong selection limit, all selected parents in year t have trait value (very close to) M_t. The cohort with index j (relative to the present) therefore has random trait mean m_j with the same distribution as M_t and variance V_0. Letting Z denote the trait in the entire population, and X_j the trait in component cohort j, we use the law of total variance (conditioning on j) to write

$$\mathrm{Var}(Z) = \mathbb{E}\left[\mathrm{Var}(X_j)\right] + \mathrm{Var}\left[\mathbb{E}(X_j)\right] = V_0 + \mathrm{Var}\left[\mathbb{E}(X_j)\right]. \tag{A.1}$$

$\mathrm{Var}\left[\mathbb{E}(X_j)\right]$is random, depending on the sequence of past M values, and our goal is to compute its expectation. Writing $\mathrm{Var}\left[\mathbb{E}(X_j)\right] = \sum_j p_j m_j^2 - \left(\sum_j p_j m_j\right)^2$, we have

$$\mathbb{E}\left(\sum_j p_j m_j^2\right) = \sum_j p_j \mathbb{E}(m_j^2) = \sum_j p_j \mathrm{Var}(M_t) = \sigma_M^2 \tag{A.2}$$

and, assuming that trait optima in different years are independent,

$$\mathbb{E}\left(\sum_j p_j m_j\right)^2 = \mathbb{E}\left(\sum_{i,j} p_i p_j m_i m_j\right) = \sum_{i,j} p_i p_j \mathbb{E}(m_i m_j) = \sum_j p_j^2 \mathbb{E}(m_j^2) = \sum_j p_j^2 \sigma_M^2. \tag{A.3}$$

As $1 - \sum_j p_j^2 = 2\gamma/(1+\gamma)$, combining the three equations above we get that $\mathbb{E}\left[\mathrm{Var}(Z)\right]$ is given by equation (**??**).

A.3 Equation (??) and its derivation

This Appendix provides the derivation of eqn. (**??**) referred to in the main text, which gives the conditions under which the approximate rate of response to selection, eqn. (**??**), is an increasing function of γ at $\gamma = 0$.

Eqn. (**??**) is proportional to $f = (1-\gamma)\bar{a}/\sqrt{2-(1-\gamma)\bar{a}}$. Our goal is to find when the derivative of f with respect to γ is positive at $\gamma = 0$, which implies that the mean selection response magnitude is maximized at some γ strictly between 0 and 1. The calculations described below were done using Maxima (online SI script `QG-IBM-noTrend.max`).

We substitute into f the definition of a in terms of σ_z^2, write

$$\sigma_z^2(\gamma) = d_0 + d_1\gamma + \frac{d_2\gamma^2}{2} + \cdots$$

where d_k is the k^{th} derivative of σ_z^2 at $\gamma = 0$, and take the derivative of f with respect to γ. The result is a product of always-positive terms with the factor

$$d_1\sigma_w^2 - d_0\sigma_w^2 - d_0^2. \tag{A.4}$$

To see when (**??**) is positive we need to find d_0 and d_1.

d_0 is $\sigma_z^2(0)$ which (in our approximations) is the solution to $(1+a)\sigma_z^2 = 2V_0$. Using the definition of a, this is found to be

$$d_0 = \frac{\sqrt{\sigma_w^4 + 12\,V_0\,\sigma_w^2 + 4\,{V_0}^2} - \sigma_w^2 + 2\,V_0}{4}. \tag{A.5}$$

To find d_1 we define G to be the difference between the left- and right-hand sides of (**??**), and apply the Implicit Function Theorem to $G(\gamma, \sigma_z^2(\gamma)) \equiv 0$:

$$d_1 = \frac{d\sigma_z^2}{d\gamma} = -\frac{\partial G/\partial\gamma}{\partial G/\partial\sigma_z^2}. \tag{A.6}$$

