On deleterious mutations in perennials: inbreeding depression, mutation load and life-history evolution

Thomas Lesaffre*,[1] and Sylvain Billiard[1]

[1] Univ. Lille, CNRS, UMR 8198 - Evo-Eco-Paleo, F-59000 Lille, France

* Corresponding author. E-mail: thomas.lesaffre@univ-lille.fr

---

Abstract. In Angiosperms, perennials typically present much higher levels of inbreeding depression than annuals. Two hypotheses have been proposed to explain this pattern. Because plants do not have a segregated germline, the first hypothesis states that more long-lived species may accumulate more somatic mutations as they grow, which they could then transmit to their offspring and thereby generate higher inbreeding depression. The second hypothesis, which does not contradict the first, stems from the observation that inbreeding depression is expressed across multiple life stages in Angiosperms. It posits that increased inbreeding depression in more long-lived species could be explained by differences in the way mutations affect fitness in these species, through the life stages at which they are expressed. In this study, we investigate the second hypothesis, setting aside somatic mutations accumulation. We combine a physiological growth model and multi-locus population genetics approaches to describe a full genotype-to-phenotype-to-fitness map, where the phenotype relates to fitness through biological assumptions. We study the behaviour of mutations affecting growth or survival, and explore their consequences in terms of inbreeding depression and mutation load. Then, we discuss the role deleterious mutations maintained at mutation-selection balance may play in the coevolution between growth and survival strategies.

Keywords: age-structured population, fitness landscape, mutation-selection balance, epistasis, size-structured population

1 Introduction

Perennials, which make up the majority of Angiosperms ($\sim$ 70%, Munoz et al., 2016), typically present much higher levels of inbreeding depression than annuals. Indeed, meta-analyses found inbreeding depression to span from $\delta \approx 0.2$ on average in short-lived herbaceous species to $\delta \approx 0.5$ in long-lived herbaceous species and shrubs, and $\delta \approx 0.6$ in woody species (Duminil et al., 2009; Angeloni et al., 2011). Inbreeding depression, defined as the reduction in fitness of inbred relative to outbred individuals, is thought to be mainly due to the increased homozygosity of inbred individuals for recessive deleterious mutations segregating at low frequencies in populations (Charlesworth and Charlesworth, 1987; Charlesworth and Willis, 2009). Why inbreeding depression is higher in more long-lived species is still poorly understood, despite the potential significance of this pattern for important evolutionary questions, such as mating systems or dispersal rates evolution (Barrett and Harder, 1996; Roze and Rousset, 2005; Epinat and Lenormand, 2009; Duputié and Massol, 2013), and for more applied issues, as many cultivated species are perennial (e.g. fruit trees in general) and efforts are being made to develop perennial grain crops (DeHaan and Van Tassel, 2014).

Two hypotheses have been proposed to explain this pattern. The first hypothesis was formally put forward by Scofield and Schultz (2006). In plants, mutations occurring in somatic tissues may be passed onto the offspring, because they do not have a segregated germline. Thus, Scofield and Schultz (2006) proposed that more long-lived species may accumulate more somatic mutations as they grow and transmit them to their offspring, thereby generating the increase in inbreeding depression observed in such species. Phenotypic data in a long-lived clonal shrub (*Vaccinium angustifolium*, Bobiwash et al., 2013),

and genomic results in *Quercus robur* (Plomion et al., 2018) demonstrated that some somatic mutations can indeed be passed onto the offspring. However, the number of detected heritable somatic mutations is low (Schmid-Siegert et al., 2017; Plomion et al., 2018), and recent studies have concluded that the number of cell divisions from embryonic cells to gametes production may be much lower than previously thought (Burian et al., 2016; Watson et al., 2016; Burian et al., 2016; Schmid-Siegert et al., 2017; Lanfear, 2018), due for instance to early specification and quiescence mechanisms of axillary meristems cells, resulting in little opportunity for heritable mutations to accumulate during plant growth. Furthermore, intraorganismal selection is expected to efficiently purge deleterious somatic mutations, resulting in little to no somatically generated mutation load at the population level (Otto and Orive, 1995). Hence, although somatic mutations can be inherited and contribute to the mutation load in plants, their relative significance compared with meiotic mutations remains unclear (Schoen and Schultz, 2019).

The second hypothesis stems from the observation that inbreeding depression is typically expressed across multiple life stages in Angiosperms (Husband and Schemske, 1996; Winn et al., 2011; Angeloni et al., 2011). It posits that increased inbreeding depression in more long-lived species could be explained by the fact that mutations, regardless of whether they are produced during mitosis or meiosis, may differ in the way they affect fitness in annual and perennial populations, through the life stages at which they are expressed. Most theoretical studies of the mutation load focused on the case of mutations affecting fitness on a strictly linear fitness landscape (that is, fitness is the trait, e.g. Charlesworth et al., 1990; Roze, 2015), or through an abstract trait (or set of traits) under stabilising selection on a gaussian fitness landscape (e.g. Roze and Blanckaert, 2014;

Abu Awad and Roze, 2018). In these cases, inbreeding depression in annual and perennial populations is not expected to differ (Charlesworth, 1980). On the other hand, the dynamics of mutations affecting other aspects of individuals' life cycle, such as survival or growth, were seldom investigated. Morgan (2001) investigated the dynamics of mutations affecting survival between mating events in a perennial populations. They concluded that inbreeding depression should sharply decrease as life expectancy increases, and even become negative for long-lived species (outbreeding depression). However, Morgan (2001) studied mutations with a strong effect on fitness, and assumed no age-structure, that is, individuals did not differ in fecundity or survival probability with age. Strong variations in both survival and fecundity with respect to age are yet observed in perennials. Indeed, while juveniles typically suffer from very high mortality rates, established individuals tend to experience rather low mortalities with particularly slow senescence (Petit and Hampe, 2006). Furthermore, fecundity usually increases dramatically with age in perennials (Franco and Silvertown, 1996), due to the positive scaling of reproductive output with size in plants (Klinkhamer et al., 1985; Weiner et al., 2009). Mutations slowing their bearer's growth could therefore play a role in generating higher inbreeding depression in more long-lived species, as growth delays may impact individuals' fecundities differently depending on their age or size. This latter aspect of age-structuration in perennials was, to our knowledge, never tackled theoretically.

The present study aims to study the second hypothesis, that is, we investigate the behaviour of mutations affecting fitness differently with respect to life-history, putting aside somatic mutations accumulation. Namely, we study meiotic mutations affecting growth or survival in a partially selfing population, in which individuals grow as they

age and fecundity is proportional to size, but survival between mating events is assumed to not depend on age. We combine a physiological growth model (West et al., 2001) and multilocus population genetics approaches (Barton and Turelli, 1991; Kirkpatrick et al., 2002) in order to describe a full genotype-to-phenotype-to-fitness map, where the fitness landscape emerges from biological assumptions instead of being assumed *a priori*. We study the behaviour of different types of mutations affecting growth or survival, and explore their consequences in terms of inbreeding depression and mutation load (Crow, 1958). Then, we show the role deleterious mutations maintained at mutation-selection balance may play in the coevolution between growth and survival strategies.

2 Model outline and methods

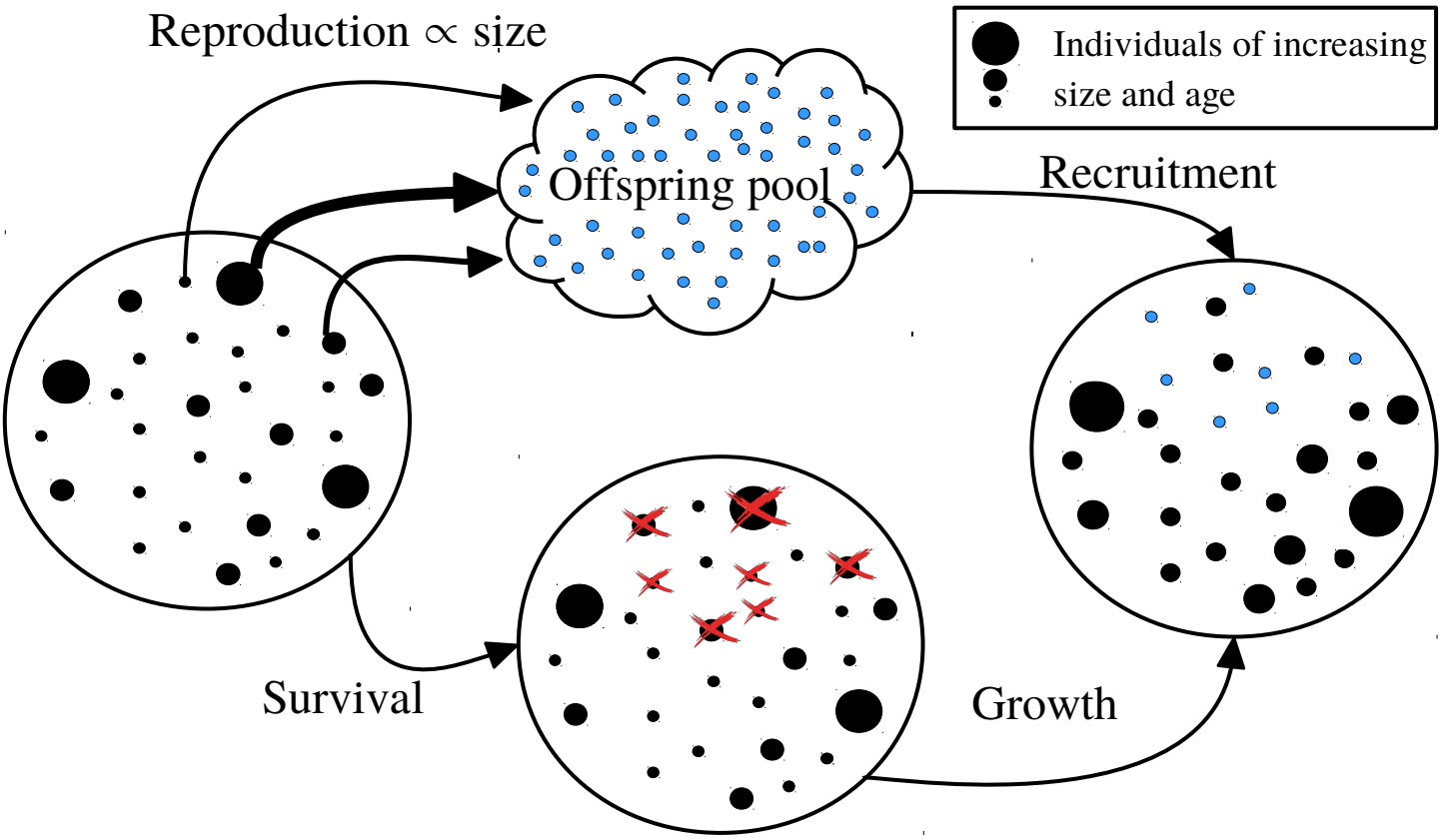

FIGURE 1: Life cycle and demography. Deceased individuals are marked by a red cross, juveniles are depicted in light blue. Larger dots depict larger and older individuals.

Life cycle. We consider a large population of diploid hermaphrodites, which may survive from one mating season to another with a probability S, assumed to be constant with respect to age. If they survive, individuals grow between mating events, following a physiological growth model described briefly in the next paragraph. If they die, juveniles, which are assumed to be produced in large excess compared with the resources available for establishment, are recruited to replace the dead, so that population size is kept constant (FIG. 1). Each juvenile has a probability J of being recruited. During reproduction,

individuals are assumed to contribute to the gamete pool in proportion to their size (the larger an individual, the larger its contribution to the gamete pool), and to self-fertilise at a fixed rate α.

Growth model. We consider the growth model developed by West et al. (2001) (see details in Appendix I). The energy available for growth and maintenance at age t, B_t, is assumed to scale as a $^3/_4$-power law of body size as a result of allometry (Peters, 1983), so that

$$B_t = B_0 \, G_t^{3/4}, \tag{1}$$

where B_0 is the basal metabolic rate and G_t is body size at age t. This energy can be subdivided into the energy required to maintain the existing body, controlled by a maintenance cost c, and the energy available to produce new body parts, controlled by a production cost ε, so that growth is fully described by the following differential equation

$$\frac{dG}{dt} = \frac{B_0}{\varepsilon} G_t^{3/4} - \frac{c}{\varepsilon} \, G_t. \tag{2}$$

Under this model, individual size naturally saturates when the energy required to maintain the existing body equals the available energy (Fig. 2a-2b).

Genetic assumptions. Mutations are assumed to occur at rate U (per haploid genome) at a large number of loci, which recombine at rate $0 \leqslant r \leqslant \frac{1}{2}$. In three separate models, we consider mutations affecting three different traits. Mutations may affect growth by increasing either their bearer's maintenance cost (c) or production cost (ε), or they may affect its survival. When mutations affect survival, they are assumed to decrease both their

bearer's probability of being recruited as a juvenile (J) and its adult survival probability (S). The effect of mutations is denoted s, with a dominance coefficient $h \leqslant \frac{1}{2}$. Loci affect traits multiplicatively, so that for any trait z ($z \in \{c, \varepsilon, S\}$), we have

$$z = z_0 \left(1 \pm s\right)^{H_o} \left(1 \pm sh\right)^{H_e},$$

where H_o (resp. H_e) is the number of homozygous (resp. heterozygous) mutations born by the considered individual.

Approximation of the expected number of mutations, inbreeding depression and mutation load. We use two approaches to study our model. In the first approach, we make the assumption that selective pressures acting on mutations at various life-stages can be summarised into a single lifetime fitness expression, so that the population can be studied as an adequately rescaled annual population (the Lifetime Fitness approach, LF). For each type of mutation, this approach allows us to gain insights into the way selection acts on mutations, by summarising it into a single lifetime selection coefficient. This coefficient is denoted $\bar{s}_z$ (where $z = c$, ε or S depending on the considered trait). Besides, reasoning in terms of lifetime fitness is paramount to compute key quantities such as inbreeding depression or the mutation load. However, the LF approach fails to account for genetic associations correctly. Thus, in order to obtain approximations of the expected number of mutations per haploid genome accounting for genetic associations, we also use a second approach where we study each step of the life cycle successively (the Life Cycle approach, LC). We do so under the assumption that the phenotypic effect of mutations is weak and that the number of segregating mutations is large, following the work of Roze

(2015) that we adapted to the case of many mutations affecting a trait rather than fitness directly, in an age- and size-structured population.

