Life history and deleterious mutation rate coevolution

Piret Avila* and Laurent Lehmann

Department of Ecology and Evolution, University of Lausanne

Biophore, 1015 Lausanne, Switzerland

*Corresponding author: piret.avila@gmail.com.

Author contributions: All authors gave final approval for publication and each agreed to be held accountable for all the parts of the work performed therein.

Acknowledgements: -

Data Accessibility Statement: The proofs of the analysis written in Mathematica and the code for individual-based simulations are available in the S.I. as a Mathematica notebook.

Conflict of Interest Statement: All authors declare that they have no conflicts of interest.

Abstract

The cost of germline maintenance gives rise to a trade-off between lowering the deleterious mutation rate and investing in life history functions. Therefore, life history and the mutation rate evolve jointly, but this coevolution is not well understood. We develop a mathematical model to analyse the evolution of resource allocation traits affecting simultaneously life history and the deleterious mutation rate. First, we show that the invasion fitness of such resource allocation traits can be approximated by the basic reproductive number of the least-loaded class; the expected lifetime production of offspring without deleterious mutations born to individuals without deleterious mutations. Second, we apply the model to investigate (i) the joint evolution of reproductive effort and germline maintenance and (ii) the joint evolution of age-at-maturity and germline maintenance. This analysis provides two biological predictions. First, under higher exposure to environmental mutagens (e.g. oxygen), selection favours higher allocation to germline maintenance at the expense of life history. Second, when exposure to environmental mutagens is higher, life histories tend to be faster with individuals having shorter life spans and smaller body sizes at maturity. Our results suggest that mutation accumulation via the cost of germline maintenance is a major force shaping life-history traits.

Keywords: life-history evolution, mutation accumulation, adaptive dynamics, cost of fidelity, mutation rate evolution

1 Introduction

Maintaining and accurately transmitting genetically encoded information is central for every living organism. Mutations induce errors in the processing of genetic information and their effects on fitness are generally (mildly) deleterious (Eyre-Walker and Keightley, 2007). Therefore, it is likely that selection primarily favours a reduction of the mutation rate of organisms (Sniegowski et al., 2000). Yet, investing resources into germline maintenance is physiologically costly (Kirkwood, 1986; Maklakov and Immler, 2016; Monaghan and Metcalfe, 2019; Chen et al., 2020). Therefore, the balance between selection against deleterious mutations driving for lower mutation rates and selection for reduced physiological cost increasing mutation rates is expected to lead to a positive evolutionary equilibrium rate of mutation. This argument has been formalised in a number of classical population genetic models assuming semelparous reproduction (e.g. Kimura, 1967; Kondrashov, 1995; Dawson, 1998, 1999; Johnson, 1999*b*; André and Godelle, 2006; Gervais and Roze, 2017) as well iteroparous reproduction (Lesaffre, 2021) to show that evolution indeed favours an evolutionary stable non-zero mutation rate. These studies emphasise the role of the physiological cost of germline fidelity in explaining the patterns of mutation rates but do not connect the cost of germline fidelity explicitly to life-history evolution, which depends on underlying physiological trade-offs.

Indeed, a central premise made in life-history theory is that life-history trade-offs are mediated through the allocation of resources to different life-history functions, such as growth, reproduction, and maintenance of soma, or information gathering and processing (Stearns, 1992; Roff, 2008). Since germline maintenance takes a toll on available resources, mutation rate evolution and life-history evolution are tied together through a resource allocation trade-off. This implies that the rate of deleterious mutations should evolve jointly with life history and affects various life-history traits, such as reproductive effort, age-at-maturity, adult body size, and expected lifespan. Yet the bulk of models about the evolution of mutation rates, which often go under the heading of modifier models, consider physiologically neutral mutation rate (e.g. Leigh, 1970; Gillespie, 1981; Holsinger and Feldman, 1983; Liberman and Feldman, 1986; Gerrish et al., 2007; Altenberg et al., 2017*a*; Baumdicker et al., 2020). And while the effect of fixed mutation rate on life-history trade-offs has been studied before (e.g. Charlesworth, 1990; Dańko et al., 2012), these models suggest that a relatively high level of mutation rates is needed for mutation accumulation to alter life-history trade-offs. This led to the conclusion that mutation accumulation is a minor force in shaping life-history traits (Dańko et al., 2012). But these studies view mutation rates as fixed traits acting only as a hindrance to adaptive life-history evolution. Hence, no study has yet investigated the joint evolution of both life-history and deleterious mutation rates via allocation to germline maintenance.

The aim of this paper is to start to fill this gap by formally extending evolutionary invasion analysis–"ESS" theory–(e.g., Maynard Smith, 1982; Eshel and Feldman, 1984; Metz et al., 1992; Charlesworth, 1994), which has been routinely applied to life-history evolution (e.g., León, 1976; Michod, 1979; Schaf-

fer, 1982; Iwasa and Roughgarden, 1984; Stearns, 1992; Perrin, 1992; Perrin and Sibly, 1993; Cichon and Kozlowski, 2000; Iwasa, 2000; Day and Taylor, 2000), to the case where life-history trait(s) evolving by selection also control the rate of deleterious mutations. This covers the situation where life-history resource allocation schedules evolve on a background where deleterious mutation accumulation can occur. Our formalisation thus aims to integrate both the standard theories of life-history evolution and deleterious mutation rate accumulation.

The rest of this paper is organised into two parts. First, we characterise the invasion process of a mutant life-history trait affecting the load of deleterious mutation accumulation in an age-structured population. We show that ascertaining the joint evolutionary stability of life-history schedules and mutation rates of deleterious mutations is usually complicated, but under certain biologically feasible conditions; in particular, when the zero mutation class (least-loaded class) dominates the population in frequency, evolutionary stability can be ascertained from the basic reproductive number of the least-loaded class alone (expected lifetime production of offspring with no mutations born to individuals with no mutations). Second, we analyse and solve two concrete biological scenarios: (i) joint evolution between reproductive effort and germline maintenance, where individuals face a trade-off between allocating resources to survival, reproduction and germline maintenance and (ii) joint evolution between age-at-maturity and germline maintenance, where individuals face a trade-off between allocating resources to growth, reproduction and germline maintenance. These scenarios allow us to illustrate how our model can be applied to analyse questions in life history and mutation rate evolution and provide predictions about how life history and mutation rate evolve jointly. It also allows us to verify that the analysis based on the basic reproductive number of the least-loaded class as an invasion fitness proxy is a useful approximation as analytical predictions match closely with the results from individual-based stochastic simulations.

2 Model

2.1 Main biological assumptions

We consider a panmictic population of haploid individuals reproducing asexually. The population size is assumed to be large enough to neglect the effect of genetic drift and is regulated by density-dependent competition. Individuals in the population are structured into age classes and each individual undergoes birth, possibly development, reproduction, and death. Age classes can be either continuous or discrete, and in the discrete case, an individual in the age range $[a-1, a]$ for $a = 1, 2, 3, ...$, will, by convention, belong to the a-th age class and so we begin counting discrete age classes with 1 (one of the two possible conventions to count discrete age classes, Fig. 3.1 Case, 2000). An individual of age class a is characterised by a type $\theta(a) = (\boldsymbol{u}(a), n_{\mathrm{m}}(a))$, which consists of two genetically determined components (see Table 1 for a list of symbols and more formalities). The first component, $\boldsymbol{u}(a)$, is the individual's life-history trait expressed in age class a; namely, a resource allocation decision to different life-history functions (e.g. growth, reproduction, or somatic maintenance, see e.g. Stearns, 1992; Perrin, 1992; Perrin and Sibly,

1993; Day and Taylor, 2000). We denote by $\boldsymbol{u} = \{\boldsymbol{u}(a)\}_{a\in\mathcal{T}}$ the whole life-history schedule or a path over all possible age classes $\mathcal{T}$ an individual can be in (formally, $\boldsymbol{u} \in \mathcal{U}[\mathcal{T}]$, where $\mathcal{U}[\mathcal{T}]$ is the set of all admissible life-history schedules with domain $\mathcal{T}$, which for for discrete age classes is $\mathcal{T} = \{1, 2, .., T\}$ and for continuous age classes is $\mathcal{T} = [0, T]$ when the maximum lifespan is T). The second component, $n_{\mathrm{m}}(a)$, represents the number of deleterious germline mutations of an individual belonging to the a-th age class, which, by definition, affects negatively viability traits (e.g. physiology, reproduction). Hence, $n_{\mathrm{m}}(a)$ can be thought of as the load of deleterious mutations as considered in classical population genetic models of mutation accumulation (e.g., Kimura and Maruyama, 1966; Haigh, 1978; Dawson, 1999; Bürger, 2000), but here extended to an age-structured model (see also Steinsaltz et al., 2005 for aged-structured mutation accumulation model). As such, $n_{\mathrm{m}} = \{n_{\mathrm{m}}(a)\}_{a\in\mathcal{T}}$ represents the profile of deleterious mutations across the lifespan.

We envision that the genotype determining the type $\theta = \{\theta(a)\}_{a\in\mathcal{T}} = (\boldsymbol{u}, n_{\mathrm{m}})$ of an individual consists of two separate positions (or loci) that are necessarily linked under asexual reproduction, one locus determining $\boldsymbol{u}$ and the other n_{m} (see Fig. 1). The mutation rate at the life-history trait $\boldsymbol{u}$ locus is assumed to be fixed and is thus exogenously given. However, this trait is assumed to control allocation to germline maintenance and other life-history functions, and thus to control the mutation rate at the locus where the n_{m} deleterious mutations accumulate whose number are thus endogenously determined (Fig. 1).

We assume that the effective number of births $b(a)$ and the death rate $d(a)$ of an individual of age a can depend on both the number of deleterious mutations $n_{\mathrm{m}}(a)$ at age a and the life-history schedule $\boldsymbol{u}$. The life history schedule $\boldsymbol{u}$ thus not only affects the vital rates, as usually assumed in life-history theory but also the rate of deleterious mutation accumulation (see also Fig. 1). The vital rates may further depend on properties of the population, such as age class densities and allele frequencies, but we leave this dependence implicit. We note that when measured on an exponential scale, the death and mutation rates define a survival probability $\exp(-d(a))$ and an immutability probability $\exp(-\mu(a))$ (probability that no mutations occur in an individual of age a). Note that for a discrete-time process, the birth function $b(a)$ gives the expected number of offspring produced by an individual of age a, while for a continuous time process, $b(a)$ is defined as the rate at which an individual produces a single offspring. Finally, we make the following assumptions that are central to our analysis.

1. Mutations at the locus determining n_{m} can only be deleterious or neutral. The effective birth rate $b_i(a)$, survival probability $\exp(-d_i(a))$, and immutability probability $\exp(-\mu_i(a))$ of an individual at age a with $n_{\mathrm{m}}(a) = i$ mutations are non-increasing functions of the number of mutations. Formally, $b_i(a) \geq b_{i+1}(a)$, $d_i(a) \leq d_{i+1}(a)$, and $\mu_i(a) \leq \mu_{i+1}(a)$.

2. Deleterious mutations can only accumulate within an individual. There are no back mutations and an individual with i mutations can only mutate towards having $i + 1$ mutations.

These assumptions are standard in population genetics (e.g. Kimura, 1967; Leigh, 1970; Haigh, 1978; Dawson, 1998; Johnson, 1999*a*; Gillespie, 2004) and we are here endorsing these assumptions in formal-

ising selection at the life-history locus. Namely, the objective of our analysis is to develop a tractable evolutionary invasion analysis to evaluate candidate evolutionary stable life-history trait $u^* \in \mathcal{U}[\mathcal{T}]$ that will be favoured by long-term evolution.

2.2 Invasion analysis

For a moment let us ignore the the effect of deleterious mutations. Then, evolutionary invasion analysis (e.g., Eshel and Feldman, 1984; Parker and Maynard Smith, 1990; Metz et al., 1992; Charlesworth, 1994; Ferrière and Gatto, 1995; Eshel, 1996; Otto and Day, 2007; Mullon et al., 2016; Avila and Mullon, 2023) can be applied straightforwardly to our model as follows. Mutations at the life-history trait are postulated to be rare so that one can focus on the invasion fitness $\rho(\boldsymbol{u}, \boldsymbol{v})$ of a mutant trait $\boldsymbol{u}$ introduced in a population monomorphic for some resident life-history trait $\boldsymbol{v}$ and that has reached its demographic equilibrium. The invasion fitness can be interpreted as the per capita number of mutant copies produced per unit time by the mutant lineage descending from a single initial mutation when the mutant reproductive process has reached a stationary state distribution (over ages and the number of deleterious mutations) in a resident population and the mutant remains overall rare (i.e., the geometric growth ratio of the mutant). To invade the population, the invasion fitness must exceed unity $\rho(\boldsymbol{u}, \boldsymbol{v}) > 1$; otherwise, the mutant will go extinct with probability one in an (infinitely) large population (Harris, 1963; Mode, 1971). A trait value $\boldsymbol{u}^*$ will be called uninvadable if it is resistant to invasion by all trait values in the trait space: $\rho(\boldsymbol{u}, \boldsymbol{u}^*) \leq 1$ for all $\boldsymbol{u} \in \mathcal{U}[\mathcal{T}]$. Hence, an uninvadable trait $\boldsymbol{u}^*$ maximises the invasion fitness of the mutant holding the resident at the uninvadable population state; namely, $\boldsymbol{u}^*$ solves the maximization problem: $\max_{\boldsymbol{u} \in \mathcal{U}[\mathcal{T}]} \rho(\boldsymbol{u}, \boldsymbol{u}^*)$, whereby $\boldsymbol{u}^*$ defines a candidate endpoint of evolution.

To be a meaningful endpoint of evolution, an uninvadable trait needs to be an attractor of the evolutionary dynamics and thus be convergence stable (Eshel and Motro, 1988; Geritz et al., 1998; Leimar, 2009*a*). When mutant and resident traits are sufficiently close to each other so that selection is weak, mutants increasing invasion fitness will not only invade the resident population but also become fixed in it (Eshel et al., 1997, a result applying to the present and more complicated demographic scenarios, Rousset, 2004; Priklopil and Lehmann, 2021). Therefore, invasion fitness alone under weak phenotypic deviations allows to determine whether gradual evolution under the constant influx of mutations will drive a population towards the uninvadable population state, regardless of patterns of frequency-dependence. An evolutionary invasion analysis thus generally consists of both using invasion fitness to (i) characterise uninvadable trait values and (ii) determine whether these are attractors of the evolutionary dynamics, thus allowing to make definite statements about joint evolutionary dynamics. Useful summaries of these concepts can be found in Geritz et al. (1998); Leimar (2009*b*) and individual-based stochastic simulations have repeatedly validated the conclusions of this approach in genetic explicit contexts (e.g., Mullon et al., 2018; Mullon and Lehmann, 2019, see also Otto and Day, 2007; Dercole and Rinaldi, 2008 for textbook discussions). This approach to characterise stable and attainable life histories has indeed been in use more or less explicitly in standard life-history theory for decades and underlies most life-history evolution

results (e.g., Hamilton, 1967; León, 1976; Michod, 1979; Schaffer, 1982; Iwasa and Roughgarden, 1984; Perrin, 1992; Perrin and Sibly, 1993; Charlesworth, 1994; Day and Taylor, 2000; Cichon and Kozlowski, 2000; Iwasa, 2000; Irie and Iwasa, 2005; Rueffler et al., 2013; Metz et al., 2016; Avila et al., 2019). We next push forward this approach into the realm where life-history evolution interacts with mutation accumulation and thus relax the standard life-history theory assumption that the rate of deleterious mutations is zero.

