Mating system and speciation I: accumulation of genetic incompatibilities in allopatry

Lucas Marie-Orleach[1,2*], Christian Brochmann[1], Sylvain Glémin[2,3],

1 Natural History Museum, University of Oslo, NO-0318 Oslo, Norway
2 CNRS, Université Rennes 1, ECOBIO (écosystèmes, biodiversité, évolution) – UMR 6553, F-35042 Rennes, France
3 Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Norbyvägen 14-18, SE-752 36 Uppsala, Sweden

* lucas.marie-orleach@univ-rennes1.fr

Abstract

Self-fertilisation is widespread among hermaphroditic species across the tree of life. Selfing has many consequences on the genetic diversity and the evolutionary dynamics of populations, which may in turn affect macroevolutionary processes such as speciation. On the one hand, because selfing increases genetic drift and reduces migration rate among populations, selfing may be expected to promote speciation. On the other hand, because selfing reduces the efficacy of selection, selfing may be expected to hamper ecological speciation. To better understand under which conditions and in which direction selfing affects the build-up of reproductive isolation, an explicit population genetics model is required. Here, we focus on the interplay between genetic drift, selection and genetic linkage by studying speciation without gene flow. We test how fast populations with different rates of selfing accumulate mutations leading to genetic incompatibilities. When speciation requires the population to pass through a fitness valley caused by underdominant and compensatory mutations, selfing reduces the depth and/or breadth of the valley, and thus overall facilitates the fixation of incompatibilities. When speciation does not require the population to pass through a fitness valley, as for Bateson-Dobzhanzky-Muller incompatibilities (BDMi), the lower effective population size and higher genetic linkage in selfing populations facilitates the fixation of incompatibilities. Interestingly, and contrary to intuitive expectations, local selection does not always accelerate the build-up of reproductive isolation in outcrossing relative to selfing populations. Our work helps to clarify how selfing lineages may speciate and diversify over time, and emphasizes the need to account for interactions among segregating mutations within populations to better understand macroevolutionary dynamics.

Author summary

Hermaphroditic organisms may use their male gametes to fertilise their own female gametes, and species vary greatly in how much they self-fertilise. Self-fertilisation induces many genetic modifications in the population, which may ultimately affect the rates at which lineages diversify. Here we aim to build predictions on how self-fertilisation affects the rate at which reproductive isolation arises between geographically isolated populations. Specifically, we develop theoretical models in which populations varying in their rates of self-fertilisation may fixate mutations leading to reproductive isolation.

We first explored scenarios in which reproductive isolation is made by mutations whose fixations necessitate the population to experience temporally deleterious effects (i.e., a fitness valley), and found that self-fertilisation reduces the breadth and depth of the fitness valley and thereby overall facilitates the accumulation of such mutations. Second, we explored scenarios in which genetic incompatibilities are caused by interactions between derived alleles of different genes (i.e., BDMi). By allowing the BDMi to occur within populations, we found that self-fertilisation reduces the manifestation of BDMi within population, and thereby facilitates their fixation. This effect prevails even in the face of local adaptation. Thus, our study clarifies how fast species are expected to arise in self-fertilisation lineages.

Introduction

> *Species belonging to the same section or species group usually cross freely in woody plants and perennial herbs, but are usually separated by incompatibility barriers in annual herbs. [...] It is possible, therefore, that the correlation [...] between sterility and life form is a reflection of a more fundamental relationship between type of breeding system and the formation of sterility barriers"* [1].

The wide variety of mating systems observed in animals and plants, and also fungi and algae, has multiple ecological and evolutionary consequences that might impact higher-level evolutionary processes such as species extinction and speciation. For instance, hermaphroditic species vary in their rate of self-fertilisation – spanning from obligate outcrossing to predominant selfing species, and including all degrees of mixed mating [2–4] – which has long been argued to affect macroevolutionary processes [5–9]. Because selfing tends to reduce both the genetic diversity and the population adaptive potential, selfing lineages have been argued to be 'evolutionary dead-ends' as they are expected to go extinct at faster rates than outcrossing lineages [5,8,10,11]. The study of the macroevolutionary effects of selfing has however mostly focused on species extinction, while the effects of selfing on speciation has received relatively less attention, both empirically and conceptually (but see [7,9]).

The effects of selfing on speciation have been studied based on phylogenies. Phylogenetic trees with variation in mating system may allow us to estimate and compare the rates of species diversification, and potentially speciation and extinction, in selfing *vs.* outcrossing lineages (*e.g.*, [12–15]). For instance, in the Solanaceae plant family, outcrossing is enforced by a self-incompatibility mechanism that has broken down several times, leading to multiple independent self-compatible lineages. Compared to the self-incompatible lineages (*i.e.*, obligate outcrossers), the self-compatible ones (*i.e.*, potential selfers) show lower diversification rates [16] which, interestingly, are likely to be caused by higher rates of both speciation and species extinction in the selfing lineages [12]. Other phylogenetic studies carried out in the Primulaceae [13] and in the Onagraceae [15] plant families suggest that young selfing taxa experience a burst of speciation that fade away with time (*i.e.*, a 'senescing diversification rate' [17]). In contrast, mixed evidence are reported in the Polemoniaceae plant family, in which alternative phylogenetic methods provide positive or no associations between selfing and speciation rates [14].

The effects of selfing on speciation have also to some degree been studied based on experimental crosses, addressing whether reproductive isolation (RI) between populations evolves at different rates in selfing *vs.* outcrossing species. To our knowledge however, there are only a few of such studies. For instance, in the Arctic flora, intraspecific crosses between geographically isolated populations of eight predominantly selfing species resulted in F1 hybrids with low pollen fertility and seed set, whereas no reduced fertility was observed in the single outcrossing species for which successful crosses could be made [18,19]. In the selfing species, it was estimated that RI may have developed over

just a few millennia. The experimental crosses in one of these Arctic species, *Draba nivalis*, were also used to address the genetic architecture of RI. Quantitative trait loci analyses of F2 populations of this predominantly selfing species showed that the post-zygotic incompatibilities are due to single-locus underdominance, a putative chromosomal translocation, and nuclear-nuclear and cyto-nuclear epistatic incompatibilities [20,21].

Theoretical expectations on the effects of selfing on speciation are poorly studied, and not straightforward. This is because selfing impacts several interconnected population genetics parameters that have opposite effects on speciation. First, selfing decreases gene flow within and among populations, enhancing the isolation of populations and thus possibly facilitating speciation [7]. Second, the non-random sampling of gametes used for reproduction by selfing individuals reduces the effective population size. For instance, the effective population size is expected to be halved in purely selfing species compared to a randomly mating outcrossing species of the same population size [22,23]. A reduction of effective population size has cascading effects. It elevates genetic drift, reduces genetic polymorphism, and overall weakens selection. Third, selfing increases homozygosity, which makes recombination less efficient because homologous chromosomes tend to be identical [23]. Thus, recombination breaks down linkage disequilibrium less efficiently in strongly selfing populations, thereby reducing the evolutionary advantages of recombination [24], and overexposing the populations to the deleterious effects of linked selection, such as background selection [25], further reducing effective population size in selfing populations [26,27].

Here, we developed analytical and simulation models of population genetics to better understand the effects of selfing on speciation. We did not consider the effects of gene flow, but focused on the interplay between genetic drift, genetic linkage, and selection efficacy. We studied how mutations leading to RI accumulate within populations differing in selfing rates, and asked if selfing affects (i) the pace of speciation, and (ii) the genetic architecture of RI.

We sequentially explored three types of mutations: underdominant mutations, compensatory mutations, and Bateson-Dobzhansky-Muller incompatibility mutations. Underdominant mutations have deleterious effects in the heterozygous state, but have no deleterious effects in either homozygous state (which may for instance be due to structural variants [28]). Compensatory mutations are a pair of mutations that are both deleterious when they occur alone in a genome, but are neutral when they occur together [29] (*e.g.*, compensatory evolution of *cis*- and *trans*-regulation of gene expression [30–32]). Finally, Bateson-Dobzhansky-Muller incompatibility (BDMi) mutations are a pair of mutations that have no deleterious effects when they occur alone in a genome, but cause genetic incompatibilities when they occur together [33–36]. Importantly, the fixation of underdominant and compensatory mutations requires the population to pass through a fitness valley in which mutations may be counter selected. In contrast, the fixation of BDMi mutations may be neutral, or even positively selected when the mutations are advantageous [37]. Because homozygotes are formed more readily in selfing species, selfing has previously been shown to facilitate the fixation of underdominant mutations [38]. It is however unknown if and how selfing modulate the accumulation of mutations with epistatic effects, such as compensatory and BDMi mutations.

Overall, we hypothesised that the effects of selfing on speciation depend on the mode of speciation. If RI evolves through genetic drift, selfing should promote speciation because underdominant and compensatory mutations are more likely to get fixed through genetic drift in selfing lineages. In contrast, if RI evolves as a by-product of selection (*e.g.*, ecological speciation), we expect genetic incompatibilities to arise more readily in outcrossing lineages because selfing populations have an overall lower efficacy of selection.

Methods

So far, the accumulation of BDMi in allopatry has mostly been modelled as a combinatorial process of substitutions, each predicted by single-locus theory [36, 37]. Extending such results to selfing populations would be straightforward but partly misleading as these models do not explicitly consider the underlying multi-loci population genetics dynamics and the possible interactions among alleles that can be affected by selfing. Instead, we studied the effects of selfing on speciation by modelling – in a single population – the fate of different types of mutation that create genetic incompatibilities among populations. Especially, we determined the probability of and the time to fixation of the different categories of incompatibilities. The fixation of a single incompatibility mutation is, in most cases, not sufficient to complete speciation but determines the overall pace at which RI builds. We first studied the dynamics of a single (pair of) incompatibility and then consider multi-loci models where mutations can occur recurrently throughout the genome.

For all models, we considered a population of N hermaphroditic individuals reproducing by selfing at rate $0 \leq \sigma \leq 1$. The effective size of partially selfing populations is given by Pollak [22]:

$$N_e = \frac{N}{1+F} \tag{1}$$

where F, the Wright's fixation index, is:

$$F = \frac{\sigma}{2-\sigma} \tag{2}$$

Note that the effective population size, N_e, can be further reduced in selfing populations due to background selection, which we included in the single-locus and two-loci simulations (see Simulations). For multi-loci models, we assumed a recombination rate r between adjacent loci and an equal mutation rate, μ, for all loci. We first analysed the single-locus and two-loci models to characterize the underlying mechanisms. In particular, we focused on the mean time to fix the first incompatibility allele or haplotype under recurrent mutations, which can be decomposed into the mean waiting time of occurrence of the first mutation destined to be fixed (T_{wait}) and the mean fixation time conditioned on fixation (T_{fix}):

$$T = T_{wait} + T_{fix} \tag{3}$$

Then we performed multi-loci simulations to assess how the results scale up the genome scale.

Single-locus incompatibility

Underdominant mutations

This model has already been studied by Charlesworth [38] but we summarized it for completeness and provided additional results. We considered a single bi-allelic locus, with the ancestral allele A_1 that can mutate to the derived allele A_2 at rate μ. The fitness of genotypes A_1A_1, A_1A_2, and A_2A_2 are 1, $1-s_u$, and $1+s$, respectively, and the frequency of allele A_2 is noted x.

The change in allele frequencies in one generation is given by:

$$\begin{aligned}\Delta x &= x(1-x)\big((1-F)(2(2s_u+s)x-2s_u)+Fs\big)/\overline{W} \\ &\approx x(1-x)\big((1-F)(2(2s_u+s)x-2s_u)+Fs\big)\end{aligned} \tag{4}$$

Table 1. Glossary of the main notations

Symbol	Biological meaning
N	Population size
N_e	Effective population size
σ	Selfing rate
F	Wright's fixation index
μ	Mutation rate
r	Recombination rate
L	Genome length (used in multi-loci simulation models only)
h	Coefficient of dominance (besides genetic incompatibilities)
s	Strength of selection (besides genetic incompatibilities)
h_c, h_B	Coefficient of dominance of the compensatory and BDMi mutations
s_u, s_c, s_B	Strength of selection of the underdominant, compensatory and BDMi mutations
k_c, k_B	Coefficient of dominance in double heterozygotes for compensatory and BDMi mutations

where $\overline{W}$ is the mean fitness of the population, approximately equal to 1 when selection is weak. Equation (4) can also be written as:

$$\begin{aligned}\Delta x &\approx 2(1-F)(2s_u+s)x(1-x)(x-x_{eq}) \quad \textbf{if } F<1\\ &\approx sx(1-x) \quad \textbf{if } F=1\end{aligned} \tag{5}$$

with

$$x_{eq} = \frac{2s_u - F(2s_u+s)}{2(1-F)(2s_u+s)}$$

According to diffusion theory, the probability of fixation of a single A_2 mutant is given by:

$$P_{fix} = \frac{\int_0^{1/2N} \exp(-\frac{2M_{\delta x}}{V_{\delta x}})dx}{\int_0^1 \exp(-\frac{2M_{\delta x}}{V_{\delta x}})dx} \tag{6}$$

where $M_{\delta x} = \Delta x$ is the expected infinitesimal change in allele frequency, and $V_{\delta x} = \frac{x(1-x)}{2N_e}$ is the expected infinitesimal variance.

Two-loci incompatibilities

We also considered models with two bi-allelic loci. A_1 and A_2, B_1 and B_2 denote the ancestral and derived alleles at the first and second locus, respectively, and their respective frequencies are x and $1-x$, y and $1-y$. We assumed the same mutation rate, μ, at the two loci. The recombination rate between the two loci is $0 \leq r \leq 0.5$, where 0 corresponds to fully linked loci and 0.5 corresponds to loci located on different chromosomes with Mendelian segregation. The frequency of the four haplotypes $\{A_1B_1, A_1B_2, A_2B_1, A_2B_2\}$ are noted as $\{X_1, X_2, X_3, X_4\}$, and the frequency of the ten genotypes, the combination of haplotypes X_i and X_j, are noted as G_{ij} with $i \leq j$ (for example $G_{12} = [A_1A_1; B_1B_2]$). Note that we must distinguish G_{14} from G_{23}, which correspond to identical genotypes $[A_1A_2; B_1B_2]$ but may differ in the haplotypes produced through gametogenesis. Changes in genotype frequencies can be obtained as follows [38]:

After meiosis, haplotype frequencies are given by:

$$X'_1 = G_{11} + \frac{1}{2}(G_{12} + G_{13} + (1-r)G_{14} + rG_{23})$$
$$X'_2 = G_{22} + \frac{1}{2}(G_{12} + G_{24} + (1-r)G_{23} + rG_{14})$$
$$X'_3 = G_{33} + \frac{1}{2}(G_{13} + G_{34} + (1-r)G_{23} + rG_{14})$$
$$X'_4 = G_{44} + \frac{1}{2}(G_{24} + G_{34} + (1-r)G_{14} + rG_{23})$$

After syngamy, genotypic frequencies are given by:

$$G'_{11} = (1-\sigma)X'^2_1 + \sigma\big(G_{11} + \frac{1}{4}(G_{12} + G_{13} + r^2G_{23} + (1-r)^2G_{14})\big)$$
$$G'_{22} = (1-\sigma)X'^2_2 + \sigma\big(G_{22} + \frac{1}{4}(G_{12} + G_{24} + r^2G_{14} + (1-r)^2G_{23})\big)$$
$$G'_{33} = (1-\sigma)X'^2_3 + \sigma\big(G_{33} + \frac{1}{4}(G_{13} + G_{34} + r^2G_{14} + (1-r)^2G_{23})\big)$$
$$G'_{44} = (1-\sigma)X'^2_4 + \sigma\big(G_{44} + \frac{1}{4}(G_{24} + G_{34} + r^2G_{23} + (1-r)^2G_{14})\big)$$
$$G'_{14} = (1-\sigma)2X'_1X'_4 + \sigma\frac{1}{2}(r^2G_{23} + (1-r)^2G_{14})$$
$$G'_{23} = (1-\sigma)2X'_2X'_3 + \sigma\frac{1}{2}(r^2G_{14} + (1-r)^2G_{23})$$
$$G'_{ij} = (1-\sigma)2X'_2X'_3 + \sigma\frac{1}{2}(G_{ij} + r(1-r)(G_{14} + G_{23})) \text{ for other cases}$$

Finally, after selection:

$$G^{sel}_{ij} = \frac{w_{ij}G'_{ij}}{\sum_{i\leq j} w_{ij}G'_{ij}} \tag{7}$$

where w_{ij} is the fitness of genotype G_{ij}, which depends on the type of two-loci incompatibility studied (see Compensatory mutations and BDMi mutations below). Note that in either scenarios, we do not distinguish *cis* and *trans* effects on fitness (*i.e.*, $w_{14} = w_{23}$).

