The role of pseudo-overdominance in maintaining inbreeding depression

Diala Abu Awad[*1,‡] and Donald Waller[†2]

[1]Professorship for Population Genetics, Department of Life Science Systems, Technical University of Munich, Germany

[2]Department of Botany, University of Wisconsin, Madison, WI USA

[‡]Current affiliation: Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE—Le Moulon, Gif-sur-Yvette, France

1   **Abstract**

2   Classical models ~~ignoring~~ that ignore linkage predict that deleterious recessive mu-

3   tations should purge or fix within inbred populations, yet ~~these~~ inbred populations of-

4   ten retain moderate to high segregating load. True overdominance ~~generates balancing~~

5   ~~selection that sustains~~ could generate balancing selection strong enough to sustain in-

6   breeding depression even ~~in inbred populationsbut is rare. In contrast~~ within inbred

7   populations, but this is considered rare. However, arrays of ~~mildly~~ deleterious recessives

8   linked in repulsion ~~may occur commonly enough to generate~~ could generate appreciable

9   pseudo-overdominance ~~and~~ that would also sustain segregating load. We used sim-

10   ulations to explore how long pseudo-overdominant ~~regions (POD's) persist following~~

11   ~~their creation via~~ (POD) zones persist once created (e.g., by hybridization between

12   populations fixed for alternative ~~mutations at linked loci. Balancing~~ mildly deleterious

[*]diala.abu-awad@universite-paris-saclay.fr

[†]dmwaller@wisc.edu

mutations). Balanced haplotype loads, tight linkage, and moderate to strong cumulative selective effects ~~all~~ all serve to maintain POD ~~'s~~zones. Tight linkage is key, suggesting that ~~POD's may most often~~ such regions are most likely to arise and persist in low recombination regions (~~e.g.,~~ like inversions). Selection and drift unbalance the load, eventually eliminating POD ~~'s~~zones, but this process is ~~very slow when~~ quite slow under strong pseudo-overdominance~~is strong~~. Background selection ~~across the genome~~ accelerates the loss of weak POD ~~'s~~ zones but reinforces strong ~~POD's~~ ones in inbred populations by disfavoring homozygotes. ~~Further modeling and~~ Models and empirical studies of POD dynamics within populations ~~could~~ help us understand how POD ~~'s affect persistence of the load and how inbred mating systems evolve.~~ zones may allow the load to persist, greatly affecting load dynamics and mating systems evolution.

Keywords: Inbreeding; purging; fixation; drift load; pseudo-overdominance; associative overdominance; POD; linkage; recombination.

1 Introduction

~~Mechanisms exist to ensure outcrossing in most species indicating that selection consistently acts against inbred progeny (Darwin, 1876). Nevertheless, numerous taxa inbreed either inadvertently (e.g., from local dispersal and/or mating) or via evolved adaptations, e.g., facultative self-fertilization in some snails and the self-fertilizing flowers found in many Angiosperms. Since Darwin's seminal work, inbreeding depression (ID) has been accepted as the primary selective force favoring outcrossing. Inbreeding greatly depresses fitness in most higher plants and animals (Keller and Waller, 2002) including many inbred populations (Byers and Waller, 1999). Although dominance, overdominance, and certain forms of epistasis can generate ID, evolutionists generally accept that true overdominance is rare and that most ID derives from dominance, i.e., the cumulative action of~~ Inbreeding depression (δ) is defined as the lower fitness of inbred compared to outbred individuals (Darwin, 1876). It is now generally accepted that δ is mainly due to the expression of segregating deleterious recessive mutations ~~acting across many loci (Charlesworth and Charlesworth, 1987; Crow, 1993, 1999a).~~ (Charlesworth and Charlesworth, 1987; Crow, 1993, 1 As direct selection, background selection, genetic drift and inbreeding all act to reduce diversity at such loci, maintaining non-negligible levels of inbreeding depression is difficult to explain (Byers and Waller, 1999; Winn et al, 2011). Examples include inbred lines of *Zea mays* Kardos et al (2014); Larièpe et al (2012), *Arabidopsis* (Seymour et al, 2016), *Mimulus* (Brown and Kelly, 2020) and *C. elegans* (Chelo et al, 2019; Bernstein et al, 2019). Such observations led many to conclude that overdominant selection, *i.e.* a higher fitness of heterozygotes compared to either homozygote, was operating (Kimura and Ohta, 1971; Charlesworth and Charles But truly overdominant loci are rare, and most effects previously attributed to overdominance (such as heterosis and hybrid vigor) can be explained by simple dominance interactions (Crow, 1999b). Curiously, analyses of inbreeding depression often detect evidence of overdominance (see for example Baldwin and Schoen 2019). These apparent overdominant effects, however, probably reflect the effects of many deleterious recessive mutations linked in repulsion, a phenomenon termed pseudo-overdominance (hereafter POD, introduced by Ohta and Kimura 1969; reviewed by Waller 2021). We have known for half a century

3

that a single strong overdominant locus can generate enough selection against homozygotes to persist even under complete self-fertilization (Kimura and Ohta, 1971). Could such strong effects also arise and persist via pseudo-overdominance?

~~By greatly increasing homozygosity, inbreeding exposes deleterious recessive mutations, reducing inbred fitness and selecting against those mutations. This purifying selection acts to purge unlinked, highly deleterious, recessive mutations, reducing the inbred load (Lande and Schemske, 1985a). Selection is far less effective against mildly deleterious mutations, allowing many to drift and fix whenever the product of effective population size (N_e) and their selective effect in homozygotes (s) approaches one or less (Crow and Kimura, 1970; Hedrick et al, This generates an appreciable fixed~~ Pseudo-overdominant selection will only emerge in genomic regions where many deleterious alleles are clustered together and often linked in repulsion, generating complementary haplotypes that express similar inbreeding loads as homozygotes. Genomic regions with reduced recombination, such as centromeric regions and chromosomal inversions, often maintain higher than expected heterozygosity. Centromeric regions in *Zea mays*, for example, maintain heterozygosity even after repeated generations of inbreeding (McMullen et al, 2009). This has also been found in 22 centromeric regions in the human genome (Gilbert et al, 2020). Kremling et al (2018) confirmed that many rare variants in maize express deleterious effects confirming that "~~drift load~~even intensive artificial selection is insufficient to purge genetic load." ~~in many small populations (Whitlock et al, 2000; Charlesworth, 2018). Strong selection acting against deleterious~~ Brandenburg et al (2017) identified 6,978 genomic segments ($\approx 9\%$ of the genome) with unexpectedly high heterozygosity in land races of maize. These heterozygous segments contained more deleterious mutations than other parts of the genome, with several deeply conserved across multiple land races. Inversions, which halt recombination, also appear to accumulate lasting loads of deleterious mutations~~at some loci also tends to fix some deleterious mutations segregating in repulsionat linked loci. This process is termed background selection (Charlesworth et al, 1997; Zhao and Charlesworth, 2016). Purifying selection , drift, and background selection all reduce heterozygosity and thus segregating variation and ID. It is thus puzzling why appreciable ID and outcrossing~~

4

mechanisms persist in many inbred plant and animal populations (Waller, 2021, termed this "Darwin's paradox").

. Jay et al (2021) found that ancient inversions contribute greatly to heterosis in Heliconius butterflies. Kirkpatrick (2010) concluded that although the genetic basis for inversion overdominance has not yet been clearly determined, POD is plausible.

That heterozygosity, ID, and outcrossing mechanisms persist even in inbred populations suggests the existence of some force countering purging and fixation. Inbreeding reduces effective rates of recombinationwhile increasing linkage and identity disequilibria, the associations occurring between alleles and genotypes at different loci (Sved, 1971; Lewontin, 1974). These associations generate selective interference among loci, reducing the efficiency of purifying selection (Roze, 2015). Some hypothesize that this selective interference might sustain enough segregating variation to account for persistent inbreeding depression and load(Winn et al, 2011).New mutations or the regular introduction of alternative fixed mutations in crosses between sub-populations provide other mechanisms to replenish variation and sustain ID (Charlesworth, 2018).

Overdominance generating balancing selection against both homozygotes provides another mechanism that could sustain heterozygosity and ID. Single-locus overdominance, however, is considered only a rare contributor to inbreeding depression relative to the deleterious recessive mutationsthat constitute most of the load segregating in natural populations (Charlesworth and Charlesworth, 1998; Crow, 1999a). Most apparent overdominance instead appears to reflect pseudo-overdominance (Charlesworth and Willis, 2009; Hedrick et al, 2016). According to expectations from previous theoretical works, whenever deleterious recessive mutations are linked in repulsion, pseudo-overdominance will emerge (Ohta, 1971). Because pseudo-overdominance is a consequence of commonly occurring minor deleterious loci , it should be more frequent than true overdominance. Extrapolating Kimura and Ohta's (1971) results on true and associative overdominance, when the cumulative overdominant effects are strong enough, pseudo-overdominant selection should act to sustain segregating genetic variation and inbreeding depression, even within small and inbred populations, making its consequences on selection highly relevant for understanding the evolution of inversions Kirkpatrick (2010), inbreeding depression (Waller, 2021; Glémin, 2021), and mutational

~~dynamics~~ Pseudo-overdominance (POD) at many loci of small effect should mimic overdominant selection at a single locus, favouring heterozygosity for load within particular genomic regions~~(Zhao and Charlesworth, 2016; Becher et al, 2020). Waller (2021) reviewed these ideas and the evidence that localized zones of high heterozygosity exist generating balancing selection~~. This could sustain inbreeding depression even in the face of purifying selection and drift. For POD to influence species evolution, it must exist for long enough and generate enough overdominant selection to leave a signature. Recombination, however, acts to break up such regions by unbalancing haplotype loads, allowing selection and drift to purge or fix their mutations. It is thus remarkable that polymorphic inversions expressing balancing selection to date back to ancient hybridization events in Heliconius butterflies (Jay et al, 2021). Similarly, five ancient polymorphic zones predate the divergence of *Arabidopsis* from *Capsella* (Wu et al, 2017). These observations suggest that polymorphic regions may generate enough selection to sustain themselves for long periods of time. Could this selection derive from POD?

~~Here we evaluate whether the conditions necessary for maintaining pseudo-overdominance are plausible and thus might help account for persistent genetic diversity and inbreeding depression particularly in partially self-fertilizing populations. We use theory and simulations to explore the evolutionary stability and dynamics of pseudo-overdominant blocks of load loci ("POD's" here following Waller, 2021). In particular, we first link existing theory on overdominant selection to evaluate the conditions necessary for the stability of POD's made up of varying numbers of segregating mutations. Unless they are associated with truly overdominant loci (i.e., sheltered load), it seems unlikely that POD's would arise in isolated populations. New mutations are more likely to go to fixation or be purged before enough are accumulated to generate a POD capable of affecting other genetic processes. It therefore seems more plausible that POD 's would be observed in hybrid populations isolated long enough to allow for mutations in low recombination regions to fix. We therefore simulate~~ Several mechanisms might generate enough initial overdominance to create a POD zone including crosses between independently inbred lineages or sub-populations (generating high heterosis in the F1), a truly overdominant

(e.g., self-incompatibility) locus, or chromosomal inversions where recombination is strongly suppressed, allowing mutations to accumlate. Here, we use simulations to study the evolutionary dynamics of POD zones generated initially by admixture between two populations ~~, each~~ fixed for different sets of deleterious mutations~~within a specific genomic region. This allows us to estimate the parameter values needed to generate a POD with enough power to sustain heterozygosity and ID over many generations. We then~~ . In this scenario, high fitness emerges in the F1 where mutations fixed within each population are 'masked' as heterozygotes in hybrid offspring (Kim et al, 2018). We extend existing theory regarding the stable polymorphism that can exist at a single bi-allelic overdominant locus to examine the conditions necessary for POD to maintain two haplotypes containing many linked recessive deleterious mutations as heterozygotes. Because pseudo-overdominance depends on tight linkage among these loci, we expect that over time such zones will be vulnerable to being broken up by recombination. We therefore also explore how varying levels of linkage, dominance, selection ~~, and self-fertilization affect POD persistence and erosion within populations. We also examine patterns of POD erosion in response to selection and recombination. These results lead to us to discuss when and how often POD's arise, how long they persist , and thus how evolutionarily significant they may be in sustaining heterozygosity, inbreeding depression, and segregating load .~~ and selfing rates affect POD zone stability and decay. Finally, we test how selection elsewhere in the genome affects the ability of POD zones to persist and the reciprocal effects of POD zones on load dynamics elsewhere in the genome.

2 Approaches

2.1 Load needed to generate a POD

Kimura and Ohta (1971) demonstrated that when the selective effects generating true overdominance are strong enough, a stable equilibrium can exist that perpetuates the two overdominant alleles indefinitely even within a fully self-fertilizing population. Con-
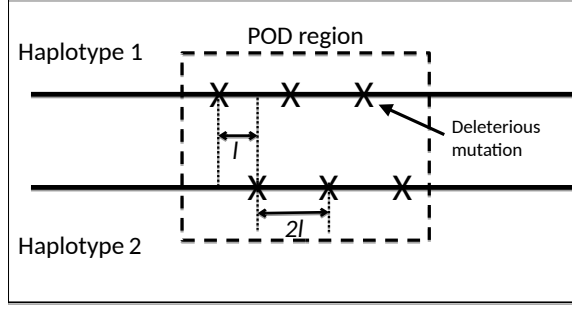
Figure 1: Genetic structure of the POD region (delimited by the dashed box). Deleterious mutations (represented by crosses) linked in cis occur at a distance 2ℓ cM from each other along the same chromosome, alternating (at a distance ℓ cM) with trans mutations on the opposite chromosome. Close, regular, and alternating spacing of recessive deleterious mutations along both haplotypes ensure linkage and pseudo-overdominance.

