Affiliations

1 Ashworth Laboratories, Institute of Evolutionary Biology, The University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh, EH9 3JT, United Kingdom
2 Department of Molecular Biology and Genetics, Cornell University, Ithaca, New York 14853
3 School of Biological Sciences, Monash University, Clayton, Victoria, 3800, Australia

Received June 10, 2014
Accepted September 9, 2014

In hermaphrodites, pleiotropic genetic trade-offs between female and male reproductive functions can lead to sexually antagonistic (SA) selection, where individual alleles have conflicting fitness effects on each sex function. Although an extensive theory of SA selection exists for dioecious species, these results have not been generalized to hermaphrodites. We develop population genetic models of SA selection in simultaneous hermaphrodites, and evaluate effects of dominance, selection on each sex function, self-fertilization, and population size on the maintenance of polymorphism. Under obligate outcrossing, hermaphrodite model predictions converge exactly with those of dioecious populations. Self-fertilization in hermaphrodites generates three points of divergence with dioecious theory. First, opportunities for stable polymorphism decline sharply and become less sensitive to dominance with increased selfing. Second, selfing introduces an asymmetry in the relative importance of selection through male versus female reproductive functions, expands the parameter space favorable for the evolutionary invasion of female-beneficial alleles, and restricts invasion criteria for male-beneficial alleles. Finally, contrary to models of unconditionally beneficial alleles, selfing decreases genetic hitchhiking effects of invading SA alleles, and should therefore decrease these population genetic signals of SA polymorphisms. We discuss implications of SA selection in hermaphrodites, including its potential role in the evolution of “selfing syndromes.”

KEY WORDS: Adaptation, balancing selection, fitness trade-off, genetic hitchhiking, intralocus sexual conflict, sexual dimorphism.

Genetic trade-offs between male and female fitness are thought to play critical roles in the evolution of reproductive traits and gender-related differences. An allele is said to be “sexually antagonistic” (hereafter “SA”; Rice 1992) when its fitness effect is beneficial when present in one sex, and deleterious when present in the other (Rice 1984; Pischedda and Chippindale 2006; Delph et al. 2010). SA alleles may contribute substantially to the maintenance of genetic variation in life-history traits and fitness (e.g., Charlesworth and Hughes 1999; Bonduriansky and Chenoweth 2009; Connallon and Clark 2012), and they play important roles in theories of adaptation and genome evolution in dioecious species (i.e., those with separate sexes; see e.g., Kidwell et al. 1977; Charlesworth and Charlesworth 1980; Lande 1980; Rice 1984, 1987; Prout 2000; Albert and Otto 2005; Bergero and Charlesworth 2008; Bonduriansky and Chenoweth 2009; Patten and Haig 2009; Fry 2010; Patten et al. 2010; Arnvist 2011; Immler et al. 2012; Jordan and Charlesworth 2012; Connallon and Clark 2014a,b). This theory serves as an important stimulant for experimental research on the evolutionary consequences of sexual antagonism (e.g., Gibson et al. 2002; Pischedda and Chippindale 2006; Foerster et al. 2007; Delph et al. 2010; Harano et al. 2010; Mokkonen et al. 2011; Long et al. 2014).
where animals represent the primary taxonomic focus (at least been developed within the context of dioecious mating systems, hermaphrodites, the theory of SA genetic variation has largely evidence for SA trade-offs in hermaphrodite animals, and provides a recent review by Abbott (2011) synthesizes experimental evi-
dence for SA trade-offs in hermaphrodite animals, and provides a considerable scope for fitness trade-offs between the male and female functions of hermaphrodites.

Hermaphrodite mating systems are common among plant and animal species (Renner and Ricklefs 1995; Jarne and Auld 2006; Bachtrog et al. 2014), and current empirical data suggest that SA trade-offs are likely common in hermaphrodite populations. For example, studies of siring success in hermaphroditic plants often report evidence for trade-offs between male and female sex functions (Elle and Meagher 2000; Hodgins and Barrett 2008; Ellis and Johnson 2010; Rymer et al. 2010), with individual plants differentially contributing genetic material to subsequent generations via pollen versus ovules (i.e., they differ in “functional gender”; Lloyd 1980; Ennos and Dodson 1987; Devlin and Ellstrand 1990; Elle and Meagher 2000; see Ross 1990 for an early review). Additional studies report conflicting selection gradients on traits that mediate male versus female fitness (reviewed by Conner 2006). Extensive sexual dimorphism in dioecious plants (Delph 1996; e.g., dimorphism in flower morphology, vigor, niche differences, leaf distributions, and reproductive effort) suggests that each sex has a distinct set of optimal (fitness maximizing) trait values (Lloyd and Webb 1977), and implies a considerable scope for fitness trade-offs between the male and female reproductive structures of hermaphrodite individuals. A recent review by Abbott (2011) synthesizes experimental evidence for SA trade-offs in hermaphrodite animals, and provides a compelling argument for expecting SA variation in hermaphrodite populations.

Despite the evolutionary relevance of fitness trade-offs in hermaphrodites, the theory of SA genetic variation has largely been developed within the context of dioecious mating systems, where animals represent the primary taxonomic focus (at least implicitly). In contrast, a variety of plant-oriented models address fitness trade-off scenarios involving hermaphroditic sex functions (Gregorius 1982; Morgan 1992), and these models represent an important component of the evolutionary theory of plant mating systems. For example, special cases of SA polymorphisms play prominent roles in evolutionary theories of monoecy, dioecy, and gynodioecy (Charlesworth and Charlesworth 1978a,b; Gregorius et al. 1982), and scenarios of balancing selection in obligate outcrossing hermaphrodite populations (e.g., Morgan 1992). Several additional models identify general evolutionary criteria for balancing selection in partially self-fertilizing hermaphrodite populations (e.g., Ross 1977, 1984, 1985; Gregorius 1982, 1984; Ross and Gregorius 1983; Ziehe 1985; note, in particular, Gregorius 1982 and Ziehe 1982, who provide general conditions for protected polymorphism in partially selfing hermaphrodites). However, to our knowledge, no model fully explores the conditions maintaining SA polymorphism in a hermaphrodite population, including an explicit analysis of the effects of selfing rates, inbreeding depression, and arbitrary dominance coefficients for each sex function. Sex-specific dominance has a particularly strong influence on the theoretical predictions for dioecious species (e.g., Rice 1984; Patten and Haig 2009; Connallon and Clark 2010, 2014b; Fry 2010; Jordan and Charlesworth 2012), yet its effects in partially selfing hermaphrodite populations remain unclear. Finally, genetic drift is expected to play an important role in the evolutionary dynamics of SA alleles in dioecious species (Connallon and Clark 2012, 2013; Mullon et al. 2012), and may additionally influence the detectability of population genetic signals of balancing selection (Charlesworth 2006; Connallon and Clark 2013). Such finite population size considerations have yet to be incorporated into SA theory for hermaphrodites.

Here, we investigate the conditions for maintaining SA polymorphism in partially selfing hermaphrodites, and consider both deterministic and finite population size models of balancing selection via sexual antagonism. We consider in detail two idealized models of self-fertilization: (1) a model in which the selfing rates of individuals are independent of the genotypes at the SA locus; and (2) a model of frequency-dependent selfing that is directly influenced by segregating alleles at the SA locus. We subsequently explore how finite population size influences the maintenance of SA alleles, and consider how self-fertilization affects molecular population genetic signals associated with recent balancing selection at an SA locus. Our analysis provides a general extension of the theory of SA alleles, and draws new connections between theories of SA variation and the balancing selection literature of partially selfing hermaphrodites. Although our biological examples and discussion primarily focus on plants, our results should apply generally to plant and animal populations of simultaneous hermaphrodites.
Table 1. Values of male and female fitness components through each genotype.\(^1\).

| Genotype | \(AA\) | \(Aa\) | \(aa\) |
|----------|-------|------|------|
| Female fitness value | 1 | \(1 - h_f t\) | \(1 - t\) |
| Male fitness value | \(1 - s\) | \(1 - h_m s\) | 1 |

\(^1\)Following convention, we assume that \(s\) and \(t\) are each positive, which requires that they fall within the range \(0 < s, t < 1\).

Models
MECHANISMS OF SEXUAL ANTAGONISM IN OUTCROSSING AND SELF-FERTILIZING POPULATIONS

Models of obligate outcrossing individuals (dioecious or hermaphroditic) can typically ignore the specific traits involved in SA trade-offs. What matters most in these models is the relative ranking of genotypes in terms of their sex-specific fitness values. For example, given a set of specified sex-specific fitness values for a single locus, two-allele system (e.g., p. 172 of Kidwell et al. 1977; our Table 1), one can develop simple equations to represent the relative contributions of each genotype to the pool of female and male gametes that make up the next generation of individuals of the population. The simple inheritance structure of obligate outcrossing populations circumvents the need to specify specific traits that mediate the SA fitness trade-off. In the absence of selfing, female and male fitness components in hermaphrodites are analogous to the relative fitness rankings of genotypes in males and females of dioecious species. In short, in the limiting cases where selfing rates are zero, our models of SA polymorphism are as generally applicable as their dioecious analogues.

When individuals of hermaphrodite populations reproduce through outcrossing and selfing, the evolutionary dynamics of SA selection at a locus can be influenced by the specific traits that underlie the SA trade-off, as well as the mechanism by which self-fertilization occurs. For example, if variation in male and female fitness components directly influences the genotype-specific rates of self-fertilization, then the relationship between the evolutionary dynamics at an SA locus, and population-wide selfing patterns, become interdependent. The potential for evolutionary feedback between self-fertilization patterns and the dynamics of SA alleles requires that we specify mechanistic details of the SA trade-off, and the form of self-fertilization that is practiced by individuals of the population.

For populations with partial self-fertilization, we model sexual antagonism as a trade-off between investment in ovule production, and traits that are involved in pollination (i.e., traits affecting overall pollen production or traits that facilitate pollen transport and performance under conditions of competitive fertilization). Thus, the female component of fitness varies among genotypes because they differ in their relative resource allocations to ovule production, whereas the male component of fitness can have a more general interpretation (with strict outcrossing, we interpret variation in female fitness with similar generality, encompassing influences on female fertility through gamete production and ability to sequester pollen for fertilization). Below, we develop generalized evolutionary recursions for a biallelic SA locus with alleles affecting such an allocation trade-off. We then analyze the model under two different forms of self-fertilization (each inspired by common forms of selfing in plant species): (1) a fixed selfing model in which selfing rates are independent of the genotype at the SA locus; and (2) a “mass-action” selfing mechanism (Holsinger 1991) in which selfing rates become frequency dependent with respect to alleles at the SA locus. These selfing scenarios are described in greater detail below.

DETERMINISTIC POPULATION DYNAMICS

We consider evolution at a single diploid locus with two alleles and discrete generations. The life cycle follows the order of “birth,” then viability selection due to inbreeding depression in self-fertilized individuals, then selection on male and female fitness components prior to fertilization of ovules that give rise to the next generation. We arbitrarily define \(A\) as the female-benefit allele, and \(a\) as the male-benefit allele. The relative contributions of the three genotypes (\(AA, Aa, aa\)) to male and female fitness components are presented in Table 1. These follow standard formulations for SA genotypes in a dioecious population (see Kidwell et al. 1977; Patten and Haig 2009; Fry 2010).