We then substitute into the expression for (**??**) $\gamma = 0$ and $\sigma_z^2 = d_0$ from (**??**). The resulting expression for d_1 is proportional to σ_M^2, $d_1 = \sigma_M^2 Q$. The condition for (**??**) to be positive is then that

$$\sigma_M^2 > \frac{d_0(1 + d_0/\sigma_w^2)}{Q}. \tag{A.7}$$

Define $u = \sigma_w^2/V_0$. Substituting this into the right-hand side of (**??**) (by setting $V_0 = \sigma_w^2/u$) reveals that the right-hand side has the form $\sigma_w^2 F(u)$, where

$$F(u) = \frac{\sqrt{u^2+12\,u+4}\,\left(3\,u^4 - 72\,u^3 - 168\,u^2 - 96\,u - 16\right) - 3\,u^5 - 74\,u^4 - 504\,u^3 - 592\,u^2 - 240\,u - 32}{8u^2[\sqrt{u^2+12\,u+4}\,\left(u^2 - 4\,u - 4\right) - u^3 - 2\,u^2 - 20\,u - 8]} \tag{A.8}$$

So finally, the condition we have been seeking is that $(\sigma_M/\sigma_w) > \sqrt{F(\sigma_w^2/V_0)}$. Numerically, $\sqrt{F(u)}$ remains between 2 and 4 for $0.36 \le u \le 49$. The critical ratio of σ_M/σ_w, above which the condition is satisfied, is therefore between 2 and 4 so long as $0.6 \le (\sigma_w/\sqrt{V_0}) \le 7$, as stated in the main text.

A.4 Adaptation lag in the strong selection limit of the quantitative trait model, trending environment

Here we derive the adaptation lag for the quantitative trait model with a trending trait optimum, in the steady-state situation where a steady environmental trend has been running for a long time ($\mu_t = \mathbb{E}(M_t) = \eta t$, $t \gg 1$).

As in a stationary environment, the population at any time consists of a cohort of new recruits plus the survivors from all past cohorts, in proportions $p_j = (1-\gamma)\gamma^j, j = 0, 1, \cdots$ where j is the number of years in the past. In the strong selection limit, all selected parents in year t have trait value M_t. The cohort with index j (relative to the present) therefore has random trait mean m_j and variance V_0. The mean of m_j is $\mathbb{E}(m_j) = (t-j)\eta$, and its variance is σ_M^2. The population (random) trait mean in year $t+1$ is therefore

$$\bar{z}_t = \sum_{j=0}^{\infty} (1-\gamma)\gamma^j m_j \tag{A.9}$$

which has expectation

$$\mathbb{E}(\bar{z}_{t+1}) = \sum_{j=0}^{\infty} (1-\gamma)\gamma^j (t-j)\eta = \eta t - \frac{\eta\gamma}{1-\gamma}. \tag{A.10}$$

The average adaptation lag is therefore

$$\bar{L} = \eta(t+1) - \mathbb{E}(\bar{z}_{t+1}) = \frac{\eta}{1-\gamma}. \tag{A.11}$$

A.5 Approximation of stationary σ_z^2 with trending trait optimum: quantitative trait model

In this Appendix we extend the derivation of eqn. (**??**), the implicit equation for $\bar{\sigma}_z^2$ in the quantitative trait model, to the case of a steadily trending trait optimum.

As with a stationary environment $\eta = 0$ we first need $\mathrm{Var}(\bar{z})$. Pretending that $a(t)$ is constant at its mean $\bar{a}$, we can write eqn. (**??**) as

$$\begin{aligned} \bar{z}(t+1) &= [1-(1-\gamma)\bar{a}]\bar{z}(t) + (1-\gamma)\bar{a}M_t \\ &= [1-(1-\gamma)\bar{a}]\bar{z}(t) + (1-\gamma)\bar{a}\mu_t + (1-\gamma)\bar{a}(M_t - \mu_t) \end{aligned} \tag{A.12}$$

Because $M_t - \mu_t$ and $\bar{z}_t$ are independent, taking the variance of both sides

$$\mathrm{Var}[\bar{z}(t+1)] = [1-(1-\gamma)\bar{a}]^2 \mathrm{Var}[\bar{z}(t)] + (1-\gamma)^2\bar{a}^2\sigma_M^2. \tag{A.13}$$

The steady-state $\mathrm{Var}(\bar{z})$ satisfies the last equation with time-dependence dropped, giving

$$\mathrm{Var}(\bar{z}) = \frac{(1-\gamma)\bar{a}\sigma_M^2}{2-(1-\gamma)\bar{a}}, \tag{A.14}$$

the same as with a stationary environment.