169 sider a scenario in which two haplotypes, noted H1 and H2, occur within a diploid

170 population self-fertilizing at rate σ. Each homozygote suffers a fitness reduction (s_1

171 or s_2) compared to the heterozygote fitness. In the case of true overdominance, ~~they~~

172 Kimura and Ohta (1971) showed that a stable polymorphism will persist at an ~~over-dominant~~

173 ~~locus (with both alleles at a frequency of 0.5)~~ overdominant locus when:

$$\sigma < \frac{2s_2(1-s_2)}{s_1 + s_2 - 2s_1s_2}\frac{2s_x(1-s_x)}{s_1 + s_2 - 2s_1s_2}. \tag{1}$$

174 ~~This equilibrium always exists when~~ where $s_x = min{s_1, s_2} < 0.5$. When both segregating

175 homozygotes reduce fitness by at least half ($s1$, $s2 > 0.5$)~~. Beyond this point~~, selection

176 acts to maintain overdominance even as the selfing rate approaches one, as selection re-

177 moves homozygotes faster than they are ~~created by selfing (Rocheleau and Lessard, 2000).~~

178 ~~Inequality (1) will always be true when~~ generated (Rocheleau and Lessard, 2000). For

179 situations with stable polymorphism, setting $s1 = s2$ ~~except in completely selfing populations.~~

180 ~~The higher the self-fertilization, the more balanced the load must be (see Supp. Fig.~~

181 ~~A2).~~ results in both alleles being maintained at a frequency of 0.5.

182 We use this threshold under true overdominance to estimate the number of ~~fully~~

183 ~~linked~~ load loci within ~~a block of~~ pseudo-overdominant ~~alleles~~ (POD) zone required to

8

generate the necessary level of overdominance needed to maintain a stable equilibrium (see Eq. 1). ~~This much balancing selection ensures that the POD will persist. To make the models equivalent, we~~ We assume complete linkage among matched sets of mildly deleterious mutations, all with the same coefficient of selection ~~$s = 0.01$ and plausible levels of dominance $h = 0.2$ matching empirical observations in yeast and *Drosophila* – Agrawal and Whitlock (2011). We further~~ s and dominance h. For the sake of simplicity, we assume that each haplotype carries the same number ~~n~~ n_L of deleterious mutations and that mutations within the POD zone are evenly spaced, occurring at intervals of ℓ cM between alternating trans-mutations on opposing haplotypes (Fig. 1). As fitness effects are considered multiplicative across loci, an individual's fitness is:

$$W = (1 - hs)^{he}(1 - s)^{ho} \tag{2}$$

where *he* and *ho* are the number of heterozygous and homozygous mutations, respectively, carried by the individual. ~~Assuming no cross-overs occur within the POD, homozygosity~~ In the case of complete linkage homozygosity at these loci only occurs in individuals ~~that carry~~ carrying two copies of the same haplotype (genotype $H_1 H_1$ or $H_2 H_2$). ~~In this case~~ As both haplotypes carry the same number of mutations, the coefficient of selection acting against either homozygote ($s_H = s_1, s_2$), relative to the fitness of the heterozygote $H_1 H_2$ (W_{AA}/W_{Aa}) is:

$$s_H = 1 - \frac{(1 - s)^n}{(1 - hs)^{2n}}. \tag{3}$$

This expression allows us to determine the number of deleterious alleles per haplotype necessary to sustain enough overdominance to preserve both haplotypes via stable balancing selection (see ~~Supplementary Material~~ Supp. File 1):

$$n_L = \frac{\log(1 - s_H)}{\log(1 - s) - 2\log(1 - hs)} \tag{4}$$

~~Given the~~ As expected, the number of loci required to obtain a strength of selection against homozygotes s_H decreases for higher values of s and h ~~values assumed above and the need~~. For $s = 0.01$ and $h = 0.2$, $n_L = 115$ for s_H to be at least 0.5, ~~115 or more load~~

9

indefinitely (Supp. File 1, Fig. S1). ~~As with true overdominance, strong and symmetrical selection create the most favorable conditions for sustaining both haplotypes.~~

2.2 Inbreeding depression

Inbreeding depression δ is a local variable, reflecting the number of heterozygotes maintained in a population. The general equation used to estimate inbreeding depression is:

$$\delta = 1 - \frac{W_s}{W_o} \tag{5}$$

where W_s is the fitness of selfed offspring and W_o that of outcrossed offspring (Charlesworth and Charlesworth, 1987). When mutations are deleterious, and accounting for drift, this variable depends on the haploid mutation rate U, the coefficient of selection s and the dominance of mutations h (see equation 3 from ~~Bataillon and Kirkpatrick 2000~~Bataillon and Kirkpatrick 2000):

$$\delta_s = 1 - \exp\left[-U\left(\frac{(1-2h)(1+F)}{2(h+F-hF)} - \frac{(1-2h)(1+F)(1-2hs)}{8(h+F-hF)^2sN}\right)\right], \tag{6}$$

where $F = \sigma/(2-\sigma)$ is the equilibrium inbreeding coefficient (expected deviation from Hardy-Weinberg equilibrium of genotype frequencies). Though this expression for F remains true for weak overdominance (Glémin, 2021), when there is strong overdominance, the inbreeding coefficient depends on the coefficients of selection and allelic frequencies (Appendix A4 from Kimura and Ohta, 1971). In our case with symmetrical selection against homozygotes, this term is given as:

$$\hat{F} = \frac{2 - s_H - \sigma + s_H\sigma - \sqrt{(2-s_H)^2 - 2\left(2-s_H-s_H^2\right)\sigma + (1-s_H)^2\sigma^2}}{2s_H}. \tag{7}$$

$\hat{F}$ will tend to zero with increasing s_H (see Fig. A1 in Supp. File 1)~~, eventually strengthening selection enough to remove homozygotes from the population faster than they are generated~~. Selfing populations subject to strong overdominant selection thus tend to behave like outcrossing ones as low fitness homozygotes are eliminated. At equilibrium, this inbreeding load δ_{od} for symmetrical overdominance can be written as:

$$\delta_{od} = \frac{(1+\hat{F})i}{2 - s_H} \tag{8}$$

230 where, when $s_1 = s_2 = s_H$, $i = \frac{s_H^2}{2s_H}$ - see Eq. A2from Supp. File 1 and Kimura and

231 Ohta (1971). We provide the general expressions for $\hat{F}$ and δ_{od} in Supp. File 1 (see Eq.

232 A3).

233   If we assume that selection against deleterious mutations elsewhere in the genome and

234 overdominant selection do not interfere with one another (*i.e.* no associative overdominance

235 or effects of background selection) and fitness effects remain multiplicative (see for ex-

236 ample ~~Kirkpatrick and Jarne 2000)~~Kirkpatrick and Jarne 2000, the upper limit of the

237 expected level of inbreeding depression will be:

$$\delta = 1 - (1 - \delta_{od})(1 - \delta_s), \tag{9}$$

238 setting F in Eq. 6 to $\hat{F}$.

239   ~~Depending on the strength of overdominant selection~~As previously shown, δ_{od} ~~can~~

240 ~~increase~~ increases with the selfing rate σ ~~(Charlesworth and Charlesworth, 1987, 1990).~~

241 ~~The opposite tendency is observed for~~ for strong overdominant selection and δ_s ~~, which~~

242 decreases with σ ~~as purging becomes more efficient~~(Charlesworth and Charlesworth, 1987, 1990).

243 It is therefore possible to have similar δ ~~maintained at similar levels~~ in outcrossers and

244 selfers, depending on the rates of background mutation U and the strength of ~~the POD~~

245 POD selection.

2.3 Recombination and POD's

247 Thus far, we have assumed complete linkage in order to apply one-locus overdominance

248 theory to infer the strength of selection against homozygotes necessary to sustain a

249 stable equilibrium. However, some recombination will occur, allowing the strong linkage

250 disequilibrium among loci within a POD to erode over time. ~~If recombination within~~

251 ~~the POD block is rare (reflecting tight linkage), any new haplotype H_c will be generated~~

252 ~~by a single recombination event. As the positions of deleterious alleles among different~~

253 ~~haplotypes alternate in trans relative to flanking mutations on the same chromosome in~~

254 ~~H_1H_2 heterozygotes (Fig. 1), the recombinant haplotype H_c, and its complement will~~

255 ~~also carry n deleterious homozygous mutations, except when recombination occurs at a~~

256 ~~tip of the POD. A single mutation is then cleaved, trimming the load to $n - 1$ deleterious~~

11

257 alleles. Using ~~In order to examine the effect of recombination on the stability of POD, we~~

258 ~~propose~~ a system of Ordinary Differential Equations ~~, (ODEs) representing~~ the change

259 in frequencies ~~P_1, P_2 and P_c of the three respective haplotypes H_1, H_2 and H_c can be~~

260 ~~expressed~~of the two initial haplotypes (Δ_{P_1} and Δ_{P_2}) and that of a newly introduced

261 recombinant haplotype (Δ_{P_c}):

$$
\Delta_{P_1} = \frac{P_1((1-\hat{F})(1-s_{c,1})P_c + (1-s_H)((1-\hat{F})P_1 + \hat{F}) + (1-\hat{F})P_2) - P_1\overline{W}}{\overline{W}}
$$
$$
\Delta_{P_2} = \frac{P_2((1-\hat{F})(1-s_{c,2})P_c + (1-s_H)((1-\hat{F})P_2 + \hat{F}) + (1-\hat{F})P_1) - P_2\overline{W}}{\overline{W}}
$$
$$
\Delta_{P_c} = \frac{P_c((1-\hat{F})(1-s_{c,1})P_1 + (1-s_c)((1-\hat{F})P_c + \hat{F}) + (1-\hat{F})(1-s_{c,2})P_2) - P_c\overline{W}}{\overline{W}}.
$$

$$(10)$$

262 The mean fitness of the population $\overline{W}$ is the sum of the expected genotypic frequencies

263 after selection (see ~~Supplementary Material~~Supp. File 2, Eq. (A4)), and s_c, $s_{c,1}$ and $s_{c,2}$

264 are the coefficients of selection associated respectively with haplotypes H_cH_c, H_cH_1 and

265 H_cH_2. ~~Resolving this system for arbitrary values of s_c, $s_{c,1}$ and $s_{c,2}$, one finds that the~~

266 ~~only possible equilibria are those involving the fixation of one of the three haplotypes,~~

267 ~~or maintaining only two of them. Hence a rare~~ We resolve this system of equations to

268 determine the conditions necessary for a recombinant haplotype H_c ~~should either be lost,~~

269 ~~go to fixation or replace one of the initial haplotypes and co-exist with the remaining~~

270 ~~haplotype. In order for H_c~~ to increase in frequency ~~, Δ_{P_c} must be positive at the moment~~

271 ~~it enters the population, otherwise it will be eliminated immediately. Assuming P_c is of~~

272 ~~order ϵ (ϵ being very small), the expression for Δ_{P_c} for the leading order of P_c (noted~~

273 ~~$\bar{\Delta}_{P_c}$) can be derived. In a population at equilibrium with $P_1 = P_2 = (1-\epsilon)/2$:~~

~~.~~ ($i.e. only n changes$), $but also for s, h and n chosen for a constant s_H$ (Eq. 3 and see Fig.

B2 in Supp File 2). $\Delta_{P_c} > 0$).

~~Combined with the previous expectations detailed above, we predict that if a POD is~~

~~initially stable, its eventual loss will usually occur gradually as recombination at the distal~~

~~ends of the POD cleaves off mutations. This creates haplotypes with improved relative~~

~~fitnesses. The reduced zones of stable equilibria for $s_c = s_H$ in selfing populations (Fig. A2,~~

in Supp.File 1) means that POD's will destabilise, erode, and more easily be lost in such populations, resulting in the rapid fixation of one of the original haplotypes or a recombinant. Again, this process will occur faster when fewer load loci are involved as the loss of one mutation has a greater effect on the relative fitness. Persistent POD 's only occur if mutations have small deleterious effects. In outcrossers, POD's can persist for longer as new haplotypes are not strongly counter selected when $s_c = s_H$, and when $s_c < s_H$ (*i.e.* a mutation is cleaved)selection is not as positive as in selfing populations (see Fig. B2 in Supp. File 2). Drift thus plays an important role in how recombinants increase in frequency in small populations.

3   Simulations

We use an So as to confirm expectations from the analytical model given above and explore the dynamics of POD selection, we develop an individual-based simulation program in Csimulation program in C++ to evaluate whether and for how long POD's can be maintained under various conditions. Populations consist +, uploaded to Zenodo.org (Abu Awad and Waller, 2022). We consider a scenario where POD selection arises after an admixture event between two initially isolated populations fixed for different mutations within the same genomic region (a "proto-POD" zone). Each population is made up of N sexual diploid individuals, self-fertilizing at a fixed rate, σ. Each individual is represented by two vectors, each carrying the positions (between 0 and 1) of deleterious mutations along a single chromosome with map length R Morgans. Recombination occurs uniformly through throughout the genome. Mutations in and outside the POD within and outside of the POD zone have a fixed effect, with respective coefficients of selection, s and s_d, and dominances, h and h_d. Individual fitness is calculated as shown in Eq. 2. To assess the effects of background selection we introduce new mutations New mutations are sampled from a Poisson distribution with parameter U, the haploid mutation rate . The positions of these mutations are uniformly sampled, and

~~they can be introduced anywhere in~~ and their positions are uniformly distributed along the genome (~~infinite-alleles~~ infinite-locus model). Generations are discrete (no overlap) and consist of three phases: *i*) introducing new mutations, *ii*) selection, and *iii*) recombination and gamete production.

