PHILOSOPHICAL TRANSACTIONS —OF—

THE ROYAL SOCIETY B | BIOLOGICAL SCIENCES

The demography and population genomics of evolutionary transitions to self-fertilization in plants

Spencer C. H. Barrett, Ramesh Arunkumar and Stephen I. Wright

Supplementary data	"Data Supplement" http://rstb.royalsocietypublishing.org/content/suppl/2014/06/12/rstb.2013.0344.DC1.html
References	**This article cites 47 articles, 16 of which can be accessed free** http://rstb.royalsocietypublishing.org/content/369/1648/20130344.full.html#ref-list-1
Subject collections	Articles on similar topics can be found in the following collections ecology (530 articles) evolution (693 articles) genetics (98 articles) genomics (76 articles) plant science (90 articles)
Email alerting service	Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

Review

Author for correspondence:
Spencer C. H. Barrett
e-mail: spencer.barrett@utoronto.ca

The demography and population genomics of evolutionary transitions to self-fertilization in plants

Spencer C. H. Barrett, Ramesh Arunkumar and Stephen I. Wright

Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, Ontario, Canada M5S 3B2

The evolution of self-fertilization from outcrossing has occurred on numerous occasions in flowering plants. This shift in mating system profoundly influences the morphology, ecology, genetics and evolution of selfing lineages. As a result, there has been sustained interest in understanding the mechanisms driving the evolution of selfing and its environmental context. Recently, patterns of molecular variation have been used to make inferences about the selective mechanisms associated with mating system transitions. However, these inferences can be complicated by the action of linked selection following the transition. Here, using multilocus simulations and comparative molecular data from related selfers and outcrossers, we demonstrate that there is little evidence for strong bottlenecks associated with initial transitions to selfing, and our simulation results cast doubt on whether it is possible to infer the role of bottlenecks associated with reproductive assurance in the evolution of selfing. They indicate that the effects of background selection on the loss of diversity and efficacy of selection occur rapidly following the shift to high selfing. Future comparative studies that integrate explicit ecological and genomic details are necessary for quantifying the independent and joint effects of selection and demography on transitions to selfing and the loss of genetic diversity.

1. Introduction

Hermaphroditism is the primary sexual condition of most seed plants and some animals, particularly invertebrates. The reproductive opportunities provided by cosexuality result in a range of mating systems from obligate outcrossing, through mixtures of outcrossing and self-fertilization (mixed mating), to predominant selfing (autogamy). In angiosperms, this variation is associated with diverse features of the ecology, life history and demography of populations. For example, long-lived woody species of stable habitats are usually highly outcrossing and maintain large effective population sizes, whereas those that are annual colonizers commonly practice high levels of selfing and have smaller populations [1]. Such associations implicate environmental and demographic conditions as playing important roles in the evolution and maintenance of plant mating systems, which are key components of life history and the recombination system. Because a variety of demographic and genetic processes can influence levels of recombination and effective population size, identifying their relative importance will provide a more complete understanding of the evolution of mating systems.

The evolution of predominant self-fertilization from outcrossing represents the most important reproductive transition in flowering plants. Although only approximately 10–15% of angiosperm species are highly autogamous, there is extensive evidence from numerous herbaceous taxa of multiple independent origins of autogamy from outcrossing [1–3]. The frequency with which selfing originates provides opportunities for identifying convergent patterns, and the mechanisms and consequences of this transition. Accordingly, efforts to understand how and why selfing evolves has had a long and venerable history, beginning with the early naturalists, including Charles Darwin, George Henslow

and Hermann Müller (reviewed by Lloyd [4]), and continuing today using diverse comparative, experimental and genomic approaches [5–7]. In addition, a rich body of theory (reviewed by Goodwillie *et al.* [8]) has developed on the selective mechanisms governing mating system variation and the genetic and evolutionary consequences of transitions to selfing.

There are four principal reasons why the evolution of selfing from outcrossing has attracted so much attention from evolutionary biologists: (i) transitions to selfing are associated with major changes to floral morphology, sex allocation and life history, thus contributing to taxonomic and ecological diversity and serving as valuable systems for the study of plant adaptation [9]; (ii) because selfing enables individuals to start colonies following dispersal, or to reproduce under conditions of low density, the transition to selfing can have significant demographic and biogeographic consequences [4]; (iii) predominant selfing has a profound effect on the genome, and autogamous species can provide important model systems for investigating the evolutionary consequences of restricted recombination. High rates of selfing reduce levels of heterozygosity, restrict recombination and increase selection at linked sites (linked selection) due to both deleterious and advantageous mutations, thus leading to a genome-wide loss of genetic diversity. These changes reduce the efficacy of natural selection and increase the probability of fixation of deleterious mutations owing to genetic drift [10]; and (iv) predominant selfing has important influences on macroevolutionary processes affecting speciation and extinction rates and patterns of lineage diversification [5,7]. These diverse facets of the biology of selfing present both challenges and opportunities for evolutionary biologists interested in adaptation and speciation. In particular, opportunities arise to better understand the evolutionary causes and consequences of transitions in life history and genetic system, although because many of the genetic and demographic factors involved change concurrently, disentangling their individual and joint effects presents an important challenge.

The late plant evolutionary biologist Leslie Gottlieb, to whom this volume is dedicated, made valuable contributions to our understanding of the evolution of selfing. Gottlieb's classic investigations of sympatric speciation in annual *Stephanomeria* provided evidence that the origin of *Stephanomeria malheurensis* from *Stephanomeria exigua* ssp. *coronaria* was associated with a shift from outcrossing to selfing. This transition resulted from the breakdown of self-incompatibility and the development of hybrid sterility, owing to chromosomal structural differences between the progenitor and derivative [11,12]. Today, it is widely recognized that the loss of self-incompatibility is commonly associated with shifts from outcrossing to selfing, and in many cases this also appears to promote speciation [13]. In *Clarkia*, Gottlieb also investigated the evolutionary origins of selfing from outcrossing. Using allozyme variation, he demonstrated two independent origins in *Clarkia concinna* [14] and a single origin in *Clarkia xantiana* [15]. These studies were among the first to use molecular data to trace transitions from outcrossing to selfing and to show that selfing populations contained markedly reduced amounts of genetic variation compared with outcrossing populations, a result that Gottlieb extended by surveying additional taxa [16], and is now confirmed in numerous species [17]. His pioneering efforts helped to stimulate subsequent work on the origins of selfing and its influence on patterns of genetic diversity.

