# Early-acting inbreeding depression can evolve as an inbreeding avoidance mechanism.

**Yaniv Brandvain**[1,✉], **Lia Thomson**[1,2], **and Tanja Pyhäjärvi**[3]

[1]Department of Plant & Microbial Biology. University of Minnesota - Twin Cities. St Paul, MN
[2]School of Life Sciences. Arizona State University. Temple, AZ.
[3]Department of Forest Sciences. University of Helsinki. Helsinki, Finland.

**Despite the potential for mechanical, developmental and/or chemical mechanisms to prevent self-fertilization, incidental self-fertilization is inevitable in many predominantly outcrossing species. In such cases, inbreeding can compromise individual fitness. Unquestionably, much of this inbreeding depression is maladaptive. However, we show that when reproductive compensation allows for the replacement of inviable embryos lost early in development, selection can favor deleterious recessive variants that induce "self-sacrificial" death of inbred embryos. Our theoretical results provide numerous testable predictions which could challenge the assumption that inbreeding depression is always maladaptive. Our work is applicable any species that cannot fully avoid inbreeding, exhibits substantial inbreeding depression, and has the potential to compensate embryos lost early in development. In addition to its general applicability our theory suggests that self-sacrificial variants might be responsible for the remarkably low realized selfing rates of gymnosperms with high primary selfing rates, as gymnosperms exhibit strong inbreeding depression, have effective reproductive compensation mechanisms, and cannot evolve chemical self-incompatibility.**

**Inbreeding depression | Self Incompatibility | Altruism | Reproductive compensation**

**Correspondence:** ***ybrandva@umn.edu***

## Introduction

Inbreeding depression caused by both rare, recessive, and highly deleterious mutations (1, 2), and by deviations from phenotypic optima under multi-locus stabilizing selection (3–6), is an inevitable consequence of habitual outcrossing. As such, inbreeding avoidance mechanisms, including chemically mediated self-incompatibility (SI) systems, are common across the tree of life (7, 8). The mechanisms and timing of SI are diverse (9), but the best-characterized cases involve an interaction between the diploid maternal genotype and the pollen's haploid (gametophytic SI) or diploid (sporophytic SI) genotype before fertilization.

Late Acting SI – in which self-pollen can effectively germinate, grow down the stigma, and in some cases can enter and fertilize the ovule before ovular development arrests (10–13) – is less well understood (14). In the most delayed form of Late Acting SI, "Post-zygotic" SI occurs after fertilization (15). Claims of Late Acting SI, and post-zygotic SI in particular, have attracted much attention and criticism both because this mechanism of incompatibility is unusual (14), and because it is difficult to distinguish late-acting SI from early-acting inbreeding depression (9, 16). Traditionally, the key distinction between these mechanisms is that SI is "*a mechanism operated by the maternal parent plant and controlled by the genotypes of that plant and its pollen donor*" and that inbreeding depression is "*a process acting in the progeny zygote, determined by its own genotype*" (16). Substantial effort has gone into distinguishing between early-acting inbreeding depression and late-acting SI (reviewed in 14, 17), with evidence supporting late-acting SI in some taxa (e.g. 18–21) and early inbreeding depression (e.g. 22, 23) in others.

We argue that with high reproductive compensation (i.e. the exchange of dead embryos for potential viable embryos), the distinction between early-acting inbreeding depression and late-acting SI is not necessarily evolutionarily relevant because early-acting inbreeding can adaptively evolve as a mechanism to minimize reproductive investment in low fitness selfed seed. Reproductive compensation is likely common across taxa in which offspring receive parental resources after fertilization, or even when offspring compete among one another for a local resource. The large levels of maternal investment in humans (e.g. 24, 25) and other mammals (e.g. 26) has led to a substantial body of research on reproductive compensation and its consequences in animals. Reproductive compensation is also likely common among angiosperms (27), as most species produce fewer progeny than zygotes (28, 29), a disparity that is elevated in habitually outcrossing species (29). Additionally, reproductive compensation is likely built in to the system of "simple polyembryony" found in gymnosperms, as only one of numerous genetically distinct embryos matures within a seed. As such, in the many taxa with an opportunity for reproductive compensation, an early-expressed recessive load allows mothers to trade in inbred for outbred embryos and can essentially function like an SI system.

Homozygosity for rare recessive alleles serves as a reliable indicator of inbreeding – much like the similarity of pollen and style S-locus genotype serves as a reliable indicator of self vs. non-self pollen. As such, we further argue that all forms of SI can be considered as a special case of this sort of altruistic act as rejected pollen are sacrificing their ability to sire potentially viable seed. Because in many cases early-acting inbreeding depression will reflect the standard

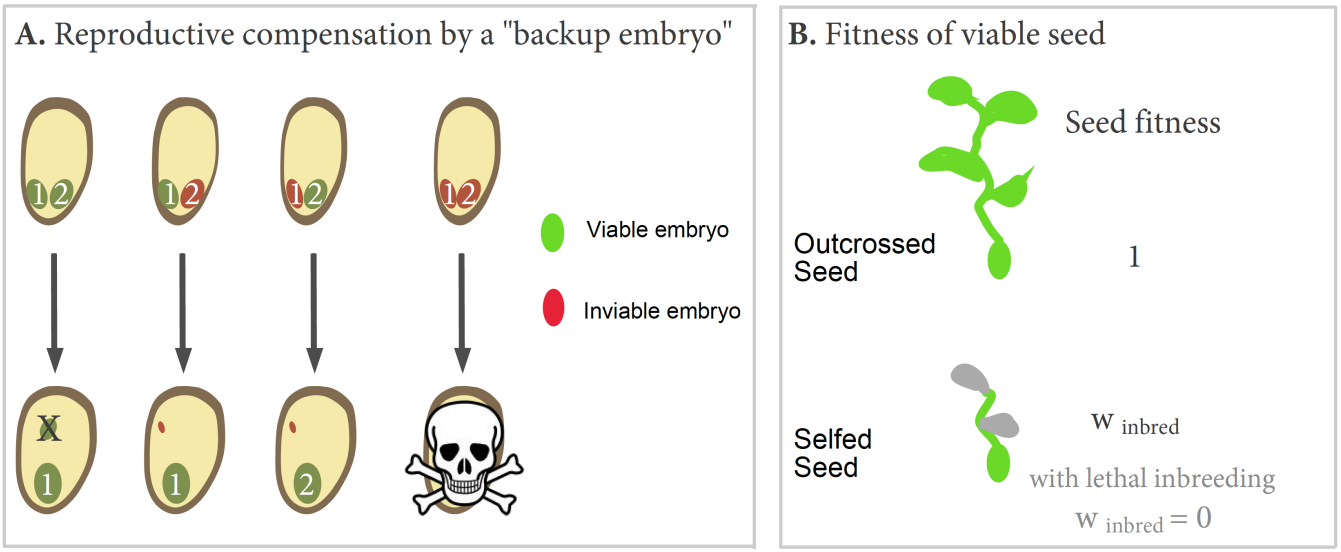


**Fig. 1. Stages of selection in the "backup embryo" model**. Figure 1A depicts our model of reproductive compensation. Here the backup embryo (embryo 2) is eliminated if both embryos are viable (noted in green), dies if inviable (noted in red), and replaces the primary embryo (embryo 1) if the primary embryo is inviable. If both embryos are inviable there is no reproductive compensation. Figure1 B Shows the second stage of selection in which the fitness of viable outcrossed seeds is one, and the fitness of viable selfed seed is $w_{\text{inbred}}$, which equals $0$ in the case of lethal inbreeding.

expectation of a balance of selection against deleterious alleles and their mutational input (i.e. traditional "mutation-selection-balance" ), rather than our novel hypothesis that selection directly favors such alleles, we provide numerous predictions which allow researchers to evaluate whether an early-expressed recessive load reflects selection for, or imperfect selection against it. Since our model makes very few critical assumptions – namely that homozygosity for a rare allele signals recent inbreeding, that inbreeding depression exists, and that early mortality can be compensated, we suggest that many taxa may express an altruistic recessive load.

## Methods

**Overview.** We model reproductive compensation as follows – each seed contains a primary and backup embryo, allowing an inviable 'primary' embryo to be replaced by another, potentially viable embryo (Figure 1A). Although this model assumes exactly two embryos per seed, our results do not depend on this model of reproductive compensation – in Figure 2 and the Supplementary Material, we show that analytical results from this "backup embryo" model are nearly identical to those obtained by using Charlesworth's (26), model of reproductive compensation, in which maternal fitness equals $\frac{1-p_m}{(1-p_m)^C}$, where $p_m$ is the probability that an embryo will not survive and $C$ is the extent of reproductive compensation (Described in more details in the Supplementary Text). All code and data associated with this manuscript are available through an associated zenodo repository (30).

Selection occurs at two distinct phases of the lifecycle. Selection first occurs within maternal families based on the recessive allele(s) underlying early embryo viability (i.e. between viable and inviable embryos in Figure 1A), and among maternal families based on number of fully developed seeds per mother (e.g. mothers with inviable primary and backup embryos will experience a reduction in fitness – Figure 1A). Next (after embryos can be compensated) selection occurs based on inbreeding status and the extent of inbreeding depression (Figure 1B). We assume that inbreeding depression is expressed only after seeds complete development (initially all embryos survive to become full seeds), that it is not purged at the population's mutation and selfing rates. This evolutionary stable load can occur when deleterious recessive alleles accumulate more rapidly than selection eliminates them (31), or when inbreeding leads to predictable deviations away from phenotypic optima in populations experiencing multi-locus stabilizing selection (4–6). Despite the ample evidence for early expressed inbreeding depression in nature, our model assumes its absence because (a) if there is more early inbreeding depression than our model predicts that a 'self-sacrificial allele' cannot invade, and if there is less, such an allele can invade with its equilibrium frequency lowered by the amount of pre-existing early inbreeding depression, and (b) Empirical cases of early acting inbreeding depression in the many taxa with a possibility for reproductive compensation could reflect the processes presented in our model.

First, to evaluate whether early expressed deleterious recessive alleles can be favored by natural selection, we analyze deterministic single-locus models without mutation or drift. Second, to evaluate the extent to which the adaptive altruistic load discovered in our single locus deterministic models can contribute meaningfully to the early expressed recessive load in a finite population experiencing recurrent mutation and genetic drift, we conduct stochastic, genome-scale simulations.

**Deterministic Single-Locus Models Without Mutation.** We initially assume that all primary embryos survive to become a seed, and that following seed development, seeds die if they are inbred and survive otherwise (we later relax the assumption that inbreeding is lethal). We then introduce a rare, highly recessive (with dominance coefficient, $h$) allele, $A$, which decreases embryo survival by $s$. Embryo survival is then converted into the expected maternal fitness following the model of reproductive compensation employed (Figure 1A,B)

We then explore the invasion of this allele under differing selection and dominance coefficients, and find the equilibrium frequency of such an allele in the absence of mutation. Throughout, we assume a population of infinite size in which mothers which incidentally self-fertilize (with probability, $\sigma$) or randomly mate (with probability $1-\sigma$). Details of our analytical setup are provided in the appendix.