3.1 ~~Introducing pseudo-overdominance~~ POD zone architecture and initiation

~~To assess how well POD's persist in populations, we introduce~~ Two types of simulation are run, one with an arbitrary ideal haplotype structure expected to favour POD persistence ~~.~~ ~~This consists of two~~ and one with a more realistic distribution of mutations within the POD zone. The former consists of constructing two perfectly complementary haplotypes, H_1 and H_2 ~~, each carrying n mutations. As before, cis-mutations~~. Cis-mutations occur at regular intervals (every 2ℓ ~~cM~~ M) along each haplotype and mutations are staggered, spreading the load evenly through the POD and ensuring pseudo-overdominance (Fig. 1). The probability that a recombination event occurs between two trans-mutations is then ℓ. The ~~total length of the POD zone is then the product of the number of loci involved and the map distances between them, or $2\,\ell\,n$ cM.~~ second type of POD zone architecture is one with randomly placed mutations in a predefined genomic region, their positions sampled from a uniform distribution, while ensuring that a locus with the same position is not sampled for both haplotypes. In both cases the center of the POD zone is kept constant for both haplotypes and the size of the POD zone is $2\ell n_L$ M, with n_L potentially different for each haplotype. The POD zone is arbitrarily positioned around the center of the genome, its exact center at position 0.5 along the chromosome.

3.2 ~~Initial conditions and variables measured~~

~~Each simulation begins with two populations of size N. Each population is fixed for either H_1 or H_2.~~ After a burn-in period of 4 000 generations, ~~a~~ allowing the two source populations

(each fixed for a given haplotype in the proto-POD zone) to reach mutation-selection-drift equilibrium, a new population of size N is created by randomly sampling individuals from ~~the two source populations, with each population contributing~~ both populations. We arbitrarily consider that each source population contributes 50% of individuals to the new population. ~~Simulations then run for another~~ The new population is then allowed to evolve for a further 4000 generations. Samples of 100 individuals are taken every 10 generations to estimate inbreeding depression, which we compare to the theoretical expectations presented above (Eqs. 6, 8 and 9). We also use these samples to estimate heterozygosity within and outside the POD zone (POD H_e and genome H_e, respectively) as:

$$H_e = \sum_{j=1}^{100} \frac{he_j}{L}. \tag{11}$$

where he_i is the number of heterozygous mutations carried by individual j (out of a sample of 100) and L is the total number of segregating sites in the genomic region of interest. At higher mutation rates, singletons will be frequent. This will reduce H_e by inflating L. ~~We thus also measure heteroyzgosity within the POD block keeping L fixed at $2n$, the number of initial segregating sites, giving us a measure of "absolute" heterozygosity $\hat{H}_e$, to accurately represent how POD's erode. As long as a POD is stable, $\hat{H}_e$ should be equivalent to the frequency of heterozygotes expected at equilibrium in a true overdominant system, given by (Kimura and Ohta, 1971, see Supp. File 1):~~

~~We initially ran simulations for three population sizes ($N = 100, 1000$~~ Simulations are run for population size $N = 100, 1000$ and 5000 ~~) with no background selection ($U = 0$) and three fixed selfing rates $\sigma = 0$,~~ and for selfing rates σ between 0 and 0.95. The haploid background mutation U is set to $0, 0.1$ and 0.5, with new mutations outside the POD zone having a fixed coefficient of selection ($s_d = 0.01$) and dominance ($h_d = 0.2$ or 0.5). The general map length $R = 1$ and ~~0.95. In these simulations we fix the number of mutations to be $n = 25, 57$ and 100, corresponding to relative~~ 10 Morgans and we examine different

strengths of linkage between loci in the POD zone, with $\ell = 10^{-4}, 10^{-5}$ and 10^{-6}. We consider both weak and strong selection against homozygotes, setting s_H to $s_H = 0.14$, ~~0.29~~ 0.26 and 0.45. These ~~result in overdominant selection sufficient to ensure stability~~ correspond to stable (polymorphic) overdominant selection when $\sigma = 0$, 0.5 or even (with a narrow range of stability) 0.95 (Fig. A2, dotted lines). ~~We also compare three levels of linkage between cis-mutations (inter-mutaion lengths of $\ell = 5.10^{-3}$ cM, 1.10^{-3} cM and 5.10^{-4} cM). Finally, we explore the consequences of background selection on POD dynamics for $N = 1000$~~To determine the effects of POD selection on heterozygosity elsewhere in the genome, ~~$0 \leq \sigma \leq 0.95$, $n = 25$ and 100 and $\ell = 5.10^{-4}$ cM. In these simulations, we set the haploid mutation rate U to either 0.1 or 0.5 and assume a total genomic map length of either $R = 1$ or 10 Morgans. To determine whether the heterozygosity maintained within the POD reflects selection rather than drift,~~ we also ~~ran control~~ run simulations where all alleles within the initial POD zone are neutral for all parameter sets mentioned above (achieved by setting s and $h = 0$ within the POD). We ~~ran~~ run 100 repetitions for each parameter set.

4 Results

4.1 POD ~~dynamics~~persistence and degradation

We first ~~explore how various parameter values affect the stability of pseudo-overdominance, then determine the role of fixation and purging in destabilising and eroding POD's. We use heterozygosity~~ examine how recombination, the strength of selection against linked load loci, and their arrangement within the POD zone, influence POD persistence.

4.1.1 Recombination and POD degradation

Under the assumption that recombination within the POD ~~as a measure of POD stability, with any decline from that expected under overdominant selection (Eq. ??) implying erosion of the POD. Rates of decline in mean absolute heterozygosity $\hat{H}_e$ (see Eq. 11~~

16

~~and accompanying text)~~ strongly depend on how closely load loci are linked and levels of ~~pseudo-overdominance (reflecting the coefficient~~ block is rare (reflecting tight linkage), any new haplotype H_c will be generated by a single recombination event. This is reflected in the ODEs introduced in Eq. (10) which compute changes in frequency of the two initial haplotypes (H_1 and H_2) and a recombinant (H_c). For simplicity, we initially assume an ideal case where mutations are arranged alternately within the POD zone (see Fig 1). Positions of deleterious alleles in $H_1 H_2$ heterozygotes alternate in trans relative to flanking mutations on the same chromosome (Fig. 1). Each haplotype carries n_L deleterious mutations. Consider two cases: 1) the recombinant haplotype H_c (and its complement) each carry n_L deleterious mutations; 2) H_c carries $n_L - 1$ mutations because recombination has cleaved one from one end of the POD zone.

Given arbitrary values of s_c, $s_{c,1}$ and $s_{c,2}$ (the coefficients of selection against ~~homozygotes s_H). The number of load loci on a haplotype, n, has stronger effects as linkage increases. For tight linkage ($\ell = 5.10^{-4}$ cM) and a sufficiently large n, POD zones behave like true overdominant loci for 500+ generations (red lines in Fig. ??c and d). This reflects the strength of overdominant selection even under high rates of self-fertilisation, σ. When comparing POD dynamics to control simulations ($s = 0$, blue lines in Fig. ??)~~ $H_c H_c$, we ~~confirm that POD selection greatly slows the decrease in $\hat{H}_e$ relative to the neutral case. However, once $\hat{H}_e$ drops to a threshold reflecting the erosion of the structured haplotypes, selection quickly eliminates heterozygosity much faster than in the neutral case. This reflects POD instability occurring faster when linkage is weak (i.e. $\ell = 5.10^{-3}$ cM, red lines in Figs. ??a and b). As expected from the~~ $H_c H_1$ and $H_c H_2$ genotypes, respectively), the only possible equilibria involve fixing one of the three haplotypes or maintaining only two of them. Hence any rare haplotype, H_c, should either be lost, go to fixation, or replace one of the initial haplotypes (co-existing with the other). For H_c to increase in frequency, Δ_{P_c} (Eq. (10)) must be positive when it enters the population (or it would be eliminated). Assuming

the frequency of a recombinant P_c is of order ϵ (ϵ being very small), the expression for Δ_{P_c} for the leading order of P_c (noted $\bar{\Delta}_{P_c}$) can be derived. In a population at equilibrium with $P_1 = P_2 = (1 - \epsilon)/2$ and setting $s_1 = s_2 = s_H$:

$$\bar{\Delta}_{P_c} = \frac{2((1 + \hat{F})s_H - s_{c,1} - s_{c,2} - \hat{F}(2s_c - s_{c,1} - s_{c,2}))}{2 - s_H - \hat{F}s_H}. \tag{12}$$

The denominator of this expression is always greater than 0 for $s_H < 1$. To understand the behavior of $\bar{\Delta}_{P_c}$ ~~(Eq. 12), increased selfing tends to favor recombinants whether or not a mutation has been cleaved (Fig.~~ we simplify the above equation by setting $\hat{F}$ to 0 (no self-fertilisation or very strong overdominant selection with $s_H \approx 1$, see Supp Fig. A1). In this case Eq. 12 simplifies to $2(s_H - s_{c,1} - s_{c,2})/(2 - s_H)$. If no mutations have been cleaved off by recombination (*i.e* H_c carries n mutations), the numerator $2(s_H - s_{c,1} - s_{c,2}) \leq 0$ (see Eq. B1 in Supp. File 2 for expressions of $s_{c,1}$ and $s_{c,2}$) making $\bar{\Delta}_{P_c}$ negative (Fig. B2 in Supp. File 2). Hence H_c haplotypes will be selected against. If instead H_c carries $n - 1$ mutations, the resulting coefficients of selection (Eq. B2 , Supp. File 2) lead to a positive $\bar{\Delta}_{P_c}$ (the numerator in this case can be positive). The larger $\hat{F}$ (or the selfing rate σ) the more positive the resulting $\bar{\Delta}_{P_c}$. ~~B2~~

This result leads us to predict that if a POD is initially stable, its eventual loss will usually occur gradually as recombination events near the distal ends of the POD cleave off mutations creating haplotypes with improved relative fitness. The reduced zones of stable equilibria for $s_c = s_H$ in selfing populations (Fig. A2, in ~~Supp.~~ Supp. File ~~2), destabilising POD's faster in selfing than in outcrossing populations.~~ 1) means that selection will more easily act to destabilise the POD zone by eroding mutations. This should fix one of the original haplotypes or a recombinant with the strength of selection affecting the rate at which this occurs.

~~Declines in heterozygosity within the POD zone due to drift and selection. Graphs show declines in absolute heterozygosity $\hat{H}_e$ as a function of different levels of linkage (ℓ)~~

18

and differing numbers of load mutations within the POD zone (n), with resulting lengths of the POD zones given as bold numbers in each panel. Blue lines represent control cases where all POD mutations are neutral ($s = 0$). Red lines represent simulations run with deleterious mutations ($s = 0.01, h = 0.2$) within the POD zone. Full lines represent completely outcrossing populations ($\sigma = 0$) while lower dashed lines represent mostly self-fertilizing populations ($\sigma = 0.95$). Light shading around each line shows standard errors. Black lines represent heterozygosity expected with true overdominance (Eq. **??**). Population size $N = 1000$ with no new mutations introduced anywhere in the genome ($U = 0$).

~~Declines in $\hat{H}_e$ mainly reflect fixation at the load~~ Using simulations, we confirm results from single locus overdominance that stronger selection is more likely to result in stable polymorphism even for high selfing rates (Supp. Fig. S2). Drift and selection can both act to erode POD (shown by the rate of decrease of heterozygosity in Supp. Fig. S2). Strong drift renders selection neutral when $N_e s_H << 1$, accelerating the loss of supposedly stable POD selection ($N = 100$ in Supp Fig. S2). Increasing the efficacy of selection will also favour the loss of POD selection, but unlike for strong drift, this is due to a more efficient purging (and higher effective recombination rate) of loci contributing to ~~the POD~~ (POD selection ($N = 5000$ in Supp Fig. S2). As the differences between population sizes are quantitative, and s_H is a good predictor of mid/long-term stability of POD zones, in the following, we examine simulations only for $N = 1000$, for which both drift and selection act on POD stability, and $s_H = 0.45$, for which overdominant selection is stable for all self-fertilisation rates simulated.

4.1.2 Effect of the strength of selection against individual loci

As mutations are progressively lost from POD zones, recombinants can go to fixation. This will eventually destabilize the POD zone. We next assess how varying the coefficients of selection s and dominance h against individual loci affects POD persistence. For a fixed

value of selection against homozygotes, s_H, varying s, h and n_L (obtained using Eq. (4)), we calculate the expected increase in frequency a recombinant haplotype $\bar{\Delta}_{P_c}$ using Eq. (12). If no mutation is lost (H_c also carries n_L mutations), $\bar{\Delta}_{P_c}$ remains negative except under high rates of self-fertilisation when they can be positive (though close to 0). However, a mutation lost through recombination generates a positive $\bar{\Delta}_{P_c}$ that increases with increasing strengths of selection and dominance of the mutations for all rates of self-fertilisation (Figs. 2 a and b for $s_H = 0.45$). We confirm this prediction via simulations. These show that most losses of diversity (fixation or loss of mutations) occur at the ends of the POD zone (Figs.2c and d for selfing rate $\sigma = 0.95$). Losses of diversity within the POD zone intensify as s and h increase.