To use (**??**) in this case we approximate

$$\begin{aligned}
\mathbb{E}\left\{[M_t - \bar{z}(t)]^2\right\} &= \mathbb{E}\left\{[(M_t - \mu_t) + (\mu_t - \bar{z}(t))]^2\right\} \\
&= \mathbb{E}\left[(M_t - \mu_t)^2\right] + \mathbb{E}\left[(\mu_t - \bar{z}(t))^2\right] \text{ (as } M_t - \mu_t \text{ is indep. of } \bar{z}_t \text{ with mean 0)} \\
&= \sigma_M^2 + \left(\mathbb{E}\left[\mu_t - \bar{z}(t)\right]\right)^2 + \mathrm{Var}[\mu_t - \bar{z}(t)] \\
&\approx \sigma_M^2 + \bar{L}^2 + \mathrm{Var}[\bar{z}(t)] \\
&\approx \sigma_M^2 + \frac{\eta^2}{(1-\gamma)^2\bar{a}^2} + \frac{(1-\gamma)\bar{a}\sigma_M^2}{2-(1-\gamma)\bar{a}} \\
&= \frac{2\sigma_M^2}{2-(1-\gamma)\bar{a}} + \frac{\eta^2}{(1-\gamma)^2\bar{a}^2}.
\end{aligned} \tag{A.15}$$

Using (**??**) in (**??**), the stationary environment approximation (**??**) becomes

$$(1+a)\,\sigma_z^2 = 2V_0 \,+\, \frac{4\gamma\bar{a}^2\sigma_M^2}{2-(1-\gamma)\bar{a}} + \frac{2\eta^2\gamma}{(1-\gamma)^2} \tag{A.16}$$

A.6 Generation overlap and adaptation lag in the quantitative trait model

Here we report a small simulation experiment that was conducted to test our observation that adaptation lag in the quantitative trait model increases with generation overlap, all else being equal, unless the trend is very slow (script file `QG-IBM-Trend-Random.R`). We generated random parameter sets from the following distributions:

$$\begin{aligned}
\eta &\sim \mathrm{Uniform}(0.02, 0.1) \\
\sigma_w &\sim \mathrm{Uniform}(0.2, 2) \\
\sqrt{V_0} &\sim \mathrm{Uniform}(0.05, 0.5) \\
s_d &\sim \mathrm{Uniform}(0.7, 0.95) \\
s_a &\sim \mathrm{Uniform}(0.05, 0.3)
\end{aligned} \tag{A.17}$$

with $\sigma_M = 1$ (recall that the trait z can always be rescaled so as to make $\sigma_M = 1$). For each such parameter set, 6000-year simulations with a trend starting in year 1000 were done for two H values chosen from a $\mathrm{Uniform}(0.05, 0.95)$ distribution subject to the constraint that their difference

was at least 0.2. We computed the mean adaptation lag over the last 4000 years of each simulation, and checked whether the higher-γ simulation also had the higher adaptation lag. In 500 replicates there was one exception, which had $\eta = 0.023$ so that lags were very short relative to their variability. To check this exception we ran longer simulations at the exceptional parameter set, 15000 years with average lag computed over the last 13000 years. In two such replicates, the higher-γ simulation then had the higher adaptation lag.

A.7 Selection on trait variance in the quantitative trait model with trending optimum

In this Appendix we consider how trait skew affects selection on trait variance in the quantitative trait model.

? showed (their eqn. 23) that the quadratic selection differential $C = Cov[w, (z - \bar{z})^2]$ can be approximated, through Taylor expansion of fitness w, as

$$C = w_1 m_3 + \frac{w_2}{2}(m_4 - m_2^2) + \frac{w_3}{6}(m_5 - m_2 m_3) + \frac{w_4}{24}(m_6 - m_2 m_4) + \ldots \tag{A.18}$$

where $w_k = \partial^k w / \partial z^k$ evaluated at $\bar{z}$ and m_k is the kth central moment of z.

The first term on the right-hand side is the product of the fitness gradient (at the trait mean) and the skew. When the population mean lags behind a moving trait optimum, skew is negative and the fitness gradient will be positive on average (not necessarily at all times, because of the random fluctuations in M_t), resulting in a negative contribution to the quadratic selection differential. Countering that, in the tail of the fitness function the second derivative is positive, so the second term will be positive, especially when the kurtosis is large.

The fitness derivatives and terms in (**??**) are plotted in Figs. **??** and **??**. These come from the simulations plotted in Fig. **??**. Prior to the trend onset a positive first term is offset by a negative second term (the third and fourth are relatively unimportant most of the time). After the trend onset the second term becomes much smaller (presumably due to the variance shrinking, and the fourth moment shrinking even more), while the first term becomes strongly negative due to combination of negative skew in the trait distribution, and a positive first derivative of fitness because the population lags the moving optimum phenotype.