Simulations. Individual-centered simulations were run assuming individuals depicted by two linear chromosomes of length λ (in cM), along which mutations occur stochastically (Roze and Michod, 2010). The number of mutations occurring is sampled from a Poisson distribution with mean U, and their position on chromosomes are sampled from a uniform distribution. Recombination is modeled by exchanging segments between chromosomes. Similar to mutations, the number of crossing-overs is sampled from a Poisson distribution with mean λ, while their position along chromosomes are randomly drawn in a uniform distribution. The population has a constant size N. At each timestep, an individual survives with probability S, which can depend on its genotype in the case of mutations affecting survival. If the individual does survive, it grows deterministically depending on its age and individual physiological growth costs (Fig. 2, Equation A4). If it does not, it is replaced by an offspring generated from the population, in which parents are chosen with a probability proportional to their size. The offspring is produced by self-fertilisation with probability α, and by random mating otherwise. We measure the average of the trait affected by mutations, the average number of mutations per haploid genome in the population, and inbreeding depression. Inbreeding depression is measured as the relative difference in lifetime reproductive success of selfed and outcrossed individuals. Individuals lifetime reproductive success is obtained by counting the number of times they are chosen as parents before they die. The mutation load in the sense of Crow (1958), that is the decrease in mean fitness of the population compared with a population with no mutations, is measured using the average of the trait

affected by mutations. This average is used to compute the expected mean lifetime fitness in the population (Equation (A16)). Then, this quantity is compared to the expected mean lifetime fitness when mutations are absent. All programs are available from GitHub (https://github.com/Thomas-Lesaffre/On_deleterious_mutations_in_perennials).

3 Results

3.1 Mutation-selection equilibrium

The intensity of selection acting on mutations depends on how they modify the fitness of their bearer, and how the fitness of their bearer then compares with the fitnesses of all the individuals they are competing with, that is the distribution of fitness in the population. Hence, the intensity of selection faced by a mutation segregating at a given locus may depend on the presence of mutations at other loci, since said mutations may alter both the distribution of fitness in the population and how a mutation at the considered locus affects its bearer's fitness. This second aspect is referred to as epistasis. There are numerous ways by which epistasis may be generated between loci (Phillips, 2008), which can roughly be split into two categories: genotype-to-phenotype epistasis on the one hand, and phenotype-to-fitness epistasis on the other hand. Genotype-to-phenotype epistasis refers to loci which influence each other's expression in the determination of a phenotype, say genes involved in a common metabolic pathway, while phenotype-to-fitness epistasis refers to situations where loci, even if they influence phenotype independently, influence each other because the way the phenotype affects fitness is non-linear.

In our model, we assume that mutations affect a phenotype multiplicatively, so that genotype-to-phenotype epistasis does not occur. On the other hand, we do not impose

any constraint on the shape of the fitness landscape. Instead, we let the genotype-to-phenotype-to-fitness map arise from biological assumptions. The resulting fitness landscape is non-linear, convex, and has singularities at the optima (Equation A16 in Appendix III.1, Fig. S1). Thus, phenotype-to-fitness epistatic interactions are generated, and have to be accounted for to obtain accurate mutation-selection balance approximations. Quantifying the intensity of selection acting on a mutation at a given locus requires accounting for the average number of mutations it may be in epistatic interaction with, which can be done by following the change in the trait average as mutations segregate when the population is large enough. We do so through Equation (3) below, the derivation of which is detailed in Appendix II.

$$\begin{cases} \Delta p_i\left(\bar{z}\right) = 0 \\ \bar{z}\left(p_i\right) = z_0 \times e^{\pm s \sum_i \overline{Z}_i} \left(1 + s^2 \sum_{j \neq i} \frac{\overline{\Delta Z_i \times \Delta Z_j}}{2}\right). \end{cases} \tag{3}$$

In the first line of Equation (3), Δp_i is the change in the frequency of the mutant allele at the ith locus, which depends on the trait average $\bar{z}$. On the second line, $\bar{z}$ is the average of the trait in the population, which depends on the number of mutation per haploid genome $\sum_i p_i$. On this line, the term $\sum_i \overline{Z}_i$ quantifies the effect on the trait of the $\sum_i p_i$ mutations per haploid genome born on average by individuals, neglecting the effects of genetic associations between loci, while the term $\overline{\Delta Z_i \times \Delta Z_j}$ quantifies the effect of pairwise associations between loci on the trait average. Associations of higher order are neglected. Solving Equation (3) for p_i and $\bar{z}$ allows us to obtain predictions for the average of the trait and for the average number of mutations per haploid genome maintained at mutation-selection equilibrium, which we then use to compute the inbreeding depression

and mutation load expected at equilibrium.

3.2 Fitness effect of mutations neglecting genetic associations

To leading order, genetic associations can be neglected so that Equation (3) simplifies into

$$\begin{cases} U - \bar{s}_z\left(\bar{z}\right)\left[h\sum_i p_i + (1-h)F\sum_i p_i\right] = 0 \\ \bar{z}\left(p_i\right) = z_0 \times \exp\left[\pm s\left(2h\sum_i p_i + (1-2h)F\sum_i p_i\right)\right], \end{cases} \tag{4}$$

where $F = \frac{\alpha}{2-\alpha}$ is the inbreeding coefficient. The first line of Equation (4) shows that when we neglect genetic associations between loci, the selective pressures acting on mutations are encapsulated in a single lifetime selection coefficient $\bar{s}_z$ (Appendix III.1.2).

The phenotypic effect of each type of mutation (Fig. 2, top row), and the resulting lifetime selection coefficients are presented in Fig. 2 (bottom row). These coefficients depend on the population average of the trait, owing to the epistatic mechanisms described in the former section. Although it is not possible to fully disentangle the effects of epistasis from the effects of direct selection at the locus, one may gain insight into the sign of epistasis by studying the direction in which $\bar{s}_z$ coefficients are changed when the number of segregating mutations increases in the population. The $\bar{s}_z$ coefficients plotted in black were obtained assuming mutations are absent (*i.e.* $\bar{z} = z_0$), while those plotted in red were obtained assuming $n = \sum_i p_i = 10$ mutations were segregating at surrounding loci. In other words, the coefficients plotted in black in Fig. 2 (bottom row) represent the intensity of selection a mutation would face if a single locus was modeled, while the coefficients plotted in red depict the intensity of selection a mutation would face when in interaction with ten mutations at other loci.

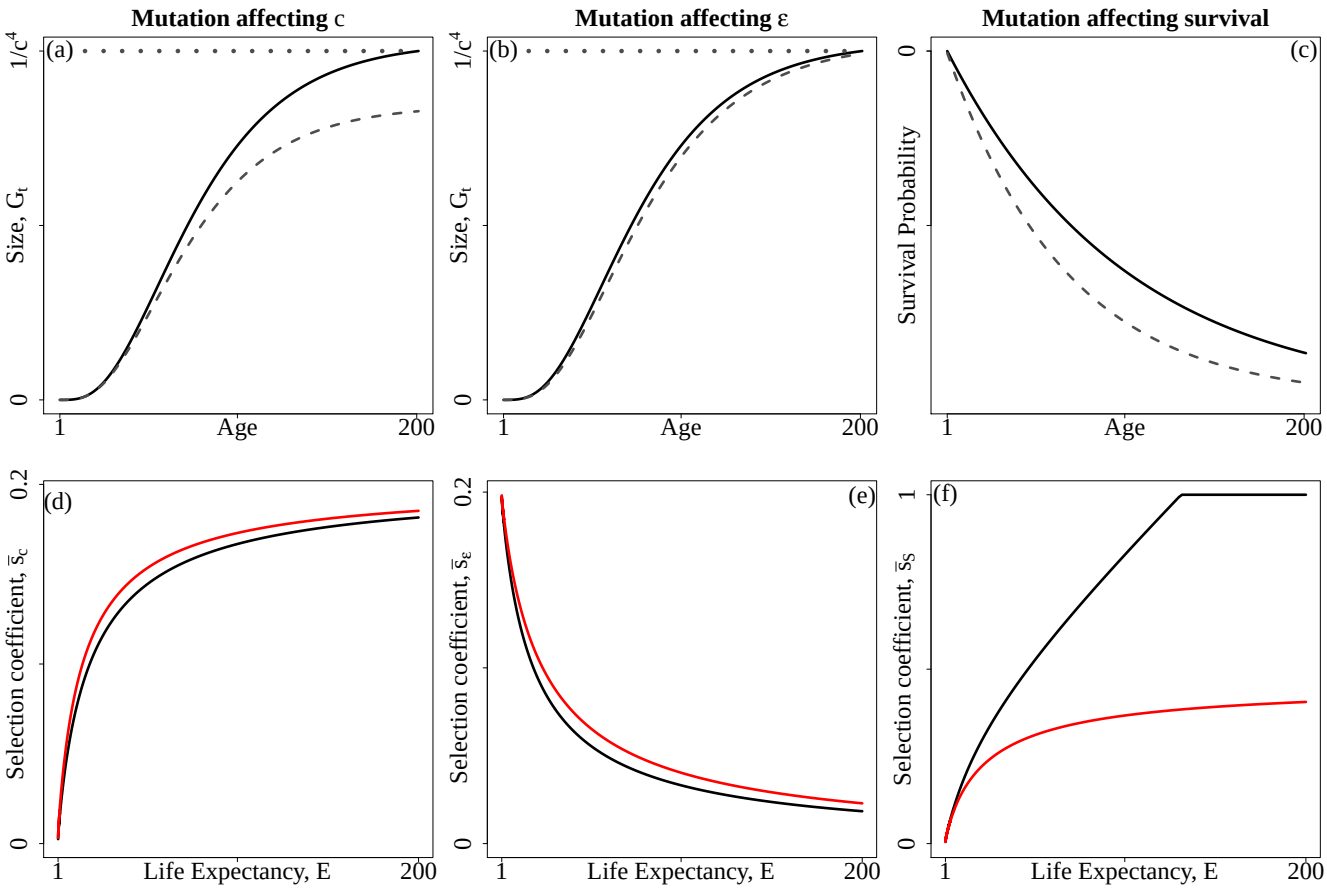

FIGURE 2: Phenotypic and leading order fitness effect of a mutation. Top row: phenotypic effects on growth or survival are presented as a function of age, the dashed line depicts the mutated phenotype while the solid line depicts the unmutated ($c = 0.001$, $\varepsilon = 0.01$ and $s = 0.05$ for mutations affecting growth, and $S = 0.99$, $s = 0.005$ for mutations affecting survival). The dotted gray lines depict maximal size. Bottom row: Resulting effects of mutations on lifetime fitness as a function of life expectancy in a single locus case (black lines, $n = 0$), and 10 mutations segregate at other loci (red lines, $n = 10$). Note that the phenotypic effect of mutations (s) differs between mutations affecting growth and survival.

Mutations affecting the maintenance cost c cause growth delays to increase as individuals age (FIG. 2a). Hence, their fitness effect increases with life expectancy (FIG. 2d). Furthermore, epistatic interactions cause selection against mutations to increase that is, we observe negative epistasis. On the contrary, mutations affecting the production cost ε do not affect individuals' maximal size, as they asymptotically tend to the same size, but rather the speed at which they reach it (FIG. 2b). Therefore, the growth delay mu-

tated individuals accumulate in early years fades away in older individuals. This causes selection against mutations affecting ε to decrease with respect to life expectancy, because they become gradually neutral in older age-classes. However, epistasis also causes selection against mutations to increase, so that it is negative in this case as well. Mutations affecting survival cause individuals to perform less mating events in a lifetime. They also cause mutated individuals to perform less well during mating events, because they tend to be younger than unmutated individuals (Fig. 2c). Thus, age-structure increases selection against mutations affecting survival. Furthermore, selection against mutations affecting survival strongly increases as life expectancy increases (Fig. 2f), contrary to mutations affecting growth costs, whose fitness effects remain moderate and of the same order of magnitude (Fig. 2d-e). Besides, epistasis causes selection to sharply decrease when the number of segregating mutations increases, that is, strong positive epistasis is observed. Overall, the results described in Fig. 2 show that mutations affecting different traits on the same genotype-to-phenotype-to-fitness map face very different selective pressures, both in magnitude and in the way they vary with life expectancy.

3.3 Average number of mutations, inbreeding depression and mutation load

The intensity of selection acting on mutations does not depend on the values of c and ε, but on the ratio $\frac{c}{\varepsilon}$ (Appendix III.1.2). Thus, in what follows, results are described using this ratio. Figure 3 presents results for the average number of mutations per haploid genome maintained at equilibrium, and the resulting inbreeding depression and mutation load for $h = 0.25$ and $\frac{c}{\varepsilon} = 1$ (other parameter sets are shown in appendix as results are qualitatively

similar, Fig. S3 to S10). Analytical predictions, which are depicted by solid lines, are obtained by solving Equation (3) numerically using the LC approach and accounting for pairwise genetic associations, and dots depict simulations results. For mutations affecting the maintenance cost and survival, the average number of mutations per haploid genome maintained at equilibrium ($n = \sum_i p_i$) decreases as life expectancy increases (Fig. 3a-c), because selection against mutations increases. This effect is more marked for mutations affecting survival because selection is stronger in this case (Fig. 2d-f). Conversely, n increases as life expectancy increases for mutations affecting the production cost, because selection against these mutations weakens as life expectancy increases (Fig. 2e). In every case, n decreases as the selfing rate increases due to the purging effect of self-fertilisation.

Large differences in the number of deleterious mutations maintained per genome do not translate into strong variations in inbreeding depression with respect to life expectancy (Fig. 3, middle row). However, these variations differ between mutation types. Indeed, inbreeding depression always decreases with life expectancy when mutations affect survival, irrespective of the selfing rate (Fig. 3f), contrary to mutations affecting growth. Mutations affecting c generate higher inbreeding depression in more short-lived species at low selfing rate, but this pattern is reversed at higher selfing rates (Fig. 3d). Conversely, mutations affecting ε generate higher inbreeding depression in more long-lived species for low α and this pattern is reversed for high α (Fig. 3e). This is the result of the interaction between life expectancy, epistasis, the magnitude of selection and the selfing rate.

