Title page

The evolutionary fate of divergent genomes in hybrid zones between selfing and outcrossing species

Fréderic Fyon[1, *], Waldir M. Berbel-Filho[2]

[1] Department of Biology, Royal Holloway University of London, Egham, TW20 0EX, UK.

[2] Department of Biology, University of Oklahoma, Norman, Oklahoma, USA.

* Corresponding author: frederic.fyon@gmail.com

Keywords: Hybridization – Mating systems – Mutational load – Recombination – Mutation rate – Self-fertilization

Abstract

Hybridization is a natural process whereby two diverging evolutionary lineages reproduce and create offspring of mixed ancestry. Differences in mating systems (e.g., self-fertilization and outcrossing) are expected to affect the direction and extent of hybridization and introgression in hybrid zones. Among other factors, selfers and outcrossers are expected to differ in their mutation loads. This has been studied both theoretically and empirically; however, conflicting predictions have been made on the effects mutation loads of parental species with different mating systems can have on the genomic composition of hybrids. Here we develop a multi-locus, selective model to study the effects of selective interference and homozygosity on (i) the mutation load of selfers and outcrossers, and (ii) the long-term genetic composition of resulting hybrid populations. Notably, our results emphasize that in the long term, genes from the parental population with lesser mutation load will be over-represented in hybrid genomes, regardless of the hybrids own mating system. Under a large portion of the parameter space, we predict that selfers' genomes are of higher quality and will prevail. However, this will ultimately depend on the intensity of genetic drift and selective interference in the system considered. Taken together, these results provide a "null" hypothesis for the genetic composition of hybrids in hybrid zones containing species with different mating systems.

1. Introduction

Hybrid individuals are defined as progeny of mixed ancestry; they bear genes inherited from divergent evolutionary lineages. Introgressive hybridization (i.e., formation of hybrids with backcrossing) functions as bridges between divergent lineages, allowing gene flow and shared evolutionary history. Identifying the processes shaping the evolutionary fate of hybrid populations is a fundamental task for evolutionary biology, as well as for conservation, and wildlife management.

The possibility of hybridization happening, along with its ecological and genomic consequences, depends, among other factors, on the mating systems and reproductive biology of the parental species and their hybrids. Most plants and many animals are simultaneous hermaphrodites (JARNE AND AULD 2006; THE TREE OF SEX CONSORTIUM 2014). In these species, self-fertilization (selfing) – offspring produced by the fusion of gametes stemming from different meiosis in the same individual – is a possibility. Selfing contrasts with outcrossing as in the latter the progeny is produced by the fusion of gametes proceeding from different meiosis in different individuals (hermaphrodites or not). Mating systems are usually defined in three main categories according to the selfing rate: obligate outcrossing (selfing rate ≤ 20%), predominantly selfing (selfing rate > 80%), and mixed-mating (selfing rate > 20% and ≤ 80%) (SHIMIZU AND TSUCHIMATSU 2015). It is only recently that attention has been brought on the effects that different mating systems have on the occurrence of hybridization and on the genomic composition of hybrid populations (HU 2015; KIM *et al.* 2018; PICKUP *et al.* 2019).

Hybrid zones are regions of contact between divergent species/lineages that often produce individuals of admixed ancestry through interspecific mating. Hybrid zones between species with

different mating systems (e.g., predominantly selfing and obligate outcrossing) are commonly reported in plants (RUHSAM *et al.* 2011; PICKUP *et al.* 2019; OSTEVIK *et al.* 2021), but examples are scarce in animals (BERBEL-FILHO *et al.* 2021). Some predictions on how reproductive systems (BRANDVAIN AND HAIG 2005) and reproductive timing (MARTIN AND WILLIS 2007; BERBEL-FILHO *et al.* 2021), along with other ecological factors (BUSCH *et al.* 2022), can influence the occurrence and outcome of hybridization in these cases have been extensively provided. On the other hand, few studies have tried to address how mating systems differences influence (i) the quality of each parental genomes inherited by first-generation hybrids (F1), and (ii) how the genetic composition (defined in terms of genes' species-of-origin) of hybrids evolves in the long-term. Disparate, sometimes contradictory predictions on these points have been made; the present work was designed to disentangle these, provide a unified framework, and introduce clear and testable predictions.

Predictions concerning hybrids' genetic composition and evolution have focused on mutation load. They rely on the idea that species with different mating systems are expected to harbor different mutation loads (ARUNKUMAR *et al.* 2015; PICKUP *et al.* 2019). High homozygosity caused by long-term selfing means that recombination is – most of the time – inefficient at shuffling genetic combinations, which results in higher rates of hitch-hiking, background selection, and selective interference (WRIGHT *et al.* 2013). That is, selection is expected to be less efficient in selfing populations, which are thus predicted to have more slightly deleterious alleles become fixed by random genetic drift (CHARLESWORTH *et al.* 1993). This relates to the classic argument that the absence of effective recombination in asexual and self-fertilizing species should lead them to extinction due to mutational meltdown (GABRIEL *et al.* 1993; LYNCH *et al.* 1993; LYNCH *et al.* 1995). This predicts that genomes of self-fertilizing species tend to be

of worse quality. On the other hand, however, highly deleterious recessive alleles are expected to be purged by selection in selfers as they are "exposed" to selection by high homozygosity, while in outcrossers they can be maintained at low frequencies as they are partly hidden from selection in heterozygotes (WANG *et al.* 1999). In other words, here selection is predicted to be more efficient in self-fertilizing species, resulting in genomes of better quality in selfers. Overall, selfers' genomes are expected to harbor more fixed slightly deleterious, codominant mutations and less strongly deleterious, recessive mutations (ARUNKUMAR *et al.* 2015). How are these effects expected to combine to determine genetic composition and evolution of hybrid populations?

The answer to this question is likely to depend on several parameters of the selection processes in parental species, but also on the hybrid mating system. Highly deleterious, recessive mutations coming from outcrossing ancestors may rapidly be purged by selection if the hybrids self-fertilize (as these mutations rapidly become homozygous and visible to selection), taking together any linked genes. This process is likely to limit the introgression of outcrossers' genomes into selfers' (Pickup et al., 2019). On the other hand, the many fixed slightly deleterious, codominant mutations coming from the selfing ancestors should be purged out in an outcrossing hybrid, this time restricting introgression of self-fertilizers' genomes into outcrossers' (PICKUP *et al.* 2019). Under this argument, everything happens as if each parental genome was "adapted" to the mating system it has evolved in and tends to fare worse in alternative mating systems. This raises the question, does the mating system of the hybrid population have a critical influence on the evolution of hybrid genome composition, or is this simply determined by the respective quality of the parental populations' genomes?

Here, we aim at answering these two questions by building a multi-locus model allowing for selection, recombination, mutation, and drift to interplay and modify simple genomes. Running numerical simulations for a diversity of scenarios, we were able to study the effects of selective interference and homozygosity on (1) the relative mutation loads of self-fertilizing and outcrossing populations, and (2) the long-term genetic composition of hybrid populations deriving from parental populations with different mating systems. First, our work allows to confirm some of the classic arguments on mutation load in self-fertilizing species. However, we argue that the mutational meltdown argument, as it relies on selective interference, is a weak force that is most of the time not expected to have a strong influence on the mutation load of self-fertilizing species (FRANKHAM 2005). Only in systems plagued with strong genetic drift (e.g., population bottlenecks, colonization events) should selective interference be strong enough to provide selfers with a stronger mutation load than outcrossers. Second, we were able to determine the parameters that have a notable influence on the genetic composition and evolution of the hybrid population. Interestingly, we show that the mating system of the hybrid population does not largely influence the outcome of the simulations. Hybrid genetic composition in the long term appears to be primarily dictated by the relative mutation load of the parental species: genomes of worse quality get preferentially eliminated. This prediction provides a "null hypothesis" for the study of genetic composition of hybrid species deriving from ancestors with different mating systems.

