Hidden genetic variance contributes to increase the short-term adaptive potential of selfing populations

Josselin CLO [§], Joëlle RONFORT [§], Diala ABU AWAD [§,*]

[§] AGAP, Univ Montpellier, CIRAD, INRAE, Institut Agro, 34000 Montpellier, France.

[*] Department of Population Genetics, Technische Universität München, 85354 Freising, Germany.

Running title: Adaptation from standing genetic variation under selfing.

Corresponding author: Josselin CLO, Arthur Young street, 34000 Montpellier, +33432722335, josselin.clo@gmail.com

Acknowledgment: We would like to thank D. Roze for helpful discussions about the analyses and interpretations of the results, and E. Noël, P. David, O. Ronce, S. Glémin and C. Devaux

for insightful discussions on the results. We also thanks editors and reviewers for improving the manuscript. This work has been conducted with the help of the data and calculation center South Green of the CIRAD- UMR AGAP. This work was supported by project SEAD (ANR-13-ADAP-0011), the TUM University Foundation Fellowship and the Alexander von-Humboldt Foundation.

INTRODUCTION

Natural populations typically harbour genetic variation, especially at loci governing polygenic traits (Mittell *et al.*, 2015; Wood *et al.*, 2016; Clo *et al.*, 2019). This variation, known as standing genetic variation, has been considered to be an important predictor for the adaptive potential of populations (Orr & Betancourt, 2001; Hermisson & Pennings, 2005; Barrett & Schluter, 2008; Pritchard *et al.*, 2010; Glémin & Ronfort, 2013; Matuszewski *et al.*, 2015). Indeed, standing variation represents an easily accessible source of genetic variation, that is readily available for adaptation to changing or heterogeneous conditions (Hermisson & Pennings, 2005; Barrett & Schluter, 2008). Compared to adaptation from *de novo* mutations, the probability of adapting from standing variation is higher simply because mutations already segregating in a population are expected to be present at higher frequencies (Innan & Kim, 2004; Barrett & Schluter, 2008). It has also been suggested that populations adapting from standing genetic variation can cope with more severe and more rapid environmental change, as they are able to cross larger distances in the phenotype space (Matuszewski *et al.*, 2015). The amount of standing variation available in a population is thus expected to play a key role in adaptation, and any forces affecting it may greatly influence whether or not populations are able to survive new environments.

The mating system is a population characteristic that is known to greatly affect the amount and structure of genetic variation. For instance, both theoretical models (Charlesworth & Charlesworth, 1995; Lande & Porcher, 2015; Abu Awad & Roze, 2018) and empirical data (Charlesworth & Charlesworth, 1995; Geber & Griffen, 2003; Clo *et al.*, 2019) have shown that, compared to outcrossing populations, self-fertilization reduces, on average, the amount of additive genetic variance for polygenic or quantitative traits under stabilizing selection.

This diminution is due to more efficient purifying selection under selfing and to negative linkage disequilibria maintained between alleles at different loci: the so-called Bulmer effect (Lande & Porcher, 2015; Abu Awad & Roze, 2018). Due to this reduction in genetic variability, predominant selfing has been described as an evolutionary dead-end (Stebbins, 1957; Takebayashi & Morrell, 2001; Igic & Busch, 2013). However, recent theoretical work and empirical data have highlighted that cryptic genetic variability could contribute to the adaptive potential of natural populations. Cryptic genetic variation has been defined as the part of a population's standing genetic variation that has no effect on the phenotypic variation in a stable environment but can increase heritable variation in environmental conditions rarely experienced (Gibson & Dworkin, 2004; Paaby & Rockman, 2014). Such variability has been detected in both outcrossing (in sticklebacks, McGuigan *et al.*, 2011, cavefish, Rohner *et al.*, 2013, dung flies, Berger *et al.*, 2011, gulls, Kim *et al.*, 2013 or spadefoot toads, Ledon-Rettig *et al.*, 2010) and selfing species (*Caenorhabditis elegans*, Milloz *et al.* 2008; *Arabidopsis thaliana*, Queitsch *et al.* 2002). Two main mechanisms could explain the accumulation and the release of such variance: interactions between loci (Badano & Katsanis, 2002; Shao *et al.*, 2008), and phenotypic plasticity (Anderson *et al.*, 2013). In this paper, we focus on interactions between loci maintained at stabilizing selection.

In maintaining the population close to the phenotypic optimum, stabilizing selection disfavors genetic and phenotypic diversity (Lande & Porcher, 2015; Abu Awad & Roze, 2018). However, the structure of the additive variance also strongly depends on the trait mutation rate and the prevalence of pleiotropy (Charlesworth, 1990; Lande & Porcher, 2015; Abu Awad & Roze, 2018). When the per-trait mutation rate is weak, associations between loci are negligible (mutations of strong effect arise slowly and are easily purged), but when the rate increases, the creation and maintenance of co-adapted gene complexes structure the additive variance into positive within-loci components and negative among-loci components

of variance, reducing the observed level of additive variance (Abu Awad & Roze, 2018). The remobilization through recombination of this among-loci component of variance could boost the adaptability of populations undergoing an environmental shift (Le Rouzic and Carlborg 2008). Indeed, if associations between loci are broken, segregating alleles could express some or all of their additive effects in new genetic backgrounds. Such remobilization is only possible if residual allogamy occurs (*i.e.* if the selfing rate is not equal to 1), which is expected to be common in natural plant populations (Kamran-Disfani & Agrawal, 2014; Clo *et al.*, 2019). Classical models analyzing the effect of selfing on adaptation from standing genetic variation have considered a single locus (Glémin & Ronfort, 2013), thus neglecting interactions among loci that could contribute to an important fraction of standing genetic variation. As self-fertilization reduces the effective recombination rate (Nordborg, 2000), allowing the maintenance of co-adapted gene complexes, the storage of genetic diversity through genetic associations should be more prevalent in selfing populations (as suggested in Charlesworth, 1990; Lande & Porcher, 2015; Abu Awad & Roze, 2018).

In this paper, we explore this hypothesis, using a quantitative genetics framework. We describe and quantify how, to what degree, and under which conditions populations accumulate genetic variation at polygenic traits. Because the development of tractable analytical predictions of the phenotype and genetic variance dynamics is not trivial and often requires strong hypotheses to be made, notably random mating and considering the system to be at linkage equilibrium (Barton & Turelli, 1987; Keightley & Hill, 1989; Bürger, 1993), we chose to use a simulation approach. Although it is known that directional dominance plays an important role in the fate of mutations, especially in the presence of self-fertilization, the interaction of both directional dominance and epistasis with mating systems are not well known. In order to avoid further complicating the interpretation of our results, we have considered a fully additive model, which despite its simplicity, has previously led to

surprisingly accurate predictions (Martin *et al.*, 2007; Manna *et al.*, 2011). We show that, in models allowing genetic diversity to be stored through genetic associations and when adaptation is only possible from pre-existing standing genetic variation, predominantly selfing populations can adapt as well as their mixed-mating and outcrossing counterparts, when mutation rate is high (but realistic), despite initially low levels of observable genetic diversity. If the mutation rate is low, purifying selection impedes the storage of diversity through genetic associations, and our simulations confirm the more classical results that selfing populations adapt less well than their outcrossing counterparts do.

MATERIAL AND METHODS

General assumptions

We consider the evolution of a quantitative trait Z in a population of size N, made of diploid individuals reproducing through partial self-fertilization, with a constant selfing rate, (and hence constant inbreeding level) σ. The phenotypic value z of an individual is determined by the additive action of L loci each with an infinite possible number of alleles and is given by

$$z = g + e, \quad (1)$$

where g is the genetic component of the individual's phenotype, and is given by $g = \sum_j^L (g_j^M + g_j^P)$, with $g^M{}_j$ (respectively $g^P{}_j$) the additive allelic effect at locus j inherited from the maternal (respectively paternal) gamete. There are no dominance or epistasis at the phenotypic scale in this model, but both of which arise naturally at fitness scale, when considering stabilizing

selection and a mean phenotype close to the optimum, the mean dominance h ≈ 0.25 and epistasis effects are on average null for fitness. The random environmental effect, *e*, is drawn from a Gaussian distribution of mean 0 and variance V_E, and is considered to be independent from the genetic components of fitness. The trait initially undergoes stabilizing selection around an optimal phenotypic value, denoted Z_{opt}. The fitness value W_Z of an individual with phenotype *z* is thus described by the Gaussian function:

$$W_Z = e^{-d^2/2\omega^2}, (2)$$

where *d* is the distance between the individual's phenotype *z* and the optimum trait value and ω^2 is the width of the fitness function, and represents the strength of selection.

