Background selection in recombining genomes and its consequences for the maintenance of variation in complex traits

Xinyi Li[1] and Jeremy J Berg[1,2]

Affiliations

[1]Committee on Genetics, Genomics, & Systems Biology, University of Chicago, Chicago, IL, USA

[2]Department of Human Genetics, University of Chicago, Chicago, IL, USA

Abstract

Background selection (BGS)—the reduction of linked neutral diversity via the purging of deleterious mutations—is a pervasive force in genomic evolution. However, its impact on complex phenotypes remains poorly understood because classical theory treats fitness effects as fixed rather than emerging from a phenotype-to-fitness map. Here, we investigate the impact of BGS across three phenotypic selection frameworks: exponential directional, a liability threshold model, and stabilizing selection. First, we develop an effectively non-recombining block approximation for the site frequency spectrum (SFS) and show that this framework accurately describes the skew in the site frequency spectrum in the weak mutation regime typical of humans. Second, we show that phenotypic impacts of BGS depend on how selection is coupled across loci. In the liability threshold model, strong synergistic epistasis generates a global compensation mechanism—driven by tiny shifts in the mean phenotype—that propagates BGS effects to strongly selected variants otherwise immune to linked selection. This coupling reduces genetic variance across almost the entire effect-size distribution by an amount determined by a non-linear average of local effective population size reductions across the genome. Conversely, under stabilizing selection, BGS can counter-intuitively increase genetic variance. This occurs because BGS shifts strongly selected sites into the weakly selected underdominant regime where they persist at intermediate frequencies longer than in the equivalent directional selection model. Our results inform both longstanding evolutionary conversations regarding synergistic epistasis and efforts to model the impact of background selection on individual variants in recombining genomes.

Significance

Natural selection against deleterious mutations reshapes genetic variation through background selection, a process typically modeled as a local reduction in effective population size. We show that the consequences of this reduction for genetic architecture are mediated by the relationship between phenotype and fitness. Extending background selection theory to complex traits, we show that in models with strong synergistic epistasis, a global compensation mechanism propagates the effects of selection to strongly selected variants. Furthermore, we find that background selection can counter-intuitively increase genetic variation under stabilizing selection. Our work shows that the total impact of background selection on functional variation thus depends on the answers to longstanding questions about the ultimate sources of variance in fitness.

Introduction

Natural selection acting on deleterious mutations reduces genetic variation at linked neutral sites, a process known as background selection (BGS; Charlesworth et al. 1993). In the decades since its initial description, BGS has become a standard component of population genetic null models. To a first approximation, BGS acts as a local reduction in the effective population size ($B = {}^{N_e}/_{N}$), which reduces the expected number of segregating sites and increases the rate of genetic drift (Hudson and Kaplan 1995; Nordborg et al. 1996; McVicker et al. 2009). The magnitude of this reduction depends largely on the local density of functional sites relative to the recombination rate and varies substantially across species. For example, in humans, BGS effects are generally modest ($B \approx 0.6$ in the most affected regions, with $B \approx 0.83$ on average; Murphy et al. (2022) and Buffalo and Kern (2024)), whereas in the more compact *Drosophila* genome, the average reduction is much stronger (mean $B \approx 0.3$; Comeron (2014) and Elyashiv et al. (2016)). The existence of broad-scale correlations between recombination rate and diversity across taxa suggests that these linked selection effects are a pervasive feature of genomic evolution across the tree of life (Corbett-Detig et al. 2015; Coop 2016; Buffalo 2021).

Despite this progress, our understanding of how BGS impacts the evolution of complex phenotypes remains limited. Empirical studies have suggested that BGS shapes the genetic architecture of traits. For example, Rockman et al. (2010) observed that in *C. elegans*, loci in gene-poor, high-recombination regions explain significantly more variance in gene expression than those in regions with higher gene density and lower recombination rate. Similarly, in humans complex traits, per-SNP heritability appears enriched in regions of strong background selection (Gazal et al. 2018; Pardiñas et al. 2018). However, theoretical work has largely been restricted to single-site models in which the fitness effects of individual loci are fixed (Charlesworth 1994; Stephan et al. 1999). However, assuming that the fitness effects of variants are largely mediated by their contributions to polygenic traits, their fitness

effects are not fixed, but a emerge from the distribution of phenotypes in the population and the relationship between phenotype and fitness (Simons et al. 2018; Berg et al. 2025). Consequently, determining the full impact of background selection requires understanding how these local linkage effects interact with the global, polygenic constraints that ultimately generate fitness.

Even at the level of individual sites, the standard "reduced local N_e" approximation provides an incomplete picture. While it accurately predicts the total heterozygosity and fixation rate, it fails to capture the distortion of the site frequency spectrum (SFS) relative to a simple population size reduction (Gordo et al. 2002). This distortion arises because BGS does not act instantaneously; alleles must persist long enough to "feel" the purging of linked deleterious backgrounds. This delay creates a skew in the SFS that can confound both demographic inference and inferences about the distribution of fitness effects (Johri et al. 2021). Prior work has provided an analytical description of this skew for non-recombining regions in the strong mutation regime (Cvijović et al. 2018) . Here, we develop an effectively non-recombining block approximation for the dynamics of BGS in recombining genomes, and use it describe the skew in the frequency spectrum for in the weak mutation limit, and show that is valid broadly in the parameter regime relevant to humans.

We then study the impact of BGS in three models of phenotypic mutation-selection-drift balance. We first contrast two models of directional selection: an exponential fitness model, where loci evolve independently, and a liability threshold model, where fitness effects are tightly coupled across loci. We show that in the threshold model, the system responds to BGS through a global compensation mechanism driven largely by the genomic regions experiencing the strongest background selection. Unlike simple local N_e rescaling—which impacts only weakly selected sites—this global mechanism propagates the effects of BGS to strongly selected variants that would otherwise be immune to linked selection. Finally, we consider a model of stabilizing selection, the ubiquity of which is motivated by both direct (Karn and Penrose 1951; Sanjak et al. 2018) and indirect (Smith 1983; Simons et al. 2022; Koch et al. 2024) lines of evidence. We find that due to the specific dynamics of underdominant selection maintained by stabilizing selection, BGS can, counter-intuitively, lead to an *increase* in the genetic variance of quantitative traits.

Results

Our results are presented in three parts. First, we consider a model of BGS acting on a single focal site in a recombining background. We focus on modeling the skew in the frequency spectrum under moderate BGS strengths and resolving the distinct effects of BGS on alleles under weak versus strong direct selection. Second, we apply results from this single site analysis to models of polygenic directional selection. We contrast a 'decoupled' exponential fitness model, where BGS effects are entirely local, with a 'coupled' threshold model, where strong epistasis forces a global, compensatory response. Finally, we analyze a stabilizing selection model to demonstrate how a different form of

epistasis leads to a distinct, local-only BGS dynamic, which can lead to increases in genetic variance under some conditions.

Background selection in single site models

The frequency spectrum

Neutral alleles We begin by considering a diploid population of size N, focusing on an individual site embedded in a recombining chromosome of length M. Adjacent sites recombine with rate r, and each produces deleterious mutations (with selection coefficient s_b) at rate ν. We further assume selection against these mutations is strong relative to drift (i.e., $2Ns_b \gg 1$, ensuring they are purged and never fix), but weak relative to the chromosome's total recombination rate (i.e., $s_b \ll Mr$). For the moment, we suppose this site is neutral, but will relax this shortly.

Under these assumptions, the impact of BGS can be understood through an "effectively non-recombining linkage block" approximation (see Supplementary Text A; Neher et al. 2013; Good et al. 2014; Weissman and Hallatschek 2014). This model approximates the background as a single non-recombining block of characteristic length $M_{s_b} = {}^{2s_b}/_{r}$, a scale set by the balance between the purging of deleterious mutations and the erosion of the haplotype by recombination. Because the expected frequency of a deleterious allele at any background site is ${}^{\nu}/_{s_b}$, this characteristic block carries an average of

$$\lambda = M_{s_b}\frac{\nu}{s_b} = 2\frac{\nu}{r}. \tag{1}$$

deleterious alleles (Eq. (A.4)).

While it is the burden of deleterious mutations within the characteristic length governs the classic BGS process, newly arising mutations also exert an influence in early generations. Immediately upon arising, the focal allele becomes associated with an average of λ additional deleterious mutations distributed over a physical distance much larger than the characteristic length-scale, an expectation that remains approximately constant for short times due to the balance between mutation and recombination. In Supplementary Text A, we develop a heuristic "shrinking block" approximation for the impact of this additional fitness drag, and describe how it differs from the analogous process of fitness decline in a non-recombining model.

We find that if mutation is weak relative to recombination (more precisely, if $\lambda \lesssim {}^{1}/_{2}$), then the characteristic haplotype block does not accumulate enough fitness drag to have an impact before selection can act against the existing background. Consequently, if a focal allele is eliminated by selection against a linked deleterious allele, it is almost certainly due to an allele that was already present when it arose, and so mutations outside the characteristic length-scale are irrelevant. Under these assumptions, the expected deleterious load on a haplotype block of length M_{s_b} at frequency q is

exponentially suppressed to $\lambda_q = \lambda e^{-4Ns_b q}$. To reach frequency q, a focal allele must therefore reside on a background that carries, on average, $\lambda - \lambda_q = \lambda(1 - e^{-4Ns_b q})$ fewer deleterious mutations than a background chosen at random. This purging process reduces the number of alleles that persist to frequency q by a factor equal to the survival probability of the background:

$$B\left(q\right) = e^{-\lambda\left(1-e^{-4Ns_b q}\right)} \tag{2}$$

(see Eq. (A.26)), which converges to the standard "reduced effective population size" approximation, $B = e^{-\lambda}$, for for $q \gg 1/4Ns_b$ (Hudson and Kaplan 1995; Nordborg et al. 1996). Thus, in the weak mutation limit, the skew in the frequency spectrum mirrors the exponential suppression of frequencies among the deleterious alleles responsible for BGS (i.e. $e^{-4Ns_b q}$). This approximation can be straightforwardly extended to a distribution over s_b by averaging the exponential suppression term, giving the frequency-dependent reduction factor:

$$B\left(q\right) \approx e^{-\lambda\left(1-\mathbb{E}_{s_b}\left[e^{-4Ns_b q}\right]\right)} \tag{3}$$

(see Eq. (A.27)). While the assumptions justifying these approximations are only formally satisfied in the limit as $\lambda \to 0$, in practice we find that they provide a good fit for the $\lambda \lesssim 1/2$ (see Figure S1 and Figure 1 below), which corresponds to $B \geq 0.6$, encompassing roughly the full range of background selection effects observed in humans (Murphy et al. 2022; Buffalo and Kern 2024).

Directly selected alleles Now, suppose the focal site itself produces alleles experiencing additive direct selection with coefficient s, and let $\xi\left(q \mid \gamma\right) = \frac{2}{q(1-q)} \frac{1-e^{-2\gamma(1-q)}}{1-e^{-2\gamma}}$ be the number of generations that such an allele spends at frequency q during its transit through the population, where $\gamma = 2Ns$ is the scaled selection coefficient (Wright 1938; Bustamante et al. 2001).

When direct selection on the focal allele is weak relative to background selection ($|\gamma| \ll 2Ns_b$), it influences dynamics only at high frequencies ($q \gg 1/4Ns_b$), where the initial cull of alleles arising on deleterious backgrounds is complete ($B(q) = e^{-\lambda}$). This convergence reflects a separation of timescales: because haplotypes carrying deleterious alleles are rapidly purged, any focal allele reaching high frequency must have arisen on a background that is free of deleterious mutations over the characteristic length-scale (Charlesworth 1994; Stephan et al. 1999). A new, weakly selected mutation is thus expected to spend

$$\xi_B\left(q \mid \gamma\right) \approx B\left(q\right) \xi\left(q \mid \gamma B\right) \tag{4}$$

generations at frequency q (see Eq. (A.30)). Consequently, for deleterious alleles under direct selection weaker than the background intensity, BGS reduces sojourn times at low frequencies but increases them at high frequencies. Notably, this effect persists even for alleles that are strongly selected relative to drift ($|\gamma| \gg 1$), provided they satisfy $|\gamma| \ll 2Ns_b$ (Figure 1B). For beneficial alleles, these effects

compound, further suppressing the number of alleles that reach high frequencies (Figure S2).

Finally, when direct selection is strong ($|\gamma| \gtrsim 2Ns_b$), direct and background selection act on similar timescales. The allele's fate is determined by its own fitness effect before the background is purged, suggesting the alternative approximation

$$\xi_B\left(q \mid \gamma\right) \approx B\left(q\right)\xi\left(q \mid \gamma\right). \tag{5}$$

(Eq. (A.35)). In practice, however, these strongly selected alleles rarely reach the high frequencies where background selection takes effect ($q \gg 1/4Ns_b$). Consequently, for the vast majority of variants in this regime, $B(q) \approx 1$ and BGS is irrelevant. While we do observe some distortions in the frequency spectrum for $q \gtrsim 1/4Ns_b$ (Figure 1C), these rare variants contribute minimally to standing variation.

Phenotypes under directional selection

Heterozygosity and fixation rates at individual sites

We now turn to models of mutation-selection-drift balance (MSDB) for polygenic phenotypes. We analyze these models using a mean-field approach, approximating the evolution of the phenotypic distribution by the average behavior of a single site. The key features of this behavior are captured by two quantities: an allele's fixation probability and its expected lifetime contribution to heterozygosity.

As shown by Kimura (1969), these two quantities are intimately related, as the lifetime contribution to heterozygosity is:

$$h_d(\gamma) = 2\left(\frac{\rho(\gamma) - 1}{\gamma}\right), \tag{6}$$

where $\rho(\gamma) \approx 2\gamma/(1-e^{-2\gamma})$ is the expected number of fixations per $2N$ new mutations, and the quantity $\rho(\gamma) - 1$ is the "selective fixation rate" (the number of fixations driven or prevented by selection). To model the impact of background selection, we employ the standard reduced effective size approximation ($B = e^{-\lambda}$). While equation (4) breaks down for $B \lesssim 0.6$, this approximation remains accurate for the heterozygosity provided that $2NBs_b \gg 1$—a condition that ensures local effective size reductions are not severe enough to push sites into the interference selection regime (Good et al. 2014). This means that our phenotypic results should apply in the presence of much greater effective size reductions than our simple approximation for the skew.

Applying this approximation, the reduction in the lifetime contribution to heterozygosity due to BGS is precisely equal to the effect on the selective fixation rate:

$$B_\gamma = B\frac{h_d(\gamma B)}{h_d(\gamma)} = \frac{\rho(\gamma B) - 1}{\rho(\gamma) - 1} \tag{7}$$

where the factor of B represents the probability of surviving the initial low-frequency cull, and $h_d(\gamma B)/h_d(\gamma)$ represents the effect on the lifetime contribution to heterozygosity conditional on sur-

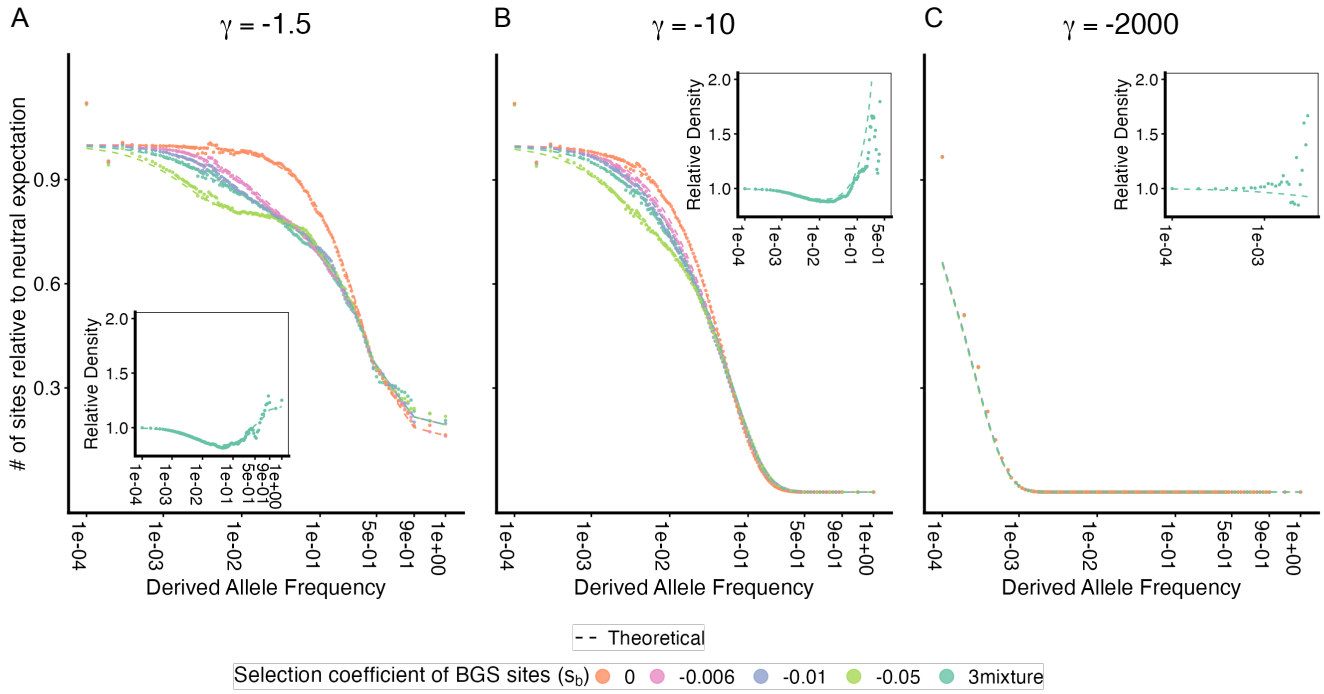


Figure 1: **The frequency spectrum of focal selected alleles in the weak mutation limit ($\lambda \ll 1$).** The site frequency spectrum observed in simulations (points) are compared to theoretical predictions (lines) for focal alleles under varying strengths of direct selection, assuming a background selection intensity of $B \approx 0.82$ (the human average). The site frequency spectrum is scaled relative to the standard neutral expectation (θ/q). Focal alleles are (A) nearly neutral ($2Ns = -1.5$), (B) moderately deleterious ($2Ns = -10$), and (C) strongly deleterious ($2Ns = -2000$). Different colors represent simulations where the same total reduction in effective population size (B) is achieved using different selection coefficients or a distribution with an equal mixture of all three for the background deleterious mutations (s_b). The dashed lines represent the approximation developed in Equation 4 for panels A and B, and Equation (5) for panel C. Insets show the relative density of alleles across frequency bins, comparing scenarios with background selection to those without, with theoretical prediction plotted in dashed line.

viving this cull (Figure 2A). Notably, this relationship holds even for strongly selected sites ($\gamma \ll -1$) where the physical mechanism differs. In the regime where direct selection is strong relative to drift but weak relative to background selection ($2Ns_b \gg |\gamma| \gg 1$; Figure 1B), BGS significantly distorts the frequency spectrum: it reduces the number of alleles that survive the initial purge, but allows the survivors to drift to higher frequencies due to weakened effective selection. These opposing effects cancel out, leaving the total heterozygosity unchanged ($B_\gamma \approx 1$). Conversely, when direct selection dominates background selection ($|\gamma| \gg 2Ns_b$; Figure 1C), the focal allele is purged by its own fitness

cost before the background can be removed. In this limit, BGS is irrelevant, and the heterozygosity is naturally unaffected. We can therefore rely on this reduced local N_e approximation to model the underlying allelic dynamics across the full range of selection coefficients.

Phenotype model and chromosome structure

To apply these insights in the context of phenotypic evolution, we assume a simple additive model where each site ℓ has two alleles differing by effect a_ℓ. Individual i's trait value is

$$Z_i = G_i + E_i = \sum_\ell a_\ell g_{i\ell} + E_i \tag{8}$$

where $g_{i\ell} \in \{0, 1, 2\}$ is the genotype count and $E_i \sim N(0, V_E)$. We assume that causal sites are sufficiently spaced such that their local evolutionary dynamics can be treated independently. Each site is embedded in a recombining background that generates unconditionally deleterious alleles according to the model described above; however, we allow the density of these mutations per unit recombination (i.e. $\lambda = 2\nu/r$)—and thus the magnitude of the local effective population size reduction (i.e. $B = e^{-\lambda}$)—to vary across causal sites.

Exponential selection

The simplest model of directional selection is one in which fitness declines exponentially with increasing phenotype, i.e., $W(Z_i) \propto e^{-Z_i/\eta}$ (Figure 2B). In this model, sites evolve independently with an approximately constant directional selection coefficient $s \approx a\eta^{-1}$, corresponding to a scaled selection coefficient $\gamma = 2Ns \approx 2Na\eta^{-1}$ (see Supplementary Text B).

At mutation-selection-drift equilibrium, fixations of trait-increasing and trait-decreasing mutations reach a detailed balance (Li 1987; Sella and Hirsh 2005; Rice et al. 2015). Among sites with scaled selection coefficient γ, the fraction of sites fixed for the trait-increasing (deleterious) allele is determined by the ratio of fixation rates,

$$p^+(\gamma) = \frac{\rho(-\gamma)}{\rho(\gamma) + \rho(-\gamma)} = \frac{1}{1 + e^{2\gamma}}. \tag{9}$$

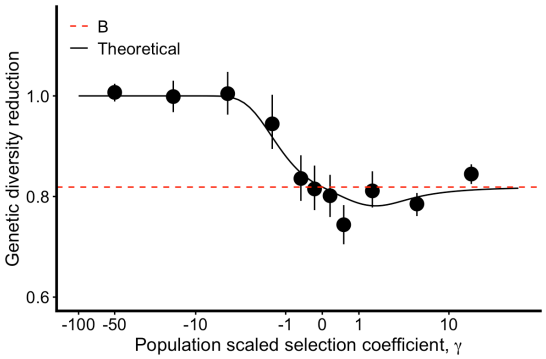
The resulting genetic variance refelcts a combination of contributions from alleles under positive and negative selection

$$V_G = \theta \left[p^-(\gamma) h_d(-\gamma) + p^+(\gamma) h_d(\gamma) \right] = a^2 \frac{\theta}{\gamma} b(\gamma), \tag{10}$$

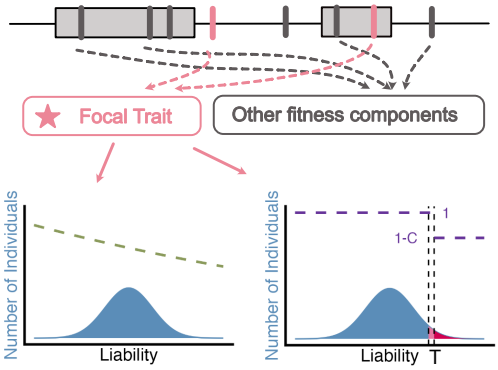
where $b(\gamma) = {}^{(\rho(\gamma)-\rho(-\gamma))}/_{(\rho(\gamma)+\rho(-\gamma))} = \tanh(\gamma)$ measures the asymmetry of the mutational input at equilibrium.

Because of the independence among sites, the impact of background selection on each site is

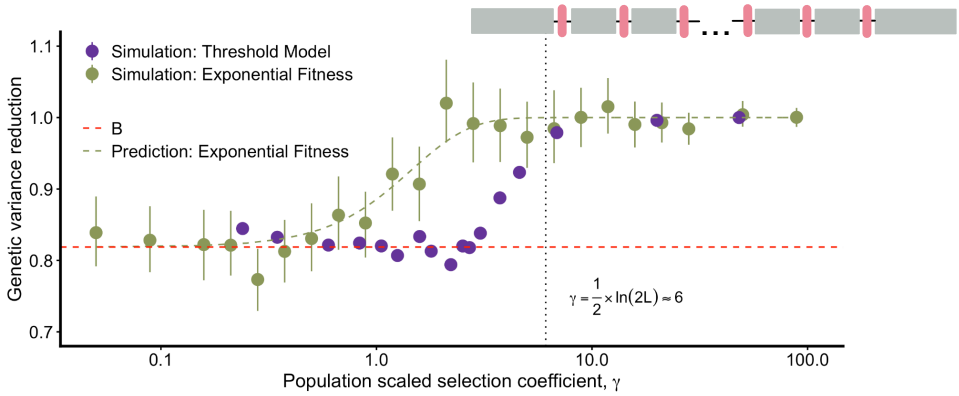
A. Genetic diversity reduction for single-site model



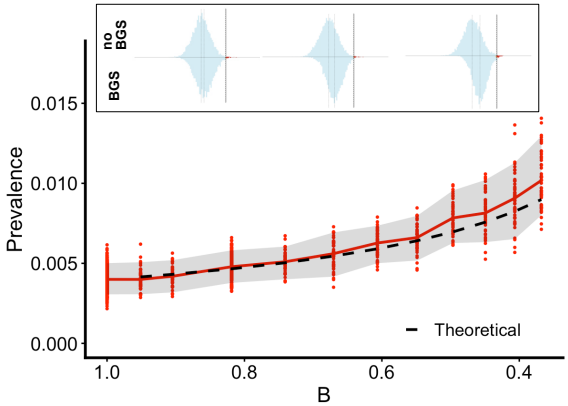
B. Schematic of focal traits vs. linked deleterious selection



C. Genetic variance reduction for two trait models



D. Disease prevalence under BGS ($p_T \gg 0$)



E. Disease prevalence under BGS ($p_T \approx 0$)

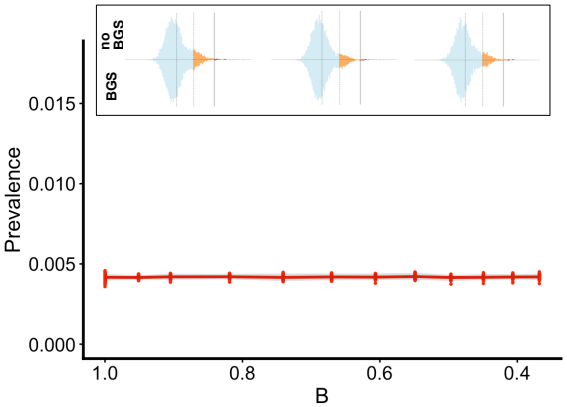


Figure 2: The impact of background selection on genetic variation under directional selection. (A) The reduction in genetic diversity for single-site additive directional selection model, represented as the ratio of expected heterozygosity with BGS to that without BGS ($h_d(\gamma B)/h_d(\gamma)$), is plotted against the population-scaled selection coefficient (γ). Simulation results (points) are shown alongside the theoretical prediction from equation (7) (solid line). The horizontal red dashed line indicates the expected diversity reduction for a neutral allele, $B \approx 0.82$. (B) A schematic depicting the relationship between sites that are causal to a focal trait (pink bars) and linked deleterious mutations that contribute to other fitness components (gray bars), which are responsible for BGS effects. (C) We compare the genetic variance reduction ($V_{G,B}/V_G$) for the exponential fitness model (green) and the single effect size liability threshold model (purple). Simulation results are plotted as points, and the prediction for the exponential fitness function from equation 11 is plotted as a green dashed line. The horizontal red dashed line indicates the expected diversity reduction for a neutral allele, $B \approx 0.82$. The vertical dotted line marks $\gamma \approx 6$, which corresponds roughly to $p_T \approx 1/2L$ given that we use $L = 10^5$ to simulate. The chromosome structure we used to simulate the threshold model is depicted at right top, showing trait loci (pink bars) interspersed among BGS sites (gray regions). (D-E) Disease prevalence under BGS is shown for scenarios where (D) the equilibrium fraction of liability increasing fixations is large ($p_T \gg 0$), versus (E) scenarios where liability increasing fixations are absent ($p_T \approx 0$). In (D), the prediction from equation 17 is plotted in dashed line. In the inset, the liability distribution in one simulation is plotted comparing BGS to no BGS at background selection intensities of approximately $B \approx (0.95, 0.67, 0.45)$ from left to right.

captured entirely by the local reduction in effective population size. This leads to a rescaled fixation asymmetry $p^+(\gamma B) = (1 + e^{2\gamma B})^{-1}$, which drives a large shift in the mean phenotype. For weakly selected sites, this shift is substantially larger than the scale of the standing genetic variation (i.e., $(\bar{G}_B - \bar{G})/\sqrt{V_G} \gg 1$), resulting in a significant decline in the mean fitness of the population (see Supplementary Text B).