Stronger selection against individual mutations sustains heterozygosity more effectively as fewer mutations suffice to generate the same amount of balancing selection. However, the loss of a stronger mutation as a result of recombination will more likely unbalance and destabilise the POD zone. This accelerates the fixation or loss of mutations (Fig.~~??a for $\sigma = 0.95$ and Supp. Fig. ??a for $\sigma = 0$). Generally, when the POD is unstable (weak linkage and /or a small number of load loci n), fixation occurs throughout the POD. For more stable POD's, mutations are more commonly fixed on either end of the POD. Self-fertilising populations with low load (small n) quickly fix single recombinant haplotypes (Fig.3b). For larger n, mutations fix more slowly, as recombination events cleave off single mutations at either end of the POD under tight linkage ($\ell = 5.10^{-4}$cM, Fig.~~??~~2c). In outcrossing populations, where fixation takes longer, we do not observe single recombinant haplotypes being fixed, though the pattern of fixation at the ends is maintained (Supp Fig. ~~??~~Increasing the dominance of load loci has similar effects as increasing s but requires more mutations to reach the same s_H (i.e. $n_L = 60$ and 150 for $h = 0$ and 0.3 respectively, Fig. 2f). This ~~reflects the greater predicted stability of POD's~~ is because increased dominance increases the relative fitness of both the fitter homozygote (i.e. the haplotype with one less mutation due

to recombination) and the heterozygote, increasing the overall fitness advantage of losing a mutation. The same patterns are observed in outcrossing populations ~~despite weaker selection (smaller n; Supp. File 1~~ to a lesser extent (Supp. Fig. A4). Increased linkage within the POD zone reduces the rate at which these higher fitness recombinants occur, slowing this process (dashed lines, Figs. 2e and f; see Supp. Fig. S4 for patterns of mutation loss within the POD zone).

~~A surprising result emerges when we compare PODpersistence $\hat{H}_e$ in populations of different size N reflecting effects of genetic drift. Counter-intuitively, POD heterozygosity persists much longer (relative to neutral heterozygosity) in smaller populations than in larger populations (compare the red and blue lines in Fig. ?? where $N = 100$ to Fig. ?? and Fig. ?? where $N = 1000$~~

4.1.3 POD region architecture

So far, we have considered only an ideal genetic architecture that favours maintaining POD, namely homozygotes of both haplotypes having identical fitness disadvantages relative to the heterozygote and equally spaced cis and trans mutations within the POD zone. We now relax these assumptions by considering initial haplotypes carrying different numbers of mutations, n_L, within the POD region (while maintaining equal spacing) and then by placing randomly spaced mutations within the POD zone.

To unbalance the segregating homozygotes, consider alternative POD zone haplotypes with $n_L = 80, 100, or 120$ mutations paired with a haplotype H_1 with $n_L = 100$ mutations (denoted by relative lengths of 0.8 1 and 1.2 respectively in Figs. 3a and c). These generate substantial fitness differentials with relative selection coefficients against homozygotes $s_1 = 0.47$ and $s_2 = 0.35$ (blue lines), $s_1 = s_2 = 0.45$ (black lines), or $s_1 = 0.43$ and ~~5000). While POD's are always strong and stable when many load loci are tightly linked (panels c and d in each figure), POD's that erode quickly after ≈ 1000 generations in $N = 5000$ populations (Fig.~~
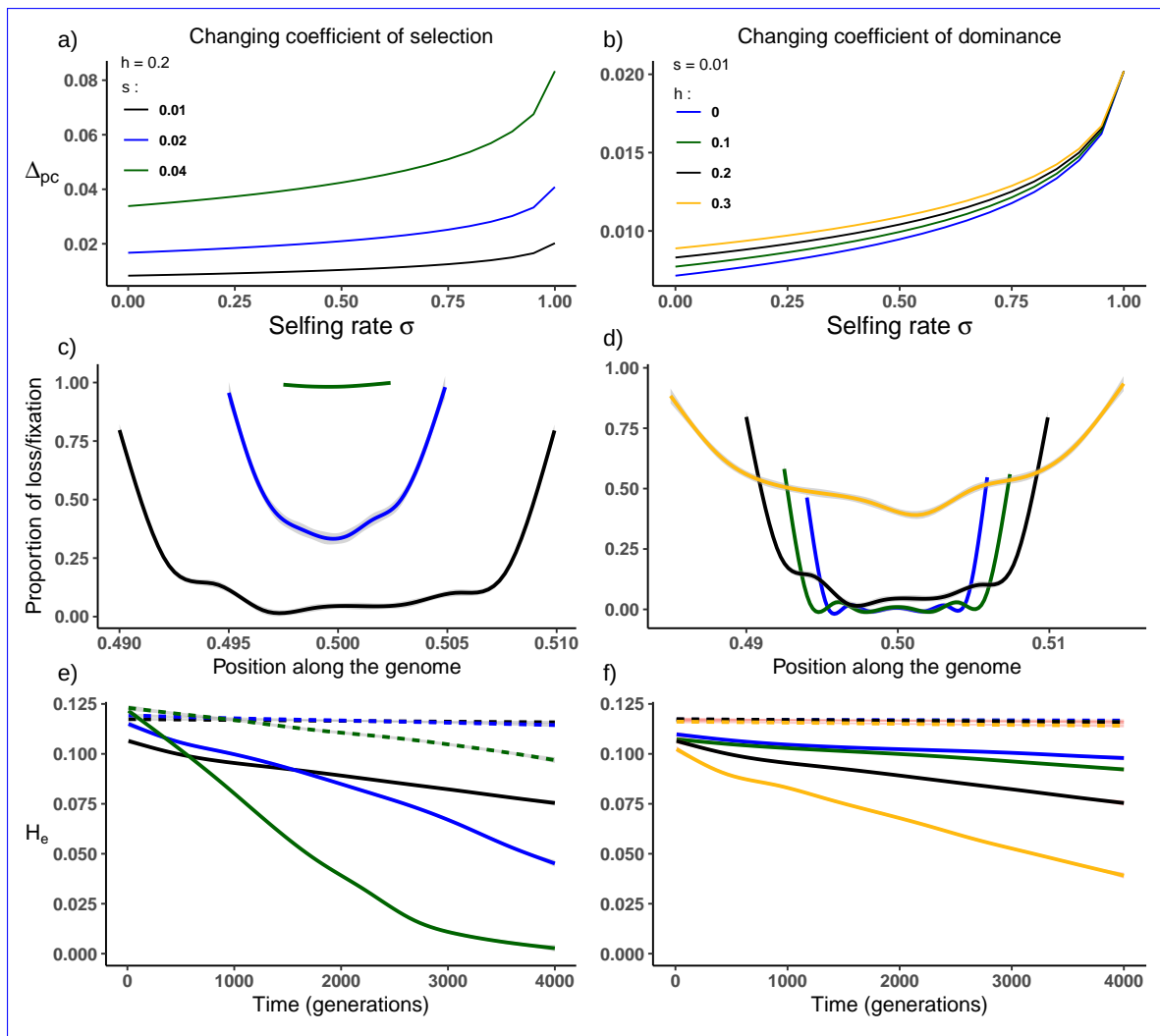
Figure 2: ~~Fixation~~ Effects of levels of selection and dominance on selection dynamics within ~~the~~ a POD zone ~~in highly selfing populations ($\sigma = 0.95$). a) Proportion~~ Left panels show the effects of ~~mutations fixed within~~ varying the ~~POD zone over time containing~~ coefficient of selection at ~~a POD with either loose or tight linkage (dotted or solid lines) reflecting longer ($\ell = 5.10^{-3}$ cM) or shorter ($\ell = 5.10^{-4}$ cM) POD regions. Weaker POD's ($n = 25$~~ load locus s ($s = 0.01, 0.02$ and 0.04, corresponding to $n_L = 100, 50$ and 25 loci ~~per haplotype) decay relatively rapidly losing most of their heterozygosity through fixation within 2000 generations (upper lines). Stronger POD's ($n = 100$) are robust persisting via strong balancing selection for many thousands~~ Dominance is fixed at $h = 0.2$ and $s_H = 0.455$. Right panels show the effects of ~~generations~~ varying dominance (~~lower coincident lines~~ $h = 0, 0.1, 0.2$ and 0.5 with $n_L = 60, 75, 100$ and 150) with selection fixed at $s = 0.01$. ~~b~~ Panels a) and ~~c~~ b) ~~Maps showing distributions for 100 replicate simulations~~ show theoretical rates of ~~where loci are fixed (blue and yellow pixels representing each~~ increase in frequency for a recombinant haplotype ~~that~~

22

~~**??**, panels a, b) remain more heterozygous than expected out to > 2500 generations if N is 1000~~ $s_2 = 0.53$ (green lines). In outcrossing populations, selection trims down longer, more loaded haplotypes as recombination makes variants available. This shrinks more loaded haplotypes to sizes close to the smaller haplotype (Fig. ~~**??**, a, b). At $N = 100$, POD heterozygosity lasts 1000 to 1500 generations longer than neutral heterozygosity (FIg. **??**, a, b), though both decline greatly by 3000 generations unless many loci are strongly linked. In the presence of strong genetic drift, selection against or for new recombinant haplotypes is weakened (equivalent to bringing $\bar{\Delta}_{P_c}$ closer to 0). Another factor contributing to this longer-lived stability is that recombinant haplotypes occur more rarely in smaller populations as there are fewer individuals (reducing the effective recombination rate). In selfing populations, increasing drift mostly increases the variance of $\hat{H}_e$ (shading around~~ 3a, solid lines). Overdominant selection, however, sustains the core POD region's heterozygosity, H_e(Fig. 3b, solid lines). Self-fertilising populations, in contrast, show less POD zone stability under asymmetric selection despite the fact that populations with balanced loads showed only slight observed losses or fixations of mutations (dashed black lines in Figs. 3a and c). When the alternative haplotype has less load (a relative size of 0.8), it quickly goes to fixation (dashed blue lines in Figs. ~~**??**, **??** and **??**). Although mean $\hat{H}_e$ with POD selection initially remains higher than for control (neutral)simulations, it does so for a much shorter time-span. This increases the overlap between the control simulations and those with POD selection, making it difficult to distinguish between a decrease of heterozygosity due to drift and that due to selection against the POD~~ 3a and c). This result matches the theoretical expectation that no overdominant polymorphism can be maintained with these coefficients of selection against homozygotes when the selfing rate is 0.95 (see Fig.A2 in the Supp. File 1).

~~The simulations also allow us to track the dynamics of mutation fixation and loss, illustrating further how stronger loads and tighter linkage promote POD stability~~ When

the total load of the second haplotype increases to a relative size of 1.2, the POD zone is more commonly sustained as mutations are trimmed off the ends of the POD zone (Fig. ~~??). With limited load ($n = 25$ load loci per haplotype), selection and drift act to dissolve the POD within a few thousand generations. The optimal structure of the POD, designed to maximize pseudo-overdominance andminimize effects of crossing over, may account for the initial $\approx$ 300-generation stability even in these weak POD 's (lower left, Fig. ??a). After that, however, mutations fix or purge quickly in~~ 3a, c). This difference in behavior reflects the need for segregating load to exceed a threshold to sustain a POD zone. As for outcrossing, most mutations of the larger haplotype will be trimmed off the edges, but there is some fixation and/or loss of mutations along the whole POD region (dashed green line in Fig. 3a), lowering the mean observed H_e (dashed green line in Fig. 3c). This is most probably due to a larger range of recombinants having a higher selective advantage, provided that they trim the larger haplotype and thus help destabilize POD selection.

When the mutations are not in an ideal configuration, but randomly positioned throughout the designated POD zone, stability of the ~~absence of strong balancing selection, resulting in a fairly even distribution of fixation across the length of the POD (Fig. ??b). When selection is strong ($n = 100$)~~POD zone is barely affected in outcrossing populations (solid lines in Figs. 3b and d), even when the haplotypes are initially uneven. Selfing populations, however, ~~the power of drift and selection to fix and purge these minor mutations is almost eliminated (lower curves, Fig. ??a). Nevertheless, some fixation of mutationsoccurs via crossing-over, mostly at the far ends of the POD (Fig. ??c). In this case, drift generally appears to be more potent than purifying selection particularly in selfing populations.~~ require stronger linkage to retain the POD zone (compare dashed lines in Fig. 3 for $\ell = 10^{-6}$ M to Fig. S5 for $\ell = 10^{-5}$). Despite more frequent fixations/losses of mutations, some heterozygosity nonetheless persists for approximately 1000 generations even with lower linkage (Supp. Fig. S5).

4.2 ~~Background selection and POD~~

~~Background selection acts in two opposing ways on the observable heterozygosity~~

4.2 Background mutations

Mutations introduced elsewhere in the genome influence POD selection dynamics and persistence and vice versa as POD's affect purifying selection across the genome. In general, when a POD zone is stable, background mutations will not destabilise it. Background selection does, however, affect heterozygosity within and outside the POD zone. Let us compare heterozygosity within the POD zone in simulations with background mutations to simulations lacking it (*i.e.* $U > 0$ vs. $U = 0$; Fig. 4a). Interestingly, in self-fertilising populations, H_e ~~, depending on the strength of POD . When fewer loci (n) contribute to a POD and linkage is loose (larger ℓ), recurrent deleterious mutations speed the decay of heterozygosity . For more stable POD 's, however, background selection has little effect. In fact, when selfing and mutation rates are high (~~within the POD zone rises when background selection occurs elsewhere in the genome. These effects increase when mutation rates rise (green *vs.* blue lines, $U = 0.5$ ~~), POD heterozygosity actually increases (Fig. ??). This effect is more pronounced in earlier generations following POD formation via admixture (see results after 100 and 1000 generations in Supp. Fig. ??). Conversely, POD selection increases~~ and 0.1 respectively) and linkage increases (full *vs.* dashed lines reflecting map lengths of $R = 1$ and 10 Morgans respectively).