Simulation model

We implement the model described above into an individual based simulation model written in C++, a modified version of the "continuum of alleles" program provided in Abu Awad and Roze (2018) available in supplementary materials and online (https://github.com/dialaAbAw/SelfingAdaptation).

We consider a population of *N* diploid individuals, each represented by two linear chromosomes with *L* multi-allelic loci, coding for a single quantitative trait under selection. At the beginning of each simulation, all individuals are genetically identical and are at the phenotypic optimum (all loci carry alleles with effect 0 and $Z_{OPT} = 0$). The life cycle can be summarized by five successive events. First, there is a phenotype-dependent choice of the first parent (selection), followed by mating-type choice (selfing versus outcrossing at rates σ and (1- σ) respectively) and then, in the case of outcrossing, phenotype-dependent choice of

the second parent. Selection takes place as follows: if the ratio of the selected parent's fitness over the highest recorded fitness value in the current generation is higher than a number sampled in a uniform law comprised between 0 and 1, the individual is allowed to reproduce. Once the two parents are chosen, they each contribute a gamete, produced through uniformly recombining the parental chromosomes. The number of cross-overs is sampled from a Poisson distribution with parameter R, the map length. From Haldane's mapping function, the recombination rate between two adjacent loci is $r = \frac{1}{2}\left[1 - exp\left(\frac{-2R}{L-1}\right)\right]$. We choose parameters that ensure that $r \approx 0.5$, such that loci remain, at least physically, unlinked. This phase is then followed by the introduction of new mutations, the number of which is sampled from a Poisson distribution with parameter U (with $U = \mu L$, μ being the per locus mutation rate). The additive value of a new mutant allele is drawn from a Normal distribution of mean 0 and variance a^2.

Each simulation consists of two phases, the first being burn-in time to allow the population to attain Mutation-Selection-Drift equilibrium (M-S-D). The population is considered to be at M-S-D equilibrium when the average fitness value calculated over the last thousand generations does not differ by more than one percent from the mean fitness calculated over the previous thousand generations. The second phase consists of following the population after a shift in the phenotypic optimum. After the shift, the haploid genomic mutation rate U is set to 0 so that the only source of genetic variability to reach the new optimum is the standing variation accumulated before the environmental change.

Simulation parameter values:

Simulations were run for several parameter sets in order to evaluate the effects of each parameter on both the equilibrium conditions and on the ability of populations to adapt after

the environmental change. The values chosen for the mutation rate U range from 0.005 to 0.1, reflecting the *per*-trait haploid genomic mutation rate found in the literature (Keightley & Bataillon, 2000; Shaw *et al.*, 2002; Haag-Liautard *et al.*, 2007). We use parameter set values similar to those explored in Bürger *et al.*, (1989) and Ronce *et al.*, (2009), with the number of freely recombining loci under selection L = 50, and a^2 = 0.05, V_E =1, ω^2 = 1, 9 or 99. In this model, the mean deleterious effect of mutations $\bar{s}$ is therefore equal to 0.0125, 0.0025 or 0.00025 ($\bar{s} = a^2 / 2(\omega^2 + V_E)$, Martin and Lenormand 2006).

Before the shift in the optimum, Z_{OPT} is set to 0, and after M-S-D equilibrium it is set to 2.5, so that the shift is of order $L.a^2$, and ensure that adaptation is due to genetic changes and not environmental effects (Ronce *et al.*, 2009). We follow populations only over 20 generations after the shift in the optimum, as we chose to follow the possibility of adaptation without any contribution from *de novo* mutations. After this time limit, genetic variation starts to erode, and the observed results are mainly due to the effects of loss of diversity due to drift and not to the selection process. Although simulations were run over a large range of selfing rate values, throughout the manuscript we show results run principally for three rates of self-fertilization, σ = 0, 0.5 and 0.95, representing outcrossing, mixed-mating and predominantly selfing respectively. These three values were chosen because they were representative of the main patterns of adaptation we observed over the whole range of selfing rates (σ from 0 to 1). We also considered three population sizes N = 250, 1000 or 10000, which allow to test the effect of drift.

Analysis of the effect of selfing on the equilibrium genetic variance

Following Turelli & Barton (1990), we decompose the genetic variance of a polygenic trait using the following equation:

$$V_A = 2 \sum_j^L (C_{jj} + C_{j,j}) + 2 \sum_{j \neq k}^L (C_{jk} + C_{j,k}), \quad (3)$$

with

$$C_{jj} = \frac{1}{2} E[\,(g^M{}_j - \bar{g}_j)^2 + (g^P{}_j - \bar{g}_j)^2] \quad (4)$$

and

$$C_{j,j} = E[\,(g^M{}_j - \bar{g}_j) \,.\, (g^P{}_j - \bar{g}_j)], \quad (5)$$

where $\bar{g}_j$ is the mean allelic effect on the phenotype at locus j and $g^M{}_j$ (respectively $g^P{}_j$) is the allelic effect at locus j inherited from the maternal (respectively paternal) gamete. The sum of of C_{jj} over loci represents the genic variance (noted V_{genic}, which is the genetic variance of a trait in a population harboring the same allelic frequencies as the population under study, but without any genetic associations between loci), and is computed from simulation outputs following equation (4). The sum of all values of $C_{j,j}$ represents the covariance in allelic effects at locus j on the maternally and paternally inherited chromosomes, and represents the fraction of the genetic variance due to excess of homozygosity (noted V_{inbred}); we compute it following equation (5). This quantity represents $F.V_{genic}$, where F is the inbreeding coefficient of the population. The first term of equation (3) ($2 \sum_j^L (C_{jj} + C_{j,j})$) represents the genetic variance due to within locus variation. The second term ($2 \sum_{j \neq k}^L (C_{jk} + C_{j,k})$) represents the component of the variance due to associations between loci (noted COV_{LD}), and is obtained by computing the sum of covariances between genotypic values among all pair of loci. This component is proportional to linkage disequilibrium (LD), and tends to be negative under stabilizing selection due to associations between alleles at different loci with compensatory effects (*i.e.* the allele on one locus is positive, the other negative, their effects on the phenotype thus

cancel out when both are present). At any time in the simulated populations, $V_A = V_{genic} + V_{inbred} + COV_{LD}$.

As a check, estimates of additive genetic variance observed at M-S-D equilibrium in our simulations were compared to analytical predictions obtained by Bürger et al. (1989) (Stochastic House of Cards approximate) and by Abu Awad and Roze (2018) under similar assumptions. Comparisons and details are shown in Appendix A.

Analyses of the response to an environmental change

To analyse the population's response to an environmental change, we focus on the 20 first generations after the environmental change. We follow the temporal dynamics of the additive genetic variance of the trait and of its components. We also analyze changes with time of the mean trait value and population fitness, as well as the dynamics of the number of haplotypes and the proportion of new haplotypes generated after the environmental change.

In addition, and in order to determine if the remobilization of COV_{LD} plays a role in the adaptive process of selfing populations, we compute for each mutation rate, the slope of the trait mean variation shortly after the environmental change (during the first five generations) as a function of the amount of additive variance available at M-S-D equilibrium. If remobilization of COV_{LD} is involved in the adaptive process of selfing populations, the initial slope, for an initially similar amount of additive variance, should be higher in selfing populations than in mixed mating and outcrossing ones.

RESULTS

LEVEL AND STRUCTURE OF ADDITIVE GENETIC VARIANCE AT MUTATION SELECTION DRIFT EQUILIBRIUM

In agreement with expectations from previous works, the additive genetic variance present at M-S-D equilibrium in our simulations is overall negatively correlated with the selfing rate (Figure 1), in part due to the higher efficiency of selfing in purging deleterious mutations. In general, reducing the strength of stabilizing selection increases the overall amount of additive variance, as mutations of small effect are not as easily purged from the population (Figure 1). The population size, however, has little qualitative effect on the relationship between selfing and V_A, at least as long as the mutation rate is not too high (Figure S1). As detailed in Appendix A, the values of V_A obtained from our simulations were compared to expectations from analytical approximations from Abu Awad & Roze (2018) and the Stochastic House of Cards approximation (SHC, Bürger *et al.*, 1989). We generally find a good agreement between our results and the SHC approximation for the smaller population sizes and outcrossing populations (σ = 0 or 0.5). Approximations from Abu Awad & Roze (2018) only hold for predominantly selfing populations, considering the whole genome as a single locus (σ = 0.95, equation 44 in Abu Awad & Roze, 2018), or for all mating systems when population size is large (N=10.000) and the strength of selection weak (ω^2=99, equation D23). The lack of fit between our results and approximations from Abu Awad & Roze (2018) is probably due to the small number of loci considered in our simulations (n = 50 in our parameter set) and to the fact that we consider only a single trait, both parameters violate the assumptions for these analytical approximations to hold.