2.3 Invasion analysis with mutation accumulation

2.3.1 Mutation-selection balance in the resident population

Let us now consider that deleterious mutations can accumulate. We still assume that mutations at the life-history locus are rare enough so that whenever a mutant trait $\boldsymbol{u}$ arises, it does so in a resident population monomorphic for some resident life-history trait $\boldsymbol{v}$. But owing to the occurrence of deleterious mutations, the resident population will be polymorphic for the number of deleterious mutations in the n_{m} locus, and this polymorphism will depend on $\boldsymbol{v}$. The resident population is then assumed to have reached a mutation-selection equilibrium for deleterious mutations and the resident trait $\boldsymbol{v}$ thus determines a resident probability distribution $\mathbf{p}(\boldsymbol{v})$ over the different number of deleterious mutations carried by individuals across the different age-classes. This assumption is nothing else than the usual assumption of the internal stability of the resident population used in invasion analysis (see e.g. Altenberg et al., 2017*a*; Eshel and Feldman, 1984; Metz et al., 1992). Here, it entails that the resident population has reached an equilibrium for both demographic and genetic processes.

In the absence of age classes, $\mathbf{p}(\boldsymbol{v})$ is the equilibrium probability distribution for the number of deleterious mutations in standard selection-mutation balance models (see Bürger, 2000 for a general treatment and where the sample space of $\mathbf{p}(\boldsymbol{v})$ reduces to $\mathbb{N}$). For instance, when the number of novel (deleterious) mutations follows a Poisson distribution with the mean rate μ and each additional mutation decreases baseline fecundity by a constant multiplicative factor σ in a semelparous population, then $\mathbf{p}(\boldsymbol{v})$ is Poisson distributed with mean μ/σ (Haigh, 1978, Bürger, 2000, p. 300). This holds in an age-structured population across age classes under a certain but limited number of conditions (Steinsaltz et al., 2005). More generally, $\mathbf{p}(\boldsymbol{v})$ will depend on the details of the model.

The key feature of the invasion process of a mutant $\boldsymbol{u}$ in a resident population $\boldsymbol{v}$ is that it depends on which mutational background it arises, which in turn depends on the distribution $\mathbf{p}(\boldsymbol{v})$. Therefore characterising the invasion fitness of a mutant is not straightforward (see Appendix A.1 for the details of the invasion process with mutation accumulation). This is because under the assumptions about the mutation process of our model, the invasion fitness of the life-history trait depends on how many deleterious mutations the individual carrying that trait has (formally, the invasion process is reducible, see Appendix A.1 and also Altenberg, 2009, p. 1278). In Appendix A.1 we show that the invasion fitness of other classes can not be higher than the invasion fitness of the least-loaded class. And thus, as a first step, it is reasonable to consider a situation where the mutation-selection process is such that the least-

loaded class–individuals with zero mutations–dominates the population in frequency (i.e. the frequency of the zero-class individuals is close to one). If selection is stronger than mutation, then deleterious alleles will tend to be purged and the mutation-selection balance will be far away from the error threshold of mutation accumulation or meltdown of asexual populations (e.g., Eigen, 1971; Lynch et al., 1993; Szathmary and Maynard Smith, 1997). For instance, in the aforementioned classical mutation-selection equilibrium model (Haigh, 1978, Bürger, 2000), the frequency of the zero mutation class is $e^{-\mu/\sigma}$. So when $\mu \ll \sigma$, say for definiteness the selection coefficient is one order of magnitude larger than the mutation rate (e.g. for $\mu = 0.01$ and $\sigma = 0.1$, $\mu/\sigma = 0.1$), then the least loaded class dominates in frequency ($e^{-\mu/\sigma} \approx 0.9$). Under these conditions, the click rate of Muller's ratchet, which is the rate at which individuals with the least amount of deleterious mutations in the population become extinct, is small for finite but sufficiently large population sizes. For instance, in a population of size $N = 1000$, the click rate is 8.4×10^{-34} (obtained from $1/\tau$ where $\tau = \sigma\sqrt{2\pi N/\mu} \times \exp\Big(N\left(\sigma - \mu\left(1 - \log\left(\frac{\mu}{\sigma}\right)\right)\right)\Big)/(\sigma - \mu)^2$ is the inverse of the click rate, see eq. 23 Metzger and Eule, 2013, where $\sigma = s$ and $\mu = u$). Hence, the click rate of Muller's ratchet can be considered negligible compared to the scale of mutation rates. Thus, whenever the selection coefficient is one order of magnitude larger than the mutation rate, whenever a mutant life-history trait $\boldsymbol{u}$ appears in a resident $\boldsymbol{v}$ population, it is likely to arise on a zero mutation background.

2.3.2 Invasion analysis for dominating least loaded class

Now fully endorsing the assumption that the least-loaded class dominates in frequency the resident population, which allows to characterise uninvadability and convergence stability directly from the basic reproductive number of the least-loaded class $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$. This is the expected number of offspring without mutations produced by an individual without mutations over its lifespan and is a fitness proxy for invasion fitness $\rho(\boldsymbol{u}, \boldsymbol{v})$ whereby an uninvadable strategy $\boldsymbol{u}^*$ solves the maximization problem $\max_{\boldsymbol{u}\in\mathcal{U}[\mathcal{T}]} \tilde{R}_0(\boldsymbol{u}, \boldsymbol{u}^*)$ (see Appendix A for a proof). In a discrete age-structured population, the basic reproductive number of the least-loaded class is given explicitly in terms of vital rates as

$$\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v}) = \sum_{a=1}^{T} \tilde{b}_0(a, \boldsymbol{u}, \boldsymbol{v})\tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v}), \tag{1}$$

where

$$\tilde{b}_0(a, \boldsymbol{u}, \boldsymbol{v}) = b_0(a, \boldsymbol{u}, \boldsymbol{v}) \times \exp\left(-\mu_{\mathrm{f}}(a, \boldsymbol{u}, \boldsymbol{v})\right) \tag{2}$$

and

$$\tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v}) = l_0(a, \boldsymbol{u}, \boldsymbol{v}) \times \exp\left(-\sum_{t=1}^{a} \mu_{\mathrm{s}}(t, \boldsymbol{u}, \boldsymbol{v})\right) \quad \text{with} \quad l_0(a, \boldsymbol{u}, \boldsymbol{v}) = \prod_{t=1}^{a} s_0(t, \boldsymbol{u}, \boldsymbol{v}). \tag{3}$$

In eq. (2), $b_0(a, \boldsymbol{u}, \boldsymbol{v})$ is the effective fecundity of an individual in the a-th age class with trait $\boldsymbol{u}$ and $\exp(-\mu_{\mathrm{f}}(a, \boldsymbol{u}, \boldsymbol{v}))$ is the probability that such an individual does not produce a mutated offspring with $\mu_{\mathrm{f}}(a, \boldsymbol{u}, \boldsymbol{v})$ being the mutation rate during reproduction in the a-th age class (recall that $a \in \{1, 2, .., T\}$). In eq. (3), $l_0(a, \boldsymbol{u}, \boldsymbol{v})$ is the probability that an individual with trait $\boldsymbol{u}$ survives to age a, which depends on its probability $s_0(a, \boldsymbol{u}, \boldsymbol{v}) = \exp(-d_0(a, \boldsymbol{u}, \boldsymbol{v}))$ of survival over the age interval $[a-1, a]$, where $d_0(a, \boldsymbol{u}, \boldsymbol{v})$ is its death rate. Finally, $\exp(-\sum_{t=1}^{a} \mu_{\mathrm{s}}(t, \boldsymbol{u}, \boldsymbol{v}))$ is the probability that an individual with trait $\boldsymbol{u}$ does not acquire any germline mutation until age a with $\mu_{\mathrm{f}}(a, \boldsymbol{u}, \boldsymbol{v})$ being the germline mutation rate when being in the a-th age class. In eqs. (2)–(3), we have thus distinguished between the germline mutation rate $\mu_{\mathrm{f}}(a, \boldsymbol{u}, \boldsymbol{v})$ when an offspring is produced and that rate $\mu_{\mathrm{s}}(a, \boldsymbol{u}, \boldsymbol{v})$ in the parent in the a-th age class. When $\mu_{\mathrm{f}}(a, \boldsymbol{u}, \boldsymbol{v}) = \mu_{\mathrm{s}}(a, \boldsymbol{u}, \boldsymbol{v}) = 0$ for all $a \in \{1, 2, .., T\}$, eq. (1) reduces to the standard basic reproductive number for age-structured populations (e.g. Charlesworth, 1994). We emphasise that we allowed for fecundity, survival and mutation rate to be dependent on the whole life history schedule because the evolving traits may affect physiological state variables (e.g. body size). As long as there is a direct correspondence between age and physiological state (see e.g. the discussion in de Roos, 1997), then the extension of current formalisation to physiologically-structured populations is direct (see also section 3.2 for an example). Further, since all rates in eqs. (2)–(3) depend on the resident trait $\boldsymbol{v}$, individuals can be affected by the trait of others and our thus covers frequency- and density-dependent interactions.

For a continuous set of age classes $\mathcal{T} = [0, T]$, the basic reproductive number of the least-loaded class is

$$\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v}) = \int_0^T \tilde{b}_0(a, \boldsymbol{u}, \boldsymbol{v}) \tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v}) \, \mathrm{d}a, \tag{4}$$

(see Appendix A.1 for a proof)) where $\tilde{b}_0(a, \boldsymbol{u}, \boldsymbol{v})$ takes the same functional form as in eq. (2) but is now interpreted as the effective birth rate (of offspring with no mutations) at age a, and $\tilde{l}(a, \boldsymbol{u}, \boldsymbol{v})$ satisfies the differential equation:

$$\frac{\mathrm{d}\tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v})}{\mathrm{d}a} = -\left[d_0(t, \boldsymbol{u}, \boldsymbol{v}) + \mu_{\mathrm{s}}(t, \boldsymbol{u}, \boldsymbol{v})\right] \tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v}) \quad \text{subject to} \quad \tilde{l}_0(0, \boldsymbol{u}, \boldsymbol{v}) = 1. \tag{5}$$

We now make four observations on the use of $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ to characterise long-term coevolution for life-history traits and mutation rates. (1) Because $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ depends on the amount of deleterious mutations in the population solely via $\boldsymbol{v}$, the distribution $\mathbf{p}(\boldsymbol{v})$ is needed only under frequency-dependent selection. This makes life-history evolution in the presence of deleterious mutations tractable even if the underlying evolutionary process of mutation is not (see section eq. 3.2 for an example). The characterisation of uninvadability using $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ (and thus applying eqs. 1–5) generalises the results of Leigh, 1970 and Dawson (1998, p. 148) to overlapping generations and an explicit life-history context. (2) Because $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ takes the standard form of the basic reproductive number, the results of optimal control and dynamic game theory can be applied to characterise uninvadability. This is useful in particular for reaction norm and developmental evolution and formalising different modes of trait expressions (see Avila et al., 2021).

(3) While low mutation rates relative to selection are presumed to be able to use $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ as a proxy for invasion fitness, these mutation rates are endogenously determined by the uninvadable strategy. It is thus plausible that the uninvadable mutation rate generally entails a low mutation rate. So the assumption of a low mutation rate may not appear so drastic and the extent to which this assumption is limiting depends on investigating explicit evolutionary scenarios. (4) If deleterious mutations are such that regardless in which background the life-history mutation appears, the invasion process is fully characterised by $\tilde{R}_0$, which is the case for the standard mutation accumulation models with the multiplicative effect of (deleterious) mutations (see Appendix A.1 for more details), then using $\tilde{R}_0$ does no rely on making the assumption of low (deleterious) mutation rates relative to selection.

All this gives good reasons to use $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ as a proxy of invasion fitness to understand life-history evolution in the context of mutation accumulation and as such, in the rest of this paper we consider two scenarios of the joint evolution of life history and mutation rate which we analyse by using $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$. This allows us to illustrate the different concepts, demonstrate the usefulness of focusing on $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ to get insights about how life-history evolution interacts with mutation accumulation, and check how closely our analysis matches with individual-based stochastic simulations.

3 Examples of life-history and mutation rate coevolution

3.1 Coevolution of reproductive effort and germline maintenance

3.1.1 Biological scenario

Our first scenario considers the evolution of reproductive effort when resources can be allocated to (germline) maintenance in an iteroparous population. To that end, we assume a population with a large but fixed number N of individuals undergoing the following discrete-time life-cycle (see Table 2 for a summary of key symbols for this example). (1) Each of the N adult individuals produces a large number of juveniles and either survives or dies independently of other individuals. Juveniles and surviving adults acquire mutations at the deleterious allele locus at the same rate. (2) Density-dependent competition occurs among juveniles for the vacated breeding spots (left by the dead adults) and the population is regulated back to size N.

We postulate that an individual has a life-history trait consisting of two components $\boldsymbol{u} = (u_{\mathrm{g}}, u_{\mathrm{s}})$ ($\boldsymbol{u} \in \mathcal{U}[\mathcal{T}] = [0,1]^2$), which determines how a fixed amount of resources available to each individual is allocated between three physiological functions: (i) a proportion $(1-u_{\mathrm{g}})(1-u_{\mathrm{s}})$ of resources is allocated to reproduction, (ii) a proportion $(1-u_{\mathrm{g}})u_{\mathrm{s}}$ of resources is allocated to survival, and (iii) a proportion u_{g} of resources is allocated to germline maintenance.

We assume that an individual with trait $\boldsymbol{u}$ and i deleterious mutations has the following fecundity $f_i(\boldsymbol{u})$, survival probability $s_i(\boldsymbol{u})$, and mutation rate $\mu(\boldsymbol{u})$ [at giving birth and when surviving to the next

generation, i.e., $\mu(\boldsymbol{u}) = \mu_{\mathrm{f}}(\boldsymbol{u}) = \mu_{\mathrm{s}}(\boldsymbol{u})$]:

$$\begin{aligned} f_i(\boldsymbol{u}) &= f_0(\boldsymbol{u}) \times (1-\sigma_{\mathrm{f}})^i \\ s_i(\boldsymbol{u}) &= s_0(\boldsymbol{u}) \times (1-\sigma_{\mathrm{s}})^i \\ \mu(\boldsymbol{u}) &= \mu_{\mathrm{b}}\,(1-u_{\mathrm{g}})^{\alpha_\mu}\,, \end{aligned} \tag{6}$$

where σ_{f} and σ_{s} are, respectively, the reductions in fecundity and survival from carrying an additional deleterious mutation (that are assumed to act multiplicatively), μ_{b} is the baseline mutation rate (mutation rate when allocation to germline maintenance is at its minimum, $u_{\mathrm{g}} = 0$), and α_μ is the maintenance scaling factor (a parameter tuning how investing a unit resource into maintenance translates into reducing the mutation rate). We assume that $\alpha_\mu > 1$, such that $\mu(\boldsymbol{u})$ has decreasing negative slopes in u_{g} and hence exhibits diminishing returns from investment into germline maintenance. The parameter α_μ can thus be interpreted as the "efficiency" of converting resources to lowering the mutation rates (since higher values of α_μ correspond to lower mutation rates). The quantities $f_0(\boldsymbol{u})$ and $s_0(\boldsymbol{u})$ are, respectively the fecundity and survival of the least-loaded class and they are written as

$$\begin{aligned} f_0(\boldsymbol{u}) &= f_{\mathrm{b}}\Big((1-u_{\mathrm{s}})(1-u_{\mathrm{g}})\Big)^{\alpha_{\mathrm{f}}} \\ s_0(\boldsymbol{u}) &= s_{\mathrm{b}}\Big(u_{\mathrm{s}}(1-u_{\mathrm{g}})\Big)^{\alpha_{\mathrm{s}}}. \end{aligned} \tag{7}$$

Here, f_{b} and s_{b} are, respectively, the baseline fecundity and baseline probability of survival; α_{f} and α_{s} are, respectively, the fecundity and survival scaling factors (parameters tuning how a unit resource translates into fecundity and survival). We assume that $\alpha_{\mathrm{f}}, \alpha_{\mathrm{s}} \leq 1$ whereby both survival and fecundity have decreasing positive slopes in the net amount of resources allocated to them and thus exhibit diminishing returns. Smaller values of α_{f} and α_{s} correspond to more efficient returns of investing resources into reproduction and survival, respectively.