Similarly, the presence of a stable POD zone affects the heterozygosity of deleterious mutations observed elsewhere in the genome~~(see Supp. Fig. ??)suggesting these two forms of selection reinforce each other. That is, POD 's appear to decrease the efficiency of purging deleterious mutations throughout the genome while selection against mutations segregating outside the POD zone appear to strengthen the balancing selection maintaining heterozygosity by further disfavoring homozygotes.~~
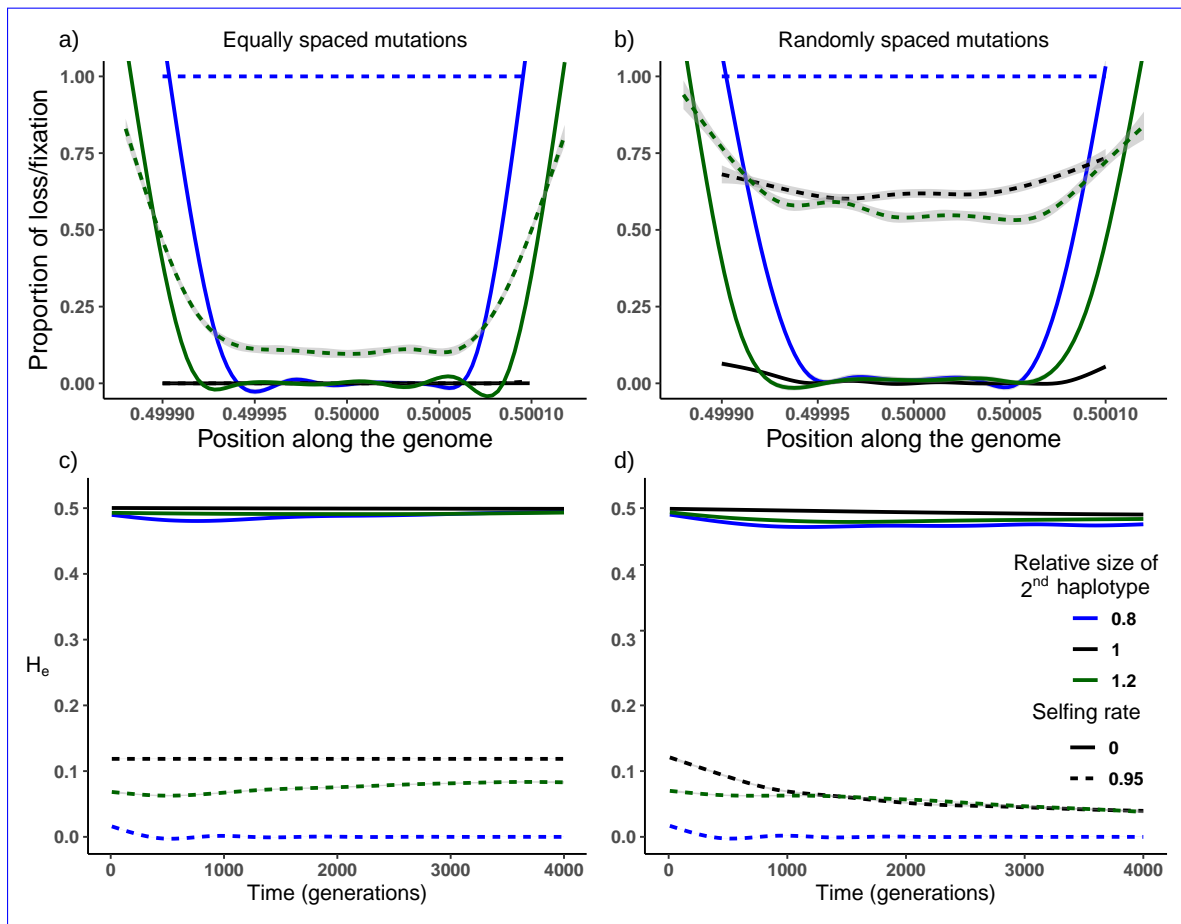
Figure 3: Effects of relaxing the assumptions of symmetric overdominance and evenly spaced mutations. Upper panels show locations within the POD zone where load mutations are most likely to be lost (a, b) and how this depends on whether mutations are evenly spaced (a) or randomly distributed (b). Results are shown for both symmetric (black) and asymmetric (green and blue) loads. Outcomes under both outcrossing and high selfing (solid vs. dotted lines) are shown. Note erosion of mutations via recombination and selection at both ends of the POD zone. Lower panels show overall stability of the POD zone (shown as heterozygosity, He) over time. As in the upper panels, graphs show results for both symmetric (black) and asymmetric (green and blue) loads and for evenly and randomly placed mutations (panel c vs. d). The coefficients of selection and dominance are $s = 0.01$ and $h = 0.2$ respectively, linkage within the POD zone is $\ell = 10^{-6}$ and population size $N = 1000$.

~~Background selection acting at other loci~~. When mutation rates are low ($U = 0.1$), POD selection slightly decreases the mutational heterozygosity elsewhere in the genome ~~thus affect POD dynamics in two distinct ways. Weaker POD's (small n and s_H) with loose linkage (larger ℓ)are vulnerable to the effects of recurrent deleterious mutations elsewhere in the genome, speeding the decay of POD heterozygosity~~ (blue lines Fig. 4b). Conversely, a higher genomic mutation rate ($U = 0.5$, green lines) results in increased heterozygosity, especially in highly selfing populations with small genomes (implying tight linkage - solid green line in Fig. 4b). Effects of POD selection on effective population size are complex but in most cases, POD selection tends to decrease N_e (Supp. Fig. ~~??). This ressembles classical effects~~ S6).

To confirm that these effects derive from overdominance rather than some other effect of background selection~~on neutral variation (Fig. ??a). In contrast, stronger, more stable POD's in predominantly outcrossing populations are barely affected by strong background selection~~(~~i.e.~~ for $U = 0.5$, Fig. ~~??b). Yet strong background selection in mostly selfing populations strengthens the POD by favoring heterozygotes. This effect is evident early on, after the admixture event (see~~, we simulated effects of co-dominant background mutations ($h_d = 0.5$). Because such mutations are expressed in heterozygotes and thus easily removed by selection, they generate few associations with other loci. Co-dominant background mutations have little effect on within-POD zone heterozygosity in contrast to simulations with more recessive mutations ($h_d = 0.2$). This is true even within selfing populations (Supp. Fig. ~~?? a~~and b at 100 generations, and c and d at 1000 generations). These two forms of selection mutually reinforce each other as POD selection increases the heterozygosity of deleterious mutations elsewhere in the genome (SuppFigS7a). ~~??). This reduces the efficiency with which selection acts against segregating deleterious mutations in other regions (by reducing the frequency of homozygotes) while selection acting at those load loci either barely affect the POD or strengthen the balancing selection maintaining the POD. Although~~
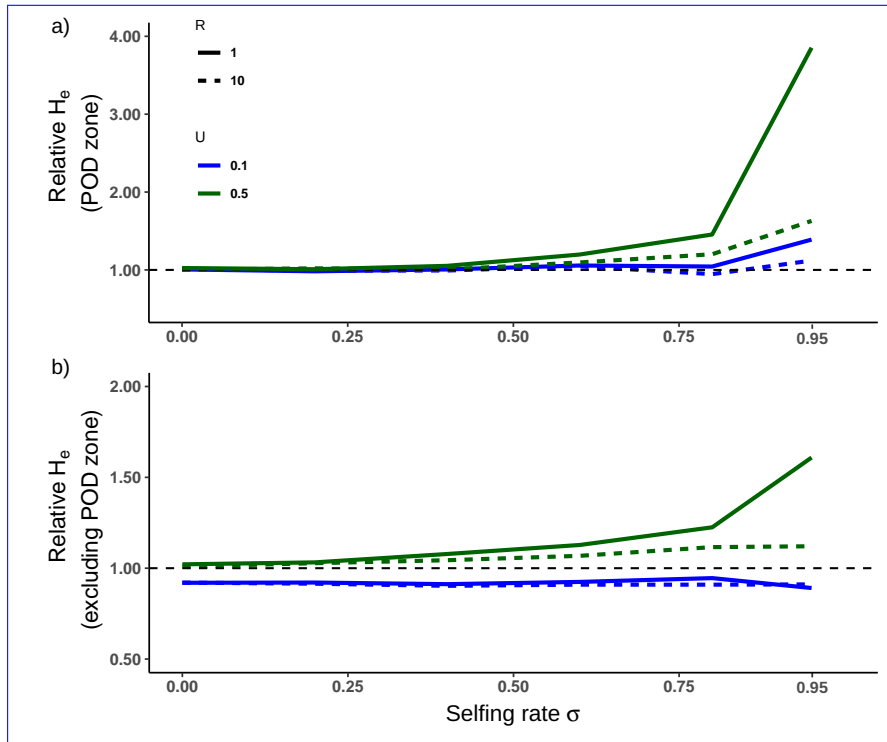
Figure 4: Background mutations affect POD selection and vice versa. Graph (a) shows heterozygosity, H_e, within the POD zone with background mutations relative to H_e in the absence of background mutations and graph (b) H_e elsewhere in the genome with a POD zone relative to without, both as a function of the selfing rate. Populations are subject to different background mutation rates (U) and shorter and longer map lengths (R in Morgans). These simulations use 100 POD load loci ($n_L = 100$) and a map length of $\ell = 10^{-6}$ Morgans. Mutations within the POD zone are randomly placed. Selection coefficients in- and outside the POD zone (s and s_d respectively) are 0.01 with dominances h and $h_d = 0.2$.

background selection amplifies the selection maintaining the POD (especially in highly selfing populations), it has little effect on fixation dynamics within the POD This confirms that it is associative overdominance between the POD zone and other load loci that increases heterozygosity (Supp. Fig. S7b). Varying rates of background mutation and POD zone length also have complex effects on effective population size N_e (Supp. Fig. ??S7c).

Opposing effects of background selection on heterozygosity dynamics in the presence and absence of a POD. Observed heterozygosity (H_e) after 4000 generation as a function of the rate of self-fertilization (σ) in the population. Total genome size R is 10 cM. The POD initially contains n = 100 hetrozygous mutations per haplotype. a) H_e decay for loci lacking selection (control, $s = 0$). b) Results for loci subject to selection ($s = 0.01$, $h = 0.2$). Background mutation rates are $U = 0$, 0.1 and 0.5 mutations per haploid genome per generation, and all new mutations are under selection with $s_d = 0.01$ and $h_d = 0.2$. Linkage between trans-mutations $\ell = 5.10^{-4}$ cM and population size is $N = 1000$.

4.3 Inbreeding depression

As expected, the overdominance generated by a POD increases the amount of in a POD zone increases the inbreeding depression, δ, expressed in a population (compare lower lines in Fig. ??a and ??b). However, it also reduces the decline in populations express (Supp. Fig. S8). Observed δ with σ relative to unlinked expectations. Background selection ($U > 0$), again as expected, accentuates inbreeding depression in highly selfing populations relative to classic expectations especially soon after admixture, with or without a POD (Supp. Fig. ??). Fixation of the load via drift and purging of the load via purifying selection both reduce levels of inbreeding depression. In opposition to these, both background selection and the balancing selection associated with a POD instead sustain or increase inbreeding depression, even at high selfing rates (right side, Fig. ?? and in outcrossing populations can be predicted using Eq. (9), which accounts for overdominant selection and unlinked

deleterious mutations. In selfing populations variable erosion of the POD zone and POD selection dynamics generate bimodal distributions of δ (see Supp. Fig. ~~??). The marked departure of simulation results from theoretical expectations (Eq. 9) reflects how assuming independence among load loci actively misleads us especially at high selfing rates (where strong identity and linkage disequilibria occur). Distributions of observed values of δ in the simulations also become highly variable as background selection increases (Fig. ??c) . This presumably reflects the cumulative effects of random mutations accruing over 4000 generations. Notably, the presence of a POD accentuates this variability (Fig. ??d~~S9 for clearer representations). Some simulations generate values of δ close to those predicted by Eq. (9) (dashed lines in Fig. 5) while others generate values predicted when selection acts only against the unlinked recessive deleterious mutations (Eq. (6), dotted lines in Fig. 5). This ~~underlines how variable POD erosion dynamics are, even to the point of generating bimodal distributions of~~ may reflect loss of the POD zone. Smaller genomes (*e.g.*, $R = 1$ Morgans) generally increase the observed δ ~~(Fig. ??d~~, especially in selfing populations (see Supp. Figs. S8 and A4).