By decomposing the observed additive variance V_A into the components described in the Methods above (V_{genic}, V_{inbred}, COV_{LD}), we find that for the range of parameters we chose to explore, the strong decrease in V_A observed under predominant selfing is mostly driven by a large but negative COV_{LD} (Figure 2). This confirms the expectation that recurrent selfing favours the maintenance of associations among loci. Under such associations the efficiency of purifying selection is reduced, as the selection pressure on a single mutation is weakened by its positive association with other mutations at other loci, which in turn contributes to higher V_{genic} and V_{inbred} in selfers compared to outcrossers, as observed in Figure 2. Our analysis also showed that the higher the mutation rate, the more genetic associations are maintained and the weaker the efficiency of purifying selection. The increase of both V_{inbred} and V_{genic} can, in some cases, overcome the negative effect of COV_{LD}, as is the case for large populations (N=10.000), leading to selfing populations that have close to, and sometimes even greater amounts of genetic variance than outcrossing or mixed mating populations (Figure S2). As can be expected from classical population genetics models, smaller population size leads to a decrease in V_A, whereas higher mutation rates tend to increase it (Figure S1).

PATTERNS OF ADAPTATION THROUGH STANDING GENETIC VARIATION:

After populations have reached mutation-selection-drift balance, an environmental change is induced and no new mutations are introduced. Population dynamics are followed over 20 generations and, if during this time the population reaches a similar level of fitness as that observed before the environmental change, then we consider that the population is able to adapt to the new optimum. When the strength of purifying selection is moderate to weak (ω^2 = 9 or 99), all populations seem to harbour enough genetic variation to quickly and efficiently

respond to an environmental change, independently of their selfing rate. As there were no observable dynamic changes in V_A and its components for these parameter values over the course of the first 20 generations following the environmental change, the results for these simulations are only presented in Supplementary Figures S3 to S8. In the following sections we will only be concentrating on the case of $\omega^2 = 1$, representing the strongest selection pressure we explored for.

DYNAMICS OF PHENOTYPE AND FITNESS DURING ADAPTATION UNDER STRONG SELECTION:

As illustrated in Figure 3, when the mutation rate is small ($U = 0.005$), and hence the amount of diversity stored through genetic associations for selfing population is also small, none of the populations are able to fully adapt during the 20 first generations, and population fitness remains low. In this case the amount of within-loci genetic variation at equilibrium is the only source of diversity, and as outcrossing and mixed mating populations have higher V_A, they reach phenotypic values closer to the new optimum, and have a significantly higher fitness than their selfing counterparts (Figure 3 for U=0.005). For higher mutation rates (U = 0.1), dynamics of the phenotypic trait and fitness are extremely similar for all mating systems, despite lower genetic diversity in selfing populations. As shown in the following sections, when the mutation rate increases, the variance stored in selfing populations through genetic associations is partially released, increasing the adaptive potential of these populations. In such cases, selfing populations are able to reach (1) the new phenotypic optimum, and (2) level of fitness similar to those observed at M-S-D equilibrium, thus performing as well as outcrossing populations (Figure 3).

The above-mentioned patterns are similar irrespective of the population size. As mentioned in the first section of the results, the greater the population size the greater the genetic diversity (V_A and COV_{LD} for selfing species) for a given mutation rate, increasing the range of mutation rates for which selfing populations are at least as well, if not better, adapted as outcrossing and mixed mating populations (Figures S9 and S10). This means that population size has a quantitative but not a qualitative effect.

DYNAMICS OF THE ADDITIVE VARIANCE AND OF ITS COMPONENTS

As the population undergoes directional selection, the dynamics of the observable additive variance and each of its components are telling of the processes driving adaptation to the new optimum. In small outcrossing populations (N = 250), the additive variance exhibits very small changes during the adaptation process, as do its components, whatever the strength of selection (Figures S9 to S11). For larger population sizes (N=1000 or 10.000), the additive variance increases during the first generations when the mutation rate is small to medium ($U \leq 0.05$, Figure 4A for N=1000, see supplementary materials for N=10.000). This increase of V_A stems from an increase in V_{genic}, indicative of the change in the phenotype being driven by an increased frequency of alleles that may have been neutral or deleterious before the environmental change (Bürger 1993). For high mutation rates (U = 0.1), as for small populations, the additive genetic variance and its components remain constant through time (Figure 4). V_{inbred} does not vary (Figure 4C). Overall patterns are similar for mixed mating populations.

For predominantly selfing populations, the initial dynamics strongly depend on “how much” diversity the population has stored at M-S-D equilibrium (represented by COV_{LD}). In

all cases, additive genetic variance increases during the adaptation process (Figure 5A). When genetic associations are rare (Figure 5, U = 0.005), adaptation is driven solely through the increase in frequency of beneficial alleles, leading to an increase of the genic variance and inbred variance (Figures 5B and C). In this case, COV_{LD} slightly decreases during the first generations (Figure 5D), probably due to the formation of linkage disequilibrium between alleles rising up in frequency. When among-loci associations are more substantial (U > 0.005), the dynamics observed are qualitatively different, and reflect another process of adaptation. The increase in additive variance is due to the "release" of genetic diversity (COV_{LD} increases, figure 5D) over the first five generations, whereby the genic and inbred variances remain constant or decrease (Figure 5B and C). The variation of V_{inbred} remains equal to $F.V_{\mathrm{genic}}$, indicating that, as observed for outcrossing populations, the inbreeding coefficient does not vary during adaptation.

The initial response to selection (increase in trait mean over the first five generations) tightly correlates with the initial amount of additive variance (Figure 6). Nevertheless, for a theoretical equal amount of additive variance at equilibrium (over all parameter sets combined), selfing populations have a stronger initial response at the phenotypic level compared to more allogamous ones (Figure 6). This is due the remobilization of the diversity stored by genetic associations during the first few generations, increasing the additive variance and thus the initial response to selection in selfing populations. The larger the population size, the more the pattern observed for selfing populations differs from that observed in outcrossing and mixed mating ones (Figure 6). When selection is moderate to weak, the release of the stored genetic variation is not necessary for selfing populations to adapt, thus this increase in the strength of adaptation for selfers is no longer observed.

DISCUSSION

In accordance with Stebbins' definition of the dead-end hypothesis (Stebbins, 1957), single-locus models predict that adaptation is less likely in selfing populations compared to outcrossing ones, notably due to the reduced standing genetic variation resulting from purging (Glémin & Ronfort, 2013). Considering a polygenic trait, and among loci associations, we find that this is not always the case. Previous works have highlighted that stabilizing selection favours the build-up of associations between loci and, through this, favours the storage of genetic diversity especially in self-fertilising populations (Charlesworth, 1990; Lande & Porcher, 2015; Abu Awad & Roze, 2018). If the selective regime were to change from stabilising to directional selection, due to a strong environmental effect for example, our simulations show that some of this variance can be released through residual allogamy. Residual allogamy is primordial, as selfing populations are organized in multi-locus genotypes. The release of the diversity stored through genetic associations is only possible through rare outcrossing events between two lineages, resulting in fully heterozygote hybrids, which can potentially generate D^3 new genotypes (D being the number of differentiated loci among two selfing lineages) after one generation of selfing. Although the strength of the response to selection is positively correlated with the amount of observable standing genetic variation, it is possible to observe similar levels of adaptation in selfing compared to outcrossing or mixed mating populations, for high but realistic mutation rates (Shaw *et al.*, 2002). When the strength of selection is moderate to weak ($\omega^2 = 9$ or 99), populations seem to harbour enough genetic variation to quickly and efficiently respond to an environmental change, independently of the self-fertilization rate, and thus without remobilizing any variance accumulated through genetic associations.

