 doi:10.20944/preprints202010.0634.v1

# Why do inbreeding depression and load resist both purging and fixation?

Donald M. Waller
Department of Botany
University of Wisconsin - Madison
Madison, WI 53706
dmwaller@wisc.edu
608 698 6495
ORCID ID: 0000-0001-5377-3929

Version: Nov. 1, 2020

**Length**: Words – Abstract: 223 Main text: 7457

Figures: 3 Tables: 1 References: 114

**Running Title:** "Recalcitrant inbred load"

**Data Accessibility Statement:** I share the Excel spreadsheet used to generate Table 1 and Figure 3 as Supplemental Information.

**Short summary:**

Selection within inbred populations could generate a residue of load loci linked in repulsion. Such blocks create pseudo-overdominance perpetuating the load and inbreeding depression by limiting both purging and fixation. Genomic signatures left by such blocks allow us to test this hypothesis.

**Abstract**

Upon inbreeding, the architecture of the inbreeding load shifts as selection purges strongly deleterious recessive mutations and drift fixes many milder ones. Most small inbred populations show limited genetic variation while crosses between such populations commonly express pronounced heterosis, confirming fixation. In contrast, purging appears to be limited in that inbred populations often retain substantial inbreeding depression. In addition we have the enigma Darwin noted: purely selfing taxa are unknown. Because both purging and fixation reduce inbreeding depression and load, another mechanism must exist to sustain these. Background selection and the associations that develop among alleles in small inbred populations will shift the architecture of the load potentially creating blocks of recessive mutations linked in repulsion. This would generate pseudo-overdominance that could sustain these "PODs" and inbreeding load. Recombination and crosses between lineages could erode PODs. Crosses between populations fixed for different mutations would generate high pseudo-overdominance, enhancing heterosis and potentially POD formation. New recessive mutations arising within PODs would reinforce overdominance. PODs should generate clear genetic signatures including genomic hotspots of heterozygosity and linkage disequilibrium containing alleles at intermediate frequency generating segregating load. Results from several simulation and empirical studies match these predictions. Further simulations and comparative genomic analyses are needed to rigorously test whether PODs exist in sufficient strength and number to generate persistent inbreeding depression and load in inbred lineages.

**Keywords**: genetic load; purging; drift load; heterozygosity-fitness correlation; genomic architecture.

 doi:10.20944/preprints202010.0634.v1

48 "Nature abhors perpetual self-fertilization" C. Darwin, 1876

49 Charles Darwin investigated the effects of self- and cross-fertilization in 52 taxa of
50 flowering plants over 10 years. He concluded "It is as unmistakably plain that innumerable
51 flowers are adapted for cross-fertilisation, as that the teeth and talons of a carnivore are adapted
52 for catching prey" and that "It is a general law of nature that flowers are adapted to be crossed, at
53 least occasionally, by pollen from a distinct plant" (p. 3 and 6, Darwin 1876). Although Darwin
54 usually found substantial inbreeding depression, he also noted occasional exceptions like the
55 individual he dubbed "Hero" that emerged in the sixth generations of inbreeding in *Ipomea*
56 *purpurea*. This plant had high fitness, a trait it shared with its progeny. We now interpret its high
57 fitness as the likely result of strong selection eliminating recessive deleterious alleles, reducing
58 inbreeding depression (Crow 1999; Charlesworth and Charlesworth 1999; Charlesworth & Willis
59 2009). But if selection can purge inbred lines of their genetic load, the question arises: why
60 don't high fitness "Heroes" emerge regularly? Once they arise, why don't such individuals
61 generate purely selfing populations and taxa? Echoing the "Knight-Darwin Law" reflected in the
62 opening quote, Stebbins (1957, p. 348) noted that "no species of plant . . studied intensively
63 throughout its entire range . . has been found to be exclusively self-pollinated." Recent work
64 confirms that obligate outcrossers enjoy faster rates of diversification than selfing taxa (Goldberg
65 et al. 2010) and that selfing lineages suffer a long-term disadvantage (Cheptou 2019).

66 The twin enigmas of persistent inbreeding depression (ID) in inbred populations and the
67 lack of fully selfing species motivate this paper. I argue here that the same selective processes
68 acting to purge and fix deleterious recessive mutations within small or inbreeding populations
69 also restructure the load creating blocks of mutations maintained by pseudo-overdominance (see
70 Glossary). I describe these dynamics and illustrate the idea with a numerical example. I then

 doi:10.20944/preprints202010.0634.v1

make specific predictions about the genomic signatures such pseudo-overdominant blocks (or "PODs") should generate, indicating how these ideas might be tested. I conclude by considering implications of PODs and the extent to which simulation and empirical studies support their existence. For additional background on inbreeding concepts, see Waller and Keller (2020).

*Why doesn't inbreeding eliminate inbreeding depression?*

Inbreeding dramatically increases the rate at which deleterious recessive alleles are expressed in homozygotes. This exposes them to selection, purging at least some of the inbreeding load. This could potentially allow fitness to recover following population bottlenecks or episodes of inbreeding. Lande and Schemske's (1985) theoretical model confirmed how efficiently selection purges unlinked, fully recessive, lethal mutations within inbred populations, reducing ID and increasing inbred fitness (Fig. 1). They predicted that this should favor run-away selection for increased self-fertilization, resulting in disruptive selection favoring either purely outcrossing or purely selfing populations and species (with high and low ID, respectively). This led them to argue that mixed mating systems exist only as a transient condition (Schemske and Lande 1985). Hopes that purging could efficiently reduce or eliminate the load also led some to advocate intentional inbreeding in captive breeding programs to reduce inbreeding load and boost population viability (Templeton and Reed 1984).

Contrary to these predictions, empirical studies often find inbred populations retain substantial inbreeding depression (e.g., Willis 1999; vanOosterhout et al. 2000). Levels of ID are often unrelated to population size, the amount of inbreeding, or amounts of genetic variability within populations (e.g. Paland and Schmid 2003; Willi et al. 2013; Toczyldowski and Waller, submitted). Indeed, both individual studies (e.g., Spigler et al. 2017; Baldwin and Schoen 2019)

and meta-analyses (Byers and Waller 1999; Winn et al. 2011) find persistent ID in inbred and mixed-mating populations and taxa. Keller and Waller (2002) concluded that purging usually fails to substantially reduce inbreeding depression in field populations. Substantial ID in inbred populations could explain why mixed mating systems persist and appear evolutionarily stable.

Whether deleterious recessive alleles are purged rapidly as predicted by the Lande-Schemske model hinges critically on assuming that mutations are lethal, fully recessive, and unlinked. In fact, most deleterious mutations are mild (i.e., $s<0.03$), expressed partially in heterozygotes, and often linked to other selected loci, greatly weakening the efficacy of purging (Charlesworth and Charlesworth 1987; Willis 1999). It is difficult to purge mutations of small effect from small, inbred populations (Wang 2000). Incompletely recessive mutations are primarily selected against as heterozygotes in outbred populations, greatly reducing their frequency even before inbreeding occurs. Linkage interferes with the ability of selection to act independently across loci by generating associations (the Hill-Robertson effect), augmenting drift (Charlesworth et al. 1992). Sophisticated models that omit linkage often conclude that purging is effective (e.g., Lande and Porcher 2015). However, models that include linkage often generate different conclusions (e.g., Abu Awad and Roze 2018). Identity disequilibrium (the correlation in homozygosity among loci) increases with inbreeding making individuals either omozygous or heterozygous at many loci. This promotes the purging of recessive mutations, but linkage disequilibrium, pseudo-overdominance, and higher-level associations interfere with purging, modifying selective dynamics in complex ways (Roze 2015).

Persistent inbreeding depression within inbred populations and evolutionarily stable mixed mating systems have generated efforts to explain these that invoke various mechanisms (e.g., Johnston and Schoen 1996; Goodwillie et al. 2005; Charlesworth and Willis 2009). Some

 doi:10.20944/preprints202010.0634.v1

interpret the lack of purging as evidence that most genetic load reflects mutations of small effect against which selection is weak (Byers and Waller 1999). Others hypothesize that inbreeding load persists because mutation rates to partially recessive deleterious alleles are high and continuously replenish the load (Barrett and Charlesworth 1991). Others invoke complex fluctuations in population size (demographic disequilbrium – Spigler et al. 2017) or selective interference among multiple load loci (Winn et al. 2011). We lack a general consensus on how to explain these enigmas.

Models for the evolution of sex recognize recombination as a key force, favored because it reduces linkage disequilibrium among alleles at different loci (Hill and Robertson 1968; Maynard-Smith 1978). Some recent studies on the effects of inbreeding also include effects of linkage and interactions among loci. For example, selective sweeps in partially selfing populations increase fixation rates for deleterious alleles, reducing polymorphism and effects of recombination (Hartfield and Glémin 2014). Viability loci also develop associations with alleles that modify selfing rates, substantially altering evolutionary dynamics relative to simple models that ignore linkage (Uyenoyama et al. 1993). Recombination facilitates selection, as found in selfing lines of maize where deleterious mutations purge slowly and variation is maintained in genomic regions of high recombination (McMullen et al. 2009; Roessler et al. 2019).

Identity disequilibrium (the correlation in homozygosity among loci) arising in inbred populations could also enhance purging as killing one highly homozygous individual eliminates several mutations. Linkage and identity disequilibrium among fully recessive mutations could also reduce purging, however, if selfed progeny simply do not survive (Lande et al. 1994; Bierne et al. 2000). Glémin (2003) distinguished purging dynamics resulting from inbreeding from those

139 caused by drift. He noted that drift only purges highly recessive mutations and that both types of
140 purging become ineffective once populations drop below some threshold size.