In the absence of deleterious mutations, the model defined by eq. (6) reduces to the standard model of reproductive effort of life-history theory with trade-off between reproduction and survival (Charnov, 1993; Pen, 2000; Case, 2000). Conversely, with no overlapping generations and no life-history evolution, the model reduces to the classical model of mutation accumulation (Haigh, 1978; Bürger, 2000), and with zero survival and resource allocation evolution, it is equivalent to the asexual model of Dawson (1998). The model thus combines an unexplored trade-off between life-history traits (survival and reproduction) and immutability (germline maintenance).

3.1.2 Uninvadable and convergence stable strategies

We now carry out the invasion analysis. To that end, we use eqs. (6)–(7) to first evaluate the basic reproductive number $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ for this model (see Appendix B.1 where we also show that the equilibrium distribution $\mathbf{p}(\boldsymbol{v})$ of individuals with different numbers of mutations follows a Poisson distribution with mean $\lambda(\boldsymbol{v}) = \mu(\boldsymbol{v})/\sigma$). From eq. (B.4) we obtain that the selection gradient (e.g., Parker and May-

nard Smith, 1990; Frank, 2008; Geritz et al., 1998; Rousset, 2004) on maintenance can be written as

$$\left.\frac{\partial \tilde{R}_0(\boldsymbol{u},\boldsymbol{v})}{\partial u_\mathrm{g}}\right|_{\substack{u_\mathrm{g}=v_\mathrm{g}\\ u_\mathrm{s}=v_\mathrm{s}}} = \frac{1}{(1-\bar{s}(\boldsymbol{v}))\,(1-v_\mathrm{g})}\left(\alpha_\mu \mu(\boldsymbol{v}) - \left[\alpha_\mathrm{s}\bar{s}(\boldsymbol{v}) + \alpha_\mathrm{f}\,(1-\bar{s}(\boldsymbol{v}))\right]\right), \tag{8}$$

which displays a trade-off between allocating resources to maintenance vs into the two vital rates. The first term in eq. (8) is the marginal benefit of investment into repair, which is an increasing function of μ_b and α_μ. The second term is the marginal cost of investment into maintenance, which depends on the weighted sum over average survival and open breeding spots. Thus we find that an increase in the baseline mutation rate and more efficient returns from investment into germline maintenance and vital rates promote allocation to maintenance (increasing α_μ and decreasing α_s and α_f). Meanwhile, the selection gradient on survival can be written as

$$\left.\frac{\partial \tilde{R}_0(\boldsymbol{u},\boldsymbol{v})}{\partial u_\mathrm{s}}\right|_{\substack{u_\mathrm{g}=v_\mathrm{g}\\ u_\mathrm{s}=v_\mathrm{s}}} = \frac{1}{1-\bar{s}(\boldsymbol{v})}\left(\frac{\alpha_\mathrm{s}\bar{s}(\boldsymbol{v})}{v_\mathrm{s}} - \frac{\alpha_\mathrm{f}(1-\bar{s}(\boldsymbol{v}))}{(1-v_\mathrm{s})}\right), \tag{9}$$

where the terms in the parenthesis display the trade-off between allocating resources into survival vs fecundity (i.e. the classical reproductive effort trade-off, e.g. Pen, 2000, eq. 4) with the difference that it is here affected by the mutation rate. The first term in eq. (9) is the marginal benefit of investment into survival, which is a decreasing function of v_g. In contrast, the second term is the marginal benefit of investment into fecundity, and it is an increasing function of v_g. Thus, all else being equal, higher allocation to germline maintenance promotes fecundity over survival. We also find that an increase in the baseline mutation rate $\mu_\mathrm{b}(v_\mathrm{g})$ favours higher allocation to survival (by increasing $\bar{s}(\boldsymbol{v})$).

A necessary condition for $(u_\mathrm{g}^*, u_\mathrm{s}^*) = \boldsymbol{u}^*$ to be an evolutionary equilibrium is that the selection gradients vanish at this point (for an interior equilibrium, i.e. $0 < u_\mathrm{g}^*, u_\mathrm{s}^* < 1$, $\partial\tilde{R}_0(\boldsymbol{u},\boldsymbol{v})/\partial u_\mathrm{s} = 0$ and $\partial\tilde{R}_0(\boldsymbol{u},\boldsymbol{v})/\partial u_\mathrm{g} = 0$ at $\boldsymbol{v} = \boldsymbol{u} = \boldsymbol{u}^*$). Without further assumptions on eqs. (8)–(9), we were unable to find such analytical solutions. However, when assuming that the efficiency of investing resources into fecundity and survival is the same (i.e. setting $\alpha_\mathrm{s} = \alpha_\mathrm{f} = \alpha$), we find that the solution takes two forms, depending on the parameter values. First, when $\mu_\mathrm{b} \leq \alpha/\alpha_\mu$, then $u_\mathrm{g}^* = 0$ and $u_\mathrm{s}^* = (\exp(\mu_\mathrm{b})/s_\mathrm{b})^{1/(\alpha-1)}$. This means that when the baseline mutation rate μ_b is lower than the threshold α/α_μ given by the efficiency parameters, germline maintenance does not evolve. Hence, when investment into vital rates and maintenance is more efficient (i.e. the threshold α/α_μ is smaller), then allocation to maintenance evolves for lower values of baseline mutation rate (recall that α is the inverse efficiency of investment into vital rates, while α_μ is the efficiency of investment into maintenance). The equilibrium mutation rate, in this case, is equal to the baseline mutation rate $\mu(\boldsymbol{u}^*) = \mu_\mathrm{b}$ and and the mean number of novel mutations is $\lambda(\boldsymbol{u}^*) = \mu_\mathrm{b}/\sigma$, as expected when there is no germline maintenance. Second, when $\mu_\mathrm{b} > \alpha/\alpha_\mu$, there is a

unique interior equilibrium:

$$u_{\mathrm{g}}^* = 1 - \left(\frac{\alpha}{\alpha_\mu \mu_{\mathrm{b}}}\right)^{\frac{1}{\alpha_\mu}} \qquad u_{\mathrm{s}}^* = \left(\frac{\exp(\frac{\alpha}{\alpha_\mu})\left(\frac{\alpha}{\alpha_\mu \mu_{\mathrm{b}}}\right)^{-\frac{\alpha}{\alpha_\mu}}}{s_{\mathrm{b}}}\right)^{\frac{1}{\alpha-1}} \tag{10}$$

with corresponding expressions for the mutation rate $\mu(\boldsymbol{u}^*)$ and mean number of novel (deleterious) mutations $\lambda(\boldsymbol{u}^*)$ taking the following form

$$\mu(\boldsymbol{u}^*) = \alpha/\alpha_\mu \qquad \lambda(\boldsymbol{u}^*) = \alpha/(\alpha_\mu \sigma). \tag{11}$$

Fig. (2) illustrates these equilibrium strategies $\boldsymbol{u}^* = (u_{\mathrm{g}}^*, u_{\mathrm{s}}^*)$ (panels a and b), the corresponding mutation rate $\mu(\boldsymbol{u}^*)$ (panel c) and the mean number of novel mutations $\lambda(\boldsymbol{u}^*)$ (panel d) as a function of the baseline mutation rate μ_{b} for different values of the scaling parameter α of vital rates. Fig. (2) shows that investment into maintenance is higher when the baseline mutation rate μ_{b} increases and when investment into vital is more efficient ($\alpha \ll 1$).

Three main conclusions can be drawn from this analysis. First, selection favours physiologically costly germline maintenance at the expense of lowering investment into vital rates (survival and reproduction) when the mutation rate is high enough (whenever $\mu_{\mathrm{b}} > \alpha/\alpha_\mu$). Investment into maintenance evolves for lower values of μ_{b}, but then levels off faster when investment into maintenance is more efficient (for larger α_μ; see section 1.5.5. in S.I. for graphical illustration). Second, when germline maintenance evolves, the mutation rate, $\mu(\boldsymbol{u}^*)$, depends only on the efficiency parameters (α and α_μ) and is independent of the baseline mutation rate μ_{b} (see Fig. 2c and eq. 11). This is so in this model because the effect of μ_{b} on the cost of germline maintenance via the expected survival $\bar{s}(\boldsymbol{u}^*)$ cancels out due to the nature of density-dependence (decrease in expected survival is cancelled out by the increase in the expectation of acquiring a breeding spot; see eq. (8) when taking $\alpha_{\mathrm{s}} = \alpha_{\mathrm{f}} = \alpha$). Third, higher allocation to reproduction at the expense of survival occurs as μ_{b} increases (Fig. 2b). This is so because the effect of the mutation rate on fitness is similar to that of external mortality and thus decreases the value of allocating resources to survival. As a result, reproduction is prioritised when μ_{b} is large. This effect is more pronounced when investment into vital rates and maintenance is less efficient (α is larger or α_μ is smaller; see also section 1.5.5. in S.I.). We also find that immortality (complete survival, $\bar{s}(\boldsymbol{u}^*) = s_0(\boldsymbol{u}^*) = 1$) can evolve in our model only in the absence of external mortality ($s_{\mathrm{b}} = 1$) and zero baseline mutation rate ($\mu_{\mathrm{b}} = 0$, see eq. (10)). In section 1.1.4. of S.I., we numerically checked that our results are qualitatively robust when relaxing the assumption that the scaling factors of investment into reproduction and survival are not equal $\alpha_{\mathrm{f}} \neq \alpha_{\mathrm{s}}$.

Using the standard local analysis (Eshel, 1983; Taylor, 1989; Geritz et al., 1998; Mullon et al., 2016 and see section 1.5.3. in S.I.), we checked that the strategy $\boldsymbol{u}^*$ defined by eq. (10) is also convergence stable and locally uninvadable for biologically realistic parameter values (e.g. for the parameter values in Fig. 2). Since there is a unique solution, the local analysis implies that $\boldsymbol{u}^*$ is also the global maximum

(Sydsaeter et al., 2008) and thus the globally uninvadable strategy. The results from individual-based stochastic simulations, Fig. (3) confirms that the co-evolutionary dynamics indeed converges towards the uninvadable strategy $\boldsymbol{u}^*$ (eq. 10) predicted by the analytical model. We can observe from Fig. (2) that the analytically obtained results (eq. 10–11) correspond very closely to those obtained by carrying out individual-based stochastic simulations of the full process, which implements the life-cycle and assumptions of the present biological scenarios but allows for mutation at the life-history locus (e.g. Fig. 1 and see Appendix C.1 for the description of the simulations, Table C-4 for the standard deviations around the mean traits, and the S.I. for the Mathematica code of the simulations). We observed that simulation results generally matched well with the analytical predictions when the selection coefficient is one order of magnitude larger than the baseline mutation rate (e.g., recall the first paragraph of section 2.3.2).

3.2 Coevolution of age at maturity and germline maintenance

3.2.1 Biological scenario

Our second scenario considers the evolution of age-at-maturity when mutation accumulation can occur during growth and reproduction. To that end, we consider that age is continuous, and each individual undergoes the following events. (1) An individual is born and grows in size until it reaches maturity (growth phase). (2) At maturity, an individual starts to reproduce at a constant rate, and fecundity is assumed to be density- and size-dependent (reproductive phase). (3) Throughout their lives, individuals can die at some constant rate and acquire germline mutations. We postulate that individuals have again a life-history trait consisting of two components $\boldsymbol{u} = (u_{\mathrm{g}}, u_{\mathrm{m}})$, where u_{g} is the allocation to germline maintenance (lowering the mutation rate) and u_{m} is the age-at-maturity (see Table 3 for a summary of key symbols for this example). The life-history trait $\boldsymbol{u}$ determines how resources are allocated between three physiological functions: (i) a proportion u_{g} of resources is allocated to the maintenance of the germline at any age a, (ii) a proportion $(1 - u_{\mathrm{g}})$ of resources are allocated to growth when an individual is of age $a < u_{\mathrm{m}}$, (iii) a proportion $(1 - u_{\mathrm{g}})$ of resources is allocated to reproduction when an individual is at age $a \geq u_{\mathrm{m}}$, (hence $\boldsymbol{u} \in \mathcal{U}[\mathcal{T}] = [0, 1] \times \mathbb{R}^+$).

We assume that the effective fecundity, death, and mutation rate of an individual with trait $\boldsymbol{u}$ and i deleterious mutations in a population with resident trait $\boldsymbol{v}$ is given by

$$\begin{aligned} \tilde{b}_i(\boldsymbol{u}, \boldsymbol{v}) &= \tilde{b}_0(\boldsymbol{u}, \boldsymbol{v}) - i\sigma_{\mathrm{b}} \quad \text{if age } a \geq u_{\mathrm{m}}\text{, zero otherwise} \\ d_i &= d_{\mathrm{b}} + i\sigma_{\mathrm{d}} \\ \mu_{\mathrm{s}}(u_{\mathrm{g}}) &= \mu_{\mathrm{b}}(1 - u_{\mathrm{g}})^{\alpha_\mu}, \end{aligned} \tag{12}$$

where σ_{b} and σ_{d} are, respectively, the additive effects on fecundity and death rate from carrying deleterious mutations. The death rate of an individual of the least-loaded class is given by the baseline death

rate d_{b} and the effective birth rate of zero class individuals of the least-loaded class is given by

$$\tilde{b}_0(\boldsymbol{u},\boldsymbol{v}) = \underbrace{B(x_{\mathrm{m}}(\boldsymbol{u}))(1-u_{\mathrm{g}}^{\alpha_{\mathrm{b}}})(1-\gamma N(\boldsymbol{v}))}_{b_0(\boldsymbol{u},\boldsymbol{v})}\exp(-\mu_{\mathrm{f}}), \tag{13}$$

where $B(x_{\mathrm{m}}(\boldsymbol{u})) = ax_{\mathrm{m}}(\boldsymbol{u})^c$ is the surplus energy rate, i.e., the rate of energy available to considered life-history functions, where $x_{\mathrm{m}}(\boldsymbol{u})$ is size-at-maturity and $a > 0$ and $0 < c < 1$ are parameters. Here, $(1-u_{\mathrm{g}}^{\alpha_{\mathrm{b}}})$ represents how reproduction depends on the allocation strategy and $u_{\mathrm{g}}^{\alpha_{\mathrm{b}}}$ represents the cost to reproduction when allocating a proportion u_{g} of resources to germline maintenance. The parameter α_{b} is a scaling factor ($\alpha_{\mathrm{b}} > 1$ corresponds to diminishing returns of investing resources into reproduction), α_{b} can be interpreted as the efficiency parameter. The term $(1-\gamma N(\boldsymbol{v}))$ accounts for density-dependent regulation of reproduction, where $N(\boldsymbol{v})$ is the total population size of the resident population, which can be solved analytically (see eq. (B.8) of Appendix (B.2)), and γ tunes the intensity of density dependence. Finally, $\exp(-\mu_{\mathrm{f}})$ is the probability that the offspring do not acquire new mutations during reproduction where the mutation rate at giving birth μ_{f} is assumed constant. In order to close the expression for the birth rate, we need an explicit expression for size at maturity $x_{\mathrm{m}}(\boldsymbol{u})$. During the growth phase, we postulate that size follows the differential equation

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = \beta B(x(t))(1-u_{\mathrm{g}}^{\alpha_{\mathrm{b}}}) \qquad \text{with i.c. } x(0) = x_0, \tag{14}$$

where $B(x(t))$ is the surplus energy rate and $(1-u_{\mathrm{g}}^{\alpha_{\mathrm{b}}})$ represents the proportional allocation of resources devoted towards growth (instead of repair). For tractability, we assume that $(1-u_{\mathrm{g}}^{\alpha_{\mathrm{b}}})$ has the same functional form as the proportional allocation towards reproduction (eq. 13) and β allows to tune how much resources are needed to grow one unit, compared to the resources needed to produce one offspring. We further assume that the surplus energy rate is given by the power law $B(x(t)) = ax(t)^c$, which is considered to be appropriate for modelling size/age-at-maturity under determinate growth (see Day and Taylor (1996) for a justification). It follows from integrating eq. (14) that the size at maturity takes the form

$$x_{\mathrm{m}}(\boldsymbol{u}) = \left(\beta a(1-c)(1-u_{\mathrm{g}}^{\alpha_{\mathrm{b}}})u_{\mathrm{m}} + x_0^{1-c}\right)^{\frac{1}{1-c}}. \tag{15}$$

In the absence of mutation rate, the model reduces to the standard model of age-at-maturity (Kozlowski, 1992; Day and Taylor, 1997; Stearns, 1992; Roff, 2008). The model thus combines unexplored trade-offs between life-history traits (growth and reproduction) and immutability (germline maintenance and repair).