Compensatory mutations

We extended previous models of compensatory mutations [39, 40] by including the effects of partial selfing. In brief, compensatory mutations at two loci can be viewed as a generalization of the one-locus underdominant model presented above where two haplotypes are equally fit, A_1B_1 and A_2B_2, but the intermediate paths, A_1B_2 and A_2B_1 are deleterious. Thus, alike underdominant mutations, the evolution of pairs of compensatory mutations requires to cross a fitness valley.

In the compensatory mutation models, we set the fitness of genotypes as follows:

$$w_{11} = w_{44} = 1$$
$$w_{22} = w_{33} = 1 - s_c$$
$$w_{12} = w_{13} = w_{24} = w_{34} = 1 - h_c s_c$$
$$w_{14} = w_{23} = 1 - h_c k_c s_c$$

where $s_c \geq 0$ and $0 \leq h_c \leq 1$ are the strength and the coefficient of dominance of the deleterious effects of each mutation respectively, and k_c is the coefficient of dominance for the double heterozygous genotype $A_1A_2B_1B_2$.

BDMi mutations

We extended the model of Kimura and King [41] by including the effects of partial selfing. BDMi mutations generate genetic incompatibility only when an individual carries both derived alleles. Thus, individuals carrying either derived alleles do not experience deleterious effects, meaning that the fixation of BDMi mutations does not require to cross a fitness valley.

In the BDMi mutations models, we set the fitness of the genotypes as follows:

$$\begin{aligned} w_{11} &= 1 \\ w_{12} = w_{13} &= 1 + hs \\ w_{22} = w_{33} &= 1 + s \\ w_{44} &= (1+s)(1+s)(1-s_B) \\ w_{24} = w_{34} &= (1+hs)(1+s)(1-h_B s_B) \\ w_{14} = w_{23} &= (1+hs)(1+hs)(1-h_B k_B s_B) \end{aligned}$$

where $s_B \geq 0$ and $0 \leq h_B \leq 1$ are the strength and the coefficient of dominance of the genetic incompatibility between the two derived alleles A_2 and B_2 respectively, and k_B is the coefficient of dominance for the double heterozygous genotype $A_1A_2B_1B_2$. $s \geq 0$ and $0 \leq h \leq 1$ are the strength of selection and the coefficient of dominance of the local adaptation. For the standard BDMi model s is set to 0, and $s > 0$ only when we assume that BDMi mutations are also driven by local adaptation.

Simulations

We also developed, first, single-locus (underdominance mutations) and two-loci (compensatory and BDMi mutations) simulation models to check for analytical and numerical predictions and, second, multi-loci simulation models (underdominant, compensatory, and BDMi mutations) to test for the possible effect of interactions between segregating mutations.

Single-locus and two-loci simulations

We developed C++ programs to simulate populations of N hermaphroditic individuals producing gametes with mutations from ancestral to derived alleles at a rate μ and, for the two-loci simulations, recombinations between the two loci at a rate $0 \leq r \leq 0.5$. Syngamy may occur through self-fertilisation, at a rate $0 \leq \sigma \leq 1$. Then, selection on offspring occurs differently in the underdominant, compensatory, and BDMi mutations simulations (as described above). Finally, genetic drift was included by sampling offspring from the genotype frequencies using a multinomial distribution using the *gsl_ran_multinomial* function from the GNU Scientific Library [42]. For each iteration, we measured the number of generations required to fixate the derived allele (underdominant mutation), a pair of derived alleles (compensatory mutations), and either derived allele (BDMi mutations).

Importantly, selfing is known to increase the strength of background selection, which further reduces effective population size [26, 27, 43]. The strength of this effect critically depends on genomic recombination rate. When the genomic recombination rate is high, only highly selfing populations suffer from background selection. In contrast, when the genomic recombination rate is low, the effect of background selection rise linearly with selfing [26, 27]. To account for this effect, we used a Dirichlet-multinomial distribution in the genetic drift function (instead of a multinomial distribution) that allowed us to tailor the effective size of the population to its selfing rate. For this we used analytical approximations in Roze [27] to simulate two background selection scenarios

corresponding to two levels of genomic recombination. The first scenario corresponds to a low rate of genomic recombination and leads to a linear decrease in N_e with selfing rate (hereafter 'linear BG effects'). The second scenario corresponds to a high rate of genomic recombination and leads to an accelerating decrease in N_e with selfing rate (hereafter 'curved BG effects'). Specifically, using the BS1 function from the supplemental material of Roze [27] (File S2), we modelled the background selection effects assuming that deleterious alleles with selection and dominance coefficients of $s = 0.05$ and $h = 0.2$ occurred with a genomic mutation rate of $U = 0.1$ in a genome map length of either $R = 0.5$ ('linear BG effects') or $R = 40$ ('curved BG effects'). This informed us by how much N_e was reduced due to background selection for different selfing rates. Then, we used these background selection effects on N_e to compute the parameter vector α of the Dirichlet-multinomial distribution given that:

$$\frac{N_e}{N} = \frac{1+\alpha}{N+\alpha} \tag{8}$$

Multi-loci simulations

We used individual based forward simulations in which the individuals have diploid genomes of length L elements representing loci on which mutations can occur at rate μ and between which recombination can occur at rate r. As above, we modelled a single population of N individuals reproducing through selfing at rate $0 \leq \sigma \leq 1$ and on which selection depends on the type of genetic incompatibility considered (see above).

The single-locus and two-loci simulation models were performed using C++ scripts, and the multi-loci simulations models were performed on the software SLiM [44], both using GNU parallel [45] (see Code availability).

Results

Underdominant mutations

When underdominant, a mutation arising in a population is first counter-selected and it is well known that either genetic drift or selfing facilitates crossing the fitness valley [38]. For completeness we first summarize previous results (Fig. S1, Appendix S1). When selfing rate increases, heterozygous individuals become rarer in the population and so does the selection against underdominant mutations, increasing the probability of fixation and decreasing the time to fixation. When $s > 0$, a new underdominant mutation arising in a population may be directly positively selected if selfing rate is high enough, that is if:

$$\sigma > \sigma_{lim} = \frac{2s_u}{s + 2s_u} \tag{9}$$

Including the effects of background selection slightly shortens the time to fixation of symmetrical underdominant mutations ($s = 0$), but substantially increases the time to fixation of asymmetrical mutations (> 0) in highly selfing populations (Fig. S1 E). This is because background selection reduces the effective population size, N_e, and thereby the efficacy of selection. Such a low efficacy of selection virtually reduces both the fitness valley in the heterozygotes, A_1A_2 and the fitness peak in the homozygotes, A_2A_2. But, because background selection mostly impact highly selfing populations – in which selection on heterozygotes is irrelevant – the effects of background selection mostly manifests as a lower probability of fixation in highly selfing populations because the effect of positive selection on the homozygotes is reduced.

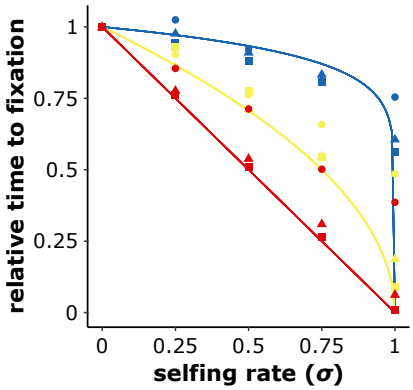


Fig 1. Underdominant mutations accumulate more rapidly in selfing populations. Analytical predictions (lines) and outcomes of multi-loci simulations (symbols) of the averaged number of generations required to fixate a symmetrical underdominant mutation compared to an outcrossing population ($\sigma = 0$). $4Ns_u$ follows a Gamma distribution with a shape parameter β of 0.1 (green), 0.5 (blue), or 1 (red), and mean parameter γ of 10 (circles), 100 (triangles), or 1,000 (squares). $L = 100$, $N = 1,000$, $\mu = 10^{-6}$, $r = 0.01$. The analytical predictions are $(1-\sigma)^{\beta}$. $1,000$ *iterations*.

Previous studies have only considered a single mutation with constant effect [38]. Interesting properties can be obtained if we assume a distribution of deleterious effect. Assuming that $4Ns_u$ follows a Gamma distribution with mean γ and shape β, the proportion of underdominant mutations that can fixate in partially selfing populations compared to outcrossing populations is well approximated by:

$$(1-\sigma)^{-\beta} \tag{10}$$

which is independent of γ. Thus, because the time to fixation is mostly determined by the waiting time that a mutation that will get fixed arises in a population, the relative time to fixation in partially selfing compared to outcrossing populations is simply given by:

$$(1-\sigma)^{\beta} \tag{11}$$

which our multi-loci simulation results confirm as long as γ is not too small (Fig. 1). It shows that when $\beta < 1$ there is a non-linear accelerating effect of selfing on accumulation of underdominant mutations and the lower β the stronger the effect.

Compensatory mutations

Compensatory mutations can also contribute to RI, and their fixation also requires crossing a fitness valley. How selfing affects the fixation of such mutations depends first on the strength of the deleterious effect relative to drift ($N_e s_c$ of the order of 1 or lower) (Fig. 2). When it is low, fixation occurs in two steps. First, a primary mutation goes to fixation as a weakly deleterious mutation and, second, the compensatory mutation restoring fitness goes to fixation as a weakly beneficial mutation. As the total time mainly depends on the waiting time for mutations that will be fixed arise, it can be approximated by:

$$T_{weak} \approx \frac{1}{4Nsu}\left(\frac{e^{4N\tilde{h}_c s_c}-1}{2\tilde{h}_c} + \frac{1-e^{-4N(1-\tilde{h}_c)s_c}}{1-\tilde{h}_c}\right) \tag{12}$$

where $\tilde{h}_c = \frac{h_c(1-F)+F}{1+F}$, which corresponds to the effective dominance level $(h_c(1-F)+F)$ scaled by the increase in drift due to selfing $(1+F)$. The factor 2 in the first step corresponds to the two potential initial paths.

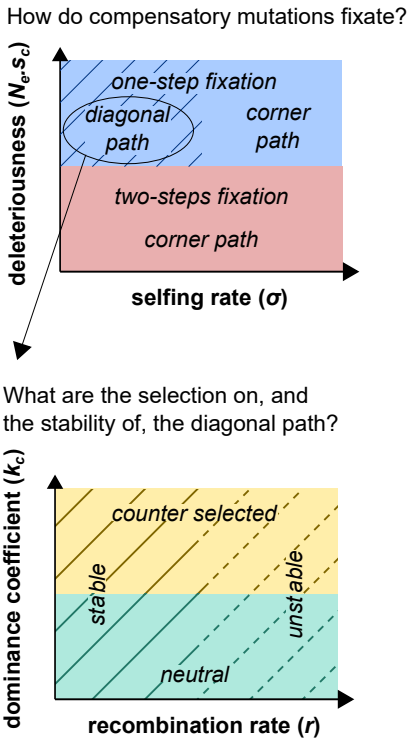


Fig 2. Fixation regime of compensatory mutations as a function of the selfing rate (σ), the strength of deleterious effect (deleteriousness), the recombination rate between the two loci (r), and the coefficient of dominance in double heterozygotes (k_c). Populations with high selfing rates always take the corner path on the fitness landscape (*i.e.*, the deepest point of the fitness valley represented by genotypes that are homozygous for the ancestral allele at one locus, and homozygous for the derived allele at the other locus), while populations with low selfing rates can take the diagonal path (*i.e.*, double heterozygous genotype). The diagonal path may however be counter selected (when k_c is relatively high) and unstable (when the recombination rate is relatively high). Note that, for the sake of simplicity, all effects are here presented as qualitative categories while they are in fact gradual.

Selfing may affect both steps but, depending on the coefficient of dominance (h), fixation is either faster or slower in selfing populations than in outcrossing population because highly selfing populations are less affected by the selection at the heterozygous state (Fig. S2). For instance, recessive deleterious mutations (*e.g.*, $h_c = 0$), do not experience fitness costs in heterozygotes, which facilitates the fixation of both the primary deleterious and the secondary beneficial mutations in outcrossing populations [46]. In contrast, dominant deleterious mutations (*e.g.*, $h_c = 1$), do experience fitness costs in heterozygotes, which both hamper the fixation of the primary deleterious mutation and hide the beneficial effects of the second one in outcrossing populations. Overall, selfing speeds up the fixation of compensatory mutations if the mutation is dominant ($h_c > 1/2$), and slows it down if the mutation is recessive ($h_c < 1/2$), and has no effects when mutation is codominant ($h_c = 1/2$ hence $\tilde{h}_c = 1/2$, independent of F) (Fig. S2).

In contrast, when the deleterious effect is too high ($N_e s_c >> 1$) to allow the fixation of a singly mutation, then the two compensatory mutations must segregate together in the population and fixate together. In this case, the key parameters are the recombination rate between loci (r), and the coefficient of dominance in the double heterozygote (k_c) (Fig. 2). When $r = 0$ and $k_c = 0$, we found that the fixation time of a pair of

compensatory mutations ($T_{0,0}$) may be approximated by (see Appendix S1):

$$T_{0,0} \approx \frac{(h_c(1-F)+F)s_c}{2\mu^2} \tag{13}$$

which shows that selfing always increases the fixation time (Fig. 3B and Fig. S3). This is because evolution can follow the diagonal path of the fitness landscape (*i.e.*, through the double heterozygous genotype $A_1A_2B_1B_2$), and thus avoid the corner path made of genotypes with low fitness value (*i.e.*, $A_1A_1B_2B_2$ or $A_2A_2B_1B_1$), more easily in outcrossing than in selfing populations (Fig. 2, Fig. 3, Fig. S4). In comparison to the corner path, the diagonal path represents a lower fitness valley, which flattens out as k approaches 0. Thus, when r and k_c are both equal to 0, outcrossing populations fixate compensatory mutations by forming the double mutated haplotype (A_2B_2), which is stable because it is not broken down by recombination and can spread in the population because it is less counter selected.

When $k_c > 0$ and/or $r > 0$, the dynamics is very different. Either selection against double heterozygotes or recombination that breaks down double-mutated haplotypes considerably reduces the probability of crossing the fitness valley. When r is small, we can show that:

$$T_{r,k} \approx T_{0,0}\frac{\sqrt{\pi}e^R Erf(\sqrt{R})}{2\sqrt{R}} \tag{14}$$

where Erf is the error function and $R = N(1-\sigma)(r+hks)$. The right-hand function increases very rapidly with R. These approximations and our simulations show that both recombination and selection against heterozygotes strongly impedes the fixation of compensatory mutations, both effects being reduced by selfing (Fig. 3B and Fig. S3). When $k_c > 0$, the diagonal path represents a fitness valley that selects against the double mutated haplotype (A_2B_2). And, when $r > 0$, the double mutated haplotype (A_2B_2) may break, forming genotypes that are off the diagonal path and eliminating the derived alleles (A_2 and B_2) (Fig. 2). In contrast, the fixation time in highly selfing populations is less influenced by r and k_c because their high level of homozygosity (i) makes selfing populations not following the diagonal path of the fitness landscape, and (ii) makes the double mutated haplotype (A_2B_2) more stable over time due to the inefficient recombination in selfers [23].