The impact on the variance follows directly from the relationship between heterozygosity and fixation rates established in Eq. (7). Specifically, the variance is reduced by a factor

$$\frac{V_{G,B}}{V_G} = \frac{b(\gamma B)}{b(\gamma)} = \frac{\tanh(\gamma B)}{\tanh(\gamma)} \tag{11}$$

(Figure 2C; McVean and Charlesworth 1999). This reduction is greatest for effectively neutral alleles ($V_{G,B}/V_G \approx B$ for $\gamma \ll 1$) but weakens as selection strength increases, vanishing entirely for strongly selected sites ($\gamma \gg 1$). This fading reflects the fact that strong negative selection purges alleles faster than BGS can remove the background, rendering the reduction in N_e irrelevant. Critically, because the selection coefficient s in this model is fixed by the fitness function parameter η, it is unaffected by the shift in the mean phenotype. As a result, all sites respond to their local BGS effects independently, and the effect on the phenotype in any given case is simply found by averaging these effects over the joint distribution of effect sizes and local effective size reductions.

Threshold epistasis

In this model, fitness depends on whether the phenotype exceeds a threshold, T. Individual i's relative fitness is $W(Z_i) \propto 1 - \mathbb{1}[Z_i > T]C$ where C is the cost of exceeding the threshold and $\mathbb{1}[Z_i > T]$ indicates whether $Z_i > T$ (Figure 2B). We treat this as a model of complex disease, where Z_i represents an individual's liability (Berg et al. 2025). Equivalent threshold models have appeared in other contexts, where the phenotype is simply termed "fitness potential" (Milkman 1978; Kondrashov 2018). We adopt the complex disease framework here, though our results could easily be applied in other contexts by relabeling terms.

A single effect size and a single B value The threshold model differs from the exponential model in two fundamental ways, which are easiest to see if we first assume a single effect size, a, and constant local reduction in effective population size, B, across sites. First, the threshold imposes strong epistasis: a site's selection coefficient depends on the variation at all other sites. However, because the background phenotypic variation smoothes out the sharp epistasis of the threshold, the marginal effect of an allele on fitness is nonetheless constant, allowing individual sites to evolve as if under additive selection. If the liability effect a is small relative to the variance in liability ($a \ll \sqrt{V_P}$), this dependency is mediated solely by the liability density at the threshold, $f(T)$:

$$s = \delta C \approx a f(T) C, \tag{12}$$

where $\delta \approx a f(T)$ is the site's average effect on disease risk (Berg et al. 2025).

Second, because the threshold has a fixed position along the liability continuum, it determines the fraction of sites fixed for the liability-increasing allele $p^+(\gamma) \approx p_T = \frac{T}{2La}$. Consequently, the threshold position determines the scaled selection coefficient via the relationship between the fixation asymmetry, i.e., $p^+(\gamma)$, the fixation rates $\rho(\gamma)$ and γ (Eq. (9)). In a single-effect model, we can solve directly for:

$$\gamma \approx \frac{1}{2} \ln \frac{1 - p_T}{p_T} \tag{13}$$

(Berg et al. 2025). Combining this with the relationship between threshold density $f(T)$ and the selection coefficient (Eq. (12)), equilibrium is established when the mean liability is positioned at a distance just below the threshold, such that $f(T) \approx \frac{1}{4NCa} \ln \frac{1-p_T}{p_T}$. Because sites evolve under additive selection, genetic variance depends on γ exactly as in the exponential model (Eq. (10)); however, γ is now determined by the threshold position rather than the rate of fitness decline.

In the threshold model, strong selection against individuals with liability greater than T prevents the sort of large shift in the mean that occurs in the exponential model. Consequently, the system must maintain the equilibrium fraction of liability increasing fixations ($p^+(\gamma) \approx p_T$). As a result, BGS must reduce the selective fixation rates for liability increasing and liability decreasing mutations by the same fraction. This constraint imposes a specific requirement on the dynamics: the reduction in

genetic variance must arise *solely* from the initial culling of new mutations (i.e., the factor B). Recall from Equation (7) that the total impact of BGS is the product of this initial cull and the change in the per-mutation contribution to variance ($h_d(\gamma_B)/h_d(\gamma)$). Therefore, for the system to maintain equilibrium via the initial cull alone, the per-mutation contribution must remain unchanged. This implies that the scaled selection coefficient itself must be invariant, i.e. $\gamma_B = \gamma$, which then satisfies the initial requirement that the selective fixation rates are modified by the same factor, i.e.

$$\frac{B\left(\rho(\gamma_B) - 1\right)}{(\rho(\gamma) - 1)} = \frac{B\left(\rho(-\gamma_B) - 1\right)}{(\rho(-\gamma) - 1)}. \tag{14}$$

The system satisfies this requirement through a global compensation mechanism. In response to the local increase in genetic drift, the mean liability shifts closer to the threshold by a tiny amount, increasing the threshold density to $f_B(T) \approx f(T)/B$. This global adjustment increases the unscaled selection coefficients ($s_B \approx s/B$), which exactly offsets the local reduction in effective population size ($N \to NB$):

$$\gamma_B = 2N_B s_B = 2(NB)\frac{s}{B} = 2Ns = \gamma. \tag{15}$$

Because γ is invariant, the per-mutation variance $h_d(\gamma)$ remains unaffected, as predicted. The total reduction in genetic variance is therefore driven entirely by the reduced supply of mutations ($\theta \to \theta B$), resulting in a simple, γ-independent reduction:

$$\frac{V_{G,B}}{V_G} = B. \tag{16}$$

Thus, the threshold model avoids the large decrease in mean fitness characteristic of the exponential model through a global compensation of selection coefficients that maintains the mean phenotype. This compensation impacts the liability distribution by simultaneously reducing the variance ($V_G \to BV_G$) and increasing the threshold density ($f(T) \to f(T)/B$) (Figure 2C, Figure S3). These two effects combine to increase the disease prevalence (Figure 2D) by a factor of approximately

$$\frac{\bar{R}_B}{\bar{R}} \approx \frac{\sqrt{1 - h^2(1 - B)}}{B} \tag{17}$$

where h^2 is the heritability of liability in the absence of BGS (We also use simulations to explore whether the skew in the frequency spectrum impacts the prevalence, but observed no effect; see Supplementary Text C, Figure S12).

However, both the invariance of γ and the increase in prevalence depend entirely on the global compensation mechanism, which is itself predicated on the existence of deleterious fixations to balance (i.e., $p_T > 0$; Eq. (13)). This assumption breaks when the threshold is so low that p_T is driven to effectively zero. In a genome with L contributing sites, this occurs roughly when the expected number of fixed deleterious alleles drops below one ($p_T \lesssim 1/2L$), corresponding to a critical selection coefficient

of $\gamma \approx \frac{1}{2}\ln(2L)$. If γ exceeds this value, there are no deleterious fixations left to maintain. The global compensation mechanism shuts down, γ is no longer forced to be invariant, and the system reverts to standard local dynamics. Consequently, in this strong selection regime, the distinctive signatures of the threshold model disappear: the γ-independent variance reduction fades away (Figure 2C), and the increase in disease prevalence predicted by Equation (17) does not occur (Figure 2E).

.

Variation in effects and variation in B values The threshold model described above assumes a single effect size and constant N_e reduction, allowing for a straightforward population size rescaling. However, if effect sizes (a) and local N_e reductions (B) vary across sites, the dynamics are complicated by the fact that all sites remain coupled by a single threshold density, $f(T)$. The BGS-driven increase in $f(T)$ must therefore represent a global compromise set by the varying impacts of local effective size reductions across sites. This single compensatory change couples back to all sites, impacting their individual selection coefficients. To capture these dynamics, we must first determine how the global compensatory evolution of $f(T)$ is determined, and then resolve how this common shift interacts with local effective size reductions across the distribution of effect sizes.

For simplicity, suppose a and B vary independently across sites with densities $g_a(a)$ and $g_B(B)$. Then, the equilibrium threshold density, $f(T)$, must solve:

$$p_T = \int g_B(B) \int g_a(a) \frac{a}{\bar{a}} p^+ \left(2NC \cdot f(T) \cdot aB\right) \mathrm{d}a\mathrm{d}B. \tag{18}$$

In Supplement F, we find analytical approximations for this integral in tractable limits assuming a gamma distribution on effect sizes $g_a(a)$. Here, we summarize the resulting behaviors by categorizing the distinct regimes controlled by the value of p_T (Figure 3B). We first consider the effectively neutral limit ($p_T \approx \frac{1}{2}$) before examining the transition toward the strong selection limit ($p_T \to 0$). Crucially, as selection strengthens, the solution depends heavily on the shape of the effect size distribution; we therefore contrast distributions with a low coefficient of variation (CV) against those with a high CV (e.g., "L-shaped" distributions) to illustrate how the tail of the distribution shapes the global compensation.

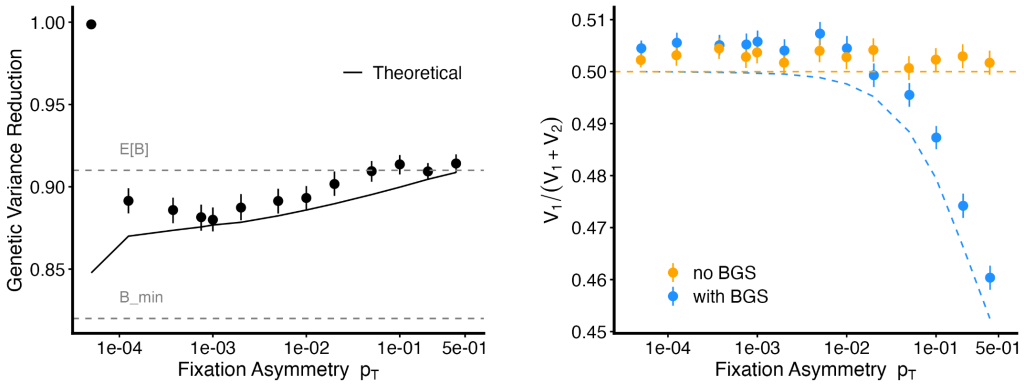
When $p_T \approx \frac{1}{2}$, neutral dynamics and local N_e effects dominate. The global compensation required to maintain equilibrium is set by the arithmetic mean of B:

$$\frac{f_B(T)}{f(T)} \approx \frac{1}{\mathbb{E}[B]}. \tag{19}$$

However, because most sites are effectively neutral, they are insensitive to this global change in $f(T)$. Their dynamics are instead dominated by their local N_e reduction, meaning their individual variance reductions simply depend on their local B value.

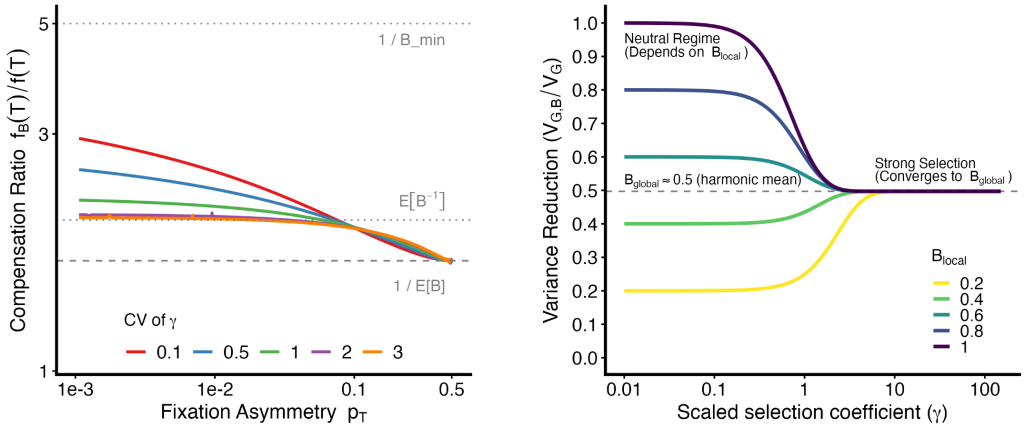
Now consider the limit where we decrease p_T toward zero. The solution to Equation (18) depends

A. Simulations of two chromosomes with different background selection intensities



B. Uniform BGS intensity (B~U(0.2,1))

C. Transition from local to global control of variance reduction



D. Two-effect-size threshold model

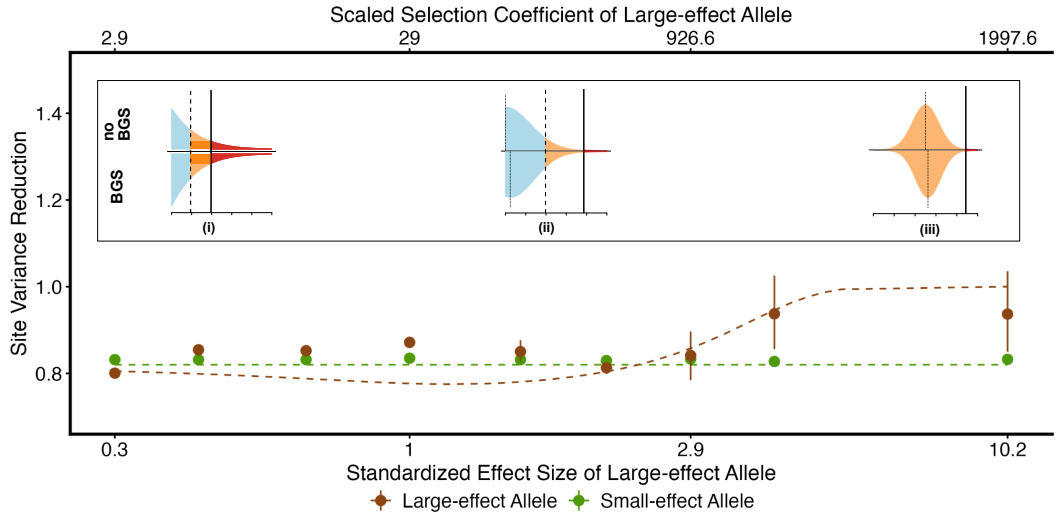


Figure 3: Global coupling of background selection effects in the liability threshold model.
(A) We simulated using a single effect size model with causal sites evenly distributed across two chromosomes, and varied the fixation asymmetry by varying the relative position of the threshold, p_T. In one set of simulations, there is no BGS. In the second set of simulations, one chromosome experiences an effective size reduction due to BGS of $B \approx 0.82$ while the other experiences no BGS. The left panel reports the total reduction in genetic variance (across both chromosomes)in the simulations with BGS versus those without. The solid black line plots the numerical solution of equation 18. The right panel reports the fraction of genetic variance contribute by chromosome one in the presence (blue) and absence of BGS (orange). The theoretical prediction is obtained from equation 22.
(B) The ratio of threshold density, $f_B(T)/f(T)$, is plotted against the fixation asymmetry (p_T) for different coefficients of variation (CV) of the effect size distribution. Curves represent numerical solutions to equation 18, assuming a Gamma distribution for effect sizes and a uniform distribution of local B values, $B \sim U(0.2, 1)$. Horizontal dashed lines indicate theoretical limits derived in the text.
(C) The variance reduction due to BGS as a function of scaled selection coefficient γ and local B value (i.e. Eq. (22)), assuming $B \sim U(0.2, 1)$ and the high CV limit such that $B_{global} = 1/\mathbb{E}[B^{-1}]$. Curves illustrate the transition from dependence on B_{local} to the global average governed by B_{global}.
(D) We simulated a two effect size architecture where one set of sites has a small effect size and belongs to the the weak selection regime, while the other has a large effect size and belongs to the strong selection regime (all sites have the same local effective size reduction). We varied the effect sizes of the large-effect alleles. The x axis on the bottom of the figure measures the size of the large effect relative to the phenotypic standard deviation, while the axis on the top of the figure measures it in terms of the population scaled selection coefficient, γ. The pattern of genetic variance reduction for all but the largest effect sites illustrates the spreading of BGS effects to strongly selected sites via the global compensation mechanism. The theoretical predictions are obtained by solving the two effect model numerically, see Supplementary section D (i) – (iii) illustrate how the risk effect size, δ, for the large-effect-size allele changes with BGS depending on its effect size and drive this effect (recall that $s \approx \delta C$). The region illustrates the proportion of the population who are pushed across the threshold by the allele, and thus its effect on disease risk and fitness. (i) The effect size is relatively small and the risk effect size can be linearly approximated via the dark orange rectangle. BGS increases the area of this rectangle by a factor of $1/B$. (ii) This linearity breaks down, and the effects on risk and therefore fitness increase by more than $1/B$. (iii) When the effect size is large enough to push the entire population across the threshold, BGS no longer has any impact.

heavily on the shape of the effect size distribution, $g_a(a)$. Suppose $g_a(a)$ has a low coefficient of variation (CV), meaning effect sizes are tightly clustered around the mean. As p_T decreases, the bulk of the sites transition into the strong selection regime simultaneously. In this limit, the fixation asymmetry decays exponentially (i.e., $p^+(\gamma) \approx e^{-2\gamma}$). The fixed liability burden at equilibrium is therefore sustained almost entirely by the few sites where genetic drift is strong enough to overpower selection—specifically, those with the smallest scaled selection coefficients. In a model where effect sizes are all roughly equal to one another, these sites are the ones with the strongest local BGS effects (i.e. the smallest B values). Consequently, the global compensation converges to depend on the

minimum B value in the genome:

$$\frac{f_B(T)}{f(T)} = \frac{1}{B_{min}}. \tag{20}$$

Thus, in this limit, the specific local N_e reductions at most sites become irrelevant, as the system's global equilibrium is determined by the region experiencing the strongest background selection.

However, as in the single-effect-size case, this compensation mechanism shuts down when $p_T \lesssim \frac{a_{min}}{2\bar{a}L} \approx \frac{1}{2L}$. This occurs because at this point, even sites with the minimum effect size, a_{min}, can no longer be fixed for the deleterious allele. Thus, for a fixed low-CV distribution, decreasing p_T from $\frac{1}{2}$ to 0 causes the ratio $f_B(T)/f(T)$ to first increase from $\mathbb{E}[B]^{-1}$ toward $\frac{1}{B_{min}}$, and then abruptly drop to 1 as p_T approaches $\frac{1}{2L}$ (Figure 3A).

Alternatively, if the effect size distribution has a high coefficient of variation (CV)—corresponding to the "L-shaped" limit of the gamma distribution—sites will be spread across all selective regimes simultaneously. While the long tail extends into the strong selection regime, the bulk of sites remain in the effectively neutral regime near zero. The resulting global compensation factor, $f_B(T)/f(T)$, must therefore represent a compromise between the limits set by these different regimes: the large mass of neutral sites pulls the compensation toward the neutral limit ($\mathbb{E}[B]^{-1}$), while the exponential sensitivity of the sites in the tail pulls it toward the strong selection limit (B_{min}^{-1}). In Supplement F, we show that in this high-CV limit, this compromise converges to the reciprocal of the harmonic mean B value:

$$\frac{f_B(T)}{f(T)} = \mathbb{E}\left[B^{-1}\right], \tag{21}$$

which satisfies $1/\mathbb{E}[B] \leq \mathbb{E}\left[B^{-1}\right] \leq 1/B_{min}$ for any distribution g_B. In Figure 3B, we illustrate the convergence to this limit by solving equation (18) numerically assuming the local effective size reductions experienced by causal sites are uniformly distributed between 0.2 and 1.

In contrast to the low-variance case, this global compensation persists even as $p_T \to 0$. This difference stems from the high density of small-effect variants in distributions with a mode at zero (e.g., the high-CV Gamma). In the low-CV case, all sites have approximately the same effect ($a \approx \bar{a}$), so they all transition into the strong selection regime simultaneously, causing the compensation mechanism to shut down abruptly. In the high-variance case, there is no such simultaneous shutdown. As the threshold density $f(T)$ increases, sites with large effects are indeed pushed into the strong selection regime, where they stop contributing to the fixed liability burden. However, the mode at zero ensures a continuous reservoir of sites with even smaller effects. As sites with larger effects are all pushed into the strong selection regime, the burden of maintaining the total equilibrium fixation asymmetry dictated by the threshold shifts to these smaller effect variants, which—precisely because of their small size—remain in the weakly selected regime where they are sensitive to changes in local effect size. Thus, baring pathological cases (see Supplementary Text F), this supply of weakly selected

sites is effectively inexhaustible, maintaining the global compensation factor at $f_B(T)/f(T) \approx \mathbb{E}\left[B^{-1}\right]$ as $p_T \to 0$.

While the specific result in Equation (21) arises from the power-law tail of the high-CV gamma distribution, the broader insight is that in biologically realistic scenarios—where both effect sizes (a) and local B values vary—the phenotype is shaped by a combination of two forces: a single global compensation factor and site-specific local BGS effects. The total impact on any given site depends on its selective regime. For a site with scaled selection coefficient γ and local N_e reduction B_{local}, the total reduction in genetic variance is given by:

$$\frac{V_{G,B}}{V_G} = B_{global} \frac{b\left(\gamma \frac{B_{local}}{B_{global}}\right)}{b\left(\gamma\right)} \tag{22}$$

where $b(\gamma) = \tanh(\gamma)$ is the mutational asymmetry and $B_{global} = f(T)/f_B(T)$ is the global compensation factor determined by the genome-wide distributions of effects and B values (e.g., in the high-CV gamma limit, B_{global} is the harmonic mean, $1/\mathbb{E}[B^{-1}]$). This implies that the system behaves in two distinct ways depending on the selection regime: effectively neutral sites are governed by their local effective size reductions ($V_{G,B}/V_G \approx B_{local}$), while strongly selected sites are pinned to the global average ($V_{G,B}/V_G \approx B_{global}$), regardless of the local B value. We illustrate this transition in Figure 3C for the same uniform distribution of local effective size reductions used in Figure 3B.

Notably, this result demonstrates that BGS reduces genetic variance even at sites under very strong selection. In the single-effect model, such sites were immune to BGS because the global compensation mechanism shut down entirely at the $p_T \to 1/2L$ boundary. By contrast, in a model with a high-variance distribution of effect sizes, the global compensation persists, and effectively subjects these strongly selected sites to a variance reduction dictated by the genome-wide background. This globally driven effect, however, has its own limit. The underlying approximation $s \approx af(T)C$ eventually breaks down for sites with very large effects, as their selection coefficients begin to scale non-linearly with their liability contributions (Berg et al. 2025). For sites with effects comparable to the standard deviation of liability ($a \sim \sqrt{V_P}$), this non-linearity can even drive variance reductions that are greater than B_{global}, as the selection coefficients increase faster than linearly with increasing threshold density in this range. As a increases further, however, the coupling breaks down entirely: sites with effects so large that carriers cross the threshold regardless of their genetic background have selection coefficients determined solely by the fitness cost ($s \approx C$), independent of the threshold density $f(T)$. Thus, for sites under the very strongest selection, the impact of BGS does ultimately fade away. Crucially, the point at which this occurs is not determined by population-genetic limits on fixation dynamics, but by the fitness cost of the threshold itself. We illustrate these effects in Figure 3D by comparing numerical solutions (see Supplementary Text D) with simulations of models with two effect sizes, varying the magnitude of the larger effect relative to the background variance in liability.

This result demonstrates that in a realistic model with a high-variance distribution of effect sizes,

BGS reduces genetic variance even at sites under very strong selection. Because the wide distribution of effects ensures that global compensation (B_{global}) is active, even strongly selected sites remain coupled to the genome-wide background and experience a variance reduction dictated by it. This stands in contrast to the single-effect model, where such sites are immune to BGS because the entire compensation mechanism shuts down abruptly once the unique effect size transitions into the strong selection regime (at the $p_T \to 1/2L$ boundary).

