Interplay between high-drift and high-selection limits the genetic load in small selfing maize populations.

Arnaud Desbiez-Piat[*], **Arnaud Le Rouzic**[†], **Maud I. Tenaillon**[*,1] **and Christine Dillmann**[*,1]

[*]Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, France, [†]Université Paris-Saclay, CNRS, IRD, UMR Évolution, Génomes, Comportement et Écologie, France.

ABSTRACT Population and quantitative genetics provide useful approximations to predict the evolution of populations and their multilocus adaptive dynamics. They are not supposed to hold under extreme parameter combinations, for which deviations need to be further quantified to provide insights into specific population dynamics. Here we focused on small selfing populations evolving under an under-explored High Drift-High Selection (HDHS) regime. We combined experimental data from the Saclay divergent selection experiments on maize flowering time, forward individual-based simulations, and theoretical predictions to dissect the evolutionary mechanisms at play in the observed selection response for a highly complex trait. We asked two main questions: How do mutations arise, spread, and reach fixation in populations evolving under HDHS? How does the interplay between drift and selection influence the response to selection ? We showed that the long-lasting response to selection in populations whose estimated effective population size ranged between 2.5 to 4 is due to the rapid fixation of *de novo* mutations. Among all fixed mutations, we found a clear signal of enrichment for beneficial mutations revealing a limited cost of selection in these populations. We argue that environmental stochasticity and variation in selection coefficients contribute to exacerbate mutational effects, thereby facilitating selection grasp and fixation of small-effect mutations. Hence the HDHS regime with non-limiting mutation highlights an interesting interplay between drift and selection that sustains a continuous response to selection. We discuss our results in the context of breeding populations and long-term survival of small selfing populations.

KEYWORDS Truncation selection, Experimental evolution, Adaptive dynamics, Distribution of fitness effects, Selection cost, Effective population size, Environmental stochasticity

Understanding the evolutionary processes sustaining phenotypic shifts is at the core of quantitative genetic models. Empirical description of such shifts takes its roots in the breeding literature where truncation selection generates significant and sustainable responses (Hill and Caballero 1992; Walsh and Lynch 2018). Truncation selection is known to be the most effective form of directional selection (Crow and Kimura 1979). Under truncation selection, limits to the evolution of phenotypes are rarely reached as heritable variation persists through time (Odhiambo and Compton 1987; Moose *et al.* 2004; Weber and Diggins 1990; Caballero *et al.* 1991; Mackay 2010; Lillie *et al.* 2019). Such observations fit well with the Fisher's infinitesimal model (Fisher Ronald Aylmer 1930) and the derivatives of the breeder equation (Lush 1943; Lande 1979; Lande and Arnold 1983), which predict a continuous and linear response with no finite limits. The rate of response is however expected to decline with selection-induced linkage disequilibrium (Bulmer 1971; Hospital and Chevalet 1996). Hence under finite population size, selection response is predicted to reach an asymptotic finite limit (Robertson 1960) as exemplified in mice (Roberts 1967; Falconer 1971). Results from other species are more equivocal (e.g. drosophila (Weber 1990; Weber and Diggins 1990; Weber 1996), or maize (Odhiambo and Compton 1987; Moose *et al.* 2004; Dudley and Lambert 2010; De Leon and Coors 2002; Lamkey 1992). Incorporation of *de novo* mutations indeed predicts a slower rate of response instead of a hard limit (Hill 1982b,a; Weber and Diggins 1990; Wei *et al.* 1996; Walsh and Lynch 2018). These models point to a sub-optimal average selection response in two situations: when population size, N is below 10^4 and the genetic variance correspondingly small at mutation-drift equilibrium $\widehat{V_G}$ (Hill 1982b; Houle 1989); and when $\widehat{V_G}$ is reduced due to strong selection (Houle 1989). More

doi: 10.1534/genetics.XXX.XXXXXX
Manuscript compiled: Tuesday 22nd December, 2020
[1]Corresponding authors: Université Paris-Saclay, INRAE, CNRS, AgroParisTech, GQE - Le Moulon, 91190, Gif-sur-Yvette, France

generally, quantitative genetic models that include selection, drift and mutation (Houle 1989) are well-suited for predicting observed selection responses in a broad range of parameters (Hill and Rasbash 1986) — providing appropriate corrections (Walsh and Lynch 2018). Nevertheless, these models have been often developed under the general assumption of random mating, and a probability of fixation of new mutations determined by the product of population size by their selection coefficient, Ns, to be either $\ll 1$ or $\gg 1$. Mathematical models for the intermediate regime $Ns \approx 1$ and non-random mating still remains unsatisfactory. And the description of mechanisms of long-term selection response and whether it can be understood and predicted by existing equations has yet to be explored for polygenic traits under a High-Drift High-Selection regime and selfing.

Both the Distribution of mutational Fitness Effects (DFE) and the associated mutation rate are central to such predictions. Selection makes the DFE of fixed mutations different from that of incoming mutations (Kassen and Bataillon 2006). In large populations, a high proportion of incoming beneficial mutations are predicted to reach fixation, together with vanishing small effect deleterious mutations (Crow and Kimura 1971; Kimura 1983). In small populations and/or small selection intensity instead, frequent loss of beneficial mutations due to drift together with the fixation of moderately strong deleterious mutations is expected. Hence Kimura's equation that links the fixation probability (p) of a mutation to the population size (N) and selective coefficient (s) — $P_{fix}(s,p,N) = \frac{1-e^{-4spN}}{1-e^{-4sN}}$ — applies to a vast range of parameters including s values as high as 0.1 and N as small as 10 individuals Carr and Nassar (1970). An additional layer of complexity to DFE prediction comes from the mating system. Adaptation of very large asexual populations (such as microbes) is indeed affected by competition between alternative beneficial mutations occurring in different genetic background, a process referred to as clonal interference (Gerrish and Lenski 1998). Here the absence of recombination favors enrichment of the DFE in large mutational effect (Gerrish and Lenski 1998). However, if selection overpowers drift, *i.e.* $Ns \gtrapprox 1$, or if the rate of beneficial mutation (μ_B) is small enough, the expected time lag between two successive mutations is sufficiently large for the first beneficial mutation to fix without interference of the second. While such behavior is expected when $N\mu_B \gg 1/ln(Ns)$, for $N\mu_B \gtrapprox 1/ln(Ns)$ beneficial mutations evolves under clonal interference (Desai and Fisher 2007). Altogether these results highlights how the interplay of key parameters - N, s, μ, effective recombination — determine the DFE and in turn, the long-term selection response.

Genomic footprints of selection have considerably enriched our vision of allele adaptive trajectories sustaining selection responses. Observed genomic footprints include hard selective sweeps characterized by strong decrease in genomic diversity at the selected loci and its surrounding region through genetic hitchhiking (Hermisson and Pennings 2017); and soft sweeps associated with a weaker signature either because recombination on standing variation occurs so that a given advantageous mutation is associated with multiple haplotypes, or because recurrent *de novo* mutations are associated with multiple haplotypes. Classically, population genetic models describe adaptation as a succession of sweeps at loci encoding a trait. They have been challenged by quantitative genetics that rather posits a collective response at many loci translating into simultaneous subtle shifts in allele frequencies, the so-called polygenic selection model (Berg and Coop 2014; Wellenreuther and Hansson 2016; Walsh and Lynch 2018). Whether adaptation proceeds through hard/soft sweeps or polygenic model primarily depends on the population-scaled mutation rate (θ) as well as the number of redundant loci that offer alternative ways for adaptation (L) — the mutational target. Adaptation proceeds by sweeps for small $\theta \times L$ ($\leq$ 0.1) while polygenic adaptation require large $\theta \times L$ ($\geq$ 100) — in compliance with the infinitesimal model — with partial/soft sweeps in between (Höllinger *et al.* 2019; Messer and Petrov 2013). Extension of the hitchhiking model to a locus affecting a quantitative trait with an infinitesimal genetic background predicts that, under the hypothesis of a Gaussian fitness function, the fixation of a favorable mutation critically depends on the initial mutation frequency and the distance to the optimum (Chevin and Hospital 2008). Interestingly, while demographic parameters play a relatively small role in the speed of adaptation compared to standing and mutational variance, they change its qualitative outcome. Population bottlenecks diminish the number of segregating beneficial alleles, favoring hard sweeps from *de novo* mutations over soft sweeps from standing variation (Stetter *et al.* 2018).

Experimental evolution tracing short-term temporal dynamics of adaptation have complemented and validated models of adaptation, providing further hints into allele frequency changes, and into the extent of polymorphism and competition among beneficial mutations under various drift/selection/recombination regimes. Temporal dynamics may be obtained either through pedigree information or time series samples. This last approach has been employed successfully in microorganisms where complex patterns of mutation spreading have been observed during the course of adaptation. These include clonal interference, reduction of the benefit of a mutation in fit *versus* less fit genotypes (diminishing-return epistasis), and evidence for the same favorable mutation being selected in multiple independent evolved clones (genetic parallelism) (Good *et al.* 2017; Neher 2013; Good *et al.* 2012; Desai and Fisher 2007; Gerrish and Lenski 1998; Spor *et al.* 2014). However, in asexually reproducing microbes, adaptation proceeds through *de novo* mutations, which may reveal specific patterns not found in sexually-reproducing eukaryotes. In yeast, for instance, most adaptive changes correspond to the fixation of initial standing variation (Burke *et al.* 2014; Burke 2012). Patterns of allele frequency changes depend crucially on both N_e and the frequency of sex, that are themselves intimately linked (see Hartfield *et al.* (2017)). Considering a single locus, fixation time decreases correlatively with the level of self-fertilization (Haldane 1927). At the same time, multilocus simulations have shown that selfing reduces effective population size through background selection, and in turn beneficial mutations are less likely to fix (Kamran-Disfani and Agrawal 2014; Roze 2016). In addition, as selection interference reduces the efficiency of selection in low-recombining regions, high selfing rates also increase the fixation of deleterious mutations through genetic hitchhiking (Hartfield and Glémin 2014). These insights are together in line with the low selection approximation that posits that reduction in effective recombination decreases selection efficiency.

In the current paper, we aimed at investigating the dynamics of the response to selection of selfing populations evolving under High Drift-High Selection regime. Situated at the parameters boundaries of current models, this regime is of particular interest to understand the limits of adaptation and long-term survival of small selfing populations undergoing strong selection. We relied here on two Divergent Selection Experiment (DSEs) conducted

for 18 generations on Saclay plateau (Saclay DSEs). These Saclay DSEs are ideal settings to address those issues: selection-by-truncation has been applied in a higher organism (maize), on a highly polygenic and integrated trait (flowering time, (Buckler *et al.* 2009; Tenaillon *et al.* 2018)) that directly affects fitness. Previous results indicate continuous phenotypic responses — values of mutational heritability ranged from 0.013 to 0.025 — sustained by a constant mutational input (Durand *et al.* 2010, 2012, 2015). We asked two main questions: How do mutations arise, spread, and reach fixation in populations evolving under HDHS? How does the interplay between drift and selection influence the response to selection ? To answer those questions, we confronted the observed phenotypic response in Saclay DSEs to forward individual-based simulations that explicitly modeled the same selection (selection of 1% of the most extreme) and demographic scheme, and used theoretical predictions to measure deviations from expectations.

Materials and Methods

Plant material and historical field evaluation

We have conducted two independent divergent selection experiments (Saclay DSEs) for flowering time from two commercial maize inbred lines, F252 and MBS847 (thereafter MBS). These experiments were held in the field at Université Paris-Saclay (Gif-sur-Yvette, France). The selection procedure is detailed in Fig. S1 and Durand *et al.* (2010). Briefly, within each Saclay DSE, the ten earliest (resp. ten latest) flowering individuals were selfed at each generation to produce each 100 offspring used for the next generation of selection within the Early (resp. Late) populations. Within each population, we evaluated offspring of a given progenitor in four rows of 25 plants randomly distributed in a four-block design, so that each block contained 10 rows. We applied a truncation selection of 10/1000=1%. We conditioned selection on the maintenance of two families, *i.e.* two sub-pedigrees derived from two separate G_0 ancestors. Thus, each family was composed of three to seven individuals at each generation with the additional condition that at least two different G_{n-1} progenitors were represented. Furthermore we applied a two-steps selection procedure, so that among the 100 offspring, we recorded the flowering time of 12 individuals per family, *i.e.* 24 per population. To ensure the maintenance through time of minimal fitness, we selected among the 24 earliest (resp. latest) individuals, the 10 earliest (resp. latest) individuals with the highest kernel weight. Seeds from selected progenitors at all generations were stored in cold chambers.

We traced back the F252 and MBS pedigrees from generation 20 (G_20) to the start of the divergent selection experiments, G_0. The initial MBS pedigrees encompassed four families: ME1 and ME2 for the MBS Early (ME) population, and ML1 and ML2 for the MBS Late (ML) population (Fig. S2. F252 Early (FE) population was composed of FE1 and FE2 families (Fig. S2). F252 Late populations genealogies were more complex: FVL families (F252 Very Late in Durand *et al.* (2015)) ended at generation 14 with the fixation of a strong effect allele at the *eIF-4A* gene (Durand *et al.* 2015). To maintain two families in F252 Late population, two families FL2.1 and FL2.2 were further derived from the initial FL2. These two families pedigrees are rooted in FL2 from a single G_3 progenitor (Fig. S2).

Phenotypic evaluation and observed selection response analysis

The same approach as Durand *et al.* (2015) was applied. Briefly, progenitor flowering dates, measured here as the number of days to flowering after sowing equivalent to 20°C days of development (Parent *et al.* 2010), were recorded as the 12 earliest or latest plants in their progeny at each generation of the Saclay DSEs. We used these records to investigate the response to selection treating each family independently. After correction for block effects, and year effects according to equation (1) of Durand *et al.* (2015), the linear component b_{jk} of the within-family response to selection was estimated using the following linear model:

$$Y_{ijklm} = \mu_0 + b_{jk} \times \text{gener}_i + \varepsilon_{ijklm} \quad (1)$$

where μ_0 is the intercept corresponding to the average flowering time at generation G_0, i stands for the year and corresponding generation of selection, j for the population, k for the family within population, l for progenitor within family, and m for the plant measurements within progenitor.

Family means and standard errors were computed at each generation to represent families selection responses presented Fig. 1 (a). All the values were centered around 100 for comparison purposes with the simulated responses.

Model framework

We used forward individual-based simulations that explicitly modeled the same selection — proportion of selected individuals=1% — and demographic scheme — variations in population size — as Saclay DSEs. This regime is referred to High-Drift High Selection (HDHS). **Initial G_0 simulation**: We obtained our initial population by mimicking a classical selection scheme used to produce fixed maize inbred lines in industry. To do so, we started from an heterozygous individual that was selfed for eight generations in a single-seed descent design. An additional generation of selfing produced 60 offspring that were reproduced in panmixia for two generations to constitute the 60 individuals of the G_0 initial population. Therefore, we started our simulations with a small initial residual heterozygosity ($\leq$ 0.5%). **G_1 simulation**: Considering one Saclay DSE, we selected from the initial population (60 individuals), the two earliest and the two latest flowering parents on the basis of their average phenotypic value measured over 12 offspring. Each of these individuals constituted the ancestor of each of the four families. They were selfed to produce 100 offspring. **Subsequent generations n**: From there, we simulated the exact same selection scheme that included a two-steps procedure (Fig. S1). First, we selected the 12 earliest (within each early family) and the 12 latest individuals (within each late family) from the 100 offspring of each parent. We next selected the five earliest (within each early family) and five latest (within each late family). In other words, at each step we retained 5 out of 500 individuals within each of the four families. Note that we imposed that the five selected individuals did not share the same parent.