**Stochastic Genome-Wide Models With Mutation.**

**Simulation framework:** We developed genome-scale, individual-based stochastic simulations to test whether (and to what extent) the adaptive altruistic load discovered in our analytical model could contribute to the early expressed load in a finite population experiencing recurrent mutation and genetic drift. Specifically, we ran non-Wright-Fisher forward simulations in SLiM 3.0 (32) in which early-acting recessive deleterious mutations recurrently arise (at rate $\mu$) at any location in the genome. We followed these simulations until populations achieved mutation-selection balance. We specify a genome of one hundred thousand sites, such that e.g. a mutation rate of $0.5 \times 10^{-6}$ corresponds to an expectation of one mutation per diploid genome per generation. These sites are uniformly spread over ten chromosomes, each of which is one Morgan in length with a uniform recombination map. We present results in the main text for a population of size ten thousand, and explore how changing this size impacts results in the supplementary material.
We explore results for a variety of selfing rates, and for modest ($\mu = 5 \times 10^{-8}$) high ($\mu = 5 \times 10^{-6}$) mutation rates, and for lethal (($w_{\text{inbred}}$ initial $= 0.00$) and sub-lethal (($w_{\text{inbred}}$ initial $= 0.25$) inbreeding depression. Because deterministic results from our "Backup embryo" model are nearly identical to results from those obtained using Charlesworth's (26) parameterization of reproductive compensation, we limit our computationally intensive whole genome simulations to our primary model of reproductive compensation.

**Evolving an altruistic load:** Each generation, 10,000 survivors were randomly sampled with replacement to act as mothers of two half-sib progeny. Each embryo was inbred (i.e. the product of selfing) with probability $\sigma$, the product of outcrossing (i.e. outbred) with probability $1-\sigma$. Next, early-acting deleterious mutations were expressed and embryos experienced viability selection based on their embryonic load. Finally, if both siblings survived, one was randomly considered "primary" and became a seed. The fitness of this seed was independent of its genotype and depended exclusively on its inbreeding status.

**A control:** As a control, we developed a second model where the early expressed load was not favored by natural selection. Again, two maternal siblings were produced from 10,000 mothers (sampled with replacement) every generation and embryo fitness depended on early-acting deleterious mutations. Unlike the first model, replacement did not occur between siblings; instead occurring between random embryo pairs. If both embryos in the pair survived, one was randomly considered "primary" and became a seed, this removes the opportunity for altruistic selection to favor an early-expressed genetic load. We finally instituted viability selection on surviving embryos as described above.

## Results

### Selection can favor recessive deleterious alleles.

**Invasion of a self-sacrificial mutation:** To evaluate invasion conditions in the "Backup embryo" model, we assume the $A$ allele is so rare that $AA$ homozygotes can be ignored, so we consider sibling pairs from $aa$ or $aA$ mothers. Outbred ovules are pollinated by $A$ or $a$ bearing pollen, with probabilities $p_A$ and $1-p_A$, respectively.

During the invasion of the recessive $A$ allele, (i.e. $\lim_{p_A \to 0}$) $aa$ mothers produce $aa$ offspring by selfing with probability $\sigma$ and produce $aa$ and $Aa$ offspring by outcrossing with probabilities $(1-p_A)(1-\sigma)$, and $(p_A)(1-\sigma)$, respectively. The probability that the primary embryo in an $aa$ maternal family is replaced by a backup embryo, $\psi_{aa}$, equals the product of the probability that an embryo is fertilized by $A$ pollen and dies in early development i.e. $h\,s\,p_A\,(1-\sigma)$ (See our mathematical appendix for details).

At invasion, $Aa$ mothers produce $aa$, $Aa$, and $AA$ offspring by selfing with probability $\frac{\sigma}{4}$, $\frac{\sigma}{2}$ and $\frac{\sigma}{4}$, respectively, and produce $aa$, $Aa$, and $AA$ offspring by outcrossing with probabilities $(1-p_A)\frac{(1-\sigma)}{2}, \frac{1-\sigma}{2}$ , and $p_A\frac{(1-\sigma)}{2}$, respectively. The probability that the primary embryo in an $Aa$ maternal family is replaced by a backup embryo is $\psi_{Aa}$, and equals $\frac{s}{2}(h+\sigma/2+p_A(1-\sigma))$.

At invasion, the population mean fitness is $1-\sigma$ (as we are assuming lethal inbreeding) and the mean fitness of rare $Aa$ maternal families, $\overline{w}_{\text{Aa:fam}}$ is

$$\overline{w}_{\text{Aa:fam},\lim_{p_A \to 0}} = (1-h\,s)(1-\sigma)(1+\frac{s}{4}(h+\frac{\sigma}{2}))$$

Assuming full recessivity (i.e. $h=0$), $\overline{w}_{\text{Aa:fam},\lim_{p_A \to 0}}$ is $\approx \frac{1}{8}(1-\sigma)(8+s\,\sigma)$, which always exceeds mean population fitness at invasion, and therefore a recessive $A$

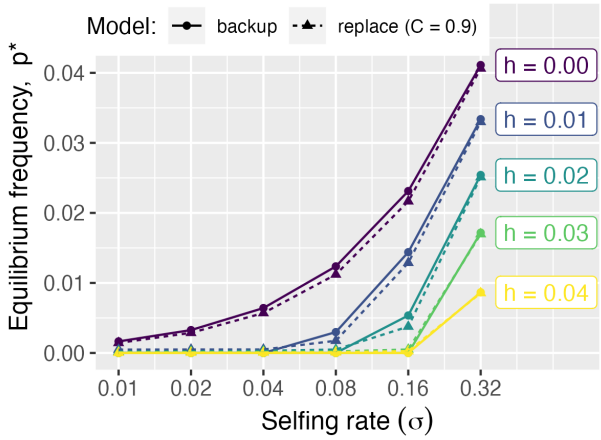


**Fig. 2.** The equilibrium frequency, $p^*$, of a self-sacrificial (partially) recessive allele for which lost embryos can be compensated, under a variety of selfing rates and dominance coefficients. Results are presented from our "backup embryo" model are presented with circles connected by solid lines). These results are nearly identical to those from Charlesworth's parameterization of reproductive compensation (triangles connected by dotted lines). in which maternal fitness equals $\frac{1-p_m}{(1-p_m)^C}$, where $p_m$ is the probability that an embryo will not survive and $C$ is the extent of reproductive compensation. This figure displays results for $C = 0.9$ and with $s = 0.6$. The difference between results with our and Charlesworth's parameterization of reproductive compensation widens as $C$ diverges from $0.9$, with $p^*$ decreasing as $C$ decreases (Figure S4). In this figure, we assume that the allele is lethal when homozygous, however, results are remarkably insensitive to $s$ (Figure S5).

is never disfavored when rare. The expected increase in frequency of a rare recessive allele that acts in this manner is $p_A(1-\sigma)(s\sigma)/(4\overline{w})$. Our results do not depend on the assumption of lethal inbreeding depression – – assuming that all outbred seed survive, and that the load cannot be purged, a recessive allele that decreases embryo fitness invades when the probability that an inbred offspring survives is less than $\frac{s(1-\sigma)}{4-s\sigma}$.

Similarly, complete recessivity is not required for the deterministic spread of this self-sacrificial allele. With lethal inbreeding, a partially recessive self-sacrificial allele is favored when the primary selfing rate exceeds a critical selfing level, which equals

$$\text{crit.self} = \frac{2h(3+hs)}{1-hs}$$

Figure S1 shows that at first approximation, the critical selfing rate required for the invasion of a self-sacrificial allele is linear with its recessivity ($h$), such that it will spread when its selfing rate, $\sigma$, exceeds $6h$.

Invasion criteria for a model with Charlesworth's parameterization of reproductive compensation are nearly identical to those obtained from our backup embryo model (Figures S2 and S3). Specifically, when $s$ and $h$ are small, invasion criteria from the backup model resembles cases in which reproductive compensation is very effective (between $C = 0.90$ and $C = 0.99$), while the backup model more closely resembles less effective reproductive compensation as the dominance coefficient is larger.

*Equilibrium frequency of a self-sacrificial mutation:* As the self-sacrificial allele increases in frequency, it begins to be exposed commonly in the allozygous, rather than the autozygous, state. As such, as $A$ increases in frequency, it becomes a less reliable indicator of inbreeding. To demonstrate this phenomenon, we apply Bayes' theorem to show that the probability that an embryo was inbred, given that it is homozygous for $A$ equals $\frac{\sigma p_A/2}{\sigma p_A/2+p_A^2(1-\sigma)}$ under the assumption of complete recessivity and lethality of the self-sacrificial allele. This results in frequency dependent selection in which the self-sacrificial allele is favored when rare but disfavored when common.

Because our model contains both family structure and departures from Hardy-Weinberg equilibrium, analytically deriving stable intermediate frequency, $p_A^*$, is quite challenging. Rather than deriving $p_A^*$, we iterate the genotype frequency recursion equations until genotype frequencies are stable. In both models, $A$'s equilibrium frequency is highest when it is highly recessive and in a frequently self-fertilizing population (Figures 2 and S4), and its equilibrium frequency is lower as its dominance increases or the population's selfing rate decreases. The selection coefficient, $s$ has a nearly no effect on this equilibrium frequency (Fig S5).

**A genome-wide altruistic load.** Is this deterministic selective benefit of self-sacrifice strong enough to overcome the sampling process of random genetic drift? If so, when will the benefit of replacing oneself with a viable half sibling leave an observable pattern on the genetic load, such that it can be separated from the standard process of mutation selection balance? What should the frequency spectrum of such self-sacrificial alleles look like – should it be marked by an excess of relatively rare or relatively common variants? Answers to such questions – which we derive by genome-wide simulation – are essential for our ability to test the adaptive load hypothesis in natural populations.

Figure 3 compares the mean frequency of early-expressed deleterious recessive alleles ($h = 0$) per individual across three selfing rates ($\sigma = 0.1$, $\sigma = 0.2$, and $\sigma = 0.3$, x-axis), two mutation rates ($\mu = 5 \times 10^{-8}$, or $\mu = 5 \times 10^{-6}$, rows), three selection coefficients ($s = 0.2$, $s = 0.6$, and $s = 1.0$, columns), and two levels of inbreeding depression ($w_{\text{inbred}}$ initial $= 0.00$ Fig. 3A or ($w_{\text{inbred}}$ initial $= 0.25$, Fig. 3B). Although we present results only for mean allele frequency, we find qualitatively similar results for different summaries of the mutational load (e.g. the mean number of mutations per individual, Fig S6). We find that the mean mutation frequency only weakly depends on the selection coefficient (Fig. 3), a result also found in our single-locus analytical model (Fig S5). The efficacy of selection for an altrusitic load increases with population size – most notably the number of deleterious recessive variant per individual and the mean frequency of such variants is comparable between our model and the control when populations are small (e.g. $N = 100$ or $N = 500$) but these values become more distinct as the population size increases (Figures S7 and S8).