This globally driven effect, however, has its own limit. The underlying approximation $s \approx af(T)C$ eventually breaks down for sites with very large effects, as their selection coefficients begin to scale non-linearly with their liability contributions (Berg et al. 2025). For sites with effects comparable to the standard deviation of liability ($a \sim \sqrt{V_P}$), this non-linearity can drive variance reductions even greater than B_{global}. As a increases further, however, the coupling breaks down entirely: sites with effects so large that carriers cross the threshold regardless of their genetic background have selection coefficients determined solely by the fitness cost ($s \approx C$), independent of the threshold density $f(T)$. Thus, for sites under the very strongest selection, the impact of BGS does ultimately fade away. However, the point at which this occurs is not determined by population-genetic limits on fixation dynamics, but by the fitness cost of the threshold itself.

Stabilizing selection

As we have shown, fitness epistasis in the form of threshold selection alters the impact of BGS via a global compensation mechanism, which is necessary to maintain the specific fixation asymmetry imposed by the threshold. We now turn to a model of stabilizing selection, defined by a Gaussian fitness function with an optimum phenotype of $L\bar{a}$ (i.e., $W(Z_i) = e^{\frac{-(Z_i - L\bar{a})^2}{2\omega^2}}$; Lande 1975). This model contrasts with the threshold case in two key ways. First, it ensures mutational symmetry at equilibrium, eliminating the net fixation asymmetry that drove the global coupling in the threshold model. Second, the local dynamics differ: this form of epistasis causes individual alleles to evolve as if underdominant, with a selection coefficient $t \approx \frac{a^2}{2\omega^2}$ against heterozygotes (Fisher 1930; Robertson 1956). This allows us to isolate the impact of BGS on these purely local dynamics.

Defining the scaled coefficient of underdominance as $\tau = 2Nt$, Hayward and Sella 2022 showed that the expected lifetime contribution to heterozygosity is

$$h_u(\tau) \approx 2\int_0^{\frac{1}{2}} x(1-x)\Psi(x|\tau)dx = 2\frac{D_+(\sqrt{\frac{\tau}{2}})}{\sqrt{\frac{\tau}{2}}}, \tag{23}$$

where $D_+(y)$ is the Dawson function, and $\Psi(x|\tau) = 2e^{-2\tau x(1-x)}/x(1-x)$ is the sojourn time at minor allele frequency x. The behavior of this function reveals a crucial deviation from simple directional selection. While $h_u(\tau)$ converges to the directional result $h_d(-\tau)$ at the extremes ($\tau \ll 1$ and $\tau \gg 1$), it is significantly elevated in the intermediate regime ($1 \lesssim \tau \lesssim 30$). In this intermediate regime,

alleles that drift to intermediate frequency "slow down" and persist longer than predicted by negative selection alone, causing the ratio $h_u(\tau)/h_d(-\tau)$ to peak at ≈ 1.29 when $\tau^* \approx 4.42$ (Figures S7 and S8).

Applying the reduced effective size approximation ($N_e \to N_e B$), the modification to heterozygosity is given by

$$B_\tau = B\frac{h_u(\tau B)}{h_u(\tau)} = \sqrt{B}\frac{D_+(\sqrt{\frac{\tau B}{2}})}{D_+(\sqrt{\frac{\tau}{2}})}. \tag{24}$$

This equation demonstrates that BGS does not simply reduce variance by a factor of B or cancel out as it does under strong directional selection. Rather, for some combinations of τ and B, the increased rate of drift leads sites that would otherwise be held at low frequencies to drift up to higher frequencies where the selection against them is weaker, allowing them to continue drifting even further. The increased contribution to heterozygosity conditional on surviving the low frequency cull due to BGS therefore overcompensates for heterozygosity lost that is eliminated by it.

This variance inflation is maximized when BGS shifts alleles from the strong selection regime down to the intermediate regime where the impact of underdominance is most pronounced ($\tau^* \approx 4.42$). For strong BGS ($B \ll 1$), this peak net increase occurs at $\tau \approx {}^{\tau^*}/_{B}$ and converges to the theoretical maximum of ≈ 1.29, while the range of affected sites scales with $1/B$. Figure 4A shows that multi-locus simulations of stabilizing selectino faithfully replicate these single-site predictions.

In Figures 4C and 4D, we apply the frequency dependent factor $B\,(q)$ from equation (2) and the rescaling $\tau \to \tau B$ to the frequency spectrum for underdominant sites for $\tau = -1.5$ and 10 under modest background selection (i.e., $B = 0.83$, the human average). In this regime, the rescaling of τ is not dramatic enough to move sites from the strong into the weak selection regime, so the effects on the frequency spectrum are modest. In contrast, under larger effective population size reductions, as BGS moves sites fully between regimes, the inflation of the frequency spectrum at intermediate frequencies becomes substantial, although here our simple approximation for the skew in the low frequency regime breaks down (Figure S9).

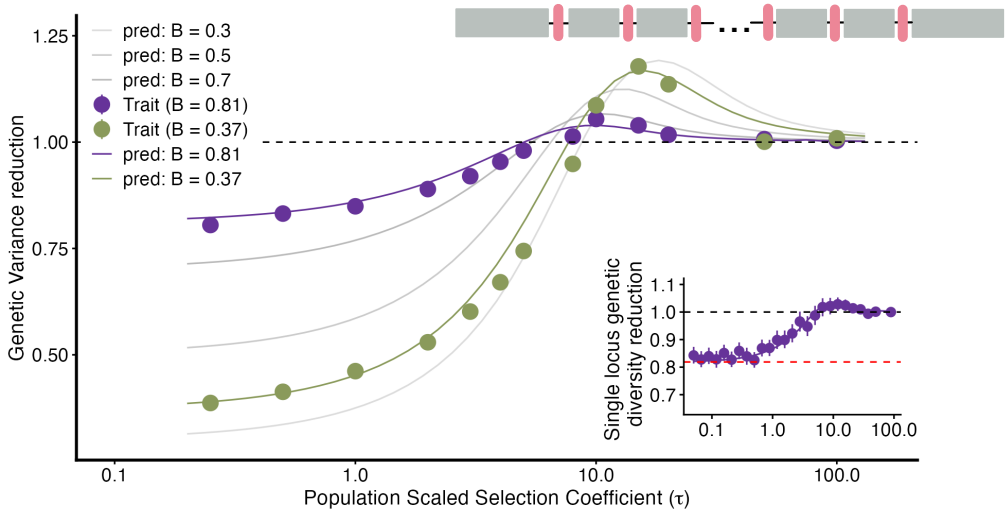
Discussion

In this paper, we have used a combination of analytical theory and simulations to study the impacts of background selection in recombining genomes, both at the level of individual variants and the level of phenotypes, in several long-studied models.

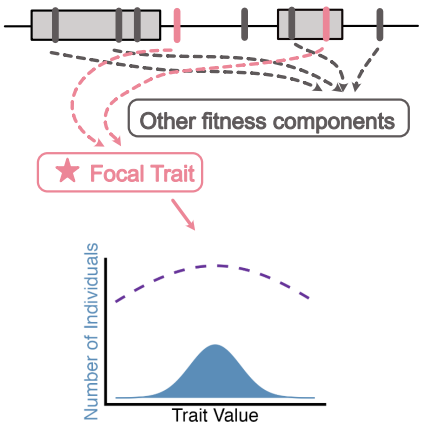
Background selection and the frequency spectrum in recombining genomes

While the impact of background selection on aggregate diversity is often well-approximated by a simple reduction in effective population size, characterizing the full shape of the site frequency spectrum

A. Genetic variance reduction of traits under stabilizing selection



B. Focal traits under stabilizing selection



C. SFS at a single locus with different selection coefficients

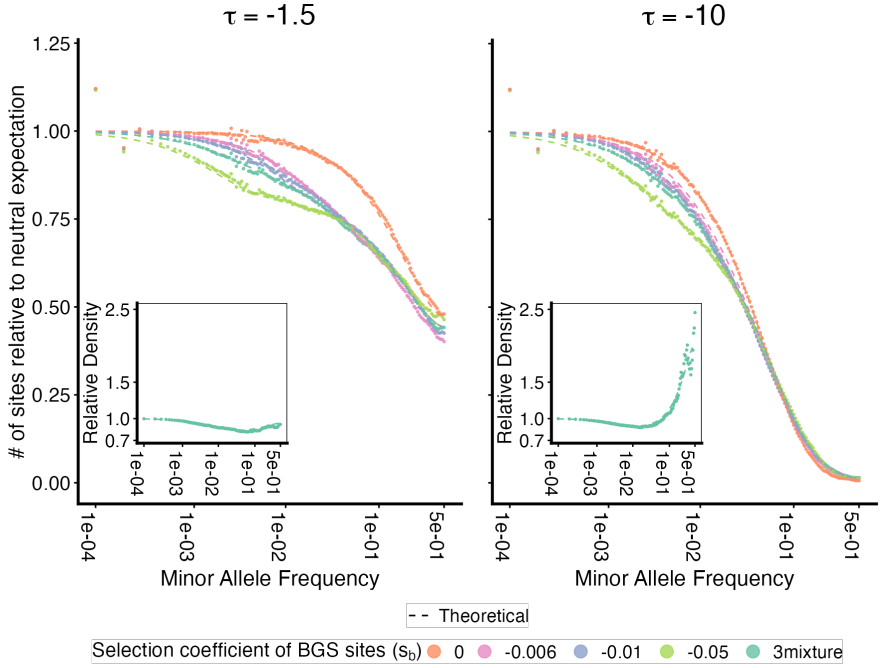


Figure 4: Stabilizing Selection. (A) We simulate traits with a single effect size under stabilizing selection with two different effective population size reductions. For each effective size reduction, we plot the change in genetic variance against the population-scaled selection coefficient ($\tau = 2N\frac{a^2}{2\omega^2}$). The chromosome structure used for simulation is depicted in the top right, showing trait loci interspersed among BGS regions. Simulations results (points) are compared with the theoretical prediction, obtained from equation 23, with the B values estimated from the diversity reduction of neutral sites simulated in the same genome. The inset shows the genetic diversity of underdominant alleles in a single locus simulations with $B = 0.82$, the human average.
(B) Schematic describing the relationship between sites contributing to focal traits (pink bars) and linked deleterious mutations (gray bars).
(C) The site frequency spectrum observed in simulation (points) are compared to the theoretical predictions (dashed lines) for focal selected sites under weak underdominant selection ($\tau = -1.5$, left) and stronger underdominant selection ($\tau = -10$, right). Different colors represent different selection coefficients or a distribution with an equal mixture of all three at the background mutations (s_b). Insets show the relative density of alleles at each frequency bin with background selection vs. no background selection.

in recombining genomes presents a more difficult challenge. Here, we describe an effectively non-recombining block approximation (Neher et al. 2013; Good et al. 2014; Weissman and Hallatschek 2014) for the classic model of background selection with recombination by defining the characteristic length scale of the process, and demonstrate that this approximation recovers classic results previously derived by summing contributions across many loci in pairwise models (Hudson and Kaplan 1995; Nordborg et al. 1996). This re-framing allows us to make progress in understanding the impact of BGS on the skew in the frequency spectrum in recombining genomes and to derive a simple expression for the effect that is accurate in the weak mutation regime which characterizes humans ($\lambda \lesssim 1/2$).

In Supplementary Text A, we use heuristic arguments to extend the effectively non-recombining block approximation into a "shrinking block" model. In this framework, the length of the original haplotype remaining associated with a focal allele of age t is $M_t \approx 2/rt$. This approximation allows us to map the dynamics onto the non-recombining model of Cvijović et al. (2018) and extend their piecewise expression for the frequency spectrum skew under $\lambda \gg 1$ to recombining genomes (Figure S5).

Notably, this shrinking block perspective reveals a key difference in dynamics between the two models: because M_t is large for small t, an allele in a recombining genome becomes associated with newly arising deleterious mutations far beyond the length $M_{s_b} = 2s_b/r$ that characterizes classical background selection (BGS) against pre-existing variation. Over the short timescales that characterize this process, the balance between mutation and recombination ensures that the number of new deleterious mutations associated with the focal allele remains constant on expectation and is also equal to λ (Eq. (A.9)). Consequently, the time required for an allele to feel the effect of selection against new mutations on its background is reduced by a factor of $1/\sqrt{\lambda}$ relative to a non-recombining model with the same local reduction in effective size (Eq. (A.17)). This implies that the frequency at

which substantial skew begins to accumulate is lowered by the same factor (Eq. (A.18)), and that the blocks eliminated by this process are longer than the characteristic length M_{s_b} by a factor of λ (Eq. (A.12)).

Developing a complete analytical description of the frequency spectrum for $\lambda \gtrsim 1/2$ (or a numerical method to compute it accurately) remains a priority for future work. For $1/2 \lesssim \lambda \lesssim 1$, progress might be made by assuming the focal allele accumulates at most one additional deleterious mutation within the characteristic length M_{s_b}, though this approximation would clearly fail for larger λ. For the $\lambda \gg 1$ case, the shrinking block perspective is promising: the impact on an allele at frequency q could be described by a characteristic length $M_q = (Nrq)^{-1}$, treating the region as a non-recombining chromosome of that length. While this approach requires a more robust solution for the frequency spectrum in non-recombining genomes, we hope these insights can help provide a bridge toward that solution.

Models of phenotypic selection and linkage in recombining genomes

In our phenotypic analyses, we assume $2NBs_b \gg 1$, ensuring that the skew in the frequency spectrum is fully "priced in" at relatively low frequencies. This permits us to assume that BGS preserves the standard functional forms of heterozygosity and fixation rates under Wright-Fisher diffusion, allowing us to apply classic effective population size rescaling in our phenotypic models while retaining the standard . However, the relationship between an allele's contribution to heterozygosity and its fixation rate expressed in equation (6) is more general than the diffusion limit, as it simply represents an averaging of Fisher's fundamental theorem over an allele's transit through the population (Fisher 1930; Santiago and Caballero 2016). Consequently, the behaviors we describe for directional models are expected to be extremely general, and should equally apply in models that relax the strong selection assumption (Good et al. 2014; Santiago and Caballero 2016; Barroso and Ragsdale 2025). Under the threshold model in particular, introducing essentially any form of linked selection should force the population distribution to evolve such that the selective fixation rates in both directions are reduced by the same factor. Even if the form of linked selection alters the functional form of the fixation rate, a reduction in the selective fixation rate by a given multiplicative factor must imply a corresponding reduction in heterozygosity by the same factor, independent of the underlying details (although changes in the form of the fixation rate could alter the precise form of B value averaging that drives the global effect).

That said, different models of linked selection nonetheless result in different relationships between fundamental parameters and patterns of genetic diversity, and there remains considerable uncertainty about the underlying source of linked selection signals (Murphy et al. 2022; Buffalo and Kern 2024). While we have primarily focused on understanding the impact of BGS on the equilibrium evolution of complex traits, it is also interesting to ask the converse question: what forms of phenotypic variation underlie linked selection signals that have been measured empirically, and what model of selection

applies to the underlying loci? This is particularly interesting to consider in light of arguments for the ubiquity of stabilizing selection (Simons et al. 2018; Koch et al. 2024; Simons et al. 2025).

Intuitively, alleles with effects large enough to fall into the strong selection regime (i.e., $2Nt \approx 2N\frac{a^2}{2\omega^2} \gtrsim 30$) should impact linked sites precisely as predicted by classic BGS theory, because their local dynamics mirror the standard strong selection model. A more complex question is what effect we should expect from loci falling into the intermediate weak selection regime, where underdominant dynamics become important. These loci are more likely to drift to intermediate frequencies and persist than their counterparts under weak *directional* selection (which can be modeled using the quantitative genetic framework of Santiago and Caballero 2016). However, they also convert additive variance into various forms of interaction variance as they change frequency, and it is unclear to what extent this non-additive variance might remain correlated with the focal allele across generations, or how the conversion of variance between different statistical partitions impacts linked selection dynamics. This remains an interesting avenue for future work.

Of course, the linked selection effects we study are not the only mechanism through which linkage influences complex trait evolution, particularly under stabilizing selection. By selecting against extreme phenotypes, stabilizing selection generates negative LD between alleles with similar effects, causing a short-term reduction in genetic variance known as the Bulmer effect (Bulmer 1971; Bulmer 1974). From the perspective of a single focal site, the Bulmer effect manifests as a negative covariance between the trait-increasing allele and the genetic background. This attenuates the allele's marginal phenotypic association and, by extension, its marginal fitness effect (i.e., it decreases the site's "effective" selection coefficient; Negm and Veller 2024). Although not the primary focus of our study, we found while calibrating our simulations that this reduction in selection strength leads to a long-term increase in *genic* variance that exceeds the short-term reduction caused by negative LD. Thus, if stabilizing selection is sustained consistently over the timescales on which individual alleles make their transits through the population, then the net impact of linkage may actually be to increase the genetic variance of quantitative traits, rather than to decrease it (see Supplementary Text E), even before accounting for the potential variance increasing effects of BGS that we describe.

Implications for natural populations

The extent to which background selection shapes genetic architecture in nature depends critically on the underlying model of phenotypic selection and/or the marginal fitness effects of individual loci. Naturally, for effectively neutral loci, the specific selection model is irrelevant, and the impact of BGS follows standard neutral expectations. Thus, the observation by Rockman et al. (2010) in *C. elegans* that loci in gene-poor, high-recombination regions explain significantly more gene expression variance implies that if BGS is responsible, the genetic variance must be effectively neutral. The most interesting potential effects are thus for strongly selected loci, where different models diverge in their predictions.

Much of the literature on complex trait evolution assumes that stabilizing selection is the norm (Kingsolver et al. 2001; Sella and Barton 2019). While this assumption has historically been difficult to assess, several recent analyses in human genetics support the hypothesis, a finding that appears to hold for diseases as well as quantitative traits (Sanjak et al. 2018; Simons et al. 2018, 2022; Koch et al. 2024; Berg et al. 2025; Ragsdale 2025). Suppose we accept the maximalist interpretation of this inference and assume that all functional loci contribute to fitness variance via their effects on traits under stabilizing selection. What, then, would be the effect of BGS on functional variation? In humans, the consequences would be straightforward and relatively minor, given the relatively modest local reductions in effective population size. However, in species experiencing stronger reductions, the impact is more significant, as we would expect the contribution to variance to increase across a wide range of sites that would otherwise by strongly selected. While the increase in heterozygosity due to BGS is capped at ≈ 1.29—and thus can never provide more than a small boost to total genetic variance—the impact on genetic architecture could be substantial, as strongly selected loci, where alleles would typically be confined to low frequencies, would instead have a significant portion of their variance contribution driven by variants drifting to intermediate frequencies. Notably, this dynamic would be expected to operate only to the extent that $\tau < 2Ns_b$, so that the effects of background selection have time to take effect.

While it seems clear that stabilizing selection plays an important role in shaping much functional variation, it is reasonable to question whether it is the whole picture. It could be that axes of phenotypic integration which experienced a sustained directional mutation-selection balance are simply more difficult to measure, and have thus evaded detection. If a substantial fraction of functional variation is under directional selection, and if most directional selection were derived from traits subject to exponential fitness functions (or some other model in which individual loci have independent fitness effects), then there is no effect on the contribution to variance from strongly selected sites, although the frequency spectra may still be shifted toward large effect loci is $|\gamma| < 2Ns_b$, though in a less extreme way than under stabilizing selection because of the absence of the slow down dynamic at intermediate frequencies.

However, models in which selection acts independently predict unreasonably large genetic load. This occurs for the same reason that the exponential model predicts large changes in mean fitness due to BGS: if loci fix independently, the accumulation of weakly selected fixations drives mean fitness down by several orders of magnitude—a result inconsistent with the persistence of natural populations. Stabilizing selection offers one resolution to this "load paradox" (Charlesworth 2013; Barton 2017). Strong synergistic epistasis of the kind present in the liability threshold model, offers another, for precisely the reason that the threshold model differs from the exponential model in our analysis: the sharp increase in the rate of mean fitness decline with increasing phenotype. (Although we frame our analysis in terms of disease, any model exhibiting this pattern will behave similarly; Kondrashov 1995; Berg et al. 2025).

Thus, to the extent that individual variants derive their fitness effects from directional mutation-selection balance, there has been a strong theoretical expectation that synergistic epistasis must be involved. Despite this, empirical support for its importance remains mixed, for a variety of technical reasons (Mukai 1964; Mukai and Cockerham 1977; Fry 2004; Halligan and Keightley 2009; Sohail et al. 2017; Garcia and Lohmueller 2021; Sandler et al. 2021). However, efforts to detect synergistic epistasis are often underpowered because they typically test for interactions between random sets of alleles, when we might expect *a priori* that it exists only among sets of variants which integrate into the same phenotype, of which there are presumably many (Rice (1998); but see Lee (2022) for a counterexample). Consequently, it remains difficult to determine how much "evidence of absence" should be inferred from the current scarcity of empirical evidence.

Thus, while the relative importance of synergistic epistasis in evolution remains unclear, our results suggest that to the extent that it is important, then it is likely that BGS plays a substantial role in reducing contributions to genetic variance from strongly selected loci which are otherwise largely immune to its local effects. These effects would be most pronounced in species where the average genome-wide reduction in effective size is large (i.e. species with a high density of deleterious mutations per unit of recombination), or where it varies substantially across different parts of the genome, given that the magnitude of the global compensation is driven more by sites with the smallest B values. Notably, because this phenomenon manifests through a genome-wide change in the marginal selection coefficients themselves, it would be difficult to detect empirically, other than by demonstrating that both synergistic fitness epistasis and BGS are operating.

Methods

Simulation of single-site selection models

To validate the predictions on diversity reduction by BGS in our single-site models, we performed SLiM simulations (Haller and Messer 2019) on a single site under directional or underdominant selection across a range of selection coefficients. In each case, we simulated $5,000$ recombining diploid genomes, with each genome having $M = 4 \times 10^5$ background sites, and a single focal site in the middle of the chromosome. The per-site recombination rate is set to 2.5×10^{-6}, such that the map length is equal to 1 Morgan. The mutation rate at each of the M background sites is 2.5×10^{-7}. When we simulate without background selection, all mutations at the M background sites other than the focal one are neutral. When we simulate with background selection, all mutations at these M background sites have a selection coefficient $s = -0.005$. With $\nu = 2.5 \times 10^{-7}$ and $r = 2.5 \times 10^{-6}$, this produces a B of 0.82. To vary the background selection strength, we tune the mutation rate (ν) with a fixed recombination rate (r) to produce a specified B ($B = e^{-\lambda} = e^{-2\frac{\nu}{r}}$).

The mutation rate at the focal site is set to zero, and we introduce mutations at this site "manually".

We first let the simulation of the background sites run for $2N$ generations so that the population reaches mutation-selection balance. After this burn-in period, we introduced a single selected allele at the focal site in the center of the genome.

To measure the reduction of heterozygosity by BGS, we adopt the method used by Charlesworth et al. (1993) to measure the expected total contribution to heterozygosity of a new mutation during its transit through the population (H_{tot}). This is given by the sum of heterozygosity at the focal site during the T generations that the mutation segregated in the population,

$$H_{tot} = 2\sum_{i=1}^{T} x_i(1-x_i), \tag{25}$$

which has a theoretical expectation of $\mathbb{E}\left[H_{tot}|\gamma\right] = h_d(\gamma)$, or $\mathbb{E}\left[H_{tot}|\tau\right] = h_u(\tau)$, depending on the model, and is thus proportional to the expected heterozygosity for a single site at a random sampling time in the low mutation limit. For each selection coefficient, we average the value of H_{tot} across 10^7 separate focal sites. These mutations were simulated across 10^4 replicate burn-ins, with 10^3 mutations considered per burn-in. We then take the ratio of H_{tot} under BGS relative to no BGS as our estimator of the impact of background selection on the expected heterozygosity. To obtain a standard error, we treat the value of H_{tot} for a single mutation's transit through the population as a single data point, and bootstrap across these values.

To construct the site frequency spectrum (SFS), we record the allele frequency of each new mutation at every generation during its sojourn in the population. To validate the frequency-dependent B approximation, we generate multiple SFS datasets in which ν and r are held constant, ensuring the same expected reduction in neutral diversity across datasets. We then vary the selection coefficients of background mutations, $s_b \in \{-0.006, -0.01, -0.05\}$, as well as an equal mixture of these three coefficients.

Simulation of complex Traits

Exponential selection model

To plot the reduction in variance due to background selection under the exponential model, we take advantage of the theoretical equivalence between the exponential model and the single-site additive model. To this end, we simulate a single site under directional selection with a fixed selection coefficient, as described above, except that each time a mutation ends its transit through the population and we revert the simulation back to the end of the burn-in, we choose the sign of the effect of the next mutation to be opposite to that of the allele that fixed at the end of the previous mutation's transit; i.e. if it was the deleterious allele that fixed, then the next mutation would be beneficial, and vice versa. This scheme simulates the long-term fixation dynamics under the exponential model. Calculations regarding the reduction in heterozygosity due to BGS are then otherwise performed as

described above in the single-site case.

Threshold selection model

In this case, we simulate under the full threshold model, again with $5,000$ diploid individuals. For each choice of $p_T = \frac{T}{2La}$ in the single-effect-size model, we simulated $L = 10^5$ causal sites which contribute to phenotypic variation, set $a = 1$, and choose T so as to attain the correct value of p_T. We set the mutation rate for causal sites to $\mu = 10^{-7}$. All causal sites are equally spaced along the chromosome. Between each pair of causal sites, there are 25 BGS sites, as well as a single neutral site (these neutral sites are used to empirically check the strength of the BGS effect, i.e. the B value, against the expectation). For most of our simulations, the recombination rate among adjacent sites is set to $r = 2 \times 10^{-7}$, and the mutation rate at BGS sites is set to $\nu = 2 \times 10^{-8}$, to yield an expected B value of $e^{-0.2} \approx 0.82$, approximately equal to the estimated human average. To ensure that all causal sites experience a similar background selection effect, we added 6.5×10^5 negatively selected sites on each end of the chromosomes, such that the chromosome has a total map length of 0.8 Morgan. In the simulations where we vary the strength of the BGS effect, we do so by varying the mutation rate for BGS sites, ν, with the recombination rate held constant.