We further found that, regardless of the coefficient of dominance in the double heterozygote (k_c) and the recombination rate (r), background selection shortens fixation time in highly selfing populations (Fig. S5). By increasing drift, background selection reduces the efficacy of selection against the primary deleterious mutation, which may thus segregate at higher frequency and be more likely to be associated with the secondary compensatory mutation. Moreover, because of background selection, N_es can be small enough under selfing for the fixation to occur in two steps (which is rapid see (13)) whereas, for the same population size, N_es can be too high under outcrossing so that fixation can occur in a single step (which is much longer (14)).

The outcome of our two-loci simulations concur with the multi-loci simulations. Specifically, our multi-loci simulations indicate that the fixation of two compensatory mutations depends on the coefficient of dominance in the double heterozygote (k_c), and the recombination rate (r). Selfing speeds up the fixation of compensatory mutations, except when the fitness of the double heterozygote is neutral (k_c=0) and the recombination rate is low enough ($r < 0.0001$) for the haplotype carrying both derived alleles to be stable (Fig. S6).

BDMi mutations

Contrary to our models of underdominant and compensatory mutations, the BDMi model does not require crossing a fitness valley for RI to accumulate. Therefore, there is

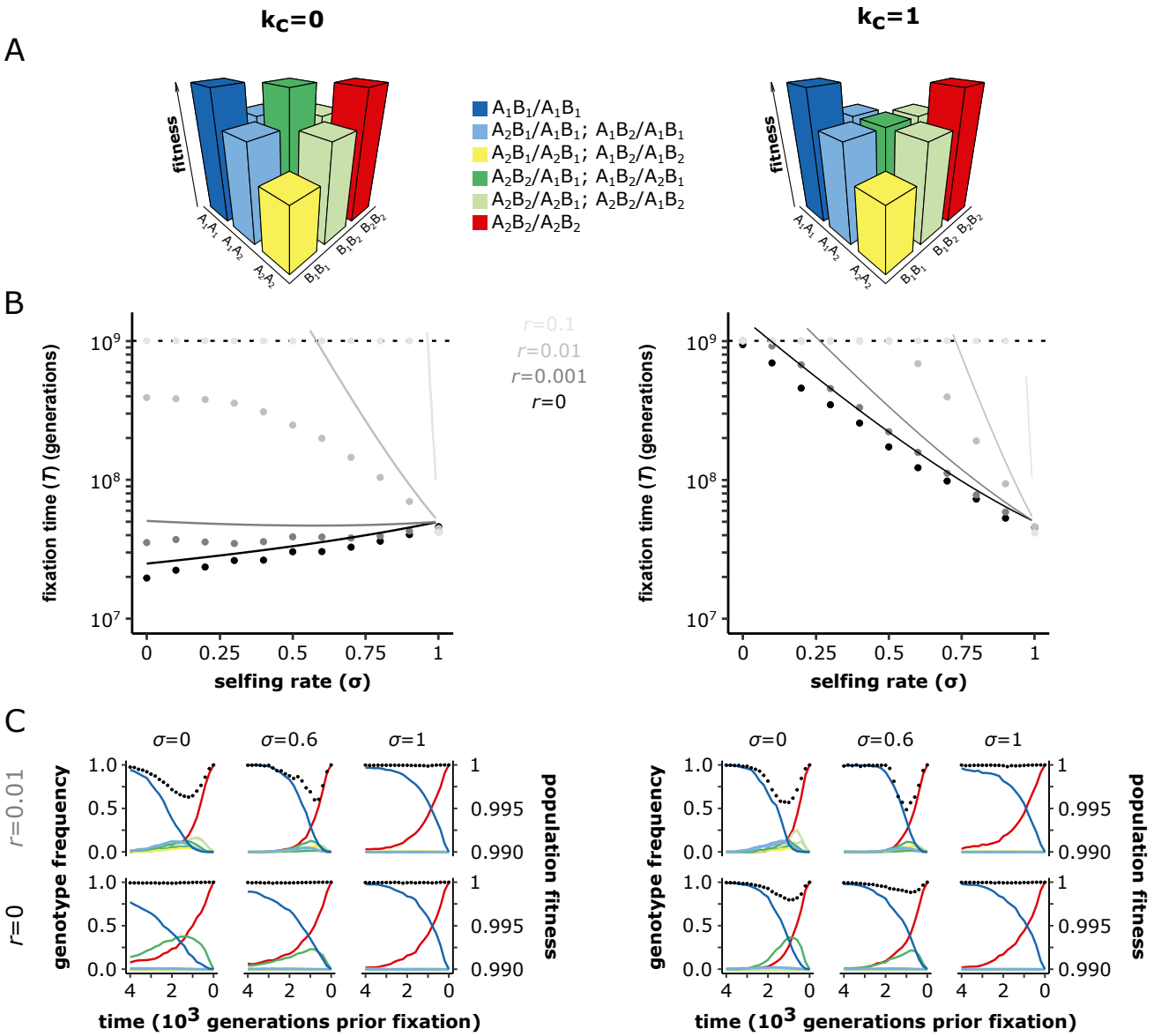


Fig 3. The effect of selfing on the fixation of compensatory mutations depends on the coefficient of dominance in the double heterozygote (k_c), and the recombination rate(r). (A) Schematic representations of the fitness landscapes for compensatory mutations with $k_c = 0$ (left) or $k_c = 1$ (right). A_1 and B_1 are the ancestral alleles. A_2 and B_2 are the derived alleles. (B) Time to fixation of a pair compensatory mutations with $k_c = 0$ (left) or $k_c = 1$ (right). The lines correspond to the analytical approximations (eq. 14) for $r = 0$ (black), $r = 0.001$ (dark grey), $r = 0.01$ (grey) and $r = 0.1$ (light grey). Dots show the corrected mean times to fixation of a pair of compensatory mutations from our two-loci simulations. We corrected the raw means because the threshold made our data right censored (*i.e.*, missing estimates above 10^9 generations), which we accounted for by first estimating the full distribution by fitting gamma distributions on our simulation outputs using the *fitdistriplus* R package [47], from which we estimated the corrected mean times to fixation. The dashed horizontal lines indicate the generation threshold after which simulations stop. $N = 1,000$, $\mu = 10^{-5}$, $h_c = 0.5$, $s_c = 0.01$. $1,000$ *iterations*. (C) Population fitness (black dots – right y axis) and the frequencies of the 10 possible genotypes on the two loci fitness landscapes (solid lines – left y axis) over the last 4200 generations preceding the fixation of the pair of compensatory mutations. The line colours match the genotype colours on the fitness landscapes. $N = 1,000$, $\mu = 10^{-5}$, $h_c = 0.5$, $s_c = 0.01$. 100 *iterations*. See Fig. S3 and Fig. S4 for the visualisation of additional parameter combinations.

no obvious reason for selfing to accelerate RI. We first consider mutations that behave neutrally in isolation ($s = 0$). If mutation rate is low compared to drift ($4N_e\mu < 1$), the two incompatible alleles (A_2 and B_2) rarely segregate at the same time in the population and one can consider that they fix independently. According to basic results of the neutral theory, Equation 3 reduces to:

$$T = \frac{1}{2\mu} + 4N_e \tag{15}$$

The 2μ term comes from the fact that mutation can occur at the two loci. In this case, selfing only shortens the time a mutation needs to spread through the population and get fixed ($4N_e$), but does not affect the waiting time ($1/2\mu$). Because the former may be negligible compared to the latter (as implicitly assumed in [37]), selfing barely impacts the fixation time as confirmed by simulations (see below).

However, when the mutation rate is high ($4N_e\mu > 1$), the mean time a mutation needs to get fixed is shorter than $4N_e$ because mutations of the same type may arise on different individuals of the populations, so that multiple mutations can get fixed simultaneously. Taking this effect into account, Kimura [48] showed that, at a single locus, the mean time to fixation under continuous mutation pressure was:

$$T_{Kimura} = \frac{4N_e}{4N_e\mu - 1} \times (\gamma + \psi(4N_e\mu)) \tag{16}$$

where γ is Euler's constant and ψ is the digamma function. Note that (16) converge to (15) for small $N_e\mu$. Because our BDMi models have two mutation types (for the A and B loci), we cannot use equation (16) as such because mutations of different types cannot get fixed simultaneously in the populations. As a heuristic argument, we can decompose $T_{Kimura} = T_{wait} + T_{fix}$ as in equation (3), where $T_{wait} = 1/\mu$. Thus, by using $T_{fix} = T_{Kimura} - 1/\mu$ instead of $4N_e$ in (15), a more accurate expression under high mutation rate is:

$$T = T_{Kimura} - \frac{1}{2\mu} \tag{17}$$

However, mutations can segregate at the two loci at the same time and be jointly selected against, which is not taken into account in (17). Equation 17 thus serves as a neutral reference to assess whether incompatibilities between mutations segregating within populations affect the fixation time of BDMi mutations, and whether it does so independently of the population selfing rate. Simulations showed that several BDMi mutations could often segregate together in the population (Fig. S8), and cause incompatibilities within populations (Fig. 5A, Fig. S7). Accordingly, simulations showed that high mutation rate and/or high effective size slowed down the fixation of BDMi mutations (Fig. 4). Counter selection of such segregating incompatibilities hamper the fixation of either derived alleles, and stronger (high s_B; Fig. 6A) or more dominant incompatibilities ($h_B > 0.5$; Fig. 6B) increase fixation time. Finally, recombination helps forming incompatible A_2B_2 haplotypes and also reduces the accumulation of BDMi mutations (Fig. 4). Thus, selfing has opposing effects on these dynamics. On the one hand, it increases selection by exposing the incompatible haplotype in homozygouste state (especially when h_B is low). On the other hand it increases drift and reduces genetic shuffling, which reduce the occurrence of the incompatible haplotype. Overall the second effects dominate and selfing globally reduces the time to fixation of BDMi mutations.

When there is direct selection on mutations ($s > 0$), the effect of selfing on the accumulation of BDMi mutations depends on the interaction between the mutation rate (μ) and the coefficient of dominance (h) (Fig. 5B, Fig. S7). When the mutation rate is low, BDMi mutations get fixed like beneficial mutations, whose probability of

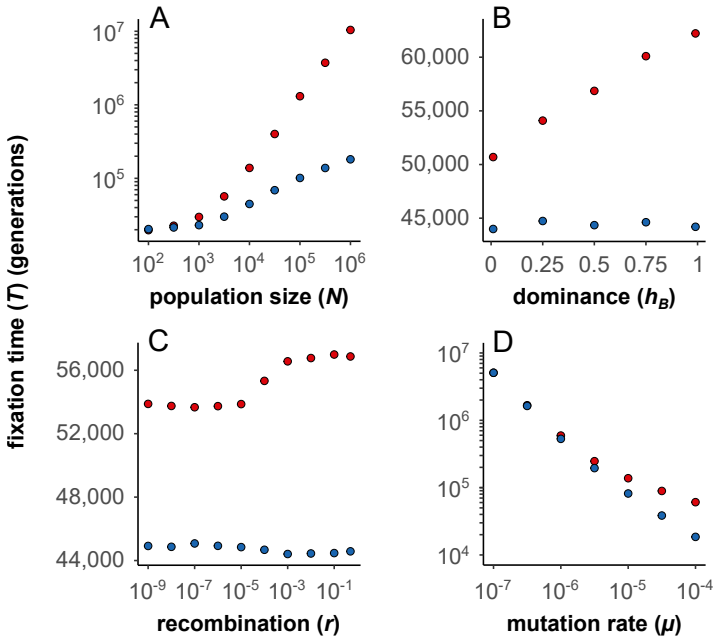


Fig 4. Fixation time of BDMi mutations (red) compared to neutral mutations (blue) in outcrossing populations ($\sigma = 0$) (two-loci model). (A) Higher population sizes (N) decelerate the fixation of BDMi mutations ($s_B = 2.5.10^{-2}$, $h_B = 0.5$; red) compared to neutral mutations ($s_B = 0$; blue). $r = 0.5$, $\mu = 2.5.10^{-5}$. $1,000$ *iterations*. (B) Higher coefficients of dominance (h_B) decelerate the fixation of BDMi mutations ($s_B = 2.5.10^{-4}$; red) compared to neutral mutations ($s_B = 0$; blue). $N = 10,000$, $r = 0.5$, $\mu = 2.5.10^{-5}$. $10,000$ *iterations*. (C) Higher rates of recombination (r) decelerate the fixation of BDMi mutations ($s_B = 2.5.10^{-4}$, $h_B = 0.5$; red) compared to neutral mutations ($s_B = 0$; blue). $N = 10,000$, $\mu = 2.5.10^{-5}$. $100,000$ *iterations*. (D) Higher rates of mutation (μ) accelerate the fixation of BDMi mutations ($s_B = 2.5.10^{-3}$, $h_B = 0.5$; red) compared to neutral mutations ($s_B = 0$; blue). $N = 10,000$, $r = 0.5$. $10,000$ *iterations*.

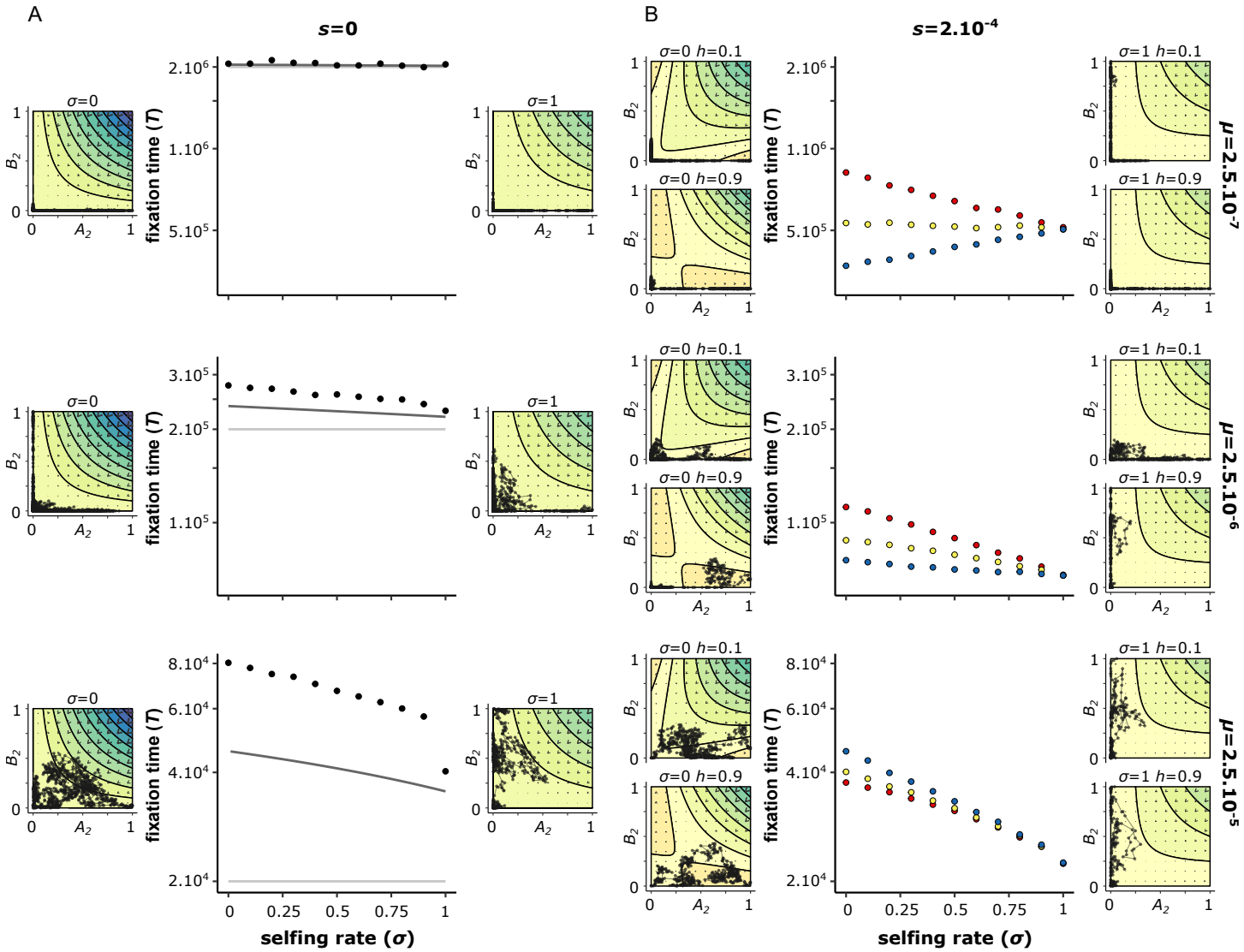


Fig 5. Selfing rate accelerates the fixation of BDMi mutations for high mutation rates The fixation time estimated from the two loci models with mutation rates (μ) for A_2 and B_2 of either $2.5.10^{-7}$ (top), $2.5.10^{-6}$ (middle), or $2.5.10^{-5}$ (bottom), and the strength of selection on the derived alleles (s) is either 0 (left), or $2.10,^{-4}$ (right). When there is selection, the coefficient of dominance (h) is either recessive ($h = 0.1$, red), co-dominant ($h = 0.5$, yellow), or dominant ($h = 0.9$, blue). The solid lines in the neutral scenario correspond to analytical approximations: $1/2\mu$ (light grey), and equation (17) (dark grey) (see BDMi Results section for details on the approximations). Each phase portrait shows, for a single simulation, the change in allele frequencies of A_2 and B_2 plotted from the beginning, and then every 100 generations until the fixation of a derived allele. The isoclines represent the expected benefits (warm colour) and costs (cold colours) on population fitness (multiplied by N_e, and with a increment of 1). The direction of the arrow indicate the expected allele change (which is the balance between mutation rates and selection), and their size indicates the strength of the change. $N = 10,000$, $h_B = k_B = 0.5$, $s_B = 10^{-3}$, $r = 0.5$. $10,000$ *iterations*.