Simulated genetic and phenotypic values

Because maize flowering time is a highly polygenic trait (Buckler *et al.* 2009; Tenaillon *et al.* 2018), we imposed the haploid number of loci $L = 1000$. The genome of one individual was composed of 10 chromosomes. In each simulation: (i) we randomly assigned each locus to a chromosome so that genome composition varied from one simulation to another; (ii) the position of each

locus within each chromosome was uniformly drawn between 0 and 1.5, 1.5 Morgan being the total genetic length of each chromosome; (iii) the crossing-over positions along chromosomes were drawn in an exponential law of parameter 1, which corresponded to an effective crossing-over every Morgan. The initial population (G_0) consisted of 60 individuals polymorphic for a small fraction of loci (residual heterozygosity). Let G_g^i be the genotype of the individual i of the generation g. Let $a_l^{f(i,g)}$ the allelic effect at the locus l of the paternal chromosome f of the individual i at the generation g and $a_l^{m(i,g)}$ the allelic effect at locus l of maternal chromosome m of individual i at generation g. This allows us to model the genotype of an individual as :

$$G_g^i = [(a_1^{f(i,g)}, a_2^{f(i,g)}, ..., a_l^{f(i,g)}, ..., a_L^{f(i,g)}), (a_1^{m(i,g)}, a_2^{m(i,g)}, ..., a_l^{m(i,g)}, ..., a_L^{m(i,g)})] \quad (2)$$

The initial allelic effects were drawn in an reflected exponential distribution, that is to say :

$$\forall l \wedge \forall (f(i,g) \vee m(i,g)), a_l \sim \text{Reflected exp}(\lambda) \quad (3)$$

Hence the probability density:

$$f(a_l, \lambda) = \frac{1}{2}\lambda e^{-|\lambda a_l|} \quad (4)$$

which implied that:

$$\mathbb{E}[a_l] = 0 \text{ and } \mathbb{V}[a_l] = \frac{2}{\lambda^2} \quad (5)$$

Starting from a hybrid heterozygote at all loci, we showed that after g generations of selfing and two generations of bulk, for L loci without linkage disequilibrium or mutation, we expected:

$$\mathbb{E}(\sigma^2_{A_0} = \sigma^2_{g+2}) = \mathbb{E}(\sigma^2_g) = \frac{1}{2^g} \times L \times \frac{2}{\lambda^2} \quad (6)$$

Therefore, to match the field estimation $\widehat{\sigma^2_{A_0}}$, one could let

$$\lambda = \sqrt{2L\frac{1}{2^g}\frac{1}{\widehat{\sigma^2_{A_0}}}}. \quad (7)$$

However, drift, linkage disequilibrium and mutation can lead to deviations from the expected value of the initial genetic variance. We therefore recalibrated all the allelic effects at generation 0 to match the initial $\widehat{\sigma^2_{A_0}}$ additive variance. To do so, we multiplied by a corrective factor $k = \sqrt{\frac{\widehat{\sigma^2_{A_0}}}{\mathbb{V}(A_0)}}$, where $\mathbb{V}(A_0)$ was the additive variance of our population G_0, calculated in multiallelic as $2 \times \sum_{i=1}^{L} p_i \alpha_i^2$ with p_i the frequency of the allele i and α_i its effect. So at G_0, $\mathbb{V}(A_0) = \widehat{\sigma^2_{A_0}}$.

Mutations occurred at each reproduction event. We drew the number of mutation per locus in a Poisson distribution of mean $L \times \mu$ where μ was the mutation rate per locus. We drew the effect of a mutation at a locus in a reflected exponential distribution of parameter $\lambda_{mut} = 2\sqrt{\frac{L\mu}{\sigma^2_M}}$. We computed phenotypic values as the sum of all allelic effects (100×2) plus an environmental effect randomly drawn in a normal distribution of mean 0 and variance σ^2_E.

Selection and drift regimes

As a control, we considered a neutral model without selection (the No Selection regime, NS) where the same selection scheme as in regime with selection was applied, but individual genotypic values were drawn in a normal distribution of mean 100 and variance σ^2_E, independently of the previous progenitors. In other words, we attributed as phenotypic values non-heritable environmental values. Genotypes were recorded for mutation-tracking purpose only.

In addition, we considered an alternative drift regime, where we increased the census population size by a factor 10, *ceteris paribus*, thereafter referred as the Low Drift regime (LD). Under this regime, we selected in each Saclay DSE the fifty earliest/latest individuals within early/late families out of five thousand individuals (instead of five hundred).

We performed 2000 simulations for each of the four families in each of the four regimes.

Parameter calibration

Clark *et al.* (2005) estimated the nucleotidic substitution rate to be in the range of 30×10^{-9}. We estimated roughly the mean mRNA length from maize reference genome V4 (Jiao *et al.* 2017) to be equal to 6000 (median=5197, mean=7314). Hence we used a mutation rate per loci of: $\mu = 6000 \times 30 \times 10^{-9} = 1.8 \times 10^{-4}$. Other variance parameters were chosen such that the HDHS simulated cumulative response falls in the same range of values as the MBS observed one. To account for some flexibility in these variance parameters, we drew for each simulation the variance in an inverse-gamma distribution prior. More precisely,

$$\sigma^2_E \sim \Gamma^{-1}(17.6700, 37.5075) \implies \mathbb{E}(\sigma^2_E) = 2.25, \mathbb{V}(\sigma^2_E) = 0.32$$

$$\sigma^2_{A_0} \sim \Gamma^{-1}(3.540, 5.715)) \implies \mathbb{E}(\sigma^2_{A_0}) = 2.25, \mathbb{V}(\sigma^2_{A_0}) = 3.28$$

$$\sigma^2_M \sim \Gamma^{-1}(17.67, 0.5626125) \implies \mathbb{E}(\sigma^2_M) = 3.38 \times 10^{-2}$$
$$\mathbb{V}(\sigma^2_M) = 7.27 \times 10^{-5}$$

Expected response, effective population size and time to the most recent common ancestor

We computed the expected cumulative response after t generations for haploid population as (Hill 1982b; Wei *et al.* 1996; Weber and Diggins 1990; Walsh and Lynch 2018):

$$R(t) \approx N_e \frac{i}{\sigma_P}\left[t\sigma^2_m + \left(1 - e^{-\frac{t}{N_e}}\right)\left(\sigma^2_A(0) - N_e\sigma^2_m\right)\right] \quad (8)$$

The effective population was the only parameter not explicitly defined in our simulations and is of crucial importance in the response to selection. We estimated N_e following two approaches. First using the Time to the Most Recent Common Ancestor (TMRCA) from the standard coalescence theory for a haploid sample of size k at generation g (Walsh and Lynch 2018):

$$\mathbb{E}(\text{TMRCA}_g) = 2N^{Coal}_{e(g)} \times (1 - \frac{1}{k}) \quad (9)$$

Second, from the variance in offspring number Crow and Kimura (1971); Durand *et al.* (2010), where N_e can be computed as

$$N^{Var(o)}_{e(g)} = \frac{N-1}{\text{Var}_{(g)}(\text{OffspringNumber})} \quad (10)$$

In the simulations, $N_{e(g)}^{Coal}$ and $N_{e(g)}^{Var(o)}$ were computed at generation G_{20}. We also computed the harmonic means between generations G_2 and G_{20} and computed the whole distribution (in $2N_e$ generations) of the Kingman coalescent TMRCA as (Tavaré 1984):

$$f_{\text{TMRCA}}(t) = \sum_{i=2}^{n} \frac{(2i-1)(-1)^i(n(n-1)...(n-i+1))}{n(n+1)...(n+i-1)} \binom{i}{2} e^{-\binom{i}{2}t} \tag{11}$$

Fitness function and Kimura's expected fixed mutational DFE

Using diffusion equations, Kimura (Kimura 1962) predicts the fixation probability of a mutation of selective value s and initial frequency p:

$$P_{fix}(s(a), p, N_e) = \frac{1 - e^{-4s(a)pN_e}}{1 - e^{-4s(a)N_e}}. \tag{12}$$

We considered a mutation occurring during meiosis in one plant among the 500 of a family observed at a given generation. When occurring, its effect on the phenotypic variance is negligible. Therefore, initially this plant was selected independently of the mutation. The 5 selected individuals comprised one heterozygote (Aa) bearing the mutation, and 4 homozygotes (aa). Each selected individual produced 100 progenies, so that the fitness effect of the mutation was evaluated at the next generation in a population of 500 plants where the frequency of the mutant allele was $p = 1/10$ (Table 1).

Table 1 Fitness model

Genotype	AA	Aa	aa
Genotypic frequency	1/20	2/20	17/20
Fitness value	w_{AA}	w_{Aa}	w_{aa}
Additive mutational effect	$a_{AA} = 2a$	$a_{Aa} = a$	$a_{aa} = 0$

In this population, the distribution of flowering time resulted from a mixture of gaussian distributions.

$$f(x) = \sum_k \Pi_k f_k(x) \tag{13}$$

where $f_k(x)$ is the flowering time distribution for plants with genotype $k \in AA, Aa, aa$. As we selected 1% of the latest (resp. earliest) flowering plants, all selected plants did flower after the date z, computed as the 1% quantile of the mixture distribution. The selection effect $s(a)$ depended on the effect a of the mutation on flowering time (Table 1). Indeed, the relative weight of homozygous mutants AA among selected individuals was computed as:

$$w_{AA} = \frac{1 - F_{AA}(z)}{\sum_k 1 - F_k(z)} \tag{14}$$

Which leads to:

$$s(a) = \frac{F_{aa}(z) - F_{AA}(z)}{1 - F_{aa}(z)} \tag{15}$$

The fixation probability $P_{fix}(s(a), p, N_e)$ was computed as in (Eq. 12) using $s(a)$ (Eq. 15), $p = 1/10$, and $N_{e(g)}^{Coal}$ for N_e. The mutational effect a was drawn in a reflected exponential distribution of parameter λ_{mut} and density function $g_{\lambda_{mut}}(a)$. Hence, the density of fixed mutations $h(a)$ was computed as:

$$h(a) = \frac{g_{\lambda_{mut}}(a) P_{fix}(s(a), p, N_e)}{\int g_{\lambda_{mut}}(x) P_{fix}(s(x), p, N_e) dx}. \tag{16}$$

Moreover, we recorded the simulated a_{sim} of each fixed mutation and computed the realized distribution, using kernel estimation methods, $h_{obs}(a)$.

Results

In order to examine the evolution and fate of small selfing populations submitted to strong selection, we investigated the dynamics of the response to selection under a High Drift-High Selection (HDHS) regime imposed on two divergent artificial selection experiments for flowering time in maize (Saclay DSEs). We compared experimental data to results of a simulation model specifically devised to mimic our experiments; and further computed when possible expectations from population and quantitative genetics theory.

Empirical response after 20 generations of selection In line with previous observations for the first 16 generations, we observed significant responses (Fig. 1 a, Tab. 2a, 2b) to selection after 20 generations in all families. Marked differences among families nevertheless characterized these responses. This is well exemplified in the Late F252 families where one family (FVL) responded very strongly with a mean shift of 11.32 Days to Flowering (DTF) after 13 generations, corresponding to a linear regression coefficient of 0.86 DTF/generation (Tab. 2a). This family fixed a deleterious allele at G_{13} and could not be maintained further (Durand *et al.* 2012). We examined two derived families from G_{11}, the FL2.1 and FL2.2. These families were shifted by 3.19 DTF and 2.60 DTF from the G_0 FL2 mean value for FL2.1 and FL2.2, respectively. These corresponded to a linear regression coefficient of 0.11 DTF/generation for FL2.1 and 0.12 DTF/generation for FL2.2 (Tab. 2a). The selection response were more consistent for the two Early F252 families, with a shift after 20 generations of -4.27 DTF for FE1, and a shift of -5.34 DTF for FE2 (Tab. 2a). Considering MBS genetic background, the late (resp. early) MBS families were shifted by 8.64 DTF for ML1, and 11.05 DTF for ML2 (resp. -9.34 DTF for ME1 and -11.72 DTF for ME2), with linear regression coefficient of 0.24 DTF/generation for ML1, and 0.46 DTF/generation for ML2 (resp. -0.41 DTF/generation for ME1 and -0.42 DTF/generation for ME2) DTF (Tab. 2b).

Simulation model validation We used simulations both to validate our model and to explore two drift intensities, High and Low. We used corresponding negative controls with No Selection (NS) which lead to four regimes: High Drift-High Selection (HDHS, the default regime), High Drift-No Selection (HDNS), Low Drift-High Selection (LDHS) and Low Drift-No Selection (LDNS). In order to validate the parametrization of our model, we compared the observed MBS response in all families to the simulated selection responses. Because of the symmetry in the model construction and for simplicity, simulated results are described for late populations only. Considering the HDHS regime, we recovered a simulated response with a mean genetic gain of 0.49 DTF/generation (Fig. 1, Tab. 2c). Starting from a mean genotypic value of 100 DTF, the mean genotypic value was shifted by 13.0 DTF (SD: 5.2) after 20 generations. Our simulated response therefore closely matched the observed response indicating an accurate parametrization of our simulation model (Fig. 1. We

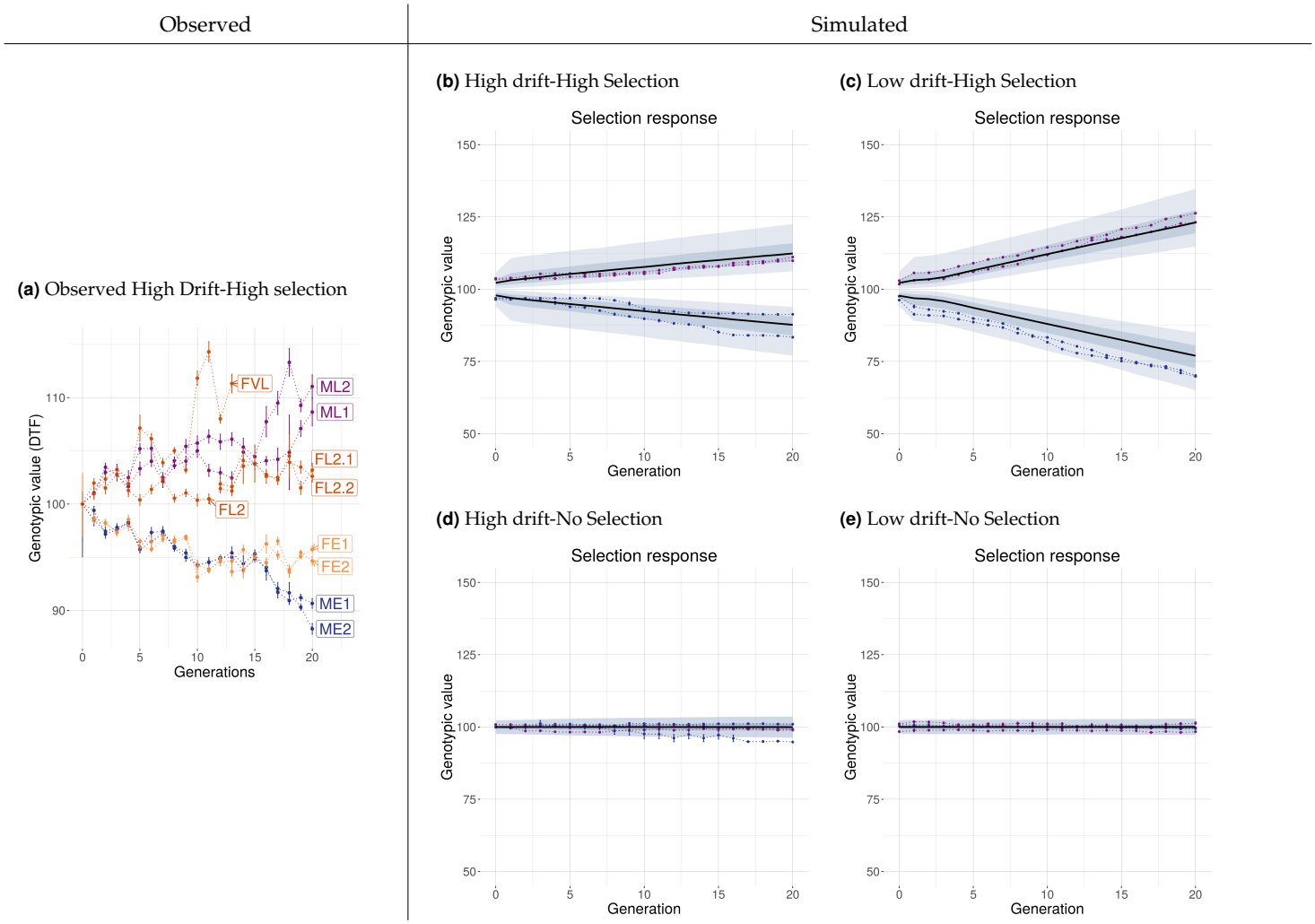

Figure 1 Observed and simulated selection response. Selection response is visualized by the evolution of the mean genotypic values of the selected progenitors per family (expressed in Days To Flowering, DTF) across generations based on observed (a) and simulated (b-e) data. In (a), red (resp. orange) corresponds to late (resp. early) flowering F252 families, while violet (resp. blue) corresponds to late (resp. early) flowering MBS families. All families were centered around 100, and Vertical bars correspond to $\pm$ 1 genotypic standard error around the mean. We simulated four regimes with the parameters calibrated from the MBS observed response: High Drift-High Selection (b), Low Drift-High Selection (c), High Drift-No Selection (d), Low Drift-No Selection (e). Violet (resp. blue) color identifies late (resp. early) population. In each population, the black line represents the evolution of the median value over 2000 simulations of the family genotypic mean. The shaded area corresponds to the 5^{th}-95^{th} percentiles (light blue) and to the 25^{th}-75^{th} percentiles (dark blue). In addition, two randomly chosen simulations are shown with dotted lines

formally tested the significance of our simulated response by comparing the linear response under HDHS to that obtained under HDNS. We were able to reject the null hypothesis of no selection response in 96.4% of the simulations under HDHS (P-value<0.05).

To investigate the impact of a ten-fold increase of the census population size on selection response, we contrasted HDHS to LDHS. Just like for HDHS, we obtained a significant response under LDHS with a mean genetic gain of 1.10 DTF/generation (Fig. 1, Tab. 2c). This gain was greater than the +0.035 DTF/generation (SD: 0.035) obtained for the LDNS control model, and we were able to reject the null hypothesis of no selection response in 100% of the simulations. The gain under LDHS corresponded to a shift of +24 DTF (SD: 6.2), which was substantially higher than that observed under HDHS. Hence multiplying the census population size of HDHS by 10 (LDHS) resulted in roughly doubling the selection response.

In sum, we validated the accuracy of our model by showing that the simulated response closely matched the observed response. We further demonstrated that selection triggered the response in all populations under both low and High Drift. Finally, we confirmed our expectation that the selection response was higher in a Low Drift than in a High Drift regime.