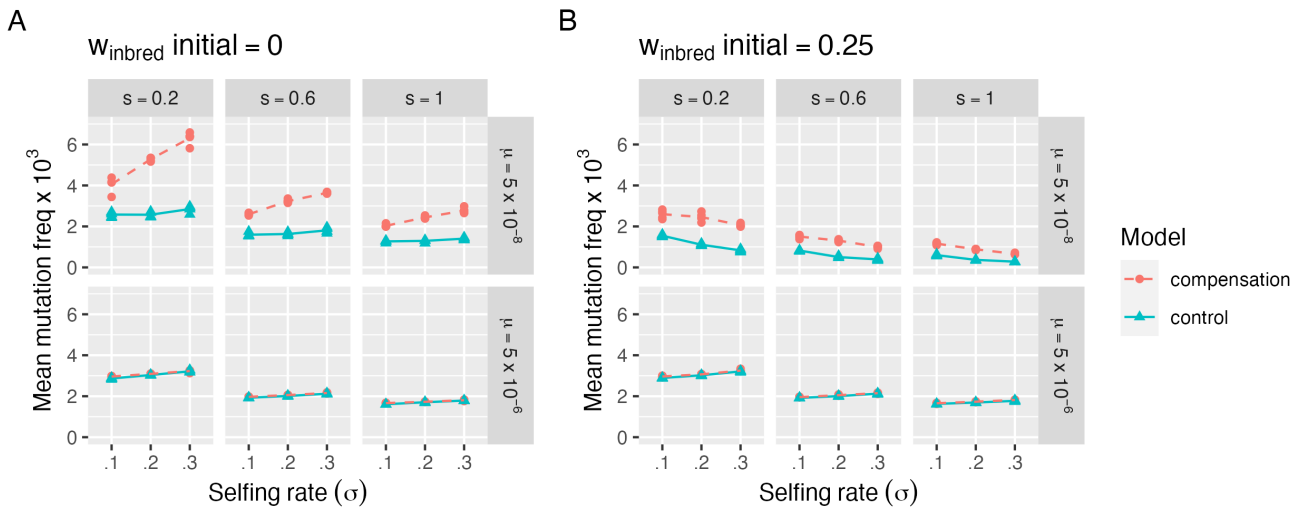


**Fig. 3.** The mean frequency of early-acting, recessive alleles, when inbred embryos have zero or low ($w = 0.25$) fitness before such alleles are introduced (panels A and B, respectively). We present results of a stochastic genome-wide SLiMulation with $N = 10,000$ over a combination of selfing rates ($\sigma$), mutation rates ($\mu$), and selection ($s$) coefficients. Red dotted lines display results for cases in which primary embryos are replaced by siblings (our model), while blue solid lines display cases in which primary embryos are replaced by a non-sibling (a control). Simulations were run in a population of size $10000$.

With lethal inbreeding depression and low mutation rates, there is a clear elevation in the frequency of early-expressed deleterious recessive mutations when reproductive compensation occurs in comparison to when it does not (Top panel of Fig. 3A), consistent with expectations from our analytical theory. This pattern intensifies as the selfing rate increases, presumably because a higher selfing rate increases the necessity of removing selfed embryos. These patterns disappear as the mutation rate increases (Bottom panel of Figure 3A), presumably because the mutational input outpaces what is required to ensure the death of selfed embryos. As such, selection acts to remove low fitness allozygotes (i.e. homozygotes formed by outcrossing). This explanation is consistent with the increased number of such mutations per individual with this higher mutation rate (Fig S6) and their lower frequency when segregating.

With sublethal inbreeding depression ($w_{\text{inbred}}$ initial $= 0.25$, 3B), we find a similar elevation in the frequency of these recessive alleles with reproductive compensation as compared to controls. However, in this case, the frequency decreases, rather than increases with the selfing rate. This difference in the relationship between the selfing rate and the frequency of a self-sacrificial recessive allele with lethal versus sublethal inbreeding likely reflects the fact that with sublethal inbreeding depression there is a cost to having two inviable selfed embryos. As such, this result may be idiosyncratic to our model of reproductive compensation.

## Discussion

The opportunity for reproductive compensation is likely pervasive across the tree of life. Egrets lay two eggs per season with only one destined to survive (33). Humans can make up for embryo mortality by increasing reproductive effort (34, but see (35)). Birds tend to lay more eggs during plagues of their rodent predators, and ultimately clutch size in many birds is determined by how many young can be fed, not how many can be laid (36). Similarly, angiosperms produce many more ovules than seeds (37). The evolutionary consequences of reproductive compensation have received much attention. For example, Hurst (38) recently argued that reproductive compensation provides selfish centromeres novel opportunities to functionally distort meiosis. In Hurst's model, rather than directly driving during meiois, selfish centromeres inducing nondisjunction to sabotage embryos destined no to inherit this centromere, and if such lost embryos can be compensated this sabotage gives the selfish centromere an additional opportunity to solely end up in the egg, rather than the polar body. Additionally many researchers have followed up on the suggestion R.A. Fisher made to Race (39) that reproductive compensation in humans allows for evolutionary stability of polymorphism at the underdominant Rh locus (40, 41). More generally studies in plants (27, 42, 43) and mammals (24, 44) have shown that reproductive compensation results in an increase in the early-expressed recessive load. Our study shows that reproductive compensation not only relaxes selection against deleterious recessives, but actually favors an early-expressed recessive load when inbreeding is unavoidable and inbreeding depression is severe.

Our finding that selection can favor an early expressed recessive load makes sense of Sakai's (45, 46) observation that reproductive compensation plus both early and late acting deleterious recessive mutations allowed for both mutational types to segregate at higher frequencies than they would on their own. We interpret this "two-stage effect" as an increase in the early expressed load favored by the mechanisms proposed above, followed by the relaxation of selection against late expressed recessive load as the loss of embryos formed by self-fertilization and biparental prevents recessive alleles

expressed later in life from being exposed when rare.

We focused on cases of unavoidable self-fertilization in plants. However the logic of our model is more broadly applicable. Although quantitative results will differ based on details of the systems of mating, we hypothesize that an adaptive early-expressed recessive load could evolve in plants and animals with the potential for reproductive compensation and in which dioecy or self-incompatibility cannot eliminate incidental biparental inbreeding and subsequent inbreeding depression (e.g. 47–50).

**Precedent for the adaptive evolution of self-sacrificial alleles.** Ours is not the first study to examine the possibility of the evolution of self-sacrificial alleles. For example, cytoplasmic male killers can adaptively spread in maternally transmitted organelles in taxa with reproductive compensation and/or sibling competition (51), and more broadly males are thought to be "cursed" by mitochondrial alleles that can harm male fitness at a benefit to female fitness (52, 53). Similarly, natural selection also favors maternal effect dominant embryonic arrest alleles (MEDEA) that epigenetically program the arrest of embryos that do not inherit this allele (54). These scenarios are somewhat similar to the adaptive load model presented above, in that these alleles induce self harm based on contextual information of their situation to increase their inclusive fitness. However, contrary to the case of the adaptive load modelled here, both cytoplasmic male killers and MEDEA alleles increase their inclusive fitness at the expense of unlinked loci – setting up a genetic conflict between these selfish alleles and the rest of the genome. In our model of the adaptive load, the self-sacrificial allele increases the inclusive fitness of alleles at unlinked loci (at least for the maternally-derived genome). As such, rather than representing a genomic conflict, ours is a model of cooperation, and is consistent with Hamilton's (55) idea that "infant mortality may evolve when the early death of one infant makes more likely the creation or survival of a close relative" (see also (24)). A closer analogue to our model is Charlesworth's (26) model for the adaptive evolution of recessive lethal alleles linked to a driving t-haplotype, which would otherwise lead to sterility when homozygous.

Our results are reminiscent of models for the evolution of organismal robustness and sensitivity to deleterious mutations under differing forms of selection (56–58). For example, Archetti (57) showed that soft selection within families (akin to reproductive compensation in our model) favors alleles hypersensitive to deleterious mutations, and Lan et al (58) showed that asexual lineages that resist Mueller's Ratchet by amplifying the fitness impacts of additional mutations by displaying synergistic epistasis can be favored by selection. However, in our model, alleles displaying sensitivity to homozygosity rather than mutation are favored.

**The efficacy of an altruistic load in preventing inbreeding.** An adaptive recessive load can effectively prevent self-fertilization. For example, with a selfing rate of ten percent, lethal inbreeding, and complete recessivity and lethality of self-sacrificial alleles, we found about five such alleles per individual. Assuming these loci are unlinked, we expect an only 0.3125% of selfed seed to give rise to embryos for this parameter combination. However, the location, size, and selection coefficient of the mutational target for loci potentially underlying such early-acting inbreeding depression are unknown, so we have no firm predictions.

**Implications for the evolution of Self Incompatibility.** Although we model the altruistic evolution of an early-expressed genetic load (rather than the spiteful load represented by cytoplasmic male killing and MEDEA), this evolutionary model is not without conflict. Here, the altruistic load clearly provides a greater patrilineal than matrilineal fitness cost, because an incidentally sacrificed embryo can be replaced by another from the same mother, while this pollen grain is replaced by one from a randomly chosen father. Although not directly modelled here, this conflict between maternally and paternally derived genes could favor the differential expression of the early embryonic load (i.e. genomic imprinting in embryos) depending on parent of origin (especially once the deleterious mutation is at high frequency), much like conflict over maternal resource provisioning favors parent-of-origin expression under the kinship theory of genomic imprinting (59). This view of the early adaptive load as a conflict between parental genomes over a self-sacrificial act provides a potential connection between self-incompatibility and the early expressed recessive load. This view is quite similar to Uyenoyama's models that show that the differing costs and benefits of self-incompatibility alleles to pollen and style can favor pollen with novel SI specificities, (60, 61), consistent with the idea that pollen benefit more from having potentially low-fitness offspring than do mothers. This tension between pollen (or pollen parent) and style (or seed parent) over fertilization and development is not unique to self-fertilization – for example upon secondary contact the benefit to pollen in overcoming stylar incompatibilities and siring potentially low-fitness hybrids can frustrate the evolution of reinforcement (62).

**Testable predictions.**

> Adaptation is a special and onerous concept that should not be used unnecessarily.
> — George C. Williams. 1966.
> *Adaptation and Natural Selection* (63).

Our model predicts that inadvertent selfing in predominantly outcrossing species with reproductive compensation can favor the adaptive evolution of early expressed recessive deleterious variation. However, our genome-scale simulations point to a more parsimonious explanation for an early expressed load. That is, when mutation rates are sufficiently high, the early expressed load that evolves by standard mutation-selection balance can equal or exceed the load favored by the adaptive process presented here, obviating the need for an adaptive load. Given that a recessive load can be

efficiently generated by the process of mutation rather than selection, we urge for careful tests of the ideas developed here before accepting the adaptive load hypothesis. We present some testable predictions below.

*We expect an elevated recessive embryonic load in taxa with effective reproductive compensation.* The benefit of early embryonic death is that this death can be compensated by redirected resources that would go to this embryo to a current or future sibling. As such, we predict the load of early expressed recessive deleterious variants to increase in taxa with more reproductive compensation. However, while this observation is consistent with our adaptive load model, it is also consistent with the more parsimonious explanation that reproductive compensation relaxes selection against an early-expressed load (25, 27), counteracting the purging process in inbred populations (44).

*Early-acting inbreeding depression should increase with the primary selfing rate (up to a point).* The greater the risk of inadvertent selfing, the greater the benefit of adaptive early-acting inbreeding depression. As such, we predict a larger early expressed recessive load in species with a higher selfing risk than those that effectively prevent the risk of self pollination. In such comparisons, the post-embryonic load – which should be unaffected by the process modelled above – could serve as a useful control. We note however, that we assumed that inbreeding depression was extreme and could not evolve. This occurs when selfing is rare and recessive mutations are common (31), has been observed in self-incompatible populations (e.g. 64), and is likely true of most primarily outcrossing populations. Because we do not model the co-evolution of mating system and the early genetic load, our predictions are only valid for taxa in which selfed seeds have very low fitness.