To compare traits with and without background selection, we first simulate a trait with the environmental variance tuned so that the heritability is $h^2 = 0.5$ in the absence of background selection (i.e. setting $s = 0$ for all BGS sites). We then rerun the same simulation with background selection turned on (i.e. with $s = -0.005$ for all BGS sites), holding the environmental variance constant at the same value used in the no background selection case. In all cases, we set the fitness cost of being across the threshold at $C = 0.2$. An individual's fitness is then computed as a product of their fitness effect due to the trait together with independent multiplicative contributions from the BGS sites.

To compare the single-effect-size trait model to the single-site results, we assign a scaled selection coefficient to each set of trait simulations. This quantity measures the average selection coefficient that causal variants experience in the trait model. For each choice of $p^+(a)$ in the simulations, we estimated the unscaled selection coefficient by regressing individual fitness against genotypes in the absence of background selection. This regression was performed for each causal site segregating in a given generation and repeated across multiple generations, with the average slope serving as our estimate. We then obtained the scaled selection coefficient by multiplying this estimate by $2N$. The corresponding simulation run, conducted with the same $p^+(a)$ but with background selection included, was assigned the same scaled selection coefficient. In principle, in the small effect regime we could obtain the scaled selection coefficient analytically using Equation (13), but this method breaks down for large effects, so we use the empirical regression-based estimate across the full range.

Stabilizing Selection

Our simulations under stabilizing selection are similar to those under the threshold model, with a few key differences. Notably, to avoid the Bulmer effect in our simulations of stabilizing selection (see Supplementary section E and Figure S11), we had to increase the harmonic mean recombination rate between pairs of causal sites (Bulmer 1971; Bulmer 1974; Negm and Veller 2024; Veller and Coop 2024). In these simulations, we set $L = 5.4 \times 10^4$, and split the causal sites across 6 independent chromosomes, so that each chromosome has 9000 causal sites. Between each pair of adjacent causal sites on a given chromosome, there are 98 BGS sites and 1 neutral site. The recombination rate is set to $r = 10^{-6}$ per base pair, while the mutation rate is $\mu = \nu = 10^{-7}$ for both causal and BGS sites. We set the width of the stabilizing selection surface $\omega^2 = 2N$ and vary the effect size for each site a.

To compare the single-effect-size trait model to the single-site results, for each choice of a, we measure the average heterozygosity of all the causal sites in the simulations without background selection over multiple generations and solved for the τ with Equation (23). We then multiply this estimate by $2N$ and assign it to the simulations with the same a with and without background selection.

Use of generative AI tools

We used generative AI tools, including Gemini 2.5 Pro and Gemini 3.0 Pro, as well as ChatGPT-5, to help draft and edit text in this manuscript, and to produce and edit code for plotting figures. All text and code drafted or edited in this manner was thoroughly checked and edited further before inclusion in the final manuscript. We take full responsibility for the contents of this manuscript.

Resource

The documented code, including simulations, numerical solutions of the model, and figure-geneartion scripts, is publicly available at `https://github.com/xinyli/BGS_msdb.git`.

Acknowledgements

We would like to thank John Novembre for many helpful conversations, and for reading and commenting extensively on several drafts of the manuscript. We would also like to thank members of the Berg, Novembre and Steinrücken labs for feedback on the work at various stages. This work was completed in part with resources provided by the University of Chicago's Research Computing Center. XL was supported by NIH grant R35 GM149521 to John Novembre and R01 HG010773 to John Novembre and Xin He. JJB was supported by R35 GM151257.

References

Fisher, RA (1930). *The Genetical Theory of Natural Selection*. Oxford University Press.

Wright, S (July 1938). "The Distribution of Gene Frequencies Under Irreversible Mutation". In: *Proceedings of the National Academy of Sciences of the United States of America* 24.7, pp. 253–259. ISSN: 0027-8424. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1077089/ (visited on 09/17/2020).

Karn, MN and L Penrose (1951). "Birth weight and gestation time in relation to maternal age, parity and infant survival". In: *Annals of eugenics* 16.1, pp. 147–164.

Robertson, A (1956). "The effect of selection against extreme deviants based on deviation or on homozygosis: With Two Text-figures". In: *Journal of Genetics* 54, pp. 236–248.

Mukai, T (July 1964). "The Genetic Structure of Natural Populations of Drosophila Melanogaster. I. Spontaneous Mutation Rate of Polygenes Controlling Viability". In: *Genetics* 50.1, pp. 1–19. ISSN: 0016-6731. DOI: 10.1093/genetics/50.1.1. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC1210633/ (visited on 12/18/2025).

Kimura, M (1969). "The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations". In: *Genetics* 61.4, p. 893.

Bulmer, M (1971). "The effect of selection on genetic variability". In: *The American Naturalist* 105.943, pp. 201–211.

Bulmer, M (1974). "Linkage disequilibrium and genetic variability". In: *Genetics Research* 23.3, pp. 281–289.

Lande, R (Dec. 1975). "The maintenance of genetic variability by mutation in a polygenic character with linked loci". en. In: *Genetics Research* 26.3, pp. 221–235. ISSN: 1469-5073, 0016-6723. DOI: 10.1017/S0016672300016037. URL: https://www.cambridge.org/core/journals/genetics-research/article/maintenance-of-genetic-variability-by-mutation-in-a-polygenic-character-with-linked-loci/D12E61BA54B9B4BE2480C13E15757432 (visited on 10/15/2024).

Mukai, T and CC Cockerham (June 1977). "Spontaneous mutation rates at enzyme loci in Drosophila melanogaster." In: *Proceedings of the National Academy of Sciences* 74.6. Publisher: Proceedings of the National Academy of Sciences, pp. 2514–2517. DOI: 10.1073/pnas.74.6.2514. URL: https://www.pnas.org/doi/abs/10.1073/pnas.74.6.2514 (visited on 12/18/2025).

Milkman, R (1978). "Selection differentials and selection coefficients". In: *Genetics* 88.2, pp. 391–403.

Smith, JM (1983). "The genetics of stasis and punctuation". In: *Annual review of genetics* 17.1, pp. 11–25.

Li, WH (Apr. 1987). "Models of nearly neutral mutations with particular implications for nonrandom usage of synonymous codons". en. In: *Journal of Molecular Evolution* 24.4, pp. 337–345. ISSN: 1432-1432. DOI: 10.1007/BF02134132. URL: https://doi.org/10.1007/BF02134132 (visited on 10/06/2025).

Charlesworth, B, M Morgan, and D Charlesworth (1993). "The effect of deleterious mutations on neutral molecular variation." In: *Genetics* 134.4, pp. 1289–1303.

Charlesworth, B (1994). "The effect of background selection against deleterious mutations on weakly selected, linked variants". In: *Genetics Research* 63.3, pp. 213–227.

Hudson, RR and NL Kaplan (1995). "Deleterious background selection with recombination." In: *Genetics* 141.4, pp. 1605–1617.

Kondrashov, AS (Aug. 1995). "Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over?" en. In: *Journal of Theoretical Biology* 175.4, pp. 583–594. ISSN: 0022-5193. DOI: 10.1006/jtbi.1995.0167. URL: https://www.sciencedirect.com/science/article/pii/S0022519385701671 (visited on 07/15/2022).

Nordborg, M, B Charlesworth, and D Charlesworth (1996). "The effect of recombination on background selection". In: *Genetics Research* 67.2, pp. 159–174.

Rice, WR (Mar. 1998). "Requisite mutational load, pathway epistasis, and deterministic mutation accumulation in sexual versus asexual populations". en. In: *Genetica* 102.0, pp. 71–81. ISSN: 1573-6857. DOI: 10.1023/A:1017055429755. URL: https://doi.org/10.1023/A:1017055429755 (visited on 12/18/2025).

McVean, GAT and B Charlesworth (Oct. 1999). "A population genetic model for the evolution of synonymous codon usage: patterns and predictions". en. In: *Genetics Research* 74.2, pp. 145–158. ISSN: 1469-5073, 0016-6723. DOI: 10.1017/S0016672399003912. URL: https://www.cambridge.org/core/journals/genetics-research/article/population-genetic-model-for-the-evolution-of-synonymous-codon-usage-patterns-and-predictions/785E506112E3C9505F6EFFF7D262 (visited on 08/28/2024).

Stephan, W, B Charlesworth, and G McVean (1999). "The effect of background selection at a single locus on weakly selected, partially linked variants". In: *Genetics Research* 73.2, pp. 133–146.

Bustamante, CD, J Wakeley, S Sawyer, and DL Hartl (Dec. 2001). "Directional Selection and the Site-Frequency Spectrum". en. In: *Genetics* 159.4. Publisher: Genetics Section: INVESTIGATIONS, pp. 1779–1788. ISSN: 0016-6731, 1943-2631. URL: https://www.genetics.org/content/159/4/1779 (visited on 09/23/2020).

Kingsolver, JG, HE Hoekstra, JM Hoekstra, D Berrigan, SN Vignieri, C Hill, A Hoang, P Gibert, and P Beerli (2001). "The strength of phenotypic selection in natural populations". In: *The American Naturalist* 157.3, pp. 245–261.

Gordo, I, A Navarro, and B Charlesworth (June 2002). "Muller's Ratchet and the Pattern of Variation at a Neutral Locus". In: *Genetics* 161.2, pp. 835–848. ISSN: 1943-2631. DOI: 10.1093/genetics/161.2.835. URL: https://doi.org/10.1093/genetics/161.2.835 (visited on 12/11/2025).

Fry, JD (Feb. 2004). "On the Rate and Linearity of Viability Declines in Drosophila Mutation-Accumulation Experiments: Genomic Mutation Rates and Synergistic Epistasis Revisited". In:

Genetics 166.2, pp. 797–806. ISSN: 1943-2631. DOI: 10.1093/genetics/166.2.797. URL: https://doi.org/10.1093/genetics/166.2.797 (visited on 12/18/2025).

Sella, G and AE Hirsh (July 2005). "The application of statistical physics to evolutionary biology". In: *Proceedings of the National Academy of Sciences* 102.27. Publisher: Proceedings of the National Academy of Sciences, pp. 9541–9546. DOI: 10.1073/pnas.0501865102. URL: https://www.pnas.org/doi/abs/10.1073/pnas.0501865102 (visited on 10/08/2024).

Halligan, DL and PD Keightley (Dec. 2009). "Spontaneous Mutation Accumulation Studies in Evolutionary Genetics". en. In: *Annual Review of Ecology, Evolution, and Systematics* 40.Volume 40, 2009. Publisher: Annual Reviews, pp. 151–172. ISSN: 1543-592X, 1545-2069. DOI: 10.1146/annurev.ecolsys.39.110707.173437. URL: https://www.annualreviews.org/content/journals/10.1146/annurev.ecolsys.39.110707.173437 (visited on 12/18/2025).

McVicker, G, D Gordon, C Davis, and P Green (May 2009). "Widespread Genomic Signatures of Natural Selection in Hominid Evolution". en. In: *PLOS Genetics* 5.5. Publisher: Public Library of Science, e1000471. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1000471. URL: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1000471 (visited on 09/23/2020).

Rockman, MV, SS Skrovanek, and L Kruglyak (2010). "Selection at linked sites shapes heritable phenotypic variation in C. elegans". In: *Science* 330.6002, pp. 372–376.

Charlesworth, B (2013). "Why We Are Not Dead One Hundred Times Over". en. In: *Evolution* 67.11. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/evo.12195, pp. 3354–3361. ISSN: 1558-5646. DOI: 10.1111/evo.12195. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/evo.12195 (visited on 07/15/2022).

Neher, RA, TA Kessinger, and BI Shraiman (Sept. 2013). "Coalescence and genetic diversity in sexual populations under selection". en. In: *Proceedings of the National Academy of Sciences* 110.39. Publisher: National Academy of Sciences Section: Biological Sciences, pp. 15836–15841. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1309697110. URL: https://www.pnas.org/content/110/39/15836 (visited on 09/23/2020).

Comeron, JM (June 2014). "Background Selection as Baseline for Nucleotide Variation across the Drosophila Genome". en. In: *PLOS Genetics* 10.6. Publisher: Public Library of Science, e1004434. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1004434. URL: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004434 (visited on 12/11/2025).

Good, B, A Walczak, R Neher, and M Desai (Mar. 2014). "Genetic Diversity in the Interference Selection Limit". en. In: *PLOS Genetics* 10.3. Publisher: Public Library of Science, e1004222. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1004222. URL: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004222 (visited on 12/11/2025).

Weissman, DB and O Hallatschek (Apr. 2014). "The Rate of Adaptation in Large Sexual Populations with Linear Chromosomes". In: *Genetics* 196.4, pp. 1167–1183. ISSN: 1943-2631. DOI: 10.1534/

genetics.113.160705. URL: https://doi.org/10.1534/genetics.113.160705 (visited on 12/16/2025).

Corbett-Detig, RB, DL Hartl, and TB Sackton (2015). "Natural selection constrains neutral diversity across a wide range of species". In: *PLoS biology* 13.4, e1002112.

Rice, DP, BH Good, and MM Desai (May 2015). "The Evolutionarily Stable Distribution of Fitness Effects". In: *Genetics* 200.1, pp. 321–329. ISSN: 1943-2631. DOI: 10.1534/genetics.114.173815. URL: https://doi.org/10.1534/genetics.114.173815 (visited on 07/06/2022).

Coop, G (2016). "Does linked selection explain the narrow range of genetic diversity across species?" Preprint at bioRxiv https://www.biorxiv.org/content/10.1101/042598v1.

Elyashiv, E, S Sattath, TT Hu, A Strutsovsky, G McVicker, P Andolfatto, G Coop, and G Sella (2016). "A genomic map of the effects of linked selection in Drosophila". In: *PLoS genetics* 12.8, e1006130.

Santiago, E and A Caballero (Nov. 2016). "Joint Prediction of the Effective Population Size and the Rate of Fixation of Deleterious Mutations". In: *Genetics* 204.3, pp. 1267–1279. ISSN: 1943-2631. DOI: 10.1534/genetics.116.188250. URL: https://doi.org/10.1534/genetics.116.188250 (visited on 12/18/2025).

Barton, NH (Jan. 2017). "How does epistasis influence the response to selection?" en. In: *Heredity* 118.1. Publisher: Nature Publishing Group, pp. 96–109. ISSN: 1365-2540. DOI: 10.1038/hdy.2016.109. URL: https://www.nature.com/articles/hdy2016109 (visited on 10/18/2024).

Sohail, M, OA Vakhrusheva, JH Sul, SL Pulit, LC Francioli, Genome of the Netherlands Consortium, Alzheimer's Disease Neuroimaging Initiative, LH van den Berg, JH Veldink, PIW de Bakker, GA Bazykin, AS Kondrashov, and SR Sunyaev (May 2017). "Negative selection in humans and fruit flies involves synergistic epistasis". In: *Science* 356.6337. Publisher: American Association for the Advancement of Science, pp. 539–542. DOI: 10.1126/science.aah5238. URL: https://www.science.org/doi/abs/10.1126/science.aah5238 (visited on 12/18/2025).

Cvijović, I, BH Good, and MM Desai (2018). "The effect of strong purifying selection on genetic diversity". In: *Genetics* 209.4, pp. 1235–1278.

Gazal, S, PR Loh, HK Finucane, A Ganna, A Schoech, S Sunyaev, and AL Price (Nov. 2018). "Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations". en. In: *Nature Genetics* 50.11. Bandiera_abtest: a Cg_type: Nature Research Journals Number: 11 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Functional genomics;Genome-wide association studies;Population genetics;Statistics Subject_term_id: functional-genomics;genome-wide-association-studies;population-genetics;statistics, pp. 1600–1607. ISSN: 1546-1718. DOI: 10.1038/s41588-018-0231-8. URL: https://www.nature.com/articles/s41588-018-0231-8 (visited on 09/09/2021).

Kondrashov, AS (2018). "Through sex, nature is telling us something important". In: *Trends in Genetics* 34.5, pp. 352–361.

Pardiñas, AF, P Holmans, AJ Pocklington, V Escott-Price, S Ripke, N Carrera, SE Legge, S Bishop, D Cameron, ML Hamshere, et al. (2018). "Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection". In: *Nature genetics* 50.3, pp. 381–389.

Sanjak, JS, J Sidorenko, MR Robinson, KR Thornton, and PM Visscher (2018). "Evidence of directional and stabilizing selection in contemporary humans". In: *Proceedings of the National Academy of Sciences* 115.1, pp. 151–156.

Simons, YB, K Bullaughey, RR Hudson, and G Sella (2018). "A population genetic interpretation of GWAS findings for human quantitative traits". In: *PLoS biology* 16.3, e2002985.

Haller, BC and PW Messer (2019). "SLiM 3: forward genetic simulations beyond the Wright–Fisher model". In: *Molecular biology and evolution* 36.3, pp. 632–637.

Sella, G and NH Barton (2019). "Thinking about the evolution of complex traits in the era of genome-wide association studies". In: *Annual review of genomics and human genetics* 20.1, pp. 461–493.

Buffalo, V (2021). "Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain Lewontin's Paradox". In: *Elife* 10, e67509.

Garcia, JA and KE Lohmueller (July 2021). "Negative linkage disequilibrium between amino acid changing variants reveals interference among deleterious mutations in the human genome". en. In: *PLOS Genetics* 17.7. Publisher: Public Library of Science, e1009676. ISSN: 1553-7404. DOI: `10.1371/journal.pgen.1009676`. URL: `https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1009676` (visited on 12/18/2025).

Johri, P, K Riall, H Becher, L Excoffier, B Charlesworth, and JD Jensen (July 2021). "The Impact of Purifying and Background Selection on the Inference of Population History: Problems and Prospects". In: *Molecular Biology and Evolution* 38.7, pp. 2986–3003. ISSN: 1537-1719. DOI: `10.1093/molbev/msab050`. URL: `https://doi.org/10.1093/molbev/msab050` (visited on 09/25/2025).

Sandler, G, SI Wright, and AF Agrawal (Oct. 2021). "Patterns and Causes of Signed Linkage Disequilibria in Flies and Plants". In: *Molecular Biology and Evolution* 38.10, pp. 4310–4321. ISSN: 1537-1719. DOI: `10.1093/molbev/msab169`. URL: `https://doi.org/10.1093/molbev/msab169` (visited on 12/18/2025).

Hayward, LK and G Sella (2022). "Polygenic adaptation after a sudden change in environment". In: *Elife* 11, e66697.

Lee, YCG (Feb. 2022). "Synergistic epistasis of the deleterious effects of transposable elements". In: *Genetics* 220.2, iyab211. ISSN: 1943-2631. DOI: `10.1093/genetics/iyab211`. URL: `https://doi.org/10.1093/genetics/iyab211` (visited on 12/18/2025).

Murphy, DA, E Elyashiv, G Amster, and G Sella (2022). "Broad-scale variation in human genetic diversity levels is predicted by purifying selection on coding and non-coding elements". In: *Elife* 12, e76065.

Simons, YB, H Mostafavi, CJ Smith, JK Pritchard, and G Sella (2022). "Simple scaling laws control the genetic architectures of human complex traits". Preprint at bioRxiv https://www.biorxiv.org/content/10.1101/2022.10.04.509926v1.

Buffalo, V and AD Kern (2024). "A quantitative genetic model of background selection in humans". In: *Plos Genetics* 20.3, e1011144.

Koch, E, N Connally, N Baya, M Reeve, M Daly, B Neale, E Lander, A Bloemendal, and S Sunyaev (2024). "Genetic association data are broadly consistent with stabilizing selection shaping human common diseases and traits". Preprint at bioRxiv https://www.biorxiv.org/content/10.1101/2024.06.19.599789v2.

Negm, S and C Veller (2024). "The effect of long-range linkage disequilibrium on allele-frequency dynamics under stabilizing selection". Preprint at bioRxiv https://www.biorxiv.org/content/10.1101/2024.06.27.601075v1.

Veller, C and GM Coop (2024). "Interpreting population-and family-based genome-wide association studies in the presence of confounding". In: *PLoS Biology* 22.4, e3002511.

Barroso, GV and AP Ragsdale (Sept. 2025). *A model for background selection in non-equilibrium populations*. en. ISSN: 2692-8205 Pages: 2025.02.19.639084 Section: New Results. DOI: 10.1101/2025.02.19.639084. URL: https://www.biorxiv.org/content/10.1101/2025.02.19.639084v5 (visited on 12/18/2025).

Berg, JJ, X Li, K Riall, LK Hayward, and G Sella (Dec. 2025). "Mutation–selection–drift balance models of complex diseases". In: *Genetics* 231.4, iyaf220. ISSN: 1943-2631. DOI: 10.1093/genetics/iyaf220. URL: https://doi.org/10.1093/genetics/iyaf220 (visited on 12/12/2025).

Ragsdale, AP (Mar. 2025). "Archaic introgression and the distribution of shared variation under stabilizing selection". en. In: *PLOS Genetics* 21.3. Publisher: Public Library of Science, e1011623. ISSN: 1553-7404. DOI: 10.1371/journal.pgen.1011623. URL: https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1011623 (visited on 12/17/2025).

Simons, YB, H Mostafavi, H Zhu, CJ Smith, JK Pritchard, and G Sella (Oct. 2025). "Simple scaling laws control the genetic architectures of human complex traits". en. In: *PLOS Biology* 23.10. Publisher: Public Library of Science, e3003402. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.3003402. URL: https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3003402 (visited on 10/16/2025).

A. The frequency spectrum under background selection in a recombining background

Table 1: Notation table for Supplementary Text A

Symbol	Description
Model Parameters	
N	Diploid population size.
M	Total chromosome length.
r	Per-site recombination rate.
ν	Per-site deleterious mutation rate.
s_b	Selection coefficient against a single background deleterious allele.
s	Direct selection coefficient of the focal allele.
γ	Population-scaled direct selection coefficient (e.g., $2Ns$).
$g_b(s_b)$	Distribution of fitness effects (DFE) for background mutations.
t	Time in generations.
Frequency Spectrum & Allele Notation	
q	Allele frequency.
$q(t)$	Expected allele frequency at time t.
θ	Population-scaled neutral mutation rate (i.e., $4N\mu$).
$\xi(q)$	Standard neutral frequency spectrum (proportional to θ/q).
$\xi(q \mid \gamma)$	Frequency spectrum for a directly selected allele with coefficient γ.
$\pi(p \mid \gamma)$	Fixation probability of a selected allele at frequency p.
Classic BGS Model (Pre-existing Mutations)	
T_{s_b}	Purging timescale for a haplotype with one deleterious allele ($\approx 1/s_b$).
M_{s_b}	Characteristic block length for classic BGS ($2s_b/r$).
λ	Expected number of **pre-existing** mutations on a block of length M_{bgs} ($2\nu/r$).
B_i	Proportion of haplotypes carrying i deleterious alleles in an M_{s_b} block.
B_0	Proportion of haplotypes free of deleterious alleles ($e^{-\lambda}$).
$\xi_{B_0}(q)$	Classic BGS approximation for the frequency spectrum ($\approx B_0\xi(q)$).
$q(T_{s_b})$	Frequency threshold for the classic BGS approximation ($\approx 1/(2Ns_bB_0)$).

Continued on next page

Table 1 – *Continued from previous page*

Symbol	Description
$B(q)$	Frequency spectrum skew in weak mutation limit ($e^{-\lambda(1-e^{-4Ns_bq})}$).
$\xi_B(q \mid \gamma)$	Frequency-dependent BGS approximation for a directly selected allele.
Fitness Drag Model (New Mutations)	
M_t	Expected length of the unrecombined block at time t ($2/(rt)$).
D_t	Expected number of new mutations accumulated on a block of length M_t (also $= \lambda$).
s_{drag}	Effective selection coefficient due to new mutations (λs_b).
T_{drag}	Purging timescale due to fitness drag ($\approx 1/s_{drag} = 1/(\lambda s_b)$).
M_{drag}	Characteristic block length for the fitness drag process ($2s_{drag}/r = 2\lambda s_b/r$).
$q(T_{drag})$	Frequency threshold for the fitness drag model ($1/(2Ns_e)$).
Fitness Variance & Cvijović et al. (2018) Comparison	
σ_W	Standard deviation of fitness.
$\sigma_W(t)$	SD of fitness for equilibrium blocks of length M_t ($\sqrt{M_t s_b \nu}$).
$\sigma_{W,M_{drag}}$	SD of fitness for equilibrium blocks of length M_{drag} (also $= s_{drag}$).
U_D	Total deleterious mutation rate per block (in the non-recombining model).
$T_{decline}$	Purging timescale in the non-recombining model ($\approx 1/(s_b\sqrt{\lambda})$).
$q(T_{decline})$	Frequency threshold in the non-recombining model ($\approx 1/(2Ns_b\sqrt{\lambda})$).

Classic background selection: removal due to pre-existing mutations

We consider a chromosome of length M recombining at rate r per site, on which deleterious alleles with selection coefficient s_b arise at rate ν per site. We are interested in the fate and frequency distribution of a focal allele that arises on a random chromosome in this population at time $t = 0$.

