Balancing selection in self-fertilizing populations

Sylvain Glémin[1,2,*]

[1]University of Rennes, CNRS, ECOBIO [(Ecosystèmes, biodiversité, évolution)] - UMR 6553, Rennes, France

[2]Department of Ecology and Genetics, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden

[*]Corresponding author: sylvain.glemin@univ-rennes1.fr

Running head: balancing selection with selfing

1

Abstract

Self-fertilization commonly occurs in hermaphroditic species, either occasionally or as the main reproductive mode. It strongly affects the genetic functioning of a population by increasing homozygozity and genetic drift and reducing the effectiveness of recombination. Balancing selection is a form of selection that maintains polymorphism, which has been extensively studied in outcrossing species. Yet, despite recent developments, the analysis of balancing selection in partially selfing species is limited to specific cases and a general treatment is still lacking. In particular, it is unclear whether selfing globally reduced the efficacy of balancing selection as in the well-known case of overdominance. I provide a unifying framework, quantify how selfing affects the maintenance of polymorphism and the efficacy of the different form of balancing selection, and show that they can be classified into two main categories: overdominance-like selection (including true overdominance, selection variable in space and time, and antagonistic selection), which is strongly affected by selfing, and negative frequency-dependent selection, which is barely affected by selfing, even at multiple loci. I also provide simple analytical results for all cases under the assumption of weak selection. This framework provides theoretical background to analyze the genomic signature of balancing selection in partially selfing species. It also sheds new light on the evolution of selfing species, including the evolution of selfing syndrome, the interaction with pathogens, and the evolutionary fate of selfing lineages.

1 Introduction

Self-fertilization is common in hermaphroditic organisms. In flowering plants, up to 50% of species could be capable of self-fertilization to some extant (Igic and Kohn, 2006) and predominant selfing mating systems recurrently evolve from outcrossing (For example, up to 40 independent transitions in the Polemoniaceae family, Landis et al., 2018). High or partial selfing is also observed in many hermaphrodite animals (Jarne and Auld, 2006), fungi (Billiard et al., 2012) and algae (Hanschen et al., 2018). Even among obligatory outcrossers such as self-incompatible plants, selfing can incidentally occur as a results of self-incompatibility breakdown (Stone, 2002).

Selfing strongly affects the genetic functioning of a population, especially by increasing homozygosity, genetic drift and genetic linkage (Burgarella and Glémin, 2017). Because the two uniting gametes come from the same parent, homozygote offspring increases every generation, which enhances selection on recessive alleles: selfing helps purge strongly deleterious recessive alleles and favors adaptation from new recessive beneficial mutations. In reverse, non-independent gamete sampling enhanced genetic drift (by a factor two under full selfing, Pollak, 1987), which reduces the efficacy of selection. Therefore, the final outcome of selection depends on the balance between drift and homozygosity effects. For example, a classical result is that, considering a single locus, selfing increases the probability of fixation of partially recessive beneficial mutations but decreases it for partially dominant ones, with no effect of selfing for exactly co-dominant mutations (Caballero and Hill, 1992; Charlesworth, 1992). However, selfing also strongly enhances genetic linkage: although crossing-overs do occur in selfers, they mostly exchange identical genomic fragments (due to the high homozygosity), which has no effect on genetic mixing. As a consequence, selfing reduces the efficacy of selection far beyond the predictions of single-locus models (Roze, 2016; Kamran-Disfani and Agrawal, 2014; Hartfield and Glémin, 2014, 2016). Reduced adaptive potential and accumulation of weakly deleterious mutations genome wide are thus expected in highly selfing species. These predictions have been supported by various comparative population genomic studies: selfing species usually exhibit lower genetic diversity, higher linkage disequilibrium and lower selection efficacy

when assessed through the comparison of non-synonymous and synonymous polymorphism (but much less when divergence between species is compared) (see Glémin et al., 2019 for a recent review). On the long term, these effects could contribute to increasing extinction rate of selfing lineages, making self-fertilization an evolutionary dead-end reproductive strategy (Stebbins, 1957; Igic and Busch, 2013).

In contrast to purifying and positive selection, balancing selection is a form of selection that maintains polymorphism (Delph and Kelly, 2014). After an initial appeal as a possible mechanism to explain observed polymorphism patterns and the maintenance of variation for fitness, balancing selection has long been considered as relatively anecdotal (Charlesworth and Hughes, 1999). Theoretical models have shown that conditions for maintenance of balancing selection (at least some forms, see below) are rather limited and it has proved difficult to identify robust examples of balancing selection, in contrast to the widespread evidence for directional selection. However, regain of interest for balancing selection has come from both new theoretical studies and the accumulation of convincing empirical examples (Delph and Kelly, 2014; Hedrick, 2012). Balancing selection can take different forms. Polymorphism can be maintained through direct heterozygote advantage (true overdominance). More generally, overdominance can emerge from antagonistic pleiotropy between fitness components throughout the life cycle, between sexes, or even among generations when selection varies in time. If so, dominance reversal, where an allele dominant for a component becomes recessive for another one, makes conditions for the emergence of overdominance and the maintenance of balanced polymorphism easier (Curtsinger et al., 1994). Although initially thought to be rather restrictive, this condition can actually be easily generated for realistic life cycles (Brown and Kelly, 2018; Connallon and Chenoweth, 2019) and it has been observed experimentally in recent studies (Gemmell and Slate, 2006; Grieshop and Arnqvist, 2018; Kellenberger et al., 2019). Negative frequency-dependence is another form of balancing selection where rare alleles or genotypes are selected for (Delph and Kelly, 2014), as at the self-incompatibility locus of flowering plants (Charlesworth et al., 2005). It could also be a general form of selection for genes involved in defense against pathogens (Bergelson et al., 2001; Tiffin and Moeller,

3

2006) and in some pollinator-mediated floral polymorphisms (Gigord et al., 2001).

Selfing can affect the strength and outcome of balancing selection. In addition to the effects mentioned above, selfing reduces the role of heterozygotes' fitness in the dynamics of allelic frequencies and alters selection balance on the different fitness components throughout the life cycle (e.g.: male vs female traits, Damgaard, 2000). Selfing has been shown to globally reduce selection in a few well-studied cases, such as overdominance (ex: Kimura and Ohta, 1971) and sex-antagonistic selection (ex: Jordan and Connallon, 2014). However, it is not clear if those conclusions can be extended to any kind of balancing selection. Assessing the importance of the different forms of balancing selection in (partially) selfing species is needed to better understand how variation in fitness is maintained or lost compared to outcrossing species and to properly interpret the associated genomic patterns. It is also important to understand the evolution of mating systems. For example, changes in the dynamics of antagonistic pleiotropy (especially male/female antagonism) can affect trait evolution associated with the so-called selfing syndrome - the set of traits associated with selfing, such as the reduction in resource allocation to male function (Sicard and Lenhard, 2011). It has also been proposed that co-evolution with pathogens could influence mating system transitions by promoting recombination and outcrossing (Levin, 1975; Agrawal and Lively, 2001) or that inefficient response to pathogen pressures could contribute to higher extinction risk in selfers (Wright et al., 2013). These predictions should depend on the way selfing affects polymorphism at genes involved in host-pathogen interactions.

Some theoretical results on the effect of selfing on the maintenance of balanced polymorphism have already been obtained for specific cases. However, important cases has not been addressed yet, and a general comparison among models is also lacking. In the present study, I use a single framework to synthesize and generalize previous results - multiallelic overdominance, variable selection in space and time - and present new ones - negative frequency-dependent selection, various forms of antagonistic selection. I quantify how selfing affects the maintenance of polymorphism and the efficacy of selection. Using weak selection approximations, I show that the different forms of balancing se-

4

118 lection can be classified in two main categories: overdominance-like selection, which is strongly affected by selfing, and negative frequency-dependent selection, which is barely

120 affected by selfing. These results shed new light on the evolution of selfing species and may help analyze and interpret genomic patterns in such species.

2 Models

2.1 General equations

124 Consider a single hermaphroditic diploid population of size N, with non-overlapping generations and reproducing partially through selfing at a fixed rate σ. Except when

126 specified, I consider single locus models with $n \geq 2$ alleles, noted A_i with $i \in \{1, ..., n\}$. The general life cycle is as follows, starting from adult individuals (Figure 1): first diploid

128 fecundity selection occurs for production (through meiosis) of all female gametophytes and of exported male gametophytes, *i.e;* involved in outcrossing; if we assume no male

130 gamete limitation there is no male diploid fecundity selection for the selfing fraction because only one male parent contributes to the next generation (Lloyd, 1979). Then,

132 haploid selection occurs on both female and male gametophytes before producing gametes, followed by fertilization. Note that for zygotes produced by selfing, male haploid selection

134 can only occur in heterozygote parents. After zygote formation, diploid viability selection occurs, then population regulation to N adults. Here, all stages of a haplo-diploid life

136 cycle are described but the model also applies for animals where meiosis directly produces gametes and where haploid selection operates on gamete viability before fertilization. The

138 list of notations used in the derivation of the followng equationsa regiven in Table 3.

Figure 1: General life cycle. The diploid phase is represented by bold-lined ovals and the haploid phase by thin-lined circles. σ is the probability of selfing.

The expected (deterministic) changes in genotype frequencies in one generation is ob-

140 tained as follows. Noting X_{ij} the frequency of genotype A_iA_j, where alleles are unordered $(X_{ij} = X_{ji})$, and x_i the frequency of allele A_i in adults, one has $x_i = X_{ii} + \frac{1}{2}\sum_{j \neq i} X_{ij}$.