141 Even a small amount of outcrossing helps to break up the negative associations that arise
142 among alleles within selfing populations (Kamran-Disfani and Agrawal 2014). More elaborate
143 models assuming linkage, moderate interference, and pairwise genic interactions within large
144 populations show ID to be sustained even under moderately high rates of selfing (Roze 2015). In
145 Roze's models, identity disequilibrium reduces ID directly (as noted above) and indirectly (by
146 decreasing homozygosity at each locus) but also decreases the efficiency with which selection
147 eliminates mutations, increasing ID. The last effect dominates, so overall inbreeding load
148 increases. Deviations of simulations from analytical models led him to conclude that "higher-
149 order genetic associations . . must have important effects."

150

151 *How do we infer purging?*

152 Confusion remains over how to infer purging. Many interpret declines in inbreeding
153 depression after inbreeding as evidence of purging. This is incorrect, however, as ID only
154 measures relative fitness (typically defined as the proportional decline in fitness of inbred
155 relative to outcrossed progeny: $\delta = 1 - \boldsymbol{Ws} / \boldsymbol{Wo}$). Drift and background selection also reduce ID,
156 not by eliminating deleterious mutations in inbred progeny but by fixing them in outcrossed ones
157 (Fig. 1). Fixation thus mimics purging as segregating load is fixed (documented in *Daphnia* by
158 Lohr and Haag 2015). To demonstrate true declines in the load would require us to show that
159 absolute inbred fitness improves upon inbreeding (Keller and Waller 2002).

160

161 *Drift load and heterosis*

 doi:10.20944/preprints202010.0634.v1

Population declines and inbreeding enhance the power of drift even as they expose deleterious recessive mutations to purging. Drift greatly affects the frequency of mildly deleterious alleles as selection loses power in small and/or inbred populations. In particular, as the scaled selection coefficient $\boldsymbol{S} = 2\boldsymbol{N_e s}$ (where $\boldsymbol{s}$ is the fitness decrement in mutant homozygotes and $\boldsymbol{N_e}$ is the genetically effective population size) declines below one, selection becomes unable to counter drift (Wright 1931; Haldane 1932; Charlesworth 2018). These mutations then start to fix in proportion to their frequency (Kimura and Crow 1964), creating a fixed "drift load" (Whitlock et al. 2000) or "finite population load" (Crow and Kimura 1970). As mutations continue fix and accumulate, small populations may undergo "mutational meltdown" (Gabriel et al. 1993; Schultz and Lynch 1997).

Progeny from crosses within highly inbred populations may show little fitness recovery if individuals fix for the same mutations (as Oakley et al. 2019 found in wild populations of *Arabipsosis thaliana*). In contrast, crosses between separately inbred lineages or populations often result in high fitness. Such heterosis emerges as different deleterious recessive mutations tend to fix in each population. This allows alleles from each population to "mask" the deleterious effects of alleles in the other. Not surprisingly, small or selfing populations usually display more heterosis in crosses than outcrossing populations of the same species (e.g., Oakley and Winn 2012; Oakley et al. 2015). Heterosis generated excitement among plant and animal breeders in the mid 20$^{th}$ century when it was ascribed to true overdominance. As genetic maps became more precise, however, most heterotic effects turned out to reflect associative overdominance (Crow 1999; Hedrick 2012).

Small or inbred populations often express appreciable drift loads as in small populations of the perennial *Hypericum cumulicola* (Oakley and Winn 2012) and *Gentianella germanica*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

(Paland and Schmidt 2003). In both species, small and large populations had similar ID, leading them to conclude that purging was weaker than drift in these populations. Willi (2013), Willi et al. (2013), and Oakley et al. (2015) similarly found high drift loads in small or selfing populations of *Arabidopsis lyrata* yet also appreciable ID rather than the significant declines in ID expected under purging. Oakley et al. (2015) found heterosis and ID to be uncorrelated, leading them to conclude that they have different genetic bases.

In sum, the appreciable drift load and high heterosis we find in crosses between small or inbred populations support predictions from classical drift models. In contrast, we fail to find consistent declines in ID as expected with purging. What could account for this discrepancy? Charlesworth (2018) extended classic 1-locus load theory to predict levels of fixed drift load and ID under various assumed parameter values. He compared these theoretical predictions to results from empirical studies (cited above). He also concluded that the consistent increases in between-population heterosis seen in smaller or more selfing populations agree well with theory. In contrast, the strong ID in many inbred populations suggested to him that selected loci within small or inbred populations likely show associations, violating the independence assumed in classic models. He interpreted recent theory (Kamran-Disfani and Agrawal, 2014, Roze 2015, and Bersabé et al. 2016) to argue that such associations should be modest, but noted that heterozygote advantage (balancing selection) emerges regularly in genomic surveys (Charlesworth 2006; Gao et al. 2015; Hedrick et al. 2016; Siewert and Voight 2017).

*Hypothesis*

If both purging and drift act to erode inbreeding (segregational) load in bottlenecked and/or inbred populations, why does inbreeding depression persist? Is there a mechanism that

 doi:10.20944/preprints202010.0634.v1

might limit both purging and fixation of deleterious recessive mutations within inbred populations? Selection and drift act together in such populations to modify the architecture of the load – could this sustain inbreeding depression? "Architecture" here refers to the number of deleterious mutations, their joint distributions of deleterious and dominance effects, and linkage relations within the genome (cf. Remington and O'Malley 2000; Glémin 2003). Any mechanism that perpetuates the inbreeding load would help to explain the persistence of mixed mating systems although other factors also affect selection on the mating system. These include associations among loci (Uyenoyama et al. 1993), correlations among fitness components (Johnston et al. 2008), and pollen discounting (Holsinger 1991; Chang and Rausher 1998).

Recessive deleterious mutations exist as both highly deleterious alleles (lethals and sub-lethals) and mildly deleterious alleles. Even if these arise at similar rates, selection ensures that mild mutations will be far commoner at equilibrium (Crow and Kimura 1970). This is true in large outcrossing populations and truer in inbred populations as highly deleterious recessive mutations are purged. Kimura et al. (1963) first explored how equilibrium frequencies of mutations of variable effect shift as population sizes shrink. Large populations undergoing efficient selection eliminate many highly deleterious mutations (which may already be at low frequency if they are expressed partly within heterozygotes). Their residual load thus consists mostly of milder and more recessive mutations. Their high number could generate high ID. As population sizes shrink, mutations of progressively larger effect become invisible to selection and start to drift and fix (Kimura et al. 1963; Crow and Kimura 1970; Glémin 2003). This increases the total load as some mutations fix (creating drift load) and others continue to segregate (generating inbreeding load).

As outcrossed populations start to inbreed, the architecture of the load shifts as highly deleterious recessive alleles are lost to purging and many mild mutations fix. Purging generates "background selection" and "selective sweeps", reducing variation at linked loci (Charlesworth et al. 1993; Barton 1995). Purifying selection also acts to flatten the distribution of heritability effects across the genome as rarer large-effect variants are quickly selected, leaving commoner alleles with smaller effects (O'Connor et al. 2019). These processes are especially strong in species with few chromosomes or short map length (Nordborg et al. 1996). Many mildly deleterious mutations linked in repulsion to lethals or sub-lethals thus fix and favorable mutations linked in coupling may be lost. Genomic regions with low recombination increase the power of this background selection.

For linked deleterious recessive mutations, two outcomes emerge. Mutations linked in coupling enhance purging as one selective death eliminates two or more mutations (amplifying background selection and sweeps). Mutations linked in repulsion, however, interfere with purging by reducing the fitness of both segregating homozygotes (Fig. 2a). This generates pseudo-overdominance, favoring the heterozygote via balancing selection (Ohta 1971). This resembles associative overdominance where observed (presumed neutral) genetic markers appear to show overdominance because nearby deleterious mutations are linked in repulsion (Frydenberg 1963). Such sets of mutations showing negative linkage disequilibrium (Hill-Robertson effects) tend to persist and segregate load until recombination occurs to break up dual (or multiple) heterozygotes. Tight linkage and the extended linkage groups generated by inbreeding reduce the frequency of such recombination events, extending their persistence.

