A gene-based model of fitness and its implications for genetic variation: Genetic and inbreeding loads

Parul Johri[1,2,3,*] and Brian Charlesworth[4,*]

[1]Department of Biology, University of North Carolina at Chapel Hill, NC 27599, United States
[2]Department of Genetics, University of North Carolina at Chapel Hill, NC 27599, United States
[3]Integrative Program for Biological and Genome Sciences, University of North Carolina at Chapel Hill, NC 27599, United States
[4]Institute of Ecology and Evolution, School of Biological Sciences, University of Edinburgh, EH9 3FL, United Kingdom

*correspondence may be addressed to these authors: pjohri@unc.edu; brian.charlesworth@ed.ac.uk

Keywords: complementation, dominance, epistasis, fitness models, genetic load, inbreeding load

Running title: Fitness models and mutational load

Abstract

In the companion paper to this, we examined the consequences for patterns of linkage disequilibrium of the "gene" model of fitness, which postulates that the effects of recessive or partially recessive deleterious mutations located at different sites within a gene fail to complement each other. Here, we examine the consequences of the gene model for the genetic and inbreeding loads, using both analytical and simulation methods, and contrast it with the frequently used "sites" model that allows allelic complementation. We show that the gene model results in a slightly lower genetic load, but a much smaller inbreeding load, than the sites model, implying that standard predictions of mutational contributions to inbreeding depression may be overestimates. Synergistic epistasis between pairs of mutations was also modeled, and shown to considerably reduce the inbreeding load for both the gene and sites models. The theoretical results are discussed in relation to data on inbreeding load in *Drosophila melanogaster*.

Introduction

A large volume of recent work on the properties of diploid populations subject to the input of deleterious mutations has been stimulated by concerns about their implications for the survival of small, endangered populations, and by the availability of population genetic data on non-model organisms that shed light on the prevalence of deleterious mutations in natural populations (e.g., Teixeira and Huber 2021; Kardos et al. 2021; Bertorelli et al. 2023; Kyriazis, Robinson and Lohmueller 2023). There are, however, sharp disagreements about the interpretation of the data (Kyriazis et al., 2021, 2023; García-Dorado and Hedrick 2023), emphasizing the need for a secure theoretical basis for data analyses.

In the companion paper to this (Johri and Charlesworth 2025), we analysed the consequences for patterns of linkage disequilibrium (LD) of the differences between two models of the fitness effects of deleterious mutations located within the same gene, the "gene" and "sites" models. The gene model assumes that two heterozygous mutations in *trans* have a larger effect on fitness than two mutations in *cis*, due to a lack of complementation, whereas the sites model assumes that there is no such difference. Theoretical predictions about the properties of deleterious mutations in populations, and their effects on features such as population mean fitness, genetic variance in fitness and inbreeding depression have largely been based on the sites model, together with the assumption of multiplicative fitnesses when the effects of mutations at different sites are combined (for a recent review, see Kyriazis, Robinson and Lohmueller 2023).

But the difference between the gene and sites models can affect the effective level of recessivity experienced by deleterious mutations, as can be seen from the following, highly simplified, example. Consider a single coding sequence that is segregating for mutations at two different nucleotide sites (1 and 2) with the same selection coefficient s against homozygotes at each site. Assume that the double mutant haplotypes are sufficiently rare that they can be ignored. Let the frequency of mutant variants at site i be q_i. Then haplotypes carrying a mutation at site 1 will meet a haplotype that is wild-type at site 1 and mutant at site 2 with an approximate frequency of p_1q_2; the fitness of this genotype is $1 - s$ if there is no allelic complementation (*i.e*, under the gene model), but the sites model would assign it a fitness of $1 - 2hs$ (assuming additive effects across sites), where h is the dominance coefficient. It will meet haplotypes that are mutant at site 1 but wild-type at site 2 with a frequency of approximately q_1p_2. Both models would assign fitness $1 - s$ to this genotype. A similar argument applies to a haplotype carrying a mutation at site 2, with the appropriate switch between subscripts 1 and 2.

All of the double mutant genotypes are classed as homozygotes under the gene model, which is biologically realistic for the majority of genes, for which sites with allelic complementation are relatively rare (Fincham et al. 1979, pp.348-353; Hawley and Gilliland 2006). The mean difference between the gene and sites models in the fitness reduction assigned to haplotypes that each carry a mutation at one or other site is thus:

$$-\ \delta w = (1-2h)s\frac{2(q_1p_2p_1q_2)}{(q_1p_2+p_1q_2)} \qquad (1)$$

If the frequencies of the deleterious mutations at the two sites are similar, so that the subscripts can be dropped, $-\ \delta w \approx 2(1-2h)spq$, which is positive if $h << 0.5$. If $h > 0$ and the population is at deterministic mutation-selection balance, $q \approx u/(hs)$, (Haldane 1927), and $-\ \delta w \approx 2(1-2h)u/h$. For completely recessive deleterious mutations at equilibrium, $q \approx \sqrt{u/s}$ and $-\delta w \approx 2(1-2h)\sqrt{us}$. This implies that the sites model can underestimate the true homozygous fitness effects of deleterious mutations, although the effect is small for a single pair of rare deleterious variants. It follows that deleterious mutations are more likely to segregate at higher allele frequencies under the sites model than the gene model, leading to a lower mean fitness of the population, a higher genetic load, and a larger effect of inbreeding on fitness. In addition to this effect, our previous paper showed that the gene model can generate positive LD among deleterious variants under circumstances where the sites model causes zero or slightly negative LD (Johri and Charlesworth 2025). Since positive LD enhances the efficacy of selection (e.g., Barton 1995; Roze 2021), this effect would also help to reduce the frequency of deleterious mutations expected under the gene model compared with the sites model.

The purpose of the present paper is to examine the consequences of the differences between the gene and sites models for the load statistics, using deterministic and stochastic simulation models of a single coding sequence subject to deleterious mutations, which were described in Johri and Charlesworth (2025). We show that the difference between the two models has especially significant consequences for the level of inbreeding depression caused by deleterious mutations, suggesting that current predictions based on the sites model could greatly overestimate the expected level of inbreeding depression contributed by deleterious mutations. We also note that the sites model is used in current methods for estimating the distribution of fitness effects of deleterious mutations, and that this will lead to

overestimation of the mean strength of selection against such mutations if the gene model is more appropriate.

Methods

Fitness models and simulation methods

The fitness models and simulation results used here are described in the Methods section of Johri and Charlesworth (2025). Recursion relations for haplotype frequencies for the two-site deterministic model are described in the Appendix of that paper; use of these expressions allows the numerical values of the equilibrium haplotype frequencies to be determined for a given set of mutation and fitness parameters.

Estimating the load statistics from the simulations

The state of a population with *N* diploid individuals (where *N* was usually equal to 1000) with respect to the load statistics was summarized using four quantities – the genetic load (*L*), the inbreeding load (*B*), the variance in fitness between individuals (*V*), and the mean frequency of mutant alleles per site ($\bar{q}$). The genetic load for each simulation replicate was calculated as $L = 1 - \bar{w}$, where $\bar{w}$ is the mean fitness of the population for all individuals in the simulated population, relative to a value of 1 for individuals homozygous for the wild-type allele at all sites; $\bar{w}$ was calculated using all individuals sampled after burn-in. The variance *V* was calculated in a similar way. Allele frequencies (q) were estimated from a sample of 100 genomes and included monomorphic sites.

To calculate B, which is defined here as the mean difference in fitness between completely outbred and completely inbred individuals, 100 genomes were sampled randomly from the population, from which 100 diploid inbred individuals were artificially created (*i.e*, each haploid genome was duplicated from a single diploid individual). The inbreeding load was calculated as $B = \bar{w} - \bar{w}_I$, where $\bar{w}_I$ is the mean fitness of the subsampled inbred population. Because *L* and *B* are both small, this expression for *B* differs little from the usual expression for inbreeding depression, $1 - \bar{w}_I/\bar{w}$. Under the assumption of small multiplicative fitness effects across sites, it is equivalent to the negative of the regression coefficient of fitness on inbreeding coefficient (Morton et al. 1956). Each simulation replicate yields a single data point for each variable, and the distributions, means and standard errors shown in the figures and tables are derived from 1000 replicate simulations.

Data availability

The scripts used to determine the properties of two-site equilibrium populations, perform all the simulations, and calculate population genetic statistics are provided at https://github.com/paruljohri/Gene_vs_sites_model/tree/main.

Results

Deterministic results for two selected sites

To aid the reader, Table 1 of Johri and Charlesworth (2025) is reproduced here. The fact that the two sites have identical fitness effects means that the four haplotype frequencies, x_1 (+ +), x_2 (– +), x_3 (+ –) and x_4 (– –) (where + and – indicates wild-type and mutant alleles, respectively) can be replaced with $x = x_2 = x_3$, $y = x_4$, and $z = 1 - 2x - y$. The frequency of the – allele at a site is $q = x + y$. The load statistics L, B, and V (the genetic load, the inbreeding load, and the genetic variance, respectively) can easily be determined once the equilibrium haplotype frequencies are known, using the formulae presented in the Appendix, Equations (A1) – A(8). As in Johri and Charlesworth (2025), we assume no recombination between the two sites, which should provide a good approximation to what is expected for a single gene.

For the purpose of illustrating the results for the two-site model, L and B are represented by the ratios of their values to those for an additive model of two independent loci, when $h > 0$ and $u << hs$. In this case, $L \approx 4u$ (see Haldane 1937) and $B \approx 2u[\,(1/(h) - 2]$ (see Morton et al. 1956). To avoid large values when h is small, the variance is represented by $Var = hV/(4uhs)$, the denominator being the equilibrium value of V for the purely additive case with no LD and $u << hs$. This means that results for the variances with different h values are not directly comparable, but the sites and gene models with the same h can be compared. The frequency of deleterious mutations is calculated from the ratio of the exact deterministic equilibrium value to the approximate value for a single site subject to mutation and selection, given by Equation (A4) of Johri and Charlesworth (2025). This ratio is denoted by q_r in Table 2, which shows the results for the case of $u = 0.0001$, $s = 0.01$, which was previously used for the LD statistics presented in Johri and Charlesworth (2025). As noted there, these are not realistic parameter values, but produce relatively large equilibrium allele frequencies, enhancing the contrast between the two models. Since the denominator q_r is independent of the strength of epistasis, and is the same for the gene and sites model, differences in q_r between the two models and between different values of the

epistasis coefficient ϵ for a given h and s reflect differences in the absolute deleterious allele frequency, q.

