Genetic load may increase or decrease with selfing depending upon the recombination environment

Shelley A. Sianta⋆, Stephan Peischl†, David A. Moeller⋆, Yaniv Brandvain⋆

⋆ Department of Plant and Microbial Biology; University of Minnesota; St. Paul, Minnesota, United States of America, 55108

† Interfaculty Bioinformatics Unit; University of Bern; Bern, Switzerland, 3012

Running title: Effect of selfing rate on genetic load

Key words: Background selection, Efficacy of selection, Genetic load, Mating system, Pseudo-overdominance, Purging, Selfing

Corresponding author: Shelley Sianta, Department of Plant and Microbial Biology; 1445 Gortner Avenue; St. Paul, Minnesota, USA 55108; email: ssianta@umn.edu

ABSTRACT:

Theory predicts that the ability for natural selection to remove deleterious mutations from a population, and prevent the accumulation of genetic load, is a function of the effective population size (N_e). Shifts from random mating to self-fertilization ("selfing") are predicted to decrease N_e through a variety of genomic changes - including a reduction in effective recombination and an increase in homozygosity. While a long history of theory suggests that the efficacy of selection, particularly against non-recessive mutations, should decrease with selfing rate, comparisons of genomic-based estimates of the efficacy of selection between related outcrosser-selfer pairs have revealed conflicting results. We address this paradox by simulating the evolution of strongly deleterious recessive and weakly deleterious additive mutations across a range of recombination, mutation and selective parameter combinations. We find that the genetic load of a population can either increase, decrease, or not vary with selfing rate. Genetic load is higher in selfers only when recombination rates are greater than mutation rates. When recombination rates are lower than mutation rates, an accumulation of recessive mutations leads to pseudo-overdominance, a type of balancing selection, in outcrossing populations. Using both simulations and analytical theory, we show that pseudo-overdominance has strong negative effects on the efficacy of selection against linked additive mutations and that a threshold level of selfing prevents pseudo-overdominance. Our results show that selection can be more or less effective in selfers as compared to outcrossers depending on the relationship between the deleterious mutation rate and gene density, and therefore different genomic regions in different taxa could show differing results.

INTRODUCTION:

Populations are constantly battling an influx of deleterious mutations. The ability of natural selection to remove deleterious mutations from a population is a function of how well it can distinguish the fitness effect of a mutation among the genetic backgrounds that mutation resides in. Large sample sizes of independent genetic backgrounds (i.e. a large effective population size, N_e) allow natural selection to be more effective. The inability of natural selection to remove deleterious mutations results in genetic load, or a decrease in average population fitness, and this load can consequently affect population viability in small populations (Lynch *et al.* 1995), patterns of introgression (Sankararaman *et al.* 2014; Kim *et al.* 2018), and can further reduce selection efficacy (Charlesworth *et al.* 1993a; Charlesworth 1994).

Variation in the mating system – from random mating (“outcrossing”) to full self-fertilization (“selfing”) – is a model for the study of the efficacy of selection . It is currently understood that selfing has two differing effects on the removal of deleterious mutations. First, by exposing rare mutations as homozygotes, an increase in selfing increases the rate at which deleterious recessive mutations are “purged” from a population. Second, the consensus in the theoretical literature is that by decreasing the effective population size and lowering effective rates of recombination, selfing decreases the efficacy of selection against mildly deleterious additive mutations (Charlesworth and Wright 2001; Wright *et al.* 2008).

Despite the theoretical consensus, empirical evidence supporting the hypothesis that selfers carry an elevated additive genetic load is quite mixed (Haudry *et al.* 2008; Escobar *et al.* 2010; Slotte *et al.* 2010, 2013; Qiu *et al.* 2011; Ness *et al.* 2012; Hazzouri *et al.* 2013; Gioti *et al.* 2013; Brandvain *et al.* 2014). The ambivalent association between mating system and deleterious mutations, has been attributed to the recency of selfing lineages and/or to differences between focal taxa unrelated to selfing *per se* (Haudry *et al.* 2008; Glémin and Galtier 2012). Here we propose a novel situation -- that, depending on the relationship between recombination and mutation rates in the genome, selfing populations can actually be more effective at removing additive mutations, and have less genetic load, than outcrossing populations.

It is widely understood that the effective population size, N_e, determines the efficacy of selection. However, the existence of several concepts of N_e in population genetics, each associated with the population genetic question at hand (Kimura 1983), can generate some confusion. For example, the fusion of gametes from the same individual inherently reduces the number of independent genetic backgrounds in a population. This is known as the two-fold effect, because coalescent derivations of neutral diversity showed that $N_e = N/(1+F)$, where F is the inbreeding coefficient (Nordborg and Donnelly 1997). Full selfing (F=1) halves N_e relative to N. It is therefore commonly assumed that the strength of selection on a given additive mutation (N_e*s) will be halved in a selfing compared to outcrossing population given that they have equal census sizes (N). However, at a single locus, selfing does not affect the efficacy of selection in removing a deleterious mutation with an additive effect on fitness, and has nearly no effect on the removal of recessive or dominant mutations (Charlesworth 1992). This is because the measure of N_e that decreases with selfing determines the coalescent rate, not the efficacy of selection.

By contrast, the decrease in the effective recombination rate associated with selfing can decrease the efficacy of selection (Kamran-Disfani and Agrawal 2014). That is, the increase in homozygosity with selfing results in lower effective recombination rates because crossover events between two homozygous chromosomes result in parental-like, rather than recombinant, chromosomes (Glémin and Galtier 2012). The reduction in the effective recombination rate is simply the actual recombination rate times one minus the selfing rate; thus, there is eventually no effective recombination with full selfing (Nordborg 2000). Consequently, forms of linked selection, particularly background selection, are stronger in selfers, reducing neutral diversity and the efficacy of selection (Charlesworth *et al.* 1993a; b; Kamran-Disfani and Agrawal 2014; Slotte 2014).

As such, slightly deleterious mutations that are selected against in outcrossing populations are more likely to drift to fixation in selfing populations. These slightly deleterious mutations are often additive (Agrawal and Whitlock 2011; Huber *et al.* 2018), and thus are expressed in both outcrossing and selfing populations, contributing directly to genetic load. It is currently believed that, among the genomic consequences of selfing, this increase in the reach of background selection plays a dominant role in decreasing the efficacy of selection in selfers relative to outcrossers (Glémin 2007; Wright *et al.* 2008; Kamran-Disfani and Agrawal 2014; Arunkumar *et al.* 2015).

On the other hand, the purging of recessive mutations is associated with a *greater* efficacy of selection in selfers than outcrossers (Glémin 2007). Strongly deleterious mutations are often recessive (Agrawal and Whitlock 2011; Huber *et al.* 2018), and thus are masked in the heterozygous state in outcrossing populations. Selfing populations are able to remove recessive deleterious alleles from the population through exposure in the homozygous state, a.k.a "purging" (Crnokrak and Barrett 2002) or the "segregation effect" (Uyenoyama and Waller 1991). Thus, while selfing populations are predicted to accumulate more additive, slightly deleterious mutations, outcrossing populations should accumulate more recessive, strongly deleterious mutations.

These consequences of selfing on the efficacy of selection against additive versus recessive mutations have been studied in isolation. It is therefore unclear if and/or how mutations with different dominance coefficients interact with one another and the mating system to affect total genetic load. However, it seems likely that different architectures of genetic load in selfers and outcrossers can result in different types of linked selection predominating in parts of the genome with low recombination. The additive load accumulates in selfing populations because their low effective recombination rate increases background selection, preventing the precise selection against mildly deleterious mutations (Charlesworth *et al.* 1993a). Conversely, in outcrossing populations, deleterious recessive mutations can be induce either background selection, when selection can maintain haplotypes with no recessive deleterious mutations, or a balancing selection known as pseudo-overdominance, when it cannot (Ohta and Kimura 1970; Gilbert *et al.* 2020). With pseudo-overdominance, fit individuals have complementary haplotypes in repulsion -- that is, heterozygotes with recessive alleles at different loci are most fit. Pseudo-overdominance acts as a form of balancing selection, increasing variation at linked neutral loci (Gilbert *et al.* 2020). The effect of pseudo-overdominance on the efficacy of selection against linked *additive* mutation is unknown, but any such impact will be disproportionately felt in outcrossing populations that do not purge their recessive load.

Theoretical expectations for reduced efficacy of selection in selfers have mixed empirical support. There is strong empirical evidence of reduced neutral diversity (π) in selfing taxa relative to outcrossing taxa, suggesting that selfers do experience lower Ne than their outcrossing relatives (Glémin *et al.* 2006; Foxe *et al.* 2009; Koelling *et al.* 2011; Brandvain *et al.* 2014). However, the evidence for differences in selection efficacy (primarily molecular differences in π_{NS}/π_S and/or differences in premature stop codons) is equivocal (Haudry *et al.* 2008; Escobar *et al.* 2010; Slotte *et al.* 2010, 2013; Qiu *et al.* 2011; Ness *et al.* 2012; Hazzouri *et al.* 2013; Gioti *et al.* 2013; Brandvain *et al.* 2014). The lack of consistent evidence for reduced efficacy of selection in selfers is often explained by the different histories and population dynamics of most selfers compared to outcrossers. Notably, selfing lineages are often of recent origin and, thus, haven't had enough time to accumulate more deleterious mutations than their outcrossing sister taxa (Haudry *et al.* 2008; Glémin and Galtier 2012). In addition, selfing lineages are more prone to recurrent and/or extreme population bottlenecks, which substantially reduce N_e beyond the expected two-fold decline caused by inbreeding alone (Ingvarsson 2002; Brandvain *et al.* 2013, 2014; Barrett *et al.* 2014; Pettengill *et al.* 2016). Non-equilibrial demography, which can affect all populations regardless of mating system, can have complex effects on selection efficacy and the architecture of genetic load, confounding inference of selection efficacy from molecular variation (Brandvain and Wright 2016).

Simulation studies have also found equivocal evidence of reduced selection efficacy in selfers, suggesting that the genomic consequences of selfing may not consistently affect the efficacy of selection. For example, Kim *et al*. (2018) found that increases in the selfing rate led to a greater accumulation of deleterious mutations, but Arunkumar *et al*. (2015) only found evidence for small reductions in selection efficacy against weakly deleterious mutations in selfers under a restricted set of parameter combinations. These explicit forward-time simulations were often built to model the specific biology of the organism of interest and/or explored a limited part of parameter space, and therefore results to date limit our ability to generalize more broadly.

This paper aims to solve this paradox of reduced efficacy of selection in selfers – i.e., most theory suggests that the genomic consequences of selfing should have a negative effect on selection efficacy, whereas empirical and simulation studies provide equivocal evidence. We used individual-based forward simulations to explore how selfing affects the efficacy of selection against recessive and additive deleterious mutations. In particular, we examined how the selfing rate interacts with the action of linked selection and purging to differentially shape the genetic load of selfers and outcrossers. We performed these simulations under a range of mutation and recombination environments. Our results indicate that recombination rate modulates the direction and strength of the relationship between selection efficacy and selfing rate. These qualitative outcomes are the result of a balance between different forms of linked selection that predominate in selfing vs outcrossing genomes. Because we find that the transition from background selection to pseudo-overdominance interacts with the selfing rate to produce unexpected patterns of genetic load, we extend analytical models from Gilbert *et al.* (2020) to determine if analytical theory predicts similar shifts from background selection to pseudo-overdominance as a function of the selfing rate.

MATERIALS AND METHODS

To evaluate how linked selection, purging and inbreeding interact to affect the efficacy of purifying selection, we used individual-based forward simulations in SLiM v3.3.2 (Haller and Messer 2019), and developed analytical theory to complement these simulations.

Individual-based forward simulations in SLiM

Fixed model parameters:

Demography: All simulations consisted of a census size of 10,000 diploids with non-overlapping generations and multiplicative fitness across loci. Fitness at each locus is defined by $1 - hs$. All simulations ran for 6N generations.

Genome size and structure: Genomes were 45 Mb and split into 6-7.5Mb chromosomes, as in Gilbert *et al*. (2020). In each simulation, the recombination rate was constant for each chromosome, with free recombination between chromosomes. For simplicity we used a uniform genomic architecture, i.e., each type of mutation (see below) could occur anywhere in the genome with equal probability.

Mutational effects: In SLiM, specific "mutation types" are simulated in the genome. Each mutation type is characterized by a fixed dominance coefficient (h) and distribution of selection coefficients (s). All simulations included four mutation types – one (fully to partially) recessive mutation type and three additive mutation types. There were no neutral or beneficial mutations in the simulations, such that the overall genome-wide mutation rate (U) of a simulation represented the deleterious genome-wide mutation rate (U_{del}). Each of the four mutation types contributed equally to U_{del} such that the total additive deleterious mutation rate was three times that of the recessive deleterious mutation rate.

The three additive mutation types all had dominance coefficients of 0.5 and each of the additive mutation types had a different fixed selection coefficient. These three selection coefficients were parameterized to hover around the nearly neutral boundary ($4N_es > 1$), which differentiates where natural selection can and cannot effectively remove deleterious mutations (Kondrashov 1995). The nearly neutral boundary is where we expect to see a difference in selection efficacy against additive load in selfers and outcrossers. The selection coefficients ranged from $s = 0.0005$ ($Ns = 5$, $4Ns = 20$), $s = 0.00025$ ($Ns = 2.5$, $4Ns$=10) to $s = 0.00005$ ($Ns = 0.5$, $4Ns$=2). Given that N_e is likely less than the census population size (N) in all multilocus simulations, we assume the $4N_es$ values will be less than the $4Ns$ values listed above. We chose these fixed selection coefficients, rather than the more biologically realistic distribution of fitness effects, because they provide theoretical insight into when and how selection becomes less effective.

Variable model parameters:

On top of these fixed parameters, we investigated all factorial combinations of five variables: (1) selfing rate, (2) deleterious mutation rate, (3) recombination rate, (4) fitness cost of strongly deleterious recessive mutations ($s_{recessive}$) and (5) recessivity of strongly deleterious mutations ($h_{recessive}$). These factors modulate the extent of background selection and purging. Across all unique parameter combinations, we ran 10 replicates.

Selfing rate: We examined populations with selfing rates ranging from obligate outcrossing ($\alpha = 0$) to near obligate selfing ($\alpha = 0.99$) and values between these extremes ($\alpha = 0.05, 0.1, 0.25, 0.5, 0.75$, or 0.9). This allowed us to evaluate not only the effects of relatively rare selfing or outcrossing events but also that of mixed mating systems on selection efficacy. We note that we do not model an evolutionary transition between selfing rates, and therefore do not consider the initial purging, or lack thereof, of the load required for the transition to selfing (see Wang *et al.* 1999; Bataillon and Kirkpatrick 2000; Waller 2021).

Deleterious mutation rate: We varied the genome-wide deleterious mutation rate ($U_{del} = \mu_{del}$ * genome size), as U_{del} changes the expected background in which a new mutation arises, modulating background selection (Charlesworth *et al.* 1993a; Kamran-Disfani and Agrawal 2014). Each other the four mutation types occurred at equal frequency, such that U-per-mutation type was one quarter of U_{del} and the additive mutation rate was three times that of the recessive mutation rate. We chose U_{del} values of 0.04, 0.16, and 0.48 to span a range of empirical estimates of genome-wide deleterious mutation rates estimated in multicellular eukaryotes (Schultz *et al.* 1999; Willis 1999; Cutter and Payseur 2003; Haag-Liautard *et al.* 2007; Lynch 2010; Slotte 2014).

Recombination rate: To further probe how background selection interacts with the selfing rate to modulate the efficacy of selection, we varied the recombination rate as the per-base-pair recombination rate divided by the per-base-pair mutation rate. We call this the relative recombination rate which took values of 0.01, 0.1, 1 and 10. Per-base-pair recombination rates ranged from $8.89\text{x}10^{-12}$ to $1.07\text{x}10^{-7}$ across all mutation rates (see Table S1 for per-base pair mutation and recombination rates for each parameter combination).

Selection against recessive deleterious mutations: We varied the intensity of selection against strongly deleterious mutations from $s_{recessive} = 0.015$, $s_{recessive} = 0.3$, and $s_{recessive} = 0.9$. All of these values prevent the chance fixation of such mutations, but they have different effects on the expected rate of removal from populations and on organismal fitness.

Recessivity of strongly deleterious mutations: We varied the recessivity of the strongly deleterious recessive mutation ($h_{recessive} = 0$, $h_{recessive} = 0.1$ or $h_{recessive}=0.25$), as we expect recessivity to affect the degree to which an increase in selfing rate causes purging of the recessive load.

Quantifying the efficacy of selection and consequences of linked selection:

We explored how the efficacy of selection changes as a function of selfing rate for every combination of the genome-wide deleterious mutation rate (U_{del}), the relative recombination rate, and the selection ($s_{recessive}$) and dominance ($h_{recessive}$) coefficients of recessive mutations. We focused first on patterns of selection efficacy against recessive mutations and then against additive mutations. We assessed the efficacy of selection both in terms of prevalence (the average number of mutations per individual) and fitness for each of the four mutation types, as well as overall population fitness. Because the degree of homozygosity increases as a function of selfing rate, we expected that the translation of prevalence to fitness of recessive mutations would vary across selfing rates. By contrast, additive mutations are expressed to some extent across all populations and thus patterns of prevalence across selfing rates should translate more directly to patterns of overall fitness. Prevalence is meant to capture common genomic summaries of the number of derived deleterious mutations, while the genetic load is meant to quantify the fitness consequences of these mutations (Lohmueller 2014; Do *et al.* 2015).

We quantified both neutral diversity (π) and the unfolded allele frequency spectrum (AFS) in eight replicates runs per simulation to assess how selection dynamics and mating system interact to affect patterns of sequence diversity. Because explicitly modeling neutral mutations in simulations drastically increases the computational resources needed per run, we used the tree sequence recording function within SLiM (Haller *et al.* 2019) to output a tree sequence at the final generation of each simulation. Tree sequences record genealogies across the genome, and features of a tree sequence (e.g., the number of unique trees recorded across the genome, tree height, and tree topology) are affected by the selection and recombination dynamics in the simulation. We subsequently overlaid neutral mutations on each tree sequence at a mutation rate of μ = 1e-7 in msprime (Kelleher *et al.* 2016), sampled one genome from 200 individuals, and calculated Tajima's π and the AFS with the 200 samples using tskit (https://tskit.dev/tskit/docs/stable/).

In the main text we only present data for the most deleterious of the three additive deleterious mutation types (s = 0.0005, Ns = 5). For any given parameter combination, the way that selfing rate affected selection efficacy was qualitatively similar among the three additive mutation types. With smaller selection coefficients the effect of selfing was less pronounced, as the deleterious mutations became effectively "nearly neutral" regardless of selfing rate. We present data for all three additive mutation types in the supplement.

Analytical Model for the transition to pseudo-overdominance

Our SLiMulation results suggested that the transition from background selection to pseudo-overdominance plays a major role in mediating the effect of mating system on the efficacy of selection We therefore extended the multi-locus model of this transition from Gilbert *et al.* (2020) to include selfing. We considered $n$biallelic loci and denoted wild-type and derived alleles alleles at locus i by a_i and A_i, respectively. Derived alleles were assumed deleterious and fully recessive with fitness $1 - s$when homozygous and1otherwise, with multiplicative fitness effects across loci. Following the approach in Gilbert *et al.* (2020), we derived the frequency of the zero-mutation haplotype at mutation-selection balance. When this frequency approaches zero, we expect a transition from background selection to pseudo-overdominance. We assumed that haplotypes with more than one deleterious mutation are vanishingly rare at equilibrium and hence ignored them. This has the important consequence that any genotype can be polymorphic at most at two deleterious loci. This assumption is quite different than what we find in our SLiMulation model, but we show below that it still allows for a reasonable guide to our major qualitative results.

We used p_i, $i = 1, \ldots, n$, to denote the frequency of the haplotype carrying a derived deleterious allele at locus i, and p_0 the frequency of the haplotype without any deleterious mutations. Loci were equidistantly distributed over a region of length r cM, such that the recombination rate between adjacent loci is $r/(n - 1)$. This model can be studied by following the frequencies of four different categories: genotypes without any deleterious mutations (p_{00}), genotypes that are heterozygous at exactly one locus ($p_{01}^{(1)}$), genotypes that are heterozygous at exactly two loci ($p_{01}^{(2)}$) and genotypes that are derived homozygous at exactly one locus (p_{11}). Fitness of these genotype-classes were set to $w_{00} = w_{01}^{(1)} = w_{01}^{(2)} = 1$ and $w_{11} = 1 - s$. Mean fitness is then simply $1 - p_{11}\, s$. In the Appendix we derived equations for how the frequencies of these haplotype classes change over time under the joint action of selection, recombination, mutation and selfing. The frequency of the zero-mutation haplotype could be readily determined in the absence of recombination, and was approximated for weak recombination and selfing rates, see eqs. (1) and (2) in the Appendix.