5

142 I also note W_{ij}^F and W_{ij}^M, the number of female and male gametophytes produced by adults A_iA_j, w_i^F and w_i^M, the number of female and male gametes produced by gameto-

144 phytes A_i, and V_{ij}, the survival probability of zygote A_iA_j to adulthood. The frequency of female (superscript F) gametophytes A_i produced after diploid selection is:

$$x_i^F = \left(X_{ii}W_{ii}^F + \frac{1}{2}\sum_{j\neq i} X_{ij}W_{ij}^F \right) \Big/ \overline{W}^F \tag{1}$$

146 where $\overline{W}^F = \sum_i \sum_{j\leq i} X_{ij}W_{ij}^F$. Similarly, the frequency of male (superscript M) gameto-phytes A_i involved in outcrossing produced after diploid selection is:

$$x_i^M = \left(X_{ii}W_{ii}^M + \frac{1}{2}\sum_{j\neq i} X_{ij}W_{ij}^M \right) \Big/ \overline{W}^M \tag{2}$$

148 where $\overline{W}^M = \sum_i \sum_{j\leq i} X_{ij}W_{ij}^M$. After haploid selection, the frequencies of female gametes and of male gametes contributing to outcrossing are, respectively:

$$o_i = x_i^F w_i^F / \overline{w}^F \tag{3}$$

150 and

$$p_i = x_i^M w_i^M / \overline{w}^M \tag{4}$$

where $\overline{w}^F = \sum_i x_i^F w_i^F$ and $\overline{w}^M = \sum_i x_i^M w_i^M$. After reproduction (superscript R), geno-

152 type frequencies are given by:

$$X_{ii}^R = (1-\sigma)o_i p_i + \sigma\left(X_{ii} \underbrace{\frac{w_i^F}{\overline{w}^F}}_{\text{hap. } \female \text{ sel.}} \overbrace{\frac{W_{ii}^F}{\overline{W}^F}}^{\text{dip. } \female \text{ sel.}} + \underbrace{\frac{w_i^F}{\overline{w}^F}}_{\text{hap. } \female \text{ sel.}} \sum_{j\neq i}\frac{X_{ij}}{2} \overbrace{\frac{W_{ij}^F}{\overline{W}^F}}^{\text{dip. } \female \text{ sel.}} \underbrace{\frac{w_i^M}{w_i^M + w_j^M}}_{\text{hap. } \male \text{ sel.}} \right)$$

$$X_{ij}^R = (1-\sigma)(o_i p_j + o_j p_i) + \sigma\frac{X_{ij}}{2}\frac{W_{ij}^F}{\overline{W}^F}\left(\frac{w_i^F}{\overline{w}^F}\frac{w_j^M}{w_i^M + w_j^M} + \frac{w_j^F}{\overline{w}^F}\frac{w_i^M}{w_i^M + w_j^M} \right) \tag{5}$$

The first terms correspond to random mating, with probability $1-\sigma$ and the second ones

154 to selfing, with probability σ. For the outcrossing part, diploid fecundity and haploid

6

selection are directly encapsulated within the o_i and p_i terms. For the selfing part,

156 the term in parenthesis includes female fecundity selection, W_{ij}^F, (no male fecundity

selection under selfing), and male haploid selection acting only in heterozygotes, where

158 $\frac{w_i^M}{w_i^M + w_j^M}$ is the relative fitness of male gametes A_i among the "population" of male gametes

produced by a heterozygous individual $A_i A_j$ (that is one half of A_i and one half of

160 A_j). So, without male haploid selection this term reduces to $1/2$ and we get back the

classical $1/4$ term for the production of a homozygote from the selfing of a heterozygote.

162 Female haploid selection on A_i allele occurs for all individuals. For simplicity I do not

consider competing selfing ("mass-action" mating, see Holsinger, 1991) where selfing rate

164 depends on genotypes if there is fecundity and/or haploid selection. However, as already

explored for sexually antagonistic selection, the difference between mass-action mating

166 and fixed selfing is rather weak and only quantitative (Jordan and Connallon, 2014). Note

also that the effect of inbreeding depression that reduces the fitness of selfed individuals

168 (independently of their genotype at the focal locus) can simply be included in the model

by considering that σ represents the "effective" selfing rate that is the proportion of

170 adults issued from selfing, instead of the primary selfing rate (Jordan and Connallon,

2014). Finally, after reproduction and viability selection, the allelic frequencies in adults

172 at the next generation are given by :

$$x_i' = \left(X_{ii}^R V_{ii} + \frac{1}{2} \sum_{j \neq i} X_{ij}^R V_{ij} \right) \bigg/ \overline{V} \tag{6}$$

where $\overline{V} = \sum_i \sum_{j \leq i} X_{ij}^R V_{ij}$. While mean fitness is computed at several steps in the

174 recursion, including for male fitness components, what mainly matters for population

demography is viability and female fecundity. So, neglecting male fitness components,

176 the mean "demographic" fitness of the population is:

$$\mathfrak{W} = \overline{V} \, \overline{w}^F \overline{W}^F \tag{7}$$

7

and the genetic load is simply given by:

$$L = 1 - \mathfrak{W} \qquad (8)$$

Table 1: Glossary of main notations

Symbol	Usage
N	Population size
σ	Selfing rate
F	Wright's fixation index, probability of identity at one locus.
	$F = \sigma/(2 - \sigma)$ at steady state
Φ	Probability of identity at two loci
r	Recombination rate
W_{ij}^F, W_{ij}^M	Female, Male fecundity of genotype $A_i A_j$
V_{ij}	Viability of genotype $A_i A_j$
w_i^F, w_i^M	Female, Male fitness of gametophyte A_i
$\overline{W}, \overline{V}, \overline{w}$	Corresponding mean fitnesses
$\mathfrak{W}$	Overall mean fitness of the population
L	Genetic load. $L = 1 - \mathfrak{W}$
n	Number of alleles
s_k	Diploid selection coefficients, depending on the model (see each part)
s_H	Harmonic mean of selection coefficients
h_k	Diploid dominance coefficients, depending on the model (see each part)
$s_{k,i}$	Diploid selection coefficients in niche i
t_k	Haploid selection coefficients, depending on the model (see each part)
b_f, b_m	Female, male distortion bias
X_{ij}	Frequency of genotype A_{ij} in adults
X_{ij}^R	Frequency of genotype A_{ij} in zygotes
x_i	Frequency of allele A_i in adults
x_i^F, x_i^M	Frequency of allele A_i in female, male gametophytes
o_i, p_i	Frequency of allele A_i in female gametes, in exported male gametes
$\Omega(\mathbf{z})$	Efficacy of selection at point $\mathbf{z}$, where $\mathbf{z}$ is a vector of $n-1$ frequencies

2.2 Approximations for weak selection

The full model can be described by a set of difference equations on the $n(n + 1)/2$ genotype frequencies as described above. Alternatively, after variable transformation, one can write equations for the n alleles frequencies plus the $n(n - 1)$ inbreeding coefficients, F_{ij}, measuring departure from frequencies expected under random mating. Under weak selection one can use a separation of time-scale argument to reduce the system to

8

184 only $n-1$ independent allelic frequencies. F_{ij} equilibrate in very few generations (Weir and Cockerham, 1973), much more quickly than changes in allelic frequency, so that

186 equilibrium F_{ij} values can be used (for a detailed treatment with two alleles see Pollak and Sabran, 1992). These argument and approximation also hold at equilibrium for any

188 strength of selection because allelic frequencies do not change. In addition, due to the weak selection assumption, all F_{ij} can be approximated by the neutral expectation (Weir,

190 1970):

$$F = \frac{\sigma}{2-\sigma} \tag{9}$$

To obtain difference equations on allele frequencies only, one can thus replace genotype

192 frequencies at the previous generation by their expected equilibrium frequencies given x_i and F:

$$
\begin{aligned}
X_{ii} &= x_i^2(1-F) + x_i F \\
X_{ij} &= 2x_i x_j(1-F)
\end{aligned}
\tag{10}
$$

194 For n alleles, noting $\mathbf{x} = \{x_1, x_2, ...x_{n-1}\}$ with $x_n = 1 - \sum_{i=1}^{n-1} x_i$, the general recursion after one generation is thus given by: $\mathbf{x}' = \mathbf{f}(\mathbf{x})$ where $f_i(\mathbf{x}) = x_i'(\mathbf{x})$ is obtained by plugging

196 (10) in equation (6). Equivalently, the change in allelic frequencies in one generation is given by $\mathbf{\Delta x} = \mathbf{f}(\mathbf{x}) - \mathbf{x}$.

198 ## 2.3 Measuring the effect of selfing

To assess the effect of selfing on the maintenance of polymorphism I consider (i) the range of conditions allowing the maintenance of a stable polymorphism, and (ii) the efficacy of selection, relative to drift, on the maintenance of internal equilibrium or on the emergence of new alleles. The dynamics of the system around a focal point, $\mathbf{z}$ (a vector of $n-1$ frequencies), is driven by the leading eigenvalue (that is with the highest absolute value) of the Jacobian matrix of $\mathbf{f}(\mathbf{x})$ evaluated at this point: $\lambda_1(\mathbf{z})$. For two alleles it reduces to $\lambda_1(z) = 1 + \frac{\partial \Delta x}{\partial x}\big|_{x=z}$ (see Figure 2 for illustration). An equilibrium point, is locally stable if $|\lambda_1(\mathbf{z})| < 1$ (Otto and Day, 2007). The efficacy of selection on

9

the emergence of a rare allele can be measured as the asymptotic rate of spread when the frequency tends towards zero, which is given by $\lambda_1(0) - 1$ (Otto and Day, 2007). Similarly, near the deterministic equilibrium, the efficacy of selection can be defined as the rate at which it moves allelic frequencies back to the equilibrium, which is given by $1 - \lambda_1(\mathbf{x_{eq}})$ (the opposite sign is taken because $\lambda_1 < 1$ when the equilibrium is stable, see above). A natural measure of the efficacy of selection relative to genetic drift, is the

Figure 2: Δx as a function of x for the classical two-allele overdominant model in a random mating population: $\Delta x = x(1 - x)(s_2(1 - x) - s_1 x)/\overline{W}$, where $\overline{W}$ is the mean fitness of the population. It illustrates how to compare the dynamics of more or less balanced forms of selection. Coefficients of selection against homozygotes: for the upper curve $s_1 = 0.005$ and $s_2 = 0.015$. For the lower curve $s_1 = 0.02$ and $s_2 = 0.015$. The slope of the lines give the strength of selection in $x = 0$, dashed line, in $1 - x = 0$, dotted lines, and at equilibrium, plain lines. In this example, the strength of selection is the same for the two cases in $x = 0$ but stronger for the lower curve at equilibrium and in $1 - x = 0$.

product $\pm N_e(\lambda_1 - 1)$, where $N_e = N/(1 + F)$ is the effective population size. Because I am interested in the relative effect of selfing, I can drop out the N term. I define the efficacy of selection relative to drift on the maintenance of internal equilibrium, Ω_{eq}, and on the emergence of a new $n + 1^{th}$ allele when the n others are previously at equilibrium, Ω_0, as:

$$\Omega_{eq} = -\frac{\lambda_1(\mathbf{x_{eq}}) - 1}{1 + F} \tag{11a}$$

and

$$\Omega_0 = \frac{\lambda_1(\{\mathbf{x_{eq}}, 0\}) - 1}{1 + F} \tag{11b}$$

which reduces to:

$$\Omega_{eq} = -\frac{1}{1 + F} \frac{\partial \Delta x}{\partial x}\bigg|_{x=x_{eq}} \tag{12a}$$

and

$$\Omega_0 = \frac{1}{1 + F} \frac{\partial \Delta x}{\partial x}\bigg|_{x=0} \tag{12b}$$

10

Because of linked selection, N_e can be reduced further in selfing species (Agrawal and
Hartfield, 2016; Roze, 2016), so one can always find arbitrary conditions for which selfing
decreases the efficacy of selection. Equation (12) tells whether selfing automatically
lessens selection or whether additional reduction in N_e is required for selection to be less
efficient under selfing than under outcrossing.

3   Results

For completeness, previous results are given in addition to new ones. For clarity, only
main results are given in the main text and details of derivations can been found in the
appendices.

3.1   Overdominance

Overdominance is the classical illustrative example of balancing selection. It will
serve as a reference model as I show below that many other models reduce to (some form
of) overdominance after appropriate rescaling. The initial result for a bi-allelic locus
was obtained by Hayman (1953) (see also Appendix A4 in Kimura and Ohta, 1971) and
generic equations for multi-allelic selection were given by Weir (1970) but he did not
provide any simple analytical approximation.