We track genotype frequencies of adults in the population (i.e., immediately following viability selection among self-fertilized and outcrossed individuals), with \(F_{ij}\) representing the adult frequency of genotype \(ij\) (\(ij = \{AA, Aa, aa\}\)). The frequency of \(A\) among adults is \(F_{AA} + F_{Aa}/2 = q\), and the frequency of \(a\) is \(F_{aa} + F_{Aa}/2 = 1 - q\). The relative contribution of each genotype to the production of ovules is given by:

\[
F_{AA}^f = \frac{F_{AA}}{W_f},
\]

\[
F_{Aa}^f = \frac{F_{Aa}(1 - h_f t)}{W_f},
\]

\[
F_{aa}^f = \frac{F_{aa}(1 - t)}{W_f},
\]

where \(W_f = 1 - h_f F_{aa} - t F_{aa}\). The frequency of \(A\) among the total pool of ovules is \(q_f = F_{AA}/A + F_{Aa}/2\). Similarly, the relative contribution of each genotype to the production of exported pollen (i.e., the pool of pollen available for outcross fertilization) is as follows:

\[
F_{AA}^m = \frac{F_{AA}(1 - s)}{W_m},
\]

\[
F_{Aa}^m = \frac{F_{Aa}(1 - h_m s)}{W_m},
\]

\[
F_{aa}^m = \frac{F_{aa}}{W_m}.
\]
where $W_m = 1 - b_m sF_{AA} - sF_{AA}$. The frequency of $A$ among in
this pool of pollen is $q_m = F_{AA}^m + F_{AA}^m / 2$.

Let $\theta_i$ represent the relative proportion of ovules of genotype
$i j$ that are self-fertilized ($1 - \theta_{ij}$ are outcrossed), and $\delta$ represent
the viability cost suffered by self-fertilized individuals due to
inbreeding depression (Charlesworth and Charlesworth 1987).

The frequency of each genotype in the adult population of the
next generation will then be:

\[
F'_{AA} = \frac{(1 - \theta_{AA}) F'_{AA} + \left(1 - \theta_{AA} \right) \frac{F_{AA}^f}{2}}{W_{TOT}} q_m + \left[ \theta_{AA} F'_{AA} + \frac{\theta_{AA} F_{AA}^f}{4} \right] \left(1 - \delta \right)
\]

\[
F'_{Aa} = \frac{(1 - \theta_{AA}) F'_{Aa} + \left(1 - \theta_{AA} \right) \frac{F_{Aa}^f}{2}}{W_{TOT}} (1 - q_m) + \left[ (1 - \theta_{aa}) F'_{aa} + \left(1 - \theta_{aa} \right) \frac{F_{aa}^f}{2} \right] q_m + \frac{\theta_{AA} F_{AA}^f}{2} \left(1 - \delta \right)
\]

\[
F'_{aa} = \frac{(1 - \theta_{aa}) F'_{aa} + \left(1 - \theta_{aa} \right) \frac{F_{aa}^f}{2}}{W_{TOT}} (1 - q_m) + \left[ \theta_{aa} F'_{aa} + \frac{\theta_{AA} F_{AA}^f}{4} \right] \left(1 - \delta \right)
\]

where $W_{TOT} = 1 - \delta \left( \theta_{AA} F'_{AA} + \theta_{Aa} F'_{Aa} + \theta_{aa} F_{aa}^f \right)$. Note that
in the absence of selfing ($\theta_{ij} = 0$), the recursion equations be-
come identical to those of dioecious models of SA selection (e.g.,
Kidwell et al. 1977), a point that we expand on in the Results. The
allele frequency change across a single generation is therefore:

\[
\Delta q = \frac{q_f + q_m - 2q - (1 - q_m) \left[ \theta_{AA} F'_{AA} + \theta_{Aa} \frac{F_{AA}^f}{2} \right] - q_m \left[ \theta_{aa} F'_{aa} + \frac{\theta_{AA} F_{AA}^f}{2} \right]}{2W_{TOT}}
\]

\[
-2\delta \left[ \theta_{AA} F'_{AA} + \theta_{Aa} \frac{F_{AA}^f}{2} \right] - q \left[ \theta_{aa} F'_{aa} + \frac{\theta_{AA} F_{AA}^f}{2} \right] \]

\[
\frac{1}{2W_{TOT}}
\]

To determine the criteria for evolutionary invasion of each allele (i.e., from low initial frequency), and for the maintenance of
polymorphism, we determined the stability of the equilibria with
$A$ and $a$ alleles each fixed in the population. Invasion is favored,
for each equilibrium, when the leading eigenvalue of the Jacobian
matrix exceeds 1.

MODELS OF SELF-FERTILIZATION

The selfing rate among individuals of the population takes two
basic forms. In our first model (which we refer to as “fixed
selfing”), selfing rates are independent of genotype at the SA
locus, and in this case, the selfing rate is defined as $\theta = \theta_{AA} = \theta_{Aa} = \theta_{aa}$. This characterization corresponds to resource-limited
reproduction (Knight et al. 2005), and assumes self- and outcross
male gametes do not directly compete for access to female ga-
metes. Such scenarios occur under conditions of “prior selfing”
in plants (Lloyd 1992; Lloyd and Schoen 1992; where self-pollen
fertilizes ovules before outcross pollen arrives; see Jordan and
Otto 2012) and “delayed selfing,” where selfing occurs following
opportunities for fertilization via outcross pollen (Lloyd 1992;
Lloyd and Schoen 1992). Lloyd (1979, 1992) noted that prior
and delayed selfing involve contrasting benefits and costs that
affect the evolution of selfing rate. However, these relative benefits
are unimportant here because we do not consider the evolution of
selfing rate per se; rather, we address the population genetic
consequences of selfing for the evolution of a SA locus.

In the second model, self- and outcross pollen compete for
female gametes (as in so-called “mass-action” mating models;
Holsinger 1991), and frequency-dependent selfing rates arise
when genotypes vary in their export of male gametes (Ross 1990).
Self- and outcross male gametes frequently arrive at a female
organ simultaneously, at least in flowering plants (Goodwillie
et al. 2005), so competition between them may commonly occur.
We assume that individuals of all genotypes allocate an identical
fraction of pollen to selfing, that genotypes differ with respect to the number of male and female gametes produced, and in contrast to the fixed-selfing model, SA selection specifically arises when there is a genetic trade-off between female and male gamete production. Under this mass-action framework, the selfing rate at genotype $ij$ is given by:

$$
\theta_{ij} = \frac{e^{\pi} P_{ij}}{e^{\pi} P_{ij} + P_{ij}(1 - x)(1 - F_{AA}s - F_{Aa}h_{as})}
$$

$$
= \frac{x P_{ij}}{(1 - x) W_{m}}.
$$

where $P_{AA} = 1 - s$, $P_{Aa} = 1 - h_{as}$, $P_{aa} = 1$, $x$ is the fraction of pollen allocated to selfing ($1 - x$ is allocated to outcrossing), $\pi$ is the fraction of selfing-allocated pollen that is available for self-fertilization (e.g., due to differences in competitive ability of self versus outcross pollen tubes), and $\pi$ is the fraction of outcross pollen that successfully reaches a female reproductive organ ($1 - \pi$ is lost during transport between plants; see Harder and Wilson 1998). Equation (2) can be incorporated into equation (1) to evaluate evolutionary dynamics of the SA locus under mass-action self-fertilization.

**FINITE POPULATION EXTENSION**

With finite population size, eventual fixation of one of the alleles (in a population segregating for both) is guaranteed. Genetic variation is then maintained by recurrent mutation, balanced against its loss due to the combined effects of genetic drift and selection. Using standard diffusion theory, the equilibrium distribution of $q$ in a finite population will be:

$$
f(q) = \frac{k}{q(1 - q)} e^{2 \int \frac{M}{V} dq}.
$$

where $k$ is chosen so that $\int f(q) dq = 1$ (Wright 1945), $V$ is the variance of allele frequencies due to binomial sampling in a Wright–Fisher population, and $M$ represents the expected frequency change due to selection and mutation. For a diploid (autosomal) locus with effective population size $2N_e$, $V = q(1 - q)/(2N_e)$. $M = \Delta q + (1 - q)u_A - q_uo$, where $u_A$ is the mutation rate from $A$ to $a$, $u_o$ is the mutation rate from $A$ to $a$, and $\Delta q$ is the expected change in frequency due to selection (i.e., eq. (1)). Equation (3) can be used to characterize the degree to which finite populations approach frequency states predicted by deterministic models, or in our case of interest, the “efficacy” with which balancing selection at an SA locus will maintain alleles near their intermediate deterministic equilibrium frequencies (see Ewens 2004, pp. 26–27; Connallon and Clark 2014b).

Assuming that selection coefficients are small ($s, t << 1$), and selfing rates are approximately equal among the genotypes at the SA locus, we were able to obtain a general expression for the expected allele frequency change due to selection:

$$
\Delta q \approx \frac{q(1 - q)}{2(1 - C)} \frac{\hat{h}f + q(1 - 2h_f)}{W_f} - \frac{q(1 - q)}{2(1 - C)} \frac{h_{as} + q(1 - 2h_{as})}{W_m},
$$

where $C = \alpha(1 - \hat{h})(1 - s\theta)$ represents the proportion of adults in the population that are produced by selfing (i.e., it represents an “effective” rate of selfing; Supporting Information Materials). Equation (4) simplifies substantially for the two biologically interesting cases of dominance that we evaluate below. With additive fitness effects of the SA locus ($h_{as} = h_f = \frac{1}{2}$), the allele frequency change is approximately:

$$
\Delta q \approx \frac{\hat{q}(1 - q)(\hat{q} - q)}{(2 - C)},
$$

where $\hat{q} = \frac{1}{2} - \frac{1}{2} C(1 - s\theta) (1 - 2s\theta)$. Note that, when criteria for balancing selection are met, $\hat{q}$ represents the equilibrium frequency of $A$. When fitness costs of each allele are partially recessive (i.e., there is a beneficial “dominance reversal”): $h_{as} = h_f = h < \frac{1}{2}$; see Fry 2010), equation (4) simplifies to:

$$
\Delta q \approx \frac{(1 - C)(1 - 2h)[s + 2s\gamma(1 - 2h)]}{(2 - C)} q(1 - q)(\hat{q} - q),
$$

where:

$$
\hat{q} = \frac{2s(1 - C) - [s + C(1 - C)]}{2(1 - C)(1 - 2h)(1 - 2s)}.
$$

Note that equations (4a) and (4b) take the general form $\Delta q = f(s, t, C)(1 - q)(\hat{q} - q)$, where $f(s, t, C) = st(2 - C)$ when SA alleles have additive effects ($h_{as} = h_f = 0.5$), $f(s, t, C) = [s + t + C(t - s)](1 - 2h)(1 - C)/(2 - C)$ for the dominance reversal case ($h_{as} = h_f = h < \frac{1}{2}$), and terms of $\hat{q}$ are specified above for each dominance scenario. These can alternatively be expressed as: $f(s, t, C; h_{as} = h_f = \frac{1}{2}) = (S^2 - \gamma)^2/(2 - C)$ and $f(s, t, C; h_{as} = h_f = \frac{1}{2}) = (2S - \gamma C)(1 - 2h)(1 - C)/(2 - C)$, where $S = s/2 + t/2$ is the sex-averaged selection coefficient, and $\gamma = s - t$ is the fitness effect differential (or asymmetry) between the sexes. The latter can be calculated under the assumption of weak selection ($s, t << 1$):

$$
\gamma \approx \left\{ \begin{array}{ll}
S[C(2 - S) + S(1 - 2\hat{q})] & \text{for } h_{as} = h_f = \frac{1}{2} \\
2S \frac{1 - 2(1 - C)(1 - 2\hat{q})}{(1 - C)(1 - C)(1 - 2\hat{q})} & \text{for } h_{as} = h_f = \frac{1}{2}.
\end{array} \right.
$$

These expressions become useful in our subsequent contrasts between obligate outcrossing and partially selfing populations.