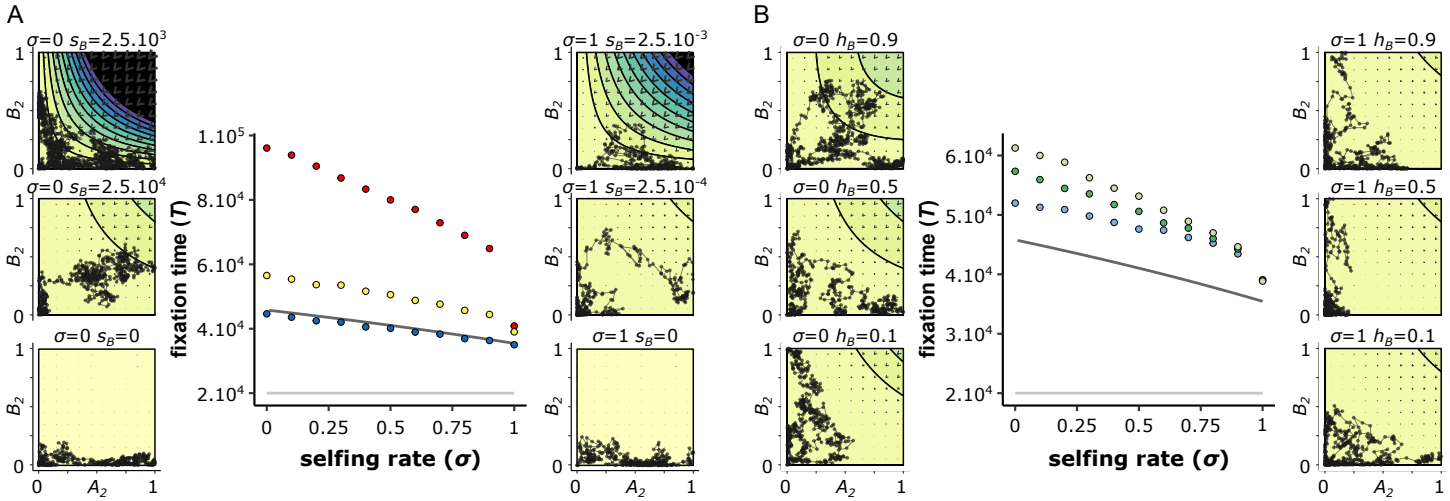


Fig 6. Within-population interferences between BDMi mutations slow down their fixation, and occur less in highly selfing populations (two loci model). The A panel shows the fixation time of co-dominant BDMi mutations ($h_B = 0.5$), and with a strength of selection (s_B) of either 0 (blue), $2.5.10^{-4}$ (yellow), or $2.5.10^{-3}$ (red). The B panel shows the fixation time of BDMi mutations with a strength of selection (s_B) of $2.5.10^{-4}$ that are either recessive ($h_B = 0.1$; light blue), co-dominant ($h_B = 0.5$; green), or dominant ($h_B = 0.9$; light green). The solid lines in the neutral scenario correspond to analytical approximations: $1/2\mu$ (light grey), and equation (17) (dark grey) (see BDMi Results section for details on the approximations). Each phase portrait shows, for a single simulation, the change in allele frequencies of A_2 and B_2 plotted from the beginning, and then every 100 generations until the fixation of a derived allele. The isoclines represent the expected benefits (warm colour) and costs (cold colours) on population fitness (multiplied by N_e, and with a increment of 1). The direction of the arrow indicate the expected allele change (which is the balance between mutation rates and selection), and their size indicates the strength of the change. $N = 10,000$, $k_B = h_B$, $s = 0$, $r = 0.5$. $10,000$ *iterations*.

fixation depends on selfing rate and the coefficient of dominance (h) [46]. When the mutation is recessive (*i.e.*, $h < 0.5$), the probability of fixation increases with selfing. When the mutation is dominant (*i.e.*, $h < 0.5$), the probability of fixation decreases with selfing. And, when the mutation is codominant (*i.e.*, $h = 0.5$), selfing does not affect the probability of fixation [46]. Therefore, when mutation rate is low, selfing speeds up fixation of recessive BDMi mutations, and slows down fixation of dominant BDMi mutations and can be approximated by equation (13) in [46] (without standing variation, $P_{sv} = 0$, their notations) (Fig. 5B, Fig. S7). In contrast, when mutation rate increases, selfing speeds up fixation of BDMi mutations regardless of the coefficient of dominance (h). This is because it is more likely that multiple BDMi mutations segregate in the population (Fig. S8) and cause genetic incompatibilities within populations, hampering more the fixation of BDMi mutations in outcrossing population than in selfing populations, as described above.

The fixation time of BDMi mutations is also affected by background selection, which has two main effects: higher drift makes selection less efficient when selfing increases and the occurrence of segregating mutations at both loci less likely. When there is no direct selection ($s = 0$), BDMi are less counter selected within populations and accumulate faster under selfing than under outcrossing. So, background selection reinforces the effect of selfing. When there is direct selection ($s > 0$) selfing reduces both selection on the beneficial allele and against the incompatible haplotype. When mutation rate is low and selection against the incompatible haplotype limited even in outcrossing populations, reducing selection on the beneficial allele predominate and selfing slows down the accumulation of BDMi mutations (Fig. 7, Fig. S9). However, for higher mutation rates, reduced selection against the incompatible haplotype predominates and selfing speeds up fixation. Overall, under a wide range of conditions, even with direct

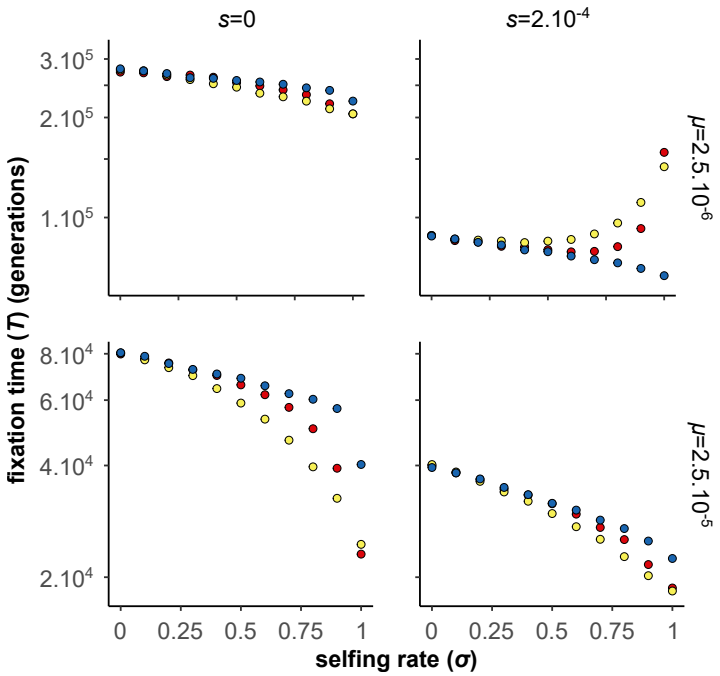


Fig 7. Effects of background selection on the accumulation of BDMi mutations (two-loci model). The graph displays the fixation time, either without (left) or with (right) selection on the derived alleles, and under different scenarios of background selection, 'linear BG effects' (yellow) or 'curved BG effects' (red), which we compared to a scenario without background selection (blue) (see methods for details on the implementation of background selection effects in our simulations). $N = 10,000$, $h = 0.5$, $s = 2.5.10^{-4}$, $h_B = k_B = 0.5$, $s_B = 10^{-3}$, $r = 0.5$. $10,000$ *iterations*.

selection, selfing facilitates RI.

In the previous analyses we considered that the direct selection on the A_2 or B_2 alleles is independent of selfing (s is a constant), as for adaptation to a new environment. Selfing only modulated the efficacy of selection through its effect on N_e and homozygosity. However, selfing can also directly affects selection, in particular in case of genetic or sexual conflicts [49]. If such conflicts play an important role in RI, it has been proposed that selfing may slow down speciation [9]. We did not explore an explicit model involving conflicts, but to mimic such a situation we considered the simple case where selfing directly reduced the selection coefficient: $s = s_0(1-\sigma)$. This is a strong effect as selection vanishes under full selfing. We consider the case of high mutation rate with background selection that corresponds to the best conditions under which selfing promotes speciation. If "conflict" selection is weak (s_0 here), selfing still facilitates the accumulation of BDMi. However, for stronger selection, despite negative interaction among segregating BDMi, fixation of BDMi is faster under outcrossing than selfing (Fig. 8).

Discussion

The role of mating systems in speciation is an old question, in particular among plant evolutionary biologists [1, 50, 51]. Depending on the underlying mechanisms, selfing has been proposed to either promote or hamper speciation [8, 9]. Surprisingly, despite this long-standing interest, specific models on the role of selfing on RI are scarce and mainly concerns the build-up of RI caused by underdominant mutations [38]. We filled this gap by expanding previous theoretical work on underdominant mutations and by

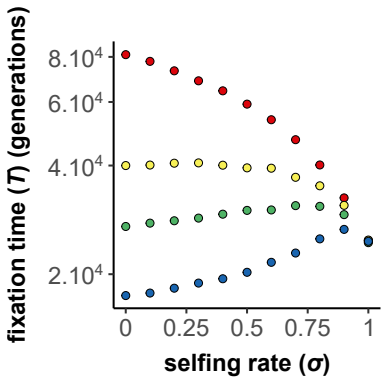


Fig 8. Effects of selfing on the fixation time of BDMi mutations when the strength of selection (s) varies with selfing rate (σ) (two-loci model). The strength of selection (s) decreases with selfing rate as follows: $s = s_0(1-\sigma)$, with $s_0 = 0$ (red), $s_0 = 2.10^{-4}$ (yellow), $s_0 = 4.10^{-4}$ (green), and $s_0 = 8.10^{-4}$ (blue). $N = 10,000$, $\mu = 2.5.10^{-5}$, $h = 0.5$, $h_B = k_B = 0.5$, $s_B = 10^{-3}$, $r = 0.5$. $10,000$ *iterations*.

considering RI caused by epistatic mutations, such as compensatory mutations and Bateson-Dobzhansky-Muller incompatibility mutations. Overall, we showed that selfing promotes allopatric speciation for a wide range of parameters. In addition, our results predicts that mating systems should affect the genomic architecture of reproductive isolation.

Selfing helps crossing fitness valleys

The Bateson-Dobzhansky-Muller model was initially proposed as a possible solution of the puzzling question of the evolution of hybrid incompatibilities as it does not require crossing fitness valleys [52]. Alternatively, some mechanisms can help crossing fitness valleys, and it is well known that selfing facilitates the fixation of underdominant mutations [38]. We extended this model to a two-loci fitness landscape for which selfing also helps crossing the valley under most conditions. Selfing has two main effects: first it increases drift (in particular if background selection is strong); second, because of reduced recombination and heterozygozity, it limits the breakdown of the new fittest genotype, once it has been created. At the genome scale, assuming several loci where underdominant fitness landscape may occur, we also showed that the effect of selfing is stronger when there is a highly skewed distribution of the depth of the valleys to be crossed (11). As similar conclusion is also likely for the compensatory mutation model although we did not obtain an equivalent analytical result.

The role of interferences among mutations segregating within populations

In the simplest form of the Bateson-Dobzhansky-Muller model of speciation, genetic incompatibilities occur between derived alleles that are supposed to arise and become fixed independently in different populations [7,53], and the phase during which BDMi mutations emerge, spread through the population and eventually get fixed is often dismissed (e.g. [36]). This inevitable phase has recently been argued to have important implications in speciation genetics [54]. Considering that BDMi alleles may segregate in natural populations at polymorphic frequencies allows for instance to better explain (i) why hybrid incompatibility may be variable between different pairs of individuals originating from the same two populations (reviewed in [54]), and (ii) why genetic

incompatibilities is widespread within species, as found in *Drosophila melanogaster* [55], *Caenorhabditis elegans* [56], *Arabidopsis thaliana* [57], or in the genus *Draba* [18].

To mimic genetic incompatibilities arising from mutations at multiple sites in the genome, we used elevated mutation rates ($4N_e\mu > 1$) in the two-loci model, which allowed us to dissect the dynamics of multiple incompatible alleles segregating within populations. Under these conditions, we showed that epistatic interactions among segregating BDMi delays their fixations, especially when they are unlinked and when they are not too recessive. These results were validated by explicit multi-loci simulations. This effect was not predicted by previous approximations that showed that, when mutations are rare enough, RI only depended on mutation rate, independently on population parameters and reproductive mode [37]. However, they are in agreement with some phenomenological models that assumed that incompatibilities was function of the genetic distance between individuals, so can be counter selected in large polymorphic populations [58].

The purging of segregating BDMi mutations bears implications for the effect of selfing on RI. First, selfing increases drift and reduces polymorphism (including for BDMi mutations) and second, selfing reduces genetic shuffling thus limits the possibility for two mutations arose in different individuals and genetic background to be gathered in a same genotype. The combination of both effects facilitates RI compared to selfing, even under certain conditions where outcrossing favours the fixation of locally selected BDMi mutations.

Mating systems and the pace of speciation

It was previously unknown if and how selfing affects the pace of speciation. Our results overall suggest that selfing reduces the fixation time of underdominant, compensatory, and BDMi mutations in allopatry, making speciation overall faster in selfing populations. Remarkably, this effect may even persist in the face of local adaptation, suggesting that – contrarily to our initial hypothesis – ecological speciation may occur faster in selfing populations than in outcrossing populations. It is unclear what are the relative importance of underdominant, compensatory, and BDMi mutation in determining the pace of speciation in natural populations. In any case, selfing broaden the spectrum of incompatibilities that can fix, and so should on average shorten the waiting time to complete speciation, even for ecological speciation.