Data Availability

All scripts used to generate SLiM scripts, scripts for analysis and figures, and seeds used to run simulations will be available at: https://github.com/ssianta/Selfing_Genetic_Load.git. Supplementary Material and the Appendix (attached at end of pdf) will ultimately be uploaded to figshare.

RESULTS

Selfing does not have a consistent effect on the genetic load

Across parameter space, we found that increases in the selfing rate can have highly variable effects on the genetic load: it can either slightly or severely reduce the frequency of recessive deleterious mutations, and it can result in an increase, decrease or no change in the frequency of mildly deleterious additive mutations. We explore these alternative outcomes, and what modulates them, below.

Selfing prevents the accumulation of recessive mutations and impedes the transition to pseudo-overdominance

Selfers largely purge their recessive load, while obligate outcrossers maintain a large number of recessive mutations. When recombination rate is equal to or greater than the mutation rate, the decay in prevalence of recessive deleterious mutations with the selfing rate tends to be smooth. The decay in prevalence is rapid and steep when mutations are strongly deleterious -- populations with mixed mating systems effectively purge extremely deleterious mutations ($s_{recessive}$= 0.3 or $s_{recessive}$= 0.9), but maintain a larger number of less deleterious mutations ($s_{recessive} = 0.015$, Figure 1, Figure S1-A). The purging dynamics in our multilocus simulations are consistent with our analytical derivation of the frequency of a recessive deleterious mutation at a single locus as a function of selfing rate and with that derived by Roze and Rousset (2004) (equation A1 and Figure A1 in the Appendix; Figure S1-B).

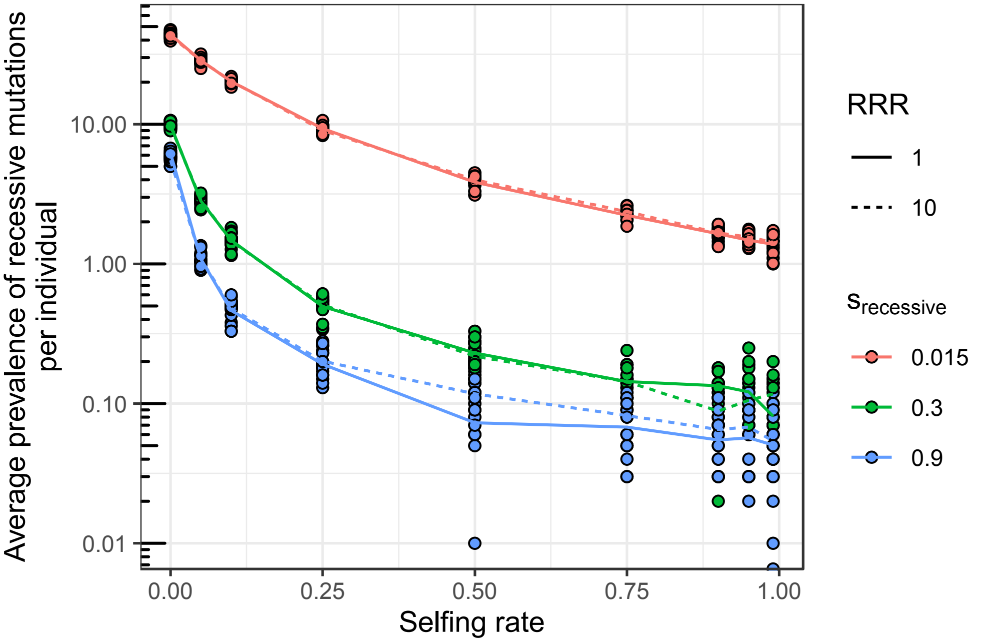

Figure 1: Purging dynamics across selfing rates are dependent on the fitness effects of recessive mutations ($s_{recessive}$). Purging patterns are similar when the recombination rate is equal to or greater than the mutation rate (relative recombination rate, RRR, >= 1). Points are simulation replicates and lines connect mean values across simulation replicates. $U_{del} = 0.04$. Note log10 scale on y-axis.

Under some parameters combinations, the number of recessive mutations appears to drastically decrease as the selfing rate increases from zero, and smoothly decrease at higher selfing rates. Such drastic shifts are mediated by the recombination rate relative to the deleterious mutation rate. For example, Figure 2A shows a dramatic decline in the number of recessive mutations between a selfing rate of 0 and 0.1 when recombination is much rarer than mutation (dark blue line), yet a smooth decay when mutation is as equal to (yellow) or rarer than (red) mutation. These drastic drops in the prevalence of recessive mutations are more common, and larger in magnitude, with smaller dominance coefficients (Figure S2).

We attribute the increase in recessive load in primarily outcrossing populations to a shift from background selection to pseudo-overdominance. Consistent with this explanation, simulations runs where there is a shift to pseudo-overdominance are associated with elevated genetic diversity (π) at linked neutral sites (see Gilbert *et al.* 2020, Figure 2B, Figure S3) and a shift in the allele frequency spectrum to more high-frequency alleles (Figure 2C, Figure S4). We use these three defining characteristics - an increase in prevalence of recessive mutations, an increase in π, and a shift in the AFS - to qualitatively determine whether or not pseudo-overdominance occurs in each simulation.

The shift to pseudo-overdominance does not occur when selfing leads to purging of the recessive load and maintains at least one haplotype free of deleterious recessive mutations. The magnitude of selfing at which populations transition from background selection to pseudo-overdominance is a function of the fitness effects of recessive mutations and the mutation rate - more selfing is required to purge the recessive load when mutations are less deleterious and when mutations occur at higher rates (Figure 2D). Once pseudo-overdominance haplotypes form homozygote fitness plummets; as homozygotes contribute relatively few offspring to the next generation these haplotypes further accumulate a large number of deleterious recessive mutations With partial selfing (or more recombination) fitness due to pseudo-overdominance haplotypes drops further, as recombination creates more non-complementary haplotypes (the recombination load) and/or as selfing exposes more haplotypes in the homozygous state (the segregation load, Figure S5).

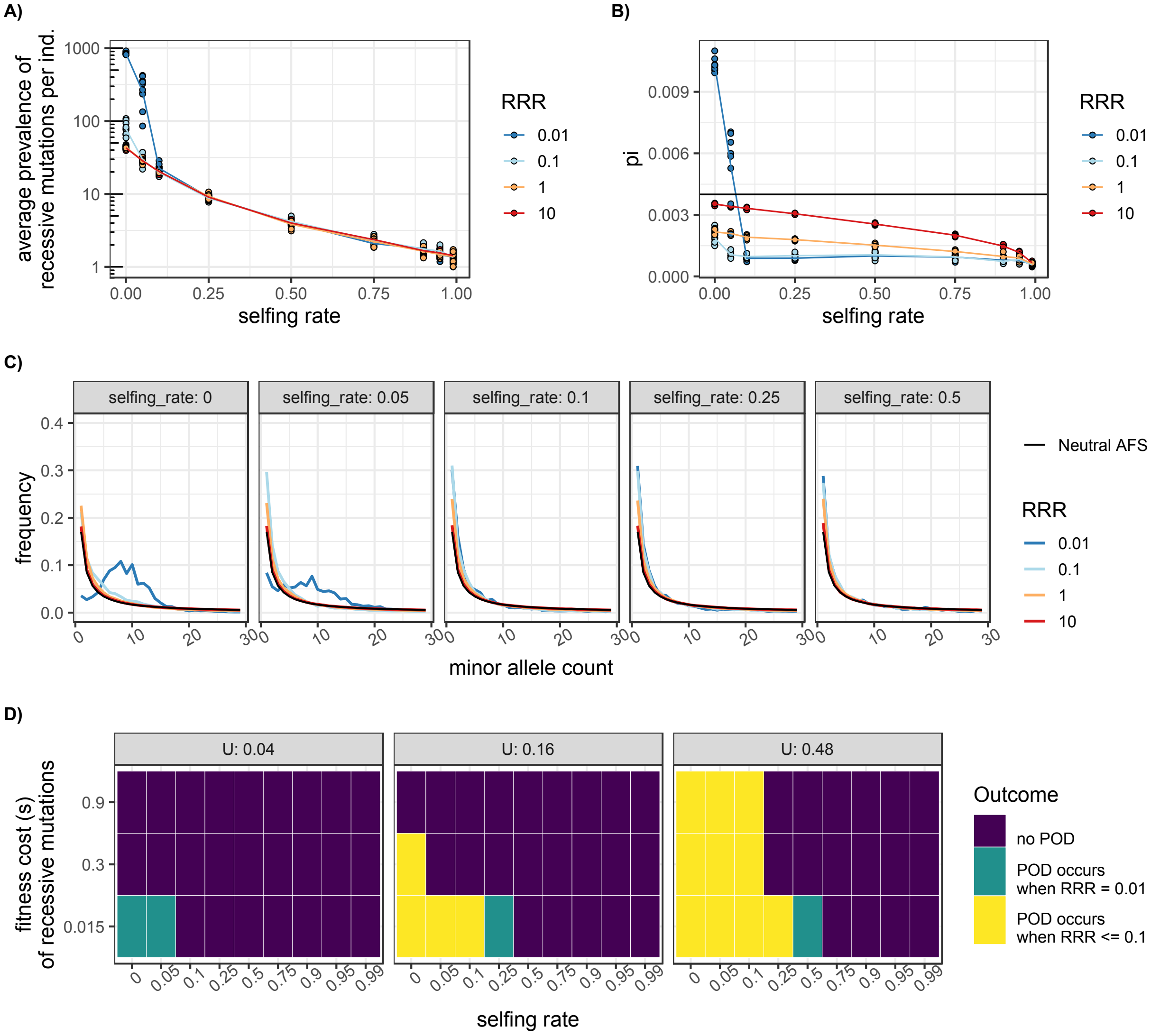

Figure 2: Pseudo-overdominance (POD) occurs in low recombination environments (relative recombination rate, RRR, < 1). (A) POD leads to a spike in the prevalence of recessive mutations in predominantly outcrossing populations. Points are simulation replicates and lines connect mean values. Genome-wide deleterious mutation rate $U_{del} = 0.04$, $s_{recessive} = 0.015$, and $h_{recessive} = 0$. (B) POD also leads to a spike in neutral diversity, driven by heterozygosity at linked neutral sites. $U_{del} = 0.04$, $s_{recessive} = 0.015$, and $h_{recessive} = 0$. Expected neutral diversity (4Nμ) is shown by the black horizontal line. (C) Allele frequency spectra (AFS) at a subset of selfing rates (different facets) for $U_{del} = 0.04$, $s_{recessive} = 0.015$, and $h_{recessive} = 0$. POD shifts the AFS to more intermediate frequency alleles. Mean AFS are in bold lines, and individual simulation replicates are in faint lines. Black lines correspond to the neutral AFS. (D) Outcome

plot of POD occurrence when $h_{recessive} = 0$. Green blocks indicate POD occurs at only the lowest relative recombination rate (RRR; RRR = 0.01); yellow blocks indicate POD occurs at the two lowest RRR (0.01, 0.1). Data in panels A-C correspond to the bottom row of the $U_{del} = 0.04$ outcome plot.

Full recessivity is not required for pseudo-overdominance. Because partial recessivity ($h_{recessive}= 0.1$ or $h_{recessive}= 0.25$) increases the ability for selection to remove recessive mutations in predominantly outcrossing populations, it substantially decreases the parameter space under which pseudo-overdominance occurs (Figure S6-S9). When $h_{recessive}= 0.1$, pseudo-overdominance occurs when $s_{recessive}$ is relatively modest ($s_{recessive}= 0.015$) and mutation rates are high (U_{del}=0.16 and U_{del}=0.48). When $h_{recessive} = 0.25$, recessive mutations accumulate in predominantly outcrossing populations at the highest mutation rate and lowest selection coefficient (Figure S2-C), but the accumulation is not enough to cause the switch to pseudo-overdominance, i.e., there is no increase in π (Figure S3-C) nor a shift in the AFS (Figure S9). Under the remainder of parameter space, increasing the dominance coefficient of recessive mutations decreases prevalence (number of mutations per individual) in the primarily outcrossing populations, as expected. Prevalence still decreases with selfing rate, but the absolute difference in prevalence between outcrossers and selfers diminishes (Figure S7). Thus, while partial recessivity limits the extent of pseudo-overdominance, it does not eliminate it.

Analytical theory can explain the transition to pseudo-overdominance with selfing

In the Appendix we derived an approximation for the frequency of the unloaded haplotype under the assumption of weak selfing and weak recombination (equation (2) in the Appendix). Roughly speaking, our approximation assumes that $F, r, u < 1/N < s$, but the exact conditions for when our approximation will become unreliable are difficult to derive (see Appendix for more details). We find that our approximation is in perfect agreement with previous single-locus theoretical results of Roze and Rousset (2004) and that the approximation for the unloaded haplotype fits very well with results obtained from two-locus simulations (see Appendix Figure A1 and A2). Importantly, the critical selfing rate for loss of the zero-mutation haplotype in two-locus simulations is accurately predicted by our model (see Appendix Figure A2).

Figure 3 shows the predicted frequency of the zero-mutation haplotype and illustrates that pseudo-overdominance is most likely for a combination of small selection coefficients and small selfing rates. The main effect of increasing recombination is that selfing becomes more efficient at preventing pseudo-overdominance, particularly at small selection coefficients (compare Figure 3A with a relative recombination rate of 0.01 to 3B with a relative recombination rate of 0.1). This can be considered as a synergistic effect of recombination and selfing on the efficacy of purging the recessive load, and thus preventing pseudo-overdominance. Recombination generally increases the genetic variation across individuals by creating novel combinations of haplotypes. At a given partial selfing rate, having higher recombination increases the variance in fitness exposed at that selfing rate, maximizing the efficacy of selection. Synergies between recombination and selfing are strongest for small selection coefficients, as seen by the large proportional decrease in parameter space resulting in pseudo-overdominance at small selection coefficients in Figure 3A (recombination rate 1/100 of mutation rate) vs 3B (recombination rate 1/10 of mutation rate). This effect is mostly observed for small selection coefficients, simply because pseudo-overdominance is unlikely if selection is strong and hence there is no opportunity for recombination to prevent it.. The synergistic effect of recombination and selfing is quite general in our model (Appendix Figure A3) and also observed in the individual-based simulations (Figures 2D and 3).

Our analytical model shows that higher mutation rates also increase the parameter range for which we observe pseudo-overdominance (Appendix Figure A3). Even though our approximation is not tailored to direct comparison with genome-wide simulations of thousands of loci in regions of low recombination, we find a tight qualitative fit of the analytical predictions with the simulation results (Figure 3). Note, however, that the parameters used for the comparison between simulations and the analytical model in

Figure 3 are not directly comparable as we assumed a much smaller number of loci and larger per locus mutation rate in the analytical model as compared to the simulations.

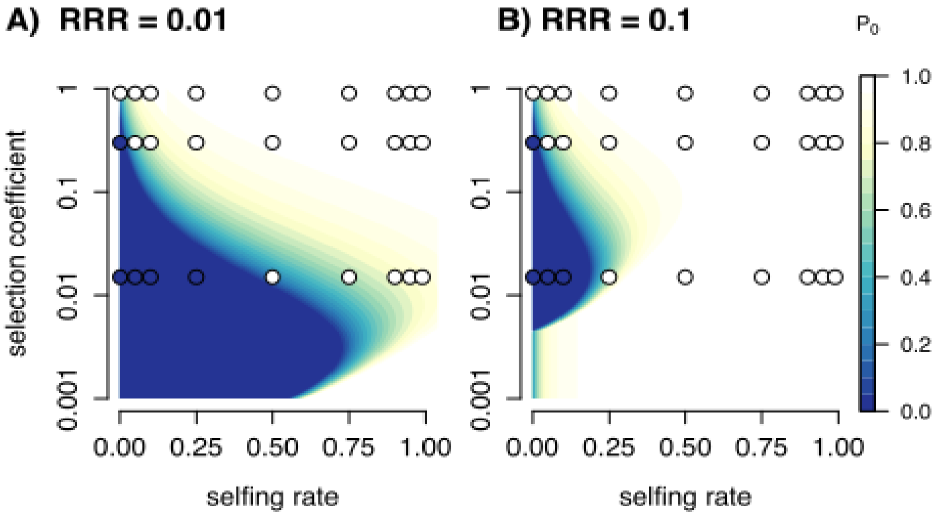

Figure 3: Qualitative comparison between analytical predictions and simulation results. Shaded areas indicate the analytical predictions for pseudo-overdominance (POD), specifically the frequency of the zero-mutation haplotype (P_0), based on equation (2; see Appendix) for various recombination rates. For the analytical model, we assumed a total of $n = 100$ loci equidistantly spaced on a chromosome with a total mutation rate of $U_{del} = 0.005$ and relative recombination rate RRR = 0.01 (A) and 0.1 (B). Circles show results from simulations when $U_{del} = 0.16$: filled circles indicate simulations where we observed POD and white circles indicate no POD.

The additive load does not necessarily increase with selfing

Previous theory, modelling mildly deleterious additive alleles in isolation, showed that the lower effective recombination rate in selfers should result in higher accumulation and fixation of additive, weakly deleterious mutations than in outcrossing populations (Glémin 2007; Wright *et al.* 2008; Glémin and Galtier 2012). By simultaneously considering the evolution of both recessive and additive deleterious variants, we show that this is not always the case. We find that the additive load can increase, decrease, or remain constant as a function of the selfing rate. These different outcomes are determined by the recombination environment and whether or not primarily outcrossing populations experience background selection or pseudo-overdominance, both of which disproportionately affect the efficacy of selection in primarily outcrossing populations. To demonstrate the effects of pseudo-overdominance on selection against additive mutations, we contrast simulations with fully recessive mutations ($h_{recessive}$= 0; where pseudo-overdominance occurs) to simulations with partially recessive mutations ($h_{recessive}$= 0.25; where pseudo-overdominance does not occur). Below, we present prevalence results for the most deleterious of the three additive mutation types that were present in each simulation ($s_{additive}$ = 0.0005, Ns = 5), as selection dynamics are qualitatively similar across the three mutation types (Figures S10-12).

Selfers evolve a larger additive load than outcrossers when recombination rates are high:

Selfing only results in higher prevalence of additive mutations in high recombination environments (recombination rate is greater than (red) or equal to (yellow) the mutation rate, Figures 4, S10-12). There is no difference in additive prevalence or genetic load between simulations with fully or partially recessive load, as the recessive load is low in both cases with high recombination and pseudo-overdominance never occurs (Figure 4 A vs B for high RRR). At the highest relative recombination rates (recombination 10x greater than mutation), an increase in the additive load only occurs with exceptionally large selfing rates; as the local effective recombination rate (and therefore local N_e) is still sufficient to remove mildly deleterious mutations in mixed mating populations. As the recombination rate gets closer to the mutation rate, the prevalence of mildly deleterious additive mutations increases more gradually with the selfing rate.

With low recombination, selfers evolve an additive load that is equal to or smaller than that of outcrossers:

When recombination is rarer than mutation the prevalence of additive mutations either decreases, does not change, or subtly increases with the selfing rate. The different outcomes depend on whether there has been a shift from background selection to pseudo-overdominance in primarily outcrossing populations. When primarily outcrossing populations experience background selection instead of pseudo-overdominance (e.g., in simulations with $h_{recessive}$= 0.25 instead of $h_{recessive}$= 0) the additive load is largely insensitive to, or slightly increases with, the selfing rate (blue lines, Figure 4B, S10-12).

As linked selection in outcrossing populations transitions from background selection to pseudo-overdominance, outcrossing populations accumulate a larger number of linked additive mutations than selfing populations. Figures 4A vs B distinguishes the influence of pseudo-overdominance on the additive load by contrasting the same set of parameter conditions (U_{del} = 0.04 and $s_{recessive}$= 0.015) between simulations with fully recessive (Figure 4A) and partially recessive (Figure 4B) mutations. Under this set of parameter conditions, pseudo-overdominance and the large increase in the prevalence of additive mutations only occurs in primarily outcrossing populations when $h_{recessive}$= 0 (Figure 4A). When the recessive mutations are partially recessive ($h_{recessive}$= 0.25), pseudo-overdominance does not occur (Figure S6) and the additive load does not vary with the selfing rate in low recombination simulations (Figure 4B).