**HITCHHIKING MODEL (SIGNS OF RECENT BALANCING SELECTION)**

Alleles under short-term balancing selection—where one of the alleles has recently spread from low to intermediate frequency
within the population—can leave a characteristic population genetic signal of a partial selective sweep (Charlesworth 2006). We consider a population that is initially fixed for one of the SA alleles (A or a), and where the other SA allele invades the population and evolves from an initial population frequency of \( q_0 = 1/(2N) \), toward its deterministic equilibrium frequency (\( \hat{q} \)). Evolution at the SA locus affects evolution at a nearby neutral locus with alleles B and b.

Following prior hitchhiking theory, we model the evolution of the four-haplotype system \((AB, Ab, aB, ab)\), assuming selection coefficients are small, and the effective strength of selection at the selected locus is sufficiently strong that successfully invading mutations (i.e., those escaping the initial, “stochastic” phase of a sweep; Barton 1998) and subsequently evolve pseudodeterministically (Barton 1998; Betancourt et al. 2004; Otto and Day 2007). The assumption of effectively strong selection (i.e., large \( N_f(s, t, C) \)) defines the parameter space over which hitchhiking effects are relevant. The parameters remain the same for the antagonistically selected locus (i.e., two alleles, A and a, and parameters, s, t, h, and C). The recombination rate between the loci is \( r \) per meiosis. Haplotype frequencies in the population are \( q_1 = [AB], q_2 = [Ab], q_3 = [aB], \) and \( q_4 = [ab] \), and the allele frequencies are \( q_1 = q_1 + q_2 \) and \( q_2 = q_1 + q_3 \). Linkage disequilibrium between A and B alleles is defined as \( D = q_1q_4 - q_2q_3 \).

With small selection coefficients and tight linkage between locus A and B \((r << \frac{1}{2})\), the expected haplotype frequency changes are given by (see Supporting Information Material):

\[
\begin{align*}
\Delta q_1 & \approx f(s, t, C)q_1(1 - q_A)(\hat{q} - q_A) - \frac{2rD(1 - C^2)}{2 - C}, \\
\Delta q_2 & \approx f(s, t, C)q_2(1 - q_A)(\hat{q} - q_A) + \frac{2rD(1 - C^2)}{2 - C}, \\
\Delta q_3 & \approx -f(s, t, C)q_3q_4(\hat{q} - q_A) + \frac{2rD(1 - C^2)}{2 - C}, \\
\Delta q_4 & \approx -f(s, t, C)q_4q_3(\hat{q} - q_A) - \frac{2rD(1 - C^2)}{2 - C}.
\end{align*}
\]

The change in linkage disequilibrium across a single generation is given by \( \Delta D = (q_1 + \Delta q_1)(q_4 + \Delta q_4) - (q_2 + \Delta q_2)(q_3 + \Delta q_3) - D \). Each sweep is initiated with a single AB haplotype introduced into a population that is otherwise fixed for ab (invasion of a female-benefit allele), or an ab haplotype introduced into a population fixed for AB (invasion of a male-benefit allele). To calculate trajectories of the four haplotypes, we iterated equation (6) with adjusted starting frequency \( 1/[4N_f(s, t, C)\hat{q}] \) when AB is the invading haplotype, and adjusted starting frequency \( 1/[4N_f(s, t, C)(1 - \hat{q})] \) when ab is the invading haplotype. This adjustment accounts for the initial stochastic phase of each sweep, in which successfully invading alleles experience a faster-than-exponential rate of increase (see Maynard Smith 1971; Barton 1998; Betancourt et al. 2004; Connallon and Clark 2013; Orr and Unkless 2014). Results from this approach match up extremely well with stochastic simulations that formally integrate random drift via multinomial sampling of genotypes (e.g., Betancourt et al. 2004; Connallon and Clark 2013).

**Results**

**CRITERIA FOR BALANCED POLYMORPHISM IN HERMAPHRODITES**

**Fixed self-fertilization model**

When selfing rates are independent of the genotype at the SA locus \((\theta = \theta_A = \theta_a = \theta_m); \) see Models section, above), the expected allele frequency change across a generation reduces to:

\[
\Delta q = \left(1 - C\right)\frac{\Delta q_f + \Delta q_m}{2} + C\Delta q_f,
\]

where \( \Delta q_m = q_m - q \), and \( \Delta q_f = q_f - q \). Terms \( \Delta q_m \) and \( \Delta q_f \) partition the allele frequency changes that arise from selection on the male versus female components of fitness, and they are proportional to relative strength of selection through these fitness functions. From its structure, equation (7) illustrates that self-fertilization generates an asymmetry in the relative importance of male and female selection on the overall allele frequency change in the population. In the limit of obligate outcrossing \((C = 0)\), male and female selection equally contributes to allele frequency change, as in dioecious models of SA selection. As the effective rate of self-fertilization increases away from zero \((C > 0)\), this symmetry is broken, and allele frequency change becomes increasingly sensitive to selection through female function. This asymmetry follows from the dampened effect of pollen competition among the self-fertilized component of reproduction. Genotypes that invest heavily in pollination traits contribute disproportionately (via male function) to offspring produced by outcrossing, whereas the genotypes that invest heavily in ovule production disproportionately contribute (via female function) to both selfed and outcrossed offspring. Selfing increases the number of routes through which female-benefit alleles are preferentially transmitted (see Fisher 1941); the net transmission rate of female-benefit alleles is enhanced because they are enriched among the maternal contribution to selfed and outcrossed progeny.

Parameter criteria for balancing selection and the selective maintenance of both SA alleles reflect the female bias that is introduced by self-fertilization. For simplicity, and ease of comparison with analogous results from dioecious theory, we focus on two biologically motivated cases of dominance: (1) the case where SA alleles have additive fitness effects \((h = h_m = h_f = \frac{1}{2})\) that often applies for alleles with small to moderate fitness effects (Phadnis and Fry 2005; Agrawal and Whitlock 2011); and (2) the case where sex-specific fitness costs of SA alleles are partially
masked within heterozygotes \((h = h_m = h_f < \frac{1}{2})\), as expected under biologically plausible fitness landscape models (see Manna et al. 2011; Conallon and Clark 2014b; for a recent empirical example, see Posavi et al. 2014). Note that these dominance cases are representative of broader dominance scenarios of “parallel dominance” \((h_m + h_f = 1)\) and “dominance reversals” \((h_m, h_f < \frac{1}{2})\;\text{see Kidwell et al. 1977; Curtsinger et al. 1994; Hedrick 1999; Prout 2000; Fry 2010; Aronvist 2011}\). With additive fitness effects, balancing selection requires that:

\[
\frac{t(1 + C)}{1 + t - C(1 - t)} < s < \frac{t(1 + C)}{(1 - t)(1 - C)},
\]

which, under obligate outcrossing, reduces to the classic dioecious result: \(t/l(1 + t) < s < t/(1 - t)\). Under dominance reversal conditions and approximating to first order in \(s\) and \(t\) (valid for small selection coefficients, e.g., \(s, t < 0.1\)), the condition for balancing selection is:

\[
I \left[ \frac{1 + C}{1 - C} \right] \left[ \frac{C + 2h(1 - C)}{2 - C - 2h(1 - C)} \right] < s < I \left[ \frac{1 + C}{1 - C} \right] \left[ \frac{2 - C - 2h(1 - C)}{C + 2h(1 - C)} \right],
\]

which, for obligate outcrossing, converges to the dioecious result: \(th(1 - h) < s < t(1 - h)/l\) (e.g., Conallon and Clark 2010).

An analysis of equations (8) and (9) (see Supporting Information Material) demonstrates three effects of self-fertilization. First, self-fertilization always constrains the criteria for invasion of male-beneficial alleles (Fig. 1). Second, selfing tends to expand invasion conditions for female-benefit alleles, but this effect is sensitive to dominance. When dominance is partially recessive (roughly \(0.18 < h < \frac{1}{2}\) when \(C\) is small, and \(\sim 0 < h < \frac{1}{2}\) as \(C\) approaches one), selfing expands the parameter space that permits invasion of female-benefit alleles. In contrast, selfing restricts female-benefit invasion criteria under strong recessivity (i.e., \(h < [3 - 2C - \sqrt{5 - 4C}]/(4 - 4C)\); see Supporting Information Material). This effect of strong recessivity makes sense in light of the following constraint within the model. In the limiting case of \(h = 0\), with complete outcrossing, female-benefit (and male-benefit) alleles will invade for any set of \(s\) and \(t\) values. Under selfing, this situation cannot possibly improve. By reducing the frequency of heterozygotes in the population, and thereby reducing the masking effect of dominance on the expression of male-deleterious alleles, selfing must somewhat dampen the ability of female-benefit alleles to invade. In contrast, with additive effects (\(h = 1/2\)), where masking does not apply, selfing benefits invasion of female-benefit alleles due to the dampened effects of pollen competition as discussed above. This tension, under selfing, between masking of recessive deleterious fitness effects and the intrinsic transmission advantage of female-benefit alleles, accounts for the threshold \(h\) at which selfing impedes or facilitates invasion of female-benefit alleles. Selfing always impedes invasion of male-benefit alleles, for all relevant values of \(h\), because it uniformly dampens the opportunity for selection through male reproductive function. Third, selfing reduces the parameter range that gives rise to a balancing selection (Fig. 2). When we consider the entire conceivable parameter space for \(s\) and \(t\) \((0 < s, t < 1\), for the additive model), modest selfing will only slightly reduce the parameter space for balancing selection (Fig. 2, solid black curve). However, if we limit our focus to mutations of small effect \((0 < s, t < 0.1\), which is probably typical; e.g., Eyre-Walker and Keightley 2007), then the balancing selection parameter space declines sharply under moderate and high rates of self-fertilization. The latter result holds qualitatively for a wide range of dominance conditions.

**Mass-action selfing model**

Under the mass-action model (Holsinger 1991; see Models section, above), rates of self-fertilization depend on population composition. In the absence of self-fertilization \((t_l = 0)\), the general recursion (in eq. (1)) again converges to \(\Delta q = (\Delta q_m + \Delta q_l)/2\), which is identical to the SA allele frequency dynamics in dioecious species (e.g., Kidwell et al. 1977). Under partial self-fertilization of individuals in the population, and assuming that polymorphism is segregating at the SA locus, self-fertilization rates of the population become frequency-dependent, and genotypes differ in their selfing rates. This genotype-specific selfing arises because mass-action permits competitive fertilization between self and outcross pollen. Genotypes that are less competitive in pollination contests (i.e., produce less pollen) are less likely to self-fertilize than those that invest heavily in pollen.
where recall that \( x \) is the fraction of pollen allocated to selfing, \( s \) is the fraction of selfing-allocated pollen that is available for self-fertilization (e.g., due to differences in competitive ability of self versus outcross pollen tubes), and \( \pi \) is the fraction of outcross-allocated pollen that successfully reaches a female reproductive organ. These criteria for balancing selection are qualitatively similar to those of the fixed selfing scenario, in that self-fertilization constrains invasion of male-beneficial alleles, facilitates invasion of female-beneficial alleles (like the fixed-selfing case, this always holds true for mildly recessive deleterious effects: approximately \( 0.3 < h < 0.5 \)), and reduces the parameter space conducive to balancing selection (Fig. 3). Relative to the fixed-selfing case, competition between self and outcross pollen increases the transmission rate of the \( a \) allele to self-fertilized offspring. This effect somewhat dampens the severity of the bias toward female-beneficial alleles, but is generally insufficient to remove the bias altogether.

