Limitations of the inference of the distribution of fitness effects of new mutations in partially selfing populations with linkage

Austin Daigle[1,2,3] and Parul Johri[1,2,4,*]

[1]Department of Biology, University of North Carolina, Chapel Hill, NC 27599

[2]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599

[3]Curriculum in Bioinformatics and Computational Biology, University of North Carolina, Chapel Hill, NC 27599

[4]Integrative Program for Biological & Genome Sciences, University of North Carolina, Chapel Hill, NC 27599

*Correspondence can be addressed to these authors

Parul Johri

pjohri@unc.edu

ABSTRACT

The accurate estimation of the distribution of fitness effects (DFE) of new mutations is critical for population genetic inference but remains a challenging task. While various methods have been developed for DFE inference using the site frequency spectrum of putatively neutral and selected sites, their applicability in species with diverse life history traits and complex demographic scenarios is not well understood. Selfing is common among eukaryotic species and can lead to decreased effective recombination rates in such populations, increasing the effects of selection at linked sites, including interference between selected alleles. We employ forward simulations to investigate the limitations of current DFE estimation approaches in the presence of selfing and linked effects of selection. We find that distortions of the site frequency spectrum due to Hill-Robertson interference in highly selfing populations lead to mis-inference of the deleterious DFE of new mutations. While accounting for the decrease in the effective population size due to linked effects of selection largely accounts for the observed bias in populations with moderate levels of selfing, this correction is unable to accurately estimate the DFE in highly selfing populations. In addition, the presence of cryptic population structure and uneven sampling across subpopulations leads to the false inference of a deleterious DFE skewed towards effectively neutral/mildly deleterious mutations. Finally, the proportion of adaptive substitutions estimated at high rates of selfing is substantially overestimated. Our results clarify the parameter space where current DFE methods might be problematic and where they remain robust in the presence of selfing and other model violations, such as departures from semidominance, population structure, and uneven sampling. Our observations apply broadly to species and genomic regions with little/no recombination.

Keywords: distribution of fitness effects, self-fertilization, Hill-Robertson interference, background selection, population structure

Running title: Inferring fitness effects in selfing populations

INTRODUCTION

The distribution of fitness effects (DFE) of new mutations is an important and fundamental parameter in population genetics. The DFE describes the continuum of selective effects of spontaneous mutations in a population and is critical for understanding the effects of natural selection on genetic variation (Eyre-Walker and Keightley 2007). Because of the considerable impact deleterious mutations can have on allele frequency patterns in a population, knowledge of the DFE is often necessary to produce accurate evolutionary models, which are required for unbiased estimates of both the demographic history of a population and the proportion of adaptive nucleotide substitutions in a lineage (Charlesworth 1994; Eyre-Walker and Keightley 2007; Bank *et al.* 2014; Johri *et al.* 2022a). While the DFE can be estimated with experimental approaches using random mutagenesis or mutation-accumulation lines (*e.g.*, Wloch *et al.* 2001; Estes *et al.* 2004; Sanjuán *et al.* 2004), it can also be inferred from the distribution of allele frequencies from population genetic data (Williamson *et al.* 2005; Eyre-Walker *et al.* 2006; Keightley and Eyre-Walker 2007; Galtier 2016). As many organisms cannot easily be cultured in laboratories, are not genetically manipulatable, and/or have large generation times, population genetic approaches remain an important and the only possible approach to estimating the DFE in some species.

Population genetic methods use putatively neutral sites to fit the demographic history or to account for other potential nuisance parameters, while selected sites are used to infer the DFE (by inferring the parameters of a prespecified probability distribution), conditional on the demographic history obtained from neutral sites (reviewed in Johri *et al.* 2022b). While most two-step population genetic approaches that infer the DFE make a series of assumptions, like random mating and the absence of effects of selection at linked sites, they appear to be robust to certain model violations like mutation rate variability (Tataru *et al.* 2017), linkage (Kim *et al.* 2017; Huang *et al.* 2021), population substructure (Kim *et al.* 2017), and some complex demographic changes (Kousathanas and Keightley 2013). This is because two-step inference methods are usually employed using interdigitated neutral and selected sites and might therefore correct for model violations that skew the neutral and selected SFS similarly (for instance, linked effects of selection). As an aside, because demographic history is unlikely to affect the SFS of neutral and very strongly deleterious mutations equally (Eyre-Walker *et al.* 2006), this might lead to slight inaccuracies in inference using methods that do not explicitly model population

size changes. In fact, realistic simulations of chromosomal segments in both *Drosophila* (Huang *et al.* 2021) and humans (Huber *et al.* 2017) suggest that inference of the DFE of new deleterious mutations remains accurate, despite the effects of background selection. Similarly, Kim *et al.* (2017) were able to accurately infer the parameters of the deleterious DFE in a simulated metapopulation that grew eight times larger and split into eight isolated populations. However, since migration was not simulated between these populations, the effects of population structure on DFE inference have not been fully explored.

Despite significant progress, not much work has been done on understanding the effects of non-random mating, in particular self-fertilization, on the inference of the DFE using population genetic approaches. Self-fertilization or selfing (*i.e.*, an individual mating with itself) is common across all eukaryotic species - about 35% of all seed plants and ~50% of all animals exhibit moderate to high rates (50-100%) of selfing (Goodwillie *et al.* 2005; Jarne and Auld 2006). It is thus important to thoroughly investigate if current methods can be used to obtain accurate parameters of the DFE in partially selfing organisms. A transition from random mating to partial selfing has several effects on the population. Firstly, the increase of nonrandom mating in a partially selfing population decreases the effective population size (N_e), due to a reduction in the amount of independently sampled gametes (Pollak 1987) . Note that while the amount of drift experienced by a mutation is larger with a higher rate of selfing, the mean change in allele frequency at a selected locus is also greater, due to the increased frequency of homozygotes (Charlesworth and Charlesworth 1987; Caballero and Hill 1992; Charlesworth 1992). There is no change in the probability of fixation of semidominant mutations, but the probability of fixation of recessive or partially recessive beneficial mutations is increased by selfing, and the probability of fixation of recessive or partially recessive deleterious mutations is slightly reduced (Caballero and Hill 1992; Charlesworth 1992). Thirdly, increased homozygosity in the genomes of partially selfing organisms results in a higher proportion of recombination events occurring on nearly identical chromosomes, which decreases the effective rate of recombination (Nordborg 2000). This effect increases the extent of background selection (BGS; Charlesworth *et al.* 1993; Charlesworth and Charlesworth 1998; Charlesworth 2003) and selective sweeps (Hedrick 1980; Hartfield and Bataillon 2020) experienced by the population. Finally, as the effective rate of recombination is lowered drastically in highly selfing populations, there can be a further decrease

105 in the efficacy of selection due to Hill-Robertson interference (HRI) between selected loci (Hill
106 and Robertson 1966; McVean and Charlesworth 2000; Comeron and Kreitman 2002).
107 Hill-Robertson interference (HRI) describes how the efficacy of selection at an allele at
108 one locus can be reduced due to the segregation of selected alleles at a second linked locus (Hill
109 and Robertson 1966; Felsenstein 1974). While BGS and selective sweeps are also a type of HRI,
110 those models assume that selection is strong and that there is no interference between selected
111 sites (reviewed in Charlesworth and Jensen 2021). However, because weakly selected mutations
112 can segregate in the population for longer times, interference between them can become
113 important especially if they are sufficiently prevalent (McVean and Charlesworth 2000;
114 Comeron and Kreitman 2002). Such interference between weakly selected alleles can result in
115 reduced efficacy of selection and has been termed as interference selection (Comeron *et al.* 1999;
116 Comeron and Kreitman 2002) or weak selection HRI (McVean and Charlesworth 2000; Kaiser
117 and Charlesworth 2009). Moreover, the effects of such interference have been shown to extend
118 to more strongly selected mutations in cases where recombination is greatly reduced (Kaiser and
119 Charlesworth 2009). We therefore refer to the effects of interference selection/weak selection
120 HRI as simply HRI in this study. Since the entire genomes of populations with high selfing rates
121 have greatly reduced effective recombination rates (Nordborg 2000) and deleterious mutations
122 are prevalent across the genome (Bank *et al.* 2014), HRI is likely to play an important role in
123 determining the dynamics of deleterious mutations in such populations.
124 Unlike the model of classical BGS, where deleterious alleles are at deterministic
125 mutation-selection equilibrium, selected sites in populations or genomic regions with pervasive
126 HRI depart from equilibrium (Comeron and Kreitman 2002; Kaiser and Charlesworth 2009;
127 Seger *et al.* 2010; O'Fallon *et al.* 2010; Nicolaisen and Desai 2013). HRI can cause an increase
128 in diversity at selected sites relative to neutral sites (as the efficacy of purifying selection is
129 reduced). Further, the site frequency spectra (SFS) of neutral and selected sites affected by HRI
130 are skewed towards rare variants and become more similar to each other as the strength of HRI
131 increases, and these changes cannot be summarized as a simple change in effective population
132 size (Comeron and Kreitman 2002; Kaiser and Charlesworth 2009; Seger *et al.* 2010; O'Fallon *et*
133 *al.* 2010; Nicolaisen and Desai 2013; Good *et al.* 2014). More recent work assuming a single
134 fixed selection coefficient has shown that when HRI is sufficiently common, the nucleotide site
135 diversity can be entirely described by the variation of fitness between individuals in the

population. However, populations with different combinations of average selection coefficients, recombination, and mutation rates can have the same nucleotide diversity in this regime (Good *et al.* 2014), suggesting difficulties in estimating a distribution of selection coefficients along with other population genetic parameters in populations with little/no recombination.

A recent study suggested that the DFE of new deleterious mutations is mis-inferred when the probability of selfing is sufficiently high in simulated populations resembling *Caenorhabditis elegans* (Gilbert *et al.* 2022). However, it remains unclear which of the evolutionary processes discussed above were responsible for their observations and whether their results would be applicable to a wider range of evolutionary scenarios. Here, we thoroughly characterize how HRI and other linked effects of selection might bias DFE inference in highly selfing species and in regions of low recombination for a variety of DFE shapes. To investigate other potential causes of DFE mis-inference in partially selfing populations, we explore the effects of changing the dominance coefficients, coding density, and number of selected sites, as well as the effects of cryptic population structure and sampling on DFE inference. Moreover, we characterize the accuracy of inference of both the DFE of new deleterious mutations as well as the proportion of adaptive substitutions. We use the population parameters of *C. elegans* nematodes because they have a high rate of selfing, which has been estimated to be 98%-99.9% using laboratory observations (Hodgkin 1983; Anderson *et al.* 2010), field work, and population genetic analyses (Barrière and Félix 2005, 2007). This confirmation of high selfing, combined with the detailed knowledge of the mutation rates (Denver *et al.* 2004; Konrad *et al.* 2019) and recombination rates (Rockman and Kruglyak 2009), makes *C. elegans* an ideal system to understand the effect of selfing on DFE inference.

METHODS

Simulations

All simulations were performed using the forward simulator SLiM (version 4.0.1; (Haller and Messer 2023). The genome architecture and population parameters were modeled to mimic genomes of *C. elegans* populations. A chromosome with an intron-exon structure was simulated such that each gene consisted of 6 exons (250 bp in length) and 5 introns (100 bp in length). Genes were interspersed with intergenic regions of length 3000 bp each. These numbers were

167 close to the median of the length of exons, introns, and intergenic regions observed in the *C.*
168 *elegans* genome with a gene density of ~1 gene every 5kb (*C. elegans* Sequencing Consortium
169 1998). The *C. elegans* genome comprises 30% coding, 25% intronic, and 45% intergenic regions
170 (*C. elegans* Sequencing Consortium 1998). In line with these genome-wide characteristics, our
171 simulated genome comprised of 30% coding, 10% intronic, and 60% intergenic regions, and a
172 total of 500 genes, resulting in a full chromosomal length of 2,503,000 bp. All introns and
173 intergenic regions were assumed to be neutral, while exons experienced selection.

174 To take synonymous sites into account, 25% ($= f_{neu}$) of all mutations in exons were
175 assumed to be fully neutral (with s=0) while the rest were under selection. Of the selected
176 mutations, a fraction f_{pos} were assumed to be beneficial, where f_{pos} was assumed to be either
177 0.0, 0.01, or 0.001, and the rest ($1 - f_{pos} - f_{neu}$) were deleterious. The DFE of deleterious
178 mutations was assumed to be a gamma distribution, parametrized by the shape parameter (β) and
179 the mean population-scaled selection coefficient ($\gamma = 2N_e\overline{s_d}$), where $\overline{s_d}$ is the mean selective
180 disadvantage of the mutant homozygote relative to wildtype. Three different deleterious DFEs
181 were modeled, representing fitness effects skewed towards mildly, moderately, and strongly
182 deleterious classes respectively (Table 1). The DFE of beneficial mutations was assumed to be
183 exponential with mean $2N_e\overline{s_a} = 200$, where $\overline{s_a}$ is the mean selective advantage of the mutant
184 homozygote relative to wildtype. All mutations were assumed to be semidominant unless
185 otherwise specified.

186 A Wright-Fisher population under equilibrium was simulated with a population size (N_e)
187 of 500,000 diploid individuals (following Gilbert *et al.* 2022) and with varying probabilities of
188 selfing – 0%, 50%, 80%, 90%, 95%, and 99%, where $N_e = N$ for 0% selfing. Mutation (μ) and
189 recombination (r) rates were assumed to be constant at 3.3×10^{-9} per site/generation (Saxena *et*
190 *al.* 2019) and 3.12×10^{-8} per site/generation (Rockman and Kruglyak 2009), respectively.
191 Simulations were conducted with a population size of 5000 diploid individuals, with a rescaling
192 factor of 100, and rates of mutation and recombination as well as selection coefficients were
193 correspondingly adjusted so that the parameters $N\mu$, Nr, and Ns were equivalent to a
194 population of 500,000. Simulations were run for a total of $14N$ generations, where $10N$
195 generations corresponded to the burn-in period and $4N$ generations were used to calculate the
196 number of substitutions. Five replicates of each evolutionary scenario were simulated and the
197 DFE was estimated for each replicate separately.

198

199 *Single-locus simulations with selfing*: To investigate the effects of selfing on DFE inference
200 without linkage, selected, semidominant mutations were simulated at single sites with
201 probabilities of selfing of 0%, 50%, and 99%. The s_d of deleterious mutations were drawn from
202 the gamma distributions shown in Table 1. As allele frequencies of neutral mutations are
203 unaffected by linkage in the absence of selected mutations, 187,500 neutral sites were simulated,
204 which is equivalent to the number of neutral sites simulated in the simulations of 500 *C. elegans*
205 genes. 50,000 selected sites were simulated for each scenario (compared to 562,500 selected sites
206 in previous simulations), then paired with neutral mutations simulated with the appropriate
207 selfing rate. The simulations of individual selected sites and the neutral sites used rescaled values
208 of N, μ, and r equal to 5,000 individuals, 3.3×10^{-7} per site/generation and 3.12×10^{-6} per
209 site/generation respectively, which is equivalent to the previous simulations with linkage. All
210 simulations were run for $14N$ generations with $10N$ burn-in generations.
211

212 *Simulations under low recombination rates with no selfing*: To clarify if the lower effective
213 recombination rates in a partially selfing population were responsible for DFE mis-inference, we
214 conducted simulations with low recombination rates and 0% selfing. As before, these simulations
215 were conducted assuming a census size of 500,000 diploid individuals and rescaled by a factor of
216 100 to 5,000 individuals. The *C. elegans* recombination rate of 3.12×10^{-8} was thus rescaled to
217 3.12×10^{-6}, and multiplied by 0.5, 0.1, 0.05, 0.01, 0.005, or 0.001 to determine a final
218 simulated recombination rate. Beneficial mutations were not included in these simulations.
219

220 *Simulations with varying dominance coefficients*: All selected mutations shared a single
221 dominance coefficient of 0.1, 0.25, or 0.75. Populations with selfing probabilities of 0%, 50%,
222 and 99% were simulated and compared to the original simulations (where $h = 0.5$).
223

224 *Simulations under varying coding density and chromosome size*: Simulations were performed
225 with varying coding densities by varying the intergenic lengths - 1500 bp and 500 bp. The total
226 number of genes was not increased in these simulations. Thus, in the simulations with 1500 bp
227 intergenic regions, the simulated genome was comprised of ~43% coding, ~14% intronic, and
228 ~43% intergenic regions, with a total length of 1,752,100 bp. In the simulations with 500 bp

229 intergenic regions, the genome was comprised of ~55% coding, ~18% intronic, and ~27%
230 intergenic regions, with a total length of 1,376,650 bp. These simulations were run with
231 probabilities of selfing equal to 0%, 50%, and 99%.
232 To study the effects of a more realistically sized chromosome on DFE inference in
233 selfers, simulations where the total number of genes was increased from 500 to 3000 were
234 conducted, corresponding to approximately the number of genes observed on the first
235 chromosome of the *C. elegans* genome. This increased the size of the simulated chromosome to
236 15,018,000 bp, which is close to the true size of *C. elegans* chromosome I (15,072,434 bp)
237 (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000002985.6/; last accessed in Dec.
238 2023).
239
240 **Quantifying the effect of BGS in simulated data**
241 To determine the theoretically expected neutral nucleotide diversity (π) in simulations with
242 partial selfing, we calculated the expected effective population size after correcting for selfing
243 (N_{self}) as $N_{self} = N/(1+F)$, where $F = P_{self}/(2-P_{self})$, and P_{self} represents the proportion
244 of selfing individuals in a population per generation (Pollak 1987). Without the linked effects of
245 selection, the neutral π is expected to be approximated by $\pi \sim \theta = 4N_{self}\mu$. In simulations with
246 deleterious mutations, the extent of BGS was quantified as the observed nucleotide diversity
247 (π_{obs}) at neutral sites relative to that expected under strict neutrality, *i.e.*,
248
249 $$B = \frac{\pi_{obs}}{4N_{self}\mu} \quad (1)$$
250
251 **Simulations with population structure**
252 To study the effects of population structure on DFE inference in selfers, a finite island model
253 (Maruyama 1970) with five demes and an equal probability of exchange of migrants between all
254 demes was considered. Note that in this model the migration rate is independent of the selfing
255 rate, which applies well to animals but is not necessarily the case in plant species because selfing
256 can prevent pollen dispersal (reviewed in Sicard and Lenhard 2011). Scenarios with different
257 migration rates were simulated such that the effective metapopulation size remained constant and
258 equivalent to the size of the panmictic population simulated above. Genome structure was

identical to that in the previous simulations mimicking *C. elegans* genome architecture. In a metapopulation with d demes and a deme effective population size N_{deme}, the mean coalescence time for a pair of alleles sampled from the same deme (T_S) is given by the following equation (Nei 1973; Slatkin and Voelm 1991; Charlesworth and Charlesworth 2010):

$$T_S \sim 2dN_{deme} \tag{2}$$

The mean coalescence time for a pair of alleles sampled from different demes (T_B) is given by (Nei 1973; Slatkin and Voelm 1991; Charlesworth and Charlesworth 2010):

$$T_B \sim 2dN_{deme} + \frac{(d-1)^2}{2m} \tag{3}$$

where the migration rate m is the probability that an allele in one deme is derived by migration from any other deme in a generation (which means that the migration rate between any two demes is equal to $\frac{m}{d}$). Combining these two equations gives the mean coalescence time (T_T) for a pair of alleles drawn from the entire metapopulation (Nei 1973; Slatkin and Voelm 1991; Charlesworth and Charlesworth 2010):

$$T_T \sim 2dN_{deme} + \frac{(d-1)}{2dm} \tag{4}$$

To keep the effective size of the metapopulation ($0.5T_T$) equivalent to the rescaled size of 5,000 individuals used in previous simulations, N_{deme} and m were varied such that T_T remained equal to 10,000. To allow a wide range of migration rates, $N_{deme}m$ was varied from 10, 2, 1, 0.5, and 0.1 in populations with 0%, 50%, and 99% selfing, with separate simulations for DFE1, DFE2, and DFE3. The exact parameters used in each simulation can be found in Supplementary Table 7. Parameters were chosen so that the predicted F_{ST} of the entire metapopulation ($F_{ST} =$

288 $\frac{T_T - T_S}{T_T}$) would cover a range of values (from 0.02 to 0.76) matching previously estimated values
289 for *C. elegans* populations in the literature (from 0.15 to 0.43; Cutter 2006). Neutral simulations
290 with identical genome structure and migration parameters were used to verify the joint effects of
291 population structure and selfing on neutral nucleotide diversity and F_{ST}, and were found to yield
292 expected results (Supplementary Table 8). The expected neutral nucleotide diversity for partially
293 selfing populations was calculated assuming that the effective size of each deme will be
294 equivalent to $N_{deme,self} = \frac{N_{deme}}{1+F}$, which was used to calculate the appropriate T_T and T_S values
295 for each evolutionary scenario. In SLiM, m represents the proportion of individuals in a deme
296 that are migrants each generation. Therefore, m still represents the probability that an allele is
297 derived from migration each generation, and $N_{deme}m$ represents the number of diploid migrants
298 per deme per generation.
299 Twenty genomes were sampled from each deme (totaling 100 genomes from the
300 metapopulation) for DFE inference. To understand the effects of uneven sampling on DFE
301 inference, two alternative sampling schemes were employed. In the first, 60 genomes were
302 sampled from one deme, while 10 genomes were drawn from the remaining four demes. In the
303 second, 35 genomes were drawn from two demes, while 10 genomes were drawn from the
304 remaining three demes.
305
306 **Calculating the SFS and the number of adaptive substitutions from simulations**
307 The full output provided by SLiM was used to generate the site frequency spectra of neutral and
308 selected mutations in exonic regions individually. The fixed class was obtained as the number of
309 substitutions that occurred after the burn-in period ($= 10N$ generations). The zero class was
310 calculated by subtracting the number of polymorphic and fixed sites from the total number of
311 sites for both neutral ($= 0.25$ $\times$ number of exonic sites) and selected ($= 0.75$ $\times$ number of
312 exonic sites) mutations respectively. Scripts to calculate the SFS are available at
313 https://github.com/JohriLab/DFESelfing/tree/main/calculateSFS.
314
315 The true fraction of beneficial substitutions (α) in simulations was obtained as follows:
316
317 $$\alpha = \frac{number\ of\ substitutions\ with\ 2N_e s_a > 5}{number\ of\ substitutions\ at\ selected\ sites} \quad (5)$$

Only substitutions that occurred after the burn-in with $2N_e s_a > 5$ were included when calculating the true value of α in order to match the assumptions of GRAPES.

DFE-alpha

Est_dfe (Keightley and Eyre-Walker 2007) was employed to infer the DFE of deleterious mutations. We fitted a two-epoch model, with variable and starting values of $t2$ (time of change), β, and s_d to be 50, 0.5, and -0.01. $N2$ (current size) was also set to be variable. In the absence of positive selection, the unfolded SFS was used as input, and the fold setting was set to 0. In the presence of beneficial mutations, the folded SFS was used instead. DFE-alpha provides the most likely estimate of the expected value of $\overline{s_d}$, the shape parameter (β) of the gamma distribution, and a weighted effective population size (N_w). These parameters were used to obtain the proportion of mutations (f_0, f_1, f_2, f_3) in the four classes- effectively neutral ($0 \leq 2N_w s_d < 1$), weakly deleterious ($1 \leq 2N_w s_d < 10$), moderately deleterious ($10 \leq 2N_w s_d < 100$), and strongly deleterious ($100 \leq 2N_w s_d$), respectively.

GRAPES

The program GRAPES (Eyre-Walker and Keightley 2009; Galtier 2016), Version 1.1.1 (https://github.com/BioPP/grapes/releases/tag/v1.1.1) was used to infer the full DFE using simulations with and without positive selection. GRAPES provides the maximum likelihood estimate of the shape parameter (β) and the expected population-scaled mean strength of selection in terms of $2N_e\overline{s_d}$, in order to characterize the gamma distribution representing the deleterious DFE. GRAPES also infers the parameters of the beneficial DFE, including the proportion of substitutions that are beneficial mutations (α) and the mean population-scaled strength of selection favoring new advantageous mutations $2N_e\overline{s_a}$, assuming that advantageous selection coefficients are exponentially distributed. The unfolded SFS was provided as input for both scenarios (with and without positive selection). In simulations without positive selection, the default parameters for the GammaZero model were used (*i.e.*, parameters of the beneficial DFE were not inferred). In the presence of positive selection, the estimated beneficial DFE was assumed to follow an exponential distribution (*i.e.*, the GammaExpo model was used), consistent

with our simulations. As GRAPES can use either the folded or the unfolded SFS to provide estimates of the beneficial DFE, we performed inference using the unfolded and folded SFS separately to compare the results of both estimation methods. These parameters were accordingly used to obtain the values f_0, f_1, f_2, f_3, representing the classes of the deleterious DFE as described above.

Comparison of the expected and estimated DFE when mutations are semidominant

In a partially selfing population, diffusion theory provides the following mean ($M_{\Delta q}$) and variance ($V_{\Delta q}$) of the change in allele frequency (q) assuming that the change (Δq) in allele frequency is small (Caballero *et al.* 1991):

$$M_{\Delta q} = q(1-q)s[h + q(1-2h) + F(1-q-h+2qh)] \tag{6}$$

$$V_{\Delta q} = \frac{q(1-q)}{2N_{self}} = \frac{q(1-q)(1+F)}{2N} \tag{7}$$

For completely additive mutations ($h = 0.5$), the mean reduces to:

$$M_{\Delta q} = \frac{q(1-q)s(1+F)}{2} \tag{8}$$

In a panmictic population (of size N) with random mating, the mean and variance of change in allele frequency (assuming semidominance) is given by:

$$M_{\Delta q} = \frac{q(1-q)s}{2} \tag{9}$$

$$V_{\Delta q} = \frac{q(1-q)}{2N} \quad (10)$$

Thus, by comparing the above set of equations, a selfing population of size N experiences more genetic drift than a randomly mating population of the same size (Caballero and Hill 1992). However, the mean change in allele frequency in selfing populations is also larger, by the same factor, $(1+F)$, and thus the fixation probability of mutations remains the same in a selfing *vs*. a randomly mating population (Charlesworth and Charlesworth 1987; Caballero and Hill 1992; Charlesworth 1992). Thus, for the case of additive mutations and not taking selection at linked sites into account, the expected DFE would be determined by $2N_{self}s_d(1+F) = 2\frac{N}{(1+F)}s_d(1+F) = 2Ns_d$. To account for the relative decrease in population size due to BGS or selective sweeps, the simulated DFE was also adjusted and presented as $2NBs_d$, where B was calculated using equation 1.

In the case of the estimated DFE, because DFE-alpha only infers a relative increase or decrease in population size ($\frac{N_2}{N_1}$), these estimates of demographic parameters should not be affected by selfing rates (where the factor $1+F$ will cancel out). We thus present our inferred DFE as $2N_w s_d$.

Comparison of the expected and estimated DFE when mutations are not semidominant

When mutations were not additive, our expected DFE was parameterized by $2N_{self}s_d h_{self}$, where

$$N_{self} = \frac{N}{1+F} \quad (11)$$

and

$$h_{self} = h + q(1-2h) + F(1-q-h+2qh) \quad (12)$$

Assuming that $q << 1$, this equation simplifies to:

$$h_{self} \sim h + F(1-h) \tag{13}$$

Note that this approximation is less accurate when selection is weak ($2Ns_d \sim 1$) and thus might not be as accurate for the DFE comprising predominantly weakly deleterious mutations. The estimated DFE (using DFE-alpha and GRAPES) was parameterized by $2N_w s_d h$ where $h = 0.5$ is assumed by both the programs.