*Recessive loci underlying late-acting inbreeding depression should show signatures of balancing selection.* When the recessive load evolves adaptively (rather than under mutation selection balance) we see that recessive alleles are maintained under balancing selection and are on average older than those maintained under mutation selection balance. Thus, if the early expressed load is an adaptation, we predict a signature of balancing selection around loci underlying it – i.e. under a model of mutation-selection balance, recessive deleterious alleles should be relatively young, while under a model of an altruistic load, such alleles should be relatively old.

**Examination of the embryo lethality in gymnosperms.** Our model is of particular relevance to gymnosperms, as they lack a self-incompatibility mechanism, cannot avoid geitonogamous selfing and display high inbreeding depression. Pines have one to four archegonia per ovule – of which all can be fertilized, with only one surviving to be included in the seed (65). Our model makes sense of Williams' (66) conclusion that embryo death in pines occurs by two processes: (i) Numerous "Embryo Viability Loci" which are expressed throughout development and which potentially allow for embryo replacement, and (ii) The Embryo Lethal System – a coordinated process of embryo death which occurs after a dominant embryo has been established with limited opportunity for replacement of a dead embryo. We argue that polyembryony initially evolves as a mechanism to compensate the loss of embryos to inbreeding (as we 67, found previosuly), and that many of the "embryo viability loci" seen in nature represent the adaptive recessive load discovered in this work. Finally, we argue that the "embryo lethal system" is an evolved mechanism to suppress the non-dominant embryo if both embryos are viable – an interpretation consistent with both our theoretical observation that embryo competition frustrates the evolution of a recessive load (67), and the claim that death induced by the embryo lethal system is regulated by the position of the embryo rather than its competitive ability (66). This interpretation of a recessive load as an inbreeding avoidance strategy, is also consistent with the observation that gymnosperms have a lower realized selfing rate than expected given their primary selfing rates (68), and an intense study of a pine orchard demonstrating that the recessive load, rather than the opportunity for self-fertilization, controlled the realized selfing rate (69)).

**What proportion of the early-expressed load is altruistic?.** We have show that an altruistic,"self-sacrificial," recessive genetic load *can* evolve by natural selection. We again reiterate that this does not mean that much of the early-expressed load is adaptive. The proportion of the load which is self-sacrificial likely varies tremendously across taxa and depends on the extent of reproductive compensation (more compensation should yield a more altruistic load), the mutation rate to deleterious mutations that allow for compensation (the lower this rate, the more of the load can be attributable to selection rather than mutation), and the extent of inbreeding and inbreeding depression.

**Summary.** We show that when mothers can compensate for inviable embryos, a self-sacrificial early expressed recessive load can adaptively evolve by natural selection. Our model shows that the line between "early-acting inbreeding depression" and "late-acting self-incompatibility" might not be as conceptually clear-cut as previous authors have insisted. As such, our model potentially refutes the orthodox idea that "Abortion of inbred embryos is not an evolved trait for ensuring outcrossing, but is considered an early expression of genetic load, a lethal 'price' a species pays for maintaining genetic variability" (22) – and the price paid is the that of incomplete reproductive compensation. However, our model also shows that such a mechanism only meaningfully operates when the rate of origin of such mutations is low, as selection does not favor such mutations when they arise so frequently that a substantial recessive load would evolve in the absence of selection. We therefore argue that the key evolutionary question is not to distinguish between early-acting inbreeding depression and late acting self-incompatibility, but rather to distinguish between adaptive and non-adaptive mechanisms that prevent the production of selfed seed.

**ACKNOWLEDGEMENTS**

We would like to acknowledge Boris Igic, Mike Wade, Husain Agha, Aidan Harrington, Brooke Kern, and Manisha Munasinghe, and other members of the Brandvain lab for comments and the Research Council of Finland for supporting our project 'Genetic determinants of embryo destiny in a polyembryonic seed – mortality, inbreeding depression and sibling rivalry in a pine nutshell' (349057). Previous drafts are available on bioRxiv (70), and code underlying the results are available on zenodo (30).

# Bibliography

1. A S Kondrashov. Deleterious mutations as an evolutionary factor. ii. facultative apomixis and selfing. Genetics, 111(3):635–653, Nov 1985. ISSN 0016-6731 (Print); 0016-6731 (Linking). doi: 10.1093/genetics/111.3.635.
2. Deborah Charlesworth and John H Willis. The genetics of inbreeding depression. Nat Rev Genet, 10(11):783–796, Nov 2009. ISSN 1471-0064 (Electronic); 1471-0056 (Linking). doi: 10.1038/nrg2664.
3. R Lande. The influence of the mating system on the maintenance of genetic variability in polygenic characters. Genetics, 86(2 Pt. 1):485–498, Jun 1977. ISSN 0016-6731 (Print); 0016-6731 (Linking).
4. Russell Lande and Douglas W. Schemske. The evolution of self-fertilization and inbreeding depression in plants. i. genetic models. Evolution, 39(1):24–40, 1985.
5. D. Charlesworth and B. Charlesworth. Inbreeding depression and its evolutionary consequences. Annual Review of Ecology and Systematics, 18(1):237–268, 1987.
6. Russell Lande and Emmanuelle Porcher. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. Genetics, 200(3):891–906, Jul 2015. ISSN 1943-2631 (Electronic); 0016-6731 (Print); 0016-6731 (Linking). doi: 10.1534/genetics.115.176693.
7. Darwin CR. On the various contrivances by which british and foreign orchids are fertilised by insects, and on the good effects of intercrossing, 1862.
8. Anne Pusey and Marisa Wolf. Inbreeding avoidance in animals. Trends in Ecology & Evolution, 11(5):201–206, 1996. ISSN 0169-5347. doi: https://doi.org/10.1016/0169-5347(96)10028-8.
9. SCH Barrett. The evolution, maintenance and loss of self-incompatibility system. In J. Lovett-Doust and L. Lovett-Doust, editors, Plant reproductive ecology: patterns and strategies, pages 98–123. Oxford University Press, Oxford, 1988.
10. Steven R. Seavey and Kamaljit S. Bawa. Late-acting self-incompatibility in angiosperms. Botanical Review, 52(2):195–219, 1986. ISSN 00068101, 18749372.
11. R. A. Brink and D. C. Cooper. Somatoplastic sterility in medicago sativa. Science, 90(2345): 545–546, 1939. ISSN 00368075, 10959203.
12. D. C. Cooper and R. A. Brink. Partial self-incompatibility and the collapse of fertile ovules as factors affecting seed formation in alfalfa. Journal of Agricultural Research, 60:453–472, 1940.
13. Nelson S. Bittencourt Jr, Peter E. Gibbs, and João Semir. Histological Study of Post-pollination Events in Spathodea campanulata Beauv. (Bignoniaceae), a Species with Late-acting Self-incompatibility. Annals of Botany, 91(7):827–834, 06 2003. ISSN 0305-7364. doi: 10.1093/aob/mcg088.
14. Peter E. Gibbs. Late-acting self-incompatibility – the pariah breeding system in flowering plants. New Phytologist, 203(3):717–734, 2014. doi: https://doi.org/10.1111/nph.12874.
15. PETER E. Gibbs and MARTA B. Bianchi. Does late-acting self-incompatibility (lsi) show family clustering? two more species of bignoniaceae with lsi: Dolichandra cynanchoides and Tabebuia nodosa. Annals of Botany, 84(4):449–457, 1999. ISSN 0305-7364. doi: https://doi.org/10.1006/anbo.1999.0933.
16. D Charlesworth. Distribution of dioecy and self-incompatibility in angiosperms. In P.J. Greenwood, P.H. Harvey, and M. Slatkin, editors, Evolution: Essays in honour of John Maynard Smith, pages 237–268. Cambridge University Press, Cambridge, 1988.
17. Tammy L Sage, Robert I Bertin, and Elizabeth G Williams. Ovarian and other late-acting self-incompatibility systems. In Elizabeth G Williams, Adrienne E Clarke, and R Bruce Knox, editors, Genetic control of self-incompatibility and reproductive development in flowering plants, chapter 7, pages 116–140. Kluwer Academic Publishers, 1994.
18. Tammy L Sage, Mary V Price, and Nickolas M Waser. Self-sterility in ipomopsis aggregata (polemoniaceae) is due to prezygotic ovule degeneration. American Journal of Botany, 93 (2):254–262, 2006. doi: https://doi.org/10.3732/ajb.93.2.254.
19. S R Lipow and R Wyatt. Single gene control of postzygotic self-incompatibility in poke milkweed, asclepias exaltata l. Genetics, 154(2):893–907, Feb 2000. ISSN 0016-6731 (Print); 0016-6731 (Linking). doi: 10.1093/genetics/154.2.893.
20. Marta B Bianchi, Thomas R Meagher, and Peter E Gibbs. Do s genes or deleterious recessives control late-acting self-incompatibility in Handroanthus heptaphyllus (Bignoniaceae)? A diallel study with four full-sib progeny arrays. Annals of Botany, 127(6):723–736, 02 2021. ISSN 0305-7364. doi: 10.1093/aob/mcab031.
21. Nelson Sabino Bittencourt Júnior. Evidence for post-zygotic self-incompatibility in handroanthus impetiginosus (bignoniaceae). Plant Reproduction, 30(2):69–79, 2017. doi: 10.1007/s00497-017-0300-7.
22. Stephen L. Krebs and James F. Hancock. Embryonic genetic load in the highbush blueberry, Vaccinium corymbosum (ericaceae). American Journal of Botany, 78(10):1427–1437, 1991. doi: https://doi.org/10.1002/j.1537-2197.1991.tb12609.x.
23. Nic Lughadha, SJ Owen, and PJ Rudall. Preferential outcrossing in gomidesia (myrtaceae) is maintained by a post-zygotic mechanism. In S.J. Owens and P. J. Rudall, editors, Reproductive Biology in systematics, conservation and economic botany, pages 363–379. Kew Royal Botanic Gardens, London, 1998.
24. IAN M. HASTINGS. Models of human genetic disease: how biased are the standard formulae? Genetical Research, 75(1):107–114, 2000. doi: 10.1017/S0016672399003961.
25. I M Hastings. Reproductive compensation and human genetic disease. Genetics Research, 77(3):277–283, Jun 2001. doi: 10.1017/s0016672301004992.
26. Brian Charlesworth. The evolution of lethals in the *t* - haplotype system of the mouse. Proceedings of the Royal Society of London. Series B: Biological Sciences, 258(1352): 101–107, 1994. doi: 10.1098/rspb.1994.0149.
27. Emmanuelle Porcher and Russell Lande. Reproductive compensation in the evolution of plant mating systems. The New Phytologist, 166(2):673–684, 2005. ISSN 0028646X, 14698137.
28. A G Stephenson. Flower and fruit abortion: Proximate causes and ultimate functions. Annual Review of Ecology and Systematics, 12(1):253–279, 1981. doi: 10.1146/annurev.es.12.110181.001345.
29. D Wiens, C L Calvin, C A Wilson, C I Davern, D Frank, and S R Seavey. Reproductive success, spontaneous embryo abortion, and genetic load in flowering plants. Oecologia, 71 (4):501–509, Mar 1987. ISSN 1432-1939 (Electronic); 0029-8549 (Linking). doi: 10.1007/BF00379288.
30. Yaniv Brandvain, Lia Thomson, and Tanja Pyhäjärvi. Data and code from: Early-acting inbreeding depression can evolve as an inbreeding avoidance mechanism. Zenodo, 2024. doi: https://zenodo.org/records/4410128.
31. Russell Lande, Douglas W Schemske, and Stewart T Schultz. High inbreeding depression, selective interference among loci, and the threshold selfing rate for purging recessive lethal mutations. Evolution, 48(4):965–978, August 1994.
32. Benjamin C Haller and Philipp W Messer. SLiM 3: Forward Genetic Simulations Beyond the Wright–Fisher Model. Molecular Biology and Evolution, 36(3):632–637, 01 2019. ISSN 0737-4038. doi: 10.1093/molbev/msy228.
33. D W Mock. Siblicidal aggression and resource monopolization in birds. Science, 225(4663): 731–733, August 1984.
34. W J Schull and J V Neel. The effects of parental consanguinity and inbreeding in hirado, japan. v. summary and interpretation. Am J Hum Genet, 24(4):425–453, Jul 1972. ISSN 0002-9297 (Print); 1537-6605 (Electronic); 0002-9297 (Linking).
35. T E Reed. Dogma disputed. does reproductive compensation exist? an analysis of rh data. Am J Hum Genet, 23(2):215–224, Mar 1971. ISSN 0002-9297 (Print); 1537-6605 (Electronic); 0002-9297 (Linking).
36. David Lack. The significance of clutch-size. Ibis, 89(2):302–352, 1947. doi: https://doi.org/10.1111/j.1474-919X.1947.tb04155.x.
37. Delbert Wiens. Ovule survivorship, brood size, life history, breeding systems,and reproductive success in plants. Oecologia, 64(1):47–53, 1984. doi: 10.1007/BF00377542.
38. Laurence D. Hurst. Selfish centromeres and the wastefulness of human reproduction. PLOS Biology, 20(7):1–17, 07 2022. doi: 10.1371/journal.pbio.3001671.
39. R. R. Race. Some recent observations on the inheritance of blood groups. British Medical Bulletin, 2(8-9):165–165, 08 1944. ISSN 0007-1420. doi: 10.1093/oxfordjournals.Bmb.A071145.
40. B R Levin. The effect of reproductive compensation on the long term maintenance of the rh polymorphism: the rh crossroad revisited. Am J Hum Genet, 19(3 Pt 1):288–302, May 1967. ISSN 0002-9297 (Print); 1537-6605 (Electronic); 0002-9297 (Linking).
41. M W Feldman, M Nabholz, and W F Bodmer. Evolution of the rh polymorphism: a model for the interaction of incompatibility, reproductive compensation, and heterozygote advantage. Am J Hum Genet, 21(2):171–193, Mar 1969. ISSN 0002-9297 (Print); 1537-6605 (Electronic); 0002-9297 (Linking).
42. Philip W Hedrick, Outi Savolainen, and Katri Kärkkäinen. Factors influencing the extent of inbreeding depression: an example from scots pine. Heredity, 82(4):441–450, 1999. doi: 10.1038/sj.hdy.6885020.
43. Edward J Klekowski. Genetic load and soft selection in ferns. Heredity, 49(2):191–197, 1982. doi: 10.1038/hdy.1982.85.
44. A D J Overall, M. Ahmad, and R A Nichols. The effect of reproductive compensation on recessive disorders within consanguineous human populations. Heredity, 88(6):474–479, 2002.
45. Satoki Sakai. Maintenance of high inbreeding depression in selfing populations: Two-stage effect of early- and late-acting mutations. Journal of Theoretical Biology, 502:110307, 2020. ISSN 0022-5193. doi: https://doi.org/10.1016/j.jtbi.2020.110307.
46. Satoki Sakai. Why are deleterious mutations maintained in selfing populations? an analysis of the effects of early- and late-acting mutations by a two-locus two-allele model. Journal of Theoretical Biology, 533:110956, 2022. ISSN 0022-5193. doi: https://doi.org/10.1016/j.jtbi.2021.110956.
47. John S Heywood. Biparental inbreeding depression in the self-incompatible annual plant *Gaillardia pulchella* (asteraceae). Am J Bot, 80(5):545–550, May 1993.
48. Joseph I Hoffman, Fraser Simpson, Patrice David, Jolianne M Rijks, Thijs Kuiken, Michael A S Thorne, Robert C Lacy, and Kanchon K Dasmahapatra. High-throughput sequencing reveals inbreeding depression in a natural population. Proc Natl Acad Sci U S A, 111(10): 3775–3780, Mar 2014.
49. Jisca Huisman, Loeske E B Kruuk, Philip A Ellis, Tim Clutton-Brock, and Josephine M Pemberton. Inbreeding depression across the lifespan in a wild mammal population. Proc Natl Acad Sci U S A, 113(13):3585–3590, Mar 2016.
50. John D. Nason and Norman C. Ellstrand. Lifetime estimates of biparental inbreeding depression in the self-incompatible annual plant *Raphanus sativus*. Evolution, 49(2):307–316, 04 1995. ISSN 0014-3820. doi: 10.1111/j.1558-5646.1995.tb02243.x.
51. Laurence D. Hurst. The incidences and evolution of cytoplasmic male killers. Proceedings of the Royal Society of London. Series B: Biological Sciences, 244(1310):91–99, 1991. doi: 10.1098/rspb.1991.0056.
52. S. A. Frank and L. D. Hurst. Mitochondria and male disease. Nature, 383(6597):224–224, 1996. doi: 10.1038/383224a0.
53. Neil J. Gemmell, Victoria J. Metcalf, and Fred W. Allendorf. Mother's curse: the effect of mtdna on individual fitness and population viability. Trends in Ecology & Evolution, 19(5): 238–244, 2021/12/14 2004.
54. M J Wade and R W Beeman. The population dynamics of maternal-effect selfish genes. Genetics, 138(4):1309–1314, 12 1994. ISSN 1943-2631. doi: 10.1093/genetics/138.4.1309.
55. W.D. Hamilton. The moulding of senescence by natural selection. Journal of Theoretical Biology, 12(1):12–45, 1966. ISSN 0022-5193. doi: https://doi.org/10.1016/0022-5193(66)90184-6.