**BALANCING SELECTION IN FINITE POPULATIONS**

To evaluate the effects of finite population size on the maintenance of SA alleles in hermaphrodites, we used diffusion equations (following Wright 1945; Charlesworth and Charlesworth 2010) to analyze the equilibrium distribution of SA alleles under three simplifying assumptions. First, selection coefficients are assumed to be small \( (s, t << 1) \); with increasingly strong selection, the evolutionary dynamics of SA alleles will approach those predicted by the exact deterministic model. Second, we assume that self-fertilization rates are the same among \( AA, Aa, \) and \( aa \) individuals (as in the fixed-selfing model), which is generally reasonable under weak SA selection. Finally, because our focus is on the role of balancing selection in maintaining genetic variation, we ignore the effects of recurrent mutation on the stationary distribution (our results about the efficacy of balancing selection nevertheless remain applicable with recurrent mutation).

Like other models of balancing selection (e.g., Robertson 1962; Ewens and Thomson 1970; Connallon and Clark 2012, 2013; Mullon et al. 2012), the expected frequency change of SA alleles can be expressed using the general form, \( f(s, t, C)q(1 - q) \) \((\dot{q} - q)\), where \( f(s, t, C) \) and \( \dot{q} \) are functions of selection and selfing parameters (see Models section, above). Because we are primarily interested in balancing selection, we focus below on cases where...
balancing selection. Under the combined effects of genetic drift and selection, the stationary distribution of \( q = [A] \) is:

\[
f(q) = \frac{k}{q(1 - q)} \exp \left[ -2N_e f(s, t, C)(\hat{q} - q)^2 \right],
\]

where the constant \( k \) is chosen so that \( \int f(q) dq = 1 \) (Robertson 1962; Ewens 2004, p. 26). The term \( 2N_e f(s, t, C) \) is proportional to the “efficacy” of balancing selection—the degree to which selection is able to maintain the SA alleles near their deterministic equilibrium values. To quantify the role of selfing on the efficacy of selection, let \( N \) be the population size of a randomly mating population with no selfing; the effective population size of a partially selfing population is \( N_e = N(2 - C)/2 \) (Charlesworth 2009).

Self-fertilization has two general consequences for the efficacy of balancing selection (see Fig. 4, which plots the relative reduction of \( 2N_e f(s, t, C) \) in a partially selfing population, relative to an obligate outcrossing (or dioecious) population with an effective size of \( N = N_e \), and with the same sex-averaged strength of selection: \( S = (s + t)/2 \)). First, selfing generally reduces the efficacy of selection, and the magnitude of this reduction is greater under dominance reversal conditions. Dominance reversals facilitate the maintenance of polymorphism because they generate net overdominance for fitness. By reducing the population fraction of \( Aa \) heterozygotes, selfing severely reduces evolutionary relevance of overdominant fitness effects. This likely accounts for the greater sensitivity of dominance reversal models to selfing. Second, reductions in the efficacy of selection are more severe when the deterministic equilibrium frequency of \( A \) is closer to zero than to one (the pattern arises for additive and dominance reversal cases, but is more apparent for the latter). This reflects an interaction between self-fertilization and asymmetries in selection between sex functions. Smaller deterministic equilibrium values are associated with a selection asymmetry between the sexes (weaker selection in female than male function), which reduces the population’s ability to increase the frequency of \( A \) alleles when they are rare.

**Figure 4.** Self-fertilization reduces the efficacy of balancing selection at an SA locus. Results plot the reduction of \( 2N_e f(s, t, C) \) in a partially selfing population, relative to an obligate outcrossing (or dioecious) population with an effective size of \( N = N_e \). Outcrossing and partially selfing populations are assumed to have the same sex-averaged strength of selection at the SA locus: \( S = (s + t)/2 \). Representative results are shown for \( S = 0.01 \), and are based on equations developed in the Supporting Information Material.

\( 0 < \hat{q} < 1 \), where \( \hat{q} \) represents the deterministic equilibrium under balancing selection. Under the combined effects of genetic drift and balancing selection, the stationary distribution of \( q = [A] \) is:

**POLYMORPHISM IN HERMAPHRODITES**

For recently established balanced polymorphisms, the invasion of one of the alleles from low initial frequency can generate a detectable “hitchhiking” signature of extended homozygosity (absence of neutral diversity) among haplotypes that carry the invading (derived) allele (Charlesworth 2006; Connallon and Clark 2013). To examine the strength of this signal for recently established SA alleles, we developed a two-locus hitchhiking model and evaluated the coevolutionary dynamics between the SA locus and a linked, neutral locus with alleles \( B \) and \( b \) (the SA and neutral loci are separated by a recombinational distance of \( r \); see Models section). Let \( q_A \) and \( q_B \) represent the frequencies of \( A \) and \( B \) alleles, respectively, and \( D \) represent the linkage disequilibrium between them. Assuming tight linkage and weak selection \((r, s, t \text{ small})\), the expected rates of change in \( q_A, q_B, \) and \( D \) are as follows:

\[
\begin{align*}
\Delta q_A & \approx f(s, t, C)q_A(1 - q_A)(\hat{q} - q_A), \\
\Delta q_B & \approx f(s, t, C)D(\hat{q} - q_A), \\
\Delta D & \approx f(s, t, C)D(\hat{q} - q_A)(1 - 2q_A) - 2rD(1 - C^2)/2 + O(\Delta q_A, \Delta q_B),
\end{align*}
\]

where terms \( f(s, t, C) \) and \( \hat{q} \) are the same as defined above (in the absence of selfing, these expressions are identical to those for dioecious populations; see eq. (7) of Connallon and Clark 2013).

The strength of the hitchhiking effect primarily depends on the initial rate of increase of the invading allele relative to the decay rate of initial LD between the selected and neutral loci (Maynard Smith and Haigh 1974; Coop and Ralph 2012). In the context of our model, invasion rates of SA alleles are dampened by self-fertilization, but so too is the rate of recombination between \( ABlab \) and \( aBlAb \) haplotypes. For the additive model of balancing selection \((h_m = h_f = 5\%)\), we find that these factors roughly cancel out, and the magnitude of the hitchhiking effect is relatively
Figure 5. Selfing reduces the hitchhiking signal of SA alleles maintained by net overdominant selection. The x-axis represents the distance (based on crossover frequency per meiosis) between the SA locus (at position $r = 0$) and linked neutral loci. The y-axis shows the scaled linkage disequilibrium between the SA locus (at position $r = 0$) and linked neutral loci, given by the correlation, $D[q_A(1 - q_A)q_B(1 - q_B)]^{p^2}$. Representative results are shown, for the specific case of invading female-benefit alleles, with intermediate deterministic equilibrium at $q = 0.5$, and sex-specific dominance of $h_m = h_f = \frac{1}{2}$. Results are based on iterations of equation (6), with adjusted initial haplotype frequencies as described in the Models section (directly following eq. (6)). Allele frequency values and LD were assessed after the derived SA allele had evolved 95% of the distance to equilibrium (i.e., to $0.95\hat{q}$).

Insensitive to the selfing rate. In contrast, for a SA balancing selection model with beneficial dominance reversal ($h = h_m = h_f < \frac{1}{2}$), hitchhiking effects are reduced by self-fertilization (Fig. 5). This makes sense for the dominance reversal case because balancing selection is generated by a net heterozygote advantage at the SA locus. By reducing the frequency of $Aa$ heterozygotes in the population (relative to Hardy–Weinberg expectations), selfing greatly reduces the invasion rate of rare SA alleles. The reduction of invasion rate is more pronounced than the accompanying reduction of recombination events between $AB/ab$ individuals. Consequently, partial sweeps generate weaker hitchhiking effects in partially selfing hermaphrodites.

**Discussion**

Genetic correlations between male and female reproductive organs in hermaphrodites provide a plausible biological basis for widespread fitness trade-offs between male and female functions, which are evolutionarily coupled by way of their shared genetic basis (Abbott 2011). Although theoretical and empirical research on SA genetic variation has traditionally focused on dioecious animals, the evolutionary genetics literature of simultaneously hermaphrodite plants provides many potential examples of trade-offs between male and female sex functions (see Introduction), and such trade-offs have recently been emphasized in hermaphrodite animals (e.g., Abbott 2011; Sprenger et al. 2012; Monro and Marshall 2014).

Our study integrates and extends previous population genetics theory of balancing selection under partial selfing (e.g., Kimura and Ohta 1971; Charlesworth and Charlesworth 1978a,b; Gregorius 1982; Ziehe 1982; Ross and Gregorius 1983), and the theory of SA genetic variation in species with separate sexes (e.g., Kidwell et al. 1977; Patten and Haig 2009; Connallon and Clark 2012; Mullon et al. 2012; Jordan and Charlesworth 2012). We have identified general parameter conditions of selection, dominance, and inbreeding depression that will maintain SA polymorphism in hermaphrodite species with arbitrary rates of selfing, and we have identified several specific ways in which dioecious and hermaphrodite species are differentially susceptible to the accumulation of SA alleles. We have also evaluated how finite population size and self-fertilization interact to influence the efficacy of balancing selection to maintain SA alleles, and the empirical population genetic signals of recent balancing selection at an SA locus (by way of hitchhiking signals of recent partial selective sweeps). Our analysis significantly broadens the theoretical scope of SA polymorphism beyond its traditional context in dioecious species, where its empirical importance is already widely recognized (e.g., Delph et al. 2004; Pischcheda and Chippindale 2006; Bonduriansky and Chenoweth 2009; van Doorn 2009; Pennell and Morrow 2013; Qiu et al. 2013; Wright and Mank 2013).

**SELFING AND THE MAINTENANCE OF SA ALLELES**

In the absence of selfing, balancing selection criteria are identical between dioecious and hermaphrodite population models (consistent with the intuition of Abbott 2011; Bodmer 1965 discusses reasons for the equivalence of such models). In agreement with models of dioecious species (Bodmer 1965; Kidwell et al. 1977; Fry 2010), we find that SA alleles are most likely to be maintained when there is a net overdominance for fitness (as in scenarios involving a “dominance reversal”; see Kidwell et al. 1977; Gillespie 1978; Curtsinger et al. 1994; Prout 2000; Fry 2010), though this condition is not a requirement for balancing selection (e.g., Kidwell et al. 1977).

Although many hermaphrodite species are outcrossing, a substantial fraction practice self-fertilization at nonnegligible rates (e.g., roughly half of flowering plants; Igić and Kohn 2006). Selfing introduces a “female-bias” in the net direction of selection, expands criteria for the evolutionary accumulation of female-beneficial alleles, and restricts criteria for male-benefit alleles. Charlesworth and Charlesworth (1978b) were first to recognize that such a bias should occur under partial selfing, and our analysis confirms that the biased accumulation of female-benefit alleles...
is a general feature of both genotypically fixed selfing models, and mass-action models with frequency-dependent selfing rates. The prominent female-bias under high selfing suggests a strong potential for coevolution between selfing rates and the accumulation of female-benefit/male-detriment alleles. Such a scenario may contribute to the evolution of “selfing syndromes” (Sicard and Lenhard 2011), where highly selfing species have traits that are favorable for seed production (female function) and disadvantageable for outcross siring success (male function).

Selfing also decreases the parameter space that gives rise to balancing selection at SA loci. Like heterozygote advantage models (Hayman 1953; Kimura and Ohta 1971), net overdominance at SA loci (i.e., when \( h_m, h_f < \frac{1}{2} \), the sex-averaged genotypic fitnesses can exhibit overdominance) is often insufficient for maintaining polymorphism under self-fertilization. Opportunities for balancing selection are also reduced in the absence of net overdominant effects of SA alleles (e.g., for \( h_m = h_f = \frac{1}{2} \)). This latter effect arises from the asymmetric parameter space for balancing selection under selfing (see Fig. 1).