Most of the load reflects mild mutations, particularly in inbred populations. Models accurately predict the evolution of ID (and heterosis in between-population crosses) when

 doi:10.20944/preprints202010.0634.v1

253 selection acts independently among loci (Charlesworth 2018). Such models are of less use when
254 linkage and selective interference affect selective dynamics. As selection loses its efficiency in
255 small populations, the density of segregating mild mutations will increase, enhancing linkage
256 effects and selective interference. These linkage effects could accelerate purging via two
257 mechanisms. Identity disequilibrium (the correlation in homo- or heterozygosity among loci)
258 increases in inbred populations, allowing one selective death to eliminate multiple mutations (but
259 see Roze 2015). If mutations are consistently linked in coupling, purging would also be enhanced
260 (see above and Bersabé et al. 2016). This is inherently unlikely, however, as more mild
261 mutations are involved. A chromosomal segment with many mutations at random locations is
262 likely to have a roughly even number of mutations on opposing segments (Fig. 2b). Mutations in
263 both coupling and repulsion linkage disequilibrium generate pseudo-overdominance (Lynch and
264 Walsh 1998, p. 288). Many such mutations could combine to yield appreciable fitness effects
265 and pseudo-overdominance. Heterozygotes would retain high fitness by masking these
266 deleterious effects within loci. Such pseudo-overdominant blocks (PODs) would act to
267 perpetuate ID by shielding mutations from purging and slowing fixation.
268
269 *Numerical example*
270 How commonly will pseudo-overdominance arise and persist? Extensive simulations
271 would allow us to rigorously explore the circumstances favoring the creation and persistence of
272 PODs. However, a simple example – a roll of the dice for one individual – serves to show how
273 such effects might arise. I used an Excel spreadsheet (see SI) to generate random map locations
274 for five major and 200 minor mutations distributed along each of two homolog chromosomes (a
275 and b) with a map length of 1000. I arbitrarily divided this into ten bins to represent linked sets

 doi:10.20944/preprints202010.0634.v1

of loci (Table 1) assuming alleles to be fully linked within bins but unlinked among bins. Most bins show zero or one major mutation, as expected, but bins #1 and #9 gained three and two such genes. Because these are linked in coupling with no opposing mutations on the opposite chromosome, they would quickly purge. Bin 4, in contrast, supports two major mutations in repulsion (Table 1; Fig. 3a), generating strong pseudo-overdominance ("fixed heterozygosity" if lethal - Roose and Gottlieb 1976). In general, cases of coupling and repulsion should occur in equal frequencies.

The far more common mild mutations occur in similar numbers among segments and on the two homologs (mean=20 here; Table 1; Fig. 3b). The largest asymmetry occurs in bin #3 where 23 mild mutations occur on one segment but only 10 in its homolog. The major mutation co-occurring among the 23 mild mutations on 1a assures that this segment would be rapidly purged (fixing the 10 mild mutations in repulsion). In the four segments lacking major mutations, the largest asymmetry in loads occurred in bin #7 (20 vs. 30 mutations). If we assume multiplicative fitness interactions among loci, s = 0.02, and h = 0.1, the three genotypes segregating from self-fertilization would have fitnesses of 0.668, 0.905, and 0.545. This degree of asymmetery would allow some selection, but would also exhibit overdominance, slowing purging. The other three segments lacking strong mutations show more balanced selection, reducing purging.

*Implications*

Although this example includes complete linkage within the bins/blocks, actual PODs would be vulnerable to recombination. Recombination weakens these PODs by reducing linkage disequilibrium and increasing potential asymmetry. Mutations (particularly in cis-clumps)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

become more exposed to purging. Others could fix via background selection or drift. This would reduce the balance of selective forces maintaining PODs and erode their overdominance effects. As PODs decay, these processes could accelerate, dissolving the POD. Inbreeding would accentuate selection by increasing segregation, exposing recessive mutations. However, inbreeding would also reduce recombination and diminish the power of selection, slowing the decay of PODs.

Because PODs would vary in their number of mutations and mutational effects, so would the selective effects of any one POD. Selective effects would accumulate across PODs, however, increasing their total effect. The identity disequilibrium present in inbred populations would amplify this effect by increasing the expression of overdominance and ID in multiply heterozygous individuals. These selective effects might even be strong enough to eliminate inbred progeny, blocking purging as already noted. Recombination and selection would act to erode PODs over time, but POD strength could also grow as new recessive mutations arise within these blocks, boosting overdominance while being shielded from selection.

Crosses between independently inbred lineages or sub-populations could strongly affect PODs both by eroding existing PODs and by creating conditions favoring the evolution of new ones. Recessive mutations unique to either population would be masked as heterozygotes in the F1, generating high heterosis via pseudo-overdominance. Extant PODs in two populations would add to the initial high heterosis seen in F1 crosses but would not affect heterosis much if this mostly reflects mutations independently fixed in the two populations. Existing PODs would, however, become more exposed to selection as they paired up with new haplotypes that did not mirror their load. That is, we expect much of the load initially segregating in each source population to be concentrated within PODs that would usually occur at different locations in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

two populations (given independent origins). Thus, upon crossing, load-laden POD haplotypes would become paired with low-load non-POD haplotypes from the other population. The resulting asymmetry in fitness of segregating homozygotes could quickly purge POD haplotypes. Again, self-fertilization would promote purging by increasing the frequency of homozygotes and identity disequilibrium while enhanced linkage and associations among loci would interfere with selection, making outcomes uncertain.

In addition to eroding PODs already present in the source populations, between-population crosses could also create new PODs by generating high levels of pseudo-overdominance and linkage disequilibrium (LD) in the F1. Mutations in either population would initially exist in repulsion to all mutations unique to the other, generating high linkage disequilibrium (LD). With free recombination (unlinked loci) and random mating, this LD would decay by one half each generation. However, with many load loci, linkage, small $N_e$, and the genic associations present in inbred populations, much LD would persist. This high pseudo-overdominance and LD would create conditions conducive to forming new PODs. These would be most likely to persist if the segregating load were large enough to select against most homozygous progeny, limiting purging.

A third scenario might unfold if a POD contributed by one population overlapped or lay adjacent to a POD from the other population. This could allow the PODs to fuse and reinforce each other, ensuring that both would persist. It would thus be interesting to map PODs in different populations to see whether PODs occur non-randomly, e.g., in mutation hotspots or recombination cold spots.

In sum, a history of inbreeding and drift or heterotic crosses between inbred sub-populations could generate PODs and/or increase their selective effects. The existence of

substantial pseudo-overdominance would substantially alter predictions based on classical models of purging and drift (Charlesworth 2018), allowing populations to sustain enough ID to counter selection for selfing. Populations selected to retain some outcrossing would also benefit from boosts in fitness provided by heterotic crosses to other populations. They would also benefit from increased recombination allowing selection to counter accumulations of drift load. Ironically, although selection acts to purge inbreeding load, it may also act to sustain it by creating regions of persistent pseudo-overdominance. Cycles of inbreeding alternating with between-lineage crosses and selective purging might also allow the number (or selective effects) of PODs to cycle. Selection may also favor the early expression of ID to reduce investment in inviable progeny (Husband and Schemske 1996). This, too, could help perpetuate PODs and ID by eliminating homozygotes. Over time, selection and fixation should eliminate PODs with smaller and/or more asymmetric fitness effects. Thus, the remaining inbreeding load should steadily become more balanced, stable, and recalcitrant to further purging.

*Predictions*

If pseudo-overdominance blocks, or PODs, exist, these should be evident both in the form of persistent inbreeding depression and heterogeneity in the architecture of load. I make these predictions to promote tests of the POD hypothesis. POD signatures should resemble those found for other loci subject to balancing selection like disease resistance genes (e.g., Wang et al. 2019) and self-incompatibility loci. Some studies relevant to these predictions are cited here. A more extensive discussion of relevant results follows.

1. *Genetic variation and inbreeding depression* should persist within inbred populations. Balancing selection acting on linkage blocks could sustain polymorphisms at marker loci

 doi:10.20944/preprints202010.0634.v1

(as observed by Pamilo and Pálsson 1998). Amounts of variation and ID, however, will vary depending on the particular demographic and genetic histories experienced by a population and which mutations fix, purge, or continue segregating. Lines of honeybees (*Apis mellifera capensis*) inbred for more than 100 generations in way that should have reduced heterozygosity by a third each generation retained heterozygosity at 34% of 10,884 loci and in the same genomic regions across all sub-lineages (Smith et al. 2019). This strongly suggests that PODs exist in this strain and demonstrates how strong overdominance can sustain variation even in highly inbred lines.

2. *Pseudo-overdominance effects.* Evidence for overdominance should exist in many inbred populations mapping to particular restricted genomic regions that may differ among sub-populations. PODs are difficult to distinguish from truly overdominant loci unless very fine-scale maps are available. Lynch and Walsh (1998, p. 284), drawing on Crow (1958) and Lewontin (1974), argued that dominance and overdominance can be distinguished as mechanisms of inbreeding depression using the ratio (A+B)/A where A is the sum of mortalities in an outcrossing population and B is the additional mortality seen in a fully inbred population (the intercept and slope of a log fitness vs. F regression – Morton et al. 1956). This ratio greatly exceeds 2 in nine outcrossing vertebrate species, leading them to conclude that overdominance does not cause ID in these populations. If overdominance were great enough, selfed progeny would be eliminated, sustaining high B even in inbred populations. It might thus be interesting to compare (A+B)/A ratios in related outcrossing and inbred populations.

3. *Site frequency spectra.* Without PODs, genetic diversity would erode in small inbred populations leaving few variable loci. Those that did still vary should show a fairly

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

continuous range of allele frequencies under drift. In contrast, PODs maintained by balancing selection should contain alleles with frequencies close to 0.5. Genomic scans should also reveal heterogeneity or bimodality with high fixation across much of the genome alternating with local "hot spots" of heterozygosity. Spectra from independently inbred lineages, however, should show variability at distinct sets of loci.