Calculation of summary statistics from simulated data

In order to quantify the effects of HRI, the skew in the SFS towards rare alleles was measured using the summary statistic $\Delta\theta_w = 1 - (\frac{\pi}{\theta_w})$, which has been demonstrated to be less sensitive than Tajima's *D* to the number of SNPs used in the calculation (Becher *et al.* 2020). Here, π is the nucleotide diversity (Nei and Tajima 1981) and θ_w denotes Watterson's θ (Watterson 1975). Both summary statistics were calculated in nonoverlapping 5000 bp windows using *pylibseq* v. 0.2.3 (Thornton 2003). To summarize linkage disequilibrium, D' was calculated for pairs of polymorphisms at frequencies between 1% and 5% in each window. Only low-frequency SNPs were used for D' calculations to correct for the allele frequency dependence of linkage disequilibrium statistics (Lewontin 1988). Here D and D' Lewontin (1964) were computed by *pylibseq* as follows:

$$D = p_{AB} - p_A p_B \tag{14}$$

and

$$D' = \frac{D}{D_{max}} \tag{15}$$

where

if $D > 0$

$$D_{max} = \min(p_A p_b, p_a p_B) \tag{16}$$

if $D < 0$

$$D_{max} = \min(p_A p_B, p_a p_b) \tag{17}$$

436

437 where p_A and p_B represent the minor allele frequency for each polymorphism, p_a and p_b
438 represent the major allele frequency, and p_{AB} represents the frequency of the haplotype AB. All
439 summary statistics were calculated separately for neutral and selected polymorphisms in exonic
440 regions.

441
442

443 **RESULTS**

444 **Inference of the DFE of new deleterious mutations in self-fertilizing populations**

445 We simulated a single *C. elegans*-like population with varying rates of selfing – 0, 50, 80, 90, 95,
446 and 99% and three different DFEs (following a gamma distribution; parameters specified in
447 Table 1) skewed towards mildly (DFE1), moderately (DFE2), and strongly deleterious (DFE3)
448 classes respectively. Here we assumed that all new mutations were semidominant. We used
449 DFE-alpha and GRAPES to infer the shape of the deleterious DFE, specified to be gamma-
450 distributed. As expected, when there is no selfing, inference was accurate (Figure1,
451 Supplementary Figure 1, Supplementary Table 1, Supplementary Table 2). As the selfing rate
452 increased ($\geq 90\%$), the estimated DFE was skewed strongly towards the class of mildly
453 deleterious mutations (Figure 1), when compared to the simulated DFE, with a particularly
454 strong skew in the most highly selfing population, consistent with Gilbert *et al.* (2022). This mis-
455 inference corresponded with large decreases in the mean values of population-scaled selection
456 coefficients of the inferred gamma distributions (Supplementary Figure 1). The β parameter of
457 the gamma distribution was also mis-inferred as the selfing rate increased, with the inferred β
458 becoming 2-3× smaller than the simulated value at 99% selfing (Supplementary Table 2). While
459 the overall shape and skew of the DFE were largely recovered when there was some chance of
460 outcrossing (*e.g.*, with $P_{self}\sim 90\%$; Figure 1), the mean population-scaled selection coefficients
461 remained slightly biased even when selfing rates were as low as 80% (*e.g.* for DFE2;
462 Supplementary Figure 1, Supplementary Table 2). Because BGS will lead to a decrease in the
463 effective population size, we hypothesized that the true population-scaled selection coefficients
464 in our simulated populations are likely to be smaller than those simulated (*i.e.*, $2Ns_d$). To
465 account for the reduction in the efficacy of purifying selection due to BGS, we corrected the
466 simulated DFE with the relative rate of coalescence (B) with and without BGS observed in the

467 simulated population. Note that this adjusted DFE accounts for the relative decrease in
468 population size due to linked effects of selection only, not due to selfing (see details in Methods).
469 While the inferred DFE matches the corrected expected DFE substantially well for moderate
470 levels of selfing (Figure 1), the DFE accounting for BGS slightly overcorrects at high selfing
471 levels. That is, the inferred DFE has more strongly deleterious mutations than the DFE
472 accounting for linked effects of selection. This suggests that a simple rescaling of the effective
473 population size can correct for the misinference when selfing rates are low/moderate (when the
474 population is likely experiencing classical BGS) but not when rates are high. We, therefore,
475 conclude that Hill-Robertson interference (HRI) among deleterious alleles is likely causing the
476 mis-inference of the DFE of new mutations in highly selfing populations.
477
478 **The effects of selfing on DFE inference**
479 To confirm that selfing alone does not cause DFE mis-inference, we tested the performance of
480 DFE estimation methods on simulations with no linkage. To accomplish this, 50,000 selected
481 sites were simulated individually with separate replicates for each of the three DFEs (Table 1) at
482 three selfing levels (0, 50, and 99%). As expected, the DFEs estimated by DFE-alpha and
483 GRAPES were accurately recovered even in highly selfing populations (Supplementary Figure
484 2), despite a significant reduction in the total number of segregating sites at high selfing
485 (Supplementary Table 3). As expected, the average nucleotide diversity (π) of neutral sites
486 calculated from simulations was comparable to the theoretically expected neutral π for
487 populations with varying selfing rates (Supplementary Table 4). Concordantly, the observed
488 SFSs at selected and neutral sites were highly similar across different selfing levels, confirming
489 that selfing does not drastically influence the proportions of allele frequencies in a population
490 without linkage (Supplementary Figure 3), assuming that deleterious mutations are additive.
491 Importantly, as high rates of selfing increase the levels of homozygosity, the site
492 frequency spectrum (SFS) obtained from a random sample of diploid *individuals* deviated starkly
493 from one obtained by a random sampling of *genomes* (an example shown in Supplementary
494 Figure 4A). The inferred DFEs were very similar when we sampled individuals *vs*. genomes
495 (Supplementary Figure 4B) and thus all results that follow were obtained by sampling genomes.
496
497

498 **The effects of linkage on DFE inference in regions of low recombination**

499 Because selfing by itself is unlikely to be responsible for the observed DFE mis-inference, we
500 further investigated the role of linkage. We hypothesized that we would see a similar bias in DFE
501 inference in simulations with low recombination rates even when there is no selfing. With
502 decreasing rates of recombination (0.5, 0.1, 0.05, 0.01, 0.005, and 0.001×), the extent of mis-
503 inference of the DFE increased (Figure 2). As before, the estimated DFE was strongly skewed
504 towards mildly deleterious mutations compared to the simulated DFE, consistent with a decrease
505 in the efficacy of selection due to linked effects of selection.

506 The relative decrease in nucleotide diversity ($B = \frac{\pi}{\pi_0}$) due to background selection,
507 measured in simulations of regions with low recombination ($B \sim 0.05 - 0.92$) captures the range
508 of B values measured in the simulations with selfing and linkage ($B \sim 0.08 - 0.95$). To account
509 for the decrease in the effective population size due to BGS, the inferred DFE was compared to
510 the DFE given by the distribution of $2NBs_d$. Again, while the DFE accounting for BGS
511 explained the estimated DFEs quite well for higher recombination regions, at very low
512 recombination rates, it did not match the inferred DFEs. As with high selfing, the estimated DFE
513 at low recombination was observed to be more strongly deleterious than the DFE accounting for
514 BGS. This reflects the fact that a simple rescaling effect of BGS (*i.e.*, the decrease in the rate of
515 coalescence) is unlikely to explain mis-inference observed in regions of very low recombination.
516 When recombination rates are extremely low, the probability of interference between selected
517 alleles at multiple loci increases, leading to a departure of the SFS from equilibrium. Overall,
518 increased HRI caused by severely reduced recombination rates, which are of particular
519 importance in selfing organisms, might lead to the observed DFE mis-inference.

520

521 **The quantification of HRI using summary statistics**

522 Much theoretical and computational work on HRI has suggested that the decrease in the efficacy
523 of selection results in a decrease in the nucleotide diversity at selected and linked neutral sites as
524 well as an increase in divergence at selected sites due to the increased fixation of deleterious
525 mutations (Comeron *et al.* 1999; McVean and Charlesworth 2000). Moreover, the site frequency
526 spectrum for selected and neutral sites is expected to be skewed towards low-frequency variants
527 as the strength of HRI increases, and the nucleotide diversity and site frequency spectrum of
528 selected mutations are expected to become more similar to those of neutral mutations as HRI

529 increases (Comeron and Kreitman 2002; Kaiser and Charlesworth 2009; Seger *et al.* 2010;
530 O'Fallon *et al.* 2010; Nicolaisen and Desai 2013). Because interference results in a gene
531 genealogy with longer external branches than expected under a standard neutral coalescent
532 (mimicking the genealogy of an expanding population; Walczak *et al.* 2012; Nicolaisen and
533 Desai 2013), and consistent with a skew towards low-frequency alleles, new mutations are more
534 likely to occur on separate branches of the genealogy, generating negative LD. Thus signed
535 measures of linkage disequilibrium between selected mutations and nearby neutral sequences
536 become more negative as HRI becomes stronger (Hill and Robertson 1966; McVean and
537 Charlesworth 2000; Comeron and Kreitman 2002; Garcia and Lohmueller 2021).
538 In our simulations, the SFSs of neutral and selected sites shifted towards lower frequency
539 mutations and became more similar at 90% selfing and above (Figure 3), which is consistent
540 with the extent of DFE mis-inference and with previous simulated results in highly selfing
541 populations (Charlesworth *et al.* 1993). Note that the SFS of neutral sites at 0% selfing was very
542 similar to the unlinked neutral SFS (Supplementary Figure 3), consistent with high B values.
543 Trends in $\Delta\theta_w$, a measure of the skewness of the SFS (with positive values indicating a relative
544 increase in rare alleles), were concordant with the patterns in the SFSs, with $\Delta\theta_w$ values for
545 neutral and selected sites sharply increasing and converging towards each other at 90% selfing
546 and above (Figure 4A). Values of $\Delta\theta_w$ at neutral sites nearly converge with values at selected
547 sites at 99% selfing, reflecting the reduced efficacy of selection and almost complete linkage.
548 For selfing rates between 80% and 95%, D' tended to increase at both selected and
549 neutral sites (Figure 4B). This is similar to an effect observed in simulations by Kamran-Disfani
550 and Agarwal (2014), where partially selfing populations with intermediate selfing rates showed
551 increases in linkage disequilibria. This phenomenon is expected at higher selfing levels if the
552 efficacy of selection is not overwhelmed by HRI, because of correlations in homozygosity at
553 different loci across the genome (Roze and Lenormand 2005). At 99% selfing, D' values dropped
554 considerably for selected and neutral sites, consistent with a population experiencing pervasive
555 HRI.
556
557 **Extent of mis-inference depends on chromosome length but not coding density**
558 The effects of HRI are expected to increase when there is a higher density of selected sites, as
559 deleterious mutations arising in the population will be more closely linked on average (Comeron

560 and Kreitman 2002). Gene density varies along the *C. elegans* chromosomes, with higher gene
561 densities found in the chromosome centers than in the chromosome arms (*C. elegans* Sequencing
562 Consortium 1998). To understand if increases in coding density change the extent of DFE mis-
563 inference, we decreased the length of intergenic regions in our simulations, while keeping the
564 number and length of genes constant, producing simulations with coding densities 2× and 6×
565 higher than our original coding density. We observed no obvious changes in DFE inference in
566 these scenarios (Supplementary Figure 5), suggesting that DFE mis-inference will not increase in
567 more coding-dense regions of the genome, although local gene densities could increase more
568 than we simulated. In addition, the rate of recombination in *C. elegans* is higher in the gene-poor
569 chromosome arms than in the gene-rich cores (Rockman and Kruglyak 2009), which could
570 worsen DFE mis-inference.
571 We hypothesized that a higher total number of selected sites would increase the effects of
572 BGS and possibly interference in selfing populations due to long-distance interactions between
573 loci when rates of recombination are very low (Comeron *et al.* 1999; McVean and Charlesworth
574 2000; Comeron and Kreitman 2002; Kaiser and Charlesworth 2009). We increased the total
575 length of our simulated chromosomes such that the final simulated chromosome was close to the
576 length of *C. elegans* chromosome I. Strikingly, we observed a large increase in DFE mis-
577 inference in larger chromosomes in highly selfing populations (Figure 5), while there was no
578 change in DFE inference in populations with no selfing (Supplementary Figure 6). In populations
579 with 50% selfing, DFE inference remained accurate for the weakly and moderately deleterious
580 DFEs, but the strongly deleterious DFE was mis-inferred (Supplementary Figure 6). The DFE
581 accounting for BGS ($2NBs_d$) explained the results at lower selfing rates well but showed similar
582 patterns of inaccuracy at high selfing levels. This is expected because BGS is likely to act at
583 longer genetic distances than HRI. These results have implications for populations with
584 intermediate selfing rates, as the extent of DFE mis-inference seems to depend on the density of
585 selected sites, DFE, selfing rate, and the chromosomal map length. Moreover, these results
586 highlight the importance of simulating full-length chromosomes when studying the effects of
587 selection at linked sites, especially in organisms with a slow LD decay with distance between
588 sites, like *C. elegans* (Rockman and Kruglyak 2009).
589
590

591 **The impact of dominance on DFE inference**

592 DFE-alpha and GRAPES assume that all new mutations are semidominant, but it is known that
593 dominance coefficients of new deleterious mutations vary considerably and are partially
594 recessive on average, with strongly deleterious mutations tending to be highly recessive
595 (Simmons and Crow 1977; Agrawal and Whitlock 2011; Manna *et al.* 2011). The effect of
596 dominance on DFE inference is not fully understood. Most DFE inference methods assume that
597 new mutations are semidominant and thus provide estimates of the joint distribution of the
598 selection and dominance coefficients, "$s_d h$". A recent study by Wade *et al*. (2023) suggests that,
599 if there are recessive lethal mutations, current DFE inference approaches can lead to
600 underestimating the true proportion of strongly deleterious mutations, as more strongly
601 deleterious mutations are likely to persist in the population if they are recessive. Furthermore,
602 since selfing increases the levels of homozygosity in the genome, the efficacy of selection
603 against large effect and highly recessive deleterious mutations is expected to be greater
604 (Charlesworth and Charlesworth 1998), leading to more efficient purging of deleterious
605 mutations. Thus, with higher rates of selfing one might instead expect to observe a decrease in
606 the allele frequencies of mildly and moderately deleterious mutations, leading to the false
607 inference of the presence of more strongly deleterious mutations in the population.

608 To investigate the joint effects of dominance and selfing on DFE inference, we performed
609 simulations where all selected mutations had fixed dominance coefficients equal to 0.1, 0.25, or
610 0.75. Note that the outputs of DFE-alpha and GRAPES were adjusted to represent $2N_{self}s_d h_{self}$
611 rather than $2Ns_d$, and the expected values from the simulations were adjusted to account for the
612 dominance and homozygosity of selfing genomes (Caballero and Hill 1992), as described in
613 Methods. At 0% selfing, as new deleterious mutations became more recessive (especially when
614 $h = 0.1$), the distribution of the product of the selection and dominance coefficients was
615 somewhat mis-inferred. When mutations were predominantly moderately and strongly
616 deleterious, there was an underestimation of the proportion of strongly deleterious mutations
617 (Figure 6, Supplementary Table 5), consistent with Wade *et al.* (2023). Additionally, the
618 frequency of the nearly neutral class of mutations was underestimated in all three DFEs, an effect
619 that can also be observed to a smaller degree in Wade *et al.* (2023). However, when new
620 mutations were comprised largely of mildly deleterious mutations (*i.e.*, DFE1), there was an
621 overestimation of the fraction of strongly deleterious mutations. In contrast, when more

622 dominant mutations were simulated, DFE inference resulted in mis-inference towards both
623 weakly and strongly deleterious mutations, though this effect was fairly small. As expected, mis-
624 inference was stronger when the true DFE was skewed towards mildly deleterious mutations, as
625 these mutations are more likely to be present as homozygotes in the population.
626 As selfing increases homozygosity, the inferred DFEs in partially selfing populations at
627 different dominance levels are expected to become more similar to each other, making the
628 dominance of mutations a less important factor for the efficacy of selection (Caballero and Hill
629 1992; Charlesworth 1992). As expected, at 50% selfing, the DFE was inferred more accurately
630 (Supplementary Figure 7, Supplementary Table 5). While the dominance coefficient played a
631 negligible role in DFE inference when the selfing rate was 99%, the estimated DFEs were
632 strongly skewed towards mildly deleterious mutations (Figure 6). Thus, just as expected, at high
633 rates of selfing, dominance coefficients play a relatively minor role, while HRI and BGS are
634 primarily responsible for DFE mis-inference.
635 Increased purging of recessive deleterious mutations has been hypothesized to reduce
636 their average allele frequencies, leading to DFE mis-inference (Arunkumar *et al.* 2015; Gilbert *et*
637 *al.* 2022). However, observable effects of purging are likely to occur only when mutations are
638 highly recessive (Nei 1968) and strongly deleterious (Glémin 2003). Concordantly, we observed
639 that the distribution of allele frequencies of partially recessive ($h = 0.25$) mutations at
640 intermediate selfing levels was the same as those of semidominant mutations, suggesting that
641 increased purging of recessive deleterious alleles did not have a strong effect on the average
642 allele frequencies in a population sample at equilibrium (see the SFS in Supplementary Figure
643 8).
644
645 **Inference of the full DFE of new deleterious and beneficial mutations**
646 Both BGS and HRI lead to a decrease in the efficacy of selection at linked selected sites, thus
647 reducing the probability of fixation of beneficial mutations while increasing the probability of
648 fixation of deleterious mutations (Charlesworth 1992; Kamran-Disfani and Agrawal 2014). It is
649 therefore likely that the proportion of fixed mutations at selected sites that are beneficial (α) is
650 mis-estimated when effective recombination rates are low. We conducted simulations with 0%,
651 50%, 80%, 90%, 95%, and 99% selfing, where 0.1% and 1% of new mutations in exons were
652 beneficial, following an exponential distribution with $2N\bar{s_a} = 200$. When 0.1% of new exonic

653 mutations were beneficial, the inference of the deleterious DFE by GRAPES was fairly accurate
654 at 0% and 50% selfing (Figure 7A). However, the deleterious DFE was mis-inferred at a much
655 lower selfing rate than in scenarios without beneficial mutations and mis-inference was
656 particularly strong when the deleterious DFE was highly skewed towards weakly deleterious
657 mutations. When 1% of new mutations were beneficial, the deleterious DFE estimated by
658 GRAPES was strongly skewed towards nearly neutral mutations even at 0% selfing for the
659 weakly deleterious and moderately deleterious DFEs (Supplementary Figure 9). Due to recurrent
660 selective sweeps in this scenario as well as BGS, there is likely to be a substantial decrease in
661 effective population size experienced by selected alleles (and thus a reduced $2N_e s_d$). We thus
662 accounted for this decrease again by comparing the inferred DFE to the distribution of
663 $2N(\pi/\pi_0)s_d$, where π/π_0 represents the neutral nucleotide site diversity with and without
664 selection. Again, the DFE accounting for the decrease in the effective size did not explain the
665 inferred DFE at high selfing rates, including 95% selfing, suggesting that HRI effects might be
666 responsible for the mis-inference, even at moderate levels of selfing (Figure 7A). The selected
667 SFSs resembled the neutral SFSs in simulations with DFE mis-inference, reflecting the observed
668 results of DFE inference (Supplementary Figure 10). With a high rate of beneficial mutations
669 entering the population, hitchhiking can reduce the efficacy of selection on weakly deleterious
670 alleles (Campos and Charlesworth 2019), leading to the mis-inference of the DFE for the more
671 weakly deleterious DFEs. In addition, the failure of DFE inference programs to fully correct for
672 the presence of mildly beneficial alleles in the SFS could also have contributed to the observed
673 mis-inference (Tataru *et al.* 2017).
674 As the folded SFS might provide better estimates of the deleterious DFE when beneficial
675 mutations are segregating at high frequencies (Keightley and Eyre-Walker 2007; Eyre-Walker
676 and Keightley 2009), we also tested the inference using the folded SFS. Overall, the deleterious
677 DFE inferred by GRAPES was comparable to the unfolded results when the folded SFS was
678 used (Supplementary Figures 11-12). However, the DFEs estimated by DFE-alpha were biased
679 towards weakly deleterious mutations even with no selfing, with much larger uncertainty in the
680 estimated DFE at high selfing levels, especially when 1% of all new selected mutations were
681 beneficial. This indicates that folding the SFS is inadequate to correct for the mis-inference of
682 the deleterious DFE when the rate of new beneficial mutations is extremely high.

683 As expected, the true fraction of beneficial substitutions (α) decreased as the selfing rate
684 increased (Figure 7B, Supplementary Figure 9B). Changes in this parameter were driven by both
685 increases in the number of deleterious substitutions ($\sim 2-4\times$ higher at 99% selfing relative to
686 0% selfing) as well as the decreases in the number of beneficial substitutions ($\sim 3\times$ lower at
687 99% selfing relative to 0% selfing; Supplementary Table 6), consistent with previous
688 observations in Kamran-Disfani and Agrawal (2014). In simulations with both 0.1% beneficial
689 mutations and 1% beneficial mutations, α tended to be overestimated for the weakly deleterious
690 DFE at all selfing levels, with mis-inference increasing steadily as the selfing rate increased. For
691 the moderately and strongly deleterious DFEs, α was only overestimated at high selfing levels,
692 with the strongly deleterious DFE corresponding with the most accurate inference of α. This is
693 consistent with the expectation that a decrease in the efficacy of selection will increase the
694 fixation of weakly deleterious alleles, which will be misinterpreted as beneficial fixations by
695 GRAPES, leading to an overestimation of α at high selfing levels. Very similar trends were
696 observed when the folded SFS was used to estimate α (Supplementary Figures 11-12). Overall,
697 at high selfing levels, α values were substantially overestimated by GRAPES.
698
699 **Effects of population structure and sampling on the accuracy of inference**
700 A final factor that could bias DFE inference in partially selfing populations is population
701 structure because the inference procedure of DFE-alpha and GRAPES is not designed to account
702 for the effects of population structure on the SFS. DFE inference in the context of population
703 structure has been largely unexplored, though the effects of cryptic structure (without migration)
704 were tested in Kim *et al.* (2017). Here, we simulated an island model, with five demes that
705 exchanged migrants at a total rate m each generation. Multiple values of m were used to vary the
706 degree of isolation between demes in this model, and the number of individuals in each deme
707 (N_{deme}) was varied so that the effective size of the metapopulation (N_{meta}) was equivalent to
708 500,000 individuals (scaled by a factor of 100; Supplementary Table 7). Under strict neutrality
709 and while selfing was taken into account (see Methods), the neutral nucleotide diversity of the
710 entire metapopulation, of individual subpopulations, and F_{ST} was equal to theoretical
711 expectations (Supplementary Table 8). We performed simulations at $N_{deme}m = 0.1,\ 0.5,$
712 $1,$ and $2,$ which corresponded to $F_{ST} = 0.76,\ 0.38,\ 0.24,$ and 0.14 respectively at 99% selfing.

713 These values were chosen to mimic F_{ST} values observed empirically amongst *C. elegans*
714 populations, which range from 0.44 to 0.83 (Barrière and Félix 2005; Richaud *et al.* 2018).
715 Population structure by itself did not result in mis-inference of the DFE when the rate of
716 migration was sufficiently high (*i.e.*, $N_{deme}m = 2, 1,$ and 0.5). High selfing rates with
717 population structure again led to inferring a DFE skewed almost entirely towards mildly
718 deleterious mutations, thus producing a very similar extent of mis-inference as that observed in a
719 panmictic population (Supplementary Figure 13 A, B, and C, respectively). However, at low
720 rates of migration between populations ($N_{deme}m$ of 0.1), the DFE was mis-inferred by DFE-
721 alpha and GRAPES, even at 0% selfing, with both programs tending to infer a lower efficacy of
722 selection than the simulated DFE (Figure 8; Supplementary Figure 13). This was largely because
723 the relative increase in intermediate frequency alleles due to population structure as opposed to
724 panmixia (Supplementary Figures 14-15) was falsely inferred as a bottleneck (mean fold change
725 = 0.45-0.50; similar to Chikhi *et al.* 2010) by DFE-alpha, leading to incorrect estimates of the
726 DFE. As expected, these distortions of intermediate frequency alleles in the SFS caused by
727 population structure were observed to increase at higher selfing levels and at lower migration
728 rates, with particularly strong deviations from panmixia observed at $N_{deme}m = 0.1$.
729 Empirical studies that draw from worldwide population samples may knowingly or
730 unknowingly sample unevenly from structured populations when performing DFE inference. To
731 investigate if this has unintended consequences on DFE inference, we repeated our experiments
732 with population structure, altering our sampling strategy to primarily draw genomes from one
733 deme (60%) or two demes (35% from each), with the remaining demes contributing only 10% of
734 the sampled genomes each. As expected, the site frequency spectra from both unevenly sampled
735 populations were observed to be even more distorted at intermediate frequencies than the evenly
736 sampled populations with low migration rates (Supplementary Figures 16-19). For the sampling
737 strategy primarily drawing from one deme, some DFE mis-inference towards nearly neutral and
738 strongly deleterious mutations was observed at 0% and 50% selfing even for $N_{deme}m = 1$ and
739 more strongly at $N_{deme}m = 0.5$, and this effect was stronger at the DFE with predominantly
740 weakly deleterious mutations (Supplementary Figure 20). This effect was also observed for the
741 sampling strategy primarily drawing from two demes, but to a lesser extent (Supplementary
742 Figure 21). At the lowest migration rate, the inferred DFEs for both uneven sampling cases were
743 substantially less deleterious at 0% and 50% selfing than for even sampling, particularly for the

744 DFE with predominantly weakly deleterious mutations (Figure 8). At higher selfing levels, the
745 inferred DFEs for all sampling schemes falsely inferred a larger proportion of weakly deleterious
746 mutations than predicted by the DFE corrected for BGS (Figure 8), likely due to a combination
747 of HRI and uneven sampling. The BGS-corrected DFE was unable to correct for the mis-
748 inference of the DFE for all selfing levels at the lowest migration rate, reflecting the fact that
749 DFE mis-inference is likely driven by changes in the SFS caused by population structure rather
750 than changes in the level of BGS.
751 Overall, these results indicate both DFE inference approaches are fairly robust to
752 population structure in an island model, even with uneven sampling, until the populations
753 become extremely isolated from each other, or individuals are mostly drawn from one deme,
754 especially when the true DFE has a significant proportion of mildly deleterious mutations.
755 Additionally, both low rates of migration and highly uneven sampling result in the inference of
756 an even more weakly deleterious DFE than simulated. Testing DFE inference with more models
757 of population structure and additional sampling strategies might reveal additional biases in DFE
758 inference approaches.
759
760
761 **DISCUSSION**
762 Inference of the distribution of fitness effects (DFE) of new mutations is a critical step in
763 understanding the forces that shape genetic variation in a population. Population genetic methods
764 to infer the DFE that account for nonequilibrium demography usually assume independent
765 segregation of putatively neutral and selected sites (Eyre-Walker and Keightley 2007; Keightley
766 and Eyre-Walker 2007; Galtier 2016; Tataru *et al.* 2017). However, organisms that have little/ no
767 recombination, for instance, those that undergo high rates of selfing or asexual reproduction, are
768 likely to violate this assumption. A major consequence of strong linkage is the increase in BGS,
769 selective sweeps, and Hill-Robertson effects that lower the efficacy of selection. Particularly in
770 the regime where HRI is pervasive, many linked selected polymorphisms segregate in a
771 population, reducing the population's variance in fitness and thus the efficacy of selection
772 (Comeron and Kreitman 2002; Kaiser and Charlesworth 2009; Seger *et al.* 2010; O'Fallon *et al.*
773 2010; Nicolaisen and Desai 2013). In this study, we demonstrate that in highly selfing species
774 and in regions of low recombination, the DFE of new deleterious mutations is mis-inferred due

775 to effects of selection at linked sites. Although accounting for the relative decrease in the rate of
776 coalescence due to BGS and selective sweeps can provide reasonably accurate estimates of the
777 DFE in moderately selfing populations, inference remains biased when rates of recombination
778 are very low, likely due to interference between selected alleles. We also find that while purging
779 of partially recessive deleterious mutations does not lead to mis-inference, the presence of
780 cryptic population structure where there is a low rate of migration between subpopulations
781 and/or sampling from predominantly one subpopulation leads to inferring a DFE skewed towards
782 effectively neutral/mildly deleterious class of mutations. Moreover, the proportion of beneficial
783 substitutions (α) is substantially overestimated in highly selfing species, suggesting caution
784 when performing inference in species or genomic regions with little/no recombination.
785

786 **Why is the DFE estimated from *C. elegans* populations not skewed towards mildly**
787 **deleterious mutations?**

788 While our study suggests that DFE inference in highly selfing species should yield DFEs skewed
789 towards weakly/moderately deleterious mutations, surprisingly, the previous empirical estimate
790 of the *C. elegans* DFE using DFE-alpha is heavily skewed towards the class of strongly
791 deleterious mutations ($2N_e\overline{s_d} > 10{,}000$, Gilbert *et al.* 2022). One explanation of this
792 discrepancy is that the mean $2N_e\overline{s_d}$ of the population is even greater than 10,000. Interestingly,
793 direct measurements of fitness-related traits in mutation accumulation lines suggest that *C.*
794 *elegans* has a highly deleterious DFE, with estimates of the mean deleterious selection
795 coefficient ranging from -0.09 to -0.46 (Vassilieva *et al.* 2000; Baer *et al.* 2005; Katju *et al.*
796 2015). Thus, using previously calculated *C. elegans* N_e ranging from 5,000 to 80,000 (Table 4 in
797 Teterina *et al.* 2023), the expected $2NB\overline{s_d}$ would range from 900 to 72,000, consistent with this
798 explanation. However, there are a few major caveats to these observations. (a) The experimental
799 estimates of selection coefficients are likely biased towards more deleterious selection
800 coefficients due to the difficulty of measuring small fitness changes in experimental lines
801 (Davies *et al.* 1999; Eyre-Walker and Keightley 2007), and might not reflect fitness effects in the
802 wild. (b) It is unclear whether previous estimates of the population sizes of *C. elegans*
803 populations using neutral nucleotide diversity (Barrière and Félix 2005; Cutter 2006; Rockman
804 and Kruglyak 2009; Teterina *et al.* 2023) are accurate due to the joint effects of selfing, HRI, and
805 demography. However, natural populations may have similar sizes to the ones assumed in our

simulations ($N = 500{,}000$) as our simulations resulted in neutral nucleotide diversity values close to natural *C. elegans* populations for the DFEs with more strongly deleterious mutations, with 95% ($\pi \sim 0.001$) and 99% selfing ($\pi \sim 0.0003$) populations having values close to those reported in a Hawaiian population with ~93% selfing ($\pi \sim 0.0012$) and for French populations with ~99% selfing ($\pi \sim 0.00002 - 0.0008$; Barrière and Félix 2005; Teterina *et al.* 2023). (c) Empirical estimates of the deleterious DFE in *C. elegans* (Gilbert *et al.* 2022) were obtained by sampling worldwide rather than from local populations, and thus other factors such as complex demographic history and population structure could play a role as well. For instance, selfing populations are likely to follow extinction-recolonization dynamics (Charlesworth and Wright 2001; Ingvarsson 2002; Wright *et al.* 2013) and such models remain understudied.