The mechanism of self-fertilization affects the degree to which predictions of hermaphrodite SA models diverge from dioecious model predictions. One can consider the full set of population genetic models of sexual antagonism as points along a gradient. At one extreme, we have the standard dioecious theory and models of obligate outcrossing hermaphrodites, which both predict an equal impact of male and female selection on the dynamics of SA alleles. The rate of allele frequency change at an SA locus is equally dependent on its fitness effects on males and females, and polymorphism is most likely to be maintained when SA alleles have opposing, but symmetrically strong, sex-specific fitness effects. Fixed-selfing hermaphrodite populations reside at the other extreme of the spectrum, at which female-biased selection is most pronounced, and polymorphism is most likely to be maintained when allelic fitness effects on males are stronger than fitness effects on females. Mass-action models will fall somewhere in between the endpoints of fixed selfing and obligate outcrossing or dioecy. For a given population-wide rate of self-fertilization, mass-action selfing will generate a female-bias of lesser magnitude than that of fixed-selfing, and the degree of fitness effect asymmetry between male and female functions will be less severe among balanced polymorphic alleles. These predictions could potentially be tested using plant species where fixed and mass-action selfing are most likely to apply (e.g., “prior” or “delayed selfing” as a form of fixed selfing; Lloyd 1992; Lloyd and Schoen 1992; mass-action selfing for flowering plants where self- and outcross pollen compete to fertilize individual ovules; see Holsinger 1991; Goodwillie et al. 2005). One might naively expect lesser opportunities for maintaining polymorphism, and more pronounced selfing syndromes, in plant species practicing fixed-like selfing mechanisms. Finally, because fixed-selfing requires relatively specialized fertilization mechanisms, competitive selfing likely predominates among self-compatible flowering plants (Holsinger 1991; Goodwillie et al. 2005), and establishes expectations for SA polymorphism among partially selfing species.

**POPULATION GENETIC SIGNALS OF BALANCING SELECTION TO MAINTAIN SA ALLELES**

Population genomic data can be leveraged to detect genes under balancing selection, which may leave characteristic signals associated with ancient or newly established balanced polymorphisms (Charlesworth 2006). Signals of ancient balancing selection (e.g., gene genealogies with long internal branches: Hudson and Kaplan 1988; Kaplan et al. 1988; Hudson 1990; Hey 1991; trans-species polymorphisms: Klein et al. 1998; Asthana et al. 2005; Leffler et al. 2013) will be reliably strong when polymorphisms are stably maintained at intermediate frequencies over long evolutionary intervals (i.e., strong and persistent balancing selection for \( > > 2N_e \) generations; Charlesworth and Charlesworth 2010). These stringent requirements generally limit opportunities to detect loci segregating for balanced polymorphisms, though several documented examples do exist (Andres et al. 2009; Leffler et al. 2013; DeGiorgio et al. 2014). Sexual antagonism is likely to leave a particularly weak signal of ancient polymorphism (i.e., relative to other balancing selection mechanisms; Connnallon and Clark 2012, 2013). Even if balancing selection remains temporally constant, opportunities for stably maintaining SA alleles will be hampered by their extreme susceptibility to genetic drift (Connnallon and Clark 2012; Mullon et al. 2012; Hesketh et al. 2013), and partial selfing further erodes the efficacy of balancing selection on SA alleles (as shown here; see Fig. 4). On the other hand, selfing decreases the effective rate of recombination (Nordborg 2000), and increases the length of neutral LD blocks around temporally stable balanced polymorphisms (Nordborg et al. 1996; Nordborg 1997). We therefore note that SA alleles that are stably maintained over time, despite extensive selfing, should exhibit elevated signatures of ancient balancing selection.

In populations experiencing recent balancing selection, partial selective sweeps (e.g., where a rare SA allele has recently increased from low to intermediate frequency, and neutral sites “hitchhike” along with it; Coop and Ralph 2012) can generate a characteristic short-term signal near the selected site (Charlesworth 2006). These signals of recent selection provide additional scope for detecting loci under balancing selection (for relevant examples, see Charlesworth 2006). Relatively slow invasion rates of SA alleles (Livingstone 1992) should dampen signatures of recent SA selection in dioecious populations (Connnallon and Clark 2013). By reducing the effective rate of recombination, selfing is known to exaggerate the hitchhiking effects of unconditionally beneficial alleles, leading to sharper reductions in nearby neutral diversity, and to elevated frequencies.
of deleterious alleles at sites linked to positively selected alleles (see Glémin 2012; Hartfield and Glémin 2014). In contrast, invading SA alleles instigate weaker hitchhiking effects under self-fertilization. Although selfing reduces recombination rates in double heterozygotes, and thereby slows the decay of linkage disequilibrium, this effect is more than compensated by the reduced invasion rates of rare SA alleles, particularly in cases when there is net overdominance at the SA locus.

Conclusions

SA genetic variation may persist in a population by way of balancing selection, or through recurrent mutation, which counteracts the loss of SA alleles through drift and net directional selection (Bonduriansky and Chenoweth 2009; Fry 2010; Connallon and Clark 2012; Delph and Kelly 2014). Both processes should play some role in maintaining variation, though the relative contribution of each is currently unknown. Balancing selection is least likely in species with high selfing rates, and our results therefore point to recurrent mutation as a likely mechanism for maintaining SA genetic variation in such species. In populations with low selfing rates (e.g., which includes a large fraction of hermaphroditic plant and animal species; Igic and Kohn 2006; Jarne and Auld 2006), net overdominance for fitness can play an important role in maintaining genetic variation in female and male functions. To the extent that selection at individual loci is weak, alleles with partially recessive deleterious effects ($h_m, h_f < \frac{1}{2}$, as uniformly predicted by fitness landscape models; Manna et al. 2011; Sellis et al. 2011; Connallon and Clark 2014b) can be maintained at relatively stable, intermediate frequencies, when they are involved in trade-offs with other fitness components. Among outbreeding populations, such alleles may disproportionately contribute to the maintenance of the high levels of genetic variation that are commonly observed in reproductive phenotypes, life-history traits, and other fitness components (e.g., Houle 1992; Pomiankowski and Moller 1995; Charlesworth and Hughes 1999).

Acknowledgments

We thank S. Otto for encouragement, discussion, and suggestions on the deterministic results, and D. Charlesworth, B. Charlesworth, H.-R. Gregorius, B. Nurnberger, D. Obbard, J. Pannell, and S. Wright for helpful discussion. We also thank D. Charlesworth, H.-R. Gregorius, M. Hartfield, S. Otto, J. Pannell, D. Roze, and two anonymous reviewers for comments that greatly improved the manuscript. CYJ was supported by the University, and by funds from the School of Biological Sciences at Monash University.

Literature Cited

Abbot, J. K. 2011. Intralocus sexual conflict and sexually antagonistic genetic variation in hermaphroditic animals. Proc. R. Soc. B 278:161–169.
Agrawal, A. F., and M. C. Whitlock. 2011. Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics 178:553–566.
Albert, A. Y. K., and S. P. Otto. 2005. Sexual selection can resolve sex-linked sexual antagonism. Science 310:119–121.
Andres, A. M., M. J. Hubisz, A. Indap, D. G. Torgerson, J. D. Degenhardt, A. R. Boyko, R. N. Gutenkunst, T. J. White, E. D. Green, C. D. Bustamante, et al. 2009. Targets of balancing selection in the human genome. Mol. Biol. Evol. 26:2755–2764.
Arnqvist, G. 2011. Assortative mating by fitness and sexually antagonistic genetic variation. Evolution 65:2111–2116.
Asthana, S., S. Schmidt, and S. Sunyaev. 2005. A limited role for balancing selection. Trends Genet 21:30–32.
Bachtrog, D., J. E. Mank, C. L. Peichsel, M. Kirkpatrick, S. P. Otto, T. L. Ashman, M. W. Hahn, J. Kitano, I. Mayrose, R. Ming, et al. 2014. Sex determination: why so many ways of doing it? PLoS Biol. 12:1001899.
Barrett, S. C. H. 2002. Sexual interference of the floral kind. Heredity 88:154–159.
Barton, N. H. 1998. The effect of hitch-hiking on neutral genealogies. Genet. Res. 72:123–133.
Berger, D., K. Grieshop, M. I. Lind, J. Goenaga, A. A. Maklakov, and G. Arnqvist. 2014. Intralocus sexual conflict and environmental stress. Evolution 68:2184–2196.
Bergero, R., and D. Charlesworth. 2008. The evolution of restricted recombination in sex chromosomes. Trends Ecol. Evol. 24:94–102.
Betancourt, A. J., Y. Kim, and H. A. Orr. 2004. A pseudohitchhiking model of X vs. autosomal diversity. Genetics 168:2261–2269.
Bodmer, W. F. 1965. Differential fertility in population genetics models. Genetics 51:411–424.
Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends Ecol. Evol. 24:280–288.
Charlesworth, B. 2009. Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. 10:195–205.
Charlesworth, B., and D. Charlesworth. 1978a. A model for the evolution of dioecy and gynodioecy. Am. Nat. 112:975–997.
———. 2010. Elements of evolutionary genetics. Roberts and Company Publishers, Greenwood Village, CO.
Charlesworth, B., and K. A. Hughes. 1999. The maintenance of genetic variation in life-history traits. Pp. 369–392 in R. S. Singh and C. B. Krimbas, eds. Evolutionary genetics: from molecules to morphology. Vol. 1. Cambridge Univ. Press, Cambridge, U.K.
Charlesworth, D. 2006. Balancing selection and its effects on sequences in nearby genome regions. PLoS Genet. 2:e64.
Charlesworth, D., and B. Charlesworth. 1978b. Population genetics of partial male-sterility and the evolution of monoecy and dioecy. Heredity 41:137–153.
———. 1980. Sex differences in fitness and selection for centric fusions between sex chromosomes and autosomes. Genet. Res. 35:205–214.
———. 1987. The effect of investment in attractive structures on allocation to male and female functions in plants. Evolution 41:948–968.
Connallon, T., and A. G. Clark. 2010. Sex linkage, sex-specific selection, and the role of recombination in the evolution of sexually dimorphic gene expression. Evolution 64:3417–3442.
———. 2012. A general population genetic framework for antagonistic selection that accounts for demography and recurrent mutation. Genetics 190:1477–1489.
———. 2013. Antagonistic versus nonantagonistic models of balancing selection: characterizing the relative timescales and hitchhiking effects of partial selective sweeps. Evolution 67:908–917.
———. 2014a. Evolutionary inevitability of sexual antagonism. Proc. R. Soc. B 281:20132123.
Conner, J. 2006. Ecological genetics of floral evolution. Pp. 260–277 in L. D. Harder and S. Barrett, eds. The ecology and evolution of flowers. Oxford Univ. Press, New York.

Coop, G., and P. Ralph. 2012. Patterns of neutral diversity under general models of selective sweeps. Genetics 192:205–224.

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

DeGiorgio, M., K. E. Lohmueller, and R. Nielsen. 2014. A model-based approach for identifying signatures of ancient balancing selection in genetic data. PLoS Genet. 10:e1004561

Delph, L. F. 1996. Flower size dimorphism in plants with unisexual flowers. Pp. 217–237 in D. G. Lloyd and S. C. H. Barrett, eds. Flora: studies on floral evolution in animal-pollinated plants. Chapman and Hall, New York.

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

Delph, L. F., J. L. Gehring, F. M. Frey, A. M. Arntz, and M. Levri. 2004. Genetic constraints on floral evolution in a sexually dimorphic plant revealed by artificial selection. Evolution 58:1936–1946.

Delph, L. F., A. M. Arntz, C. Scotti-Saintagne, and I. Scotti. 2010. The genomic architecture of sexual dimorphism in the dioecious plant silene latifolia. Evolution 64:2873–2886.