4. *Allele age.* If balancing selection acts to sustain mutations as masked heterozygotes, these might persist for long enough for gene genealogies to distinguish these from shorter-lived mutations in other regions (which forms the basis for genomic tests of balancing selection - Siewert and Voight 2017). Becher et al. (2020) noted that overdominance increases the lengths of internal branches in coalescent trees. However, this effect may be too modest to detect in sequence data given that inbred lineages are often short-lived and that recombination probably limits the longevity of PODs.
5. *Inbreeding estimates.* Inbreeding is usually estimated from heterozygosity at genetic marker loci. PODs, however, act to sustain heterozygosity both locally within blocks and generally in surviving offspring. This could generate estimates of inbreeding below those calculated from pedigrees. Viewed another way, PODs should retard the degree of inbreeding attained within inbred pedigrees. PODs could also cause estimates of inbreeding to vary with marker loci located within PODs showing low values and markers outside showing higher values.
6. *Runs of homozygosity*. Fisher (1965) developed the theory of junctions in the second edition of his book on inbreeding to predict the length of runs of homozygosity expected in lineages with various histories of inbreeding. These have been used to infer inbreeding in human populations (McQuillan et al. 2008). Stable PODs maintained by balancing

414 selection could interrupt such runs, however, reducing their upper limit. This should
415 reduce the frequency of longer run lengths.
416 7. *Linkage disequilibrium*. Balancing selection acting within PODs would generate strongly
417 positive and negative LD among alleles in or near these blocks. Genomic regions of high
418 LD, high heterozygosity, and intermediate allele frequencies should all coincide.
419 Estimates of LD should also differ when calculated from markers in or outside PODs.
420 Tenaillon et al. (2008) simulated heterozygosity and LD among neutral loci in
421 populations emerging from a bottleneck. The found discrete hotspots of polymorphism
422 and high levels of LD across long map distances with strikingly discrete modes. These
423 high LD values reflected long coalescent times, resembling expectations under balancing
424 selection. One might extend their model (lacking selection) by simulating PODs with
425 selection and assess how levels of heterozygosity, LD, and their heterogeneity respond.
426 8. *Dynamics*. Crosses between separately inbred populations usually generate high levels of
427 heterozygosity, LD, and pseudo-overdominance in the F1. Does their decline in
428 subsequent generations match predictions from classic models? PODs should slow these
429 declines. We also expect levels of inbreeding depression to increase following
430 hybridizations as some of the drift loads fixed within sub-populations converts to become
431 segregating load. The presence of PODs would slow subsequent purging and fixation
432 below classic (unlinked) expectations.
433 9. *Dominance vs. additive variance*. Inbreeding substantially alters evolutionary responses
434 to selection (Kelly & Arathi 2003). Populations starting to inbreed express considerable
435 quantitative genetic variation resulting from both additive and dominance variance
436 components, allowing them to respond quickly to selection. As inbreeding proceeds,

 doi:10.20944/preprints202010.0634.v1

however, purging and fixation deplete additive variance (Clo et al. 2019). Classical models of purging and fixation suggest that dominance variance (reflected in ID) should also decline. Overdominance, however, would sustain dominance variance (Lynch and Walsh 1988, p. 604). Noël et al. (2017) found abrupt reductions in the ability of selfing lines of *Physa* snails to respond to selection after only three generations of selfing. This is striking given that these lines continued to express higher levels of phenotypic variance than outcrossed lines. PODs might explain this result if the segregating variance PODs maintain fail to maintain adaptive quantitative genetic variation.

10. *Selection for outcrossing.* As inbred lineages progressively lose fitness, either by steadily accumulating more mutations or by generating PODs, selection should favor outcrossing more as lineages become more inbred. Spigler et al. (2016) indeed found higher ID in smaller populations of *Sabatia*. This effect might favor adaptations that would allow more inbred lineages to outcross more. Evidence for this may exist in two inbred annuals that self via cleistogamous flowers. Larger individuals outcross more in both species. *Amphicarpaea bracteata* first produces a few large seeds via subterranean selfed cleistogamous flowers before producing aerial cleistogamous and (finally) chasmogamous (outcrossing) flowers. Only plants growing from large buried (selfed) seeds ever grow large enough to outcross, ensuring that outcrossed plants only self and only selfed plants can outcross (Schnee and Waller 1986). More inbred individuals of *Impatiens capensis* similarly invest more in outcrossing via chasmogamous structures (Toczydlowski 2019).

11. *Heterozygosity fitness correlations.* Positive correlations between heterozygosity and fitness occur in partially inbred populations reflecting both identity disequilibrium and

the fact that heterozygosity inversely measures inbreeding (Pamilo and Pálsson 1998; Lynch and Walsh 1998, p. 289). As inbreeding proceeds, however, any PODs present would limit declines in heterozygosity, restricting the range of heterozygosity values observed and thus heterozygosity-fitness correlations. These relationships might even reverse to become negative in highly inbred populations if heterozygosity starts to indicate the presence of PODs and their segregating load. Analyses of molecular genetic variation and fitness in 12 populations of *Impatiens capensis* generated exactly these results (Toczydlowski and Waller, submitted).

## Discussion

Given that both purging and fixation act to reduce ID, the force maintaining ID under inbreeding has remained obscure. Classic theory explored how purging reduces the load and how drift often fixes it but fewer studies have explored the fate of linked, segregating deleterious mutations in complex genetic association with each other. Likewise, few studies dissect the molecular bases for quantitative fitness variation. These gaps have left our picture of how selection, linkage, inbreeding, and drift interact to modify the architecture of the fixed and segregating load incomplete. Nevertheless, recent studies are starting to fill in these gaps.

Despite the ability of selection to swiftly purge unlinked, strongly deleterious, recessive mutations, background selection simultaneously fixes many mild mutations. Declines in population size and inbreeding also reduce the power of selection causing more mild mutations to fix via drift. This likely accounts for the low fitness of many isolated populations and fitness recovery we see following between-population crosses allowing ‘genetic rescue’ (Richards 2000; Hedrick and Garcia-Dorado 2016). Among the mutations drifting in small, inbred populations,

 doi:10.20944/preprints202010.0634.v1

483 some become linked in repulsion creating pseudo-overdominance. Several such blocks (PODs)
484 could together generate significant ID. These PODs gain strength by sheltering load and
485 accumulating additional mutations. Crosses between sub-populations fixed for different
486 mutations could foster POD creation by converting fixed drift load mutations to arrays of
487 segregating load loci, many with high LD. PODs decay as recombination dissolves linkage
488 disequilibrium (LD) exposing more mutations to selection. This may explain why PODs can
489 arise in 'cold-spots' of low recombination (see below). With PODs, ID and load persist not in
490 spite of selection but because of it.

491 The number and effects of PODs hinge on the balance that exists between how quickly
492 they are created and how long they persist before being disrupted and dissolved. Recombination
493 will at to erode the genic associations that sustain PODs. Inbreeding reduces effective
494 recombination while increasing genic associations including identity disequilibrium. The
495 resulting increase in multiply homozygous individuals should increase the efficacy of purging,
496 but increases other genic associations that may interfere. Overall, genic associations cause loads
497 to increase in parallel with levels of neutral genetic variation (e.g., Pálsson and Pamilo 1999;
498 Uyenoyama et al. 1993; Charlesworth 2012). Alleles at viability and quantitative trait loci can
499 also interfere with selection, increasing or decreasing the load so as to favor mixed mating
500 (Lande and Porcher 2017).

501 Developing theory to treat the dynamics of selected alleles at multiple loci subject to
502 linkage and associations within inbred populations is difficult, but ignoring these complications
503 has misled us. We lack knowledge on key parameters like the number of load loci, their selective
504 effects, degrees of dominance, and linkage relationships and how the architecture of the load
505 shifts with inbreeding. Nevertheless, we can explore POD dynamics using simulations to explore

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

how and why PODs emerge, how long they persist, and the effects they may exert. We can also capitalize on genomic data emerging from crops and model organisms like *Drosophila*, *Arabidopsis*, and *Mimulus* to see which predictions are met.

*Simulation results*

Zhao and Charlesworth (2016) used theory and simulations to explore how the strength of background selection (acting to deplete neutral variation at linked loci) compares with selection from pseudo-overdominance (acting to sustain variation). They found that pseudo-overdominance arose in small or inbred populations among weakly selected loci when $N_e s$ was of order one or less (initially at linkage equilibrium in their model). Small population sizes and weak selection facilitated the emergence of pseudo-overdominance. Inbreeding and population bottlenecks thus promote the formation of PODs by reducing $N_e$, as could strong background selection on alleles at linked loci. Pseudo-overdominance can then dominate background selection, greatly retarding the loss of genetic variation. Reduced recombination increases these overdominance effects, potentially causing small inbred populations to "crystalize" into two complementary haplotypes (e.g., + - + - - + vs. - + - + + -), mimicking a strong overdominant locus (cf. Charlesworth and Charlesworth 1997; Pálsson and Pamilo 1999). Becher et al. (2020) extended these results, noting that "LD is the driving force for AOD." More recessive mutations and those with more variable effects generated more LD while increasing allelic diversity ($\pi$) and reducing the asymmetry in allele frequencies.

Striving to understand why hotspots of nucleotide diversity and recombination evolve, Tenaillon et al. (2008) simulated the effects of population bottlenecks on how polymorphic sites aggregate and resulting levels of LD. Conspicuous discrete spikes in variation (50x more

 doi:10.20944/preprints202010.0634.v1

529 polymorphic sites per kilobase) emerged, resembling mutation hotspots. These occurred at
530 various locations but always supported alleles at intermediate frequencies. Their highest
531 aggregation scenario typically generated just two haplotypes, mimicking the 'crystalized'
532 scenario noted above.

533 Bersabé et al. (2016) simulated the effects of linkage on purging of the genetic load after
534 a population bottleneck (from N=1000 to 10 or 50) due to alleles of varying selective effect and
535 dominance. Linkage increased the segregating load relative to free recombination, slowing
536 purging. These outcomes reflected the accumulation of mildly deleterious mutations within
537 blocks, allowing them to be partly sheltered from selection within heterozygotes. This runs
538 counter to the negative linkage disequilibrium predicted by classical theory (Barton & Otto
539 2005) but fits with pseudo-overdominance. Balancing selection could also explain their finding
540 that linkage retarded the loss of neutral genetic diversity. As they only report mean LD levels, we
541 do not know if these values fell into distinct modes as expected with PODs.