56. David C. Krakauer and Joshua B. Plotkin. Redundancy, antiredundancy, and the robustness of genomes. Proceedings of the National Academy of Sciences, 99(3):1405–1409, 2002. doi: 10.1073/pnas.032668599.
57. M. Archetti. Survival of the steepest: hypersensitivity to mutations as an adaptation to soft selection. Journal of Evolutionary Biology, 22(4):740–750, 2009. doi: https://doi.org/10.1111/j.1420-9101.2009.01697.x.
58. Yinghong Lan, Aaron Trout, Daniel Michael Weinreich, and Christopher Scott Wylie. Natural selection can favor ratchet robustness over mutational robustness. bioRxiv, 2018. doi: 10.1101/121087.
59. David Haig. The kinship theory of genomic imprinting. Annual Review of Ecology and Systematics, 31(1):9–32, 2000. doi: 10.1146/annurev.ecolsys.31.1.9.
60. M K Uyenoyama and E Newbigin. Evolutionary dynamics of dual-specificity self-incompatibility alleles. The Plant cell, 12(3):310–315, 03 2000. doi: 10.1105/tpc.12.3.310.
61. M K Uyenoyama, Y Zhang, and E Newbigin. On the origin of self-incompatibility haplotypes: transition through self-compatible intermediates. Genetics, 157(4):1805–1817, 04 2001. doi: 10.1093/genetics/157.4.1805.
62. Catherine A. Rushworth, Alison M. Wardlaw, Jeffrey Ross-Ibarra, and Yaniv Brandvain. Conflict over fertilization underlies the transient evolution of reinforcement. PLOS Biology, 20(10):1–21, 10 2022. doi: 10.1371/journal.pbio.3001814.
63. GC Williams. Adaptation and natural selection: A critique of some current evolutionary thought. Princeton University Press, Princeton, NJ, 1966.
64. Sarah J. Baldwin and Daniel J. Schoen. Inbreeding depression is difficult to purge in self-incompatible populations of leavenworthia alabamica. New Phytologist, 224(3):1330–1338, 2019. doi: https://doi.org/10.1111/nph.15963.
65. Risto Sarvas et al. Investigations on the flowering and seed crop of pinus silvestris. Metsatieteellisen tutkimuslaitoksen julkaisuja, 53(4), 1962.
66. Claire G Williams. Re-thinking the embryo lethal system within the pinaceae. Botany, 85 (7):667–677, 2007.
67. Yaniv Brandvain, Alexander Harkness, and Tanja Pyhäjärvi. Reproductive compensation and selection among viable embryos drive the evolution of polyembryony. The American Naturalist, 0(0):000–000, 0. doi: 10.1086/723454.
68. Brian C. Husband and Douglas W. Schemske. Evolution of the magnitude and timing of inbreeding depression in plants. Evolution, 50(1):54–70, 1996. doi: https://doi.org/10.1111/j.1558-5646.1996.tb04472.x.
69. Katri Kärkkäinen and Outi Savolainen. The degree of early inbreeding depression determines the selfing rate at the seed stage: model and results from *Pinus sylvestris (Scots pine)*. Heredity, 71(2):160–166, 1993.
70. Yaniv J Brandvain, Lia Thomson, and Tanja Pyhajarvi. Early acting inbreeding depression can evolve as an inbreeding avoidance mechanism. bioRxiv, 2023. doi: 10.1101/2023.03.20.533524.

# Appendix

This mathematical appendix accompanies the manuscript, Brandvain, Y., Thomson, L., and Pyhäjärvi, T. 2024. Early-acting inbreeding depression can evolve as an inbreeding avoidance mechanism.

**Mating Table, Single locus model:.** Here we show all possible offspring genotypes, their adult and embryo fitness, and their frequency in each maternally family after early embryo selection. Maternal families occur in frequencies of their genotypes after selection. The relative fitness of a selfed plant equals $1-\delta$, which we assume to equal zero throughout the analytical portion of the manuscript.