We propose that pseudo-overdominance limits the efficacy of selection through effectively subdividing the population into haplotypic classes of complementary recessive mutations. When there are only recessive mutations driving pseudo-overdominance in a population, an equilibrium frequency is reached for each pseudo-overdominant haplotype, such that the expressed recessive load is minimized. Additional selection for or against a linked additive mutation will lead to deviations in the equilibrium haplotype frequencies and an increase in the expression of the recessive load on more-frequent pseudo-overdominant haplotypes. Thus selection on the recessive load will counteract selection on the additive mutation. In the Appendix, we show that when an additive deleterious (or beneficial) mutation falls on a haplotype maintained at equilibrium by pseudo-overdominance, selection against (or for) the new mutation will be limited by the recessive load at linked sites. Specifically, we show in a two locus model for pseudo-overdominance that the efficacy of selection is reduced by a factor of 1-$s/2$ for outcrossing populations, where s is the fitness effect of recessive mutations. Intuitively speaking, in our two-locus model a new additive mutation will have a 50% chance to be in a beneficial heterozygous genotype where the recessive load is masked, effectively reducing the strength of selection against (or for) the additive mutation by (1-s) in half of the genotypes.

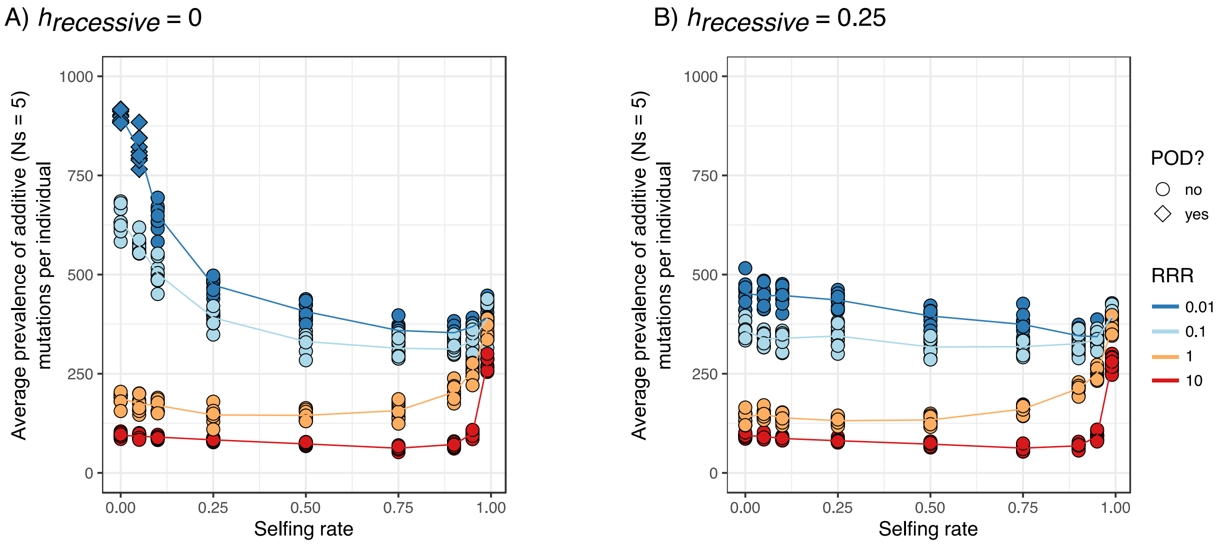

Figure 4: The accumulation of additive mutations as a function of selfing rate is heavily influenced by the relative recombination rate (RRR) and whether or not pseudo-overdominance (POD) occurs. (A) and (B)

contrast simulations where the recessive load is fully recessive (A) and partially recessive (B). $U_{del} = 0.04$ and $s_{recessive} = 0.015$.

The effect of mating system on mean population fitness

When recombination rates are high relative to mutation rates, mean population fitness is generally greatest in outcrossers and lowest in selfers, reflecting the elevated additive load accumulated by selfers (Figure 4A, Figure S13). By contrast when recombination rates are lower than mutation rates, mean population fitness either does not vary with selfing rate (in the absence of pseudo-overdominance), or increases with selfing rate (in the presence of pseudo-overdominance). The effect of pseudo-overdominance on overall fitness is primarily due to its effect on the prevalence of additive mutations.

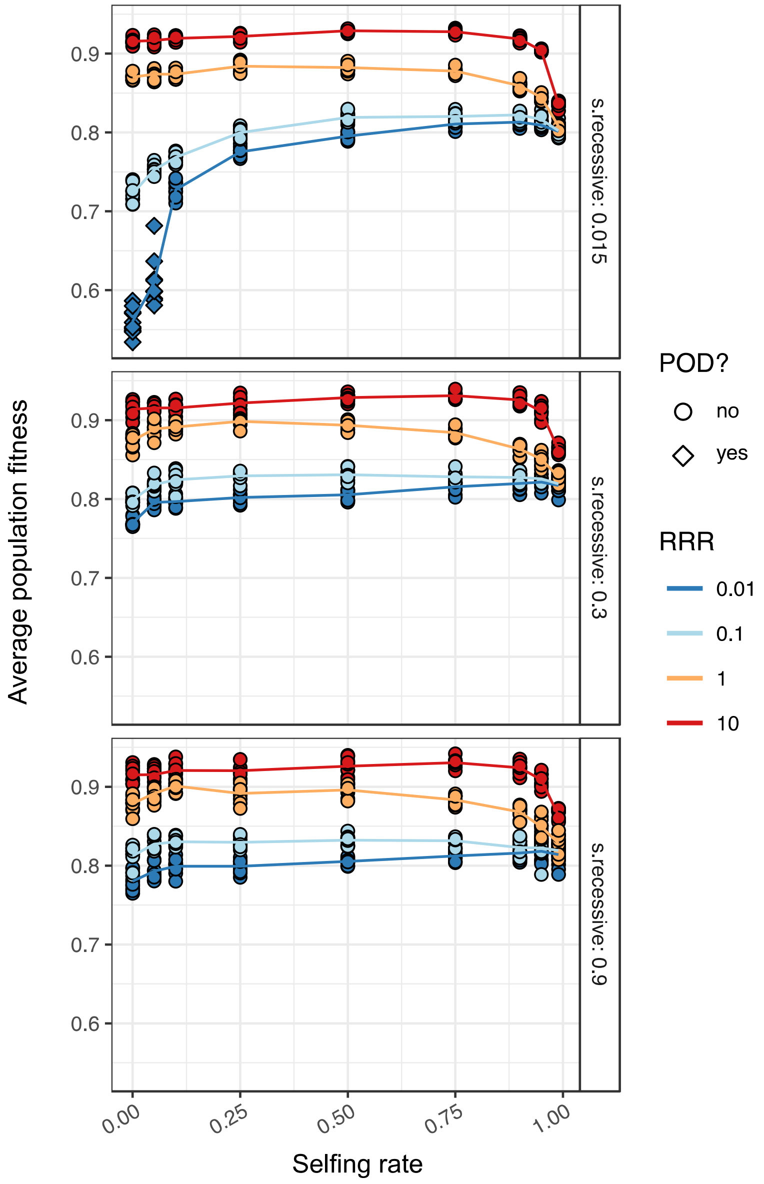

Figure 5: The relationship between mean population fitness and the selfing rate varies as a function of the recombination rate and the presence of pseudo-overdominance (POD). $U_{del} = 0.04$, $h_{recessive} = 0$. The three facets represent different fitness effects of the recessive mutations. At this U_{del}, POD only occurs when the recessive mutations are relatively mild ($s_{recessive} = 0.015$).

DISCUSSION

The evolution of self-fertilization is remarkably common across the tree of life, including many animals, fungi, plants, protozoa, and algae (Jarne and Auld 2006; Hanschen *et al.* 2018). Evolutionary transitions to selfing also provide an unusual opportunity to examine how genomic factors affect the efficacy of selection against deleterious mutations. It has long been argued that selfing should decrease the efficacy of selection. Historically, it was argued that selfing decreases the efficacy of selection by its two-fold decline in N_e (the two-fold effect). More recently, it has become appreciated that the reduced effective recombination rate (because recombination between homozygous chromosomes does not generate new allele combinations) plays a larger role in reducing selection's efficacy than does the two-fold effect (Charlesworth 1992; Charlesworth and Wright 2001; Kamran-Disfani and Agrawal 2014). Empirical evidence for a decrease in selection efficacy with a shift to selfing, however, is equivocal. Here, we simulated the accumulation of recessive and additive genetic load in populations that vary in selfing rate, under a wide array of parameter space. We found that increases in selfing rate can have positive, neutral or negative effects on selection efficacy; the direction and magnitude of the relationship between selfing rate and selection efficacy is primarily determined by the recombination environment.

Effects of selfing rate on the efficacy of selection against deleterious mutations

Selection efficacy against deleterious mutations is primarily a function of the relative recombination rate (the per-base recombination rate relative to the per-base deleterious mutation rate), which highlights the importance of linked selection. Because predominantly selfing populations have near-zero effective recombination rates and always experience background selection (Nordborg 2000), recombination rate affects patterns of selection efficacy by changing the magnitude of linked selection in outcrossing and mixed-mating populations.

A little outcrossing goes a long way when recombination is frequent:

When the recombination rate is large relative to the mutation rate, mildly deleterious additive mutations accumulate at high selfing rates (selfing rates > 0.9 in our simulations). Here, background selection limits selection's efficacy in predominant selfers, while mixed mating and outcrossing populations can recombine away their deleterious load (as seen in Charlesworth *et al.* 1993a; Glémin 2007; Glémin and Ronfort 2013; Kamran-Disfani and Agrawal 2014). As the recombination rate becomes more similar to the mutation rate, we see a more linear decrease in the efficacy of selection with selfing rate, presumably because the reduced effective recombination rate in mixed mating populations is low enough to increase background selection and reduce the efficacy of selection (as in Glémin 2007; Glémin and Ronfort 2013; Kamran-Disfani and Agrawal 2014).

With high recombination rates, even rare outcrossing events can remove the additive genetic load. With full selfing, the physical extent of linkage disequilibrium is large (Nordborg 2000), as recombination is effectively eliminated. Because of the greater extent of linkage disequilibrium, more deleterious mutations, that would otherwise be effectively unlinked, contribute to background selection (Roze 2016). While full selfing completely suppresses recombination, a fifty, ninety and ninety-nine percent selfing rate only decreases the recombination rate by a factor of two, ten, and one-hundred, respectively (Nordborg 2000). Our results predict that we should see an elevated load only in taxa that are nearly obligately selfing. Consistent with this prediction, Laenen *et al.* (2018) found elevated load in only highly selfing (~0.9 selfing rate) populations of *Arabis alpina*; with no increase in the load in mixed-mating (~0.8 selfing rate) as compared to outcrossing populations. Analytical models derived in Glémin (2007) also demonstrate that, at high recombination rates, a drop in the efficacy of selection is apparent only in highly selfing populations (selfing rate > 0.9).

The positive effects of a little outcrossing on the efficacy of selection and the dearth of fully selfing taxa may explain why empirical studies fail to find a decrease in selection efficacy in some selfing taxa. In simulations that allow outcrossing rates to evolve, small amounts of outcrossing are selected for in predominant selfers because they dissipate the extent of linkage disequilibrium created by selfing and reduce the magnitude of background selection (Kamran-Disfani and Agrawal 2014). This may explain the paucity of completely selfing taxa (Goodwillie *et al.* 2005; Moeller *et al.* 2017).

The shift from background selection to pseudo-overdominance weakens the efficacy of selection:

Gilbert *et al.* (2020) discovered that selection against deleterious recessive variants can lead to an increase in linked neutral diversity, rather than the decrease expected by background selection, when selection cannot maintain haplotypes free of a deleterious recessive alleles. In our simulations, this “transition to pseudo-overdominance”, occurs in genomic regions with low recombination rates, relative to the mutation rate. We find that the transition from background selection to pseudo-overdominance is accompanied with a large decrease in the efficacy of selection. Specifically, we find that genomic regions that experience pseudo-overdominance accumulate a large number of additive, mildly deleterious mutations, resulting in high genetic load. Although the transition to pseudo-overdominance increases population-genetic-based estimates of N_e by increasing diversity at linked neutral sites (Ohta and Kimura 1970; Gilbert *et al.* 2020), we discover that pseudo-overdominance substantially decreases the N_e that describes selection efficacy.

Thus, we propose that when there is a shift to pseudo-overdominance, the complementary haplotypes in repulsion effectively decreases the N_e that affects selection efficacy. Pseudo-overdominance is akin to negative-frequency dependent balancing selection: as haplotypes rise in frequency, it is more probable that they are exposed in the homozygous state and selection drives the haplotype’s frequency down. At equilibrium, all pseudo-overdominant haplotypes in the population should have roughly equal fitness. Similar to other genetic systems with long-term balancing selection, e.g., the self-incompatibility locus in some plants (Vekemans and Slatkin 1994) or the MHC locus in vertebrates (Garrigan and Hedrick 2003), pseudo-overdominant haplotypes form haplotype classes. Each haplotype class can accumulate its own neutral and, in our simulations, mildly deleterious additive mutations. Because the pseudo-overdominant haplotypes form in low recombination regions, there is effectively no ‘migration’ of neutral or mildly deleterious mutations between allele classes (Charlesworth *et al.* 2003; Charlesworth 2006). The consequence is that the fitness effect of an additive, mildly deleterious mutation is only relative to the fitness of other genomes of the same haplotype class. The N_e that determines selection efficacy is thus a function of the number of genomes *within* a haplotype class. Indeed, the population genetics of balancing selection suggests that the sampling of a neutral allele in two haplotype classes is akin to migration between two populations, each of size N_e, and that the expected coalescence time for a pair of neutral alleles sampled from within a haplotype class is the panmictic coalescence time of $2N_e$ (Charlesworth *et al.* 2003).

Balancing selection in other systems has consequences for genetic load that parallel those driven by pseudo-overdominance. For example, the “sheltered load” is a well-known phenomenon across classic cases of balancing selection (Mather and de Winton 1941; Glémin *et al.* 2001; van Oosterhout 2009; Jay *et al.* 2021), and refers to the accumulation of recessive mutations linked to loci under balancing selection (specifically, loci under heterozygote advantage). Pseudo-overdominance is driven by the accumulation of recessive mutations in low recombination regions, but once the shift from background selection to pseudo-overdominance has occurred, more recessive mutations accumulate on the complementary pseudo-overdominant haplotypes and become “sheltered load”. This sheltered load in pseudo-overdominant haplotypes may reinforce pseudo-overdominance. That is, as a genomic region becomes less likely to be homozygous in individuals that survive to mating, it is free to acquire additional recessive variants (Llaurens et al. 2017), which further increases the strength of selection selected against individuals homozygous in this regions. This reinforcement of pseudo-overdominance, or haplotype structuring driven by recessive mutations, can also occur in genomic regions with inversions (Berdan *et* *al.* 2021). Low initial frequency of an inversion and low recombination between heterokaryotypes leads to the accumulation of mutational load in the inversion, the division of the inversion into haplotypes with complementary recessive mutations, and the further accumulation of mutational load on those haplotypes.

We have not found evidence in the literature for high additive, weakly deleterious load in systems characterized by balancing selection. One example may be the relatively high accumulation of transposable elements at the MHC locus (van Oosterhout 2009) and in a mimicry supergene (Jay *et al.* 2021), if we assume that one copy of a transposable element at a given locus has deleterious fitness effect (Charlesworth and Charlesworth 1983). Nevertheless, our results suggest that additive genetic load linked to loci under balancing selection may accumulate to appreciable levels. The lack of empirical evidence for this may be due to the difficulty of detecting additive load in classic inbreeding depression designs. Alternative methods, such as dN/dS ratios or inference of the DFE, applied across the genome could help determine whether regions known to experience long term balancing selection have elevated genetic additive load.

Selfing prevents the shift from background selection to pseudo-overdominance:

Our study extends Gilbert *et al.'s* (2020) study of pseudo-overdominance to partially selfing populations. Whether or not partially selfing populations experience pseudo-overdominance at a given set of parameter combinations is a function of whether or not partial selfing purges the recessive load. The analytical theory derived here qualitatively matches our simulation results of 1) how selfing purges the recessive load and thus prevents a shift from background selection to pseudo-overdominance, and 2) how recombination amplifies the effects of partial selfing on preventing the transition to pseudo-overdominance. Overall, we find that as the selfing rate increases, there is a sharp decrease in the parameter space that leads to pseudo-overdominance (Figure 2E). At a given partial selfing rate (i.e., selfing rate < 0.5), pseudo-overdominance becomes more likely when U_{del} is high and $s_{recessive}$ is low, as these are parameter combinations where it is harder to purge the recessive load (Wang *et al.* 1999). As populations experience a greater influx of deleterious recessive mutations, a higher selfing rate is needed to purge the recessive load before it accumulates into complementary haplotypes and causes pseudo-overdominance.

Future directions

Extensions to realistic genomic architecture:

For simplicity, we modelled homogeneous recombination and deleterious mutation rate across all chromosomes in a genome, for each set of parameters, rather than modelling heterogeneity in these features across the genome. Future studies could explicitly model whole genomes with realistic genomic structure. Nevertheless, for now we can use our results to make qualitative predictions of how patterns of selection efficacy may vary both across the genome within individuals and for different species by assuming that variation across parameter values for a simulation translates into predictions for variability across the genome.

Our results show that the recombination rate relative to the deleterious mutation rate is the main factor determining how genetic load changes with selfing rate. Recombination rates vary dramatically across the genome, which affects linked selection (Gaut *et al.* 2007). For example, genomic regions with low recombination should experience more linked selection and lowered selection efficacy than high recombination regions (Smukowski and Noor 2011). This prediction assumes that the intensity of selection is equal across the genome, when in fact, organisms can have either positive or negative covariances between gene density (a proxy for selection intensity) and recombination rate (Slotte 2014). Thus, our work reinforces the need to consider the recombination rate relative to the deleterious mutation rate (or some proxy for it like gene density, e.g., see McVicker *et al.* 2009), when considering the effects of background selection.

Organisms with a positive covariance between recombination rate and gene density (e.g., *Mimulus guttatus* (Aeschbacher *et al.* 2017), maize (Anderson *et al.* 2006), rice (International Rice Genome Sequencing Project, Takuji Sasaki 2005), wheat (Dvorak *et al.* 2004), *A. thaliana* (Wright *et al.* 2003; Giraut *et al.* 2011), and *Populus* species (Wang *et al.* 2016; Apuli *et al.* 2020)) should primarily experience selection in high recombination regions. Our results predict that the genomic factors associated with shifts to selfing would cause a reduction in selection efficacy in these organisms given a shift to full selfing. Conversely, organisms with a negative covariance between recombination rate and gene density (e.g., *Caenorhabditis* (Barnes *et al.* 1995), and *Mimulus aurantiacus* (Stankowski *et al.* 2019)) should more often experience selection in low recombination environments. In these organisms, we would predict a shift to selfing to either have no effect on or increase selection efficacy, depending on the extent to which pseudo-overdominance occurs in the genome. This idea may explain the results of an experimental evolution on *C. elegans* where predominant selfers had higher fertility (i.e., lower genetic load) in addition to lower inbreeding depression, and lower neutral diversity than predominant outcrossers in an experimental evolution study (Chelo *et al.* 2019). While theory would predict higher neutral diversity and a higher recessive deleterious load in the outcrosser than in the selfer, these two lines of evidence, along with the higher load due to mildly deleterious additive mutations, are consistent with pseudo-overdominance.

The distribution of dominance and fitness effects:

To understand how linked selection against additive and recessive mutations shape the genetic load, we focused on mutations with fixed effects and those that span the boundary of near neutrality. By contrast, evolution does not pick such parameters - rather mutations take selective and dominance coefficients from a two-dimensional density function. There are numerous hurdles to estimating this function, especially in eukaryotes. The best methods to infer the distribution of fitness effects from polymorphism data (Keightley and Eyre-Walker 2007) provide only crude estimates of the distribution of fitness effects and cannot tell us about dominance. This joint distribution of dominance and fitness effects is even more poorly characterized; however, two of the best studies on the topic show that more recessive mutations are more deleterious (Agrawal and Whitlock 2011; Huber *et al.* 2018)

Our simulations capture the spirit of this result -- we modeled recessive mutations with selection coefficients much larger than $4Ns$, and additive mutations with selection coefficients that are closer to this boundary. However, we did not attempt to model continuous variation in either variable nor covariation between them. We interpret our result of “no effect” or “some effect” of mating system on the load as a qualitative summary that answers the question -- does the mating system shift incoming mutations to one side of the near neutrality boundary or the other? As quantitative variation in the distribution of fitness effects could cover a greater amount of fitness effect space (depending on dispersion), it could potentially reveal subtle effects missed by our model. For example, it seems likely that a broad distribution of selection coefficients would reveal an effect of the selfing rate on the load for some combinations of mutation and recombination rates which showed no effect in our study. Consider, for example, a case for which we found no effect of the mating system on the load because both selfers and outcrossers could effectively remove their additive load. It seems likely that there is some modest selection coefficient here for which outcrossers could more effectively remove their additive load than could selfers, but that our parameter choices missed this “Goldilocks” selection coefficient.