542 Harkness et al. (2018) also simulated load dynamics in independently inbred lines
543 (N=100) to understand how the heterosis resulting from crosses would affect selection for
544 outcrossing at an unlinked modifier locus. They assumed that purging and fixation were initially
545 complete within these lines, yielding monomorphic viability loci and thus no inbreeding
546 depression (or PODs). High heterosis emerged, as expected, in F1 crosses, initially favoring the
547 outcrossing modifier. These often persisted only through a transient period, however. Only
548 highly polygenic load or modifiers located on a rare background generated longer-term selection
549 for outcrossing. The increase in rare outcrossing modifiers may reflect the associations they
550 gained with multiply heterozygous mutations. This and the fact that only highly polygenic load
551 favored such modifiers agree with the POD model. A single balanced overdominant locus

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

552 anywhere in the genome can interact with a selfing modifier locus to favor persistently mixed
553 mating systems (Uyenoyama and Waller 1991).
554
555 *Empirical evidence on pseudo-dominance*
556 If PODs exist, sustaining segregating load (ID), there should be empirical evidence to
557 support the above *Predictions*. As already noted, many studies failed to find the systematic
558 declines in inbreeding depression expected if selection purges load in inbred populations. Carr et
559 al. (1997), for example, found no association between levels of inbreeding depression and
560 autogamy among populations of *Mimulus guttatus*. Carr and Dudash (2003) noted that
561 "Investigators using molecular markers to study quantitative trait loci (QTL) often find support
562 for overdominance" and that pseudo-overdominance could account for findings of segregation
563 distortion among markers and genes with major effect on viability.
564 Hedrick et al. (2016) examined patterns of heterozygosity and fitness in 28 selfed
565 progeny derived from one parent *Eucalyptus grandis* tree heterozygous at 9590 marker loci.
566 They found high inbreeding depression on 10 of 11 chromosomes, losses of particular
567 homozygous segments, and high variance in heterozygosity over chromosomes. In contrast, wild
568 trees showed low identity disequilibrium and no excess heterozygosity. They concluded that
569 100+ load loci probably interacted to generate pseudo-overdominance, generating the high
570 inbreeding depression observed.
571 Gilligan et al. (2005) experimented with small populations of *Drosophila*. They found
572 that quantitative and molecular genetic variation declined at lower rates than expected, a result
573 they attributed to overdominance. In seeking to account for slower losses of heterozygosity than
574 expected over 200-generation bottlenecks in *Drosophila* populations, Latter (1998) simulated the

evolution of linked sets of mutations of minor effect (***s***=1-2%). He found a rapid buildup in heterozygote advantage (pseudo-overdominance) of 5-10% sufficient to stabilize allele frequencies in accord with his experimental results. Charlesworth (2015) concluded that although individual coding mutations are often weakly deleterious, balancing selection generates much of the fitness variation observed in *Drosophila*.

Genomic tools are now being used to measure inbreeding and ID and to pinpoint loci with large inbreeding effects (Kardos et al. 2016). In 11 maize lines selfed over six generations, Roessler et al. (2019) found heterozygosity to decline less than expected (35-40% per generation vs. 50%), a result they attributed to "pervasive associative overdominance." Brandenburg et al. (2017) similarly used 22.2M SNPs to analyze variation in 67 maize genomes. They found 6978 segments representing 8.8% of the genome with unexpectedly high rates of heterozygosity. These contributed to inbreeding depression with deleterious variants occurring more often within the heterozygous segments than elsewhere in the genome.

Li et al. (2015) set out to rigorously assess four alternative hypotheses to explain heterosis for height in sorghum. They used genetic mapping, a genome-wide association study, and controlled crosses between recombinant inbred lines to identify QTLs for height on chromosome 7 and 9. They confirmed pseudo-overdominance between two key genes located 29 cM (3 Mb) apart. Lariépe et al. (2012) also used QTL mapping to localize QTLs for heterotic effects in maize. Tracking heterozygosity, they found most QTLs for grain yield showed apparent overdominance. These QTLs tended to be located near centromeres where low recombination would conserve repulsion linkage.

Seymour et al. (2016) set out specifically to test dominance versus overdominance as the genetic basis for heterosis in the inbred model plant, *Arabidopsis thaliana*. This annual

predominantly self-fertilizes but wild populations maintain a low level of outcrossing (estimated at 10% by Bomblies 2010). They intercrossed 30 parent plants (spanning much of the genetic diversity found across this species' natural range) in a diallel to analyze genome-wide associations with fitness traits. This assessed dominance and overdominance effects on both trait means and mid-parent heterosis. Heterosis proved to be more predictable than mean effects. They found overdominance for all traits and a general lack of positive correlations between non-additive effects and genetic distance (expected under the dominance hypothesis – Charcosset et al. 1991). The tested SNPs accounted for 35-45% of the total genetic variance of late fitness traits in the overdominant model with 35 sites individually significant. These mapped to three specific overdominant genomic regions.

Brown and Kelly (2020) used genomic tools to uncover remarkably large inbreeding load in *Mimulus guttatus* distributed mostly as rare alleles within clusters distributed in clumps across the genome, matching the POD hypothesis. Gilbert et al. (2020) derived conditions under which background selection could transition to pseudo-overdominance and used simulations to show that this could greatly increase diversity in low-recombination regions. They identified 22 low-recombination regions in the human genome consistent with multi-locus overdominance. Further work should seek to identify just how recombination cold-spots (including inversions) may generate pseudo-overdominance.

Chelo et al. (2019) compared load dynamics in more and less inbred lines of *Caenorhabditis elegans* using high-quality SNP data. Over 50 generations of selfing, populations gained the ability to survive under a salt-stress treatment, a result they interpreted as evidence of purging. They were thus surprised to find that androdieoecious lines sustained most of their genetic variability despite being inbred to a level of $\boldsymbol{F} = 0.77$. Both expected heterozygosity and

621 the effective numbers of haplotypes far exceeded what was expected under neutral drift or
622 purging dynamics. These lines also developed high and variable levels of identity disequilibrium
623 with ***g2*** increasing from 0.07 to 2.5. Even monoecious lines inbred to $\boldsymbol{F} = 0.99$ retained half their
624 allele diversity. As only mean levels were reported, it is unclear whether allele frequency spectra
625 became bimodal in inbred populations as expected if PODs were sheltering load. Chelo et al.
626 found these results hard to reconcile with current theoretical studies but noted they were
627 compatible with associative overdominance generated by closely linked deleterious recessive
628 alleles in linkage disequilibrium. In fact, recessive semi-lethal mutations linked in strong
629 repulsion occur in wild populations of *C. elegans* (Seidel et al. 2008). Bernstein et al. (2019)
630 found similar evidence for large-effect variants linked in repulsion in crossed lines of *C. elegans*.
631 Together, these studies strongly support the existence and strength of pseudo-
632 overdominant effects in many species. Our ability to characterize and map PODs means we
633 could also compare their number, location, and strength among populations to assess how these
634 affect levels of ID and selection on mating systems. We could also assess their stability among
635 diverging lineages.
636
637 *Conclusion*
638 The qualitative model outlined here could help to explain why inbreeding depression and
639 mixed mating systems persist under inbreeding and why exclusively self-fertilizing taxa are
640 vanishingly scarce. Despite accumulating evidence and frequent mention of how pseudo-
641 overdominance caused by linked deleterious mutations could be relevant, few have emphasized
642 how general this phenomenon may be or its broad relevance to mating system evolution. Many
643 of the explanations proposed to explain why ID and selection for outcrossing persists appear

644 rather *ad hoc* or depend on particular ecological or genetic circumstances. While a rich portfolio
645 of hypotheses and ideas is always useful, it would be more parsimonious to find a general
646 mechanism acting regularly under a broad set of circumstances to perpetuate inbreeding
647 depression and selection for outcrossing. Because they can persist even in highly inbred
648 populations, PODs could account for these phenomena. Focused simulation and genomic studies
649 could test these ideas more rigorously.

650 These results reviewed above suggest that PODs exist and can be important. If PODs act
651 to sustain inbreeding depression and outcrossing, they could provide a mechanism to explain
652 why mixed mating systems persist and why fully self-fertilizing species do not exist. PODs
653 might thus act as a brake to limit how often inbred lineages lose genetic variability,
654 compromising their adaptability or risking "mutational meltdown." Although clade selection
655 appears to sometimes act against predominantly selfing lineages and taxa (see *Introduction*), it
656 may occur less often than it would if PODs did not exist to interrupt the march to increased
657 selfing. PODs may thus help to explain why at least opportunistic outcrossing predominates in
658 nature.

659

660

661 **Acknowledgements:** I dedicate this paper to my late friend and mentor, James Crow, who
662 inspired deep thinking about mutations and inbreeding. I have no conflicts of interest. NSF (DEB
663 97-28855) supported my empirical research on inbreeding. R. Toczydlowski's empirical work
664 spurred me to write up 15-year old ideas. She, D. Abu-Awad, B. Charlesworth, P. David, J.
665 Doebley, P. Hedrick, K. Holsinger, P. Jarne, L. Keller, J. Kelly, M. Kirkpatrick, T. Mitchell-
666 Olds, R. Waples, and two anonymous reviewers provided useful comments on the manuscript.

## References Cited


Abu Awad, D., and D. Roze. 2018. Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. Evolution 72:751–769.