The values, $\psi_{aa}$, $\psi_{Aa}$, and $\psi_{AA}$ denote the probability that an embryo is replaced by another (i.e. compensated) in $aa$, $Aa$, and $AA$ families, respectively. The value, $\psi$ in maternal family, $k$ equals replace$_k = \sum\sum \text{freq}_{\text{Geno:i, of type: j, in Fam:k}} \times$ embryo fitness$_i$ summed across all primary embryos for each maternal family. These values are: $\psi_{aa} = phs(1-\sigma)$, $\psi_{Aa} = s(2h+2p(1-\sigma)+\sigma)/4$, and $\psi_{AA} = s(p+(1-p)(h(1-\sigma)+\sigma))$.

| maternal fam | genotype | type | adult fitness | primary | embryo fitness | freq in fam |
|---|---|---|---|---|---|---|
| aa | aa | Self | $1-\delta$ | Primary | 1 | $\sigma$ |
| aa | Aa | Self | $1-\delta$ | Primary | $1-hs$ | 0 |
| aa | AA | Self | $1-\delta$ | Primary | $1-s$ | 0 |
| aa | aa | Out | 1 | Primary | 1 | $(1-p)(1-\sigma)$ |
| aa | Aa | Out | 1 | Primary | $1-hs$ | $p(1-\sigma)$ |
| aa | AA | Out | 1 | Primary | $1-s$ | 0 |
| aa | aa | Self | $1-\delta$ | Backup | 1 | $\psi_{aa}\sigma$ |
| aa | Aa | Self | $1-\delta$ | Backup | $1-hs$ | 0 |
| aa | AA | Self | $1-\delta$ | Backup | $1-s$ | 0 |
| aa | aa | Out | 1 | Backup | 1 | $\psi_{aa}(1-\sigma)(1-p)$ |
| aa | Aa | Out | 1 | Backup | $1-hs$ | $\psi_{aa}(1-\sigma)p$ |
| aa | AA | Out | 1 | Backup | $1-s$ | 0 |
| Aa | aa | Self | $1-\delta$ | Primary | 1 | $\sigma/4$ |
| Aa | Aa | Self | $1-\delta$ | Primary | $1-hs$ | $\sigma/2$ |
| Aa | AA | Self | $1-\delta$ | Primary | $1-s$ | $\sigma/4$ |
| Aa | aa | Out | 1 | Primary | 1 | $(1-\sigma)(1-p)/2$ |
| Aa | Aa | Out | 1 | Primary | $1-hs$ | $(1-\sigma)/2$ |
| Aa | AA | Out | 1 | Primary | $1-s$ | $(1-\sigma)(p/2)$ |
| Aa | aa | Self | $1-\delta$ | Backup | 1 | $\psi_{Aa}(\sigma/4)$ |
| Aa | Aa | Self | $1-\delta$ | Backup | $1-hs$ | $\psi_{Aa}(\sigma/2)$ |
| Aa | AA | Self | $1-\delta$ | Backup | $1-s$ | $\psi_{Aa}(\sigma/4)$ |
| Aa | aa | Out | 1 | Backup | 1 | $\psi_{Aa}(1-\sigma)(1-p)/2$ |
| Aa | Aa | Out | 1 | Backup | $1-hs$ | $\psi_{Aa}(1-\sigma)/2$ |
| Aa | AA | Out | 1 | Backup | $1-s$ | $\psi_{Aa}(1-\sigma)p/2$ |
| AA | aa | Self | $1-\delta$ | Primary | 1 | 0 |
| AA | Aa | Self | $1-\delta$ | Primary | $1-hs$ | 0 |
| AA | AA | Self | $1-\delta$ | Primary | $1-s$ | $\sigma$ |
| AA | aa | Out | 1 | Primary | 1 | 0 |
| AA | Aa | Out | 1 | Primary | $1-hs$ | $(1-p)(1-\sigma)$ |
| AA | AA | Out | 1 | Primary | $1-s$ | $p(1-\sigma)$ |
| AA | aa | Self | $1-\delta$ | Backup | 1 | 0 |
| AA | Aa | Self | $1-\delta$ | Backup | $1-hs$ | 0 |
| AA | AA | Self | $1-\delta$ | Backup | $1-s$ | $\psi_{AA}\sigma$ |
| AA | aa | Out | 1 | Backup | 1 | 0 |
| AA | Aa | Out | 1 | Backup | $1-hs$ | $\psi_{AA}(1-\sigma)(1-p)$ |
| AA | AA | Out | 1 | Backup | $1-s$ | $\psi_{AA}(1-\sigma)p$ |

**Selection, Single locus model.** The frequency of each genotype after selection is the product of its frequency in a maternal family, the maternal family's frequency, its embryonic and adult fitness, weighted by mean fitness. By appropriately summing and weighting the content of the table above, and assuming complete inbreeding depression ($\delta = 1$), we find mean fitness equals

$$\overline{w} = (1-\sigma)\left(p_{\text{aa}}(\psi_{\text{aa}}+1)(1-hps)+p_{\text{Aa}}(\psi_{\text{Aa}}+1)(1-\frac{s(h+p)}{2})+p_{\text{AA}}(\psi_{\text{AA}}+1)(1-hqs-ps)\right)$$

Where $\sigma$ is the selfing rate, $p$ and $q$ are frequencies of alleles $a$ (wild type), and $A$ (self-sacrificial), respectively, and $\psi_k$ is the probability that a primary embryo in the family of the subscripted maternal genotype, $_k$ is replaced.

From the table above (see algebraic details in our Supplemental Mathematica file), we find that the frequencies of each genotype in the next generation are as follows:

$$p'_{aa} = \frac{(1-p)(1-\sigma)(2p_{aa}(\psi_{aa}+1)+p_{pAa}(\psi_{Aa}+1))}{2\overline{w}}$$

$$p'_{Aa} = \frac{((1-\sigma)(1-hs))(2pp_{\text{aa}}(\psi_{aa}+1)+2(1-p)pp_{AA}(\psi_{AA}+1)+p_{Aa}(\psi_{Aa}+1))}{2\overline{w}}$$

$$p'_{AA} = \frac{p(1-s)(1-\sigma)\left(p+\frac{pp_{Aa}\psi_{Aa}}{2}+p_{AA}\psi_{AA}\right)}{\overline{w}}$$

## Supplementary figures

These supplementary figures accompany the manuscript, Brandvain, Y., Thomson, L., and Pyhäjärvi, T. 2024. Early-acting inbreeding depression can evolve as an inbreeding avoidance mechanism.

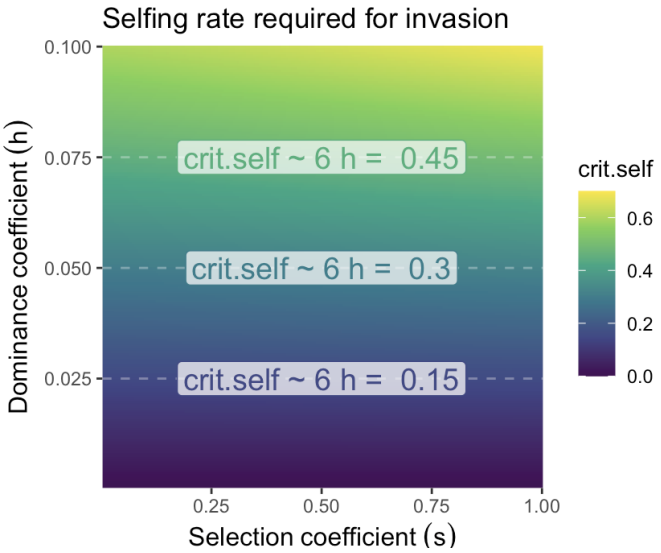


**Fig S1:** The critical selfing threshold required for the invasion of a self-sacrificial, $A$, allele as a function of its selection and dominance coefficients, under our model and assuming that all selfed seed dies. The colored text reflects the approximation that a rare $A$ allele is favored when crit.self $> 6h$. This approximation is exact as $s$ approaches zero, but crit.self must exceed $6h$ as $s$ increases (e.g. with $s = 1$ crit.self equals 0.155, 0.321, and 0.499 , when $h$ is 0.025, 0.05, and 0.075, respectively.)

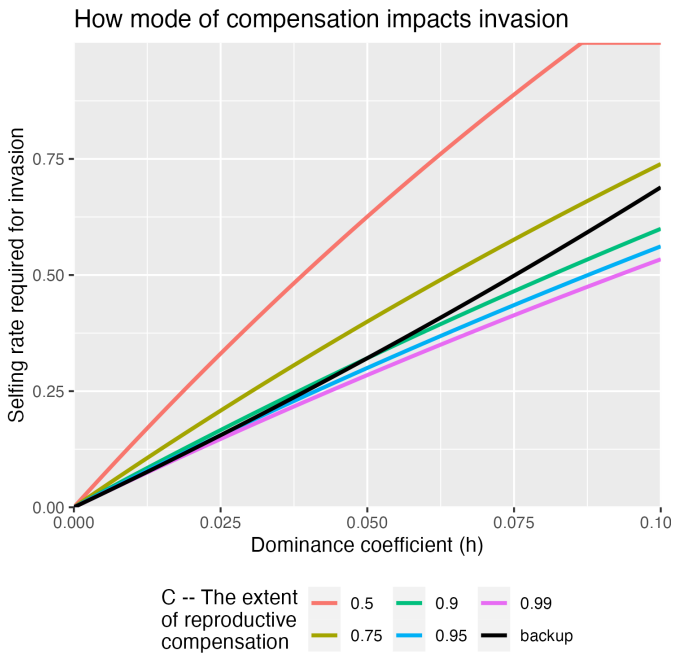


**Fig S2:** How the extent and mechanism of reproductive compensation impact the invasion criteria of a self-sacrificial (partially) recessive allele across dominance coefficients. The level of compensation, $C$, in Charlesworth's parameterization of reproductive compensation is shown in color, while results from our parameterization of reproductive compensation are provided in black as a reference. Results – reported as the critical selfing rate required for invasion (y-axis) are shown for a variety of dominance coefficients (x-axis). All lines assume a selection coefficient of $s = 0.5$.

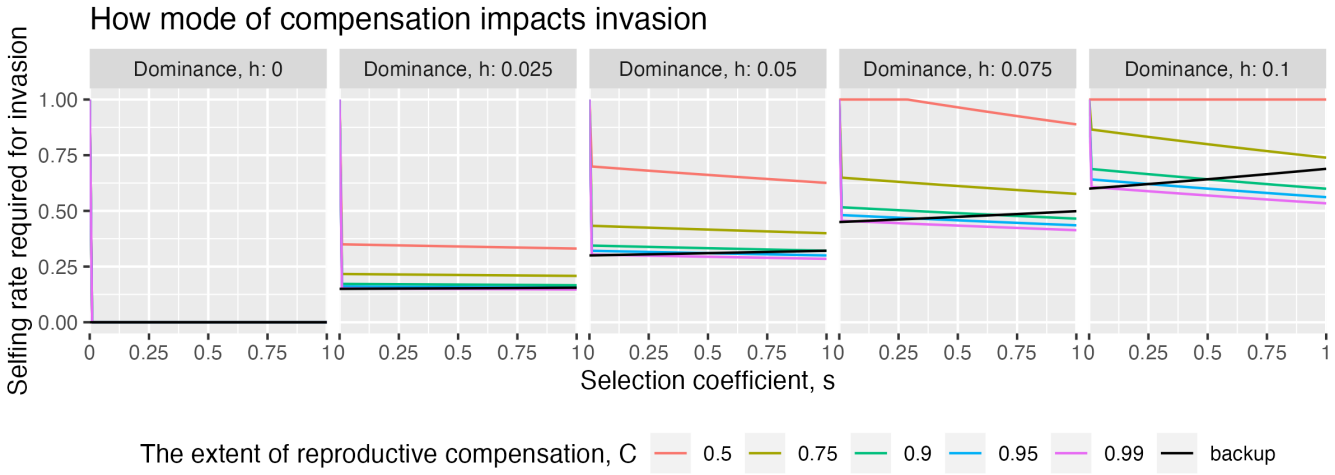


**Fig S3:** How the extent and mechanism of reproductive compensation impact the invasion criteria of a self-sacrificial (partially) recessive allele across selection and dominance coefficients. The level of compensation, $C$, in Charlesworth's parameterization of reproductive compensation is shown in color is show in color, while results from our parameterization of reproductive compensation are provided in black as a reference. Results – reported as the critical selfing rate required for invasion (y-axis) are shown for a variety of selection coefficients (x-axis) and dominance coefficients (facets).