Demographic history:

We find that genomic factors associated with selfing, particularly the relative effects of linked selection across selfing rates, can result in various outcomes for patterns of selection efficacy. As noted above, this could contribute to the contrasting empirical evidence for how shifts to selfing affect selection efficacy. However, other factors affecting selection efficacy can be associated with mating system shifts. Perhaps the most influential of these factors is differences in demography. Indeed, selfing is often associated with colonization of and rapid expansion in islands, disturbed, or other marginal habitats (Baker 1955), further decreasing effective population size of and efficacy of selection in selfers. Similarly, as many selfers have experienced more rapid post-glaciation range expansion than outcrossers (Grossenbacher *et al.* 2015), selfers may suffer a more severe expansion load (Peischl *et al.* 2013, 2015) than outcrossers. Again however, this expansion load is more influenced by recessive than by additive mutations (Peischl and Excoffier 2015), and therefore likely has limited application to selfers. Finally, the impact of demographic changes likely differs in selfers and outcrossers -- population contraction in outcrossers exposes and purges previously rare recessive deleterious mutations (Kirkpatrick and Jarne 2000; Balick *et al.* 2015). Thus, recurrent bottlenecks could facilitate purging of the recessive load in outcrossers, plausibly preventing the transition to pseudo-overdominance. Conversely, recurrent bottlenecks have little impact on the additive, mildly deleterious load.

Finally, it is possible that selfers and outcrossers have different census sizes. In plants, selfing taxa are commonly associated with marginal and stochastic environments (Baker 1955). This association may be driven by the stochasticity of pollinator population dynamics in marginal habitats or the temporal fluctuations in plant census size that cause strong pollen limitation, both of which select for selfing through reproductive assurance (Kalisz *et al.* 2004; Moeller and Geber 2005). Similarly because selfers are freed from "S-Allee Effects" (Wagenius *et al.* 2007), and because a purged load can help them avoid the extinction vortex (Gilpin and Soulé 1986), selfers may persist longer at smaller census sizes than outcrossers. Thus, demographic history may affect the efficacy of selection in selfers versus outcrossers in different ways, confounding patterns of the efficacy of selection between sister taxa that vary in mating system.

Positive selection:

We focused on how purifying selection at one site impacts the efficacy of purifying selection at a linked site. However the effects of positive selection, too, are extended over long genetic distances in selfers, occasionally covering nearly entire chromosomes (Andersen *et al.* 2012), and sweeps increase the probability of fixation of linked deleterious alleles in selfing vs outcrossing populations (Hartfield *et al.* 2017). Like a range expansion or bottleneck, such sweeps will have a more modest effect on additive than recessive mutations; however, ongoing recurrent sweeps could prevent the removal of linked deleterious mutations.

Moreover, beneficial mutations linked to a recessive deleterious mutation have a lower probability to sweep through the population as compared to beneficial mutations in regions where recombination rates are high (Assaf *et al.* 2015). As positive selection increases the frequency of the haplotype containing the beneficial and recessive mutation, it also increases the probability that the haplotype is homozygous for the deleterious allele. The exposed fitness effect of the recessive mutation can slow down or completely prevent a selective sweep, depending on the net fitness effects of the advantageous and deleterious mutation, as well as the recombination rate. Fixation of adaptive mutations should thus be particularly thwarted in regions with pseudo-overdominance, because every new beneficial mutation in such a region will fall on a haplotype with one or several recessive deleterious mutations that are maintained at intermediate frequency to minimize mutation load. Any deviation from this equilibrium will lead to an increase in the expressed recessive load and hence counter both the purging of new deleterious mutations as well as sweeps of beneficial mutations. Our analytical model confirms this intuition and shows that pseudo-overdominance decreases the rate of evolution for any co-dominant mutation, regardless of whether it is beneficial or deleterious.

Conclusions

Our simulation study and analytical models suggest that patterns of the efficacy of selection among populations that vary in mating system can be confounded by genomic factors; in particular, the recombination rate relative to the mutation rate determines whether the efficacy of selection decreases, remains constant, or increases with selfing rate. Outcrossers and populations with mixed mating systems have more efficacious selection and lower genetic load that predominant selfers in high recombination environments, highlighting the power of a little amount of outcrossing in preventing genetic load. The transition from background selection to pseudo-overdominance in low recombination environments occurs under a broad range of parameter space, particularly when the fitness effects of recessive mutations are low, and flips the relationship between mating system and the efficacy of selection. We find that pseudo-overdominance reduces the efficacy of selection against additive mildly deleterious mutations, and that a threshold level of selfing prevents pseudo-overdominance by purging the recessive load. Taken together, genomic factors associated with shifts to selfing do not predictably decrease selection efficacy either, as much classic theory predicts, and thus may also contribute to the ambiguity we see in empirical data. Future work is needed to investigate the reaches of pseudo-overdominance in genomes that vary in recombination rates and gene density, in order to better understand the limits to natural selection in natural populations.

ACKNOWLEDGMENTS

We would like to thank Stephen Wright, Kimberly Gilbert, and members of the Brandvain and McGaugh labs at the University of Minnesota for comments on early drafts. Computing resources were provided through the Minnesota Supercomputing Institute at the University of Minnesota.

FUNDING

This project was funded by NSF DEB #1754246 awarded to DM and YB.

AUTHOR CONTRIBUTIONS

SAS and YB conceived the project. SP derived the analytical models. SAS wrote the first manuscript draft, and SAS, SP, DM, and YB contributed to the final version.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Aeschbacher S., J. P. Selby, J. H. Willis, and G. Coop, 2017 Population-genomic inference of the strength and timing of selection against gene flow. Proc. Natl. Acad. Sci. USA 114: 7061–7066. https://doi.org/10.1073/pnas.1616755114

Agrawal A. F., and M. C. Whitlock, 2011 Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics 187: 553–566. https://doi.org/10.1534/genetics.110.124560

Andersen E. C., J. P. Gerke, J. A. Shapiro, J. R. Crissman, R. Ghosh, *et al.*, 2012 Chromosome-scale selective sweeps shape Caenorhabditis elegans genomic diversity. Nat. Genet. 44: 285–290. https://doi.org/10.1038/ng.1050

Anderson L. K., A. Lai, S. M. Stack, C. Rizzon, and B. S. Gaut, 2006 Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. Genome Res. 16: 115–122. https://doi.org/10.1101/gr.4249906

Apuli R.-P., C. Bernhardsson, B. Schiffthaler, K. M. Robinson, S. Jansson, *et al.*, 2020 Inferring the genomic landscape of recombination rate variation in European Aspen (Populus tremula). G3: Genes, Genomes, Genetics 10: 299–309. https://doi.org/10.1534/g3.119.400504

Arunkumar R., R. W. Ness, S. I. Wright, and S. C. H. Barrett, 2015 The evolution of selfing is accompanied by reduced efficacy of selection and purging of deleterious mutations. Genetics 199: 817–829. https://doi.org/10.1534/genetics.114.172809

Assaf Z. J., D. A. Petrov, and J. R. Blundell, 2015 Obstruction of adaptation in diploids by recessive, strongly deleterious alleles. Proc. Natl. Acad. Sci. USA 112: E2658–66. https://doi.org/10.1073/pnas.1424949112

Baker H. G., 1955 Self-compatibility and establishment after “long-distance” dispersal. Evolution 9: 347. https://doi.org/10.2307/2405656

Balick D. J., R. Do, C. A. Cassa, D. Reich, and S. R. Sunyaev, 2015 Dominance of deleterious alleles controls the response to a population bottleneck. PLoS Genet. 11: e1005436. https://doi.org/10.1371/journal.pgen.1005436

Barnes T. M., Y. Kohara, A. Coulson, and S. Hekimi, 1995 Meiotic recombination, noncoding DNA and genomic organization in Caenorhabditis elegans. Genetics 141: 159–179. https://doi.org/10.1093/genetics/141.1.159

Barrett S. C. H., R. Arunkumar, and S. I. Wright, 2014 The demography and population genomics of evolutionary transitions to self-fertilization in plants. Philos. Trans. R. Soc. Lond. B Biol. Sci. 369. https://doi.org/10.1098/rstb.2013.0344

Bataillon T., and M. Kirkpatrick, 2000 Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. Genet. Res. 75: 75–81. https://doi.org/10.1017/S0016672399004048

Berdan E. L., A. Blanckaert, R. K. Butlin, and C. Bank, 2021 Deleterious mutation accumulation and the long-term fate of chromosomal inversions. PLoS Genet. 17: e1009411. https://doi.org/10.1371/journal.pgen.1009411

Brandvain Y., T. Slotte, K. M. Hazzouri, S. I. Wright, and G. Coop, 2013 Genomic identification of founding haplotypes reveals the history of the selfing species Capsella rubella. PLoS Genet. 9: e1003754. https://doi.org/10.1371/journal.pgen.1003754

Brandvain Y., A. M. Kenney, L. Flagel, G. Coop, and A. L. Sweigart, 2014 Speciation and introgression between Mimulus nasutus and Mimulus guttatus. PLoS Genet. 10: e1004410. https://doi.org/10.1371/journal.pgen.1004410

Brandvain Y., and S. I. Wright, 2016 The limits of natural selection in a nonequilibrium world. Trends Genet. 32: 201–210. https://doi.org/10.1016/j.tig.2016.01.004

Charlesworth B., and D. Charlesworth, 1983 The population dynamics of transposable elements. Genet. Res. 42: 1–27. https://doi.org/10.1017/S0016672300021455

Charlesworth B., 1992 Evolutionary rates in partially self-fertilizing species. Am. Nat. 140: 126–148. https://doi.org/10.1086/285406

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

Charlesworth D., M. T. Morgan, and B. Charlesworth, 1993b Mutation accumulation in finite populations. J. Hered. 84: 321–325. https://doi.org/10.1093/oxfordjournals.jhered.a111351

Charlesworth B., 1994 The effect of background selection against deleterious mutations on weakly selected, linked variants. Genet. Res. 63: 213–227. https://doi.org/10.1017/s0016672300032365

Charlesworth D., and S. I. Wright, 2001 Breeding systems and genome evolution. Curr. Opin. Genet. 11: 685–690. https://doi.org/10.1016/S0959-437X(00)00254-9

Charlesworth B., D. Charlesworth, and N. H. Barton, 2003 The effects of genetic and geographic structure on neutral variation. Annu. Rev. Ecol. Evol. Syst. 34: 99–125. https://doi.org/10.1146/annurev.ecolsys.34.011802.132359

Charlesworth D., 2006 Balancing selection and its effects on sequences in nearby genome regions. PLoS Genet. 2: e64. https://doi.org/10.1371/journal.pgen.0020064

Chelo I. M., B. Afonso, S. Carvalho, I. Theologidis, C. Goy, *et al.*, 2019 Partial selfing can reduce genetic loads while maintaining diversity during experimental evolution. G3: Genes, Genomes, Genetics 9: 2811–2821. https://doi.org/10.1534/g3.119.400239

Crnokrak P., and S. C. H. Barrett, 2002 Perspective: purging the genetic load: a review of the experimental evidence. Evolution 56: 2347–2358. https://doi.org/10.1111/j.0014-3820.2002.tb00160.x

Cutter A. D., and B. A. Payseur, 2003 Rates of deleterious mutation and the evolution of sex in Caenorhabditis. J. Evol. Biol. 16: 812–822. https://doi.org/10.1046/j.1420-9101.2003.00596.x

Do R., D. Balick, H. Li, I. Adzhubei, S. Sunyaev, *et al.*, 2015 No evidence that selection has been less effective at removing deleterious mutations in Europeans than in Africans. Nat. Genet. 47: 126–131. https://doi.org/10.1038/ng.3186

Dvorak J., Z.-L. Yang, F. M. You, and M.-C. Luo, 2004 Deletion polymorphism in wheat chromosome regions with contrasting recombination rates. Genetics 168: 1665–1675. https://doi.org/10.1534/genetics.103.024927

Escobar J. S., A. Cenci, J. Bolognini, A. Haudry, S. Laurent, *et al.*, 2010 An integrative test of the dead-end hypothesis of selfing evolution in Triticeae (Poaceae). Evolution 64: 2855–2872. https://doi.org/ https://doi.org/10.1111/j.1558-5646.2010.01045.x

Foxe J. P., T. Slotte, E. A. Stahl, B. Neuffer, H. Hurka, *et al.*, 2009 Recent speciation associated with the evolution of selfing in Capsella. Proc. Natl. Acad. Sci. USA 106: 5241–5245. https://doi.org/10.1073/pnas.0807679106

Garrigan D., and P. W. Hedrick, 2003 Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. Evolution 57: 1707–1722. https://doi.org/10.1111/j.0014-3820.2003.tb00580.x

Gaut B. S., S. I. Wright, C. Rizzon, J. Dvorak, and L. K. Anderson, 2007 Recombination: an underappreciated factor in the evolution of plant genomes. Nat. Rev. Genet. 8: 77–84. https://doi.org/10.1038/nrg1970

Gilbert K. J., F. Pouyet, L. Excoffier, and S. Peischl, 2020 Transition from background selection to associative overdominance promotes diversity in regions of low recombination. Curr. Biol. 30: 101–107.e3. https://doi.org/10.1016/j.cub.2019.11.063

Gilpin M. E., and M. E. Soulé, 1986 Minimal viable populations: processes of species extinction, pp. 19–34 in *Conservation Biology: the science of scarcity and diversity*, edited by Soulé M. E. Sinauer Associates, Sunderland, Massachusetts.

Gioti A., J. E. Stajich, and H. Johannesson, 2013 Neurospora and the dead-end hypothesis: genomic consequences of selfing in the model genus. Evolution 67: 3600–3616. https://doi.org/10.1111/evo.12206

Giraut L., M. Falque, J. Drouaud, L. Pereira, O. C. Martin, *et al.*, 2011 Genome-wide crossover distribution in Arabidopsis thaliana meiosis reveals sex-specific patterns along chromosomes. PLoS Genet. 7: e1002354. https://doi.org/10.1371/journal.pgen.1002354

Glémin S., T. Bataillon, J. Ronfort, A. Mignot, and I. Olivieri, 2001 Inbreeding depression in small populations of self-incompatible plants. Genetics 159: 1217–1229. https://doi.org/10.1093/genetics/159.3.1217

Glémin S., E. Bazin, and D. Charlesworth, 2006 Impact of mating systems on patterns of sequence polymorphism in flowering plants. Proc. Biol. Sci. 273: 3011–3019. https://doi.org/10.1098/rspb.2006.3657

Glémin S., 2007 Mating systems and the efficacy of selection at the molecular level. Genetics 177: 905–916. https://doi.org/10.1534/genetics.107.073601

Glémin S., and N. Galtier, 2012 Genome evolution in outcrossing versus selfing versus asexual species, pp. 311–335 in *Evolutionary Genomics. Methods in Molecular Biology (Methods and Protocols)*, edited by Anisimova M. Human Press, Totowa, NJ.

Glémin S., and J. Ronfort, 2013 Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. Evolution 67: 225–240. https://doi.org/10.1111/j.1558-5646.2012.01778.x

Goldberg E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith, *et al.*, 2010 Species selection maintains self-incompatibility. Science 330: 493–495. https://doi.org/10.1126/science.1194513

Goodwillie C., S. Kalisz, and C. G. Eckert, 2005 The evolutionary enigma of mixed mating systems in plants: Occurrence, theoretical explanations, and empirical evidence. Annu Rev Ecol Evol Syst 36: 47–79. https://doi.org/10.1146/annurev.ecolsys.36.091704.175539

Grossenbacher D., R. Briscoe Runquist, E. E. Goldberg, and Y. Brandvain, 2015 Geographic range size is predicted by plant mating system. Ecol. Lett. 18: 706–713. https://doi.org/10.1111/ele.12449

Haag-Liautard C., M. Dorris, X. Maside, S. Macaskill, D. L. Halligan, *et al.*, 2007 Direct estimation of per nucleotide and genomic deleterious mutation rates in Drosophila. Nature 445: 82–85. https://doi.org/10.1038/nature05388

Haller B. C., J. Galloway, and J. Kelleher, 2019 Tree-sequence recording in SLiM opens new horizons for forward-time simulation of whole genomes. Mol. Ecol. 19: 552–566. https://doi.org/10.1111/1755-0998.12968

Haller B. C., and P. W. Messer, 2019 SLiM 3: Forward genetic simulations beyond the Wright-Fisher model. Mol. Biol. Evol. 36: 632–637. https://doi.org/10.1093/molbev/msy228

Hanschen E. R., M. D. Herron, J. J. Wiens, H. Nozaki, and R. E. Michod, 2018 Repeated evolution and reversibility of self-fertilization in the volvocine green algae. Evolution 72: 386–398. https://doi.org/10.1111/evo.13394

Hartfield M., T. Bataillon, and S. Glémin, 2017 The evolutionary interplay between adaptation and self-fertilization. Trends Genet. 33: 420–431. https://doi.org/10.1016/j.tig.2017.04.002

Haudry A., A. Cenci, C. Guilhaumon, E. Paux, S. Poirier, *et al.*, 2008 Mating system and recombination affect molecular evolution in four Triticeae species. Genet. Res. 90: 97–109. https://doi.org/10.1017/S0016672307009032

Hazzouri K. M., J. S. Escobar, R. W. Ness, L. Killian Newman, A. M. Randle, *et al.*, 2013 Comparative population genomics in Collinsia sister species reveals evidence for reduced effective population size, relaxed selection, and evolution of biased gene conversion with an ongoing mating system shift. Evolution 67: 1263–1278. https://doi.org/10.1111/evo.12027

Huber C. D., A. Durvasula, A. M. Hancock, and K. E. Lohmueller, 2018 Gene expression drives the evolution of dominance. Nat. Commun. 9: 2750. https://doi.org/10.1038/s41467-018-05281-7

Igic B., R. Lande, and J. R. Kohn, 2008 Loss of self-incompatibility and its evolutionary consequences. Int. J. Plant Sci. 169: 93–104. https://doi.org/10.1086/523362

Igic B., and J. W. Busch, 2013 Is self-fertilization an evolutionary dead end? New Phytol. 198: 386–397. https://doi.org/10.1111/nph.12182

Ingvarsson P. K., 2002 A metapopulation perspective on genetic diversity and differentiation in partially self-fertilizing plants. Evolution 56: 2368–2373. https://doi.org/10.1111/j.0014-3820.2002.tb00162.x

International Rice Genome Sequencing Project, Takuji Sasaki, 2005 The map-based sequence of the rice genome. Nature 436: 793–800. https://doi.org/10.1038/nature03895

Jarne P., and J. R. Auld, 2006 Animals mix it up too: The distribution of self-fertilization among hermaphroditic animals. Evolution 60: 1816–1824. https://doi.org/10.1111/j.0014-3820.2006.tb00525.x

Jay P., M. Chouteau, A. Whibley, H. Bastide, H. Parrinello, *et al.*, 2021 Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. Nat. Genet. 53: 288–293. https://doi.org/10.1038/s41588-020-00771-1

Kalisz S., D. W. Vogler, and K. M. Hanley, 2004 Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. Nature 430: 884–887. https://doi.org/10.1038/nature02776

Kamran-Disfani A., and A. F. Agrawal, 2014 Selfing, adaptation and background selection in finite populations. J. Evol. Biol. 27: 1360–1371. https://doi.org/10.1111/jeb.12343

Keightley P. D., and A. Eyre-Walker, 2007 Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. Genetics 177: 2251–2261. https://doi.org/10.1534/genetics.107.080663

Kelleher J., A. M. Etheridge, and G. McVean, 2016 Efficient coalescent simulation and genealogical analysis for large sample sizes. PLoS Comput. Biol. 12: e1004842. https://doi.org/10.1371/journal.pcbi.1004842

Kim B. Y., C. D. Huber, and K. E. Lohmueller, 2018 Deleterious variation shapes the genomic landscape of introgression. PLoS Genet. 14: e1007741. https://doi.org/10.1371/journal.pgen.1007741

Kimura M., 1983 *The Neutral Theory of Molecular Evolution*. Cambridge University Press, Cambridge.