2. Methods

In the following, we present the model we used for this work. We used SLiM3 (HALLER AND MESSER 2019) software to code and run the model. SLiM3 has been used in a diversity of population genetics studies as it provides a resource-effective and flexible computing framework.

Because the model is of a stochastic nature, each simulation is repeated 100 times, and mean values as well as standard deviations are calculated from these 100 iterations. The code used here is available on the following GitHub repository: https://github.com/FredericFyon/Mating-systems.

2.1. Theoretical populations

Our aim here is to follow the genetic alleles at different loci in populations with different mating systems, along with their hybrids. To do so, we always model three populations: (1) an exclusively or mostly self-incompatible (SI) population that reproduces via obligate outcrossing (outcrossers); (2) a self-compatible (SC) population that reproduces exclusively or mostly by means of self-fertilization (selfers); (3) a population born by hybridization of the two previous ones. Populations do not interact with each other (except at the one generation where hybridization happens) and do not compete. Each population size is set at N_{pop}. In the following, we call σ_i the rate of self-fertilization of a given population i. We refer to the self-fertilization rate of the outcrossing, self-fertilizing, and hybrid populations as $\sigma_o \leq 0.1$, $\sigma_s \geq 0.9$ and σ_h respectively.

2.2. Genetic architecture

The genome of all individuals in the three populations consists of N_{loc} loci, equally distributed along a chromosome. We note r the recombination rate between two adjacent loci. We consider here an infinite-allele model: each mutation occurring is unique, all loci potentially have an infinite number of alleles. In this model we only consider deleterious mutations and note s the detrimental effect of a particular mutation on the fitness of the host individual.

2.3. Life Cycle

At each generation, selection is modelled by choosing the parents of every offspring born at this generation. To do that, we first determine if an offspring was produced by self-fertilization or not. With probability σ, it is produced by self-fertilization. In that case, we randomly draw one parent from the distribution of individuals of the same population at the previous generation. In contrast, with probability 1 - σ, the offspring is producing by outcrossing: we randomly draw two parents. Drawn parents are then accepted with a probability equal to their fitness w_i. We assume fitness to be purely multiplicative among loci (we assume no epistasis): $w_i = \prod_j w_{ij}$, with $w_{i,j}$ being the contribution of locus $j \in [1, N_{loc}]$ to the fitness of individual i. $w_{i,j}$ depends on (1) the detrimental effect s_{ij}^1 and s_{ij}^2 of the alleles present on the two homologous chromosomes at locus j of individual i, and on (2) the dominance coefficients h_{ij}^1 and h_{ij}^2 of these two alleles. If $s_{ij}^1 \neq s_{ij}^2$ (heterozygote), $w_{i,j}$ is calculated as:

$$w_{i,j} = 1 - \sum_{k=1}^{2} h_{ij}^k \, s_{ij}^k$$

If $s_{ij}^1 = s_{ij}^2 = s_{ij}$ (homozygote), then $w_{i,j}$ simply is: $w_{i,j} = 1 - s_{ij}$.

Following empirical data suggesting that the dominance of a deleterious mutation is negatively correlated with its detrimental effect (AGRAWAL AND WHITLOCK 2011), we assume in this model a negative exponential relationship between s and h:

$$h_{ij}^k = \frac{e^{-\alpha s_{ij}^k}}{2}$$

with α being an arbitrarily chosen parameter that allows to modulate this relationship. Here, we will always assume that $\alpha = 7$ for simplicity. This allows to approximately match empirical data

of mean dominance of mildly deleterious mutations being around 0.25 (MANNA *et al.* 2011). The exponential function allows for a long-tailed asymptotic distribution of dominance effects, letting highly deleterious mutations to have a very low but strictly positive dominance. Overall, this allows the whole selective process to be determined by only one parameter: the distribution of detrimental effects of deleterious mutations.

In addition to selection, the life cycle includes recombination and mutation. Recombination happens at a rate r being any two adjacent loci. A recombination event exchanges the composition of the two homologous chromosomes in an individual from the recombination point to the downstream end on the chromosomes. Of course, multiple recombination events can take place during any meiotic event, leading to potentially mosaic inherited chromosomes.

Mutations happen with probability μ per locus per chromosome per individual. When a mutation occurs, we draw its detrimental effect from a negative exponential distribution of mean $1/\lambda$. This means that there is a majority of slightly deleterious, codominant mutations arising, along with few highly deleterious, strongly recessive mutations.

2.4. Hybridization

At first, we consider only the parental populations of outcrossers and selfers. These reproduce within themselves for a certain number of generations t_1. Then, a hybridization event happens. Hybridization is modelled by created a third, hybrid population. This population is created by drawing for each individual one parent from each parental population. As a result, at $t_1 + 1$, every individual of the hybrid generation bears one chromosome from one parental population, and the other from the other parental population. After that initial hybridization event, we consider that there is no additional gene flow from the parental populations. We let the hybrid

population reproduce within itself and evolve in the absence of additional mutation. This allows us to determine the evolutionary fate of the parental genomes within the hybrid population.

2.5. Parameter values

Unless specified otherwise, simulations throughout this work have been run with the following parameter values: $Npop = 500$, $Nloc = 100$, $\lambda = 10$. We discuss later the influence of these parameters and provide examples with some other values in the Supplementary Material.

3. Results

3.1. Mutational Load in Outcrossers and Selfers

We present on Fig. 1 the characteristics of the mutational load born by two populations, one purely outcrossing and another one purely self-fertilizing.