Effective population size:

We estimated coalescent effective population sizes N_e from the standard coalescence theory (Eq: (9)) using a Wright-Fisher population of size 5 (HD) and 50 (LD) individuals. With 5 individuals, we expected a theoretical coalescence time around 8 generations, and with 50 individuals, around 98 generations (i.e. more than the number of simulated generations). Focusing on the last generation, our simulations provided estimates of mean G_{20} TMRCA of 7.6 generations under neutrality (NS) for

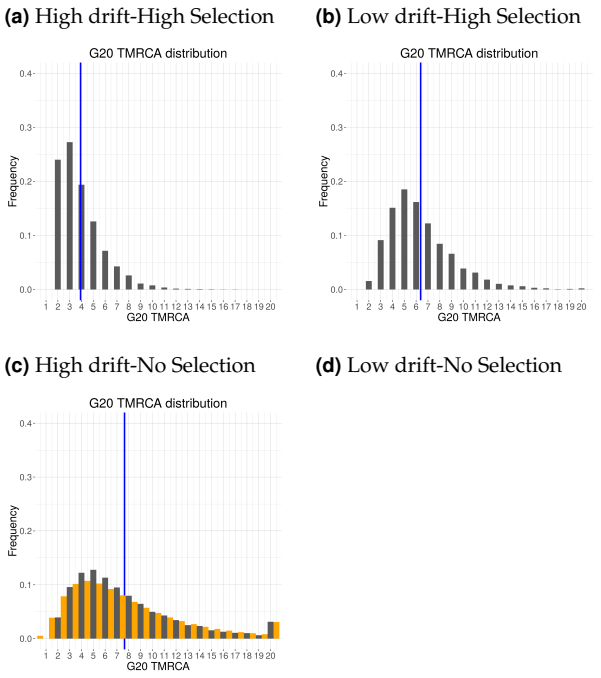

Figure 2 Frequency distribution of the Time to the Most Common Ancestor of progenitors constituting the last simulated generation. G_{20} TMRCA distribution (in grey) was obtained under HDHS (a), LDHS (b), HDNS (c) with mean TMRCA indicated as a blue vertical line. In (c), we plotted in gold the theoretical expectation of TMRCA distribution following Eq: (11). Note that under LDNS, theoretical expectations for TMRCA reached 98 generations, while our simulations were run for 20 generations. We therefore discarded the corresponding graph.

HD, closely matching the theoretical expectation of 8 (Tab. 2c). Considering the LDNS simulations, theoretical expectations (98) largely exceeded the number of generations (20). In contrary, we found mean G_{20} TMRCA of 3.9 under HDHS, and 6.4 under LDHS. Fig. 2 shows the distribution of the TMRCA estimated at G_{20} in the three regimes. Indeed, under HDNS, the distribution fits the expectation from Eq: (11). As compared to the neutral case, Fig. 2 also shows that both the high drift (HDHS) and low drift (LDHS) selection cases lead to reduced TMRCA, as expected.

We next assessed the impact of selection on N_e and compared different estimates, either based on TMRCA (Eq: (9)), or on the variance in offspring number (Eq: (10)), or on the cumulated response to selection (Eq: (8)). Values obtained are summarized in Tab. 2c. We found that in the absence of selection, N_e estimated from the mean TMRCA were close to the actual number of reproducing individuals (4.8 for HDNS and >10 for LDNS), while they were much smaller under both selection regimes (2.5 for HDHS and 3.3 for LDHS). The observed differences between $N_{e(20)}^{\text{Coal}}$ and the harmonic mean of $N_{e(g)}^{\text{Coal}}$ revealed a strong influence from a pedigree perspective, of the first generation on the adaptive dynamics. When N_e was estimated from the variance in offspring number, estimations without selection (4.1 under HDNS and 42 under LDNS) were close to the actual number of reproducing individuals, even for LDNS. Finally, N_e estimations from the cumulated response to selection fell within the same range as the ones from the variance in offspring number in both selection regimes. In summary, most N_e estimates were close to the actual number of reproducing individuals in the absence of selection. High selection strongly reduced N_e estimations, but merely in the low drift (LDHS) case. Indeed, N_e reduction due to selection is around 67%, depending on the estimations in the LDHS regime, while it is only around 50% in the HDHS regime.

Stochasticity in the response to selection: We addressed the qualitative nature of selection response focusing on its linearity. To do so, we measured in each family the average genetic gain per generation over 2000 simulations by fitting a linear regression model. The average genetic gain was 0.49 DTF/generation under HDHS, and 1.1 DTF/generation under LDHS (Tab. 2c). Associated $R^2 > 0.95$ indicated an accurate fit of the data to the linear model. Yet, large standard deviations around these estimates (0.2 and 0.27 for HDHS and LDHS, respectively) pointed either to high stochasticity or a non-linear response. Single simulations indicated non-linear response Fig. S3. Noteworthy, a strong response was observed between G_0 and G_1 (G_0G_{20} Fig. S3) with similar values in HDHS and LDHS, around 1.6 DTF/generation (Tab. 2c). Subsequently, simulations displayed discontinuities with abrupt changes of slopes at some generations, a signal compatible with the fixation of new mutations (Fig. S3). In order to characterize such discontinuities, we fitted a linear segmentation regression on individual simulations from G_1 and onwards. We estimated the number of breakpoints (i.e. slope changes), the corresponding slopes, and the first and greatest slope based on AIC minimization (Durand *et al.* 2010). The first slope described an average gain of 0.59 DTF/generation in the HDHS regime, and almost twice (0.96 DTF/generation) in the LDHS regime (Tab. 2c). These values were lower than those observed in G_0G_{20}.

Those results are consistent with a G_0G_{20} response resulting from the recruitment of initial genetic variance, independently of the population size, and a later response based on mutational variance being less effective in small than in large populations. To confirm those results, we performed a principal component analysis (PCA) and explored correlations between input parameters: initial additive genetic variance $\sigma^2_{A_0}$, mutational variance σ^2_M and residual variance σ^2_E, and descriptors of the response to selection : G_0G_{20}, number of breakpoints, first slope and greatest slope. In line with our interpretation, irrespective of the selection regime, $\sigma^2_{A_0}$ positively correlated with G_0G_{20}, and σ^2_M positively correlated with the first (after G1) and greatest slope (Fig. S4). Note that this stochastic process of mutation occurrence and fixation resulted in large differences among replicates, as illustrated by the breadth of the response (shaded areas in Fig. 1).

Evolution of genetic diversity: Because of the well established role of standing variation in selection response, we focused on its temporal dynamics. Standing variation in our experiment consisted in residual heterozygosity found in the initial inbred lines. Starting with a mean residual heterozygosity of 3.0×10^{-3} at G_0 (Tab. 2c), we observed a consistent decrease throughout selfing generations until the mutation-drift-equilibrium was reached (Fig. S5). Without selection the mean values reached $\approx 7.0 \times 10^{-4}$ at G_{20}. Considering a haploid Wright-Fisher population and the infinite allele model, the neutral prediction of equilibrium heterozygosity is $E(H) = \frac{\theta}{1+\theta}$ with $\theta = 2N_e\mu$. This translated into an estimate of $N_e = 1.94$ close to the values we obtained from the harmonic mean of N_e estimated from TMRCA (Tab. 2c).

Concerning the number of polymorphic loci, a mutation-

drift-equilibrium was reached in all cases except for the LDNS selection regime (Fig. S6). The equilibrium value depended on the census population size: around 6 polymorphic loci with high drift (HDHS and HDNS), 40 polymorphic loci under LDHS, and $>$ 66 polymorphic loci after 20 generations under LDNS (Tab. 2c (c) and Fig. S6). Altogether, our results show that the mean heterozygosity was affected neither by drift, nor by selection, but instead by the mutation rate. On the contrary, the number of polymorphic loci depended on the census population size.

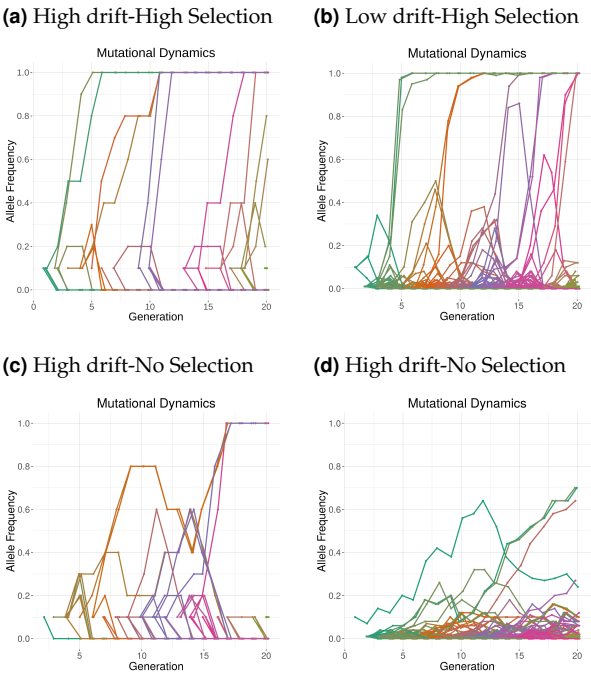

Figure 3 Evolution of allele frequencies within families under four simulated regimes. Examples of mutational fates are given for HDHS (a), LDHS (b), HDNS (c), LDNS (d). Mutations are recorded only when occurring in one of the selected progenitors, and corresponding frequencies are computed over all selected individuals. For example under High Drift regimes, the initial frequency of a mutation occurring in any given progenitor within a family is $1 \div (2 \times 5)$ as 5 diploid individuals are selected at each generation. Under Lower Drift regimes, the mutation initial frequency equals $1 \div (2 \times 50)$.

The dynamics of de novo mutations: Evolution of frequencies of new mutations revealed three fates: fixation, loss, and rare replacement by incoming mutation at the same locus. The four regimes strikingly differed in their mutational dynamics (Fig. 3). Under HDHS, most mutations quickly reached fixation (3.8 generations), with an average of 7.7 fixed mutations/population in 20 generations (Tab. 2c). The corresponding Low Drift regime (LDHS) displayed longer fixation time 5.9 generations, and an average of 10 fixed mutations/family (Tab. 2c). No selection regimes tended to exhibit a depleted number of fixed mutations, with no fixation under LDNS after 20 generations. Variation around the mean fixation time was substantial across all regimes Fig. S7. In sum, HDHS was characterized by the fast fixation of new mutations, 53% of which were fixed within 2 to 3 generations which contrasted to 15% under LDHS or 17% under HDNS. Selection therefore increased the number of fixed mutations while decreasing their fixation time.

Effects of mutations:

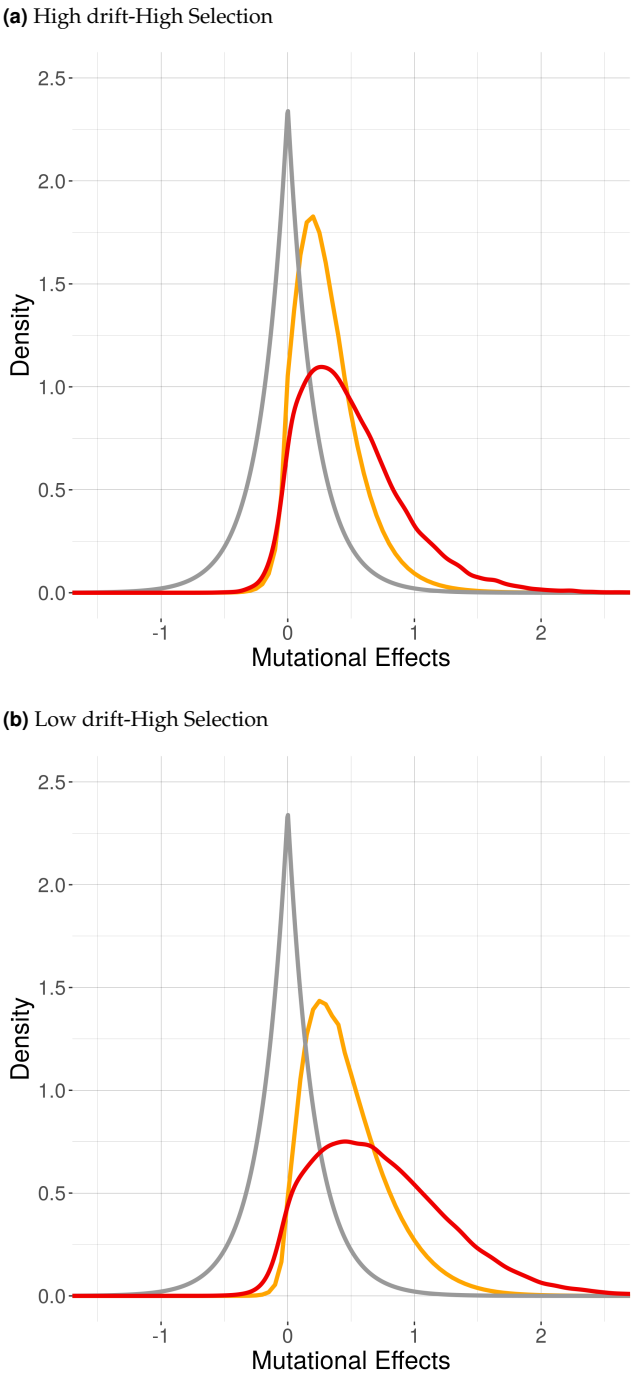

Figure 4 Distribution of effects of incoming and fixed mutations under High Selection regimes. Density distributions for the HDHS (a) and the LDHS (b) regime are shown for all incoming mutational effects in grey — reflected exponential distribution —, and fixed mutations over 2000 simulations in red. Theoretical expectations from (Eq: 16) are plotted in gold.

Beyond fixation time, a key aspect of our work was to investigate the impact of drift and selection on the type of fixed mutations, best summarized by their genotypic effects. In order

to do so, we compared the distribution of incoming mutations to that of fixed mutations. We evidenced a strong depletion of deleterious mutations together with a striking enrichment in beneficial mutations under the two Selection regimes, HDHS (quantile 5%=-0.02, median value=0.43, 95% quantile=1.3) and LDHS (quantile 5%=0.011, median value=0.66, 95% quantile=1.7) (Fig: 4). We also derived a theoretical expectation from Kimura's allele fixation probability using the selection coefficient computed in the case of truncation selection (Eq: 16). Accounting for the specificities of our selection procedure we found under both selection regimes, a slight excess of detrimental mutations, and a large excess of beneficial mutations as compared to Kimura's predictions. Note however that, comparatively, the excess of detrimental mutations was reduced under HDHS than under LDHS (Fig: 4).

As expected, selection generated a relation between the average size of a mutation and its time to fixation : the higher the effect of the mutation, the lower the time to fixation (Fig. 5 (a) and Fig. 5 (b)). Comparison between HDHS and LDHS revealed interesting features: under high drift, the average effect of mutations fixed was lower and variance around mutational effects tended to decrease correlatively with fixation time so that large size mutations were all fixed during the first generation while they persisted at subsequent generations under Low Drift (Fig. 5 (a) and Fig. 5 (b)).

In sum, our two selection regimes lead to an enrichment of beneficial mutations. Compared with LDHS, HDHS regime fixed fewer detrimental mutations but the average effect of fixed beneficial mutations was smaller.

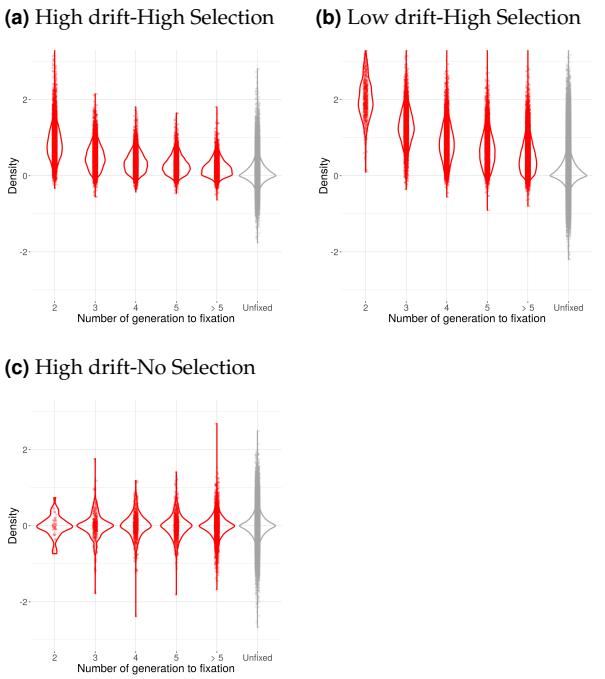

Figure 5 Violin plots of raw mutational effects according to fixation time under three simulated regimes. Plots are indicated for fixed (red) and lost (grey) mutations under HDHS (a), LDHS (b) and HDNS (c). Note that under LDNS, we obtained very few fixed mutations so that we were unable to draw the corresponding distribution.