After t generations have passed, the distance (in each direction) from the focal allele to the first recombination event is exponentially distributed with rate rt. The length of the block that remains associated with the focal mutation without any intervening recombination is on expectation

$$M_t = \frac{2}{rt} \tag{A.1}$$

(the factor of 2 comes from the fact that there is an uninterrupted block with expected length $1/rt$ in

each direction). A haplotype carrying a single deleterious allele is purged from the population in

$$T_{s_b} \approx \frac{1}{s_b} \tag{A.2}$$

generations on expectation, so the expected length of a haplotype block over which selection discriminates between focal alleles linked to one vs. zero deleterious alleles is

$$M_{s_b} = \frac{2}{rT_{s_b}} = 2\frac{s_b}{r}. \tag{A.3}$$

This defines a characteristic length scale in the classic model of background selection in a recombining background. At mutation-selection balance, the expected number of deleterious alleles in a window of this size is

$$\lambda = M_{s_b}\frac{\nu}{s_b} = 2\frac{\nu}{r}, \tag{A.4}$$

and the distribution is approximately Poisson, with the proportion of haplotypes carrying i deleterious alleles within this window given by

$$B_i = \frac{\lambda^i}{i!}e^{-\lambda}. \tag{A.5}$$

This result is the source of standard "reduced effective population size" approximation for a long recombining chromosome. In this approximation, only a neutral allele that arises on a length M_{bgs} haplotype free of deleterious mutations (which occurs with probability $B_0 = e^{-\lambda}$) has any hope of being seen at appreciable frequency in the population. Therefore, while the neutral frequency spectrum in the absence of background selection would be expected to take the form

$$\xi\left(q\right) = \frac{\theta}{q}, \tag{A.6}$$

the classic background selection approximation implies a systematic reduction by a factor of B_0 in the number of alleles reaching any given frequency

$$\xi_{B_0}(q) \approx B_0\xi\left(q\right). \tag{A.7}$$

Because the effect of selection is not instantaneous, this approximation is only valid above a certain critical frequency. To determine what this frequency is, consider the transit of a focal neutral allele linked to a haplotype block carrying a single deleterious allele. This haplotype block originated from within the population of mutation free blocks, and all other blocks carrying at least one mutation will soon be extinct, so we can think of it as evolving within a population of reduced effective size NB_0.

As we discussed above, the allele will be eliminated by selection on a timescale of $T_{s_b} = {}^1/_{s_b}$

generations, and so will be prohibited from reaching frequencies that cannot be reached via genetic drift in that amount of time. For small values of t, the expected frequency of a neutral allele in a diploid population of size NB_0 conditional on not yet having been lost is approximately $q(t) \approx t/2NB_0$. Therefore, haplotypes carrying at least one deleterious allele are limited to frequencies no greater than approximately

$$q(T_{s_b}) = \frac{T_{s_b}}{2NB_0} = \frac{1}{2Ns_bB_0}. \tag{A.8}$$

Equation (A.8) thus gives an approximate threshold for the frequency at which equation (A.7) becomes valid. This is the the threshold suggested by Cvijović et al. 2018, provided that one replaces $\lambda = \frac{U_D}{s}$ in their haploid model of a non-recombining block with $\lambda = 2\frac{\nu}{r}$ in our diploid model of a recombining chromosome.

Length- and time-scale of removal due to selective drag of new mutations

The classic background selection model considered above assumes that the most relevant process is the removal of the haplotype that the focal allele arises on, due to selection against mutations that are already present there when the allele arises. If the deleterious mutation rate is high relative to the recombination rate, then it could also be removed due to selection against deleterious mutations accumulating after it arises. After t generations have passed, the haplotype associated with the deleterious allele has expected length $M_t \approx 2/rt$ (see eq. (A.1)), and so on expectation has accumulated

$$D_t = M_t \nu t = 2\frac{\nu}{r} = \lambda \tag{A.9}$$

additional deleterious alleles since it arose. It is notable that this expectation is independent of time, as the deleterious alleles that are separated by recombination are replaced by new deleterious mutations arising on the (on expectation, slightly shortened) block that remains associated. While their expectations are numerically equal ($D_t = \lambda$), it is important to distinguish these two quantities conceptually: D_t is the expected number of new mutations accumulated on a shrinking, age-t block, whereas λ is the expected number of pre-existing mutations found on an equilibrium block of length M_{bgs}.

This time-independence represents a key distinction from the non-recombining model of Cvijović et al. 2018, in which the absence of recombination means that the number of new mutations accumulates linearly with time (i.e. $D_t \propto t$). In that case, the process is one of steady fitness decline as the haplotype becomes progressively loaded with mutations. In the recombining model, the balance between recombination and mutation establishes a constant, time-independent expected mutational load, $D_t = \lambda$, for the block associated with the focal allele. (We say more about the similarities and distinctions between the two models below.)

Each recombination event generates a new set of associations with deleterious alleles at distances greater than the expected $1/rt$ distance to the edge of the block. However, these new associations are samples from the equilibrium distribution, and therefore have contributions that are on average equal to the population mean, and thus have no net effect on the mean fitness of the focal block. The result is that in the early generations before selection has had a chance to shape the trajectory of our focal allele, its expected fitness drag due to new mutations is constant:

$$s_{drag} = D_t s_b = \lambda s_b. \tag{A.10}$$

This suggests that a typical allele at the focal sites should be removed due to this immediately accumulated selective drag on a timescale of

$$T_{drag} = \frac{1}{s_{drag}} = \frac{1}{\lambda s_b} \tag{A.11}$$

generations, and that the expected length of the haplotype block that is removed with it is

$$M_{drag} = \frac{2}{rT_{drag}} = 2\frac{s_{drag}}{r} = 2\frac{\lambda s_b}{r} = \lambda M_{s_b}. \tag{A.12}$$

This process of selection against newly mutations arising on the background of the focal allele will be important for frequencies less than

$$q\left(T_{drag}\right) = \frac{T_{drag}}{2N} = \frac{1}{2Ns_{drag}} = \frac{1}{2N\lambda s_b}, \tag{A.13}$$

but becomes unimportant above this threshold because any alleles exceeding this frequency are likely outliers that have accumulated fewer deleterious mutations than average.

As we indicated above, this process is only important to the dynamics if recombination is weak relative to mutation, in which case the timescale of removal due to the accumulation of deleterious mutations is faster than the timescale on which mutation-free haplotypes are preserved, i.e. $T_{drag} \ll T_{s_b}$. Intuitively, this condition translates into a requirement that $\lambda \gg 1$.

More precisely, we might think that in order for the effects of the immediate fitness drag to be important, the threshold frequency at which this fitness drag has substantially reshaped the frequency spectrum must be below the threshold frequency at which the classic background approximation begins to apply, i.e. $q\left(T_{drag}\right) \ll q\left(T_{s_b}\right)$. This is equivalent to the condition that $B_0 T_{drag} \ll T_{s_b}$. This leads to the more restrictive condition $\lambda \gg e^{-\lambda}$ on the value of λ at which the fitness drag becomes important. This is solved by $\lambda \gg W_0\left(1\right) \approx 1/2$, i.e. $\nu \gg r^{W_0(1)}/2 \approx r/4$, where W_0 is the main branch of the Lambert W function and, very roughly, $W_0(1) \approx 1/2$. This condition more closely matches our observations from simulations that the fitness drag begins to impact the frequency spectrum roughly when $\lambda > 1/2$ (Figure S1).

Comparison to Cvijović et al. (2018) non-recombining model This "fitness drag" process provides a close parallel to the "fitness decline" process in the non-recombining model of Cvijović et al. 2018. In that model, λ is defined as $\lambda = {}^{U_D}/_{s_b}$, where U_D is the total deleterious mutation rate for the entire non-recombining block. For an equivalent value of λ, both models predict the same overall reduction in effective population size ($B_0 = e^{-\lambda}$). Furthermore, if s_b is also held constant, the upper frequency threshold $q(T_{s_b})$—which defines the frequency below which the classic BGS approximation of equation (A.7) breaks down—is also the same. This equivalence follows directly from the validity of the "effectively non-recombining block" approximation, which simply converts the recombining problem into the corresponding non-recombining one for alleles above the relevant threshold.

The most significant difference between the two models lies in their predictions regarding the impact of selection on alleles at very low frequencies. Notably, because the "block length" is fixed in the non-recombining model, there is no contribution from the accumulation of deleterious alleles at distances beyond the characteristic block length. Instead, the block must first decline in fitness by accumulating deleterious mutations. As a result, we speak of a process of "fitness decline" and its consequences, rather than the "fitness drag" process described above for the recombining model.

In the non-recombining model, these accumulated mutations become important once the fitness of the block has declined by an amount equal to the equilibrium standard deviation in fitness, $\sigma_W = s_b\sqrt{\lambda}$. At this point, the fitness deficit relative to the rest of the population is substantial enough that selection can effectively remove the allele.

Ignoring the impact of selection on short timescales, fitness declines linearly with time in the non-recombining model (i.e., $s_e(t) \approx (U_D t)s_b$). Setting the accumulated load $(U_D t)s_b$ equal to the equilibrium standard deviation $\sigma_W = s_b\sqrt{\lambda}$ and solving for t, the timescale on which an allele declines sufficiently in fitness to be removed by selection is

$$T_{decline} \approx \frac{s_b\sqrt{\lambda}}{U_D s_b} = \frac{1}{s_b\sqrt{\lambda}} = \frac{1}{\sigma_W}. \tag{A.14}$$

Consequently, in the non-recombining model, we expect that for frequencies below a threshold of

$$q(T_{decline}) \approx \frac{T_{decline}}{2N} = \frac{1}{2Ns_b\sqrt{\lambda}} = \frac{1}{2N\sigma_W}, \tag{A.15}$$

selection has not yet had an opportunity to act, so the frequency spectrum should be approximately neutral. Once an allele reaches this frequency, however, it is likely to have accumulated a fitness deficit of approximately $s_e\left(T_{decline}\right) = \sigma_W$, and will therefore be eliminated on a timescale of roughly $T_{decline} = {}^{1}/_{\sigma_W}$ additional generations. During this time, the allele's frequency can increase by at most another factor of roughly 2. Thus, equation (A.15) provides the correct order-of-magnitude limit on the frequency at which the accumulation of new mutations dominates the dynamics.

Notably, the equivalent threshold in the recombining model, $q\left(T_{drag}\right) = {}^{1}/_{2Ns_b\lambda}$ (Eq., (A.13)), is

lower by a factor of $1/\sqrt{\lambda}$. This difference reflects the difference in dynamics between the two models which can also be understood in terms of an appropriate standard deviation of fitness. Specifically, in the recombining model, the drag s_{drag} is equal to the equilibrium fitness standard deviation of blocks with length M_{drag} (i.e., blocks of the characteristic length eliminated by this process). The equilibrium distribution of pre-existing mutations on a block of length M_{drag} is Poisson with mean $M_{drag}(\nu/s_b) = (2\lambda s_b/r)(\nu/s_b) = \lambda^2$. The standard deviation of fitness among blocks of this length is therefore:

$$\sigma_{W,M_{drag}} = \sqrt{\lambda^2 s_b^2} = \lambda s_b = s_{drag}. \tag{A.16}$$

This equality provides a dynamic interpretation of the removal timescale, T_{drag}. When an allele is young ($t \ll T_{drag}$), its associated block is long ($M_t \gg M_{drag}$). The background standard deviation of fitness for very young blocks is consequently very large ($\sigma_W(t) = \sqrt{M_t s_b \nu} \gg \sigma_{W,M_{drag}}$), and the allele's constant fitness drag is small relative to this background variation. As the allele ages and t approaches T_{drag}, its associated block shrinks toward length M_{drag}, and the background fitness variance shrinks with it. When the block reaches length M_{drag}, the background variance shrinks to equal the fitness drag ($\sigma_{W,M_{drag}} = s_{drag}$), and it is around this time that the allele is likely to be purged.

This difference in mechanism—the accumulation of fitness deficit versus a constant fitness drag—arises because, in the recombining model, deleterious mutations far beyond the characteristic block length of classic BGS exert an immediate influence. This extended range accelerates the onset of strong selection: the average fitness deficit becomes visible to selection on a timescale shorter by a factor of

$$\frac{T_{drag}}{T_{decline}} = \frac{1}{\sqrt{\lambda}} \tag{A.17}$$

compared to the equivalent non-recombining model, and consequently that BGS should skew the frequency spectrum at frequencies that are also lower by a factor of

$$\frac{q\left(T_{drag}\right)}{q\left(T_{decline}\right)} = \frac{1}{\sqrt{\lambda}} \tag{A.18}$$

This perspective on the recombining model is somewhat at odds with the idea that the allele is under strong, constant selection immediately. For early times $t \ll T_{drag}$, when the drag is much smaller than the background variance ($s_{drag} \ll \sigma_W(t)$), the argument made by Cvijović et al. (2018) in the context of the non-recombining model would suggest that the allele should behave neutrally at first due to the noise of the background. This tension is in some part real, and reflects the fact that all of these are coarse approximations for an extremely complex process. However, this tension is partially resolved by the fact that the argument from Cvijović et al. (2018) is a bit too strong, and also partially by the difference in dynamics between the two models.

In the non-recombining model, the fitness deficit grows linearly relative to a fixed standard devi-

ation, meaning the ratio also scales linearly with time:

$$\frac{s_e(t)}{\sigma_W} \approx \frac{(U_D t)s_b}{s_b\sqrt{\lambda}} = ts_b\sqrt{\lambda} = \frac{t}{T_{decline}}. \tag{A.19}$$

Consequently, the time-averaged value of $s_e(t)/\sigma_W$ over the interval $[0, T_{decline}]$—the period during which the allele's fitness deficit increases from 0 to σ_W—is $\int_0^{T_{decline}} t/T_{decline} \mathrm{d}t = 1/2$. This suggests that if one wishes to approximate the early dynamics using a constant selection coefficient, $\sigma_W/2$ may be a more useful choice than 0.

In contrast, in the recombining model, the background fitness standard deviation shrinks with $1/\sqrt{t}$, while the fitness deficit remains constant. Consequently, the ratio scales with the square root of time:

$$\frac{s_{drag}}{\sigma_W(t)} \approx \frac{\lambda s_b}{\sqrt{s_b\lambda/t}} = \sqrt{ts_b\lambda} = \sqrt{\frac{t}{T_{drag}}}. \tag{A.20}$$

The time-average of $s_{drag}/\sigma_W(t)$ over the interval $[0, T_{drag}]$ is therefore $\int_0^{T_{drag}} \sqrt{t/T_{drag}} \mathrm{d}t = 2/3$. This reflects the fact that the ratio of the fitness deficit to the background standard deviation is higher on average during the early phases of the allele's trajectory than in the non-recombining case (Figure S4). This suggests that to approximate the low-frequency dynamics of the recombining model using a single constant selection coefficient, one should use $2s_{drag}/3$, thereby discounting for the impact of background fitness variance in the earliest generations.

These arguments suggest that the frequency spectrum is more strongly distorted in recombining models than in equivalent non-recombining ones, not only because the focal allele becomes "visible" to selection faster, but also because it spends more time near this visibility threshold before reaching it. While these arguments are heuristic, the central conclusion is deeply intuitive: the suppression of the frequency spectrum extends to lower frequencies in large recombining genomes than predicted by non-recombining models with equivalent reductions in effective size, due to the action in early generations of deleterious mutations at recombination distances greater than the characteristic length-scale of the classical recombining BGS model.

The frequency spectrum at low frequencies

The strong recombination/weak mutation regime

Neutral alleles The arguments in the preceding sections allow us to understand the shape of the frequency spectrum for neutral alleles at very low and very high frequencies. Roughly, when $\lambda \ll 1/2$ (or equivalently, when $B_0 \gg 0.6$) recombination is too fast relative to mutation for the process of fitness decline to be relevant (i.e. $T_{drag} \gg B_0 T_{s_b}$), and so the frequency spectrum is shaped entirely by the effect of selection acting against deleterious alleles that are already present when the focal

neutral allele arises. Based on the approximations given above, this suggests that

$$\xi_{B_0}(q) = \begin{cases} \frac{\theta}{q} & q \ll \frac{1}{2Ns_bB_0} \\ \frac{\theta}{q}B_0 & q \gg \frac{1}{2Ns_bB_0}. \end{cases} \tag{A.21}$$

We can push this perspective further and obtain an approximation that is valid across all frequencies by taking advantage of the assumption that the focal allele is eliminated by deleterious alleles already present on its background when it arises. To this end, consider that with probability B_i the focal allele arises on a haplotype that carries i deleterious alleles within a distance of $M_{bgs}/2 = s_b/r$ in either direction. Provided that carrying *any* deleterious alleles on a haplotype of this length is rare in general (consistent with the assumption that $B_0 \gg 0.6$), then standard diffusion results tell us a neutral allele on a haplotype with i deleterious alleles will spend approximately $\frac{2}{q}e^{-4Ns_biq}$ generations at frequency q. We can therefore approximate the frequency spectrum across the full range of frequencies as

$$\xi_B(q) \approx \xi(q) \sum_{i=0}^{\infty} B_i e^{-4Ns_biq} \tag{A.22}$$

$$= \xi(q)\, e^{-\lambda} \sum_{i=0}^{\infty} \frac{\left(\lambda e^{-4Ns_bq}\right)^i}{i!} \tag{A.23}$$

$$= \xi(q)\, B(q, \lambda), \tag{A.24}$$

where

$$B(q, \lambda) = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\left(\lambda e^{-4Ns_bq}\right)^i}{i!} \tag{A.25}$$

$$= e^{-\lambda\left(1-e^{-4Ns_bq}\right)}. \tag{A.26}$$

is a frequency dependent B value, the form of which follows from the fact that the infinite sum in equation (A.25) is the full Taylor series representation of the function $e^{\lambda e^{-4Ns_bq}}$ expanded around the point $\lambda e^{-4Ns_bq} = 0$. We note that the sum in equation H3 of Cvijović et al. 2018 is essentially identical to our equation (A.22) above, although they evaluate it only to zeroth order, i.e., $\xi_B(q) = \xi(q) + \mathcal{O}(\lambda)$.

A distribution of background fitness effects This frequency dependent B value can easily be generalized for a distribution of fitness effects among the sites generating background selection. Suppose that sites responsible for background selection draw their selection coefficients from some distribution, $g_b(s_b)$, and suppose that the largest selection coefficients sampled from this distribution still conform to the simplifying assumptions that $s_b \ll Mr$ and $2Ns_b \gg 1$.

Under these assumptions, the number of linked deleterious mutations is Poisson distributed with mean λ, and the selection coefficient of each mutation is drawn independently from $g_b(s_b)$. Since the

effect of each mutation on the sojourn time is multiplicative (multiplying the neutral time by e^{-4Ns_bq}), the total frequency dependent B value is the expectation of the product of these effects, yielding:

$$\begin{aligned} B\left(q,\lambda,g_b\right) &= \sum_{k=0}^{\infty} B_i \left(\mathbb{E}_{g_b}\left[e^{-4Ns_bq}\right]\right)^k \\ &= e^{-\lambda\left(1-\mathbb{E}_{g_b}\left[e^{-4Ns_bq}\right]\right)}. \end{aligned} \tag{A.27}$$

where $\mathbb{E}_{g_b}\left[e^{-4Ns_bq}\right] = \int g_b\left(s_b\right) e^{-4Ns_bq} ds_b$. Note that this derivation appears to neglect the fact that the characteristic length scale depends on the selection coefficient. However, while we might expect this dependency to complicate the strong mutation limit, it is irrelevant in the weak mutation limit. In this regime, a haplotype is typically eliminated by a single deleterious allele; consequently, the effective length scale is determined solely by the mutation responsible for the elimination. The result therefore correctly averages over the relevant distribution of length scales.

Directly selected alleles When the focal allele undergoing background selection is itself directly selected, we can find useful approximations by straightforwardly extending the results above under neutrality. First, consider the expected frequency spectrum of a deleterious allele in the absence of linked selection, which is

$$\xi\left(q \mid \gamma\right) = 2\frac{e^{2\gamma q}}{q(1-q)}\left(1-\pi\left(q \mid \gamma\right)\right) \tag{A.28}$$

where

$$\pi\left(q \mid \gamma\right) = \begin{cases} q & \gamma = 0 \\ \frac{1-e^{-2\gamma q}}{1-e^{-2\gamma}} & \gamma \neq 0 \end{cases} \tag{A.29}$$

is the fixation probability of an allele at frequency q with scaled selection coefficient γ.

In Charlesworth (1994)'s classical approximation for weakly selected alleles (i.e. $\gamma \sim 1$), we simply ignore all variants arising on backgrounds other than the mutation free background, i.e.

$$\xi_{B_0}\left(q \mid \gamma\right) = B_0 \xi\left(q \mid \gamma B_0\right). \tag{A.30}$$

This approximation extends the "scaled mutation rate" effect of background selection by including a "scaled selection coefficient" effect. The argument for this approximation relies on a separation of timescales between background selection (which acts quickly), and direct selection (which is assumed to act slowly). Alleles reaching frequencies greater than $1/2Ns_bB_0$ must have originated on backgrounds with $i = 0$ deleterious alleles. Once these alleles reach appreciable frequency, new deleterious mutations occurring on their background spread the focal allele across all fitness classes, but because

backgrounds with $i > 0$ deleterious alleles are always quickly removed, only the population of size NB_0 is relevant to the evolution of the focal allele. As a result, the balance between drift and selection over the longer timescale on which (weak) direct selection acts is captured by a rescaled version of the population scaled selection coefficient: γB_0.

Combining this classic separation of timescales argument with the frequency dependent B value argument we articulated above, we can approximate the full frequency spectrum of directly selected alleles under weak selection as

$$\xi_B\left(q \mid \gamma\right) = \xi\left(q \mid \gamma B_0\right) \sum_{i=0}^{\infty} B_i e^{-4Ns_b iq} \tag{A.31}$$

$$= B(q, \lambda)\xi\left(q \mid \gamma B_0\right). \tag{A.32}$$

Because $B_0 = e^{-\lambda}$ regardless of the value of s_b (provided that $s_b \ll Mr$), this approximation can be straightforwardly extended to the case with a distribution of background fitness effects using equation (A.27).

Alternatively, when the strength of direct selection is similar to that of background selection, this separation of timescales does not apply. Direct and background selection act on the same timescale, leading to:

$$\xi_B(q \mid \gamma) = \sum_{i=0}^{\infty} B_i \xi(q \mid \gamma - 2Ns_b i) \tag{A.33}$$

$$= \xi(q \mid \gamma) \sum_{i=0}^{\infty} B_i e^{-4Ns_b iq} \tag{A.34}$$

$$= B(q, \lambda)\xi(q \mid \gamma) \tag{A.35}$$

The strong mutation/weak recombination regime

When recombination is weak relative to mutation (i.e., $\lambda \gg 1/2$), the frequency spectrum is shaped differently. As established, this condition implies $T_{drag} \ll T_{s_b}$, creating two distinct selective thresholds, $q(T_{drag}) \ll q(T_{s_b})$, which define three frequency regimes. In the lowest frequency regime, $q \ll q(T_{drag})$, the frequency spectrum is shaped by the emergence of the time-independent, constant expected fitness drag against the shrinking background fitness variation. In the highest frequency regime, $q \gg q(T_{s_b})$, all alleles originated on the fittest background, and the frequency spectrum is well approximated by the classic rescaled effective population size. For the intermediate regime, where $q(T_{drag}) \ll q \ll q(T_{s_b})$, Cvijović et al. (2018) provide an approximation in the non-recombining model for the form of the frequency spectrum. Here, for completeness, we use heuristic arguments to recapitulate their results in the context of the recombining model.

The key to this regime is that our constant s_{drag} approximation, while useful for describing the

behavior at frequencies below the lower threshold $q(T_{drag})$, breaks down at frequencies above it. Specifically, $s_{drag} = \lambda s_b$ represents the expected fitness drag for an allele that has not yet persisted long enough for selection to significantly impact its trajectory. For an allele to reach a frequency $q \gg q(T_{drag})$, it must have been lucky, arising on a background with $i \ll \lambda$ deleterious alleles.

Consequently, the dynamics of the intermediate regime are governed by the "classic" BGS process. While this is the same process as in the $\lambda \ll 1$ case, the required mathematical approximation is different. When $\lambda \ll 1$, the variance of the distribution of i is low, and the fittest class (i.e., $i = 0$) is essentially the same as the population mean. In contrast, when $\lambda \gg 1$, the variance is high, and the lucky alleles that survive into the intermediate regime are extreme outliers, far fitter than the population mean (i.e., $i \ll \lambda$).

When $\lambda \gg 1$, the probability of arising on a background with i deleterious alleles declines quickly as $i \to 0$. It follows that the set of backgrounds reaching frequency q will be dominated by the least fit backgrounds (i.e., those with the highest value of i) that are nonetheless capable of reaching frequency q. We can heuristically derive this least fit class, $i(q)$, via an extreme value argument. To reach a frequency q, an allele must arise on one of the fittest "available" backgrounds. We can estimate the number of "available" backgrounds as $2Ns_bq$ (the number of individuals at that frequency, $2Nq$, divided by the selection timescale, $T_{s_b} \approx 1/s_b$). The fittest available background class, i, is the one that solves the extreme value condition:

$$(2Ns_bq) \cdot B_i \approx 1 \tag{A.36}$$

where $B_i = \frac{\lambda^i}{i!}e^{-\lambda}$ is the probability of arising on a background with i deleterious alleles within the characteristic block of length M_{bgs} (Eq. (A.5)). Taking the log of both sides and applying Stirling's approximation ($\ln(i!) \approx i \ln i - i$) gives:

$$\ln(2Ns_bq) + \ln(B_i) \approx 0 \tag{A.37}$$

$$\ln(2Ns_bq) + i \ln \lambda - (i \ln i - i) - \lambda \approx 0 \tag{A.38}$$

$$\ln(2Ns_bq) \approx \lambda - i - i \ln \lambda + i \ln i \tag{A.39}$$

$$\ln(2Ns_bq) \approx \lambda - i(1 + \ln \lambda - \ln i). \tag{A.40}$$

Assuming that the $\ln \lambda$ term dominates, we can approximate $1 + \ln \lambda - \ln i \approx \ln \lambda$ (this holds when λ is large and i is effectively intermediate—i.e., $1 \ll i \ll \lambda$). This allows us to solve for i as a function of q:

$$i(q) \approx \frac{\lambda - \ln(2Ns_bq)}{\ln \lambda} = \log_\lambda \left(\frac{1}{2NB_0s_bq} \right) \tag{A.41}$$

$$= \log_\lambda \left(\frac{q(T_{s_b})}{q} \right) \tag{A.42}$$

This $i(q)$ represents a "self-consistency" condition (equivalent to $k_c(f) + 1$ in Cvijović et al. 2018), which shows that the dominant fitness class among backgrounds reaching frequency q declines logarithmically with increasing frequency. The alternative form for $i\,(q)$ in Equation (A.42) follows from the definition $q(T_{s_b}) = {}^1\!/_{2NB_0 s_b}$. This form emphasizes that as q increases, the allele must have been increasingly lucky to survive: to reach higher frequencies, it must have arisen on a background with fewer deleterious alleles, until eventually, to reach frequency $q \approx q\,(T_{s_b})$, the allele must have arisen on a background with $i = 0$ (by which point this approximation has broken down).