Genetic variance dynamics under directional selection: the role of genetic associations

Few theoretical studies have tried to predict the dynamic of the additive variance for a trait under directional selection (Barton & Turelli, 1987; Keightley & Hill 1989; Bürger, 1993). From these works, under the assumption of random mating and no association between loci, it is predicted that V_A remains unchanged in the short-term, if population size is small ($N < 500$, Bürger, 1993) and/or if the inverse of the strength of selection, given by the term ω^2, is large compared to the amount of available genetic variance (Barton & Turelli, 1987). This prediction is supported by our results for obligate outcrossing populations (Figures S11 to S15), even when some negative linkage disequilibrium contributes significantly to additive variance (Figures 2 & 4). For selfing populations, the hypothesis of linkage equilibrium is no longer valid. As shown in our analysis and known from previous results (Lande & Porcher, 2015; Abu Awad & Roze, 2018), genetic associations between loci are significant in these populations, notably when the mutation rate is high. In this case, the breakup of genetic associations is necessary for adaptation when most of the genetic variance is hidden in among-loci associations, leading to an increase of observed additive variance. The mutation rates for which such associations contribute significantly to the genetic additive variance are in the range of mutation rates observed for phenotypic traits in the literature (Keightley & Bataillon, 2000; Shaw *et al.*, 2002; Haag-Liautard *et al.*, 2007), which suggests that these associations are probably common in natural populations. The more frequent observations of transgressive segregation (progeny of a cross being outside the parental phenotypic range) in inbred compared to outbred species also give support to this mechanism (Rieseberg *et al.*, 1999; Johansen-Morris & Latta, 2006).

***De novo* mutations *vs.* standing genetic variation: rethinking adaptation in selfing species?**

It has been a long accepted paradigm that the advantage procured by selfing is the more rapid fixation of *de novo* beneficial mutations, independently of their dominance coefficient, compared to outcrossing populations, where recessive beneficial mutations can be lost through drift before selection drives them to fixation, a process known as "Haldane's sieve" (Haldane, 1927). From single locus theory, it is expected that adaptation through new mutations should be more likely in selfing species, and should be more likely than adaptation from standing genetic variation (Glémin & Ronfort, 2013). However, recent works have suggested that the reduced effective recombination rate in selfing populations adds a disadvantage even when it comes to the fixation of new mutations. Unlike what is expected in outcrossing populations, the fixation of beneficial mutations in selfing populations can be hindered if they appear during a selective sweep triggered by a beneficial allele at another locus (Hartfield & Glémin, 2016). This observation, as well as the results presented here, shows that taking interactions between loci into account can strongly modify the expectations from single-locus models. The mutation rate seems to be the major parameter to consider when studying the dynamics of adaptation from a polygenic point of view, independently of whether adaptation occurs from *de novo* mutations (Hartfield & Glémin, 2016) or standing genetic variation.

In our simulations, we have considered a quantitative trait with a simple architecture, assuming pure additivity of genetic effects on the phenotype. Epistasis, and notably its directionality, is known to play a key role in adaptation (Hansen, 2013). Positive epistasis

(several loci reinforcing each other's effects in the direction of selection), inflates the additive variance and thus the ability of a population to adapt to an environmental change (Carter *et al.*, 2005; Monnahan & Kelly, 2015). On the contrary, negative epistasis (loci suppressing effects at other loci), reduces the additive variance of the character, thus limiting adaptive potential (Carter *et al.*, 2005). Few empirical estimations of the directionality of epistasis are available in the literature (Le Rouzic 2014; Monnahan and Kelly 2015; Oakley *et al.* 2015, all detecting positive epistatic interactions), despite numerous methods and the diversity of data used to infer it (Le Rouzic, 2014). How the directionality of epistatic interactions varies in relation to the mating system remains unknown in natural populations. Knowledge on this relationship may bring us closer to the understanding of the differences in patterns of adaptation observed between selfing and outcrossing populations.

New insights into the role of standing genetic variation in the adaptation dynamics of selfing populations

The overwhelming success of selfing species in the domestication process and as invasive species has been attributed to mechanisms other than their adaptive ability from standing genetic variation. For instance, the invasive success of selfing populations is attributed to their reproductive assurance, since a single individual is able to colonize a new environment (Rambuda & Johnson, 2004; van Kleunen *et al.*, 2008), and to reduced gene flow, which is expected to limit maladapted gene exchanges between populations (Levin, 2010). Regarding domestication as an adaptation process, it has been suggested that domestication in selfing populations most probably relied on new mutations, due to the initially low genetic variance that would have been further reduced due to the bottleneck

effect of domestication (Glémin & Bataillon, 2009). This idea is reinforced by the fact that selfing species are expected to quickly fix a rare beneficial mutation, independently of its dominance level (Ross-Ibarra, 2005). In their review on mating system variation in domesticated plant species, Glémin and Bataillon (2009) have suggested that the high frequency of self-fertilizing crop species could be related to an increase in the amount of additive variance during the domestication process of selection. This idea has, however, never been tested theoretically or empirically. Here we show that this increase in additive variance could indeed be an advantage when selfing species are faced with strong directional selection and selection occurs through standing genetic variation. However, our results hold true only if the bottleneck related to domestication (or other invasion processes) is not too strong and if mutation rates are high enough to maintain enough diversity that can promote adaptation to new conditions.

CONCLUSION AND PERSPECTIVES

In this work, we show that under stabilizing selection and if mutation rates are high enough (but realistic), selfing populations are able to accumulate genetic variation through negative linkage disequilibrium. This structure of genetic variation contributes to acting as a store of genetic variability, thanks to which, adaptation under high self-fertilisation rates is no longer constrained by the reduced observable additive genetic variance due to purging. For low mutation rates, the storage of diversity is impeded by stabilizing selection in selfing populations, leading to the more classical conclusion that adaptation in selfing species is limited by the available amount of genetic diversity. Our analysis thus shows that measuring the amount of additive variance available for quantitative traits is not always sufficient to

make predictions about the adaptive potential of a population. Complementary analyses should be carried out when quantifying the short-term adaptability of a population. Such analyses could include looking for transgressive segregations or carrying out experimental evolutionary experiments in which directional selection is induced, and follow the dynamic of additive genetic variance. Following the components of additive variance is complicated, as it necessitates thorough knowledge of the genetic architecture of the quantitative trait under study (number of loci underlying phenotypic variation, their frequency, the linkage disequilibrium…), which is complicated, even for model species. More empirical evidence is required to determine how frequently diversity is stored through genetic associations in natural populations of selfing species, and whether this property is sufficient to allow selfing species to adapt to a changing environment.

Bibliography:

Abu Awad, D. & Roze, D. 2018. Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. *Evolution* **72**: 751–769.

Anderson, J.T., LEE, C.-R., Rushworth, C.A., Colautti, R.I. & MITCHELL-OLDS, T. 2013. Genetic trade-offs and conditional neutrality contribute to local adaptation. *Mol. Ecol.* **22**: 699–708.

Badano, J.L. & Katsanis, N. 2002. Human genetics and disease: Beyond Mendel: an evolving view of human genetic disease transmission. *Nat. Rev. Genet.* **3**: 779.

Barrett, R.D. & Schluter, D. 2008. Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**: 38–44.

Barton, N.H. & Turelli, M. 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. *Genet. Res.* **49**: 157–173.

Berger, D., Bauerfeind, S.S., Blanckenhorn, W.U. & Schäfer, M.A. 2011. High temperatures reveal cryptic genetic variation in a polymorphic female sperm storage organ. *Evol. Int. J. Org. Evol.* **65**: 2830–2842.

Bürger, R. 1993. Predictions of the dynamics of a polygenic character under directional selection. *J. Theor. Biol.* **162**: 487–513.

Bürger, R., Wagner, G.P. & Stettinger, F. 1989. HOW MUCH HERITABLE VARIATION CAN BE MAINTAINED IN FINITE POPULATIONS BY MUTATION–SELECTION BALANCE? *Evolution* **43**: 1748–1766.

Carter, A.J., Hermisson, J. & Hansen, T.F. 2005. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theor. Popul. Biol.* **68**: 179–196.

Charlesworth, B. 1990. Mutation-selection balance and the evolutionary advantage of sex and recombination. *Genet. Res.* **55**: 199–221.

Charlesworth, D. & Charlesworth, B. 1995. Quantitative genetics in plants: the effect of the breeding system on genetic variability. *Evolution* **49**: 911–920.

Clo, J., Gay, L. & Ronfort, J. 2019. How does selfing affect the genetic variance of quantitative traits? An updated meta-analysis on empirical results in angiosperm species. *Evolution*.

Geber, M.A. & Griffen, L.R. 2003. Inheritance and natural selection on functional traits. *Int. J. Plant Sci.* **164**: S21–S42.

Gibson, G. & Dworkin, I. 2004. Uncovering cryptic genetic variation. *Nat. Rev. Genet.* **5**: 681.