Here, we consider recent developments and ongoing challenges in understanding the evolution of selfing and the demographic context in which it occurs. The advent of high-throughput DNA sequencing provides opportunities for investigators to address a variety of questions concerned with the selective mechanisms causing transitions, as well as their genomic and evolutionary consequences. We focus in particular on the problem of distinguishing the relative importance of demographic and genetic processes affecting the selection of selfing, and the reduction in genetic diversity that accompanies this transition. To provide a general background, we begin by summarizing the two main hypothesized selective mechanisms for selfing, and the principal reproductive modifications that cause increased selfing rates and what is known about their genetic architecture. We then consider the demographic circumstances in which selfing evolves and how this may influence mechanisms of selection. Using simulation studies and comparative data from related selfers and outcrossers, we then address three main questions. (i) Can molecular evidence be used to distinguish the mechanisms of selection for selfing, as previous authors have proposed? (ii) What is the relative importance of demographic factors and linked selection in causing reductions in genome-wide diversity in recently derived selfing populations, and over what time period do these processes become evident? (iii) How rapidly does the transition to selfing lead to reductions in the efficacy of natural selection across the genome? Our principal conclusions are that the effects of linked selection on diversity and the efficacy of selection can act quite quickly following the shift to selfing, and that there is little evidence to be gleaned from molecular data for strong bottlenecks during the earliest stages of the shift to selfing.

2. Selection of self-fertilization

Diverse reproductive, demographic and genetic factors influence the transition from outcrossing to selfing, of which the magnitude of inbreeding depression is generally considered most influential for the maintenance of outcrossing [18,19]. The assured reproduction that selfing provides when pollen vectors and/or compatible mates limit cross-pollination, and the transmission advantage of genes which cause selfing are the two most general selective mechanisms that have been proposed to explain the shift to selfing. The 'reproductive assurance hypothesis' traces back to Darwin [20] and is the most common explanation for why selfing evolves. Less often examined is the 'automatic selection hypothesis', which is based on Fisher's [21] idea that a gene for selfing has a 3 : 2 transmission advantage when it arises in an outcrossing population. Determining the relative importance of these two hypotheses is not straightforward, particularly where both processes operate during the transition to selfing. Distinguishing the two hypotheses requires determining empirically the modes of self-fertilization in a population [19], and the extent to which selection of mating system modifiers occurs through pollen and/or seed, which necessitates information on pollen and seed discounting [6]. Neither of these parameters has been simultaneously measured in any plant population, and currently it is unclear how important they are in counteracting the selective advantage of selfing, particularly in comparison with inbreeding depression.

An alternative approach for investigating the relative importance of the two main hypotheses for the evolution of selfing was first suggested by Schoen *et al.* [22]. They suggested

3

that the demographic and genetic processes associated with reproductive assurance and automatic selection might be expected to result in different molecular signatures that could potentially be distinguished by examining patterns of neutral diversity among selfing populations. Specifically, they suggested that the evolution of selfing through reproductive assurance should be accompanied by a genome-wide reduction in genetic diversity due to population bottlenecks, potentially resulting in founder events. By contrast, the evolution of selfing through automatic selection involves the spread of mating system modifiers in outcrossing populations, which should cause reduced diversity at the modifier itself, but not at unlinked neutral loci. They also suggested that recurrent bottlenecks due to reproductive assurance should lead to elevated among-population variance in diversity, as populations will be at different degrees of recovery from bottlenecks.

Recently, molecular population genetic data from five taxa and a total of six relatively recent transitions to selfing (less than 10–150 ka) were used to evaluate mechanisms causing shifts to selfing [6]. In half of these cases, reductions in diversity were severe, consistent with predictions from reproductive assurance, while the remaining cases showed no evidence for a greater reduction in diversity than expected by selfing alone. However, there are several thorny conceptual and empirical issues that need to be addressed in using molecular population genetic data to distinguish the selective mechanisms driving the evolution of selfing.

The above predictions rest on various assumptions, several of which may be violated in selfing populations. First, one key assumption is that selfing evolves through mutations of small effect [22]. If major mutations causing large shifts in selfing rate arise and spread rapidly, genome-wide reductions in diversity may occur in the absence of selection for reproductive assurance. Although small-effect modifiers are reported in some systems, the breakdown of self-incompatibility and floral modifications promoting selfing can involve genes of large effect (see below). Second, subsequent gene flow from outcrossing populations following the shift to selfing could erode an initial signal of a bottleneck during the transition to selfing [6]. Third, additional factors during the origin of selfing other than bottlenecks are known to reduce diversity in selfing relative to outcrossing populations [10], as we explore in detail later in this article.

3. Modification to reproductive traits that promote selfing

The shift to high selfing rates is generally associated with a series of modifications to reproductive traits often beginning with the facility for autonomous self-pollination, and culminating in the evolution of the selfing syndrome [9]. A challenging problem, also shared with work in speciation genetics, is to identify which traits directly initiate the process of mating system change, and which are consequences of increased selfing rates, for example subsequent changes in sex allocation to male function [23]. In addition, determining the genetic basis of modifiers of the selfing rate can provide insight into the mechanisms responsible for transitions.

In self-incompatible species, the initial step usually involves the breakdown of the physiological self-recognition system and the evolution of self-compatibility. This change commonly involves major mutations producing fully self-compatible

variants, although it can be gradual involving different degrees of self-fertility [24]. In self-compatible species, three morphological modifications commonly cause increasing selfing rates: a reduction in stigma–anther separation (herkogamy), the loss of dichogamy (timing of anther dehiscence and stigma receptivity) and a decrease in flower size. The extent to which these traits function alone or in concert to initiate increased selfing is poorly understood in most lineages. This is especially the case in highly autogamous species with well-developed selfing syndromes, in which numerous substitutions have probably accumulated since their early divergence from outcrossing ancestors.