The mutation load is lower when mutations are more numerous for all three mutations types (Fig. 3, bottom row). This result is not a mere reflection of differences in the absolute strength of selection acting on mutations in populations with different life

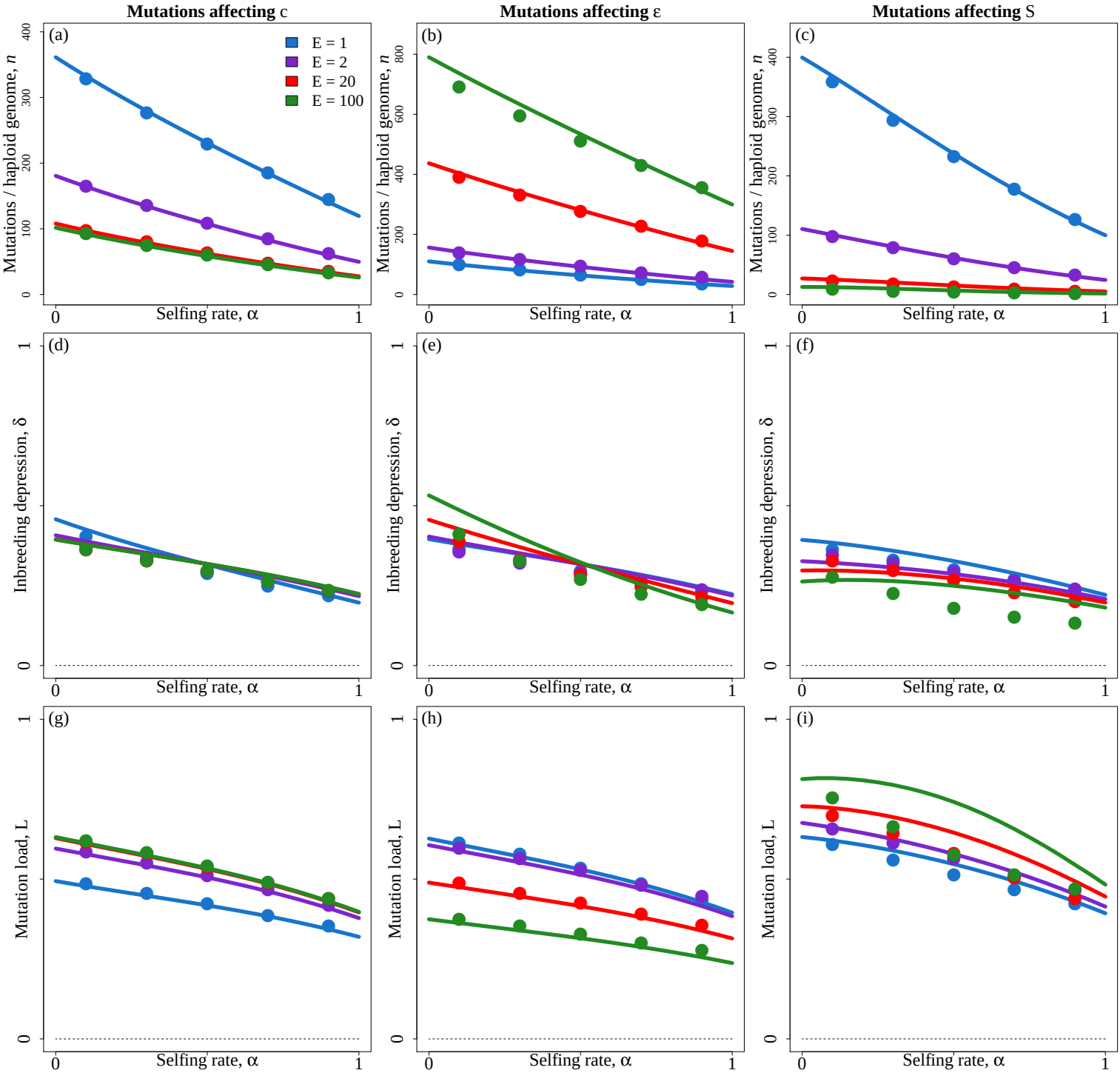

FIGURE 3: Average number of mutations per haploid genome (n, top row), inbreeding depression (δ, middle row), and mutation load (L, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lines: analytical predictions accounting for genetic associations between loci. Parameters shown here are $\frac{c}{\varepsilon} = 1$, $U = 0.5$, $s = 0.005$, $h = 0.25$.

expectancies. Otherwise, changing the phenotypic effect of mutations in a given model and for a given parameters set would also change the mean phenotypic deviation from the optimum, and therefore the mutation load. This is not what we observe. Indeed, FIGURE S2 shows that when the effect of mutations is made ten times larger, this does not cause differences in the equilibrium phenotypic deviation, although the number of segregating mutations at equilibrium was considerably lower (because selection was stronger). This observation is consistent with results obtained by previous authors (Bataillon and Kirkpatrick, 2000), who showed that when populations exceeds a particular size, the mutation load becomes independent from the strength of selection acting on mutations. We argue that the differences we observe in terms of mutation load are imputable to differences in the shape of the fitness landscapes (FIG. S1), which are not fully captured by differences in the efficacy of selection.

3.4 Consequences for life-history evolution

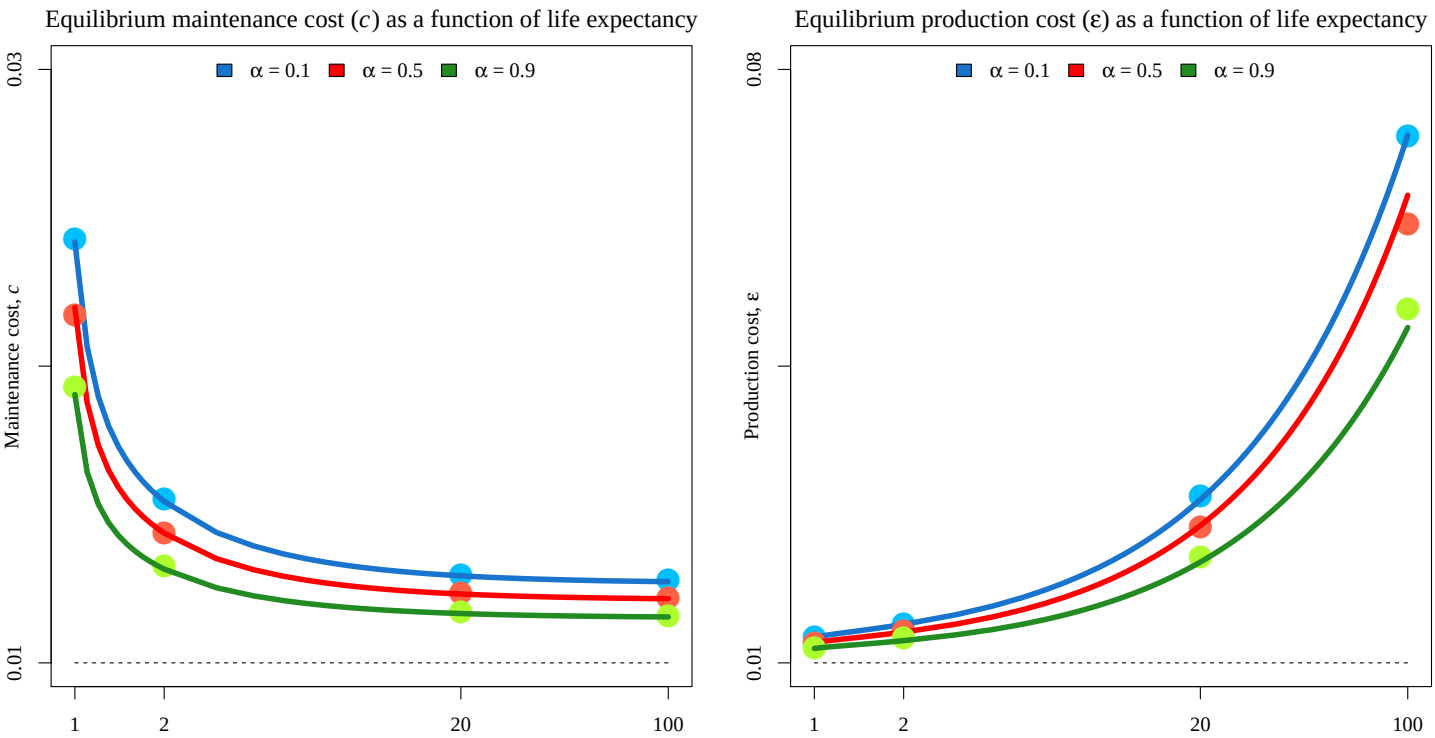

FIGURE 4: Equilibrium maintenance (left) and production (right) costs as function of life expectancy (log-scaled), for various selfing rates. Dots depict simulation results while lines depict analytical predictions. Parameters shown here are $c = \varepsilon = 0.01$, $h = 0.25$, $s = 0.005$.

The mutations segregating in the population cause the mean phenotype to deviate from its initial value. As the intensity of selection acting on mutations depends on this mean phenotype, this generates a coevolutionary process, which leads to a joint equilibrium for both the mean phenotype and the number of mutations segregating in the population (Appendix II, III). This equilibrium varies with respect to life expectancy and mating system. Increasing the selfing rate slightly decreases the deviation of the mean phenotype from its initial value, because selfing induces a better purging of mutations. More importantly, the equilibrium phenotypes vary significantly between populations with different life expectancies. Indeed, as species become more long-lived, the maintenance cost de-

creases, so that maximal size increases, and the production cost increases, so that growth is slowed down. This means that if, for any reason, life expectancy changes in a species, the growth strategy should also be changed as a consequence of the selective pressures acting on deleterious mutations maintained at mutation-selection balance being altered.

4 Discussion

Inbreeding depression. Our initial aim was to investigate the proposition that higher levels of inbreeding depression in more long-lived species may be explained by mutations affecting fitness differently in such species. To do so, we studied mutations affecting growth and survival. Mutations affecting growth, that is affecting the maintenance and the production cost, followed qualitatively different patterns, but differences in the magnitude of inbreeding depression with respect to life expectancy were quantitatively small in every investigated case, even for low dominance coefficients and high mutation rates. This result can be explained by the moderate variations in the intensity of selection on growth-related traits with respect to life expectancy, and by the fact that we only considered mutations with weak phenotypic effects, which translated into weak fitness effects. Therefore, this result is consistent with previous population genetics work that showed that inbreeding depression becomes independent of the intensity of selection when selection is weak and the population is large (e.g. Charlesworth et al., 1990; Bataillon and Kirkpatrick, 2000; Roze, 2015).

Prior to the present study, the dynamics of mutations affecting survival had been studied by Morgan (2001), in a perennial but not age-structured population. They concluded that inbreeding depression should quickly decrease as life expectancy increases, and that

significant outbreeding depression should be observed in long-lived species. They argued this result could be attributed to the greater variance in fitness observed among the offspring produced by self-fertilisation, which led to a higher mean lifetime fitness among them. Although they did not state it explicitly, this result stems from the fact that a very small number of mutations were maintained at mutation-selection balance in long-lived species under the parameter sets they investigated. Indeed, Morgan (2001) considered very large phenotypic effects of mutations ($s = 0.1$ and $s = 1.0$), which resulted in tremendous lifetime fitness effects and therefore in the maintenance of almost no mutation at equilibrium. This led variance in fitness to dominate over other aspects. Here, we investigated mutations with much lower phenotypic effects ($s = 0.005$), and did not observe the same patterns. In fact, we show that although inbreeding depression still decreases with life expectancy, differences are considerably reduced, and outbreeding depression is no longer observed. Thus, if inbreeding depression is observed on traits related to adult survival in long-lived species, we conclude that this should be caused by mutations with very weak phenotypic effects.

In summary, in our model, inbreeding depression at mutation-selection balance either varies weakly with respect to life expectancy, as it is the case with mutations affecting growth because the fitness effect of mutations remains weak and variations with respect to life expectancy are moderate, or inbreeding depression decreases as life expectancy increases in the case of mutations affecting survival, because selection sharply increases with life expectancy. In either case, our results are not in agreement with empirical data, as a strong increase in inbreeding depression is observed in more long-lived species (Duminil et al., 2009; Angeloni et al., 2011). At first glance, this could be interpreted as

pointing towards a more preponderant role for somatic mutations accumulation in generating higher inbreeding depression in more long-lived species. However, this hypothesis has been little explored theoretically and empirical investigations have led to conflicting results, so that whether somatic mutations accumulation can account for the increase in inbreeding depression in more long-lived species remains an open question. Besides, our results, together with those of Morgan (2001), can be interpreted differently. Indeed, we show that significant differences in inbreeding depression with respect to life expectancy can be generated by differences in the fitness effect of mutations, provided that selection is strong and varies extensively with life expectancy, as it is the case for mutations affecting survival for instance. This implies that mutations facing sufficiently strong selective pressures, which decrease sharply as life expectancy increases, could in principle generate the increase in inbreeding depression observed in more long-lived species. In the light of this result, we propose that asking on what traits selection is expected to be strong and to decrease significantly as life expectancy increases in plants is a track worth following in order to further our understanding of the mechanisms underlying inbreeding depression.

Mutation load. We showed that more mutations maintained at mutation-selection balance led to a lower mutation load. This result is not a straigthforward consequence of differences in the intensity of selection acting on mutations, in agreement with results obtained by previous authors (Bataillon and Kirkpatrick, 2000). Instead, it is a consequence of differences in the shape of fitness landscapes between species with different life expectancies on the one hand, and of the interaction between the mean phenotype and the intensity of selection on the other hand. This result highlights the fact that incorporating a phenotypic dimension to population genetics studies may lead to counter-intuitive results.

In particular, when fitness landscapes are obtained as the result of biological assumptions, and not arbitrarily assumed to be of a particular shape, unusual interactions with various aspects of species life-histories may result in novel predictions. For instance, in the present study, we showed that the mutation load may behave very differently depending on the trait affected by mutations and life expectancy. Therefore, we conclude that comparing mutation loads between species with contrasting life-histories may sometimes be misleading, and would likely require trait-specific approaches.

Life-history evolution. Plants vary widely in life expectancy and stature, with life expectancies ranging from a few weeks to hundreds, possibly thousands of years, and stature spanning several orders of magnitude across Tracheophytes (Ehrlén and Lehtilä, 2002). These variations are correlated. Indeed, long-lived species tend to grow slower than short-lived species (Salguero-Gómez et al., 2016). In life-history traits evolution theory, this type of correlation is usually interpreted in terms of trade-offs, with populations evolving towards the evolutionarily stable allocation of resources between growth, survival and reproduction, given a number of constraints (Stearns, 1992). In this paper, we have shown that the equilibrium maintenance and production costs differ between life expectancies, causing more long-lived species to grow slower but ultimately larger than more short-lived species, as commonly expected. However, the mechanism underlying this result is completely different. Indeed, in our model, life-history traits do not coevolve in response to trade-offs. Instead, the equilibrium growth costs are modified when life expectancy varies because the selective pressures acting on the many mutations affecting these traits are altered, leading to a more or less efficient purging of said mutations and thereby phenotypic differences at mutation-selection equilibrium. Thus, our results suggest that life-history

traits may sometimes coevolve regardless of trade-offs, because a change in a given trait may alter the efficiency of purging of deleterious mutations affecting other traits.

Epistasis emerging from the fitness landscape. The phenotypic dimension of the genotype-to-phenotype-to-fitness map is usually overlooked in mutation load dynamics studies. Indeed, most theoretical investigations interested in such matters consider mutations affecting fitness multiplicatively (e.g. Kondrashov, 1985; Charlesworth et al., 1990; Roze, 2015), so that fitness is the phenotype affected by mutations and the phenotype-to-fitness map (the fitness landscape) is strictly linear. Hence, neither phenotype-to-fitness epistasis nor genotype-to-phenotype epistasis occur in these models. Some authors however considered the effects of epistasis on mutation load dynamics.

