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]Department of Biology. Texas Tech University - Twin Cities. Lubbock, TX.
[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 (by self-fertilization or biparental inbreeding) may severely compromise individual fitness by exposing the recessive genetic load, which is otherwise hidden in the heterozygous state. Unquestionably, much of this load is an inevitable consequence of the mutational process and the masking of deleterious recessive alleles in predominantly outcrossing populations. However, we show that when reproductive compensation allows for the replacements of inviable embryos lost early in development, natural selection can favor highly deleterious recessive variants that functionally induce self-sacrificial death of inbred embryos. Our work provides numerous testable hypotheses, and suggests that the distinction between late self-incompatibility and early acting inbreeding depression may be somewhat artificial in some taxa.

Inbreeding depression | Self Incompatibility | Altruism | Reproductive compensation

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

Introduction

A recessive genetic load, caused by both rare, recessive, and highly deleterious mutations (1, 2) and by deviations from phenotypic optima under multi-locus stabilizing selection (3), is an inevitable consequence of habitual outcrossing. As such, numerous mechanisms have adaptively evolved to avoid inbreeding across plants, animals and fungi (4, 5). For example, many plant species have a chemically-mediated self-incompatibility (SI) system which rejects self-pollen. The mechanisms and timing of SI are diverse (6) 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 at the S-locus well before fertilization.

Late Acting SI – in which self-pollen can effectively germinate, grow down the stigma, and in some cases can enter and even fertilize the ovule before ovular development arrests (7–10) – is less well understood (11). In the most delayed form of Late Acting SI, "Post-zygotic" SI (which is particularly common in the Bignoniaceae (12)), self-incompatibility is realized after fertilization. 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 (11), and because it is difficult to distinguish late-acting SI from early-acting inbreeding depression (6, 13). 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 early-acting inbreeding depression is "*a process acting in the progeny zygote, determined by its own genotype*" (13).

Substantial effort has gone into distinguishing between early acting inbreeding depression and late-acting SI (reviewed in 11, 14), with evidence supporting late-acting SI in some taxa (e.g. 15–18) and early inbreeding depression (e.g. 19, 20) in others. We argue that with high reproductive compensation the distinction above is not necessarily evolutionarily relevant, as early acting inbreeding can adaptively evolve as a mechanism to minimize reproductive investment in low fitness selfed seed. Put another way, an adaptive early-expressed recessive load that allows mothers to trade in inbred for outbred seeds plays the role as traditional self-incompatibility. The key to this argument is that homozygosity for a rare recessive allele serves as a reliable indicator of inbreeding – much like similarity of pollen and style S-locus genotype serve as a reliable indicator of self vs on-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. Such altruistic early death has been previously considered in relation to the reinforcement of postyzgotic isolating mechanisms (21, 22), the evolution of recessive lethality of a selfish transmission distorter that makes homozygous males sterile (23), and earlier onset of of genetic diseases that would guarantee low fitness (24).

After introducing the plausibility of this argument, we show that, depending on the mutation rate, ongoing evolution of this early-acting "altrusitic" adaptive load will lead to patterns of variation that are difficult to distinguish from standard models of mutation selection balance. As such, while self-sacrifice may favor an elevated genetic load, Occam's razor argues for a null interpretation of early embryonic death as a simple consequence of standard mutation selection balance, rather than a panglossian (25) interpretation of embryonic death as an altruistic act. We therefore conclude with a series of specific predictions which allow for distinguishing between an adaptive recessive load and the load expected under standard mutation-selection balance.

Because our model makes very few critical assumptions – namely that homozygosity for a rare allele signals recent inbreeding, that inbreeding is bad, and that mothers can compensate for early mortality – our work suggests that an altruistic recessive load could adaptively evolve across many organisms.