3.2.2 Uninvadable and convergence stable strategies

Let us now carry out the invasion analysis for which we first evaluate the basic reproductive number for this model (see Appendix B.1). Then using eq. (B.8) along with eq. (15), we find that the selection

gradient on maintenance can be written as

$$\left.\frac{\partial \tilde{R}_0(\boldsymbol{u},\boldsymbol{v})}{\partial u_{\mathrm{g}}}\right|_{\substack{u_{\mathrm{m}}=v_{\mathrm{m}}\\ u_{\mathrm{g}}=v_{\mathrm{g}}}} = \frac{\alpha_\mu \mu_{\mathrm{s}}(v_{\mathrm{g}})}{(1-v_{\mathrm{g}})}\left(v_{\mathrm{m}} + \frac{1}{\mu_{\mathrm{s}}(v_{\mathrm{g}})+d_{\mathrm{b}}}\right) - \alpha_{\mathrm{b}} v_{\mathrm{g}}^{\alpha_{\mathrm{b}}-1}\left(\beta \frac{c v_{\mathrm{m}} B(x_{\mathrm{m}}(\boldsymbol{v}))}{x_{\mathrm{m}}(\boldsymbol{v})} + \frac{1}{(1-v_{\mathrm{g}}^{\alpha_{\mathrm{b}}})}\right), \quad (16)$$

where the two terms display the trade-off between allocating resources into maintenance vs. growth and reproduction. The first term is the marginal benefit of investing into germline maintenance and it increases with the age-of-maturity and the expected lifespan. The second term is the marginal cost of investing into maintenance, which is a weighted sum of the expected loss in growth and reproduction. The marginal cost of maintenance is smaller when size-at-maturity $x_{\mathrm{m}}(\boldsymbol{v})$ is larger (since $B(x_{\mathrm{m}}(\boldsymbol{v}))/x_{\mathrm{m}}(\boldsymbol{v})$ decreases with $x_{\mathrm{m}}(\boldsymbol{v})$ for $0 < c < 1$). This implies that all else being equal, organisms that grow larger should invest more into germline maintenance. Note that the selection gradient is independent of the mutation rate μ_{f} at giving birth. We find that the selection gradient on the age-at-maturity can be written as

$$\left.\frac{\partial \tilde{R}_0(\boldsymbol{u},\boldsymbol{v})}{\partial u_{\mathrm{m}}}\right|_{\substack{u_{\mathrm{m}}=v_{\mathrm{m}}\\ u_{\mathrm{g}}=v_{\mathrm{g}}}} = c \times \frac{\beta\left(1-v_{\mathrm{g}}^{\alpha_{\mathrm{b}}}\right)B(x_{\mathrm{m}}(\boldsymbol{v}))}{x_{\mathrm{m}}(\boldsymbol{v})} - (\mu(v_{\mathrm{g}}) + d_{\mathrm{b}}). \quad (17)$$

The first term is the marginal benefit of investment into growth and thus the benefit for maturing later, while the second term is the marginal cost of investment into growth and thus the benefit for maturing earlier. We can see that the increase in mutation rate will select for earlier age-at-maturity. This implies that organisms with lower germline mutation rate can grow larger.

By first solving $\partial \tilde{R}_0(\boldsymbol{u},\boldsymbol{v})/\partial u_{\mathrm{m}} = 0$ for u_{m}^* when evaluated at $\boldsymbol{v} = \boldsymbol{u} = \boldsymbol{u}^*$, we obtain

$$u_{\mathrm{m}}^*(u_{\mathrm{g}}^*) = \frac{1}{1-c} \times \left(\frac{c}{d_{\mathrm{b}} + \mu_{\mathrm{s}}(u_{\mathrm{g}}^*)} - \frac{x_0}{\beta[1-(u_{\mathrm{g}}^*)^{\alpha_{\mathrm{b}}}]B(x_0)}\right), \quad (18)$$

which is a function u_{g}^*. Eq. (18) says that individuals tend to mature later, when individuals growth rate at birth $\dot{x}(0)$ $(= \beta[1-(u_{\mathrm{g}}^*)^{\alpha_{\mathrm{b}}}]B(x_0))$ is higher and/or when death rate d_{b}, mutation rate $\mu_{\mathrm{s}}(u_{\mathrm{g}}^*)$, and birth size x_0 are smaller (holding everything else constant). When $\mu_{\mathrm{b}} \to 0$ and $u_{\mathrm{g}}^* \to 0$, age-at-maturity reduces to $u_{\mathrm{m}}^* = (1-c)^{-1}[c/d_{\mathrm{b}} - x_0/[\beta a x_0^c]$, which is consistent with standard results about the optimal age/size at maturity (see e.g. Day and Taylor, 1996) and it is useful to compare how an allocation to germline maintenance affects the age-at-maturity. In order to determine the joint equilibrium $\boldsymbol{u}^* = (u_{\mathrm{m}}^*, u_{\mathrm{g}}^*)$, we need to substitute eq. (18) into eq. (17) and solve for u_{m}^* and u_{g}^* at $\boldsymbol{v} = \boldsymbol{u}^*$. We were unable to obtain an analytical solution for the general case. But restricting attention to $\alpha_\mu = \alpha_{\mathrm{b}} = 2$ (i.e. assuming diminishing returns of investment into germline maintenance and reproduction) and biologically feasible set of solutions (such that $0 < u_{\mathrm{g}}^* < 1$ and $u_{\mathrm{m}}^* > 1$; see section 2.1.3 of S.I. for calculations), we find that

there is a unique interior equilibrium

$$
\begin{aligned}
u_{\mathrm{g}}^{*} &= \frac{2\mu_{\mathrm{b}} + d_{\mathrm{b}} - \sqrt{d_{\mathrm{b}}(d_{\mathrm{b}} + 4\mu_{\mathrm{b}})}}{2\mu_{\mathrm{b}}} \\
u_{\mathrm{m}}^{*} &= \frac{1}{(1-c)} \times \left[c\Big(\frac{\sqrt{d} + \sqrt{d_{\mathrm{b}} + 4\mu_{\mathrm{b}}}}{2d_{\mathrm{b}}\sqrt{d_{\mathrm{b}} + 4\mu_{\mathrm{b}}}} \Big) - \frac{x_0}{\beta c B(x_0)} \Big(\frac{d_{\mathrm{b}} + 2\mu_{\mathrm{b}} + \sqrt{d_{\mathrm{b}}(d_{\mathrm{b}} + 4\mu_{\mathrm{b}})}}{2\sqrt{d_{\mathrm{b}}(d_{\mathrm{b}} + 4\mu_{\mathrm{b}})}} \Big) \right]
\end{aligned}
\tag{19}
$$

with the corresponding mutation rate given by

$$
\mu(u_{\mathrm{g}}^{*}) = \frac{(d_{\mathrm{b}} - \sqrt{d_{\mathrm{b}}}\sqrt{d_{\mathrm{b}} + 4\mu_{\mathrm{b}}})^2}{4\mu_{\mathrm{b}}}, \tag{20}
$$

while the corresponding population size $N(\boldsymbol{u}^*)$ can also be explicitly expressed in terms of parameters but remains complicated (see section 2.1.4. in S.I. for the full expression). See Fig. 4 for a graphical depiction of the equilibrium strategy $\boldsymbol{u}^*$ as a function of the baseline mutation rate (panels (a) and (b)) and the corresponding equilibrium population size and mutation rate (panels (c) and (d)). Fig. (2) shows that that investment into maintenance is higher when the baseline mutation rate is high and when external mortality is low.

Three main results can be drawn from this analysis. First, as in the previous example, selection favours physiologically costly germline maintenance at the expense of lowering the investment into life-history functions (here, into growth and reproduction, see Fig. 4a). Here, the uninvadable mutation rate ($\mu(\boldsymbol{u}^*)$) monotonically increases with the baseline mutation rate (Fig. 4d). Second, we predict earlier age-at-maturity when the baseline mutation rate is high (Fig. 4b). In fact,baseline mutation rate and external mortality have qualitatively similar effects on fitness as they increase the marginal cost of investment into growth (see the last term in eq. (17)). Thus, we find that the shift in growth-reproduction trade-off towards reproduction is higher under: (i) high external mortality rates and (ii) high baseline mutation rates. Since maturing earlier causes the growth period to be shorter, the body size at maturity $x_{\mathrm{m}}(\boldsymbol{u}^*)$ will also be smaller with a higher baseline mutation rate μ_{b} (Fig. 6a). Smaller body size at maturity, in turn, causes the birth rate $b_0(\boldsymbol{u}^*, \boldsymbol{u}^*)$ to be smaller (Fig. 6b). Third, a higher baseline mutation rate causes a smaller equilibrium population size (Fig. 4c), which is a known result in population genetics (see e.g. Gabriel et al., 1993).

The equilibrium strategy $\boldsymbol{u}^* = (u_{\mathrm{g}}^*, u_{\mathrm{m}}^*)$ determined by eq. (19) is convergence stable and locally uninvadable (see section 2.5.4. and 2.5.5.in S.I. for derivation for the parameter values in Fig. 4 and). Since eq. (19) is a unique equilibrium for the feasible trait space (see section 2.1.3 of S.I.), then local uninvadability implies global uninvadability. Using individual-based stochastic simulations, we were able to confirm that $\boldsymbol{u}^* = (u_{\mathrm{g}}^*, u_{\mathrm{m}}^*)$ given in eq. (19) is indeed a stable attractor of the evolutionary dynamics (see Fig. 5 for a graphical depiction of convergence in the individual-based simulations for four different initial population states). Fig. (4) also illustrates the equilibrium population size $N(\boldsymbol{u}^*)$ (panel c), and the uninvadable mutation rate $\mu(u_{\mathrm{g}}^*)$ (panel d) as a function of the baseline mutation rate μ_{b}. Fig. (6) illustrates the body size at maturity $x_{\mathrm{m}}(\boldsymbol{u}^*)$ (panel a) and the effective birth rate $\tilde{b}_0(\boldsymbol{u}^*, \boldsymbol{u}^*)$ at the

uninvdable population state as a function of baseline mutation rate. Overall, Fig. (4) demonstrates, again, that the analytically obtained results (here using eqs. 19–20) correspond very closely to those obtained by carrying out individual-based simulations of the full process (see Appendix C.2 for the description of the simulations, Table C-5 for the standard deviations around the mean traits, and section 2.3. in S.I. file for the Mathematica code).

4 Discussion

We formalised selection on resource allocation traits affecting life history and the deleterious mutation rate under asexual reproduction. When selection against deleterious mutations is high compared to the deleterious mutation rate, the class of individuals having the fewest such mutations (the least-loaded class) dominates in frequency the population and thus determines the fate of any mutant modifier allele affecting the deleterious mutation rate and life history. While this is a well-known result (e.g. Kimura, 1967; Kondrashov, 1995; Dawson, 1998, 1999; Johnson, 1999*b*; André and Godelle, 2006; Gervais and Roze, 2017), we here extended this result by allowing for reproduction and survival to be age-dependent under iteroparous reproduction and allowing for modifier alleles to affect jointly life-history and mutation rates via allocation to germline maintenance. Then, the basic reproductive number of the least-loaded class (eq. 1 and eq. 4) allows to characterise the joint evolutionary stable life-history and deleterious mutation rate. We analysed two specific models to illustrate this approach: (1) joint evolution between reproductive effort and the mutation rate and (2) joint evolution between the age-at-maturity and the mutation rate. These two models confirmed the validity of using the least-loaded class as a fitness proxy by comparing results to those obtained by individual-based stochastic simulations (Figs. 2–5) and provide several insights about the joint evolution of life-history and deleterious mutation rate.

The first model shows that a positive deleterious mutation rate evolves when selection against increasing the mutation rate is balanced by the cost of germline maintenance. This extends established results (Kimura, 1967; Kondrashov, 1995; Dawson, 1998, 1999) to an explicit life history context. We find that the trade-off between reproduction and survival shifts towards higher allocation into reproduction under a high baseline mutation rate. This confirms the numerical observation of Charlesworth (1990) that a higher level of a fixed mutation rate (no germline maintenance) causes higher allocation to reproduction over survival. We further predict that the shift in survival-reproduction trade-off towards reproduction is stronger: (i) when the conversion of resources into vital and germline maintenance is less efficient (e.g. in environments where organisms have high maintenance costs, e.g., colder climates), (ii) high external mortality rates (e.g., high predation environment), and (iii) high baseline mutation rates (e.g., induced by environmental stressors). We also find that immortality (complete survival) cannot evolve even in an environment with no external mortality because the mutation rate cannot be brought down to zero. This highlights the less appreciated role of mutation accumulation, which in addition to extrinsic environmentally caused hazards emphasised previously (Medawar, 1952; Hamilton, 1966; Charlesworth, 1994), can thus prevent the evolution of immortality. This means that the forces of selection on survival and

reproduction (Hamilton, 1966; Ronce and Promislow, 2010) also decline due to mutation. Overall, this example reveals that endogenous and/or exogenous factors that increase the baseline mutation rate cause lower lifespans through higher allocation to fecundity.

The model for the coevolution of age-at-maturity and the mutation rate similarly yields that a positive evolutionary stable mutation rate, but now germline maintenance trades off against investment into growth and reproduction. This extends the observation of Dańko et al. (2012) from a numerical model that looked at the effect of fixed mutation rate on the age-at-maturity. They found that a higher fixed mutation rate (no germline maintenance) causes earlier age-at-maturity, but they concluded that this effect would be relatively small and observable only under extreme conditions. By contrast, we show that mutation rate can significantly affect life-history trade-offs since allocation to germline maintenance co-evolves with life history. We predict that higher mutation rates are expected to be correlated with smaller size-at-maturity (earlier switch to reproduction) and lower equilibrium population size. Increased baseline mutation rate thus increases the effect of genetic drift and when the population size is small enough the force of genetic drift can no longer be ignored (Lynch et al., 2016), which can eventually lead to a positive feedback between drift and mutation accumulation, i.e., the mutational meltdown of the (asexual) population (Gabriel et al., 1993). Our simulations show, however, that even for a population of about 2000 individuals, genetic drift does not significantly affect the predictions of our model (see Fig. 4). Using individual-based simulations, the joint evolution between somatic maintenance, germline maintenance, size-at-maturity, and population size has been explored by Rozhok and DeGregori (2019) who found that selection for higher body size (by imposing size-dependent mortality) can lead to higher germline mutation rate because more resources need to be invested into somatic maintenance. Thus, they found that a higher germline mutation rate and size-at-maturity are expected to be positively correlated, which is an opposite prediction from our result. Further studies clarifying the selection pressures involved in the trade-off between growth and investment into the maintenance of germline and soma, can shed light on how patterns of body size, longevity and mutation rate are expected to be correlated.

Two findings about how life history evolves jointly with deleterious mutation rate emerge from our models. First, the trade-off between lowering the rate of mutations vs investing in life-history functions affects the evolutionary outcome of life-history trade-offs (e.g. survival–vs–reproduction or growth–vs–reproduction). Hence, mutation accumulation can have a significant effect on life-history evolution through the process of joint evolution that previous models focusing on the effect of fixed mutation rates on life-history evolution have not revealed (Charlesworth, 1990; Dańko et al., 2012). Looking at the effect of fixed mutation rates on life history evolution underestimates the effect of deleterious mutation accumulation on life history evolution, as it does not consider the physiological cost of immutability on life history evolution. Second, factors that contribute to higher baseline mutation rate select for faster life histories: higher investment into current reproduction at the expense of survival and earlier age–at–maturity. Factors that could increase the baseline mutation rate include factors that increase DNA replication errors (number of germ-line cell divisions) or environmental mutagens (oxygen level, nutrition

quality, see e.g. Ferenci, 2019 for a review).