Charlesworth et al. (1991) studied uniformly deleterious mutations under synergistic selection. Because they assumed mutations to affect fitness directly too, no phenotype-to-fitness epistasis occured in their model. On the other hand, genotype-to-phenotype epistasis occured because they assumed mutations to affect fitness synergistically. Abu Awad and Roze (2019) generalised Charlesworth et al. (1991)'s results, showing that their model is equivalent to assuming pairwise fixed negative epistasis between loci. Furthermore, they showed that different forms of pairwise epistasis have different effects: while additive-by-additive epistasis lowers the frequency of mutations by increasing selection, additive-by-dominance and dominance-by-dominance epistasis tend to increase inbreeding depression by making homozygotes less fit.

Phenotype-to-fitness epistasis was also considered in the case of mutations affecting an abstract trait (or set of traits) under stabilising selection (e.g. Abu Awad and Roze, 2018, 2019). These studies assumed mutations to affect the trait(s) additively, so that no

genotype-to-phenotype epistasis occured, but assumed the fitness landscape to be gaussian (or quasi-gaussian), so that phenotype-to-fitness epistasis occured as the landscape was non-linear. However, because the fitness landscape was symetrical with respect to the optimal phenotype and mutations occured in both directions at the same rate, epistasis was null on average. Yet, it influenced mutation load dynamics through variance effects, by generating associations between alleles with compensatory effects, and decreasing the efficiency of selection on deleterious alleles when these associations stay moderate.

In this paper, we studied the dynamics of mutations affecting growth and survival using a physiological growth model (West et al., 2001). Hence, the fitness effect of mutations emerged from biological assumptions instead of being assumed a priori. The resulting fitness landscape had singularities at the optimal phenotypes, as in the absence of trade-offs, immortality ($S = 1$) is the optimal survival strategy and leads to an infinite life expectancy, and null growth costs ($c = \varepsilon = 0$) are the optimal growth strategy, which leads to infinitely large individuals. Furthermore, these optimal phenotypes were situated on the boundaries of the landscape for all three traits, as negative growth costs or survival probabilities greater than one are meaningless. As a consquence, mutations were always under directional selection in our model. We hence focused on the case of mutations increasing growth costs or decreasing survival, that is deleterious mutations, as opposite mutations would always be favoured and eventually reach fixation. Besides, the fitness landscape was convex. Hence, although we assumed mutations to affect phenotypes multiplicatively, so that no genotype-to-phenotype epistasis occured, phenotype-to-fitness epistasis was generated. Contrary to models assuming stabilising selection (e.g. Abu Awad and Roze, 2018, 2019), epistasis was non-zero on average in our model, and its direction

(positive or negative) differed depending on the phenotype affected by mutations. It was however difficult to isolate the effects of epistasis in our model because pairwise epistatic interactions are not sufficient to capture these effects, as considering only pairwise interactions is equivalent to linearising the effect of mutations on the trait average, which only works when very few mutations segregate in the population.

Acknowledgements

The authors thank Sylvain Glémin, Florence Débarre, Vincent Castric and Denis Roze for taking the time to discuss this work. This work was funded by the European Research Council (NOVEL project, grant #648321). The authors also thank the Région Hauts-de-France, and the Ministère de l'Enseignement Supérieur et de la Recherche (CPER Climibio), and the European Fund for Regional Economic Development for their financial support.

References

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

Abu Awad, D. and Roze, D. (2019). Epistasis, inbreeding depression and the evolution of self-fertilization. *BioRXiv*.

Angeloni, F., Ouborg, N., and Leimu, R. (2011). Meta-analysis on the association of popu-

lation size and life history with inbreeding depression in plants. *Biological Conservation*, 144:35–43.

Barrett, S. C. and Harder, L. D. (1996). The comparative biology of pollination and mating in flowering plants. *Philosophical Transactions of the Royal Society of London B*, 351:1271–1280.

Barton, N. and Turelli, M. (1991). Natural and sexual selection on many loci. *Genetics*, 127:229–255.

Bataillon, T. and Kirkpatrick, M. (2000). Inbreeding depression due to mildly deleterious mutations in finite populations: size does matter. *Genetics Research*, 75:75–81.

Bobiwash, K., Schultz, S., and Schoen, D. (2013). Somatic deleterious mutation rate in a woody plant: estimation from phenotypic data. *Heredity*, 111:338–344.

Burian, A., Barbier de Reuille, P., and Kuhlemeier, C. (2016). Patterns of stem cell divisions contribute to plant longevity. *Current Biology*, 26:1385–1394.

Charlesworth, B. (1980). *Evolution in age-structured populations*. Cambridge Studies in Mathematical Biology.

Charlesworth, B., Morgan, M., and Charlesworth, D. (1991). Multilocus models of inbreeding depression with synergistic selection and partial self-fertilization. *Genetics*, 57:177–194.

Charlesworth, D. and Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, 18:237–268.

Charlesworth, D., Morgan, M., and Charlesworth, B. (1990). Inbreeding depression, genetic load, and the evolution of outcrossing rates in a multilocus system with no linkage. *Evolution*, 44:1469–1489.

Charlesworth, D. and Willis, J. (2009). The genetics of inbreeding depression. *Nature reviews genetics*, 10:783–796.

Crow, J. (1958). Some possibilities for measuring selection intensities in man. *Human Biology*, 61:763–775.

DeHaan, L. and Van Tassel, D. (2014). Useful insights from evolutionary biology for developing perennial grain crops. *American Journal of Botany*, 101:1801–1819.

Duminil, J., Hardy, O., and Petit, R. (2009). Plant traits correlated with generation time directly affect inbreeding depression and mating system and indirectly genetic structure. *BMC Evolutionary Biology*, 9:177.

Duputié, A. and Massol, F. (2013). An empiricist's guide to theoretical predictions on the evolution of dispersal. *Interface Focus*, 40:20130028.

Ehrlén, J. and Lehtilä, K. (2002). How perennial are perennial plants? *Oikos*, 98:1070–1072.

Enquist, B., Brown, J., and West, G. (1998). Allometric scaling of plant energetics and population density. *Nature*, 395:163–165.

Epinat, G. and Lenormand, T. (2009). The evolution of assortative mating and selfing with in- and outbreeding depression. *Evolution*, 63:2047–2060.

Franco, M. and Silvertown, J. (1996). Life-history variation in plants: an exploration of the fast-slow continuum hypothesis. *Philosophical Transactions of the Royal Society of London B*, 351:1341–1348.

Husband, B. and Schemske, D. W. (1996). Evolution of the magnitude and timing of inbreeding depression in plants. *Evolution*, 50:54–70.

Kirkpatrick, M., Johnson, T., and Barton, N. (2002). General models of multilocus evolution. *Genetics*, 161:1727–1750.

Klinkhamer, P., Meelis, E., De Jong, T., and Weiner, J. (1985). On the analysis of size-dependent reproductive output in plants. *Functional Ecology*, 6:308–316.

Kondrashov, A. (1985). Deleterious mutations as an evolutionary factor. II. facultative apomixis and selfing. *Genetics*, 111:635–653.

Lanfear, R. (2018). Do plants have a segregated germline? *PloS Biology*, 16.

Morgan, M. (2001). Consequences of life history for inbreeding depression and mating system evolution in plants. *Proceedings of the Royal Society of London B*, 268:1817–1824.

Munoz, F., Violle, C., and Cheptou, P.-O. (2016). CSR ecological strategies and plant mating systems: outcrossing increases with competitiveness but stress-tolerance is related to mixed mating. *Oikos*, 125:1296–1303.

Otto, S. and Orive, M. (1995). Evolutionary consequences of mutation and selection within an individual. *Genetics*, 141:1173–1187.

Peters, R. (1983). *The ecological implications of body size.* Cambridge studies in ecology.

Petit, R. and Hampe, A. (2006). Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, 37:187–214.

Phillips, P. (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews Genetics*, pages 855–867.

Plomion, C., Aury, J.-M., Amselem, J., Leroy, T., Murat, F., Duplessis, S., Faye, S., Francillonne, N., Labadie, K., Le Provost, G., Lesur, I., Bartholomé, J., Faivre-Rampant, P., Kohler, A., Leplé, J.-C., Chantret, N., Chen, J., Diévrat, A., Alaeitabar, T., Barbe, V., Belser, C., Bergès, H., Bodénès, C., Bogeat-Triboulot, M.-B., Bouffaud, M.-L., Brachi, B., Chancerel, E., Cohen, D., Couloux, A., Da Silva, C., Dossat, C., Ehrenmann, F. Gaspin, C., Grima-Pettenati, J., Guichoux, E., Hecker, A., Herrmann, S., Hugueney, P., Hummel, I., Klopp, C., Lalanne, C., Lascoux, M., Lasserre, E., Lemainque, A. Desprez-Loustau, M.-L., Luyten, I., Madoui, M.-A., Mangenot, S., Marchal, C., Maumus, F., Mercier, J., Michotey, C., Panaud, O., Picault, N., Rouhier, N., Rué, O., Rustenholz, C., Salin, F., Soler, M., Tarkka, M., Velt, A., Zanne, A., Martin, F., Wincker, P., Quesneville, H., Kremer, A., and Salse, J. (2018). Oak genome reveals facets of long lifespan. *Nature Plants*, 4:440–452.

Roze, D. (2015). Effects of interference between selected loci on the mutation load, inbreeding depression, and heterosis. *Genetics*, 201:745–757.

Roze, D. and Blanckaert, A. (2014). Epistasis, pleiotropy, and the mutation load in sexual and asexual populations. *Evolution*, 86:137–149.

Roze, D. and Michod, R. (2010). Deleterious mutations and selection for sex in finite, diploid populations. *Genetics*, 184:1095–1112.

Roze, D. and Rousset, F. (2005). Inbreeding depression and the evolution of dispersal rates: a multilocus model. *The Amrican Naturalist*, 166:708–721.

Salguero-Gómez, R., Jones, O. R., Jongejans, E., Blomberg, S. P., Hodgson, D. J., Mbeau-Ache, C., Zuidema, P. A., de Kroon, H., and Buckley, Y. M. (2016). Fast–slow continuum and reproductive strategies structure plant life-history variation worldwide. *Proceedings of the National Academy of Sciences*, 113(1):230–235.

Schmid-Siegert, E., Sarkar, N., Iseli, C., Calderon, S., Gouhier-Darimont, C., Chrast, J., Cattaneo, P., Schütz, F., Farinelli, L., Pagni, M., Schneider, M., Voumard, J., Jaboyedoff, L., Fankhauser, C., Hardtke, C., Keller, L., Pannell, J., Reymond, A., Robinson-Rechavi, M., Xenarios, I., and Reymond, P. (2017). Low number of fixed somatic mutations in a long-lived oak tree. *Nature Plants*, 12:926–929.

Schoen, D. and Schultz, S. (2019). Somatic mutation and evolution in plants. *Annual Review of Ecology, Evolution and Systematics*, 50:2.1–2.25.

Scofield, D. and Schultz, S. (2006). Mitosis, stature and evolution of plant mating systems: low-ϕ and high-ϕ plants. *Proceedings of the Royal Society of London B*, 273:275–282.

Stearns, S. (1992). *The Evolution of Life Histories.* Oxford University Press.

Watson, J., Platzer, A., Kazda, A., Akimcheva, S., Valuchova, S., Nizhynska, V., Nordborg, M., and Riha, K. (2016). Germline replications and somatic mutation accumulation are independent of vegetative life span in arabidopsis. *Proceedings of the National Academy of Sciences*, 113:12226–12231.

Weiner, J., Campbell, L., Pino, J., and Echarte, L. (2009). The allometry of reproduction within plant populations. *Journal of Ecology*, 97:1220–1233.

West, G., Brown, J., and Enquist, B. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276:122–126.

West, G., Brown, J., and Enquist, B. (1999). A general model for the structure, function and allometry of plant vascular systems. *Nature*, 400:664–667.

West, G., Brown, J., and Enquist, B. (2001). A general model for ontogenetic growth. *Nature*, 413:628–631.

Winn, A., Elle, E., Kalisz, S., Cheptou, P.-O., Eckert, C., Goodwillie, C., Johnston, M., Moeller, D., Sargent, R., and Vallejo-Marìn, M. (2011). Analysis of inbreeding depression in mixed-mating plants provides evidence for selective interference and stable mixed mating. *Evolution*, 65:3339–3359.

APPENDIX

I Growth model

At any age t, we assume that an individual is composed of G_t identical units (say, branches or buds), and has a resting metabolic rate B_t. This metabolic rate can be subdivided into energies spent on maintenance and growth as follows (West et al., 2001),

$$B_t = c\ G_t + \varepsilon\ \frac{dG_t}{dt}, \tag{A1}$$

where c is the energy required to maintain a single unit, and ε is the energy required to create a new unit. Moreover, the resting metabolic rate can be expressed as

$$B_t = B_0\ {G_t}^x, \tag{A2}$$

where B_0 is the species' basal metabolic rate and x is the allometric coefficient (Peters, 1983). In vascular plants and animals, the value $x = {}^3/_4$ is generally used in growth models on the basis of both theoretical (e.g. West et al., 1997, 1999), and empirical arguments (e.g. Peters, 1983; Enquist et al., 1998; West et al., 2001), and will thus be used in what follows. Different exponent values would however likely yield qualitatively similar results, provided that they remain positive and smaller than one. Using Equation (A2), Equation (A1) can be rearranged into the following differential equation

$$\frac{dG}{dt} = \frac{1}{\varepsilon}\left(B_0 {G_t}^{\frac{3}{4}} - cG_t\right). \tag{A3}$$

Solving Equation (A3) for G_t, and setting $B_0 = 1$ for convenience, we obtain

$$G_t = \frac{e^{-\frac{c}{\varepsilon}(t+1)}}{c^4}\left(e^{\frac{c}{4\varepsilon}(t+1)} - 1\right)^4. \tag{A4}$$

This growth function naturally saturates, when the energy required to maintain existing units becomes too large for new units to be produced. The size at which it saturates is given by

$$\lim_{t\to\infty} G_t = \frac{1}{c^4}.$$

II General recursions for the effects of selection and reproduction under partial selfing

II.1 Selection

In this paper, we use two different approaches to study our model. The first approach makes the assumption that age-structured populations can be viewed as a rescaled annual population (the Lifetime Fitness approach, LF), so that individuals first undergo lifetime selection, then reproduce. The second method proceeds in a more detailed manner, by studying each step of the life cycle successively (the Life Cycle approach, LC). Irrespective of the method, one may derive a general approximation of the relative fitness for an individual as a function of its genotype during any given selection stage. In this section, we derive such approximation assuming fitness is a function of a trait z, using the theoretical framework introduced by Barton and Turelli (1991) and generalized by Kirkpatrick et al. (2002). We then compute general recursions for the effects of selection on allelic frequencies and genetic associations.