Methods

Throughout, we assume that there is an opportunity for reproductive compensation ("that is a dead embryo can be replaced by a potentially viable one" (26)). For simplicity, we model reproductive compensation by providing each seed with a single backup - such that an inviable 'primary' embryo can be replaced by another, potentially viable, embryo. This roughly follows the case of simple polyembryony in pines which we modeled previously (27), with the distinction that in this model maternal half-sibs do not necessarily inherit the same maternal allele. This model differs slightly from that of Porcher and Lande's (26), which represents the extent of reproductive compensation the proportion of seeds that are matured to ovules. However, we do not believe that this modelling decision would change our qualitative conclusions.

For tractability, we assume that inbreeding depression is expressed after seed development (such that initially all embryos live to become full seeds), and that it is not purged at the population's mutation and selfing rates (i.e. the selfing rate is too low relative to the mutation rate to allow natural selection to decrease the recessive load). Such an "unpurgeable load", which evolves when inbreeding exposes too many recessive deleterious mutations at once to allow for their purging, occurs over a broad range of parameter space (28). Our analytical model further assumes lethal inbreeding depression, but relaxing this assumption in our simulations shows that an altruistic load can adaptively evolve even with sublethal inbreeding depression.

Deterministic Models Without Mutation. We initially assume that all primary embryo survive to become a seed, and that following dispersal, seeds die if they are inbred and survive otherwise. We then introduce a rare highly recessive (with dominance coefficient, h) allele, A, which decreases embryo survival by selection coefficient, s. Primary embryos killed by exposing the effects of the A allele can be replaced by a backup embryo, which cannot be replaced if it dies. We assume that the resource transfer from primary to backup embryos is perfect – so there is no maternal cost to losing their primary embryo. 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 randomly mating population of infinite size, but make no assumptions about Hardy-Weinberg Equilibrium (HWE).

Stochastic Model With Mutation.

Simulation framework: We developed genome-scale, individual-based stochastic simulations to examine the genomic signatures of this process. Specifically, we ran non-Wright-Fisher forward simulations in SLiM 3.0 (29) in which early-acting recessive deleterious mutations recurrently arise (at rate μ) 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 0.5×10^{-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. Because early expressed recessive variants represent only a small portion of the total available mutations, the number of such mutations per a generation is likely much less than the total number of new mutations every generation.

Evolving an altruistic load: Our SLiM model matches the essence of our analytical framework for the adaptive evolution of an altruistic load. Each generation, 10,000 survivors were randomly sampled with replacement to act as mothers. Each mother produced two maternal siblings, each of which could be the product of selfing or outcrossing based on a set selfing rate, α. Next, early-acting deleterious mutations were expressed. If both siblings survived, one was randomly considered "primary", and it experienced viability selection in which survival depended on whether or not they were the product of selfing, while the other embryo did not develop further. The fitness of inbred offspring was a constant that we varied across simulations.

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 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 selected and died, 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, we assume the A allele is so rare that AA homozygotes can be ignored, so we consider sibling pairs from aa or aA mothers. Each offspring is the product of selfing or outcrossing with probabilities α and $1-\alpha$, respectively. Outcrossed ovules are pollinated by A or a bearing pollen, with probabilities p_A and $1-p_A$, respectively.

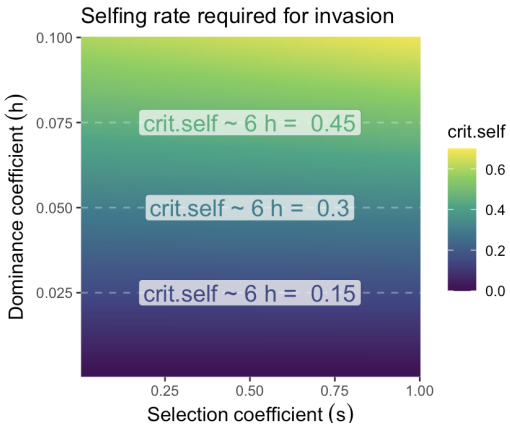

Fig. 1. The critical selfing threshold required for the invasion of a self-sacrificial, A, allele as a function of its selection and dominance coefficients. This model assumes that all selfed seed dies. The colored text reflects the approximation that a rare A allele is favored when crit.self $> 6h$. 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.)