Our analysis is based on a number of simplifying assumptions, most notably (i) a separation between life-history traits (e.g., the timing of reproduction, age and size at maturity, longevity) and traits that could be referred to as viability traits, such as morphology and physiology, (ii) that mutations are only deleterious at the viability traits, and (iii) that reproduction is asexual. The separation between life-history traits and viability traits allows us to focus on the life-history resource allocation trade-off and avoids modelling the viability traits mechanistically. This separation makes biological sense for organisms that have similar viability traits but differ in their life history. For instance, both annual and perennial plants can have similar cellular pathways for photosynthesis and oxidative phosphorylation, where deleterious mutations can affect the functionality of these pathways that are under strong selection. For such organisms, our model thus consists of characterising the optimal life history and how the mutation rate at the cellular machinery evolves. While other formulations are possible and could be explored in future work, our approach has analytical traction and is conceptually equivalent to the separation between modifier locus and loci affecting vital rates from modifier theory, where modifier alleles affect the pattern of transmission of other traits (e.g., Leigh, 1970; Altenberg, 2009). In our model, modifier alleles are also under direct selection owing to their effect on life history. From modifier theory, we know that regardless of the details of transmission; namely, regardless of whether reproduction is asexual or sexual or the exact pattern of mutation rates (whether they are reducible or not), a decrease in the mutation rate is always favoured in the absence of trade-off with reproduction (Altenberg, 2009; Altenberg et al., 2017*b*). Thus, in the presence of a trade-off in any genetic system, the evolutionary stable mutation rate will be determined by the balance between the benefit of lowering the mutation rate and the benefit of increasing the mutation rate and thus reducing the physiological cost of germline maintenance (e.g. André and Godelle, 2006; Kondrashov, 1995; Dawson, 1998). Under sexual reproduction however, fewer resources will be allocated to germline maintenance (Kondrashov, 1995; Dawson, 1998; André and Godelle, 2006; Altenberg, 2009) because deleterious mutations are linked under asexual reproduction, and this linkage is broken down by recombination under sexual reproduction whereby the benefit of lowering the mutation is smaller (Dawson, 1998; Gervais and Roze, 2017). However, any direct effect that the modifier has would be experienced in the same way regardless of the reproductive system. In summary, while genetic details might be important to compare the quantitative effect of the reproductive system on the joint evolution of mutation rate and life history, they are unlikely to qualitatively affect our main prediction that the cost of germline maintenance is a significant force affecting life-history traits.

Our model can be further extended to study several open questions in life-history theory and mutation accumulation theory. First, our model can be applied to understand the role of germline maintenance in explaining the evolution of ageing. Current theories of ageing either rely on mutation accumulation in the germline (Medawar, 1952) or are based on a trade-off between reproduction and survival (disposable soma and antagonistic pleiotropy theories of ageing; Williams, 1957; Kirkwood, 1977; Cichon and Kozlowski, 2000), but no theoretical study has explored how these processes evolve jointly when germline maintenance

is costly. Second, our model can also be extended to study the Lansing effect, a widely reported negative effect of parental age on offspring fitness (e.g. Monaghan et al., 2020). Third, extending our model for sexual reproduction would allow studying how sex-specific differences in germline maintenance can induce sex differences in life-history trade-offs (Maklakov and Lummaa, 2013). Our hope is that the formalisation proposed in this paper can be fruitfully used to these ends.

A Invasion process with mutation accumulation

In this appendix, we prove that the basic reproductive number $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ defined by eq. (1) (in discrete-time) and eq. (4) (in continuous-time) is an appropriate invasion fitness proxy when the least loaded class dominates in frequency the resident population. To that end, we first characterise the mutant invasion process by following the assumptions of the main text section 2.1 and then derive the renewal equations in dicrete-and continuous-time (section A.2), which give rise to using $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ as a fitness proxy.

A.1 Reducible invasion process

Since the resident population is assumed to have reached a mutation-selection balance and is thus at a genetic-demographic equilibrium, i.e. the standard internal stability assumptions of invasion analysis (Eshel and Feldman, 1984; Altenberg et al., 2017*b*; Metz, 2011), the mutant allele for trait $\boldsymbol{u}$ can arise in individuals carrying different numbers of deleterious mutations. Hence, the invasion process of $\boldsymbol{u}$ is contingent on the genetic background in which it arises. We refer to the initial carrier of the $\boldsymbol{u}$ trait as the progenitor (or ancestor) of $\boldsymbol{u}$. The invasion fitness of mutant $\boldsymbol{u}$ is then determined by the size of the lineage of the progenitor, which consists of all of its descendants carrying $\boldsymbol{u}$ far into the future. Namely, the immediate descendants of the progenitor, including the surviving self, the immediate descendants of the immediate descendants, etc., covering the whole family history tree *ad infinitum*. Crucially, descendants may accumulate deleterious mutations during the initial growth or extinction of the mutant lineage when this lineage is rare (referred throughout as the "invasion process"). As such, the invasion process, whether in continuous or discrete time, can be regarded as a multitype age-dependent branching process (Mode, 1968, 1971) since during the growth or extinction of the mutant lineage, novel genotypes are produced by mutation.

To analyse this invasion process, it is useful to organise individuals into *equivalence classes*. The defining feature of an equivalence class is that it is a collection of states of a process among which transitions eventually occur, so the states are said to communicate (Karlin and Taylor, 1975, p. 60). The equivalence class $\mathcal{C}_i$ will stand for all mutant individuals carrying i deleterious mutations and thus consist of individuals of all ages. This is an equivalence class because, through survival and reproduction, an individual of any age with i mutations may eventually transition to become an individual of any other age (in the absence of menopause). This follows from the fact that the process of survival and reproduction in an age-structured population in the absence of mutations (and menopause) is *irreducible* (Karlin and Taylor, 1975, p. 60), since starting in any age-class, eventually, every age class can be reached by the members of a lineage of individuals. Owing to our assumptions that mutations are deleterious and can only accumulate, however, starting in a given mutational class, it is possible to enter another class, but not transition back from that class (otherwise the two classes would form a single class) whereby the process of mutation is *reducible*, since a lineage of individuals cannot go from any mutational class to any other mutation class. Equivalence class $\mathcal{C}_{i+1}$ is then said to follow class $\mathcal{C}_i$ since individuals in equivalence

class i can only transition to class $i+1$ by acquiring mutations. This means that the invasion process is reducible and can be regarded as a reducible multitype age-dependent branching process (Nair and Mode, 1971; Mode, 1971). Reducibility typically arises in population genetics models without back-mutations but sometimes also occurs in other models in ecology and evolution, e.g. when some class of individuals do not contribute to reproduction, as is, for instance, the case under menopause (e.g. Caswell, 2000; Altenberg, 2009; Bode et al., 2006; Stott et al., 2010; McDonald, 2015).

To see why the notion of an equivalence class is useful to understand the mutant invasion process, let us first focus on a discrete-time process with T discrete age classes and denote by $n_i(t)$ the expected number of individuals at time $t = 0, 1, 2, ...$, in class $\mathcal{C}_i$ that descend from a single class $\mathcal{C}_i$ newborn progenitor born at $t = 0$ (i.e. $n_i(0) = 1$). Thus, $n_i(t)$ is the expected lineage size of class $\mathcal{C}_i$ individuals descending from a newborn progenitor of class $\mathcal{C}_i$ (including the surviving self). Accounting entails that $n_i(t)$ satisfies the renewal equation

$$n_i(t) = \tilde{l}_i(t) + \sum_{a=1}^{t} n_i(t-a)\tilde{b}_i(a)\tilde{l}_i(a), \tag{A.1}$$

where $\tilde{l}_i(a)$ is the probability that a $\mathcal{C}_i$ class newborn survives to the a-th age class and has not acquired any new mutation, and $\tilde{b}_i(a)$ is the expected number of (newborn) offspring without mutations produced by an individual belonging to the a-th age class and being of mutation class $\mathcal{C}_i$. Eq. (A.1) is derived in Appendix A.2.1 from branching process considerations and its left-hand side can be understood as follows. First, $\tilde{l}_i(t)$ accounts for the survival and immutability of the progenitor itself until age t. Second, $\tilde{b}_i(a)\tilde{l}_i(a)$ is the progenitor's expected number of offspring of class $\mathcal{C}_i$ born during the time interval corresponding to the a-th age class and each of these newborns contribute an expected number $n_i(t-a)$ of class $\mathcal{C}_i$ individuals to the progenitor's total lineage size at t. This is so because as long as the mutant is rare, each newborn starts a new independent lineage and adding all terms together, the right-hand side of eq. (A.1) thus gives the total lineage size of the progenitor (see Appendix A.2.1 for more details).

A key feature of eq. (A.1) is that it depends only on the vital rates and states of individuals of class $\mathcal{C}_i$. As such, eq. (A.1) is functionally equivalent to the standard renewal equation of population dynamics in discrete age-structured populations (Charlesworth, 1994, eq. 1.34), but recall that $n_i(t)$ counts total lineage size in class $\mathcal{C}_i$ and not zygote size as in the classical theory of age-structured populations. It then follows from the classical theory of age-structured populations (e.g., Charlesworth, 1994, p. 25-26) or the branching process formulation (Mode, 1974) that asymptotically, as $t \to \infty$, the number $n_i(t)$ grows geometrically as

$$n_i(t) \sim \rho_i^t K_i, \tag{A.2}$$

where K_i is a constant depending on the process and ρ_i is the unique root satisfying the characteristic (or Euler-Lotka) equation $\sum_{a=1}^{T} \rho_i^{-a}\tilde{b}_i(a)\tilde{l}_i(a) = 1$.

Since individuals of class i contribute to individuals of class $i+1$ through mutations, the equivalence

class $\mathcal{C}_{i+1}$ follows class $\mathcal{C}_i$, then $n_i(t)$ does not describe the total expected lineage size of the progenitor. However, owing to the assumptions that mutations are deleterious and can only accumulate, the growth ratio ρ_i is at least as large as ρ_{i+1}, i.e., $\rho_i \geq \rho_{i+1}$ for all i. This implies that when the ancestor is of type i, the expected lineage size is determined by the growth ratio ρ_i, since it dominates that of any other following equivalence class. Hence, asymptotically, the total expected lineage of an $\mathcal{C}_i$ class progenitor has a geometric growth ratio ρ_i. It further follows from the theory of multitype age-dependent branching processes that the realised lineage size of a single progenitor (a random variable) has growth ratio ρ_i if $\rho_i > 1$ and otherwise if $\rho_i \leq 1$, the lineage goes extinct with probability one (Mode, 1971, Theorem 7.2 p. 245, Corrolary 6.1 p. 280, see also Mode, 1974 for the single type case). Further, $\rho_i \leq 1$ if and only if $\tilde{R}_i \leq 1$, where $\tilde{R}_i = \sum_{a=1}^{T} \tilde{b}_i(a)\tilde{l}_i(a)$ is the expected number of offspring of a progenitor of type i produced throughout its lifespan (i.e. Mode, 1971, Theorem 7.2 p. 245, Corrolary 6.1 p. 280, see also Karlin and Taylor, 1981, p. 424, Caswell, 2000). Hence, ρ_i is an appropriate measure of invasion fitness and $\tilde{R}_i$ is an appropriate proxy of it for a type i mutant $\boldsymbol{u}$ arising in a resident $\boldsymbol{v}$ background. The same argument can be made for continuous-time processes, in which case, the invasion fitness of a mutant arising in a progenitor in class $\mathcal{C}_i$ is $\rho_i = \exp(r_i)$, where r_i is the rate of natural increase of the lineage size of a progenitor of type i, i.e., the Malthusian growth rate and for which $\tilde{R}_i = \int_0^T \tilde{b}_i(a)\tilde{l}_i(a)\,\mathrm{d}a$ is an appropriate invasion fitness proxy (see Appendix A.2.2 for a derivation).

The key feature of the invasion process in a population with distinct mutational equivalence classes is that the invasion fitness of a mutant life history trait thus depends on the class in which it appears (ρ_i for class i), which in turn depends on the distribution $\mathbf{p}(\boldsymbol{v})$. This means that there are as many growth rates as equivalence classes, since the invasion process is reducible (see also Altenberg, 2009, p. 1278). Therefore characterising long-term evolution using a single representation of invasion fitness (or proxy thereof) is at first glance unattainable as it requires to make the process irreducible, which can be achieved by introducing back mutations or recombination. Reaching irreducibility in this way, however, would make the model much more complicated. It also follows from these considerations, that when least-loaded class dominates in frequency the population, the face of mutant $\boldsymbol{u}$ appearing in a resident $\boldsymbol{v}$ population is determined from the growth ratio $\rho_0(\boldsymbol{u}, \boldsymbol{v})$ of the least-loaded class only, since any mutant will appear on the $\mathcal{C}_0$ background and so this is an appropriate overall measure of invasion fitness ($\rho(\boldsymbol{u}, \boldsymbol{v}) = \rho_0(\boldsymbol{u}, \boldsymbol{v})$). Further, since $\rho_0(\boldsymbol{u}, \boldsymbol{v}) \leq 1 \iff \tilde{R}_0(\boldsymbol{u}, \boldsymbol{v}) \leq 1$, where $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ is the basic reproductive number of the least-loaded class, i.e. the expected number of class $\mathcal{C}_0$ offspring produced by a class individual $\mathcal{C}_0$ individual over its lifespan, is sufficient to characterise the fate of the mutant. This then justifies using $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$ as an invasion fitness proxy defined by eq. (1) and eq. (4) for discrete and continuous time respectively. Further, if deleterious mutations are such that all the ρ_i's are proportional to ρ_0's, which is the case for the standard mutation accumulation models with multiplicative effect of (deleterious) mutations, then using $\tilde{R}_0$ does no rely on making the assumption of low mutation rates relative to selection, since regardless in which background the mutation appears, it will grow proportionally to ρ_0.

A.2 Renewal equations from age-dependent branching process

We here derive the renewal eq. (A.1) for a discrete time process and its continuous time analogue from underlying age-dependent branching process assumptions (Crump and Mode, 1968; Mode, 1968, 1971). As such, we consider a full stochastic model of survival and reproduction to describe the fate, invasion or extinction, of a mutation introduced as a single copy in a resident population. The defining assumption of a branching process is that a new and independent copy of the process begins each time a new individual is born (Harris, 1963; Mode, 1968, 1971).

To describe the branching process under our general assumptions of the main text section 2.1, let us introduce the following notations, which will apply to both the discrete and continuous time processes. First, let $Z_i(t)$ stand for the random number of individuals in class $\mathcal{C}_i$ at time t that descend from a single newborn (age class zero) mutant progenitor in class $\mathcal{C}_i$ at $t = 0$, whereby $Z_i(0) = 1$. Second, denote by L_i the random sojourn time of a progenitor of class $\mathcal{C}_i$ in that class; the progenitor exits class $\mathcal{C}_i$ either because of death (in which case L_i is the lifespan) and/or because a germline mutation occurred. Third, let

$$\theta(t - L_i) = \begin{cases} 1 & \text{if } t - L_i \geq 0 \\ 0 & \text{if } t - L_i < 0 \ , \end{cases} \tag{A.3}$$

by the indicator random variable describing whether a progenitor of class $\mathcal{C}_i$ has exited class $\mathcal{C}_i$ by time t (either by death and/or mutation). Namely, if $\theta(t - L_i) = 0$, then the progenitor is still in class $\mathcal{C}_i$ at t (alive without having mutated). Finally, let $n_i(t) = \mathrm{E}[Z_i(t)]$ denote the expected number of individuals at time t in class $\mathcal{C}_i$ that descend from a single class $\mathcal{C}_i$ newborn progenitor at $t = 0$, where the expectation is over all stochastic events affecting survival and reproduction (we refer to Crump and Mode, 1968; Mode, 1971 for a construction of the probability spaces for branching processes, here we simply assume they exist).