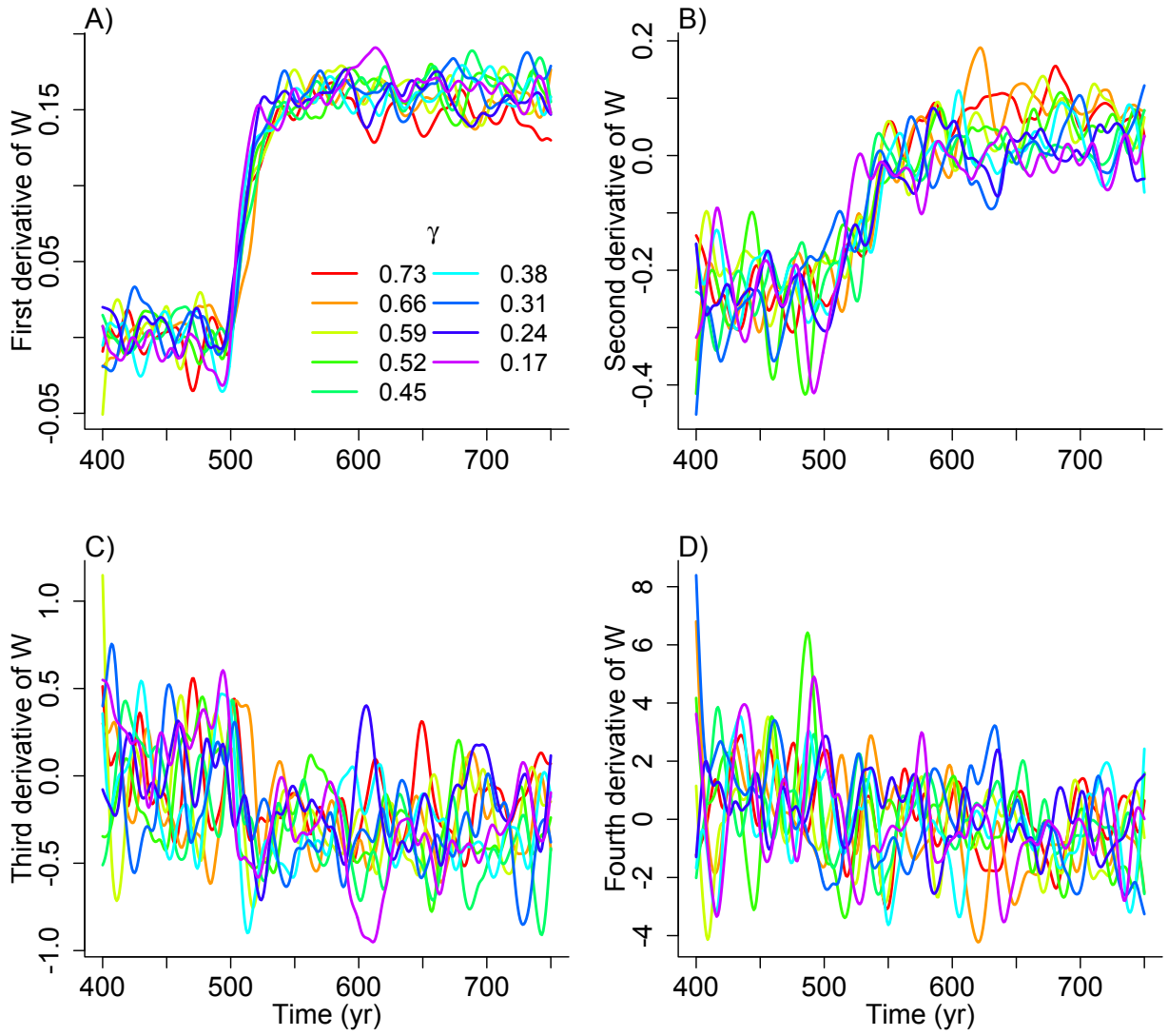

Figure A-5: Derivatives of relative fitness $R(z - M_t)$ evaluated at the trait mean $\bar{z}(t)$. These come from the same simulations plotted in Fig. **??**. As in that figure each curve is the average across 500 simulations with the same value of H, here smoothed (with a spline) to emphasize the trends. Figure produced by script `QG-IBM-Trend-Initial.R`.