In some taxa, it has been possible to identify the early changes provoking mating system evolution. For example, in *Collinsia* spp. a weakening of dichogamy appears to be the critical first step causing increased selfing [25], whereas in *Eichhornia paniculata* the loss of herkogamy through stamen elongation, governed by recessive modifiers of large effect, is of primary importance [26]. In North American populations of *Arabidopsis lyrata*, the breakdown of self-incompatibility system is associated with increased selfing rates [27], but this physiological change does not appear to be associated with obvious morphological modifications to flowers causing autonomous self-pollination. In *Capsella rubella*, flower size reduction appears to have followed the loss of self-incompatibility [28]. Interestingly, the breakdown of self-incompatibility in this system alone confers a moderate facility for autonomous selfing, which is approximately 50% that of the derived selfing species [28], indicating that the mutation causing self-compatibility represents a major-effect mating system modifier. In *Leavenworthia alabamica*, two independent losses of self-incompatibility have occurred, one of which is accompanied by floral modifications and high selfing rates, whereas the other shows no obvious changes to floral traits and populations exhibit mixed mating [29]. Although the loss of self-incompatibility provides opportunities for selfing, the extent to which this occurs will depend on several factors. In taxa with generalized pollinators and relatively small flowers (e.g. many Brassicaceae), the loss of incompatibility alone may be sufficient to permit substantial selfing, as in *C. rubella*, through various modes of selfing (e.g. prior, delayed and pollinator mediated, see Lloyd [19] and below). By contrast, in taxa with larger more specialized flowers, the availability of standing genetic variation for traits promoting autonomous self-pollination is probably necessary for high selfing rates to evolve.

Information on the genetic architecture of mating system modifiers can provide insight into the selective mechanisms driving the transition from outcrossing to selfing, its demographic context and the tempo with which it occurs. Key questions include: what is the relative importance of mutations of large versus small effect governing selfing rate modifiers? How important are new mutations versus standing genetic variation as a source of modifier alleles? Answers to these questions are important because theoretical work indicates that a mutation of large effect causing predominant selfing can spread to fixation despite high inbreeding depression [18]. If this occurs, it has the potential to result in a genome-wide selective sweep because the highly selfing variant may be reproductively isolated from outcrossing individuals immediately, adding to the diverse forces reducing diversity in recently derived selfing populations. Moreover, if bottlenecks associated with reproductive assurance play an important

role in the shift to selfing, the importance of standing genetic variation as a source of selfing modifiers may be quite limited [30]. Unfortunately, few detailed studies have been conducted on the genetics of selfing traits so at this stage there are no clear answers to these questions.

Quantitative trait locus (QTL) mapping of crosses between self-compatible, outcrossing *Mimulus guttatus* and selfing *Mimulus nasutus* [31] provided no evidence of large-effect QTLs for flower width, stigma–anther separation and several other traits associated with selfing. Although the two *Mimulus* species have undergone considerable floral divergence, the absence of major QTLs governing floral traits is consistent with the hypothesis that selfing has evolved gradually by the sequential fixation of modifiers of small effect. In self-compatible groups such as *Mimulus*, fluctuating selection on outcrossing rates, due to year-to-year variation in population size and pollinator service, may lead to relatively high levels of standing genetic variation for mating system traits and thus more often polygenic control. Indeed, a novel experimental evolution study, conducted with the presence and absence of pollinators in an outcrossing population of *M. guttatus*, revealed remarkably rapid evolution of floral traits promoting autonomous self-pollination during five generations in the 'pollinator-free' treatment [32], implying high levels of standing genetic variation for traits influencing the mating system. Alternatively, new mutations could play a more prominent role during transitions to selfing in self-incompatible species because guaranteed outcrossing may be associated with less temporal variation in selection and thus lower levels of standing genetic variation on mating system traits.

Consistent with this prediction, QTL mapping of floral traits from crosses between self-incompatible, outcrossing *Capsella grandiflora* and selfing *C. rubella* revealed a different genetic architecture than was evident in *Mimulus*. Flower size traits were governed by changes at a relative small number of loci, some with large effect, in addition to the single dominant allele that causes self-compatibility [28,33]. Thus, it appears that genes of major effect for both self-compatibility and floral traits may have been important in the spread of selfing in *C. rubella*, and the spread of these major-effect alleles could have been an important contributor to genome-wide reductions in diversity in the early stages of mating system evolution.

4. Ecological and demographic context

The observation that autogamous plant species are not randomly distributed with respect to biogeography, colonizing ability, habitat, demography and life history has often been made (reviewed by Lloyd [4]). Although there are numerous exceptions, selfing is commonly associated with island colonization and geographical marginality; weediness; ephemeral, open disturbed habitats; low density; and short-lived life forms (e.g. annuality). Baker's law and various aspects of the reproductive assurance hypothesis are most often used to explain these associations. Individuals that are self-compatible, with a capacity for autonomous self-pollination, are likely to be favoured over outcrossing individuals when pollinators and/or mates are in short supply, as might often occur at low density or following dispersal. Although such explanations seem perfectly plausible in explaining these correlations, it has proven more difficult to obtain convincing experimental evidence

that these ecological associations are directly involved in driving transitions from outcrossing to selfing, rather than simply being a consequence of the possession of an autogamous mating system. The direction of causality cannot be determined from ecological and demographic correlates of selfing alone.

Studies of the ecology of wide-ranging species with variable mating systems can provide insights into the potential role of demographic factors in causing shifts to selfing. In *E. paniculata*, selfing has evolved on multiple occasions associated with the evolutionary breakdown of tristyly [34]. Several lines of evidence indicate that repeated bottlenecks and genetic drift in small populations have played an important role in destabilizing tristyly [35]. The main concentration of populations in northeastern Brazil is largely outcrossing and primarily pollinated by specialist long-tongued bees. By contrast, in Cuba and Jamaica, where specialist bees do not visit populations, highly selfing populations predominate, consistent with Baker's law and reproductive assurance. Selfing and mixed mating populations occur in Brazil, but have not evolved the selfing syndrome evident in island populations, probably owing to their more recent origin and occasional gene flow from tristylous populations. In *E. paniculata*, selfing rates are negatively correlated with population size, and both selfing rates and population size are strongly correlated with diversity [36]. Both reproductive assurance and automatic selection probably play a role in the evolution of selfing in Brazil, with their relative importance depending on the extent to which small populations of *E. paniculata* receive pollinator service, and selfing variants gain fitness as male parents.

Insufficient outcross pollen delivery due to low density has frequently been invoked to account for the evolution of selfing, but relatively few studies have investigated this problem directly, despite extensive experimental data on pollen limitation in numerous species [37]. In *L. alabamica*, two self-compatible races have evolved independently from self-incompatible, outcrossing populations in different parts of the periphery of the species' range, and are characterized by significantly smaller populations, often of lower density [38]. Molecular data have been interpreted as suggesting that a bottleneck could have facilitated the evolution of selfing by reproductive assurance in one race, but not in the other [29], although a study of pollen limitation found no evidence that it has played a role in the evolution of self-compatibility [38]. By contrast, field experiments on pollen limitation in *Clarkia xantiana* provided evidence for the evolution of selfing by reproductive assurance in populations of small size or low density due to few mates and effective pollinators [39]. Molecular data on the evolutionary history of selfing consistent with a strong genetic bottleneck have also been interpreted to support the reproductive assurance hypothesis [40]. The field study of *Clarkia* is particularly significant because it demonstrated context-depend natural selection on selfing traits; in small populations, isolated from congeners, selection was strongest for reduced herkogamy and protandry.