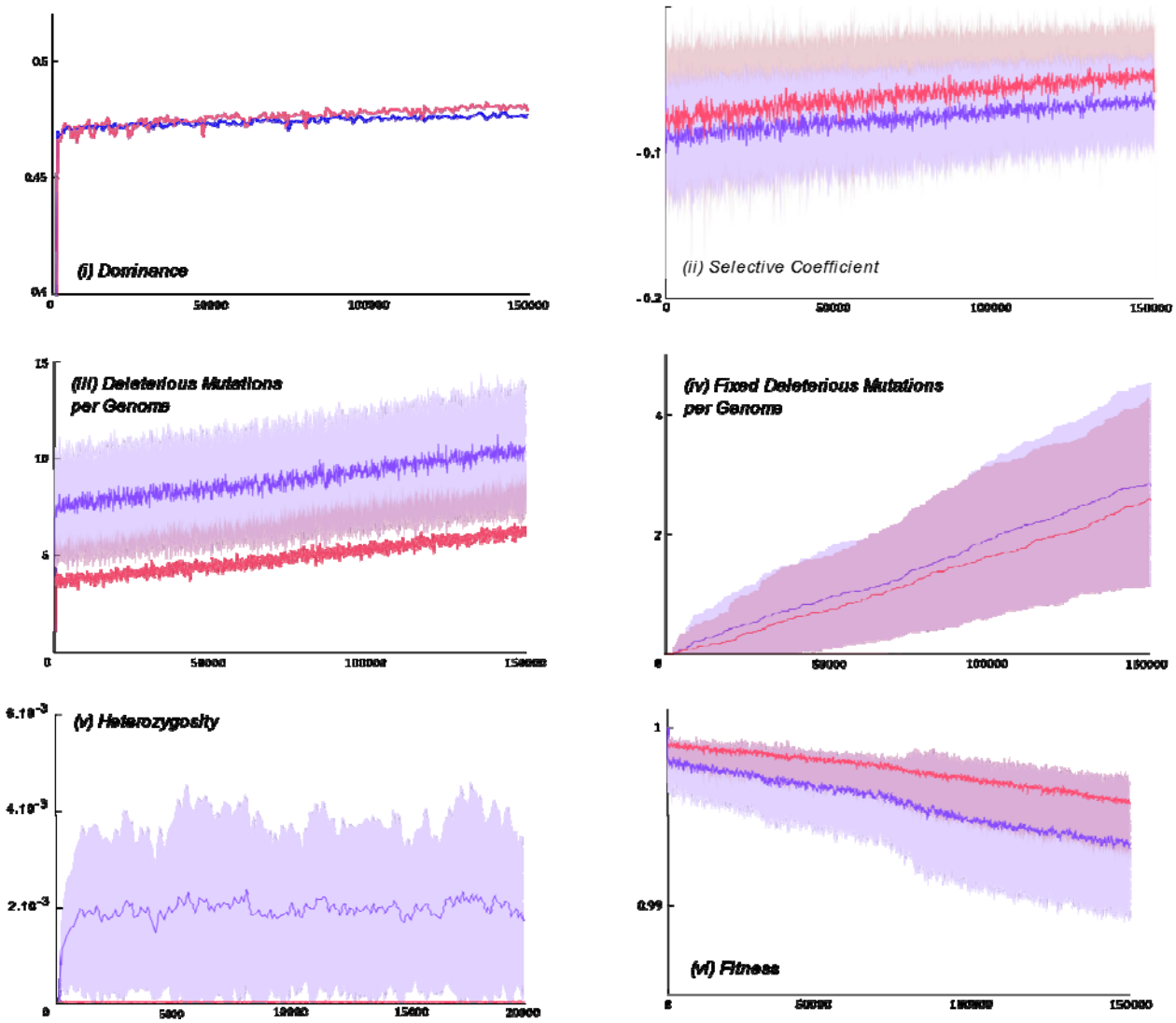


Fig. 1. Genomic composition of an outcrossing population (purple, selfing rate of $\sigma_o = 0$) and a self-fertilizing population (red, selfing rate of $\sigma_s = 1$) left to evolve during 150,000 generations (x-axis). Panels represent: (i) dominance of deleterious mutations, (ii) selective coefficients of

deleterious mutations, (iii) number of deleterious mutations per genome, (iv) number of fixed deleterious mutations per genome and (v) heterozygosity and (vi) fitness of individuals. In each graph, plain lines represent mean values, while shaded areas represent standard deviation. Results are illustrated for the following parameter values: mutation rates $\mu = 10^{-5}$, recombination rate $r = 10^{-6}$.

The results found here are consistent with classic population genetics arguments. While outcrossers' genomes bear a part of its polymorphism in a heterozygotic form, selfers´ genomes are mostly homozygous (Fig. 1.(i)). This leads to recessive mutations in selfers to be revealed to selection, in contrast with outcrosser, and thus to decrease in frequency or be eliminated altogether via purging. Because we assume a negative correlation between dominance of alleles and their detrimental effect on fitness, recessive mutations tend to be highly deleterious ones. Thus, selfers eliminate the highly deleterious mutations more than outcrossers, leading to the mean detrimental effect of deleterious mutations present in selfers' genomes to be reduced in comparison with the mean detrimental effect of deleterious mutations present in outcrossers' genome (Fig. 1.(ii)). High homozygosity in selfers strengthen the efficacy of selection, leading to a better purge of deleterious, recessive mutations and improving individuals' fitness (Fig. 1.(v)). Because of this effect, less highly deleterious mutations are maintained as standing variation (polymorphism) in the self-fertilizing population (Fig. 1.(iii)). The total number of deleterious mutations per genome tends to increase with time. This is due to the progressive fixation of slightly deleterious mutations by genetic drift (the increase in Fig. 1.(iii) corresponds to the increase in Fig. 1.(iv)).

There does not seem to be a strong difference in the rates of random fixation of slightly deleterious alleles between the two populations. One mutation becomes fixed about every 55,000

generations in the outcrossing population, and about every 62,000 generations in the self-fertilizing population. Note that this difference is much smaller than the standard deviations of the two population. In accordance with classic population genetics results, the fixation rate in our simulations is of the order of the mutation rate (10^{-5}).

Because the rate of fixation of deleterious mutations in selfers and outcrossers is similar, differences in selection and fitness are only due to the increase in selection efficacy in the self-fertilizing population associated with increased homozygosity. As a result, though the fitness of the two populations decrease at a steady, similar rate because of the fixation of deleterious mutations, the fitness of the selfers is consistently higher than the fitness of the outcrosser (Fig. 1.(v)). In the case illustrated here, the selfers have an overall better genome, a genome better purged from highly deleterious, recessive alleles.

We show on Figure 2 the accumulation of randomly fixed deleterious mutations for a diversity of parameter values.

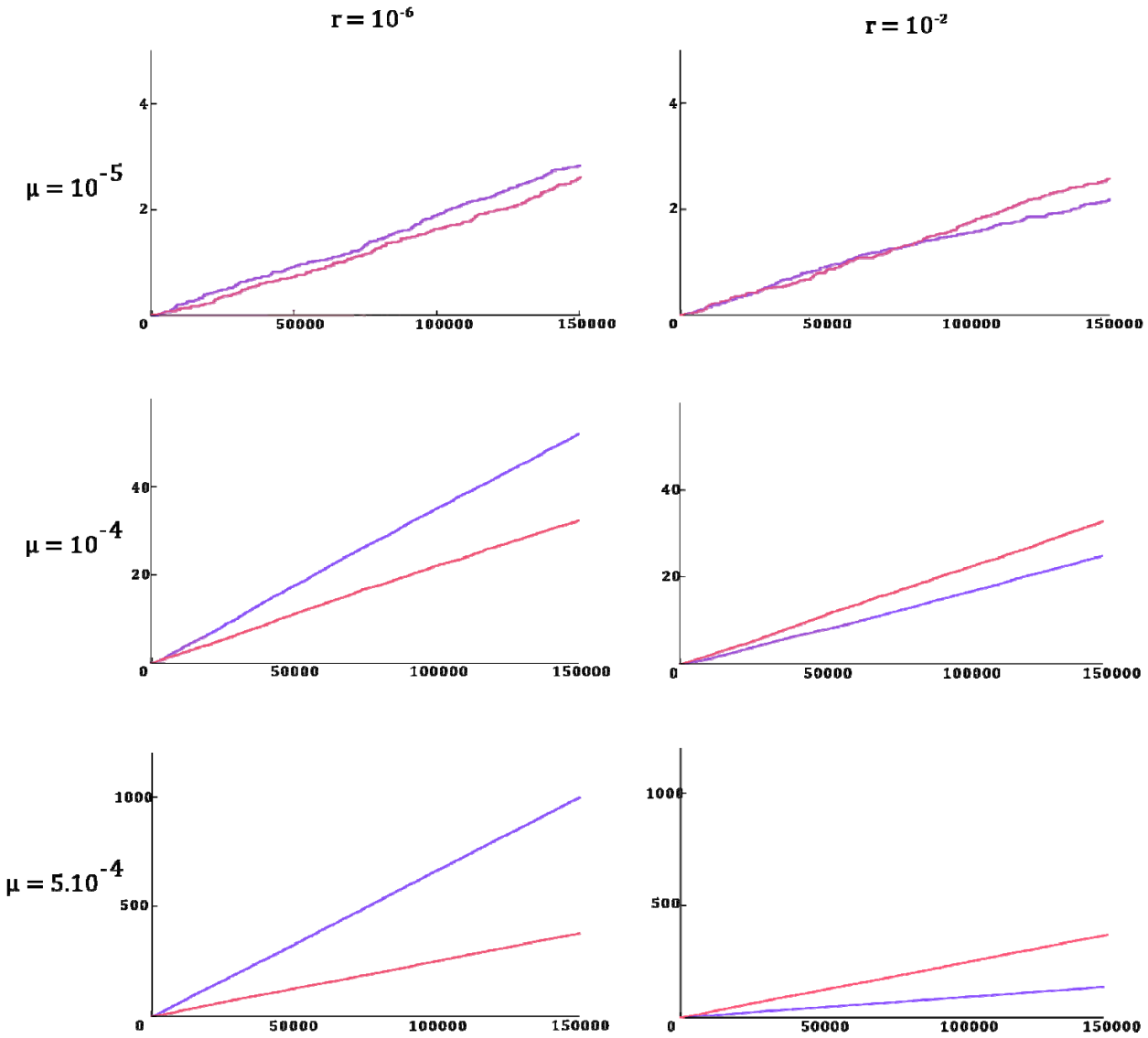


Fig. 2. Number of fixed deleterious mutations (y-axis) during 150,000 generations (x-axis) in an outcrossing population (purple, selfing rate of $\sigma_o = 0$) and a self-fertilizing population (red, selfing rate of $\sigma_s = 1$). In each graph, plain lines represent mean values, while shaded areas represent standard deviation. Values of r and μ are specified and change as columns and rows respectively.