A second explanation for the discrepancy between the expected and observed DFE in *C. elegans* is that the true selfing rate in wild *C. elegans* might be lower than 99%, thus reducing linked effects of selection (Figure 1). Though selfing rate estimates from unique populations range from 80% to ~100% using various methods (Barrière and Félix 2005; Sivasundar and Hey 2005; Teterina *et al.* 2023), in general, the strong LD in *C. elegans* populations (haplotype blocks of 2.5 Mb on average; Cutter 2006) along with low frequency of males in wild specimens (Barrière and Félix 2005; Richaud *et al.* 2018) suggest the outcrossing rate is very low in most populations. A recent study using whole-genome sequencing data focusing on Hawaiian *C. elegans* estimated a much higher selfing rate of 93%, suggesting that selfing rate can vary across populations (Teterina *et al.* 2023). Importantly, as the samples used in Gilbert *et al.* (2022) were unevenly sampled worldwide (with many samples from Hawaiian populations), variation in selfing rate across populations or through time could additionally influence DFE inference results.

A third possibility relates to the difficulty in obtaining the SFS using highly inbred strains. The SNP-calling pipeline used to obtain *C. elegans* population-genetic data converted heterozygous sites into homozygous reference or alternate alleles depending on if the mutant was above or below 50% (Cook *et al.* 2017). This method of SNP processing could lead to the inference of a more deleterious DFE if the heterozygote frequency is higher at neutral sites than selected sites by skewing the neutral SFS towards higher frequencies and the selected SFS towards lower frequencies, though the low frequency of heterozygotes in highly selfing populations makes this unlikely to be a large issue. Future studies working with inbred strains

837 should evaluate the effect of employing such SNP processing approaches on population genomic
838 inference.
839 Finally, we have only examined a population under equilibrium in the current study,
840 whereas a recent population expansion/ contraction could decrease the observed effects of
841 interference between selected alleles at multiple loci, allowing for a correct inference of selection
842 coefficients. HRI effects in non-equilibrium populations remain to be studied more thoroughly.
843 Unfortunately, as linked effects of selection will also lead to falsely inferring recent population
844 growth (Supplementary Table 1; Ewing and Jensen 2016; Schrider *et al.* 2016; Johri *et al.* 2021),
845 there is no straightforward solution to accounting for the effects of the true demography (Johri *et*
846 *al.* 2022b). Thus performing a joint inference of demography and the DFE as employed by Johri
847 *et al.* (2020) is likely to be the most promising approach to circumvent this challenge.
848
849 **How broadly do our results apply**?
850 Previous studies (*e.g*., Messer and Petrov 2013; Gilbert *et al.* 2022) that have observed similar
851 biases in inference using DFE-alpha due to linked effects of selection suggested that this might
852 be because DFE-alpha cannot account for the false signal of extreme recent growth generated by
853 hitchhiking effects. Our study shows that DFE mis-inference was similarly observed when using
854 the method GRAPES, which employs a reweighting scheme of the selected relative to the neutral
855 SFS and does not explicitly model a specific demographic history. We, therefore, conclude that
856 the observed mis-inference is a general phenomenon (likely due to a combination of BGS,
857 sweeps, and HRI effects) and not due to errors unique to a particular method like DFE-alpha.
858 Our results have implications for the population genetics of other highly selfing
859 organisms, many of which will fall into a parameter space where linked effects of selection have
860 a strong effect on genomic diversity. For instance, highly selfing species with large N_e such as
861 *Arabidopsis thaliana*, *Capsella rubella*, *Capsella bursa-pastoris*, and *Capsella orientalis* have all
862 been shown to have more weakly deleterious DFEs than comparable outcrossing species (*A.*
863 *lyrata* and *C. grandiflora;* Douglas *et al.* 2015; Chen *et al.* 2020)*,* consistent with effects of
864 selection at linked sites potentially playing an important role in determining their genomic
865 diversity. Of course, the extent of HRI in these species would likely differ due to differences in
866 outcrossing rates, population sizes, mutation and recombination rates, and the genome
867 architecture (McVean and Charlesworth 2000; Comeron and Kreitman 2002; Kaiser and

868 Charlesworth 2009; Good *et al.* 2014). For instance, the rates of LD decay in the outcrossing
869 species *C. remanei* and *A. lyrata* (1-2 kb; Ross-Ibarra *et al.* 2008; Teterina *et al.* 2023) are much
870 higher than the rates in *C. elegans* and *A. thaliana*, but the rate of decay is significantly smaller
871 in *C. elegans* where a complete decay of LD occurs over several Mb (Cutter 2006; Rockman and
872 Kruglyak 2009) compared to *A. thaliana* (10-50 kb; Ross-Ibarra *et al.* 2008; Shimizu and
873 Tsuchimatsu 2015), suggesting a higher possibility of HRI affecting DFE inference in *C.*
874 *elegans*. In addition, pathogenic species like *Plasmodium* that undergo largely asexual
875 reproduction (and thus experience reduced rates of recombination), seem to have nearly identical
876 neutral and selected site frequency spectra (Figure 3 in Parobek *et al.* 2016), suggesting that HRI
877 effects might play an important role in shaping variation across their genomes. We therefore also
878 expect DFE inference to be biased due to HRI in asexual organisms, mitochondrial genomes, and
879 the Y chromosome, all of which have low effective recombination rates (Barton and
880 Charlesworth 1998; Weinreich and Rand 2000; Charlesworth and Charlesworth 2000).
881 An interesting exception is the previous study performed in *Eichhornia paniculata*
882 populations (Arunkumar *et al.* 2015), where effects of selection at linked sites did not confound
883 inference of population-scaled selection coefficients. Their simulated populations were
884 comparable to ours – a high selfing rate (98%), a mutation to recombination rate ratio of 0.14 (*vs*.
885 0.11 in this study), a deleterious DFE (with mean $N\overline{s_d}\sim 50-100$), and 15,000,000 selected sites
886 (*vs*. 562,500 in this study). However, the expected values of NU in the full haploid genomes of
887 *E. paniculata vs. C. elegans* are ~ 25 and ~ 37,125 respectively, where U is the number of
888 mutations occurring at nonsynonymous sites per genome/generation, assuming 20% and 30%
889 coding density respectively. Higher NU will lead to stronger effects of BGS and has been shown
890 to cause populations with moderate selection coefficients that would otherwise be in the classical
891 BGS regime to have pervasive HRI (Good *et al.* 2014), so it is plausible that differences in NU
892 explain the differences between our studies.
893
894 **Considerations for inferring the DFE in partially selfing organisms**
895 Given our observations, we here provide a few points to consider when inferring the DFE of new
896 mutations in organisms with little/no recombination, where effects of BGS, selective sweeps, or
897 HRI might be pervasive:

898 (1) The DFE of new deleterious mutations estimated using the two-step population-
899 genetic methods is most closely explained by the distribution of $2NBs_d$, which is equivalent to
900 $2N_{emp}(1+F)s_d$ in partially selfing populations. Here N_{emp} represents the effective population
901 size estimated using neutral nucleotide diversity (not accounting for selfing), which includes the
902 effects of BGS and is what is usually estimated using demographic inference approaches that do
903 not account for linked effects of selection. Thus, if there exist accurate experimentally validated
904 estimates of selfing rates, it is possible to accurately infer the distribution of selection
905 coefficients in populations with moderate levels of selfing. However, in many organisms selfing
906 rates are inferred using allele frequencies and those estimates are likely to be biased by
907 population history and effects of selection at linked sites. Moreover, this raises the question of
908 how one can accurately simulate an estimated DFE when multiple linked selected sites are to be
909 simulated, which will be especially important to determine the extent of HRI effects in
910 populations with little/no recombination.
911 (2) In addition, different models of population structure and the effects of demographic
912 changes need to be considered to fully understand the robustness of DFE inference. For instance,
913 it is known that local populations of selfing organisms are prone to recurrent extinction and
914 recolonization events (Pannell and Charlesworth 1999; Ingvarsson 2002) which may bias DFE
915 inference further. Indeed, recent field studies of *C. elegans* populations suggest that 3-10
916 individuals colonize a food source and quickly reproduce to census sizes of up to 10,000
917 individuals before dispersing in search of new food sources (Barrière and Félix 2005; Richaud *et*
918 *al.* 2018). When sampling across many demes, we suggest also inferring the DFE in isolated
919 populations in order to minimize biases due to cryptic population structure or uneven sampling
920 and to downsample the number of genomes such that sampling is as even as possible across
921 demes. Sampling strategies such as a scattered sampling scheme may be useful in removing such
922 bias (Wakeley 1999; Städler *et al.* 2009), though the reduction in sample size may reduce the
923 ability to accurately infer the moderate and deleterious portions of the DFE (Kim *et al.* 2017).
924 (3) If possible, it could be extremely helpful to infer the DFE in closely related non-
925 selfing species or in different populations of the same species with known differences in selfing
926 rates. For instance, *C. elegans* and *A. thaliana* populations vary in the estimated rates of selfing
927 by ~10%, and the mostly outcrossing *A. lyrata* has several populations that have evolved partial
928 selfing (Mable *et al.* 2005; Bomblies *et al.* 2010; Teterina *et al.* 2023). Though differences such

929 as population structure and demography need to be carefully accounted for, such comparisons
930 across populations/species could also be used to approximate the extent of HRI in a selfing
931 species.
932 (4) Finally, in the cases where HRI is determined to be extensive in a partially selfing
933 population, joint inference of demography and the DFE as employed by Johri *et al.* in *D.*
934 *melanogaster* (2020) and human (2022) populations could be especially useful for DFE
935 inference. They employed an approximate Bayesian computation (Beaumont *et al.* 2002)
936 framework combined with forward simulations to jointly infer parameters of the DFE of new
937 deleterious mutations and an arbitrary size change, while accounting for the effects of selection
938 at linked sites. While highly successful, such an approach has been limited to a relatively simple
939 demographic model of a single size change in one population due to the large number of
940 parameters required to be jointly estimated. Though it would be computationally intensive, such
941 an approach could yield the most likely distribution of $2Ns_d$. More importantly, the distribution
942 of $2Ns_d$ describing the fitness effects of new mutations is likely not the distribution of $2Ns_d$
943 describing the fixation probabilities of new mutations, which will be relatively smaller.
944 Furthermore, the distortions of the SFS from a neutral model caused by interference cannot be
945 explained by any constant N_e (see discussion in Neher 2013; Good *et al.* 2014; Melissa *et al.*
946 2022). Thus, until theoretical models can accurately estimate the effects of HRI in partially
947 selfing organisms with complex DFEs, simulation-based approaches will be the most helpful in
948 inferring the DFE of new mutations in such populations, as well as obtaining the fixation
949 probability of new mutations, which will be altered by HRI.
950
951
952 **ACKNOWLEDGEMENTS**
953 We thank Brian Charlesworth for many helpful discussions related to the project and for
954 providing suggestions and comments that improved the manuscript. We are also thankful to Matt
955 Rockman for many helpful discussions and for pointing us to relevant resources available for *C.*
956 *elegans*. The research in this study was conducted using computational resources provided by
957 ITS Research Computing at the University of North Carolina at Chapel Hill. A.T.D. received
958 support from NIGMS predoctoral training grant T32-GM135123. The authors declare no
959 conflicts of interests.

DATA AVAILABILITY

The following data are publicly available on https://github.com/JohriLab/DFESelfing: (1) Data and scripts to reproduce figures; (2) Config files used for DFE inference; (3) Simulation outputs in SLiM format; (4) templates and example scripts for launching SLiM simulations.

REFERENCES

Agrawal A. F., and M. C. Whitlock, 2011 Inferences about the distribution of dominance drawn from Yeast gene knockout data. Genetics 187: 553–566.

Anderson J. L., L. T. Morran, and P. C. Phillips, 2010 Outcrossing and the maintenance of males within *C. elegans* populations. J. Hered. 101: S62–S74.

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.

Baer C. F., F. Shaw, C. Steding, M. Baumgartner, A. Hawkins, *et al.*, 2005 Comparative evolutionary genetics of spontaneous mutations affecting fitness in rhabditid nematodes. Proc. Natl. Acad. Sci. USA 102: 5785–5790.

Bank C., G. B. Ewing, A. Ferrer-Admettla, M. Foll, and J. D. Jensen, 2014 Thinking too positive? Revisiting current methods of population genetic selection inference. Trends Genet. 30: 540–546.

Barrière A., and M.-A. Félix, 2005 High local genetic diversity and low outcrossing rate in *Caenorhabditis elegans* natural populations. Curr. Biol. 15: 1176–1184.

Barrière A., and M.-A. Félix, 2007 Temporal dynamics and linkage disequilibrium in natural *Caenorhabditis elegans* populations. Genetics 176: 999–1011.

Barton N. H., and B. Charlesworth, 1998 Why sex and recombination? Science 281: 1986–1990.

Beaumont M. A., W. Zhang, and D. J. Balding, 2002 Approximate Bayesian computation in population genetics. Genetics 162: 2025–2035.

Becher H., B. C. Jackson, and B. Charlesworth, 2020 Patterns of genetic variability in genomic regions with low rates of recombination. Curr. Biol. 30: 94-100.e3.

Bomblies K., L. Yant, R. A. Laitinen, S.-T. Kim, J. D. Hollister, *et al.*, 2010 Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. PLoS Genet. 6: e1000890.

C. elegans Sequencing Consortium, 1998 Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science 282: 2012–8.

Caballero A., P. D. Keightley, and W. G. Hill, 1991 Strategies for increasing fixation probabilities of recessive mutations. Genet. Res. 58: 129–138.

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

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

Charlesworth D., and B. Charlesworth, 1987 Inbreeding depression and its evolutionary consequences. Annu. Rev. Ecol. Syst. 18: 237–268.

1002 Charlesworth B., 1992 Evolutionary rates in partially self-fertilizing species. Am. Nat. 140: 126–
1003 148.
1004 Charlesworth B., M. T. Morgan, and D. Charlesworth, 1993 The effect of deleterious mutations
1005 on neutral molecular variation. Genetics 134: 1289–1303.
1006 Charlesworth B., 1994 The effect of background selection against deleterious mutations on
1007 weakly selected, linked variants. Genet. Res. 63: 213–227.
1008 Charlesworth B., and D. Charlesworth, 1998 Some evolutionary consequences of deleterious
1009 mutations. Genetica 102: 3–19.
1010 Charlesworth B., and D. Charlesworth, 2000 The degeneration of Y chromosomes. Philos. Trans.
1011 R. Soc. Lond. B. Biol. Sci. 355: 1563–1572.
1012 Charlesworth D., and S. I. Wright, 2001 Breeding systems and genome evolution. Curr. Opin.
1013 Genet. Dev. 11: 685–690.
1014 Charlesworth D., 2003 Effects of inbreeding on the genetic diversity of populations. Philos.
1015 Trans. R. Soc. Lond. B Biol. Sci. 358: 1051–1070.
1016 Charlesworth B., and D. Charlesworth, 2010 *Elements of Evolutionary Genetics*. Roberts and
1017 Company, Greenwoord Village, Colorado, USA.
1018 Charlesworth B., and J. D. Jensen, 2021 Effects of selection at linked sites on patterns of genetic
1019 variability. Annu. Rev. Ecol. Syst. 52: 177–197.
1020 Chen J., S. Glémin, and M. Lascoux, 2020 From drift to draft: How much do beneficial
1021 mutations actually contribute to predictions of Ohta's slightly deleterious model of
1022 molecular evolution? Genetics 214: 1005–1018.
1023 Chikhi L., V. C. Sousa, P. Luisi, B. Goossens, and M. A. Beaumont, 2010 The confounding
1024 effects of population structure, genetic diversity and the sampling scheme on the
1025 detection and quantification of population size changes. Genetics 186: 983–995.
1026 Comeron J. M., M. Kreitman, and M. Aguadé, 1999 Natural selection on synonymous sites is
1027 correlated with gene length and recombination in *Drosophila*. Genetics 151: 239–249.
1028 Comeron J. M., and M. Kreitman, 2002 Population, evolutionary and genomic consequences of
1029 interference selection. Genetics 161: 389–410.
1030 Cook D. E., S. Zdraljevic, J. P. Roberts, and E. C. Andersen, 2017 CeNDR, the *Caenorhabditis*
1031 *elegans* natural diversity resource. Nucleic Acids Res. 45: D650–D657.
1032 Cutter A. D., 2006 Nucleotide polymorphism and linkage disequilibrium in wild populations of
1033 the partial selfer *Caenorhabditis elegans*. Genetics 172: 171–184.
1034 Davies E. K., A. D. Peters, and P. D. Keightley, 1999 High frequency of cryptic deleterious
1035 mutations in *Caenorhabditis elegans*. Science 285: 1748–1751.
1036 Denver D. R., K. Morris, M. Lynch, and W. K. Thomas, 2004 High mutation rate and
1037 predominance of insertions in the *Caenorhabditis elegans* nuclear genome. Nature 430:
1038 679–682.
1039 Douglas G. M., G. Gos, K. A. Steige, A. Salcedo, K. Holm, *et al.*, 2015 Hybrid origins and the
1040 earliest stages of diploidization in the highly successful recent polyploid *Capsella bursa-*
1041 *pastoris*. Proc. Natl. Acad. Sci. USA 112: 2806–2811.
1042 Estes S., P. C. Phillips, D. R. Denver, W. K. Thomas, and M. Lynch, 2004 Mutation
1043 accumulation in populations of varying size: The distribution of mutational effects for
1044 fitness correlates in *Caenorhabditis elegans*. Genetics 166: 1269–1279.
1045 Ewing G. B., and J. D. Jensen, 2016 The consequences of not accounting for background
1046 selection in demographic inference. Mol. Ecol. 25: 135–141.

1047 Eyre-Walker A., M. Woolfit, and T. Phelps, 2006 The distribution of fitness effects of new
1048 deleterious amino acid mutations in humans. Genetics 173: 891–900.
1049 Eyre-Walker A., and P. D. Keightley, 2007 The distribution of fitness effects of new mutations.
1050 Nat. Rev. Genet. 8: 610–618.
1051 Eyre-Walker A., and P. D. Keightley, 2009 Estimating the rate of adaptive molecular evolution
1052 in the presence of slightly deleterious mutations and population size change. Mol. Biol.
1053 Evol. 26: 2097–2108.
1054 Felsenstein J., 1974 The evolutionary advantage of recombination. Genetics 78: 737–756.
1055 Galtier N., 2016 Adaptive protein evolution in animals and the effective population size
1056 hypothesis. PLoS Genet. 12: e1005774.
1057 Garcia J. A., and K. E. Lohmueller, 2021 Negative linkage disequilibrium between amino acid
1058 changing variants reveals interference among deleterious mutations in the human
1059 genome. PLoS Genet. 17: e1009676.
1060 Gilbert K. J., S. Zdraljevic, D. E. Cook, A. D. Cutter, E. C. Andersen, *et al.*, 2022 The
1061 distribution of mutational effects on fitness in *Caenorhabditis elegans* inferred from
1062 standing genetic variation, (L. Moyle, Ed.). Genetics 220: iyab166.
1063 Glémin S., 2003 How are deleterious mutations purged? Drift versus nonrandom mating.
1064 Evolution 57: 2678–2687.
1065 Good B. H., A. M. Walczak, R. A. Neher, and M. M. Desai, 2014 Genetic diversity in the
1066 interference selection limit. PLoS Genet. 10: e1004222.
1067 Goodwillie C., S. Kalisz, and C. G. Eckert, 2005 The evolutionary enigma of mixed mating
1068 systems in plants: Occurrence, theoretical explanations, and empirical evidence. Annu.
1069 Rev. Ecol. Syst. 36: 47–79.
1070 Haller B. C., and P. W. Messer, 2023 SLiM 4: Multispecies eco-evolutionary modeling. Am.
1071 Nat. 201: E127–E139.
1072 Hartfield M., and T. Bataillon, 2020 Selective sweeps under dominance and inbreeding. G3 10:
1073 1063–1075.
1074 Hedrick P. W., 1980 Hitchhiking: a comparison of linkage and partial selfing. Genetics 94: 791–
1075 808.
1076 Hill W. G., and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genet.
1077 Res. 8: 269–294.
1078 Hodgkin J., 1983 Male phenotypes and mating efficiency in *Caenorhabditis elegans*. Genetics
1079 103: 43–64.
1080 Huang X., A. L. Fortier, A. J. Coffman, T. J. Struck, M. N. Irby, *et al.*, 2021 Inferring genome-
1081 wide correlations of mutation fitness effects between populations. Mol. Biol. Evol. 38:
1082 4588–4602.
1083 Huber C. D., B. Y. Kim, C. D. Marsden, and K. E. Lohmueller, 2017 Determining the factors
1084 driving selective effects of new nonsynonymous mutations. Proc. Natl. Acad. Sci. USA
1085 114: 4465–4470.
1086 Ingvarsson P. K., 2002 A metapopulation perspective on genetic diversity and differentiation in
1087 partially self-fertilizing plants. Evolution 56: 2368–2373.
1088 Jarne P., and J. R. Auld, 2006 Animals mix it up too: the distribution of self-fertilization among
1089 hermaphroditic animals. Evolution 60: 1816–1824.
1090 Johri P., B. Charlesworth, and J. D. Jensen, 2020 Toward an evolutionarily appropriate null
1091 model: jointly inferring demography and purifying selection. Genetics 215: 173–192.

Johri P., K. Riall, H. Becher, L. Excoffier, B. Charlesworth, *et al.*, 2021 The impact of purifying and background selection on the inference of population history: problems and prospects. Mol. Biol. Evol. 38: 2986–3003.

Johri P., C. F. Aquadro, M. Beaumont, B. Charlesworth, L. Excoffier, *et al.*, 2022a Recommendations for improving statistical inference in population genomics. PLoS Biol. 20: e3001669.

Johri P., A. Eyre-Walker, R. N. Gutenkunst, K. E. Lohmueller, and J. D. Jensen, 2022b On the prospect of achieving accurate joint estimation of selection with population history. Genome Biol. Evol. 14: evac088.

Kaiser V. B., and B. Charlesworth, 2009 The effects of deleterious mutations on evolution in non-recombining genomes. Trends Genet. 25: 9–12.

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

Katju V., L. B. Packard, L. Bu, P. D. Keightley, and U. Bergthorsson, 2015 Fitness decline in spontaneous mutation accumulation lines of *Caenorhabditis elegans* with varying effective population sizes. Evolution 69: 104–16.

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.

Kim B. Y., C. D. Huber, and K. E. Lohmueller, 2017 Inference of the distribution of selection coefficients for new nonsynonymous mutations using large samples. Genetics 206: 345–361.

Konrad A., M. J. Brady, U. Bergthorsson, and V. Katju, 2019 Mutational landscape of spontaneous base substitutions and small indels in experimental *Caenorhabditis elegans* populations of differing size. Genetics 212: 837–854.

Kousathanas A., and P. D. Keightley, 2013 A comparison of models to infer the distribution of fitness effects of new mutations. Genetics 193: 1197–1208.

Lewontin R. C., 1964 The interaction of selection and linkage. I. general considerations; heterotic models. Genetics 49: 49–67.

Lewontin R. C., 1988 On measures of gametic disequilibrium. Genetics 120: 849–852.

Mable B. K., A. V. Robertson, S. Dart, C. Di Berardo, and L. Witham, 2005 Breakdown of self-incompatibility in the perennial *Arabidopsis lyrata* (*Brassicaceae*) and its genetic consequences. Evolution 59: 1437–1448.

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

Maruyama T., 1970 Effective number of alleles in a subdivided population. Theor. Popul. Biol. 1: 273–306.

McVean G. A., and B. Charlesworth, 2000 The effects of Hill-Robertson interference between weakly selected mutations on patterns of molecular evolution and variation. Genetics 155: 929–944.

Melissa M. J., B. H. Good, D. S. Fisher, and M. M. Desai, 2022 Population genetics of polymorphism and divergence in rapidly evolving populations. Genetics 221: iyac053.

Messer P. W., and D. A. Petrov, 2013 Frequent adaptation and the McDonald-Kreitman test. Proc. Natl. Acad. Sci. USA 110: 8615–8620.

Neher R. A., 2013 Genetic draft, selective interference, and population genetics of rapid adaptation. Annu. Rev. Ecol. Syst. 44: 195–215.