As before, we focus attention on cases with zero epistasis ($\epsilon = 0$) and synergistic epistasis ($\epsilon > 0$). Over the range of values of ϵ considered here, its magnitude has relatively little effect on L and B, for both the sites and gene models. Its effect on q_r is more complex. For both models, q decreases with ϵ, but the decline is faster for the sites model when h is close to ½, with the result that in this case q is smaller for the sites model than the gene model when ϵ is large, despite their having the same q with no epistasis. For h sufficiently smaller than ½, q is smaller for the gene model.

The gene model, as might be expected, is associated with lower values of L and B than the sites model when h <½, but has higher values of V when h is sufficiently small. The higher values of V under the gene model reflect its larger values of LD compared with the sites model, with positive associations ($D > 0$) between deleterious variants when ϵ is sufficiently small, especially with small h (Johri and Charlesworth 2025). With h = ½, $B = 0$ (so that the ratios in Table 2 are set to one) and L is the same for the two models, as expected. When h < ½ and ϵ is small, the ratios of B for the sites model to the gene model are larger than the corresponding ratios for L. V increases more quickly with ϵ for the sites model than the gene model, the reverse of the pattern for q. This can result in a slightly higher V for the sites model with large ϵ and h, probably reflecting the larger contributions of terms in ϵhs under the sites model, as can be seen in the fitness matrices in Table 1.

Simulation results for multiple sites

Load statistics with no epistasis

The analytical results described above suggest that, relative to the sites model, the gene model should result in smaller mean frequencies of the deleterious alleles, causing a a lower genetic load and inbreeding load. The multiple site simulations with scaled parameters that are appropriate for *Drosophila* populations, *i.e*, 1000 selected sites, $Nu = 0.005$, and $Nr = 0.01$ (for details, see the *Methods* section of Johri and Charlesworth 2025), show that the equilibrium frequencies of deleterious alleles are much smaller for the gene than the sites model, except when selection is weak relative to drift ($\gamma = 2$ for the case of a fixed selection coefficient (Table 3) and $\bar{\gamma} = 2$ for the case of a gamma distribution of selection coefficients (Table 4), in which case there is little difference between the two models. There is also a drastic difference between the estimated inbreeding load (B) between

the two models: B is much smaller for the gene than the sites model when $h < ½$, even with weak selection. There is a much smaller effect on the genetic load (L), which is noticeable only when mutations are fully recessive.

As discussed in the section *Deterministic results for two selected sites*, deleterious alleles are more likely to be associated on the same haplotypes in the gene model than in the sites model, resulting in larger positive mean D values for the gene model, and inflating the variance (V). This effect was visible in the simulation results described in Johri and Charlesworth (2025). Its effect on V can be counteracted by lower allele frequencies under the gene model, an effect that is especially strong for fully recessive mutations, leading to a much lower V under the gene model when $h = 0$. However, V is significantly higher for the gene model than the sites model when $h = 0.2$ and $\gamma = 20$. With $h = ½$, there is no difference between the two models, as expected. The difference in V between the two models thus has a non-linear relation with h.

To evaluate the possible effects of rescaling on our results, simulations were also performed with 5000 diploid individuals, where Nu, Nr, and Ns were the same as for $N = 1000$. The mean equilibrium frequencies of deleterious alleles were found to be identical in both cases (compare Table 3 with Table S1 in Supplementary File S1), as is expected from the general theory of population genetic processes when all evolutionary forces are weak and their strength can be represented by single variables such as u, s and r (e.g., Ewens 2004, Chap.5). The genetic and inbreeding loads involve products of the haplotype frequencies and s, whereas the variance involves products of the haplotype frequencies and s^2 (see section 1 of the Appendix). If the haplotype frequencies are preserved after multiplying the population size by a factor C and keeping the scaled parameters constant, the simulated values of the two load statistics should thus be multiplied by C and the variance by C^2, in order to produce values that are comparable with the results for the smaller N value ($C = 5$ in the present cast). As can be seen from Tables 3 and S1, the load statistics for $N = 5000$ after this correction agree well with those for $N = 1000$, implying that we can have confidence in the rescaling procedure. These results imply that studies interested in evaluating the extent of mutational load in endangered populations using forward simulations need to apply similar corrections to the load statistics if rescaling is employed.

Because the results discussed above are most relevant to *D. melanogaster*-like population parameters, we also conducted simulations with parameters appropriate for human populations (for details, see Johri and Charlesworth 2025), with much smaller values of Nu

(0.00025) and Nr (0.0002). The overall patterns are very similar to those for *Drosophila*-like populations (Table S2 in Supplementary File S1). The mean frequencies of deleterious alleles are significantly smaller with the gene than the sites model; genetic load is slightly lower, while inbreeding load is again drastically smaller for the gene model than the sites model when $h < ½$.

Load statistics with epistasis

With synergistic epistasis between deleterious mutations, a decrease in their mean equilibrium frequencies is expected compared with the case with no epistasis, due to the reduced fitness of genotypes with multiple mutations that is caused by synergistic epistasis (Crow 1970; Kondrashov 1988; Charlesworth 1990) In the simulations with multiple linked sites under a DFE with a relatively large mean scaled selection coefficient ($\bar{\gamma} = 2N\bar{s} = 100$) and varying degrees of epistasis (with ϵ ranging from 0.02 to 0.32), there is a severe decrease in the mean equilibrium frequency of deleterious mutations with increasing ϵ, for both the gene and sites models (Figure 1). Even with the lowest strength of epistasis ($\epsilon = 0.02$), there is no significant difference in mean equilibrium frequencies of deleterious alleles between the gene and sites models. There is a corresponding decrease in the inbreeding load (B) with increasing ϵ for both the gene and sites model (Figure 2) when $h < ½$.

B is still significantly and drastically higher with the sites model than the gene model for both $h = 0$ and $h = 0.2$. There is no difference between the two models with $h = 0.5$, as expected; unexpectedly, however, B becomes increasingly slightly negative as ϵ increases. One possible explanation is that this is an artefact of the way B in which was estimated; the fitness of inbred individuals was calculated by sampling 100 haploid genomes, whereas the mean fitness was calculated using the entire population. However, even if we calculate the fitness of outbred individuals by sampling 100 diploid individuals, B remains negative for these parameter sets (see Figure S1 in Supplementary File S1), so that this explanation can be ruled out.

The probable cause of this effect is as follows. With the parameter set in question, the frequencies of deleterious alleles are sufficiently high that essentially all haplotypes carry at least one deleterious mutation. In the two-site case with $h = 0.5$, a – +/+ – individual has fitness reduction $s(1 + \epsilon)$, whereas – + or + – homozygotes have a fitness reduction s (there is no effect of epistasis). Thus, the heterozygote has a lower fitness than the homozygote. In other words, the outbred population has more opportunities for pairwise fitness interactions

that reduce fitness than the inbred population, resulting in slightly negative inbreeding depression. Epistatic interactions of this kind have been studied theoretically and empirically in connection with the biometrical genetics of crosses between inbred lines (e.g., Wright 1968, Chap. 15, pp. 391-403; Jinks 1983). A lower value of a fitness component in the F_1 compared with the mean of the two parents of the kind expected under this model appears to be uncommon in such crosses, with a prevalence of the opposite pattern (heterosis), consistent with the effects of largely recessive effects of deleterious alleles or heterozygote advantage outweighing epistatic effects of this kind.

There is a much smaller effect of epistasis on the genetic load (L in Figure 2), with a minor and steady decrease in L with an increase in ϵ for both fitness models. As before, the difference in the genetic load between the two models is relatively small and only significant when mutations are completely recessive. The variance in fitness (V) also increases with an increase in the epistasis coefficient when $h = 0.2$ or 0.5 (and only very slightly when mutations are fully recessive). However, as was also found without epistasis, the variance is much larger for the gene model compared to the sites model, especially when mutations have increased recessivity and intermediate/low levels of epistasis (Figure 1). With high levels of epistasis ($\epsilon = 0.32$), however, there is no difference in V between the two fitness models. In summary, if there were widespread epistasis between deleterious mutations, there would be no substantial difference in mean equilibrium allele frequencies between the two models, but the inbreeding and genetic loads would be lower with the gene model than the sites model if mutations have moderate to high recessivity. Note that with much lower rates of crossing over (0.1 × the standard rate used here), the behavior of the load statistics remains the same as described above (Figures S1 and S2 of Supplementary File S1).

Discussion

Load statistics with no epistasis

Differences between the gene and sites models.

The main conclusion from the analytical and simulation results for the case of no epistasis is that the inbreeding load B is always much smaller under the gene model than the sites model, being effectively zero for $\bar{\gamma} \leq 20$, unless the dominance coefficient h is close to 0.5; in this case B is close to zero for both models, as expected from basic theory (Morton, Crow and Muller 1956). A similar pattern is seen with synergistic epistasis, but B becomes very small

for both the gene and sites models, and even slightly negative, when epistasis is strong (Figure 2).

The mean frequency of deleterious alleles ($\bar{q}$) is substantially smaller for the gene model than the sites model when $\bar{\gamma} \geq 100$, but its relation to $\bar{\gamma}$ does not track the behavior of *B*; for example, with $\bar{\gamma} = 100$ and $h = 0.2$, *B* and $\bar{q}$ for the gene model are approximately 6% and 30% of the values for the sites model, respectively; when $\bar{\gamma} = 1000$ and $h = 0.2$, *B* and $\bar{q}$ for the gene model are 30% and 12% of the sites model values (see Figure 1 and Tables 3 and 4).