Glémin, S. & Bataillon, T. 2009. A comparative view of the evolution of grasses under domestication. *New Phytol.* **183**: 273–290.

Glémin, S. & Ronfort, J. 2013. Adaptation and maladaptation in selfing and outcrossing species: new mutations versus standing variation. *Evol. Int. J. Org. Evol.* **67**: 225–240.

Haag-Liautard, C., Dorris, M., Maside, X., Macaskill, S., Halligan, D.L., Charlesworth, B., *et al.* 2007. Direct estimation of per nucleotide and genomic deleterious mutation rates in Drosophila. *Nature* **445**: 82.

Haldane, J.B.S. 1927. A mathematical theory of Natural and Artificial Selection. *Math. Proc. Camb. Philos. Soc.* **23**: 607–615.

Hansen, T.F. 2013. WHY EPISTASIS IS IMPORTANT FOR SELECTION AND ADAPTATION. *Evolution* **67**: 3501–3511.

Hartfield, M. & Glémin, S. 2016. Limits to Adaptation in Partially Selfing Species. *Genetics* **203**: 959.

Hermisson, J. & Pennings, P.S. 2005. Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**: 2335–2352.

Igic, B. & Busch, J.W. 2013. Is self-fertilization an evolutionary dead end? *New Phytol.* **198**: 386–397.

Innan, H. & Kim, Y. 2004. Pattern of polymorphism after strong artificial selection in a domestication event. *Proc. Natl. Acad. Sci.* **101**: 10667–10672.

Johansen-Morris, A.D. & Latta, R.G. 2006. Fitness consequences of hybridization between ecotypes of Avena barbata: hybrid breakdown, hybrid vigor, and transgressive segregation. *Evolution* **60**: 1585–1595.

Kamran-Disfani, A. & Agrawal, A.F. 2014. Selfing, adaptation and background selection in finite populations. *J. Evol. Biol.* **27**: 1360–1371.

Keightley, P.D. & Bataillon, T.M. 2000. Multigeneration Maximum-Likelihood Analysis Applied to Mutation-Accumulation Experiments in Caenorhabditis elegans. *Genetics* **154**: 1193.

Keightley, P.D. & Hill, W.G. 1989. Quantitative genetic variability maintained by mutation-stabilizing selection balance: sampling variation and response to subsequent directional selection. *Genet. Res.* **54**: 45–58.

Kim, S.-Y., Noguera, J.C., Tato, A. & Velando, A. 2013. Vitamins, stress and growth: the availability of antioxidants in early life influences the expression of cryptic genetic variation. *J. Evol. Biol.* **26**: 1341–1352.

Lande, R. & Porcher, E. 2015. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics* **200**: 891–906.

Le Rouzic, A. 2014. Estimating directional epistasis. *Front. Genet.* **5**: 198.

Le Rouzic, A. & Carlborg, Ö. 2008. Evolutionary potential of hidden genetic variation. *Trends Ecol. Evol.* **23**: 33–37.

Ledon-Rettig, C.C., Pfennig, D.W. & Crespi, E.J. 2010. Diet and hormonal manipulation reveal cryptic genetic variation: implications for the evolution of novel feeding strategies. *Proc. R. Soc. B Biol. Sci.* **277**: 3569–3578.

Levin, D.A. 2010. Environment-enhanced self-fertilization: implications for niche shifts in adjacent populations. *J. Ecol.* **98**: 1276–1283.

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

Martin, G., Elena, S.F. & Lenormand, T. 2007. Distributions of epistasis in microbes fit predictions from a fitness landscape model. *Nat. Genet.* **39**: 555–560.

Martin, G. & Lenormand, T. 2006. A GENERAL MULTIVARIATE EXTENSION OF FISHER'S GEOMETRICAL MODEL AND THE DISTRIBUTION OF MUTATION FITNESS EFFECTS ACROSS SPECIES. *Evolution* **60**: 893–907.

Matuszewski, S., Hermisson, J. & Kopp, M. 2015. Catch me if you can: adaptation from standing genetic variation to a moving phenotypic optimum. *Genetics* **200**: 1255–1274.

McGuigan, K., Nishimura, N., Currey, M., Hurwit, D. & Cresko, W.A. 2011. Cryptic genetic variation and body size evolution in threespine stickleback. *Evol. Int. J. Org. Evol.* **65**: 1203–1211.

Milloz, J., Duveau, F., Nuez, I. & Félix, M.-A. 2008. Intraspecific evolution of the intercellular signaling network underlying a robust developmental system. *Genes Dev.* **22**: 3064–3075.

Mittell, E.A., Nakagawa, S. & Hadfield, J.D. 2015. Are molecular markers useful predictors of adaptive potential? *Ecol. Lett.* **18**: 772–778.

Monnahan, P.J. & Kelly, J.K. 2015. Epistasis is a major determinant of the additive genetic variance in Mimulus guttatus. *PLoS Genet.* **11**: e1005201.

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

Oakley, C.G., Ågren, J. & Schemske, D.W. 2015. Heterosis and outbreeding depression in crosses between natural populations of Arabidopsis thaliana. *Heredity* **115**: 73.

Orr, H.A. & Betancourt, A.J. 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics* **157**: 875–884.

Paaby, A.B. & Rockman, M.V. 2014. Cryptic genetic variation: evolution's hidden substrate. *Nat. Rev. Genet.* **15**: 247.

Pritchard, J.K., Pickrell, J.K. & Coop, G. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Curr. Biol.* **20**: R208–R215.

Queitsch, C., Sangster, T.A. & Lindquist, S. 2002. Hsp90 as a capacitor of phenotypic variation. *Nature* **417**: 618.

Rambuda, T.D. & Johnson, S.D. 2004. Breeding systems of invasive alien plants in South Africa: does Baker's rule apply? *Divers. Distrib.* **10**: 409–416.

Rieseberg, L.H., Archer, M.A. & Wayne, R.K. 1999. Transgressive segregation, adaptation and speciation. *Heredity* **83**: 363–372.

Rohner, N., Jarosz, D.F., Kowalko, J.E., Yoshizawa, M., Jeffery, W.R., Borowsky, R.L., *et al.* 2013. Cryptic variation in morphological evolution: HSP90 as a capacitor for loss of eyes in cavefish. *Science* **342**: 1372–1375.

Ronce, O., Shaw, F.H., Rousset, F. & Shaw, R.G. 2009. IS INBREEDING DEPRESSION LOWER IN MALADAPTED POPULATIONS? A QUANTITATIVE GENETICS MODEL. *Evolution* **63**: 1807–1819.

Ross-Ibarra, J. 2005. Quantitative trait loci and the study of plant domestication. In: *Genetics of Adaptation* (R. Mauricio, ed), pp. 197–204. Springer Netherlands, Dordrecht.

Shao, H., Burrage, L.C., Sinasac, D.S., Hill, A.E., Ernest, S.R., O'Brien, W., *et al.* 2008. Genetic architecture of complex traits: large phenotypic effects and pervasive epistasis. *Proc. Natl. Acad. Sci.* **105**: 19910–19914.

Shaw, F.H., Geyer, C.J. & Shaw, R.G. 2002. A COMPREHENSIVE MODEL OF MUTATIONS AFFECTING FITNESS AND INFERENCES FOR ARABIDOPSIS THALIANA. *Evolution* **56**: 453–463.

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

Takebayashi, N. & Morrell, P.L. 2001. Is self-fertilization an evolutionary dead end? Revisiting an old hypothesis with genetic theories and a macroevolutionary approach. *Am. J. Bot.* **88**: 1143–1150.

Turelli, M. & Barton, N.H. 1990. Dynamics of polygenic characters under selection. *Theor. Popul. Biol.* **38**: 1–57.

van Kleunen, M., Manning, J.C., Pasqualetto, V. & Johnson, S.D. 2008. Phylogenetically Independent Associations between Autonomous Self-Fertilization and Plant Invasiveness. *Am. Nat.* **171**: 195–201.

Waddington, C.H. 1953. Genetic assimilation of an acquired character. *Evolution* **7**: 118–126.

Wood, J.L., Yates, M.C. & Fraser, D.J. 2016. Are heritability and selection related to population size in nature? Meta-analysis and conservation implications. *Evol. Appl.* **9**: 640–657.