5 Discussion

~~We evaluated whether blocks of recessive mutations expressing~~

Given that purging, drift, and background selection all reduce segregating variation and thus inbreeding depression, we face the question of what force perpetuates these even within small and inbred populations. Waller (2021) emphasized this enigma and reviewed mechanisms that might account for it. Selective interference among loci might act to slow or block purging (Lande and Schemske, 1985a; Winn et al, 2011). Recurrent mutations might also replenish the load fast enough to regenerate δ (Fisher, 1930; Charlesworth, 2018). A third possibility is that clusters of recessive mutations linked in repulsion emerge, creating enough balancing selection via pseudo-overdominance ~~(POD's) can persist long enough~~
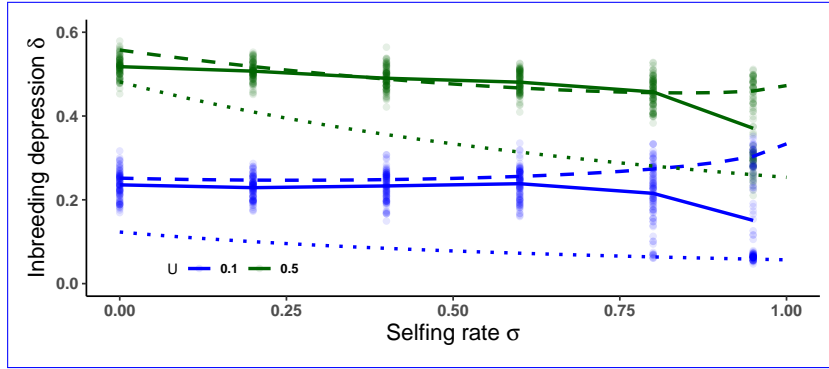
Figure 5: Inbreeding depression ~~(δ)~~ as a function of the selfing rate ~~(σ) 4000 generations after admixture~~ for ~~simulations without (s = 0, a and c) and with POD selection (s = 0.01, b and d) for~~ different ~~background~~ values of the haploid mutation ~~rates~~ rate, U. ~~Top panels: δ as a continuous function of σ. The dots represent δ for individual simulation runs while full~~ Solid lines show ~~mean observed δ and~~ means of the ~~dashed~~ simulations run. Dotted lines show ~~theoretical expectations for~~ the inbreeding depression ~~calculated using equation 9. Bottom panels: Violin plots of~~ expected in the ~~distributions~~ absence of ~~δ for different~~ overdominance (Eq. (6)) while dashed lines show increases in $delta$ expected with overdominant selection over all selfing rates ~~of σ~~(Eq. (9)). Other parameter values are ~~as~~ $n_L = 100$, $\ell = 10^{-6}$ Morgans and randomly placed mutations in ~~Fig~~the POD zone. ~~??~~Selection coefficients in- and outside the POD zone, s and s_d are set to 0.01 with dominances h and $h_d = 0.2$. The total map length (setting the recombination rate) is $R = 10$ Morgans.

~~to have significant evolutionary effects. If so, they might help to account for previously enigmatic results like the persistence of genetic diversity and inbreeding depression often observed in small or inbreeding populations. In our models, POD's did persist when linkage was tight and cumulative selective effects from both haplotypes were strong and evenly balanced. Tight linkage, in particular, creates genomic regions where cumulative pseudo-overdominance creates balancing selection strong enough to sustain POD's for long periods. When recombination occurs frequently, selection and drift become more effective, breaking up POD's by eroding mutations from their ends or by unbalancing the load. These forces then work as in classic models to purge and fix partial POD haplotypes and the mutations they carry. Whether POD 's exist in sufficient number and power to affect evolutionary dynamics thus hinges on the relative rates at which they are created and destroyed. Although we did not examine POD origins directly, our models demonstrated that POD's subject to~~ (POD) to counter purging and drift, sustaining selection for outcrossing or mixed mating systems (Waller, 2021). Our goals here were to explore the dynamic stability of POD zones (initially ignoring how they arise) using both classical one-locus overdominant theory (Kimura and Ohta, 1971) and simulations. We found that ~~strong~~ and balanced ~~selection~~ POD zones can persist for ~~long periods given tight linkage. This makes them a potentially important "player" in the evolutionary dynamics of the genetic load, both within the POD's and elsewhere in the genome (e.g., by affecting background selection and reducing homozygosity and purging, as shown in our results). Such effects are important as models that ignore linkage and genic associations predict the rapid loss of genetic variability and inbreeding depression via purging and fixation.~~ hundreds to many thousands of generations.

5.1 ~~Assumptions and structure of the model~~

Our model envisioned that isolated populations, fixed for alternative sets of recessive deleterious mutations, fused to create hybrid populations expressing heterosis (from high levels of pseudo-overdominance) . Populations regularly become isolated, diverge genetically, and then hybridize. If these populations were small or inbred, they will have fixed many mildly deleterious mutationsas selective effects of order 1Whether POD zones are fragile or robust depends critically on several genetic parameters. These include the number and severity of deleterious mutations, their proximity and cis-/2Ne or less are effectively neutral (Crow and Kimura, 1970). Inter-population crosses reveal that this fixed, or drift, load can be appreciable (Willi et al, 2013; Spigler et al, 2017). The same phenomena are observed in plant and animal breeding programs where inbred lines that may be low in fitness are hybridized to gain high fitness. This heterosis was once thought to reflect true overdominance but finer genetic characterizations reveal this to be mostly pseudo-overdominance (Crow, 1999b). Thus, our scenario for POD origins may be plausible.

The model developed by Harkness et al (2019) suggests that heterosis from between-line crosses may be short-lived. This could reflect how loads are often unbalanced, allowing rapid purging and fixation (and the former accelerate the latter via background selection). Nevertheless, particular chromosome segments in maize have unexpectedly high heterozygosity, suggesting the action of pseudo-overdominance (Brandenburg et al, 2017). The rapid and more complete decay of heterozygosity in Harkness et al's (2019) model may reflect their assumption of uniform rates of recombination along chromosomes. Our model shows this could limit POD emergence.

Strongly overdominant loci (Kimura and Ohta, 1971) persist for remarkably long periods (e. g., self-incompatibilty loci in plants and histocompatibility loci in animals)but are considered rare. For sets of recessive mutations to generate similarly strong and persistent pseudo-overdominance, they must be closely linked. Clusters of such mutations could occur by chance within small genomic regions (as assumed in our model) given mildly deleterious

mutations are common and broadly dispersed through the genome . If their cumulative selective effects were strong and balanced across haplotypes, a POD would result. Further load mutations in or near these regions would be sheltered from selection, allowing them to accumulate and reinforce the overdominance as been observed for the *Arabidopsis halleri* S-locus (Llaurens et al, 2009).

Our idealized model further assumed that mild mutations were evenly distributed along the incipient POD in repulsion to alternating mutations on the opposite chromosome. This maximized pseudo-overdominance and made it difficult to unbalance the load as internal recombinants expressed the same load (Fig. 1). Simulations that more realistically assumed randomly positioned load loci within the POD resulted in shorter-lived POD's (Supp. Figs. S10 trans- positions, and S11 vs. Fig. ??their levels of dominance/recessivity (Figs. 2 and A4). Strong and balanced selection plus tight linkage allow POD zones to persist as these conditions enhance the associations (linkage disequilibria) that generate POD effects. Recombination dissolves these associations, allowing purifying selection and drift to disrupt POD zones, purging and fixing mutations. Mutations erode from either end of the POD zone or the load becomes unbalanced enough to fix one haplotype. The importance of linkage and small mutational effects are evident in the radically enhanced purging seen in models that ignore linkage and assume major mutational effects (Lande and Schemske, 1985b). We also found that new recessive mutations that occur elsewhere in the genome generate associations with load alleles within POD zones that enhance POD zone heterozygosity and persistence (Fig. 4). Such mutations add to the segregating load, increasing heterozygote advantage. Because levels of heterozygosity are correlated across the genome in partially inbred populations (identity disequilibrium), the background selection generated by mutations outside the POD zone tend to reinforce the balancing selection favoring heterozygotes in the POD zone. POD zones also exert reciprocal effects, enhancing the heterozygosity of mutations occurring elsewhere in the genome when mutation rates are moderate (U=0.5, Fig.

4b). This ~~faster loss reflects how recombination and selection more easily unbalance the load between haplotypes when it is unevenly distributed. The linkage required for longer stability became an order of magnitude higher (*i.e.*, a map distance of $\ell = 10^{-5}$ cM) than when simulations assumed alternating load loci~~effect was amplified within selfing populations, presumably reflecting how selection against POD zone homozygotes favors heterozygosity across the genome when more identity disequilibrium occurs. These effects would be further enhanced if mutations were to have varying dominance effects, a scenario which we did not consider here. However, ~~this higher linkage resulted in more stability allowing fewer load loci (*Supp. Fig. S10c*). Recombination became so rare that linked load loci acted like a single overdominant locus. In our model, POD's needed more load loci to generate the strong balancing selection needed to stabilize the POD~~ recent work has shown that POD selection can be generated in a single population by the clustering of mutations in repulsion, even without heterogenous recombination rates along the chromosome (Sianta et al, 2021). These results coupled with ours lead us to hypothesize that any genomic region displaying reduced recombination could provide a haven for POD zones to emerge and persist.

~~Additional modeling could add more realistic variation not only in the number and spacing of mutations but also their effects and dominance levels. More variable mutational effects and spacing would likely enhance the power of selection and drift to eliminate POD's by enhancing the power of recombination to disrupt the balanced selection maintaining POD's (rather than chewing them away at their ends). Our models~~

5.1 How do POD zones originate?

Many empirical observations could be explained by the existence of POD zones (see Introduction and Waller 2021). Whether POD zones that are conserved across populations exist in sufficient number and strength to affect evolutionary dynamics hinges on the relative rates at which they are created and destroyed. We focused on POD zone erosion and loss, ~~but~~

~~understanding which mechanisms facilitate their formation would also help us understand how often POD's evolve. Notably,~~ not how they arise. As our results show, a requirement for POD stability is strong linkage within a given genomic region in which mutations can accumulate through the actions of selection and genetic drift. Inversions and centromeric regions with restricted recombination provide preconditions favoring POD zone emergence, as do genomic regions neighbouring loci currently or previously under overdominant selection, where recombination is suppressed. Examples where this has been observed include self-incompatibility loci (Takebayashi, 2003; Igic et al, 2008; Mable, 2008), MHC loci (Garrigan and Hedrick, 2003; Gemmell and S and loci with balanced polymorphisms generated by ecological selection (van Oosterhout et al, 2000; Jay et al, 2 In such regions, mutations of small effect become effectively neutral when the product of the effective population size and the selection coefficient $N_e s \ll 1$ (Crow and Kimura, 1970; Hedrick et al, 2016)). These will drift in frequency and often fix increasing the "drift load" to the ~~presence of true (one locus) overdominance and inversions could favor POD 's, extend their persistence, and thus enhance their evolutionary effects. It might also be worthwhile to explore whether modifier loci could evolve to reduce rates of recombination, particularly near the center of POD's, providing another POD-stabilizing mechanism.~~

5.2 ~~Conditions favoring POD formation and persistence~~

~~The strong selection and tight linkage needed to sustain POD's in our models suggest that such conditions should enhance POD formation. Crosses between small or inbred populations with independent drift loads would generate highly heterozygous individuals and large segregating loads in the first generation. Subsequent recombination and selection, however, likely eliminates most proto-POD's. Any POD strong enough to persist, however, would reduce the effective rate of inbreeding ($\hat{F}$, Eq. 7) limiting the rate at which deleterious recessive mutations purge. Fixation at loci outside the POD would continue in~~ point where it may compromise population viability (Whitlock et al, 2000; Charlesworth, 2018). Selection

against strongly deleterious mutations will accentuate fixation of milder mutations linked in repulsion via "background selection" (Charlesworth et al, 1997; Zhao and Charlesworth, 2016). Pairwise and higher associations (linkage disequilibria) also increase within small and inbred populations . We might therefore expect POD 's to emerge recurrently in situations where populations regularly become small, isolated , and subject to fixation, then hybridize, regenerating strong heterosis . Truly overdominant loci (like the S- and HMC loci mentioned above) might also favor POD formation. even among alleles at unlinked loci limiting selection (Hill and Robertson, 1966; Sved, 1971; Ohta and Cockerham, 1974; Lewontin, 1974).

Conditions generating tight linkage should likewise favor POD formation as high linkage can compensate for less balanced and lower loads. Chromosomal inversions that block recombination, in particular, shelter deleterious recessive mutations from selection as heterozygotes, enhancing opportunities for pseudo-overdominant effects to evolve. Kirkpatrick (2010) observes that the genetic basis for inversion overdominance has not been determined but notes that pseudo-overdominance is plausible. Jay et al (2021) dissected the structure and genealogy of inversions contributing to heterosis in *Heliconius* butterflies. These date back to hybridization events as long as 3M years ago suggesting that high heterosis following hybridization may have spawned POD's. Interestingly, subsequent mutations (primarily transposable elements)increased the size of these inversions and their load, reinforcing the pseudo-overdominance and inbreeding depression. This supports our conjecture as do recent simulations (Berdan et al, 2021) showing recurrent mutations can strengthen POD's.