These results raise the question of why *B* should be so small under the gene model when selection is relatively weak and *h* is small. The probable cause of this effect is the fact that haplotypes carrying – alleles are relatively common in these cases; for example, with $\bar{\gamma} = 100$ and $h = 0.2$ the mean number of – alleles per haploid genome is 4.67. When drift causes haplotypes carrying – alleles to reach high frequencies in the population, heterozygous genotypes with – alleles in *trans* will contribute substantially to the mean fitness of the population. Under the gene model, these genotypes have fitnesses that are much closer to homozygotes for the same haplotypes compared with the sites model, so that there will be a much closer correlation between the mean fitnesses of the outbred population and the inbred population than for the sites model, thereby reducing *B*. This is consistent with the fact that the relative difference in *B* between the two models diminishes with the strength of selection, and is much smaller for the two-locus deterministic case (Table 2) than for the multi-locus simulations.

Interestingly, when multiplied by the number of selected sites (*G*), the standard single-locus deterministic formula for *B*, $u[(1/h) - 2]$ (Morton et al. 1956), predicts a value of 0.015 for the case of $h = 0.2$ in Tables 3 and 4, which is not far from the simulation values for the sites model but not the gene model for $\gamma = 20$ (Table 3) and for $\bar{\gamma} = 1000$ (Table 4). The prediction for the sites model for the human-like case in Table S2 of Supplementary File S1, with a mutation rate of 2.5×10^{-7} and 2000 selected sites ($\bar{\gamma} = 850$) is 0.0015, which is fairly close to the simulation value of 0.00094, but much bigger than the value of 0.00056 for the gene model. Thus, if $N_e s$ is sufficiently large, standard deterministic theory predicts *B* quite well if the sites model is valid, but badly overpredicts *B* for the gene model.

Many of the older simulation studies of genetic and inbreeding loads have treated a gene as a single non-recombining unit subject to mutation from wild-type alleles to mutant alleles with identical properties, each of which causes a fitness reduction of *s* when

homozygous or in combination with each other, and *hs* when heterozygous with wild-type (e.g. Charlesworth et al. 1992; Wang et al. 1999; Kamran-Disfani and Agrawal 2014; Bersabé et al. 2016). This model is equivalent to a single locus with mutation at rate *Gu* from wild-type to a mutant allele with selection coefficient *s*. At first sight, this case seems equivalent to the gene model with identical selection coefficients for each mutation. However, with the relatively weak selection against many deleterious mutations suggested by population genomics data (see section 2 of the Appendix), the ratio of *Gu* to *hs* is likely to be such that most haplotypes segregating in the population carry at least one mutation. Even with relatively strong selection ($\bar{\gamma}$ =1000 and h = 0.2 in Table 4, the mean number of deleterious mutations per haploid genome for a sequence of 1000 selected sites is 1000 x 0.00091= 0.91.

The assumption that all deleterious mutations are only one step away from wild-type may thus be unsafe. This has important consequences for the predicted values of the genetic and inbreeding loads. Consider the case of a fixed selection coefficient with γ= 20 described in Table 3. In this case, the mean deleterious allele frequency per site is 0.00098 under the gene model, with a very small standard error. With the weak LD associated with these parameters (see Figure 3 of Johri and Charlesworth 2025), the distribution of the number of mutations per haploid genome closely follows a Poisson distribution, so that (disregarding variability between replicate simulations) the probability of a mutant-free haplotype is approximately exp(– 0.98) ≈ 0.38. The predicted frequency of haplotypes containing at least one mutation is thus 0.62. If this is used as the deleterious allele frequency q in the standard formulae for L and B ($L = [2h + (1 - 2h)q]sq$; $B = sq - L$), we obtain L = 0.0048 and B = 0.0014. These are substantially lower than the simulation values of L = 0.0074 and B = 0.0022 under the gene model, reflecting the fact that this calculation ignores the contributions of haplotypes carrying more than one mutation. One-step mutational models can therefore seriously underpredict the load statistics.

Implications for data on inbreeding load

Recent studies of genetic and inbreeding loads that have used population genomics data to simulate populations subject to deleterious mutations and predict genetic and inbreeding loads (reviewed by Kyriazis, Robinson and Lohmueller 2023) have mostly used the sites model. Multiplicative fitnesses across selected sites and genes have also usually been assumed in models of inbreeding load – but see Charlesworth (1998) and García-Dominguez

et al. (2018) for exceptions. Our results are for a single gene, so it is important to consider what happens if they are extrapolated to the whole genome. For convenience, we will assume purely autosomal inheritance, leading to a probable overestimation of the load statistics for the entire genome, given that the X chromosome is associated with lower loads than autosomes due to selection against hemizygous mutations in males (Wilton and Sved 1979).

To scale up to multiple genes, we assume that the sites model applies to between-gene effects for both models of single genes, due to complementation between non-allelic mutations. As in previous studies, starting with Haldane (1937), we assume multiplicative fitness effects across different genes; with small selective effects the differences from additive effects will be small in most cases. Our simulation results were based on a scaling of population size from 10^6 to 10^3; to generate realistic results for a *Drosophila*-like system, we would need to rescale back to the $N = 10^6$ values; this means dividing linear quantities in s by 10^3 and quadratic quantities such as the variance by 10^6 (see the discussion of rescaling in the section *Load statistics with no epistasis*). If we assume 14000 genes in the *Drosophila* genome (Misra et al. 2002), the individual gene linear quantities (the loads and inbreeding loads) in the tables should thus be multiplied by 14 and the variance by 0.014. This ignores any contributions from mutations in non-coding sequences, which will be considered below.

In general, let the linear and quadratic multipliers reflecting rescaling be f and f^2, and the number of genes be n. For the inbreeding load, defined as the difference between the natural logarithms of outbred mean fitness and fully inbred mean fitness, the B values in the tables need only be multiplied by nf to obtain the total inbreeding load, B_T. If the mean load L for each gene is small, the total genetic load with multiplicative effects across n genes is given by the following expression:

$$L_T \approx 1 - \exp\,(-nfL) \qquad (2)$$

Similarly, if there is no LD between different genes, using the rule that the expectation of a product of a set of independent variables is the product of their expectations, the total variance with multiplicative effects across genes is given by:

$$V_T \approx \exp(-2nfL)[\exp\left(-nf^2V_g\right) - 1] \qquad (3)$$

Given n and f, the values of the relevant statistics in the tables can easily be converted into values for the whole genome by using these relations. Using the mean values of the load

statistics in Table 4 with $\bar{\gamma}$ = 1000, Table 5 shows the values obtained with n = 14000 and f = 14; σ_T is the standard deviation corresponding to V_T. These results show that the difference between the two models has only a small effect on the expected genome-wide genetic load and genetic variance/standard deviation from variants with such small effects on fitness (mean s is only 5×10^{-4} after rescaling N), but that the gene model has a drastically smaller total inbreeding load than the sites model with h = 0.2 (0.050 versus 0.165), despite the fact that the sites model applies to between-locus fitness effects.

These findings suggest that, for the same distribution of selection and dominance coefficients, an equilibrium population is likely to exhibit a much smaller amount of inbreeding load than expected under previous theoretical calculations, if fitness effects of recessive or partially recessive mutations at individual genes obey the gene model rather than the usual sites model. This conclusion is important, because reconciling observed and predicted values of B has been the subject of much debate, with some workers asserting that the observed values can be explained by mutation-selection balance with multiplicative fitnesses (e.g., Pérez-Pereira et al. 2021; Robinson et al. 2023), while others have proposed that additional factors such as variation due to balancing selection are likely to be involved (e.g., Charlesworth 2015).

Our results suggest that the sites model of mutation and selection is likely to considerably *overpredict* the value of B, unless the mean strength of selection is so large and the distribution of selection coefficients is so tight that the population can be treated as fully deterministic. Population genomics studies of the DFE for nonsynonymous mutations in organisms such as humans (Kim et al. 2017) and *Drosophila melanogaster* (Campos et al. 2017; Johri et al. 2020) suggest that a substantial fraction of deleterious mutations are subject to the joint effects of drift and selection. If it can be shown that empirical estimates of B are incompatible with values predicted on the basis of the sites model, then contributions from balancing selection or synergistic epistasis of the type modeled by Charlesworth (1998) must be involved.

A difficulty with this approach is that the currently used methods for inferring the DFE from population genomic data assume the sites model for predicting variant frequency distributions at sites under selection. As pointed out in section 2 of the Appendix, this means that the mean strength of selection as measured by $\bar{\gamma}$ is likely to be overestimated, given that the gene model predicts lower deleterious variant frequencies than the sites model. Further research into the magnitude of this problem needs to be carried out, but the fact that it is

likely to inflate the predicted sizes of the genetic and inbreeding loads needs to be borne in mind.

Another problem is that reliable estimates of B for net fitness in natural populations are hard to obtain, with substantial disagreements among different workers concerning the typical magnitude of B (Robinson et al. 2023). This makes it difficult to compare the data on inbreeding load with theoretical predictions. Moreover, complex demographic histories, including strong bottlenecks and hidden population structure, are likely to play a role in the observed mismatch between theoretically expected and empirically observed estimates of B. Thus, more organism-specific forward simulations that take account demographic history into account are likely to be needed to understand the role of balancing selection and epistasis.

Inbreeding load in Drosophila

Here we take a somewhat different approach. An ingenious method for estimating B for net fitness under laboratory conditions for a single major chromosome of some *Drosophila* species was devised by John Sved, which is known as the balancer equilibration (BE) technique (Sved and Ayala 1970). This is based on using population cages to compete wild-type chromosomes extracted from natural populations (chosen to be homozygous viable and fertile) against homozygous lethal balancer chromosomes that suppress crossing over when heterozygous. The equilibrium frequencies of wild-type chromosomes can be used to estimate the ratio R of the mean fitness of chromosomal homozygotes to the mean fitness of chromosomal heterozygotes (reviewed by Frankham 2023). The mean value of R for chromosome 2 of *D. melanogaster* over two well-replicated experiments was 0.18±0.03, giving an estimate of $B = -\ln(R)$ of 1.7 (Frankham 2023). The BE results using the data on all three major chromosomes imply that B for the whole genome of *D. melanogaster* (including the contribution from homozygous lethal mutations) is approximately 5, so that the fitness of completely homozygous flies (exp –5) is effectively zero (Frankham 2023).