Consider a single locus with n alleles, A_1 to A_n for which all heterozygotes $A_i A_j$ have
the same fitness, 1, and homozygotes $A_i A_i$ the fitness $1 - s_i$. Without loss of generality
I label alleles such that $1 > s_1 \geq s_2 \geq ... \geq s_n > 0$. For a bi-allelic locus, noting x the
frequency of allele A_1, the deterministic change in allele frequency is:

$$\Delta x = x(1-x)\big((1-F)(s_2 - x(s_1 + s_2)) + F(s_2 - s_1)\big)/\overline{W} \tag{13}$$

where $\overline{W} = 1 - s_1\big(x^2 + Fx(1-x)\big) - s_2\big((1-x)^2 + Fx(1-x)\big)$. If $F = 1$, it reduces to
genic selection with $\Delta x = (s_2 - s_1)x(1-x)/\overline{W}$. Otherwise it can be written as:

$$\Delta x = (1-F)(s_1 + s_2)x(1-x)(x_{eq} - x)/\overline{W} \tag{14}$$

11

where, x_{eq}, the internal equilibrium frequency of allele A_1 is:

$$x_{eq} = \frac{s_2 - Fs_1}{(1 - F)(s_1 + s_2)} \tag{15}$$

The condition of existence $(0 \leq x \leq 1)$ and stability of this equilibrium is the same (Rocheleau and Lessard, 2000) and is given by: $F \leq s_2/s_1$. Another interesting way to write the stability condition is: $\frac{s_1 - s_2}{s_1 + s_2} \geq 1 - \sigma$, which means that the relative asymmetry between the two selection coefficients (the left-hand term) must be lower than the out-crossing rate. So, increasing selfing rate limits conditions for polymorphism to more and more symmetrical cases. While this derivation has assumed weak selection, others have analyzed the case with arbitrary selection. Equations (13) to (15) are still valid but F depends both on selfing rate and selection. They showed that for $s_1, s_2 \geq 1/2$, an internal equilibrium exist and is stable for any selfing rate (Kimura and Ohta, 1971; Rocheleau and Lessard, 2000). This is possible even with 100% selfing because selection removed homozygotes more rapidly that they are created by selfing. For this model the efficacy of selection is:

$$\Omega_{eq} = \frac{(s_2 - Fs_1)(s_1 - Fs_2)}{(1 - F)(1 + F)(s_1 + s_2)} \tag{16a}$$

$$\Omega_0 = \frac{s_2 - Fs_1}{1 + F} \quad \text{for allele } A_1$$

$$= \frac{s_1 - Fs_2}{1 + F} \quad \text{for allele } A_2 \tag{16b}$$

222 These equations show that efficacy of selection around equilibrium is stronger when more symmetrical and that selfing always reduces it, including when $s_1 = s_2$, which also con-
224 tributes to reducing the likelihood of maintaining polymorphism.

These results generalize to n alleles with:

$$\Delta x_i = \frac{x_i\big((1 - F)(J_2 - s_i x_i) + F(J_1 - s_i)\big)}{1 - (1 - F)J_2 - FJ_1} \tag{17}$$

12

where $J_1 = \sum_{i=1}^{n} s_i x_i$ and $J_2 = \sum_{i=1}^{n} s_i x_i^2$. In Appendix 1 I show that, for $F < 1$:

$$x_i^{eq} = \frac{(1-F)s_H + Fn(s_H - s_i)}{(1-F)ns_i} \tag{18}$$

where s_H is the harmonic mean of selection coefficients over the n alleles. Recalling that $s_1 = \max(s_i)$, the internal equilibrium with n alleles exists and is stable if [a proof of stability is given in Appendix 1 for the weak selection limit]:

$$F \leq \frac{s_H}{ns_1 - (n-1)s_H} \tag{19}$$

or alternatively:

$$s_1 \leq \frac{(1 + F(n-1))s_H}{Fn} \tag{20}$$

For $F = 1$ this reduces to $s_1 \leq s_H$ so that only the purely symmetrical case, where all s_i are equal, exists and is stable. Condition (20), with s_{n+1} instead of s_1, also corresponds to the condition of invasion of a new $(n+1)^{th}$ allele in a population previously at equilibrium with n alleles (Appendix 1). This condition means that an allele with a weaker negative homozygote effect can enter the population, which reduces s_H and can then lead to the lost of older alleles with stronger effect if they no longer satisfy condition (20). Partial selfing thus reduces the conditions for maintenance of polymorphism but also leads to a kind of purging of alleles with strong deleterious homozygote effects. However, despite this purging, the mean fitness of the population decreases with selfing rate until a plateau when all but the weakest allele have been lost (Figure 3). Note that in a finite population, the expected number of alleles would be lower because of drift, and the load higher.

Figure 3: Maximum (deterministic) number of overdominant alleles that can be maintained in a population (plain lines) and genetic load ($L = 1 - \mathfrak{W}$, dashed lines). The number of alleles is determined by recursively eliminating alleles with the most deleterious homozygote effect until criterion (20) is verified, using $F = \sigma/(2 - \sigma)$ in (20). In the three panels, the lowest coefficient is 0.04 and the nine others are incremented with steps: (A) 0.001 (so from 0.04 to 0.049), (B) 0.005 (so from 0.04 to 0.085), (C) and 0.02 (so from 0.04 to 0.22). As selection increases, from top to bottom, alleles are lost for lower level of selfing.

13

With n alleles, the efficacy of selection is approximately (see Appendix 1):

$$\Omega_{eq} \approx \frac{((1 + F(n-1))s_H - Fns_1)(s_1 - Fs_n)}{(1+F)(1-F)ns_1} \quad \text{with } s_H \text{ computed over } n \text{ alleles} \quad \text{(21a)}$$

$$\Omega_0 \approx \frac{(1-F)s_H - nF(s_{n+1} - s_H)}{n(1+F)} \quad \text{with } s_H \text{ computed over } n+1 \text{ alleles} \quad \text{(21b)}$$

with F satisfying (19).

3.2 Variable selection

3.2.1 Selection variable in space

Since the seminal work of Levene (1953), selection in heterogeneous environments is recognized has an important mechanism maintaining polymorphism. For example, in the tropical tree *Eperua falcata* several genes potentially involved in stress response related to edaphic conditions exhibit strong allelic divergence within a population associated with a mosaic of micro-environments at the scale of hundreds of meters (Audigeos et al., 2013). The effects of selfing on polymorphism maintained by local adaptation is complex and left for future works. Here, I simply extend the original two-allele Levene's model to partially selfing species. Using the same approach, Hedrick (1998) proposed that spatially heterogeneous selection is an effective way to maintain polymorphism in selfing species. However, he only considered rather strong and symmetrical selection.

Following Levene (1953), a partially selfing population reproduces as a single unit, then offspring grows in K different niches where viability selection occurs. Niche i is found in proportion c_i and the fitness of the three genotypes A_1A_1, A_1A_2, and A_2A_2 are $1 - s_{1,i}$, 1, and $1 - s_{2,i}$, respectively, where $s_{1,i}$ and $s_{2,i}$ can be positive or negative but of opposite sign to get directional selection in each niche. After selection, each niche contributes a proportion c_i of reproducing adults. Under weak selection, this model is equivalent to the two-allele overdominant model with the following rescaling (Appendix 3):

$$s_1 = \sum_{i=1}^{K} c_i s_{1,i} \text{ and } s_2 = \sum_{i=1}^{K} c_i s_{2,i} \tag{22}$$

14

where both s_1 and s_2 need to be positive. Using this rescaling the results presented
above for overdominance apply. Contrary to the conclusion drawn by Hedrick (1998),
polymorphism due to heterogeneous selection is difficult to maintain in selfing species,
except when selection is strong.

3.2.2 Selection variable in time

Temporal variation in selection intensity can be common, especially for traits directly affected by environmental conditions in sessile organisms, as in the flowering plant *Digitalis purpurea* where strong shift in reproductive optimum was observed between consecutive years (Sletvold and Grindeland, 2007). Such fluctuating selection can maintain polymorphism if the geometric mean fitness of the heterozygote through time is higher than for homozygotes (Haldane and Jayakar, 1963). In their seminal paper, Haldane and Jayakar (1963) already noted that "inbreeding makes the conditions for polymorphism more stringent." However, I am not aware of any formal result. This can readily be obtained as shown in Appendix 3. Noting $V_{AA}(t)$, $V_{Aa}(t)$ and $V_{aa}(t)$ the fitness of the three genotypes at time t, the condition for protected polymorphism is (Appendix 3):

$$\prod_t \left((1-F)V_{Aa}(t) + FV_{AA}(t) \right) \geq \prod_t V_{aa}(t)$$

and (23)

$$\prod_t \left((1-F)V_{Aa}(t) + FV_{aa}(t) \right) \geq \prod_t V_{AA}(t)$$

In equations (23), the terms in F corresponds to contradictory conditions ($\prod_t V_{AA}(t) \geq \prod_t V_{aa}(t)$ and $\prod_t V_{aa}(t) \geq \prod_t V_{AA}(t)$), so that increasing F decreases the range of compatible conditions (23). For $F = 1$, condition (23) reduces to the haploid case where no polymorphism can be maintained (Haldane and Jayakar, 1963).

3.3 Antagonistic pleiotropy

The importance of true overdominance is debated. Although it is supposed to be rare, it can also emerge from common forms of antagonistic pleiotropy between life-history

15

traits, an allele being beneficial for a trait but detrimental for another one (Curtsinger
276 et al., 1994; Brown and Kelly, 2018). In such models, the lifetime fitness is obtained by
multiplying the fitness of each component, so a necessary condition for polymorphism
278 is that the geometric mean fitness across the life cycle is higher for heterozygotes than
for homozygotes (see for example Hedrick, 1999). If the two traits affect equally the
280 two sexes at the diploid stage, antagonistic pleiotropy can generate apparent or emergent
overdominance (Delph and Kelly, 2014; Brown and Kelly, 2018). If selection is sexually
282 and/or ploidy antagonistic, more complex dynamics is expected. However, under weak
selection, there is still a formal equivalence with overdominance with an appropriate re-
284 scaling of selection coefficients s_1 and s_2, but they may depend on the selfing rate. Thus,
the precise effect of selfing depends on the kind of antagonism.