Barrett, S. C. H. and D. Charlesworth. 1991. Effects of a change in the level of inbreeding on the genetic load. Nature 352: 522-524.

Baldwin, S. J. and D. J. Schoen. 2019. Inbreeding depression is difficult to purge in self-incompatible populations of *Leavenworthia alabamica*. New Phytol. doi.org/10.1111/nph.15963.

Barton, N. H. 1995. Linkage and the limits to natural selection. Genetics 140: 821-841.

Barton, N. H., and S. P. Otto, 2005 Evolution of recombination due to random drift. Genetics 169: 2353–2370.

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

Bernstein, M. R., S. Zdraljevic, E. C. Andersen, and M. V. Rockman. 2019. Tightly linked antagonistic-effect loci underlie polygenic phenotypic variation in *C. elegans*. Evol. Lett. 3: 462–473.

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

Bierne, N., A. Tsitrone, and P. David, 2000 An inbreeding model of associative overdominance during a population bottleneck. Genetics 155: 1981–1990.

Bomblies K, et al. 2010. Local-scale patterns of genetic variability, outcrossing, and spatial structure in natural stands of *Arabidopsis thaliana*. PLoS Genet 6(3):e1000890.

Brandenburg, J. T., T. Mary-Huard, G. Rigaill, S. J. Hearne, H. Corti, J. Joets, C. Vitte, A. Charcosset, S. D. Nicolas, and M. I. Tenaillon. 2017. Independent introductions and admixtures have contributed to adaptation of European maize and its American counterparts. PLoS Genet. 13:1–30.

Brown, K. E., and J. K. Kelly. 2020. Severe inbreeding depression is predicted by the "rare allele load" in Mimulus guttatus. Evolution 74:587–596.

Byers, D. L., and D. M. Waller. 1999. Do plant populations purge their genetic load? Effects of population size and mating history on inbreeding depression. Annu. Rev. Ecol. Syst. 30: 479-513.

Carr, D. E. and M. R. Dudash. 2003. Recent approaches into the genetic basis of inbreeding depression in plants. Phil. Trans. R. Soc. Lond. B 358: 1071–1084.

Carr, D. E., C. B. Fenster, and M. R. Dudash. 1997. The relationship between mating-system characters and inbreeding depression in *Mimulus guttatus*. Evolution 51: 363-372.

Chang, S., and M. D. Rausher. 1998. Frequency dependent pollen discounting contributes to maintenance of a mixed mating system in the common morning glory *Ipomoea purpurea*. Am. Nat. 152:671–683.

Charcosset, A., M. Lefort-Buson, and A. Gallais. 1991. Relationship between heterosis and heterozygosity at marker loci: A theoretical computation. Theor. Appl. Genet. 81:571–575.

Charlesworth, D. 2006. Balancing selection and its effects on sequences in nearby genome regions. PLoS Genetics, *2*, e2.

Charlesworth, B. 2012. The effects of deleterious mutations on evolution at linked sites. Genetics 190: 5–22.

Charlesworth, B. 2015. Causes of natural variation in fitness: Evidence from studies of *Drosophila* populations. Proc Natl Acad Sci USA 112:1662–1669.

Charlesworth, B. 2018. Mutational load, inbreeding depression and heterosis in subdivided populations. Mol. Ecol. 27: 4991–5003.

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

Charlesworth, B. and D. Charlesworth. 1997. Rapid fixation of delerious alleles by Muller's ratchet. Genet. Res. Cambridge 70: 63-73.

Charlesworth, B. and D. Charlesworth. 1999. The genetic basis of inbreeding depression. Genet. Res. 74: 329-340.

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

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

Charlesworth, D. and J. H. Willis. 2009. The genetics of inbreeding depression. Nat. Rev. Genet. 10: 783-796.

Chelo, I. M., B. Afonso, S. Carvalho, I. Theologidis, C. Goy, A. Pino-Querido, S. R. Proulx, and H. Teotónio. 2019. Partial selfing can reduce genetic loads while maintaining diversity during experimental evolution. G3 (Bethesda). 9: 2811–2821.

Cheptou, P.-O. 2019. Does the evolution of self-fertilization rescue populations or increase the risk of extinction? Ann. Bot. 123: 337–345.

Clo, J., L. Gay, and J. Ronfort. 2019. How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species. Evolution 73:1578–1590.

Crow, J. F. 1958. Some possibilities for measuring selection intensities in man. Hum. Biol. 30:1–13.

Crow, J. F. 1999. The rise and fall of overdominance. Plant Breeding Reviews 17:225-257.

Crow, J. F. and M. Kimura. 1970. An Introduction to Population Genetics Theory. Burgess Pub. Co., Minneapolis, Minn.

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

Fisher, R. A. 1965. The Theory of Inbreeding. 2nd Edition, Oliver & Boyd, London.

764 Frydenberg, O. 1963. Population studies of a lethal mutant in *Drosophila melanogaster* I. Behaviour in
765 populations with discrete generations. Hereditas 48: 89–116.
766
767 Gabriel, W., M. Lynch, and R. Bürger. 1993. Muller's ratchet and mutational meltdowns. Evolution (N. Y). 47:
768 1744–1757.
769
770 Gao, Z., M. Przeworski, and G. Sella, G. 2015. Footprints of ancient balanced polymorphisms in genetic
771 variation: Data from closely related species. Evolution 69: 431–466. doi.org/10.1111/evo.12567
772
773 Gilbert, K. J., F. Pouyet, L. Excoffier, and S. Peischl. 2020. Transition from Background Selection to Associative
774 Overdominance Promotes Diversity in Regions of Low Recombination. Curr. Biol. 30:101-107.e3.
775 Elsevier.
776
777 Gilligan, D. M., D. A. Briscoe, and R. Frankham. 2005. Comparative losses of quantitative and molecular genetic
778 variation in finite populations of Drosophila melanogaster. Genet. Res. Cambridge 85: 47–55.
779
780 Glémin, S. 2003. How are deleterious mutations purged? Drift versus nonrandom mating. Evolution 57: 2678-.
781
782 Goodwillie, C., S. Kalisz, and C. G. Eckert. 2005. The evolutionary enigma of mixed mating systems in plants:
783 Occurrence, theoretical explanations, and empirical evidence. Annual Review of Ecology Evolution &
784 Systematics 36:47-49.
785
786 Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith et al. 2010. Species selection maintains self-
787 incompatibility. Science 330: 493– 495. https://doi.org/10.1126/science.1194513
788
789 Haldane, J. B. S. 1932. The causes of evolution. Harper & Row, New York.

790 Harkness, A., Y. Brandvain, and E. E. Goldberg. 2019. The evolutionary response of mating system to
791 heterosis. J. Evol. Biol. 32: 476–490.

792 Hartfield, M., and S. Glémin. 2016. Limits to adaptation in partially selfing species. Genetics 203:959–
793 974.

794 Hedrick PW. 2012. What is the evidence for heterozygote advantage selection? Trends in Ecology and Evolution
795 27: 698–704.
796
797 Hedrick, P. W. and A. Garcia-Dorado. 2016. Understanding inbreeding depression, purging, and genetic rescue.
798 Trends in Ecology & Evolution 31: 940-952.
799
800 Hedrick, P. W., U. Hellsten, and D. Grattapaglia. 2016. Examining the cause of high inbreeding depression:
801 Analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. New Phytol. 209:
802 600–611.
803
804 Hill, W. G. and A. Robertson. 1968. The effect of linkage on limits to artificial selection. Genet. Res. 8: 269-294.
805
806 Holsinger, K. 1991. Mass-action models of plant mating systems: the evolutionary stability of mixed mating
807 systems. Amer. Nat. 138: 606-622.
808
809 Hudson, R.R. 2015. A New Proof of the Expected Frequency Spectrum under the Standard Neutral Model. PLOS
810 ONE 10(7): 1–5.

Husband, B. C., and D. W. Schemske. 1996. Evolution of the magnitude and timing of inbreeding depression in plants. Evolution (N. Y). 50:54–70.

Johnston, M.O., E. Porcher, P-O. Cheptou, C.G Eckert, E. Elle, M.A Geber, S. Kalisz, J.K Kelly, D.A Moeller, M. Vallejo-Marin, and A.A. Winn,. 2008. Correlations among fertility components can maintain mixed mating in plants. American Naturalist 173:1-11.

Johnston, M. O., and D. J. Schoen. 1996. Correlated evolution of self-fertilization and inbreeding depression: an experimental study of nine populations of *Amsinckia* (Boragainaceae). Evolution 50: 1478–1491.

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

Kardos, M., H. R. Taylor, H. Ellegren, G. Luikart, and F. W. Allendorf. 2016. Genomics advances the study of inbreeding depression in the wild. Evolutionary Applications 9: 1205-1218.

Keller, L. F. and D. M. Waller. 2002. Inbreeding effects in the wild. Trends in Ecology and Evolution 17: 230-241.

Kelly, J. K., and H. S. Arathi. 2003. Inbreeding and the genetic variance in floral traits of *Mimulus guttatus*. Heredity (Edinb). 90: 77–83.

Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. Genetics 49:725–738.

Kimura, M., T. Maruyama, and J. Crow. 1963. The mutation load in small populations. Genetics 48: 1303–1312.

Lande, R., and E. Porcher. 2015. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. Genetics 200:891–906.

Lande, R., and D. W. Schemske. 1985. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic model. Evolution 39: 24-40.