However, the clear condition under which outcrossing should promote speciation is when it is driven by genomic or sexual conflicts, which may completely vanish under complete selfing. It is for instance known that sexual conflicts over maternal provisioning during seed development are usually stronger in outcrossers than in selfers (e.g. [59]), so that the sexually antagonistic co-evolution between male and female traits is expected to go faster in outcrossing *vs.* selfing populations, and thus to promote speciation more in outcrossing *vs.* selfing lineages [60]. A more explicit analysis remains to be done but our basic model (where selection linearly decreases with the selfing rate) confirms this prediction as soon as antagonistic selection is strong enough (say of the order of $N_e s > 5$).

So far, empirical results are still limited but tend to support these predictions, for example with the accumulation of numerous incompatibilities between recently diverged population of selfing arctic species [18,19] or with macro-analyses suggesting higher speciation rates in selfing lineages in Solanaceae [12,16] as mentioned in the introduction. However, the underlying process of speciation remains unknown. In arctic species, divergence in allopatry is likely, but in Solanaceae, selfing may have promoted speciation through the limitation of gene flow, which we did not study here.

Mating systems and the genetic architecture of reproductive isolation

Beyond the effect of mating systems on the pace of speciation, our outcomes clearly suggest that mating system should also affect the genetic architecture of speciation. In particular, underdominant and compensatory mutations are expected to be found relatively more often as reproductive barriers among selfers than among outcrossers. Classical examples of genetic modifications leading to underdominant effects include chromosomal rearrangement, which can generate and maintain RI between populations or species [28,61]. To our knowledge, there is no studies specifically comparing the occurrence of underdominant chromosomal rearrangement in selfing *vs.* outcrossing species. Reproductive isolation due to underdominant chromosomal rearrangement is however more often found in plants than in animals [62], which is possibly due to a higher frequency of selfing plant species. Our results also suggest that reproductive barriers caused by a few strongly underdominant mutations are more likely to differ between mating systems than reproductive barriers caused by many weakly underdominant mutations.

Compensatory effects are often discussed in the context of the evolution of gene expression for which stabilising selection may lead to the co-evolution of *cis*- and *trans*-regulatory mutations (*e.g.*, a *cis*-regulatory mutation increasing gene expression may be compensated by a *trans*-regulatory mutation decreasing gene expression, or *vice versa*) [63]. Although there compensatory mutations are expected to take a long time to get fixed [29,40], co-evolution of *cis*- and *trans*-regulatory mutations have been found to contribute to RI between (outcrossing) species of *Drosophila* [30] or mice [32], and between (selfing) species of nematode [31].

Finally, our models predict that in outcrossing species BDMi mutations are more likely to fix when they are clustered (but in repulsion) than when they are widespread. On the contrary, there is no specific constraint on genomic location in selfing species such that pairs of incompatible alleles could arise everywhere in a genome. This conclusion resembles the prediction that genes involved in local adaptation [64] or in domestication [65] should be less clustered under selfing than under outcrossing.

Conclusions

Our analytical and simulation models show that selfing overall fosters the accumulation of underdominant, compensatory, and BDMi mutations in allopatry. This outcome help us predicting the speciation rates as well as the architecture of RI of selfing *vs.* outcrossing species. Our results bring a theoretical background to long-standing ideas [1,5] and are tentatively supported by both phylogenetic studies and crossing experiments – though additional empirical work is needed. Future theoretical work will need to account for the effect of selfing on the stability of RI in the face of gene flow.

References

1. Grant V. Plant Speciation. Columbia University Press; 1981.

2. Goodwillie C, Kalisz S, Eckert CG. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. Annu Rev Ecol Evol Syst. 2005;36:47–79.

3. Jarne P, Auld JR. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. Evolution. 2006;60(9):1816–1824.

4. Igic B, Kohn JR. The distribution of plant mating systems: study bias against obligately outcrossing species. Evolution. 2006;60(5):1098–1103.

5. Stebbins GL. Self fertilization and population variability in the higher plants. The American Naturalist. 1957;91(861):337–354.

6. Lande R, Schemske DW. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution. 1985;39(1):24–40.

7. Coyne JA, Orr HA. Speciation. vol. 37. Sinauer associates Sunderland, MA; 2004.

8. Wright SI, Kalisz S, Slotte T. Evolutionary consequences of self-fertilization in plants. Proceedings of the Royal Society B: Biological Sciences. 2013;280(1760):20130133.

9. Cutter AD. Reproductive transitions in plants and animals: selfing syndrome, sexual selection and speciation. New Phytologist. 2019;224(3):1080–1094.

10. Takebayashi N, Morrell PL. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. American Journal of Botany. 2001;88(7):1143–1150.

11. Igic B, Busch JW. Is self-fertilization an evolutionary dead end? New Phytologist. 2013;198(2):386–397.

12. Goldberg EE, Kohn JR, Lande R, Robertson KA, Smith SA, Igić B. Species selection maintains self-incompatibility. Science. 2010;330(6003):493–495.

13. de Vos JM, Hughes CE, Schneeweiss GM, Moore BR, Conti E. Heterostyly accelerates diversification via reduced extinction in primroses. Proceedings of the Royal Society B: Biological Sciences. 2014;281(1784):20140075.

14. Landis JB, Bell CD, Hernandez M, Zenil-Ferguson R, McCarthy EW, Soltis DE, et al. Evolution of floral traits and impact of reproductive mode on diversification in the phlox family (Polemoniaceae). Molecular Phylogenetics and Evolution. 2018;127:878–890.

15. Freyman WA, Höhna S. Stochastic character mapping of state-dependent diversification reveals the tempo of evolutionary decline in self-compatible Onagraceae lineages. Systematic Biology. 2019;68(3):505–519.

16. Zenil-Ferguson R, Burleigh JG, Freyman WA, Igić B, Mayrose I, Goldberg EE. Interaction among ploidy, breeding system and lineage diversification. New Phytologist. 2019;224(3):1252–1265.

17. Ho EK, Agrawal AF. Aging asexual lineages and the evolutionary maintenance of sex. Evolution. 2017;71(7):1865–1875.

18. Grundt HH, Kjølner S, Borgen L, Rieseberg LH, Brochmann C. High biological species diversity in the arctic flora. Proceedings of the National Academy of Sciences. 2006;103(4):972–975.

19. Gustafsson ALS, Gussarova G, Borgen L, Ikeda H, Antonelli A, Marie-Orleach L, et al. Rapid evolution of post-zygotic reproductive isolation is widespread in Arctic plant lineages. Annals of Botany. 2022;129(2):171–184.

20. Skrede I, Brochmann C, Borgen L, Rieseberg LH. Genetics of intrinsic postzygotic isolation in a circumpolar plant species, *Draba nivalis* (Brassicaceae). Evolution. 2008;62(8):1840–1851.

21. Gustafsson ALS, Skrede I, Rowe HC, Gussarova G, Borgen L, Rieseberg LH, et al. Genetics of cryptic speciation within an arctic mustard, *Draba nivalis*. PloS one. 2014;9(4):e93834.

22. Pollak E. On the theory of partially inbreeding finite populations. I. Partial selfing. Genetics. 1987;117(2):353–60.

23. Nordborg M. Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. Genetics. 2000;154(2):923–929.

24. Felsenstein J. The evolutionary advantage of recombination. Genetics. 1974;78(2):737–756.

25. Barrett SC, Arunkumar R, Wright SI. The demography and population genomics of evolutionary transitions to self-fertilization in plants. Philosophical Transactions of the Royal Society B: Biological Sciences. 2014;369(1648):20130344.

26. Glémin S. Mating systems and the efficacy of selection at the molecular level. Genetics. 2007;177(2):905–916.

27. Roze D. Background selection in partially selfing populations. Genetics. 2016;203(2):937–957.

28. Zhang L, Reifová R, Halenková Z, Gompert Z. How important are structural variants for speciation? Genes. 2021;12(7):1084.

29. Phillips PC. Waiting for a compensatory mutation: phase zero of the shifting-balance process. Genetics Research. 1996;67(3):271–283.

30. Landry CR, Wittkopp PJ, Taubes CH, Ranz JM, Clark AG, Hartl DL. Compensatory cis-trans evolution and the dysregulation of gene expression in interspecific hybrids of *Drosophila*. Genetics. 2005;171(4):1813–1822.

31. Barrière A, Gordon KL, Ruvinsky I. Coevolution within and between regulatory loci can preserve promoter function despite evolutionary rate acceleration. PloS Genetics. 2012;.

32. Mack KL, Nachman MW. Gene regulation and speciation. Trends in Genetics. 2017;33(1):68–80.

33. Bateson W. Heredity and variation in modern lights. Darwin and modern science. 1909;.

34. Dobzhansky T. Studies on hybrid sterility. II. Localization of sterility factors in *Drosophila pseudoobscura* hybrids. Genetics. 1936;21(2):113.

35. Muller H. Isolating mechanisms, evolution, and temperature. In: Biol. Symp.. vol. 6; 1942. p. 71–125.

36. Orr HA. The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics. 1995;139(4):1805–1813.

37. Gavrilets S. Fitness Landscapes and the Origin of Species (MPB-41). Princeton University Press; 2004.

38. Charlesworth B. Evolutionary rates in partially self-fertilizing species. The American Naturalist. 1992;140(1):126–148.

39. Kimura M. The role of compensatory neutral mutations in molecular evolution. Journal of Genetics. 1985;64(1):7–19.

40. Stephan W. The rate of compensatory evolution. Genetics. 1996;144(1):419–426.

41. Kimura M, King JL. Fixation of a deleterious allele at one of two" duplicate" loci by mutation pressure and random drift. Proceedings of the National Academy of Sciences. 1979;76(6):2858–2861.

42. Galassi M, Davies J, Theiler J, Gough B, Jungman G, Alken P, et al. GNU scientific library. Reference Manual Edition 14, for GSL Version 14. 2003;.

43. Charlesworth B, Morgan M, Charlesworth D. The effect of deleterious mutations on neutral molecular variation. Genetics. 1993;134(4):1289–1303.

44. Haller BC, Messer PW. SLiM 3: Forward Genetic Simulations Beyond the Wright–Fisher Model. Molecular Biology and Evolution. 2019;36(3):632–637. doi:10.1093/molbev/msy228.

45. Tange O, et al. Gnu parallel-the command-line power tool. The USENIX Magazine. 2011;36(1):42–47.

46. Glémin S, Ronfort J. Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. Evolution. 2013;67(1):225–240.

47. Delignette-Muller ML, Dutang C, Pouillot R, Denis JB, Delignette-Muller MML. Package 'fitdistrplus'; 2015.

48. Kimura M. Average time until fixation of a mutant allele in a finite population under continued mutation pressure: Studies by analytical, numerical, and pseudo-sampling methods. Proceedings of the National Academy of Sciences. 1980;77(1):522–526.

49. Glémin S. Balancing selection in self-fertilizing populations. Evolution. 2021;75(5):1011–1029.

50. Stebbins GLS. Variation and Evolution in Plants. Columbia University Press; 1950.

51. Stebbins GLS. Evolution above the Species Level. Harvard University Press; 1974.

52. Orr HA. Dobzhansky, Bateson, and the Genetics of Speciation. Genetics. 1996;144:1331–1335. doi:10.1093/genetics/144.4.1331.

53. Presgraves DC. Darwin and the origin of interspecific genetic incompatibilities. The American Naturalist. 2010;176(S1):S45–S60.

54. Cutter AD. The polymorphic prelude to Bateson–Dobzhansky–Muller incompatibilities. Trends in ecology & evolution. 2012;27(4):209–218.

55. Corbett-Detig RB, Zhou J, Clark AG, Hartl DL, Ayroles JF. Genetic incompatibilities are widespread within species. Nature. 2013;504(7478):135–137.

56. Snoek LB, Orbidans HE, Stastna JJ, Aartse A, Rodriguez M, Riksen JA, et al. Widespread genomic incompatibilities in *Caenorhabditis elegans*. G3: Genes, Genomes, Genetics. 2014;4(10):1813–1823.

57. Chae E, Bomblies K, Kim ST, Karelina D, Zaidem M, Ossowski S, et al. Species-wide genetic incompatibility analysis identifies immune genes as hot spots of deleterious epistasis. Cell. 2014;159(6):1341–1351.

58. Gavrilets S. A dynamical theory of speciation on holey adaptive landscapes. The American Naturalist. 1999;154(1):1–22.

59. Raunsgard A, Opedal ØH, Ekrem RK, Wright J, Bolstad GH, Armbruster WS, et al. Intersexual conflict over seed size is stronger in more outcrossed populations of a mixed-mating plant. Proceedings of the National Academy of Sciences. 2018;115(45):11561–11566.

60. Lafon-Placette C, Köhler C. Endosperm-based postzygotic hybridization barriers: developmental mechanisms and evolutionary drivers. Molecular Ecology. 2016;25(11):2620–2629.

61. Stathos A, Fishman L. Chromosomal rearrangements directly cause underdominant F1 pollen sterility in *Mimulus lewisii–Mimulus cardinalis* hybrids. Evolution. 2014;68(11):3109–3119.

62. Rieseberg LH. Chromosomal rearrangements and speciation. Trends in ecology & evolution. 2001;16(7):351–358.

63. Signor SA, Nuzhdin SV. The evolution of gene expression in cis and trans. Trends in Genetics. 2018;34(7):532–544.

64. Hodgins KA, Yeaman S. Mating system impacts the genetic architecture of adaptation to heterogeneous environments. New Phytologist. 2019;224(3):1201–1214. doi:10.1111/nph.16186.

65. Le Thierry d'Ennequin M, Toupance B, Robert T, Godelle B, Gouyon PH. Plant domestication: a model for studying the selection of linkage. Journal of Evolutionary Biology. 1999;12(6):1138–1147. doi:10.1046/j.1420-9101.1999.00115.x.

Supporting information

Fig. S1 Effects of selfing on the accumulation of underdominant mutations.

Fig. S2 Effects of selfing on the accumulation of a pair of compensatory mutations when the strength of the deleterious effect (s_c) is low (two loci model).

Fig. S3 Effects of selfing on the accumulation of a pair of compensatory mutations (two loci model).

Fig. S4 Effects of selfing rate on the fitness of a population, and the path taken on the fitness landscape, over the 4200 generations preceding the fixation of the pair of compensatory mutation (two loci model).

Fig. S5 Effects of background selection on the accumulation of compensatory mutations (two loci model).

Fig. S6 Effects of selfing on the accumulation of compensatory mutations (multi loci model).

Fig. S7 Effects of selfing and selection on the accumulation of BDMi mutations (multi loci model).

Fig. S8 Effects of selfing and selection on the fixation time and the number of segregating of BDMi mutations in a population (multi-loci model).

Fig. S9 Effects of selfing, selection and background selection on the time to fixation fixation of BDMi mutations in a population (multi-loci model).

Appendix S1 Mathematical derivations

File S1 Mathematica notebook.

Code availability

Mathematica script, C++ script, SLiM script.