According to the relationship between the rate of random fixation of deleterious mutation and the mutation rate, we see the number of fixed mutations at any given time increasing with μ.

Increased mutation rates, as it appears, make more obvious a trend that mostly depend on the recombination rate *r*. When *r* is low, both the outcrossing and the selfing populations have effectively low recombination rates. In that case, the outcrossing population appears to fix deleterious mutations more rapidly than the selfing population. This is because recessivity partly hides the detrimental effect of mutations when heterozygotes, which makes it easier for them to randomly drift from low to intermediate frequencies. Mutations that are more likely to reach intermediary frequencies are also more likely to drift to fixation. In contrast, when *r* is high, we see that the selfing population accumulates fixed mutations more rapidly than the outcrossing population. In fact, the rate of accumulation of fixed mutations in the selfing population does not change in comparison with the previous case of low *r* values – which was to be expected since recombination is ineffective in selfers and thus should not have any effect. The increase in *r* has led to a strong reduction in the accumulation of fixed deleterious mutations in the outcrossing population. This is because increased recombination rate improves selection efficacy in this population, leading to a decrease of the probability that a given deleterious mutation can drift to fixation by chance. Here, we see two antagonistic effects. First, selection efficacy is reduced in outcrossers due to heterozygosity and recessivity. Second, selection efficacy is increased in outcrosser due to recombination and decreased selection interference. We note that the second effect can overcome the first only if (i) the mutation rate or the evolutionary time considered are large; and if (ii) recombination in the outcrossing population is high enough in the genomic segment considered so that it prevents much of selective interference effects.

3.2. Mutational Load in Hybrids between Outcrossers and Selfers

In the following, we look at the proportions of genes in the hybrid population coming from each of the parental populations. On Fig. 3, we consider a case where $\mu = 10^{-5}$ and where the parental populations have diverged during 10,000 generations prior to hybridization.

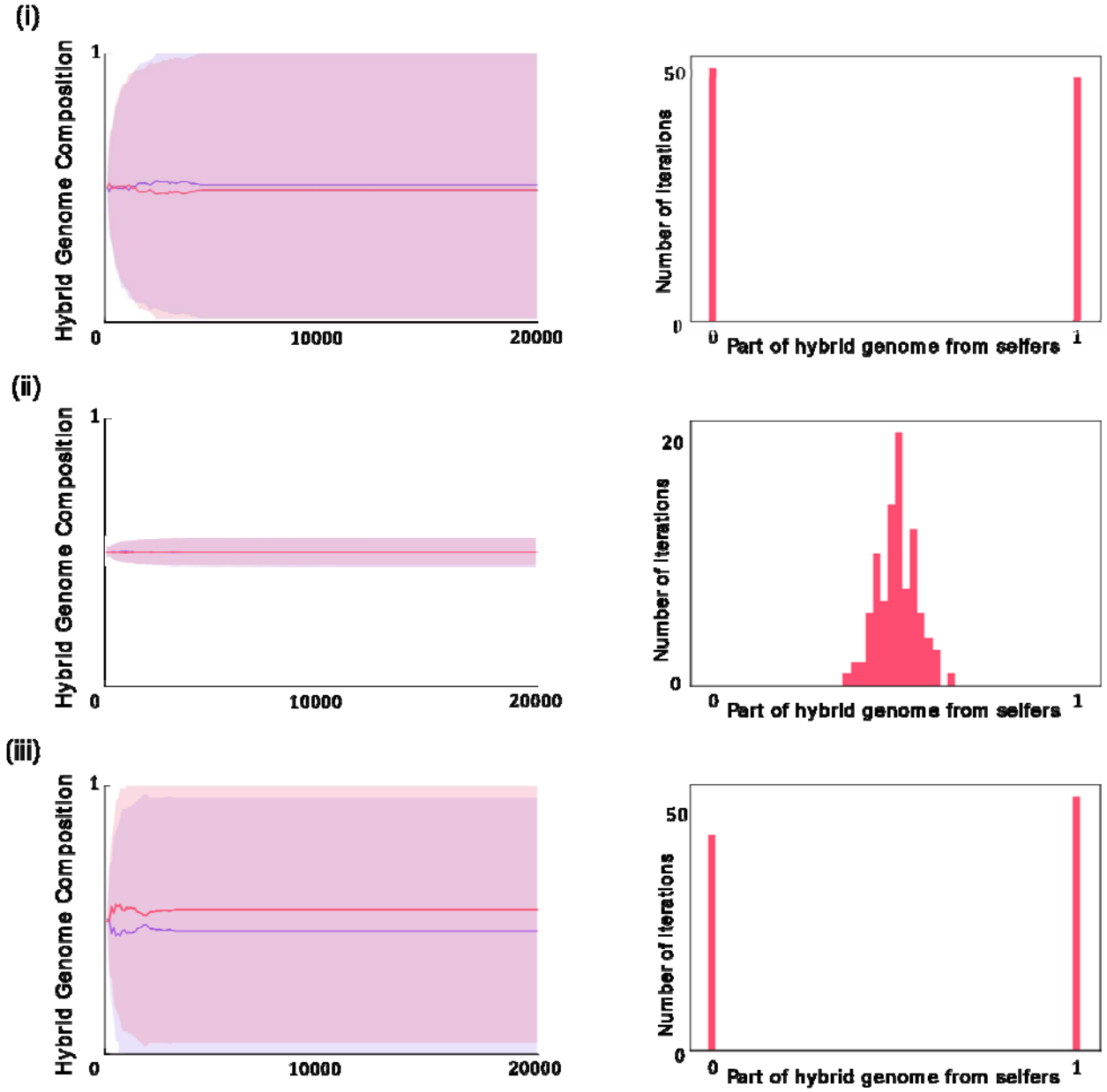


Fig. 3. Hybrid genome composition during 20,000 generations after hybridization for $\mu = 10^{-5}$. Three cases are illustrated: (i) the hybrid population reproduces by outcrossing, and

recombination rate is small: $r = 10^{-6}$; (ii) the hybrid population reproduces by outcrossing, and recombination rate is large: $r = 10^{-1}$; (iii) the hybrid population reproduces by self-fertilization, and recombination rate is large $r = 10^{-1}$. The first column represents the evolution of the genetic composition of the hybrid population. Purple represents the relative part of the genome inherited from the outcrossing population, while red represents the relative part of the genome inherited from the self-fertilizing population. Plain lines represent mean values, while shaded areas represent standard deviations. The second column represents a histogram of the mean value of the relative part of the genome that comes from the self-fertilizing population after 20,000 generations. Each case is iterated 100 times, so that the histograms represent the distribution of the results obtained at the end of the 100 iterations.

Fig. 3 shows that on average, we expect that the genes from each parental populations participate to hybrids´ genome in equal proportion. Fig. 1 shows us that the genomes of the parental populations for such parameter values differ substantially for different criteria. Relative to the outcrossing population´s genome, the selfing population´s genome show: (i) increased dominance and (ii) decreased mean selective effect of deleterious mutations present, (iii) reduce number of segregating deleterious mutations, (v) depleted heterozygosity and overall, (vi) a slightly better fitness. These do not appear to make any difference when the genomes from the two populations are crossed in hybrids. Interestingly, the only parameter that does not show any substantial difference between the genomes from the two populations is the number of fixed deleterious mutations.