II.1.1 Approximating relative fitnesses

Let X_i and $\overset{*}{X}_i$ be the indicator variables associated with the paternally and maternally inherited alleles at the ith locus. These variables are worth 1 when allele a is present at the position they are associated with, and 0 otherwise, so that for any trait z, assuming mutations affect this trait multiplicatively, an individual's trait value can be written as

$$z = z_0 \times \prod_i \left(1 \pm sh\left(X_i + \overset{*}{X}_i\right) \pm s(1-2h)X_i\overset{*}{X}_i\right), \tag{A5}$$

where $\pm$ is replaced with $+$ or $-$ depending on whether mutations are assumed to increase (e.g. physiological growth costs) or decrease (e.g. survival probability) the trait, s is the phenotypic effect of the mutant allele, and h is its dominance coefficient. The log-value of the trait is then

$$\ln z = \ln z_0 + \sum_i \ln\left(1 \pm sh\left(X_i + \overset{*}{X}_i\right) \pm s(1-2h)X_i\overset{*}{X}_i\right),$$

which, under the assumption that selection is weak, *i.e.* that s is small, can be approximated by

$$\ln z \approx \ln z_0 \pm sh\sum_i\left(X_i + \overset{*}{X}_i\right) \pm s(1-2h)\sum_i X_i\overset{*}{X}_i.$$

Let us now define the centered variables associated with the paternally and maternally inherited alleles at the ith locus, ζ_i and $\overset{*}{\zeta}_i$, as

$$\zeta_i = X_i - \mathbb{E}\left[X_i\right] = X_i - p_i \text{ and, } \overset{*}{\zeta}_i = \overset{*}{X}_i - \mathbb{E}\left[\overset{*}{X}_i\right] = \overset{*}{X}_i - p_i,$$

where $\mathbb{E}\left[X_i\right] = \mathbb{E}[\overset{*}{X}_i] = p_i$ is allele a's frequency at the ith locus. The expectation of products of these centered variables allows one to quantify genetic associations between positions in the genome. For example, the expectation $\mathbb{E}\left[\zeta_i \times \overset{*}{\zeta}_i\right]$ measures to what extent the homozygosity deviates from the panmictic expectation at locus i. To make recursions clearer, let us introduce the condensed notation

$$D_{\mathbb{U},\mathbb{V}} = \mathbb{E}\left[\zeta_{\mathbb{U},\mathbb{V}}\right] = \mathbb{E}\left[\prod_{i\in\mathbb{U}}\left(\zeta_{\mathbb{U}}\right) \times \prod_{i\in\mathbb{V}}\left(\overset{*}{\zeta}_{\mathbb{V}}\right)\right],$$

where $\mathbb{U}$ and $\mathbb{V}$ are sets of positions paternally and maternally inherited, respectively (in the previous example, $\mathbb{U} = \{i\}$ and $\mathbb{V} = \{i\}$), so that the excess in homozygotes at locus i is denoted $D_{i,i}$. Repeated indexes sometimes appear in recursions. They can be dealt with using the relationship

$$D_{\mathbb{U}ii,\mathbb{V}} = (1 - 2p_i) D_{\mathbb{U}i,\mathbb{V}} + p_i q_i D_{\mathbb{U},\mathbb{V}}.$$

Injecting ζ-variables into Equation (II.1.1), it can then be rearranged into

$$\ln z \approx \ln z_0 \pm s \sum_i Z_i$$

with $Z_i = h \sum_i \left(\zeta_i + \overset{*}{\zeta}_i + 2p_i \right) + (1 - 2h) \sum_i \left(\zeta_{i,i} + p_i \left(\zeta_i + \overset{*}{\zeta}_i \right) + p_i^2 \right)$. Hence, the mean log-value is given by

$$\overline{\ln z} = \ln z_0 \pm s \sum_i \overline{Z}_i \text{ with, } \overline{Z}_i = 2h \sum_i p_i + (1 - 2h) \sum_i \left(D_{i,i} + p_i^2 \right),$$

and one may express the trait value z as

$$z = z_0 \; e^{\pm s \sum_i Z_i} = z_0 \; e^{\pm s \sum_i \left(\overline{Z}_i + \Delta Z_i \right)},$$

with $\Delta Z_i = Z_i - \overline{Z}_i$. Therefore, the trait average, $\bar{z}$, is simply

$$\bar{z} \approx z_0 \; e^{\pm s \sum_i \overline{Z}_i}. \tag{A6}$$

All else being fixed, an individual's fitness during a given selection stage, $w(z)$, is a function

of its trait value z. The mean fitness is a constant, given by $\bar{w} = w(\bar{z})$, and the relative fitness $\frac{w(z)}{\bar{w}}$, to second order in s, is given by

$$\frac{w}{\bar{w}}(z) = 1 + \frac{w'(\bar{z})}{\bar{w}} \sum_i \Delta Z_i + \frac{w''(\bar{z})}{2\,\bar{w}} \sum_i \sum_j \Delta Z_i \times \Delta Z_j, \tag{A7}$$

where w' and w'' are the first and second derivatives of w with respect to ΔZ_i. Neglecting terms in $i = j$, and denoting $\bar{s}_z(\bar{z}) = \frac{w'(\bar{z})}{\bar{w}}$ and $\hat{\bar{s}}_z(\bar{z}) = \frac{w''(\bar{z})}{2\,\bar{w}}$, this is

$$\begin{aligned}\frac{w}{\overline{w}}(z) \approx 1 - & \bar{s}_z(\bar{z}) \sum_i \Big(h(\zeta_i + \overset{*}{\zeta}_i) + (1-2h)\,(\zeta_{i,i} - D_{i,i}) \Big) \\ & + \hat{\bar{s}}_z(\bar{z}) \Bigg[h^2 \sum_{i<j} (\zeta_i + \overset{*}{\zeta}_i)(\zeta_j + \overset{*}{\zeta}_j) + h(1-2h) \sum_{i \neq j} (\zeta_i + \overset{*}{\zeta}_i)\,(\zeta_{j,j} - D_{j,j}) \\ & + (1-2h)^2 \sum_{i<j} \Big((\zeta_{i,i} - D_{i,i})\,(\zeta_{j,j} - D_{j,j}) - (D_{ij,ij} - D_{i,i} D_{j,j}) \Big) \Bigg].\end{aligned} \tag{A8}$$

The leading order selection coefficient ($\bar{s}$) encapsulates the effects of selection acting directly on loci, while the second order selection coefficient ($\hat{\bar{s}}$) quantifies the effects of indirect selection between pairs of loci. We neglect higher order selection coefficients.

Allelic frequencies change. Using Equation (A8), the change in allelic frequencies at the ith selected locus owing to selection at a given stage is

$$\Delta^s p_i = \mathbb{E}\left[\frac{w}{\overline{w}} \; \frac{X_i + \overset{*}{X}_i}{2} \right] - p_i,$$

which, to second order in s, yields

$$\begin{aligned}\Delta^s p_i = &- \bar{s}_z\Big(hp_i + (1-h)D_{i,i}\Big) - \bar{s}_z(1-2h)\sum_{j\neq i} D_{ij,j} \\ &+ \hat{\bar{s}}_z(1-h)(1-2h)\sum_{j\neq i}\Big(D_{ij,ij} - D_{i,i}D_{j,j}\Big) + o\left(s^3\right),\end{aligned} \tag{A9}$$

Effect of selection on associations between selected loci. We consider the effect of selection on associations between selected loci to leading order in s (first line in Equation A8). Following selection, we have

$$D^s_{ij,j} = \mathbb{E}\left[\frac{w}{\overline{w}}\ \zeta_{ij,j}\right] - \Delta^s p_i D^s_{j,j} \text{ and, } D^s_{i,i} = \mathbb{E}\left[\frac{w}{\overline{w}}\ \zeta_{i,i}\right] + (\Delta^s p_i)^2\,.$$

To leading order, $\Delta^s p_i$ simplifies to

$$\Delta^s p_i = -\bar{s}_z\Big(h\ p_i - (1-h)D_{i,i}\Big) + o\left(s^2\right).$$

Hence, $D^s_{i,i}$ and $D^s_{ij,j}$ are given by

$$D^s_{i,i} = D_{i,i}\left(1-\bar{s}_z\right) - \bar{s}_z(1-2h)\sum_{j\neq i}\left(D_{ij,ij} - D_{i,i}D_{j,j}\right), \tag{A10}$$

and,

$$D^s_{ij,j} = D_{ij,j} - \bar{s}_z(1-h)\left(D_{ij,ij} - D_{i,i}D_{j,j}\right). \tag{A11}$$

We neglect the effects of selection on $D_{ij,ij}$.

II.2 Reproduction and mutation under partial selfing

As we model infinitely many loci, we may assume that mutations almost never occur two times at the same locus. Since they are not affected by recombination and self-fertilisation, allelic frequencies following reproduction are simply given by

$$p_i^r = p_i + U\left(1 - p_i\right) \approx U + p_i, \tag{A12}$$

where p_i here depicts allelic frequencies just before reproduction. On the other hand, while we neglect the effects of mutation on genetic associations, they are affected by recombination and selfing. Following reproduction, denoting $F = \frac{\alpha}{2-\alpha}$, two-way within loci associations are given by

$$D_{i,i}^r = \frac{F}{1+F}\left(D_{i,i} + p_i q_i\right), \tag{A13}$$

As for three- and four-way associations, we have

$$D_{ij,j}^r = \frac{F}{1+F}\left(D_{ij,j} + (1-r)D_{ij} + rD_{i,j}\right) \approx \frac{F}{1+F} D_{ij,j}, \tag{A14}$$

and,

$$D_{ij,ij}^r = \frac{F}{1+F}\left[(1-R)\left(D_{ij,ij} + p_i q_i p_j q_j\right) + R\left(p_i q_i D_{j,j} + p_j q_j D_{i,i}\right)\right], \text{ with } R = 2r(1-r). \tag{A15}$$

III Fitnesses and selection coefficients

In Appendix II.1, we derived general recursions for the effects of selection at any given selection stage. We showed that to second order, the intensity of selection can be summarised using selection coefficients, $\bar{s}_z$ and $\hat{\bar{s}}_z$. In this section, our aim is to derive the selection coefficients that arise from our model.

III.1 Lifetime fitness approach

III.1.1 Fitness expression

In the LF approach, we make the assumption that all the selective pressures acting on mutations can be summarised into a single lifetime fitness expression. In this section, we compute the lifetime reproductive output of an individual, given its phenotype. This quantity is used as a measure of lifetime fitness.

The size of an individual at age t (G_t) is the solution of Equation (2). Given the individual's maintenance and production costs, c and ε, it is

$$G_t = \frac{e^{-\frac{c}{\varepsilon}(t+1)}}{c^4}\left(e^{-\frac{c}{4\varepsilon}(t+1)} - 1\right)^4.$$

Hence, the contribution to reproduction of an individual up to age τ, $\bar{G}_\tau$, is given by

$$\bar{G}_\tau = \sum_{t=0}^{\tau} G_t.$$

Furthermore, the probability that, given its survival probability S, an individual lives up

to exactly age τ then dies ($\bar{S}_\tau$) is

$$\bar{S}_\tau = S^\tau(1-S).$$

Hence, the lifetime fitness of an individual (W) is given by

$$W = J \times \sum_{\tau=0}^{\infty} \bar{S}_\tau \bar{G}_\tau = \frac{J}{c^4}\left(\frac{1}{1-S} - \frac{4}{e^{\frac{c}{4\varepsilon}}-S} + \frac{6}{e^{\frac{c}{2\varepsilon}}-S} - \frac{4}{e^{\frac{3c}{4\varepsilon}}-S} + \frac{1}{e^{\frac{c}{\varepsilon}}-S}\right), \quad \text{(A16)}$$

where J is the recruitment probability of the individual as a juvenile.

III.1.2 Leading order lifetime selection coefficients

Here, we write down the leading order lifetime selection coefficients obtained with the method described in Appendix II.1. For any trait z, these coefficients are obtained by applying Equation (A7) to Equation (A16). We show that they depend on the ratio $\frac{c}{\varepsilon}$, rather than c and ε absolute values. In what follows, we set $J = 1$ for convenience.

Mutations altering c. The leading order lifetime selection coefficient for mutations altering c is given by

$$\bar{s}_c = s\frac{4V_c - c\frac{\partial V_c}{\partial c}}{V_c} \quad \text{with,} \quad V_c = \frac{1}{1-S} - \frac{4}{e^{\frac{\bar{c}}{4\varepsilon}}-S} + \frac{6}{e^{\frac{\bar{c}}{2\varepsilon}}-S} - \frac{4}{e^{\frac{3\bar{c}}{4\varepsilon}}-S} + \frac{1}{e^{\frac{\bar{c}}{\varepsilon}}-S}. \quad \text{(A17)}$$

Making the variable change $c = \varepsilon\,\frac{c}{\varepsilon}$ in Equation (A17), so that it now depends of $\frac{c}{\varepsilon}$ instead of c, we obtain that

$$\frac{\partial \bar{s}_c}{\partial \varepsilon} = 0,$$

which demonstrates that $\bar{s}_c$ depends only on the ratio $\frac{c}{\bar{\varepsilon}} = \frac{c}{\varepsilon}$. The same reasoning is followed for all three mutation types and will thus not be repeated.

Mutations altering ε. In this case, the leading order lifetime selection coefficient is

$$\bar{s}_\varepsilon = -s\frac{\bar{\varepsilon}}{V_\varepsilon}\frac{\partial V_\varepsilon}{\partial \varepsilon} \quad \text{with,} \quad V_\varepsilon = \frac{1}{1-S} - \frac{4}{e^{\frac{c}{4\bar{\varepsilon}}} - S} + \frac{6}{e^{\frac{c}{2\bar{\varepsilon}}} - S} - \frac{4}{e^{\frac{3c}{4\bar{\varepsilon}}} - S} + \frac{1}{e^{\frac{c}{\bar{\varepsilon}}} - S} \tag{A18}$$

Mutations altering S. Finally, the leading order lifetime selection coefficient for mutations affecting survival is given by

$$\bar{s}_S = s\frac{V_S + S\frac{\partial V_S}{\partial S}}{V_S} \quad \text{with,} \quad V_S = \frac{1}{1-\bar{S}} - \frac{4}{e^{\frac{c}{4\varepsilon}} - \bar{S}} + \frac{6}{e^{\frac{c}{2\varepsilon}} - \bar{S}} - \frac{4}{e^{\frac{3c}{4\varepsilon}} - \bar{S}} + \frac{1}{e^{\frac{c}{\varepsilon}} - \bar{S}} \tag{A19}$$

III.2 Life cycle approach

In the LC approach, selection at different stages is considered separately, thus we denote the selection coefficients associated with selection phase k, when mutations affect trait z, as $\bar{s}_z^k$ and $\hat{\bar{s}}_z^k$. The stages at which selection occurs in our model are reproduction (denoted by $k = g$), juvenile recruitment (denoted $k = j$) and adult survival ($k = s$).