Kirkpatrick M., and P. Jarne, 2000 The effects of a bottleneck on inbreeding depression and the genetic load. Am. Nat. 155: 154–167. https://doi.org/10.1086/303312

Koelling V. A., J. L. Hamrick, and R. Mauricio, 2011 Genetic diversity and structure in two species of Leavenworthia with self-incompatible and self-compatible populations. Heredity 106: 310–318. https://doi.org/10.1038/hdy.2010.59

Kondrashov A. S., 1995 Contamination of the genome by very slightly deleterious mutations: why have we not died 100 times over? J. Theor. Biol. 175: 583–594. https://doi.org/10.1006/jtbi.1995.0167

Laenen B., A. Tedder, M. D. Nowak, P. Toräng, J. Wunder, *et al.*, 2018 Demography and mating system shape the genome-wide impact of purifying selection in Arabis alpina. Proc. Natl. Acad. Sci. USA 115: 816–821. https://doi.org/10.1073/pnas.1707492115

Llaurens V., A. Whibley, and M. Joron, 2017 Genetic architecture and balancing selection: the life and death of differentiated variants. Mol. Ecol. 26: 2430–2448. https://doi.org/10.1111/mec.14051

Lohmueller K. E., 2014 The distribution of deleterious genetic variation in human populations. Curr. Opin. Genet. Dev. 29: 139–146. https://doi.org/10.1016/j.gde.2014.09.005

Lynch M., J. Conery, and R. Burger, 1995 Mutation accumulation and the extinction of small populations. Am. Nat. 146: 489–518. https://doi.org/10.1086/285812

Lynch M., 2010 Evolution of the mutation rate. Trends Genet. 26: 345–352. https://doi.org/10.1016/j.tig.2010.05.003

Mather K., and D. de Winton, 1941 Adaptation and counter-adaptation of the breeding system in Primula. Ann. Bot. 5: 297–311.

McVicker G., D. Gordon, C. Davis, and P. Green, 2009 Widespread genomic signatures of natural selection in hominid evolution. PLoS Genet. 5: e1000471. https://doi.org/10.1371/journal.pgen.1000471

Moeller D. A., and M. A. Geber, 2005 Ecological context of the evolution of self-pollination in Clarkia xantiana: population size, plant communities, and reproductive assurance. Evolution 59: 786–799. https://doi.org/10.1111/j.0014-3820.2005.tb01753.x

Moeller D. A., R. D. Briscoe Runquist, A. M. Moe, M. A. Geber, C. Goodwillie, *et al.*, 2017 Global biogeography of mating system variation in seed plants. Ecol. Lett. 20: 375–384. https://doi.org/10.1111/ele.12738

Ness R. W., M. Siol, and S. C. H. Barrett, 2012 Genomic consequences of transitions from cross- to self-fertilization on the efficacy of selection in three independently derived selfing plants. BMC Genomics 13: 611. https://doi.org/10.1186/1471-2164-13-611

Nordborg M., and P. Donnelly, 1997 The coalescent process with selfing. Genetics 146: 1185–1195. https://doi.org/10.1093/genetics/146.3.1185

Nordborg M., 2000 Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. Genetics 154: 923–929. https://doi.org/10.1093/genetics/154.2.923

Ohta T., and M. Kimura, 1970 Development of associative overdominance through linkage disequilibrium in finite populations. Genet. Res. 16: 165–177. https://doi.org/10.1017/s0016672300002391

Oosterhout C. van, 2009 A new theory of MHC evolution: beyond selection on the immune genes. Proc. Royal Soc. B. 276: 657–665. https://doi.org/10.1098/rspb.2008.1299

Peischl S., I. Dupanloup, M. Kirkpatrick, and L. Excoffier, 2013 On the accumulation of deleterious mutations during range expansions. Mol. Ecol. 22: 5972–5982. https://doi.org/10.1111/mec.12524

Peischl S., and L. Excoffier, 2015 Expansion load: recessive mutations and the role of standing genetic variation. Mol. Ecol. 24: 2084–2094. https://doi.org/10.1111/mec.13154

Peischl S., M. Kirkpatrick, and L. Excoffier, 2015 Expansion load and the evolutionary dynamics of a species range. Am. Nat. 185: E81–93. https://doi.org/10.1086/680220

Pettengill J. B., R. D. Briscoe Runquist, and D. A. Moeller, 2016 Mating system divergence affects the distribution of sequence diversity within and among populations of recently diverged subspecies of Clarkia xantiana (Onagraceae). Am. J. Bot. 103: 99–109. https://doi.org/ https://doi.org/10.3732/ajb.1500147

Qiu S., K. Zeng, T. Slotte, S. Wright, and D. Charlesworth, 2011 Reduced efficacy of natural selection on codon usage bias in selfing Arabidopsis and Capsella species. Genome Biology and Evolution 3: 868–880. https://doi.org/10.1093/gbe/evr085

Roze D., and F. Rousset, 2004 Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. Genetics 167: 1001–1015. https://doi.org/10.1534/genetics.103.025148

Roze D., 2016 Background selection in partially selfing populations. Genetics 203: 937–957.

https://doi.org/10.1534/genetics.116.187955

Sankararaman S., S. Mallick, M. Dannemann, K. Prüfer, J. Kelso, *et al.*, 2014 The genomic landscape of Neanderthal ancestry in present-day humans. Nature 507: 354–357. https://doi.org/10.1038/nature12961

Schultz S. T., M. Lynch, and J. H. Willis, 1999 Spontaneous deleterious mutation in Arabidopsis thaliana. Proc. Natl. Acad. Sci. USA 96: 11393–11398. https://doi.org/10.1073/pnas.96.20.11393

Slotte T., J. P. Foxe, K. M. Hazzouri, and S. I. Wright, 2010 Genome-wide evidence for efficient positive and purifying selection in Capsella grandiflora, a plant species with a large effective population size. Mol. Biol. Evol. 27: 1813–1821. https://doi.org/10.1093/molbev/msq062

Slotte T., K. M. Hazzouri, J. A. Ågren, D. Koenig, F. Maumus, *et al.*, 2013 The Capsella rubella genome and the genomic consequences of rapid mating system evolution. Nat. Genet. 45: 831–835. https://doi.org/10.1038/ng.2669

Slotte T., 2014 The impact of linked selection on plant genomic variation. Brief. Funct. Genomics 13: 268–275. https://doi.org/10.1093/bfgp/elu009

Smukowski C. S., and M. A. F. Noor, 2011 Recombination rate variation in closely related species. Heredity 107: 496–508. https://doi.org/10.1038/hdy.2011.44

Stankowski S., M. A. Chase, A. M. Fuiten, M. F. Rodrigues, P. L. Ralph, *et al.*, 2019 Widespread selection and gene flow shape the genomic landscape during a radiation of monkeyflowers. PLoS Biol. 17: e3000391. https://doi.org/10.1371/journal.pbio.3000391

Uyenoyama M. K., and D. M. Waller, 1991 Coevolution of self-fertilization and inbreeding depression. I. Mutation-selection balance at one and two loci. Theor. Popul. Biol. 40: 14–46. https://doi.org/10.1016/0040-5809(91)90045-h

Vekemans X., and M. Slatkin, 1994 Gene and allelic genealogies at a gametophytic self-incompatibility locus. Genetics 137: 1157–1165.

Wagenius S., E. Lonsdorf, and C. Neuhauser, 2007 Patch aging and the S-Allee effect: Breeding system effects on the demographic response of plants to habitat fragmentation. Am. Nat. 169: 383–397. https://doi.org/10.1086/511313

Waller D. M., 2021 Addressing Darwin's dilemma: Can pseudo-overdominance explain persistent inbreeding depression and load? Evolution 75: 779–793. https://doi.org/ https://doi.org/10.1111/evo.14189

Wang J., W. G. Hill, D. Charlesworth, and B. Charlesworth, 1999 Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. Genet. Res. 74: 165–178. https://doi.org/10.1017/s0016672399003900

Wang J., N. R. Street, D. G. Scofield, and P. K. Ingvarsson, 2016 Natural selection and recombination rate variation shape nucleotide polymorphism across the genomes of three related Populus species. Genetics 202: 1185–1200. https://doi.org/10.1534/genetics.115.183152

Willis J. H., 1999 Inbreeding load, average dominance and the mutation rate for mildly deleterious alleles in Mimulus guttatus. Genetics 153: 1885–1898. https://doi.org/10.1093/genetics/153.4.1885

Wright S. I., N. Agrawal, and T. E. Bureau, 2003 Effects of recombination rate and gene density on transposable element distributions in Arabidopsis thaliana. Genome Res. 13: 1897–1903. https://doi.org/10.1101/gr.1281503

Wright S. I., R. W. Ness, J. P. Foxe, and S. C. H. Barrett, 2008 Genomic consequences of outcrossing and selfing in plants. Int. J. Plant Sci. 169: 105–118. https://doi.org/10.1086/523366

Appendix for "Genetic load may increase or decrease with selfing depending upon the recombination environment."

Sianta, S., Peischl, S., Moeller, D., and Y. Brandvain

A model for mutation load in partially selfing populations

Genotype Frequencies with Selfing

We extend the multi-locus model from Gilbert et al. (2020) to include selfing. We consider n biallelic loci. We denote wild-type and derived alleles alleles at locus $i = 1., \ldots, n$ by a_i and A_i, respectively. Derived alleles are deleterious and fully recessive with fitness $1 - s$ when homozygous and 1 otherwise. There is no epistasis and we assume multiplicative effects across loci.

For fully linked loci (that is, there is no recombination) the model is equivalent to a single locus model with 2^n alleles where each haplotype corresponds to a single allele. As in Gilbert et al. (2020), we want to derive the frequency of the 0-mutation haplotpye p_0 at mutation selection balance. When this frequency approaches 0, we expect a transition to pseudo-overdominance (POD). We assume that haplotypes with more than one deleterious mutations are vanishingly rare at equilibrium. This has the important consequence that no genotype can be derived homozygous at more than one locus, which simplifies calcualtions considerably. Let p_i, $i = 1, \ldots, n$ be the frequency of the haplotype carrying a derived allele at locus i, and p_0 the frequency of the haplotype without any deleterious mutations.

With these haplotypes, we only need to follow four different multi genotypic categories (Caballero and Hill 1992):

1. A genotype without any deleterious mutations. This genotype is found at frequency $p_{00} = F \sum_{i=0}^{n} p_0 p_i + (1-F) p_0^2$.
2. A genotype heterozygous at one locus. This genotype is found at frequency $p_{01}^{(1)} = 2(1 - F) \sum_{i=1}^{n} p_0 p_i$.
3. A genotype heterozygous at two loci. This genotype is found at frequency $p_{01}^{(2)} = (1 - F) \sum_{i=1}^{n} \left(\sum_{j=i+1}^{n} 2 p_i p_j \right)$. And
4. Genotypes that are derived homoyzgous at one locus $p_{11} = \sum_{i=1}^{n} \left(F \sum_{j=0}^{n} p_i p_j + (1-F) p_i^2 \right)$.

Fitness of these genotype-classes is given by $w_{00} = w_{01}^{(1)} = w_{01}^{(2)} = 1$ and $w_{11} = 1 - s$. Mean fitness is given by $\overline{w} = 1 - sp_{11}$. Assuming continuous time the change in genotype frequencies after selection are then given by $dp_g/d_t = p_g(w_g - \overline{w})$, $g \in \{00, {}_{01}^{(1)}, {}_{01}^{(2)}, 11\}$. The parameter F is equivalent to $\alpha/(2-\alpha)$ where α is the selfing rate (Caballero and Hill, 1992).

Evolutionary Dynamics of p_0 with recombination

At equilibrium all haplotypes that carry a deleterious allele must have the same frequency, which we denote as p_1. That is, from now on we simply use p_1 to refer to the frequency of any haplotype carrying exactly one derived allele. At equilibrium $p_1 = (1 - p_0)/n$ such that haplotype frequencies sum up to 1. Since we are only interested in calculating haplotype frequencies at mutation-selection equilibrium we henceforth assume that $p_1 = (1 - p_0)/n$. Solving

$$\frac{dp_0}{d_t} = 0$$

then yields the equilibrium frequency of the zero-mutation haplotype $\hat{p}_0$.

We next assume that loci are equidistantly distributed over a region of length r cM, such that the recombination rate between adjacent loci is $r/(n-1)$. We assume that recombination is sufficiently rare such that we can ignore double recombination events. Since we ignore haplotypes with more than one deleterious allele, the main effect of recombination is to restore haplotypes without any deleterious alleles from genotypes heterozygous at two loci. Analogous to Gilbert et al. (2020) the equation for the equilibrium haplotype frequency then becomes:

$$\frac{dp_0}{d_t} = p_{00}(w_{00} - \overline{w}) - nup_0 + \frac{1}{2}p_{01}(w_{01} - \overline{w}) + \sum_{i=1}^{n}\left[\sum_{j=i}^{n}\left((1-F)\frac{r(j-i)}{(n-1)}p_i p_j + F\frac{r(j-i)}{(n-1)}\frac{p_{11}}{n}\right)\right] = 0.$$

Here, u is the per locus mutation rate and $\frac{r(j-i)}{(n-1)}$ is the recombination rate between loci i and j.

Special cases

Single Locus, No Selfing

For a single locus without selfing ($n = 1, F = 0$) we get

$$\frac{dp_0}{d_t} = p_0((p_0 - 1)^2 s - u)$$

and $\hat{p} = 1 - \sqrt{u/s}$, recovering the well-known result for mutation-selection balance of recessive alleles. For n loci, $\hat{p} = 1 - n\sqrt{u/s}$, in agreement with previous results (Gilbert et al., 2020).

No Recombination

For partially or fully selfing organisms without recombination we get

$$\frac{dp_0}{d_t} = -\frac{p_0[(p_0 - 1)s(F(n + p_0 - 1) - p_0 + 1) + n^2 u]}{n}$$

and

$$\hat{p}_0 = 1 - \frac{n\left(\frac{\sqrt{F^2 s - 4Fu + 4u}}{\sqrt{s}} - F\right)}{2(1 - F)}.$$

For $n = 1$ this gives

$$1 - \frac{\frac{\sqrt{F^2 s - 4Fu + 4u}}{\sqrt{s}} - F}{2(1 - F)},$$

recovering the results from Roze and Rousset (2004). In particular, the frequency of the deleterious allele at equilibrium is

$$\hat{p}_1 = (1 - \hat{p}_0) = \frac{\frac{\sqrt{F^2 s - 4Fu + 4u}}{\sqrt{s}} - F}{2(1 - F)}.$$

Using the fact that $F = \frac{\alpha}{2-\alpha}$ we get

$$\hat{p}_1 = \frac{\alpha\sqrt{s} + (\alpha - 2)\sqrt{\frac{\alpha^2 s + 8(\alpha - 2)(\alpha - 1)u}{(\alpha - 2)^2}}}{4(\alpha - 1)\sqrt{s}}, \qquad (A1)$$

which is equivalent to the formula given in Table 1 of Roze and Rousset (2004). Furthermore, the mutation load is then simply u for completely recessive mutations. Figure A1 shows the frequency of the deleterious allele at mutation selection balance.

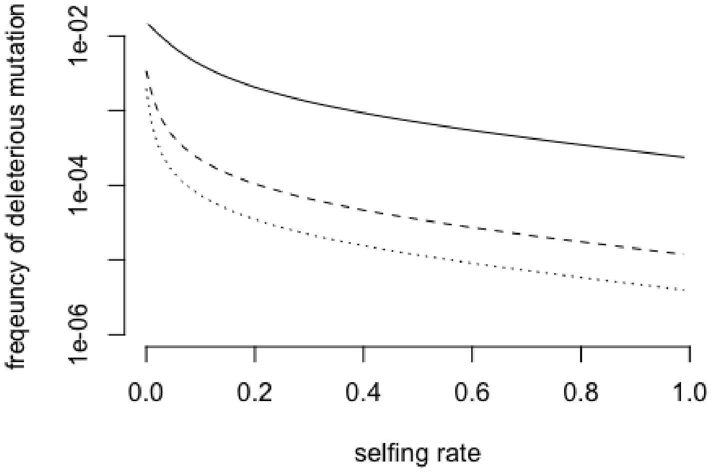

Figure A1. *Frequency of deleterious mutation at mutation selection balance. Parameter values are* $u = 3.5 * 10^{-6}$ *and* $s = 0.015, 0.3, 0.9$ *(from top to bottom).*

Multiple Loci, Full Selfing

For fully selfing populations we get:

$$\hat{p}_0 = 1 - n\frac{u}{s}, \qquad (A2)$$

and hence find that mutation load is independent of recombination. This is sensible as recombination can only act on heterozygotes, which are absent at equilibrium in fully selfing populations.

Recombination and selfing

Direct analytical progress seems unfeasible and we hence resort to approximation. We assume that evolutionary forces are weak, that is $s, u,$ and $r << 1$. This also means that n cannot be too large because with weak recombination multi-locus genotypes might experience selection coefficients of up to ns.

Furthermore, we assume that recombination rate r and mutation rate u are small relative to the strength of selection and selfing. Our approximation should be reliable if selection is weak, but larger than $1/N$ where N is the population size. Furthermore, recombination and per locus mutation rates should be weak relative to selection. The exact conditions for when our approximation will become unreliable are difficult to derive however. Roughly speaking $r, u < 1/N << s << 1$ is assumed throughout this section. Our approximation may not be useful for predictions on a genomic scale, but can accurately describe the dynamics of smaller regions of low recombination and should thus be helpful in providing qualitative explanations.

Weak recombination and selfing

Using an approach similar to that of Gilbert et al. (2020) we find that

$$\hat{\mathrm{p}}_0 = 1 - \sqrt{\mu}n + \frac{n}{2}\left(F + 2\sqrt{\mu} - \sqrt{F^2 + 4\mu}\right) + \frac{1}{12}\sqrt{\mu}(n+1)n\rho, \qquad (A3)$$

where we use the re-scaled parameters $\mu = m/s, \rho = r/s$. The solution is obtained by using the Ansatz $\hat{p}_0 = 1 + C_1\sqrt{\mu} + C_2\rho + C_3\varphi$, where $\varphi = F/s$. Then we develop $\frac{dp_0}{d_t}$ in a Taylor Series around $\rho = 0$ while keeping the ratios $\rho/\sqrt{\mu}$ and ρ/φ constant. Finally, we solve for C_1, C_2 and C_3.

Figure A2 compares the analytic approximation with 2 locus simulations and shows that the approximation is very accurate in the case without recombination. With increasing recombination or selfing rates the fit between theory and simulations becomes increasingly worse, but the transition from the region where we do not expect POD ($P_0 > 0$) to the region where we expect a transition to POD ($P_0 = 0$) is accurately predicted.

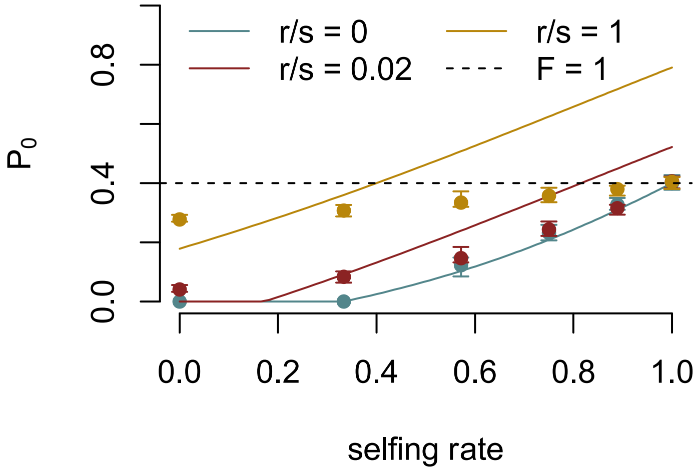

Figure A2. *Frequency of the unloaded haplotype as a function of the selfing rate. Parameter values are n = 2 loci, u/s = 0.3, s = 0.01, N = 5000 diploid individuals and recombination rate r as specified in the legend. Points show the median and error bars the 95 % CI from 50 independent simulation replicates per parameter combination. Solid lines show the analytical approximation equation (A3) and dashed line the analytical expression for F = 1, equation (A2).*

We can now use this formula to find the region in the parameter space for which we expect POD to occur. This is when $\hat{p}_0$ approaches 0. Figure A3 illustrates this region.

We note that it is difficult to directly compare our theoretical results to genome-wide simulations. First, our approximation will generally not be valid for small selection coefficients and/or large mutation rates as we have to assume weak evolutionary forces to derive the mutation-selection balance equilibrium. Second, our model assumes that mutation rates are small relative to selection coefficients at each locus, which might be violated in low recombination regions that accumulate numerous deleterious mutations with a large net fitness effect. Third, our model does not distinguish between loss of the zero-mutation haplotype and an actual transition to POD. While loss of the zero mutation-haplotype is necessary for POD to occur, it may also be the case that selection is so weak relative to mutation and recombination that there is subsequent fixation of deleterious mutations (Lynch *et al.* 1995), which would not be classified as POD in the simulations. In the main text we nevertheless show a qualitative comparison to individual based simulations and find that our model can accurately predict the qualitative patterns observed in genome wide simulations (see Figure 3 in the main text).