5.2   Empirical evidence on POD's

As noted above, balanced polymorphisms expressing overdominance exist in many species and may stably persist for long periods. These are sometimes associated with local adaptation in plants like the five ancient polymorphisms found by Wu et al (2017) that predate the divergence of *Arabidopsis* from *Capsella*. Adaptive or inversion polymorphisms represent an

37

~~alternative route to POD formation from the hybridization model outlined here. Inversions may not be associated with adaptive differences, however.~~

~~Trans-specific inversions, polymorphisms, and ancient S- and HMC loci all support the notion that overdominance can sustain polymorphisms for millions of years. Some researchers have found genomic regions supporting high heterozygosity and segregating variation suggestive of POD's. These exist in *Zea mays* (McMullen et al, 2009) and 22 low-recombination regions within~~ The scenario we suggested that might create POD zones involved drift fixing alternative sets of recessive deleterious mutations among isolated populations. When such populations hybridize, their F1 progeny experience high heterosis reflecting the cumulative effects of POD across the whole genome (Crow, 1999a). Under free recombination, this heterosis is expected to erode by 50% in the ~~human genome near centromeres (Gilbert et al, 2020). Kremling et al (2018) confirmed that many rare variants in maize have deleterious effects confirming that "even intensive artificial selection is insufficient to purge genetic load." Brandenburg et al (2017) identified 6~~F2 and each subsequent generation as recombination dissipates the associations generating the POD (Harkness et al, 2019) (ignoring the presense of epistatic Dobzhansky-Muller incompatibilities -(Ehiobu et al, 1989). However, where clumps of mutations occur within short genomic regions (or in low recombination zones), ~~978 genomic segments with unexpectedly high heterozygosity in land races of maize. Heterozygous segments were more likely to harbor deleterious mutations than the rest of the genome and several were deeply conserved across multiple land races. Brown and Kelly (2020) found an uneven distribution of load loci in~~ POD zones may be spawned. Inter-population crosses often reveal high heterosis (Willi et al, 2013; Spigler et al, 2017) as do crosses between low-fitness inbred lines in plant and animal breeding programs. Theory suggests that any incipient POD zone generating heterozygous progeny at least twice as fit as homozygous progeny will allow that POD zone to persist even in highly selfing populations. Dramatic examples of "hybrid vigor" in F1 crosses include cases where progeny have up to 35 times the fitness of

parental lineages (Tallmon et al, 2004; Hedrick and Garcia-Dorado, 2016) easily satisfying this condition.

Proto-POD zones may be fragile. Our models show that recombination and selection eliminate proto-POD zones with weak, unbalanced, or loosely linked loads. However, in some regions, cumulative selective effects from localized mutations may be large and balanced enough to allow a persistent POD zone to emerge. Such zones eliminate many homozygous progeny, reducing effective rates of inbreeding ($\hat{F}$, Eq. 7). This, in turn, reduces rates at which deleterious recessive mutations are lost both within POD zones and elsewhere in the genome (Fig. 4). Selection against low-fitness recombinants might even favor the evolution of reduced rates of recombination within POD zones providing another mechanism to stabilize POD zones (cf. Olito et al 2022). We ignore the potential of POD zones to gain strength over time by accumulating additional internal mutations sheltered from selection as heterozygotes, which would augment the overdominance as observed at the S-locus in ~~*Mimulus guttatus*~~ *Arabidopsis halleri* – ~~with deleterious alleles often localized to particular regions. These "Rare Allele Clusters" predicted fitness. Waller (2021) reviews further evidence of possible POD signatures.~~ (Llaurens et al, 2009)).

5.2 ~~ImplicationsPOD 's that arise and persist in natural populations would have several evolutionarily significant effects. Most immediately, POD's act to reinforce the persistence of inversions by increasing selection against recombinant homozygotes and enhance levels of molecular variation by sustaining deleterious recessive mutations that would otherwise be lost to selection and/or drift. Higher levels of genetic variability are usually interpreted as favorable for adaptation but POD loads are maladaptive in increasing expression of~~ Evolutionary consequences of POD selection

POD zones could affect the architecture and the dynamics of the genetic load in various ways. Most conspicuously, our simulations of background selection show how POD zones

could increase the segregational load ~~and reducing selection against newly arising deleterious recessive mutations. These effects could compromise both short- and longer-term population viability (particularly in small, isolated, and /or inbred populations that produce many homozygous progeny). They would also weaken the usually positive correlations between heterozygosity and fitness (expected when heterozygosity provides an inverse measure of individual inbreeding). The ability of POD 's to shelter and potentially increase the load might lead to progressive increases in POD size and deleterious effects with outcomes similar to~~ elsewhere in the genome and vice versa. Such effects imply that mutations both within and outside the POD zone could reinforce the selection maintaining POD zones sustaining more variability and segregating loads than otherwise expected. Such loads could favor self-incompatibility mechanisms for their ability to produce fewer low-fitness homozygous genotypes. Our scenario where population hybridization spawns POD zones suggests a mechanism whereby fixed drift loads might regularly be converted into segregating loads which then persist in regions expressing strong overdominance.

Although we expect positive heterozygosity-fitness correlations within partially inbred populations (given that heterozygosity inversely measures inbreeding), heterozygosity and variation within POD zones reflects the opposite: non-adaptive variation emerging from sustained mutational and segregational genetic loads. This may help to explain why heterozygosity-fitness correlations can be weak and inconsistent (David, 1998). POD zones might increase loads within populations by creating safe havens within which new deleterious mutations could accumulate while increasing the load of mutations segregating elsewhere in the genome. Small, inbred populations might also become vulnerable to "mutational meltdown" ~~models (Gabriel et al, 1993). Such effects might favor modifier loci that increased rates of recombination to prevent the genome from 'congealing' (Maynard Smith, 1977)~~threatening population viability(Gabriel et al, 1 Conversely, POD zones may provide individual or population advantages by sustaining inbreeding depression and favoring outcrossing in ways that better sustain adaptive genetic

variability.

POD's would also affect

5.3   POD effects on mating system evolution

The presence of POD conspicuously affects the evolution of plant and animal mating systems by sustaining segregating deleterious mutations and the inbreeding depression they generatemore segregational load and higher inbreeding depression than expected especially in small, inbred populations. Early models of mating system evolution sought to explain variable levels of self-fertilization as the outcome of selection acting on progeny expressing higher or lower levels of equilibria reflecting how selection acted on progeny with more or less inbreeding depression. These predicted that increased rates of selfing should expose deleterious mutations to selection, reducing their frequency and thus inbreeding depression. As mutationswere purged, ID would decline leading to In these simple static models, inbreeding depression less than 0.5 would result in exclusive selfing while higher levels would favor exclusive outcrossing. More dynamic simple models that allow selection make mixed mating systems even more improbable by allowing inbreeding to purge deleterious mutations, generating "run-awayselection for complete" selection for ever-increasing levels of selfing (Lande and Schemske, 1985b). Such models, however, neglected to account for If drift instead fixes many segregating mutations, similar effects emerge as this, too, causes inbreeding depression to decline. The ability of many small, inbred populations to nevertheless retain genetic variation and inbreeding depression plus the absence of purely inbreeding taxa thus pose a paradox (Byers and Waller, 1999; Winn et al, 2011). More complex and realistic models that incorporate effects of linkage, driftdynamics, and the associations among loci that arise in smalland inbred populations, all of which considerably modify classical predictions (Uyenoyama et al, 1993). More complete models display, inbred populations show far more complex dynamics and outcomes including the persistence of

41

mixed mating systems in some cases. Nevertheless, we generally expect both purging and fixation due to drift to reduce segregating variation and thus inbreeding depression in smaller and more inbred populations. The abililty of many small inbred populations to retain genetic variation and inbreeding depression has puzzled many as has the absense of purely inbreeding taxa (Byers and Waller, 1999; Winn et al, 2011). The chance presence and variable sizes and strengths of POD's among these populations might (Charlesworth and Charlesworth, 1987; Uyenoyama et One relevant model showed that a single unlinked overdominant viability locus anywhere in the genome generates positive associations with modifier alleles enhancing outcrossing (Uyenoyama and Waller, 1991). Such associations favor a persistently mixed mating system. Because POD also favors heterozygotes, we expect POD zones to exert similar effects. The presence of POD zones might thus help to account for these paradoxes .

Just how POD's influence mating system evolution remains an open question. Harkness et al (2019) suggest that the short-lived heterosis observed in hybridizing populations should result in, at most, a temporary selective advantage to outcrossing but heterosis due to POD's could also favor outcrossing. If initial levels of pseudo-overdominance decline as recombination breaks up POD haplotypes, selfing could again be favored. Persistent POD's might provide more sustained selection favoring outcrossing. Our models suggest that mutations causing heterosis are often fixed in contrast to Harkness et al (2019). This might allow recurrent hybridization events to reinitiate POD formation, creating recurrent selection that could favor outcrossing in species with the paradoxes of persistent segregating loads and populations and species that maintain mixed mating systems. If, instead, POD zones regularly arise and then deteriorate, selection could alternately favor selfing and outcrossing. This might provide an entirely different mechanism favoring mixed mating systems.

How often POD 's

5.4   Conclusions

42

Understanding the mechanisms that create and sustain POD zones cast light on how commonly POD zones may arise and persist ~~, how commonly they occur within natural populations, and how restricted they are to genomic regions with tight linkage all deserve more attention. Do most POD's occur near centromeres or within inversions? Do POD's arise and dissipate regularly following recurrent hybridization? Further comparative genomic data should cast light on all these questions~~ and the genetic and demographic circumstances that enhance their longevity. Comparative genomic data will be particularly useful for searching for POD zones and analyzing their structure and history. Our models demonstrate how several genetic, demographic, and mating system parameters may affect load dynamics within and beyond POD zones. Any POD zones that persist are likely to strongly affect mating system evolution by reducing both purifying selection and drift, sapping the power these forces would otherwise have to reduce inbreeding depression. Our models demonstrate that POD zones can persist given the right conditions. We encourage further research to extend and refine our understanding of this phenomenon.

Bibliography

Abu Awad D, Waller DM (2022) POD selection - individual based model. DOI 10.5281/ZENODO.6874327, URL https://zenodo.org/record/6874327

Agrawal AF, Whitlock MC (2011) Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics 187:553–566, DOI 10.1534/genetics.110.124560

Baldwin SJ, Schoen DJ (2019) Inbreeding depression is difficult to purge in self-incompatible populations of *Leavenworthia alabamica*. New Phytologist 224(3):1330–1338, DOI 10.1111/nph.15963, URL https://onlinelibrary.wiley.com/doi/10.1111/nph.15963

Bataillon T, Kirkpatrick M (2000) Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. Ge-

netical Research 75(1):75–81, DOI 10.1017/S0016672399004048, URL https://www.cambridge.org/core/product/identifier/S0016672399004048/type/journal$_a$rticle

Becher H, Jackson BC, Charlesworth B (2020) Patterns of genetic variability in genomic regions with low rates of recombination. Current Biology 30(1):94–100.e3, DOI 10.1016/j.cub.2019.10.047, URL https://www.cell.com/current-biology/abstract/S0960-9822(19)31387-9

Berdan EL, Blanckaert A, Butlin RK, Bank C (2021) Deleterious mutation accumulation and the long-term fate of chromosomal inversions. PLoS Genetics 17(3):e1009,411, DOI 10.1371/journal.pgen.1009411, URL https://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1009411

Bernstein MR, Zdraljevic S, Andersen EC, Rockman MV (2019) Tightly linked antagonistic-effect loci underlie polygenic phenotypic variation in *C. elegans*. Evolution Letters 3(5):462–473, DOI 10.1002/evl3.139, URL https://onlinelibrary.wiley.com/doi/10.1002/evl3.139

Brandenburg JT, Mary-Huard T, Rigaill G, Hearne SJ, Corti H, Joets J, Vitte C, Charcosset A, Nicolas SD, Tenaillon MI (2017) Independent introductions and admixtures have contributed to adaptation of european maize and its american counterparts. PLoS Genetics 13(3), DOI 10.1371/journal.pgen.1006666

Brown KE, Kelly JK (2020) Severe inbreeding depression is predicted by the "rare allele load" in *Mimulus guttatus*. Evolution 74(3):587–596, DOI 10.1111/evo.13876, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/evo.13876

Byers D, Waller D (1999) Do plant populations purge their genetic load? effects of population size and mating history on inbreeding depression. Annual Review of Ecology and Systematics 30:479–513, DOI 10.1146/annurev.ecolsys.30.1.479

Charlesworth B (2018) Mutational load, inbreeding depression and heterosis in subdivided populations. Molecular Ecology 27(24):4991–5003, DOI 10.1111/mec.14933

Charlesworth B, Charlesworth D (1998) Some evolutionary consequences of deleterious mutations. Genetica 102/103:3–19, DOI 10.1023/A:1017066304739, URL http://link.springer.com/10.1023/A:1017066304739

Charlesworth B, Nordborg M, Charlesworth D (1997) The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genetics Research 70(2):155–174, DOI 10.1017/S0016672397002954, URL https://www.cambridge.org/core/journals/genetics-research/article/effects-of-local-selection-balanced-polymorphism-and-background-selection-on-equilibrium-patterns-of-genetic-diversity-in-subdivided-populations/1F946DDC83F5277DC964D76329A6BB66

Charlesworth D, Charlesworth B (1987) Inbreeding depression and its evolutionary consequences. Annual Review of Ecology, Evolution, and Systematics 18:237–268

Charlesworth D, Charlesworth B (1990) Inbreeding Depression with Heterozygote Advantage and Its Effect on Selection for Modifiers Changing the Outcrossing Rate. Evolution 44(4):870–888, DOI 10.1111/j.1558-5646.1990.tb03811.x, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1990.tb03811.x

Charlesworth D, Willis JH (2009) Fundamental concepts in genetics the genetics of inbreeding depression. Nature Reviews Genetics 10(11):783–796, DOI 10.1038/nrg2664

Chelo IM, Afonso B, Carvalho S, Theologidis I, Goy C, Pino-Querido A, Proulx SR, Teotónio H (2019) Partial Selfing Can Reduce Genetic Loads While Maintaining Diversity During Experimental Evolution. G3

Genes|Genomes|Genetics 9(9):2811–2821, DOI 10.1534/g3.119.400239, URL https://academic.oup.com/g3journal/article/9/9/2811/6026387

Crow JF (1993) Mutation, mean fitness, and genetic load. Oxford Series in Evol Biol 9:3–42

Crow JF (1999a) The rise and fall of overdominance. Plant Breeding Reviews 17:225–257

Crow JF (1999b) The Rise and Fall of Overdominance. In: Plant Breeding Reviews, John Wiley & Sons, Ltd, pp 225–257, DOI 10.1002/9780470650134.ch5, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470650134.ch5

Crow JF, Kimura M (1970) An Introduction to Population Genetics Theory. Burgess Pub. Co., 591 pp.