A.2.1 Discrete time

Let us now focus on the discrete-time process where $t = 0, 1, 2, ...$ and further introduce the random number $W_i(a)$ of offspring without new mutations born to the progenitor of class $\mathcal{C}_i$ during the time interval corresponding to the a-th age group where $a = 1, 2, ...$ so that, recall, we begin counting discrete age classes with 1 (one of the two possible conventions to count discrete age classes, Fig. 3.1 Case, 2000). Because a branching process starts afresh with the birth of any offspring, the random variable $Z_i(t)$ satisfies the discrete time stochastic renewal equation

$$Z_i(t) = 1 - \theta(t - L_i) + \sum_{a=1}^{t} Z_i(t - a) W_i(a), \tag{A.4}$$

which can be understood as follows. First, if $\theta(t - L_i) = 0$, then the progenitor is still in class $\mathcal{C}_i$ at t and thus contributes one individual to its lineage. Second, each offspring produced without mutation during

the time interval corresponding to the a-th age class starts a new independent branching process and thus contributes $Z_i(t-a)$ individuals of class $\mathcal{C}_i$ to the progenitor's lineage at t. Summing over all offspring and all age groups the progenitors can be in, we obtain the second term in eq. (A.4). In other words, the sum $\sum_{a=1}^{t} Z_i(t-a)W_i(a)$ is the total random number of class $\mathcal{C}_i$ individuals at t that descendent from each of the offspring of the progenitor including the surviving selves (i.e., offspring, offspring from offspring, offspring from offspring from offspring, etc). Thus, the right-hand side of eq. (A.4) is the progenitor's random lineage size at t. Bearing differences of notations and context, eq. (A.4) is conceptually equivalent to eq. (5.6) of Mode (1974).

Taking the expectation over realizations of the stochastic process over both sides of eq. (A.4) yields

$$\begin{aligned} \mathrm{E}[Z_i(t)] =& 1 - \mathrm{E}[\theta_i(t-L)] + \sum_{a=1}^{t} \mathrm{E}[Z_i(t-a)W_i(a)] \\ =& 1 - \mathrm{E}[\theta_i(t-L)] + \sum_{a=1}^{t} \mathrm{E}[Z_i(t-a)]\mathrm{E}[W_i(a)] \end{aligned} \tag{A.5}$$

where the first equality follows from the linearity of the expectation operator and the second equality follows from the assumption of the independence of each new branching process. In order to further simplify this expression, we first note that the probability that the progenitor is still of class $\mathcal{C}_i$ at t is $\tilde{l}_i(t) = 1 - \mathrm{P}(L_i \leq t)$, and that by standard probability arguments, $\mathrm{P}(L_i \leq t)$ is the expectation of the indicator random variable of the event that the progenitor exits class $\mathcal{C}_i$ at t: $\mathrm{E}[\theta(t-L_i)] = \mathrm{P}(L_i \leq t)$ (Grimmett and Stirzaker, 2001). Second, we assume stochastic independence between survival and reproduction, whereby the expected number of offspring without mutations born to a progenitor of class $\mathcal{C}_i$ during the time interval corresponding to the a-th age class is $\mathrm{E}[W_i(a)] = \tilde{b}_i(a)\tilde{l}_i(a)$, which is the probability $\tilde{l}_i(a)$ that the progenitor has survived until age class a without mutating times the expected number $\tilde{b}_i(a)$ of its offspring without mutations produced when residing in the a-th age class. Substituting these quantities into eq. (A.5) and noting that, by definition, $n_i(t-a) = \mathrm{E}[Z_i(t-a)]$, eq. (A.5) can be written $n_i(t) = \tilde{l}_i(t) + \sum_{a=1}^{t} n_i(t-a)\tilde{b}_i(a)\tilde{l}_i(a)$, which is eq. (A.1). Bearing differences of notations and context, eq. (A.5) is conceptually equivalent to eq. (5.11) of Mode (1974).

A.2.2 Continuous time

Let us now focus on the continuous time process where $t \in [0, \infty)$ and for this case, let $N_i(t)$ denote the random number of offspring without mutations produced until time t by a progenitor of class $\mathcal{C}_i$, where these offspring are produced at the random times $T_{1(i)} \leq T_{2(i)} \leq ... \leq T_{N_i(t)}$ with $T_{j(i)}$ being the random time until production of the j-th offspring. Thus, $N_i(t)$ is a renewal counting (or point) process (Karlin and Taylor, 1975, p. 31) with rate function $\tilde{b}_i(t)\tilde{l}_i(t)$, where $\tilde{l}_i(t)$ is the probability that a progenitor of class $\mathcal{C}_i$ is still in that class at time t (same as in the discrete-time case), and $\tilde{b}_i(t)$ is the birth rate of offspring without mutations at t. Because a branching process starts again with the birth of each offspring, the random variable $Z_i(t)$ satisfies under the continuous time process the stochastic renewal

equation

$$Z_i(t) = 1 - \theta(t - L_i) + \sum_{j=1}^{N_i(t)} Z_i(t - T_{j(i)}), \tag{A.6}$$

where, as in the discrete case, $1 - \theta(t - L_i)$ counts the surviving progenitor without mutation and the sum counts the total number of class $\mathcal{C}_i$ individuals at t that descend from each of the offspring of the progenitor, where the progenitor's offspring are produced at random times. Bearing differences of notations and context, eq. (A.6) is conceptually equivalent to eq. (3.1) of Crump and Mode (1968).

Taking the expectation over realizations of the stochastic process over both sides of eq. (A.6) yields

$$\begin{aligned} \mathrm{E}[Z_i(t)] =& 1 - \mathrm{E}[\theta_i(t - L)] + \mathrm{E}\Big[\sum_{j=1}^{N_i(t)} Z_i(t - T_{j(i)})\Big] \\ =& 1 - \mathrm{E}[\theta_i(t - L)] + \mathrm{E}\Big[\sum_{j=1}^{N_i(t)} \mathrm{E}[Z_i(t - T_{j(i)})]\Big] \\ =& \tilde{l}_i(t) + \mathrm{E}\Big[\sum_{j=1}^{N_i(t)} n_i\left(t - T_{j(i)}\right)\Big], \end{aligned} \tag{A.7}$$

where the second line follows from the independence of each new branching process and the third from the definitions introduced above. Now, owing to Campbell's theorem relating the rate of a renewal counting process, here $\tilde{b}_i(a)\tilde{l}_i(a)$, and the expectation of a random sum of a function over the process (e.g., Kingman, 1992, p. 28), we have $\mathrm{E}\Big[\sum_{j=1}^{N_i(t)} n_i\left(t - T_{j(i)}\right)\Big] = \int_0^t n_i(t-a)\tilde{b}_i(a)\tilde{l}_i(a)\,\mathrm{d}a$ and therefore the eq. (A.7) becomes

$$n_i(t) = \tilde{l}_i(t) + \int_0^t n_i(t-a)\tilde{b}_i(a)\tilde{l}_i(a)\,\mathrm{d}a. \tag{A.8}$$

Bearing differences of notations and context, eq. (A.8) is conceptually equivalent to eq. (6.1) of Crump and Mode (1968) once the assumptions of their example 8.2 is endorsed, i.e., stochastic independence between survival and reproduction.

Eq. (A.8) is also functionally equivalent to the standard renewal equation of population dynamics for continous age-structured populations (Charlesworth, 1994, eq. 1.41). As such, and as for the discrete-time case, it then follows from the standard results of population dynamic processes in age-structured populations (Charlesworth, 1994, p. 27) or from the continuous time branching process formulation (Crump and Mode, 1968) that asymptotically, as $t \to \infty$, the number $n_i(t)$ grows geometrically as

$$n_i(t) \sim \rho_i^t K_i, \tag{A.9}$$

where K_i is some constant depending on the process and $\rho_i = \exp(r_i)$, where r_i is the mutant growth

rate (or Malthusian parameter), which is the unique root of the Euler-Lotka equation

$$\int_0^\infty \exp(-ar_i)\tilde{b}_i(a)\tilde{l}_i(a)\,\mathrm{d}a = 1. \tag{A.10}$$

Thereby $r_i \leq 0 \iff \tilde{R}_i \leq 1$ with $\tilde{R}_i = \int_0^\infty \tilde{b}_i(a)\tilde{l}_i(a)$ being the basic reproductive number of a class i individual (Karlin and Taylor, 1981, p. 424, Mode, 1971, Theorem 7.2 p. 245, Corrolary 6.1 p. 280).

B Basic reproductive numbers

In this appendix, we present the explicit expressions for $\tilde{R}_0(\boldsymbol{u},\boldsymbol{v})$ for our two biological scenarios

B.1 Coevolution of reproductive effort and germline maintenance

From the model assumptions given in main text section 3.1.1, we have that a juvenile surviving density-dependent competition, has a survival probability of one to the adult stage and that each adult of the least-loaded class survives to the next generation with probability $s_0(\boldsymbol{u})$. In the meantime, the probability that each type of individual has not acquired a germline mutation is $\exp(-\mu(\boldsymbol{u}))$. Therefore, for this model eq. (3) becomes

$$\tilde{l}_0(a,\boldsymbol{u},\boldsymbol{v}) = \begin{cases} \exp(-\mu(\boldsymbol{u})) & \text{if } a = 1 \\ s_0(\boldsymbol{u})^{a-1}\exp(-\mu(\boldsymbol{u})a) & \text{if } a > 1\ , \end{cases} \tag{B.1}$$

while the effective fecundity of the least-loaded class (eq. 2) is

$$\tilde{b}_0(a,\boldsymbol{u},\boldsymbol{v}) = \tilde{b}_0(\boldsymbol{u},\boldsymbol{v}) = \underbrace{(1-\bar{s}(\boldsymbol{v}))\,\frac{f_0(\boldsymbol{u})}{\bar{f}(\boldsymbol{v})}}_{b_0(\boldsymbol{u},\boldsymbol{v})} \tag{B.2}$$

since no mutation occurs specifically during reproduction (in addition of the germline mutation mentioned above). The effective fecundity depends on the mean survival and fecundity in the population, respectively, $\bar{s}(\boldsymbol{v}) = \sum_{k=0}^\infty s_k(\boldsymbol{v})p_k(\boldsymbol{v})$ and $\bar{f}(\boldsymbol{v}) = \sum_{k=0}^\infty f_k(\boldsymbol{v})p_k(\boldsymbol{v})$. Here, $p_i(\boldsymbol{v})$ is the probability that an individual randomly sampled from the resident population carries i deleterious mutations (and so $\mathbf{p}(\boldsymbol{v}) = \{p_i(\boldsymbol{v})\}_{i\in\mathbb{N}}$ for this model). This can be understood by noting that $(1-\bar{s}(\boldsymbol{v}))$ is the fraction of open breeding spots available to a juvenile and the probability that the offspring of a given adult acquires a breeding spot depends on the fecundity of the adult relative to the population average fecundity (as each juvenile is equally likely to acquire a breeding spot).

Since there is no fixed end to lifespan under the above life-cycle assumptions (so $T \to \infty$), the basic reproductive number of the least loaded class, eq. (1), is

$$\tilde{R}_0(\boldsymbol{u},\boldsymbol{v}) = \sum_{a=1}^\infty b_0(\boldsymbol{u},\boldsymbol{v})s_0(\boldsymbol{u})^{a-1}\exp(-\mu(\boldsymbol{u})a)\,, \tag{B.3}$$

which yields

$$\tilde{R}_0(\boldsymbol{u},\boldsymbol{v}) = \frac{b_0(\boldsymbol{u},\boldsymbol{v})}{\exp(\mu(\boldsymbol{u})) - s_0(\boldsymbol{u})}, \tag{B.4}$$

(all our mathematical computations can be followed and confirmed via an accompanying Supplementary Information, S.I. consisting of a Mathematica notebook). Since $b_0(\boldsymbol{u},\boldsymbol{v})$ is multiplicatively separable with respect to its arguments, then it follows from eq. (B.4) that the model satisfies the condition of an optimisation principle (e.g., Metz et al., 2008). Namely, $\tilde{R}_0(\boldsymbol{u},\boldsymbol{v}) = F_1(\boldsymbol{u})F_2(\boldsymbol{v})$ for the functions $F_1(\boldsymbol{u}) = f_0(\boldsymbol{u})/[\exp(\mu(\boldsymbol{u})) - s_0(\boldsymbol{u})]$ depending only on the mutant and $F_2(\boldsymbol{v}) = [1-\bar{s}(\boldsymbol{v})]/\bar{f}(\boldsymbol{v})$ depending only on the resident. It follows that maximising $F_1(\boldsymbol{u})$ is sufficient to ascertain uninvadability and uninvadability implies convergence stability when the evolutionary dynamics follow an optimization principle (Metz et al., 2008). Further, the explicit expressions for $\bar{s}(\boldsymbol{v})$ and $\bar{f}(\boldsymbol{v})$, and thus the distribution $\mathbf{p}(\boldsymbol{v})$ are not needed to carry out the invasion analysis. This allows to markedly simplify the evolutionary analysis, but we will nevertheless work out the resident distribution $\mathbf{p}(\boldsymbol{v})$ so as to have a fully worked example that allows for consistency checks and illustrating the concepts.

Since we consider a deterministic resident population process, the frequency p_k satisfies at equilibrium the equation

$$p_k(\boldsymbol{v}) = \sum_{i=0}^{k} \phi_{k-i}(\boldsymbol{v}) w_i(\boldsymbol{v}) p_i(\boldsymbol{v}), \tag{B.5}$$

where $w_i(\boldsymbol{v}) = s_i(\boldsymbol{v}) + (1-\bar{s}(\boldsymbol{v}))\, f_i(\boldsymbol{v})/\bar{f}(\boldsymbol{v})$ is the individual fitness– survival plus effective fecundity–of an individual with i deleterious mutations, and ϕ_k is the probability that k deleterious mutations are produced upon reproduction. Assuming that the mutation distribution is Poisson with mean $\mu(\boldsymbol{v})$ and $\sigma_{\mathrm{s}} = \sigma_{\mathrm{f}} = \sigma$, then eq. (B.5) becomes structurally equivalent to eq. (1) of Haigh (1978) and eq. (5.3) of Bürger (2000, p. 300) (with mean fitness $\bar{w} = 1$ since population size is constant) and as such the equilibrium distribution $\mathbf{p}(\boldsymbol{v})$ is Poisson with mean $\lambda(\boldsymbol{v}) = \mu(\boldsymbol{v})/\sigma$ (see also the section 1.1.1. in SM). This completely characterises the genetic state of the resident population and implies that

$$\bar{s}(\boldsymbol{v}) = s_0(\boldsymbol{v})e^{-\mu(\boldsymbol{v})} \quad \text{and} \quad \bar{f}(\boldsymbol{v}) = f_0(\boldsymbol{v})e^{-\mu(\boldsymbol{v})}. \tag{B.6}$$

Substituting the explicit expression for the survival and effective fecundities (eq. B.6) into eq. (B.2) and eq. (B.4) shows that in a monomorphic $\boldsymbol{v}$ population $\tilde{R}_0(\boldsymbol{v},\boldsymbol{v}) = 1$, as required for a consistent model formulation. Eq. (B.6) generalises the standard mutation-accumulation model of population genetics to overlapping generations with survival probability depending on the number of deleterious mutations (see e.g. eq. 3.3 Kimura and Maruyama, 1966).