Nei M., 1968 The frequency distribution of lethal chromosomes in finite populations. Proc. Natl. Acad. Sci. USA 60: 517–524.
Nei M., 1973 Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70: 3321–3323.
Nei M., and F. Tajima, 1981 DNA polymorphism detectable by restriction endonucleases. Genetics 97: 145–163.
Nicolaisen L. E., and M. M. Desai, 2013 Distortions in genealogies due to purifying selection and recombination. Genetics 195: 221–230.
Nordborg M., 2000 Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. Genetics 154: 923–929.
O'Fallon B. D., J. Seger, and F. R. Adler, 2010 A continuous-state coalescent and the impact of weak selection on the structure of gene genealogies. Mol. Biol. Evol. 27: 1162–1172.
Pannell J. R., and B. Charlesworth, 1999 Neutral genetic diversity in a metapopulation with recurrent local extinction and recolonization. Evolution 53: 664–676.
Parobek C. M., J. T. Lin, D. L. Saunders, E. J. Barnett, C. Lon, *et al.*, 2016 Selective sweep suggests transcriptional regulation may underlie *Plasmodium vivax* resilience to malaria control measures in Cambodia. Proc. Natl. Acad. Sci. USA 113: E8096–E8105.
Pollak E., 1987 On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics 117: 353–360.
Richaud A., G. Zhang, D. Lee, J. Lee, and M.-A. Félix, 2018 The Local Coexistence Pattern of Selfing Genotypes in Caenorhabditis elegans Natural Metapopulations. Genetics 208: 807–821.
Rockman M. V., and L. Kruglyak, 2009 Recombinational landscape and population genomics of *Caenorhabditis elegans*. PLoS Genet. 5: e1000419.
Ross-Ibarra J., S. I. Wright, J. P. Foxe, A. Kawabe, L. DeRose-Wilson, *et al.*, 2008 Patterns of polymorphism and demographic history in natural populations of *Arabidopsis lyrata*. PLoS One 3: e2411.
Roze D., and T. Lenormand, 2005 Self-fertilization and the evolution of recombination. Genetics 170: 841–857.
Sanjuán R., A. Moya, and S. F. Elena, 2004 The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. Proc. Natl. Acad. Sci. USA 101: 8396–8401.
Saxena A. S., M. P. Salomon, C. Matsuba, S.-D. Yeh, and C. F. Baer, 2019 Evolution of the mutational process under relaxed selection in *Caenorhabditis elegans*. Mol. Biol. Evol. 36: 239–251.
Schrider D. R., A. G. Shanku, and A. D. Kern, 2016 Effects of linked selective sweeps on demographic inference and model selection. Genetics 204: 1207–1223.
Seger J., W. A. Smith, J. J. Perry, J. Hunn, Z. A. Kaliszewska, *et al.*, 2010 Gene genealogies strongly distorted by weakly interfering mutations in constant environments. Genetics 184: 529–545.
Shimizu K. K., and T. Tsuchimatsu, 2015 Evolution of selfing: recurrent patterns in molecular adaptation. Annu. Rev. Ecol. Syst. 46: 593–622.
Sicard A., and M. Lenhard, 2011 The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. Ann. Bot. 107: 1433–1443.
Simmons M. J., and J. F. Crow, 1977 Mutations affecting fitness in *Drosophila* populations. Ann. Rev. Genet. 11: 49–78.

Sivasundar A., and J. Hey, 2005 Sampling from natural populations with RNAI reveals high outcrossing and population structure in Caenorhabditis elegans. Curr. Biol. 15: 1598–1602.

Slatkin M., and L. Voelm, 1991 F_{ST} in a hierarchical island model. Genetics 127: 627–629.

Städler T., B. Haubold, C. Merino, W. Stephan, and P. Pfaffelhuber, 2009 The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. Genetics 182: 205–216.

Tataru P., M. Mollion, S. Glémin, and T. Bataillon, 2017 Inference of distribution of fitness effects and proportion of adaptive substitutions from polymorphism data. Genetics 207: 1103–1119.

Teterina A. A., J. H. Willis, M. Lukac, R. Jovelin, A. D. Cutter, *et al.*, 2023 Genomic diversity landscapes in outcrossing and selfing *Caenorhabditis* nematodes. PLoS Genet. 19: e1010879.

Thornton K., 2003 libsequence: a C++ class library for evolutionary genetic analysis. Bioinformatics 19: 2325–2327.

Vassilieva L. L., A. M. Hook, and M. Lynch, 2000 The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. Evolution 54: 1234–1246.

Wade E. E., C. C. Kyriazis, M. I. A. Cavassim, and K. E. Lohmueller, 2023 Quantifying the fraction of new mutations that are recessive lethal. Evolution 77: 1539–1549.

Wakeley J., 1999 Nonequilibrium migration in human history. Genetics 153: 1863–1871.

Walczak A. M., L. E. Nicolaisen, J. B. Plotkin, and M. M. Desai, 2012 The structure of genealogies in the presence of purifying selection: a fitness-class coalescent. Genetics 190: 753–779.

Watterson G. A., 1975 On the number of segregating sites in genetical models without recombination. Theor. Popul. Biol. 7: 256–276.

Weinreich D. M., and D. M. Rand, 2000 Contrasting patterns of nonneutral evolution in proteins encoded in nuclear and mitochondrial genomes. Genetics 156: 385–399.

Williamson S. H., R. Hernandez, A. Fledel-Alon, L. Zhu, R. Nielsen, *et al.*, 2005 Simultaneous inference of selection and population growth from patterns of variation in the human genome. Proc. Natl. Acad. Sci. USA 102: 7882–7887.

Wloch D. M., K. Szafraniec, R. H. Borts, and R. Korona, 2001 Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae*. Genetics 159: 441–452.

Wright S. I., S. Kalisz, and T. Slotte, 2013 Evolutionary consequences of self-fertilization in plants. Proc. R. Soc. Lond. B 280: 20130133.

FIGURES AND TABLES

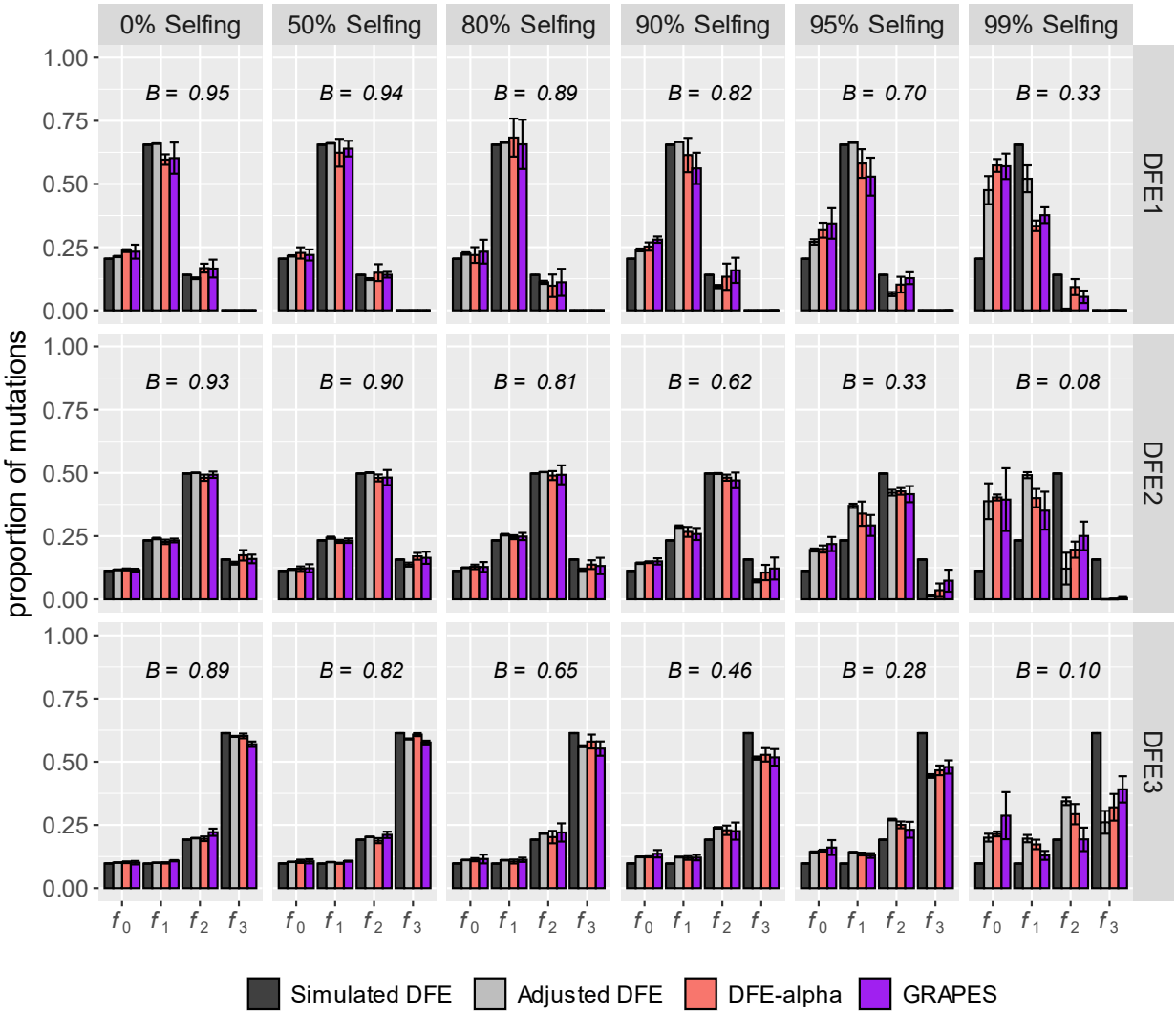


Figure 1: Effects of selfing on the inference of the DFE of new deleterious mutations using DFE-alpha (red bars) and GRAPES (purple bars). Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. Selfing rates were varied between 0 and 99 %. The nucleotide site diversity with background selection (B) relative to its expectation under strict neutrality is shown in each panel. The simulated (black bars) and adjusted (grey bars) DFE represents the distribution of $2Ns_d$ and $2NBs_d$ respectively. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.

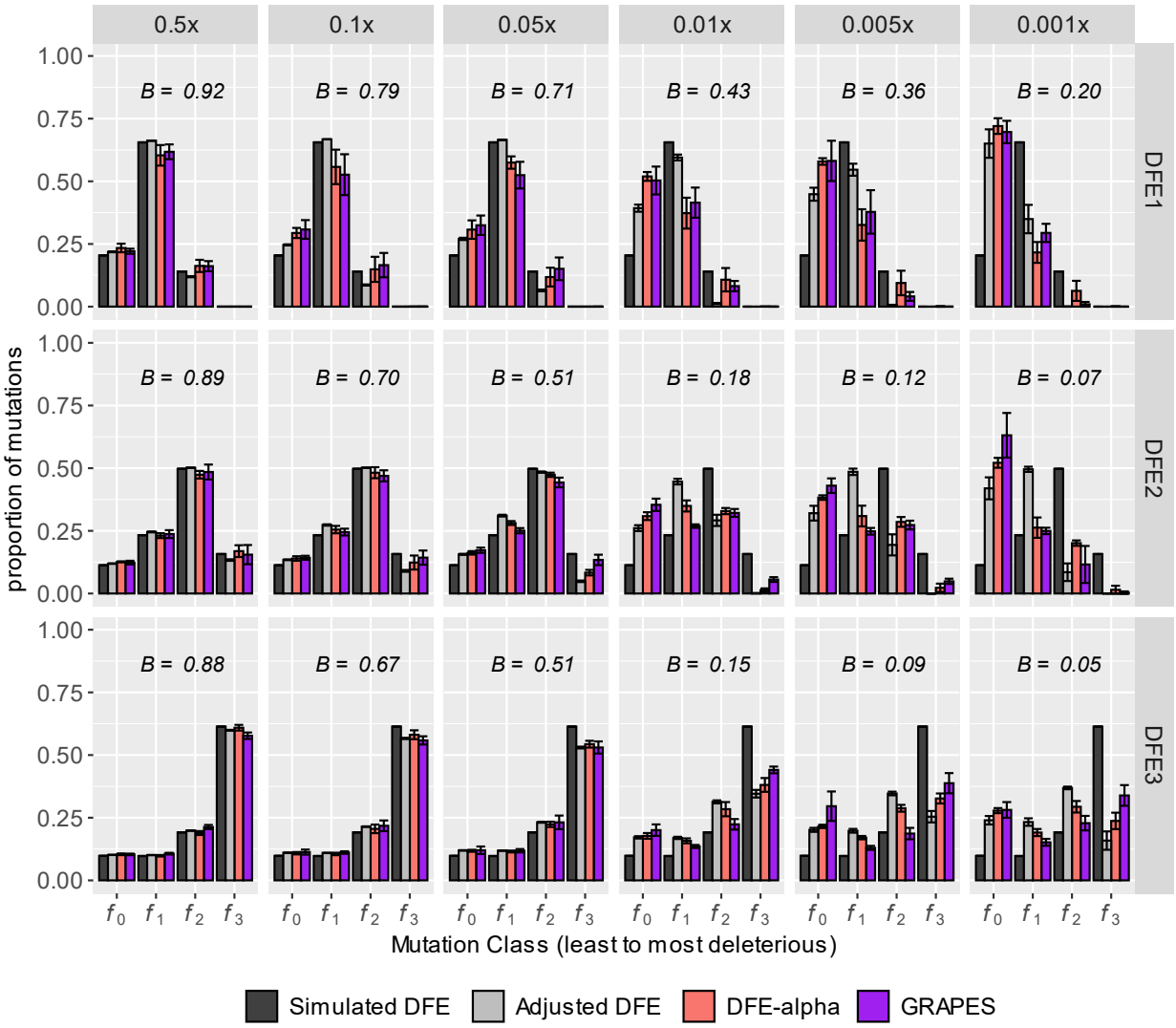


Figure 2: Effects of low recombination rate on the inference of the DFE of new deleterious mutations using DFE-alpha (red bars) and GRAPES (purple bars). Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. Recombination rates were varied between $0.5 \times$ mean and $0.001 \times$ mean, where the mean recombination rate was 3.12×10^{-8} per site/generations. The nucleotide site diversity with background selection (B) relative to its expectation under strict neutrality is shown in each panel. The simulated DFE (shown in black) and adjusted DFE (shown in grey) represents the distribution of $2Ns_d$ and $2NBs_d$ respectively. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.

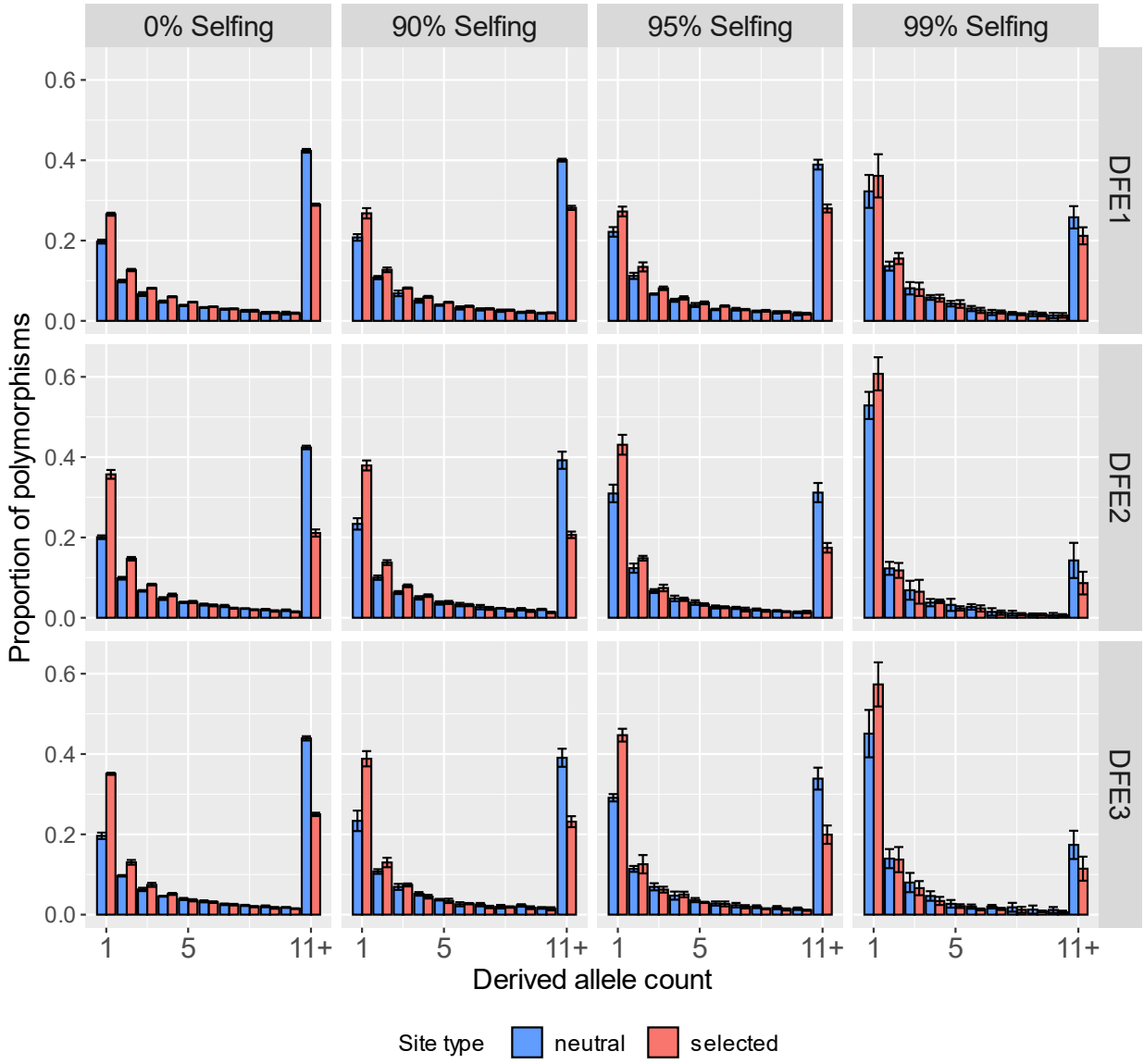


Figure 3: Site frequency spectra of neutral and selected alleles from 100 genomes sampled from simulations where all selected mutations were deleterious.The y-axis represents the proportion of segregating polymorphisms that fall into the given derived allele count. The last class (11+) refers to the derived allele counts 11-100. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.

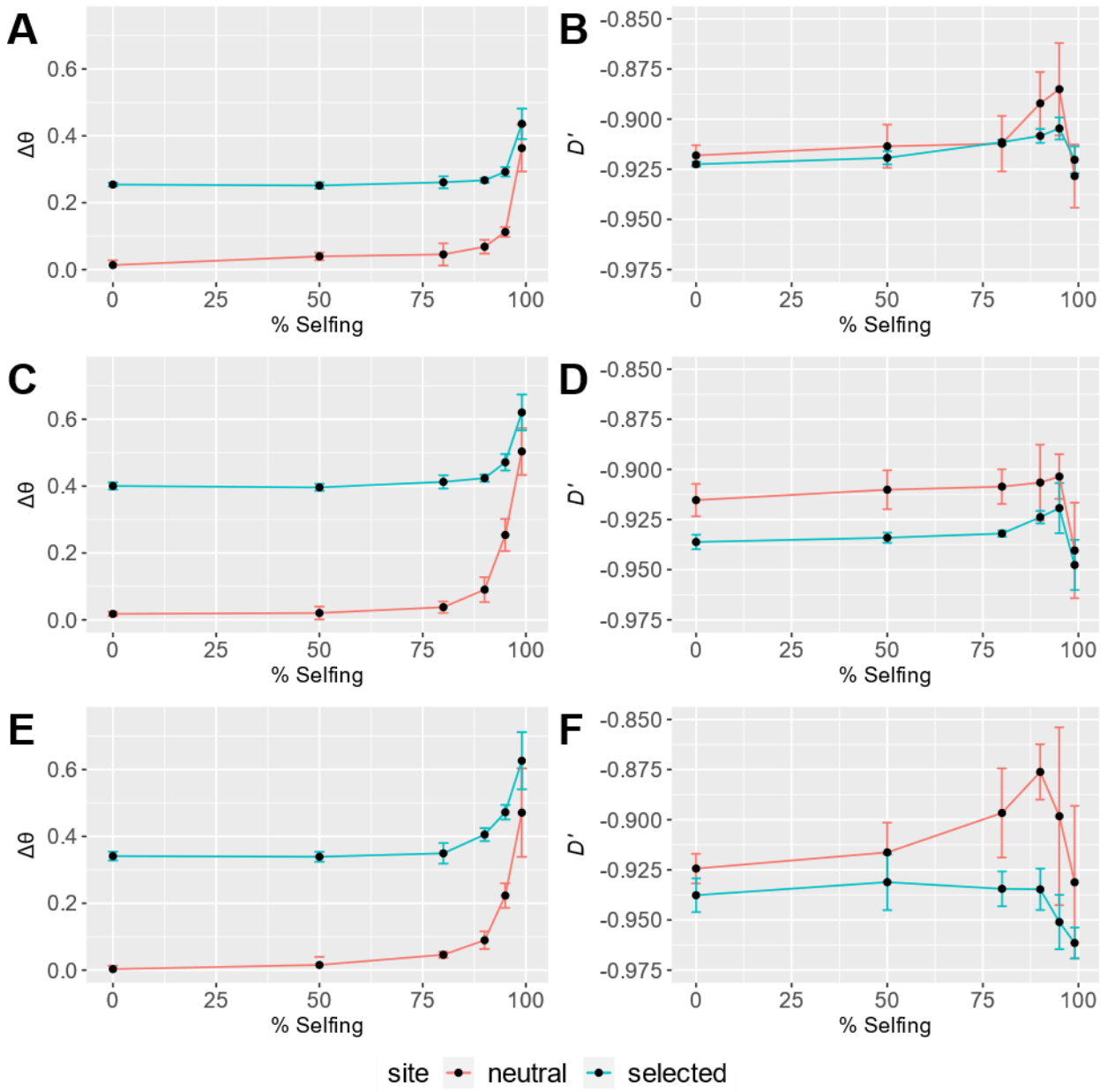


Figure 4: Effects of selfing on the SFS and linkage disequilibrium (LD) at neutral (red lines) and selected (blue lines) sites. The skew in the SFS was quantified using $\Delta\theta$, calculated using alleles at all frequencies while LD was summarized by the statistic D', calculated for alleles at frequencies of 1%-5%. The error bars denote the standard deviation of the genome-wide average of each summary statistic estimated from 5 independent replicates.

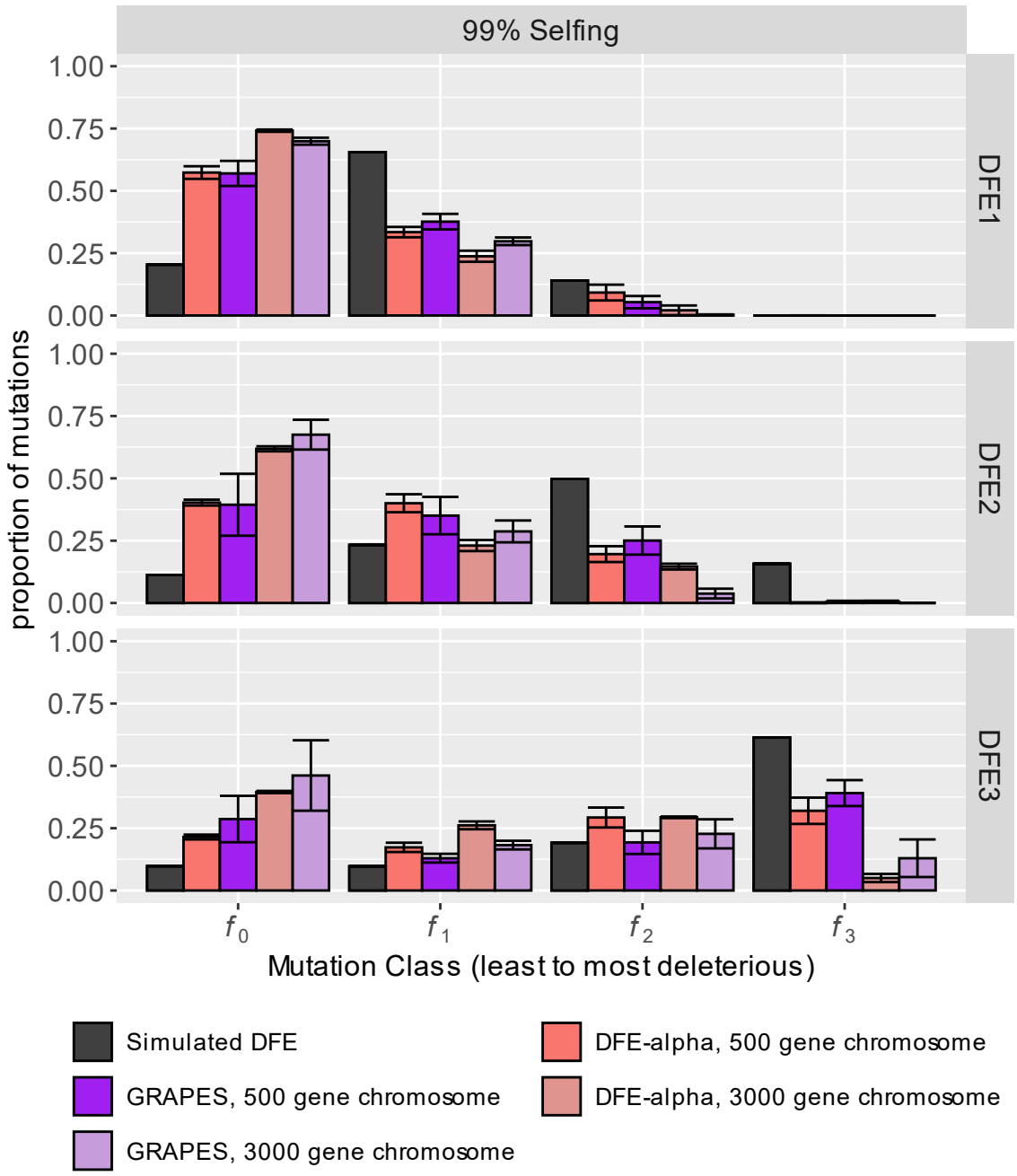


Figure 5: The inference of the DFE of new deleterious mutations when varying chromosome sizes were simulated: a chromosome with 500 genes (shown in darker shades) *vs*. a chromosome with 3000 genes (shown in lighter shades). Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. Inference was performed using DFE-alpha (red bars) and GRAPES (purple bars). The error bars denote the standard deviation of proportions estimated from 5 independent replicates.