The analysis of population genetic data for *D. melanogaster* presented in sections 2-4 of the Appendix yields a prediction of $R = 0.66$ and $B = 0.42$ for chromosome 2 under the sites model. Here, possible effects of purifying selection at non-coding sites as well nonsynonymous sites and strongly selected (but non-lethal) mutations are taken into account, using fairly liberal assumptions with respect to the magnitude of B, so that the resulting values are likely to overestimate B. This value of B is based on the deterministic expressions for R and B given by Equations (A9) and (A10), an unrealistically low h value of 0.125, and an approximate correction for the effects of drift obtained by multiplying u by the proportion

of mutations with $2N_ehs < 2.5$ under a gamma distribution with shape parameter 0.3, thus removing the contribution of nearly neutral mutations (see sections 2 and 3 of the Appendix). Given that the gene model predicts a substantially lower value of B than the sites model, B = 0.42 for chromosome 2 is almost certainly an overestimate, but is less than 25% of the observed value. There thus appears to be much more inbreeding load due to deleterious mutations under laboratory conditions in *D. melanogaster* than can be accounted for purifying selection under a multiplicative fitness model, especially considering the various biases towards overpredicting the inbreeding load caused by deleterious mutations.

Why is the inbreeding load reduced when N_es is small?

Numerical methods for obtaining an exact single-locus prediction for B in a finite population have been developed by Charlesworth (2018), and extensive multi-locus simulation results using a variety of parameter values have been described in the literature (Kyriazis et al. 2023). Both of these approaches show that the level of inbreeding load caused by deleterious mutations is negatively related to N_es under a wide range of model assumptions. As has been widely discussed in the literature (see review by Robinson et al. 2023), there are two possible reasons for this effect – the "purging" of deleterious, strongly recessive variants from small populations as a result of their exposure to selection against homozygotes (Nei 1968), and the loss of variability, including the (temporary) fixation of deleterious variants (Kimura et al. 1963).

For the results presented in Tables 3 and 4, the second factor must be involved in the reduction in B with smaller N_es, since we are comparing different strengths of selection while keeping N_e and the mutation rate constant. As expected from the effects of reduced s, $\bar{q}$ changes from 0.0092 ± 0.0028 with $\bar{\gamma}$ = 1000 to 0.0383 ± 0.0053 with $\bar{\gamma}$ = 2, the opposite of what is expected with purging; the ratio of $\bar{q}$ values (equal to 4.16) is, however, much less than the 50-fold difference in $\bar{\gamma}$. A similar pattern of decrease in $\bar{q}$ with increasing $\bar{\gamma}$ is seen with h = 0.5. The explanation for the lack of proportionality between the two variables is explored in section 5 of the Appendix.

In contrast to B, the genetic load and variance in fitness are little affected by the effects of drift when selection is weak. This is illustrated by the results on the genetic loads under the sites model for different $\bar{\gamma}$ values and h = 0.2 in Table 4; the mean loads for $\bar{\gamma}$ = 1000 and $\bar{\gamma}$ = 2 are 0.0101 ± 0.001 and 0.0095 ± 0.001, respectively, and the approximate deterministic value of $2000u$ (Haldane 1937) is 0.01. Even with $\bar{\gamma}$ = 2, the load is close to the

deterministic value, in contrast to the sizeable reduction in B. The variances in fitness under the sites model for h = 0.2 and 0.5 in Table 4 are also well approximated by the deterministic value of $4000uh\bar{s}$.

The weak effects of drift on the load and variance in fitness are at first sight surprising, given that with $\bar{\gamma}$ = 2 a substantial proportion of the distribution of selection coefficients falls below the nearly neutral threshold of $h\gamma < 2.5$ described above, where large departures from the deterministic values are expected. However, this part of the distribution involves very small selection coefficients, and thus makes an extremely small contribution to the genetic load and the variance in fitness.

The role of epistasis

Epistasis and inbreeding load

One important aspect of our simulation results with synergistic epistasis is that B for the gene and sites models is strongly reduced by even a small amount of epistasis when h < ½, despite only a small reduction in load; there is a much smaller effect on B under the gene model, where it is always small (Figure 2). This reduction in B with increasing epistasis is in sharp contrast to the model of Charlesworth (1998), which predicts that synergistic epistasis causes an *increase* in B when the logarithm of fitness is a quadratic, decreasing function of the number of deleterious mutations. As shown in section 6 of the Appendix, when only homozygous genotypes are considered, the pairwise epistasis model used here is identical to the quadratic model. The opposite behavior of the two models is thus surprising at first sight.

The explanation lies in the fact that the model of Charlesworth (1998) assumes that $h > 0$ and the quadratic term for heterozygous mutations is proportional to the negative of the product of h^2 and the square of the number of heterozygous mutations (n^2). In contrast, the pairwise epistasis model assumes that this term is proportional to $-hn^2$. This means that the fitness of the outbred population is relatively insensitive to the quadratic fitness term under the first model, since most mutations are present in heterozygotes. However, the fully inbred genotypes have strongly reduced fitnesses due to the quadratic term, resulting in an increase in B. Under the pairwise model, there is a sharp decline in $\bar{q}$ with increasing ϵ, while the load is relatively unaffected (compare Figures 1 and 2). Since B is strongly determined by $\bar{q}$, the net result is that synergistic epistasis reduces B, even when $h = 0$.

The consequences of epistasis for inbreeding depression are thus highly model dependent, and it cannot be assumed that synergistic epistasis necessarily leads to increased

inbreeding depression compared with the multiplicative fitness model, as has been done in the past (Crow 1970; Charlesworth 1998; García-Dominguez et al. 2018). The pairwise model has the advantage of greater flexibility in modeling the fitnesses of heterozygotes and differences in selection coefficients between sites, whereas the quadratic model is forced to make somewhat arbitrary assumptions about the effects of dominance (for details, see Charlesworth 1998). However, it is important to note that we have only developed this model for within-gene interactions; the nature of epistatic interactions between deleterious mutations in different genes and their consequences for the load statistics remain to be explored.

Acknowledgments

This research was conducted using computational resources provided by ITS Research Computing at the University of North Carolina at Chapel Hill. PJ was funded by the National Institute of General Medical Sciences of the National Institutes of Health under the award number R35GM154969.

Literature Cited

Assaf ZJ, Tilk S, Siegal ML, Petrov DA. 2017. Deep sequencing of natural and experimental populations of *Drosophila melanogaster* reveals biases in the spectrum of new mutations. Genome Res. 27:1988-2000.

Barton NH. 1995. A general model for the evolution of recombination. Genet. Res. 65:123-144.

Bertorelli G, Raffini M, Bortoluzzi C, Ianucci A, Trucchi E, Morales HE, Van Oosterhout C. 2022. Genetic load: genomic estimates and applications in non- model animals. Nat. Rev. Genet. 23:492-503.

Bersabé D, Caballero A, Pérez-Figueroa A, García-Dorado A. 2016. On the consequences of purging and linkage on fitness and genetic diversity. G3 6:171-181.

Campos JL, Halligan DL, Haddrill PR, Charlesworth B. 2014. The relation between recombination rate and patterns of molecular evolution and variation in *Drosophila melanogaster*. Mol. Biol. Evol. 31:1010-1028.

Campos JL, Zhao L, Charlesworth B. 2017. Estimating the parameters of background selection and selective sweeps in *Drosophila* in the presence of gene conversion. Proc Natl Acad Sci USA. 114(24):E4762–E4771.

Campos JL, Charlesworth B. 2019. The effects on neutral variability of recurrent selective sweeps and background selection. Genetics 212:287-303.

Casillas S, Barbadilla A, Bergman CM. 2007. Purifying selection maintains highly conserved noncoding sequences in Drosophila. Mol. Biol. Evol. 24:2222-2234.

Charlesworth B. 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. Genet Res. 55(3):199-221.

Charlesworth B. 1998. The effect of synergistic epistasis on the inbreeding load. Genet. Res. 71:85-89.

Charlesworth B. 2015. Causes of natural variation in fitness: evidence from studies of Drosophila populations. Proc. Natl. Acad. Sci. USA 12:1662-1669.

Charlesworth B. 2018. Mutational load, inbreeding depression and heterosis in subdivided populations. Mol. Ecol. 24:4991-5003.

Charlesworth D, Morgan MT, Charlesworth B. 1992. The effect of linkage and population size on inbreeding depression due to mutational load. Genet. Res. 59:49-61.

Crow JF. 1970. Genetic loads and the cost of natural selection. In: Kojima K, editor. Mathematical Topics in Population Genetics. Berlin: Springer-Verlag. p. 128-177.

Ewens WJ. 2004. Mathematical Population Genetics. 1. Theoretical Introduction. New York: Springer.

Eyre-Walker A, Keightley PD. 2009. Estimating the rate of adaptive mutations in the presence of slightly deleterious mutations and population size change. Mol. Biol. Evol. 26:2097-2108.

Fincham JRS, Day PR, Radford A. 1979. Fungal Genetics. Berkeley, CA: University of California Press.

Frankham R. 2023. Effects of genomic homozygosity on total fitness in an invertebrate: lethal equivalent estimates for *Drosophila melanogaster*. Cons. Genet. 24:193-201.

García-Dominguez S, García C, Quesada H, Caballero A. 2018. Accelerated inbreeding depression suggests synergistic epistasis for deleterious mutations in *Drosophila melanogaster*. Heredity 123:709-722.

García-Dorado A, Hedrick PW. 2023. Some hope and many concerns on the future of the vaquita. Heredity 130:179-182.

Haerty W, Ponting CP. 2014. No gene in the genome makes sense except in the light of evolution. Ann. Rev. Genomics Hum. Genet. 15:71-92.

Haldane JBS. 1937. The effect of variation on fitness. Am. Nat. 71:337-349.

Halligan DL, Keightley PD. 2006. Ubiquitous selective constraints in the Drosophila genome revealed by a genome-wide sequence comparison. Genome Res. 16:875-884.

Hawley RS, Gilliland WD. 2006. Sometimes the result is not the answer: The truths and the lies that come from using the complementation test. Genetics 174:5-15.

Jackson BC, Campos JL, Haddrill PR, Charlesworth B, Zeng K. 2017. Variation in the intensity of selection on codon bias over time causes contrasting patterns of base composition evolution in Drosophila. Genome Biol Evol. 9:102–123.