B.2 Coevolution of age at maturity and germline maintenance

From the model assumptions given in the main text section 3.2.1 and using eq. (4) (under $T \to \infty$ since there is also no definite end to lifepan), the basic reproductive number of the least-loaded class reduces to

$$\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v}) = \tilde{b}_0(\boldsymbol{u}, \boldsymbol{v}) \int_{u_\mathrm{m}}^{\infty} \tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v})\, \mathrm{d}a, \tag{B.7}$$

where $\tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v}) = \exp\Big(-(\mu(u_\mathrm{g}) + d_\mathrm{b})a\Big)$. Substituting the expression for eq. (13) into eq. (B.7) and integrating yields

$$\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v}) = \underbrace{B(x_\mathrm{m}(\boldsymbol{u}))\left(1 - u_\mathrm{g}^{\alpha_\mathrm{b}}\right)\exp(-\mu_\mathrm{f}) \times \frac{\exp\Big(-(\mu(u_\mathrm{g}) + d_\mathrm{b})u_\mathrm{m}\Big)}{\mu(u_\mathrm{g}) + d_\mathrm{b}}}_{F_1(\boldsymbol{u})} \times \underbrace{(1 - \gamma N(\boldsymbol{v}))}_{F_2(\boldsymbol{v})}. \tag{B.8}$$

This shows that one can again express the basic reproductive number as a product of the form $\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v}) = F_1(\boldsymbol{u})F_2(\boldsymbol{v})$ and thus the optimisation principle (e.g. Metz et al., 2008) applies also in this model. This means that evaluating $N(\boldsymbol{v})$ explicitly is not needed to ascertain uninvadability (and uninvadability will again imply convergence stability for this model). We will nevertheless work it out and in order to derive an explicit expression for $N(\boldsymbol{v})$ it suffices to note that in a monomorphic resident population at a joint demographic and genetic equilibrium, each individual belonging to the least-loaded class must leave on average one descendant with zero new mutations. Hence $\tilde{R}_0(\boldsymbol{v}, \boldsymbol{v}) = 1$ implies that that the population size at the demographic steady state is

$$N(\boldsymbol{v}) = \frac{B(x_\mathrm{m}(\boldsymbol{v}))(1 - v_\mathrm{g}^{\alpha_\mathrm{b}}) - (d_\mathrm{b} + \mu(v_\mathrm{g}))\exp\Big(\mu_\mathrm{f} + (\mu(v_\mathrm{g}) + d_\mathrm{b})v_\mathrm{m}\Big)}{\gamma B(x_\mathrm{m}(\boldsymbol{v}))(1 - v_\mathrm{g}^{\alpha_\mathrm{b}})}, \tag{B.9}$$

which holds regardless of the effects of deleterious mutations on the vital rates. This is a demographic representation and generalisation of the surprising simple result noted for unstructured semelparous populations of constant size that the nature of epistasis of deleterious mutations has no effect on the genetic load (Kimura and Maruyama, 1966; Gillespie, 2004).

C Individual-based simulations

We here describe how we carried out the individual-based (stochastic) simulations used for the two model examples in the main text. The simulation algorithms scrupulously implement the life-cycle assumption of these models with the only differences relative to the analytical model being that (i) population size is finite and (ii) the mutation rate at the life-history locus is positive $\mu_{LH} > 0$ (but kept small) in the simulations. This makes the evolutionary process in the simulations irreducible (see also discussion section 2.3.2) and subject to genetic drift along with mutation and natural selection.

C.1 Coevolution of reproductive effort and germline maintenance

The simulation algorithm for this scenario (see section 1.3. of S.I for the Mathematica code) follows a population composed of a finite and fixed number (=7500 in the simulations) of individuals, where each individual is described by its genetic state (vector of traits consisting of allocation to maintenance, allocation to survival and number of deleterious mutations the individual has). One life-cycle iteration then proceeds as follows. We start by computing the fecundity of each adult individual, which is determined by its trait values (eq. (6)). Then, we evaluate the survival probability of each adult individual according to its trait values (the survival of an individual is given by a Bernoulli random variable with mean given by its survival probability eq. (6)). After eliminating the dead individuals, we fill the "vacated breeding spots" by randomly sampling offspring from the relative fecundity of all adult individuals before survival, thus effectively implementing a Wright-Fisher process for reproduction (Mode and Gallop, 2008). Once a newborn is chosen to fill the breeding spot, each of its traits mutate independently with probability ($\mu_{\mathrm{LH}} = 0.01$ in our actual simulations). The effect size of a mutation follows a Normal distribution with zero mean and a standard deviation (=0.1 in our simulations). Finally, we allow for deleterious mutations to accumulate at the deleterious mutation locus according to a Poisson distribution with mean that depends on the life-history locus (as specified by eq. (6)). To obtain the results shown in 2, we initialised the simulation with a monomorphic population, with no deleterious mutations and life-history trait values given by the analytically predicted equilibrium. In Table C-4, we have depicted the time-averaged standard deviations from the main trait for parameter values in Fig. 2. In Fig. 3 we demonstrate the convergence stability of our simulations and we started the simulations away from the equilibrium for four different initial values of the traits.

C.2 Coevolution of age at maturity and germline maintenance

The simulation algorithm for this scenario (see section 2.3. of S.I for the Mathematica code) follows a population whose size is endogenously determined according to a continuous-time stochastic updating process using the so-called "thinning" algorithm described in Section 3.1 of Ferriere and Tran (2009), which allows to exactly implement our life-cycle assumptions. A thinning algorithm is essentially an algorithm to simulate the points in an inhomogeneous Poisson process (inhomogeneous Poisson processes can be simulated by "thinning" the points from the homogeneous Poisson process), where the points or events take place sequentially (see e.g. Chen, 2016 for a conceptual description). Hence, under this algorithm, each individual is described by a vector specifying its age, allocation to repair, the age at maturity, and the number of deleterious mutations the individual has. The events in the thinning algorithm then follow a Poisson point process whose mean is determined by the vital rates (eq. (12)) and where the occurrence of the events depends on the relative weights set by birth, death, and mutation rates of an individual. We defined as a "generation" $N(\boldsymbol{u}^*)$ iterations of the thinning algorithm, where $N(\boldsymbol{u}^*)$ is the analytical prediction of the carrying capacity of the model. This is so because during one iteration of the thinning algorithm, a maximum of one event can occur (birth, death, or mutation of an individual)

to one randomly chosen individual and so after having iterated the process $N(\boldsymbol{u}^*)$ times, on average the total population has been sampled. Thus, in order to produce a single data point in Fig. 4, we ran the six million($=N(\boldsymbol{u}^*) \times N_{\text{generations}} \approx 2000 \times 3000$) iterations of the thinning algorithm. The mutation rate in the life-history locus is set to $\mu_{LH} = 0.1$ and the effect size of the mutation follows a Normal distribution with zero mean and a standard deviation (=0.07 in our simulations). Simulating the results shown in 4, we initialised the simulation with a monomorphic population, where individual age is given by $a = 1/d_{\text{b}}$ (recall, that d_{b} is the baseline mortality, with no deleterious mutations and life-history trait values given by the analytically predicted equilibrium. To obtain the results shown in 4, we initialised the simulation with a monomorphic population, with no deleterious mutations and life-history trait values given by the analytically predicted equilibrium. In Table C-5, we have depicted the time-averaged standard deviations from the main trait for parameter values in Fig. 2. In Fig. 5 we demonstrate the convergence stability of our simulations and we started the simulations away from equilibrium for four different initial values of the traits.

References

Altenberg, L. 2009. The evolutionary reduction principle for linear variation in genetic transmission. Bull. Math. Biol. 71:1264–1284.

Altenberg, L., U. Liberman, and M. W. Feldman. 2017*a*. Unified reduction principle for the evolution of mutation, migration, and recombination. Proc. Natl. Acad. Sci. U.S.A. 114:E2392–E2400.

———. 2017*b*. Unified reduction principle for the evolution of mutation, migration, and recombination. Proceedings of the National Academy of Sciences of the United States of America 114:E2392–E2400.

André, J.-B., and B. Godelle. 2006. The evolution of mutation rate in finite asexual populations. Genetics 172:611–626.

Avila, P., L. Fromhage, and L. Lehmann. 2019. Sex-allocation conflict and sexual selection throughout the lifespan of eusocial colonies. Evolution 73:1116–1132.

Avila, P., and C. Mullon. 2023. Evolutionary game theory and the adaptive dynamics approach: Adaptation where individuals interact. Proceedings of the Royal Society of London Series B-Biological Sciences 378.

Avila, P., T. Priklopil, and L. Lehmann. 2021. Hamilton's rule, gradual evolution, and the optimal (feedback) control of phenotypically plastic traits. J. Theor. Biol. 526:110602.

Baumdicker, F., E. Sester-Huss, and P. Pfaffelhuber. 2020. Modifiers of mutation rate in selectively fluctuating environments. Stoch. Process. Their. Appl. 130:6843–6862.

Bode, M., L. Bode, and P. R. Armsworth. 2006. Larval dispersal reveals regional sources and sinks in the great barrier reef. Marine Ecology Progress Series 308:17–25.

Bürger, R. 2000. The Mathematical Theory of Selection, Recombination, and Mutation. John Wiley and Sons, New York.

Case, T. J. 2000. An Illustrated Guide to Theoretical Ecology. Oxford University Press, Oxford.

Caswell, H. 2000. Matrix Population Models. Sinauer Associates, Massachusetts.

Charlesworth, B. 1990. Optimization models, quantitative genetics, and mutation. Evolution 44:520–538.

———. 1994. Evolution in Age-Structured Populations. 2nd ed. Cambridge University Press, Cambridge.

Charnov, E. L. 1993. Life history invariants: some explorations of symmetry in evolutionary ecology. Oxford University Press, New York.

Chen, H.-y., C. Jolly, K. Bublys, D. Marcu, and S. Immler. 2020. Trade-off between somatic and germline repair in a vertebrate supports the expensive germ line hypothesis. Proc. Natl. Acad. Sci. U.S.A. 117:8973–8979.

Chen, Y. 2016. Thinning algorithms for simulating point processes. Florida State University, Tallahassee, FL .

Cichon, M., and J. Kozlowski. 2000. Ageing and typical survivorship curves result from optimal resource allocation. Evol. Ecol. Research 2:857–870.

Crump, S. C., and C. J. Mode. 1968. A general age-dependent branching process. 1. Journal of Mathematical Analysis and Applications 24:494–508.

Dańko, M. J., J. Kozlowski, J. W. Vaupel, and A. Baudisch. 2012. Mutation accumulation may be a minor force in shaping life history traits. PLoS One 7:e34146.

Dawson, K. J. 1998. Evolutionarily stable mutation rates. J. Theor. Biol. 194:143–157.

———. 1999. The dynamics of infinitesimally rare alleles, applied to the evolution of mutation rates and the expression of deleterious mutations. Theor. Popul. Biol. 55:1–22.

Day, T., and P. D. Taylor. 1996. Evolutionarily stable versus fitness maximizing life histories under frequency-dependent selections. Proc. R. Soc. B: Biol. Sci. 263:333–338.

———. 1997. Hamilton's rule meets the Hamiltonian: Kin selection on dynamic characters. Proc. R. Soc. B: Biol. Sci. 264:639–644.

———. 2000. A generalization of Pontryagin's maximum principle for dynamic evolutionary games among relatives. Theor. Popul. Biol. 57:339–356.

de Roos, A. M. 1997. A gentle intoduction to models of physiologically structured populations. *In* S. Tuljapurkar and H. Caswell, eds., Structured-Population Models in Marine, Terrestrial, and Freshwater Systems. Chapman & Hall, New York.

Dercole, F., and S. Rinaldi. 2008. Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and Its Applications. Princeton University Press, Princeton, NJ.

Eigen, M. 1971. Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften 58:465–523.

Eshel, I. 1983. Evolutionary and continuous stability. J. Theor. Biol. 103:99–111.

———. 1996. On the changing concept of evolutionary population stability as a reflection of a changing point of view in the quantitative theory of evolution. J. Math. Biol. 34:485–510.

Eshel, I., and M. W. Feldman. 1984. Initial increase of new mutants and some continuity properties of ess in two-locus systems. Am. Nat. 124:631–640.

Eshel, I., and U. Motro. 1988. The three brothers' problem: kin selection with more than one potential helper. 1. the case of immediate help. Am. Nat. 132:550–566.

Eshel, I., U. Motro, and E. Sansone. 1997. Continuous stability and evolutionary convergence. Journal of Theoretical Biology 074:222–232.

Eyre-Walker, A., and P. D. Keightley. 2007. The distribution of fitness effects of new mutations. Nature Reviews Genetics 8:610–618.

Ferenci, T. 2019. Irregularities in genetic variation and mutation rates with environmental stresses. Environ. Microbiol. 21:3979–3988.

Ferrière, R., and M. Gatto. 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. Theor. Popul. Biol. 48:126–171.

Ferriere, R., and V. C. Tran. 2009. Stochastic and deterministic models for age-structured populations with genetically variable traits. Pages 289–310 *in* ESAIM: Proceedings. Vol. 27. EDP Sciences.

Frank, S. A. 2008. All of life is social. Curr. Biol. 16:R648–R650.

Gabriel, W., M. Lynch, and R. Bürger. 1993. Muller's ratchet and mutational meltdowns. Evolution 47:1744–1757.

Geritz, S. A. H., E. Kisdi, G. Meszéna, and J. A. J. Metz. 1998. Evolutionarily singular strategies and the adaptive growth and branching of the evolutionary tree. Evol. Ecol. 12:35–57.

Gerrish, P. J., A. Colato, A. S. Perelson, and P. D. Sniegowski. 2007. Complete genetic linkage can subvert natural selection. Proc. Natl. Acad. Sci. U.S.A. 104:6266–6271.

Gervais, C., and D. Roze. 2017. Mutation rate evolution in partially selfing and partially asexual organisms. Genetics 207:1561–1575.

Gillespie, J. H. 1981. Mutation modification in a random environment. Evolution pages 468–476.

———. 2004. Population Genetics: a Concise Guide. Johns Hopkins University Press, Baltimore, Maryland.

Grimmett, G., and D. Stirzaker. 2001. Probability and Random Processes. Oxford University Press, Oxford.

Haigh, J. 1978. The accumulation of deleterious genes in a population—Muller's ratchet. Theor. Popul. Biol. 14:251–267.

Hamilton, W. D. 1966. The moulding of senescence by natural selection. J. Theor. Biol. 12:12–45.

———. 1967. Extraordinary sex ratios. a sex-ratio theory for sex linkage and inbreeding has new implications in cytogenetics and entomology. Science 156:477–88.

Harris, T. E. 1963. The Theory of Branching Processes. Springer, Berlin.

Holsinger, K. E., and M. W. Feldman. 1983. Modifiers of mutation rate: evolutionary optimum with complete selfing. Proc. Natl. Acad. Sci. U.S.A. 80:6732–6734.

Irie, T., and Y. Iwasa. 2005. Optimal growth pattern of defensive organs: the diversity of shell growth among mollusks. Am. Nat. 165:238–249.

Iwasa, Y. 2000. Dynamic optimization of plant growth. Evol. Ecol. Res. 2:437–455.

Iwasa, Y., and J. Roughgarden. 1984. Shoot/root balance of plants: optimal growth of a system with many vegetative organs. Theor. Popul. Biol. 25:78–105.

Johnson, T. 1999*a*. The approach to mutation–selection balance in an infinite asexual population, and the evolution of mutation rates. Proceedings of the Royal Society of London. Series B: Biological Sciences 266:2389–2397.

———. 1999*b*. Beneficial mutations, hitchhiking and the evolution of mutation rates in sexual populations. Genetics 151:1621–1631.

Karlin, S., and H. M. Taylor. 1975. A First Course in Stochastic Processes. Academic Press, San Diego.

———. 1981. A Second Course in Stochastic Processes. Academic Press, San Diego.

Kimura, M. 1967. On the evolutionary adjustment of spontaneous mutation rates. Genet. Res. 9:23–34.

Kimura, M., and T. Maruyama. 1966. The mutational load with epistatic gene interactions in fitness. Genetics 54:1337.

Kingman, J. 1992. Poisson Processes. Oxford University Press, Oxford.

Kirkwood, B. 1986. Accuracy in Molecular Processes: Its Control and Relevance to Living System. Chapman and Hall, New York.

Kirkwood, T. B. 1977. Evolution of ageing. Nature 270:301–304.

Kondrashov, A. S. 1995. Modifiers of mutation-selection balance: general approach and the evolution of mutation rates. Genet. Res. 66:53–69.

Kozlowski, J. 1992. Optimal allocation of resources to growth and reproduction: implications for age and size at maturity. Trends Ecol. Evol. 7:15–19.

Leigh, E. G. 1970. Natural selection and mutability. Am. Nat. 104:301–305.

Leimar, O. 2009*a*. Multidimensional convergence stability. Evol. Ecol. Research 11:191–208.