The expected number of generations an allele spends at frequency $q(T_{drag}) \ll q \ll q(T_{s_b})$ is thus approximately

$$\xi_B(q \mid q(T_{drag}) \ll q \ll q(T_{s_b})) \approx B_{i(q)} \cdot \xi_B(q \mid i = i(q), q(T_{drag}) \ll q \ll q(T_{s_b})). \tag{A.43}$$

We already assumed above that $B_{i(q)} \approx {}^1\!/_{2Ns_b q}$ (Eq. (A.36)). Given that the allele arises in this class, the amount of time spent at frequency q has the approximate form

$$\xi_B(q \mid i = i(q), q(T_{drag}) \ll q \ll q(T_{s_b})) \approx \xi(q) F_{i(q)} \tag{A.44}$$

where $\xi(q) = {}^\theta\!/_q$ is the standard neutral spectrum and $F_{i(q)}$ is a factor that accounts for the suppression of this neutral frequency spectrum by selection in the time since the allele arose. Given the fundamental timescale, $T_{s_b} = {}^1\!/_{s_b}$, of the background selection process, $F_{i(q)}$ can be approximated as

$$F_{i(q)} \approx \frac{1}{\sum_{t=0}^{T_{s_b}} s_{eff}(t)}$$

where $s_{eff}(t)$ is the effective selection coefficient experienced by the allele in generation t. Notably, $s_{eff}(0) = s_b i(q)$, but the allele will experience a range of selective environments over its transit through the population, due to a combination of mutation and recombination events shuffling it across the distribution of fitness backgrounds. This process is complicated, so we might naturally want to make an approximation $\sum_{t=0}^{T_{s_b}} s_{eff}(t) \approx T_{s_b} s_{eff,i(q)}$, for some appropriate average effective selection coefficient, $s_{eff,i(q)}$. Most backgrounds are less fit than the $i(q)$ background on which it arose, which suggests that $s_{eff,i(q)} > s_{eff}(0)$. Alternatively, in order to be found at frequency q, the allele cannot have spent too much time on much less fit backgrounds, or it would have been purged. When one conducts a rigorous analysis of the dynamics (as Cvijović et al. 2018, do for the non-recombining case), one finds that $s_{eff,i(q)} = s_b \sqrt{\lambda i\,(q)}$, which is the geometric mean of the allele's initial selection coefficient $s_{eff}(0) = s_b i(q)$, and the one that it would experience on average if it were sampling the distribution completely at random, i.e., $s_b \lambda$, and therefore nicely satisfies both of these criteria.

It follows that

$$\begin{aligned} F_{i(q)} &\approx \frac{1}{T_{s_b} s_{eff,i(q)}} \\ &= \frac{s_b}{s_{eff,i(q)}} \\ &= \frac{1}{\sqrt{\lambda i\left(q\right)}}. \end{aligned} \tag{A.45}$$

Thus, putting together Eqs. (A.41)-(A.45), we have

$$\xi_B(q \mid q(T_{drag}) \ll q \ll q(T_{s_b})) \approx \xi(q) \frac{1}{2Ns_b q\sqrt{\lambda \log_\lambda\left(\frac{1}{2NB_0 s_b q}\right)}}. \tag{A.46}$$

Combining this with the heuristic $s_{eff} = \frac{2}{3}s_{drag} = \frac{2}{3}\lambda s_b$ approximation we derived above for the $q \ll q(T_{drag})$ regime, this allows us to write an approximation for the full spectrum for neutral alleles in this regime as a three-part piecewise function, analogous to Equation I21 in Cvijović et al. 2018:

$$\xi_B(q) \approx \xi\left(q\right) \begin{cases} e^{-\frac{8}{3}N\lambda s_b q} & q \ll q(T_{drag}) \\ \frac{1}{2Ns_b q\sqrt{\lambda \log_\lambda\left(\frac{1}{2NB_0 s_b q}\right)}} & q(T_{drag}) \ll q \ll q(T_{s_b}) \\ B_0 & q \gg q(T_{s_b}). \end{cases} \tag{A.47}$$

Notably, this approximation neglects the deterministic sweep-like behavior described by Cvijović et al. (2018) at very high frequencies. We expect that an effectively non-recombining block approximation could be derived for this regime as well.

B. Derivations Supporting the Exponential Fitness Model

Selection coefficient

To begin, we divide an individual's phenotype into a contribution from the focal site ℓ, which is $a_\ell g_\ell$ (where $g_\ell \in \{0, 1, 2\}$ is the genotype at site ℓ), and the contribution from all other sites, which we write as $Z_{-\ell}$. Writing $f\left(Z_{-\ell}\right)$ for the density on this background contribution, the expected fitness of an individual with genotype g_ℓ is exactly

$$\mathbb{E}\left[W \mid g_\ell\right] = \int e^{-\frac{a_\ell g_\ell + Z_{-\ell}}{\eta}} f\left(Z_{-\ell}\right) dZ_{-\ell} \tag{B.1}$$

$$= e^{-\frac{a_\ell g_\ell}{\eta}} \int e^{-\frac{Z_{-\ell}}{\eta}} f\left(Z_{-\ell}\right) dZ_{-\ell} \tag{B.2}$$

The ability to factor the fitness effect of the focal site ℓ out of the total effect is a special property of the exponential model, which allows us to rescale the relative fitnesses so that independent of any details about the background, $Z_{-\ell}$, we can write

$$\mathbb{E}\left[W \mid g_\ell\right] = e^{-\frac{a_\ell g_\ell}{\eta}}, \tag{B.3}$$

without any loss of generality or approximation.

The expected change in frequency for an allele with effect a and frequency x can be written as

$$\mathbb{E}\left[\Delta x \mid a, x, \eta\right] = s(a,x)x(1-x) \tag{B.4}$$

where

$$s(a,x) = \frac{1}{2}\frac{1}{\overline{W}}\frac{d\overline{W}}{dx}. \tag{B.5}$$

is the marginal effect of the allele on fitness. The mean fitness is equal to

$$\overline{W} = (1-x)^2 \cdot 1 + 2x(1-x)e^{-\frac{a}{\eta}} + x^2 e^{-\frac{2a}{\eta}} \tag{B.6}$$

$$= \left(1 - x + xe^{-\frac{a}{\eta}}\right)^2 \tag{B.7}$$

so that

$$\frac{d\bar{W}}{dx} = 2\left[(1-x) + xe^{-\frac{a}{\eta}}\right]\left(e^{-\frac{a}{\eta}} - 1\right) \tag{B.8}$$

and therefore

$$s(a,x) = \frac{1}{\left[(1-x) + xe^{-\frac{a}{\eta}}\right]^2} \cdot \left[(1-x) + xe^{-\frac{a}{\eta}}\right]\left(e^{-\frac{a}{\eta}} - 1\right) \tag{B.9}$$

$$= \frac{e^{-\frac{a}{\eta}} - 1}{(1-x) + xe^{-\frac{a}{\eta}}}. \tag{B.10}$$

Equation (B.10) is still exact. Assuming that $a \ll \eta$, $xe^{-\frac{a}{\eta}} \approx x$, and $e^{-\frac{a}{\eta}} - 1 \approx -\frac{a}{\eta}$, so we can make the approximation

$$s(a,x) \approx -\frac{a}{\eta}. \tag{B.11}$$

Impact of BGS on the Mean Phenotype and Fitness Load

Here we derive the effect of background selection on the mean phenotype and the corresponding fitness load. The mean phenotype is given by the sum of contributions from sites fixed for the trait-increasing allele. At mutation-selection-drift balance, the fraction of sites fixed for the trait-increasing (deleterious) allele with scaled selection coefficient γ is

$$p^+(\gamma) = \frac{1}{1+e^{2\gamma}}. \tag{B.12}$$

In the presence of background selection, the effective population size is reduced by a factor B, rescaling the selection coefficient to γB. The change in the fraction of fixed deleterious alleles is therefore

$$\Delta p^+(\gamma B) = p^+(\gamma B) - p^+(\gamma) = \frac{1}{1+e^{2\gamma B}} - \frac{1}{1+e^{2\gamma}}. \tag{B.13}$$

The shift in the mean phenotype is the sum of the effects of these additional fixed alleles across all L loci:

$$\bar{G}_B - \bar{G} = 2La\left(p^+(\gamma B) - p^+(\gamma)\right). \tag{B.14}$$

To compare this shift to the scale of the standing genetic variation, we normalize by the genetic standard deviation $\sqrt{V_G} = \sqrt{\frac{\theta}{\gamma}b(\gamma)}$, where $b(\gamma) = \tanh(\gamma)$:

$$\frac{\bar{G}_B - \bar{G}}{\sqrt{V_G}} = 2La\left(p^+(\gamma B) - p^+(\gamma)\right)\sqrt{\frac{\gamma}{\theta b(\gamma)}} \tag{B.15}$$

$$= \gamma\frac{L}{\sqrt{\theta}}\sqrt{\frac{b(\gamma)}{\gamma}}\frac{2\left(p^+(\gamma B) - p^+(\gamma)\right)}{b(\gamma)}. \tag{B.16}$$

For weakly selected sites ($\gamma \ll 1$), we can approximate $p^+(\gamma) \approx \frac{1}{2}(1-\gamma)$ and $b(\gamma) \approx \gamma$. Substituting these approximations yields:

$$\frac{\bar{G}_B - \bar{G}}{\sqrt{V_G}} \approx \gamma\frac{L}{\sqrt{\theta}}\frac{2\left[\frac{1}{2}(1-\gamma B) - \frac{1}{2}(1-\gamma)\right]}{\gamma} = \gamma(1-B)\frac{L}{\sqrt{\theta}}. \tag{B.17}$$

Since $L \gg \sqrt{\theta}$, this standardized shift is substantial ($\gg 1$) even for moderate background selection.

This shift in the mean phenotype results in a decline in mean fitness. The ratio of mean fitness with BGS to mean fitness without BGS is:

$$\frac{W(\bar{G}_B)}{W(\bar{G})} = \frac{e^{-\bar{G}_B/\eta}}{e^{-\bar{G}/\eta}} = e^{-\frac{1}{\eta}(\bar{G}_B - \bar{G})}. \tag{B.18}$$

Substituting the expression for the shift in the mean:

$$\frac{W(\bar{G}_B)}{W(\bar{G})} = \exp\left[-\frac{2La}{\eta}\left(p^+(\gamma B) - p^+(\gamma)\right)\right]. \tag{B.19}$$

Using the relation $\frac{2Na}{\eta} \approx \gamma$ (so $\frac{a}{\eta} \approx \frac{\gamma}{2N}$), this becomes:

$$\frac{W(\bar{G}_B)}{W(\bar{G})} = \exp\left[-\frac{L\gamma}{N}\left(p^+(\gamma B) - p^+(\gamma)\right)\right]. \tag{B.20}$$

For weak selection ($\gamma \ll 1$), we use the approximation $p^+(\gamma B) - p^+(\gamma) \approx \frac{1}{2}\gamma(1-B)$. Furthermore, note that $p^+(\gamma) \approx 1/2$ and $p^-(\gamma) \approx 1/2$, so $4p^+(\gamma)p^-(\gamma) \approx 1$. Thus, we can write:

$$\frac{W(\bar{G}_B)}{W(\bar{G})} \approx 1 - \frac{L\gamma}{N}\frac{\gamma(1-B)}{2} \tag{B.21}$$

$$= 1 - 4Ls\gamma(1-B)\left(\frac{1}{4}\right) \tag{B.22}$$

$$\approx 1 - 4Ls\gamma(1-B)p^+(\gamma)p^-(\gamma). \tag{B.23}$$

This result shows that for weakly selected sites, background selection drives a reduction in mean fitness proportional to the strength of selection, the total mutation rate, and the reduction in effective population size.

C. Changes in disease prevalence in the liability threshold model due to BGS

Derivation Let V_E environmental variance, and let V_G be the genetic variance in the absence of background selection. The heritability in the absence of background selection is then

$$h^2 = \frac{V_G}{V_G + V_E}, \tag{C.1}$$

and the standardized threshold density is

$$\phi(T^*) = f(T)\sqrt{V_p}, \tag{C.2}$$

where $V_P = V_G + V_E = \frac{V_G}{h^2}$ is the total variance.

In the presence of background selection, the genetic variance is reduced to $V_{G,B} = BV_G$, and the

phenotypic variance to $V_{P,B} = V_{G,B} + V_E$, so the heritability becomes

$$h_B^2 = \frac{V_{G,B}}{V_{P,B}} = \frac{V_{G,B}}{V_{G,B} + V_E} = \frac{BV_G}{BV_G + V_E}, \tag{C.3}$$

and we can write the standardized threshold density in the presence of background selection in terms of original quantities in its absence as

$$\phi(T_B^*) = f_B(T)\sqrt{V_{P,B}} = \frac{f(T)}{B}\sqrt{BV_G + V_E}, \tag{C.4}$$

where $T_B^* = \frac{T - \bar{G}_B}{\sqrt{V_{P,B}}}$.

Now, we can eliminate the V_E by solving Equation (C.1) for $V_E = \frac{1-h^2}{h^2}V_G$ and substituting to get

$$\phi(T_B^*) = f_B(T)\sqrt{V_{P,B}} = \frac{f(T)}{B}\sqrt{V_G\left(\frac{1}{h^2} - (1-B)\right)}, \tag{C.5}$$

which is now written entirely in terms of counterfactual quantities from the equilibrium in the absence of background selection, and the B value. The increase in the standardized threshold density due to background selection is thus

$$\frac{\phi(T_B^*)}{\phi(T^*)} = \frac{\sqrt{1 - h^2(1-B)}}{B}, \tag{C.6}$$

Translating this into a statement about the effect on the prevalence requires that we make some assumption about the relationship between the standardized threshold density and the prevalence. The standard assumption is that the distribution of liability is Normal, in which case the two expressions for the standardized threshold density above can be plugged into standard normal expressions to obtain the prevalence in each case. However, Berg et al. (2025) argued that when the prevalence is low, the relationship between the standardized threshold density and the prevalence is very roughly linear. It follows that

$$\frac{\bar{R}_B}{\bar{R}} \approx \frac{\sqrt{1 - h^2(1-B)}}{B}, \tag{C.7}$$

where $\bar{R}$ is the prevalence without BGS and $\bar{R}_B$ the prevalence with BGS (i.e. equation (17) in the main text). Berg et al. (2025) also showed that if large effect sites make a substantial contribution to genetic variance in liability, the liability distribution may become skewed. However, due to the polygenic background, the far tail of the distribution will still be roughly Gaussian in shape, and it is the rapid Gaussian decay of the tail that is responsible for the fat that the change in threshold density roughly predicts the change in prevalence (because most of the density in the tail that is beyond the threshold is nonetheless bery close to it), and so we expect equation (C.7) to hold, at least approximately, even if there is some skew in the distribution.

Negligible impact of frequency spectrum skew We also used simulations to explore whether this skew in the frequency spectrum impacts the disease prevalence. To do this, we varied the selection coefficients at the background selection sites while holding the mutation and recombination rates (and therefore the local effective size reduction) constant. We found that while the reduction in heterozygosity remained constant with increasing selection coefficients, the mean derived allele frequency among rare variants ($AF < 0.1$) rose, as expected. However, this increase in rare allele frequency had no discernible impact on the prevalence (Figure S12).

D. Solving the Two Effect-size Model with Background Selection

Here, we detail the method for solving the two-effect-size liability threshold model in the presence of background selection. Berg et al. (2025) describes the method for solving this model in the absence of background selection in their Supplementary Section S7. Here, for completeness, we largely reproduced that description, with the addition of the BGS effect (as well as a few minor differences in notation).

In our two-effect model, we assume that a fraction $p_S = 1 - p_L$ of sites have small effects, while the remaining fraction p_L have large effects. We model the distribution of liability in the population as a convolution of two distributions. The first is a Normal component with variance $V_{G,B,S} + V_E$, where $V_{G,B,S}$ is the variance from small effect sites (in the presence of background selection), and V_E is the environmental variance. The second is a Poisson distribution on the number of large effect alleles that an individual carries. This Poisson distribution has mean

$$\lambda_L = \frac{2Lp_L\mu}{\delta(a_L)C}, \tag{D.1}$$

where

$$\delta(a_L) = F(T^* - a_L) - F(T^*) \tag{D.2}$$

is the risk effect of the large effect sites, $T^* = T - \bar{Z}$ is the distance between the mean and the threshold,

$$F(Z) = \sum_i P(i|\lambda_L) Q(Z|a_L(i - \lambda_L), V_{G,B,S} + V_E) \tag{D.3}$$

is the probability that an individual's liability exceeds a value of Z, where $Q(Z|u, \sigma^2) = 1 - \Phi(Z|u, \sigma^2)$ is the complementary CDF of a Normal distribution with mean u and variance σ^2, and $P(i|\lambda_L)$ is the probability that a Poisson random variable with mean λ_L takes a value of i. The density on total liability, in turn, is

$$f(Z) = \sum_i P(i|\lambda_L) \phi(Z|a_L(i - \lambda_L), V_{G,B,S} + V_E) \tag{D.4}$$

where $\phi(Z|u, \sigma^2)$ is the Normal PDF.

We solve two-effect model in the following way. First, we can solve for the fraction of sites fixed for the liability increasing allele at small effect sites. Because all large effect sites are fixed for the liability decreasing allele, this fraction is equal to

$$p^+(\gamma(a_S)) = \frac{\bar{a}p_T}{a_S p_S} \tag{D.5}$$

where

$$\bar{a} = a_S p_S + a_L p_L \tag{D.6}$$

is the mean effect size. The scaled selection coefficient of the small effect sites is determined entirely by the threshold induced fixation asymmetry and the resulting long-term fixation dynamics, and is therefore given by

$$\gamma(a_S) = \frac{1}{2} \ln \frac{1 - p^+(\gamma(a_S))}{p^+(\gamma(a_S))}, \tag{D.7}$$

and the threshold density by

$$f_B(T) = \frac{\gamma(a_S)}{2NBCa_S} = \frac{1}{4NBCa_S} \ln \frac{1 - p^+(\gamma(a_S))}{p^+(\gamma(a_S))}. \tag{D.8}$$

The genetic variance due to small effect sites is

$$V_{G,B,S} = \frac{\theta_L p_S B a_S b(\gamma(a_S))}{\gamma(a_S)}, \tag{D.9}$$

where $b(\gamma(a_S)) = 1 - 2p^+(\gamma(a_S))$ is the degree of mutational asymmetry at small effect sites, $\theta_L = 4NL\mu$ is the total population scaled mutation rate of the trait.

We can then solve for the value of $\delta(a_L)$ via a line search. First, we plug T^* into the right hand side of Equation (D.4) and set it equal to the right hand side of Equation (D.8), yielding

$$\frac{1}{4NBCa_S} \ln \frac{1 - p^+(\gamma(a_S))}{p^+(\gamma(a_S))} = \sum_i P(i|\lambda_L)\phi(T^*|a_L(i - \lambda_L), V_{G,B,S} + V_E). \tag{D.10}$$

Then, for a proposed value of $\delta(a_L)$, we compute λ_L using Equation (D.1), plug it into the RHS of Equation (D.10), and numerically solve for the value of T^* that satisfies Equation (D.10). We can then conduct a line search for the value of $\delta(a_L)$ that satisfies Equation (D.2), given the value of T^* obtained from solving Equation (D.10).

To constrain the liability-scaled heritability to a specific value, we add one additional step. Given the proposed value of $\delta(a_L)$, we compute the large effect contribution to the genetic variance as

$$V_{G,B,L} = a_L^2 \lambda_L \tag{D.11}$$

and then compute the environmental variance as

$$V_E = \frac{1 - h^2}{h^2}. \tag{D.12}$$

Note that the inclusion of the B value in Equations (D.8) and (D.9) to account for the impact of BGS on weakly selected sites represents the only substantial difference between the algorithm presented in Berg et al. (2025) and the one presented here. The solution in the case of no BGS is therefore obtained simply by setting $B = 1$. To compare cases with and without background selection, we first solve the above system of equations for a fixed value of h^2 with $B = 1$, to obtain the solution in the absence of background selection. We then solve in the presence of background selection, holding the environmental variance (as opposed to the heritability), constant at the value we obtained in the no BGS case.

E. An Incidental Finding That the Bulmer Effect Increases the Genetic Variance in the Long Term

In our multi-locus stabilizing selection simulations, when we used the same recombination rate between causal loci as used in the simulation of the liability threshold model (5×10^{-6}), the genetic diversity reduction of causal sites under background selection initially did not match our expectations based on our theory and our single-locus simulations (Figure S11). We hypothesized that this might be related to the Bulmer effect (Bulmer 1971; Bulmer 1976), i.e. the accumulation of negative linkage disequilibrium between causal sites due to selection. More precisely, the genetic variance can be composed into two components:

$$Var\left(\sum_j a_j g_j\right) = \sum_j a_j^2 Var(g_j) + \sum_j \sum_{k \neq j} a_j a_k Cov(g_j, g_k), \tag{E.1}$$

where the first term is the additive genic variance and the second term captures the contribution of linkage disequilibrium among causal sites. Under neutrality, the expectation of the LD term is equal to zero. However, because stabilizing selection acts most strongly against individuals carrying combinations of alleles that result in extreme phenotypes, it causes negative covariance among alleles of like sign, leading to a negative expectation for the contribution of the LD term. Thus, while stabilizing selection is acting, it leads to a reduction in the total genetic variance due to this negative LD. The reduction in genetic variance due to this negative LD is what is known as the "Bulmer effect". Because recombination will quickly eliminate this negative LD if stabilizing selection ceases, the Bulmer effect is generally understood as a "short term" phenomenon.

For any given focal allele, the negative LD induced by the Bulmer effect leads to an attenuation

of the allele's marginal correlation with the phenotype, relative to its causal effect (Bulmer 1971; Veller and Coop 2024). In the modern statistical genetics literature, this effect is sometimes known as "linkage masking" (Brown et al. 2016). Because the expected change in the frequency of an allele from a given generation to the next depends on its marginal correlation with fitness in that generation, rather than its causal effect on fitness, the attenuation of the marginal phenotypic correlation in turn leads to an attenuation of the selection coefficient, and this in turn leads to a reduction in the magnitude of the expected change in frequency due to selection. More precisely, whereas in the absence of any Bulmer effect, the expected change in the minor allele frequency is given by:

$$\mathbb{E}(\Delta x) \approx -t(a)x(1-x)(1-2x), \tag{E.2}$$

with $t(a) = \frac{a^2}{2\omega^2}$, in its presence we expect that $t(a) < \frac{a^2}{2\omega^2}$, where the magnitude of the reduction in $t(a)$ depends on the amount of negative LD. This slowdown in the pace of frequency change is studied in greater depth by Negm and Veller (2024).

This effect has an intriguing implication. That is, while the negative LD induced by stabilizing selection causes a short-term reduction in the genetic variance relative to the genic variance, over long timescales, we would expect it to increase in the genic variance, due to the weaker selection on individual variants. Whether the net effect is to increase or decrease the genetic variance then depends on which effect is larger.

We hypothesized that this phenomenon may be responsible for the difference that we initially observed between our predictions and our simulation results. To investigate this hypothesis, we performed simulations under the stabilizing selection model without background selection and varied the rate of recombination between the neighboring causal sites (Figure S11). When the recombination rate is low, we find that the genetic variance is reduced relative to the genic variance, but is nevertheless increased relative to what we observed with higher recombination rates (which match the theoretical predictions based on unlinked theory; Simons et al. (2018)). Thus, our simulation results suggest that over long time scales, the net impact of the Bulmer effect is to increase the genetic variance, as the increase in the genic variance caused by the slowdown in the pace of allele frequency change is larger than the reduction in variance due to the negative LD. Moreover, when we increased the recombination rate in our study of background selection to a level at which there was no longer any Bulmer effect or inflation of the genic variance, we found that our theoretical prediction for the impact of background selection then matched our theoretical predictions (Figure 4A), indicating that the Bulmer effect was indeed the culprit.

Our primary focus was on understanding the impact of background selection, so we did not study this effect any further. However, it is worth noting that, in addition to Negm and Veller (2024), at least two prior publications have also studied this effect. Specifically, Lande (1975) found that in his Gaussian allelic model, the decrease in genetic variance due to the Bulmer effect is exactly offset by the increase in genic variance, leading to no net effect. In contrast, Turelli and Barton (1990), employing

their "rare alleles" approximation, do predict an increase in the genetic variance when recombination rates are low, but their predictions do not match our simulation results. A more complete analysis of this phenomenon would be valuable.