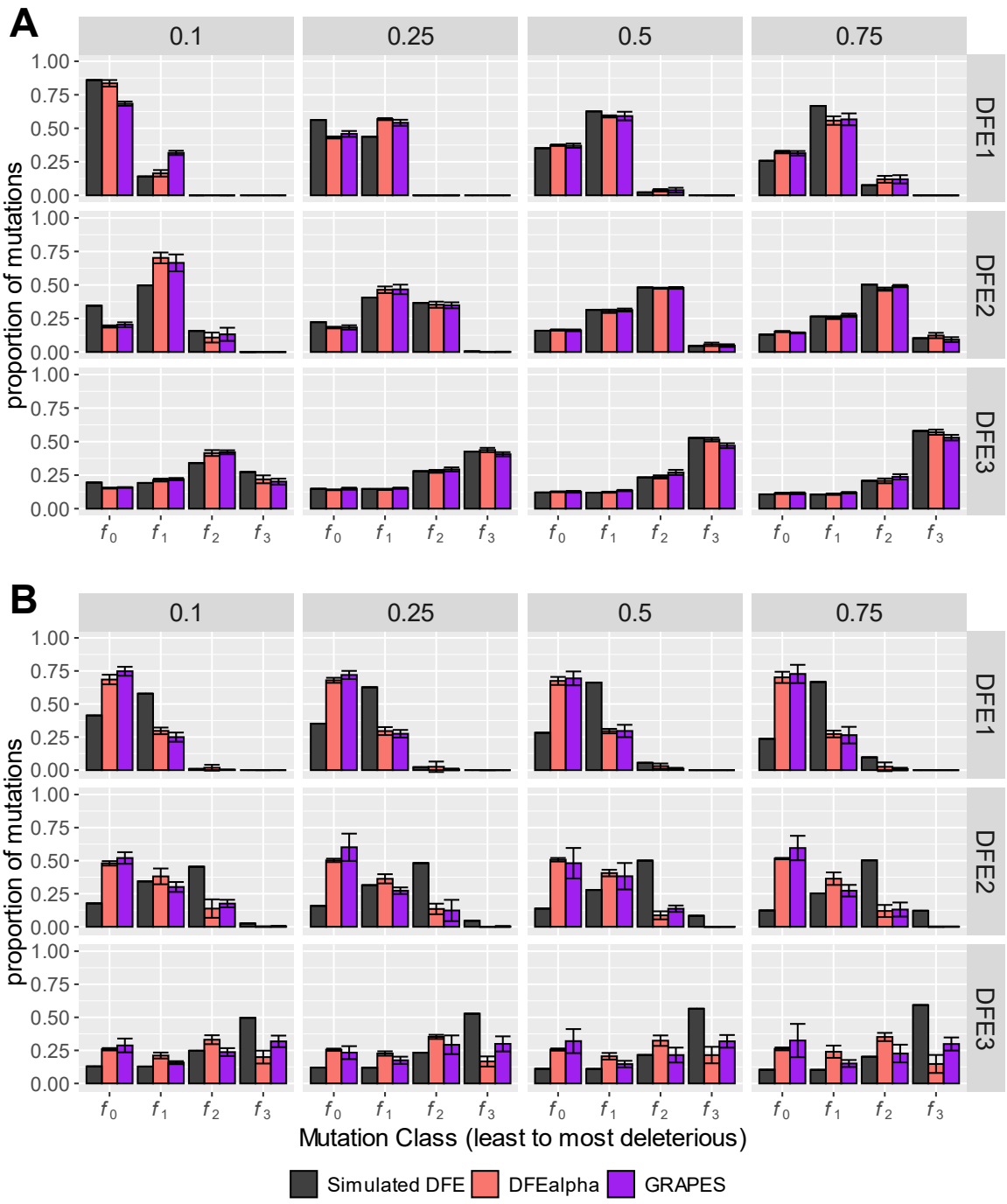


Figure 6: Effects of selfing on DFE inference of new deleterious mutations with varying dominance coefficients using DFE-alpha (red bars) and GRAPES (purple bars). The dominance of selected mutations is noted at the top of each column. Results are shown for populations with **(A)** 0% selfing and **(B)** 99% selfing. Here the effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes represent $0 < 2N_{self}s_d h_{self} < 1$, $1 < 2N_{self}s_d h_{self} < 10$, $10 < 2N_{self}s_d h_{self} < 100$, and $100 < 2N_{self}s_d h_{self} < 2N_{self}$, respectively. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.

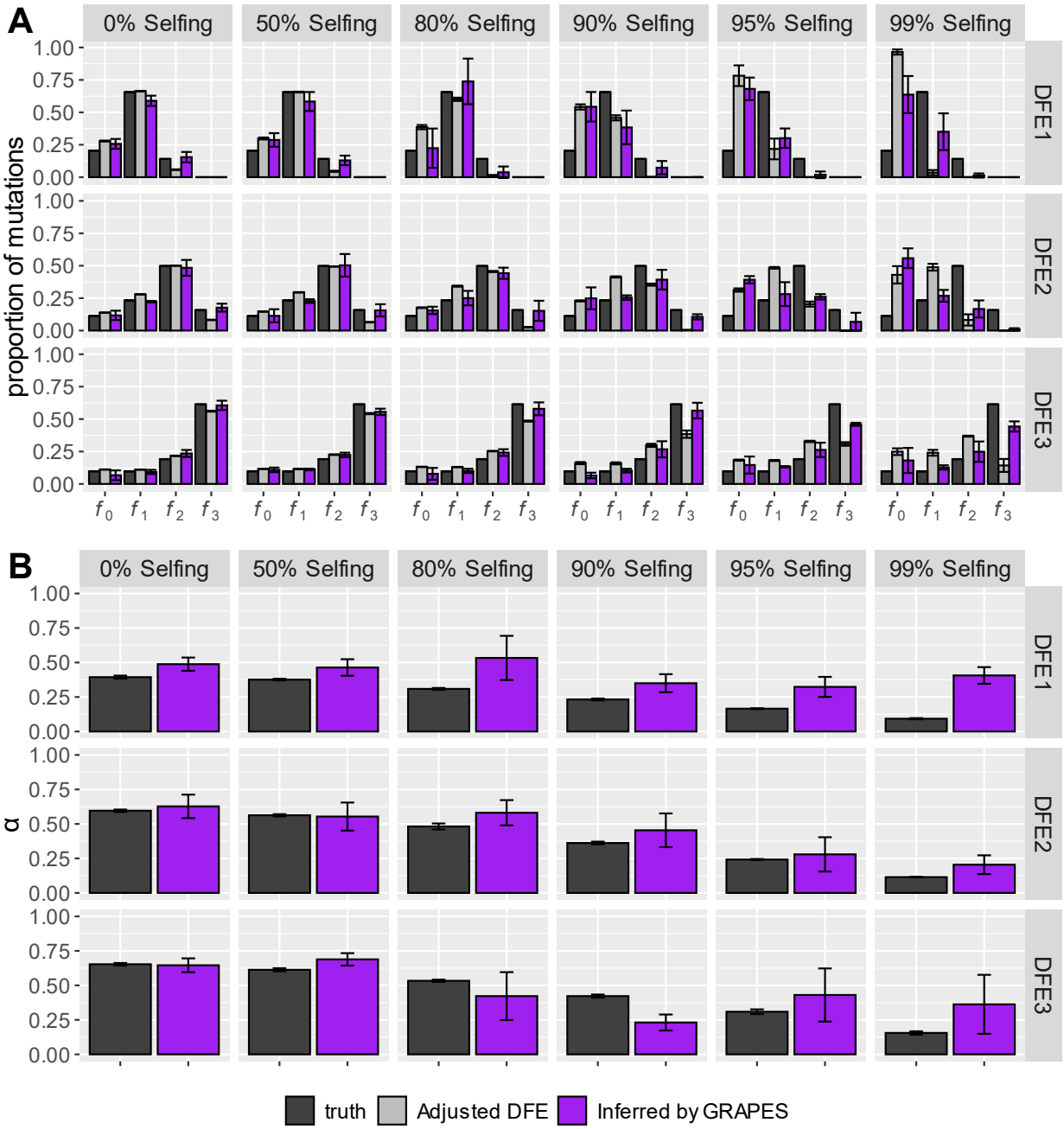


Figure 7. Effects of selfing on the inference of (A) the DFE of new deleterious mutations and (B) the proportion of beneficial fixations (α) using GRAPES (shown in purple bars). (A) Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. The adjusted DFE represents the distribution of $2N(\pi_{obs}/\pi_{exp})s_d$, where π_{obs} is the neutral nucleotide site diversity observed in each scenario and π_{exp} is the neutral nucleotide site diversity expected in a population of size N at strict neutrality. For both (A) and (B), 0.1% of new exonic mutations were beneficial and exponentially distributed with a mean $2Ns_a = 200$. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.

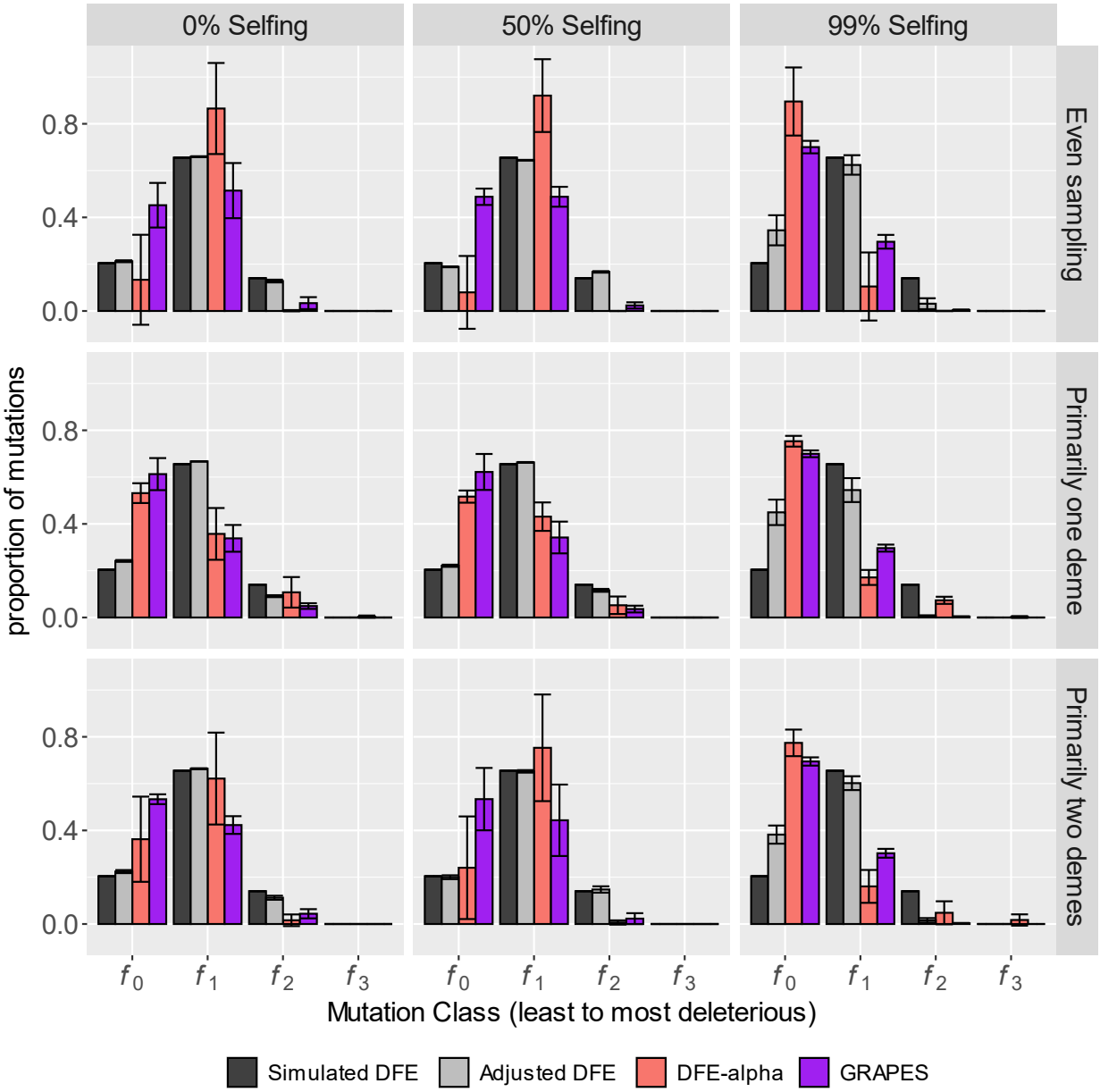


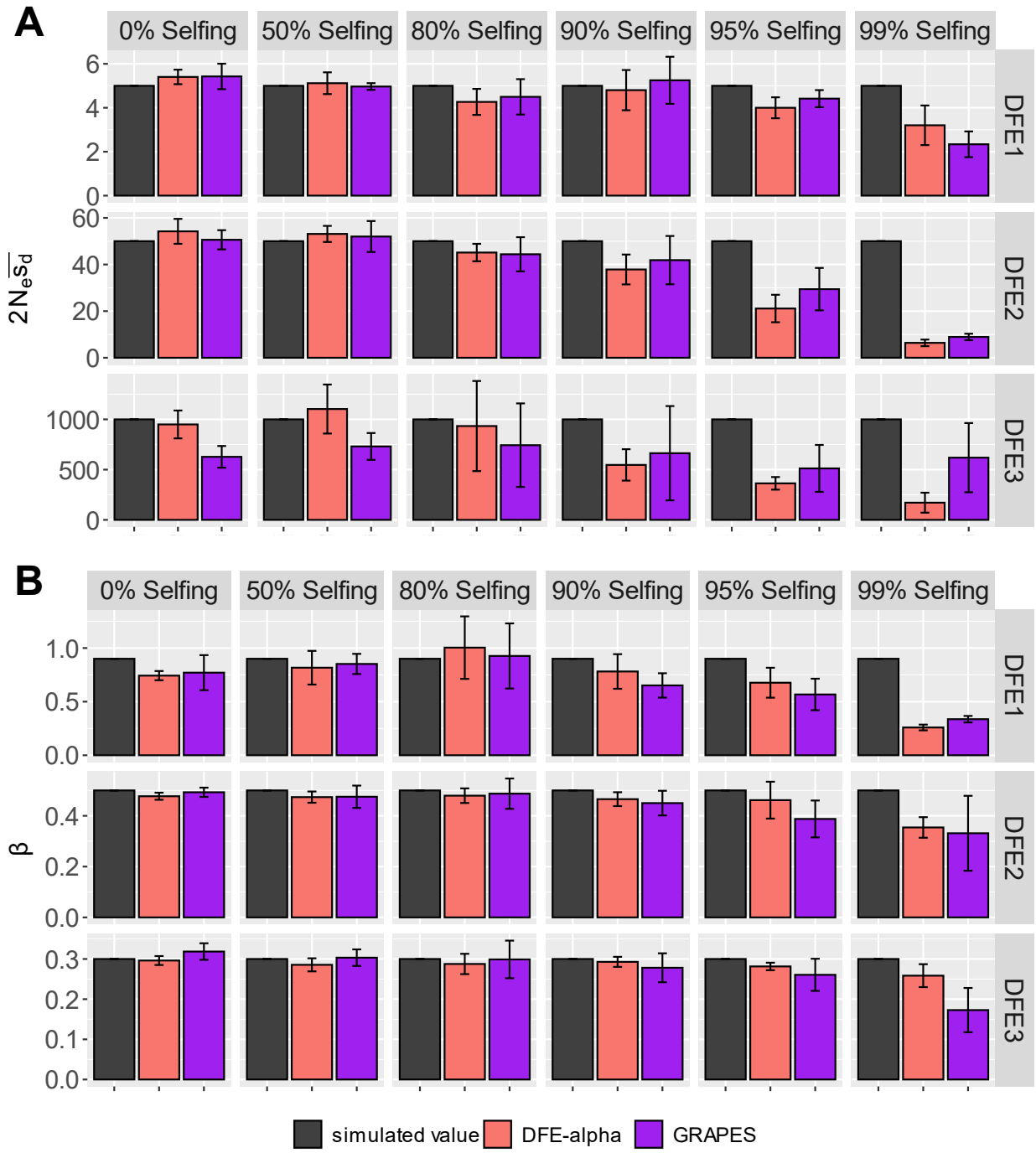
Figure 8 Effects of uneven sampling, population structure, and selfing on the inference of the DFE of new deleterious mutations using DFE-alpha (red bars) and GRAPES (purple bars). Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. Populations were simulated under an island model with five demes such that the metapopulation effective size was equal to 5000 at 0% selfing and $N_{deme}m$ was 0.1. Results are shown for the most weakly deleterious DFE (*i.e.*, DFE1). The error bars denote the standard deviation of proportions estimated from 5 independent replicates. Rows represent the sampling scheme for genomes, where sampling was even, sampling was primarily from one deme, or sampling was primarily from two demes.

1299 **Table 1**: Parameters of the simulated DFEs of deleterious mutations in terms of the proportion of
1300 mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and
1301 strongly deleterious (f_3) classes. DFEs consisted of predominantly mildly deleterious (DFE1),
1302 moderately deleterious (DFE2), or strongly deleterious (DFE3) mutations.

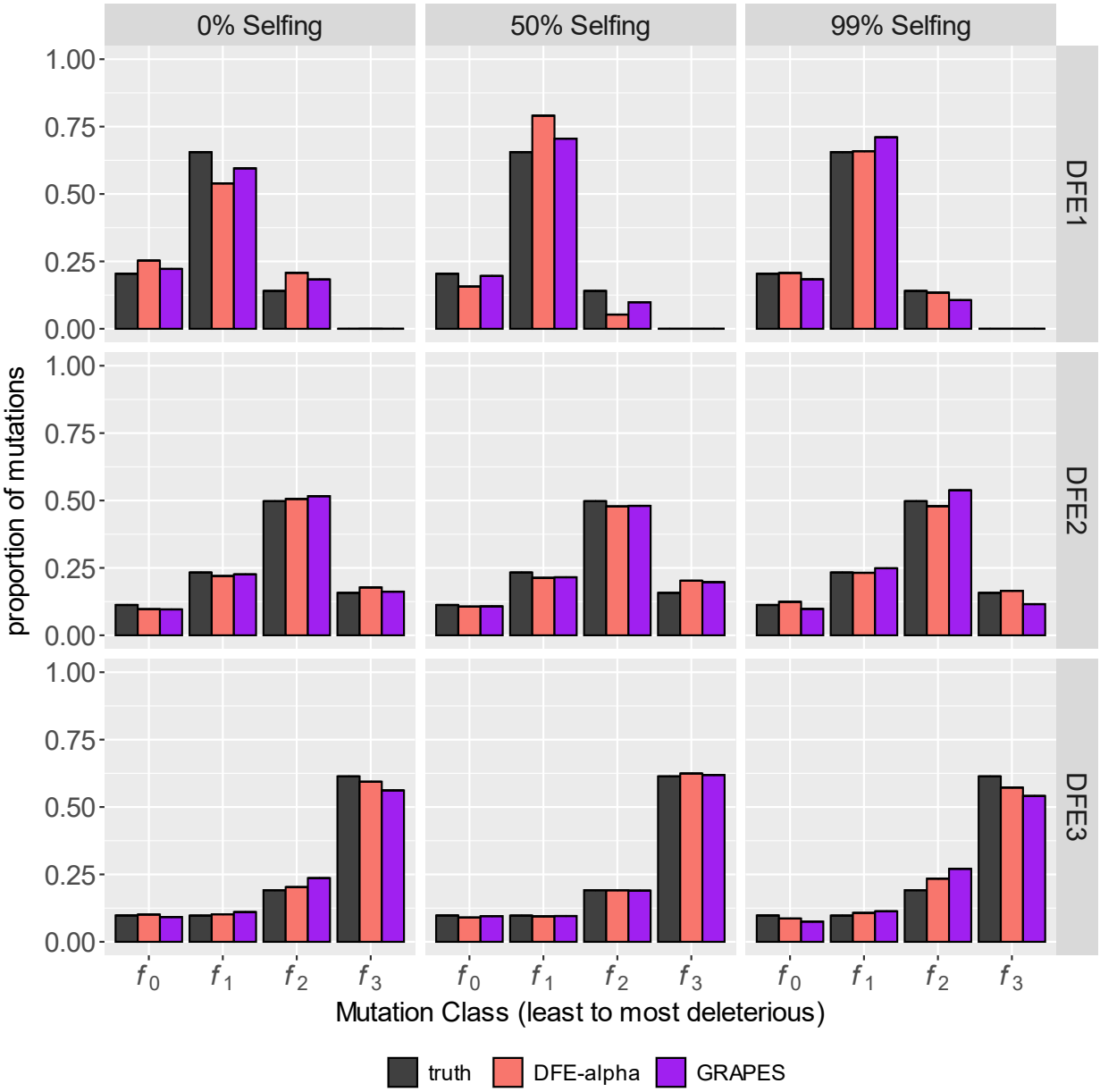
	β	$2Ns_d$	f_0	f_1	f_2	f_3
DFE1	0.9	5	0.204	0.655	0.140	0.000
DFE2	0.5	50	0.113	0.233	0.497	0.157
DFE3	0.3	1000	0.098	0.097	0.192	0.614

1303

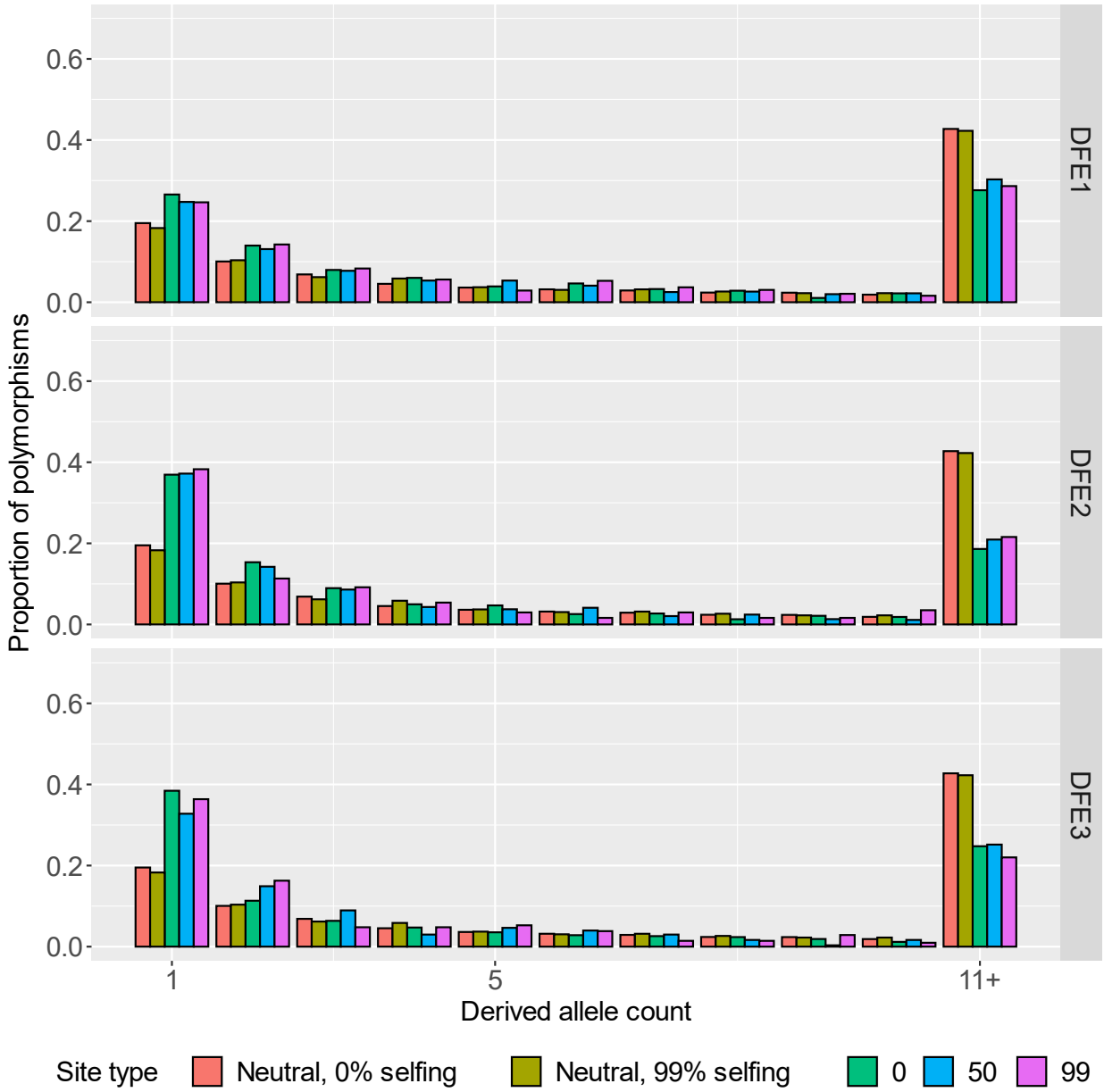
SUPPLEMENTARY FIGURES AND TABLES



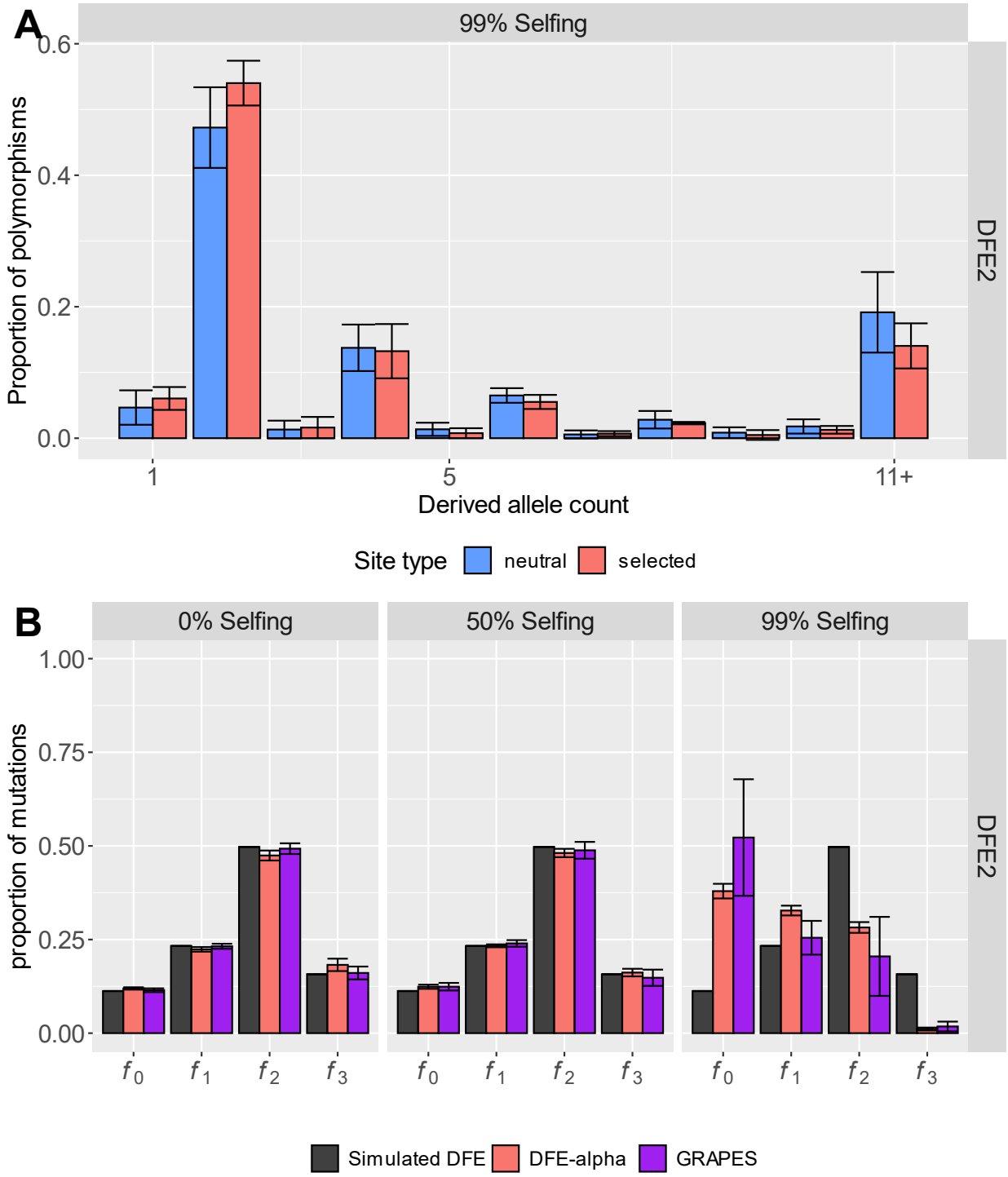
Supplementary Figure 1 Accuracy of inference for (**A**) mean selective effects ($\gamma = 2N_e\overline{s_d}$) and (**B**) shape parameters (β) of deleterious mutations in partially selfing populations. Here black bars represent simulated values, while red and purple bars show the values estimated by DFE-alpha and GRAPES respectively. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



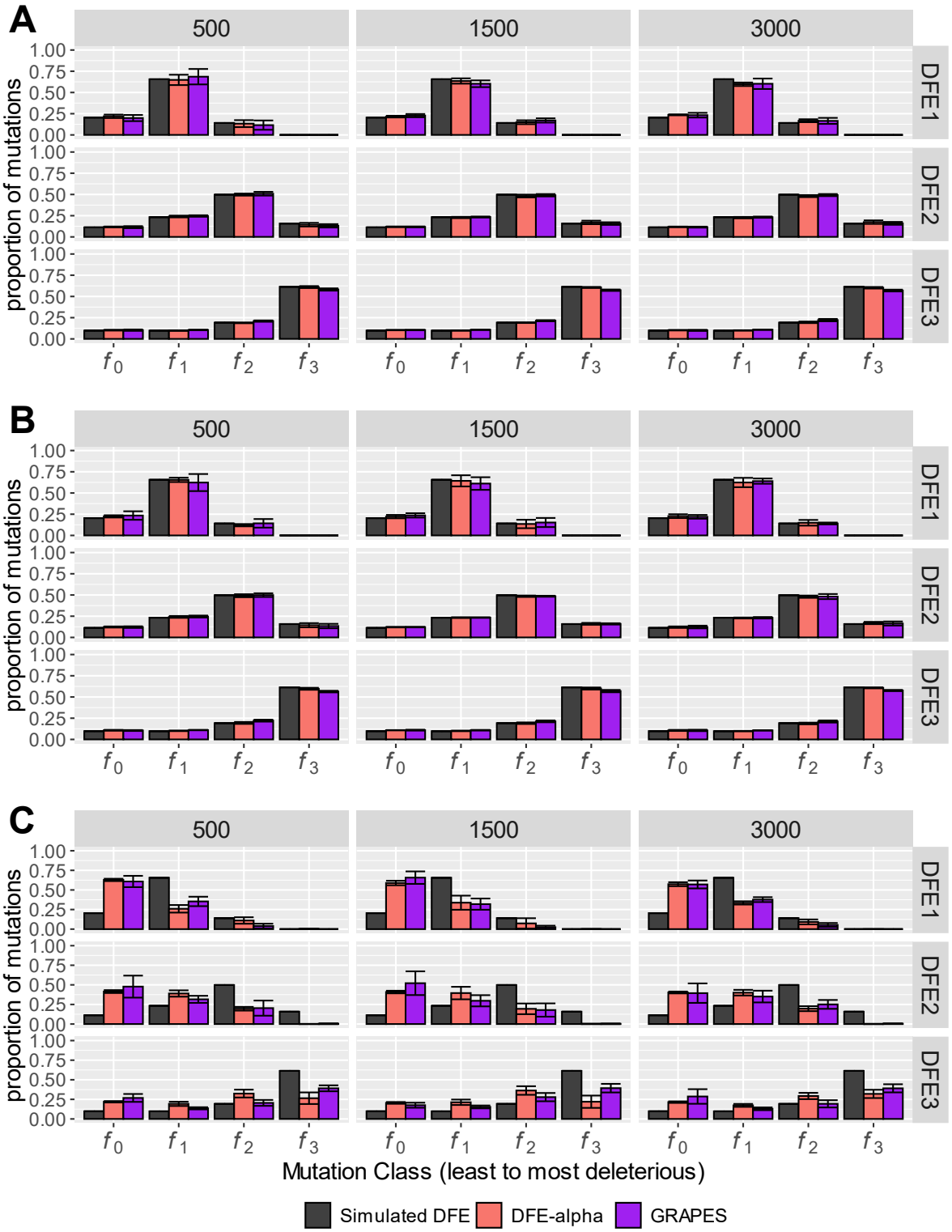
Supplementary Figure 2 Inference of the DFE of new deleterious mutations in partially selfing populations using DFE-alpha and GRAPES for simulations with no linkage. Here, 50,000 selected mutations were simulated individually (as single locus simulations) in various selfing rates without linkage to other selected mutations. Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes.