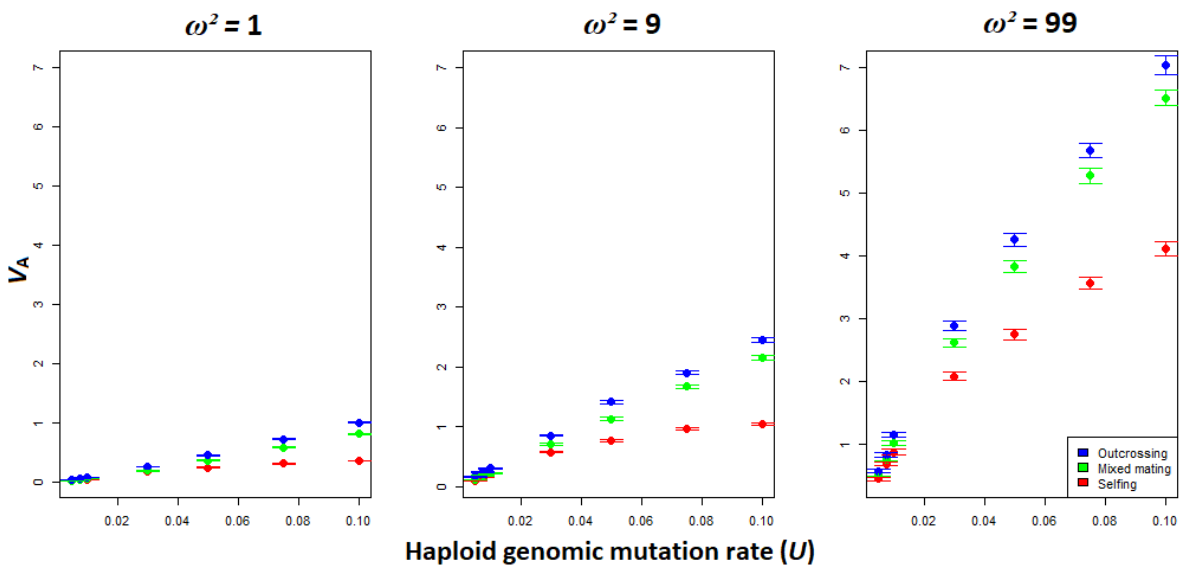

Figure 1. Additive genetic variance observed at mutation-selection-drift equilibrium in the simulated populations, for different mating systems, N=1000, and different strengths of selection (from ω^2=1 (strong selection) to ω^2=100 (weak selection), from the left to the right). Error bars stand for the 95% confidence interval (n = 100 simulations).

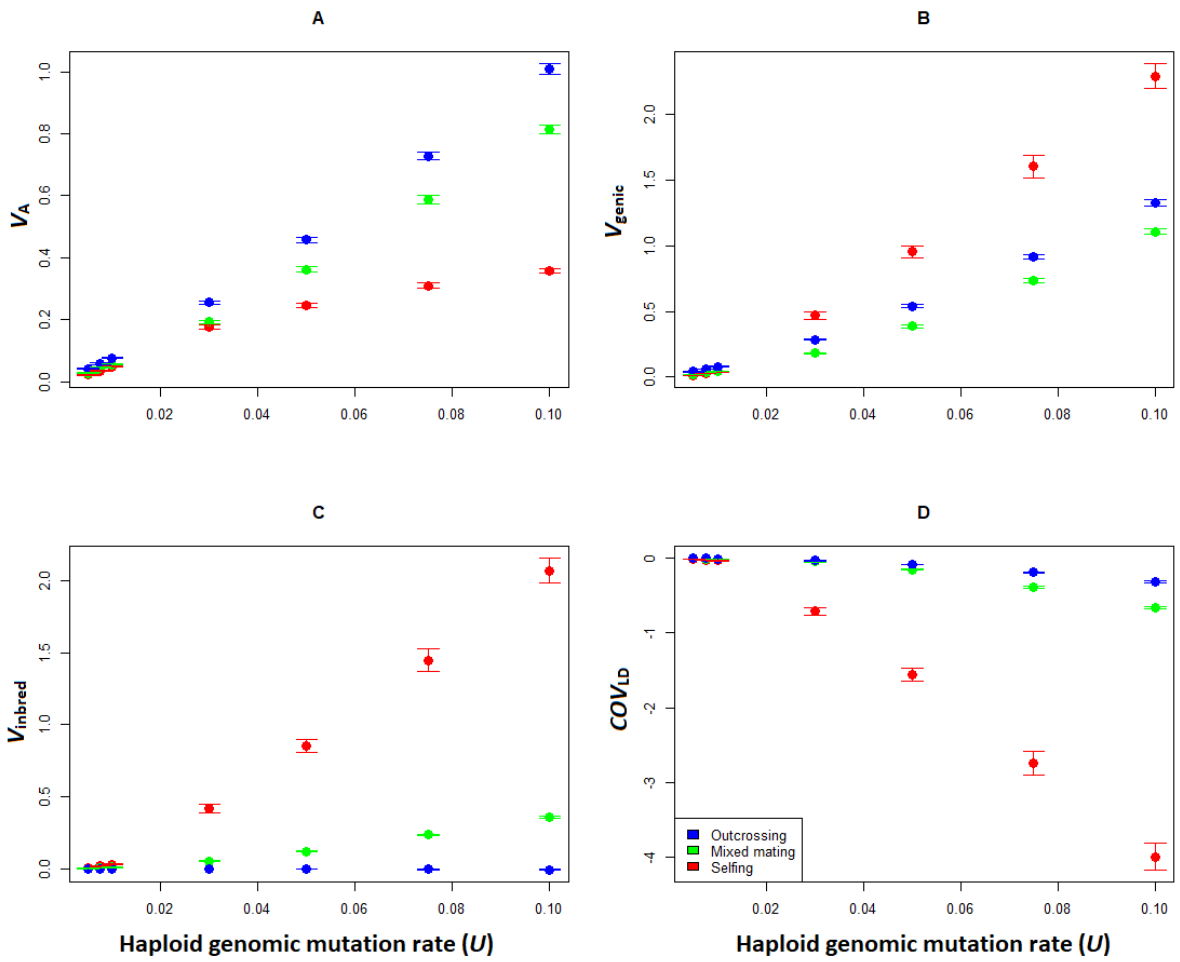

Figure 2. Total additive genetic variance and its three components estimated at MSD equilibrium, for different genomic mutation rates and mating systems, for N=1000 and ω^2=1. **A.** Total additive variance of the phenotypic trait, identical to Figure 1 first panel. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}). Error bars stand for 95% confidence interval (n = 100).

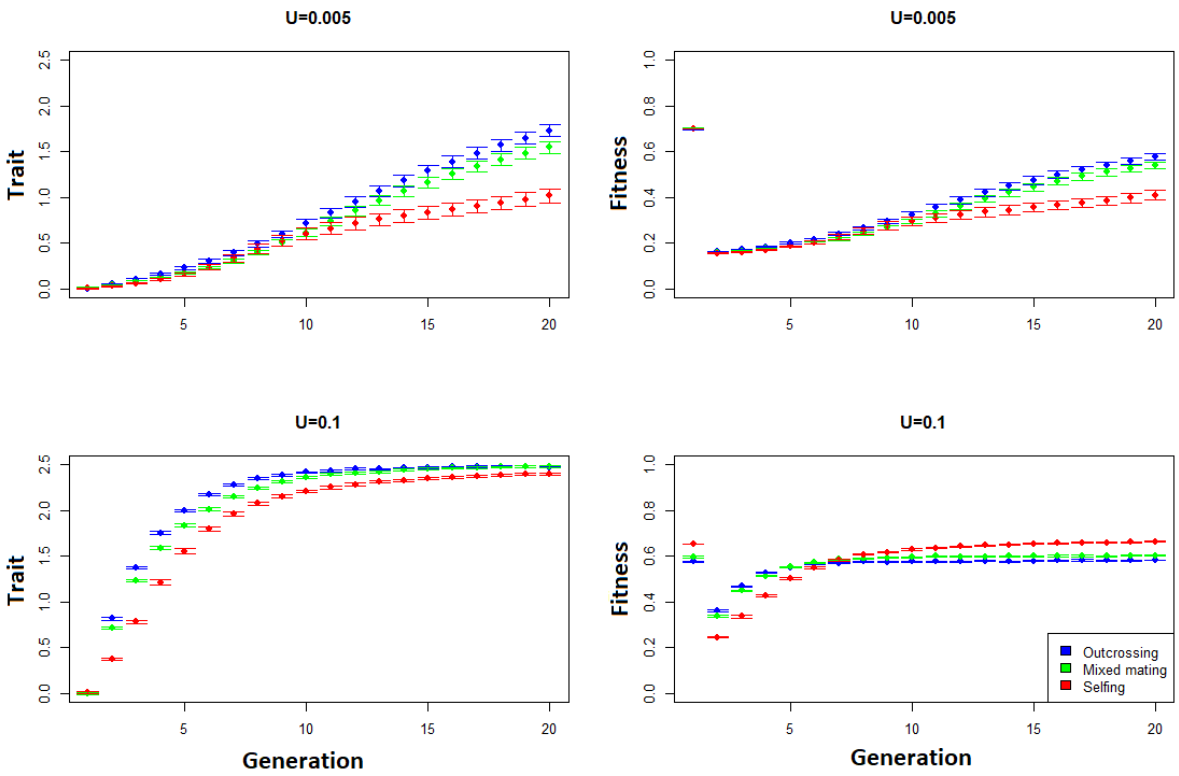

Figure 3. Dynamics of the mean trait value and of the mean fitness value of the population during the 20 generations following the environmental change, for different haploid mutation rates (Top-down: U=0.005; U=0.1) and as a function of the mating systems. Other parameter values are N=1000 and ω^2=1. Error bars stand for 95% confidence interval (n=100).