Acknowledgments

We acknowledge the GenOuest bioinformatics core facility (https://www.genouest.org) for providing the computing infrastructure. Funding was obtained from the Research Council of Norway (https://www.forskningsradet.no) to CB (project 274607: Speciation-Clock—How fast does the 'speciation clock' tick in selfing versus outcrossing lineages?). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

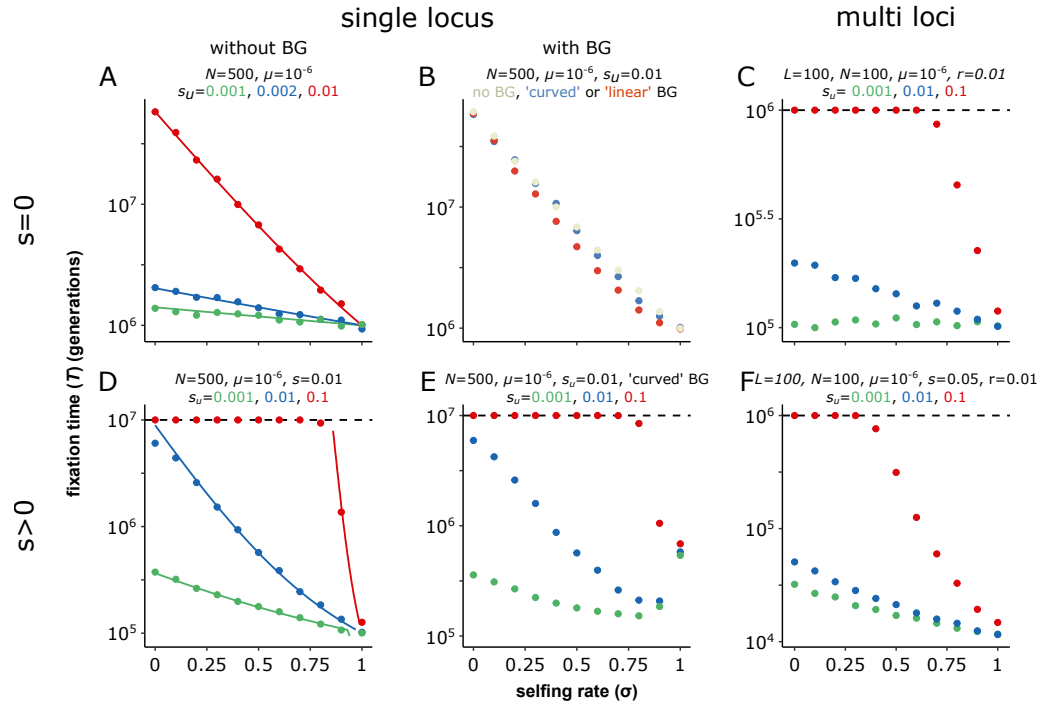


Fig S1. Effects of selfing on the accumulation of underdominant mutations. Outcomes of the simulations (circles) for the number of generations to fixate either one derived allele (single locus; A, B, D, E), or ten derived alleles (multi loci models; C, F). The top row corresponds to symmetrical underdominant mutations (*i.e.*, $s = 0$). The bottom row corresponds to asymetrical underdominant mutations (*i.e.*, $s > 0$). Single locus simulations were performed either with (A, D) or without (B, E) background selection, BG. The dashed horizontal lines correspond to the threshold after which simulations stopped to avoid computing burden. $1,000$ *iterations*. Lines on A and D correspond to numerical approximations.

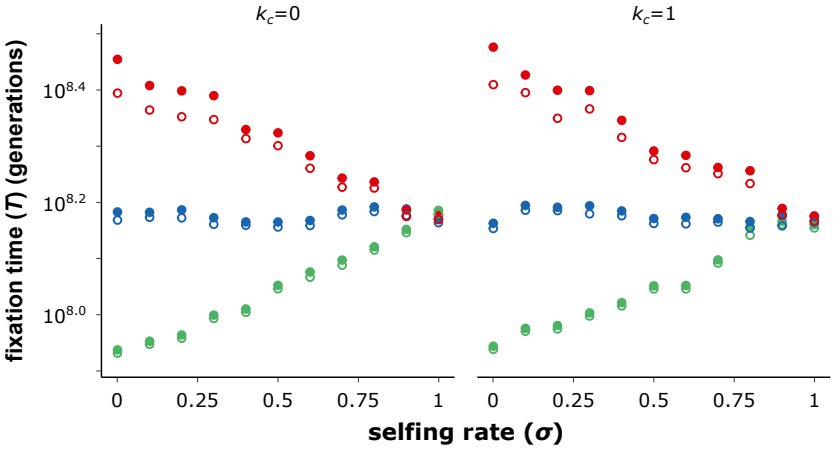


Fig S2. Effects of selfing on the accumulation of a pair of compensatory mutations when the strength of the deleterious effect (s_c) is low (two loci model). Means of the two loci simulations for the number of generations needed to fixate the first (deleterious) mutation (empty dots), and the second (compensatory) mutation (filled dots). The dot color stands for the coefficient of dominance of the deleterious effect, which is either recessive ($h_c = 0$, green), codominant ($h_c = 0.5$, blue) or dominant ($h_c = 1$, red). The coefficient of dominance of the double heterozygotes (k_c) is set to 0 (left) or 1 (right). Selfing rate (σ) ranges from 0 to 1, with a 0.1 increment. $N = 1,000$, $\mu = 10^{-7}$, $s_c = 0.0025$, $r = 0.5$. $1,000$ *iterations*.

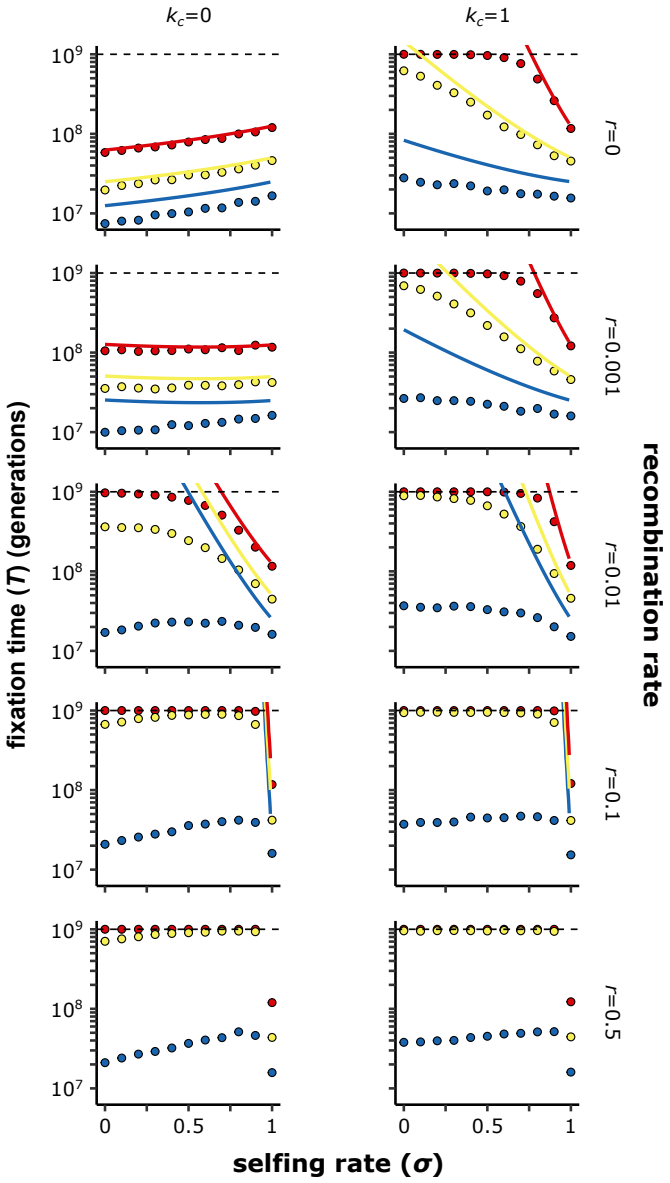
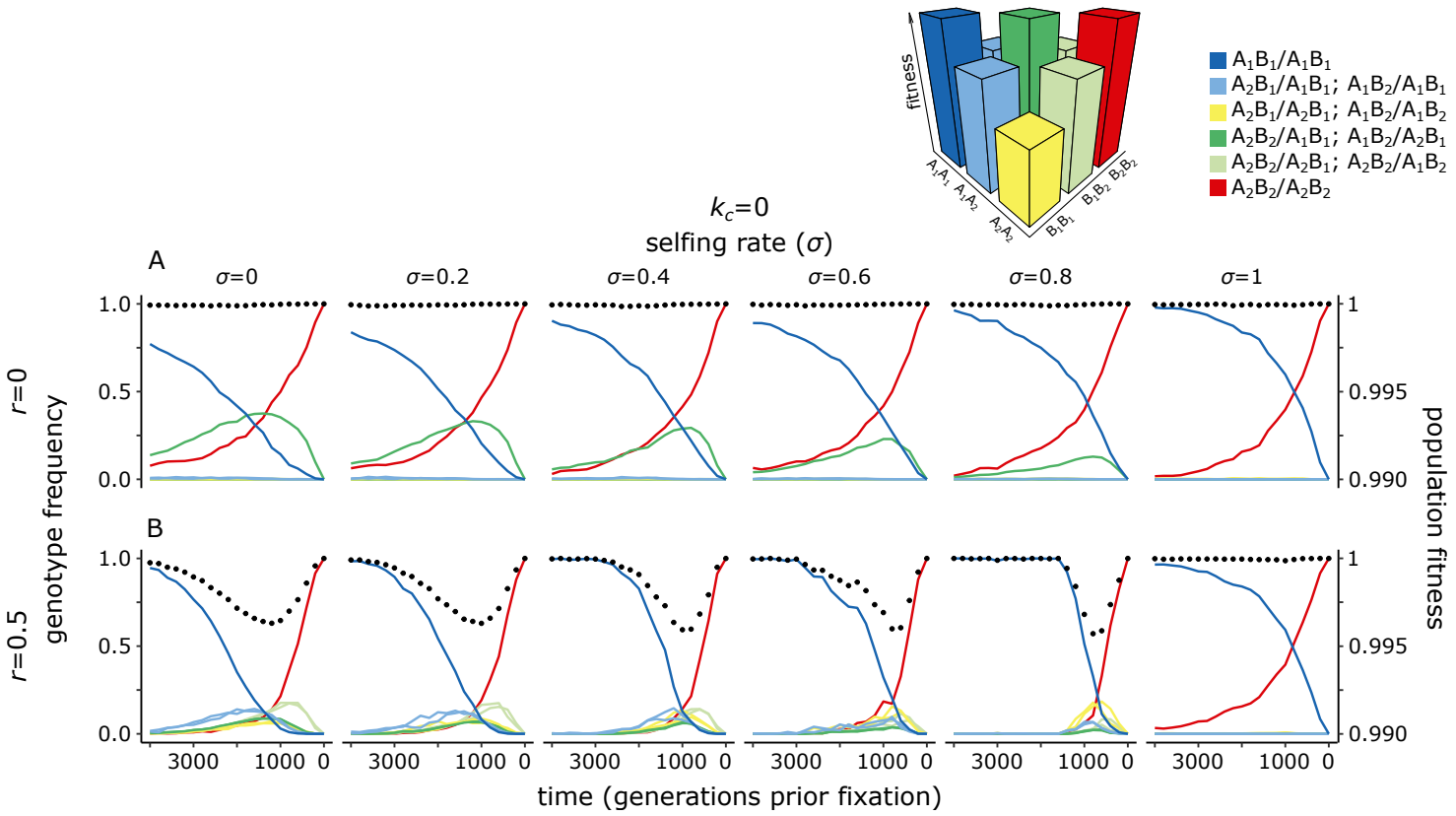


Fig S3. Effects of selfing on the accumulation of a pair of compensatory mutations (two loci model). Analytical approximations (lines; eq. 14) and outcomes of the two loci simulations (dots) for the mean number of generations needed to fixate a pair of compensatory mutations. The recombination rate between the two loci (r) ranges from 0 (top) to 0.5 (down), and the coefficient of dominance of the double heterozygotes (k_c) is set to 0 (left) or 1 (right). The strength of the deleterious effect (s_c) is set to 0.005 (blue), 0.010 (yellow), or 0.025 (red). The threshold after which simulations terminate was set to 10^9*generations* (dashed line). Selfing rate (σ) ranges from 0 to 1, with a 0.1 increment. $N = 1,000$, $\mu = 10^{-5}$, $h_c = 0.5$. 1000 *iterations*.


fitness
A1A1
A1A2
A2A2
B1B1
B1B2
B2B2
A1B1/A1B1
A2B1/A1B1; A1B2/A1B1
A2B1/A2B1; A1B2/A1B2
A2B2/A1B1; A1B2/A2B1
A2B2/A2B1; A2B2/A1B2
A2B2/A2B2
kc=0
selfing rate (σ)
A
B
σ=0
σ=0.2
σ=0.4
σ=0.6
σ=0.8
σ=1
r=0
r=0.5
genotype frequency
population fitness
1.0
0.5
0.0
1
0.995
0.990
3000
1000
0
time (generations prior fixation)

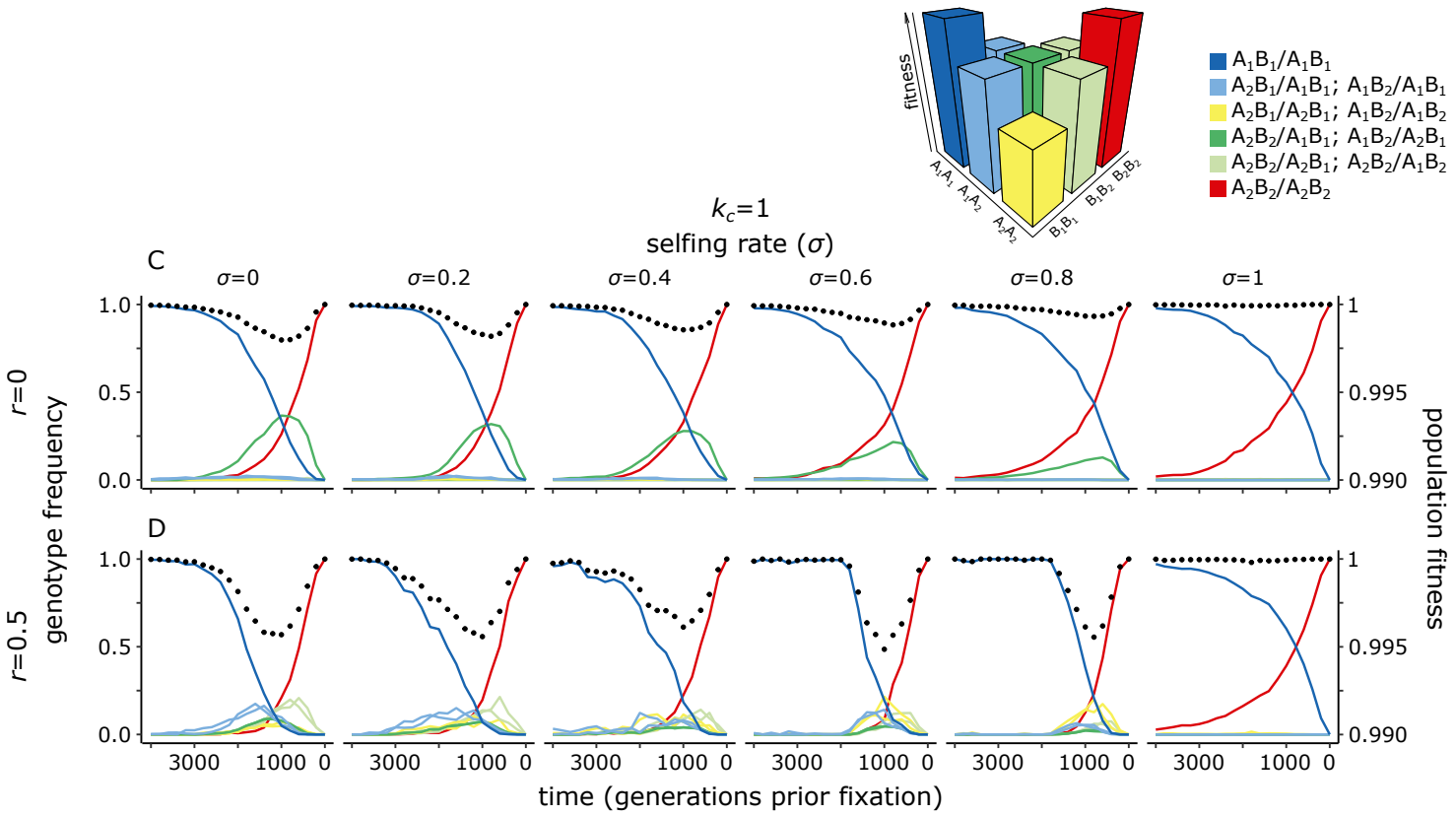


Fig S4. Effects of selfing rate on the fitness of a population, and the path taken on the fitness landscape, over the 4200 generations preceding the fixation of the pair of compensatory mutation (two loci model). Outcomes of the two loci simulations showing population fitness (black dots, right y axis) and the frequencies of the 10 possible genotypes on the two loci fitness landscapes (solid lines, left y axis). Selfing rate (σ) ranges from 0 (left) to 1 (right), with a 0.2 increment. The coefficient of dominance of the double heterozygotes (k_c) is set to 0 (A, B) or 1 (C, D), and the recombination rate between the two loci is set to 0 (A, C) or to 0.5 (B, D). $N = 1,000$, $\mu = 10^{-5}$, $h_c = 0.5$, $s_c = 0.01$. 100 *iterations*.