Table 2: Fitness in antagonistic models

A: Viability/fecundity antagonism

	AA	Aa	aa
Male fecundity	1	$1 - h_m s_m$	$1 - s_m$
Female fecundity	1	$1 - h_f s_f$	$1 - s_f$
Viability	$1 - s$	$1 - hs$	1

B: Diploid sexual antagonism

	AA	Aa	aa
Male fecundity	1	$1 - h_m s_m$	$1 - s_m$
Female fecundity	$1 - s_f$	$1 - h_f s_f$	1

C: Haploid sexual antagonism

	A	a
Male gametophyte	$1 - t_m$	1
Female gametophyte	1	$1 - t_f$

D: Ploidy antagonisms

	AA	Aa	aa
Male fecundity	$1 - s_m$	$1 - h_m s_m$	1
Female fecundity	$1 - s_f$	$1 - h_f s_f$	1
Viability	$1 - s$	$1 - hs$	1

	A	a
Male gametophyte	1	$1 - t_m$
Female gametophyte	1	$1 - t_f$

3.3.1 Fecundity-viability antagonistic selection

Viability / fecundity trade-off is a classical form of antagonism underlying, many theoretical prediction on the evolution of life-history traits (ex: de Jong and Klinkhamer, 2005). A nice empirical example was analyzed in the monkey flower (*Mimulus guttatus*) where a polymorphic QTL for flower size exhibits antagonistic pleiotropy, with the allele increasing flower size increases fecundity but reduces viability (Mojica et al., 2012). Under this model, the rescaled coefficients of selection are:

$$s_1 = s - \left(hs + \frac{h_f s_f (1 + \sigma) + h_m s_m (1 - \sigma)}{2} \right) \tag{24a}$$

$$s_2 = \frac{s_f (1 + \sigma) + s_m (1 - \sigma)}{2} - \left(hs + \frac{h_f s_f (1 + \sigma) + h_m s_m (1 - \sigma)}{2} \right) \tag{24b}$$

with conditions $s_1 > 0$ and $s_2 > 0$. Scaled selection coefficients corresponds to the difference between fitness of heterozygotes and homozygotes. Fecundity selection coefficients are averaged over both sexes. Importantly, the selfing rate, σ, is included in fecundity selection coefficients because selfing does not affect selection in the same way for male and female fecundity, and viability (Damgaard, 2000): no selection occurs on the male function but the strength of selection on the female function is doubled because two gametes are transmitted for one selected parent (see also equation 7 in Jordan and Connallon, 2014 for a similar interpretation). When fecundity selection is the same for both sexes ($s_m = s_f$ and $h_m = h_f$), scaled selection coefficients become independent of selfing and the model is strictly equivalent to overdominance. Otherwise, the effect of selfing cannot be directly deduced from the overdominant model.

Selfing has two effects. On the one hand, it affects genotype frequencies and destabilizes polymorphism, as for true overdominance. On the other hand it changes the strength of fecundity selection. So, contrary to true overdominance, some degree of selfing can favour polymorphism for some ranges of parameters. This can be illustrated by plotting the conditions for polymorphism as a function of selfing rate and the relative difference between viability and fecundity selection, $\delta = (2s - (s_m + s_f))/(2s + s_m + s_f)$ (Figure 4). The conditions also depend on dominance coefficients (h_m and h or h_f and h): on

17

Figure 4, the darker the region, the broader the conditions in the dominance space. When

306   $\delta > 0$ (resp. $\delta < 0$) viability selection is stronger (resp. weaker) than fecundity selection.

For a male-fecundity/viability trade-off, partial selfing can partly broaden the dominance

308   conditions for stable polymorphism if male-fecundity selection is stronger than viability

selection ($\delta < 0$). In reverse, for a female-fecundity/viability trade-off, selfing can stabi-

310   lize polymorphism if female-fecundity selection is weaker than viability selection ($\delta > 0$).

Globally, selfing reduces the conditions for which polymorphism can be maintained (Fig-

312   ure 4) and under full selfing stability is possible only if fecundity and viability selection

balance exactly.

Figure 4: Proportion of the dominance parameter space allowing balancing selection as a function of selfing rate, σ and relative difference in viability and fecundity selection $\delta = (2s - (s_m + s_f))/(2s + s_m + s_f)$. A): Male fecundity selection only, $s_f = 0$. B: Female fecundity selection only, $s_m = 0$. The grey scales indicate the proportion of the dominance parameter space where stable polymorphism is maintained: $\{h, h_m\} \in \{0, 1\}^2$ or $\{h, h_f\} \in \{0, 1\}^2$. The darker the region, the broader the dominance conditions allowing the maintenance of polymorphism. In the white areas no polymorphism can be maintained. At max the proportion of the dominance space that allows stable polymorphism is 0.5. A: when $\delta < 0$ there are conditions where (partial) selfing broadens the dominance space allowing for stable polymorphism as illustrated by the white dashed line ($\delta = -0.3$) for which the proportion of the dominance space increases from 27% for $\sigma = 0$ to 50% for $\sigma \approx 0.45$. B: The same is true for $\delta > 0$ (white dashed line: $\delta = 0.3$) for which the proportion increases from 27% for $\sigma = 0$ to 50% for $\sigma \approx 0.85$.

314   Another interesting conclusion of the model is that, under increasing selfing, female-

fecundity alleles fix at the detriment of viability alleles, which fix at the detriment of

316   male alleles. So, independently of selection for resource allocation that would affect

the respective values of s_f, s_m, and s, the effect of selfing on selection dynamics, by

318   itself, also favours female fecundity more than viability. As a consequence, because

male component does not affect demographic fitness, selfing reduces the load under a

320   male/viability antagonism but increases it for a female/viability antagonism (Figure 5).

Figure 5: Effect of selfing rate on the load due to viability fecundity antagonism. A: Male fecundity selection only, $s_m = 0.05$ and $s_f = 0$. B: Female fecundity selection only, $s_m = 0$ and $s_f = 0.05$. For all curves, h_m or $h_f = 0$. Thick line: $h = 0$ and $s = 0.05$; Thin line: $h = 0$ and $s = 0.02$; Thick line: $h = 0.3$ and $s = 0.05$; Thick line: $h = 0.3$ and $s = 0.02$.

3.3.2 Sexually antagonistic selection

Sex antagonism has been intensively studied in dioecious species but it can also occur in hermaphrodites. For example, in the hermaphroditic freshwater snail *Lymnaea stagnalis* experimental manipulation revealed a strong cost of copulation behavior on egg production (De Visser et al., 1994). Two recent works have shown that selfing reduces the conditions for maintenance of sexually antagonistic polymorphism (Jordan and Connallon, 2014; Tazzyman and Abbott, 2015), with slightly different assumptions on how selection acts under selfing. Tazzyman and Abbott (2015) used the strong assumption that there is no selection at all when offspring are produced by selfing whereas Jordan and Connallon (2014) considered that male fecundity selection alone vanishes. I follow Jordan and Connallon (2014) and give their results in a slightly different form for completeness with the rest of the study. Genotype fitnesses are given in Table 2B. Then rescaled selection coefficients are:

$$s_1 = \frac{s_m(1-\sigma) - h_f s_f(1+\sigma) - h_m s_m(1-\sigma)}{2} \tag{25a}$$

$$s_2 = \frac{s_f(1+\sigma) - h_f s_f(1+\sigma) - h_m s_m(1-\sigma)}{2} \tag{25b}$$

The main conclusion is that selfing reduces the range of parameters allowing the maintenance of sex-antagonist polymorphism and the efficacy of selection at polymorphic equilibrium (when it exists) (Jordan and Connallon, 2014). Under this scenario, selfing also decreases the load by optimizing female fecundity.

Previous studies have not considered the impact of selfing rates on haploid sexually antagonistic selection. At the haploid level, only strong and rather symmetrical selection allows the maintenance of polymorphism. Selfing has no effect on female haploid selection but suppresses male haploid selection. Globally, the effect of selfing at the haploid level is similar to the effect for sexual antagonism at the diploid level with two main effects: i) selfing skews and reduces the range of parameters for which polymorphism is maintained and ii) selfing favors female-beneficial over male-beneficial alleles (Appendix 4).

3.3.3 Ploidally antagonistic selection

Another form of antagonism may arise between the haploid and diploid phases of a life cycle. In plants, trade-off can occur between pollen competitive ability (haploid selection) and diploid fitness component, as shown in the self-compatible plant *Collinsia heterophylla* where delayed stigma receptivity increases female fitness but intensifies pollen competition whereas early fertilization forced by pollen reduces seed set (Hersh et al., 2015; Lankinen et al., 2016).

First, consider ploidally antagonistic selection without sex difference (Viability selection and $t_m = t_f = t$ in Table 2D). The rescaled coefficients of selection are:

$$s_1 = s(1-h) - t(1-\sigma/2) \tag{26a}$$

$$s_2 = t(1-\sigma/2) - hs \tag{26b}$$

In random mating populations, polymorphism can be maintained only if selection for the beneficial allele in the haploid phase is stronger than counter selection in a diploid heterozygote but weaker than counter selection in a diploid homozygotes (Immler et al., 2012). Equations (26) show that selfing reduces haploid selection by two ($\sigma/2$ term) because male haploid selection is ineffective under selfing and female haploid selection is unchanged (see also Peters and Weis, 2018, for more specific selection schemes).

Immler et al. (2012) showed that the combination of both ploidally and sexually antagonistic selection enlarged the range for maintenance of polymorphism, especially with negative ploidy-by-sex interactions. Selfing should also reduces this form of selection. For antagonism between haploid male and diploid female selection the rescaling leads to:

$$s_1 = ((1-h_f)s_f(1+\sigma) - t_m(1-\sigma))/2 \tag{27a}$$

$$s_2 = ((t_m(1-\sigma) - h_f s_f(1+\sigma))/2 \tag{27b}$$

The 1/2 term simply arises because selection only acts on half fitness components in both

the haploid and the diploid phase. As for sexually antagonistic selection, selfing reduces selection on the male but increases it on the female component. Similarly, for antagonism between diploid male and haploid female selection, we have:

$$s_1 = (t_f - s_m(1 - h_m)(1 - \sigma))/2 \tag{28a}$$

$$s_2 = (h_m s_m(1 - \sigma) - t_f)/2 \tag{28b}$$

Here, only the male component is affected as haploid female selection is independent of selfing.

Overall, selfing globally reduces ploidally antagonistic selection, sometimes also shifting the range of parameters allowing stable polymorphism. The antagonism between viability and haploid female selection is the least affected and for sex-by-ploidy interactions, female haploid/male diploid antagonism is more likely to persist under partial selfing than female diploid/male haploid antagonism.

3.3.4 Balanced segregation distorter

Finally, another form of antagonism is segregation distortion at meiosis where an allele experiences a transmission bias, counteracted by selection at the diploid level. In the monkey flower (*M. guttattus*) a centromere-associated female meiotic drive element is maintained polymorphic because it entails several male fitness costs (Fishman and Saunders, 2008). This is similar but not fully equivalent to haploid-diploid antagonism. Meiotic drive causes selection at the haploid level but the two alleles do not have fixed intrinsic fitness. Fitness differences only express in heterozygous parents at meiosis. Under random mating this is formally equivalent to genic selection (Nagylaki, 1983). In contrast, heterozygosity, hence the possibility of distortion, vanishes under selfing. Distortion is thus formally equivalent to male haploid selection, which also vanishes under selfing, but not to female haploid selection. As a minimum condition for polymorphism, the distorter allele must be partially recessive at the diploid level.