5. The causes for loss of genetic diversity in selfing populations

As discussed earlier, researchers have used data on molecular variation from recently derived selfing populations and their outcrossing progenitors to test for evidence of bottlenecks
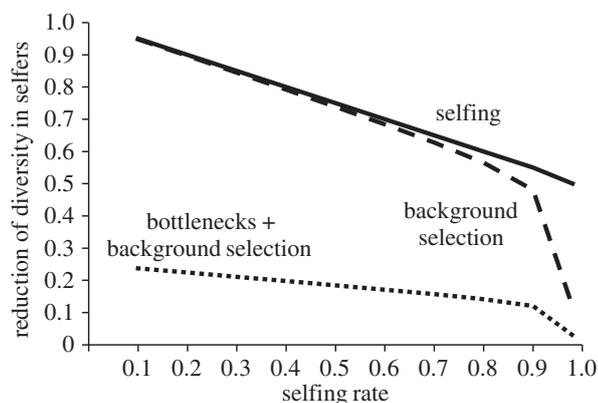
Figure 1. Expected reduction in genetic diversity in selfing compared to outcrossing populations with selfing (solid line), selfing and background selection (long dashed line) and with selfing, background selection and population bottlenecks (short dashed line). The expected reduction in diversity due to selfing alone was $1/(1 + F)$, where F is the inbreeding coefficient. The expected reduction in diversity due to background selection was modelled after Glémin & Ronfort [41], with selection and mutation parameters taken from estimates in the Brassicaceae, as described in the electronic supplementary material. The effect of a bottleneck was assumed to be a reduction in diversity to 25% of the original value, regardless of selfing rate.

and founder events and assess the role of reproductive assurance in the evolution to selfing [6]. However, as illustrated in figure 1, multiple factors can contribute to reduced diversity in selfing populations, and distinguishing their relative contributions to overall reductions in diversity following the transition to selfing represents a major challenge. Under neutrality, selfing lowers the effective population size up to twofold due to the reduced effective number of alleles in a population. Also, because of reduced effective rates of recombination, neutral diversity will be lowered further due to the action of linked selection, including background purifying selection and selective sweeps of advantageous mutations. If strong enough and with sufficient time to reach the new equilibrium, these effects will erode historical signals associated with the mechanisms of selection for selfing. Thus, in making historical inferences on the selective mechanisms for selfing, it is first necessary to assess the effect that different strengths of linked selection have, relative to other factors, in reducing variation in selfing populations.

Previously, the lack of estimates of deleterious mutation rates and their selection coefficients has made it difficult to assess the impact that background selection might have on plant genomes [42]. This has been particularly problematic because both background selection and population bottlenecks will cause genome-wide reductions in diversity [10]. Although our understanding of mutation rates and their selection coefficients is still limited, it is now possible to more explicitly assess the role of background selection in plant populations to predict the extent to which it can explain the loss of variation in selfing populations. In particular, genome-wide estimates are now available in the Brassicaceae for the per-base pair mutation rate due to nucleotide substitutions [43], the number of base pairs in the genome under selection [44] and the strength of selection on functional mutations [45]. Whereas both mutation rates and estimates of the proportion of the genome under selection are likely underestimates, these values allow us to obtain a lower bound prediction of the reduction in within-population diversity expected due to background selection in selfing populations.

Using the approximation of Glémin & Ronfort [41], and available estimates of mutation and recombination parameters, our simulations indicate that background selection under high selfing rates is expected to reduce the equilibrium effective population size to about 10% that of an obligate outcrosser (figure 1), well below the 50% reduction expected in the absence of background selection. The reductions in diversity as a result of background selection are primarily restricted to populations with rates of selfing more than 90–95%. The effects of background selection illustrated in figure 1 apply to long-term equilibrium, and thus a critical question concerns how rapidly these effects are likely to occur following the transition to selfing. If reductions in diversity due to linked selection occur quickly following the transition, it will be difficult to use molecular data to infer the role of reproductive assurance because most selfing populations will have rapidly reached the new equilibrium of low diversity.

To determine how rapidly the loss in diversity occurs following a shift to selfing, we simulated mating system evolution in the presence and absence of deleterious mutations. We kept demographic variables constant and used mutation and recombination parameters resembling the *Arabidopsis* genome. We also used a more realistic genome structure in the simulations than in analytical approximations, incorporating spatial patterning of neutral and selected sites, rather than equal spacing of selected sites across the genome. In our simulations, we varied the spatial patterning and strength of selection against deleterious mutations. This effectively varies the strength of background selection, because greater spacing between neutral and selected sites and stronger purifying selection should reduce the intensity of linked selection. Figure 2 illustrates the ratio of diversity in the derived selfing population relative to the outcrossing population as a function of time since the origin of selfing. When we did not include deleterious mutations, we observed a twofold reduction of diversity in selfing populations as expected, and the loss of diversity occurred over a timescale of approximately $2N$ generations, where N is the effective size of the ancestral outcrossing population (figure 2a). This result fits with coalescent predictions, because it takes approximately $4N$ generations to reach equilibrium [46], and the 50% reduction in effective population size causes the decline in diversity to be more rapid. With purifying selection acting on deleterious mutations, we observed a greater than twofold reduction in diversity (figure 2b–f), and the time to reach equilibrium was more rapid with increasing strengths of background selection owing to the lowered effective population sizes. Although not considered here, the effect of selective sweeps of beneficial mutations could also further reduce diversity in recently derived selfing populations. Given emerging evidence for high rates of positive selection in some plant populations [47], and the possibility of selective sweeps associated with new fitness optima due to the mating system shift, positive selection may also be a significant factor contributing to reductions in genetic diversity in selfing populations. These results imply that reductions in diversity due to linked selection in highly selfing populations can occur relatively rapidly, particularly under conditions of strong background selection where the effective population size is small.