Figure A3 explores the parameter combinations for which we expect a transition to POD. We find that like recombination, selfing can efficiently prevent the occurrence of POD. The reason is that purging of recessive mutations is more efficient under partial or full selfing, thus increasing the frequency of the unloaded haplotype P_0.

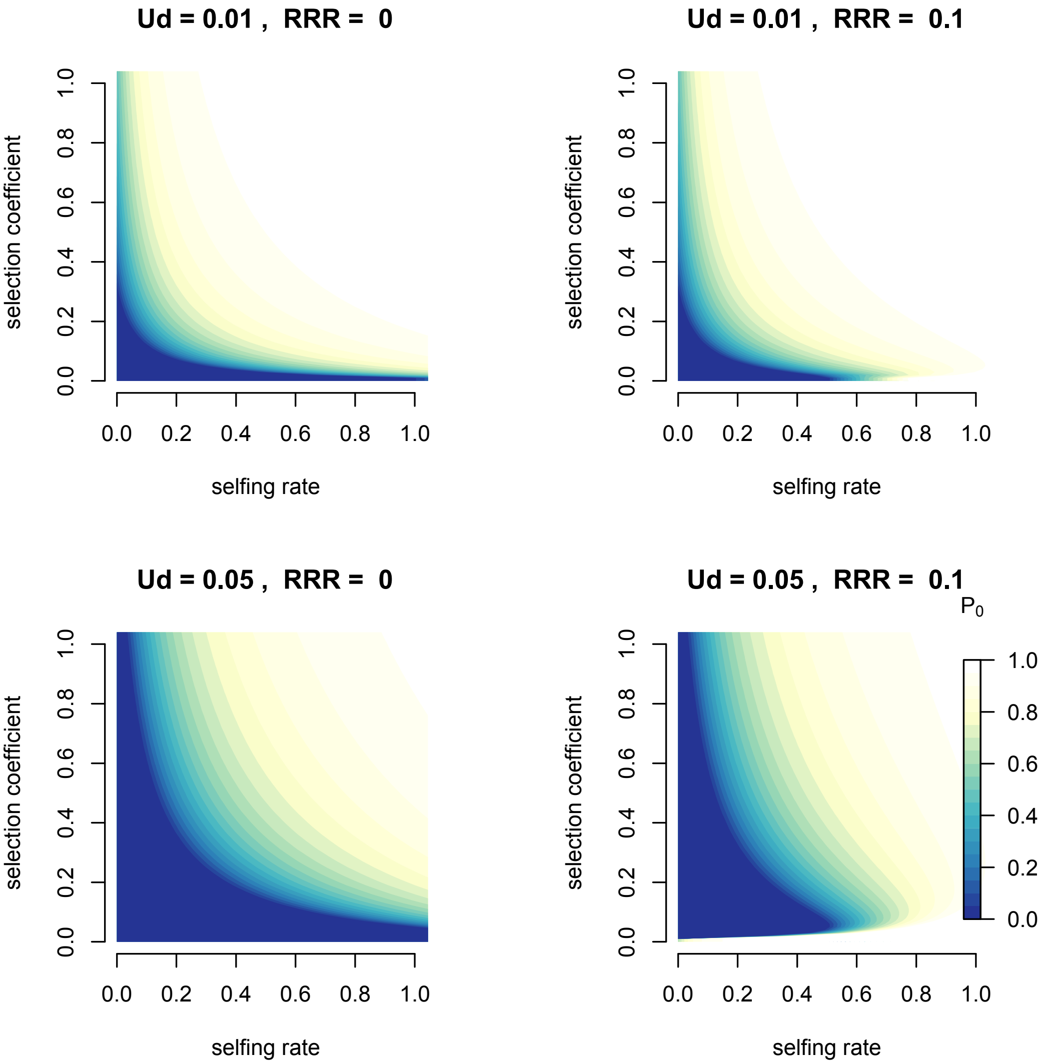

Figure A3. *Frequency of the unloaded haplotype as a function of the selfing rate and the selection coefficient. Parameter values are n = 100 loci, and the relative recombination rate RRR = u/r and total mutation rate Ud = u*n as specified in the figure legend.*

Efficacy of selection with pseudo-overdominance (POD)

We calcualte the effect of POD, that is stabilizing selection that maintains complementary haplotpes, on the efficacy of selection at linked sites. For this we assume POD occurring in a two-locus system without recombination. We consider three biallelic loci. For simplicity, we only consider the following three haplotypes: Abc, aBc, aBC, where capital letters indicate a derived mutation and each position correspond to a locus. Alleles A and B a recessive deleterious with selection coefficient $s_1 \in [-1,0]$. The allele C has selection coefficient $s_2 > -1$ and is co-dominant.

We denote the frequencies of the haplotype Abc, aBc and aBC by p_1, p_2 and p_3, respectively. We assume that the system is at POD equilibrium with $\hat{p}_1 = \hat{p}_2 = 1/2$ and allele C is introduced at low frequency on the aB background.

Genotype frequencies after one generation of selfing are given by

$$p_{ii} = (1-F)p_i + Fp_i$$

and

$$p_{ij} = (1-F)2p_ip_j$$

for $i \neq j$, where $i,j = \{1,2,3\}$ denote the haplotypes Abc, aBc and aBC, respectively. The change in genotype frequencies after selection is then given by

$$\frac{dp_{ij}}{d_t} = (W_{ij} - \overline{W})p_{ij},$$

where W_ij is the fitness of genotype ij and $\overline{W}$ is the mean fitness. We can simplify this system of ordinary differential equations by following the change in allele frequencies instead:

$$\frac{dp_i}{d_t} = \frac{dp_{ii}}{d_t} + \sum_{j\neq i} \frac{1}{2}\frac{dp_{ij}}{d_t}.$$

We further reduce the dimension of the system by setting $p_1 = 1 - p_2 - p_3$. To study the fate of the mutation C linked to the aB haplotype we calculate the rate of increase of the aBC haplotype introduced close to the POD equilibrium $\hat{p}_1 = \hat{p}_2 = 1/2$. We calculate the Jacobian of the system, that is the matrix with the entries $\frac{dp_{ij}}{d_t}$ at position i,j. The eigenvalues of this matrix evaluated at $p_1 = p_2 = 1/2$ give us the rate of change of the frequency of haplotype aBC when it is introduced at infinitesimal frequency. The relevant eigenvalue is given by

$$\lambda(s_1, s_2) = -\frac{1}{2}s_2(F(2(s_1-1)s_2 + 3s_1 - 2) + s_1 - 2).$$

Our goal is to compare the rate of change of the mutation C in the case of POD to the case without POD. We can do this by calculating the ratio $eff_{sel} = \lambda(s_1,s_2)/\lambda(0,s_2)$. If this number is 1, then the initial rate of frequency change of the mutation is not affected POD. If it is smaller than 1, the efficacy of selection is reduced by the presence of POD at linked loci. We find that

$$eff_{sel} = 1 - s_1\left[1 - \frac{1-F}{2(Fs_2 + 1) + 1}\right]. \qquad (A5)$$

For $F = 0$ this is equal to $1 - s_1/2$ and for $F = 1$ this is equal to $1 - s_1$, thus showing that the presence of POD reduces the efficacy of selection for linked additive mutations. This effect is amplified by selfing because the deleterious effects of the recessive mutations linked to the additive mutation become amplified by selfing. Note however, that this effect is counteracted by the decreases likelihood of observing POD in selfing populations (see Figure A3).

References:

Caballero A., and W. G. Hill, 1992 Effects of partial inbreeding on fixation rates and variation of mutant genes. Genetics 131: 493–507.

Gilbert K. J., F. Pouyet, L. Excoffier, and S. Peischl, 2020 Transition from background selection to associative overdominance promotes diversity in regions of low recombination. Curr. Biol. 30: 101-107.e3. https://doi.org/10.1016/j.cub.2019.11.063

Lynch M., J. Conery, and R. Bürger, 1995 Mutational meltdowns in sexual populations. Evolution 49: 1067–1080. https://doi.org/10.1111/j.1558-5646.1995.tb04434.x

Roze D., and F. Rousset, 2004 Joint effects of self-fertilization and population structure on mutation load, inbreeding depression and heterosis. Genetics 167: 1001–1015. https://doi.org/10.1534/genetics.103.025148

Supplementary Material for "Genetic load may increase or decrease with selfing depending upon the recombination environment."

Sianta, S., Peischl, S., Moeller, D., and Y. Brandvain

Table S1: Translation of the genome-wide deleterious mutation rate (U_{del}) and the relative recombination rate (RRR) parameters to other measures of the mutation rate (per-base pair) and recombination rate (per-base pair, and cM/Mb).

U_{del}	**RRR**	**mutation rate**	**recombination rate**	
		per-base-pair	per-base-pair	cm/Mb
0.04	0.01	8.89E-10	8.89E-12	0.0089
0.04	0.1	8.89E-10	8.89E-11	0.0889
0.04	1	8.89E-10	8.89E-10	0.8889
0.04	10	8.89E-10	8.89E-09	8.8889
0.16	0.01	3.56E-09	3.56E-11	0.036
0.16	0.1	3.56E-09	3.56E-10	0.356
0.16	1	3.56E-09	3.56E-09	3.556
0.16	10	3.56E-09	3.56E-08	35.556
0.48	0.01	1.07E-08	1.07E-10	0.11
0.48	0.1	1.07E-08	1.07E-09	1.07
0.48	1	1.07E-08	1.07E-08	10.67
0.48	10	1.07E-08	1.07E-07	106.67

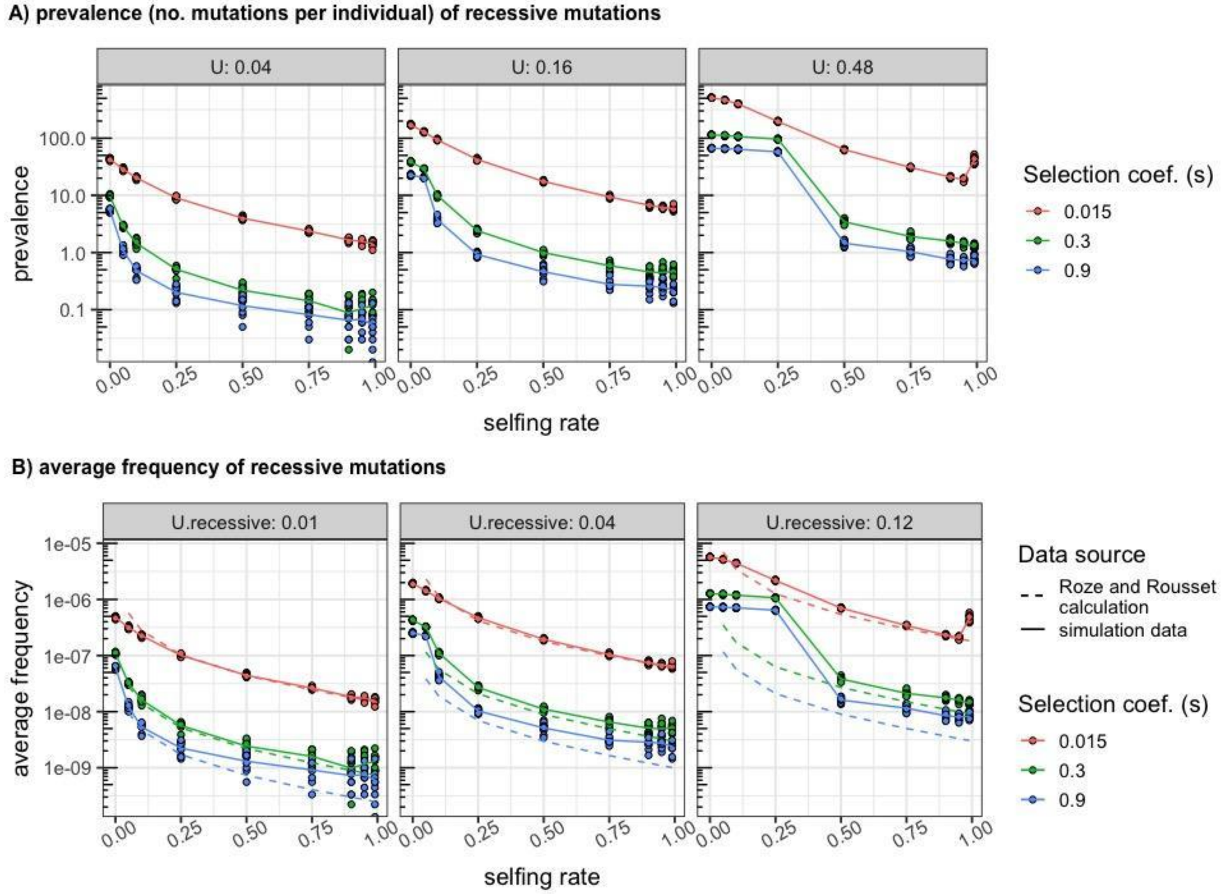

Figure S1: Purging dynamics of fully recessive mutations ($h_{recessive} = 0$) as a function of selfing rate, the fitness effects of recessive mutations ($s_{recessive}$; line colors), and U_{del} (facets) at a high relative recombination rate (RRR; RRR=10). (A) Average prevalence per individual from simulation results. Points represent simulation replicates and lines connect mean values. (B) A qualitative comparison of our simulation results with the expected equilibrium frequency of a recessive mutation at a single locus from Roze and Rousset (2004). We approximate the average frequency of recessive mutations in our simulation by dividing the average prevalence per individual by a quarter of the genome size, as the recessive mutation rate is one quarter of the total deleterious mutation rate.

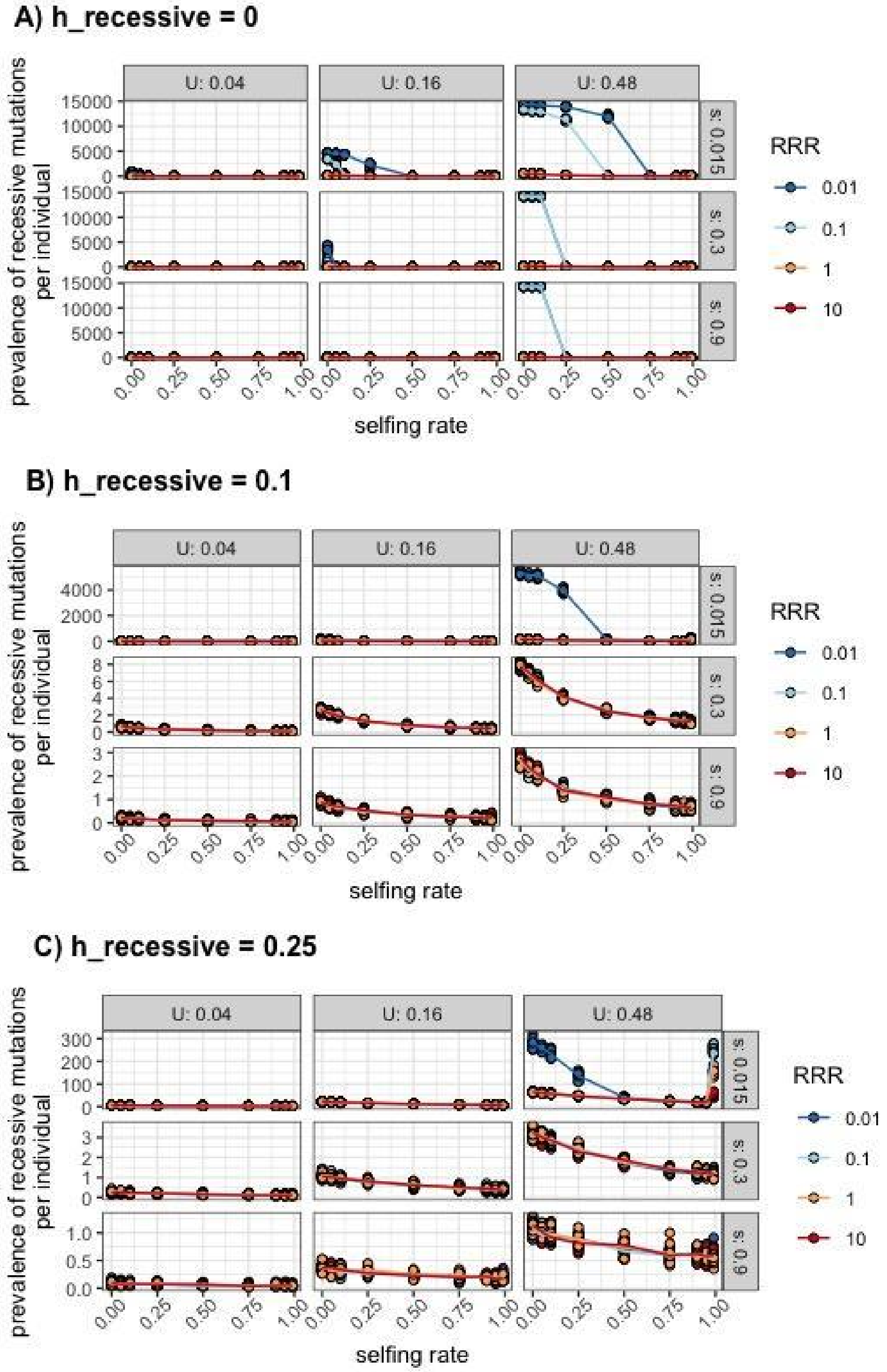

Figure S2: Purging dynamics of recessive mutations as a function of selfing rate in simulations with high (warm colors) vs low (cool colors) relative recombination rates (RRR) highlights the qualitative spike in recessive load due to pseudo-overdominance. Points indicate the average prevalence of recessive mutations per individual in a given simulation replicate, and lines connect the mean value among simulation replicates. A) $h_{recessive}$=0, B) $h_{recessive}$=0.1, C) $h_{recessive}$=0.25. Within each panel, columns indicate different genome-wide mutation rates (U_{del}) and rows indicate the deleterious fitness effects of recessive mutations ($s_{recessive}$).

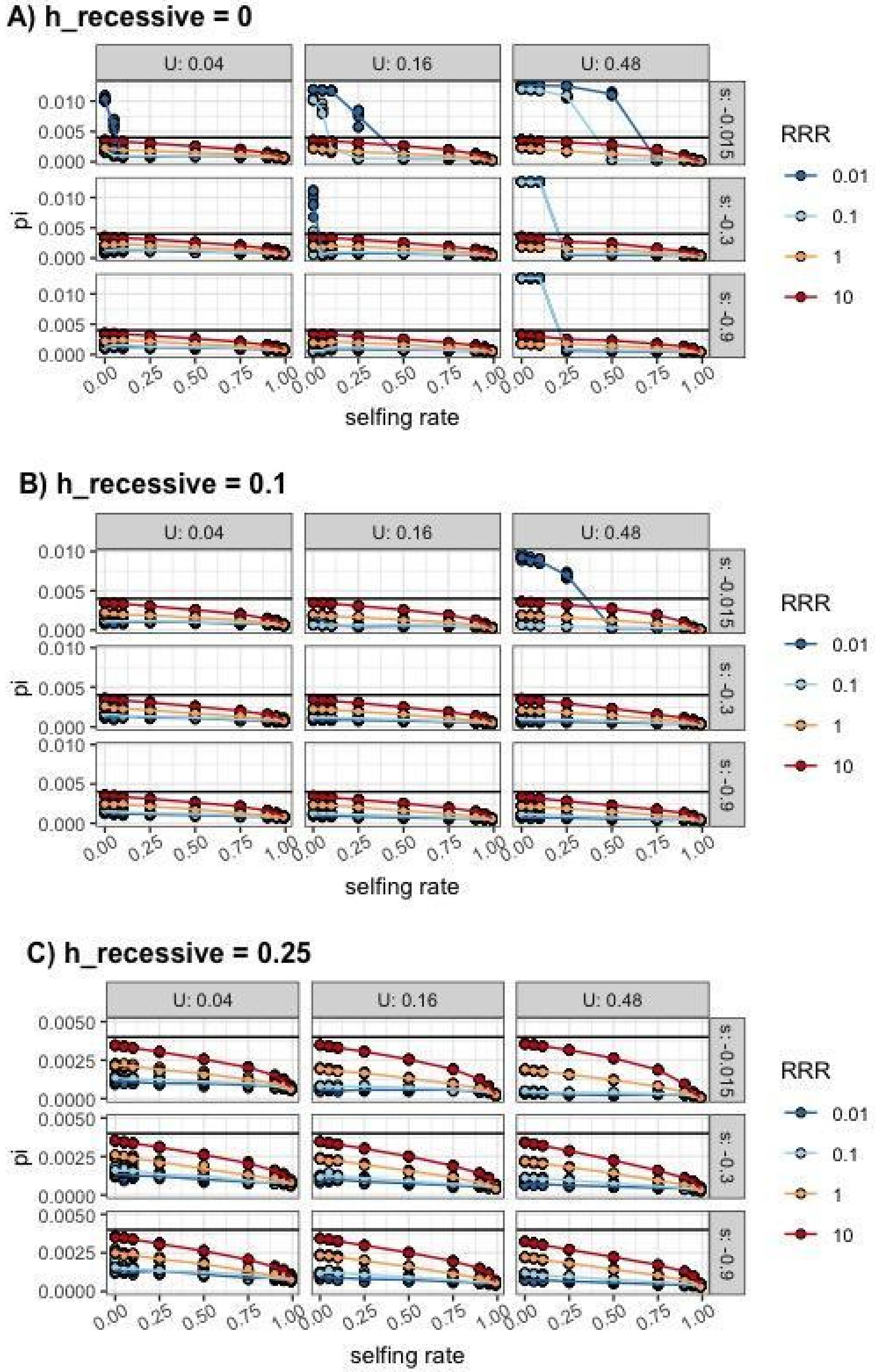

Figure S3: Neutral diversity (π) spikes in simulations with pseudo-overdominance, and otherwise decreases with selfing rate and the relative recombination rate (RRR). Black horizontal lines indicate the neutral expectation of π (= 4Nμ). A) $h_{recessive}$=0, B) $h_{recessive}$=0.1, C) $h_{recessive}$=0.25. Within each panel, columns indicate different genome-wide mutation rates (U_{del}) and rows indicate deleterious fitness effects of recessive mutations ($s_{recessive}$).