Darwin C (1876) The effects of cross and self fertilization in the vegetable kingdom. J. Murray and Co.

David P (1998) Heterozygosity–fitness correlations: new perspectives on old problems. Heredity 80(5):531–537, DOI 10.1046/j.1365-2540.1998.00393.x, URL http://www.nature.com/doifinder/10.1046/j.1365-2540.1998.00393.x

Ehiobu NG, Goddard ME, Taylor JF (1989) Effect of rate of inbreeding on inbreeding depression in Drosophila melanogaster. Theoretical and Applied Genetics 77(1):123–127, DOI 10.1007/BF00292326, URL http://link.springer.com/10.1007/BF00292326

Fisher S Ronald A (1930) The Genetical Theory of Natural Selection. Oxford University Press, Oxford

Gabriel W, Lynch M, Bürger R (1993) Muller's Ratchet and Mutational Meltdowns. Evolution 47(6):1744–1757, DOI 10.1111/j.1558-5646.1993.tb01266.x, URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1993.tb01266.x

Garrigan D, Hedrick PW (2003) Perspective: Detecting Adaptive Molecular Polymorphism: Lessons from the MHC. Evolution 57(8):1707–1722, DOI 10.1111/j.0014-3820.2003.tb00580.x, URL https://onlinelibrary.wiley.com/doi/10.1111/j.0014-3820.2003.tb00580.x

Gemmell NJ, Slate J (2006) Heterozygote Advantage for Fecundity. PLoS ONE 1(1):e125, DOI 10.1371/journal.pone.0000125, URL https://dx.plos.org/10.1371/journal.pone.0000125

Gilbert KJ, Pouyet F, Excoffier L, Peischl S (2020) Transition from Background Selection to Associative Overdominance Promotes Diversity in Regions of Low Recombination. Current Biology 30(1):101–107.e3, DOI 10.1016/j.cub.2019.11.063, URL https://www.cell.com/current-biology/abstract/S0960-9822(19)31565-9

Glémin S (2021) Balancing selection in self-fertilizing populations. Evolution 75(5):1011–1029, DOI 10.1111/evo.14194, URL https://onlinelibrary.wiley.com/doi/10.1111/evo.14194

Harkness A, Brandvain Y, Goldberg EE (2019) The evolutionary response of mating system to heterosis. Journal of Evolutionary Biology 32(5):476–490, DOI 10.1111/jeb.13430, URL https://onlinelibrary.wiley.com/doi/10.1111/jeb.13430

Hedrick PW, Garcia-Dorado A (2016) Understanding Inbreeding Depression, Purging, and Genetic Rescue. Trends in Ecology & Evolution 31(12):940–952, DOI 10.1016/j.tree.2016.09.005, URL https://linkinghub.elsevier.com/retrieve/pii/S0169534716301586

Hedrick PW, Hellsten U, Grattapaglia D (2016) Examining the cause of high inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of eucalyptus grandis. New Phytologist 209(2):600–611, DOI 10.1111/nph.13639

Hill WG, Robertson A (1966) The effect of linkage on limits to artificial selection. Genetical Research 89(5-6):311–336, DOI 10.1017/S001667230800949X

Igic B, Lande R, Kohn J (2008) Loss of self-incompatibility and its evolutionary consequences. International Journal of Plant Sciences 169(1):93–104, DOI 10.1086/523362, URL https://doi.org/10.1086/523362, https://doi.org/10.1086/523362

Jay P, Chouteau M, Whibley A, Bastide H, Parrinello H, Llaurens V, Joron M (2021) Mutation load at a mimicry supergene sheds new light on the evolution of inversion polymorphisms. Nature Genetics 53(3):288–293, DOI 10.1038/s41588-020-00771-1, URL https://www.nature.com/articles/s41588-020-00771-1

Kardos M, Allendorf FW, Luikart G (2014) Evaluating the role of inbreeding depression in heterozygosity-fitness correlations: how useful are tests for identity disequilibrium? Molecular Ecology Resources 14(3):519–530, DOI 10.1111/1755-0998.12193, URL https://onlinelibrary.wiley.com/doi/10.1111/1755-0998.12193

Keller LF, Waller DM (2002) Inbreeding effects in the wild. Trends in Ecology and Evolution 17:230–241

Kim BY, Huber CD, Lohmueller KE (2018) Deleterious variation shapes the genomic landscape of introgression. PLOS Genetics 14(10):e1007,741, DOI 10.1371/journal.pgen.1007741, URL https://dx.plos.org/10.1371/journal.pgen.1007741

Kimura M, Ohta T (1971) Theoretical Aspects of Population Genetics, monographs edn. Princeton University Press

Kirkpatrick M (2010) How and Why Chromosome Inversions Evolve. PLoS Biology 8(9):e1000,501, DOI 10.1371/journal.pbio.1000501, URL https://dx.plos.org/10.1371/journal.pbio.1000501

Kirkpatrick M, Jarne P (2000) The Effects of a Bottleneck on Inbreeding Depression and the Genetic Load. The American Naturalist 155(2):154–167, DOI 10.1086/303312, URL https://www.journals.uchicago.edu/doi/10.1086/303312

Kremling KAG, Chen SY, Su MH, Lepak NK, Romay MC, Swarts KL, Lu F, Lorant A, Bradbury PJ, Buckler ES (2018) Dysregulation of expression correlates with rare-allele burden and fitness loss in maize. Nature 555(7697):520–523, DOI 10.1038/nature25966, URL http://www.nature.com/articles/nature25966

Lande R, Schemske DW (1985a) The evolution of self-fertilization and inbreeding depression in plants. i. genetic model. Evolution 39(1):24

Lande R, Schemske DW (1985b) The Evolution of Self-Fertilization and Inbreeding Depression in Plants. I. Genetic Models. Evolution 39(1):24, DOI 10.2307/2408514, URL https://www.jstor.org/stable/2408514?origin=crossref

Larièpe A, Mangin B, Jasson S, Combes V, Dumas F, Jamin P, Lariagon C, Jolivot D, Madur D, Fiévet J, Gallais A, Dubreuil P, Charcosset A, Moreau L (2012) The Genetic Basis of Heterosis: Multiparental Quantitative Trait Loci Mapping Reveals Contrasted Levels of Apparent Overdominance Among Traits of Agronomical Interest in Maize (*Zea mays* L.). Genetics 190(2):795–811, DOI 10.1534/genetics.111.133447, URL https://academic.oup.com/genetics/article/190/2/795/6064102

Lewontin RC (1974) The Genetic Basis of Evolutionary Change. Columbia University Press

Llaurens V, Gonthier L, Billiard S (2009) The Sheltered Genetic Load Linked to the S Locus in Plants: New Insights From Theoretical and Empirical Approaches in Sporophytic Self-Incompatibility. Genetics 183(3):1105–1118, DOI 10.1534/genetics.109.102707, URL https://academic.oup.com/genetics/article/183/3/1105/6063166

Mable BK (2008) Genetic causes and consequences of the break-

down of self-incompatibility: case studies in the Brassicaceae. Genetics Research 90(1):47–60, DOI 10.1017/S0016672307008907, URL https://www.cambridge.org/core/product/identifier/S0016672307008907/type/journal$_a$$rticle$

Maynard Smith J (1977) Why the genome does not congeal. Nature 268(5622):693–696, DOI 10.1038/268693a0, URL https://www.nature.com/articles/268693a0

McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li H, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, Brown P, Browne C, Eller M, Guill K, Harjes C, Kroon D, Lepak N, Mitchell SE, Peterson B, Pressoir G, Romero S, Rosas MO, Salvo S, Yates H, Hanson M, Jones E, Smith S, Glaubitz JC, Goodman M, Ware D, Holland JB, Buckler ES (2009) Genetic Properties of the Maize Nested Association Mapping Population. Science 325(5941):737–740, DOI 10.1126/science.1174320, URL https://www.science.org/doi/10.1126/science.1174320

Ohta T (1971) Associative overdominance caused by linked detrimental mutations. Genetical Research (Cambridge) 18:277–286

Ohta T, Cockerham CC (1974) Detrimental genes with partial selfing and effects on a neutral locus. Genetical Research 23(2):191–200, DOI 10.1017/S0016672300014816, URL https://www.cambridge.org/core/product/identifier/S0016672300014816/type/journal$_a$$rticle$

Ohta T, Kimura M (1969) Linkage desequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics 63(1):229–238, DOI 10.1093/genetics/63.1.229, URL https://academic.oup.com/genetics/article/63/1/229/5989253

Olito C, Ponnikas S, Hansson B, Abbott JK (2022) Consequences of partially recessive deleterious genetic variation for the evolution of inversions suppressing recombination between sex chromosomes. Evolution 76(6):1320–1330, DOI https://doi.org/10.1111/evo.14496,

URL https://onlinelibrary.wiley.com/doi/abs/10.1111/evo.14496, https://onlinelibrary.wiley.com/doi/pdf/10.1111/evo.14496

van Oosterhout C, Zulstra WG, van Heuven MK, Brakefield PM (2000) Inbreeding depression and genetic load in laboratory metapopulations of the butterfly *Bicyclus anynana*. Evolution 54(1):218–225, DOI 10.1111/j.0014-3820.2000.tb00022.x, URL https://onlinelibrary.wiley.com/doi/10.1111/j.0014-3820.2000.tb00022.x

Rocheleau G, Lessard S (2000) Stability analysis of the partial selfing selection model. Journal of Mathematical Biology 40(6):541–574, DOI 10.1007/s002850000030, URL https://doi.org/10.1007/s002850000030

Roze D (2015) Effects of interference between selected loci on the mutation load, inbreeding depression, and heterosis. Genetics 201:745–757

Seymour DK, Chae E, Grimm DG, Martín Pizarro C, Habring-Müller A, Vasseur F, Rakitsch B, Borgwardt KM, Koenig D, Weigel D (2016) Genetic architecture of nonadditive inheritance in *Arabidopsis thaliana* hybrids. Proceedings of the National Academy of Sciences 113(46), DOI 10.1073/pnas.1615268113, URL https://pnas.org/doi/full/10.1073/pnas.1615268113

Sianta SA, Peischl S, Moeller DA, Brandvain Y (2021) Genetic load may increase or decrease with selfing depending upon the recombination environment. preprint, Genetics, DOI 10.1101/2021.05.20.445016, URL http://biorxiv.org/lookup/doi/10.1101/2021.05.20.445016

Spigler RB, Theodorou K, Chang SM (2017) Inbreeding depression and drift load in small populations at demographic disequilibrium. Evolution 71(1):81–94, DOI 10.1111/evo.13103

Sved JA (1971) Linkage disequilibrium and homozygosity of chromo-

some segments in finite populations. Theoretical Population Biology 2(2):125–141, DOI https://doi.org/10.1016/0040-5809(71)90011-6, URL https://www.sciencedirect.com/science/article/pii/0040580971900116

Takebayashi N (2003) Patterns of Variation Within Self-Incompatibility Loci. Molecular Biology and Evolution 20(11):1778–1794, DOI 10.1093/molbev/msg209, URL https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msg209

Tallmon DA, Luikart G, Waples RS (2004) The alluring simplicity and complex reality of genetic rescue. Trends in Ecology & Evolution 19(9):489–496, DOI 10.1016/j.tree.2004.07.003, URL https://www.sciencedirect.com/science/article/pii/S0169534704001843

Uyenoyama M, Holsinger K, Waller D (1993) Ecological and genetic factors directing the evolution of self-fertilization. Oxford Surveys in Evolutionary Biology 9:327 – 381

Uyenoyama MK, Waller DM (1991) Coevolution of self-fertilization and inbreeding depression II. Symmetric overdominance in viability. Theoretical Population Biology 40(1):47–77, DOI 10.1016/0040-5809(91)90046-I, URL https://linkinghub.elsevier.com/retrieve/pii/004058099190046I

Waller DM (2021) Addressing Darwin's dilemma: Can pseudo-overdominance explain persistent inbreeding depression and load? Evolution 75(4):779–793, DOI 10.1111/evo.14189, URL https://onlinelibrary.wiley.com/doi/10.1111/evo.14189

Whitlock M, Ingvarsson PK, Hatfield T (2000) Local drift load and the heterosis of interconnected populations. Heredity 84:452–457

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, URL https://ggplot2.tidyverse.org

Willi Y, Griffin P, Van Buskirk J (2013) Drift load in populations of small size and low density. Heredity 110(3):296–302, DOI 10.1038/hdy.2012.86

Winn AA, Elle E, Kalisz S, Cheptou PO, Eckert CG, Goode C, Johnston MO, Moeller DA, Ree RH, Sargent RD, et al (2011) Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. Evolution 65:3339–3359

Wu Q, Han TS, Chen X, Chen JF, Zou YP, Li ZW, Xu YC, Guo YL (2017) Long-term balancing selection contributes to adaptation in Arabidopsis and its relatives. Genome Biology 18(1):217, DOI 10.1186/s13059-017-1342-8, URL https://doi.org/10.1186/s13059-017-1342-8

Zhao L, Charlesworth B (2016) Resolving the conflict between associative overdominance and background selection. Genetics 203(3):1315–1334, DOI 10.1534/genetics.116.188912, URL https://www.genetics.org/content/203/3/1315, https://www.genetics.org/content/203/3/1315.full.pdf