Although the average hybrid genome composition in the three cases illustrated on Fig. 3 is the same, its variance varies widely. In case (i), the hybrid population outcrosses but its recombination rate is set to a low value; in case (iii), the hybrid population self-fertilizes. In the

two cases, recombination is effectively absent or rare in hybrids (even though recombination rate is set to a large value in (iii)). Consequently, haplotypes inherited from parental populations remain unchanged over time. At the end of the simulations, there is always one haplotype that managed to be fixed. Because a lot of genomes from the two populations do not have any deleterious mutations ($2N\mu = 0.01$), the haplotype that eventually gets fixed, which most of the time will be a haplotype with no deleterious mutation, can come from any of the two populations with almost equal probabilities. In case (ii), the hybrid population outcrosses and its recombination rate is set to a high value. Haplotypes from the two populations recombine often; the haplotype that will be fixed, as a result, presents a mix of genes from the two parental populations.

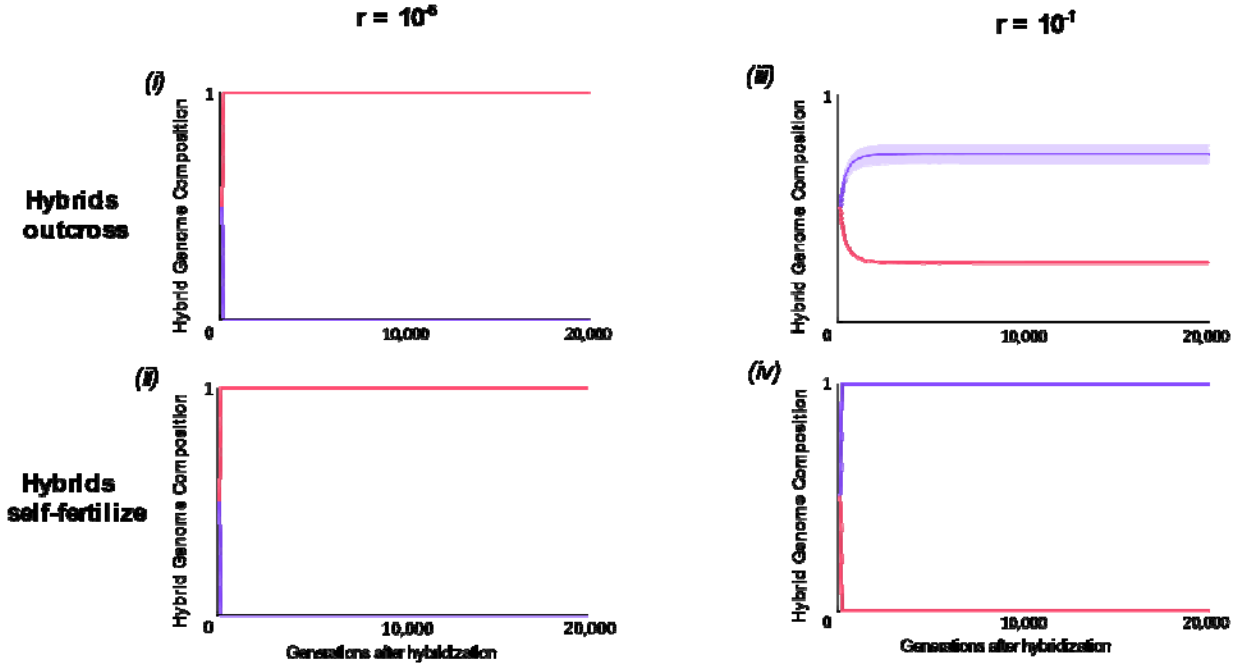


Fig. 4. Hybrid genome composition during 20,000 generations after hybridization for $\mu = 5.10^{-4}$. Four cases are illustrated, depending on the value of r used (columns) and on whether the hybrids outcross (first row) or self-fertilize (second row). Purple represents the relative part of

the genome inherited from the outcrossing population, while red represents the relative part of the genome inherited from the self-fertilizing population. Plain lines represent mean values, while shaded areas represent standard deviation.

The results of Figure 3 strongly rely on the fact that deleterious mutations are rare, so that many genomes are completely free of them or harbor very few of them. In the following, we consider the alternative case where there are many detrimental mutations around, either fixed or segregating. Here, we use mutation rate values that are not realistic but allow us to simulate a scenario with many mutations around, which is arguably more realistic than the scenario illustrated on Fig. 3 where most genomes are "perfect".

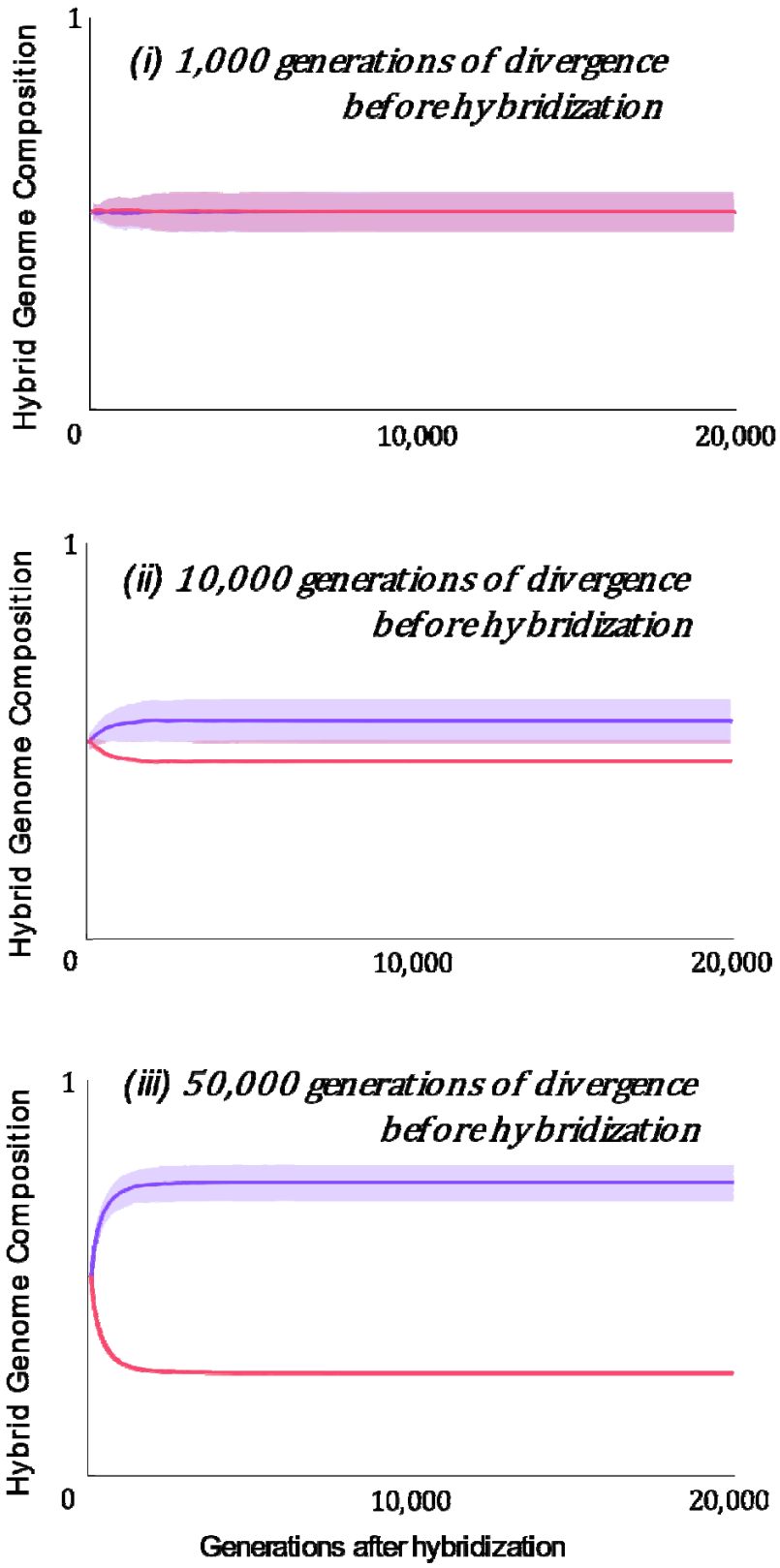
On Figure 4 we illustrate this scenario of many detrimental mutations accumulating in genomes, in a diversity of cases. In cases (i) and (ii), the recombination rate between loci is small ($r = 10^{-6}$). As illustrated on Fig. 2, for such a mutation and recombination rates the outcrossing population accumulates detrimental mutations faster than the selfing population. When the hybridization event happens, outcrossers´ genomes harbor more detrimental mutations than the selfers´ genomes, and thus are of lesser quality. Because in case (i) recombination in hybrids is rare and in case (ii) hybrids reproduce via self-fertilization so that recombination is even less efficient, the genome that fix in the hybrid population tends to only harbor genes from one population or the other (unshuffled haplotypes); we see that it is always a selfer genome that goes to fixation because it tends to harbor less deleterious mutations.