The effect of selfing on sex-independent balance between segregation distortion and

diploid selection has been investigated by Glémin (2010, 2011) and the effect of selfing on the dynamics of the centromere drive in *M. guttattus* has been modelled (Fishman and Kelly, 2015). These two cases can be encapsulated in the general case given by:

$$s_1 = \frac{b_f + b_m}{2}(1 - \sigma) - \left((1 - h)s + (1 - h_f)s_f(1 + \sigma)/2 + (1 - h_m)s_m(1 - \sigma)/2\right) \quad (29\text{a})$$

$$s_2 = hs + h_f s_f(1 + \sigma)/2 + h_m s_m(1 - \sigma)/2 - \frac{b_f + b_m}{2}(1 - \sigma) \quad (29\text{b})$$

366 with diploid fitness as in Table 2D and distortion bias b_f and b_m at female and male meiosis respectively. By preventing the spread of deleterious distorters, selfing always
368 reduces the genetic load, as for male vs female fecundity or viability antagonism but differently from female fecundity/viability antagonism.

370 ## 3.4 Negative frequency-dependent selection

The results presented above illustrate that polymorphisms maintained by various
372 forms of antagonistic selection should be rare in selfing species. Negative frequency-dependent selection is an alternative mechanism for the maintenance of polymorphism
374 by selection. In random mating populations, the dynamics is indistinguishable from over-dominance (at least for linear frequency-dependent selection, Takahata and Nei, 1990).
376 This equivalence is no longer true under partial selfing where polymorphism can be more easily maintained under negative frequency-dependent selection than under overdomi-
378 nance. This has been numerically explored (Jain and Jain, 1970) or used in simulations for a related question with a single specific setting (Nordborg et al., 1996), but to my
380 knowledge, no general and analytical result has been obtained. I only give results for linear frequency dependence in the main text but they can be generalized to any forms
382 of selection in the case of allelic frequency dependence (Appendix 5).

3.4.1 Allelic frequency-dependent selection at a single locus

384 I assume additive allelic effects with allele A_i having fitness $1 - s_i x_i$. Thus, the fitness of a diploid genotype $A_i A_j$ is given by: $1 - (s_i x_i + s_j x_j)$. Change in allelic frequencies is

22

given by:

$$\Delta x_i = (1 + F)\frac{x_i(J_2 - s_i x_i)}{(1 - 2J_2)} \tag{30}$$

where $J_2 = \sum_{i=1}^{n} s_i x_i^2$, and equilibrium frequencies by:

$$x_i^{eq} = \frac{s_H}{n s_i} \tag{31}$$

where s_H is the harmonic mean of selection coefficients. The efficacy of selection is:

$$\Omega_{eq} = \Omega_0 = \frac{s_H}{n} \tag{32}$$

Where Ω_0 stands for the efficacy of selection on a new $n+1^{th}$ allele arising in a population at equilibrium for n resident alleles. From these two equations, the main conclusion is that selfing rate has no direct effect on this form of selection. This is also true for any form of allelic frequency-dependent selection (Appendix 5).

3.4.2 Genotypic frequency-dependent selection at a single locus

Selection can also depend on genotypic frequencies. Noting X_{ij} the frequency of genotype $A_i A_j$, its fitness can be written as $1 - s_{ij} X_{ij}$. However, this general model leads to formidable equations that are hardly tractable. I thus focus on the much simpler symmetrical case where all s_{ij} are equal to s (see details in Appendix 5). If so, all allelic frequencies are equal to $1/n$ at equilibrium. Around this equilibrium, the efficacy of selection is:

$$\Omega_{eq} = \frac{s\left(F^2\left(n^2 - 2n - 3\right) + 6F + 2n - 3\right)}{(1 + F)n^3} \tag{33}$$

Equation (33) is not monotonic in F but globally, selfing increases the efficacy of selection around equilibrium and the efficacy of selection is always greater under full selfing than under outcrossing (Figure 6A). The reason is that selection decreases with the number of common genotypes. By making heterozygote genotypes rarer, selfing thus increases selection. At the extreme, under pure selfing, there are only n genotypes instead of $n(n + 1)/2$. Interestingly, the effect of selfing is different on the efficacy of selection on

new rare alleles (Figure 6B). Starting from an equilibrium population with n resident alleles, the efficacy of selection on a new $(n+1)^{th}$ is given by:

$$\Omega_0 = s\frac{2n - 1 - (F(n-1)(n(1+F) - 2)}{(1+F)n^3} \tag{34}$$

For $n = 1$, so for the emergence of an initial polymorphism, equation (34) reduces to $\frac{s}{1+F}$, hence selfing reduces the efficacy of selection: starting from a monomorphic population, selfing does not change the number of common genotypes, which is one. The only effect of selfing is the reduction in N_e by $1 + F$. For $n > 1$, selfing reduces the number of common genotype so increases selection, which partly balances the reduction in N_e and overwhelms it for large n. Up to three resident alleles, the efficacy of selection is maximal under random mating but for $n \geq 4$ the efficacy of selection is greater under full selfing (Appendix 5). For large n, equation (34) reduces to $\frac{F^2}{(1+F)n}$ and selfing always increases the efficacy of selection. Overall, selfing strengthens selection to maintain such polymorphisms, but reduces the initial chance of establishing them.

Figure 6: Efficacy of selection (normalized by s) under symmetrical genotypic frequency dependent selection as a function of the number of alleles already present in the population, n. A: at equilibrium, Ω_{eq}. B: on a new allele,Ω_0. Equation (33) and (34) divided by s are used with $F = \sigma/(2 - \sigma)$.

3.4.3 Frequency-dependent selection at two loci

Another important effect of selfing is to reduce the efficacy of recombination (Golding and Strobeck, 1980; Nordborg, 2000). To study its potential effect on balancing selection dynamics, I extended the two previous models to two bi-allelic loci with either haplotypic or genotypic frequency-dependent selection.

Consider a first locus with two alleles A_1 and A_2 in frequency x and $1-x$, and a second one with alleles B_1 and B_2 in frequency y and $1-y$. The recombination rate between the two loci is r. I denote the four haplotypes $\{A_1B_1, A_1B_2, A_2B_1, A_2B_2\}$ as $\{X_1, X_2, X_3, X_4\}$. From this four haplotypes, ten genotypes are possible that I denote G_{ij} as the combination of haplotype X_i and X_j; for example the genotype $\{A_1A_1; B_1B_2\}$ is noted G_{12}. Note that

24

428 the double heterozygote $\{A_1A_2; B_1B_2\}$ corresponds to two haplotype combinations, G_{14} and G_{23}, that must be distinguished for recursions. The general and exact recursions lead

430 to quite tedious algebra. But close to equilibrium and under weak selection, approximate equations can be written as a function of only four variables: the two allelic frequencies,

432 x and y, and two identity coefficients, F and Φ, the probability of identity by descent at one and two loci (Appendix 5). At equilibrium (Roze, 2009):

$$\Phi = \frac{1-F}{1-2Fr(1-r)} + 2F - 1 \tag{35}$$

For haplotypic frequency-dependent selection, the fitness of a genotype is the mean of the fitness of its haplotypes. The fitness of haplotype i is $1 - s_i X_i$, so the fitness of genotype G_{ij} is $1 - s_i X_i - s_j X_j$. The change in allelic frequencies is approximately given by:

$$\Delta x = (1+F)x(1-x)\left(y^2(s_1 x + s_3(1-x)) + (1-y)^2(s_2 x + s_4(1-x))\right) \tag{36a}$$

$$\Delta y = (1+F)y(1-y)\left(x^2(s_1 y + s_2(1-y)) + (1-x)^2(s_3 y + s_4(1-y))\right) \tag{36b}$$

From equation (36) the equilibrium allelic frequencies are solutions of:

$$x_{eq} = \frac{s_3 y_{eq}^2 + s_4(1-y_{eq})^2}{(s_1+s_3)y_{eq}^2 + (s_2+s_4)(1-y_{eq})^2} \tag{37a}$$

$$y_{eq} = \frac{s_2 x_{eq}^2 + s_4(1-x_{eq})^2}{(s_1+s_2)x_{eq}^2 + (s_3+s_4)(1-x_{eq})^2} \tag{37b}$$

434 There is no explicit analytical solution for x_{eq} and y_{eq} but they are independent of F and Φ, hence of selfing rate. Second, plugging Δx and Δy given by equations (36) into

436 equation (12) also shows that the efficacy of selection is exactly balanced by drift, so selfing has no effect on haplotype frequency-dependent selection at two loci.

438 The general case of genotypic frequency-dependent selection leads to tedious algebra. As for a single locus, I thus only consider symmetrical selection where the fitness of

440 genotype G_{ij} is $1 - sG_{ij}$. Here, I consider that G_{14} and G_{23} correspond to the same phenotype, hence the same fitness: $1 - s(G_{14} + G_{23})$. Note, that assuming that the two

442 genotypes correspond to different phenotypes gives a very similar conclusion. With this

25

simplifying assumption, all alleles have equilibrium frequency 1/2, which greatly simplifies

444 equations (Appendix 5). Around the polymorphic equilibrium the efficacy of selection is:

$$\Omega_{eq} = s \frac{(1 + F(3 + 8r(r-1)))(3 + F(1 + 8Fr(r-1)))}{64(1 + F)(1 + 2Fr(1-r))^2} \tag{38}$$

446 Ω_{eq} increases with F and depends very weakly on r. For $r = 0$ it reduces to:

$$\Omega_{eq} = s \frac{(1 + 3F)(3 + F)}{64(1 + F)} \tag{39}$$

Now consider an initial population at equilibrium with a polymorphic B locus but the A

448 locus fixed for A_1. The strength of selection for a rare A_2 allele is given by:

$$\Omega_0 = s \frac{(3 - 2F + 3F^2)}{8(1 + F)} \tag{40}$$

Ω_0 is independent of r and decreases with F until $F \approx 0.633$ and then increases. But

450 the maximum is reached for $F = 0$. Interestingly, this is the same equation as (34) with

$n = 2$. As for genotypic selection at a single locus, selfing can strengthen selection on

452 established polymorphism, but reduces it on establishment.

The main conclusion of the two-locus model analysis is that selfing does not modify

454 the outcome of frequency dependent selection because of its effect on genetic linkage.

4   Discussion

456 In this general survey I considered the main cases of balancing selection. I provided

a series of new results and re-derived previous ones for completeness under the weak

458 selection approximation. Despite the limits of those approximations, which I discuss first,

this leads to simpler and more insightful formula, easily interpretable from a biological

460 point of view. Using the same formalism and approximations also makes comparison

among models easier and allows a general classification (Table 3).

4.1 Interest and limits of the weak selection approximation

The weak selection approximation has several merits. In addition to providing simple and easily interpretable analytical results, it highlights the similarity between overdominance and other forms of selection, which is much less obvious from the full models. Thus, it allows a general classification of balancing selection in regard to the effect of selfing. Approximated results can also be easily incorporated into a diffusion framework, which is useful for statistical inferences in population genomics.

However, strong balancing selection is possible in natural populations, especially if it involves genomic regions with multiple variants, and not only single nucleotide polymorphism. This is possible when haplotype blocks of multiple recessive deleterious allele are linked in repulsion, generating the so-called pseudo-overdominance (Ohta and Kimura, 1969), as suspected to explain inbreeding depression in the tree *Eucalyptus grandis* (Hedrick et al., 2016) and the maintenance of heterozygosity upon enforced selfing in three outcrossing *Caenorhabditis* nematodes (Barriere et al., 2009). Structural variants, such as chromosomal inversions, can also capture multiple alleles with collectively strong effect. The fitness component of such an inversion polymorphism was studied in detail in the seaweed fly *Coelopa frigida* (Mérot et al., 2020). The authors detected and estimated both true overdominance (on male viability) and viability/fecundity antagonism (in females), resulting in sexual antagonism for whole fitness, with some selective effects of the order of 20%.