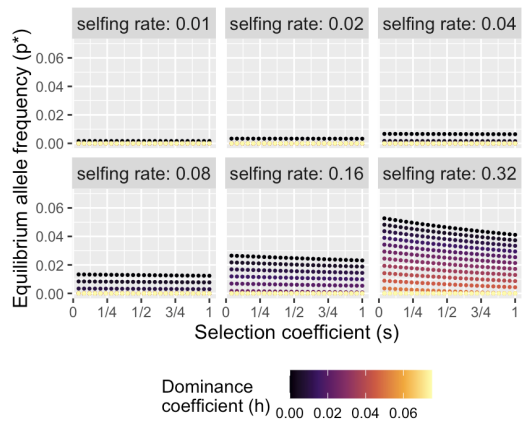

Fig. 2. The equilibrium frequency, p^*, of a self-sacrificial (partially) recessive allele for which lost embryos can be compensated, under a variety of selfing rates and dominance coefficients.

At invasion (i.e. $\lim_{p_A \to 0}$) aa mothers produce aa offspring by selfing with probability α and produce aa and Aa offspring by outcrossing with probabilities $(1-p_A)(1-\alpha)$, and $(p_A)(1-\alpha)$, respectively. The probability that the primary embryo in an aa maternal family is replaced by a backup embryo, ψ_{aa}, equals the product of probability that an embryo is fertilized by A pollen and dies in early development i.e. $h\,s\,p_A\,(1-\alpha)$.

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

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

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

Assuming full recessivity (i.e. $h = 0$), $\overline{w}_{\text{Aa:fam},\lim_{p_A\to 0}}$ is $\approx \frac{1}{8}(1-\alpha)(8+s\,\alpha)$, which always exceeds mean population fitness at invasion, and therefore a recessive A is never disfavored when rare. The expected increase in frequency of a rare recessive allele that acts in this manner is $p_A(1-\alpha)(s\,\alpha)/(4\overline{w})$. The underlying logic to our argument holds even if selfed embryos do not all die – 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 a selfed offspring survives is less than $\frac{s(1-\alpha)}{4-s\,\alpha}$.

Complete recessivity is not required for the deterministic spread of this self-sacrificial allele. This self-sacrificial allele is favored when the primary selfing rate exceeds a critical selfing level, which equals

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

Figure 1 shows that at first approximation, crit.self is linear with its recessivity (h), such that a rare self-sacrificial A allele will spread when $\alpha > 6h$. That is, a greater probability of generating a low fitness selfed seed (i.e. a higher primary selfing rate) is required to favor the evolution of a self-sacrificial allele that can decrease fitness as a heterozygote.

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 recent inbreeding. To demonstrate this phenomenon, we apply Bayes' theorem to show that the probability that an embryo was the product of selfing, given that it is homozygous for A equals $\frac{\alpha\,p_A/2}{\alpha\,p_A/2+p_A^2(1-\alpha)}$ 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 cease to change (i.e. after 1000 generations and the sum of the change in absolute difference in genotype frequencies across generations is less than 10^{-14}). A's equilibrium frequency is highest when it is highly recessive and in a frequently self-fertilizing population (Figure 2), and it decreases as its dominance increases or the population's selfing rate decreases.