In cases (iii) and (iv), the recombination rate between loci is large ($r = 10^{-1}$). In that case, as illustrated on Fig. 2, the selfing population accumulates deleterious mutations much faster than the outcrossing population. This effect rapidly overcome the selection efficacy reduction in outcrossers due to recessivity, and self-fertilizer´s genomes become of worse quality than

outcrossers´ genomes. In case (iii), recombination happens in hybrids such that the genome that will eventually fix tends to be composed of a mix of genes from each parental population. Because genomes from the outcrossing population are overall better, their genes tend to be over-represented: the genes from the self-fertilizing population are selectively eliminated as they present many more deleterious mutations. In case (iv), the hybrid population self-fertilizes, which means that recombination is effectively absent in it. Consequently, the genome that goes to fixation tends to be composed only of genes from one parental population. Because outcrossers' genomes are of better quality, Fig. 4 shows that selfers´ genes are rapidly wiped out after hybridization.

Overall, we see that the genomic composition of hybrids between species with different mating systems mostly depends on the recombination rate between loci. If recombination is low, outcrossers accumulate more mutations than selfers and their genome tend to disappear in hybrids. In contrast, if recombination rate is high, selfers' genomes accumulate more deleterious mutations than outcrossers and their genes tend to disappear in hybrids. Overall, hybrid genomic composition seems primarily directed by parental populations' mutation load and genome quality.

The only case where hybrids' genomes become a mix of genes from the two populations is when the hybrids outcross and recombination is frequent. In that case, we illustrate on Fig. 5 how the hybrid genome composition changes depending on time of divergence prior to hybridization. As the populations have gone through longer divergence, the selfers´ genome is of worst quality compared with outcrossers´ genomes (see Fig. 2). Consequently, and as illustrated on Fig. 5, the part of hybrid genome descending from self-fertilizing population ancestry tend to decrease with time prior to hybridization.


(i) 1,000 generations of divergence before hybridization
(ii) 10,000 generations of divergence before hybridization
(iii) 50,000 generations of divergence before hybridization
Hybrid Genome Composition
1
0
10,000
20,000
Generations after hybridization

338

Fig.5. Hybrid genome composition during 20,000 generations after hybridization for $\mu = 5.10^{-4}$. Here, different divergence times between the parental populations prior to hybridization are illustrated (specified above each graph). Purple represents the relative part of the genome inherited from the outcrossing population, while red represents the relative part of the genome inherited from the self-fertilizing population. Plain lines represent mean values, while shaded areas represent standard deviation.

4. Discussion

The transition from obligate outcrossing to predominantly selfing has happened multiple times in plants and animals (JARNE AND CHALESWORTH 1993; JARNE AND AULD 2006; THE TREE OF SEX CONSORTIUM 2014). This transition comes with a combination of phenotypic and genomic changes collectively known as the "selfing syndrome" (SHIMIZU AND TSUCHIMATSU 2015; CUTTER 2019). Among those changes, outcrossers and selfers are expected to differ in their mutation load. Theory predicts that higher levels of heterozygosity in outcrossing species allows to better hide mutations that are simultaneously highly deleterious and highly recessive from selection, resulting in these reaching higher frequencies in outcrossers when compared to selfers (WANG *et al.* 1999). In contrast, theory also predicts that higher levels of homozygosity in self-fertilizing species result in less efficient recombination events, and thus stronger selective interference (WRIGHT *et al.* 2013; HARTFIELD *et al.* 2017). As a result, slightly deleterious, co-dominant alleles are more likely to escape purging out by selection and increase in frequency by random genetic drift (CHARLESWORTH *et al.* 1993). Our results generally agree with these predictions. As it turns out, there seems to be pros and cons for both outcrossing and self-fertilization in terms of mutation load. An important question is, which of the two processes is the strongest and most influences the genetic content of outcrossing and selfing species?

Our results showed that selfers are expected to display overall as good or even better genomes (in terms of mutation load) when compared to outcrossers, except under certain conditions. Only with unreasonably high mutation rates, or after unreasonably long evolutionary times, we see selfers accumulating more fixed deleterious mutations and evolve worse genomes. We show here that selection interference is generally not expected to lead to a greater accumulation of mutations in selfers because it is a relatively weak (slow). In addition, this effect depends on a reduction of efficiency of recombination between loci in selfers: it should not result in relatively weaker selection if recombination is also small in the outcrossing population. If recombination is rare in both selfers and outcrosser, we see that mutation accumulation can even be greater in outcrossers. This is because heterozygosity and recessivity allow deleterious mutations to reach higher frequencies in an outcrosser, and thus get more easily fixed by drift. As such, we predict that relatively greater mutation accumulation by selective interference in selfers may concern preferentially genomic areas that exhibit high recombination rates in closely-related outcrossers. This outcome recovers the results of a recent multi-locus simulation study that showed that at low recombination rates outcrossers accumulate deleterious mutations in a selective regime called pseudo-overdominance (SIANTA *et al.* 2022). Overall, we argue that all else being equal, self-fertilizing species may often have genomes of higher quality than outcrossing species.

Mutation accumulation by selective interference highly depend on genetic drift and the effective population size N_e. On one hand, increasing population size limits the fixation of deleterious mutations. In simulations shown in Supplementary Material Figure S1, with $\mu = 10^{-5}$ and $r = 10^{-6}$, increasing population size from 500 to 5000 decreases the deleterious mutation fixation rate approximately by a factor 10. On the other hand, there are particular demographic events that leads to N_e values much lower than the actual population size such as population bottlenecks or

colonization events. Importantly, self-fertilizing species may be prone to this kind of events as they can in principle more easily colonize new environments and found new populations (THEOLOGIDIS *et al.* 2014; NOEL *et al.* 2016; GROSSENBACHER *et al.* 2017). This may explain empirical patterns of relaxed purifying selection observed in some self-fertilizing species (SLOTTE *et al.* 2010; BURGARELLA *et al.* 2015; WANG *et al.* 2021). However, evidence for purifying selection in selfers is often mixed (GLÉMIN *et al.* 2006; HAUDRY *et al.* 2008; ESCOBAR *et al.* 2010). This mixed empirical evidence may be caused by the demographic history of the species as mentioned above, but also due to other factors such as divergence time between selfers and outcrossers, biased gene conversion, and stability of selfing as a reproductive strategy (ESCOBAR *et al.* 2010),

Although the fields of mating system evolution and hybridizations have largely developed independently, they overlap substantially on questions regarding on how mating system differences (and their consequences) can act as pre and/or postzygotic barriers for introgression. They are also intertwined with respect to the effects of increased or decreased inbreeding levels on species evolutionary histories (PICKUP *et al.* 2019). Following previous arguments, we also wanted to investigate here the consequences that the differences of mutation load build-up that outcrossing and self-fertilizing species exhibit could have on the genetic composition of hybrids between species with these different mating systems.