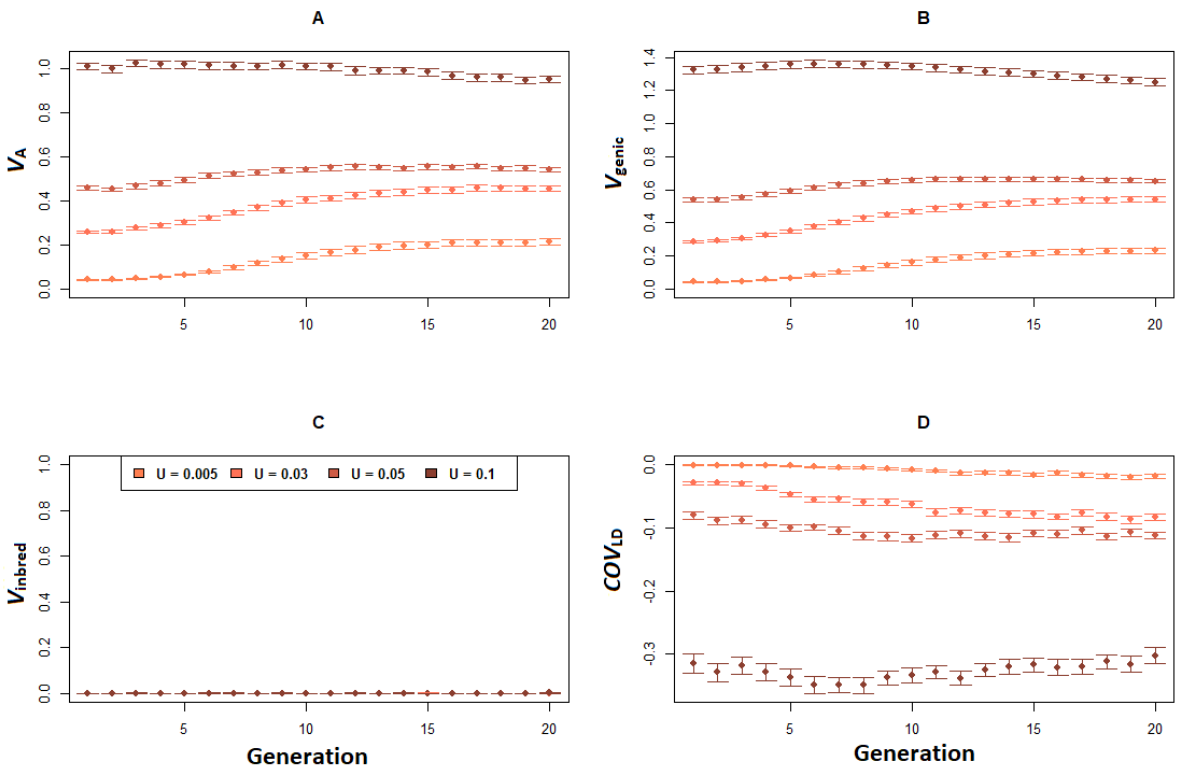

Figure 4. Dynamics of the additive genetic variance and its components in outcrossing populations (σ=0), mutation rates are set at zero after generation 0. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}), for different mutation rates (U=0.005; U=0.03; U=0.05; U=0.1). Error bars stand for 95% confidence interval (n = 100). Other parameter values are N=1000 and ω^2=1.

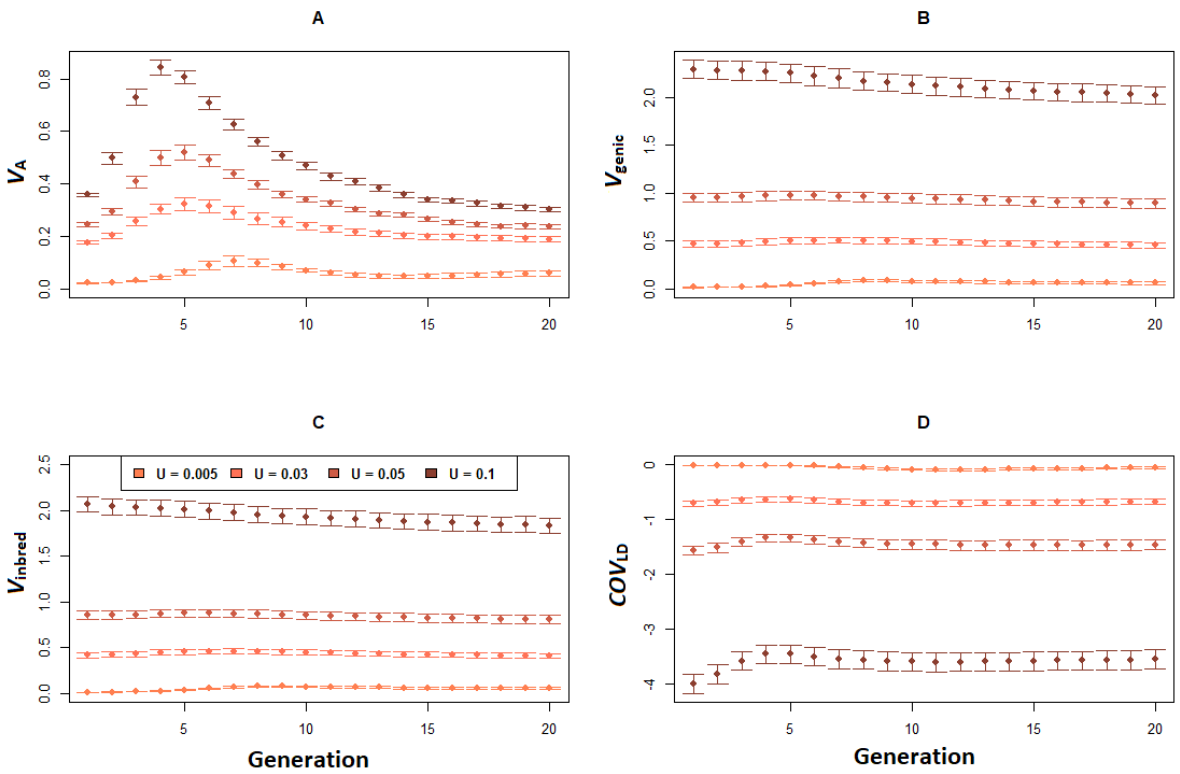

Figure 5. Dynamics of the additive genetic variance and its components in selfing populations (σ=0.95), mutation rates are set at zero after generation 0. **A.** Observed additive variance for the phenotypic trait. **B.** Genic variance for the phenotypic trait (V_{genic}). **C.** Genetic variance due to inbreeding (V_{inbred}). **D.** Genetic covariance due to linkage disequilibrium (COV_{LD}), for different mutation rates (U=0.005; U=0.03; U=0.05; U=0.1). Error bars stand for 95% confidence interval (n = 100). Other parameter values are N=1000 and ω^2=1.