III.2.1 Contribution to reproduction

Selection on contribution to reproduction occurs in all three models ($k = g$). Size at age t (G_t) is the solution of Equation (2). It is given by

$$G_t = \frac{e^{-\frac{c}{\varepsilon}(t+1)}}{c^4}\left(e^{-\frac{c}{4\varepsilon}(t+1)} - 1\right)^4.$$

Thus, the average size of an individual given its genotype, which gives its fitness during this selection phase, is

$$w_g = \sum_{t=0}^{\infty} S^t (1-S)\ G_t = \frac{1-S}{c^4}\left(\frac{1}{1-S} - \frac{4}{e^{\frac{c}{4\varepsilon}} - S} + \frac{6}{e^{\frac{c}{2\varepsilon}} - S} - \frac{4}{e^{\frac{3c}{4\varepsilon}} - S} + \frac{1}{e^{\frac{c}{\varepsilon}} - S}\right). \tag{A20}$$

For any trait z, the selection coefficients at this stage, $\bar{s}_z^g$ and $\hat{\bar{s}}_z^g$ are obtained by applying Equation (A7) to Equation (A20).

III.2.2 Survival and recruitment

Selection on survival occurs in juveniles ($k = j$) and in adults ($k = s$). Individual fitnesses during survival selection stages are given by their survival probability. Hence, when mutations affect the maintenance or the production cost (*i.e.* not survival), we have $\bar{s}_c^j = \bar{s}_c^s = \hat{\bar{s}}_c^j = \hat{\bar{s}}_c^s = 0$ and $\bar{s}_\varepsilon^j = \bar{s}_\varepsilon^s = \hat{\bar{s}}_\varepsilon^j = \hat{\bar{s}}_\varepsilon^s = 0$. On the other hand, when mutations affect survival given by $\bar{s}_S^j = s$ and $\hat{\bar{s}}_S^j = s^2$ in juveniles, and $\bar{s}_S^s = s$ and $\hat{\bar{s}}_S^s = s^2$ in adults.

IV Analytical equilibrium results

At mutation-selection balance, we may obtain leading order approximations for equilibrium genetic associations. These associations differ between our two approaches.

IV.1 QLE associations in the LF approach

Using the LF approach, selection is encapsulated into a leading and second order selection coefficients, $\bar{s}_z$ and $\hat{\bar{s}}_z$. The genetic associations at QLE are computed to leading order. Excess in homozygotes at the ith locus is given by

$$D^*_{i,i} = F\Big[1 - \bar{s}_z\Big(2F + h(1-F)\Big)\Big]p_i, \tag{A21}$$

while three- and four-way associations are given by

$$D^*_{ij,j} = -\bar{s}_z(1-h)\frac{F^2(1-F)(1-R(1-F))}{1+RF}p_ip_j, \tag{A22}$$

and,

$$D^*_{ij,ij} = \frac{F}{1+RF}\Big[1-R(1-2F)-2\bar{s}_z\Big(RF(1+F)+[1+R(1-2F)]\left(F+h(1-F)+\frac{F(1-R)}{1+RF}\right)\Big)\Big]p_ip_j. \tag{A23}$$

IV.2 QLE associations in the LC approach

In the LC approach, we consider selection occurring at the different stages of the life cycle separately. This allows us to incorporate the effects of selection more precisely. In particular, selection occurring before syngamy (*i.e.* in adults), impacts genetic associations

differently than following syngamy (in juveniles). The excess in homozygotes at QLE, at the ith locus, is given by

$$D^*_{i,i} = F\left[1 - \bar{s}^g_z\left(F + h(1-F)\right)\right]p_i, \tag{A24}$$

while other associations are given by

$$D^*_{ij,j} = -(1-h)G\left(\bar{s}^s_z\,\frac{1+F}{1-\bar{S}} + \bar{s}^g_z\,F\right)p_i p_j \tag{A25}$$

and,

$$D^*_{ij,ij} = \frac{F}{1+RF}\Big[1 - R(1-2F) - 2\bar{s}^g_z\Big(h + F(1-h)\Big)\Big(1 - R(1-F)\Big)\Big]p_i p_j, \tag{A26}$$

where $\bar{S}$ is the average survival probability in the population.

IV.3 Solution at mutation-selection balance

The intensity of selection is dependent on the ratio $\frac{c}{\varepsilon}$ and on the survival probability S. Hence, when mutations are numerous, their effect on the population average of the trait they affect has to be accounted for in order to obtain accurate approximations. To do so, one has to find the joint equilibrium of the number of mutations and the average trait value. This can be done by solving

$$\begin{cases} \Delta p_i\left(\bar{z}\right) = 0 \\ \bar{z} = z_0 \times e^{\pm s\sum_i \overline{Z}_i}\left(1 + s^2\sum_{j\neq i}\frac{\overline{\Delta Z_i \times \Delta Z_j}}{2}\right), \end{cases} \tag{A27}$$

for p_i and $\bar{z}$. However, the solution of this system cannot be written explicitly. It can however be approached numerically, by injecting a leading order approximation for $\bar{z}$ into $\Delta p_i(\bar{z})$, and looking for the smallest value of p_i for which $\Delta p_i(\bar{z})$ becomes negative. Technical details are not given here, but the Mathematica scripts we use to obtain numerical approximations are available from GitHub and commented (https://github.com/Thomas-Lesaffre/On_deleterious_mutations_in_perennials). If the reader has any question regarding these methods, feel free to contact the corresponding author.

In the main text, the analytical approximations obtained using the LF method are not presented. they are presented in a lighter color in all the figures of Appendix VII, which present results for a variety of dominance coefficient and growth costs values. The two methods yield the same results to leading order but differ when accounting for genetic associations when mutations affect survival. In this case, the LF approach is less precise than the LC method. This has to do with the timing of selection. Indeed, in this case, selection occurs both before and after syngamy, that is in adults and juveniles, respectively, or in other words before and after homozygosity is changed by self-fertilisation. Hence, the effects of selection on interactions are not accounted for correctly in the LF approach, and the population cannot simply be assimilated to a rescaled annual population.

V Fitness landscapes

In FIGURE S1, lifetime fitness is presented as a function of phenotypic deviation ΔP. For each trait z, we have $z = z_0 (1 + \Delta P)$. The resulting fitness landscape for each set of parameters is rescaled by its maximal value, which is always obtained for $\Delta P = 0$.

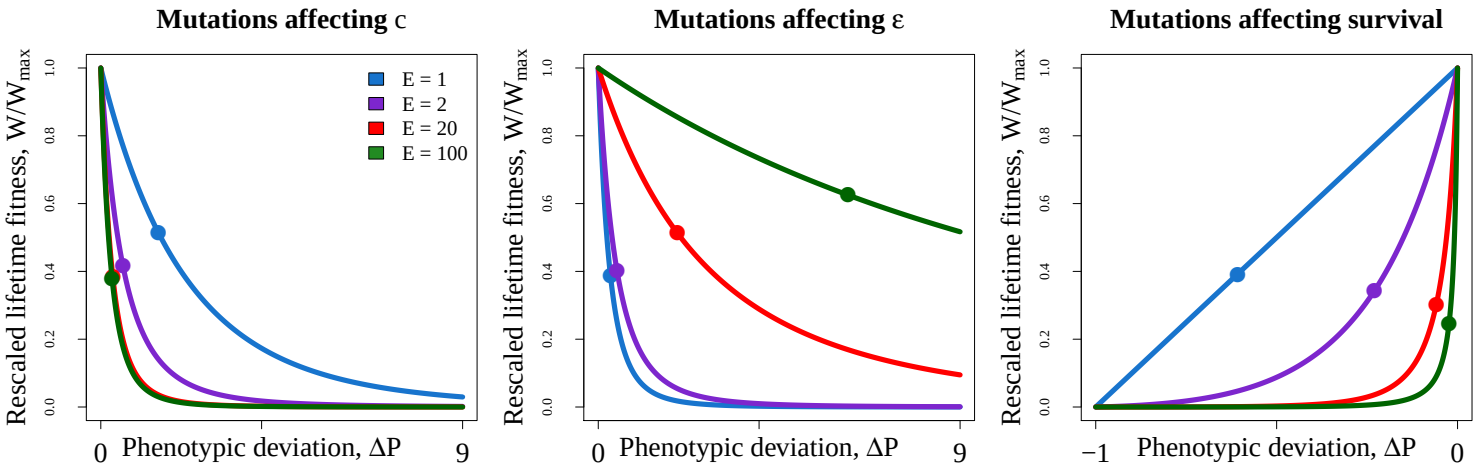

FIGURE S1: Rescaled fitness landscapes for each mutation type, for initial life expectancies ($E = 1;\ 2;\ 20;\ 100$). Lines depict fitness, while dots depict the mutation-selection equilibrium fitness reached for $\frac{c}{\varepsilon} = 1$, $h = 0.25$, $s = 0.005$, and $U = 0.5$.

VI Comparison of mean phenotypic deviation for different magnitude of effect of mutations

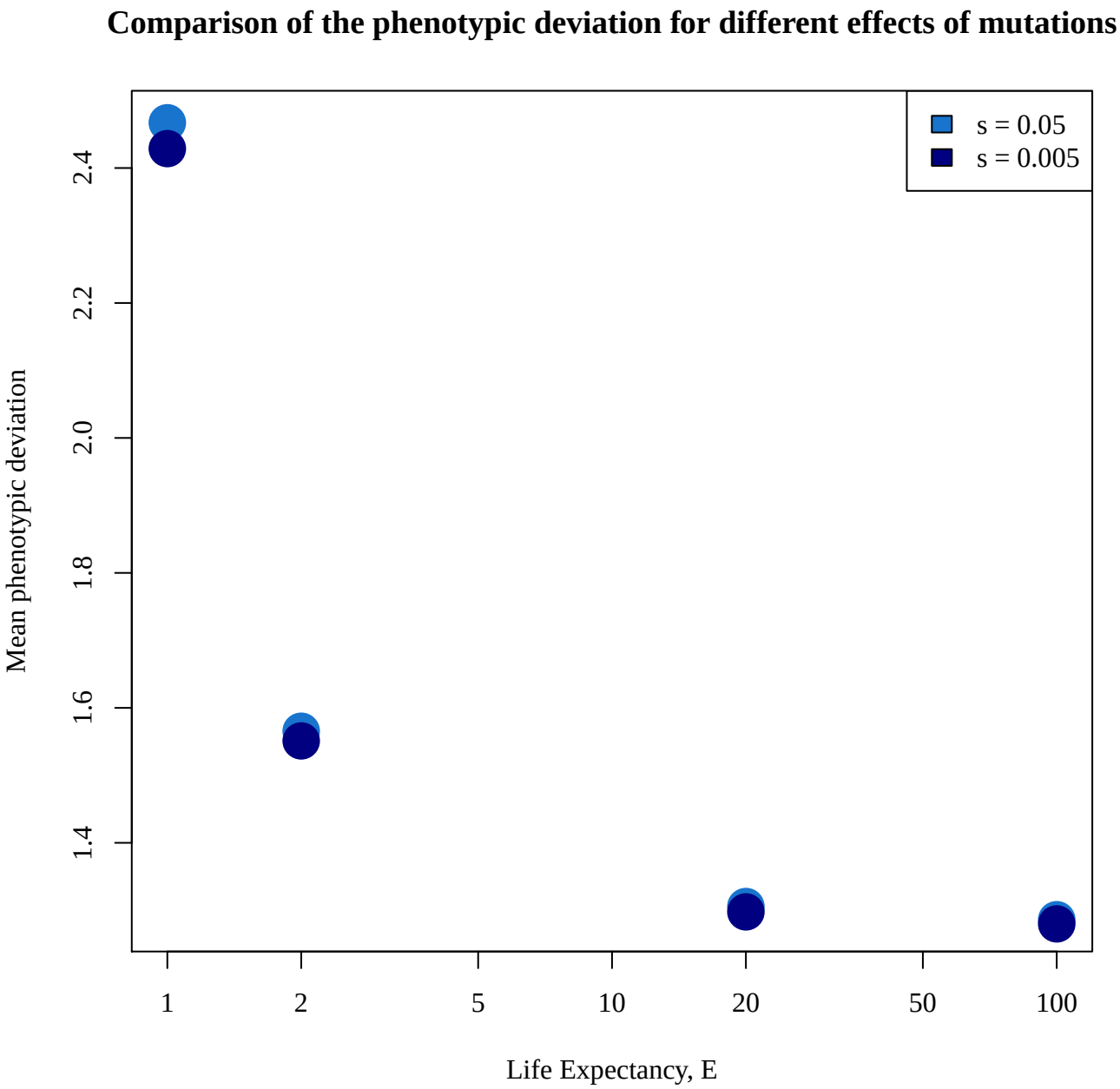

FIGURE S2: Comparison of the mean phenotypic deviation, for mutations affecting the maintenance cost (c), for two different magnitude of effect of mutations ($s = 0.05$ and $s = 0.005$). Parameters used are $\frac{c}{\varepsilon} = 1$, $h = 0.25$, $U = 0.5$.

VII Mutations per haploid genome, inbreeding depression and mutation load for other parameter values

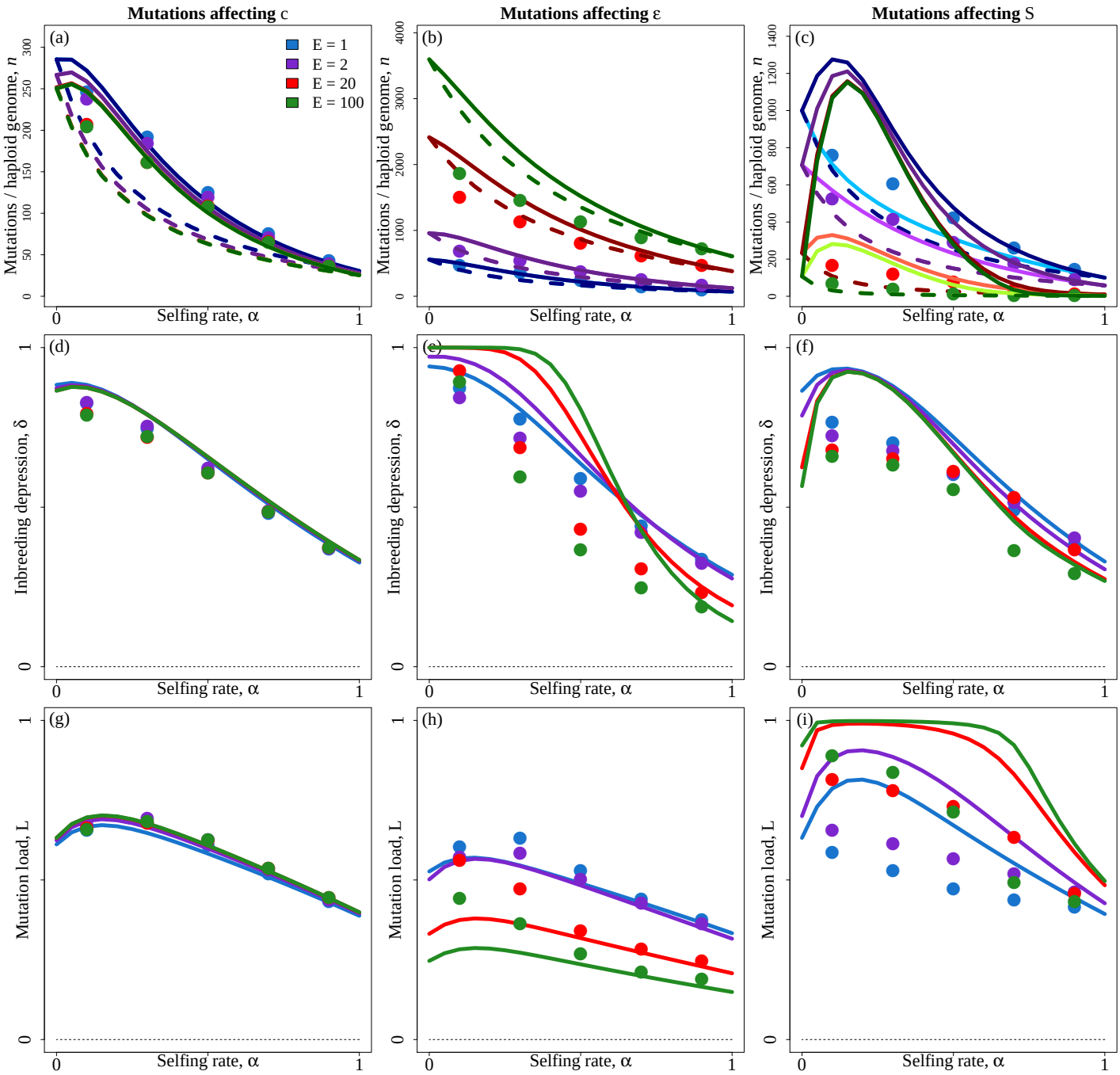

FIGURE S3: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 10$, $U = 0.5$, $s = 0.005$, $h = 0.10$.