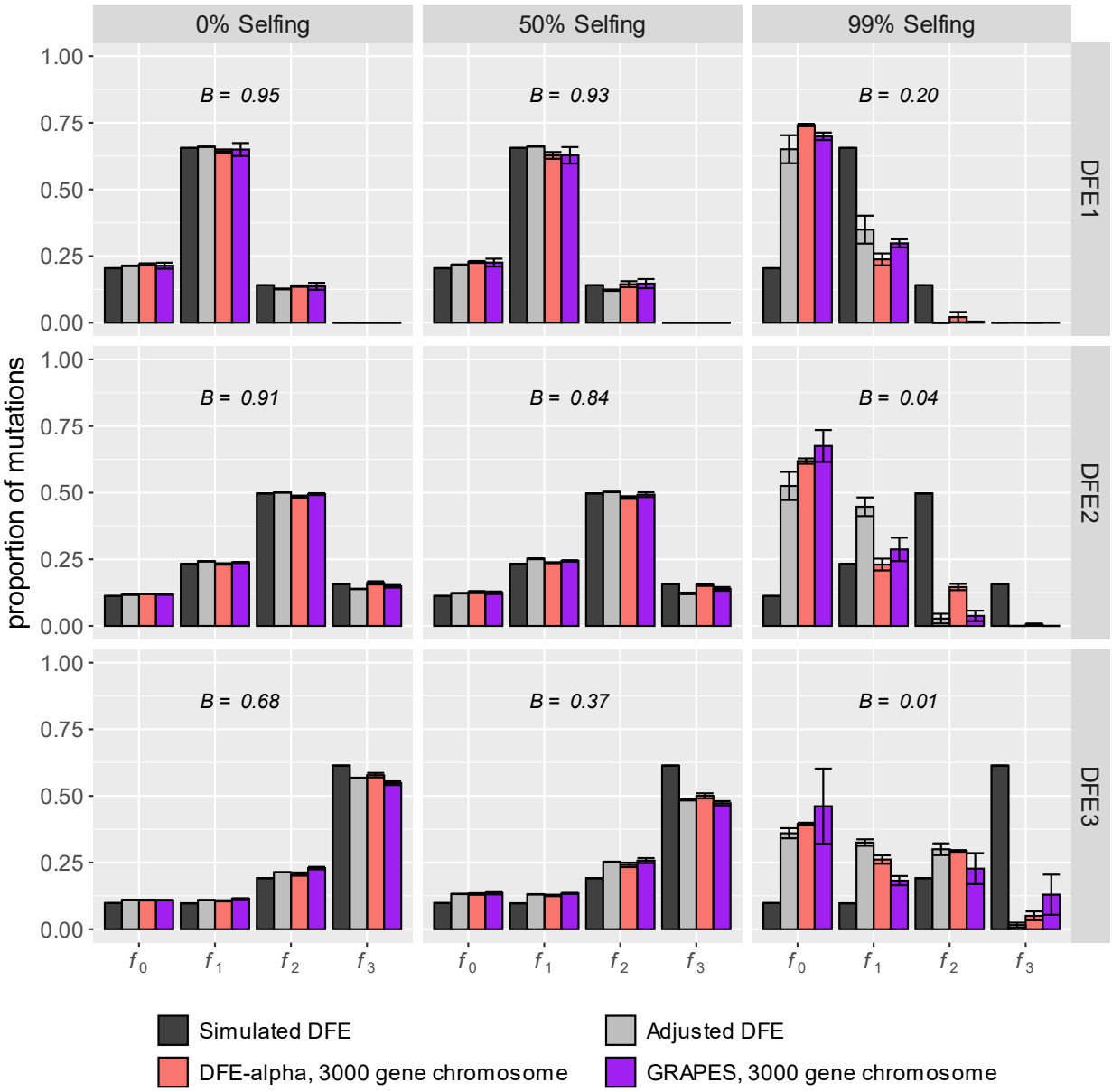
Supplementary Figure 3 Site frequency spectra of neutral and selected alleles in partially selfing populations from single-site simulations. 100 genomes were sampled and all selected mutations were additive and deleterious The *y*-axis represents the proportion of segregating polymorphisms that fall into the given allele-count class. The last class (11+) refers to the derived allele counts 11-100. The SFS of neutral sites is shown in red, while the SFS of selected sites in simulations with 0%, 50%, and 99% selfing are shown as green, blue, and purple bars respectively.



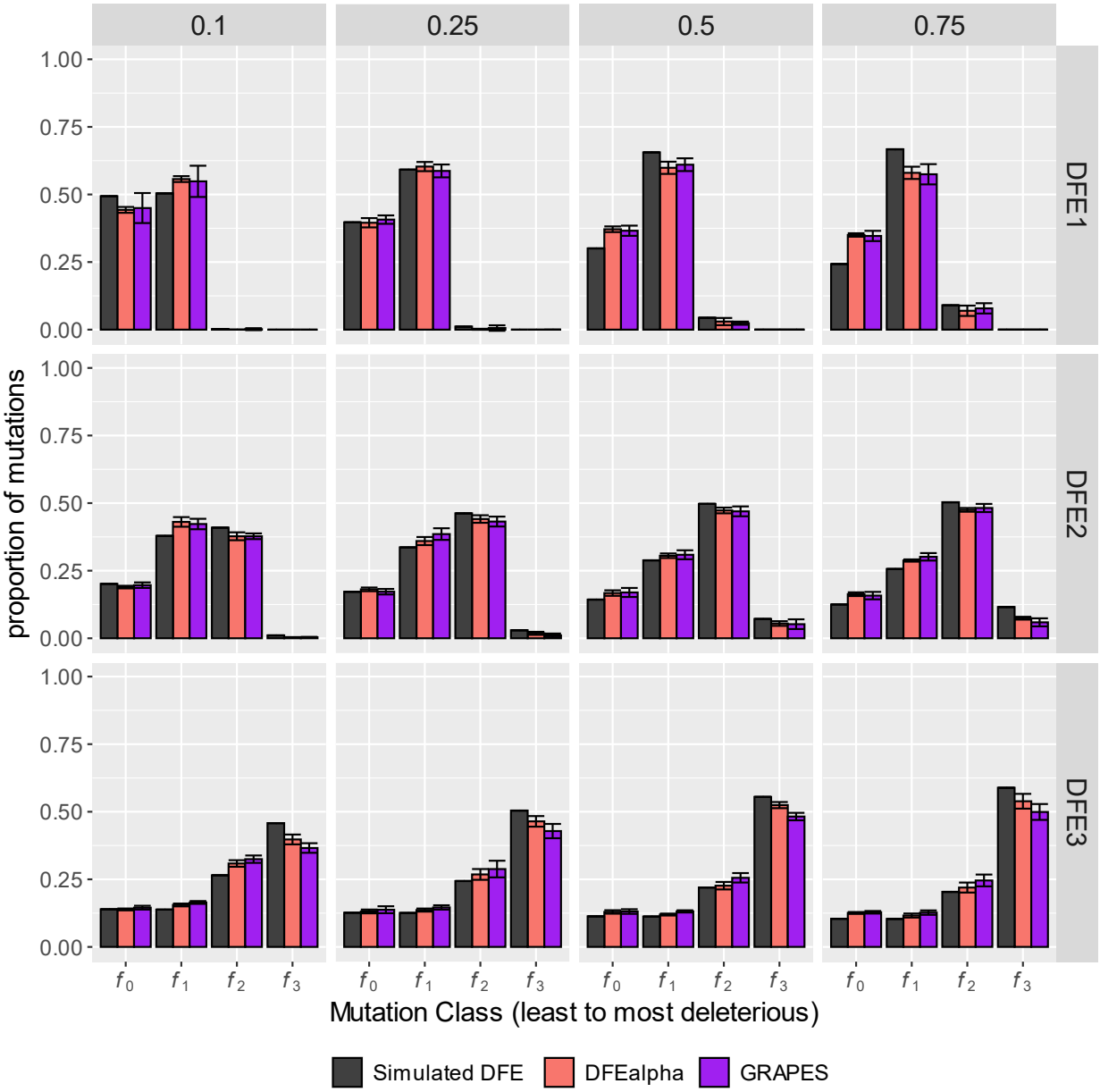
Supplementary Figure 4: (A) The SFS at neutral (blue) and selected (red) sites when 50 diploid individuals are sampled (instead of 100 random genomes). (B) The accuracy of DFE inference when 50 diploid individuals are sampled (instead of 100 random genomes). Results are shown for a predominantly moderately deleterious DFE (DFE2). Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



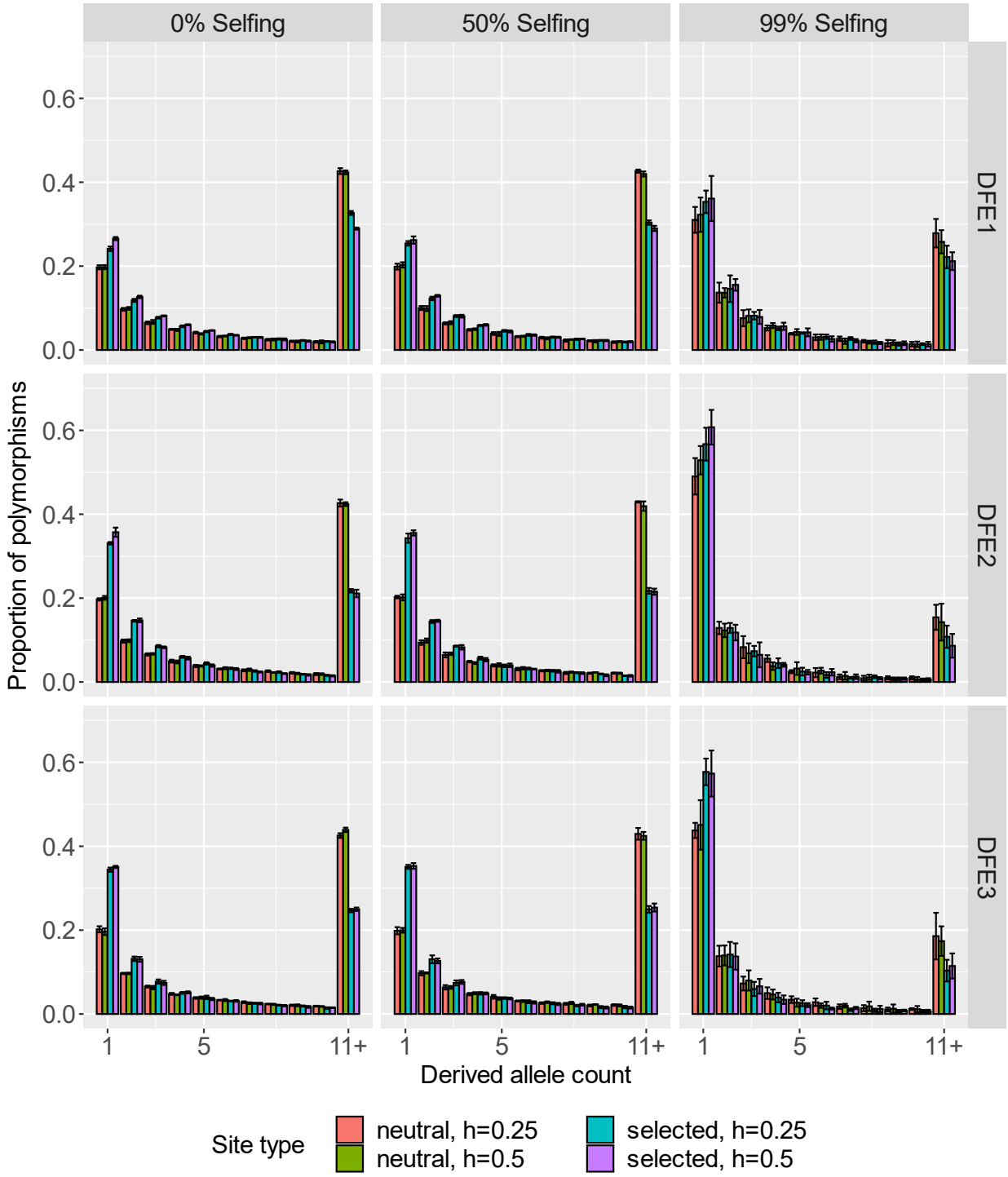
Supplementary Figure 5 Effects of varying the coding density on the estimated DFE of new deleterious mutations in partially selfing organisms using DFE-alpha and GRAPES. The intergenic regions were reduced from the original 3000 bp to 1500 bp and 500 bp (denoted on top of the columns). The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



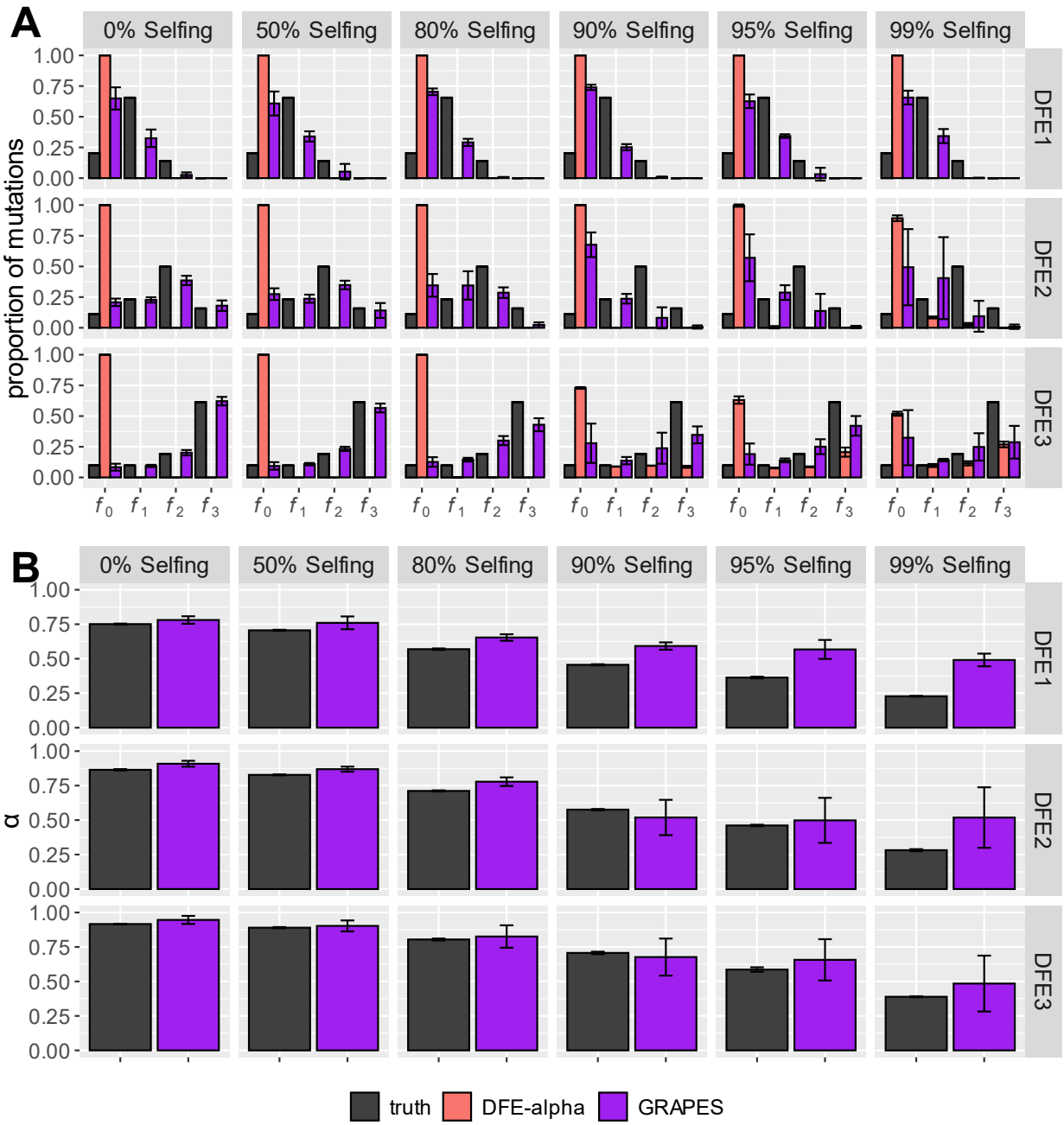
Supplementary Figure 6 The inference of the deleterious DFE when varying chromosome sizes were simulated: a chromosome with 500 genes (shown in darker shades) vs a chromosome with 3000 genes (shown in lighter shades). Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. The nucleotide site diversity with background selection (B) relative to its expectation under strict neutrality is shown in each panel. The adjusted DFE represents the distribution of $2NBs_d$. Inference was performed using DFE-alpha (red bars) and GRAPES purple bars). The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



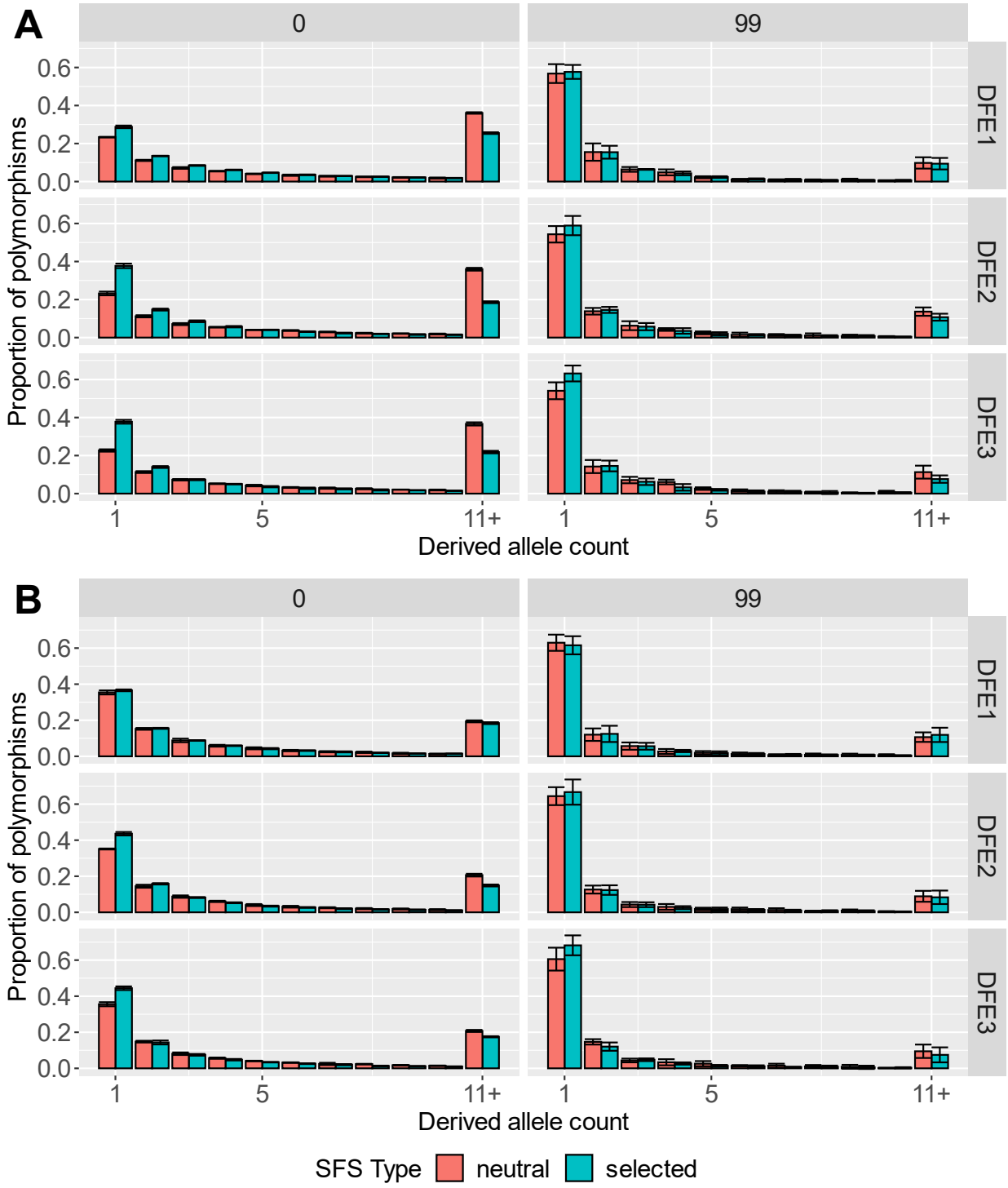
Supplementary Figure 7 Effects of selfing (50%) on the inference of the DFE of new deleterious mutations with varying dominance coefficients using DFE-alpha and GRAPES. The dominance of all new mutations is noted at the top of each column. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



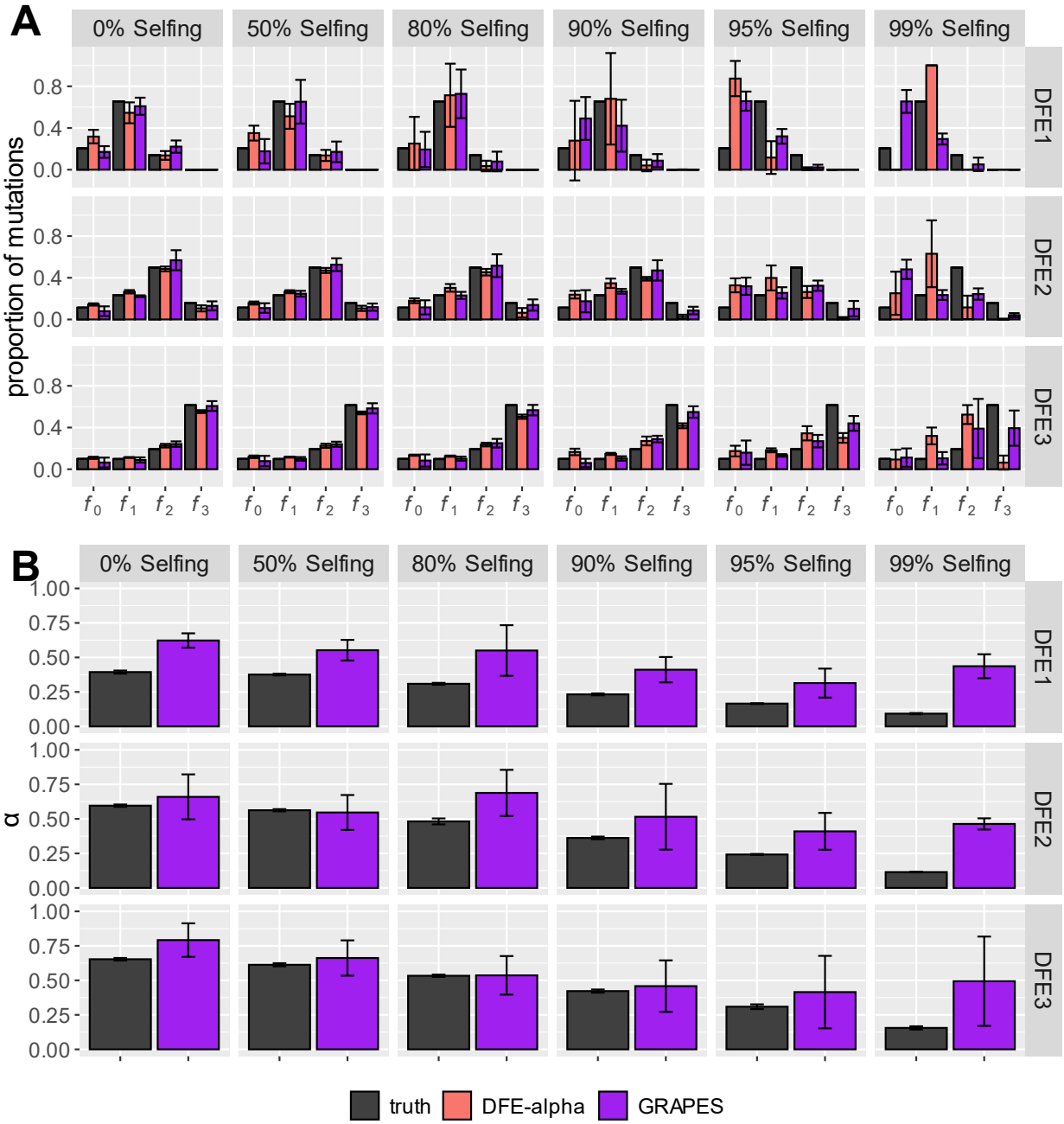
Supplementary Figure 8 Site frequency spectra of neutral and selected alleles for varying rates of selfing, dominance coefficients, and deleterious DFEs. 100 genomes were sampled from simulations. Selected mutations had a dominance coefficient $h = 0.25$ or 0.5.The y-axis represents the proportion of segregating polymorphisms that fall into the given allele-frequency class. The last class (11+) refers to the derived allele counts 11-100.