Jinks JL. 1983. Biometrical genetics of heterosis. In: Frankel R, editor. Heterosis: Reappraisal of Theory and Practice. Berlin: Springer-Verlag. p. 1-46.

Johri P, Charlesworth B. 2025. A gene-based model of fitness and its implications for genetic variation: linkage disequilibrium. (*companion paper*).

Johri P, Charlesworth B, Jensen JD. 2020. Toward an evolutionarily appropriate null model: Jointly inferring demography and purifying selection. Genetics 215:173-192.

Kamran-Disfani A, Agrawal AF. 2014. Selfing, adaptation and background selection in finite populations. J. Evol. Biol. 27:1360-1371.

Kardos M, Armstrong EE, Ftzpatrick SW, S. H, Hedrick PW, Miller DE, Tallmon DA, Funk WC. 2021. The crucial role of genome-wide genetic variation in conservation. Proc. Natl. Acad. Sci. USA 118:e2104642118.

Kim BY, Huber CD, Lohmueller KE. 2017. Inference of the distribution of selection coefficients for new nonsynonymous mutations using large samples. Genetics. 206:345–361.

Kimura M, Maruyama T, Crow JF. 1963. The mutation load in small populations. Genetics 48:1303-1312.

Kondrashov AS. 1988. Deleterious mutations and the evolution of sexual reproduction. Nature 336:435-440.

Kyriazis C. C., R. K. Wayne, and K. E. Lohmueller, 2021 Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. Evolution Letters 5: 33–47.

Kyriazis CC, Robinson JA, Nigenda-Morales SF, Beichman AC, Rojas-Bracho L, Robertson KM, et al. 2023. Models based on best-available information support a low inbreeding load and potential for recovery in the vaquita. Heredity 130:183-187.

Kyriazis CC, Robinson JA, Lohmueller KE. 2023. Using computational simulations to model deleterious variation and genetic load in natural populations. Am. Nat. 302:737-752.

Manna F, Martin G, Lenormand T. 2011. Fitness landscapes: An alternative theory for the dominance of mutation. Genetics 189:923-937.

McVean GAT, Charlesworth B. 1999. A population genetic model for the evolution of synonymous codon usage: patterns and predictions. Genet. Res. 74:145-158.

Misra S, Crosby MA, Mungall CJ, Matthews BB, Campbell KS, et al. 2002. Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. Gen. Biol. 3:Research0083.

Morton NE, Crow JF, Muller HJ. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. Proc. Natl. Acad. Sci. USA 42:855-863.

Nei M. 1968. The frequency distribution of lethal chromosomes in finite populations. Proc. Natl. Acad. Sci. USA 60:517-524.

Pérez-Pereira N, Pouso R, Rus A, López-Cortegano E, García-Dorado A, Quesada H, Caballero A. 2021. Long-term exhaustion of the inbreeding load in *Drosophila melanogaster*. Heredity 127:373-383.

Robinson JA, Kyriazis CC, Yuan SC, Lohmueller KE. 2023. Deleterious variation in natural populations and implications for conservation genetics. Annu. Rev. Animal Biosci. 11:93-114.

Roze D. 2021. A simple expression for the strength of selection on recombination generated by interference among mutations. Proc. Natl. Acad. Sci. USA 118:e2022805118.

Sved JA, Ayala FJ. 1970. A population cage test for heterosis in *Drosophila pseudoobscura*. Genetics 66:97-113.

Teixeira JC, Huber CD. 2021. The inflated significance of neutral genetic diversity in conservation genetics. Proc. Natl. Acad. Sci. USA 118:e2015096118.

Wang J, Hill WG, Charlesworth D, Charlesworth B. 1999. Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. Genet. Res. 74:165-178.

Wang Y, McNeil P, Abdulazeez R, Pascual M, Johnston SE, Keightley PD, Obbard DJ. 2023. Variation in mutation, recombination, and transposition rates in *Drosophila melanogaster* and *Drosophila simulans*. Genome Res. 33:587-598.

Wilton AN, Sved JA. 1979. X-chromosomal heterosis in *Drosophila melanogaster*. Genet. Res. 34:303-315.

Wright S. 1968. Evolution and the Genetics of Populations. Vol.1. Genetic and Biometric Foundations. Chicago, IL: University of Chicago Press.

Appendix

1. The load statistics for the two-site deterministic model

Using the fitness model of Table 1, and the notation for the haplotype frequencies described in the main text, the genetic load under the gene model is given by:

$$L_g = \{4zh[x + y(1+\epsilon)] + 4x\left[x + \tfrac{3}{2}y(1+\epsilon)\right] + 2y^2(1+\epsilon)\}s \qquad \text{(A1)}$$

The inbreeding load is:

$$B_g = 2[x + y(1+\epsilon)]s - L_g \qquad \text{(A2)}$$

and the variance in fitness is:

$$V_g = \left\{4zh^2[x + y(1+\epsilon)^2] + 4x\left[x + \tfrac{9}{4}y(1+\epsilon)^2\right] + 4y^2(1+\epsilon)^2\right\}s^2 - L_g^2 \qquad \text{(A3)}$$

Under the sites model, the corresponding formulae are:

$$L_s = \{4zh[x + y(1+\epsilon)] + 2x[x + 2xh(1+\epsilon) + 2y(1+h)(1+\epsilon)] + 2y^2(1+\epsilon)\}s \qquad \text{(A4)}$$

$$B_s = 2[x + y(1+\epsilon)]s - L_s \qquad \text{(A5)}$$

$$V_s = \{4zh^2[x + y(1+\epsilon)^2] + 2x[x + 4xh^2(1+\epsilon)^2 + 2y(1+h)^2(1+\epsilon)^2] + 4y^2(1+\epsilon)^2\}s^2 - L_s^2 \qquad \text{(A6)}$$

In the case of complete recessivity, Equations (A1) and (A4) for the genetic loads under the gene and sites models without epistasis reduce to the following expressions:

$$L_g = \{4x(x+y) + 2y^2\}s \qquad \text{(A7)}$$

$$L_s = \{2x(x+2y) + 2y^2\}s \qquad \text{(A8)}$$

Since $y << x$, the dominant terms in these equations are $4x^2s$ and $2x^2s$, respectively. It follows that, when $h = 0$, the load is higher for the gene than the sites model. By continuity, this will also be true when h is sufficiently small, as can be seen in Table 2 for the case of $h = 0.01$.

2. Estimating the level of nonsynonymous diversity in *Drosophila melanogaster* contributed by significantly non-neutral mutations

Campos et al. (2017) fitted a gamma distribution to data on the site frequency spectra of NS variants in a sample of genomes from a Rwandan population of *D. melanogaster*, together with divergence from *D. yakuba*. The method that was used (DFE-alpha of Eyre-Walker and

Keightley 2009) assumes that mutations are semi-dominant (h = 0.5), for which the sites and gene models are equivalent in the absence of epistasis. The selection coefficient against a heterozygous mutation was thus assumed to be $t = ½s$, implying that their $\gamma = 4N_et$ is approximately equivalent to $4N_ehs$, given that selection against strongly selected mutations is predominantly against heterozygotes when mating is random. Their Figure S3 suggested an overall value of approximately 1250 for $2N_ehs$ for newly arising NS mutations, with a shape parameter of 0.3.

This estimate needs to be treated with caution, given that the method implicitly uses the sites model. If all or most mutations in coding sequences obey the gene model, then the fact that it predicts a substantially lower mean allele frequency than the sites model when $\gamma \geq 20$ (Table 3) implies that the strength of selection will be substantially overestimated by this method, and other similar population genomics approaches. The following analysis is therefore likely to overestimate the abundance of mutations that are under sufficiently strong selection that they behave quasi-deterministically, and hence will overestimate their contribution to the genetic and inbreeding loads.

A conservative threshold for mutations to behave deterministically is $2N_ehs \geq 2.5$ (Campos and Charlesworth 2019). Using numerical integration of the incomplete gamma function and the fact that $\Gamma(0.3) = 2.886$, this implies that the proportion of significantly selected new NS mutations is 0.91 (i.e., the proportion of NS mutations that are nearly neutral is 0.09), and yields a value of 0.0132 for the mean of $1/(2N_ehs)$ for such mutations.

From Campos et al. (2014), the mean synonymous site diversity, π_S, for autosomal loci in regions with normal recombination rates was 0.0141; with a mutation rate per basepair of approximately 4×10^{-9}, which is an approximate consensus value from the results in Table 1 of Assaf et al. (2017; see also Wang et al. 2023). The corresponding value of N_e is 8.81×10^5. There is probably some weak selection on these sites due to codon usage bias, and a value of $N_e = 10^6$ is a reasonable approximation to correct for this (Jackson et al. 2017). Using this value of N_e, together with the above estimate of the mean of $1/(2N_ehs)$ the mean of $1/(hs)$ for significantly selected NS sites is 2.64×10^4 (i.e., the harmonic mean of hs at these sites is 3.78×10^{-5}). Under mutation-selection balance, the expected frequency of the mutant allele at a site is $q = u/(hs)$, and the expected mean diversity at significantly selected NS sites is then $\pi_{Ns} \approx 2E\{q\} = 2 \times 4 \times 10^{-9} \times 2.64 \times 10^4 = 2.11 \times 10^{-4}$.

A check on these calculations is as follows. The observed mean NS site diversity, π_N, was 0.00143. If the proportion of NS mutations that are neutral or nearly neutral is p_n, then π_N

$\approx p_n\pi_S + (1 - p_n)\,\pi_{Ns}$. This gives $p_n = (\pi_N - \pi_{Ns})/(\pi_S - \pi_{Ns})$. The predicted value of π_{Ns} is 1.06×10^{-4}, giving $p_n \approx (0.00143 - 2.11 \times 10^{-4})/(0.0141 - 2.11 \times 10^{-4}) = 0.088$, which agrees well with the population genomics estimate of 0.09 obtained above. A similar value is given in Figure 6 of Haerty and Ponting (2014).