Lande, R., D. W. Schemske, and S. T. Schultz. 1994. High inbreeding depression, selective interference among loci, and the threshold selfing rate for purging recessive lethal mutations. Evolution (N. Y). 48: 965–978.

Larièpe, A., B. Mangin, S. Jasson, V. Combes, P. J. F. Dumas, C. Lariagon, D. Jolivot, D. Madur, J. Fiévet, A. Gallais, P. Dubreuil, A. Charcosset, and L. Moreau. 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 190:795–811.

Latter, B. D. H. 1998. Mutant alleles of small effect are primarily responsible for the loss of fitness with slow inbreeding in *Drosophila melanogaster*. Genetics 148: 1143–1158.

Li, X., X. Li, E. Fridman, T. T. Tesso, J. Yu, and R. L. Phillips. 2015. Dissecting repulsion linkage in the dwarfing gene Dw3 region for sorghum plant height provides insights into heterosis. Proc. Natl. Acad. Sci. U. S. A. 112:11823–11828.

Lewontin, R. C. 1974. The Genetic Basis of Evolutionary Change. Columbia University Press, New York.

Lohr, J. N., and C. R. Haag. 2015. Genetic load, inbreeding depression, and hybrid vigor covary with population size: An empirical evaluation of theoretical predictions. Evolution 69:3109–3122.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Assoc., Sunderland, MA.

Maynard Smith, J. 1978. The Evolution of Sex. Cambridge Univ. Press, New York.

McMullen, M. D., S. Kresovich, H. S. Villeda, P. Bradbury, H. Li, Q. Sun, S. Flint-Garcia, J. Thornsberry, C. Acharya, C. Bottoms, P. Brown, C. Browne, M. Eller, K. Guill, C. Harjes, D. Kroon, N. Lepak, S. E. Mitchell, B. Peterson, G. Pressoir, S. Romero, M. O. Rosas, S. Salvo, H. Yates, M. Hanson, E. Jones, S. Smith, J. C. Glaubitz, M. Goodman, D. Ware, J. B. Holland, and E. S. Buckler. 2009. Genetic properties of the maize nested association mapping population. Science 325:737–740.

McQuillan, R., A. L. Leutenegger, R. Abdel-Rahman, et al. 2008. Runs of homozygosity in European populations. American Journal of Human Genetics 83: 359–372.

Morton, N. E., J. F. Crow, and H. J. Muller. 1956. An estimate of the mutational damage in man from data on consanguineous marriages. Proceedings of the National Academy of Science 42: 855-863.

Noel, E., P. Jarne, S. Glémin, A. MacKenzie, A. Segard, V. Sarda, and P. David. 2017. Experimental evidence for the negative effects of self-fertilization on the adaptive potential of populations. Curr. Biol. 27:237–242.

Nordborg, M. B., B. Charlesworth, and D. Charlesworth. 1996. The effect of recombination on background selection. Genet. Res. Camb. 67: 159-174.

Oakley, C. G., J. P. Spoelhof, and D. W. Schemske. 2015. Increased heterosis in selfing populations of a perennial forb. AoB Plants 7: plv122.

Oakley, C. G., S. Lundemo, J. Ågren, and D. W. Schemske. 2019. Heterosis is common and inbreeding depression absent in natural populations of *Arabidopsis thaliana*. J. Evol. Biol. 32:592–603.

Oakley, C. G., and A. A. Winn. 2008. Population and family level inbreeding depression in a cleistogamous perennial. Int. J. Plant Sci. 169: 523–530.

O'Connor, L. J., A. P. Schoech, F. Hormozdiari, S. Gazal, N. Patterson, and A. L. Price. 2019. Extreme polygenicity of complex traits is explained by negative selection. Am. J. Hum. Genet. 105:456–476.

Ohta, T., 1971 Associative overdominance caused by linked detrimental mutations. Genet. Res. 18: 277–286.

Paland, S., and B. Schmidt. 2003. Population size and the nature of genetic load in *Gentianella germanica*. Evolution 57: 2242–2251.

Pálsson, S., and P. Pamilo. 1999. The effects of deleterious mutations on linked, neutral variation in small populations. Genetics 153: 475–483.

Pamilo, P., and S. Pálsson. 1998. Associative overdominance, heterozygosity and fitness. Heredity. 81: 381–389.

Remington, D. L. and D. O. O'Malley. 2000. Whole-genome characterization of embryonic stage inbreeding depression in a selfed Loblolly Pine family. Genetics 155: 337-348.

912 Richards, C. M. 2000. Inbreeding depression and genetic rescue in a plant metapopulation. Amer. Nat. 155:383–
913 394.
914
915 Roessler, K., A. Muyle, C. M. Diez, G. R. J. Gaut, A. Bousios, M. C. Stitzer, D. K. Seymour, J. F. Doebley, Q.
916 Liu, and B. S. Gaut. 2019. The genome-wide dynamics of purging during selfing in maize. Nature Plants
917 5: 980–990.
918
919 Roose, M.L. and L.D. Gottlieb. 1976. Genetic and Biochemical Consequences of Polyploidy in *Tragopogon*.
920 Evolution 30: 818–830. doi:10.2307/2407821
921
922 Roze, D. 2015. Effects of interference between selected loci on the mutation load, inbreeding depression, and
923 heterosis. Genetics 201: 745–757.
924
925 Schemske, D. W., and R. Lande. 1985. The evolution of self-fertilization and inbreeding depression in plants.
926 II. Empirical observations. Evolution 39: 41-52.
927
928 Schnee, B. K. and D. M. Waller. 1986. Reproductive behavior of *Amphicarpaea bracteata* (Leguminosae), an
929 amphicarpic annual. Am. J. Bot. 73: 376-386.
930
931 Schultz, S. T., and M. Lynch. 1997. Mutation and extinction: the role of variable mutational effects, synergistic
932 epistasis, beneficial mutations, and degree of outcrossing. Evolution 51: 1363–1371.
933
934 Seidel, H. S., M. V. Rockman, and L. Kruglyak. 2008. Widespread genetic incompatibility in *C. elegans*
935 maintained by balancing selection. Science 319: 589–594. https://doi.org/10.1126/science.1151107
936
937 Seymour, D. K., E. Chae, D. G. Grimm, C. M. Pizarro, A. Habring-Müller, F. Vasseur, B. Rakitsch, K. M.
938 Borgwardt, D. Koenig, and D. Weigel. 2016. Genetic architecture of nonadditive inheritance in
939 *Arabidopsis thaliana* hybrids. Proc. Natl. Acad. Sci. U. S. A. 113:E7317–E7326.
940
941 Siewert, K. M., and B.F. Voight. 2017. Detecting long-term balancing selection using allele frequency correlation.
942 Molecular Biology and Evolution 34: 2996–3005. doi.org/10.1093/molbev/msx209
943
944 Smith, N. M. A., C. Wade, M. H. Allsopp, B. A. Harpur, A. Zayed, S. A. Rose, J. Engelstädter, N. C. Chapman,
945 B. Yagound, and B. P. Oldroyd. 2019. Strikingly high levels of heterozygosity despite 20 years of
946 inbreeding in a clonal honey bee. J. Evol. Biol. 32:144–152.
947
948 Spigler, R. B., K. Theodorou, and S.-M. Chang. 2017. Inbreeding depression and drift load in small populations at
949 demographic disequilibrium. Evolution 71: 81–94.
950
951 Stebbins, G. L. 1957. Self fertility and population variability in the higher plants. American Naturalist 91 :337-
952 354.
953
954 Templeton, A. R., and B. Read. 1984. Factors eliminating inbreeding depression in a captive herd of Speke's
955 gazelle (*Gazella Spekei*). Zoo Biol. 3:177–199.
956
957 Tenaillon, M. I., F. Austerlitz, and O. Tenaillon. 2008. Apparent mutational hotspots and long distance linkage
958 disequilibrium resulting from a bottleneck. J. Evol. Biol. 21: 541–550.
959
960 Toczydlowski, R.H. 2019. Genetic and phenotypic differentiation in *Impatiens capensis* in riverine networks:
961 identifying patterns and potential drivers of gene flow, local adaptation, and inbreeding. PhD dissertation,
962 Univ. of Wisconsin-Madison.

963
964 Toczydlowski, R.H., and D.M. Waller. Inbreeding depression and drift load interact to affect fitness and
965 heterozygosity-fitness correlations in natural populations of *Impatiens capensis*. Submitted, Evolution.
966
967 Uyenoyama, M. K. and D. M. Waller. 1991. Coevolution of self-fertilization and inbreeding depression II.
968 Symmetric overdominance in viability. Theoretical Population Biology 40: 47-77.
969
970 Uyenoyama, M. K., K. E. Holsinger, and D. M. Waller. 1993. Ecological and genetic factors directing the
971 evolution of self-fertilization. Oxford Surv. Evol. Biol. 9:327–381.
972
973 vanOosterhout, C. et al. 2000. Inbreeding depression and genetic load in laboratory metapopulations of the
974 butterfly *Bicyclus anynana*. Evolution 54: 218–225.
975
976 Waller, D. M. 1986. Is there disruptive selection for self-fertilization? American Naturalist 128: 421-426.
977
978 Waller, D.M., and L.F. Keller. 2020. Inbreeding and inbreeding depression. Oxford Encyclopedia of
979 Evolutionary Biology. DOI: 10.1093/obo/9780199941728-0124
980
981 Wang, J. L. 2000. Effects of population structures and selection strategies on the purging of inbreeding depression
982 due to deleterious mutations. Genetical Research (Cambridge) 76: 75-86.
983
984 Wang, B., J. Mojica, N. Perera, C.-R. Lee, J. Lovell, A. Sharma, C. Adam, A. Lipzen, K. Barry, D. Rokhsar, J.
985 Schmutz, and T. Mitchell-Olds. 2019. Ancient polymorphisms contribute to genome-wide variation by
986 long-term balancing selection and divergent sorting in *Boechera stricta*. Genome Biol. 20: 126.
987
988 Whitlock, M., P. K. Ingvarsson, and T. Hatfield. 2000. Local drift load and the heterosis of interconnected
989 populations. Heredity 84: 452-457.