A genome-wide altruistic load. Is this deterministic selective benefit of self-sacrifice strong enough to overcome

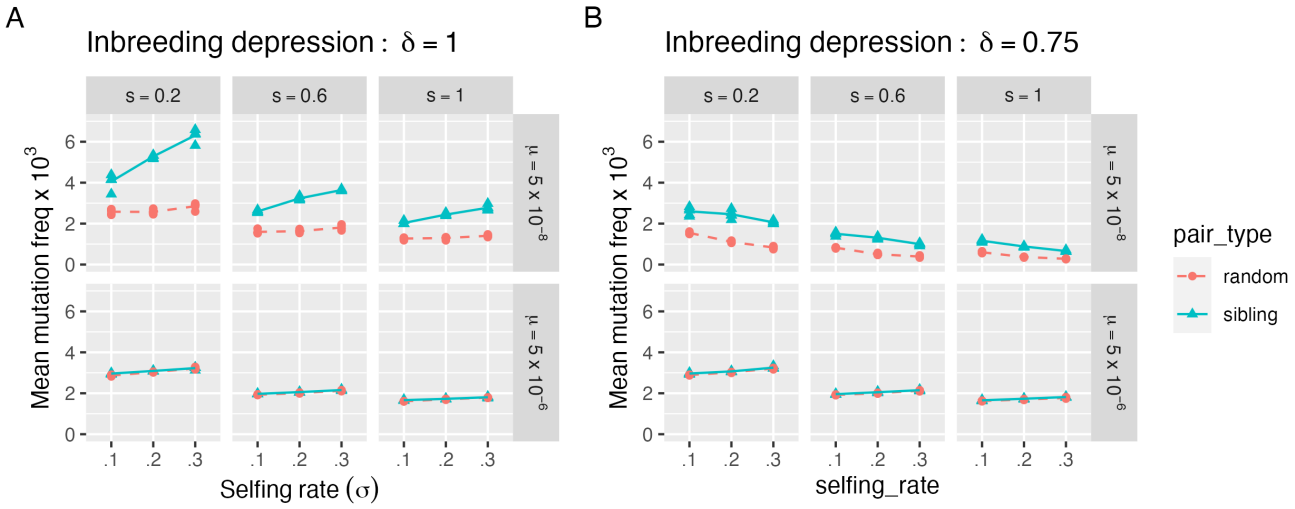

Fig. 3. The mean frequency of early acting, recessive lethal ($\delta = 1$, A), and sublethal ($\delta = 0.75$, B) alleles, over a combination of selfing rates (σ), mutation rates (μ), selection (s) and dominance (h) coefficients. Blue solid lines display results for cases in which primary embryos are replaced by siblings (our model), while red dashed lines display cases in which primary embryos are replaced by a non-sibling (a control).

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 ($\delta = 1$ Fig. 3A or $\delta = 0.75$, 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 S1). 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.

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 S1) and their lower frequency when segregating.

With sublethal inbreeding depression ($\delta = 0.75$, 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 (30). Humans can make up for embryo mortality by increasing reproductive effort (31, but see (32)). 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 (33). Similarly, angiosperms produce many more ovules than seeds (34). The evolutionary consequences of reproductive compensation have received much attention. For example, Hurst (35) recently argued that reproductive compensation provide selfish centromeres to distort meiosis. Additionally many researchers have followed up on the suggestion R.A. Fisher made to Race (36) that reproductive compensation in humans allows for evolutionary stability of polymorphism at the underdominant Rh locus (37, 38). More generally studies in plants (26, 39, 40) and mammals

(41, 42) 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 As such, our result settles an old debate between George Williams (43) and Richard Lewontin (44) as to whether reproductive compensation can maintain deleterious variants in the absence of mutation.

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.

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 (47), and more broadly males are thought to be "cursed" by mitochondrial alleles that can harm male fitness at a benefit to female fitness (48, 49). 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 (50). 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 (24) 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 (42)). A closer analogue to our model is Charlesworth's (23) 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 (51–53). For example, Archetti (52) showed that soft selection within families (akin to reproductive compensation in our model) favors alleles hypersensitive to deleterious mutations, and Lan et al (53) 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 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 (54). 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, (55, 56), 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 (57).

Testable predictions.

> Adaptation is a special and onerous concept that should not be used unnecessarily.
> — George C. Williams. 1966.
> *Adaptation and Natural Selection* (58).