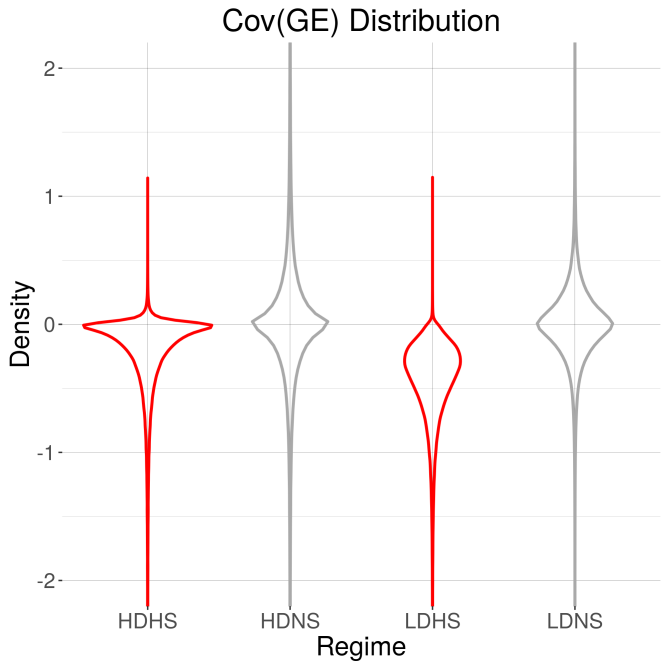

Figure 6 Violin plots of $\mathrm{Cov}(G_{|\mathrm{selected}}, E_{|\mathrm{selected}})$ under four simulated regimes. Violin plots were computed over 2000 simulations and 4 families, four families and across all generations under regimes with High Selection in red (HDHS, LDHS), and regimes with No Selection in grey (HDNS, LDNS).

Covariation between mutational and environmental effects A puzzling observation was that normalizing raw mutational effects by the environmental standard deviation translated into a distortion of the distribution so that the median value of fixed effects increased by 0.3 (from 0.4 to 0.72) under HDHS and by 0.2 under LDHS (Tab. 2 and Fig. S8). Similarly, 95% quantile increased by 1.2 (from 1.3 to 2.5) under HDHS and 0.66 (from 1.7 to 2.4) under LDHS. Hence, normalization distortion resulted in much more similar fixed mutations effects distribution under HDHS and HDNS. This was due to a non-zero negative genetic-environment covariance in selected individuals. Indeed, conditioning on the subset of selected individual, we obtained negative estimate of $\mathrm{Cov}(G_{|\mathrm{selected}}, E_{|\mathrm{selected}})$ both under HDHS and LDHS, with a median value (resp. 5% and 95% quantile) of -0.11 (resp. -0.88 and 0.029) under HDHS and -0.37 (resp. -1.1 and -0.072) under LDHS. In contrast, with no selection, values of randomly chosen individual $\mathrm{Cov}(G_{|\mathrm{random}}, E_{|\mathrm{random}})$ were centered around 0 as expected. The evolution of $\mathrm{Cov}(G_{|\mathrm{selected}}, E_{|\mathrm{selected}})$ through time (Fig. S9) evidenced a high stochasticity among generations but no temporal autocorrelation Fig. S9. In other words, this means that because of the negative correlation between residual environmental effects and genetic effects induced by selection, mutational effects depend on the environment at the generation they have been selected for.

Discussion

Population and quantitative genetics provide theoretical frameworks to investigate selection responses and underlying multilocus adaptive dynamics. Here, we focused on Saclay DSEs which were specifically designed to depict the evolutionary mechanisms behind the response to selection of a highly complex trait

Table 2 Descriptive statistics of the selection response dynamics in observed F252 genetic background (a), observed MBS genetic background (b) and the 4 simulated regimes (c).

(a) HDHS observed in F252 genetic background

F252 families:	FE1	FE2	FVL (G13)	FL2.1	FL2.2
Cumul. Resp. in DTF	-4.27	-5.34	11.32	3.19	2.60
Linear Regression Coefficient (SD)	-0.21 (0.048)	-0.22 (0.037)	0.86 (0.17)	0.11 (0.04)	0.12 (0.035)
Adjusted R-squared	0.49	0.63	0.65	0.23	0.34
Linear regression p-value	0.000269	1.074 e-05	0.000305	0.016	0.00353

(b) HDHS observed in MBS genetic background

MBS families:	ME1	ME2	ML1	ML2
Cumul. Resp. in DTF	-9.34	-11.72	8.64	11.05
Linear Regression Coefficient (SD)	-0.41 (0.03)	-0.42 (0.04)	0.24 (0.05)	0.46 (0.06)
Adjusted R-squared	0.89	0.84	0.57	0.76
Linear regression p-value	1.01 e-10	3.52 e-09	4.46 e-05	1.56 e-07

(c) Simulated regimes [a]

Simulated regimes:	HDHS	HDNS	LDHS	LDNS
Simulated Cumul. Resp.	13 (5.2)	1.7 (1.7)	24 (6.2)	1.3 (1.3)
Linear Regression Coefficient (SD)	0.49 (0.2)	0.067 (0.066)	1.1 (0.27)	0.035 (0.035)
R2 Linear Response	0.95	0.47	0.99	0.44
G_0G_{20} Response (SD)	1.6 (1.9)	0.26 (0.38)	1.7 (2.1)	0.3 (0.42)
First Slope (SD)	0.59 (0.52)	0.2 (0.3)	0.96 (0.46)	0.067 (0.091)
Greatest Slope (SD)	0.56 (0.49)	0.16 (0.24)	1 (0.46)	0.08 (0.098)
G_{20} TMRCA (SD)	3.9 (1.9)	7.6 (4.3)	6.4 (2.8)	$>$ 20 (0.098) [b]
$N_{e(G20)}^{Coal}$ (Ne from G_{20} TMRCA) (SD)	2.5 (1.2)	4.8 (2.7)	3.3 (1.4)	$>$ 10 (0.05) [b]
$N_{e(G1-20)}^{Coal}$ Ne (Harmonic Ne from all TMRCA) (SD)	1.8 (0.22)	2.5 (0.37)	2 (0.22)	2.8 (0.00096)
$N_{e(G1-20)}^{Var(o)}$ (Harmonic Ne from Var Off) (SD)	3 (0.44)	4.1 (0.6)	16 (3)	42 (4.6)
Ne required for the Simulated Cumul. Resp. (SD)	3.3 (2.0)	0.4 (0.4)	9.0 (21.1)	0.3 (0.3)
Heterozygosity at G_0 (SD)	0.003 (0.003)	0.003 (0.004)	0.003 (0.003)	0.003 (0.004)
Heterozygosity at G_{20} (SD)	0.00083 (0.00047)	0.00073 (0.00043)	0.00087 (0.00021)	0.00072 (0.00014)
Number Of Polymorphism at G_0 (SD)	3 (3)	3.2 (4.1)	3.1 (3.3)	3.3 (4.1)
Number Of Polymorphism at G_{20} (SD)	5.1 (2.5)	6 (3.1)	40 (8.8)	66 (9.2)
Simulation Fraction Without Any Fixed Mutation	0	0.0025	0	0.999 [c]
Fixation Time in generations (SD)	3.8 (1.7)	7.2 (3)	5.9 (2.3)	NA [c]
Number Of Fixed Mutation Per Family (SD)	7.7 (2.6)	2.3 (1.4)	10 (3)	NA [c]
Q5 Fixed Mut. effect (Non Normalized)	-0.019	-0.5	0.011	NA [c]
Q50 Fixed Mut. effect (Non Normalized)	0.43	-0.00086	0.66	NA [c]
Q95 Fixed Mut. effect (Non Normalized)	1.3	0.5	1.7	NA [c]
Q5 Fixed Mut. effect (Normalized)	-0.034	-0.44	0.015	NA [c]
Q50 Fixed Mut. effect (Normalized)	0.72	-0.00054	0.86	NA [c]
Q95 Fixed Mut. effect (Normalized)	2.5	0.42	2.4	NA [c]
CovGE (SD)	-0.23 (0.35)	-0.0035 (0.68)	-0.45 (0.34)	-0.0011 (0.36)
Q5 CovGE	-0.88	-0.87	-1.1	-0.5
Q50 CovGE	-0.11	0	-0.37	-0.00017
Q95 CovGE	0.029	0.86	-0.072	0.49

[a] All values are computed as the mean (resp. SD or quantile, when indicated) over 2000 simulations.

[b] Under LDNS, we expected a neutral coalescent time around 98 generations well beyond the 20 simulated generations, which provided highly biased G_{20} TMRCA and N_e estimators.

[c] Under LDNS, we obtained very few fixed mutation so that we were unable to compute the corresponding statistics.

— with a high mutational target — in small populations evolving under truncation selection (1% of selected individual), limited recombination (total selfing regime) and limited standing variation. Our main motivation was to explore how such a combination of unusual conditions, at the limits of parameters boundaries of classic models, can sustain the long-term maintenance of additive genetic variation and a significant selection response with no observed load (annual field observations). In this purpose we devised forward individual-based simulations that explicitly modeled our Saclay DSEs, and relied on theoretical predictions to investigate the interplay of evolutionary forces and patterns associated with fixation of mutations.

Mutation accounts for the maintenance of small but sufficient additive genetic variation The three determinants of the observed selection response were best summarized by three variance components namely, the initial additive variance σ^2_{A0}, the environmental variance σ^2_E), and the mutational variance σ^2_M (Fig: S4). Quantitatively, we demonstrated the importance of both initial standing variation and the necessity of a constant mutational input to explain the significant selection response in the two Saclay DSEs (Fig: 1 & S4). This result was consistent with previous reports (Durand *et al.* 2010, 2015) and showed that the first selection response between G_0 and G_1 was correlated with σ^2_{A0}, while response in subsequent generations was correlated with σ^2_M (Fig: S4). In our simulations, we chose initial values for variance components that closely matched previous estimates in the Saclay DSE derived from the MBS inbred line (Durand *et al.* 2010). The small value for initial additive variance came from the use of commercial inbred liens in our experimental evolution setting. It sharply contrasted with more traditional settings where distant genetic material and crosses are often performed to form an initial panmictic population from which selection is applied (Kawecki *et al.* 2012). While crucial in the first generation (Fig: S4), σ^2_{A0} was quickly exhausted. The long-term selection response was sustained by $\mathbb{E}(\sigma^2_M) = 3.38 \times 10^{-2}$ which corresponded to an expected mutational heritability of $\mathbb{E}(\frac{\sigma^2_M}{\sigma^2_E}) \approx \frac{\mathbb{E}(\sigma^2_M)}{\mathbb{E}(\sigma^2_E)} = 1.5 \times 10^{-2}$ (in units of residual variance per generation). These values stand as higher bounds to what was previously described in other species/complex traits (Keightley 2010; Walsh and Lynch 2018).

We further implemented an additive incremental mutation model (Clayton and Robertson 1955; Kimura 1965; Walsh and Lynch 2018). This model assumed non-limiting mutational inputs, and has been shown to be particularly relevant in systems where, just like ours, effective recombination is limited (Charlesworth 1993; Walsh and Lynch 2018). Alternative non-additive model such as the House Of Cards (HoC) that sets random allelic effect upon occurrence of a new allele (Kingman 1978; Turelli 1984) — rather than adding effects incrementally — would have likely resulted in smaller estimate of σ^2_M (Hodgins-Davis *et al.* 2015). Whether the incremental model or the HoC or a combination of both such as the regression mutation model Zeng and Cockerham (1993) was better suited to mimic our Saclay DSEs is an open question. However several lines of evidence argue in favour of a non-limiting mutational input in our setting. First, the architecture of maize flowering time is dominated by a myriad of QTLs of small additive effects (Buckler *et al.* 2009). Over 100 QTLs have been detected across maize lines (Buckler *et al.* 2009), and over 1000 genes have been shown to be involved in its control in a diverse set of landraces (Romero Navarro *et al.* 2017). Second, in Saclay DSEs alone, transcriptomic analysis of apical meristem tissues has detected 2,451 genes involved in the response to selection between early and late genotypes, some of which being interconnected within the complex gene network that determines the timing of floral transition (Tenaillon *et al.* 2018). This suggests that not only the number of loci is considerable, but also that their connection within a network further enhances the number of genetic combinations, and in turn, the associated phenotypic landscape. The breadth of the mutational target is a key parameter for adaptation (Höllinger *et al.* 2019). Together, our results suggest that our large mutational target compensates for the small population sizes, and triggers the long-term maintenance of heterozygosity, and genetic diversity at the population level after the selection-drift-mutation equilibrium is reached, i.e. after three to five generations. Noteworthy the expected level of heterozygosity in our controls (No Selection models, NS) corresponded to neutral predictions (Crow and Maruyama 1971; Kimura 1969).

Quick fixation of *de novo* mutations drive Saclay DSEs selection response The observed fixation time of mutations without selection is expected under standard neutral theory. The Kingman coalescent indeed predicts a TMRCA around 8 generations for a population size of 5 which matched closely our observed value of 7.6 obtained under HDNS. With selection, instead, we observed a quick fixation of mutations in three to four generations under HDHS. Likewise, the number of fixed mutation increased from 2.3 in HDNS to 7.7 in HDHS (Tab. 2). Note that while one would expect emerging patterns of hard sweeps following such rapid mutation fixation, our selfing regime which translated into small effective recombination likely limits considerably genetic hitchhiking footprints, so that such patterns may be hardly detectable.

Short fixation times made the estimate of effective population sizes challenging. We used two estimates of N_e to shed light on different processes entailed in HDHS stochastic regime. These estimates were based on expected TMRCA and on the variance in the number of offspring (Crow and Kimura 1971), respectively. We found the latter to be greater than the former. This can be explained by the fact that selection is known to substantially decrease effective population on quantitative trait submitted to continuous selection, because part of the selective advantage of an individuals accumulates in offsprings over generations (Santiago and Caballero 1995), and because selection on the phenotypic value acts in parts on non-heritable variance (*i.e.*, on the environmental variance component of V_P (Chantepie and Chevin 2020)). Note that estimate of N_e based on the known genealogical structure allow to compute a "realized" estimate that accounts for these effects. However, this is not without drawback, as TMRCA were much shorter than expected, a result consistent with the occurrence of multiple merging along pedigrees, *i.e.* multiple individuals coalescing into a single progenitor. Multiple merger coalescence may actually be better suited to describe rapid adaptation than the Kingman coalescent (Neher 2013).

Both fixation time and probability depend on the selection coefficient s and the initial frequency of the mutation in the population. In our setting, conditioning on its appearance in the subset of selected individuals, the initial frequency of a mutation was 0.10, which was unusually high and translated into selection and drift exerting greater control over mutations. Indeed, in more traditional drift regimes, even when an allele is strongly selected ($2N_e s \gg 1$), drift dominates at mutation occurrence,

i.e. with two absorbing states for allele frequency near zero and one (Walsh and Lynch 2018). In other words, in HDHS regime, selection induced repeated population bottlenecks so that it can not be decoupled from drift.

High stochasticity promotes the fixation of small effect mutations Interplay between drift and selection promoted stochasticity in our setting, which manifested itself in various ways : (i) through the selection response, with different families exhibiting contrasting behaviors, some responding very strongly and others not, Fig. 1; (ii) through the dynamics of allele fixation (Fig. 2 & 3); and (iii) through the distribution of $\mathrm{Cov}(GE)$ Fig. 6. Stochasticity tightly depends on census population size (Hill 1982a,b). Unexpectedly, however, we found a benefice to stochasticity as illustrated by a bias towards the fixation of advantageous mutations compared with the expectation (Fig. 4). Comparison of the distributions of the mutational raw effects indicated that, among advantageous mutations, a greater proportion of those with small effects were fixed under the High Drift than under the Low Drift regime (Fig. 4 (a) versus (b)). This result echoes those of Silander *et al.* (2007), who showed — using experimental evolution with bacteriophage — that fitness declines down to a plateau in populations where drift overpower selection. The authors note: "If all mutations were of small effect, they should be immune to selection in small populations. This was not observed; both deleterious and beneficial mutations were subject to selective forces, even in the smallest of the populations."

What are the underlying mechanism behind this fixation bias? We found a negative covariance between selected genotypes and their corresponding environmental values, that modified the mutational effect to an apparent mutational effect perceived by the environment. The negative $\mathrm{Cov}(GE)$ arose mechanically from selection of two independent random variables, whatever the sampling size as illustrated in Fig. S10 and Fig. S11. This effect reminds the so-called Bulmer effect (Bulmer 1971), that causes a reduction of genetic variance due to the effect of selection on the covariance between unlinked loci. Interestingly, under the High Drift regime, we observed a less negative $\mathrm{Cov}(GE)$ on average than with a 10 times higher census size (Low Drift). This translated, after dividing by the environmental standard deviation, to a greater apparent effect of small mutations under the High Drift regime. In other words, High Drift-High Selection tends to magnify mutational effects from an environmental perspective. In support of this explanation, normalization by the environmental standard deviation actually erased the difference between the two distributions of mutational effect (under low and High Drift, Fig. S8). Unlike the Bulmer effect however, this one was restricted to the generation of mutation occurrence, but favored long-term fixation of slightly advantageous mutations by a transient increase of their frequency. Because of a significant variance of $\mathrm{Cov}(GE)$, this effect on small effect mutation fixation was mostly stochastic. Therefore, we interpreted the fixation of a high proportion of slightly beneficial mutations, and their significant contribution to selection response, by the less efficient exploration of the initial distribution per simulation (increasing their prevalence) but the stochastic "help" of a lesser negative $\mathrm{Cov}(GE)$.