990 Willi, Y. 2013. Mutational meltdown in selfing *Arabidopsis lyrata*. Evolution 67: 806- 815.

991 Willi, Y., Griffin, P. & Van Buskirk, J. 2013. Drift load in populations of small size and low density.
992 *Heredity (Edinb)*. 110: 296–302.

993 Willis, J. H. 1999. Inbreeding load, average dominance and the mutation rate for mildly deleterious alleles in
994 *Mimulus guttatus*. Genetics 153: 1885-1898.
995
996 Winn, A. A., E. Elle, S. Kalisz, P.-O. Cheptou, C. G. Eckert, C. Goodwillie, M. O. Johnston, D. A. Moeller, R. H.
997 Ree, R. D. Sargent, and M. Vallejo-Marin. 2011. Analysis of inbreeding depression in mixed-mating
998 plants provides evidence for selective interference and stable mixed mating. Evolution 65: 3339-3359.
999
1000 Wright S. 1931. Evolution in Mendelian populations. Genetics 16: 97–159.
1001
1002 Zhao, L., and B. Charlesworth. 2016. Resolving the conflict between associative overdominance and
1003 background selection. Genetics 203: 1315-1332.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

**Table 1**. Simulation results showing the number of major and minor mutations occurring along a single pair of chromosomes. Here, map locations from 1-1000 on each of a pair of chromosomes were randomly assigned to 5 major and 200 minor mutations subsequently sorted into 10 bins. Alleles are assumed to be tightly linked within bins but unlinked among bins. Selection would act efficiently against major mutations (***s*** between 0.5 and 1), especially when these occur in coupling on the same chromosome (bins 1 and 9). Those linked in repulsion (here, in bin 4 where mutations occurred at locations 347 and 369) express pseudo-overdominance, sustaining variation by strongly selecting against both homozgotes. Minor mutations (***s*** ~ 0.02), being more common, occur in more consistent numbers among bins and between the two chromosomes (Fig. 3). This balanced occurrence makes it unlikely that multiple linked mutations will only occur on one of a pair of chromosomes, allowing efficient purging, while making it more likely that many will be linked in repulsion (generating pseudo-overdominance).

| | **Chromosome 1a mutations** | | **Chromosome 1b mutations** | |
|---|---|---|---|---|
| **Map Bin** | **Major** | **Minor** | **Major** | **Minor** |
| 1 | **3** | 20 | 0 | 18 |
| 2 | 0 | **16** | 1 | 20 |
| 3 | 1 | **23** | 0 | **10** |
| **4** | **1** | 20 | **1** | **12** |
| 5 | 0 | 24 | 1 | 21 |
| 6 | 0 | 24 | 0 | 20 |
| 7 | 0 | 20 | 0 | 30 |

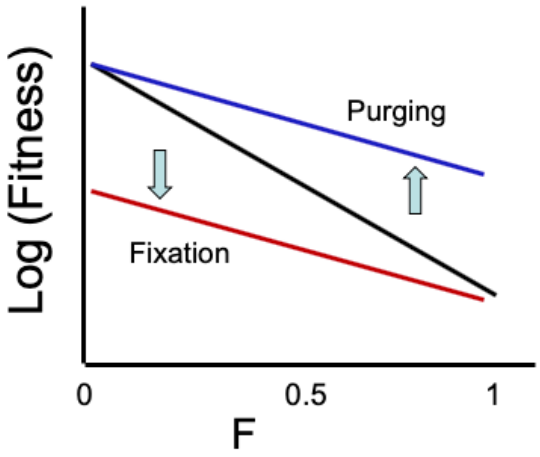
*Preprints* (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

| | | | | |
|---|---|---|---|---|
| 8 | 0 | 16 | 0 | 23 |
| 9 | 0 | 20 | **2** | 20 |
| 10 | 0 | 17 | 0 | 26 |

1017

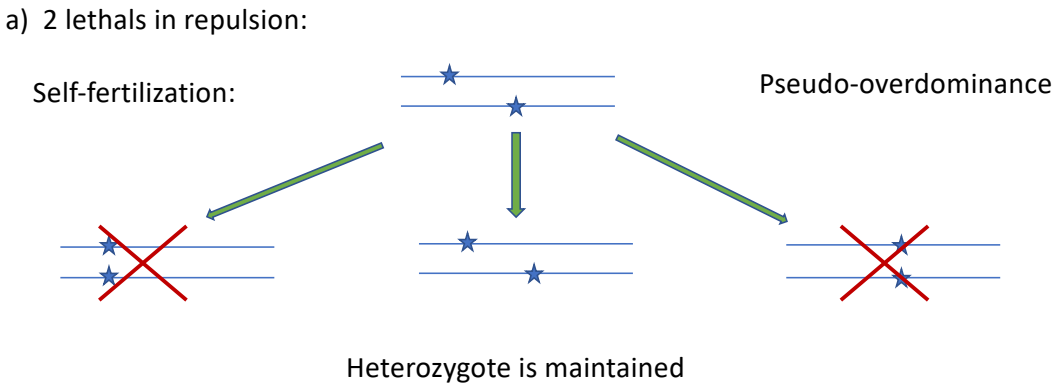
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

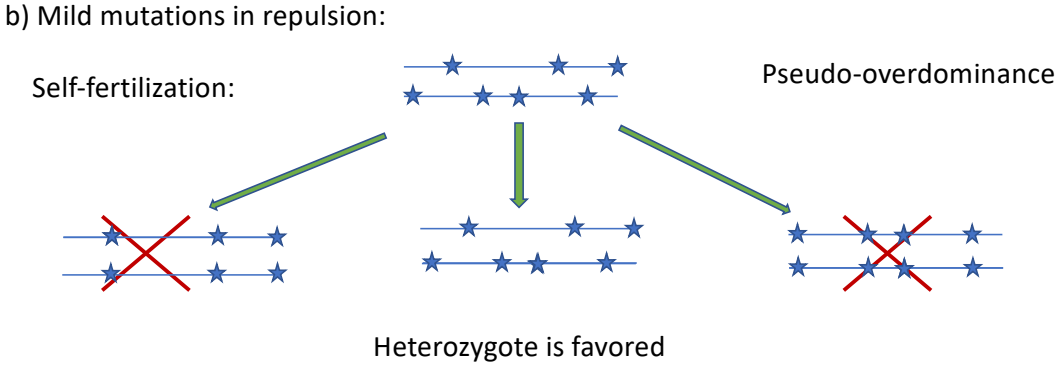
**Figure 1**. Expected relationship between (log) fitness and an individual's coefficient of inbreeding, ***F***, within a population. Multiplicative fitness interactions among mutations cause linear declines in log fitness with ***F*** (Morton et al. 1956). Selection against deleterious recessive alleles purges the genetic load, increasing fitness especially of inbred progeny (blue line hinges up). In contrast, fixing deleterious alleles via drift reduces outbred fitness (red line hinges down). Because both purging and fixation reduce inbreeding load (the slope ***-B***), reductions in inbreeding depression cannot distinguish between these mechanisms.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

**Figure 2**. Two pathways to pseudo-overdominance and a recalcitrant load. a) If two recessive lethal mutations are linked in proximity to each other along the same chromosome but occur in repulsion, selection among the self-fertilized progeny generated eliminates any progeny homozygous for either allele, maintaining the heterozygote and any closely linked alleles at an intermediate frequency. b) Selection against strongly deleterious mutations could still leave a residue of several mildly deleterious mutations linked in repulsion between chromosomal segments. Effects of individual recessive mutations would be masked in the heterozygous parent but could express significant combined deleterious effects within selfed progeny as homozygotes. The resulting pseudo-overdominance could sustain these mutations at intermediate frequency and perpetuate inbreeding load.

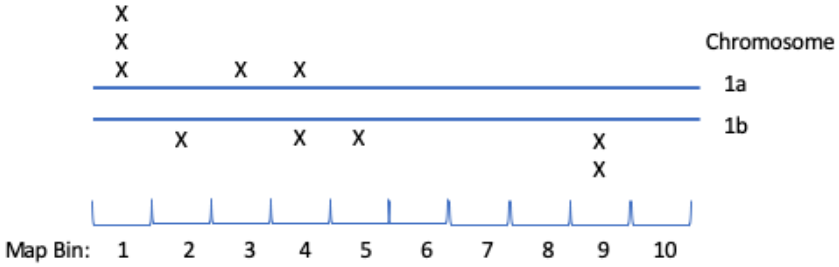


Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 October 2020 doi:10.20944/preprints202010.0634.v1

**Figure 3**. The numbers of randomly distributed major (a) and minor (b) mutations found in each of 10 bins matching map distance along a pair of chromosomes (simulated data from Table 1). a) Note the balanced major mutation load at bin #4 where 2 lethal recessive mutations could be indefinitely maintained even under close inbreeding. b) The higher numbers and density of mildly deleterious mutations ensure that most chromosomal segments contain balanced sets of mildly deleterious mutations.

a) Major mutations




b) Minor mutations