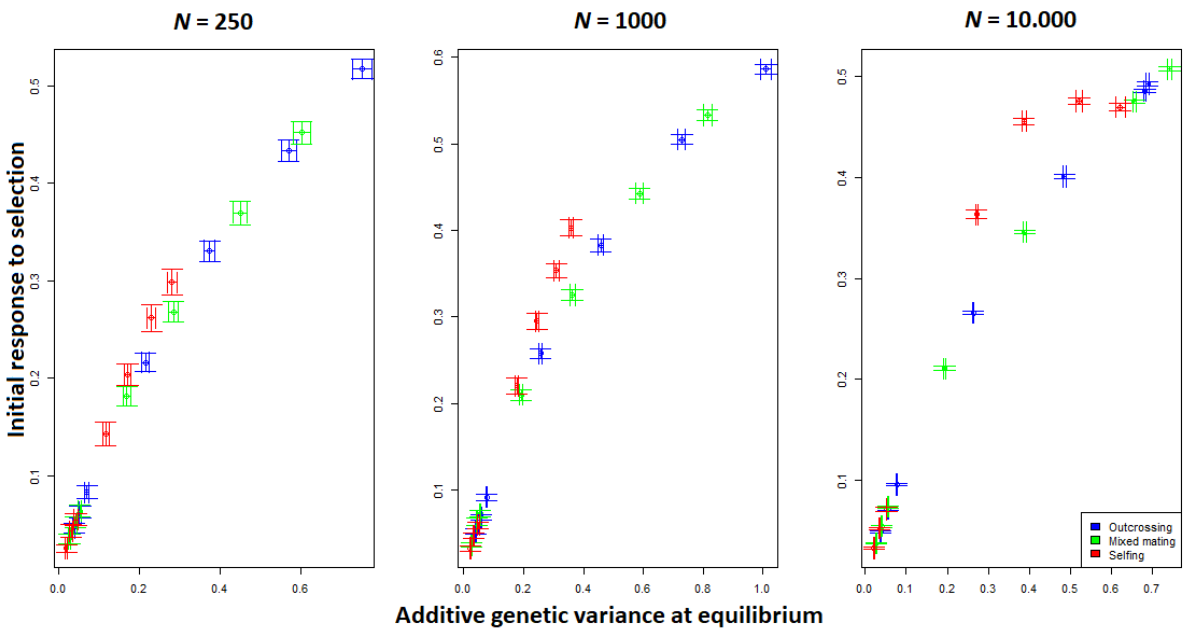

Figure 6. Initial response to selection in function of the amount of additive variance at equilibrium, as a function of the mating system, each point correspond to a haplotypic mutation rate (*U* ranging from 0.005 to 0.1, from the left to the right), and for the three population sizes, when selection is strong (ω^2 = 1). Error bars stand for 95% confidence interval (n = 100).

Appendix A: Impact of the mating system on genetic variance at mutation-selection-drift equilibrium: comparison of simulation results with analytic predictions

Several theoretical works have been developed to explicitly quantify the effect of selfing on the amount of additive genetic variance maintained in a population (Lande, 1977; Charlesworth & Charlesworth, 1995; Lande & Porcher, 2015; Abu Awad & Roze, 2018). However, these approaches often do not take the effects of genetic drift due to small population sizes into account. The Stochastic House of Cards approximation (SHC) provided in Bürger *et al.*, (1989), an extension to Turelli's (1984) House of Cards, explicitly considers the effective population size N_E, which can be chosen so as to reflect the population's reproductive strategy (see Bürger *et al.*, 1989, for the case of dieocious reproduction). Considering these different constructions of models, we briefly compare the analytical expectations of some of the above-mentioned works to our simulation results at mutation-selection-drift equilibrium.

In the SHC framework (Bürger *et al.* 1989), the amount of additive variance at equilibrium is given by:

$$V_{A-SHC} = \frac{4UV_S}{1+\frac{V_S}{N_E\alpha^2}}. \qquad \text{(A.1)}$$

Where U, V_S and α are the haploid mutation rate, the strength of selection and the variance in mutational effects, respectively. Following Caballero & Hill (1992), and ignoring background selection and linkage disequilibrium, the effective population size of an inbreeding population is $N_E = N/(1+F)$. F is the inbreeding coefficient, which, for a given selfing rate σ, is equal to σ/(2- σ). As mentioned above, this approximation takes into account the population and the strength of selection. However the inbreeding coefficient F does not account for consequences of background selection on N_E, which are not known in this specific class of population genetics

models (i.e. quantitative traits and stabilising selection) and could thus limit the accuracy of the SHC approximation.

Abu Awad & Roze (2018) provided approximations for the expected level of V_A under the continuum of alleles model, neglecting genetic drift and associations between loci:

$$V_{A-CA} = 4V_S\sqrt{\frac{\bar{s}UL}{2n}\frac{2}{2+\sigma(1-\sigma)}}\,. \qquad \text{(A.2)}$$

n is the number of quantitative traits under consideration (n = 1 in our case). In order to account for cases in which linkage is complete (i.e. for very high self-fertilisation rates), setting the number of loci L to 1 gives

$$V_{A-DL} = 4V_S\sqrt{\frac{\bar{s}U}{2n}\frac{2}{2+\sigma(1-\sigma)}}\,. \qquad \text{(A.3)}$$

This whole genome is thus modelled as a single locus, with a large number of possible alleles, whose effects follow a Gaussian distribution (Lande, 1977).

To determine the accuracy of these equations under our model conditions, we ran one hundred repetitions per parameter set for three selfing rate values (σ = 0, 0.5 and 1), as well as for the seven mutation rates tested in the main text. The mean additive variance from the simulation runs are then compared to the expectations from equations A.1, A.2 and A.3, by using the sum of absolute values of the differences between observations and expectations over the seven mutation rates. Results are presented in Table A1.

Table A1. Comparison of analytical predictions for the amount of V_A under different quantitative genetics models and results obtained in our simulations at MSD equilibrium, for different sets of parameters. The best model for the prediction of V_A (i.e. closest to our simulation results) is highlighted in green for each parameter set.

Parameters		N = 250 V_S			N = 1000 V_S			N= 10 000 V_S		
		1	9	99	1	9	99	1	9	99
$V_{A\text{-}SHC}$	σ=0	0.16	0.89	1.56	0.49	1.92	14.58	0.31	4.38	67.81
	σ=0.5	0.11	0.28	2.81	0.18	2.05	4.71	0.22	4.65	57.87
	σ=0.95	0.80	1.76	1.13	0.86	4.10	8.10	0.36	4.93	56.71
$V_{A\text{-}CA}$	σ=0	3.49	7.16	28.57	2.94	5.11	16.93	3.28	5.95	14.66
	σ=0.5	3.92	7.34	26.89	3.16	5.55	16.53	3.14	5.65	12.74
	σ=0.95	4.61	9.79	31.77	4.30	8.47	24.37	3.59	6.55	16.29
$V_{A\text{-}DL}$	σ=0	1.30	3.52	5.57	1.85	5.57	16.85	1.51	4.74	19.12
	σ=0.5	0.95	2.75	5.21	1.39	4.52	15.31	1.42	4.42	19.06
	σ=0.95	0.25	0.86	2.17	0.53	2.09	9.02	1.23	4.01	17.01

From Table 1, we find that for outcrossing and mixed-mating populations, the Stochastic House of Cars approximation (equation A.1) is the best predictor of the mean additive genetic variance for a wide range of parameters for outcrossing and mixed mating populations. The continuum of alleles approximation (equation A.2) is accurate only for very large population sizes and weak selection, conditions for which the underlying assumptions made when deriving this approximation are met (low allelic frequencies, weak genetic drift and few among-loci associations). For predominantly selfing populations, the approximation considering full linkage (equation A.3) is accurate in conditions where strong genetic associations are maintained throughout the genome *i.e.* selection is strong or moderate (table A1).

Bibliography

Abu Awad, D. & Roze, D. 2018. Effects of partial selfing on the equilibrium genetic variance, mutation load, and inbreeding depression under stabilizing selection. *Evolution* **72**: 751–769.

Bürger, R., Wagner, G.P. & Stettinger, F. 1989. HOW MUCH HERITABLE VARIATION CAN BE MAINTAINED IN FINITE POPULATIONS BY MUTATION–SELECTION BALANCE? *Evolution* **43**: 1748–1766.

Caballero, A. & Hill, W.G. 1992. Effective size of nonrandom mating populations. *Genetics* **130**: 909–916.

Charlesworth, D. & Charlesworth, B. 1995. Quantitative genetics in plants: the effect of the breeding system on genetic variability. *Evolution* **49**: 911–920.

Lande, R. 1977. The influence of the mating system on the maintenance of genetic variability in polygenic characters. *Genetics* **86**: 485–498.

Lande, R. & Porcher, E. 2015. Maintenance of quantitative genetic variance under partial self-fertilization, with implications for evolution of selfing. *Genetics* **200**: 891–906.