Previous studies have hypothesized that hybrid individuals' mating system should have a critical importance in the evolution of the genetic composition of hybrids and in the levels of introgression to be expected in a hybrid zone. PICKUP *et al.* (2019) predicts that in outcrossing hybrids, the many slightly deleterious, co-dominant mutations coming from the self-fertilizing ancestry should be selected against and purged while the highly deleterious, recessive mutations

coming from the outcrossing ancestry should stay hidden in heterozygosity and be maintained. That is, the authors predict a purge of selfing ancestry in outcrossing hybrids. Similarly, they predict that in selfing hybrids, the highly deleterious, recessive mutations coming from the outcrossing ancestry should rapidly become homozygote and be selected against, while the many slightly deleterious, co-dominant mutations should keep on evading selection and be maintained. This time, the outcrossing ancestry is expected to be purged in selfing hybrids. Everything happens as if the mutation load was adapted to a certain reproductive mode and should be maintained in a hybrid that shares the same reproductive mode. This should result in limited introgression between outcrossing and self-fertilizing species, in both directions.

Though appealing, this verbal argument does not appear to be verified in our simulations. Our results emphasize that what matters is merely which parental species has the strongest mutation load. The genome ancestry with the worst fitness will most likely be purged in hybrids, independently of the hybrid mating system itself. Instead of hybrid mating system, we show that mutation and recombination rates are the parameters with the most critical influence on hybrid genetic composition. A few predictions can be made depending on parameters' values:

1) When mutation and recombination rates allow selection to be overall more efficient in outcrossing populations (i.e., are high), selfers' genomes accumulate more fixed deleterious mutations and become of worst quality. When outcrossers' and selfers' genomes are put together in hybrids, selfing ancestry tends to be purged because of their reduced fitness. In this case, we predict that, if hybrids can backcross, little introgression of selfers' genes in outcrossers' genomes, and greater introgression of outcrossers' genes in selfers' genomes.

2) When mutation and recombination rates are not sufficiently high, outcrossers' genomes overall accumulate more mutations (fixed slightly deleterious mutations plus highly deleterious,

recessive mutations at low frequencies) and become of worst quality. Outcrossing ancestry is thus expected to be purged in hybrids: we predict little introgression of outcrossers' genes in selfers' genomes, and greater introgression of selfers' genes in outcrossers' genomes. We argue that this should be the more general case, as overall there is little evidence of putatively deleterious mutation accumulation in self-fertilizing species (BRANDVAIN *et al.* 2014; ARUNKUMAR *et al.* 2015).

Overall, we show that fitness comparison between self-fertilizing and outcrossing species is a good predictor of the long-term genetic composition of their hybrid. We predict that selfers may often have a better fitness; but careful assessment of real-life selfers and outcrossers may say otherwise depending on a diversity of ecological, demographical, and life-history parameters.

KIM *et al.* (2018) produced a similar model of introgression between outcrosser and selfer parental populations. In their simulations, they obtained that genomes from outcrossing populations are of better quality because selfer genomes accumulate fixed deleterious mutations by selective interference, predicting limited introgression of selfer into outcrosser genomes and greater introgression of outcrosser into selfer genomes. A few genomic parameters differ between our work and this paper. First, the authors use a gamma distribution for the selective coefficient of new deleterious mutations whose shape and scale parameters are such that overall selection is much weaker in their study. This may alter the mutation load balance by increasing genetic drift. However, it increases drift in both outcrossers and selfers, and simulations run with $s = 0.01$ (which provides a selective coefficient distribution much closer to the one in KIM *et al.* (2018)) show an acceleration of deleterious mutation accumulation similar in the two parental species (see Supplementary Material Figure S2). KIM *et al.* (2018) also consider a genomic structure based on chromosome 1 of *Arabidopsis thaliana*. This notably means that they consider

a greater number of loci under purifying selection, which is expected to increase the relative importance of selective interference and mutation accumulation. However, simulations with 1,000 and 10,000 instead of 100 loci (Supplementary Material Figure S3) again emphasize that mutation accumulation increases in both selfers and outcrossers; in both cases, outcrossers still accumulate more deleterious mutations.

Focusing on the case of *Arabidopsis thaliana* KIM *et al.* (2018) acknowledge that they look at a scenario with rather high exon density (more non-synonymous mutations) and high recombination rates. As such, they may simply fall in a scenario where selective interference may play a stronger role in selfers (high mutation and recombination rates). How their result is generalizable to other self-fertilizing plant and animals remains to be assessed.

Mutation load is not the only factor that may influence introgression patterns between self-fertilizing and outcrossing species. Behavioral and physiological constraints are also major factors that are known to affect hybridization and introgression directions. In a hybrid zone between species with different mating systems, introgression is expected to happen more frequently from the selfing species to the outcrosser, than vice-versa (KIM *et al.* 2018; PICKUP *et al.* 2019). This expectation relies on the premise that selfing (especially prior selfing (TIAN-BI *et al.* 2008; BRYS *et al.* 2016; BERBEL-FILHO *et al.* 2021)) provides a very limited window of opportunity for outcrossing, by conspecific, heterospecific, or potential male hybrids (PICKUP *et al.* 2019; BERBEL-FILHO *et al.* 2021). This pattern of higher levels of introgression from selfers into outcrossers has been commonly found in plant systems (RUHSAM *et al.* 2011; RUHSAM *et al.* 2013; BRANDVAIN *et al.* 2014). Our findings extend that expectation, but from a different perspective. Our results indicate that often one can expect self-fertilizing species to exhibit lower

mutation load relative to outcrossing species, which would also favor higher levels of introgression from selfers to outcrossers.

We argue that our model provides a "null model" for the direction of introgression in hybrids zones between selfers and outcrossers: after looking at the parental species' fitness, one should be able to predict hybrids' genomic composition. Of course, we acknowledge that this is an "all else being equal" argument, and any deviations from the predictions presented here should stimulate further research that some biological, behavioral, and genetic phenomena may be guiding the major direction of introgression. For instance, the weak inbreeder/strong outbreeder hypothesis (WISO) (BRANDVAIN AND HAIG 2005) states that due to the enhanced opportunity for genomic conflict, outcrossers' gametes are more competitive than selfers' ones. Given equal fertilization opportunities, if hybrids mostly outcross, higher introgression is expected from outcrossers to selfers under the WISO premises. Another factor that can contribute for deviations from the expectations of our model is selfing timing. Selfing can happen before (prior selfing), during (competitive), or after (delayed selfing) the opportunity for outcrossing (LLOYD 1979). The parental selfing time may limit the window of opportunity from an outcrossing hybrid to the selfing parental species, making the major direction of hybridization more common from selfer to outcrosser, than vice-versa (BRYS *et al.* 2016; BERBEL-FILHO *et al.* 2021). Finally, unilateral genetic incompatibilities may bias the direction of introgression (TURELLI AND MOYLE 2007) in both directions, regardless of the hybrids mating system and recombination rates. In our work, we focused solely on the effect of mating systems via mutation loads, and have shown that parental species' loads are good predictors of long-term genomic composition of hybrid species. Altogether, the examples presented above represent interesting avenues of research whenever deviations from this prediction are found.

5. Authors Contribution

FF and WMB-F designed the study. FF coded the model and ran the simulations. FF and WB-F interpreted the results and wrote the manuscript.

6. Acknowledgement

FF and WMB-F are supported by a NSFDEB-NERC Research Grant NE/T009322/1. We thank Christelle Fraisse, along with editors and reviewers, for comments and suggestions that helped us improve the quality of the manuscript.

References

Agrawal, A. F., and M. C. Whitlock, 2011 Inferences About the Distribution of Dominance Drawn From Yeast Gene Knockout Data. Genetics 187**:** 553-566.

Arunkumar, R., R. W. Ness, S. I. Wright and S. C. H. Barrett, 2015 The Evolution of Selfing Is Accompanied by Reduced Efficacy of Selection and Purging of Deleterious Mutations. Genetics 199**:** 817-829.

Berbel-Filho, W. M., A. Tatarenkov, G. Pacheco, H. M. V. Espírito-Santo, M. G. Lira *et al.*, 2021 Against the Odds: Hybrid Zones between Mangrove Killifish Species with Different Mating Systems. Genes 12**:** 1486.