3. The mean number of slightly deleterious variants per *D. melanogaster* chromosome 2 and the mutation rate to non-lethal deleterious alleles

Using the above estimates, the mean frequency per nucleotide site of significantly deleterious NS variants is approximately one-half of π_{Ns}, i.e. 1.05×10^{-4}. Assuming a mean of 1500 exonic sites per gene, and a fraction of exonic sites that are NS equal to 0.7 (Campos and Charlesworth 2019), the mean number of significantly deleterious NS variants per gene in a random sample from the population is $1500 \times 0.7 \times 0.815 \times 10^{-4} = 0.110$. Given that chromosome 2 is 37% of the genome, and that there are approximately 14000 genes in the fly genome (Misra et al. 2002), the estimated mean number of significantly deleterious NS variants per 2nd chromosome is approximately 571.

This is much smaller than the estimate using π_N itself, because the majority of NS variants segregating in the population are nearly neutral. If significantly selected non-coding variants are included, this number could be increased by a factor of two or even three, given that these contribute 3.8 times as much DNA as coding sequences and that approximately 50% of noncoding sites in *Drosophila* are under selective constraint (Halligan et al. 2006), although there is evidence for weaker selection against many non-coding sequence mutations than against NS mutations (Casillas et al. 2007; Haerty and Ponting 2014; Campos et al. 2017), reducing the fraction of segregating variants that are under significant selection.

The results of mutation accumulation experiments on *D. melanogaster* seem to require a small but significant contribution of mutations with homozygous selection coefficients that are much larger than those estimated from the population genomic data (Charlesworth 2015). Consensus values of the data from mutation accumulation experiments suggest a per 2nd chromosome mutation rate of 0.032 for large effect (but non-lethal) viability mutations. From Equation (A10) below, the corresponding value of R, the ratio of the mean fitness of non-lethal chromosome 2 homozygotes to that of heterozygotes, would be approximately 0.936 with a mutation rate of 0.032 and $h = 0.25$, not substantially different from 1.

4. Relating the population genomics approach to the classical expression for inbreeding depression

The ratio of the mean fitnesses of chromosomal homozygotes and heterozygotes under the multiplicative fitness model with deterministic mutation selection balance is:

$$R = \exp - n_s \mathrm{E}\{qs(1 - 2h)\} \qquad \text{(A9)}$$

where n_s is the mean number of sites on the chromosome that are subject to significant selection, q is the frequency of a deleterious variant at a given NS site, s and h are the selection and dominance coefficients, and E{} denotes the expectation over all NS sites.

We have $q \approx u/(hs)$ for strongly selected mutations, so that:

$$R \approx \exp - n_s \mathrm{E}\{u[(1/h) - 2]\} \qquad \text{(A10)}$$

This is the exponential of the negative of the inbreeding load (B), under mutation-selection balance with partial recessivity, assuming multiplicative fitnesses across loci (Morton et al. 1956). This formula applies even when there are significant effects of drift on q, provided that $2N_ehs$ is sufficiently large and mating is random, since the distribution of q then follows a gamma distribution with mean $u/(hs)$ and variance $u/[4N_e(hs)^2]$ (Nei 1968).

If 91% of NS sites are under significant selection, as estimated in section 1, for chromosome 2 of *D. melanogaster* we have n_s = 0.37 x 0.91 x 14000 x 0.70 x 1500 = 4.95 x 10^6. With a mutation rate of $u = 4 \times 10^{-9}$ and the value of h = 0.25 suggested by Manna *et al.* (2011), this gives R = exp – 0.0396 = 0.961. If we multiply n_s by three, to include selection on non-coding sites, R = 0.888. The strength of selection is irrelevant to this calculation, so that the mean of $2N_ehs$ could be very different from that used above but the value of R would be the same.

Variation in h around 0.25 would, of course, increase these estimates. If there were a two-fold increase in $1/h$, corresponding approximately to a squared coefficient of variation of h of 1 (as under an exponential distribution), or to a fixed value of h = 0.125, then $B = -\ln R$ would be increased to three times that corresponding to the above value, giving B = 0.356, R = 0.700 (this is likely to be an underestimate, since small values of h will reduce the proportion of significantly selected mutations). If the above estimate of the contribution from major effect mutations is also included, R becomes 0.656, B = 0.421.

A check on the use of the cut-off of $2N_ehs < 2.5$ is provided by the simulation results in Table 4. With $\bar{\gamma} = 1000$ and $h = 0.2$, the gamma distribution with shape parameter 0.3 yields a proportion of mutations of 0.83 that exceed the cut-off. With a mutation rate of 0.5 10^{-5} per site and 1000 sites in a gene, the mutation rate per gene adjusted for effectively neutral mutations is 0.005 x 0.83 = 0.00415. Using Equation (A10) with $h = 0.2$, the B value for a single gene of 0.00415 x 3 = 0.0124, compared with the simulation value of 0.0118 ± 0.0004 for the sites model. Similar calculations for $\bar{\gamma} = 100$, 20 and 2 give the respective predicted B values of 0.0091, 0.0055 and 0.0040; the corresponding simulation values for the sites model are 0.0098 ± 0.0001, 0.0071 ± 0.0001 and 0.00314 ± 0.0001. This simple approximation thus gives good fits to the simulation values for B, even for small mean values of $\bar{\gamma}$, accurately capturing the reduction in B for partially recessive mutations when N_es is reduced, as expected from previous theory (e.g., Charlesworth 2018).

This calculation can be made even simpler by noting that the cumulative gamma distribution using the scaled variable $y = \alpha\gamma/\bar{\gamma}$ for $y << 1$ between 0 and y_0 is approximated by:

$$\Phi(y_0) \approx \Gamma(\alpha)^{-1} \int_0^{y_0} y^{\alpha-1}\,\mathrm{d}y = \Gamma(\alpha)^{-1}\alpha^{-1}y_0^{\alpha} \qquad \text{(A11)}$$

If $\alpha << 1$, $\Gamma(\alpha) \approx \alpha^{-1}$, so that:

$$\Phi(y_0) \approx y_0^{\alpha} \qquad \text{(A12)}$$

With $\alpha = 0.3$, this approximation is quite accurate even for y_0 as large as 0.25, so that $1 - y_0^{0.3}$ can be used for most practical purposes to determine the proportion of sites that are under significant selection, with the appropriate choice of y_0.

The remaining question is whether nearly neutral sites, which are ignored in these calculations, contribute significantly to inbreeding depression. As just described, the effect of drift away from mutation-selection equilibrium is to reduce B, so that use of the deterministic formula will overestimate it for nearly neutral sites. If 9% of new mutations fall into this category, their maximal contribution to B would be 0.09/0.91 = 0.10 of the values calculated above, i.e, a fairly trivial amount.

5. The relation between mean q and γ

The fact that the mean frequencies of deleterious alleles are not inversely proportional to the strength of selection when γ is relatively small can be understood by examining the case of h = 0.5 and a fixed value of γ (see Table 3). In the absence of LD, the predicted value of $\bar{q}$ with our scaled mutation rate of $4Nu$ = 0.02 with γ = 2 should be closely approximated by the Li-Bulmer equation, which predicts the proportion of sites fixed for the favorable allele in a diallelic mutation-selection model (McVean and Charlesworth 1999). With no mutational bias, this gives:

$$\bar{q} \approx \frac{\exp(-\gamma)}{1+\exp(-\gamma)} \qquad \text{(A13)}$$

With γ = 2, $\bar{q} \approx 0.12$, compared with the simulation value of 0.0236 ± 0.0039 for the sites model. For $\gamma >> 1$, $\bar{q}$ is approximated by the deterministic value $u/(hs) = 2u/s$, which is equal to 0.0013 with γ = 20, compared with 0.00079 ± 0.00014 for the sites model. The observed ratio of $\bar{q}$ values is 30, compared with a predicted range of 0.12/0.0013 = 92. This suggests that selection is more effective in the multi-site context than for a single site in isolation when selection is relatively weak. At first sight, this is puzzling, because Table 2 of Johri and Charlesworth (2025) shows no evidence for significant LD with these parameters.

The probable explanation is as follows, which is based on the two-sites model (Table 1). Most of the contributions to $\bar{q}$ with these parameters come from fixed rather than segregating sites (McVean and Charlesworth 1999). Any effects of LD on $\bar{q}$ come from transitions between fixed pairs of sites where two variants are segregating simultaneously. Transitions of the type + – to – – and + + (or vice versa) are likely to only temporarily involve two simultaneous segregating sites, and can be ignored as far as the effects of LD on the outcome of selection are concerned, although they are the most common type of transition and source of variability. Transitions of the type + + to – – must involve a first mutation of the type + + to + – , followed by a + – to – – mutation; in the absence of recombination, this creates a two-site polymorphism during the transition of the type discussed in the section *Effects of weighting towards low frequencies on LD statistics* in the companion paper Johri and Charlesworth (2025). This has $D > 0$ when the sign of D is defined by selective status. This enhances the efficacy of selection, and hence reduces the chance of this transition. Similarly, transitions of the type – – to + + have $D > 0$, and their probability of success is thereby increased over the single-site expectation. Thus, despite the lack of overall positive LD, the tendency for selection to favor + versus – alleles is enhanced by linkage, thereby

reducing $\bar{q}$ below single-locus expectation. This effect is likely to be largest when selection is relatively weak, when transitions of these two types are relatively common, explaining why the ratio of $\bar{q}$ values for weak versus strong selection is less than might be expected.

6. Relating two models of epistasis

A widely used model of epistasis among deleterious mutations assumes that either fitness or the logarithm of fitness has a quadratic relation to the number of deleterious mutations (Crow 1970; Charlesworth 1990, 1998). This raises the question of its relation to the pairwise interaction model used in the present paper, which we now examine. For simplicity, consider the fitness of homozygotes only, with fixed selective effects across sites. Under the quadratic model, the fitness of an individual carrying *n* mutations can be described by a function of the following form:

$$w_n = 1 - (\alpha n + \beta n^2) \qquad \text{(A14)}$$

Under the pairwise model, we have:

$$w_n = ns + n(n-1)\epsilon s = ns(1-\epsilon) + n^2\epsilon s \qquad \text{(A15)}$$

Equating the coefficients of *n* and n^2 in these two equations gives the relations

$$\alpha = (1-\epsilon)s, \quad \beta = \epsilon s \qquad (\epsilon \leq 1) \qquad \text{(A16)}$$

This result implies that the quadratic model inherently involves pairwise epistasis, at least in its simplest realization, when only homozygotes are considered. The pairwise model has the advantage of greater flexibility in modeling the fitnesses of heterozygotes and differences in selection coefficients between sites, whereas the quadratic model is forced to make somewhat arbitrary assumptions about the effects of dominance (Charlesworth 1998).