When selection is strong, the conditions for protected polymorphism are less stringent than predicted by the weak selection approximation, as discussed above for true overdominance where polymorphism can be maintained for any selfing rate if heterozygotes are at least twice better than homozygotes (Kimura and Ohta, 1971; Rocheleau and Lessard, 2000). For antagonistic or variable selection, the conditions in an outcrossing population is that the geometric mean fitness of heterozygotes must be higher than the geometric mean fitness of the best homozygotes, which is less stringent than the overdominant condition that involves the arithmetic mean fitnesses (but as expected the two conditions converge under weak selection). For example, when selection is additive for

27

each fitness component, the weak approximation predicts that polymorphism cannot be maintained whereas it is possible under strong selection. Strictly speaking, *i.e.* without approximation, it is also true for weak selection but in a very narrow range of parameters where selection is symmetrical. As selfing makes selection more additive, the general conclusion holds that selfing limits overdominant-like polymorphism to either symmetrical conditions or to strong selection.

4.2 Maintenance of variation for fitness: two categories of balancing selection

The main result of the study is that balanced polymorphism can be classified into two main categories depending on the effect of selfing. First, true and emergent overdominance is strongly affected by selfing that destabilizes polymorphism by removing heterozygotes from the population. This occurs for all forms of overdominance. In addition, when emergent overdominance results from sex antagonism, selfing also differently alters the apparent selection coefficients on male and female components: selection on both haploid and diploid male components is reduced whereas diploid female selection is enhanced and haploid female selection remains unaffected. Except for a limited range of conditions, this process tends to increase the asymmetry of selection against the two homozygotes, hence destabilizing polymorphism further (see Figure 4). True or emergent overdominance is thus unlikely to contribute to fitness variation in highly selfing species. However, all mechanisms are not equivalent as far as the mean fitness of the population is concerned. For overdominance not generated by sex antagonism, selfing always decreases the mean fitness of the population (see Figure 3 for example). The picture is different with sex-antagonistic selection if we assume that the mean fitness of the population only depends on viability and female fecundity (Equation 7). Depending on selection and dominance parameters, selfing can either increase or decrease the population mean fitness under female fecundity/viability antagonism (see Figure 5B for examples of increasing load). However, when antagonism involves the male component, selfing always increases the mean fitness of the population (Figure 5A). Antagonistic selection on a male com-

28

Table 3: Classification of the effects of selfing on the different forms of balancing selection. $+,-$, and 0 indicate that selfing increases, decreases or has no effect on the range of parameters for which stable polymorphism is maintained (Stability), on the efficacy of selection on emergence of new alleles (Ω_0) or at equilibrium (Ω_{eq}), and on the load. NA: not applicable. References - 1: Hayman (1953). 2: Weir (1970). 3: Kimura and Ohta (1971). 4: Rocheleau and Lessard (2000). 5: Hedrick (1998). 6: Jordan and Connallon (2014). 7: Tazzyman and Abbott (2015). 8: Peters and Weis (2018). 9: Glémin (2010). 10: Fishman and Kelly (2015). 11: Jain and Jain (1970). ts: this study (indicated when new results are provided, not only re-derived). *: More precisely, the case studied in 8 is male haploid/sex-independent diploid antagonism.

Form of selection	Stability	Ω_0	Ω_{eq}	Load	Ref
Emergent overdominance					
True overdominance	$-$	$-$	$-$	$+$	1,2,3,4,ts
Selection variable in space	$-$	$-$	$-$	$+$	5,ts
Selection variable in time	$-$	Depends on dominance	NA	NA	ts
Viability/fecundity antagonism					
Global fecundity	$-$	$-$	$-$	$+$	ts
Male fecundity	$-$	$-$	$-$	$-$	ts
Female fecundity	$-$	$-$	$-$	$+$	ts
Sex antagonism					
Haploid	$-$	$-$	$-$	$-$	ts
Diploid	$-$	$-$	$-$	$-$	6,7
Ploidy antagonism					
Sex independent	$-$	$-$	$-$	$+$	ts
Male haploid/female diploid	$-$	$-$	$-$	$-$	8*
Female haploid/male diploid	$-$	$-$	$-$	$-$	ts
Balanced segregation distorter					
Sex independent	$-$	$-$	$-$	$-$	9
Male distorter	$-$	$-$	$-$	$-$	ts
Female distorter	$-$	$-$	$-$	$-$	10
Negative frequency-dependence					
At one locus					
Allelic	0	0	0	NA	ts
Genotypic	0	$-$	$+$	NA	11, ts
At two loci					
Haplotypic	0	0	0	NA	ts
Genotypic	0	$-$	$+$	NA	ts

ponent prevent the optimization of viability or female fecundity, hence coming at a cost

520 for population mean fitness. By removing selection on male components, selfing allows

optimizing viability and female fecundity. In some ecological scenarios, male fecundity

522 can also contribute to the population mean fitness, for example in plants under pollen

limitation. If so, in addition to the above-mentioned effect, selfing improves further the
524 population mean fitness by ensuring fertilization.

Second, balancing selection generated by negative frequency-dependent selection is, on
526 the contrary, mostly unaffected by selfing. When selection depends on allele or haplotype
frequencies, selfing has strictly no effect. When selection depends on genotype frequencies,
528 selfing does have an effect but rather weak. At equilibrium, selfing increases the efficacy of
selection and so does it for the emergence of new alleles when they are already numerous
530 in the population. However, when there are only a few alleles, selfing slightly decreases
the strength of selection on the emergence of new alleles (Figure 6). Overall, selfing
532 can strengthen selection on established polymorphisms but reduces the initial chance of
establishing them. However, in no case does selfing affect the stability of this category of
534 balanced polymorphism, in contrast with overdominant-like selection.

The two-locus model also shows that selfing does not affect the selection dynam-
536 ics through the reduction of effective recombination. This also differs from what can
be expected for overdominant-like selection. For example, although selfing reduces the
538 maintenance of single-locus sexually antagonistic polymorphism (Jordan and Connallon,
2014; Tazzyman and Abbott, 2015), the reduction is less pronounced in multilocus sys-
540 tems (Olito, 2016): increased genetic linkage generated by selfing helps maintain male-
beneficial alleles, hence polymorphism (but still in restricted conditions compared to
542 outcrossing). Increased linkage under selfing could also generate pseudo-overdominance,
as discussed above (Ohta and Kimura, 1969). Whether these conclusions generalizes to
544 all forms of overdominance remains to be studied.

4.3 Signature of balancing selection in self-fertilizing species

546 Balancing selection leaves specific genomic signatures because it strongly affects the
life span of an allele. When selection is not too asymmetrical, *i.e.* x_{eq} not too close to
548 0 or 1, balancing selection can considerably increase the sojourn time of an allele in a
population (Takahata, 1990). As a consequence, higher polymorphism level, higher ratio
550 of non-synonymous over synonymous polymorphism (π_N/π_S), and higher occurrence of

trans-specific polymorphism are expected at, or close to, sites under balancing selection. However, for strongly asymmetrical selection, balancing selection can speed up the loss of polymorphism, especially in small populations (Robertson, 1962; Nei and Roychoudhury, 1973; Zhao and Waxman, 2016). In a random mating population, overdominance and linear negative frequency-dependent selection generate the same allele-frequency dynamics, so identical genomic signature (Takahata and Nei, 1990; Zhao and Waxman, 2016). This is no longer true in a partially selfing population. Selfing reduces the range of stability of overdominance and reinforces the asymmetry of selection, pushing equilibrium frequency close to 0 or 1, so in the zone where overdominance selection favours the loss of polymorphism instead of its maintenance. Such polymorphisms can thus hardly be maintained, even in populations with limited selfing, except under strong or strict symmetrical selection. On the contrary, selfing has almost no effect on frequency-dependent selection, except for additional reduction in N_e below the two-fold automatic effect (such as caused by linked selection), and polymorphism can still be maintained even in populations with small N_e.

In primarily selfing species, contrary to outcrossing ones, molecular signature of balancing selection can thus be more likely attributed to negative frequency-dependent selection. So far, genomic regions under balancing selection have been detected in selfing species, mainly on genes involved in defense against pathogens. A general survey of Resistance genes (R-genes) in *Arabidopsis thaliana* revealed a global excess of amino-acid polymorphism compared to random control genes and evidence of balancing selection for some of them (Bakker et al., 2006). Signature of balancing selection has also been found in the selfing nematode *Caenorabditis elegans* in a gene involved in defense against pathogenic bacteria (Ghosh et al., 2012). In the *Capsella* plant genus, signature of balancing selection at immunity genes has been maintained after the shift from outcrossing (*C. grandiflora*) to selfing (*C. rubella*) (Gos et al., 2012; Koenig et al., 2019), despite massive loss of polymorphism in the rest of the genome (Slotte et al., 2013; Brandvain et al., 2013). Interestingly, this is in agreement with the finding that under genotypic negative frequency-dependent selection, selfing can strengthen the selection on already established

31

580 polymorphism despite slightly lowering the probability of emergence of new haplotypes. More simply, polymorphisms can be maintained because selfing has little effect on their

582 stability. Following these studies, more systematic comparison between closely related outcrossing and selfing species or populations could be a way to distinguish the two main

584 forms of balancing selection. Taking advantage of the contrasted effect of selfing on selection dynamics has recently be used to infer the dominance of deleterious mutations

586 by comparing the selfer *A. thaliana* and its outcrossing relative *A. lyrata* (Huber et al., 2018). Similar comparisons focusing on balancing selection should help understand the

588 mechanisms maintaining polymorphism in those genomic regions.

4.4 Consequences for the evolution selfing species

590 Transitions from outcrossing to selfing occur recurrently in many groups of organisms, including plants, invertebrates, algae and fungi. Such transitions should not strongly

592 affect the dynamics of genes under negative frequency-dependent selection. On the contrary, the loss of variation initially maintained by overdominance-like selection in the

594 outcrossing ancestor may have strong consequences on the derived selfing species, both at short and long time scale.