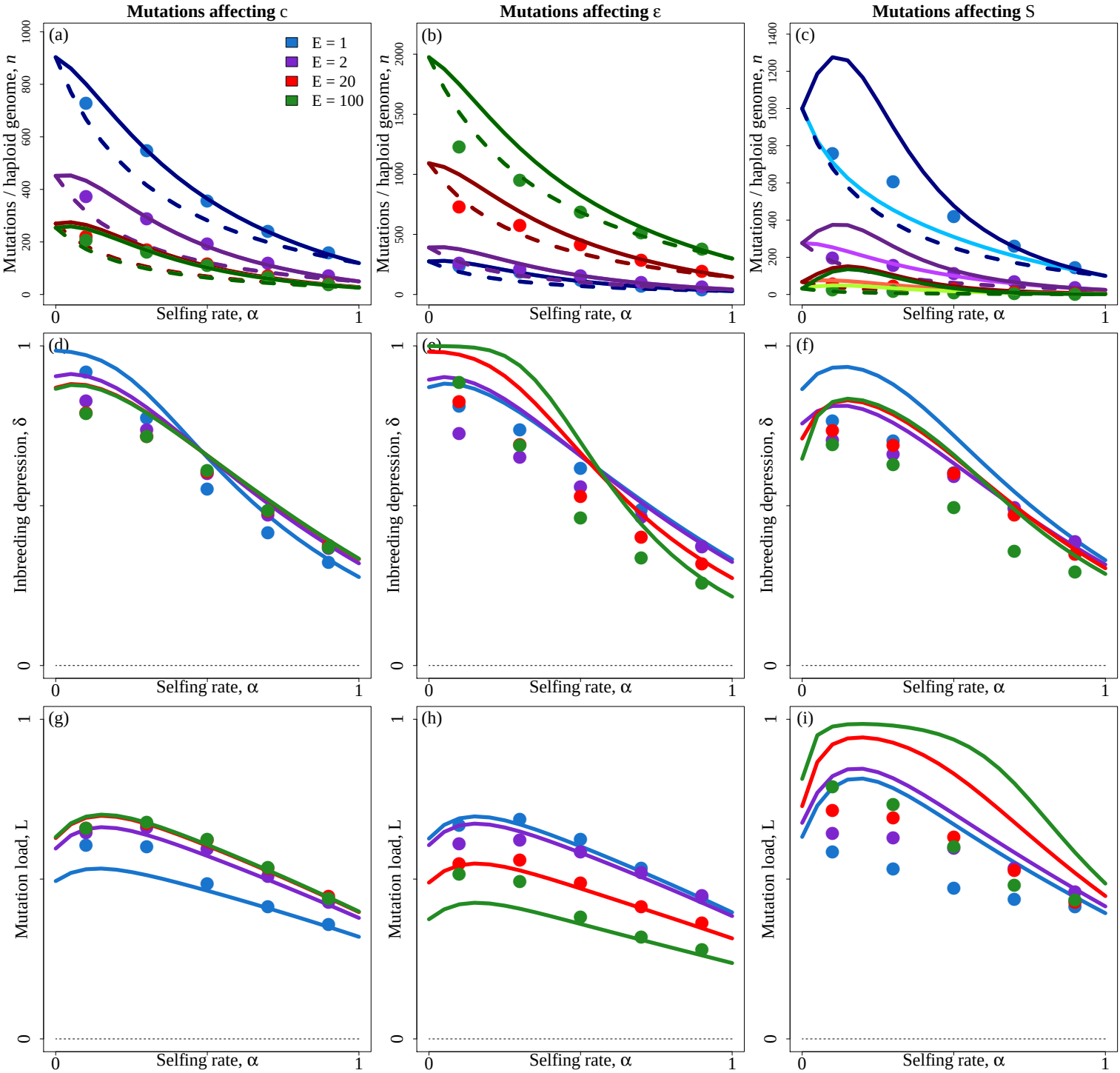

FIGURE S4: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 1$, $U = 0.5$, $s = 0.005$, $h = 0.10$.

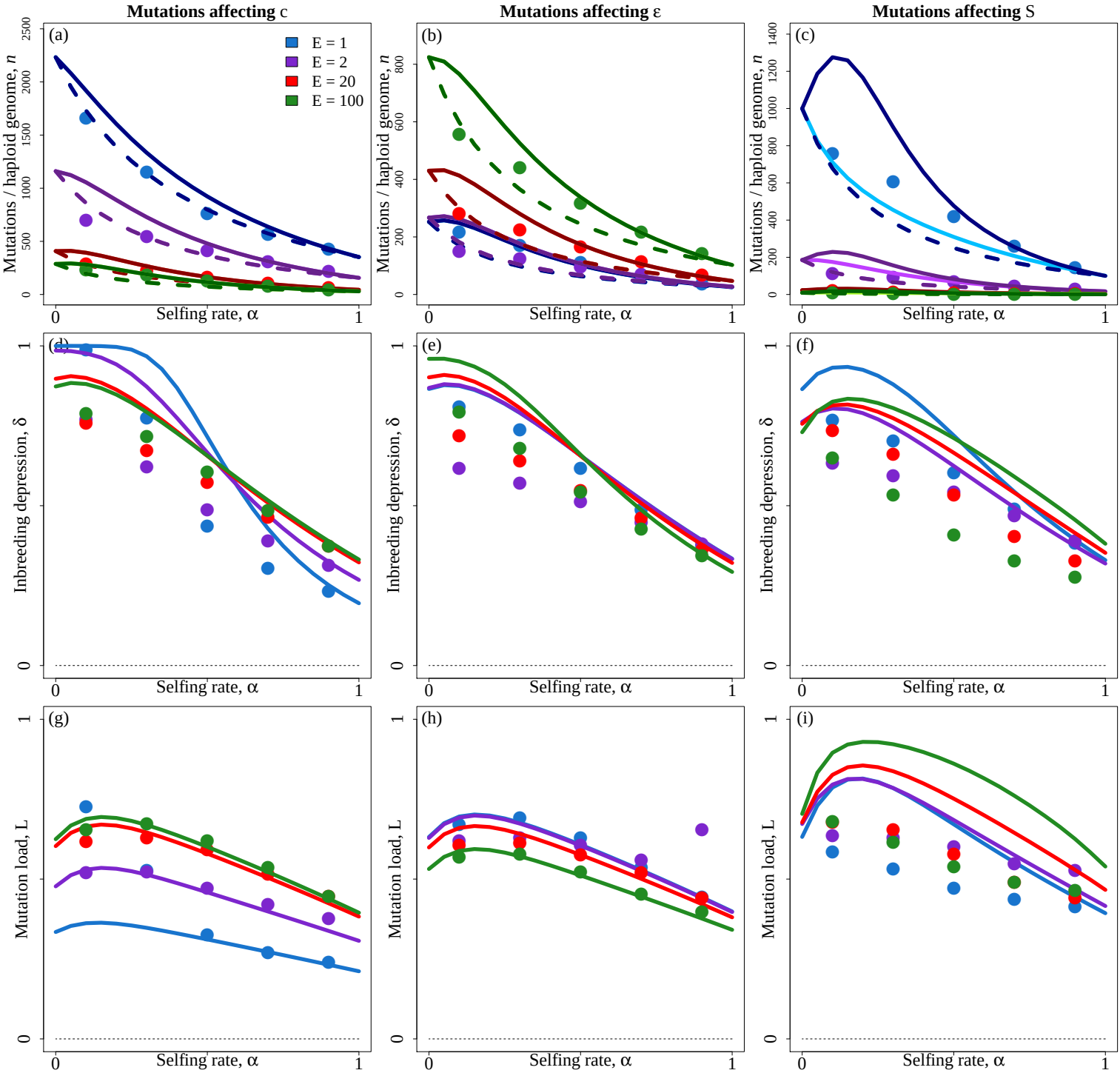

FIGURE S5: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 0.1$, $U = 0.5$, $s = 0.005$, $h = 0.10$.

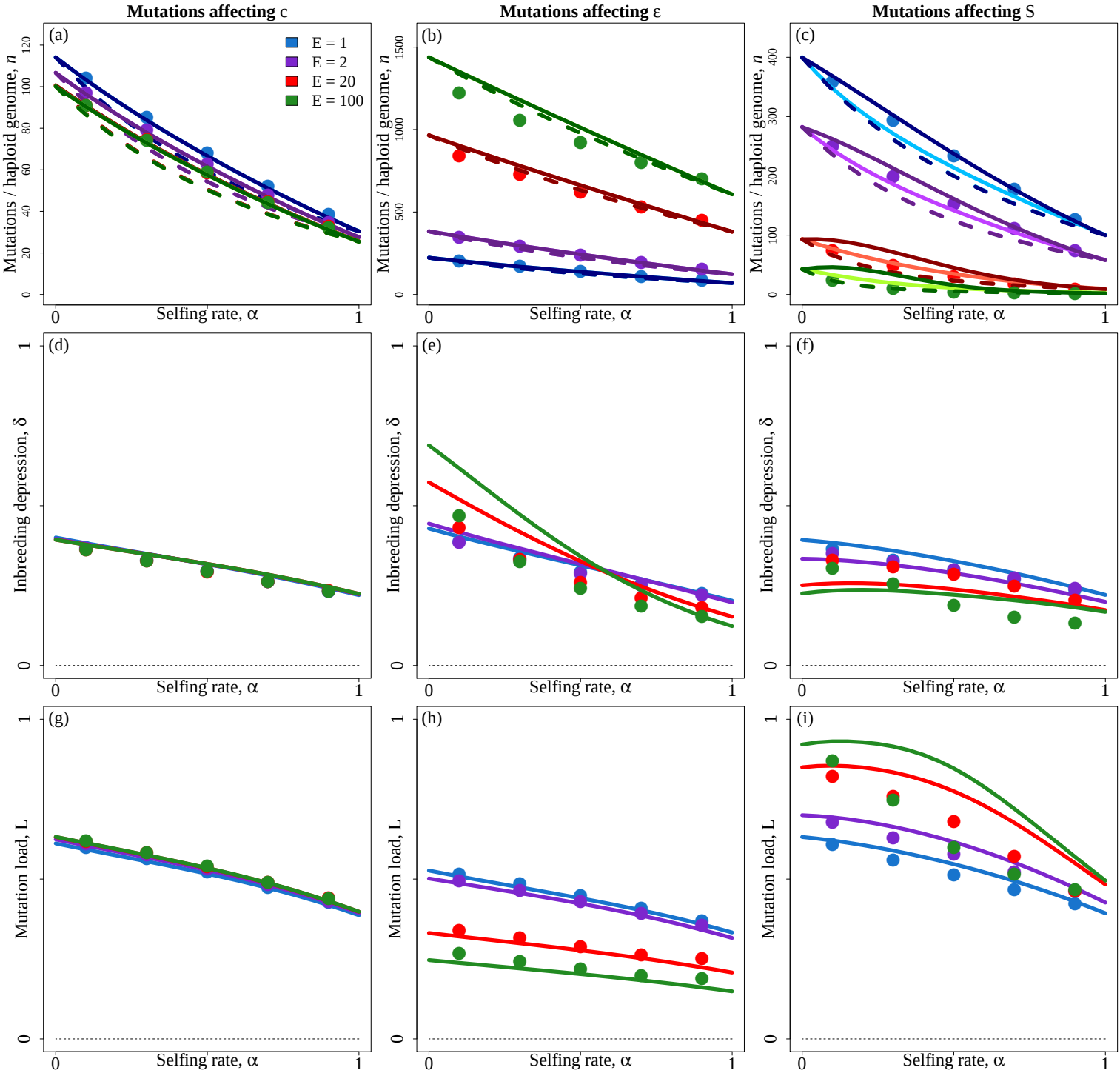

FIGURE S6: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 10$, $U = 0.5$, $s = 0.005$, $h = 0.25$.