If the effects of linked selection are a major cause of diversity loss in recently derived selfing populations, we can make several predictions. First, reductions in diversity should be primarily restricted to populations that initially transition to high selfing rates ($s > 90\%$), and be less evident where the early
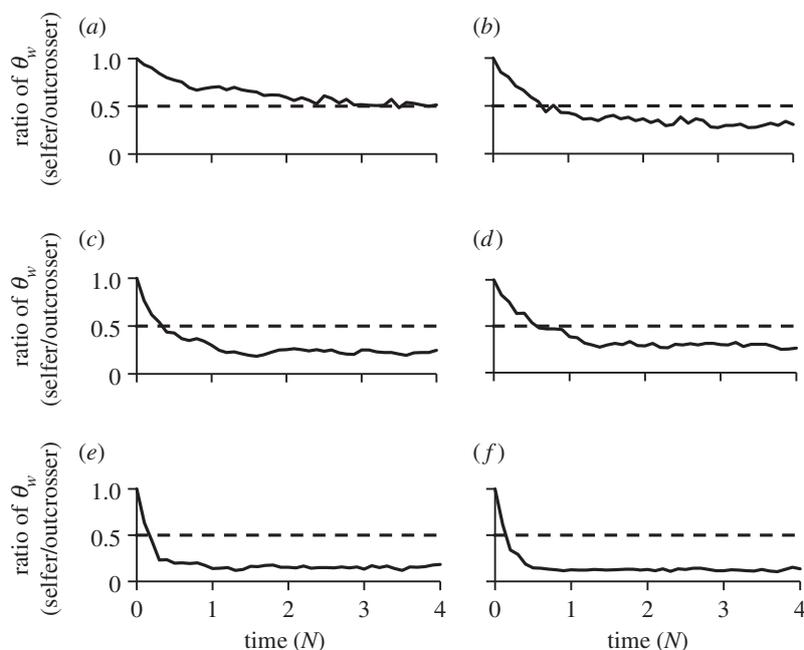
Figure 2. The reduction in synonymous diversity (θ_w) as a function of time since the transition to selfing, in units of N generations, where N is the effective population size of the ancestral outcrossing populations. We performed forward population genetic simulations using 300 individuals with 100 Mb genomes comprising alternating non-coding (NC) and coding (C) regions of various sizes. All NC sites were neutral, whereas for C sites 25% were neutral and 75% were deleterious under variable selection pressures as measured by s, the mean selection coefficient acting on deleterious mutations. Details of simulations are in the electronic supplementary material. Shown are plots for: (a) $s = 0$, C = 200 bp, NC = 800 bp; (b) $s = -0.1$, C = 200 bp, NC = 800 bp; (c) $s = -0.05$, C = 200 bp, NC = 800 bp; (d) $s = -0.005$, C = 200 bp, NC = 800 bp; (e) $s = -0.05$, C = 350 bp, NC = 650 bp; and (f) $s = -0.05$, C = 500 bp, NC = 500 bp. The ratio of the means across five runs for each interval (solid line) and the expected twofold reduction due to selfing alone (dashed line) are plotted. The plots are ordered by parameters that led to increasingly stronger background selection. Without background selection, the loss of diversity occurred within approximately $2N$ (a). As s was decreased, the loss of diversity occurred within $1 - 2N$ generations (b–d). Reduction in the size of NC, even under moderate s, led to the strongest effect of background selection where the loss of diversity occurred within $0.5N$ (e,f).

stages of the spread of selfing variants are associated with mixed mating at the population level (figure 1). Second, reductions in diversity should be more prominent in relatively older selfing populations, which have had more time for the effects of linked selection to operate, and are more likely to have reached the new equilibrium diversity due to linked selection. Alternatively, if selection for reproductive assurance causes the initial spread of selfing modifiers, we might expect populations with smaller increases in selfing rate to show reductions in diversity beyond neutral expectation, owing to the effects of population bottlenecks, as illustrated in the dotted line in figure 1 (also see [22]). Furthermore, bottleneck effects associated with the transition to selfing should be apparent even in the most recent transitions to higher selfing rates.

To assess these predictions, we conducted a literature survey to obtain estimates of synonymous nucleotide diversity for related outcrossing and selfing species pairs. We only considered studies that reported diversity estimates for species pairs in which selfing species or populations were recently derived. We excluded more distantly related pairs, as there was a greater chance that forces unrelated to mating system shifts might have influenced their patterns of diversity. We estimated the ratio of nucleotide diversity in selfing populations relative to their outcrossing progenitor and excluded studies that estimated diversity using less than five genes, because small numbers of loci will not give reliable estimates of a reduction in diversity. Consistent with the expectations of linked selection, we found that only shifts to high rates of selfing led to substantial reductions in diversity relative to outcrossing progenitors (figure 3a). Mixed mating populations maintained moderate levels of diversity, indicating that significant population bottlenecks have not occurred and/or that

gene flow from outcrossing populations after bottlenecks has masked the initial demographic effects on the genome. Our survey results also indicate that, as might be expected, the reduction of diversity in selfing populations in the early stages of the transition was generally less compared with older lineages (figure 3b). Although the age estimates of selfing lineages are highly dependent on assumptions about mutation rates and generation times, the populations with the strongest reductions in genetic diversity do appear to be among the most highly diverged and show the most extreme changes in floral morphology [6]. These results are consistent with predictions from linked selection, and/or if older selfing lineages with high selfing rates have experienced recurrent bottlenecks and reduced gene flow. However, they are inconsistent with the prediction that the earliest stages of mating system shifts are typically associated with strong genetic bottlenecks.

Given the evidence for longer term effects of linked selection reducing effective population size, another key question concerns whether an explicit coalescent framework can still be developed to infer the number of founding lineages. In a recent study [48] examining founding haplotypes in the highly selfing *C. rubella*, the authors concluded that, because of a long-term reduction in effective population size, the data were consistent with the number of founding individuals ranging from three to infinite, despite the very severe reduction in diversity and effective population size. Thus, in this example, the authors pointed out that there was little ability to distinguish modes of selection on mating system, despite the relatively recent shift approximately 50 000 years ago. While similar analyses have not yet been conducted in other systems, the overall patterns to date suggest that selfing populations experience longer term reductions in effective
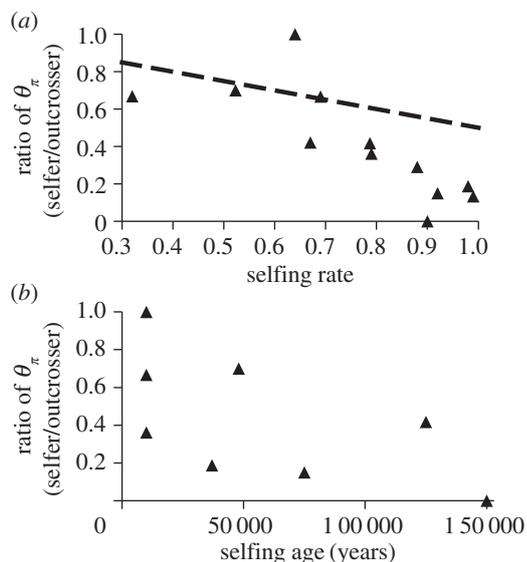
Figure 3. Ratio of synonymous diversity (θ_π) for related selfing and outcrossing populations and species in relation to: (a) the rate of selfing in populations with the higher degree of self-fertilization and (b) the estimated age (in years) of the transition to higher rates of selfing. In (a), the expected reduction in diversity due to selfing alone was $1/(1 + F)$, where F is the inbreeding coefficient (dashed line).

population size owing to linked selection, restrictions of gene flow from outcrossing populations and/or subsequent bottlenecks, with limited evidence for a direct collapse in diversity during the transition itself.