Devlin, B., and N. C. Ellstrand. 1990. Male and female fertility variation in wild radish, a hermaphrodite. Am. Nat. 136:87–107.

Elle, E., and T. R. Meagher. 2000. Sex allocation and reproductive success in the andromonoecious perennial Solanum carolinense (Solanaceae). II. Paternity and functional gender. Am. Nat. 156:622–636.

Ellis, A. G., and S. D. Johnson. 2010. Gender differences in the effects of floral spur length manipulation on fitness in a hermaphrodite orchid. Int. J. Plant Sci. 171:1010–1019.

Ennos, R. A., and R. K. Dodson. 1987. Pollen success, functional gender and assortative mating in an experimental plant population. Heredity 58:119–126.

Ewens, W. J. 2004. Mathematical population genetics. I. Theoretical introduction. 2nd ed. Springer, New York.

Ewens, W. J., and G. Thomson. 1970. Heterozygote selective advantage. Ann. Hum. Genet. 33:365–376.

Eyre-Walker, A., and P. D. Keightley. 2007. The distribution of fitness effects of new mutations. Nat. Rev. Genet. 8:610–618.

Fisher, R. 1941. Average excess and average effect of a gene substitution. Ann. Eugen. 11:53–63.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. Nature 447:1107–1110.

Fry, J. D. 2010. The genomic location of sexually antagonistic variation: some cautionary comments. Evolution 64:1510–1516.

Gibson, J. R., A. K. Chippindale, and W. R. Rice. 2002. The x chromosome is a hot spot for sexually antagonistic fitness variation. Proc. R. Soc. B 269:499–505.

Gillespie, J. H. 1978. A general model to account for enzyme variation in natural populations. v. the SAS-CFF model. Theor. Pop. Biol. 14:1–45.

Glémin, S. 2012. Extinction and fixation times with dominance and inbreeding. Theor. Pop. Biol. 81:310–316.

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

Gregorius, H. R. 1982. Selection in plant populations of effectively infinite size: II. Protectednedd of a biallelic polymorphism. J. Theor. Biol. 96:689–705.

Gregorius, H. R., M. D. Ross, and E. Gillett. 1982. Selection in plant populations of effectively infinite size: III. the maintenance of females among hermaphrodites for a biallelic model. Heredity 48:329–343.

Harano, T., K. Okada, S. Nakayama, T. Miyatake, and D. J. Hosken. 2010. Intralocus sexual conflict unresolved by sex-limited trait expression. Curr. Biol. 20:2036–2039.

Harder, L. D. 2000. Pollen dispersal and the floral diversity of monocotyledons. Pp. 243–257 in K. Wilson and D. Morrison, eds. Monocots: systematics and evolution. CSIRO, Melbourne.

Harder, L. D., and S. D. Johnson. 2008. Function and evolution of aggregated pollen in angiosperms. Int. J. Plant Sci. 169:59–78.

Harder L. D., and W. G. Wilson. 1998. A clarification of pollen discounting and its joint effects with inbreeding depression on mating system evolution. Am. Nat. 152:684–695.

Hartfeld, M., and S. Glémin. 2014. Hitchhiking of deleterious alleles and the cost of adaptation in partially selfing species. Genetics 196:281–293.

Hayman, B. 1953. Mixed selfing and random mating when homozygotes are at a disadvantage. Heredity 7:185–192.

Hedrick, P. W. 1999. Antagonistic pleiotropy and genetic polymorphism: a perspective. Heredity 82:126–133.

Heskeijth, K., K. Fowler, and M. Reuter. 2013. Genetic drift in antagonistic genes leads to divergence in sex-specific fitness between experimental populations of Drosophila melanogaster. Evolution 67:1503–1510.

Hey, J. 1991. A multi-dimensional coalescent process applied to multi-allelic selection models and migration models. Theor. Pop. Biol. 39:30–48.

Hodgins, K. A., and S. C. H. Barrett. 2008. Natural selection on floral traits through male and female function in wild populations of the heterostyloous daffodil Narcissus triandrus. Evolution 62:1751–1763.

Holsinger, K. E. 1991. Mass-action models of plant mating systems: the evolutionary stability of mixed mating systems. Am. Nat. 138:606–622.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130:195–204.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. Pp. 1–44 in D. Futuyma and J. Antonovics, eds. Oxford surveys in evolutionary biology. Vol. 7. Oxford Univ. Press, Oxford, U.K.

Hudson, R. R., and N. L. Kaplan. 1988. The coalescent process in models with selection and recombination. Genetics 120:831–840.

Igic, B., and J. R. Kohn. 2006. The distribution of plant mating systems: study bias against obligately outcrossing species. Evolution 60:1098–1103.

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

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

Jordan, C. Y., and D. Charlesworth. 2012. The potential for sexually antagonistic polymorphism in different genome regions. Evolution 66:505–516.

Jordan, C. Y., and S. P. Otto. 2012. Functional pleiotropy and mating system evolution in plants: frequency-independent mating. Evolution 66:957–972.

Kaplan, N. L., T. Darden, and R. R. Hudson. 1988. The coalescent process in models with selection. Genetics 120:819–829.

Kidwell, J. F., M. T. Clegg, F. M. Stewart, and T. Prout. 1977. Regions of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. Annu. Rev. Ecol. Evol. Syst. 36:47–79.

Gregorius, H. R. 1982. Selection in plant populations of effectively infinite size: II. Protectednedd of a biallelic polymorphism. J. Theor. Biol. 96:689–705.
Klein, J., A. Sato, S. Nagl, and C. O’Higgins. 1998. Molecular trans-species polymorphism. Annu. Rev. Ecol. Syst. 29:1–21.

Knight, T. M., I. A. Steets, J. C. Vamosi, S. J. Mazier, M. Burd, D. R. Campbell, M. R. Dudash, M. O. Johnston, R. J. Mitchell, and T. L. Ashman. 2005. Pollen limitation of plant reproduction: pattern and process. Annu. Rev. Ecol. Evol. Syst. 36:467–497.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution 34:292–305.

Leffler, E. M., Z. Gao, S. Pfeifer, L. S. Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution 34:292–305.

Knight, T. M., J. A. Steets, J. C. Vamosi, S. J. Mazer, M. Burd, D. R. Campbell, K. Klein, J., A. Sato, S. Nagl, and C. O’hUiginn. 1998. Molecular trans-species polymorphism. Annu. Rev. Ecol. Syst. 29:1–21.

C. Y. JORDAN AND T. CONNALLON

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

Patten, M. M., and D. Haig. 2009. Maintenance or loss of genetic variation under sexual and parental antagonism at a sex-linked locus. Evolution 63:2888–2895.

Patten, M. M., D. Haig, and F. Ubeda. 2010. Fitness variation due to sexual antagonism and linkage disequilibrium. Evolution 64:3638–3642.

Pennell, T. M., and E. H. Morrow. 2013. Two sexes, one genome: the evolutionary dynamics of intralocus sexual conflict. Ecol. Evol. 3:1819–1834.

Phadnis, N., and J. D. Fry. 2005. Widespread correlations between dominance and homozygous effects of mutations: implications for theories of dominance. Genetics 171:385–392.

Pishedda, A., and A. K. Chippindale. 2006. Intralocus sexual conflict diminishes the benefits of sexual selection. PLoS Biol. 4:e356.

Pomiankowski, A., and A. P. Moller. 1995. A resolution of the lek paradox. Proc. R. Soc. B 260:21–29.

Posavi, M., B. Larget, G. W. Gelembiuk, and C. E. Lee. 2014. Testing for beneficial reversal of dominance during salinity shifts in the invasive copepod *Eurytemora affinis*, and implications for the maintenance of genetic variation. Evolution 61:3166–3183.

Prout, T. 2000. How well does opposing selection maintain variation. Pp. 157–203 in R. S. Singh and C. B. Krimbas, eds. Evolutionary genetics from molecules to morphology. Cambridge Univ. Press, Cambridge, U.K.

Qiu, S., R. Bergero, and D. Charlesworth. 2013. Testing for the footprint of sexually antagonistic polymorphisms in the pseudoautosomal region of a plant sex chromosome pair. Genetics 194:663–672.

Renner, S. S., and R. E. Ricklefs. 1995. Dioecy and its correlates in the flowering plants. Am. J. Bot. 82:596–606.

Rice, W. R. 1984. Sex chromosomes and the evolution of sexual dimorphism. Evolution 38:735–742.

———. 1985. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. Evolution 41:911–914.

———. 1992. Sexually antagonistic genes: experimental evidence. Science 256:1436–1439.

Robertson, A. 1962. Selection for heterozygotes in small populations. Genetics 47:1291–1300.

Ross, M. 1977. Behaviour of a sex-differential fertility gene in hermaphrodite populations. Heredity 38:279–290.

———. 1990. Sexual asymmetry in hermaphroditic plants. Trends Ecol. Evol. 5:43–47.

Ross, M. D. 1984. Frequency-dependent selection in hermaphrodites: the rule rather than the exception. Biol. J. Linnean. Soc. 23:145–155.

———. 1985. Evolution of outbreeding systems. Pp. 180–197 in H. R. Gregorius, ed. Population genetics in forestry, lecture notes in biomathematics. Springer-Verlag, New York.

Ross, M. D., and H. R. Gregorius. 1983. Hermaphrodites: a resource-allocation model. Am. Nat. 121:204–222.

Rymer, P. D., S. D. Johnson, and V. Savolainen. 2010. Pollinator behaviour and plant speciation: can assortative mating and disruptive selection maintain distinct floral morphs in sympathy? New Phytol. 188:426–436.

Sellis, D., B. J. Callahan, D. A. Petrov, and P. W. Messer. 2011. Heterozygote advantage as a natural consequence of adaptation in diploids. Proc. Nat. Acad. Sci. USA 108:20666–20671.

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

Sprenger, D., N. J. Dingemans, N. A. Dochterman, J. Theobald, and S. P. W. Walker. 2012. Aggressive females become aggressive males in a sex-changing reef fish. Ecol. Lett. 15:986–992.

van Doorn, G. S. 2009. Intralocus sexual conflict. Ann. NY Acad. Sci. 1168:52–71.
Webb, C. J., and D. G. Lloyd. 1986. The avoidance of interference between the presentation of pollen and stigmas in angiosperms ii. Herkogamy. NZ J. Bot. 24:163–178.

Wright, A. E., and J. E. Mank. 2013 The scope and strength of sex-specific selection in genome evolution. J. Evol. Biol. 26:1841–1853.

Wright, S. 1945. The differential equation of the distribution of gene frequencies. Proc. Nat. Acad. Sci. USA. 31:382–389.

Ziehe, M. 1982. Zygotic genotypic frequencies under selection on female or male gamete production in partially self-fertilizing plant populations. Heredity 49:271–279.

———. 1985. Polymorphic equilibria under inbreeding effects and selection on components of reproduction. Pp. 92–106 in H. R. Gre- gorius, ed. Population genetics in forestry. Springer-Verlag, New York.