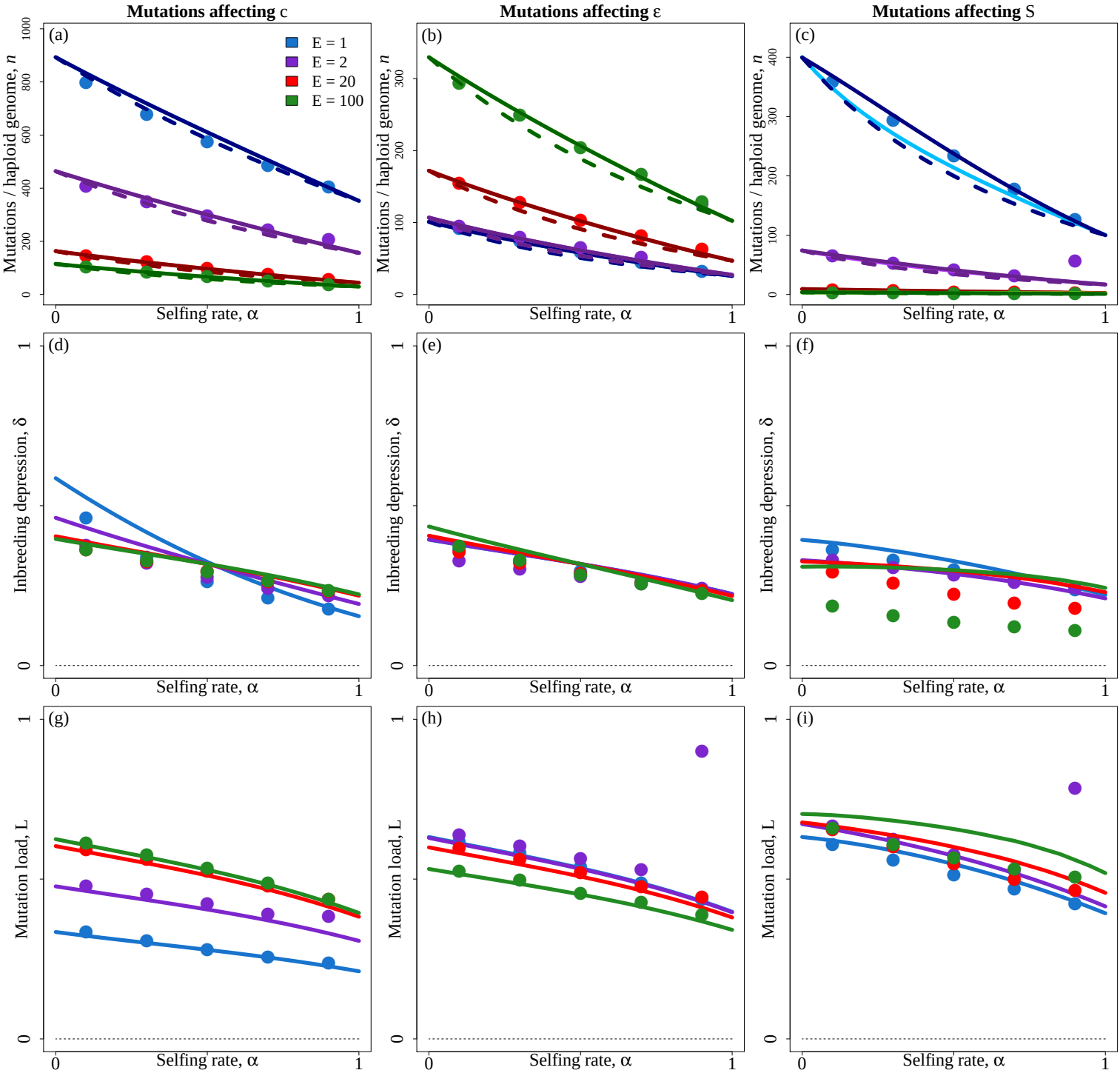

FIGURE S7: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 0.1$, $U = 0.5$, $s = 0.005$, $h = 0.250$.

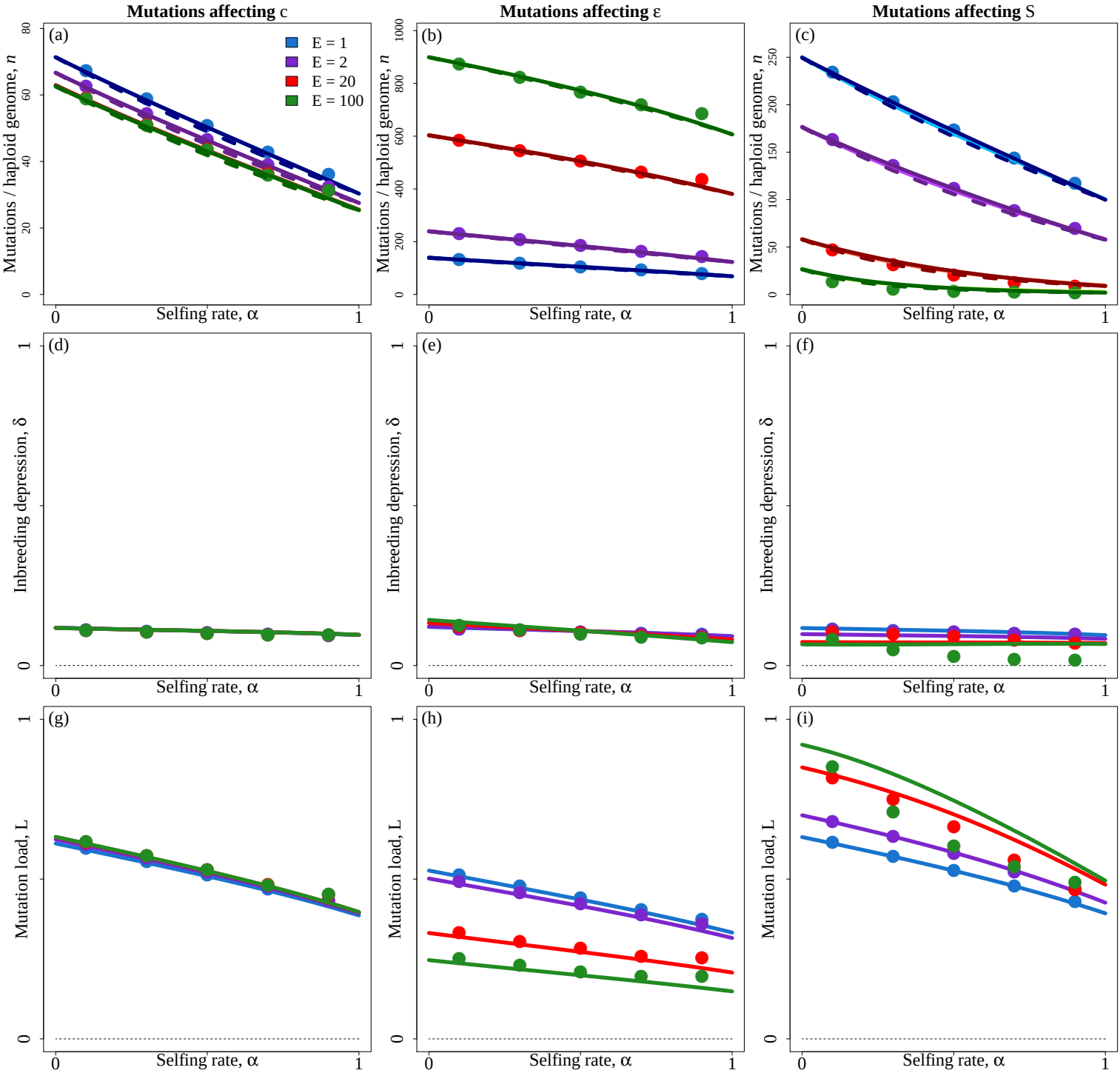

FIGURE S8: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 10$, $U = 0.5$, $s = 0.005$, $h = 0.40$.

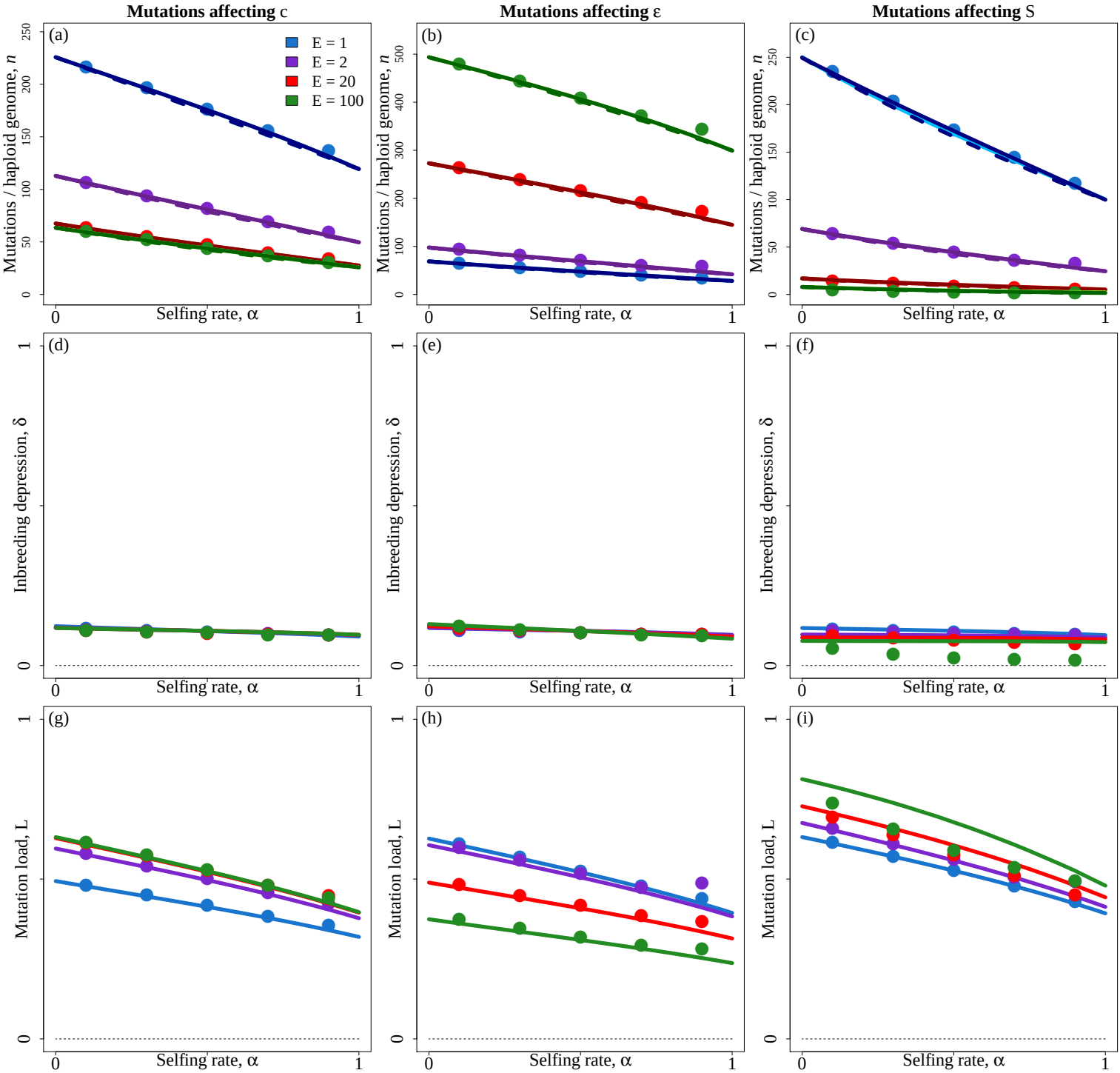

FIGURE S9: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 1$, $U = 0.5$, $s = 0.005$, $h = 0.40$.

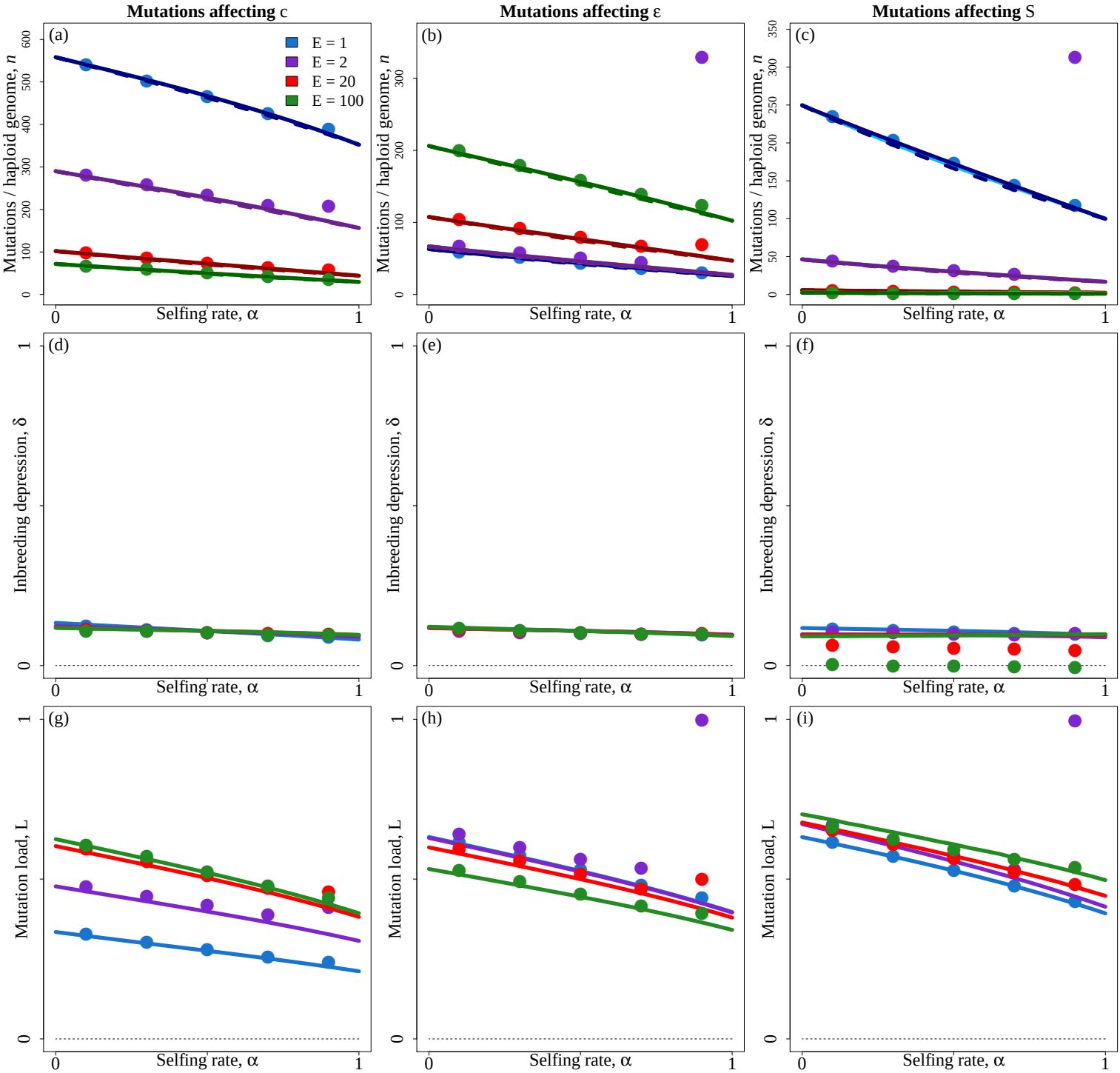

FIGURE S10: Average number of mutations per haploid genome (n, top row) and inbreeding depression (δ, bottom row) as a function of the selfing rate (α), for various life expectancies (colors). Each column corresponds to one type of mutation. Dots: simulation results for $\alpha = 0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9$. Lighter lines: LF approach predictions. Darker lines: LC approach predictions. Dashed lines: leading order approximations (neglecting interactions). Solid lines: approximations accounting for pairwise interactions between loci. Parameters shown here are $\frac{c}{\varepsilon} = 0.1$, $U = 0.5$, $s = 0.005$, $h = 0.40$.