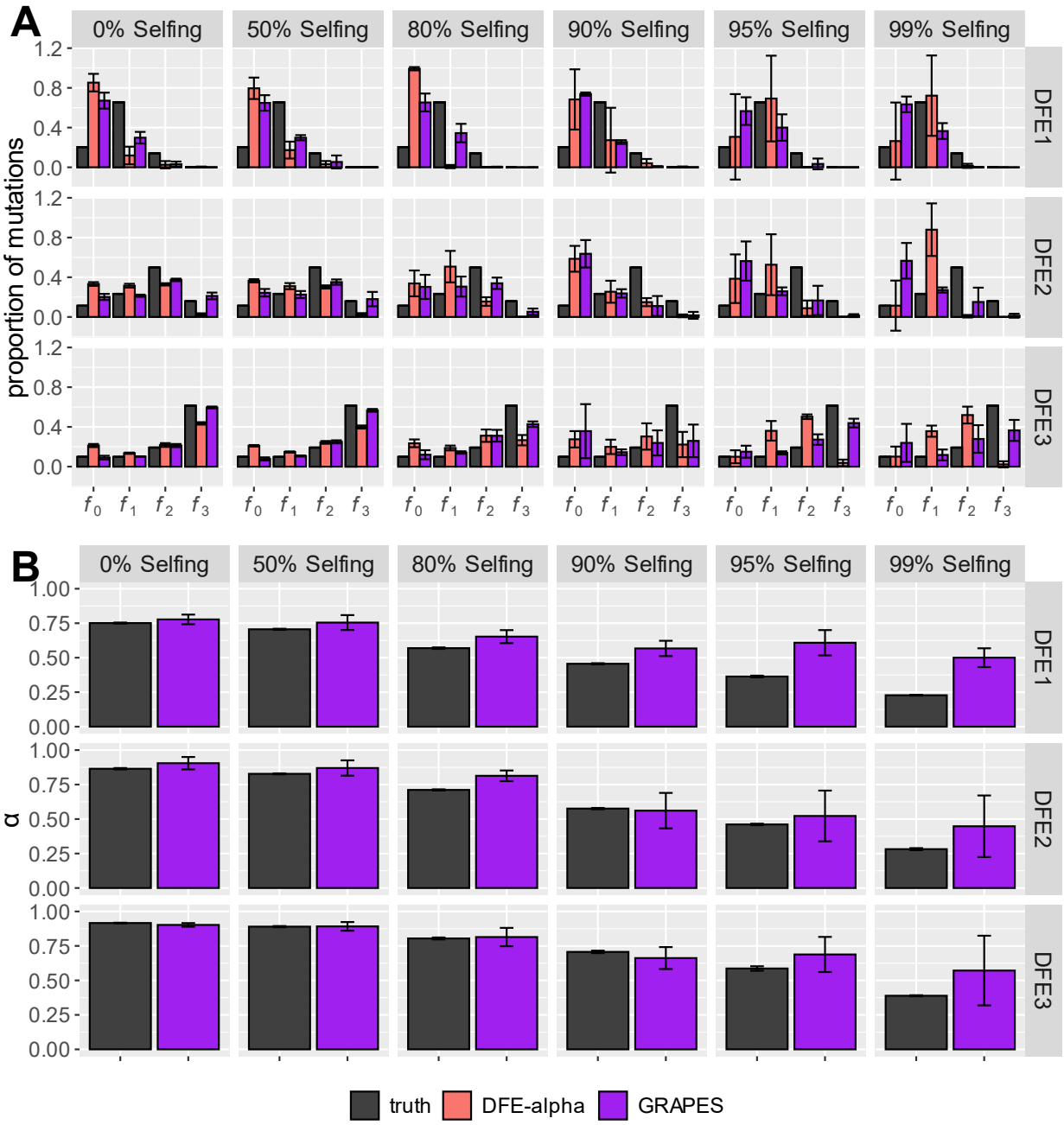
Supplementary Figure 9 Effects of selfing on the inference of (A) the DFE of new deleterious mutations and (B) the proportion of beneficial fixations (α) using DFE-alpha and GRAPES when beneficial mutations are common (1%). Results for A are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. Here, 1.0% of new exonic mutations were beneficial and exponentially distributed with a mean $2Ns_a = 200$. Here, the unfolded SFS was used for inference. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



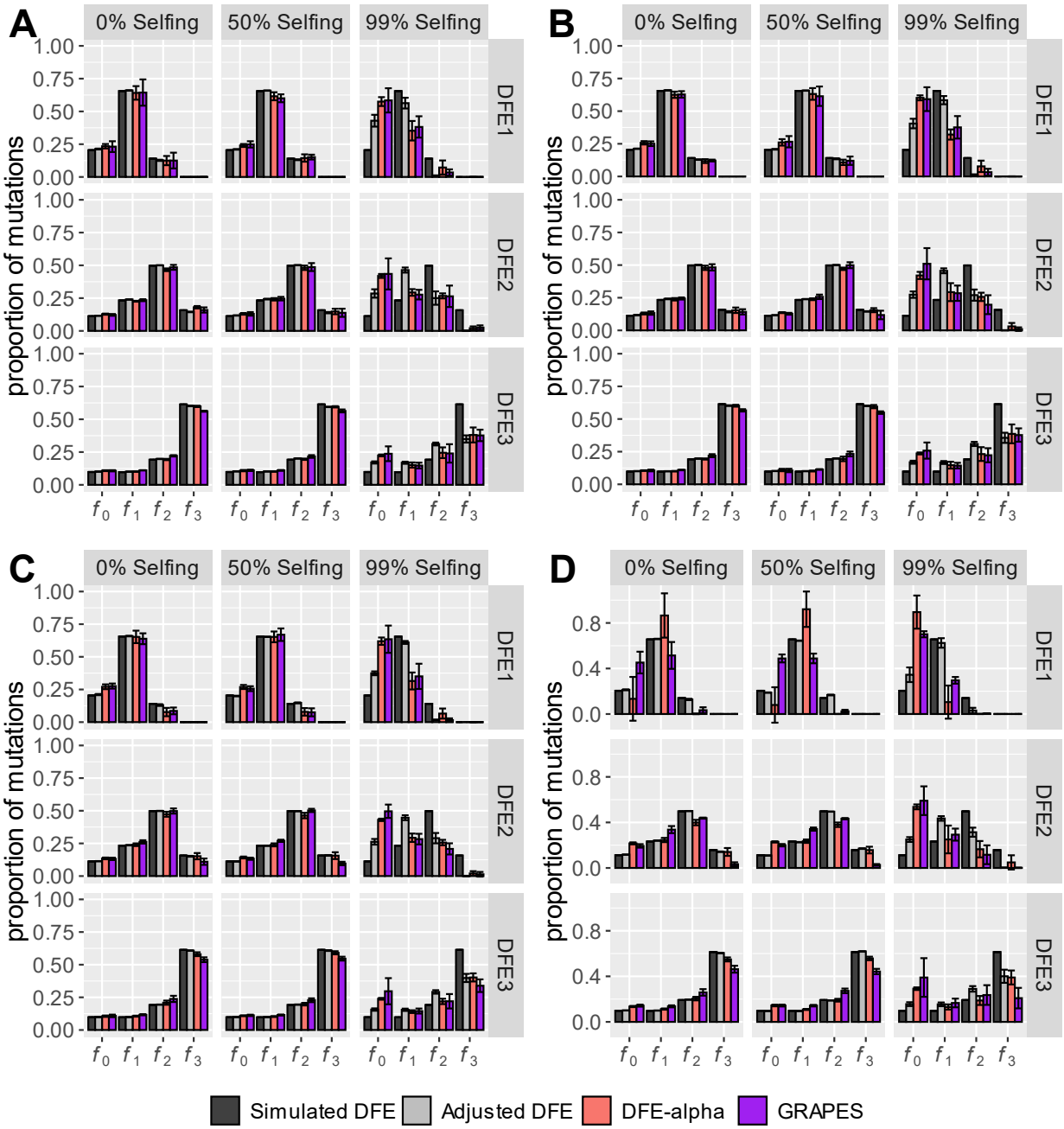
Supplementary Figure 10 Site frequency spectra of neutral vs selected sites when a full DFE was simulated (*i.e.*, both deleterious and beneficial mutations were present). 100 genomes were sampled and the y-axis depicts the proportion of segregating polymorphisms that fall into the given allele-count class. Panel **A** shows simulations where 0.1% of new mutations at selected sites were beneficial and Panel **B** shows simulations where 1% of new mutations at selected sites were beneficial. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



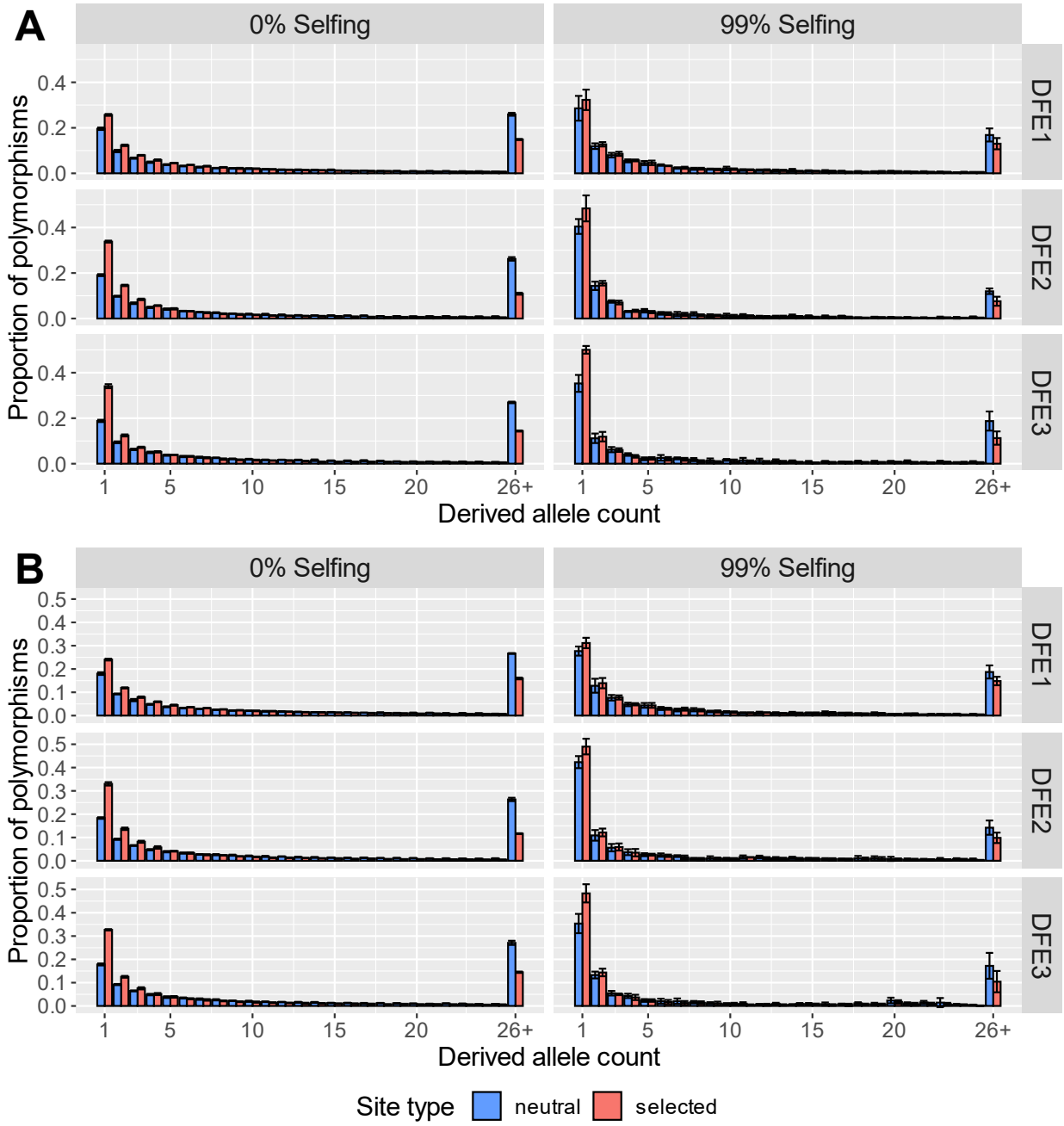
Supplementary Figure 11 Effects of selfing on the inference of (A) the DFE of new deleterious mutations and (B) the proportion of beneficial fixations (α) using DFE-alpha and GRAPES, when beneficial mutations were rare. Results for A are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. Here, 0.1% of new exonic mutations were beneficial and exponentially distributed with a mean $2Ns_a = 200$. Here, the folded SFS was used for inference. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



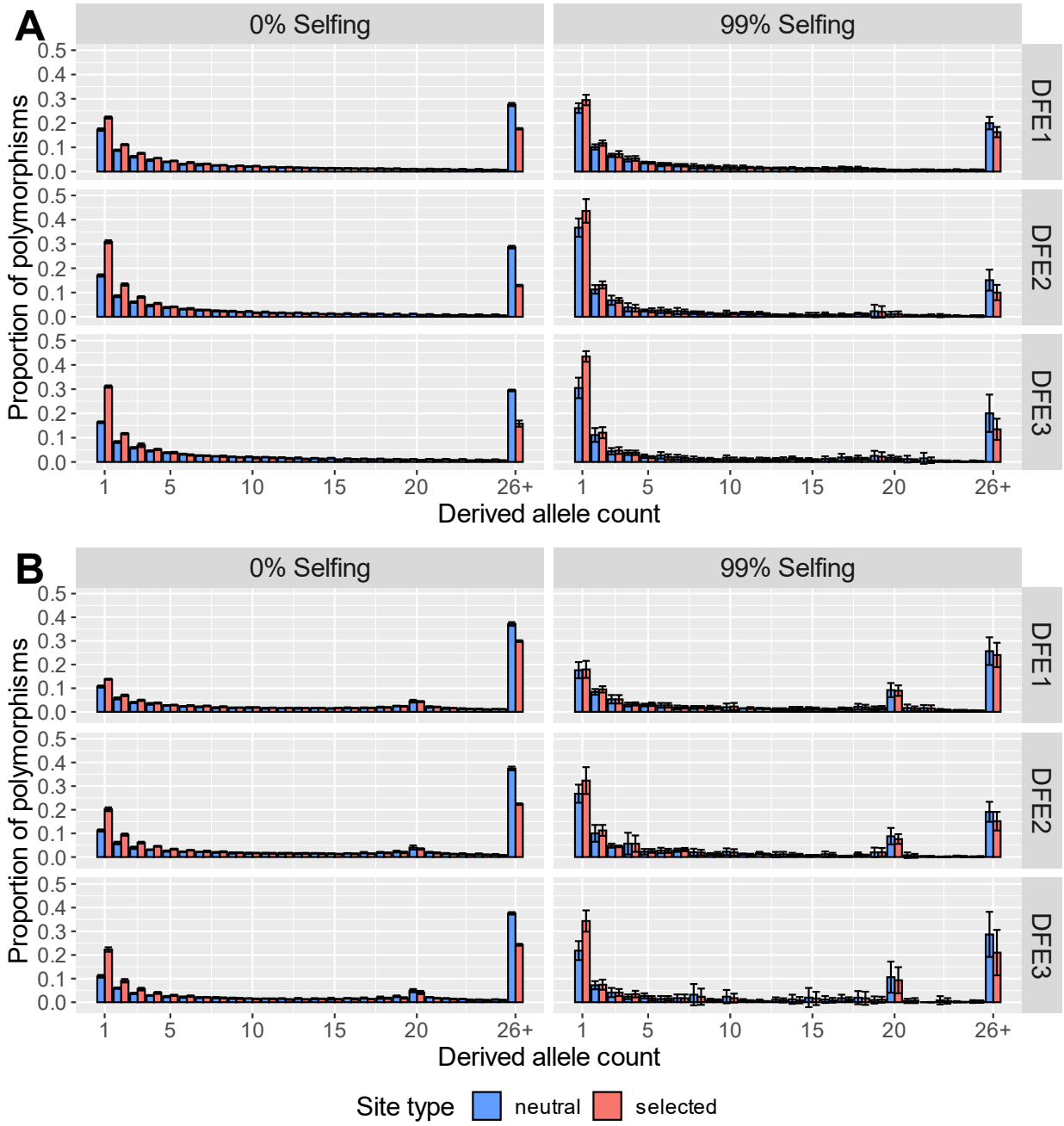
Supplementary Figure 12 Effects of selfing on the inference of (A) the DFE of new deleterious mutations and (B) the proportion of beneficial fixations (α) using DFE-alpha and GRAPES, when beneficial mutations were common. Results for A are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. Here, 1.0% of new exonic mutations were beneficial and exponentially distributed with a mean $2Ns_a = 200$. Here, the folded SFS was used for inference. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



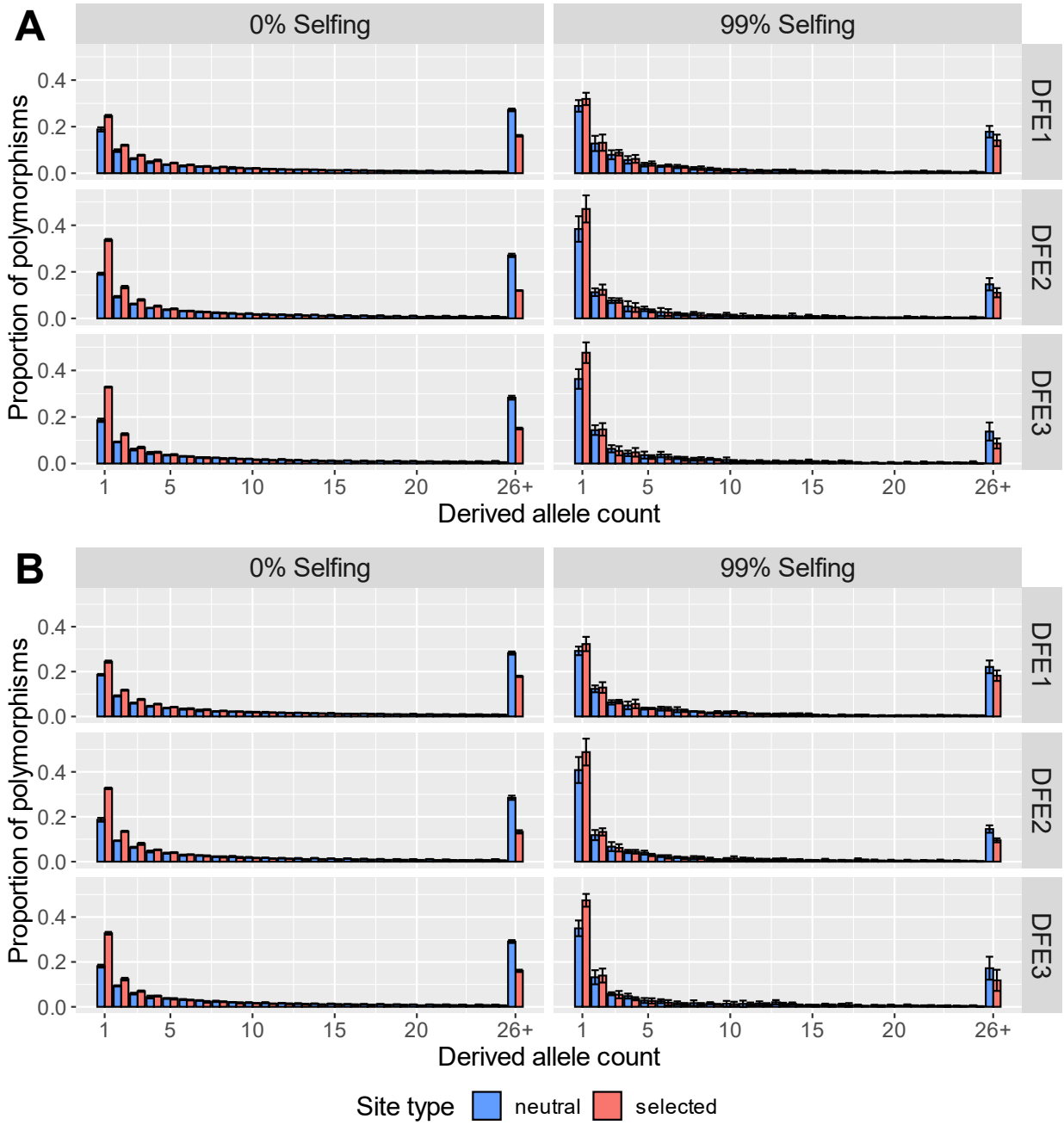
Supplementary Figure 13 Effects of population structure and selfing on the inference of the DFE of new deleterious mutations using DFE-alpha and GRAPES. Populations were simulated under an island model with five demes, and genomes were sampled evenly from each deme. Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was (**A**) 2 (**B**) 1, (**C**) 0.5, and (**D**) 0.1. Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. The nucleotide site diversity with background selection (B) relative to its expectation under strict neutrality is shown in each panel. The adjusted DFE represents the distribution of $2NBs_d$. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



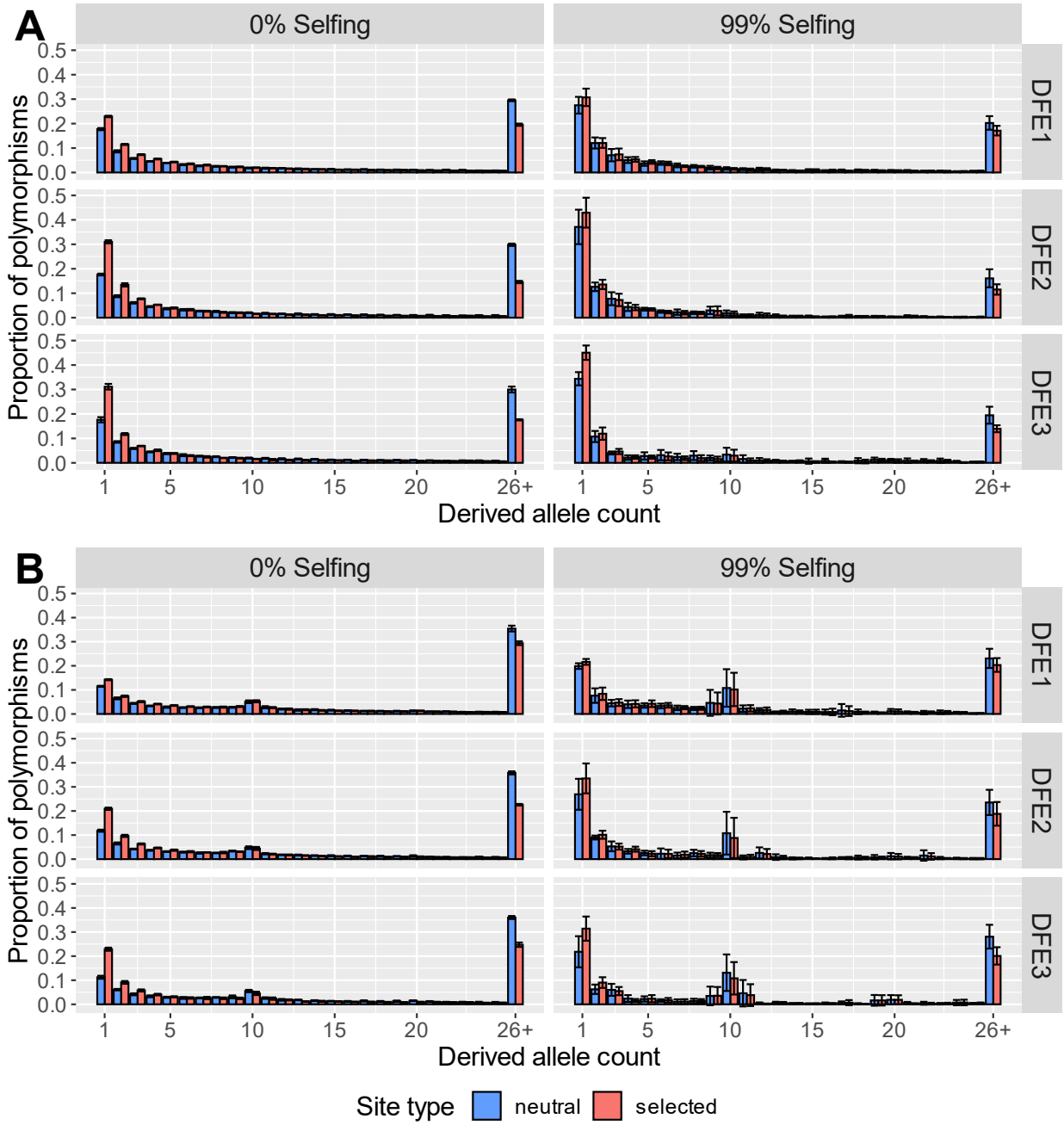
Supplementary Figure 14 Effects of population structure and selfing on the SFS. Populations were simulated under an island model with five demes, and genomes were sampled evenly from each deme. Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was **(A)** 2 **(B)** 1. The error bars denote the standard deviation of proportions estimated from 5 independent replicates. The last class (26+) refers to the derived allele counts 26-100.



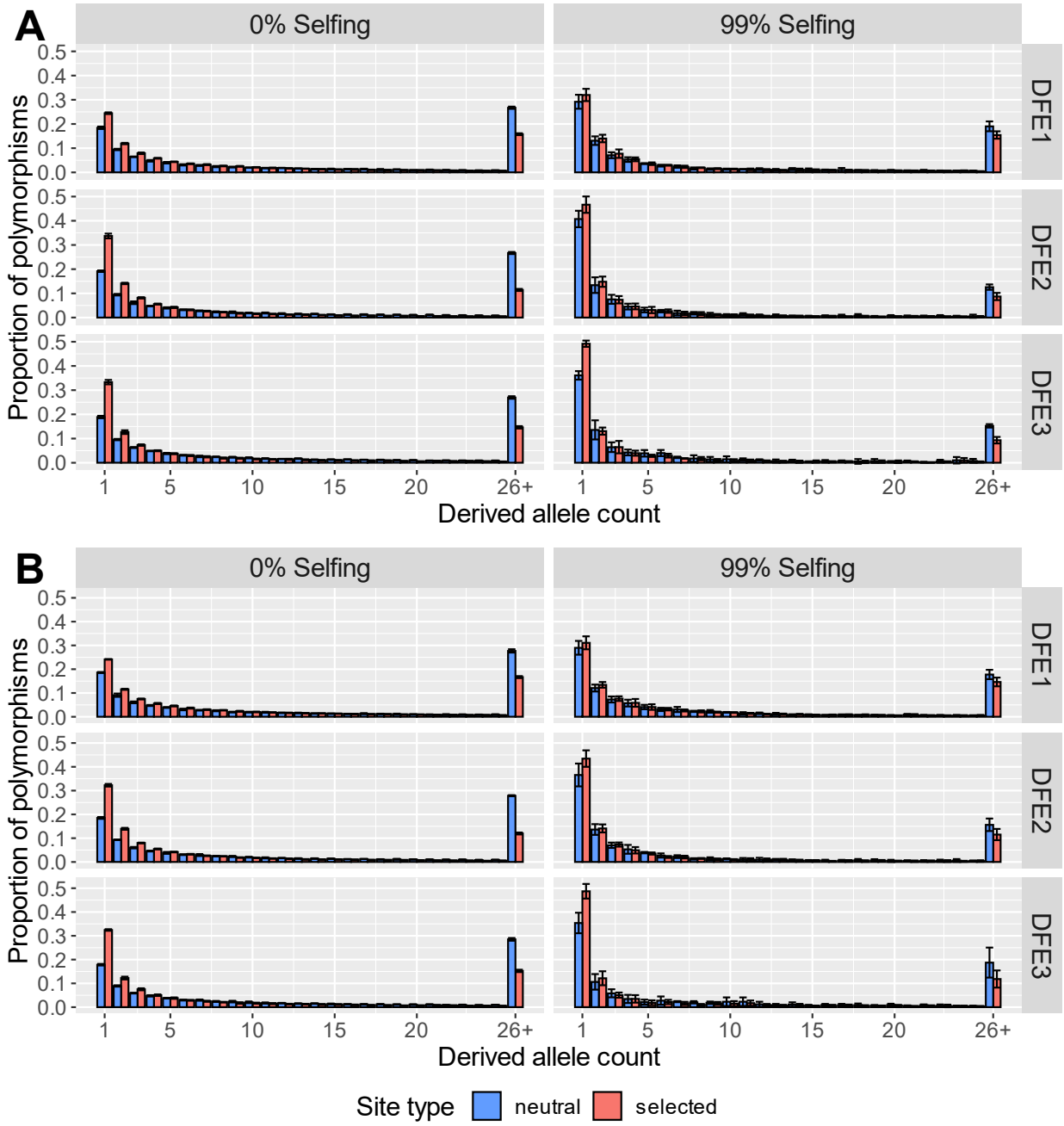
Supplementary Figure 15 Effects of population structure and selfing on the SFS with even sampling. Populations were simulated under an island model with five demes, and genomes were sampled evenly from each deme. Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was **(A)** 0.5 **(B)** 0.1. The error bars denote the standard deviation of proportions estimated from 5 independent replicates. The last class (26+) refers to the derived allele counts 26-100.