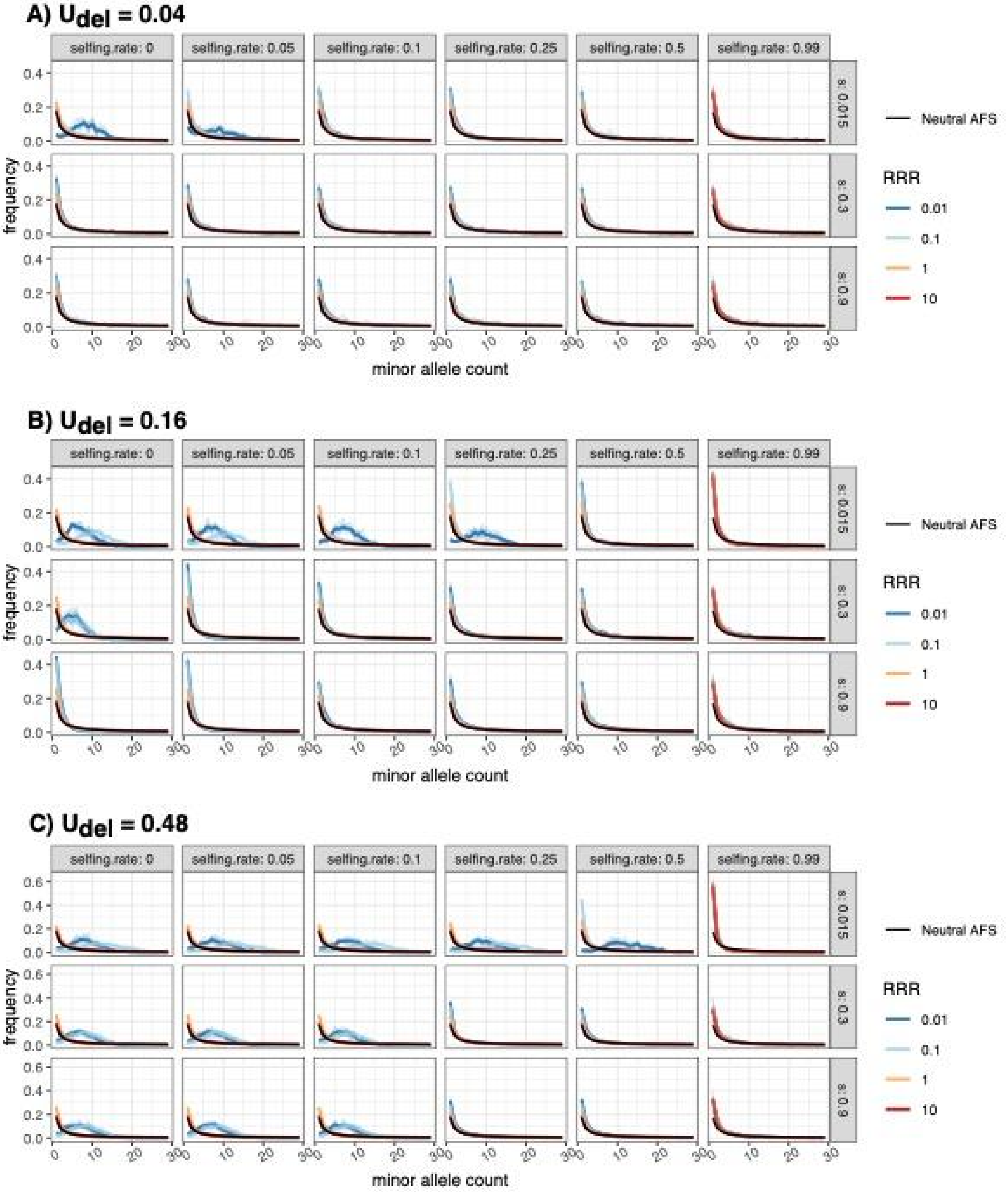

Figure S4: A snapshot of allele frequency spectra (AFS) in simulations with fully recessive ($h_{recessive}$=0) mutations. Within each panel (A-C; varying mutation rates), each graph is the AFS for a given selfing rate (column) and fitness effect of recessive mutations ($s_{recessive}$, row). Black lines indicate expected neutral AFS. We show the frequency of alleles that occur up to 30 times in a sample of 200 genomes.

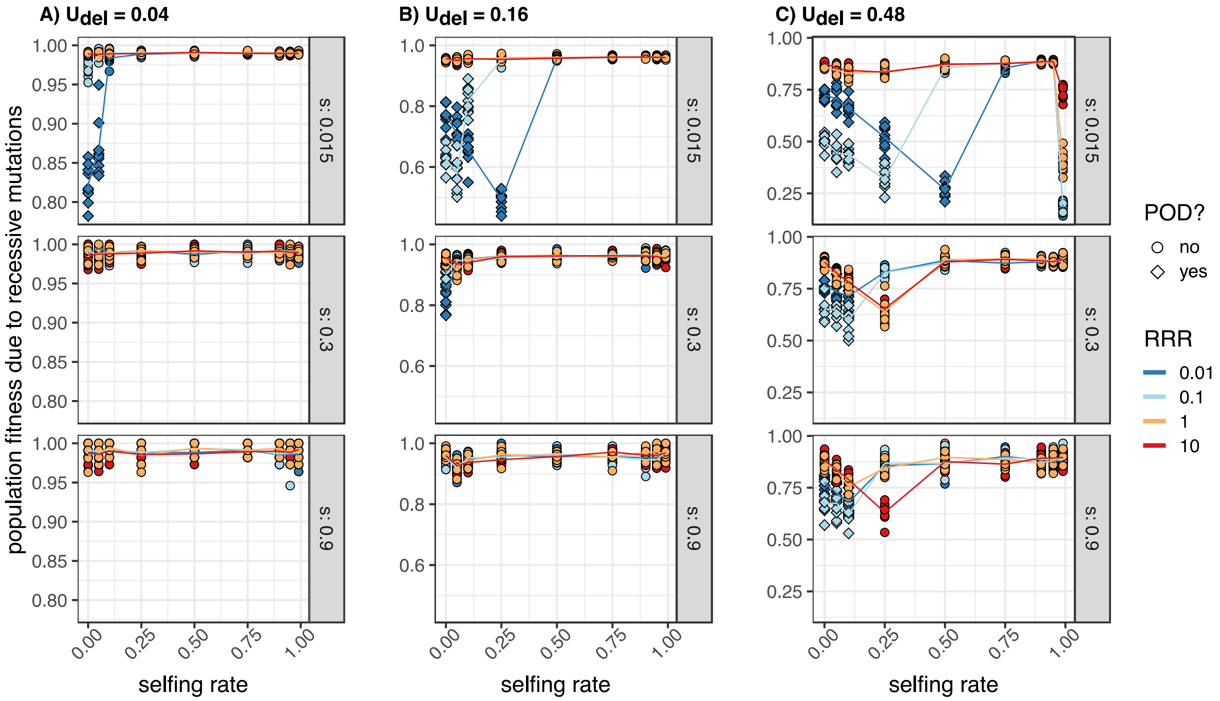

Figure S5: Average population fitness due to recessive mutations across the three genome-wide deleterious mutation rates (A-C) and the selection coefficient of the recessive mutations ($s_{recessive}$, rows). Drops in fitness in low recombination environments (RRR <1) are due to the presence of pseudo-overdominance.

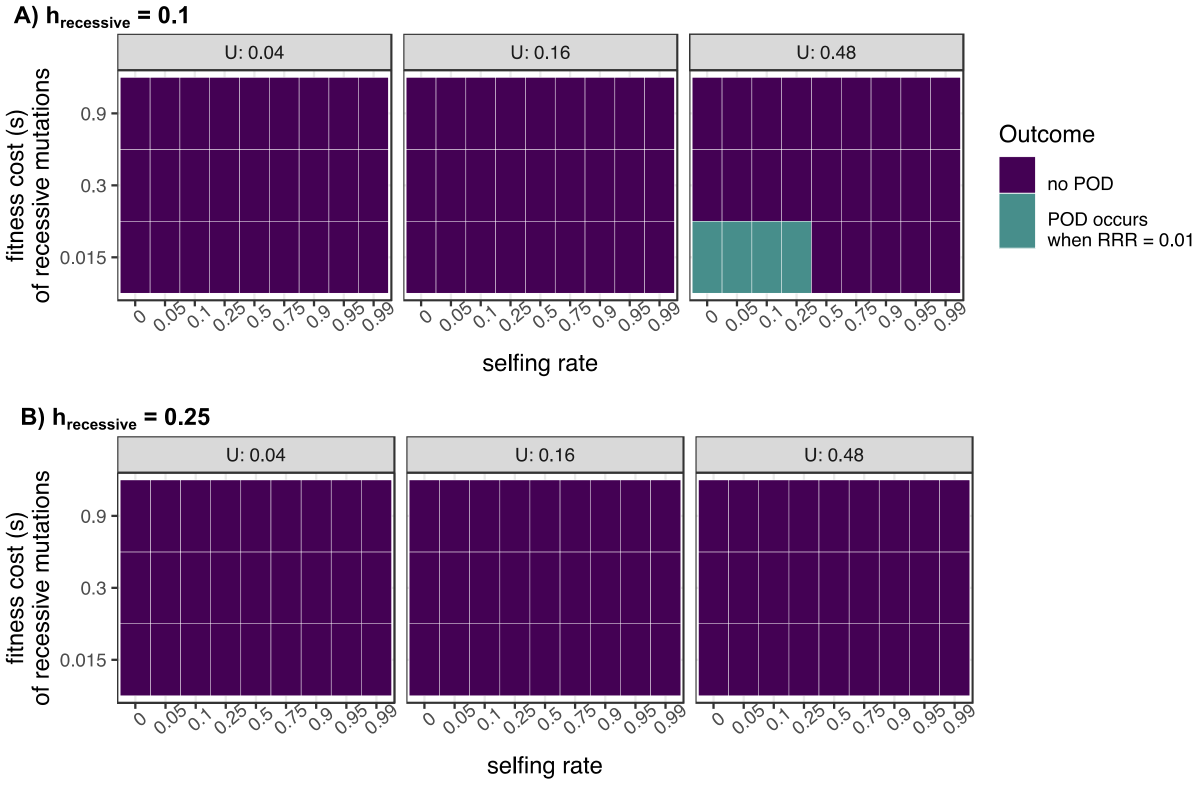

Figure S6: Outcome plot indicating when pseudo-overdominance (POD) occurs across parameter space for partially recessive mutations. (A) Dominance of recessive mutations ($h_{recessive}$) = 0.1. (B) Dominance of recessive mutations ($h_{recessive}$) = 0.25.

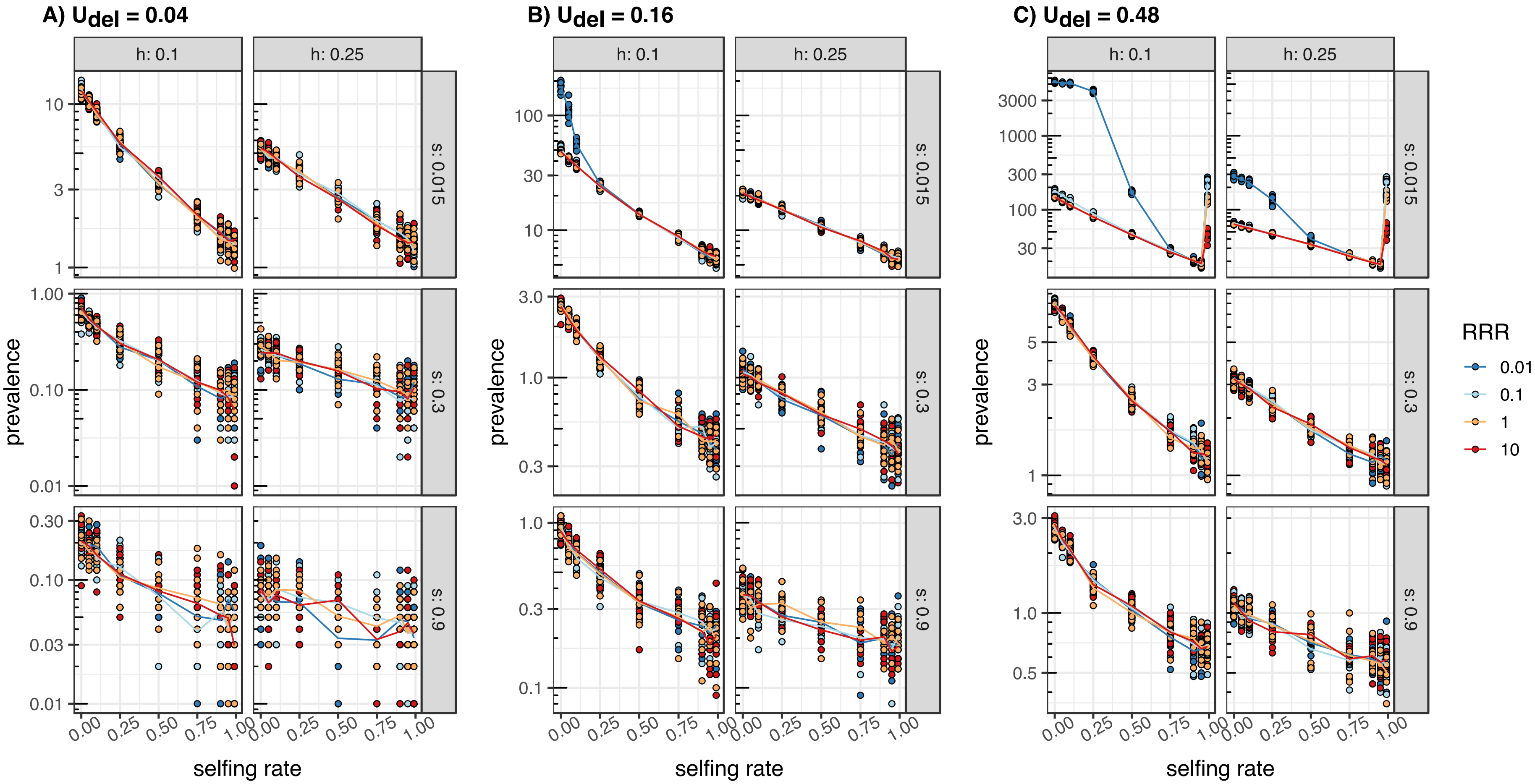

Figure S7: Prevalence of recessive mutations (average prevalence per individual) as a function of selfing rate when mutations are partially recessive ($h_{recessive}$ = 0.1 or $h_{recessive}$ = 0.25) at the three genome-wide deleterious mutation rates (A-C). Rows within panels show prevalence for different fitness effects of the recessive mutations ($s_{recessive}$).

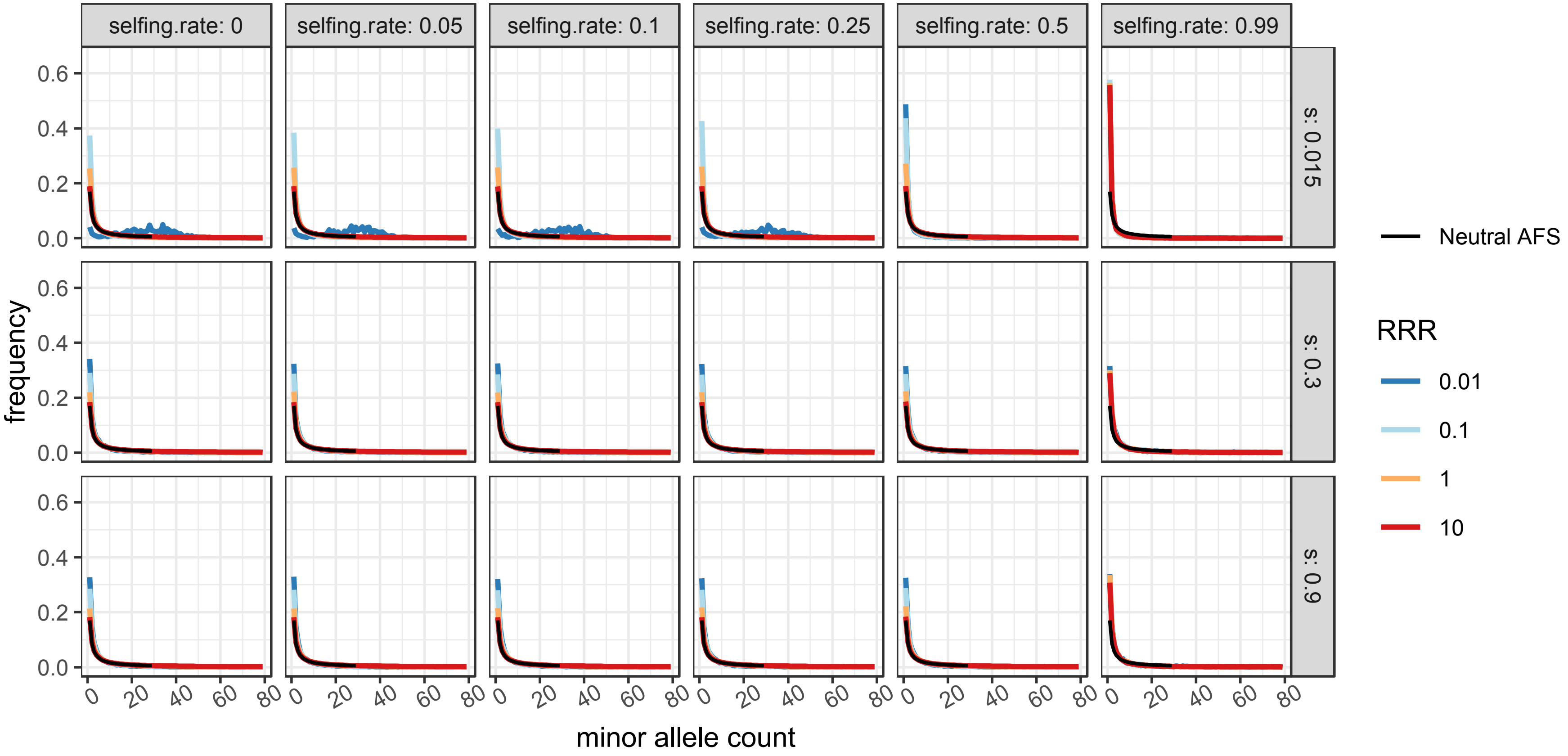

Figure S8: Allele frequency spectra when $h_{recessive}$= 0.1 and U_{del} = 0.48. We show only U_{del} = 0.48, because at this $h_{recessive}$, pseudo-overdominance only occurs at U_{del} = 0.48. The shift to pseudo-overdominance is evidenced by the shift to more intermediate allele frequencies at the lowest relative recombination rate (RRR) when the selfing rate (columns) is less than 0.5 and the fitness effect of recessive mutations (rows) is 0.015. Black lines indicate expected neutral AFS. We show the frequency of alleles that occur up to 30 times in a sample of 200 genomes.