FIGURES AND TABLES

Table 1. Models of the fitness effects of two sites with a fixed selection coefficient

The "sites" model

	++	+ -	- +	- -
+ +	1	$1 - hs$	$1 - hs$	$1 - 2hs - e_1$
- +	$1 - hs$	$1 - 2hs - e_1$	$1 - s$	$1 - s - hs - e_2$
+ -	$1 - hs$	$1 - s$	$1 - 2hs - e_1$	$1 - s - hs - e_2$
- -	$1 - 2hs - e_1$	$1 - s - hs - e_2$	$1 - s - h\,s - e_2$	$1 - 2s - e_3$

$e_1 = 2\epsilon hs$; $e_2 = \epsilon(1 + h)s$; $e_3 = 2\epsilon s$.

The "gene" model

	+ +	+ -	- +	- -
+ +	1	$1 - hs$	$1 - hs$	$1 - 2hs - e_1$
- +	$1 - hs$	$1 - s$	$1 - s$	$1 - 1.5s - e_2$
+ -	$1 - hs$	$1 - s$	$1 - s$	$1 - 1.5s - e_2$
- -	$1 - 2hs - e_1$	$1 - 1.5s - e_2$	$1 - 1.5s - e_2$	$1 - 2s - e_3$

$e_1 = 2\epsilon hs$; $e_2 = 1.5\epsilon s$; $e_3 = 2\epsilon s$.

Table 2. Population genetic statistics for the gene and sites models with different values of the dominance coefficient (h) and epistasis coefficient (ϵ)

$h = 0.01$

	Gene model				Sites model			
ϵ	q_r	L	Var	B	q_r	L	Var	B
0	0.768	0.535	0.496	0.0577	1.045	0.548	0.456	0.0821
0.02	0.767	0.535	0.498	0.0577	1.045	0.548	0.458	0.0821
0.04	0.765	0.535	0.500	0.0577	1.043	0.548	0.459	0.0821
0.08	0.763	0.535	0.503	0.0578	1.039	0.548	0.462	0.0821
0.16	0.759	0.535	0.511	0.0578	1.032	0.548	0.468	0.0820
0.32	0.751	0.535	0.525	0.0578	1.021	0.548	0.480	0.0820
0.64	0.742	0.535	0.554	0.0578	1.004	0.548	0.504	0.0818

$h = 0.2$

	Gene model				Sites model			
ϵ	q_r	L	Var	B	q_r	L	Var	B
0	0.919	0.904	0.281	0.716	1.005	0.944	0.241	0.811
0.02	0.919	0.904	0.281	0.716	1.002	0.941	0.242	0.810
0.04	0.918	0.904	0.282	0.716	1.000	0.940	0.242	0.809
0.08	0.916	0.904	0.283	0.716	0.996	0.939	0.243	0.806
0.16	0.914	0.904	0.284	0.716	0.988	0.936	0.246	0.800
0.32	0.910	0.904	0.288	0.716	0.974	0.932	0.250	0.789
0.64	0.904	0.904	0.294	0.716	0.949	0.924	0.262	0.768

$h = 0.5$

	Gene model				Sites model			
ϵ	q_r	L	Var	B	q_r	L	Var	B
0	1.000	1.000	0.480	1.000	1.000	1.000	0.480	1.000
0.02	0.999	1.000	0.481	1.000	0.999	0.999	0.481	1.000
0.04	0.999	1.000	0.481	1.000	0.997	0.999	0.482	1.000
0.08	0.998	1.000	0.481	1.000	0.995	0.998	0.483	1.000
0.16	0.997	1.000	0.482	1.000	0.991	0.997	0.486	1.001
0.32	0.995	1.000	0.484	1.000	0.983	0.994	0.492	1.001
0.64	0.992	1.000	0.487	1.000	0.969	0.988	0.506	1.001

The mutation rate is $u = 0.0001$ and the selection coefficient is $s = 0.01$ ($u/s = 0.01$), The genetic load L and inbreeding load B are expressed as their ratios with respect to their single-locus deterministic values when $u >> hs$, i.e. when the frequency of A_2 is $u/(hs)$; note that the deterministic value of B for $h = 0.5$ is $B = 0$. The genetic variance (Var) is expressed as the ratio of hV/s^2 to $4u/s$, to ensure that it lies between 0 and 1. The equilibrium frequency of deleterious mutations is expressed as its ratio to the value for a single locus with general u and hs (Equation A4 of Johri and Charlesworth 2025), denoted here by q_r.

Table 3. The load statistics for the gene and sites models with two fixed values of $\gamma = 2Ns$. 1000 selected sites were simulated, with $Nu = 0.005$ and $Nr = 0.01$, where N is the population size ($N = 1000$), u is the mutation rate per site/generation, and r is the rate of crossing over between adjacent sites. Details of how to estimate the load statistics are given in the Methods section. The means and standard errors (SEs) for 1000 replicate simulations are shown.

$\gamma = 2$		***h*=0.0**		***h*=0.2**		***h*=0.5**	
		Gene model	**Sites model**	**Gene model**	**Sites model**	**Gene model**	**Sites model**
genetic load	mean	0.02238	0.01826	0.02260	0.02122	0.02289	0.02273
	SE	0.00012	0.00010	0.00012	0.00011	0.00012	0.00012
inbreeding load	mean	0.00000	0.01201	-0.00001	0.00539	-0.00001	0.00000
	SE	0.00001	0.00007	0.00001	0.00004	0.00001	0.00001
variance in fitness	mean	2.98×10^{-6}	2.15×10^{-5}	2.84×10^{-6}	5.97×10^{-6}	2.93×10^{-6}	2.98×10^{-6}
	SE	4.16×10^{-8}	2.53×10^{-7}	3.72×10^{-8}	7.97×10^{-8}	4.32×10^{-8}	4.23×10^{-8}
allele frequency	mean	0.01100	0.02246	0.01122	0.01639	0.01115	0.01124
	SE	0.00008	0.00015	0.00009	0.00012	0.00008	0.00009
$\gamma = 20$							
genetic load	mean	0.00563	0.01296	0.00741	0.01140	0.01068	0.01077
	SE	0.00006	0.00016	0.00007	0.00012	0.00009	0.00009
inbreeding load	mean	0.00385	0.05479	0.00220	0.01256	0.00001	-0.00004
	SE	0.00004	0.00041	0.00003	0.00010	0.00003	0.00003
variance in fitness	mean	7.20×10^{-5}	2.89×10^{-4}	5.67×10^{-5}	3.87×10^{-5}	4.88×10^{-5}	4.95×10^{-5}
	SE	7.27×10^{-7}	4.62×10^{-6}	5.52×10^{-7}	4.69×10^{-7}	4.11×10^{-7}	4.56×10^{-7}
allele frequency	mean	0.00096	0.00707	0.00098	0.00241	0.00109	0.00107
	SE	0.00001	0.00006	0.00001	0.00002	0.00001	0.00001

Table 4. The load statistics under the gene and sites models for a gamma distribution of scaled selection coefficients with shape parameter 0.3 and mean $\bar{\gamma} = 2N\bar{s}$. 1000 selected sites were simulated, with $Nu = 0.005$ and $Nr = 0.01$, where N is the population size ($N = 1000$), u is the mutation rate per site/generation, and r is the rate of crossing over between adjacent sites. Details of how to estimate the load statistics are given in the Methods section. The means and standard errors (SEs) for 1000 replicate simulations are shown.

$\bar{\gamma} = 2$		***h*=0.0**		***h*=0.2**		***h*=0.5**	
		Gene model	**Sites model**	**Gene model**	**Sites model**	**Gene model**	**Sites model**
genetic load	mean	0.00976	0.00860	0.00978	0.00951	0.00980	0.00978
	SE	0.00008	0.00008	0.00008	0.00008	0.00008	0.00008
inbreeding load	mean	0.00001	0.00758	0.00001	0.00314	0.00000	0.00002
	SE	0.00001	0.00007	0.00001	0.00003	0.00001	0.00001
variance in fitness	mean	4.33×10^{-6}	1.15×10^{-5}	4.26×10^{-6}	4.21×10^{-6}	4.20×10^{-6}	4.15×10^{-6}
	SE	7.64×10^{-8}	2.31×10^{-7}	7.55×10^{-8}	8.27×10^{-8}	8.02×10^{-8}	6.79×10^{-8}
allele frequency	mean	0.01340	0.01933	0.01347	0.01600	0.01356	0.01352
	SE	0.00011	0.00014	0.00011	0.00012	0.00010	0.00011
$\bar{\gamma} = 20$							
genetic load	mean	0.01098	0.01063	0.01109	0.01157	0.01141	0.01121
	SE	0.00009	0.00010	0.00009	0.00010	0.00009	0.00009
inbreeding load	mean	0.00002	0.02960	-0.00003	0.00714	0.00003	-0.00002
	SE	0.00003	0.00026	0.00003	0.00007	0.00003	0.00003
variance in fitness	mean	4.89×10^{-5}	1.05×10^{-4}	4.85×10^{-5}	2.75×10^{-5}	4.93×10^{-5}	4.84×10^{-5}
	SE	7.22×10^{-7}	2.02×10^{-6}	7.24×10^{-7}	4.53×10^{-7}	7.88×10^{-7}	7.43×10^{-7}
allele frequency	mean	0.00739	0.01231	0.00748	0.00907	0.00741	0.00745
	SE	0.00007	0.00010	0.00007	0.00008	0.00007	0.00007

$\bar{\gamma} = 100$							
genetic load	mean	0.00768	0.00968	0.01026	0.01110	0.01089	0.01087
	SE	0.00009	0.00010	0.00008	0.00009	0.00009	0.00009
inbreeding load	mean	0.00158	0.06449	0.00006	0.00977	-0.00008	-0.00008
	SE	0.00010	0.00062	0.00007	0.00013	0.00006	0.00006
variance in fitness	mean	2.57×10^{-4}	3.78×10^{-4}	2.48×10^{-4}	1.13×10^{-4}	2.43×10^{-4}	2.45×10^{-4}
	SE	4.28×10^{-6}	8.76×10^{-6}	3.96×10^{-6}	1.87×10^{-6}	4.01×10^{-6}	3.85×10^{-6}
allele frequency	mean	0.00298	0.00777	0.00471	0.00568	0.00482	0.00470
	SE	0.00007	0.00007	0.00005	0.00006	0.00005	0.00005
$\bar{\gamma} = 1000$							
genetic load	mean	0.00563	0.00851	0.00828	0.01014	0.01061	0.01066
	SE	0.00008	0.00013	0.00008	0.00008	0.00009	0.00009
inbreeding load	mean	0.01992	0.25320	0.00359	0.01176	-0.00005	0.00035
	SE	0.00059	0.00332	0.00031	0.00035	0.00027	0.00026
variance in fitness	mean	0.00398	0.00754	0.00264	0.00102	0.00345	0.00347
	SE	0.00020	0.00045	0.00011	0.00002	0.00013	0.00012
allele frequency	mean	0.00027	0.00394	0.00091	0.00299	0.00249	0.00260
	SE	0.00000	0.00004	0.00003	0.00004	0.00004	0.00004

Table 5. The values of the load statistics obtained from Table 4 with $\bar{\gamma} = 1000$ when scaled up to the whole *D. melanogaster* genome.