Brandvain, Y., and D. Haig, 2005 Divergent mating systems and parental conflict as a barrier to hybridization in flowering plants. The American Naturalist 166**:** 330-338.

Brandvain, Y., A. M. Kenney, L. Flagel, G. Coop and A. L. Sweigart, 2014 Speciation and introgression between *Mimulus nasutus* and *Mimulus guttatus*. PLoS Genetics 10**:** e1004410.

Brys, R., J. Van Cauwenberghe and H. Jacquemyn, 2016 The importance of autonomous selfing in preventing hybridization in three closely related plant species. Journal of Ecology 104**:** 601-610.

Burgarella, C., P. Gayral, M. Ballenghien, A. Bernard, P. David et al., 2015 Molecular evolution of freshwater snails with contrasting mating systems. Molecular Biology and Evolution 32: 2403-2416.

Busch, J. W., S. Bodbyl□Roels, S. Tusuubira and J. K. Kelly, 2022 Pollinator loss causes rapid adaptive evolution of selfing and dramatically reduces genome□wide genetic variability. Evolution.

Charlesworth, D., M. T. Morgan and B. Charlesworth, 1993 Mutation accumulation in finite outbreeding and inbreeding populations. Genetical Research 61**:** 39-56.

Cutter, A. D., 2019 Reproductive transitions in plants and animals: selfing syndrome, sexual selection and speciation. New Phytol 224**:** 1080-1094.

Escobar, J. S., A. Cenci, J. Bolognini, A. Haudry, S. Laurent et al., 2010 An integrative test of the dead-end hypothesis of selfing evolution in triticeae (Poaceae). Evolution: 2855-2872.

Frankham, R., 2005 Genetics and extinction. Biological conservation 126**:** 131-140.

Gabriel, W., M. Lynch and R. Bürger, 1993 Muller's ratchet and mutational meltdowns. Evolution 47**:** 1744-1757.

Glémin, S., E. Bazin and D. Charlesworth, 2006 Impact of mating systems on patterns of sequence polymorphism in flowering plants. Proceedings of the Royal Society B: Biological Sciences 273: 3011-3019.

Grossenbacher, D. L., Y. Brandvain, J. R. Auld, M. Burd, P. O. Cheptou et al., 2017 Self□compatibility is over□represented on islands. New Phytologist 215: 469-478.

Haller, B. C., and P. W. Messer, 2019 SLiM 3: forward genetic simulations beyond the Wright–Fisher model. Molecular Biology and Evolution 36**:** 632-637.

Hartfield, M., T. Bataillon and S. Glémin, 2017 The evolutionary interplay between adaptation and self-fertilization. Trends in Genetics 33**:** 420-431.

Haudry, A., A. Cenci, C. Guilhaumon, E. Paux, S. Poirier et al., 2008 Mating system and recombination affect molecular evolution in four *Triticeae* species. Genetics Research 90: 97-109.

Hu, X. S., 2015 Mating system as a barrier to gene flow. Evolution 69**:** 1158-1177.

Igic, B., and J. W. Busch, 2013 Is self fertilization an evolutionary dead end? New Phytologist 198: 386-397.

Jarne, P., and J. R. Auld, 2006 Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals Evolution 60: 1816-1824.

Jarne, P., and D. Chalesworth, 1993 The evolution of the selfing rate in functionally hermaphrodite plants and animals. Annual Review of Ecology, Evolution, and Systematics 24: 441-466.

Kim, B. Y., C. D. Huber and K. E. Lohmueller, 2018 Deleterious variation shapes the genomic landscape of introgression. PLos Genetics 14: e1007741.

Lloyd, D. G., 1979 Some reproductive factors affecting the selection of self-fertilization in plants. The American Naturalist 113: 67-79.

Lynch, M., R. Bürger, D. Butcher and W. Gabriel, 1993 The Mutational Meltdown in Asexual Populations. Journal of Heredity 84: 339-344.

Lynch, M., J. Conery and R. Bürger, 1995 Mutational meltdowns in sexual populations. Evolution 49: 1067-1080.

Manna, F., G. Martin and T. Lenormand, 2011 Fitness Landscapes: An Alternative Theory for the Dominance of Mutation. Genetics 189: 923-937.

Martin, N. H., and J. H. Willis, 2007 Divergence associated with mating system causes nearly complete reproductive isolation between sympatric Mimulus species. Evolution 61: 68-82.

Noel, E., Y. Chemtob, T. Janicke, V. Sarda, B. Pelissie et al., 2016 Reduced mate availability leads to evolution of self-fertilization and purging of inbreeding depression in a hermaphrodite. Evolution 70: 625-640.

Ostevik, K. L., J. L. Rifkin, H. Xia and M. D. Rausher, 2021 Morning glory species co occurrence is associated with asymmetrically decreased and cascading reproductive isolation. Evolution Letters 5**:** 75-85.

Pickup, M., Y. Brandvain, C. Fraisse, S. Yakimowski, N. H. Barton *et al.*, 2019 Mating system variation in hybrid zones: facilitation, barriers and asymmetries to gene flow. New Phytologist 224**:** 1035-1047.

Poon, A., and S. P. Otto, 2000 Compensating for our load of mutations: freezing the meltdown of small populations. Evolution 54**:** 1467-1479.

Rand, D. M., 2008 Mitigating Mutational Meltdown in Mammalian Mitochondria. PLoS Biology 6**:** e35.

Ruhsam, M., P. Hollingsworth and R. Ennos, 2011 Early evolution in a hybrid swarm between outcrossing and selfing lineages in Geum. Heredity 107**:** 246-255.

Ruhsam, M., P. M. Hollingsworth and R. A. Ennos, 2013 Patterns of mating, generation of diversity, and fitness of offspring in a Geum hybrid swarm. Evolution 67**:** 2728-2740.

Sianta, S. A., S. Peischl, D. A. Moeller and Y. Brandvain, 2022 The efficacy of selection may increase or decrease with selfing depending upon the recombination environment. Evolution.

Shimizu, K. K., and T. Tsuchimatsu, 2015 Evolution of selfing: recurrent patterns in molecular adaptation. Annual Review of Ecology, Evolution, and Systematics 46**:** 593-622.

Slotte, T., J. P. Foxe, K. M. Hazzouri and S. I. Wright, 2010 Genome-wide evidence for efficient positive and purifying selection in *Capsella grandiflora*, a plant species with a large effective population size. Molecular Biology and Evolution 27: 1813-1821.

Takebayashi, N., and P. L. Morrell, 2001 Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. American Journal of Botany 88**:** 1143-1150.

The Tree of Sex Consortium, 2014 Tree of Sex: A database of sexual systems. Scientific Data 1.

Theologidis, I., I. M. Chelo, C. Goy and H. Teotónio, 2014 Reproductive assurance drives transitions to self-fertilization in experimental *Caenorhabditis elegans*. BMC Biology 12: 1-21.

Tian-Bi, Y.-N. T., E. K. N'GORAN, S.-P. N'GUETTA, B. Matthys, A. Sangare *et al.*, 2008 Prior selfing and the selfing syndrome in animals: an experimental approach in the freshwater snail Biomphalaria pfeifferi. Genetics research 90**:** 61-72.

Turelli, M., and L. C. Moyle, 2007 Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. Genetics 176**:** 1059-1088.

Wang, J., W. G. Hill, D. Charlesworth and B. Charlesworth, 1999 Dynamics of inbreeding depression due to deleterious mutations in small populations: mutation parameters and inbreeding rate. Genetical Research 74**:** 165-178.

Wang, X.-J., S. C. H. Barrett, L. Zhong, Z.-K. Wu, D.-Z. Li et al., 2021 The genomic selfing syndrome accompanies the evolutionary breakdown of heterostyly. Molecular Biology and Evolution 38: 168-180.

Wright, S. I., S. Kalisz and T. Slotte, 2013 Evolutionary consequences of self-fertilization in plants. Proceedings of the Royal Society B: Biological Sciences 280**:** 20130133.