596 The evolution of selfing is often associated with a set of phenotypic (Sicard and Lenhard, 2011; Cutter, 2008) and ecological traits (Aarssen, 2000; Snell and Aarssen,

598 2005; Petrone Mendoza et al., 2018), collectively referred to as the "selfing syndrome". It is likely that some of these traits were initially under sexual selection and conflicts in

600 the outcrossing ancestor (see Cutter, 2019, for a recent review). If, in the outcrossing ancestor, a polymorphism was maintained by antagonistic forms of balancing selection,

602 the evolution of selfing would rapidly destabilize such a polymorphism and drive the evolution towards female fecundity traits. This can explain the rapid evolution of selfing

604 syndrome as observed between closely related species (for example in *Capsella rubella* Sicard et al., 2011; Slotte et al., 2012) or even between populations of the same species,

606 as in *Arabidopsis alpina* (Tedder et al., 2015). Glémin and Ronfort (2013) suggested that positive selection on selfing-syndrome traits was more likely than accumulation of

608 deleterious mutations following relaxed selection on outcrossing traits to explain the pace

at which selfing syndrome evolved. Evolution from a balanced polymorphism maintained

610 by antagonistic selection combines both relaxation on one trait and positive selection on

another trait. Interestingly, destabilization of antagonistic polymorphism under selfing

612 not only predicts the evolution of female traits at the expense of male traits (a typi-

cal characteristic of the selfing syndrome) but also female fecundity at the expense of

614 viability, so typically towards the evolution of shorter life span and more ruderal strate-

gies. Importantly, these predictions do not rely on modification of selection pressures

616 such as in models where selfing selects for higher reproductive effort (Waller and Green,

1981; Lesaffre and Billiard, 2020) - that is increase in s_f and decrease in s with current

618 notations (see Table 2). It only comes from the destabilizing effect of selfing, which is

expected to lead to very rapid fixation of the female advantageous allele. Indeed, under

620 this scenario, selection automatically proceeds from standing variation, starting with al-

leles at potentially high frequency. Using equation (21) in Glémin (2012), the time to

622 fixation of the female allele starting from frequency x_0 in a fully selfing population is

$\frac{1}{s}\big(\gamma + \ln(4Ns(1 - x_0)/x_0)\big)$, where $\gamma \approx 0.577$ is the Euler's constant. Using $N = 20,000$

624 and $s = 0.01$, the time to fixation range from 946 to 506 generations for an initial fre-

quency ranging from 0.1 to 0.9. As a comparison, with the same N and s, evolution from

626 initial mutation-selection-drift balance takes about 2,600 to 26,000 generations if the fe-

male allele is selected for, and from one to ten millions generations if the female allele

628 is neutral and evolves only under relaxed selection (see Table 2 in Glémin and Ronfort,

2013). Evolution of selfing is expected to alter selection pressures, for example to favor

630 resources reallocation to female versus male or attraction functions (Charnov, 1987) or

to shorten the life cycle (Baker and Stebbins, 1965; Aarssen, 2000; Snell and Aarssen,

632 2005). Destabilization of balanced polymorphism can speed up the process as soon as

selfing rate increases.

634   It has long been proposed that co-evolution with pathogens could select for outcrossing

(Levin, 1975; Agrawal and Lively, 2001) and several comparative studies found a positive

636 association between outcrossing rates and pathogen attacks or herbivory (Busch et al.,

33

2004; Johnson et al., 2009; Campbell and Kessler, 2013; Rivkin et al., 2018). However, the underlying reasons are not completely obvious. If it mainly involves negative frequency-dependent selection dynamics, defence against pathogens should be similarly challenging in selfing than in outcrossing species. If N_e is reduced well beyond the twofold automatic effect, fewer alleles should be maintained at resistance genes, but polymorphism should persist as observed in several species. In addition, as immunity genes often belongs to multigenic families, combinations of distinct homozygotes at several loci provide fixed heterozygosity in highly selfing species. So, it is not clear whether the empirical observations of reduced resistance in selfing species result from lower diversity at recognition genes rather than from other negative consequences of selfing. For example, a survey of R-genes in *A. thaliana* revealed that, although upstream genes in resistance pathways exhibited clear signature of balancing selection, downstream genes appeared mainly under purifying selection (Bakker et al., 2008). It is thus possible that pathogens and herbivores select for the maintenance of outcrossing by mediating higher inbreeding depression (ex: Carr and Eubanks, 2002; Campbell et al., 2013) rather than or at least in addition to, directly selecting for diversification of resistance genes. More generally, how ecological interactions with enemies may affect or be affected by mating systems can be complex (Carr and Eubanks, 2014; Campbell, 2015). Results presented here can be used as a starting point but a general theoretical background still needs to be developed.

Acknowledgments

I thank Matthew Hartfield and Lucas Marie-Orleach for their feedbacks and two anonymous reviewers for their careful reading and helpful comments on the manuscript. I am supported by the CNRS. I declare no conflict of interest.

5 title

References

Aarssen, L., 2000. Why are most selfers annuals? a new hypothesis for the fitness benefit of selfing. Oikos 89:606–612.

Agrawal, A. F. and M. Hartfield, 2016. Coalescence with background and balancing selection in systems with bi- and uniparental reproduction: Contrasting partial asexuality and selfing. Genetics 202:313–26.

Agrawal, A. F. and C. M. Lively, 2001. Parasites and the evolution of self-fertilization. Evolution 55:869–879.

Audigeos, D., L. Brousseau, S. Traissac, C. Scotti-Saintagne, and I. Scotti, 2013. Molecular divergence in tropical tree populations occupying environmental mosaics. J Evol Biol 26:529–44.

Baker, H. and G. Stebbins, 1965. The Genetics of Colonizing Species. Academic Press.

Bakker, E. G., C. Toomajian, M. Kreitman, and J. Bergelson, 2006. A genome-wide survey of r gene polymorphisms in *Arabidopsis*. Plant Cell 18:1803–18.

Bakker, E. G., M. B. Traw, C. Toomajian, M. Kreitman, and J. Bergelson, 2008. Low levels of polymorphism in genes that control the activation of defense response in *Arabidopsis thaliana*. Genetics 178:2031–43.

Barriere, A., S. P. Yang, E. Pekarek, C. G. Thomas, E. S. Haag, and I. Ruvinsky, 2009. Detecting heterozygosity in shotgun genome assemblies: Lessons from obligately outcrossing nematodes. Genome Res 19:470–80.

Bergelson, J., M. Kreitman, E. A. Stahl, and D. Tian, 2001. Evolutionary dynamics of plant r-genes. Science 292:2281–5.

Billiard, S., M. Lopez-Villavicencio, M. E. Hood, and T. Giraud, 2012. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. Journal of Evolutionary Biology 25:1020–38.

Brandvain, Y., T. Slotte, K. M. Hazzouri, S. I. Wright, and G. Coop, 2013. Genomic identification of founding haplotypes reveal the history of the selfing species *Capsella rubella*. PLoS Genetics 9.

Brown, K. E. and J. K. Kelly, 2018. Antagonistic pleiotropy can maintain fitness variation in annual plants. J Evol Biol 31:46–56.

Burgarella, C. and S. Glémin, 2017. Population Genetics and Genome Evolution of Selfing Species. John Wiley and Sons.

Busch, J. W., M. Neiman, and J. M. Koslow, 2004. Evidence for maintenance of sex by pathogens in plants. Evolution 58:2584–90.

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

Campbell, S. A., 2015. Ecological mechanisms for the coevolution of mating systems and defence. New Phytol 205:1047–53.

Campbell, S. A. and A. Kessler, 2013. Plant mating system transitions drive the macroevolution of defense strategies. Proc Natl Acad Sci USA 110:3973–8.

Campbell, S. A., J. S. Thaler, and A. Kessler, 2013. Plant chemistry underlies herbivore-mediated inbreeding depression in nature. Ecol Lett 16:252–60.

Carr, D. E. and M. D. Eubanks, 2002. Inbreeding alters resistance to insect herbivory and host plant quality in *Mimulus guttatus* (scrophulariaceae). Evolution 56:22–30.

———, 2014. Interactions between insect herbivores and plant mating systems. Annu Rev Entomol 59:185–203.

Charlesworth, B., 1992. Evolutionary rates in partially self-fertilizing species. Am Nat 140:126–148.

Charlesworth, B. and K. Hughes, 1999. The maintenance of genetic variation in life history traits, P. 702. Cambridge University Press, Cambridge.

Charlesworth, D., X. Vekemans, V. Castric, and S. Glemin, 2005. Plant self-incompatibility systems: a molecular evolutionary perspective. New Phytol 168:61–9.

Charnov, E., 1987. On sex allocation and selfing in higher plants. Evol Ecol 1:30–36.

Connallon, T. and S. F. Chenoweth, 2019. Dominance reversals and the maintenance of genetic variation for fitness. PLoS Biol 17:e3000118.

Curtsinger, J. W., P. M. Service, and T. Prout, 1994. Antagonistic pleiotropy, reversal of dominance, and genetic-polymorphism. Am Nat 144:210–228.

Cutter, A. D., 2008. Reproductive evolution: symptom of a selfing syndrome. Curr Biol 18:R1056–8.

———, 2019. Reproductive transitions in plants and animals: selfing syndrome, sexual selection and speciation. New Phytol 224:1080–1094.

Damgaard, C., 2000. Fixation of advantageous alleles in partially self-fertilizing populations. the effect of different selection modes. Genetics 154:813–21.

De Visser, J. A. G. M., A. Ter Maat, and C. Zonneveld, 1994. Energy budgets and reproductive allocation in the simultaneous hermaphrodite pond snail, lymnaeae stagnalis (l.): a trade-off between male and female function. Am. Nat. 144:861–867.

Delph, L. F. and J. K. Kelly, 2014. On the importance of balancing selection in plants. New Phytol 201:45–56.

Fishman, L. and J. K. Kelly, 2015. Centromere-associated meiotic drive and female fitness variation in *Mimulus*. Evolution 69:1208–18.

Fishman, L. and A. Saunders, 2008. Centromere-associated female meiotic drive entails male fitness costs in monkeyflowers. Science 322:1559–62.

Gemmell, N. J. and J. Slate, 2006. Heterozygote advantage for fecundity. Plos One 1.

Ghosh, R., E. C. Andersen, J. A. Shapiro, J. P. Gerke, and L. Kruglyak, 2012. Natural variation in a chloride channel subunit confers avermectin resistance in *C. elegans*. Science 335:574–8.

Gigord, L. D., M. R. Macnair, and A. Smithson, 2001. Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid *Dactylorhiza sambucina*. Proc Natl Acad Sci U S A 98:6253–5.

Glémin, S., 2010. Surprising fitness consequences of gc-biased gene conversion: I. mutation load and inbreeding depression. Genetics 185:939–59.

———, 2011. Surprising fitness consequences of gc-biased gene conversion. ii. Heterosis. Genetics 187:217–227.

———, 2012. Extinction and fixation times with dominance and inbreeding. Theor Popul Biol 81:310–6.

Glémin, S., C. M. Francois, and N. Galtier, 2019. Genome evolution in outcrossing vs. selfing vs. asexual species. Methods Mol Biol 1910:331–369.

Glémin, S. and J. Ronfort, 2013. Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. Evolution 67:225–40.

Golding, G. B. and C. Strobeck, 1980. Linkage disequilibrium in a finite population that is partially selfing. Genetics 94:777–89.

752 Gos, G., T. Slotte, and S. I. Wright, 2012. Signatures of balancing selection are maintained at disease resistance loci following mating system evolution and a population bottleneck
754 in the genus *Capsella*. BMC Evol Biol 12:152.

Grieshop, K. and G. Arnqvist, 2018. Sex-specific dominance reversal of genetic variation
756 for fitness. PLoS Biol 16:e2006810.