Deficit of fixation of deleterious mutations suggests a limited cost of selection As expected, we observed that selection decreased the number of segregating polymorphic loci at equilibrium compared to regimes without selection (Tab. 2). Interestingly however, this effect was reduced for small population size. Under High Drift, selection induced an average loss of a single polymorphism at equilibrium (HDHS vs. HDNS, Tab. 2) while under the Low Drift regime over 20 polymorphisms were lost (LDHS vs. LDNS, Tab. 2). A similar trend was recovered at the mutation fixation level where on average 7.7 mutations were fixed under the High Drift-High Selection and only 10 under Low Drift-High Selection. In other words, the 10-fold population increase did not translate into a corresponding increase in the number of segregating and fixed mutations, as if there was a diminishing cost with decreasing population size. Under High Drift (resp. Low Drift), at each generation 500 (resp. 5000) offspring of 2×1000 loci were produced. Considering a mutation rate per locus of $6000 \times 30 \times 10^{-9}$, (*i.e.* (Clark *et al.* 2005)), it translated into 180 mutations events (resp. 1800 mutations events). However most mutations are lost as only mutations occurring in the subset of selected individuals survive. The initial frequency of a mutation in this subset, *i.e* of size 5 or 50, is $\frac{1}{10}$ under High Drift and $\frac{1}{100}$ under Low Drift. In the former, the interplay between the initial frequency and selection intensity allows a better retention of beneficial mutations of small effect (Fig. 4) than in the latter. Interestingly at equilibrium, we also observed a higher level of residual heterozygosity with selection than without, irrespective of population size, suggesting a small impact of selection in the long-term heterozygosity maintenance. Overall, our High Drift-High Selection regime maintains a small, but sufficient number of polymorphisms for the selection response to be significant.

Our selection response evidenced a deficit of fixation of deleterious mutations and hence a modest genetic load (Fig. 4 and S8). We identified three reasons behind this observation. Firstly, in our design, the selection intensity of 1% was applied on the trait. Hence, in contrast to the infinitesimal model for which a high number of polymorphic loci are expected to individually experience a small selection intensity, selection intensity was "concentrated" here on a restricted number of loci, i.e. those for which polymorphisms were segregating. Secondly, we applied truncation selection whose efficiency has been demonstrated (Crow and Kimura 1979). The authors noted: "It is shown, for mutations affecting viability in Drosophila, that truncation selection or reasonable departures therefrom can reduce the mutation load greatly. This may be one way to reconcile the very high mutation rate of such genes with a small mutation load." Thirdly, the lack of interference between selected loci in our selection regime may further diminish the selection cost (Hill and Robertson 1966). Reduced interference in our system is indeed expected from reduced initial diversity and quick fixation of *de novo* mutations. Whether natural selection proceeds through truncation selection or Gaussian selection is still a matter of debate (Crow and Kimura 1979). Measuring the impact of these two types of selection on the genealogical structure of small populations including on the prevalence of multiples mergers will be of great interest to better predict their fate.

This under-representation of deleterious variant echoes with empirical evidence that in crops, elite lines are impoverished in deleterious variants compared to landraces owing to a recent strong selection for yield increase (Gaut *et al.* 2015). Likewise, no difference in terms of deleterious variant composition were found between sunflower landraces and elite lines (Renaut and Rieseberg 2015). Hence, while the dominant consensus is that the domestication was accompanied by a genetic cost linked to the combined effects of bottlenecks, limited effective recombination reducing selection efficiency, and deleterious allele surfing by rapid population expansion (Moyers *et al.* 2018), recent breeding

highlights a distinct pattern. We argue that our results may help to understand this difference because under High Drift-High Selection, a regime likely prevalent in modern breeding, genetic load is reduced. More generally, our results may provide useful hints to explain the evolutionary potential of selfing populations located at the range margins. Just like ours, such populations are generally small, display both, inbreeding and reduced reduced standing variation (Pujol and Pannell 2008) and are subjected environmental and demographic stochasticity.

Conclusion In conclusion, our High Drift-High Selection regime with non-limiting mutation highlights an interesting interplay between drift and selection that together promote the quick fixation of adaptive *de novo* mutations fueling a significant but stochastic selection response. Interestingly, such selection response is not impeded by the fixation of deleterious mutations but displays instead a limited cost. Our results provide an explanation for patterns highlighted during recent breeding as well as the high colonization ability of small selfing populations located at species range margins. They also call for a better mathematical description of the multilocus adaptive process sustaining the evolution of small populations under intense selection.

Literature Cited

Berg, J. J. and G. Coop, 2014 A Population Genetic Signal of Polygenic Adaptation. PLoS Genetics **10**: e1004412.

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, *et al.*, 2009 The genetic architecture of maize flowering time. Science **325**: 714–718.

Bulmer, M. G., 1971 The Effect of Selection on Genetic Variability. The American Naturalist **105**: 201–211.

Burke, M. K., 2012 How does adaptation sweep through the genome? Insights from long-term selection experiments. Proceedings of the Royal Society B: Biological Sciences **279**: 5029–5038.

Burke, M. K., G. Liti, and A. D. Long, 2014 Standing genetic variation drives repeatable experimental evolution in outcrossing populations of saccharomyces cerevisiae. Molecular Biology and Evolution **31**: 3228–3239.

Caballero, A., M. A. Toro, and C. Lopez-Fanjul, 1991 The response to artificial selection from new mutations in Drosophila melanogaster. Genetics **128**: 89–102.

Carr, R. N. and R. F. Nassar, 1970 Effects of Selection and Drift on the Dynamics of Finite Populations. I. Ultimate Probability of Fixation of a Favorable Allele. Biometrics **26**: 41.

Chantepie, S. and L. Chevin, 2020 How does the strength of selection influence genetic correlations? Evolution Letters **4**: 468–478.

Charlesworth, B., 1993 Directional selection and the evolution of sex and recombination. Genetical Research **61**: 205–224.

Chevin, L. M. and F. Hospital, 2008 Selective sweep at a quantitative trait locus in the presence of background genetic variation. Genetics **180**: 1645–1660.

Clark, R. M., S. Tavaré, and J. Doebley, 2005 Estimating a nucleotide substitution rate for maize from polymorphism at a major domestication locus. Molecular Biology and Evolution **22**: 2304–2312.

Clayton, G. and A. Robertson, 1955 Mutation and Quantitative Variation. The American Naturalist **89**: 151–158.

Crow, J. F. and M. Kimura, 1971 *An Introduction to Population Genetics Theory*, volume 26. Blackburn Press.

Crow, J. F. and M. Kimura, 1979 Efficiency of truncation selection. Proceedings of the National Academy of Sciences of the United States of America **76**: 396–399.

Crow, J. F. and T. Maruyama, 1971 The number of neutral alleles maintained in a finite, geographically structured population. Theoretical Population Biology **2**: 437–453.

De Leon, N. and J. G. Coors, 2002 Twenty-four cycles of mass selection for prolificacy in the Golden Glow maize population. Crop Science **42**: 325–333.

Desai, M. M. and D. S. Fisher, 2007 Beneficial mutation-selection balance and the effect of linkage on positive selection. Genetics **176**: 1759–1798.

Dudley, J. W. and R. J. Lambert, 2010 100 Generations of Selection for Oil and Protein in Corn. In *Plant Breeding Reviews*, volume 24, pp. 79–110, John Wiley and Sons, Inc., Oxford, UK.

Durand, E., S. Bouchet, P. Bertin, A. Ressayre, P. Jamin, *et al.*, 2012 Flowering time in maize: Linkage and epistasis at a major effect locus. Genetics **190**: 1547–1562.

Durand, E., M. I. Tenaillon, X. Raffoux, S. Thépot, M. Falque, *et al.*, 2015 Dearth of polymorphism associated with a sustained response to selection for flowering time in maize. BMC Evolutionary Biology **15**: 103.

Durand, E., M. I. Tenaillon, C. Ridel, D. Coubriche, P. Jamin, *et al.*, 2010 Standing variation and new mutations both contribute to a fast response to selection for flowering time in maize inbreds. BMC Evolutionary Biology **10**: 2.

Falconer, D. S., 1971 Improvement of litter size in a strain of mice at a selection limit. Genetical Research **17**: 215–235.

Fisher Ronald Aylmer, S., 1930 *The genetical theory of natural selection.*. OxfordClarendon Press.

Gaut, B. S., C. M. Díez, and P. L. Morrell, 2015 Genomics and the Contrasting Dynamics of Annual and Perennial Domestication. Trends in Genetics **31**: 709–719.

Gerrish, P. J. and R. E. Lenski, 1998 The fate of competing beneficial mutations in an asexual population. Genetica **102-103**: 127–144.

Good, B. H., M. J. McDonald, J. E. Barrick, R. E. Lenski, and M. M. Desai, 2017 The dynamics of molecular evolution over 60,000 generations. Nature **551**: 45–50.

Good, B. H., I. M. Rouzine, D. J. Balick, O. Hallatschek, and M. M. Desai, 2012 Distribution of fixed beneficial mutations and the rate of adaptation in asexual populations. Proceedings of the National Academy of Sciences of the United States of America **109**: 4950–4955.

Haldane, J. B. S., 1927 A Mathematical Theory of Natural and Artificial Selection, Part V: Selection and Mutation. Mathematical Proceedings of the Cambridge Philosophical Society **23**: 838–844.

Hartfield, M., T. Bataillon, and S. Glémin, 2017 The Evolutionary Interplay between Adaptation and Self-Fertilization. Trends in Genetics **33**: 420–431.

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

Hermisson, J. and P. S. Pennings, 2017 Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. Methods in Ecology and Evolution **8**: 700–716.

Hill, W. G., 1982a Predictions of response to artificial selection from new mutations. Genetical Research **40**: 255–278.

Hill, W. G., 1982b Rates of change in quantitative traits from fixation of new mutations. Proceedings of the National Academy

of Sciences of the United States of America **79**: 142–145.

Hill, W. G. and A. Caballero, 1992 Artificial Selection Experiments. Annual Review of Ecology and Systematics **23**: 287–310.

Hill, W. G. and J. Rasbash, 1986 Models of long-term artificial selection in finite population with recurrent mutation. Genetical Research **48**: 125–131.

Hill, W. G. and A. Robertson, 1966 The effect of linkage on limits to artificial selection. Genetical Research **8**: 269–294.

Hodgins-Davis, A., D. P. Rice, J. P. Townsend, and J. Novembre, 2015 Gene expression evolves under a house-of-cards model of stabilizing selection. Molecular Biology and Evolution **32**: 2130–2140.

Höllinger, I., P. S. Pennings, and J. Hermisson, 2019 Polygenic adaptation: From sweeps to subtle frequency shifts. PLoS Genetics **15**: e1008035.

Hospital, F. and C. Chevalet, 1996 Interactions of selection, linkage and drift in the dynamics of polygenic characters. Genetical Research **67**: 77–87.

Houle, D., 1989 The Maintenance of Polygenic Variation in Finite Populations. Evolution **43**: 1767.

Jiao, Y., P. Peluso, J. Shi, T. Liang, M. C. Stitzer, *et al.*, 2017 Improved maize reference genome with single-molecule technologies. Nature **546**: 524–527.

Kamran-Disfani, A. and A. F. Agrawal, 2014 Selfing, adaptation and background selection in finite populations. Journal of Evolutionary Biology **27**: 1360–1371.

Kassen, R. and T. Bataillon, 2006 Distribution of fitness effects among beneficial mutations before selection in experimental populations of bacteria. Nature Genetics **38**: 484–488.

Kawecki, T. J., R. E. Lenski, D. Ebert, B. Hollis, I. Olivieri, *et al.*, 2012 Experimental evolution. Trends in Ecology and Evolution **27**: 547–560.

Keightley, P. D., 2010 Mutational Variation and Long-Term Selection Response. In *Plant Breeding Reviews*, pp. 227–247, John Wiley and Sons, Inc., Oxford, UK.

Kimura, M., 1962 On the probability of fixation of mutant genes in a population. Genetics **47**: 713–719.

Kimura, M., 1965 A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proceedings of the National Academy of Sciences of the United States of America **54**: 731–736.

Kimura, M., 1969 The number of heterozygous nucleotide sites maintained in a finite population due to steady flux of mutations. Genetics **61**: 893–903.

Kimura, M., 1983 *The Neutral Theory of Molecular Evolution*, volume 54. Cambridge University Press.

Kingman, J. F. C., 1978 A simple model for the balance between selection and mutation. Journal of Applied Probability **15**: 1–12.

Lamkey, K., 1992 Fifty years of recurrent selection in the Iowa stiff stalk synthetic maize population. Maydica **37**: 19–28.

Lande, R., 1979 Quantitative Genetic Analysis of Multivariate Evolution, Applied to Brain: Body Size Allometry. Evolution **33**: 402.

Lande, R. and S. J. Arnold, 1983 The Measurement of Selection on Correlated Characters. Evolution **37**: 1210.

Lillie, M., C. F. Honaker, P. B. Siegel, and Ö. Carlborg, 2019 Bidirectional selection for body weight on standing genetic variation in a chicken model. G3: Genes, Genomes, Genetics **9**: 1165–1173.

Lush, J. L., 1943 Animal breeding plans. Animal breeding plans .

Mackay, T. F., 2010 Mutations and quantitative genetic variation: Lessons from Drosophila. Philosophical Transactions of the Royal Society B: Biological Sciences **365**: 1229–1239.

Messer, P. W. and D. A. Petrov, 2013 Population genomics of rapid adaptation by soft selective sweeps. Trends in Ecology and Evolution **28**: 659–669.

Moose, S. P., J. W. Dudley, and T. R. Rocheford, 2004 Maize selection passes the century mark: A unique resource for 21st century genomics. Trends in Plant Science **9**: 358–364.

Moyers, B. T., P. L. Morrell, and J. K. McKay, 2018 Genetic costs of domestication and improvement. Journal of Heredity **109**: 103–116.

Neher, R. A., 2013 Genetic Draft, Selective Interference, and Population Genetics of Rapid Adaptation. Annual Review of Ecology, Evolution, and Systematics **44**: 195–215.

Odhiambo, M. O. and W. A. Compton, 1987 Twenty Cycles of Divergent Mass Selection for Seed Size in Corn 1 . Crop Science **27**: 1113–1116.

Parent, B., O. Turc, Y. Gibon, M. Stitt, and F. Tardieu, 2010 Modelling temperature-compensated physiological rates, based on the co-ordination of responses to temperature of developmental processes. Journal of Experimental Botany **61**: 2057–2069.

Pujol, B. and J. R. Pannell, 2008 Reduced responses to selection after species range expansion. Science **321**: 96.

Renaut, S. and L. H. Rieseberg, 2015 The accumulation of deleterious mutations as a consequence of domestication and improvement in sunflowers and other compositae crops. Molecular Biology and Evolution **32**: 2273–2283.

Roberts, R. C., 1967 The limits to artificial selection for body weight in the mouse: IV. Sources of new genetic variance—irradiation and outcrossing. Genetical Research **9**: 87–98.

Robertson, A., 1960 A theory of limits in artificial selection. Proceedings of the Royal Society of London. Series B. Biological Sciences **153**: 234–249.

Romero Navarro, J. A., M. Willcox, J. Burgueño, C. Romay, K. Swarts, *et al.*, 2017 A study of allelic diversity underlying flowering-time adaptation in maize landraces. Nature Genetics **49**: 476–480.

Roze, D., 2016 Background selection in partially selfing populations. Genetics **203**: 937–957.

Santiago, E. and A. Caballero, 1995 Effective size of populations under selection. Genetics **139**: 1013–1030.

Silander, O. K., O. Tenaillon, and L. Chao, 2007 Understanding the evolutionary fate of finite populations: The dynamics of mutational effects. PLoS Biology **5**: 922–931.

Spor, A., D. J. Kvitek, T. Nidelet, J. Martin, J. Legrand, *et al.*, 2014 Phenotypic and genotypic convergences are influenced by historical contingency and environment in yeast. Evolution **68**: 772–790.

Stetter, M. G., K. Thornton, and J. Ross-Ibarra, 2018 Genetic architecture and selective sweeps after polygenic adaptation to distant trait optima. PLoS Genetics **14**: 1–24.

Tavaré, S., 1984 Line-of-descent and genealogical processes, and their applications in population genetics models. Theoretical Population Biology **26**: 119–164.

Tenaillon, M. I., K. Seddiki, M. Mollion, M. L. Guilloux, E. Marchadier, *et al.*, 2018 Transcriptomic response to divergent selection for flowering times reveals convergence and key players of the underlying gene regulatory network. bioRxiv **7**: 461947.

Turelli, M., 1984 Heritable genetic variation via mutation-selection balance: Lerch's zeta meets the abdominal bristle.

Theoretical Population Biology **25**: 138–193.

Walsh, B. and M. Lynch, 2018 *Evolution and Selection of Quantitative Traits*, volume 1. Oxford University Press.

Weber, K. E., 1990 Increased selection response in larger populations. I. Selection for wing-tip height in Drosophila melanogaster at three population sizes. Genetics **125**: 579–84.

Weber, K. E., 1996 Large genetic change at small fitness cost in large populations of Drosophila melanogaster selected for wind tunnel flight: rethinking fitness surfaces. Genetics **144**: 205–13.