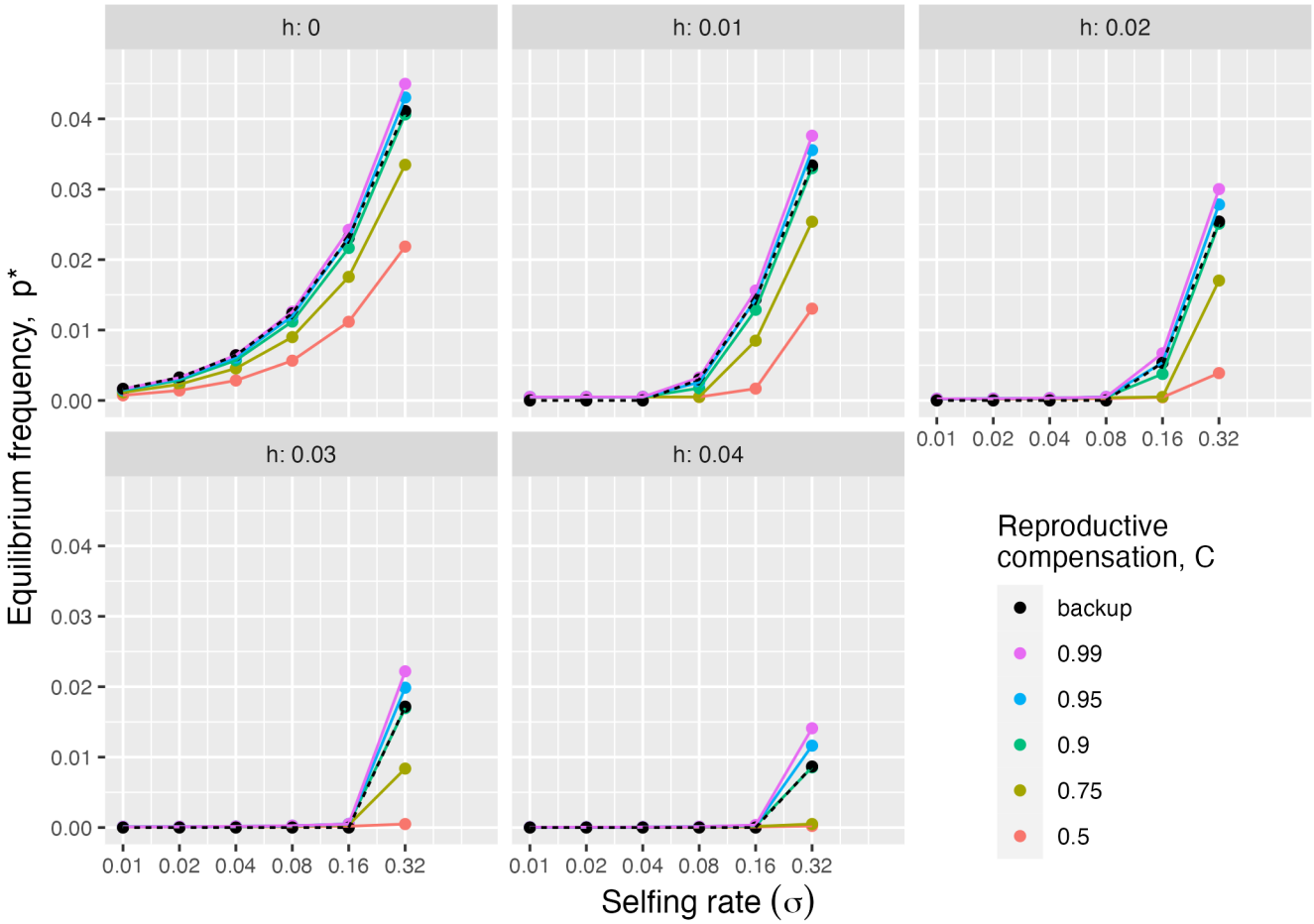


**Fig S4:** How the extent and mechanism of reproductive compensation impact the equilibrium frequency, $p^*$, of a self-sacrificial (partially) recessive lethal allele. The level of compensation, $C$, in Charlesworth's parameterization of reproductive compensation is shown in color is show in color, while results from our parameterization of reproductive compensation are provided in black as a reference. Results are shown for a variety of selfing rates (x-axis) and dominance coefficients (facets), and assume lethal inbreeding.

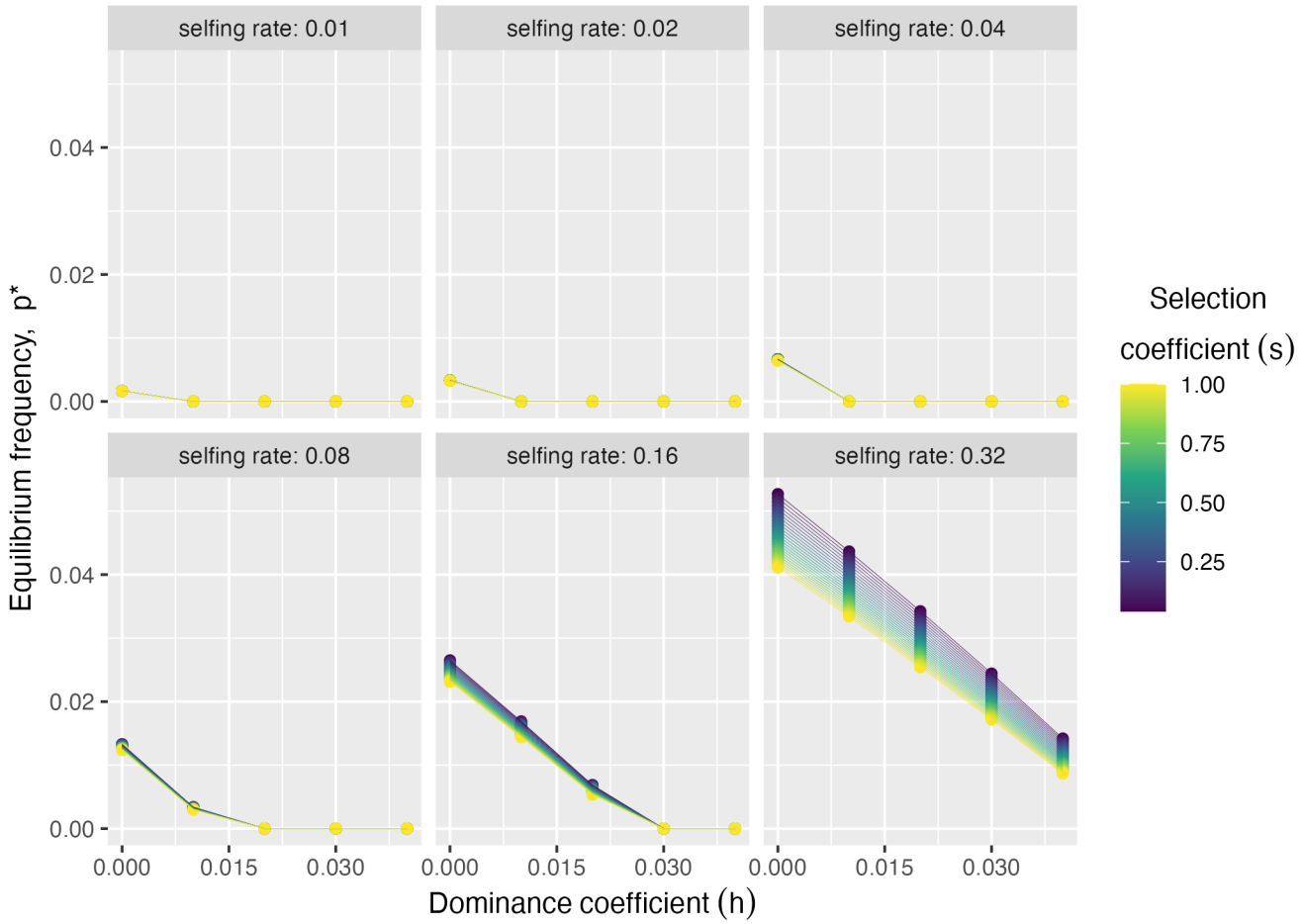


**Fig S5** The equilibrium frequency, $p^*$, of a self-sacrificial (partially) recessive allele is largely insensitive to the strength of selection (in color), and depends more on the selfing rate (facet) and dominance coefficient (x-axis).

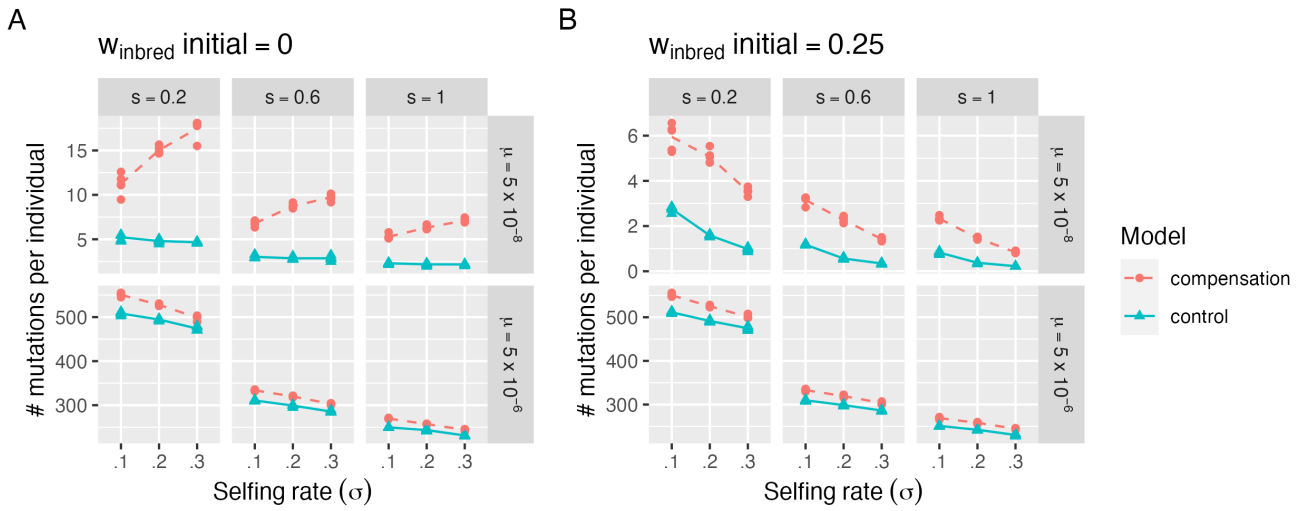


**Fig S6:** The mean number of early-acting, recessive lethal ($w_{\text{inbred}}$ initial $= 0$, A), and sublethal ($w_{\text{inbred}}$ initial $= 0.25$, B) alleles per individual, over a combination of selfing rates ($\sigma$), mutation rates ($\mu$), selection ($s$ )and dominance (h) coefficients. Blue lines display results for cases in which primary embryos are replaced by siblings (our model), while blue lines display cases in which primary embryos are replaced by a non-sibling (a control). Simulations were run in a population of size $10000$.