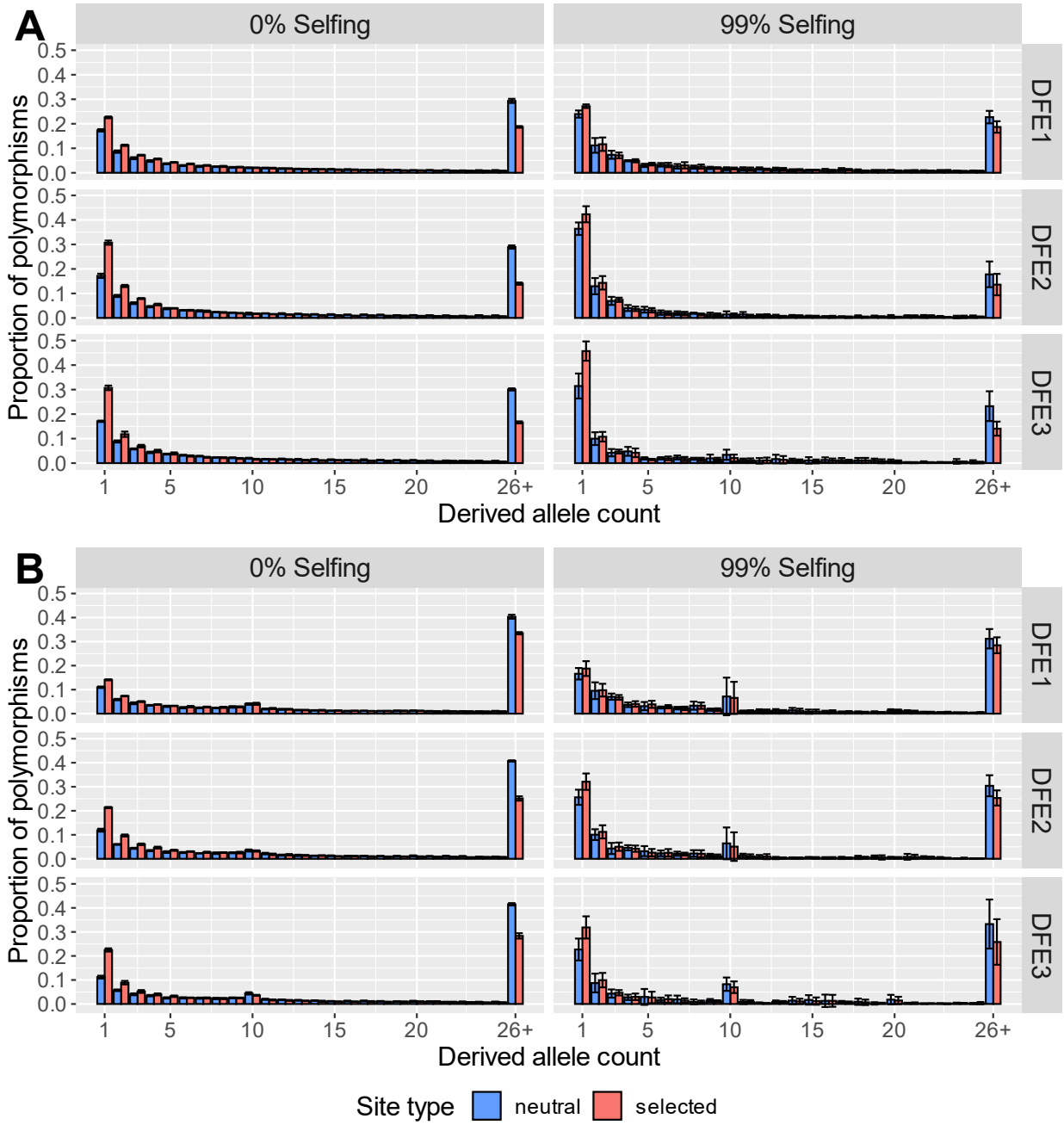
Supplementary Figure 16 Effects of uneven sampling, population structure, and selfing on the SFS at high migration rates. Populations were simulated under an island model with five demes, and genomes were sampled unevenly (60 genomes from one deme, and 10 from the remaining four demes). Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was (**A**) 2 (**B**) 1. The error bars denote the standard deviation of proportions estimated from 5 independent replicates. The last class (26+) refers to the derived allele counts 26-100.



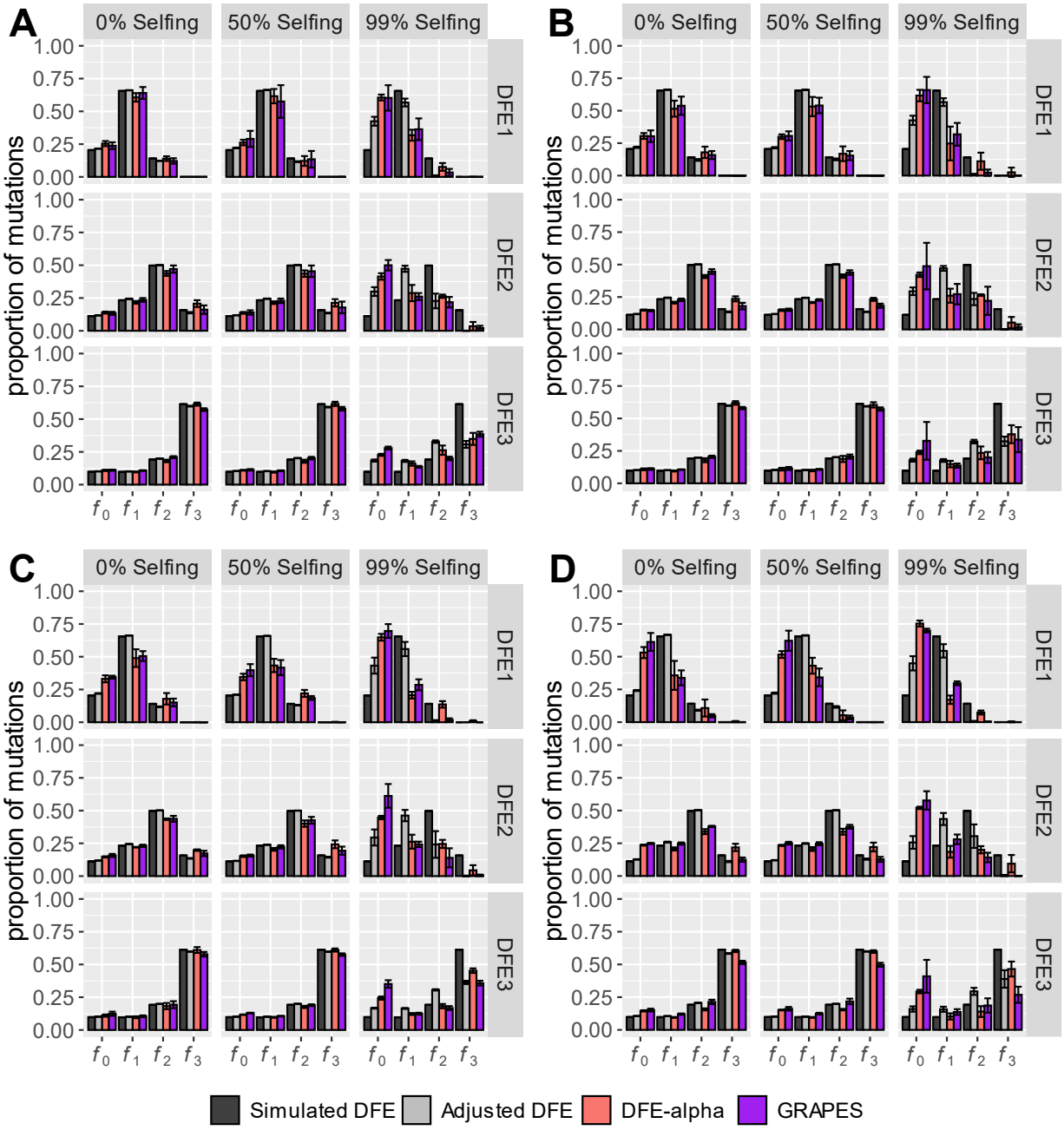
Supplementary Figure 17 Effects of uneven sampling, population structure, and selfing on the SFS at lower migration rates. Populations were simulated under an island model with five demes, and genomes were sampled unevenly (60 genomes from one deme, and 10 from the remaining four demes). Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was (**A**) 0.5 (**B**) 0.1. The error bars denote the standard deviation of proportions estimated from 5 independent replicates. The last class (26+) refers to the derived allele counts 26-100.



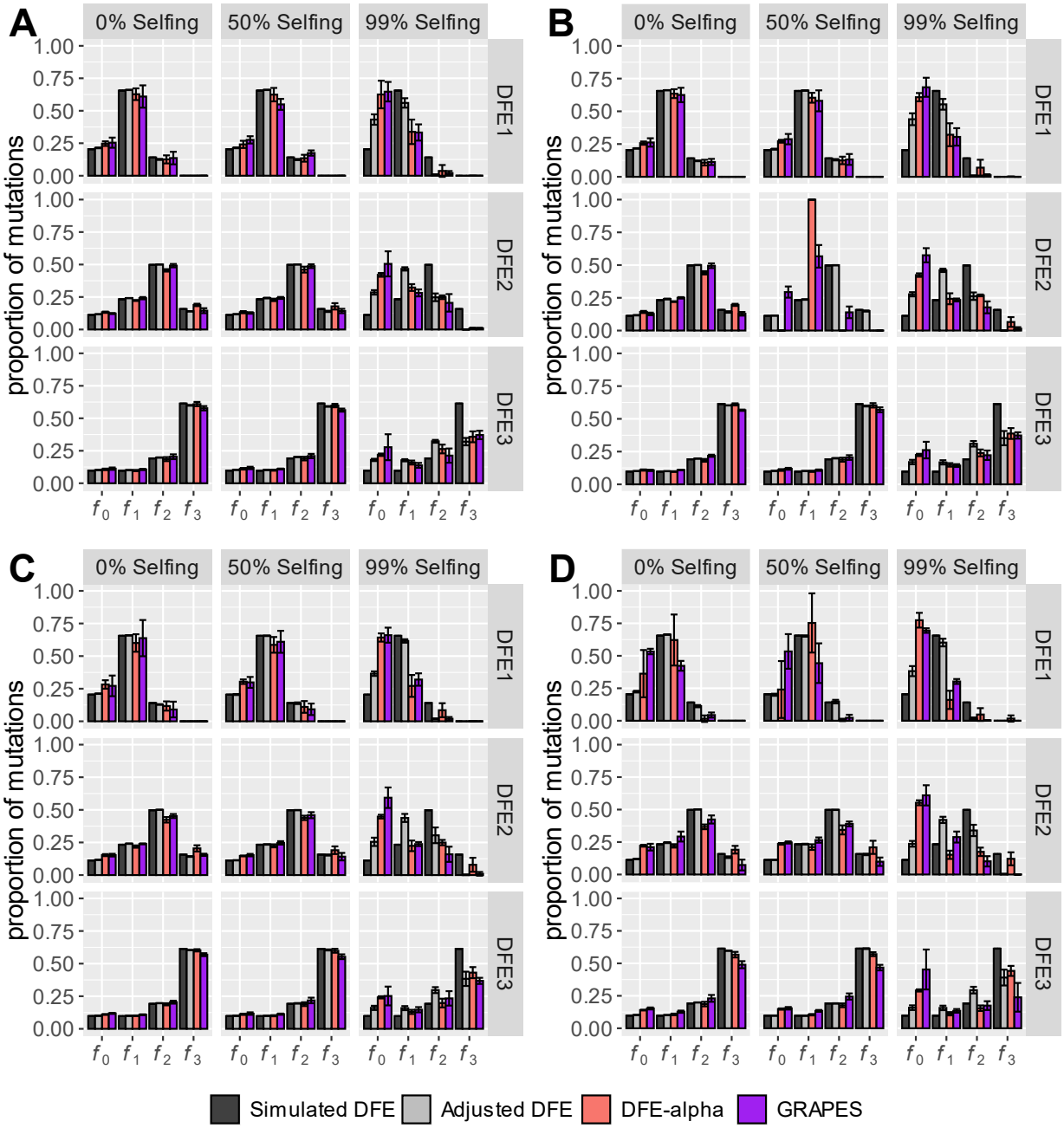
Supplementary Figure 18 Effects of uneven sampling, population structure, and selfing on the SFS at high migration rates. Populations were simulated under an island model with five demes, and genomes were sampled unevenly (35 genomes from two demes, and 10 from the remaining four demes). Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was (**A**) 2 (**B**) 1. The error bars denote the standard deviation of proportions estimated from 5 independent replicates. The last class (26+) refers to the derived allele counts 26-100.



Supplementary Figure 19 Effects of uneven sampling, population structure, and selfing on the SFS when rates of migration are low. Populations were simulated under an island model with five demes, and genomes were sampled unevenly (35 genomes from two demes, and 10 from the remaining four demes). Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was **(A)** 0.5 **(B)** 0.1. The error bars denote the standard deviation of proportions estimated from 5 independent replicates. The last class (26+) refers to the derived allele counts 26-100.



Supplementary Figure 20. Effects of uneven sampling, population structure, and selfing on the inference of the DFE of new deleterious mutations using DFE-alpha and GRAPES. Populations were simulated under an island model with five demes, and genomes were sampled unevenly (60 genomes from one deme, and 10 from the remaining four demes). Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was **(A)** 2 **(B)** 1, **(C)** 0.5, and **(D)** 0.1. Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. The nucleotide site diversity with background selection (B) relative to its expectation under strict neutrality is shown in each panel. The adjusted DFE represents the distribution of $2NBs_d$. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.



Supplementary Figure 21. Effects of uneven sampling, population structure and selfing on the inference of the DFE of new deleterious mutations using DFE-alpha and GRAPES. Populations were simulated under an island model with five demes, and genomes were sampled unevenly (35 genomes from two demes, and 10 from the remaining three demes). Here the metapopulation effective size was equal to 5000 at 0% selfing (see methods). Results are shown when $N_{deme}m$ was **(A)** 2 **(B)** 1, **(C)** 0.5, and **(D)** 0.1. Results are shown in terms of the proportion of mutations in effectively neutral (f_0), weakly deleterious (f_1), moderately deleterious (f_2) and strongly deleterious (f_3) classes. The nucleotide site diversity with background selection (B) relative to its expectation under strict neutrality is shown in each panel. The adjusted DFE represents the distribution of $2NBs_d$. The error bars denote the standard deviation of proportions estimated from 5 independent replicates.

Supplementary Table 1 Effects of selfing on the mean values of parameters of the DFE of new deleterious mutations, estimated by DFE-alpha. Here Es represents the mean efficacy of selection, before being scaled by a 2-epoch model of population size change, where $N1$ represents the ancestral population size, $N2$ represents the relative population size after the first epoch, and $t2$ represents the duration of time after the population size change (in generations).

	Selfing rate	Es	$N1$	$N2$	$t2$
DFE1	Truth	-0.025	100	NA	NA
	0%	-0.027	100	145	16
	50%	-0.025	100	153	7
	80%	-0.021	100	162	5
	90%	-0.023	100	150	17
	95%	-0.019	100	179	29
	99%	-0.013	100	394	77
DFE2	Truth	-0.25	100	100	NA
	0%	-0.269	100	158	5
	50%	-0.263	100	150	6
	80%	-0.223	100	168	5
	90%	-0.187	100	238	5
	95%	-0.099	100	322	22
	99%	-0.026	100	1000	57
DFE3	Truth	-5	100	100	NA
	0%	-4.712	100	144	5
	50%	-5.465	100	156	5
	80%	-4.62	100	162	5
	90%	-2.665	100	198	11
	95%	-1.722	100	319	17
	99%	-0.687	100	798	59

Supplementary Table 2 Estimated parameter values of the deleterious DFE by GRAPES and DFE-alpha for simulations with only semi-dominant deleterious mutations.

		β				$2N_e s_d$			
		GRAPES		DFE-alpha		GRAPES		DFE-alpha	
	Selfing rate	mean	SD	mean	SD	mean	SD	mean	SD
DFE1	Truth	0.9	0	0.9	0	5	0	5	0
	0%	0.77	0.16	0.74	0.04	5.43	0.58	5.4	0.33
	50%	0.85	0.09	0.82	0.16	4.97	0.15	5.12	0.5
	80%	0.93	0.30	1	0.29	4.5	0.81	4.27	0.59
	90%	0.65	0.11	0.78	0.16	5.25	1.07	4.8	0.92
	95%	0.57	0.15	0.68	0.14	4.42	0.39	4	0.48
	99%	0.34	0.03	0.26	0.03	2.34	0.59	3.2	0.9
DFE2	Truth	0.5	0	0.5	0	50	0	50	0
	0%	0.50	0.01	0.48	0.01	50.6	4.12	54.23	5.35
	50%	0.48	0.04	0.47	0.02	52	6.64	53.12	3.45
	80%	0.49	0.06	0.48	0.03	44.4	7.33	45.14	3.73
	90%	0.45	0.05	0.47	0.03	41.87	10.37	37.87	6.4
	95%	0.39	0.07	0.46	0.07	29.44	9.1	21.12	5.89
	99%	0.33	0.15	0.35	0.04	8.93	1.38	6.39	1.41
DFE3	Truth	0.3	0	0.3	0	1000	0	1000	0
	0%	0.32	0.02	0.3	0.01	627.15	107.89	949.56	138.64
	50%	0.30	0.02	0.29	0.02	730.72	133.39	1102.76	243.83
	80%	0.30	0.05	0.29	0.03	743.11	415.73	933.55	448.17
	90%	0.28	0.04	0.29	0.01	663.41	468.89	546.89	156.22
	95%	0.26	0.04	0.28	0.01	512.23	233.45	363.49	62.58
	99%	0.17	0.06	0.26	0.03	619.02	344.14	171.07	99.41

Supplementary Table 3 The number of segregating sites and reduction in population size due to BGS and selfing for varying selfing rates and DFEs as observed in our simulations. Here N=5000 and $B = \pi_{neu}/(4N_{self}\mu)$, where μ is the mutation rate per site/generation.

	Selfing rate	Number of deleterious segregating sites		π_{neu}	N_{emp}	N_{self}	B
		mean	SD				
DFE1	0%	13564	137	0.00626	4746	5000	0.949
	50%	10077	121	0.00464	3512	3750	0.936
	80%	7950	112	0.00352	2664	3000	0.888
	90%	6997	131	0.00299	2264	2750	0.823
	95%	6260	190	0.00244	1850	2625	0.705
	99%	4281	293	0.0011	835	2525	0.331
DFE2	0%	7930	103	0.00614	4654	5000	0.931
	50%	5837	105	0.00448	3392	3750	0.905
	80%	4463	95	0.00322	2440	3000	0.813
	90%	3559	137	0.00225	1706	2750	0.620
	95%	2656	244	0.00114	865	2625	0.329
	99%	1628	135	0.00028	212	2525	0.084
DFE3	0%	4273	49	0.0059	4470	5000	0.894
	50%	3003	36	0.00406	3078	3750	0.821
	80%	2048	68	0.00257	1946	3000	0.649
	90%	1602	25	0.00165	1254	2750	0.456
	95%	1213	54	0.00097	737	2625	0.281
	99%	818	69	0.00032	241	2525	0.096

Supplementary Table 4 Nucleotide site diversity (π) from single-site simulations of neutral and semidominant deleterious mutations in populations with varying selfing rates.

	Selfing rate	Theoretical π at neutral sites	Simulation π at neutral sites
DFE1	0	0.00660	0.00653
DFE2	0	0.00660	0.00653
DFE3	0	0.00660	0.00653
DFE1	0.50	0.00495	0.00499
DFE2	0.50	0.00495	0.00499
DFE3	0.50	0.00495	0.00499
DFE1	0.99	0.00333	0.00321
DFE2	0.99	0.00333	0.00321
DFE3	0.99	0.00333	0.00321

Supplementary Table 5 Estimates of the parameters of the deleterious DFE (a gamma distribution with mean $\gamma = 2N_e\overline{s_d}$ and shape parameter β) inferred by DFE-alpha and GRAPES. Since DFE-alpha and GRAPES assume semidominance, the mean strength of selection γ was multiplied by 0.5 to be equivalent to $2N_{self}s_dh_{self}$.

DFE	Selfing rate	$2N_{self}s_dh_{self}$ (expected)	DFE-alpha γh	GRAPES γh	β (expected)	DFE-alpha β	GRAPES β	h
DFE1	0	3.75	4.29	4.33	0.9	0.62	0.64	0.75
DFE1	0.50	4.04	3.32	3.49	0.9	0.68	0.66	0.75
DFE1	0.99	4.17	1.40	1.01	0.9	0.42	0.34	0.75
DFE2	0	37.5	41.45	34.99	0.5	0.44	0.48	0.75
DFE2	0.50	40.38	30.68	27.77	0.5	0.46	0.49	0.75
DFE2	0.99	41.67	4.23	4.54	0.5	0.30	0.19	0.75
DFE3	0	750.00	791.63	493.84	0.3	0.29	0.31	0.75
DFE3	0.50	807.69	630.82	429.23	0.3	0.28	0.30	0.75
DFE3	0.99	833.32	50.93	222.49	0.3	0.29	0.18	0.75
DFE1	0	2.50	2.70	2.71	0.9	0.74	0.77	0.50
DFE1	0.50	3.08	2.56	2.48	0.9	0.82	0.85	0.50
DFE1	0.99	3.33	1.60	1.17	0.9	0.26	0.34	0.50
DFE2	0	25.00	27.11	25.30	0.5	0.48	0.49	0.50
DFE2	0.50	30.77	26.56	26.00	0.5	0.47	0.48	0.50
DFE2	0.99	33.33	3.19	4.46	0.5	0.35	0.33	0.50
DFE3	0	500.00	474.78	313.57	0.3	0.30	0.32	0.50
DFE3	0.50	615.38	551.38	365.36	0.3	0.29	0.30	0.50
DFE3	0.99	666.64	85.53	309.51	0.3	0.26	0.17	0.50
DFE1	0	1.25	1.28	1.33	0.9	2.66	2.07	0.25
DFE1	0.50	2.12	1.66	1.74	0.9	1.54	1.49	0.25
DFE1	0.99	2.50	1.56	0.98	0.9	0.32	0.39	0.25
DFE2	0	12.50	10.73	10.62	0.5	0.63	0.64	0.25
DFE2	0.50	21.15	18.06	15.95	0.5	0.51	0.56	0.25
DFE2	0.99	25.00	4.50	4.65	0.5	0.29	0.20	0.25
DFE3	0	250.00	263.78	217.20	0.3	0.31	0.31	0.25
DFE3	0.50	423.08	316.34	253.83	0.3	0.31	0.32	0.25
DFE3	0.99	499.97	54.93	183.24	0.3	0.28	0.25	0.25
DFE1	0	0.50	0.91	0.88	0.9	100	16.45	0.10
DFE1	0.50	1.54	1.24	1.31	0.9	3.44	9.15	0.10
DFE1	0.99	2.00	1.32	0.84	0.9	0.40	0.36	0.10
DFE2	0	5.00	4.54	4.92	0.5	1.05	0.94	0.10
DFE2	0.50	15.38	12.19	12.42	0.5	0.58	0.56	0.10
DFE2	0.99	20.00	5.02	6.41	0.5	0.32	0.22	0.10
DFE3	0	100.00	69.53	63.27	0.3	0.38	0.39	0.10
DFE3	0.50	307.69	192.87	159.62	0.3	0.33	0.33	0.10
DFE3	0.99	399.96	72.07	213.35	0.3	0.26	0.19	0.10

Supplementary Table 6 Effects of selfing on the number of fixations of deleterious and beneficial mutations. Here, deleterious mutations were drawn from the DFEs defined in Table 1, while 0.1% of new exonic mutations were beneficial and exponentially distributed with a mean $2Ns_a = 200$.

Selfing rate	DFE	Deleterious substitutions avg.	Deleterious substitutions SD	Beneficial substitutions avg.	Beneficial substitutions SD	α avg.	α SD
0	DFE1	1438.4	31.48	935.8	32.09	0.39	0.01
50	DFE1	1514.2	56.92	912.2	30.19	0.38	0.01
80	DFE1	1820.8	28.03	813.4	17.9	0.31	0.01
90	DFE1	2253.6	68.2	681.8	14.75	0.23	0.01
95	DFE1	2785	77.11	552.4	19.55	0.17	0
99	DFE1	3363.8	56.91	343.4	15.53	0.09	0
0	DFE2	618.2	7.82	912	40.21	0.6	0.01
50	DFE2	682	28.57	878.2	33.25	0.56	0.01
80	DFE2	810	46.98	753.4	33.07	0.48	0.02
90	DFE2	1087.4	31.5	617.6	20.02	0.36	0.01
95	DFE2	1451.2	28.01	464	5.7	0.24	0
99	DFE2	2133	26.54	276.6	7.83	0.11	0
0	DFE3	460.8	11.76	869.2	24.38	0.65	0.01
50	DFE3	509	16.39	806.6	25.75	0.61	0.01
80	DFE3	543.2	27.22	620.6	18.99	0.53	0.01
90	DFE3	635.6	34.25	464.8	25.96	0.42	0.01
95	DFE3	762.2	24.94	341.4	20.95	0.31	0.02
99	DFE3	1040.8	36.62	191.6	18.77	0.16	0.01

Supplementary Table 7 Parameters used for five-deme island model simulations.

$N_{deme}m$	m	N_{deme}
10	0.0102	984
2	0.00216	926
1	0.00116	862
0.5	0.00066	785
0.1	0.00026	385

Supplementary Table 8 Table showing how theoretical expectations for T_T, T_S, and F_{st} match values estimated from neutral simulations.

$N_{deme}m$	Selfing rate	π_{meta} expected	π_{meta} observed	π_{deme} expected	π_{deme} observed	F_{st} expected	F_{st} observed
0.1	0	0.00660	0.00659	0.00254	0.00259	0.61515	0.60713
0.1	0.50	0.00597	0.00599	0.00191	0.00211	0.68063	0.64825
0.1	0.99	0.00534	0.00567	0.00128	0.00140	0.75991	0.76400
0.5	0	0.00678	0.00661	0.00518	0.00507	0.23595	0.23265
0.5	0.50	0.00549	0.00537	0.00389	0.00375	0.29166	0.30229
0.5	0.99	0.00422	0.00425	0.00262	0.00291	0.37947	0.31402
1.0	0	0.00660	0.00659	0.00569	0.00564	0.13794	0.14426
1.0	0.50	0.00518	0.00519	0.00427	0.00434	0.17584	0.16317
1.0	0.99	0.00378	0.00402	0.00287	0.00305	0.24062	0.24016
2.0	0	0.00660	0.00657	0.00611	0.00611	0.07407	0.06951
2.0	0.50	0.00507	0.00508	0.00458	0.00457	0.09638	0.09950
2.0	0.99	0.00358	0.00345	0.00309	0.00295	0.13674	0.14377