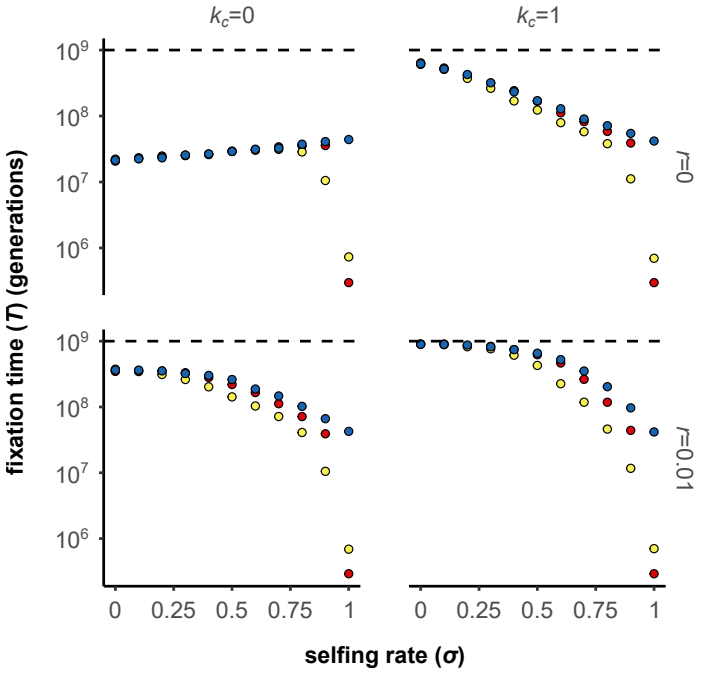


Fig S5. Effects of background selection on the accumulation of compensatory mutations (two loci model). The graph displays the fixation time estimated from the two loci models with a coefficient of dominance of the double heterozygotes (k_c) set to 0 (left) or 1 (right), and under different scenarios of background selection, 'linear BG effects' (yellow) or 'curved BG effects' (red), which we compared to a scenario without background selection scenario (blue) (see methods for details on the implementation of background selection effects in our simulations). $N = 1,000$, $\mu = 10^{-5}$, $h_c = 0.5$, $s_c = 10^{-2}$, $r = 0.5$. $10,000$ *iterations*

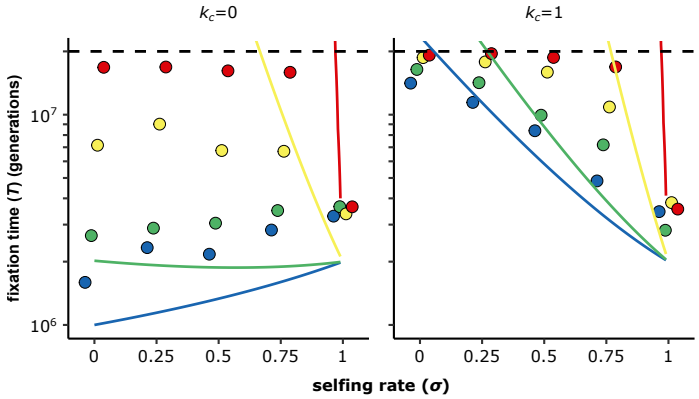


Fig S6. Effects of selfing on the accumulation of compensatory mutations in multi loci models (multi loci model). The graph shows the two loci analytical approximations (lines; eq. 14) and the outcomes of the multi loci simulations (dots) for the mean number of generation needed for the fixation of at least two compensatory mutations. The threshold after which the simulation terminates is set to 2.10^7 generations (dashed lines). The coefficient of dominance of the double heterozygotes (k_c) is set to 0 (left) or 1 (right), and the recombination rate between each locus is set to 0 (blue), 0.0001 (green), 0.001 (yellow), or 0.01 (red). Selfing rate (σ) ranges from 0 to 1, with a 0.25 increment. $L = 100$, $N = 200$, $\mu = 10^{-6}$, $h_c = 0.5$, $s_c = 0.04$. 100 *iterations*.

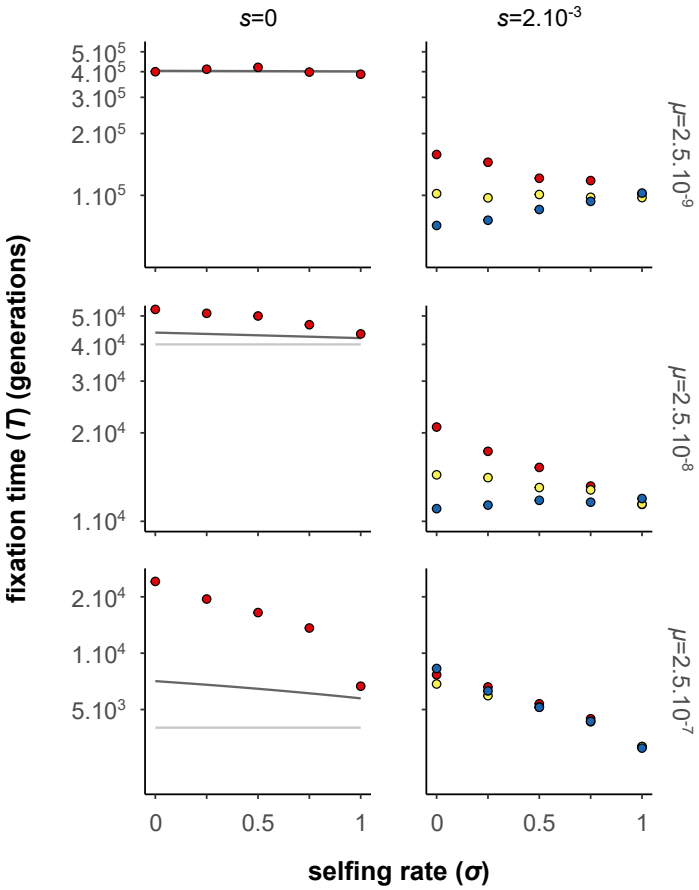


Fig S7. Effects of selfing and selection on the accumulation of BDMi mutations (multi loci model). The panel displays the fixation time estimated when the mutation rate, μ, is either $2.5.10^{-9}$ (top), $2.5.10^{-8}$ (middle), or $2.5.10^{-7}$ (bottom). The strength of selection on the derived alleles, s, is either 0 (left), or 2.10^{-3} (right). When there is selection, the dominance coefficient of the derived alleles, h, is either recessive (0.1, red), codominant (0.5, yellow), or dominant (0.9, blue). The solid lines in the neutral scenario correspond to analytical approximations: $1/2\mu$ (light grey), and equation (17) (dark grey) (see BDMi Results section for details on the approximations). $L = 1,000$, $N = 1,000$, $h_B = k_B = 0.5$, $s_B = 10^{-2}$, $r = 10^{-3}$. $1,000$ *iterations*.

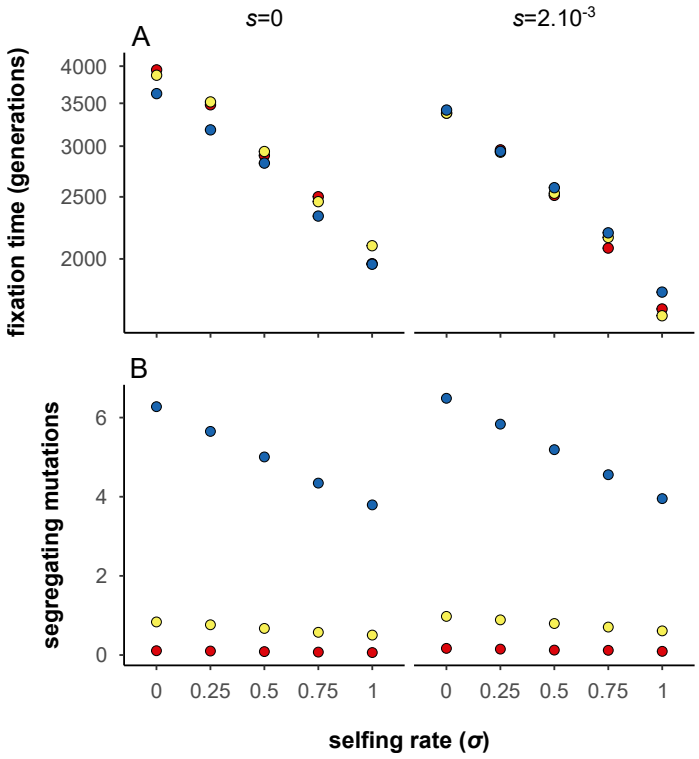


Fig S8. Effects of selfing and selection on the fixation time and the number of segregating of BDMi mutations in a population (multi-loci model). The panel displays (A) the time a mutation takes to fixate (fixation time) and (B) the mean number of BDMi mutations segregating in a population sampled every 100 generations. The mutation rate, μ, is either $2.5.10^{-9}$ (red), $2.5.10^{-8}$ (yellow), or $2.5.10^{-7}$ (blue). The strength of selection on the derived alleles, s, is either 0 (left), or 2.10^{-3} (right). $L = 1,000$, $N = 1,000$, $h = 0.5$, $h_B = k_B = 0.5$, $s_B = 10^{-2}$, $r = 10^{-3}$. $1,000$ *iterations*.

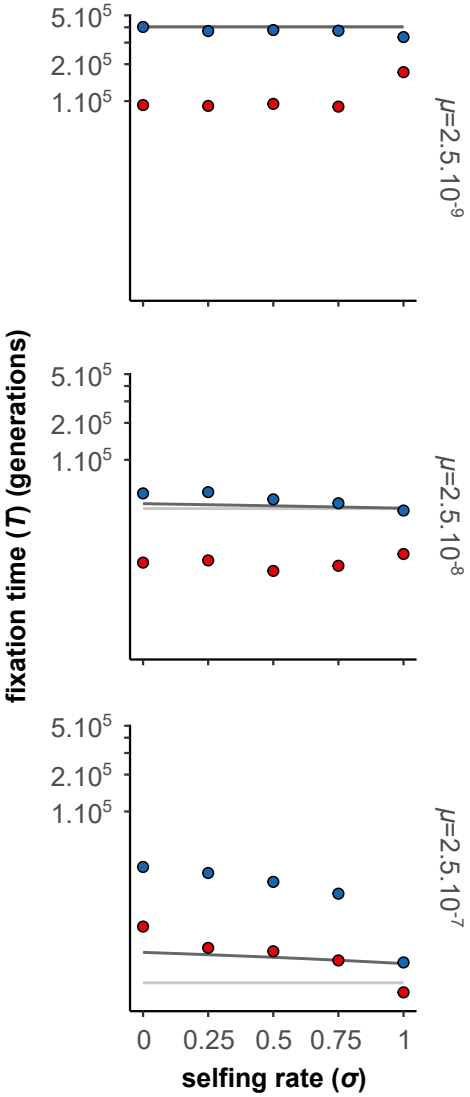


Fig S9. Effects of selfing, selection and background selection on the time to fixation fixation of BDMi mutations in a population (multi-loci model). The graph displays the mean fixation time of BDMi mutations when the strength of selection on the derived alleles, s, is either equal to 0 (blue) or to 2.10^{-3} b(red). The mutation rate, μ, is either $2.5.10^{-9}$ (top), $2.5.10^{-8}$ (middle), or $2.5.10^{-7}$ (bottom). In addition to the BDMi mutations, deleterious mutations with a coefficient of dominance of 0.1 and a strength of selection of -10^{-3} occured at a $5.10^{-}3$ rate. The solid lines in the neutral scenario correspond to analytical approximations: $1/2\mu$ (light grey), and equation (17) (dark grey) (see BDMi Results section for details on the approximations). $L = 1,000$, $N = 1,000$, $h = 0.5$, $h_B = k_B = 0.5$, $s_B = 10^{-2}$, $r = 10^{-3}$. 100 *iterations*.

Mating system and speciation I: the accumulation of genetic incompatibilities in allopatry

- Supplementary Material -

The corresponding Mathematica notebook is also provided

Lucas Marie-Orleach, Christian Brochmann and Sylvain Glémin

22nd July 2022

A1 Underdominance

A1.1 Model

We consider a population of size N reproducing with partial selfing in proportion σ. The effective size of the population is (Pollak 1987; Caballero and Hill 1992):

$$N_e = \frac{N}{1+F} \tag{A1}$$

where F is the Wright's fixation index which neutral expectation is:

$$F = \frac{\sigma}{2-\sigma} \tag{A2}$$

We consider a single bi-allelic locus, with the ancestral allele A_1 that can mutate in the derived allele A_2 at rate μ. We note the fitness of genotypes A_1A_1, A_1A_2, and A_2A_2, 1, $1-s_u$, and $1+s$, respectively, and x the frequency of allele A_2. The change in allelic frequencies in one generation is given by:

$$\begin{aligned}\Delta x &= x(1-x)\Big((1-F)(sx - s_u(1-2x) + Fs\Big)/\bar{W} \\ &\approx x(1-x)\Big((1-F)(sx - s_u(1-2x) + Fs\Big)\end{aligned} \tag{A3}$$

where $\bar{W}$ is the mean fitness of the population. Under weak selection $\bar{W} \approx 1$ (second line of A3) and F can be equated to its neutral expectation (A2). Equation

(A3) can also be written as:

$$\begin{aligned}\Delta x &= (1-F)Sx(1-x)(x-x_{eq}) \quad \text{if } F<1 \\ &= sx(1-x) \quad \text{if } F=1\end{aligned} \tag{A4}$$

where $S = s + 2s_u$, representing the total amount of selection (s_u between A_1A_1 and A_1A_2 and $s_u + s$ between A_1A_2 and A_2A_2), and x_{eq} is the internal (unstable) equilibrium:

$$x_{eq} = \frac{(1-F)s_u - Fs}{(1-F)(2s_u+s)} \tag{A5}$$

This internal equilibrium exist ($0 \leq x_{eq} \leq 1$) if $F \leq s/(s+s_u)$. Above this threshold, selection becomes directional and positive, and for F exactly equal to this threshold, equation (A3) becomes:

$$\Delta x = \frac{s}{s+s_u}(s+2s_u)x^2(1-x) \tag{A6}$$

which is equivalent to selection on a fully recessive allele with selective advantage $s(s+2s_u)/(s+s_u)$.