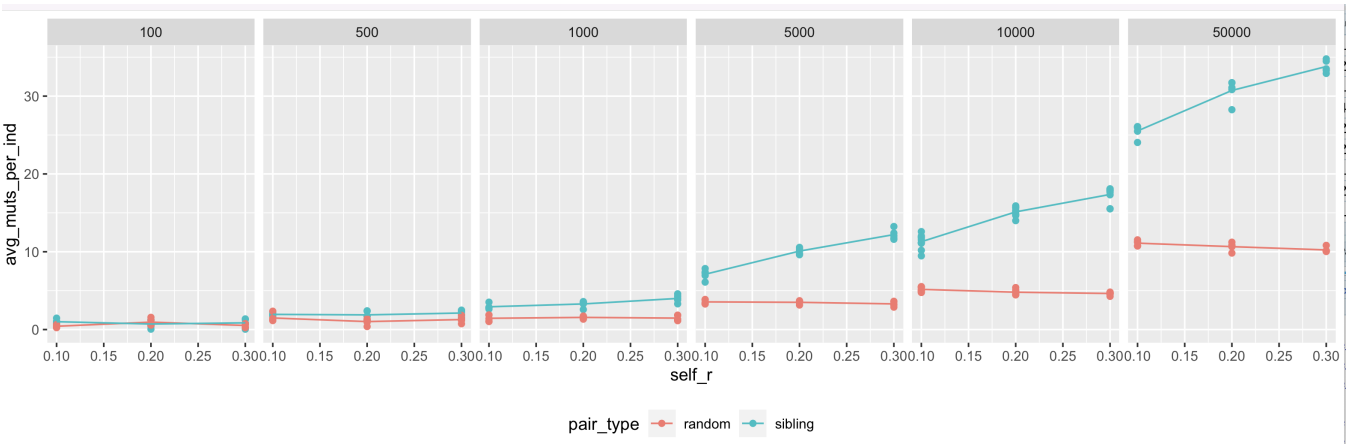


**Fig S7:** The mean frequency of early-acting, recessive alleles over a combination of selfing rates ($\sigma$, x-axis), and population sizes (facets). Blue lines display results for cases in which primary embryos are replaced by siblings (our model), while blue lines display cases in which primary embryos are replaced by a non-sibling (a control). Simulations assume lethal inbreeding and a selection coefficient of -0.2 on the recessive allele.

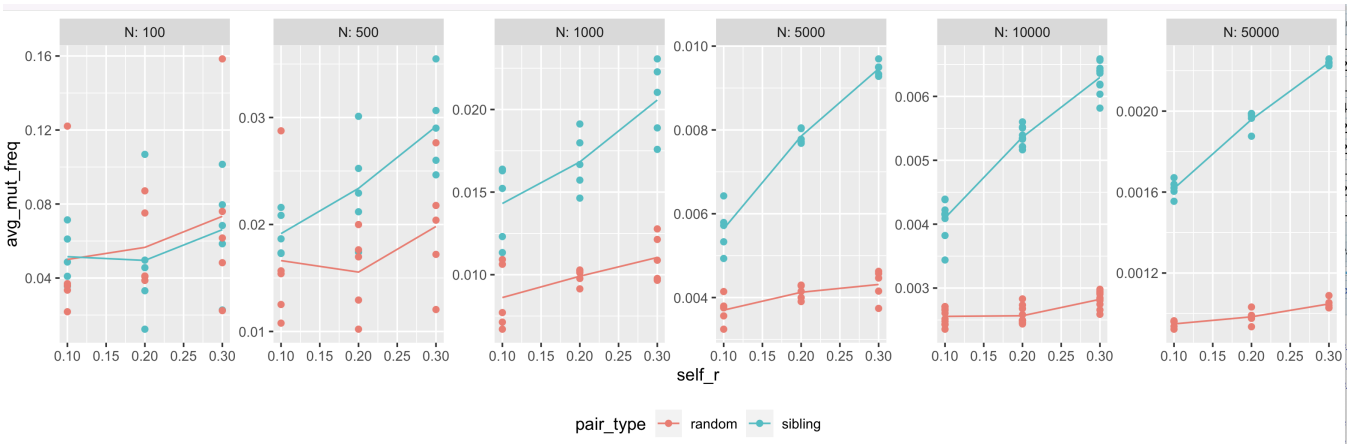


**Fig S8:** The mean number recessive alleles per individual over a combination of selfing rates ($\sigma$, x-axis), and population sizes (facets). Blue lines display results for cases in which primary embryos are replaced by siblings (our model), while blue lines display cases in which primary embryos are replaced by a non-sibling (a control). Simulations assume lethal inbreeding and a selection coefficient of -0.2 on the recessive allele.

# Supplementary text

This text accompanies the manuscript, Brandvain, Y., Thomson, L., and Pyhäjärvi, T. 2024. Early-acting inbreeding depression can evolve as an inbreeding avoidance mechanism.

**An alternative parameterization of reproductive compensation.** To evaluate the robustness of our results to our model of reproductive compensation, we developed a second model with an alternative form of reproductive compensation. Specifically we implemented reproductive compensation following Charlesworth's (26), model, in which maternal fitness equals $\frac{1-p_m}{(1-p_m)^C}$, where $p_m$ is the probability that an embryo will not survive and $C$ is the extent of reproductive compensation (Fig Supp Text 1). Each mother produces $\sigma$ of its children by selfing and $1-\sigma$ of its children by outcrossing (receiving pollen at random). Each mother produces (up to) six type of embryos – selfed or outcrossed embryos of all three genotypes – in frequencies given in Table 1:

| | Selfed | | | Outcrossed | | |
|---|---|---|---|---|---|---|
| Mom | AA | Aa | aa | AA | Aa | aa |
| AA | $\sigma$ | 0 | 0 | $p(1-\sigma)$ | $(1-p)(1-\sigma)$ | 0 |
| Aa | $\sigma/4$ | $\sigma/2$ | $\sigma/4$ | $p\frac{(1-\sigma)}{2}$ | $\frac{1-\sigma}{2}$ | $(1-p)\frac{(1-\sigma)}{2}$ |
| aa | 0 | 0 | $\sigma$ | 0 | $p(1-\sigma)$ | $(1-p)(1-\sigma)$ |

**Table 1.** Frequencies of types of embryos in each maternal family before selection within families.

We then select within families, which in our model depends only on genotype at the $\mathcal{A}$ locus. Frequencies within families after selection equal frequencies before selection multiplied by the fitness of each genotype (during early development - i.e. depending on genotype, not selfed status) divided by mean family fitness. Recall that genotypic fitnesses during within family selection are $(1-s)$, $(1-hs)$, and 1 for AA, Aa, and aa genotypes, respectively. The denominators (i.e. mean family fitnesses before accounting for compensation) are therefore

| Mom | mean fitness |
|---|---|
| AA | $1-s(\sigma+(1-\sigma)(p+h(1-p)))$ |
| Aa | $1-\frac{s}{2}(h+p(1-\sigma+\sigma/2))$ |
| aa | $1-hps(1-\sigma)$ |

**Table 2.** Mean fitness of each family before reproductive compensation.

After finding the relative frequencies of each combination of selfing and genotype in each family after selection, we introduce reproductive compensation following the model of Charlesworth (26). Here, the fitness of each maternal family $j$ after compensation, $w_{\text{Fam:j,AC}}$, becomes its value before compensation, $w_{\text{Fam:j,AC}}$ divided by this value raised to the $C$ (Eq. 1).

$$w_{\text{Fam:j, after compensation}} = \frac{w_{\text{Fam:j, before compensation}}}{(w_{\text{Fam:j, before compensation}})^C} \tag{1}$$

where $C$ is a measure of the degree of compensation (if $C=1$, early mortality has no effect on final offspring number, while if $C=0$ there is no reproductive compensation).

We then find the frequency of each of the six offspring types after selection within and among families (i.e. in dispersed seed in our model) as the product of its frequency within a family after within family selection multiplied by the contribution of this family to the seed pool (i.e. after among family after selection, where this frequency is simply the family frequency before selection multiplied by its fitness after compensation divided by the mean fitness of all families after compensation), summed across all families. So, the frequency of offspring type, $i$, after selection within and among families equals

$$p_j^{'} = \sum_{\times} p_{i,j} \times \frac{w_{i,\text{early}}}{w_{j\text{before compensation}}} \times p_j \times \frac{w_{j,\text{ after compensation}}}{\overline{w_{\text{after compensation}}}} \tag{2}$$

where $p_{i,j}$ is the frequency of offspring type, $i$ in family $j$ before selection (i.e. values in Table 2), $w_{i,\text{early}}$ is the fitness of this offspring (during within family selection), based on its genotype, $w_{j\text{before compensation}}$ corresponds to family fitness values before compensation presented in Table 2, while $w_{j,\text{ after compensation}}$ refers to values derived from Eq. 1, and $\overline{w_{\text{after compensation}}}$ is the mean of Eq. 2 (i.e. this value for each family multiplied by the family frequency before selection, summed across all families).

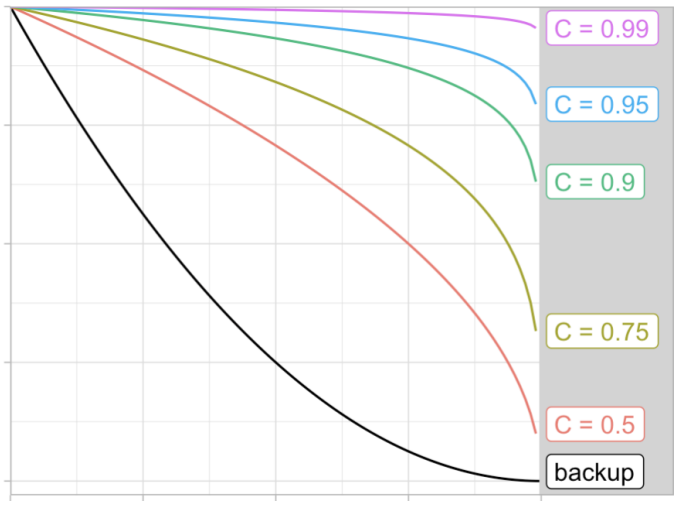


**Fig Supp Text 1:** The extent of reproductive compensation as a function of $C$ (color) under Charlesworth's parameterization of reproductive compensation, in which maternal fitness equals $(1-p_m)/((1-p_m)^C)$.

Finally, surviving seeds are subjected to later selection, which we base exclusively on their inbreeding status. So the offspring types after this stage of selection, $p_i^{''}$ equal $\frac{p_i^{'} \times w_{i,\,\text{late}}}{\overline{w}_{\text{late}}}$. Because this math is quite cumbersome, we find results for invasion and equilibrium allele frequencies by numerical iteration in R.

**Invasion analysis for Charlesworth's parameterization of reproductive compensation:.** Because this parameterization model was not analytically tractable, we used a bisection optimization approach to identify invasion criteria. For each combination of parameters, we found critical selfing rate required for invasion as follows:

1. We kept track of three values – our focal selfing rate (initially 0.5), our lower selfing rate (initially 0), and our upper selfing rate (initially 1).
2. We evaluated invasion by introducing (at frequency .000005) a deleterious early expressed allele, and asked if it deterministically increased in frequency from generation one to ten given the focal selfing rate. And then
   - If the allele invaded, we updated our values as follows:
     - *The focal selfing rate* takes a value halfway between the previous focal selfing rate and the lower selfing rate.
     - *The upper selfing rate* takes the value of the previous focal selfing rate.
     - *The lower selfing rate* does not change.
   - If the allele failed to invade, we updated our values as follows:
     - *The focal selfing rate* takes a value halfway between the previous focal selfing rate and the upper selfing rate.
     - *The upper selfing rate* does not change.
     - *The lower selfing rate* takes the value of the previous focal selfing rate.
3. After a test at one focal selfing rate we then evaluate invasion at the new focal selfing rate. We continue this process until the difference between the upper and lower selfing rates is less than $10^{0-4}$ – a precision usually achieved in about fifteen iterations. We consider the focal selfing rate at this point to be the lowest selfing rate which allows for invasion (i.e. the "critical selfing rate").