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 (26, 59), counteracting the purging process in inbred populations (41).

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 (28), has been observed in self-incompatible populations (e.g. 60), and is likely 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, display high inbreeding depression, and have two embryos per archegonium, one of which dies before seed maturation (61). Our model makes sense of Williams' (62) 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 27, 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 (27), and the claim that death induced by the embryo lethal system is regulated by the position of the embryo rather than its competitive ability (62). 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 (63), 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 (64)).

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" (19) – 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 members of the Brandvain lab for comments and the Academy 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'. A draft of is available on bioRxiv (65), and code under-

lying the results are available on github https://github.com/ybrandvain/altrustic_load.

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. Darwin CR. On the various contrivances by which british and foreign orchids are fertilised by insects, and on the good effects of intercrossing, 1862.
5. 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.
6. 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.
7. Steven R. Seavey and Kamaljit S. Bawa. Late-acting self-incompatibility in angiosperms. *Botanical Review*, 52(2):195–219, 1986. ISSN 00068101, 18749372.
8. R. A. Brink and D. C. Cooper. Somatoplastic sterility in medicago sativa. *Science*, 90(2345): 545–546, 1939. ISSN 00368075, 10959203.
9. 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.
10. 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.
11. 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.
12. 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.
13. 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.
14. 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.
15. 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.
16. 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.
17. 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.
18. 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.
19. 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.
20. 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.
21. Norman A Johnson and M J Wade. Conditions for soft selection favoring the evolution of hybrid inviability. *Journal of Theoretical Biology*, 176(4):493 – 499, 1995. ISSN 0022-5193. doi: https://doi.org/10.1006/jtbi.1995.0215.
22. Jerry A Coyne. The evolutionary origin of hybrid inviability. *Evolution*, 28(3):505–506, Sep 1974. doi: 10.1111/j.1558-5646.1974.tb00781.x.
23. 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.
24. 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.
25. S. J. Gould and R. C. Lewontin. The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 205(1161):581–598, 1979. doi: 10.1098/rspb.1979.0086.
26. 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.
27. 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.
28. 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.
29. 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.
30. D W Mock. Siblicidal aggression and resource monopolization in birds. *Science*, 225(4663): 731–733, August 1984.
31. 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).
32. 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).
33. 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.
34. 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.
35. 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.
36. 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.
37. 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).
38. 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).
39. 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.
40. Edward J Klekowski. Genetic load and soft selection in ferns. *Heredity*, 49(2):191–197, 1982. doi: 10.1038/hdy.1982.85.
41. 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.
42. 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.
43. George C Williams. The genetic effects of reproductive compensation. *Human Biology*, 31 (4):344–351, 1959.
44. R. C. Lewontin. The effect of compensation on populations subject to natural selection. *The American Naturalist*, 87(837):375–381, 1953. ISSN 00030147, 15375323.
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. 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.
48. S. A. Frank and L. D. Hurst. Mitochondria and male disease. *Nature*, 383(6597):224–224, 1996. doi: 10.1038/383224a0.
49. 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.
50. 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.
51. 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.
52. 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.
53. 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.
54. 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.
55. 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.
56. 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.
57. Catherine A. Raiment, 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.
58. GC Williams. *Adaptation and natural selection: A critique of some current evolutionary thought*. Princeton University Press, Princeton, NJ, 1966.
59. I M Hastings. Reproductive compensation and human genetic disease. *Genetics Research*, 77(3):277–283, Jun 2001. doi: 10.1017/s0016672301004992.