A1.2 Probability and time to fixation

Noting $M_{\delta x} = \Delta x$ the expected infinitesimal change in allelic frequency and $V_{\delta x} = \frac{x(1-x)}{2N_e}$ is the infinitesimal variance we define the so-called Green function as:

$$G(x) = e^{-\int 2M_{\delta x}/V_{\delta x}dx} \tag{A7}$$

The probability of fixation of a single A_2 mutant is then given by (Kimura 1962):

$$P_{fix} = \frac{\int_0^{1/2N} G(x)dx}{\int_0^1 G(x)dx} \tag{A8}$$

This solves to:

$$P_{fix} = \frac{\text{erf}\left(x_{eq}\sqrt{2N_eS(1-F)}\right) - \text{erf}\left(\left(x_{eq} - \frac{1}{2N}\right)\sqrt{2N_eS(1-F)}\right)}{\text{erf}\left((1-x_{eq})\sqrt{2N_eS(1-F)}\right) + \text{erf}\left(x_{eq}\sqrt{2N_eS(1-F)}\right)} \tag{A9}$$

where erf is the error function. Under recurrent mutations, fixation is certain and we are interested in the time to ultimate fixation. It can be obtained using Kimura (1980). The infinitesimal mean change is now given by $M_{\delta x} = \Delta x + \mu(1-x)$ where the additional term corresponds to recurrent mutation. We then plug $M_{\delta x}$ into the Green function (A7) and the time to ultimate fixation under recurrent mutation is given by:

$$T_{fix} = \int_0^1 \int_0^x 4N_e \frac{G(z)}{(1-z)z} dz \frac{1}{G(x)} dx \tag{A10}$$

No close form solution exist for (A10) so numerical integration must be carried out (see *Mathematica* notebook). However, we can obtain an approximation as follows (see also Glémin and Ronfort 2013). The time to ultimate fixation can be decomposed into two parts: the waiting time for the appearance of the mutation destined to be fixed plus the time to fixation conditioned on the fact that fixation will occur. Because of underdominance the waiting time is expected to be much longer than the conditioned fixation time, which can be neglected. The time to

ultimate fixation can thus be approximated by:

$$T_{fix} \approx \frac{1}{2NuP_{fix}} \tag{A11}$$

For most parameters, (A11) is very accurate.

A1.3 Thresholds for near-neutrality

Assuming $x << 1$ in equation (A3) shows that a rare underdominant mutation behaves almost like a deleterious allele with a deleterious heterozygote effect $(1 - F)Sx_{eq}$ if $x_{eq} > 0$. For $x_{eq} < 0$ the mutation is positively selected for and can easily fix, which corresponds to the threshold $F \geq s_u/(s + s_u)$ (see above). Alternatively, for a given selfing rate, mutations with an heterozygote effects lower than the following threshold can easily fix:

$$s_u^{lim} = s\frac{F}{1 - F} \tag{A12}$$

This threshold vanishes to 0 when $s = 0$ and all mutants initially behave as deleterious.

When $s = 0$ we can consider a less stringent threshold as follows. When a mutant arises, $x << 0$ so:

$$\Delta x \approx -(1 - F)s_u x \tag{A13}$$

This is equivalent to negative genic selection with an effective selection coefficient $s_e = (1 - F)s_u$. We can consider that the mutation behaves almost neutraly when $2N_e s_e \leq 1$, which can be expressed as: $2N(1 - \sigma)s_u \leq 1$. So the nearly neutral

threshold is simply given by:

$$s_u^{nn} = \frac{1}{2N(1-\sigma)} \tag{A14}$$

A1.4 Distribution of deleterious effects

We now consider that the scaled heterozygote effects, $2Ns_u$ are not fixed but follow a gamma distribution with mean $\gamma = 2N\overline{s}_u$ and shape β, with *pdf* given by:

$$\phi(z) = \frac{\left(\frac{\gamma}{\beta}\right)^{-\beta} z^{\beta-1} e^{-\frac{\beta z}{\gamma}}}{\Gamma(\beta)} \tag{A15}$$

The proportion of symmetrical underdominant mutations ($s = 0$) that can fix (with a reasonable chance), ρ, is thus given by:

$$\begin{aligned} p(\sigma) &= \int_0^{1/2N(1-\sigma)} \phi(z)dz \\ &= 1 - \frac{\Gamma\left(\beta, \frac{\beta}{\gamma-\gamma\sigma}\right)}{\Gamma(\beta)} \end{aligned} \tag{A16}$$

Assuming that $\beta << \gamma$ and taking Taylor expansion of (A16) in β/ρ close to 0 we get:

$$p(\sigma) \approx \frac{\left(\frac{\beta}{\gamma(1-\sigma)}\right)^{\beta}}{\Gamma(\beta+1)} \tag{A17}$$

The ratio $\rho(\sigma) = p(\sigma)/p(0)$ gives the relative excess of the proportion of mutants that can fix compared to an outcrossing population:

$$\rho(\sigma) = \frac{\Gamma(\beta) - \Gamma\left(\beta, \frac{\beta}{\gamma-\gamma\sigma}\right)}{\Gamma(\beta) - \Gamma\left(\beta, \frac{\beta}{\gamma}\right)} \tag{A18}$$

Using the same approximation that $\beta << \gamma$ we obtain the very simple formula:

$$\rho(\sigma) \approx (1-\sigma)^{-\beta} \tag{A19}$$

Numerical and simulations results show that these two expressions also very accurately approximate the excess of probability of fixation (or reduction in fixation time) compared to an outcrossing population: $\frac{\int_0^\infty P_{fix}\phi(S)dS}{\int_0^\infty P_{fix}\phi(S)dS\Big|_{\sigma=0}}$ (or the inverse for fixation time).

A2 Compensatory mutations

A2.1 Model

We now consider a model of compensatory mutations at two loci with two alleles, where two haplotypes are equally fit, A_1B_1 (haplotype 1) and A_2B_2 (haplotype 4), but the intermediate paths, A_1B_2 (haplotype 2) and A_2B_1 (haplotype 3) are deleterious. Alike in the single underdominant model described above, the evolution of pairs of compensatory mutations requires to cross a fitness valley. For simplicity we consider a symmetrical model and we set the fitness of the genotypes as follows:

$$\begin{aligned} w_{11} &= w_{44} = 1 \\ w_{22} &= w_{33} = 1-s \\ w_{12} &= w_{13} = w_{24} = w_{34} = 1-hs \\ w_{14} &= w_{23} = 1-hks \end{aligned} \tag{A20}$$

where the subscript ij denotes the genotype formed with haplotypes i and j, and $s \geq 0$ and $0 \leq h \leq 1$ are respectively the strength and the coefficient of dominance of the deleterious effects of each mutation, and k is the coefficient of dominance for the double heterozygote genotype $A_1A_2B_1B_2$. In the main text we added the subscript c to these coefficient to distinguish with the coefficients in the BDMi model. Here we remove them to ease the reading.

We build upon previous haploid models (Kimura 1985; Stephan 1996) and extend it to diploidy with partial selfing. Previous works have shown that little recombination strongly prevent the fixation of compensatory mutations by breaking down double mutants. Under the assumption of weak recombination, we can only follows the four haplotypes and assume that genotype frequencies are obtained using multi-allelic single locus theory. We note X_i the frequency of haplotype i, and G_{ij} the frequency of genotype ij. This makes the system more tractable than the general system of ten equations presented in the main text. After meiosis, haplotype frequencies are given by:

$$X_i = G_{ii} + \frac{1}{2}\left(\sum_{j \neq i} G_{ij} + r\delta_i(G_{23} - G_{14})\right) \tag{A21}$$

where $\delta_i = 1$ for $i = 1, 4$ and $\delta_i = -1$ for $i = 2, 3$. We consider unidirectionnal

mutation from A_1 to A_2 and B_1 to B_2 at the same rate, u, so after mutation:

$$X_1^u = X_1(1-2u) \tag{A22a}$$

$$X_2^u = X_2(1-u) + X_1 u \tag{A22b}$$

$$X_3^u = X_3(1-u) + X_1 u \tag{A22c}$$

$$X_4^u = X_4 + (X_2 + X_3)u \tag{A22d}$$

After syngamy, we assume that genotype frequencies directly equilibrate to:

$$G_{ii}^r = (X_i^u)^2(1-F) + FX_i^u \tag{A23a}$$

$$G_{ij}^r = 2X_i^u X_j^u(1-F) \quad \text{for } i \neq j \tag{A23b}$$

And finally, after selection:

$$G'_{ij} = w_{ij}G_{ij}^r/\overline{W} \tag{A24}$$

where $\overline{W}$ is the mean fitness of the population. To simplify the system further, we can consider that intermediate haplotypes (A_1B_2 and A_2B_1) are maintained in approximate equilibrium at low frequency. This is true if $s >> u$ and $s >> r$. We also assume weak selection $s << 1$. Given the symmetry of the model $X_2 = X_3$ and are noted χ and we note $X_4 = x$, the frequency of the compensated haplotype, for which we want to calculate the probability and time to fixation. We can write:

$$\Delta\chi(\chi, x) = \Delta X_2 = \Delta X_3$$

$$\Delta x(\chi, x) = \Delta X_4$$

with the change of variables proposed above. We can use a separation of time

scale argument and consider that χ equilibrates much more rapidly than x.Thus we first solve $\Delta\chi(\chi, x) = 0$ for a given x and then plug the equilibrium χ value into $\Delta x(\chi, x)$. We thus obtain a an equation with a single variable that can be treated with classical diffusion theory. The full equation is not analytically tractable, however, noting that χ must be small, we can perform a Taylor expansion of $\Delta\chi(\chi, x)$ in χ at the first order and solve the resulting linear equation in χ. With the help of *Mathematica* we obtained:

$$\chi_{eq}(x) = \frac{(1-x)(u + x(1-F)(r - (h + F - hF)rs - hksu))}{s(h(1-2kx(1-x)) + F(1-h+2hkx(1-x)))} \tag{A25}$$

For $h = 0$ and $F = 0$ the first order term in χ vanishes so we need expansion at the second order, which gives:

$$\chi_{eq}(x) = \sqrt{\frac{(1-x)(rx+u)}{s}} \tag{A26}$$

Then we plug either (A25) or (A26) into $\Delta x(\chi, x)$. The full expression is rather cumbersome but it can be approximated as follows. As we assumed that all parameters are small: $u, s, r = O(\epsilon)$ with $\epsilon << 1$, we can only kept first order terms, which correspond to terms in s, r, u and u^2/s. With the help of *Mathematica* we obtained for $h > 0$ or $F > 0$:

$$\Delta x = \underbrace{2\frac{u^2}{(F + (1-F)h)s}(1-x)C_1}_{\text{Mutational input}} + \underbrace{x(1-x)\left(2uC_2 - (1-F)(khs(1-2x) + rC_3)\right)}_{\text{selection-like dynamics}} \tag{A27}$$

where C_1, C_2 and C_3 are expressions independent of s, r and u:

$$C_1 = \frac{1}{(1-2kx(1-x)))}$$
$$C_2 = 1 - \frac{(1-F)hk(1-2x^2)}{F+(1-F)h(1-2kx(1-x))}$$
$$C_3 = 1 - 2x\frac{(1-F)(F+(1-F)h(1+k(1-2x)))}{F+(1-F)h(1-2kx(1-x))}$$

The first term in equation (A27) corresponds to the input of the second mutation on the deleterious haplotypes, either A_1B_2 or A_2B_1 (hence the factor 2), which are both at mutation-selection balance equilibrium, $\frac{u}{(F+(1-F)h)s}$. The second term corresponds to selection-like dynamics of the form $Sx(1-x)$ where S has a complex form here. First, as the mean fitness of the population is of the order of $1-2u$ (see classical load theory, ex Charlesworth and Charlesworth (2010)), the fitness of the double mutant is simply of the order of $2u$ (but also depends on k and F). The second term corresponds to selection against double heterozygotes $(1-F)khs$ and breakdown of the double mutant by recombination $r(1-F)$. When $r=0$ and $k=0$, the double mutant A_2B_2 simply behaves as a beneficial mutations. On the contrary, the double mutant behaves as a deleterious mutations when recombination or selection overwhelm mutation:

$$(1-F)(khs+rC_3) > 2uC_2 \tag{A28}$$

So just a little recombination or selection against double heterozygotes greatly reduce the probability of fixation of the double mutant. From (A28) it is also clear that selfing increases the conditions of fixation of the double mutant.

A2.2 Probability and time to fixation

Equation (A27) can be injected in a classical one dimensional diffusion equation and numerically solved to obtain the time to fixation as in Kimura (1980). There is no analytical solution to the full equation but we can obtained a rather simple analytical approximation as follows. As in the main text, the time until ultimate fixation can be decomposed into the waiting time of the mutation destined to fixate and the time to fixation, conditional to fixation. The first term is usually much larger than the first one so we can only consider the waiting time and we can apply diffusion theory using (A7) and (A8) and modified version of (A11):

$$T_{fix} \approx \frac{1}{4Nu\chi_{eq}P_{fix}} \tag{A29}$$

because we only consider mutation arising on deleterious haplotypes, whose number is $2N\chi_{eq}$ in the population.

When $r = 0$ and $k = 0$, $C_1 = C_2 = 1$ and the selection-like term reduces to $2u$. So we have:

$$T_{0,0} \approx \frac{(F + h - hF)s}{2u^2} \frac{1 - e^{-8Nu/(1+F)}}{8Nu/(1 + F)} \tag{A30}$$

which reduces to:

$$T^*_{0,0} \approx \frac{(F + h - hF)s}{2u^2} \tag{A31}$$

when $4Nu < 1$ as given in the main text.

When $k = 0$ but $r > 0$, we still have $C_1 = C_2 = 1$ and $C_3 = 1 - 2x$ if $r > u$ we can neglect the mutation term so we obtain a simple selection like term:

$r(1-F)x(1-x)(1-2x)$. This leads to the following solution:

$$T_{r,0} \approx \frac{(F+h-hF)s}{2u^2} \frac{1}{N\left(1-\frac{Erf\left((1-1/N)\sqrt{R}\right)}{Erf\left(\sqrt{R}\right)}\right)} \tag{A32}$$

with $R = N(1-\sigma)r$. Similarly, when $r = 0$ and $k > 0$ the selection-like term takes the same form, $hks(1-F)x(1-x)(1-2x)$ hence the same result as equation (A32) with $R = N(1-\sigma)hks$. This illustrates that recombination and selection against double heterozygotes play the same role. The general equation is more difficult to solve but as recombination and selection have the same form, so a heuristic argument is to use equation (A32) with $R = N(1-\sigma)(r+hks)$. Finally, to simplify the expression we can take the limit o (A32) when $N \to \infty$ (but $R \to cte$), which leads to:

$$T_{r,k} \approx \frac{(F+h-hF)s}{2u^2} \frac{\sqrt{\pi}e^R Erf(\sqrt{R})}{2\sqrt{R}}$$
$$\text{with} \quad R = N(1-\sigma)(r+hks) \tag{A33}$$

as given in the main text. Simulations show that this general approximation is rather accurate and allows a clear interpretation of the effect of recombination, selection against heterozygotes and selfing. It is important to note that these results are valid when effective recombination is low ($r(1-F)$). In the main text, simulations show that they quantitatively breakdown when recombination is too high. However these approximations are useful to characterize the effect of selfing on the fixation of compensatory mutations.

References

Caballero, A. and W. G. Hill, 1992 Effects of partial inbreeding on fixation rates and variation of mutant genes. Genetics **131**: 493–507.

Charlesworth, B. and D. Charlesworth, 2010 *Elements of Evolutionary Genetics*.

Glémin, S. and J. Ronfort, 2013 Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. Evolution: International Journal of Organic Evolution **67**: 225–240.

Kimura, M., 1962 On the probability of fixation of mutant genes in a population. Genetics **47**: 713–9.

Kimura, M., 1980 Average time until fixation of a mutant allele in a finite population under continued mutation pressure: Studies by analytical, numerical, and pseudo-sampling methods. Proceedings of the National Academy of Sciences **77**: 522–526.

Kimura, M., 1985 The role of compensatory neutral mutations in molecular evolution. J. Genetics **64**: 7–19.

Pollak, E., 1987 On the theory of partially inbreeding finite populations. i. partial selfing. Genetics **117**: 353–60.

Stephan, W., 1996 The rate of compensatory evolution. Genetics **144**: 419–426.