Associate Editor: D. Roze

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher’s website:

Appendix A. Weak selection approximations of allele frequency change.
Appendix B. Hitchhiking at a linked neutral locus.
Appendix C. Selfing and the invasion criteria for female- and male-benefit alleles.
Appendix D. The fraction of parameter space that generates balancing selection.
Appendix E. Efficacy of selection in balancing selection model of SA polymorphism.
SUPPLEMENTARY MATERIAL (C. Y. Jordan and T. Connallon)

Appendix A. Weak selection approximations of allele frequency change

With weak selection, allele frequency change across generations will be slow compared to the rate of convergence of genotype frequencies to equilibrium under partial selfing (Nagylaki 1997). Using a separation of timescales (Otto and Day 2007, chapter 9), we can calculate the quasi-equilibrium genotype frequencies in the absence of selection (i.e., for fixed $q$), and then use these frequencies in our calculation of the slower allele frequency changes across generations. The quasi-equilibrium genotype frequencies (given allele frequency $q$ and selfing proportion $C$) are:

\[
F_{AA}^* = q^2 + Cq(1 - q)/(2 - C)
\]

\[
F_{Aa}^* = 2q(1 - q) - 2Cq(1 - q)/(2 - C)
\]

\[
F_{aa}^* = (1 - q)^2 + Cq(1 - q)/(2 - C)
\]

(Caballero and Hill 1992). Substituting these into the general equations for allele frequency change between adults and gametes (see the main text), we get:

\[
\Delta q_m = -\frac{sq(1 - q) C + 2(1 - C) [h_m + q(1 - 2h_m)]}{2 - C} W_m
\]

\[
\Delta q_f = \frac{tq(1 - q) (2 - C) - 2(1 - C) [h_f + q(1 - 2h_f)]}{2 - C} W_f
\]

These can be substituted into the general allele frequency recursion, leading to:

\[
\Delta q = (1 - C) \frac{\Delta q_m + \Delta q_f}{2} + C \Delta q_f
\]

\[
= t(1 + C) \frac{q(1 - q) (2 - C) - 2(1 - C) [h_f + q(1 - 2h_f)]}{2(2 - C)} W_f
\]

\[
- s(1 - C) \frac{q(1 - q) C + 2(1 - C) [h_m + q(1 - 2h_m)]}{2(2 - C)} W_m
\]

For the additive case ($h_m = h_f = \frac{1}{2}$), this simplifies to:
\[
\Delta q \approx \frac{stq(1-q)(\hat{q} - q)}{W_f W_m (2-C)}
\]

\[
\hat{q} = \frac{t(1+C) - s(1-t)(1-C)}{2st}
\]

For dominance reversal case \((h_m = h_f = h < \frac{1}{2})\), we can approximate to first order in \(s\) and \(t\), which simplifies to:

\[
\Delta q \approx \frac{(1-C)(1-2h)[t(1+C) + s(1-C)]q(1-q)(\hat{q} - q)}{(2-C)}
\]

\[
\hat{q} = \frac{2t(1+C) - [t(1+C) + s(1-C)][C + 2h(1-C)]}{2(1-C)(1-2h)[t(1+C) + s(1-C)]}
\]

**Additional References**

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

Nagylaki, T. 1997. The diffusion model for migration and selection in a plant population. *J Math Biol*. 35:409-431.

**Appendix B: Hitchhiking at a linked, neutral locus**

We can extend the weak selection approach from Appendix A to approximate the trajectories of haplotypes in a two-locus system, with a SA locus and a linked neutral locus. The recombination rate between the loci is \(r\), per meiosis. We assume weak selection and tight linkage, thereby permitting us to ignore terms of \(r^2\), \(rs\), and \(rt\). Adult genotype frequencies involving the \(i\)th and \(j\)th haplotypes are represented as \(F_{ij}\), and haplotype frequencies in the population are denoted as:

\[
q_1 = F_{11} + F_{12}/2 + F_{13}/2 + F_{14}/2
\]

\[
q_2 = F_{22} + F_{12}/2 + F_{23}/2 + F_{24}/2
\]

\[
q_3 = F_{33} + F_{13}/2 + F_{23}/2 + F_{34}/2
\]

\[
q_4 = F_{44} + F_{14}/2 + F_{24}/2 + F_{34}/2
\]
The haplotype frequencies in ovules are \(x_1 = [AB], x_2 = [Ab], x_3 = [aB],\) and \(x_4 = [ab],\) and the corresponding frequencies in pollen are \(y_1, y_2, y_3, y_4.\) Let \(F_{ij}\) represent the contribution of genotype \(F_{ij}\) to the production of fertilized ovules that make up the next generation. From these terms, the genotype frequencies in the next generation are:

\[
\begin{align*}
F_{11}' &= (1 - C)x_1'y_1 + C[F_{11} + F_{12}/4 + F_{13}/4 + F_{14}/4(1 - r)^2/4 + F_{23}/r^2/4] \\
F_{12}' &= (1 - C)(x_1'y_2 + x_2'y_1) + C[F_{12}/2 + F_{14}/r(1 - r)/2 + F_{23}/r(1 - r)/2] \\
F_{13}' &= (1 - C)(x_1'y_3 + x_3'y_1) + C[F_{13}/2 + F_{14}/r(1 - r)/2 + F_{23}/r(1 - r)/2] \\
F_{14}' &= (1 - C)(x_1'y_4 + x_4'y_1) + C[F_{14}/(1 - r)^2/2 + F_{23}/r^2/2] \\
F_{22}' &= (1 - C)x_2'y_2 + C[F_{22}/4 + F_{14}/r(1 - r)/2 + F_{23}/r(1 - r)/2] \\
F_{23}' &= (1 - C)(x_2'y_3 + x_3'y_2) + C[F_{14}/r^2/2 + F_{23}/(1 - r)^2/2] \\
F_{24}' &= (1 - C)(x_2'y_4 + x_4'y_2) + C[F_{24}/2 + F_{14}/r(1 - r)/2 + F_{23}/r(1 - r)/2] \\
F_{33}' &= (1 - C)x_3'y_3 + C[F_{33}/4 + F_{13}/r(1 - r)/2 + F_{23}/r(1 - r)/2 + F_{34}/r/2] \\
F_{34}' &= (1 - C)(x_3'y_4 + x_4'y_3) + C[F_{14}/r(1 - r)/2 + F_{23}/r(1 - r)/2 + F_{34}/r/2] \\
F_{44}' &= (1 - C)x_4'y_4 + C[F_{44}/4 + F_{14}/r(1 - r)/2 + F_{23}/r^2/4 + F_{24}/4 + F_{34}/4] \\
\end{align*}
\]

Ignoring terms of \(r^2,\) these reduce to:

\[
\begin{align*}
F_{11}' &= (1 - C)x_1'y_1 + C[F_{11} + F_{12}/4 + F_{13}/4 + F_{14}/4] - rCF_{14}/2 \\
F_{12}' &= (1 - C)(x_1'y_2 + x_2'y_1) + C[F_{12}/2 + rC(F_{14} + F_{23}/r)/2] \\
F_{13}' &= (1 - C)(x_1'y_3 + x_3'y_1) + C[F_{13}/2 + rC(F_{14} + F_{23}/r)/2] \\
F_{14}' &= (1 - C)(x_1'y_4 + x_4'y_1) + C[F_{14}/2 - rCF_{14}]/r \\
F_{22}' &= (1 - C)x_2'y_2 + C[F_{22}/4 + F_{12}/4 + F_{23}/4 + F_{24}/4] - rCF_{23}/2 \\
F_{23}' &= (1 - C)(x_2'y_3 + x_3'y_2) + C[F_{23}/2 - rCF_{23}]/r \\
F_{24}' &= (1 - C)(x_2'y_4 + x_4'y_2) + C[F_{24}/2 + rC(F_{14} + F_{23}/r)/2] \\
F_{33}' &= (1 - C)x_3'y_3 + C[F_{33}/4 + F_{13}/r(1 - r)/2 + F_{23}/r(1 - r)/2 + F_{34}/r/2] \\
F_{34}' &= (1 - C)(x_3'y_4 + x_4'y_3) + C[F_{34}/2 + rC(F_{14} + F_{23}/r)/2] \\
F_{44}' &= (1 - C)x_4'y_4 + C[F_{44}/4 + F_{14}/r(1 - r)/2 + F_{23}/r^2/4 + F_{24}/4 + F_{34}/4] - rCF_{14}/2
\end{align*}
\]
The haplotype frequencies in the next generation are:

\[ q_1 = F_{11}' + F_{12}'/2 + F_{13}'/2 + F_{14}'/2 = (1 - C)(x_1 + y_1)/2 + Cx_1 - rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

\[ q_2 = F_{22}' + F_{12}'/2 + F_{23}'/2 + F_{24}'/2 = (1 - C)(x_2 + y_2)/2 + Cx_2 + rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

\[ q_3 = F_{33}' + F_{13}'/2 + F_{23}'/2 + F_{34}'/2 = (1 - C)(x_3 + y_3)/2 + Cx_3 + rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

\[ q_4 = F_{44}' + F_{14}'/2 + F_{24}'/2 + F_{34}'/2 = (1 - C)(x_4 + y_4)/2 + Cx_4 - rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

The haplotype frequency change across a generation is:

\[ \Delta q_1 = (1 - C)(\Delta x_1 + \Delta y_1)/2 + C\Delta x_1 - rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

\[ \Delta q_2 = (1 - C)(\Delta x_2 + \Delta y_2)/2 + C\Delta x_2 + rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

\[ \Delta q_3 = (1 - C)(\Delta x_3 + \Delta y_3)/2 + C\Delta x_3 + rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

\[ \Delta q_4 = (1 - C)(\Delta x_4 + \Delta y_4)/2 + C\Delta x_4 - rC(F_{14} - F_{23})(1 - h_f)/(2W_f) \]

where \( \Delta x_i \) and \( \Delta y_i \) are given by:

\[ \Delta x_1 = \frac{2F_{11} + F_{12} + F_{13}(1 - h_f)t + F_{14}(1 - h_f)t - 2q_1W_f}{2W_f} - \frac{r(F_{14} - F_{23})(1 - h_f)t}{2W_f} \]

\[ \Delta x_2 = \frac{2F_{22} + F_{12} + F_{23}(1 - h_f)t + F_{24}(1 - h_f)t - 2q_2W_f}{2W_f} + \frac{r(F_{14} - F_{23})(1 - h_f)t}{2W_f} \]

\[ \Delta x_3 = \frac{2F_{33}(1 - t) + F_{34}(1 - t) + F_{13}(1 - h_f)t + F_{23}(1 - h_f)t - 2q_3W_f}{2W_f} + \frac{r(F_{14} - F_{23})(1 - h_f)t}{2W_f} \]

\[ \Delta x_4 = \frac{2F_{44}(1 - t) + F_{34}(1 - t) + F_{14}(1 - h_f)t + F_{24}(1 - h_f)t - 2q_4W_f}{2W_f} - \frac{r(F_{14} - F_{23})(1 - h_f)t}{2W_f} \]

\[ W_f = 1 - h_f t(F_{13} + F_{14} + F_{23} + F_{24}) - t(F_{33} + F_{34} + F_{44}) \]

and

\[ \Delta y_1 = \frac{2F_{11}(1 - s) + F_{12}(1 - s) + F_{13}(1 - h_m s) + F_{14}(1 - h_m s) - 2q_1W_m}{2W_m} - \frac{r(F_{14} - F_{23})(1 - h_m s)}{2W_m} \]

\[ \Delta y_2 = \frac{2F_{22}(1 - s) + F_{12}(1 - s) + F_{23}(1 - h_m s) + F_{24}(1 - h_m s) - 2q_2W_m}{2W_m} + \frac{r(F_{14} - F_{23})(1 - h_m s)}{2W_m} \]

\[ \Delta y_3 = \frac{2F_{33} + F_{34}(1 - h_m s) + F_{23}(1 - h_m s) - 2q_3W_m}{2W_m} + \frac{r(F_{14} - F_{23})(1 - h_m s)}{2W_m} \]

\[ \Delta y_4 = \frac{2F_{44} + F_{34}(1 - h_m s) + F_{14}(1 - h_m s) - 2q_4W_m}{2W_m} - \frac{r(F_{14} - F_{23})(1 - h_m s)}{2W_m} \]

\[ W_m = 1 - s(F_{11} + F_{12} + F_{22}) - h_m s(F_{13} + F_{14} + F_{23} + F_{24}) \]
Using the separation of timescales approach from Appendix A, we get quasi-equilibrium (QE) genotype frequencies ($\Delta x_i, \Delta y_i, r \to 0$):

\[
F_{11}^* = q_1^2 + Cq_1(1 - q_1)/(2 - C)
\]

\[
F_{12}^* = 4q_1q_2(1 - C)/(2 - C)
\]

\[
F_{13}^* = 4q_1q_3(1 - C)/(2 - C)
\]

\[
F_{14}^* = 4q_1q_4(1 - C)/(2 - C)
\]

\[
F_{22}^* = q_2^2 + Cq_2(1 - q_2)/(2 - C)
\]

\[
F_{23}^* = 4q_2q_3(1 - C)/(2 - C)
\]

\[
F_{24}^* = 4q_2q_4(1 - C)/(2 - C)
\]

\[
F_{33}^* = q_3^2 + Cq_3(1 - q_3)/(2 - C)
\]

\[
F_{34}^* = 4q_3q_4(1 - C)/(2 - C)
\]

\[
F_{44}^* = q_4^2 + Cq_4(1 - q_4)/(2 - C)
\]