Haldane, J. and S. Jayakar, 1963. Polymorphism due to selection of varying direction. J
758 Genet 58:237–242.

Hanschen, E. R., M. D. Herron, J. J. Wiens, H. Nozaki, and R. E. Michod, 2018. Repeated
760 evolution and reversibility of self-fertilization in the volvocine green algae. Evolution 72:386–398.

762 Hartfield, M. and S. Glémin, 2014. Hitchhiking of deleterious alleles and the cost of adaptation in partially selfing species. Genetics 196:281–93.

764 ———, 2016. Limits to adaptation in partially selfing species. Genetics 203:959–74.

Hayman, B., 1953. Mixed selfing and random mating when homozygotes are at disad-
766 vantage. Heredity 7:185–192.

Hedrick, P. W., 1998. Maintenance of genetic polymorphism: Spatial selection and self-
768 fertilization. Am Nat 152:145–150.

———, 1999. Antagonistic pleiotropy and genetic polymorphism: a perspective. Heredity
770 82:126–133.

———, 2012. What is the evidence for heterozygote advantage selection? Trends Ecol
772 Evol 27:698–704.

Hedrick, P. W., U. Hellsten, and D. Grattapaglia, 2016. Examining the cause of high
774 inbreeding depression: analysis of whole-genome sequence data in 28 selfed progeny of *Eucalyptus grandis*. New Phytol. 209:600–611.

Hersh, E., J. A. Madjidian, S. Andersson, M. Strandh, W. S. Armbruster, and A. Lankinen, 2015. Sexual antagonism in the pistil varies among populations of a hermaphroditic mixed-mating plant. J Evol Biol 28:1321–34.

Holsinger, K., 1991. Mass-action models of plant mating systems: the evolutionary stability of mixed mating systems. Am Nat 138:602–622.

Huber, C. D., A. Durvasula, A. M. Hancock, and K. E. Lohmueller, 2018. Gene expression drives the evolution of dominance. Nat Commun 9:2750.

Igic, B. and J. W. Busch, 2013. Is self-fertilization an evolutionary dead end? New Phytol 198:386–97.

Igic, B. and J. R. Kohn, 2006. The distribution of plant mating systems: study bias against obligately outcrossing species. Evolution Int J Org Evolution 60:1098–103.

Immler, S., G. Arnqvist, and S. P. Otto, 2012. Ploidally antagonistic selection maintains stable genetic polymorphism. Evolution 66:55–65.

Jain, K. and S. Jain, 1970. Frequency-dependent selection under mixed selfing and random mating. Heredity 25:217–221.

Jarne, P. and J. R. Auld, 2006. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. Evolution 60:1816–24.

Johnson, M. T., S. D. Smith, and M. D. Rausher, 2009. Plant sex and the evolution of plant defenses against herbivores. Proc Natl Acad Sci USA 106:18079–84.

de Jong, T. and P. Klinkhamer, 2005. Evolutionary ecology of plant reproductive strategies. Cambridge University Press, Cambridge.

Jordan, C. Y. and T. Connallon, 2014. Sexually antagonistic polymorphism in simultaneous hermaphrodites. Evolution 68:3555–69.

Kamran-Disfani, A. and A. F. Agrawal, 2014. Selfing, adaptation and background selection in finite populations. J Evol Biol 27:1360–71.

Kellenberger, R. T., K. Byers, R. M. De Brito Francisco, Y. M. Staedler, A. M. LaFountain, J. Schonenberger, F. P. Schiestl, and P. M. Schluter, 2019. Emergence of a floral colour polymorphism by pollinator-mediated overdominance. Nat Commun 10:63.

Kimura, M. and T. Ohta, 1971. Theoretical aspects of population genetics. Monographs in population biology. Princeton University Press, Princeton.

Koenig, D., J. Hagmann, R. Li, F. Bemm, T. Slotte, B. Neuffer, S. I. Wright, and D. Weigel, 2019. Long-term balancing selection drives evolution of immunity genes in *Capsella*. eLife 8.

Landis, J. B., C. D. Bell, M. Hernandez, R. Zenil-Ferguson, E. W. McCarthy, D. E. Soltis, and P. S. Soltis, 2018. Evolution of floral traits and impact of reproductive mode on diversification in the phlox family (polemoniaceae). Mol Phylogenet Evol 127:878–890.

Lankinen, A., H. G. Smith, S. Andersson, and J. A. Madjidian, 2016. Selection on pollen and pistil traits during pollen competition is affected by both sexual conflict and mixed mating in a self-compatible herb. Am J Bot 103:541–52.

Lesaffre, T. and S. Billiard, 2020. The joint evolution of lifespan and self-fertilization. J Evol Biol 33:41–56.

Levene, H., 1953. Genetic equilibrium when more than one ecological niche is available. Am. Nat. 87:311–313.

Levin, D. A., 1975. Pest pressure and recombination systems in plants. Am Nat 109:437–457.

Lloyd, D., 1979. Some reproductive factors affecting the selection of self-fertilization in plants. Am Nat 113:67–79.

Mojica, J. P., Y. W. Lee, J. H. Willis, and J. K. Kelly, 2012. Spatially and temporally varying selection on intrapopulation quantitative trait loci for a life history trade-off in mimulus guttatus. Mol Ecol 21:3718–28.

Mérot, C., V. Llaurens, N. Normandeau, L. Bernatchez, and M. Wellenreuther, 2020. Balancing selection via life-history trade-offs maintains an inversion polymorphism in a seaweed fly. Nature comm. 11.

Nagylaki, T., 1983. Evolution of a large population under gene conversion. Proc Natl Acad Sci USA 80:5941–5.

Nei, M. and A. K. Roychoudhury, 1973. Probability of fixation and mean fixation time of an overdominant mutation. Genetics 74:371–80.

Nordborg, M., 2000. Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. Genetics 154:923–9.

Nordborg, N., B. Charlesworth, and D. Charlesworth, 1996. Increased levels of polymorphism surounding selectively maintained sites in highly selfing species. Proc Biol Sci 263:1033–1039.

Ohta, T. and M. Kimura, 1969. Linkage disequilibrium at steady state determined by random genetic drift and recurrent mutation. Genetics 63:229–238.

Olito, C., 2016. Consequences of genetic linkage for the maintenance of sexually antagonistic polymorphism in hermaphrodites. Evolution 71:458–464.

Otto, S. P. and T. Day, 2007. A biologist's guide to mathematical modeling in ecology and evolution. Princeton University Press.

Peters, M. A. E. and A. E. Weis, 2018. Selection for pollen competitive ability in mixed-mating systems. Evolution 72:2513–2536.

Petrone Mendoza, S., M. Lascoux, and S. Glémin, 2018. Competitive ability of *Capsella* species with different mating systems and ploidy levels. Ann Bot 121:1257–1264.

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

Pollak, E. and M. Sabran, 1992. On the theory of partially inbreeding finite populations. iii. fixation probabilities under partial selfing when heterozygotes are intermediate in viability. Genetics 131:979–85.

Rivkin, L. R., S. C. H. Barrett, and M. T. J. Johnson, 2018. The effects of plant sexual system and latitude on resistance to herbivores. Am J Bot 105:977–985.

Robertson, A., 1962. Selection for heterozygotes in small populations. Genetics 17:1291–1300.

Rocheleau, G. and S. Lessard, 2000. Stability analysis of the partial selfing selection model. J Math Biol 40:541–74.

Roze, D., 2009. Diploidy, population structure, and the evolution of recombination. Am Nat 174:S79–S94.

———, 2016. Background selection in partially selfing populations. Genetics 203:937–57.

Sicard, A. and M. Lenhard, 2011. The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. Ann Bot 107:1433–1443.

Sicard, A., N. Stacey, K. Hermann, J. Dessoly, B. Neuffer, I. Baurle, and M. Lenhard, 2011. Genetics, evolution, and adaptive significance of the selfing syndrome in the genus *Capsella*. Plant Cell 23:3156–71.

Sletvold, N. and J. M. Grindeland, 2007. Fluctuating selection on reproductive timing in digitalis purpurea. Oikos 116:476–481.

Slotte, T., K. M. Hazzouri, J. A. Agren, D. Koenig, F. Maumus, Y. L. Guo, K. Steige, A. E. Platts, J. S. Escobar, L. K. Newman, W. Wang, T. Mandakova, E. Vello, L. M. Smith, S. R. Henz, J. Steffen, S. Takuno, Y. Brandvain, G. Coop, P. Andolfatto, T. T. Hu, M. Blanchette, R. M. Clark, H. Quesneville, M. Nordborg, B. S. Gaut, M. A. Lysak, J. Jenkins, J. Grimwood, J. Chapman, S. Prochnik, S. Shu, D. Rokhsar, J. Schmutz, D. Weigel, and S. I. Wright, 2013. The *Capsella rubella* genome and the genomic consequences of rapid mating system evolution. Nat Genet 45:831–5.

43

Slotte, T., K. M. Hazzouri, D. Stern, P. Andolfatto, and S. I. Wright, 2012. Genetic architecture and adaptive significance of the selfing syndrome in *Capsella*. Evolution .

Snell, R. and L. W. Aarssen, 2005. Life history traits in selfing versus outcrossing annuals: exploring the 'time-limitation' hypothesis for the fitness benefit of self-pollination. BMC Ecol 5:2.

Stebbins, G. L., 1957. Self fertilization and population variability in higher plants. Am Nat 91:337–354.

Stone, J., 2002. Molecular mechanisms underlying the breakdown of gametophytic self-incompatibility. The Quarterly review of biology 77:17–32.

Takahata, N., 1990. A simple genealogical structure of strongly balanced allelic lines and trans-species evolution of polymorphism. Proc Natl Acad Sci USA 87:2419–2423.

Takahata, N. and M. Nei, 1990. Allelic genealogy under overdominant and frequency-dependent selection and polymorphism of major histocompatibility complex loci. Genetics 124:967–978.

Tazzyman, S. J. and J. K. Abbott, 2015. Self-fertilization and inbreeding limit the scope for sexually antagonistic polymorphism. J Evol Biol 28:723–9.

Tedder, A., S. Carleial, M. Golebiewska, C. Kappel, K. K. Shimizu, and M. Stift, 2015. Evolution of the selfing syndrome in *Arabis alpina* (brassicaceae). PLoS One 10:e0126618.

Tiffin, P. and D. A. Moeller, 2006. Molecular evolution of plant immune system genes. Trends Genet 22:662–70.

Waller, D. M. and D. Green, 1981. Implications of sex for the analysis of life histories. Am Nat 117:810–8130.

Weir, B. S., 1970. Equilibria under inbreeding and selection. Genetics 65:371–8.

Weir, B. S. and C. C. Cockerham, 1973. Mixed self and random mating at two loci. Genet Res 21:247–62.

Wright, S. I., S. Kalisz, and T. Slotte, 2013. Evolutionary consequences of self-fertilization in plants. Proc Biol Sci 280:20130133.

Zhao, L. and D. Waxman, 2016. The influence of genetic drift on the formation and stability of polymorphisms arising from negative frequency-dependent selection. J Theor Biol 391:51–64.