6. Selective consequences of transitions to selfing

There is growing evidence that the transition from outcrossing to selfing has important evolutionary consequences influencing speciation and extinction rates and the diversification of lineages [5,7]. In common with the causes of reduced genetic diversity in selfing populations, both genetic and demographic factors probably influence these evolutionary processes. Thus, a major challenge is to tease apart their relative importance. Here, we restrict our discussion to the question of why selfing lineages often appear to be ephemeral compared with outcrossing lineages. Because highly autogamous species are often short-lived and occupy unpredictable environments, their populations are more likely than long-lived outcrossing species to be susceptible to environmental catastrophes leading to local extinctions. Thus, demographic factors associated with small population size may, in part, contribute to higher extinction rates. However, genetic processes may also contribute towards higher extinction rates because the reduced effective population size of selfing populations is expected to reduce not only neutral diversity but also the efficacy of purifying selection [30]. This reduction in the intensity of natural selection can lead to a long-term accumulation of deleterious mutations, potentially contributing to increased extinction risk [7]. One key question is how rapidly deleterious mutations accumulate following the shift to selfing, as this may, in part, determine the relative importance of demographic versus genetic effects in driving species' extinction. Our simulation results highlight that the effects of linked selection can happen rapidly
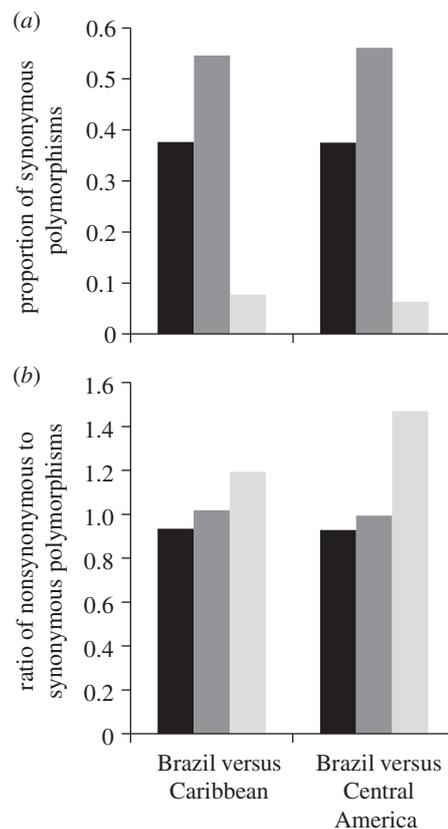

Figure 4. Shared (black) and unique sequence polymorphisms between outcrossing (dark grey) and selfing (light grey) populations of *E. paniculata* sampled from different parts of the geographical range: (a) proportion of synonymous polymorphisms and (b) ratio of nonsynonymous relative to synonymous polymorphisms. Shown are pairwise comparisons between a single diploid chromosome from an outcrossing individual and haploid chromosomes from two selfing individuals. Two selfers were used as the high homozygosity of selfing genomes reduces the effective number of chromosomes sampled from selfing individuals to approximately half of that of outcrossing individuals. Details of sampling and methods to estimate the proportion of polymorphisms are provided in the electronic supplementary material.

following the shift to selfing, implying that the early stages of deleterious mutation accumulation may be apparent soon after the transition.

Early tests of the hypothesis of a reduced efficacy of selection in selfers generally showed limited support for the prediction, but these studies were limited by the extent of the genome that was surveyed [49–51]. Furthermore, given the recent origin of most selfing populations, using substitution rates to test for relaxed selection may have limited power because of the difficulty in identifying the changes that have occurred since the shift to selfing. The use of polymorphism data from selfing populations and their outcrossing progenitors can provide a more powerful means to detect evidence for the reduced efficacy of selection in the genome of selfing populations [47,48], and to assess how rapidly reductions in the efficacy of selection can occur.

We used this approach in *E. paniculata* to identify polymorphisms in populations of two independently derived selfing lineages from the Caribbean and Central America, and their outcrossing progenitors from Brazil. The results indicate that many of the polymorphisms in populations of the selfing lineages were shared with outcrossing populations (figure 4a). Moreover, these shared polymorphic sites were under similar strengths of selection as polymorphic sites

unique to the outcrosser, as indicated by comparable ratios of nonsynonymous relative to synonymous polymorphisms (figure 4b). By contrast, relaxation of selective constraints in selfing populations was most prominent for polymorphic sites unique to their populations, as indicated by elevated ratios of nonsynonymous relative to synonymous polymorphisms, when compared with the ratios obtained for shared polymorphic sites, and polymorphic sites unique to outcrossing populations. The signal was strongest for selfing populations from Central America, a result consistent with previous work indicating increased nonsynonymous mutations in these populations [52]. Overall, these results combined with similar studies in *Capsella* [48] and *Collinsia* [53] suggest that the reduced efficacy of purifying selection in selfing lineages is detectable even following quite recent transitions to selfing. These population genomic consequences appear to occur quickly following a shift to selfing, consistent with our earlier conclusions that the effects of linked selection can occur rapidly owing to strong reductions in effective population size.

7. Unresolved questions and future work

The simulations and comparative evidence from related selfers and outcrossers we report in this article indicate that current molecular data provide little evidence for bottlenecks during the transition from outcrossing to selfing, and that the effect of linked selection is likely to be a major contributor to the loss of genetic diversity in recently derived selfing populations. Of course these results do not rule out an important role for reproductive assurance in the transition to selfing, nor do they reject the possibility that the shift to high selfing rates leads to recurrent and severe bottlenecks. However, they do call into question the power to test hypotheses involving selective mechanisms using molecular data. Selection for reproductive assurance may not necessarily cause a signal of strong founder events, and low diversity in recently diverged selfing populations can result from a number of non-mutually exclusive causes. Given that partially selfing populations also show little molecular evidence for bottlenecks, the results to date do not support the hypothesis that early transitions to partial selfing are accompanied by strong bottlenecks due to selection for reproductive assurance.