Substituting QE frequencies into the general equations above leads to:

\[
\Delta q_1 = (1 - C)(\Delta x_1 + \Delta y_1)/2 + C\Delta x_1 - 2rDC(1 - C)(1 - h_f)/(2 - C)W_f
\]

\[
\Delta q_2 = (1 - C)(\Delta x_2 + \Delta y_2)/2 + C\Delta x_2 + 2rDC(1 - C)(1 - h_f)/(2 - C)W_f
\]

\[
\Delta q_3 = (1 - C)(\Delta x_3 + \Delta y_3)/2 + C\Delta x_3 + 2rDC(1 - C)(1 - h_f)/(2 - C)W_f
\]

\[
\Delta q_4 = (1 - C)(\Delta x_4 + \Delta y_4)/2 + C\Delta x_4 - 2rDC(1 - C)(1 - h_f)/(2 - C)W_f
\]

where $D = q_1q_4 - q_2q_3$, and:

\[
\Delta x_1 = \frac{tq_1(1 - q_A)C + 2(1 - C)[1 - h_f - q_A(1 - 2h_f)]}{W_f(2 - C)} - \frac{2rD(1 - h_f)T(1 - C)}{W_f(2 - C)}
\]

\[
\Delta x_2 = \frac{tq_2(1 - q_A)C + 2(1 - C)[1 - h_f - q_A(1 - 2h_f)]}{2 - C} - \frac{2rD(1 - h_f)T(1 - C)}{W_f(2 - C)}
\]

\[
\Delta x_3 = \frac{tq_3q_A C + 2(1 - C)[1 - h_f - q_A(1 - 2h_f)]}{2 - C} + \frac{2rD(1 - h_f)T(1 - C)}{W_f(2 - C)}
\]

\[
\Delta x_4 = \frac{tq_4q_A C + 2(1 - C)[1 - h_f - q_A(1 - 2h_f)]}{2 - C} + \frac{2rD(1 - h_f)T(1 - C)}{W_f(2 - C)}
\]

\[
W_f = 1 - \frac{t(1 - q_A)}{2 - C} \{ C + 2(1 - C)[1 - q_A(1 - 2h_f)] \}^2
\]
\[ \Delta y_1 = -sq_1(1 - q_A) \frac{C + 2(1 - C)[h_m + q_A(1 - 2h_m)]}{W_m(2-C)} - 2rD(1 - h_m s)(1 - C) \]
\[ \Delta y_2 = -sq_2(1 - q_A) \frac{C + 2(1 - C)[h_m + q_A(1 - 2h_m)]}{W_m(2-C)} + 2rD(1 - h_m s)(1 - C) \]
\[ \Delta y_3 = sq_3 q_A \frac{C + 2(1 - C)[h_m + q_A(1 - 2h_m)]}{W_m(2-C)} + 2rD(1 - h_m s)(1 - C) \]
\[ \Delta y_4 = sq_4 q_A \frac{C + 2(1 - C)[h_m + q_A(1 - 2h_m)]}{W_m(2-C)} - 2rD(1 - h_m s)(1 - C) \]
\[ W_m = 1 - \frac{sq_A}{2 - C} \left\{ C + 2(1 - C)[2h_m + q_A(1 - 2h_m)] \right\} \]

By dropping the higher-order recombination terms \((r^2, rs, rt)\), these equations reduce to those in eq. (6) and eq. (12) of the main text.

**Appendix C. Selfing and the invasion criteria for female-benefit and male-benefit alleles**

Analysis of eq. (8) is straightforward and is not presented. From Eq. (9), the criterion for invasion of a male-beneficial allele is:

\[ t \frac{(1 + C)[2h + C(1-2h)]}{(1-C)[2 - C - 2h(1-C)]]} < s \]

Selfing makes invasion more difficult when the following inequality holds:

\[ \frac{h}{1-h} < \frac{(1+C)[2h+C(1-2h)]}{(1-C)[2 - C - 2h(1-C)]} \iff 0 < (1+C) + 2(1 - 2C)h - 4(1-C)h^2 \]

This is always true for the relevant dominance range of \(0 < h < \frac{1}{2}\). The female-benefit allele invasion criterion is:

\[ s \frac{(1-C)[2h_m + C(1-2h_m)]}{(1+C)[2 - C - 2h_m(1-C)]} < t \]

which is easier under selfing when:

\[ \frac{h}{1-h} > \frac{(1-C)[2h+C(1-2h)]}{(1+C)[2 - C - 2h(1-C)]} \iff 0 > (1-C) - 2(3-2C)h + 4(1-C)h^2 \]
This inequality only holds for some of the relevant parameter space for $h$. Selfing makes invasion more difficult within the dominance range:

$$0 < h < \frac{(3 - 2C) - \sqrt{5 - 4C}}{4(1 - C)}$$

Invasion is made easier by selfing when dominance falls into the range:

$$\frac{(3 - 2C) - \sqrt{5 - 4C}}{4(1 - C)} < h < \frac{1}{2}$$

**Appendix D: The fraction of parameter space that generates balancing selection**

We are interested in the fraction of parameter space of $s$ and $t$ that permits balancing selection. Let $a$ represent the length of a square that we are interested in finding this fraction for. In other words, for the parameter space $0 < s, t < a$, we wish to calculate the fraction of the cube of area $a^2$ that gives rise to balancing selection.

For the additive case, general balancing selection criteria are provided in eq. (8) of the main text, and these apply under both weak and strong selection (i.e., for the entire space of $0 < s, t < 1$). The fraction of the cube that corresponds to balancing selection is as follows. For the case of sufficiently low selfing, $C < a/(2 - a)$, the fraction is:

$$F = a^2 - \int_0^a \frac{x(1 + C)}{(1 - C) + x(1 + C)} \, dx - \int_0^a \frac{x(1 - C)}{(1 + C) + x(1 - C)} \, dx$$

$$= \frac{1 - C}{1 + C} \ln \left[ 1 + a \frac{1 + C}{1 - C} \right] + \frac{1 + C}{1 - C} \ln \left[ 1 + a \frac{1 - C}{1 + C} \right] - a(2 - a)$$

For higher rates of selfing, $C > a/(2 - a)$, we get the fraction:

$$F = \int_0^a \frac{s(1 - C)}{(1 - s)(1 + C)} \, ds - \int_0^a \frac{s(1 - C)}{(1 + C) + s(1 - C)} \, ds$$

$$= \frac{1 - C}{1 + C} \ln \left[ \frac{1}{1 - a} \right] + \frac{1 + C}{1 - C} \ln \left[ 1 + a \frac{1 - C}{1 + C} \right] - \frac{2a}{1 + C}$$

For the dominance reversal case ($h_m = h_f = h < \frac{1}{2}$), eq. (9) provides balancing selection criteria that are valid for small selection coefficients (roughly, $s, t < 0.1$). Using these, we can
similarly calculate the balancing selection fraction of an a x a square. For a sufficiently low selfing rate, \( C < (1 - 2h)/(2 - 2h) \), we get:

\[
F = a^2 - \frac{(1+C)[1-(1-C)(1-2h)]}{(1-C)[1+C(1-2h)]} \int_0^t dt - \frac{(1-C)[2h+C(1-2h)]}{(1+C)[1+C(1-2h)]} \int_0^t ds
\]

\[
= a^2 \left( 1 - \frac{(1+C)[1-(1-C)(1-2h)]}{2(1-C)[1+C(1-2h)]} \right)
\]

For higher selfing rates, we get:

\[
F = \frac{(1-C)[C+2(1-C)(1-h)]}{(1+C)[1-(1-C)(1-2h)]} \int_0^t ds - \frac{(1-C)[2h+C(1-2h)]}{(1+C)[1+C(1-2h)]} \int_0^t ds
\]

\[
= a^2 \left( 1 - \frac{(1-C)[C+2(1-C)(1-h)]}{2(1+C)[1-(1-C)(1-2h)]} \right) - \frac{2h+C(1-2h)}{1+(1-C)(1-2h)}
\]

These results are used to calculate the curves in Fig. 2 of the main text.

**Appendix E. Efficacy of selection in balancing selection model of SA polymorphism**

We use the general formula for the stationary distribution, from eq. (3) of the main text, the expected change of allele frequency per generation, \( M = f(s, t, C)q(1-q)(\hat{q} - q) \), and the binomial sampling variance of alleles in a Wright-Fisher population, \( V = q(1-q)/(2N_c) \), where \( N_c = N(2 - C)/2 \). The stationary distribution then reduces to:

\[
f(q) = \frac{k}{q(1-q)} \exp \left[ -N(2-C)f(s, t, C)(\hat{q} - q)^2 \right]
\]

as presented in eq. (11) of the main text.

We are interested in determining the efficacy of selection to maintain SA polymorphisms near a given deterministic equilibrium frequency, and for specified values of the male and female selection coefficients. The proportion of the stationary distribution’s density that is near the deterministic equilibrium value is a monotonically increasing function of \( (2 - C)f(s, t, C) \). Thus, for a given value of \( \hat{q} \), we can quantify the effect of selfing on the efficacy of selection by quantifying the effect of selfing on the value of \( (2 - C)f(s, t, C) \). Letting \( S = (s + \)
\[
t/2, \text{ we can calculate } (2 - C)f(s, t, C) \text{ for each of our models of dominance. For the additive case, we get:}
\]
\[
(2 - C)f(s, t, C) = st = (S^2 - \gamma^2/4),
\]
where:
\[
\gamma = \frac{2 - 2\sqrt{1 + S[1 - C - 2\hat{q}][2(C - \hat{q}) + S(1 - C)]}}{2\hat{q} - (1 - C)} = S[2(C - \hat{q}) + S(1 - C)]
\]

The efficacy of selection with versus without selfing is therefore:
\[
\frac{4 - [2(C - \hat{q}) + S(1 - C)]^2}{4 - S^2[1 - 2\hat{q}]^2}
\]

Under dominance reversal conditions \( h = h_m = h_f < \nicefrac{1}{2} \), we get:
\[
(2 - C)f(s, t, C) = (1 - 2h)(1 - C)(2S - \gamma C)
\]
where:
\[
\gamma = \frac{2S[1 - 2(1 - C)[h + (1 - 2h)\hat{q}]]}{1 + C(1 - C)(1 - 2h)(1 - 2\hat{q})}
\]

Thus, the efficacy of selection with versus without selfing is:
\[
(1 - C) \left(1 - C \frac{1 - 2(1 - C)[h + \hat{q}(1 - 2h)]}{1 + C(1 - C)(1 - 2h)(1 - 2\hat{q})}\right)
\]

These expressions were used to calculate the curves in Figure 4 of the text.