	***h* = 0**		***h* = 0.2**		***h* = 0.5**	
Variable	**Gene model**	**Sites model**	**Gene model**	**Sites model**	**Gene model**	**Sites model**
L_T	0.0752	0.112	0.109	0.132	0.138	0.139
B_T	0.279	3.54	0.0503	0.165	0	0
V_T	4.76×10^{-5}	8.34×10^{-5}	2.93×10^{-5}	1.08×10^{-5}	3.59×10^{-5}	3.60×10^{-5}
σ_T	0.00690	0.00913	0.00541	0.00329	0.00628	0.00603

Figure 1: The variance in fitness (Var) and mean deleterious allele frequency ($\bar{q}$) for the gene and sites models with the *Drosophila*-like simulation parameters, and varying values of the epistasis coefficient (ϵ). Scaled selection coefficients followed a gamma distribution with $\bar{\gamma} = 100$ and shape parameter 0.3. 1000 selected sites were simulated with Nu = 0.005 and Nr = 0.01, where N is the population size (N = 1000), u is the mutation rate per site and r is the rate of crossing over between adjacent sites. The means and standard errors (SEs) for 1000 replicate simulations are shown.

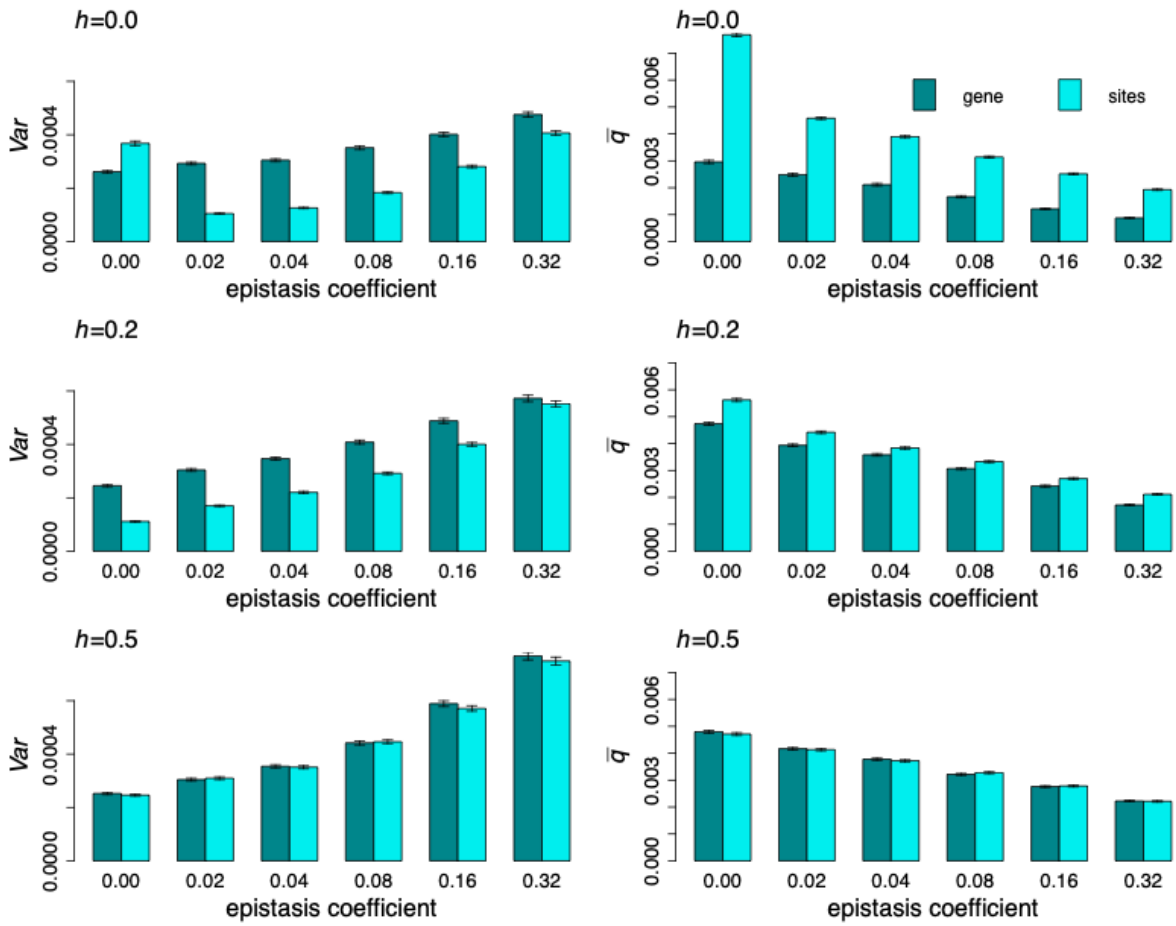
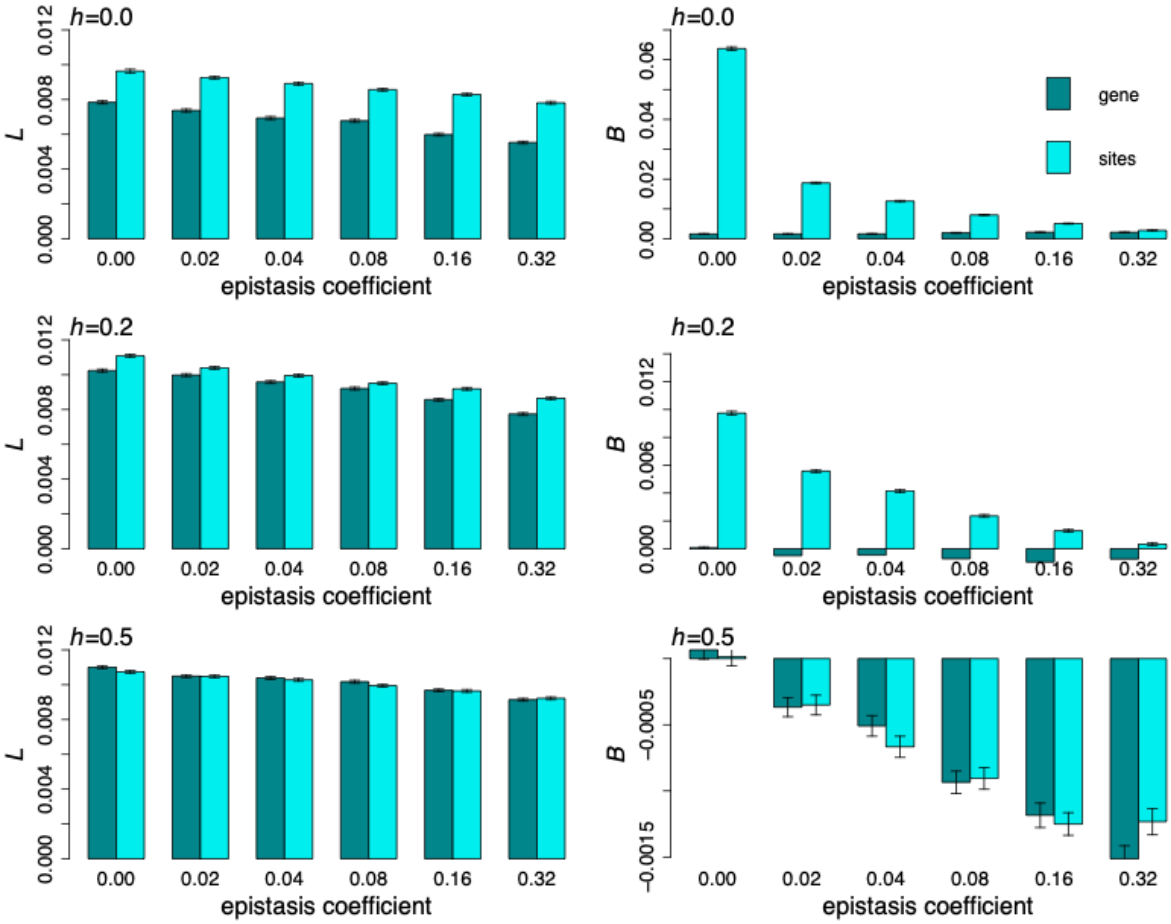
Figure 2: The means across replicate simulations of the genetic load (L) and inbreeding load (B) for the gene and sites fitness models, with varying values of the epistasis coefficient (ϵ). The selection coefficients follow a gamma distribution with $\bar{\gamma} = 100$ and shape parameter 0.3. 1000 selected sites were simulated, with $Nu = 0.005$ and $Nr = 0.01$, where N is the population size ($N = 1000$), u is the mutation rate per site/generation, and r is the rate of crossing over between adjacent sites. The error bars represent the SEs across 1000 replicate simulations.