Future work integrating detailed demographic and genomic information with variation in selfing rates should help to better determine the relative importance of the diverse factors affecting levels of diversity. With growing amounts of genomic data, explicit parameterization of background selection and positive selection will be important for predicting the effects of linked selection alone, and for assessing how important demographic factors may be in also contributing to reductions in diversity. In some systems, such as *E. paniculata*, demographic and mating system parameters are so strongly linked [37] that it may be difficult to uncouple their relative importance in natural populations. In such cases, experimental approaches will be required. Field studies of experimental populations of different size and density that measure the strength of selection on traits directly promoting selfing would also be valuable in providing insights into the demographic context in which mating system shifts occur and their effects on genetic diversity.

References

1. Barrett SCH, Harder LD, Worley AC. 1996 Comparative biology of pollination and mating. *Phil. Trans. R. Soc. Lond. B* **351**, 1271–1280. (doi:10.1098/rstb.1996.0110)

2. Stebbins GL. 1974 *Flowering plants: evolution above the species level*. Cambridge, MA: Harvard University Press.

3. Jain SK. 1976 The evolution of inbreeding in plants. *Annu. Rev. Ecol. Syst.* **7**, 469–495. (doi:10.1146/annurev.es.07.110176.002345)

4. Lloyd DG. 1980 Demographic factors and mating patterns in angiosperms. In *Demography and evolution in plant populations* (ed. OT Solbrig), pp. 67–88. Oxford, UK: Blackwell.

5. Igic B, Busch JW. 2013 Is self-fertilization an evolutionary dead end? *New Phytol.* **198**, 386–397. (doi:10.1111/nph.12182)

6. Busch JW, Delph LF. 2012 The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Ann. Bot.* **109**, 553–562. (doi:10.1093/aob/mcr219)

7. Wright SI, Kalisz S, Slotte T. 2013 Evolutionary consequences of self-fertilization. *Proc. R. Soc. B* **280**, 20130133. (doi:10.1098/rspb.2013.0133)

8. Goodwillie C, Kalisz S, Eckert CG. 2005 The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical expectations, and empirical evidence. *Annu. Rev. Ecol. Evol. Syst.* **36**, 47–79. (doi:10.1146/annurev.ecolsys.36.091704.175539)

9. Sicard A, Lenhard M. 2011 The selfing syndrome: a model for studying the genetic and evolutionary basis of morphological adaptation in plants. *Ann. Bot.* **107**, 1433. (doi:10.1093/aob/mcr023)

10. Charlesworth D, Wright SI. 2001 Breeding systems and genome evolution. *Curr. Opin. Genet. Dev.* **11**, 685–690. (doi:10.1016/S0959-437X(00)00254-9)

11. Gottlieb LD. 1973 Genetic differentiation, sympatric speciation, and the origin of a diploid species of *Stephanomeria*. *Am. J. Bot.* **60**, 545–553. (doi:10.2307/2441378)

12. Brauner S, Gottlieb LD. 1987 A self-compatible plant of *Stephanomeria exigua* subsp. *coronaria* (Asteraceae) and its relevance to the origin of its self-pollinating derivative *S. malheurensis*. *Syst. Bot.* **12**, 299–304. (doi:10.2307/2419325)

13. Goldberg EE, Igic B. 2012 Tempo and mode in plant breeding system evolution. *Evolution* **66**, 3701–3709. (doi:10.1111/j.1558-5646.2012.01730.x)

14. Allen GA, Gottlieb LD, Ford VS. 1991 Electrophoretic evidence for the independent origins of two self-pollinating subspecies of *Clarkia concinna* (Onagraceae). *Can. J. Bot.* **69**, 2299–2301. (doi:10.1139/b91-290)

15. Gottlieb LD. 1984 Electrophoretic analysis of the phylogeny of the self-pollinating populations of *Clarkia xantiana*. *Plant Syst. Evol.* **147**, 91–102. (doi:10.1007/BF00984582)

16. Gottlieb LD. 1977 Electrophoretic evidence and plant systematics. *Ann. Miss. Bot. Gard.* **64**, 161–180. (doi:10.2307/2395330)

17. Hamrick JL, Godt MJW. 1996 Effects of life history traits on genetic diversity in plant species. *Phil. Trans. R. Soc. Lond. B* **351**, 1291–1298. (doi:10.1098/rstb.1996.0112)

18. Lande R, Schemske DW. 1985 The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. *Evolution* **39**, 24–40. (doi:10.2307/2408514)

19. Lloyd DG. 1992 Self- and cross-fertilization in plants. II. The selection of self-fertilization.

9

rstb.royalsocietypublishing.org   Phil. Trans. R. Soc. B **369**: 20130344

Int. J. Plant Sci. **153**, 370–380. (doi:10.1086/297041)

20. Darwin C. 1876 *The effects of cross- and self-fertilization in the vegetable kingdom*. London, UK: John Murray.

21. Fisher RA. 1941 Average excess and average effect of a gene substitution. *Ann. Eugen.* **11**, 53–63. (doi:10.1111/j.1469-1809.1941.tb02272.x)

22. Schoen DJ, Morgan MT, Bataillon T. 1996 How does self-pollination evolve? Inferences from floral ecology and molecular genetic variation. *Phil. Trans. R. Soc. Lond. B* **351**, 1281–1290. (doi:10.1098/rstb.1996.0111)

23. Charnov EL. 1982 *The theory of sex allocation*. Princeton, NJ: Princeton University Press.

24. Good-Avila SV, Mena-Alí JI, Stephenson AG. 2008 Genetic and environmental causes and evolutionary consequences of variation in self-fertility in self-incompatible species. In *Self-incompatibility in flowering plants: evolution, diversity, and mechanisms* (ed. VE Franklin-Tong), pp. 33–51. Berlin, Germany: Springer.