Weber, K. E. and L. T. Diggins, 1990 Increased selection response in larger populations. II. Selection for ethanol vapor resistance in Drosophila melanogaster at two population sizes. Genetics **125**: 585–597.

Wei, M., A. Caballero, and W. G. Hill, 1996 Selection response in finite populations. Genetics **144**: 1961–1974.

Wellenreuther, M. and B. Hansson, 2016 Detecting Polygenic Evolution: Problems, Pitfalls, and Promises. Trends in Genetics **32**: 155–164.

Zeng, Z. B. and C. C. Cockerham, 1993 Mutation models and quantitative genetic variation. Genetics **133**: 729–736.

Supplementary material

Saclay DSE's selection scheme

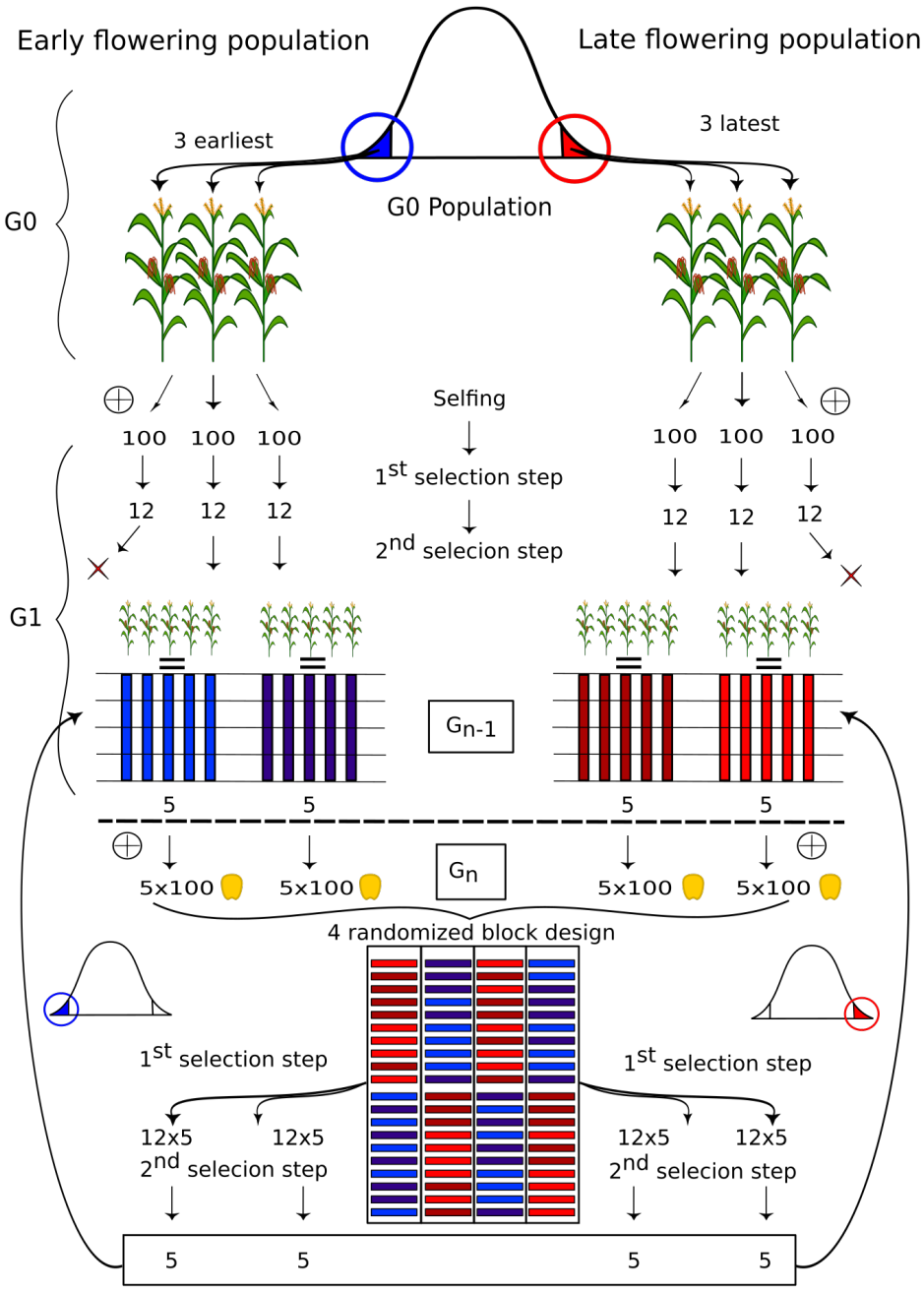

Figure S1 Experimental scheme of Saclay DSEs. For clarity a single scheme is shown but was replicated for the two DSEs. Starting from an inbred G_0 population with little standing variation ($<$ 1% residual heterozygosity (Durand *et al.* 2015)), the three earliest (resp. latest) flowering individuals represented in blue (resp. red) were chosen based on their offspring phenotypic values as the founders of two families forming the early (resp. late) population. For the subsequent generations, 10 ($\approx$ 5 per family) extreme progenitors were selected in a two step selection scheme among 1000 plants. More specifically, 100 seeds per progenitor were evaluated in a four randomized-block design, *i.e.* 25 seeds per block in a single row. In a first selection step, the $3 \times 4 = 12$ earliest (resp. latest) flowering plants among the 100 plants per progenitor were selected in a first step. Then in a second selection step, 10 (≈ 5 per family) individuals were selected within each population based on both flowering time and kernel weight and the additional condition of preserving two progenitors per family from the previous generation.

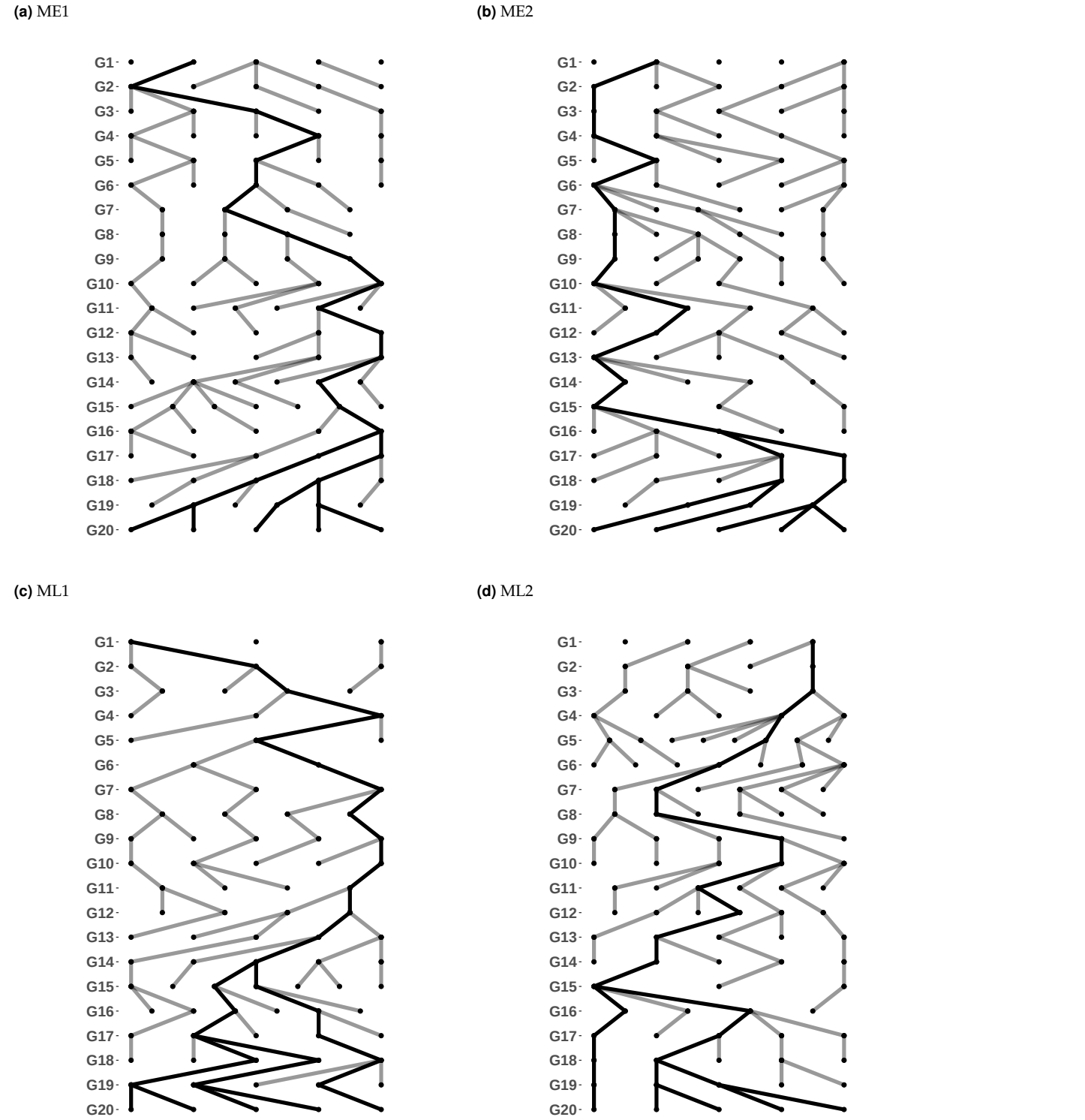

Figure S2 MBS family pedigrees from G1 to G_{20}. The two early families ME1 (a) and ME2 (b), and the two late families ML1 (c) and ML2 (d) are presented. Each node corresponds to a progenitor selected at a given generation. Each edge corresponds to a filial relationship between a progenitor and its offspring. Thick black lines indicate the ancestral path of the last generation (G_{20}).

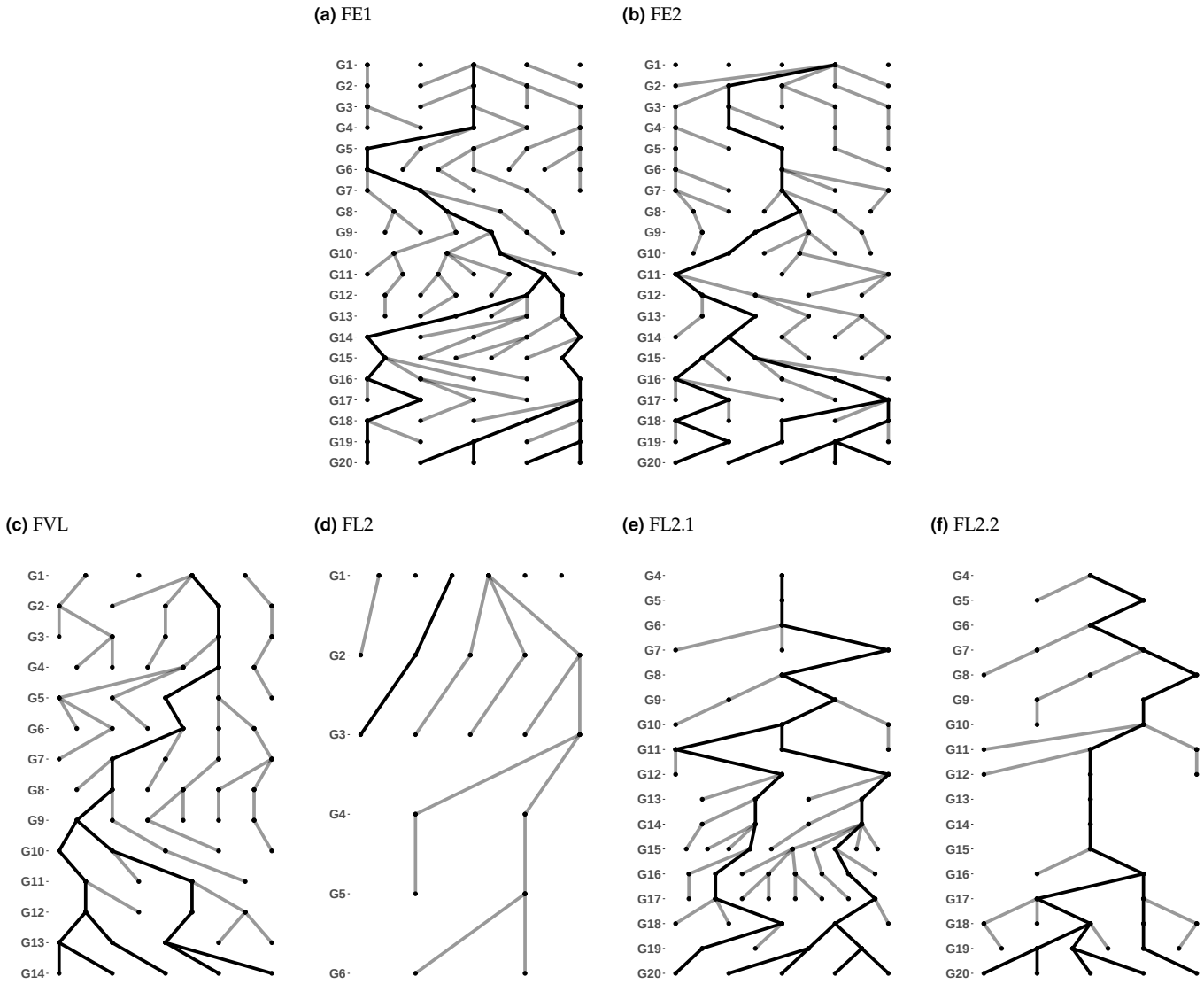

Figure S2 (Continued) F252 family pedigrees from G1 to G_{20}. Two early families FE1 (a), FE2 (b) and two late families FVL (c) & FL2 (d), are represented. FVL (c) could not be maintained after G14 as flowering occurred too late in the season for seed production. Both FL2.1 (e) and FL2.2 (f) were derived from a same individual from FL2 (d) at G3, after FVL was discarded. Each node corresponds to a progenitor selected at a given generation. Each edge corresponds to a filial relationship between a progenitor and its offspring. Thick black lines indicate the ancestral path of the last generation. (G_{20})

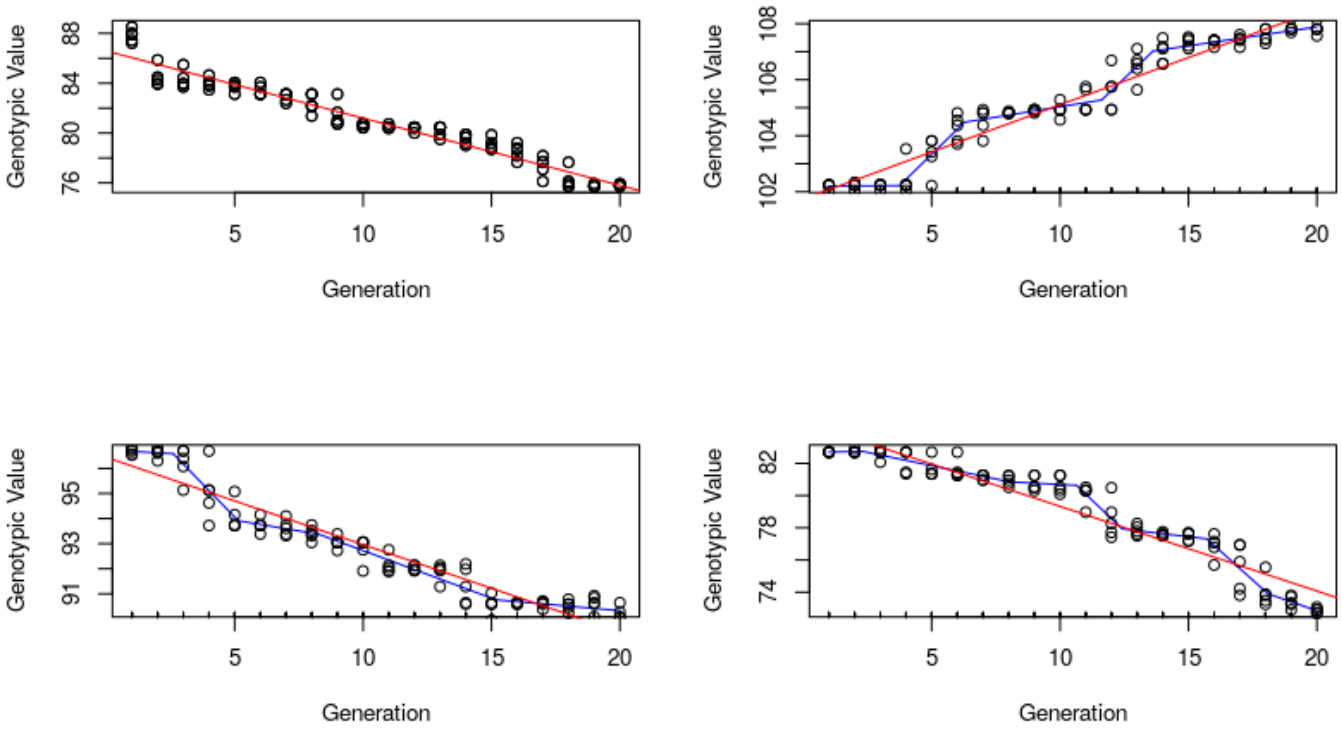

Figure S3 Illustration of simulated non-linear selection response in MBS. Each panel presents the evolution through time (x axis) of the genotypic value (y axis) of the 5 selected individual per family (empty dots). The red lines shows the linear regression of the selected genotypic values through times, while blue lines correspond to the best (AIC criterion) segmented linear model. The top left panel is an example for which a simple linear model fitted best the selection response, while the three others show a diversity of non-linear behaviors.