F. Variation in a and B in the Threshold Model

In this section, we provide more explicit mathematical support for the arguments in the main text for the cases with variation in local B value. Here, we start from a slightly more general position than in the main text, by assuming that the effects and the local N_e reductions have some joint distribution $g(a, B)$. Then analogous to equation (18) in the main text, equilibrium is established by evolving the value of $f(T)$ that solves

$$p_T = \int\int g(a, B)\frac{a}{\bar{a}}p^+\left(2NC \cdot f(T) \cdot aB\right) \mathrm{d}a\mathrm{d}B. \tag{F.1}$$

Effectively neutral regime

When $p_T \approx \frac{1}{2}$, the fixation asymmetry is approximately linear, $p^+(\gamma) \approx \frac{1}{2}(1-\gamma)$. Inserting this approximation and evaluating the resulting integral, we have

$$p_T \approx \frac{1}{2} - \frac{NCf_B(T)}{\bar{a}}\mathbb{E}[a^2B], \tag{F.2}$$

which we can solve for

$$f_B(T) \approx \left(\frac{\frac{1}{2} - p_T}{NC}\right)\left(\frac{\bar{a}}{\mathbb{E}[a^2B]}\right). \tag{F.3}$$

In the absence of background selection, $B = 1$, so

$$f(T) \approx \left(\frac{\frac{1}{2} - p_T}{NC}\right)\left(\frac{\bar{a}}{\mathbb{E}[a^2]}\right). \tag{F.4}$$

The global compensation factor is

$$\frac{f_B(T)}{f(T)} \approx \frac{\mathbb{E}[a^2]}{\mathbb{E}[a^2B]}, \tag{F.5}$$

and is thus an average of B that is weighted by the squared effect sizes, reflecting the fact that variance contributions scale with a^2 in the neutral regime. If a and B are independent, then

$$\frac{f_B(T)}{f(T)} \approx \frac{\mathbb{E}[a^2]}{\mathbb{E}[a^2]\mathbb{E}[B]} = \frac{1}{\mathbb{E}[B]}. \tag{F.6}$$

Strong selection limit

In the strong selection limit, the behavior depends on whether the coefficient of variation of the effect distribution is large or small. Let us first consider the small coefficient of variation case.

$g_a(a)$ has a small coefficient of variation

In this case, the distribution of scaled selection coefficients γ is centered over a relatively narrow range. Most sites will have similar scaled selection coefficients, and the bulk of their distribution moves from the weak into the strong selection regime as $p_T \to 0$. In this regime, $p^+(\gamma) = (1 + e^{2\gamma})^{-1} \approx e^{-2\gamma}$, so equation (F.1) becomes

$$p_T = \frac{1}{\bar{a}} \mathbb{E}\left[a e^{-4NCf(T)\cdot aB}\right]. \tag{F.7}$$

The exponential decay dominates over the linear factor of a, and so the primary contribution to the expectation is from sites with the minimum value of aB, i.e. the sites with the smallest effect sizes experiencing the strongest background selection. This implies that

$$p_T \approx K_1 e^{-4NCf(T)\cdot(aB)_{min}} \tag{F.8}$$

where K_1 is a constant, and $(aB)_{min}$ is the minimum of the product of aB across sites. Solving for $f(T)$, we have

$$f_B(T) \approx \frac{\ln\left(\frac{p_T}{K_1}\right)}{4NC} \frac{1}{(aB)_{min}} \tag{F.9}$$

and

$$f(T) \approx \frac{\ln\left(\frac{p_T}{K_1}\right)}{4NC} \frac{1}{a_{min}}, \tag{F.10}$$

so that the global compensation factor is

$$\frac{f_B(T)}{f(T)} \approx \frac{a_{min}}{(aB)_{min}}. \tag{F.11}$$

If a and B are independent, then $(aB)_{min} = a_{min} B_{min}$, and so

$$\frac{f_B(T)}{f(T)} \approx \frac{1}{B_{min}}. \tag{F.12}$$

$g_a(a)$ has a large CV

Alternatively, if there is high variance in the effect distribution, then the scaled selection coefficients of sites will not all cluster together in the same selection regime, but rather will be spread across all of the selection regimes, with a large subset of small effect sites belonging to the effectively neutral regime, even as $p_T \to 0$. As a result, we must consider contributions from the whole distribution, rather than just a narrow slice of it. Here, it is not clear *a priori* how to obtain any general results for an arbitrary joint distribution, $g(a, B)$, so we will assume they are independent (i.e., $g(a, B) = g_a(a)g_B(B)$). As a tractable choice for a distribution which can model the high variance scenario of interest, suppose that the effects follow a gamma distribution with shape and scale parameters k and ω, i.e. $g_a(a) = \frac{1}{\Gamma(k)\omega^k} a^{k-1} e^{-\frac{a}{\omega}}$. The coefficient of variation of the gamma distribution is equal to $k^{-1/2}$, so we will be interested in the $k \to 0$ limit, though we proceed generally for now, and consider this limit below.

Although the entire distribution contributes to the integral, the exponential suppression of the transition into the strongly selected regime still dominates, so we can still make the same $p^+(\gamma) \approx e^{-2\gamma}$ approximation that we made in the low CV case. Equation (F.1) becomes

$$p_T \bar{a} = \frac{1}{\Gamma(k)\,\omega^k} \int_B g_B(B) \left[\int_a a^k e^{-4NCf_B(T)aB - \frac{a}{\omega}} da \right] dB. \tag{F.13}$$

Because we are interested in the strong selection limit (i.e. $p_T \to 0$), we assume that the threshold density $f_B(T)$ is large enough that $4NCf_B(T)\,B \gg \frac{1}{\omega}$. In physical terms, this inequality implies that the fixation asymmetry $p^+(\gamma)$ decays with increasing effect size a much more rapidly than the mutational input $g_a(a)$ does. That is, selection effectively prevents the fixation of deleterious alleles with moderate effect sizes well before such alleles become rare in the mutational distribution. Consequently, the integral is dominated by the interaction between selection and the power-law portion of the effect size distribution, allowing us to ignore the exponential suppression of the tail of the gamma distribution. Mathematically, this allows us to approximate the exponential function in equation (F.13): $e^{-4NCf_B(T)aB - \frac{a}{\omega}} \approx e^{-4NCf_B(T)aB}$, so that we can write

$$p_T \bar{a} = \frac{1}{\Gamma(k)\,\omega^k} \int_B g_B(B) \left[\int_a a^k e^{-4NCf_B(T)aB} da \right] dB. \tag{F.14}$$

We can then use u substitution (set $u = 4NCf_B(T)\,aB$ so that $a = \frac{u}{4NCf_B(T)B}$ and $da = \frac{du}{4NCf_B(T)B}$) to rewrite the inner integral as

$$\int_a a^k e^{-4NCf(T)aB} da \approx \left[\frac{1}{(4NCf_B(T)\,B)^{k+1}} \right] \int_0^\infty u^k e^{-u} du, \tag{F.15}$$

where the integral in u is the gamma function, i.e. $\int_0^\infty u^k e^{-u} du = \Gamma(k+1)$. We can therefore write

equation (F.14) as

$$p_T \bar{a} \approx \frac{k}{\omega^k [4NCf_B(T)]^{k+1}} \int_B g_B(B) \left[\frac{1}{B^{k+1}} \right] dB, \tag{F.16}$$

which we can solve to find

$$f_B(T) \approx \left[\frac{k^{\frac{1}{k+1}}}{4NC(\omega^k p_T \bar{a})^{\frac{1}{k+1}}} \right] \mathbb{E}\left[B^{-(k+1)}\right]^{\frac{1}{k+1}} \tag{F.17}$$

implying that

$$\frac{f_B(T)}{f(T)} \approx \mathbb{E}\left[B^{-(k+1)}\right]^{\frac{1}{k+1}}. \tag{F.18}$$

Now, let us consider the large coefficient of variation limit. As $k \to 0$, the gamma distribution follows a power law that decays at rate a^{-1} until cutoff by the exponential suppress term $e^{-\frac{a}{\omega}}$. In this case, the effect size distribution is very broad, and $\frac{f_B(T)}{f(T)}$ is equal to the expectation of B^{-1}:

$$\frac{f_B(T)}{f(T)} \approx \mathbb{E}\left[B^{-1}\right]. \tag{F.19}$$

Alternatively, as $k \to \infty$, the variance of the effect size distribution shrinks, and $\frac{f_B(T)}{f(T)}$ converges toward the minimum,

$$\frac{f_B(T)}{f(T)} \approx \frac{1}{B_{min}} \tag{F.20}$$

consistent with our previous results.

Finally, we must address the validity of the continuous approximation in the high-CV limit. In our continuous mathematical model, the Gamma distribution provides an effectively infinite number of sites with vanishingly small effect sizes. Consequently, the compensation process never terminates because there are always sites small enough to remain in the effectively neutral regime. In a physical genome with a finite number of sites L, there is necessarily a single site with the minimum effect size, a_{min}. Technically, the compensation process described above would terminate once $f(T)$ becomes large enough that even this smallest effect is pushed into the strong selection regime. However, this termination occurs only at the limit where selection is so strong that carrying a single allele at any site with any effect size is sufficient to cross the threshold (i.e., $s \approx C$). Thus, the limit where this global compensation ultimately stops acting is precisely the limit where the epistasis becomes unimportant, and the model collapses onto one where all loci impact fitness independently in a Mendelian fashion. Thus, by invoking a model with threshold epistasis and wide variation in effect sizes in the context of a finite genome, we are effectively invoking the $p_T \to 0$ limit, but assuming that p_T nonetheless

remains large enough that $a_{min} \ll p_T$. As shown in Figure 3B, explicit numerical solutions of equation (F.1) exhibit convergence to the $f_B(T)/f(T) \approx \mathbb{E}\left[B^{-1}\right]$ limit at relatively large values of p_T once the CV becomes moderately large, so this assumption is entirely appropriate.

Impact on the genetic variance

Having determined the global compensation factor, we can now derive the impact of background selection on the genetic variance contributed by specific sites. Let $B_{global} = \frac{f(T)}{f_B(T)}$ be the effective genome-wide reduction in the threshold density (note that this is the inverse of the compensation ratio derived above, e.g., $B_{global} = \frac{1}{\mathbb{E}[B^{-1}]}$ in the high-CV gamma limit).

Consider a specific site i with effect size a and a local effective population size reduction B_{local}. In the absence of background selection, the site has a scaled selection coefficient $\gamma = 2Naf(T)C$. In the presence of background selection, the unscaled selection coefficient increases due to the global shift in threshold density:

$$s_B \approx af_B(T)C = af(T)\frac{1}{B_{global}}C = \frac{s}{B_{global}}. \tag{F.21}$$

However, the local effective population size decreases: $N \to NB_{local}$. The new scaled selection coefficient is therefore:

$$\gamma_B = 2(NB_{local})s_B = 2NB_{local}\frac{s}{B_{global}} = \gamma\frac{B_{local}}{B_{global}}. \tag{F.22}$$

This ratio $\frac{B_{local}}{B_{global}}$ determines whether the site experiences a net increase or decrease in the efficiency of selection relative to drift.

The genetic variance contributed by this site is proportional to the mutation rate and the function $b(\gamma)/\gamma$, where $b(\gamma) = \tanh(\gamma)$. Accounting for the reduction in the effective mutation rate ($\theta \to \theta B_{local}$), the variance in the presence of background selection is:

$$\begin{aligned} V_{G,B} &= \frac{\theta B_{local}}{\gamma_B} b(\gamma_B) \\ &= \frac{\theta B_{local}}{\gamma \frac{B_{local}}{B_{global}}} b\left(\gamma \frac{B_{local}}{B_{global}}\right) \\ &= B_{global}\frac{\theta}{\gamma} b\left(\gamma \frac{B_{local}}{B_{global}}\right). \end{aligned} \tag{F.23}$$

Comparing this to the original variance $V_G = \frac{\theta}{\gamma}b(\gamma)$, we obtain the reduction factor:

$$\frac{V_{G,B}}{V_G} = B_{global}\frac{b\left(\gamma \frac{B_{local}}{B_{global}}\right)}{b(\gamma)}, \tag{F.24}$$

i.e. equation (22) in the main text.

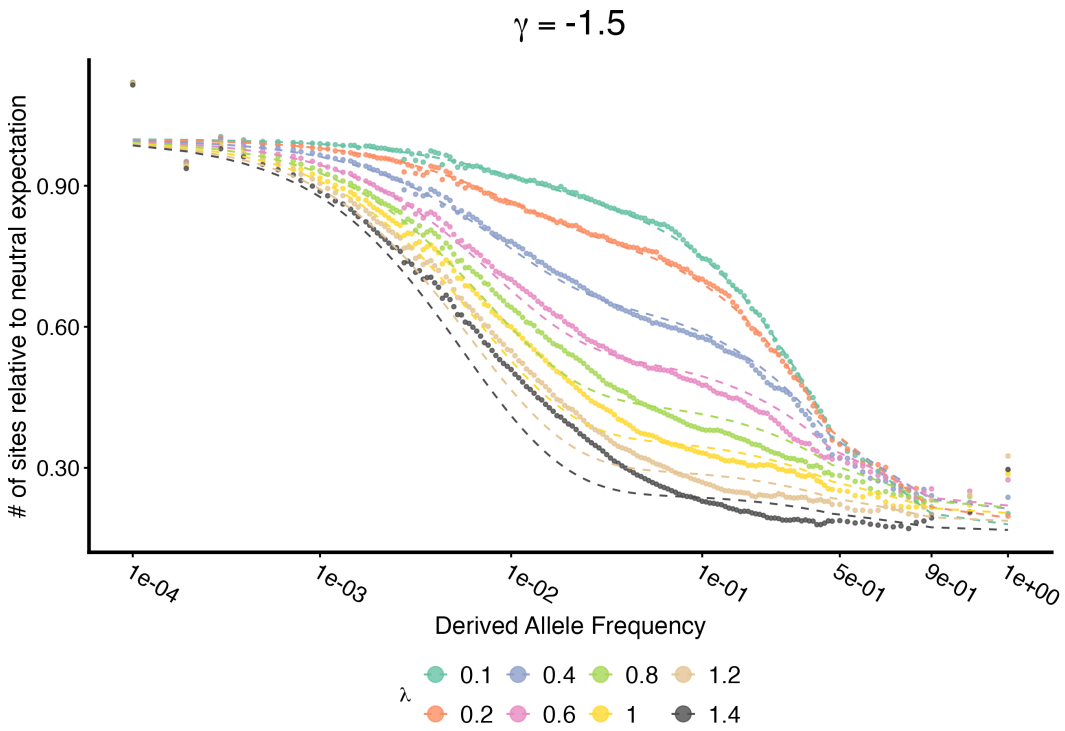


Figure S1: **Robustness of the frequency-dependent BGS approximation across varying mutation intensities (λ).** The site frequency spectrum (SFS) of a focal site with $\gamma = -1.5$, scaled relative to the standard neutral expectation, is shown for simulations with varying values of the background mutation flux $\lambda = 2\nu/r$. Points represent simulation results, while solid lines show the analytical prediction using the $B(q)$ approximation derived in the weak mutation limit (Equation (A.27)). Although the analytical model is formally derived for the limit where mutation is weak relative to recombination ($\lambda \ll 1$), it accurately captures the distortion of the SFS for $\lambda < 1/2$. Departures from the prediction for $\lambda > 1/2$ indicate the onset of the "fitness drag" regime driven by the accumulation of new mutations, which we describe in Supplementary Text A.

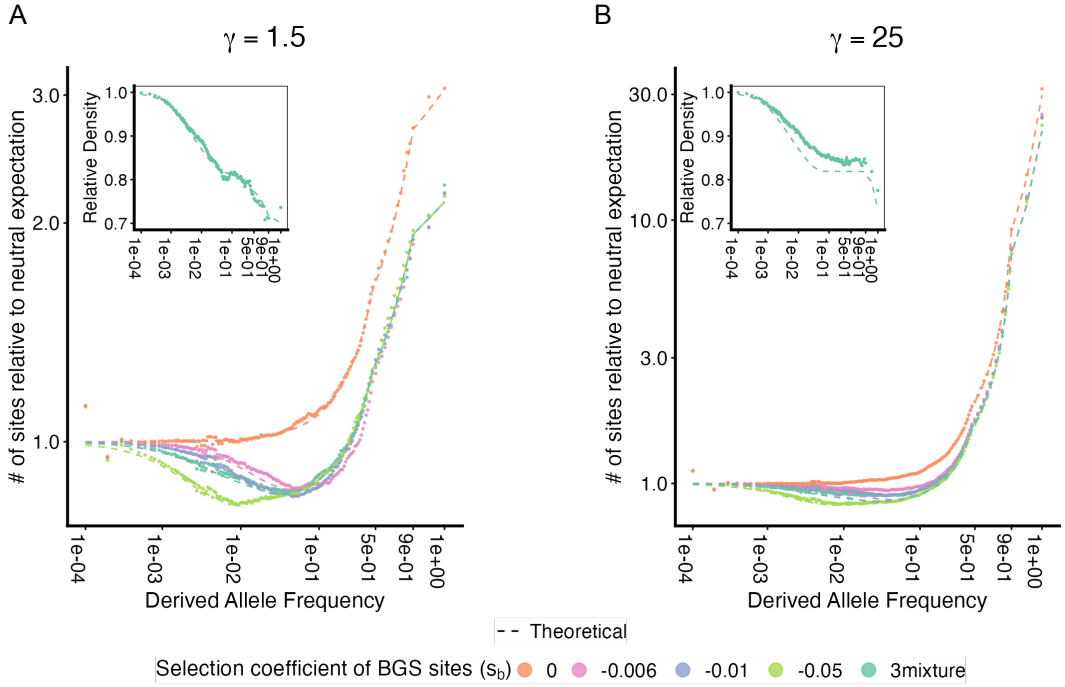


Figure S2: The site frequency spectrum (SFS) for positively selected focal alleles, scaled to the neutral allele expectation, is shown for a BGS intensity of $B \approx 0.82$. Points represent simulation results, while dashed lines indicate theoretical predictions based on the frequency-dependent $B(q)$ approximation (Equation A.32). We show results for three different selection coefficients of background selection sites ($s_b = -0.006, -0.01, -0.05$), as well as an equal mixture of all three. The insets show the relative density of alleles across frequency bins, comparing scenarios with BGS to those without. Focal selected alleles are under (A) weak positive selection ($\gamma = 1.5$) and (B) intermediate positive selection ($\gamma = 25$). When $\gamma = 25$, simulations with weaker BGS ($|s_b| < 0.05$) exhibit a deviate from the theoretical line. This occurs because when the strength of direct selection on the focal allele is similar to the strength of background selection ($|\gamma| \approx |2Ns_b|$), the two processes act on similar timescales. This transits into the regime where the focal allele is dominated by its own fitness effect before the background is purged (Equation A.35).

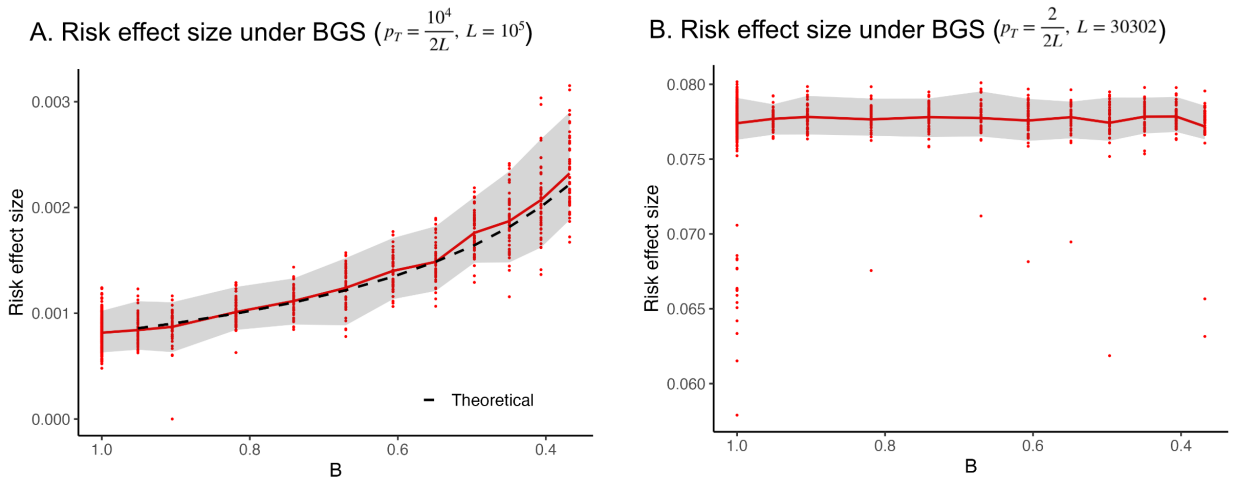


Figure S3: The impact of BGS on risk effect size in the single-effect liability threshold model. (A) When $p_T \gg \frac{1}{2L}$, the risk effect size increases with increasing strength of BGS. The black dashed line shows the theoretical prediction of $f(T) \to f(T)/B$. (B) When $p_T \sim \frac{1}{2L}$, the risk effect size does not change under BGS.

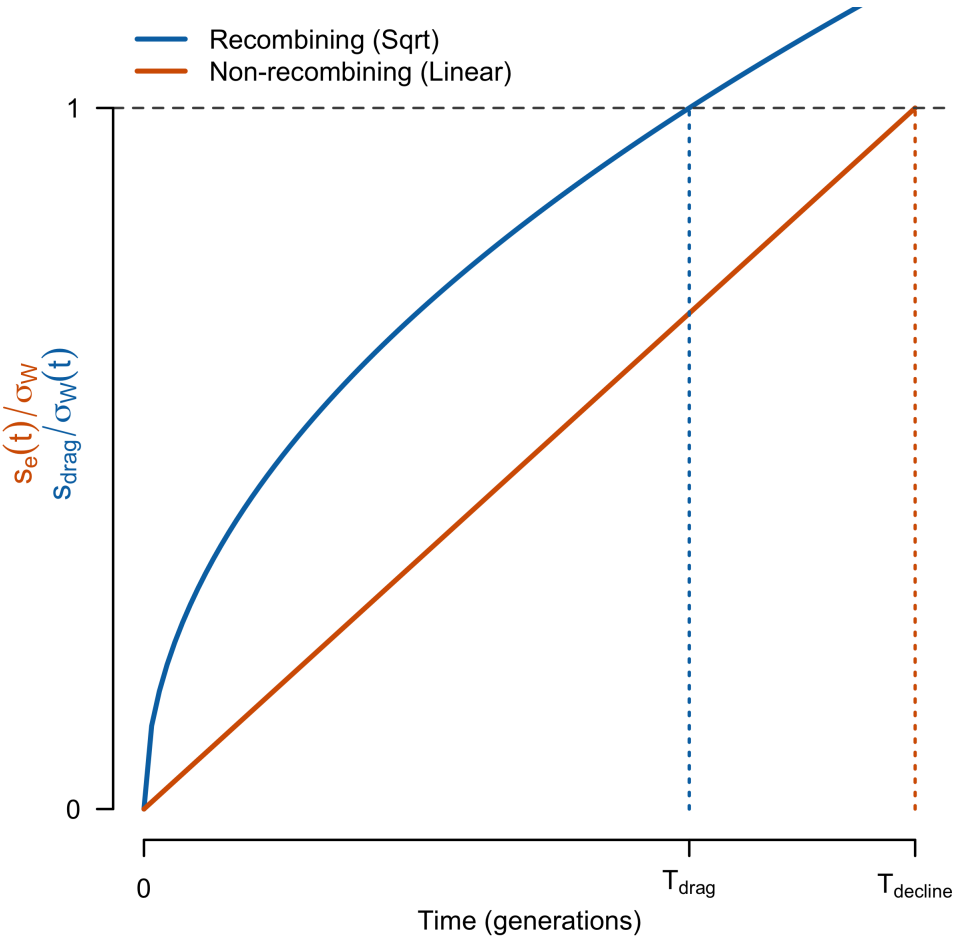


Figure S4: **Dynamics of the onset of effective selection in recombining vs. non-recombining models.** This figure compares the time evolution of the ratio of the effective selection coefficient (s_{eff}) to the relevant standard deviation of fitness (σ_W) for a newly arisen allele. In the non-recombining model (orange), the fitness deficit accumulates linearly ($s_{eff}(t) \propto t$) against a constant background variation, leading to a linear increase in the ratio. The threshold for effective selection ($T_{decline}$) is reached when this ratio equals 1. In the recombining model (blue), the fitness deficit is constant ($s_{drag} = \lambda s_b$), but the relevant background variation shrinks as the associated haplotype block shortens ($\sigma_W(t) \propto t^{-1/2}$), causing the ratio to grow with $\sqrt{t}$. This dynamic leads to a more rapid onset of effective selection, both because $T_{drag} < T_{decline}$ and because the functional form of the ratio results in faster growth of the relative fitness deficit in early generations. Together, these factors imply that the frequency spectrum is distorted at significantly lower frequencies in recombining genomes. For the example plotted here, $\lambda = 2$.