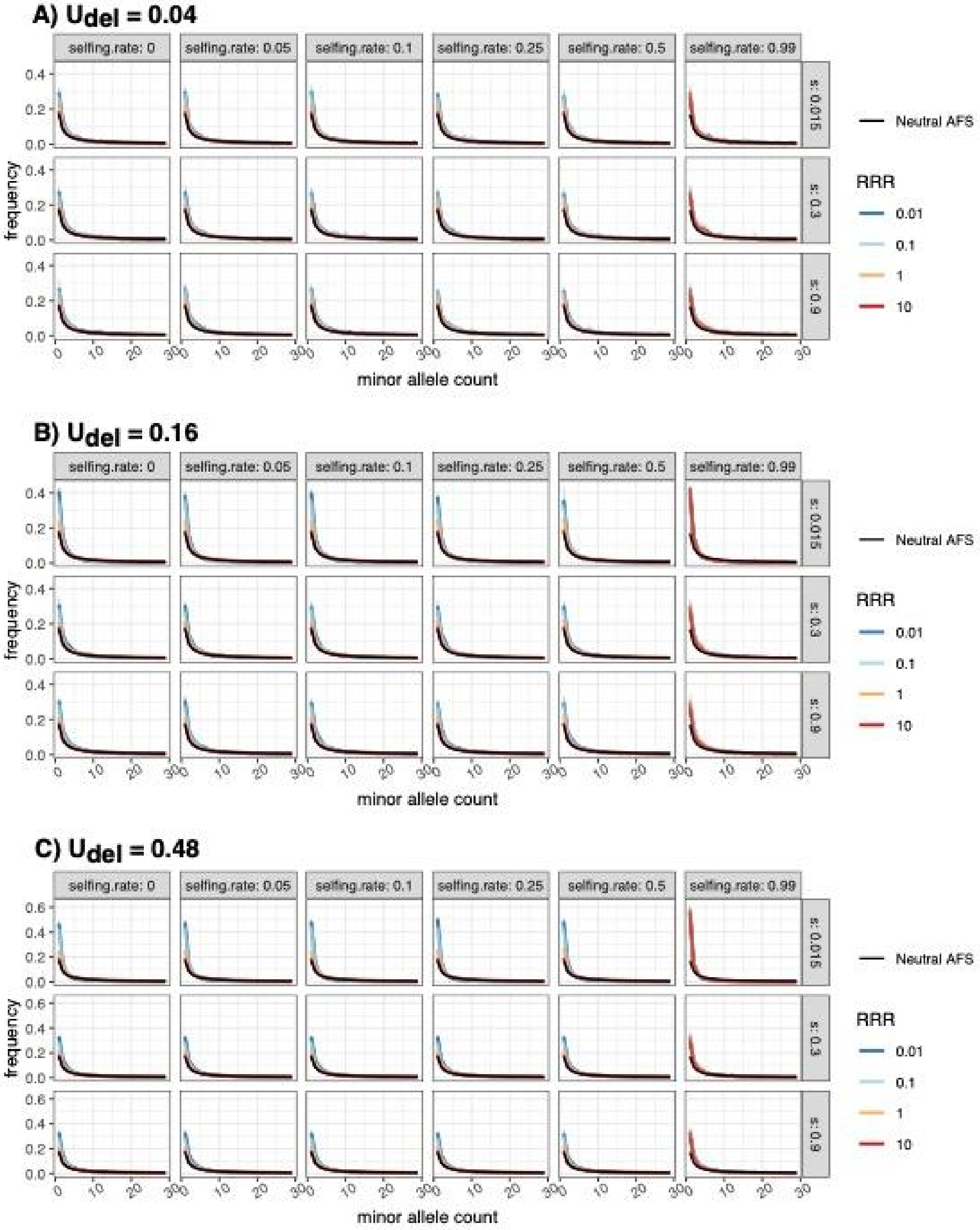

Figure S9: A snapshot of allele frequency spectra (AFS) in simulations with partially recessive ($h_{recessive}$= 0.25) mutations. Within each panel (A-C; varying mutation rates), each graph is the AFS for a given selfing rate (column) and fitness effect of recessive mutations ($s_{recessive}$, row). Black lines indicate expected neutral AFS. We show the frequency of alleles that occur up to 30 times in a sample of 200 genomes.

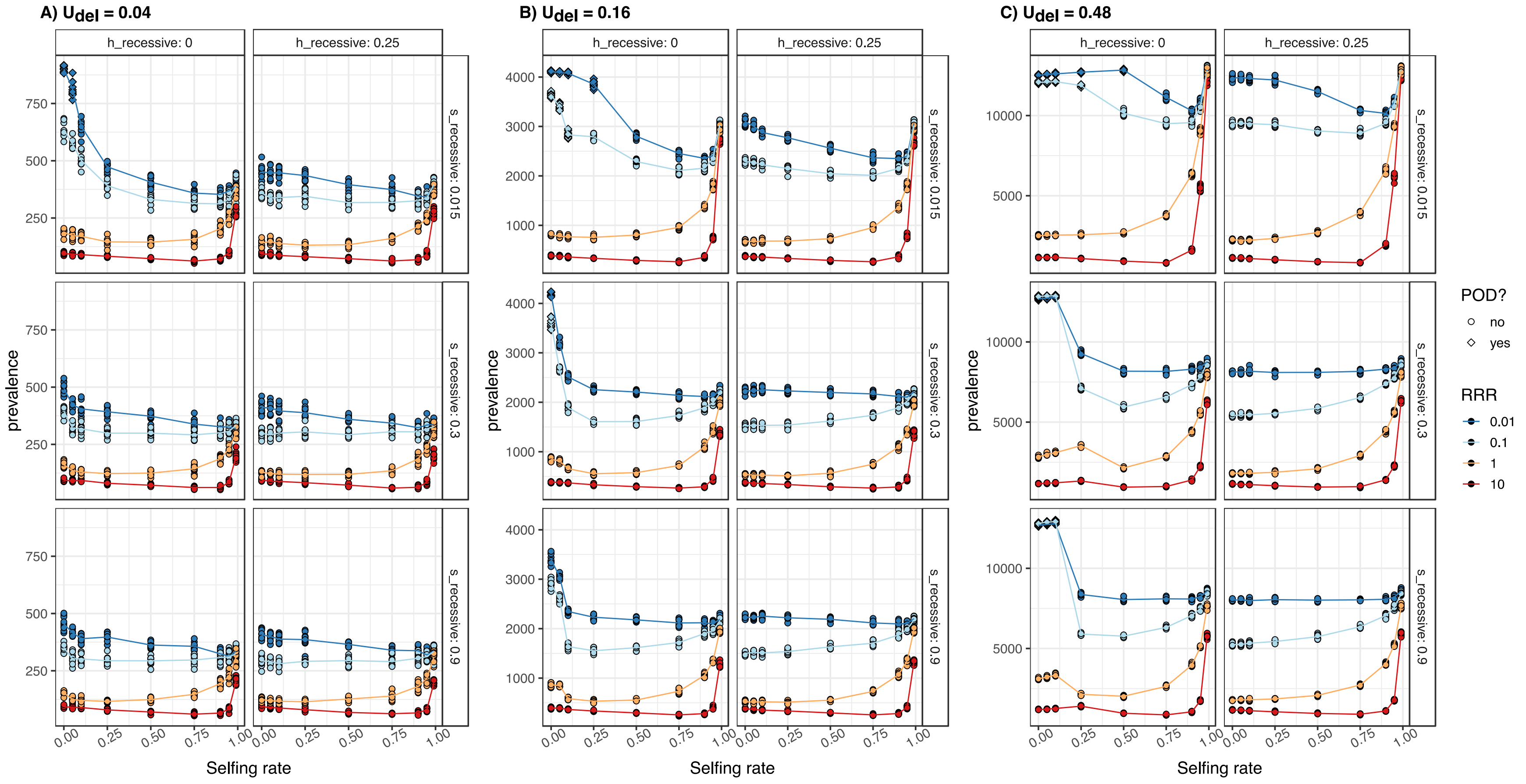

Figure S10: Prevalence of additive mutations characterized by s = -0.0005 across the three genome-wide deleterious mutation rates (A-C). Columns within each panel characterize the dominance of recessive mutations present in the simulation. The presence of pseudo-overdominance affects the accumulation of additive load, as seen by the differences in prevalence between simulations with fully and partially recessive mutations, particularly in low recombination simulations. Rows within each panel characterize the fitness effect of recessive mutations.

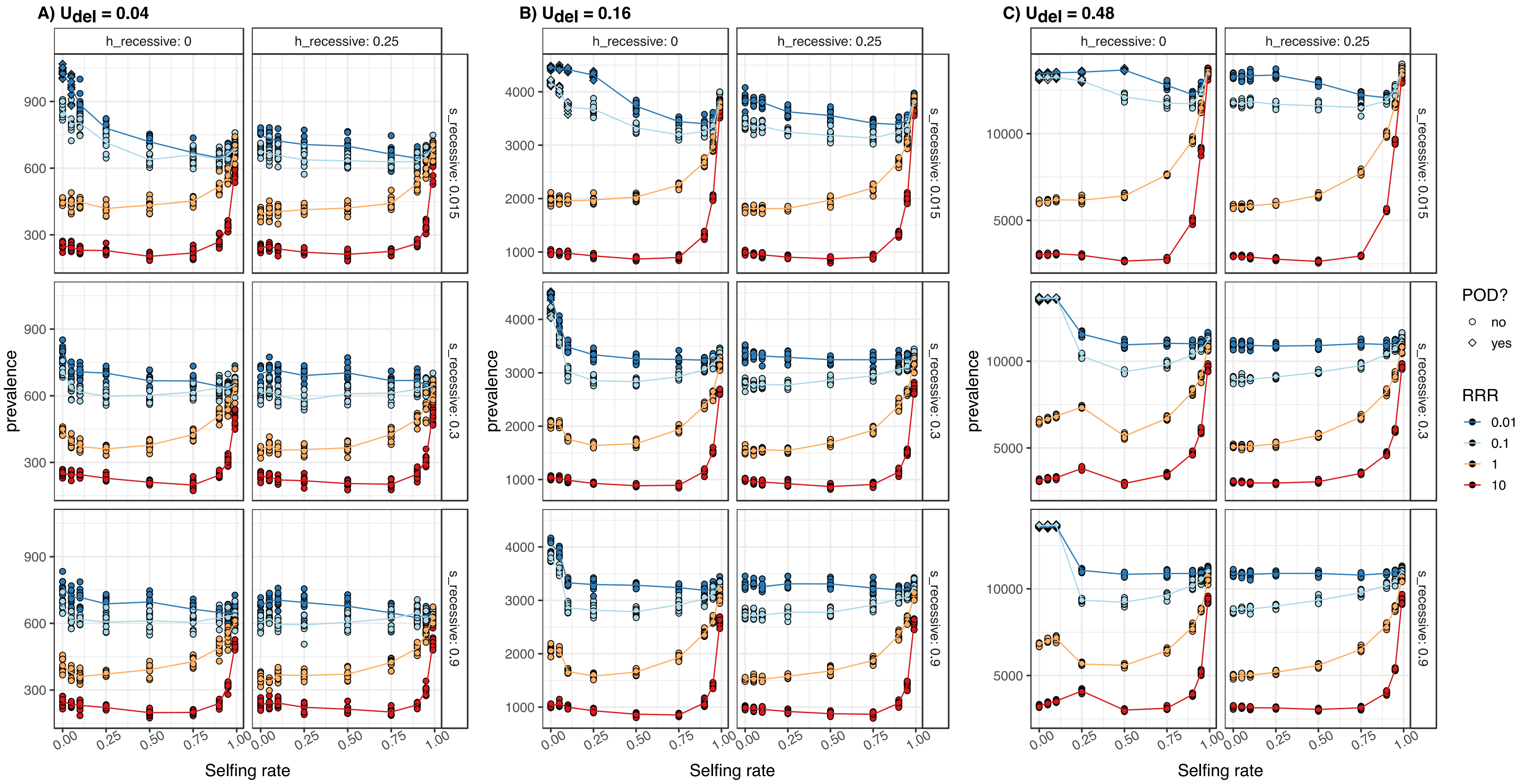

Figure S11: Prevalence of additive mutations characterized by s = -0.00025 across the three genome-wide deleterious mutation rates (A-C). Columns within each panel characterize the dominance of recessive mutations present in the simulation. The presence of POD affects the accumulation of additive load, as seen by the differences in prevalence between simulations with fully and partially recessive mutations, particularly in low recombination simulations. Rows within each panel characterize the fitness effect of recessive mutations.

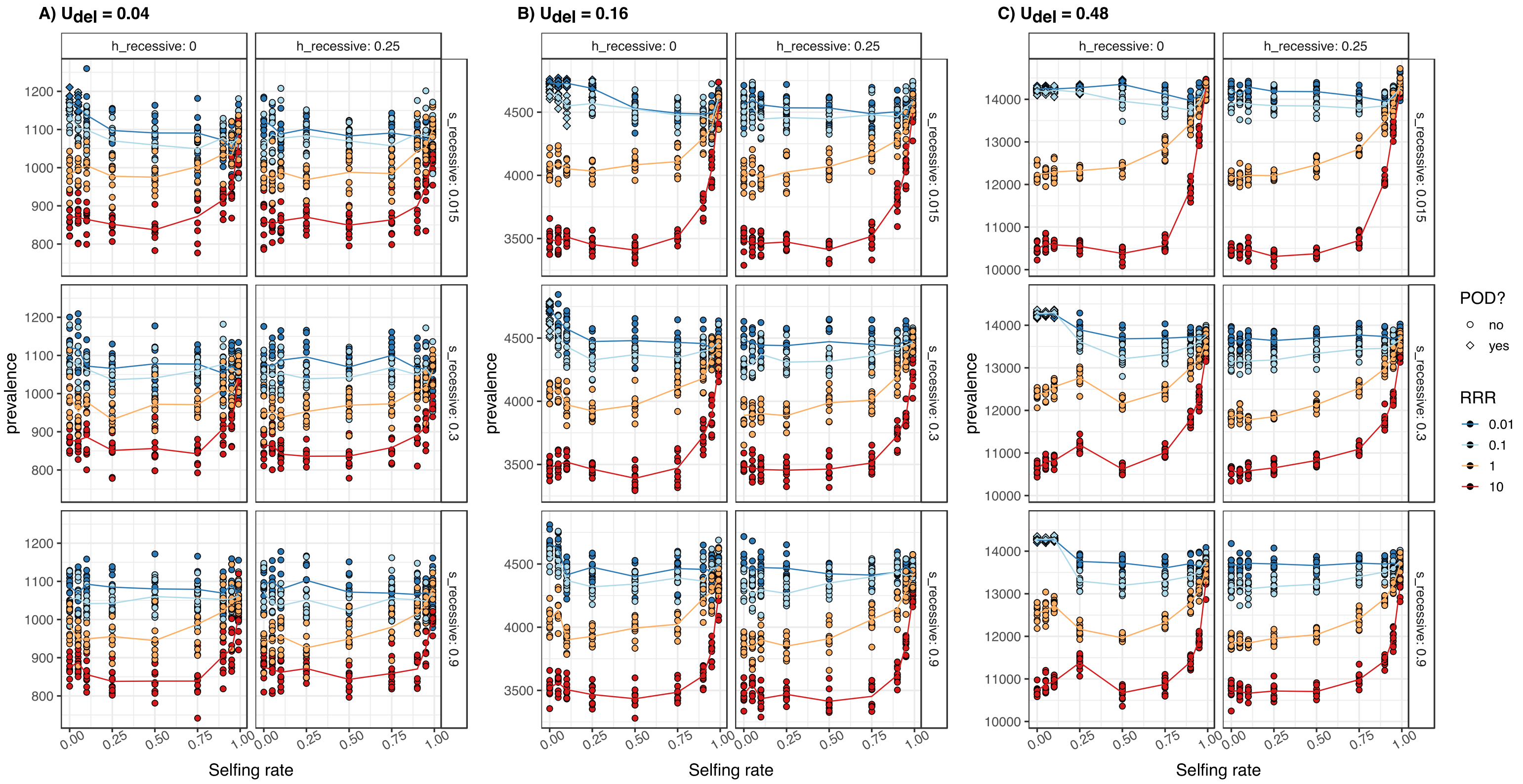

Figure S12: Prevalence of additive mutations characterized by s = 0.00005 across the three genome-wide deleterious mutation rates (A-C). Columns within each panel characterize the dominance of recessive mutations present in the simulation. The presence of POD affects the accumulation of additive load, as seen by the differences in prevalence between simulations with fully and partially recessive mutations, particularly in low recombination simulations. Rows within each panel characterize the fitness effect of recessive mutations.

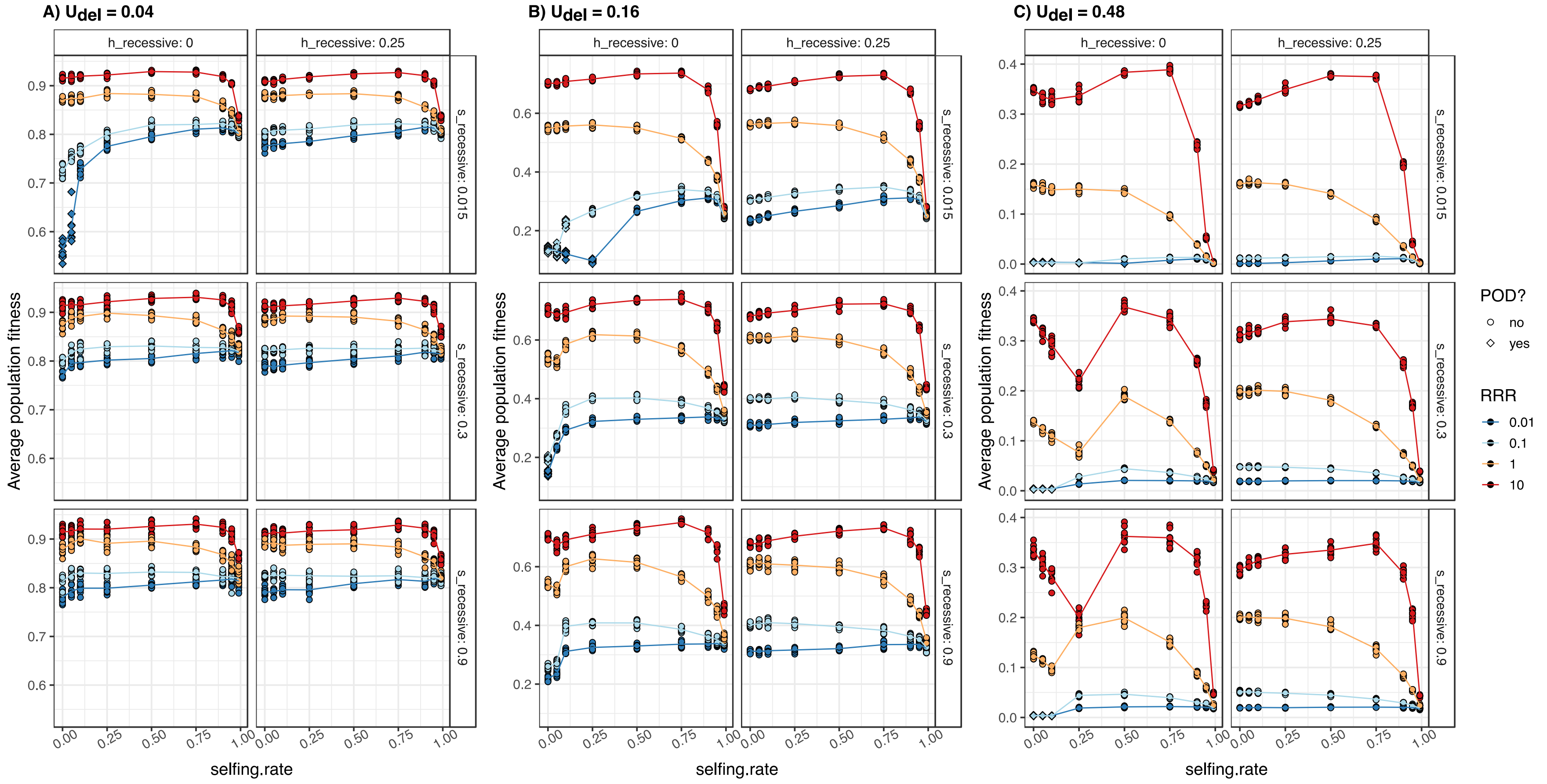

Figure S13: Mean population fitness across the three genome-wide mutation rates (A-C). Columns within each panel characterize the dominance of recessive mutations present in the simulation. The presence of POD affects population fitness, as seen by the differences in prevalence between simulations with fully and partially recessive mutations, particularly in low recombination simulations. Rows within each panel characterize the fitness effect of recessive mutations.