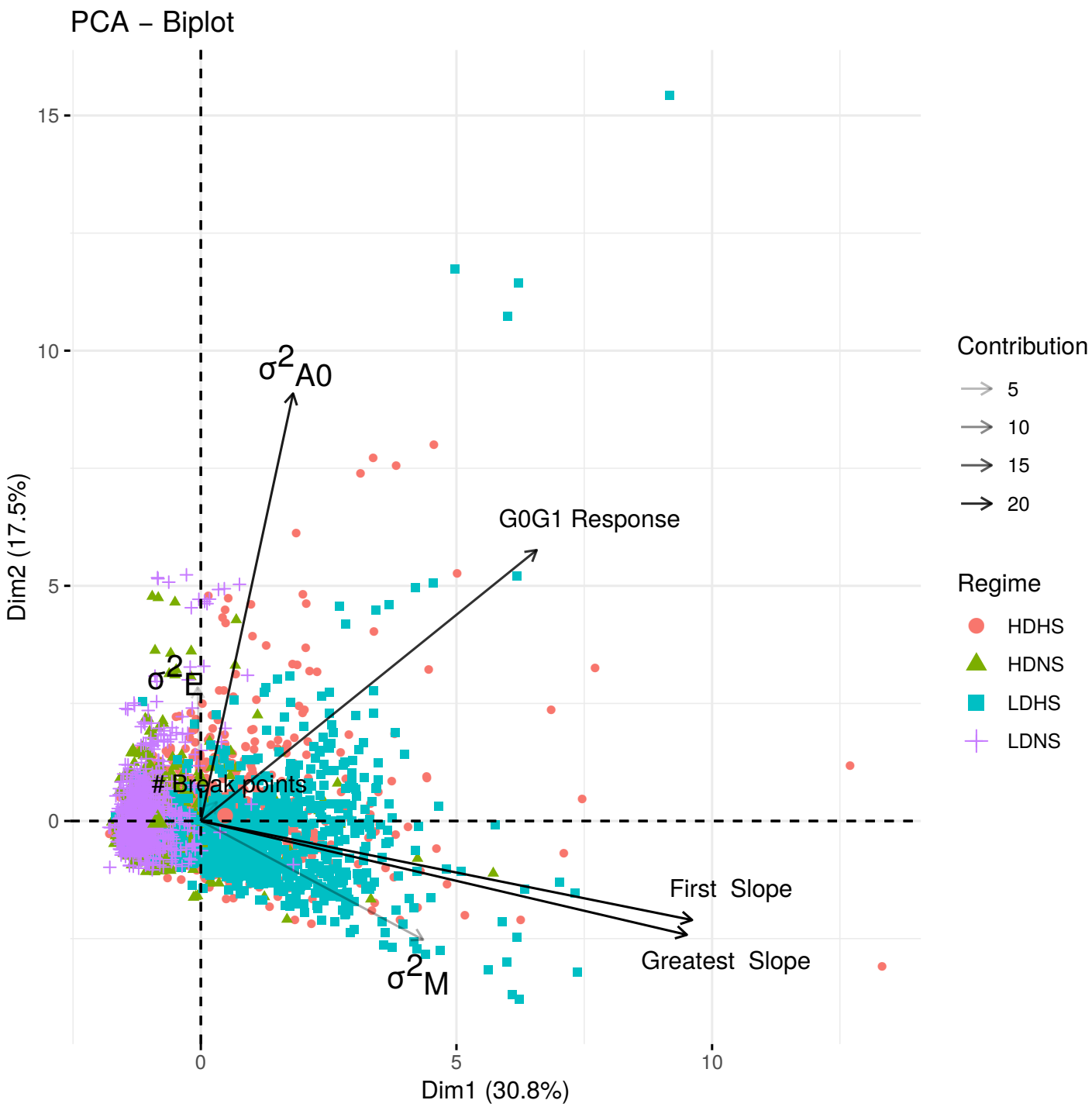

Figure S4 Correlation between model input variables (σ^2_{A0}, σ^2_M and σ^2_E) and output variables (G_0G_{20} Response, # Breakpoints, First Slope and Greatest Slope). We obtained the output variables by fitting a segmented linear regression to the selection response from G_1 to G_{20} in individual. We estimated the number of breakpoints, the corresponding slopes, as well as the first & greatest slope by AIC maximization. In addition we determined the G_0G_{20} response. A Principle Component analysis was carried out on a subset of 200 independent simulations per regime (HDHS, LDHS, HDNS, LDNS). The darker the arrow representing a variable, the higher the intensity of its correlation to the axes.

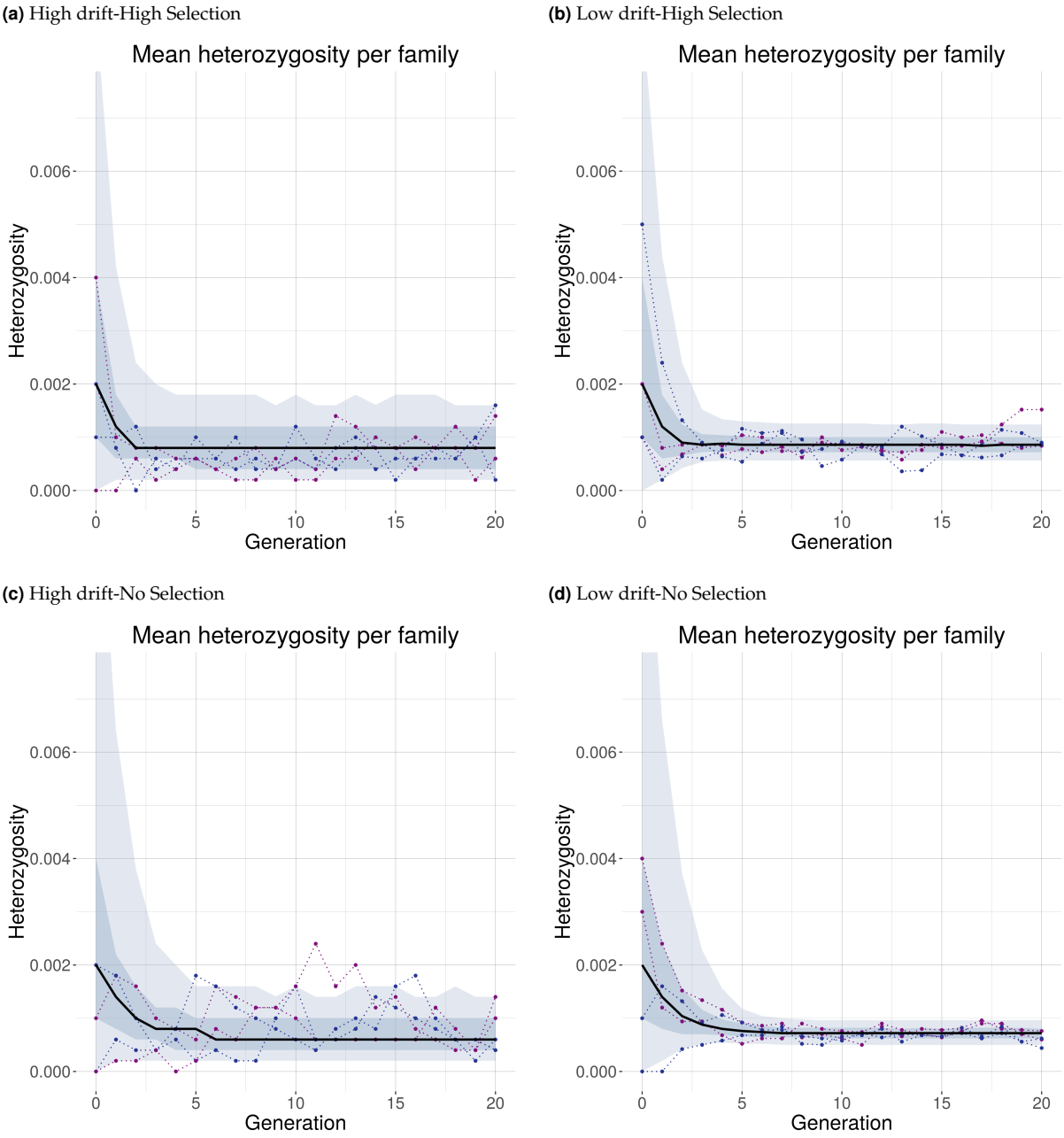

Figure S5 Evolution through time of the per-family mean heterozygosity over all loci, under HDHS (a), LDHS (b), HDNS (c), LDNS (d). The black line represents the median value of the per-family mean heterozygosity. The shaded area corresponds to the 5^{th}-95^{th} percentile (light blue) and to the 25^{th}-75^{th} percentile (dark blue). Four randomly chosen simulated families are represented with dotted line.

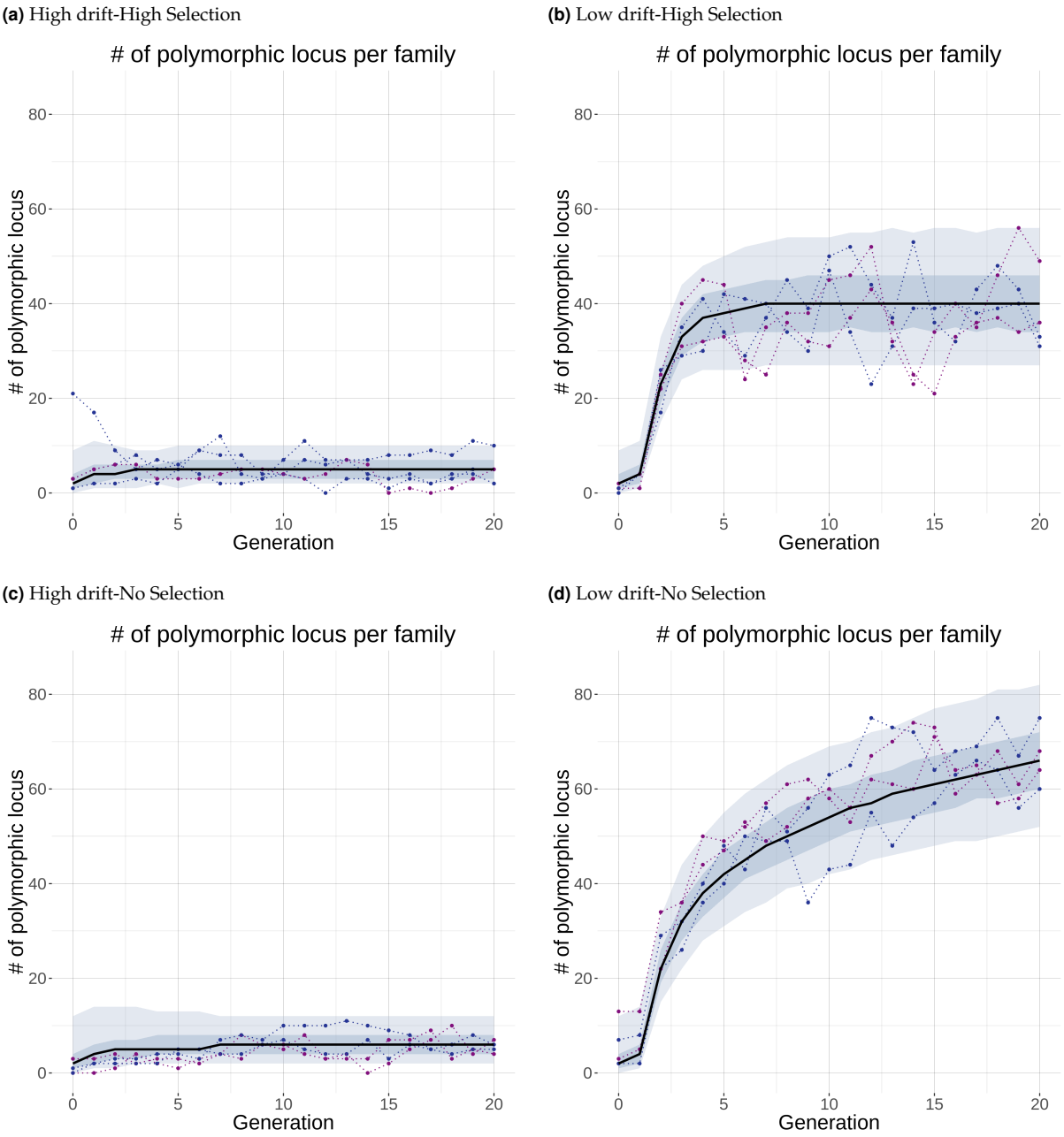

Figure S6 Evolution through time of the per-family mean number of polymorphic loci, under HDHS (a), LDHS (b), HDNS (c), LDNS (d). The black line represents the median value over 2000 simulations. The shaded area corresponds to the 5^{th}-95^{th} percentile (light blue) and to the 25^{th}-75^{th} percentile (dark blue). Four randomly chosen simulated families are represented with dotted line.

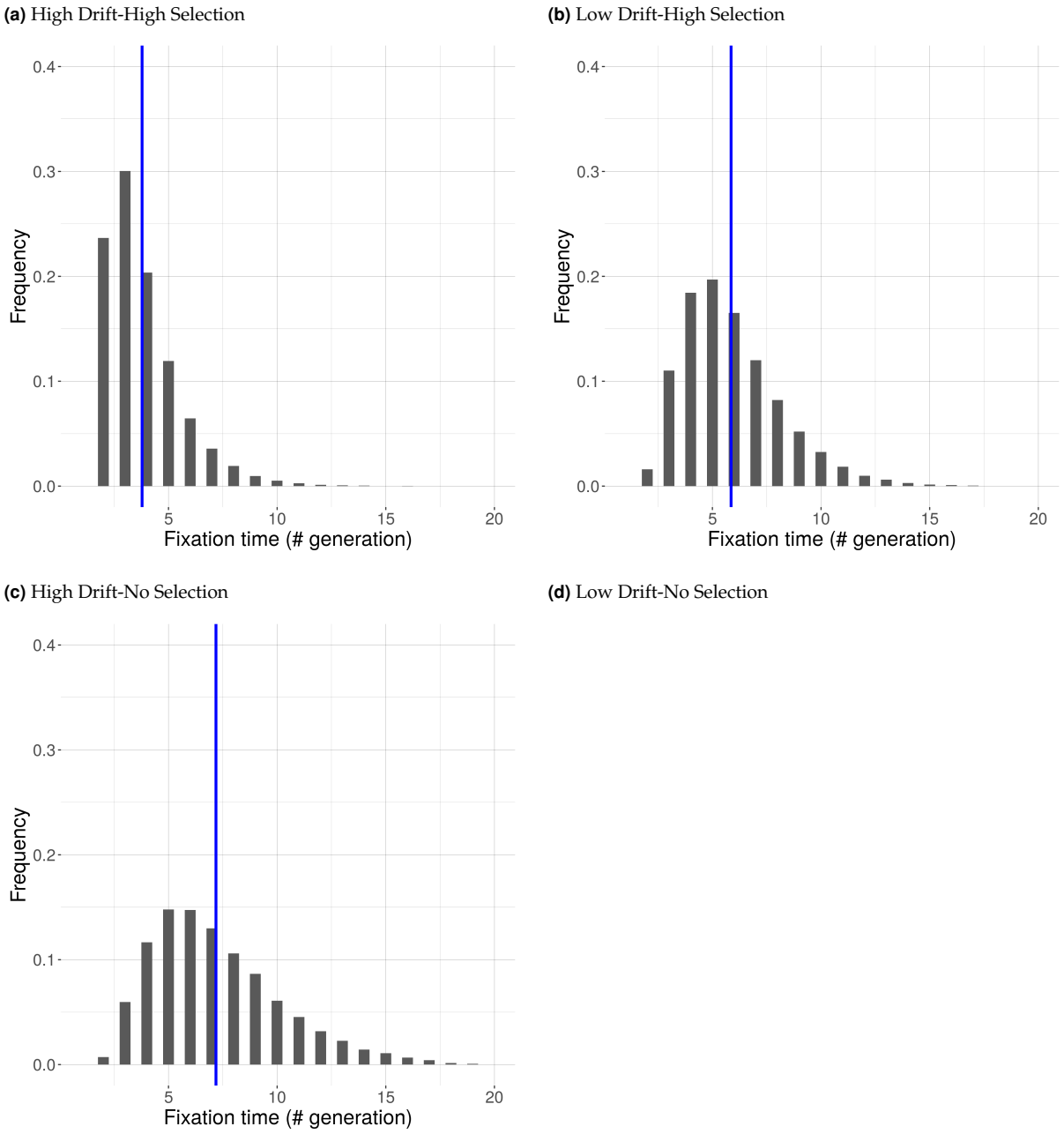

Figure S7 Frequency distribution of mutation fixation times over all simulated families under HDHS (a), LDHS (b), HDNS (c), LDNS (d). Note that under LDNS, we obtained very few fixed mutation so that we were unable to draw the corresponding distribution. Blue vertical lines represent the interpolated median.