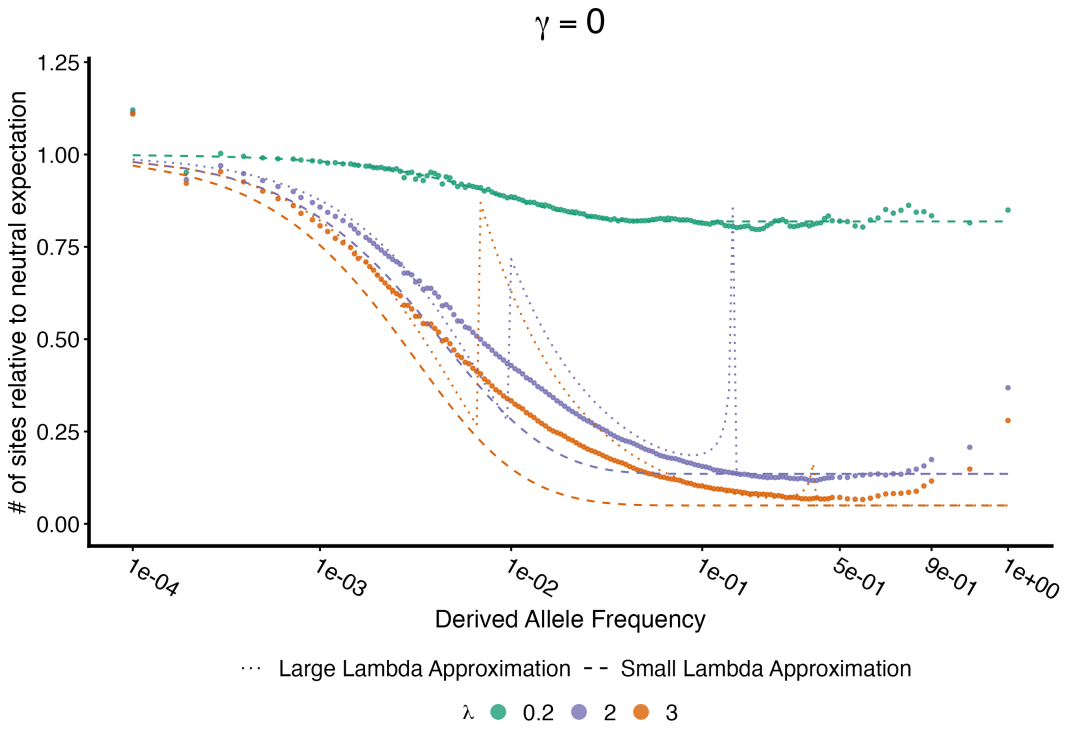


Figure S5: The site frequency spectrum (SFS) of a neutral focal site ($\gamma = 0$) is shown relative to the standard neutral expectation for three different background selection intensities. Points represent simulation results for various levels of background mutation flux ($\lambda = 2\nu/r$), ranging from weak ($\lambda = 0.2$) to strong ($\lambda = 2, 3$). For low mutation flux ($\lambda = 0.2$), the distortion of the SFS is well-captured by the weak mutation/strong recombination (small lambda) approximation with Equation A.24 (dashed lines), which accounts for selection acting primarily against deleterious alleles already present on the initial background. As the mutation flux increases ($\lambda > 1/2$), "fitness drag" regime kicks in and the dynamics are instead better described by the strong mutation/weak recombination (large lambda) approximation with (Equation A.47)(dotted lines), which accounts for the continuous accumulation of fitness drag on the shortening associated haplotype block. Notably, the spikes in the large lambda approximation represent logarithmic divergences near the transitions between frequency regimes in Equation A.47, where Cvijović et al. (2018) applied a smoothing technique to resolve these boundaries.

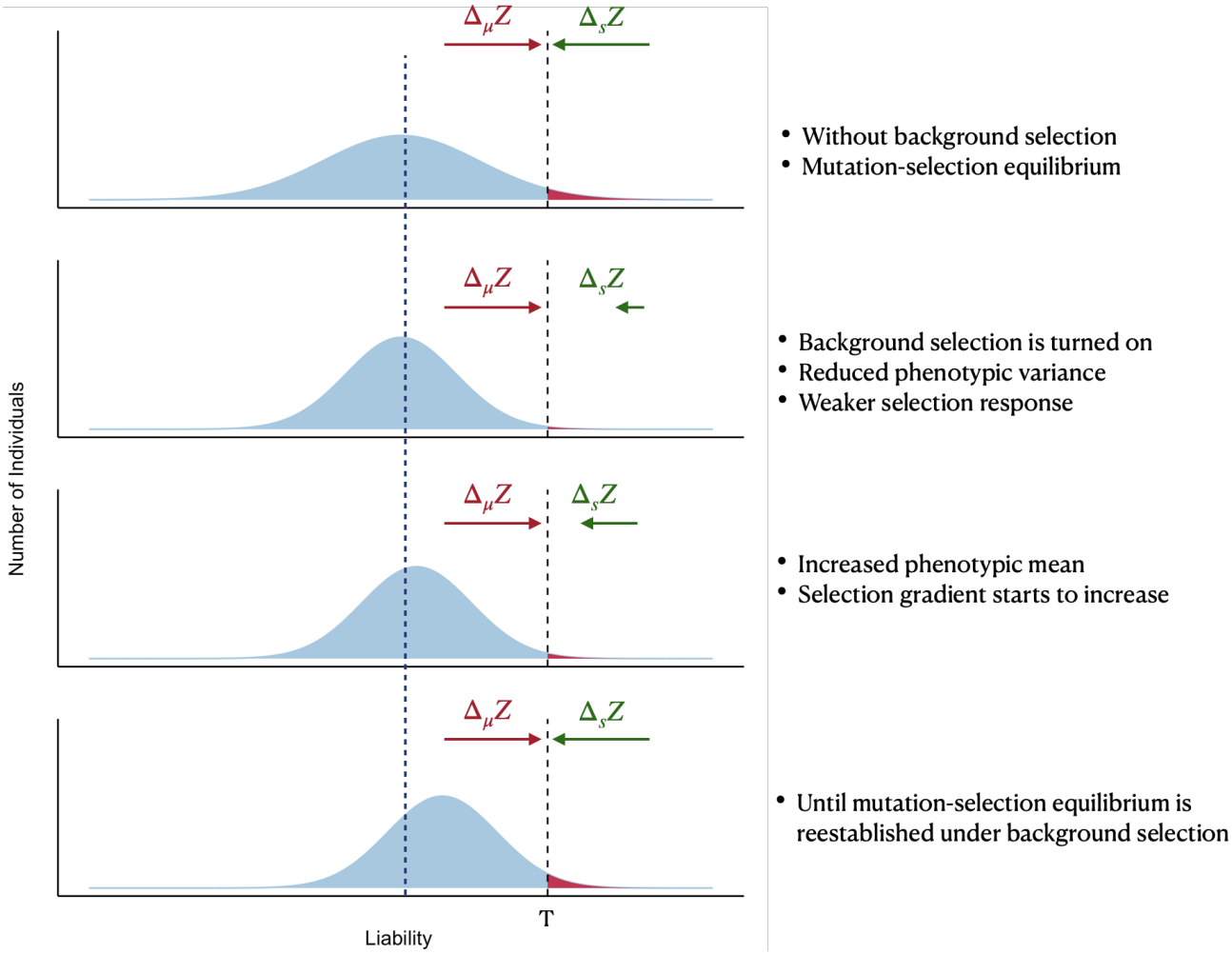


Figure S6: A cartoon illustration that explains how background selection impacts disease prevalence. This illustration imagines a dynamic process where the population is first at equilibrium in the absence of background selection, before it is "turned on", and the population evolves to the new equilibrium. In reality, top and bottom panels should be viewed as counterfactual comparison in a model without vs with background selection.

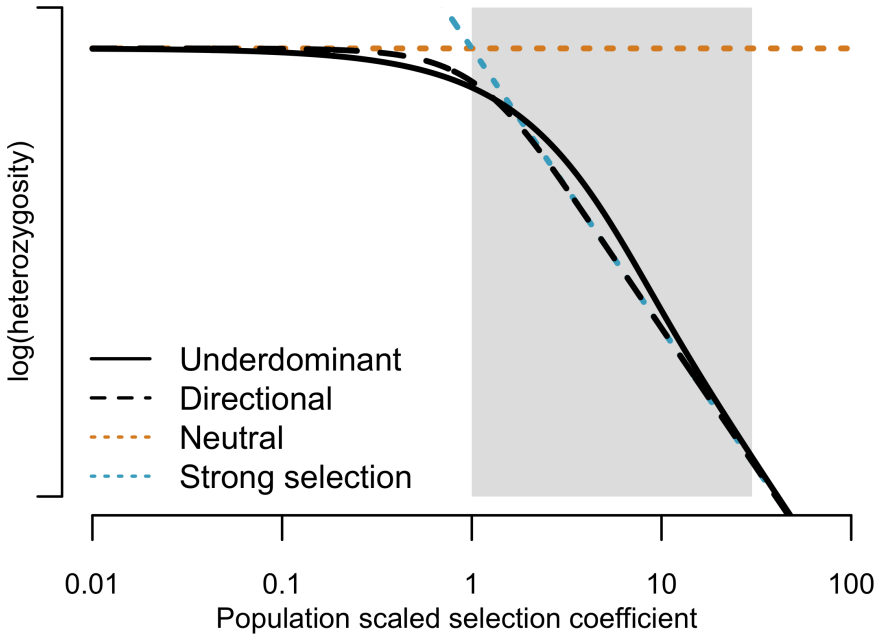


Figure S7: A comparison of expected heterozygosity for alleles with directional selection (Equation (6), dashed black line) vs. under dominant selection (Equation (23), solid black line) at different scaled selection coefficients. The expected neutral heterozygosity is θ, and the expected heterozygosity for strong selection is $\frac{\theta}{\gamma}$.

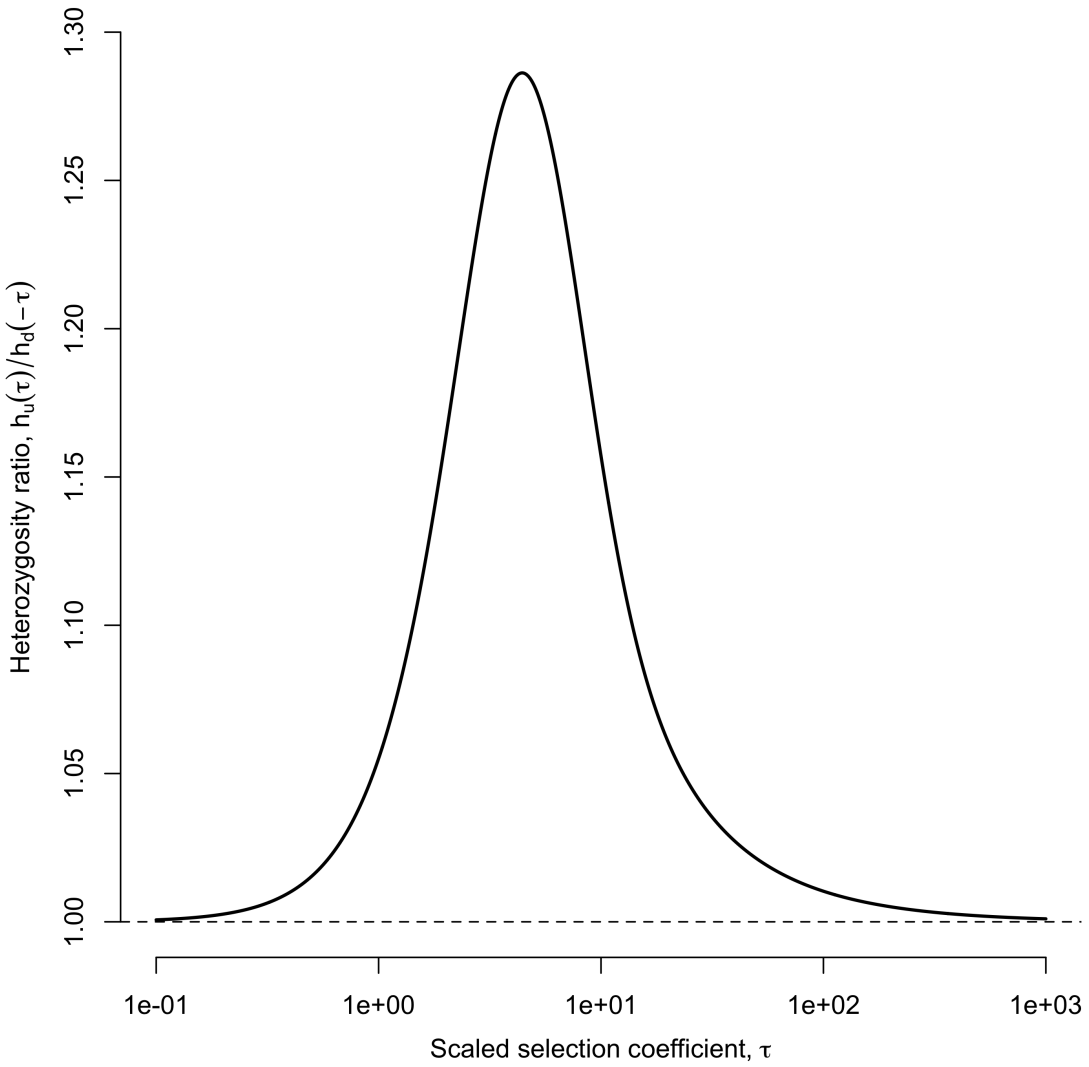


Figure S8: The ratio $h_u(\tau)/h_d(-\tau)$, measuring the increase in heterozygosity due to the underdominance-induced slow down in allele frequency change, as a function of τ.

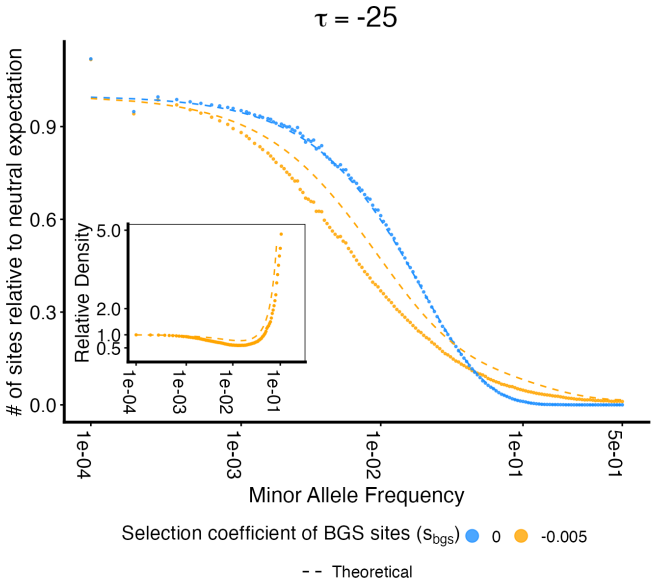


Figure S9: Site frequency spectrum (upper panel) and relative density (lower panel) of a selected allele with under-dominant selection $2Nt = -25$, with a background selection intensity $B \approx 0.2$.The theoretical results (dashed line) obtained with frequency dependent B approximation using equation A.26 breaks down in this large mutation/small recombination regime and does not predict the simulation well.

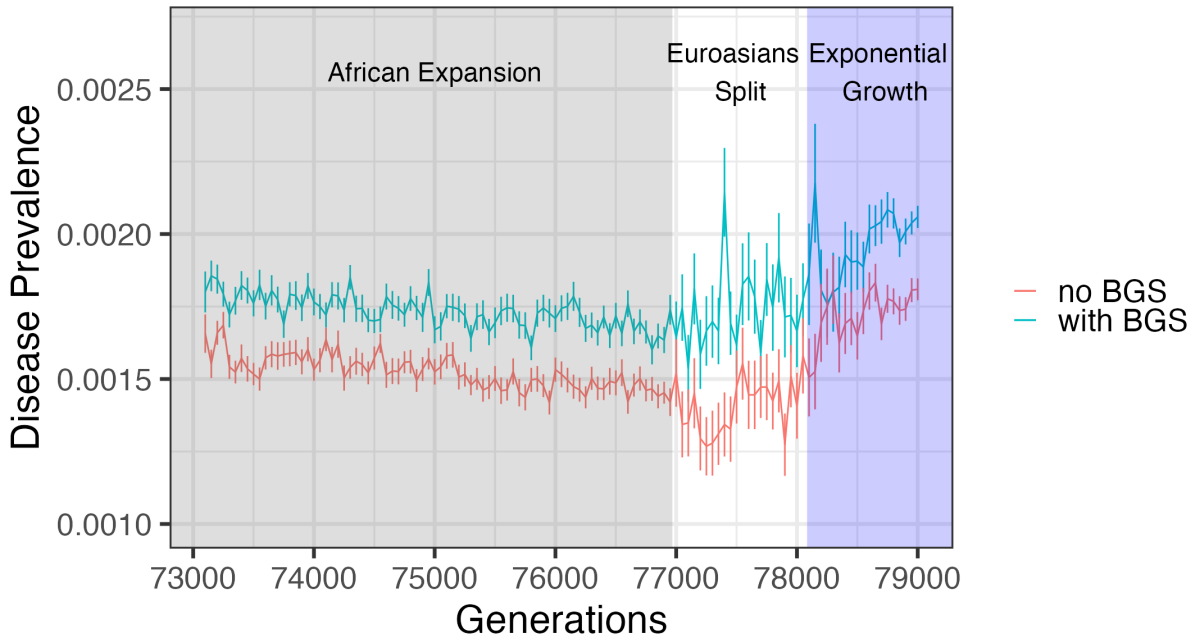


Figure S10: We simulated a "CEU-like" demographic history with the out-of-Africa demographic history (Gravel et al. 2011) with the strength of background selection set to reduce neutral variance by an average 17%, as estimated in humans (Murphy et al. 2022). We simulated a disease under the liability threshold model with heritability $h^2 = 0.5$, and a mutational asymmetry $p^+(a) = 0.05$. We reported the disease prevalence of the African population before the Euroasians split and of the European population after the split. We compared the disease prevalence with no background selection (red) and with background selection (green) intensity $B \approx 0.83$ (average estimated in humans). The error bar represents the standard error of the mean among 100 replications. In these simulations, the prevalence of a polygenic disease with heritability 0.5 and weakly selected causal loci in the presence day (generation 79000) in the absence of background selection was 0.18% (95% CI: 0.176% - 0.191%), and its presence 0.21% (95% CI: 0.203% - 0.219%), an average increase of 14.9%.

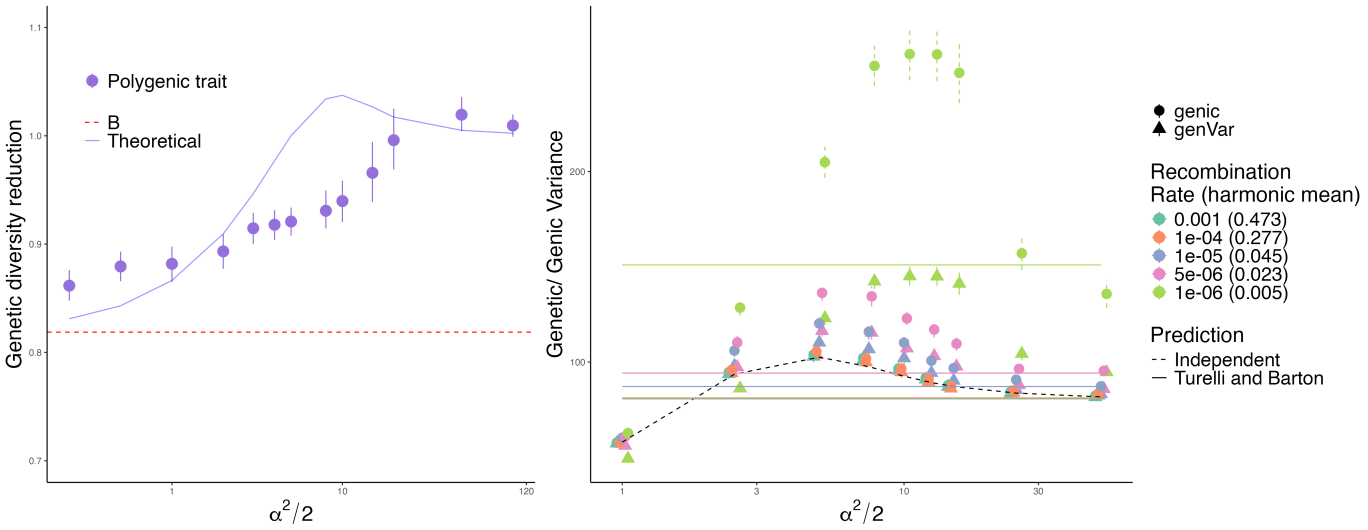


Figure S11: Left panel: we simulated polygenic traits evolving under stabilizing selection and set the recombination rate between causal loci at 5×10^{-6} and mutation rate at 2×10^{-8}. With these parameters choices, we found that the effect of BGS on the genetic variance of the trait was different from the single-site prediction. Right panel: we simulated traits under stabilizing selection without background selection and varied the recombination rates between causal loci. The mutation rate is fixed at 2×10^{-8}. We plot the genic variance (circle) and genetic variance (triangle) and compare them to the predicted genetic variance that assumes independent causal loci (dashed line), and to the "rare alleles" approximation from Turelli and Barton 1990 (solid lines). While the LD component of genetic variance is negative, as expected, due to the slow down mechanism described by Negm and Veller (2024) the genic variance is increased by more than the negative LD subtracts, so that the net effect on the genetic variance is to increase it.

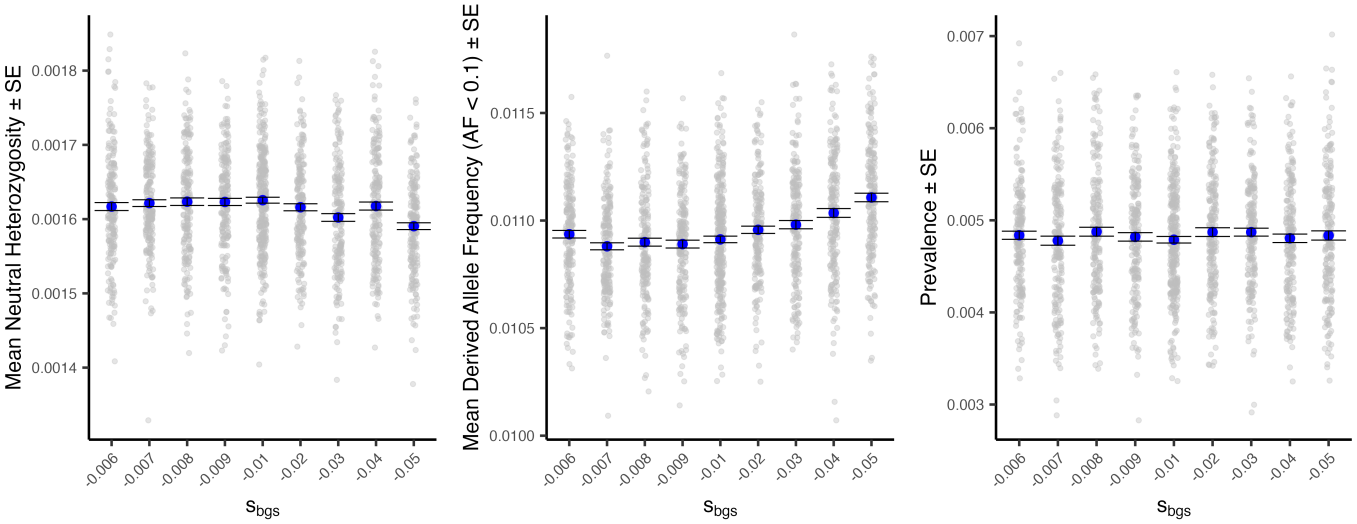


Figure S12: Simulations were designed to isolate the effect of background selection–induced distortion of the site-frequency spectrum (SFS), particularly the excess of rare variants, on disease prevalence under the liability-threshold model. Mutation rate ($\nu = 2 \times 10^{-8}$) and recombination rate ($r = 2 \times 10^{-7}$) at BGS sites were held fixed, while only the selection coefficient (s_b) was varied. The left panel shows that mean neutral heterozygosity is insensitive to s_b. The center panel shows that weakening selection at BGS sites leads to a decrease in the mean derived-allele frequency among rare variants. Despite this shift in the SFS, the disease prevalence is not minimally affected, as shown in the right panel.

Supplementary References

Bulmer, M (1971). "The effect of selection on genetic variability". In: *The American Naturalist* 105.943, pp. 201–211.

Lande, R (1975). "The maintenance of genetic variability by mutation in a polygenic character with linked loci". In: *Genetics Research* 26.3, pp. 221–235.

Bulmer, M (1976). "The effect of selection on genetic variability: a simulation study". In: *Genetics Research* 28.2, pp. 101–117.

Turelli, M and NH Barton (1990). "Dynamics of polygenic characters under selection". In: *Theoretical Population Biology* 38.1, pp. 1–57.

Charlesworth, B (1994). "The effect of background selection against deleterious mutations on weakly selected, linked variants". In: *Genetics Research* 63.3, pp. 213–227.

Gravel, S, BM Henn, RN Gutenkunst, AR Indap, GT Marth, AG Clark, F Yu, RA Gibbs, 1G Project, CD Bustamante, et al. (2011). "Demographic history and rare allele sharing among human populations". In: *Proceedings of the National Academy of Sciences* 108.29, pp. 11983–11988.

Brown, BC, AL Price, NA Patsopoulos, and N Zaitlen (July 2016). "Local Joint Testing Improves Power and Identifies Hidden Heritability in Association Studies". In: *Genetics* 203.3, pp. 1105–1116. ISSN: 1943-2631. DOI: `10.1534/genetics.116.188292`. URL: `https://doi.org/10.1534/genetics.116.188292` (visited on 05/28/2025).

Cvijović, I, BH Good, and MM Desai (2018). "The effect of strong purifying selection on genetic diversity". In: *Genetics* 209.4, pp. 1235–1278.

Simons, YB, K Bullaughey, RR Hudson, and G Sella (2018). "A population genetic interpretation of GWAS findings for human quantitative traits". In: *PLoS biology* 16.3, e2002985.

Murphy, DA, E Elyashiv, G Amster, and G Sella (2022). "Broad-scale variation in human genetic diversity levels is predicted by purifying selection on coding and non-coding elements". In: *Elife* 12, e76065.

Negm, S and C Veller (2024). "The effect of long-range linkage disequilibrium on allele-frequency dynamics under stabilizing selection". Preprint at bioRxiv `https://www.biorxiv.org/content/10.1101/2024.06.27.601075v1`.

Veller, C and GM Coop (2024). "Interpreting population-and family-based genome-wide association studies in the presence of confounding". In: *PLoS Biology* 22.4, e3002511.

Berg, JJ, X Li, K Riall, LK Hayward, and G Sella (Dec. 2025). "Mutation–selection–drift balance models of complex diseases". In: *Genetics* 231.4, iyaf220. ISSN: 1943-2631. DOI: `10.1093/genetics/iyaf220`. URL: `https://doi.org/10.1093/genetics/iyaf220` (visited on 12/12/2025).