25. Kalisz S, Randle A, Chaiffetz D, Faigeles M, Butera A, Beight C. 2012 Dichogamy correlates with outcrossing rate and defines the selfing syndrome in the mixed mating genus *Collinsia*. *Ann. Bot.* **109**, 571–582. (doi:10.1093/aob/mcr237)

26. Vallejo-Marín M, Barrett SCH. 2009 Modification of flower architecture during early stages in the evolution of self-fertilization. *Ann. Bot.* **103**, 951–962. (doi:10.1093/aob/mcp015)

27. Foxe JP, Stift M, Tedder A, Haudry A, Wright SI, Mable BK. 2010 Reconstructing origins of loss of self-incompatibility and selfing in North American *Arabidopsis lyrata*: a population genetic context. *Evolution* **64**, 3495–3510. (doi:10.1111/j.1558-5646.2010.01094.x)

28. Sicard A, Stacey N, Hermann K, Dessoly J, Neuffer B, Bäurle I, Lenhard M. 2011 Genetics, evolution, and adaptive significance of the selfing syndrome in the genus *Capsella*. *Plant Cell* **23**, 3156–3171. (doi:10.1105/tpc.111.088237)

29. Busch JW, Joly S, Schoen DJ. 2011 Demographic signatures accompanying the evolution of selfing in *Leavenworthia alabamica*. *Mol. Biol. Evol.* **28**, 1717–1729. (doi:10.1093/molbev/msq352)

30. Glémin S, Ronfort J. 2013 Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. *Evolution* **67**, 225–240. (doi:10.1111/j.1558-5646.2012.01778.x)

31. Fishman LA, Kelly AJ, Willis JH. 2002 Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution* **56**, 2138–2155. (doi:10.1111/j.0014-3820.2002.tb00139.x)

32. Bodbyl Roels SA, Kelly AJ. 2011 Rapid evolution caused by pollinator loss in *Mimulus guttatus*. *Evolution* **65**, 2541–2552. (doi:10.1111/j.1558-5646.2011.01326.x)

33. Slotte T, Hazzouri KM, Stern D, Andolfatto P, Wright SI. 2012 Genetic architecture and adaptive significance of the selfing syndrome in *Capsella*. *Evolution* **66**, 1360–1371. (doi:10.1111/j.1558-5646.2011.01540.x)

34. Barrett SCH, Ness RW, Vallejo-Marín M. 2009 Evolutionary pathways to self-fertilization in a tristylous plant. *New Phytol.* **183**, 546–556. (doi:10.1111/j.1469-8137.2009.02937.x)

35. Barrett SCH, Morgan MT, Husband BC. 1989 The dissolution of a complex genetic polymorphism: the evolution of self-fertilization in tristylous *Eichhornia paniculata* (Pontederiaceae). *Evolution* **43**, 1398–1416. (doi:10.2307/2409456)

36. Ness RW, Wright SI, Barrett SCH. 2010 Mating-system variation, demographic history and patterns of nucleotide diversity in the tristylous plant *Eichhornia paniculata*. *Genetics* **184**, 381–392. (doi:10.1534/genetics.109.110130)

37. Knight TM *et al*. 2005 Pollen limitation of plant reproduction: patterns and processes. *Annu. Rev. Ecol. Evol. Syst.* **36**, 467–497. (doi:10.1146/annurev.ecolsys.36.102403.115320)

38. Busch JW. 2005 The evolution of self-compatibility in geographically peripheral populations of *Leavenworthia alabamica* (Brassicaceae). *Am. J. Bot.* **92**, 1503–1512. (doi:10.3732/ajb.92.9.1503)

39. Moeller DA, Geber MA. 2005 Ecological context of the evolution of self-pollination in *Clarkia xantiana*: population size, plant communities, and reproductive assurance. *Evolution* **59**, 786–799.

40. Pettengill JB, Moeller DA. 2011 Tempo and mode of mating system evolution between incipient *Clarkia* species. *Evolution* **66**, 1210–1225. (doi:10.1111/j.1558-5646.2011.01521.x)

41. Glémin S, Ronfort J. 2013 Correction for Glémin and Ronfort (2012). *Evolution* **67**, 3381. (doi:10.1111/evo.12198)

42. Charlesworth B. 2013 Background selection 20 years on: the Wilhelmine E. Key 2012 invitational lecture. *J. Hered.* **104**, 161–171. (doi:10.1093/jhered/ess136)

43. Ossowski S, Schneeberger K, Lucas-Lledó JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M. 2010 The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. *Science* **327**, 92–94. (doi:10.1126/science.1180677)

44. Haudry A *et al*. 2013 An atlas of over 90,000 conserved noncoding sequences provides insight into crucifer regulatory regions. *Nat. Genet.* **45**, 891–898. (doi:10.1038/ng.2684)

45. Slotte T, Foxe JP, Hazzouri KM, Wright SI. 2010 Genome-wide evidence for efficient positive and purifying selection in *Capsella grandiflora*, a plant species with a large effective population size. *Mol. Biol. Evol.* **27**, 1813–1821. (doi:10.1093/molbev/msq062)

46. Kimura M. 1983 *The neutral theory of molecular evolution*. Cambridge, UK: Cambridge University Press.

47. Hough J, Williamson RJ, Wright SI. 2013 Patterns of selection in plant genomes. *Annu. Rev. Ecol. Evol. Syst.* **44**, 31–49. (doi:10.1146/annurev-ecolsys-110512-135851)

48. Brandvain Y, Slotte T, Hazzouri KM, Wright SI, Coop G. 2013 Genomic identification of founding haplotypes reveals the history of the selfing species *Capsella rubella*. *PLoS Genet.* **9**, e1003754. (doi:10.1371/journal.pgen.1003754)

49. Cutter AD, Wasmuth JD, Washington NL. 2008 Patterns of molecular evolution in *Caenorhabditis* preclude ancient origins of selfing. *Genetics* **178**, 2093–2104. (doi:10.1534/genetics.107.085787)

50. Haudry A, Cenci A, Guilhaumon C, Paux E, Poirier S, Santoni S, David J, Glémin S. 2008 Mating system and recombination affect molecular evolution in four *Triticeae* species. *Genet. Res.* **90**, 97–109. (doi:10.1017/S0016672307009032)

51. Escobar JS, Cenci A, Bolognini J, Haudry A, Laurent S, David J, Glémin S. 2010 An integrative test of the dead-end hypothesis of selfing evolution in *Triticeae* (Poaceae). *Evolution* **64**, 2855–2872.

52. Ness RW, Siol M, Barrett SCH. 2012 Genomic consequences of transitions from cross-to self-fertilization on the efficacy of selection in three independently derived selfing plants. *BMC Genomics* **13**, 611. (doi:10.1186/1471-2164-13-611)

53. Hazzouri KM, Escobar JS, Ness RW, Randle AM, Kalisz S, Wright SI. 2013 Comparative population genomics in *Collinsia* sister species reveals evidence for reduced effective population size, relaxed selection, and evolution of biased gene conversion with an ongoing mating system shift. *Evolution* **67**, 1263–1278.