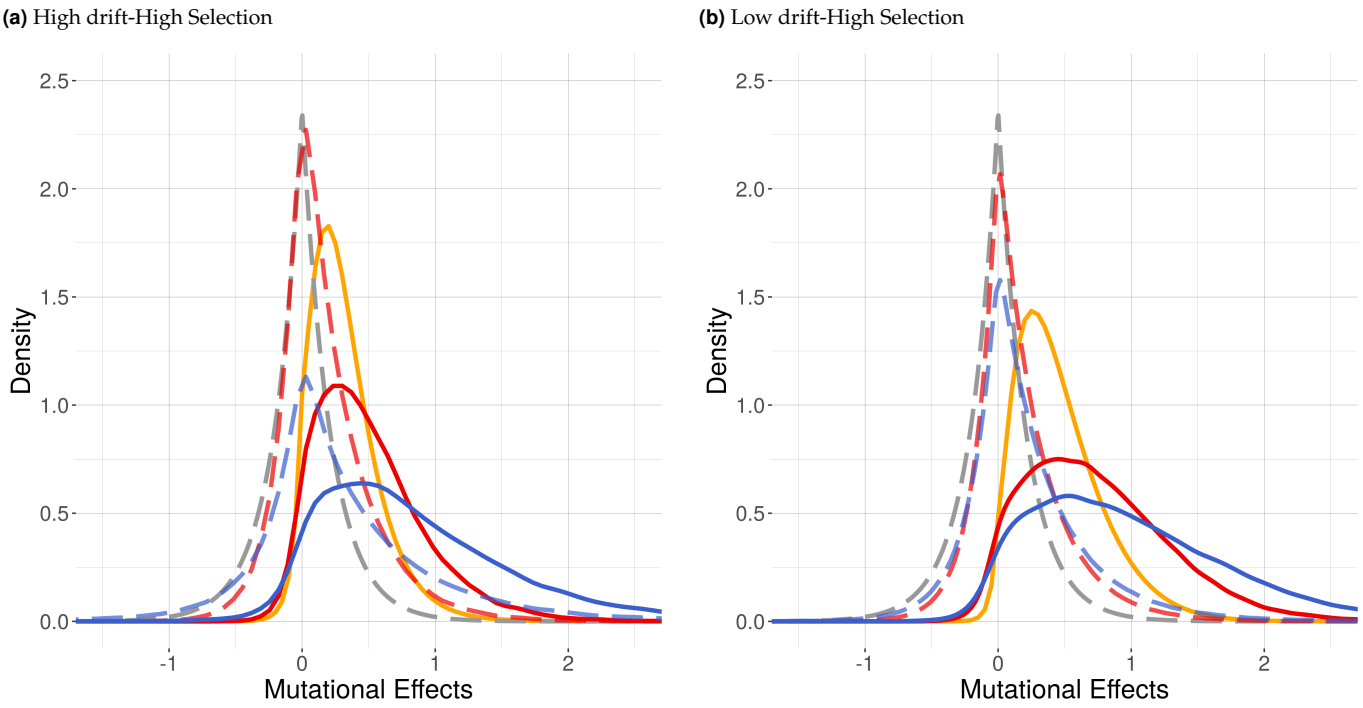

Figure S8 Distribution of mutation effects under HDHS (a), LDHS (b). The dotted lines indicate the distribution of effects (DFE) of incoming mutations considering raw effects in all individuals (grey), in selected individuals (red), and effects normalized by environmental variation in selected individuals (blue). The plain lines indicate DFE of fixed mutations following the same colour code. The golden line represents the expected DFE of fixed mutations according to Eq: 16.

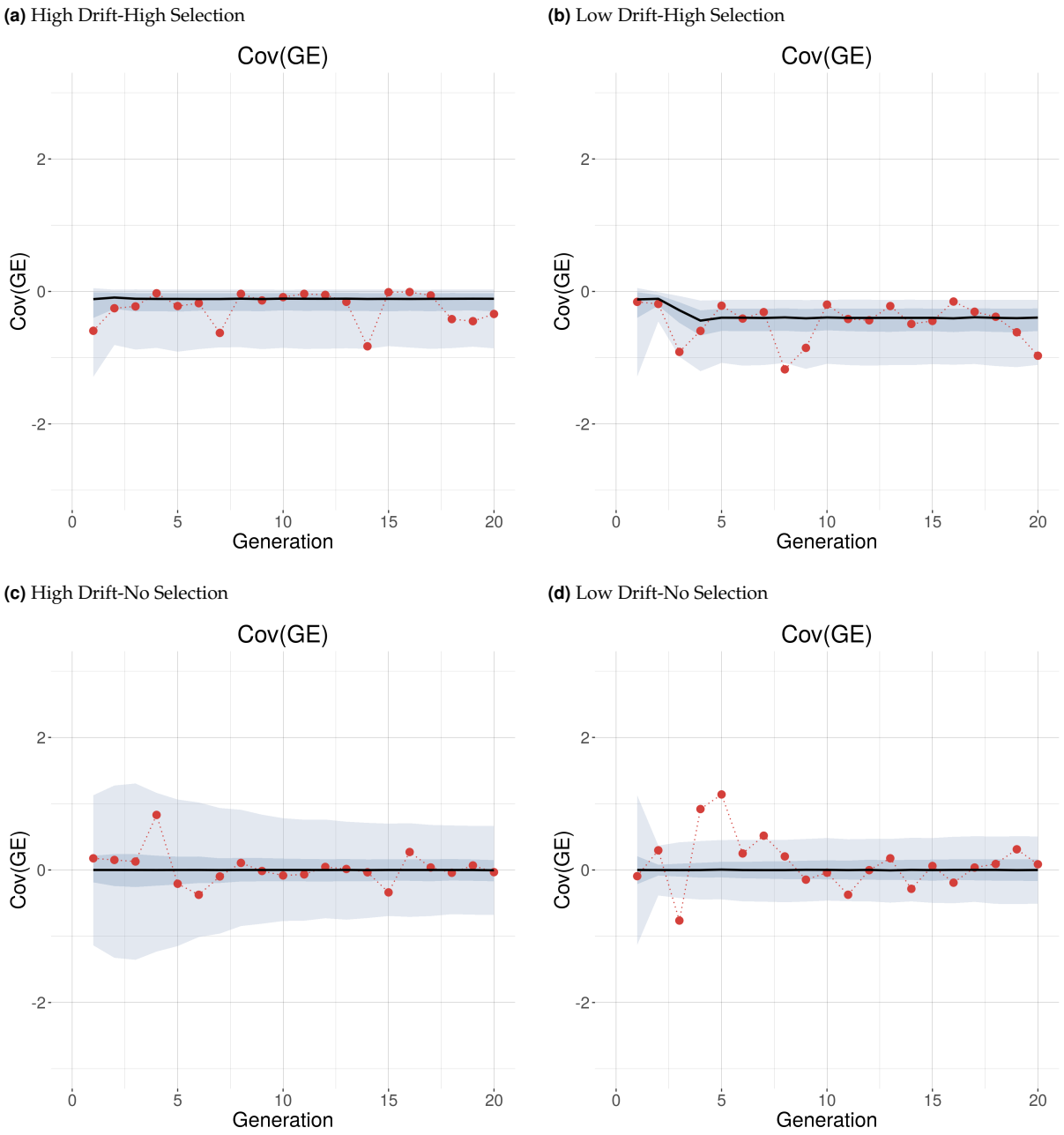

Figure S9 Evolution through time of the per-family covariance between environmental and genotypic values of the selected individuals, under our four simulated regimes. The black line represents the evolution of the median value over 2000 simulations in HDHS (a), LDHS (b), HDNS (c), LDNS (d). The shaded area corresponds to the 5^{th}-95^{th} percentile (light blue) and to the 25^{th}-75^{th} percentile (dark blue). One randomly chosen simulated family is represented with red dotted line, to highlight the inter-generation stochasticity. No significant autocorrelation was found.

Negative $\mathrm{Cov}(G_{|selected}, E_{|selected})$ ***schematic***

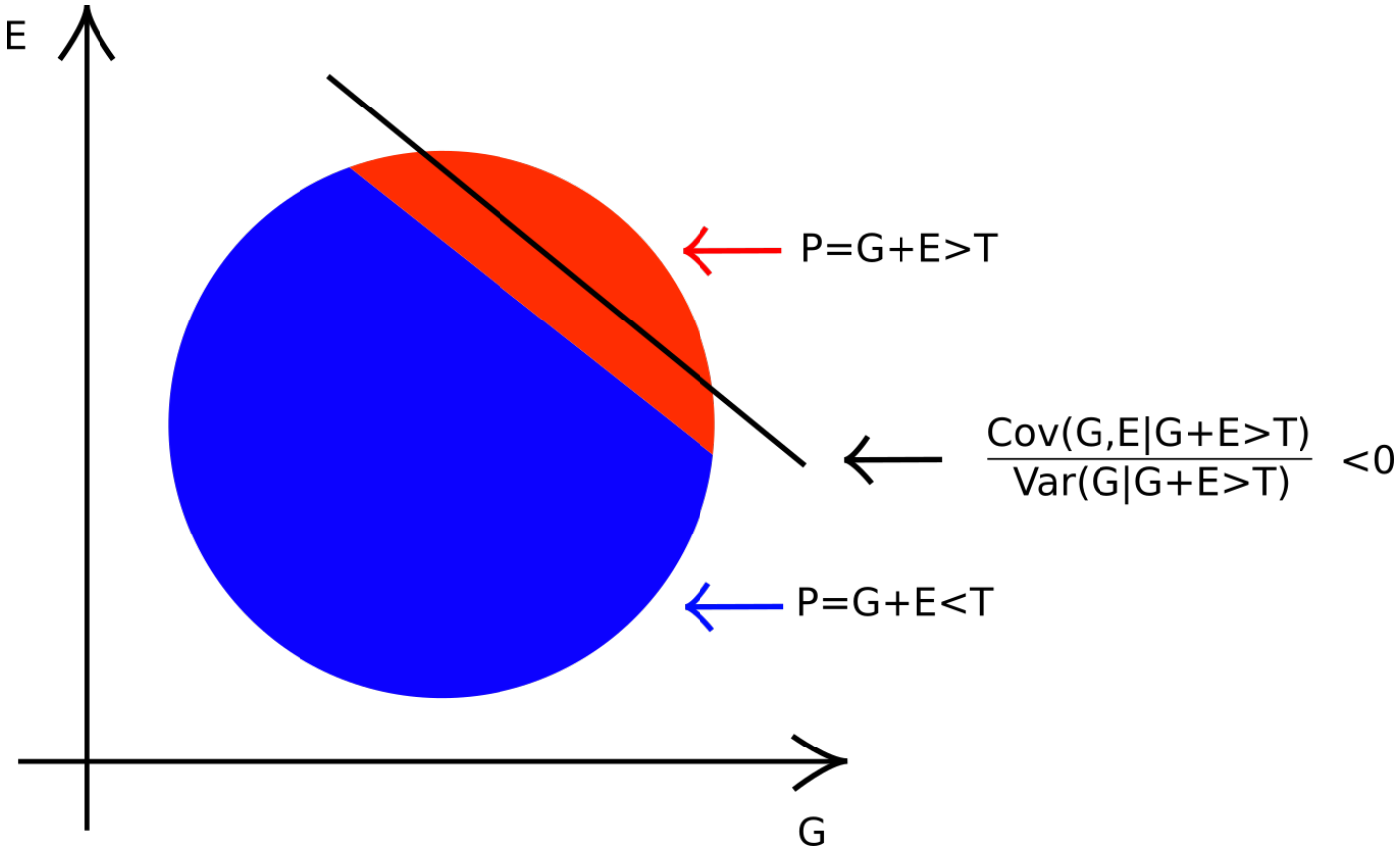

Figure S10 Schematic representation of the impact of selection on $\mathrm{Cov}(G, E)$**.** For illustration purposes, let P the sum of two independent and identically distributed random variables, G and E, such that both G and E follow a standard normal distribution, *i.e.* $P = G + E$ with $G \sim \mathcal{N}(0,1)$ and $E \sim \mathcal{N}(0,1)$. The black line represent the regression of $E_{|selected}$ on $G_{|\text{selected}}$ with a negative slope $\frac{\mathrm{Cov}(G_{|\text{selected}}, E_{|selected})}{\mathrm{Var}(G_{|\text{selected}})} \leq 0$.

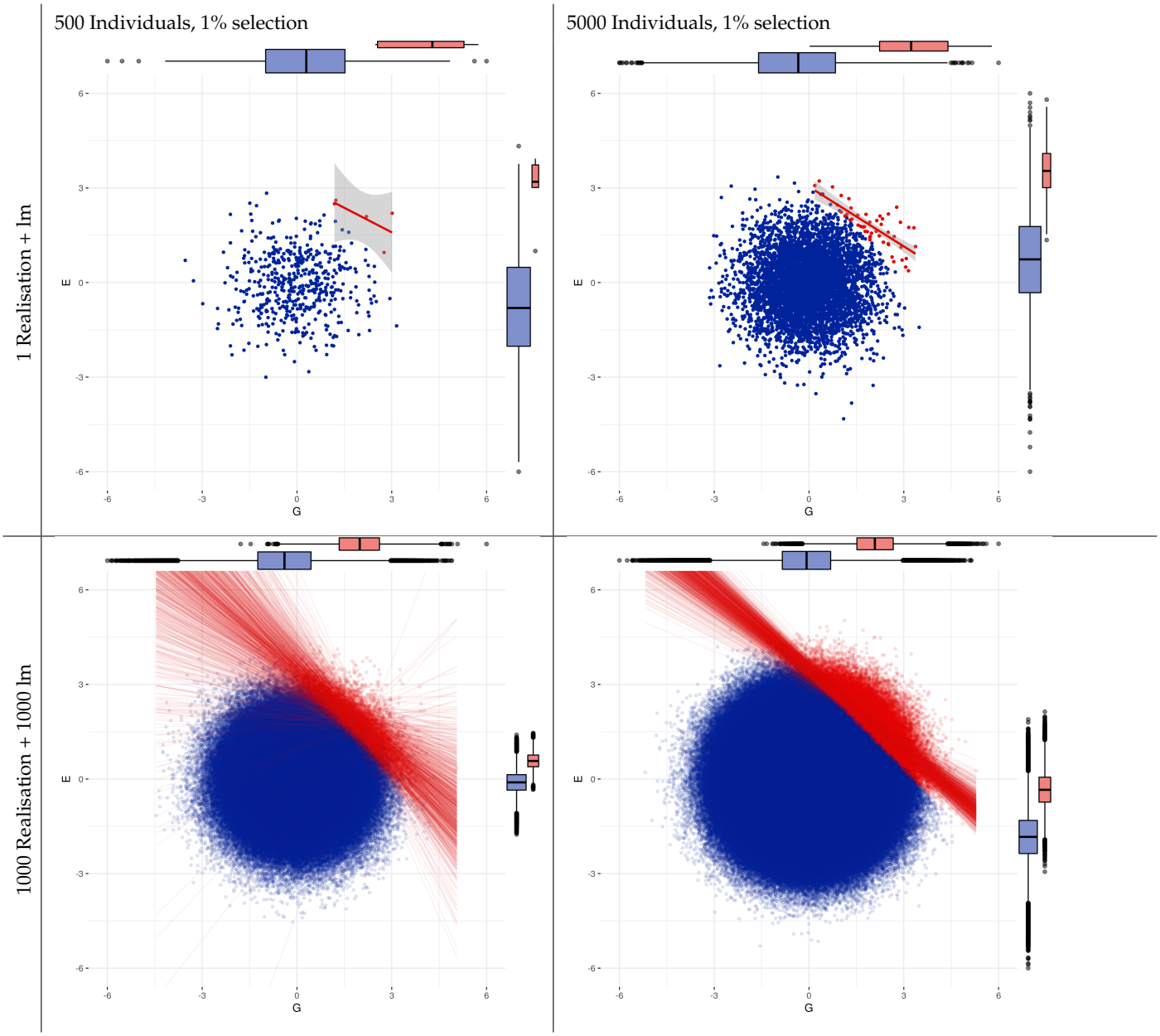

Figure S11 Schematic representation of the impact of selection and drift on $\mathrm{Cov}(G, E)$. Let P the sum of two independent random variables, G and E, such that both G and E follow a standard normal distribution, *i.e.* $P = G + E$ with $G \sim \mathcal{N}(0,1)$ and $E \sim \mathcal{N}(0,1)$. Let sample 500 individuals from P and plot $E = f(G)$ (right columns), resp. 5000 (left columns) and select (red dots) the best 1% based on P. The upper row represents one realisation, with the red line corresponding to the regression of $E_{|\text{selected}}$ on $G_{|\text{selected}}$ with a negative slope $\frac{\mathrm{Cov}(G_{|\text{selected}}, E_{|\text{selected}})}{\mathrm{Var}(G_{|\text{selected}})} \leq 0$. The lower row represents the realisation of 1000 independent sampling of 500 and 5000 individuals, with the corresponding linear regressions. We observe a lower lesser exploration of possible values (red plus blue area) under low population size and a high stochasticity in the values of $\mathrm{Cov}(G_{|\text{selected}}, E_{|\text{selected}})$

.