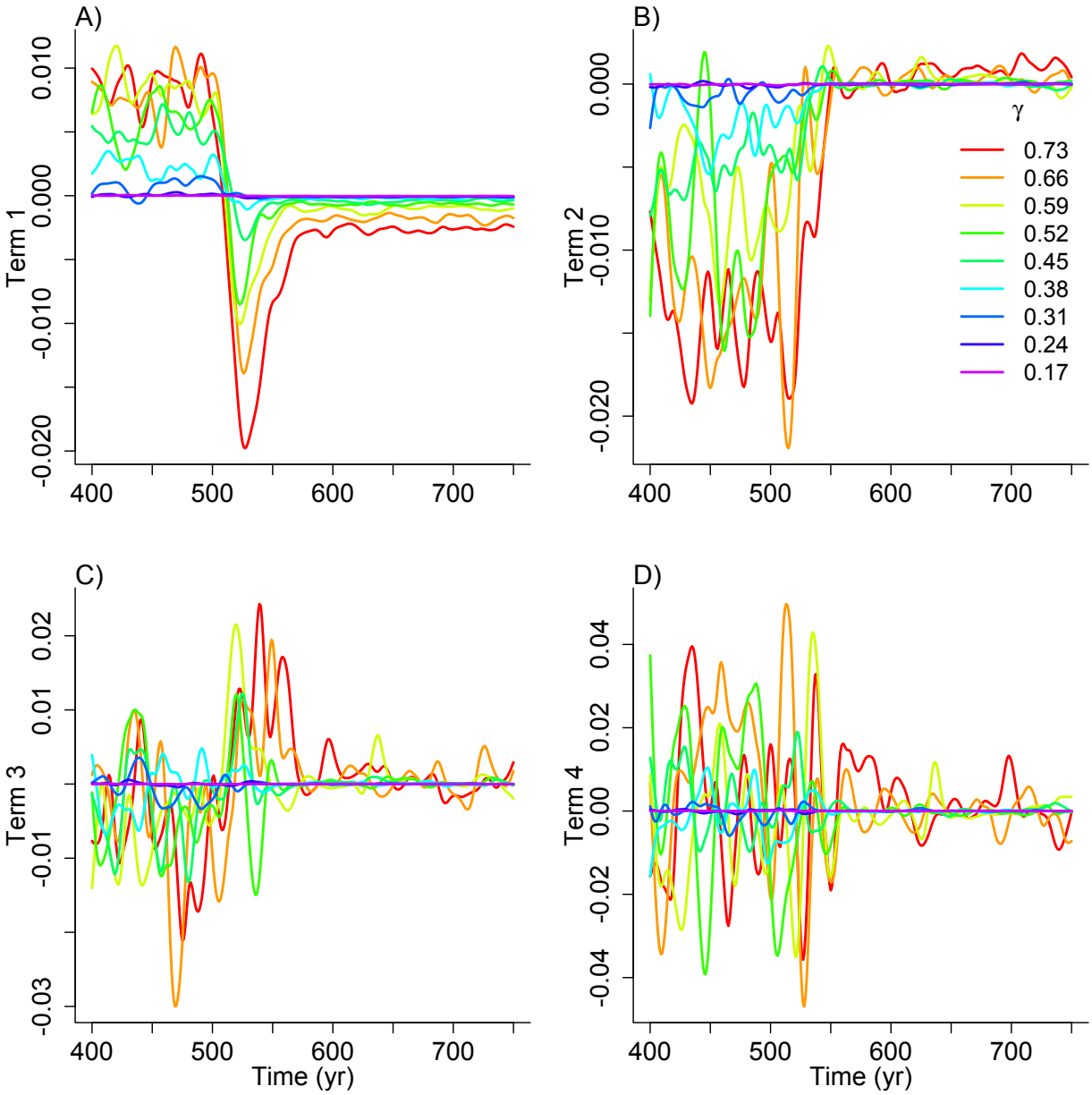

Figure A-6: The 4 terms in (**??**). These come from the same simulations plotted in Fig. **??**. As in that figure each curve is the average across 500 simulations with the same value of H, here smoothed (with a spline) to emphasize the trends. Figure produced by script `QG-IBM-Trend-Initial.R`.