60. 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.
61. Veikko Koski. On self pollination, genetic load and subsequent inbreeding in some conifers. *Communicationes Instituti Forestalis Fenniae*, 78:1–42, 1973.
62. Claire G Williams. Re-thinking the embryo lethal system within the pinaceae. *Botany*, 85 (7):667–677, 2007.
63. 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.
64. 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.
65. 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

Mating Table, Single locus model. Here we show all possible offspring genotype, its adult and embryo fitness, and its frequency in each maternally family after early embryo selection. Maternal families occur in frequencies of their genotype 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, ψ_{aa}, ψ_{Aa}, and ψ_{AA} denote the probability that an embryo is replaced by another (i.e. compensated) in aa, Aa, and AA families, respectively. The value, replace in maternal family, k equals $\text{replace}_k = \sum\sum \text{freq}_{\text{Geno:i, of type: j, in Fam:k}} \times$ embryo fitness$_i$ summed across all primary embryos for each maternal family. These values are: $\psi_{aa} = phs(1-\alpha)$, $\psi_{Aa} = s(2h+2p(1-\alpha)+\alpha)/4$, and $\psi_{AA} = s(p+(1-p)(h(1-\alpha)+\alpha))$.

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

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. The results of these calculations are somewhat long and tedious and are therefore presented in the supplemental Mathematica file.
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 σ is the selfing rate, p and q are frequencies of alleles a (wild type), and A (self-sacrificial), respectively, and subscripted ψ is the probability that a primary embryo in the subscripted maternal genotype is replaced.
From the table above, we find that the frequencies of each genotype in the next generation are as follows:

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

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

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

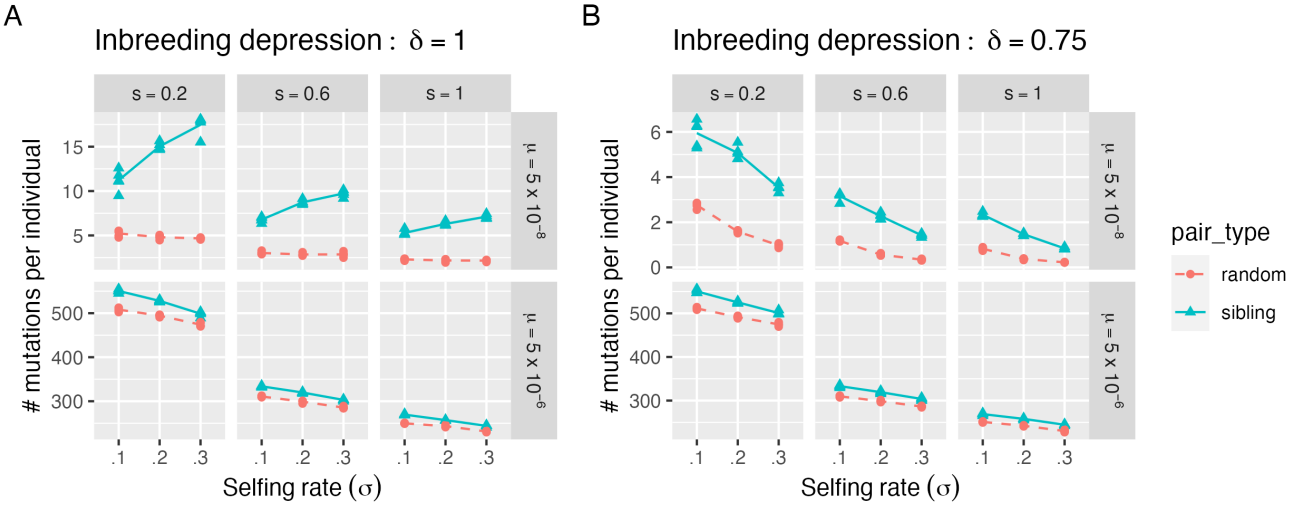

Fig S1: The mean number of early acting, recessive lethal ($\delta = 1$, A), and sublethal ($\delta = 0.75$, B) alleles per individual, over a combination of selfing rates (σ), mutation rates (μ), selection (s)and dominance (h) coefficients. Blue lines display results for cases in which primary embryos are replaced by siblings (our model), while blue lines display cases in which primary embryos are replaced by a non-sibling (a control).