———. 2009*b*. Multidimensional convergence stability. Evolutionary Ecology Research 11:191–208.

León, J. A. 1976. Life histories as adaptive strategies. J. Theor. Biol. 60:301–335.

Lesaffre, T. 2021. Population-level consequences of inheritable somatic mutations and the evolution of mutation rates in plants. Proc. Royal Soc. B 288:20211127.

Liberman, U., and M. W. Feldman. 1986. Modifiers of mutation rate: a general reduction principle. Theor. Popul. Biol. 30:125–142.

Lynch, M., M. S. Ackerman, J.-F. Gout, H. Long, W. Sung, W. K. Thomas, and P. L. Foster. 2016. Genetic drift, selection and the evolution of the mutation rate. Nature Reviews Genetics 17:704–714.

Lynch, M., R. Bürger, D. Butcher, and W. Gabriel. 1993. The mutational meltdown in asexual populations. J. Hered. 84:339–344.

Maklakov, A. A., and S. Immler. 2016. The expensive germline and the evolution of ageing. Curr. Biol. 26:R577–R586.

Maklakov, A. A., and V. Lummaa. 2013. Evolution of sex differences in lifespan and aging: causes and constraints. BioEssays 35:717–724.

Maynard Smith, J. 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge.

McDonald, T. O. 2015. Modeling clonal evolution with branching processes. Ph.D. thesis. Rice University.

Medawar, P. B. 1952. An unsolved problem of biology. An Inaugural Lecture Delivered at University College, London.

Metz, J. A. J. 2011. Thoughts on the geometry of meso-evolution: collecting mathematical elements for a post-modern synthesis. Pages 193–231 *in* F. A. C. C. Chalub and J. Rodrigues, eds. The mathematics of Darwin's legacy, Mathematics and biosciences in interaction. Birkhäuser, Basel.

Metz, J. A. J., S. D. Mylius, and O. Diekmann. 2008. When does evolution optimize? Evol. Ecol. Res. 10:629–654.

Metz, J. A. J., R. M. Nisbet, and S. A. H. Geritz. 1992. How should we define fitness for general ecological scenarios? Trends Ecol. Evol. 7:198–202.

Metz, J. A. J., K. Staňková, and J. Johansson. 2016. The canonical equation of adaptive dynamics for life histories: from fitness-returns to selection gradients and Pontryagin's maximum principle. J. Math. Biol. 72:1125–1152.

Metzger, J. J., and S. Eule. 2013. Distribution of the fittest individuals and the rate of muller's ratchet in a model with overlapping generations. PLoS Comput. Biol. 9:e1003303.

Michod, R. E. 1979. Evolution of life histories in response to age-specific mortality factors. Am. Nat. 113:531–550.

Mode, C. J. 1968. A multidimensional age-dependent branching process with applications to natural selection. i. Math. Biosci. 3:1–18.

———. 1971. Multitype branching processes: theory and applications. 34. American Elsevier Publishing Company, New York.

———. 1974. Discrete time age-dependent branching processes in relation to stable population theory in demography. Mathematical Biosciences 10:73–100.

Mode, C. J., and R. J. Gallop. 2008. A review on monte carlo simulation methods as they apply to mutation and selection as formulated in wright–fisher models of evolutionary genetics. Math. Biosci. 211:205–225.

Monaghan, P., A. A. Maklakov, and N. B. Metcalfe. 2020. Intergenerational transfer of ageing: parental age and offspring lifespan. Trends Ecol. Evol 35:927–937.

Monaghan, P., and N. B. Metcalfe. 2019. The deteriorating soma and the indispensable germline: gamete senescence and offspring fitness. Proc. R. Soc. B: Biol. Sci. 286:20192187.

Mullon, C., L. Keller, and L. Lehmann. 2016. Evolutionary stability of jointly evolving traits in subdivided populations. Am. Nat. 188:175–195.

———. 2018. Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. Nat. Ecol. Evol. 2:132–140.

Mullon, C., and L. Lehmann. 2019. An evolutionary quantitative genetics model for phenotypic (co)variances under limited dispersal, with an application to socially synergistic traits. Evolution 73:1695–1728.

Nair, K. A., and C. J. Mode. 1971. The reducible multidimensional age-dependent branching processes. J. Math. Anal. Appl. 33:131–139.

Otto, S. P., and T. Day. 2007. A biologist's Guide to Mathematical Modeling in Ecology and Evolution. Princeton University Press, Princeton, NJ.

Parker, G. A., and J. Maynard Smith. 1990. Optimality theory in evolutionary biology. Science 349:27–33.

Pen, I. 2000. Reproductive effort in viscous populations. Evolution 54:293–297.

Perrin, N. 1992. Optimal resource allocation and the marginal value of organs. Am. Nat. 139:1344–1369.

Perrin, N., and R. M. Sibly. 1993. Dynamic models of energy allocation and investment. Annu. Rev. Ecol. Syst. 24.

Priklopil, T., and L. Lehmann. 2021. Metacommunities, fitness and gradual evolution. Theor. Popul. Biol. 142:12–35.

Roff, D. A. 2008. Defining fitness in evolutionary models. J. Genet. 87:339–348.

Ronce, O., and D. Promislow. 2010. Kin competition, natal dispersal and the moulding of senescence by natural selection. Proc. R. Soc. B 277:3659–67.

Rousset, F. 2004. Genetic Structure and Selection in Subdivided Populations. Princeton University Press, Princeton, NJ.

Rozhok, A., and J. DeGregori. 2019. Somatic maintenance impacts the evolution of mutation rate. BMC Evol. Biol. 19:1–17.

Rueffler, C., J. A. Metz, and T. J. Van Dooren. 2013. What life cycle graphs can tell about the evolution of life histories. Journal of mathematical biology 66:225–279.

Schaffer, W. M. 1982. The application of optimal control theory to the general life history problem. Am. Nat. 121:418–431.

Sniegowski, P. D., P. J. Gerrish, T. Johnson, and A. Shaver. 2000. The evolution of mutation rates: separating causes from consequences. Bioessays 22:1057–1066.

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

Steinsaltz, D., S. N. Evans, and K. W. Wachter. 2005. A generalized model of mutation–selection balance with applications to aging. Adv. Appl. Math. 35:16–33.

Stott, I., S. Townley, D. Carslake, and D. J. Hodgson. 2010. On reducibility and ergodicity of population projection matrix models. Methods in Ecology and Evolution 1:242–252.

Sydsaeter, K., P. Hammond, A. Seierstad, and A. Strøm. 2008. Further Mathematics for Economic Analysis. 2nd ed. Prentice Hall, Essex.

Szathmary, E., and J. Maynard Smith. 1997. From replicators to reproducers: the first major transitions leading to life. J. Theor. Biol. 187:555–571.

Taylor, P. D. 1989. Evolutionary stability in one-parameter models under weak selection. Theor. Popul. Biol. 36:125–143.

Williams, G. C. 1957. Pleiotropy, natural selection, and the evolution of senescence. Evolution 11:398–411.

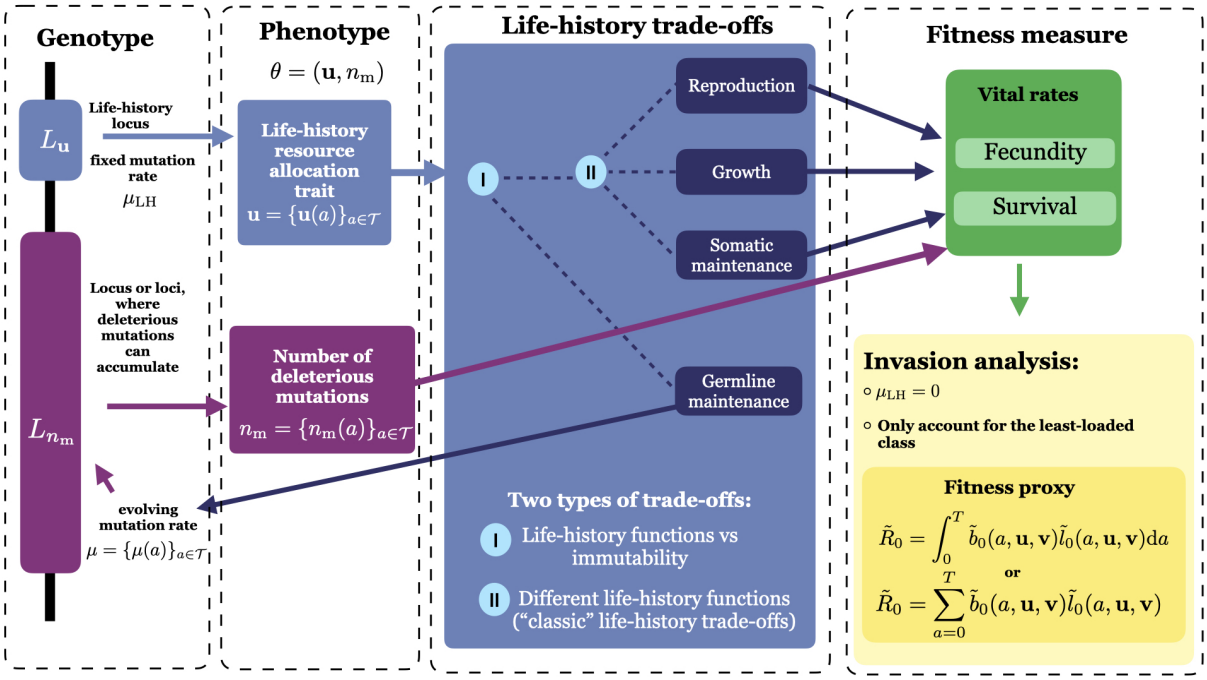

Figure 1: Key components of the life-history model with mutation accumulation. An individual's genotype is characterised by a life-history locus $L_{\boldsymbol{u}}$ and a deleterious mutation locus (purple rectangle) L_{n_m}. The mutation rate at the life-history locus μ_LH is considered to be fixed, while the mutation rate μ at the deleterious mutation locus depends on the life-history trait $\boldsymbol{u}$ and is evolving. Individuals can be characterised by the life-history allocation trajectory $\boldsymbol{u} = \{\boldsymbol{u}(a)\}_{a\in\mathcal{T}}$ (life-history trait) and the number $n_\mathrm{m} = \{n_\mathrm{m}(a)\}_{a\in\mathcal{T}}$ of deleterious mutations accumulated in the germline throughout lifespan, where a denotes the age of an individual. The resource allocation trait captures two different types of trade-offs: (i) between immutability vs life-history and (ii) between different life-history functions themselves ("classic" life history trade-offs, e.g. Stearns, 1992; Roff, 2008). Hence, the life-history locus affects the vital rates and thus fitness directly via resource allocation to life-history functions and indirectly through allocation to germline maintenance since vital rates depend on the number of deleterious mutations. For the invasion analysis we use the basic reproductive number of the least-loaded class $\tilde{R}_0$ as a fitness proxy (eqs. 1 and 4 as detailed in section 2.3.2).

Table 1: List of key symbols of the general model.

Key symbols of the model	
a	Individual age; age a can take either discrete ($a \in \{0, 1, 2, ...\}$) or continuous ($a \in [0, \infty]$) values over all possible age classes $\mathcal{T}$ (e.g. $\mathcal{T} = [0, \infty]$ in many continuous age life-history models, as maximum age is often not a fixed number).
$\boldsymbol{u}(a)$	Individual life-history trait expressed at age a (e.g. proportional allocation fecundity, survival, germline maintenance); formally, $\boldsymbol{u} : \mathcal{T} \to \mathbb{R}^n$
$\boldsymbol{u} = \{\boldsymbol{u}(a)\}_{a \in \mathcal{T}}$	Full life-history schedule over all age classes (e.g. proportional allocation of resources to fecundity from birth to death); formally, $\boldsymbol{u} \in \mathcal{U}[\mathcal{T}]$, where $\mathcal{U}[\mathcal{T}]$ is a set of all admissible life-history schedules; namely, a set of discrete or continuous real-valued functions over domain $\mathcal{T}$.
$n_{\mathrm{m}}(a)$	Number of deleterious mutations at age a in the locus where deleterious mutations can accumulate; formally, $n_{\mathrm{m}} : \mathcal{T} \to \mathbb{N}$. Since we assume asexual reproduction, the genetic details of the locus for trait $n_{\mathrm{m}}(a)$ is irrelevant (i.e. it may consist of many underlying loci).
$n_{\mathrm{m}} = \{n_{\mathrm{m}}(a)\}_{a \in \mathcal{T}}$.	Profile of deleterious mutations across all age classes; formally, $n_{\mathrm{m}} \in \mathbb{N}[\mathcal{T}]$ is an element of the space $\mathbb{N}[\mathcal{T}]$ of all possible discrete functions of range $\mathbb{N}$ over domain $\mathcal{T}$.
$\mathbf{p}(\boldsymbol{v})$	Equilibrium probability distribution for the number of deleterious mutations in the resident population carried by individuals across the different age-classes, formally $\mathbf{p}(\boldsymbol{v}) \in \Delta(\mathbb{N} \times \mathcal{T})$, where $\Delta(A)$ is the set of probability measure over set A.
$\rho_0(\boldsymbol{u}, \boldsymbol{v})$	Invasion fitness (per-capita growth rate)of zero-class individuals with mutant allele $\boldsymbol{u}$ in the population resident to trait $\boldsymbol{v}$; if the least-loaded class dominates the population.
$\tilde{R}_0(\boldsymbol{u}, \boldsymbol{v})$	Basic reproductive number of the least-loaded class, i.e. the expected number of offspring with zero deleterious mutations produced by an individual with zero deleterious mutations
$\tilde{b}_0(a, \boldsymbol{u}, \boldsymbol{v})$	Effective number of newborns with zero mutations produced by zero-class mutant individuals age a in a resident population (discrete time model); effective birth rate of newborns with zero mutations of zero class mutant individual of age a in a resident population (continuous time model).
$d_0(a, \boldsymbol{u}, \boldsymbol{v})$	Death rate of zero-class mutant individual at age a in the resident population
$\tilde{l}_0(a, \boldsymbol{u}, \boldsymbol{v})$	Probability of survival of a mutant zero-class individual to age a in a resident population.
$\mu_{\mathrm{f}}(a, \boldsymbol{u}, \boldsymbol{v})$	Rate at which germline mutations appear in an offspring when giving birth at age a.
$\mu_{\mathrm{s}}(a, \boldsymbol{u}, \boldsymbol{v})$	Rate at which germline mutations appear in an individual at age a.

Table 2: List of key symbols of "Coevolution of reproductive effort and the mutation rate model" .

Symbols for "Coevolution of reproductive effort and the mutation rate".	
u_{g}, v_{g}	Proportional allocation of resources to germline maintenance of mutant and resident individual, respectively.
u_{s}, v_{s}	Proportional allocation of resources to survival of mutant and resident individual, respectively.
N	Total population size; exogeneously fixed ($N = 7500$ in simulations).
f_{b}, s_{b}	Baseline (maximal) fecundity and probability of survival, respectively ($f_{\mathrm{b}} = 5$, $s_{\mathrm{b}} = 0.5$).
$\mu(\boldsymbol{u}) = \mu_{\mathrm{s}}(\boldsymbol{u}) = \mu_{\mathrm{f}}(\boldsymbol{u})$	Rate at which germline mutations appear in an offspring when giving birth and when surviving from generation to the next.
μ_{b}	Baseline mutation rate at which germline mutations appear; mutation rate, when no resources are allocated into germline maintenance.
α_{f}, α_{s}	Efficiency parameters (or scaling factors) of investing resources into fecundity and survival, respectively; lower values of α_{f}, α_{s} corresponds to a higher efficiency of investment; analytical results obtained for $\alpha_{\mathrm{f}} = \alpha_{\mathrm{s}} = \alpha$ ($\alpha = 0.02$, $\alpha = 0.1$, $\alpha = 0.2$ in simulations).
α_{μ}	Efficiency parameter (or scaling factor) of investing resources into germline maintenance; higher value of α_{μ} corresponds to a higher efficiency of investment; ($\alpha_{\mu} = 2$).

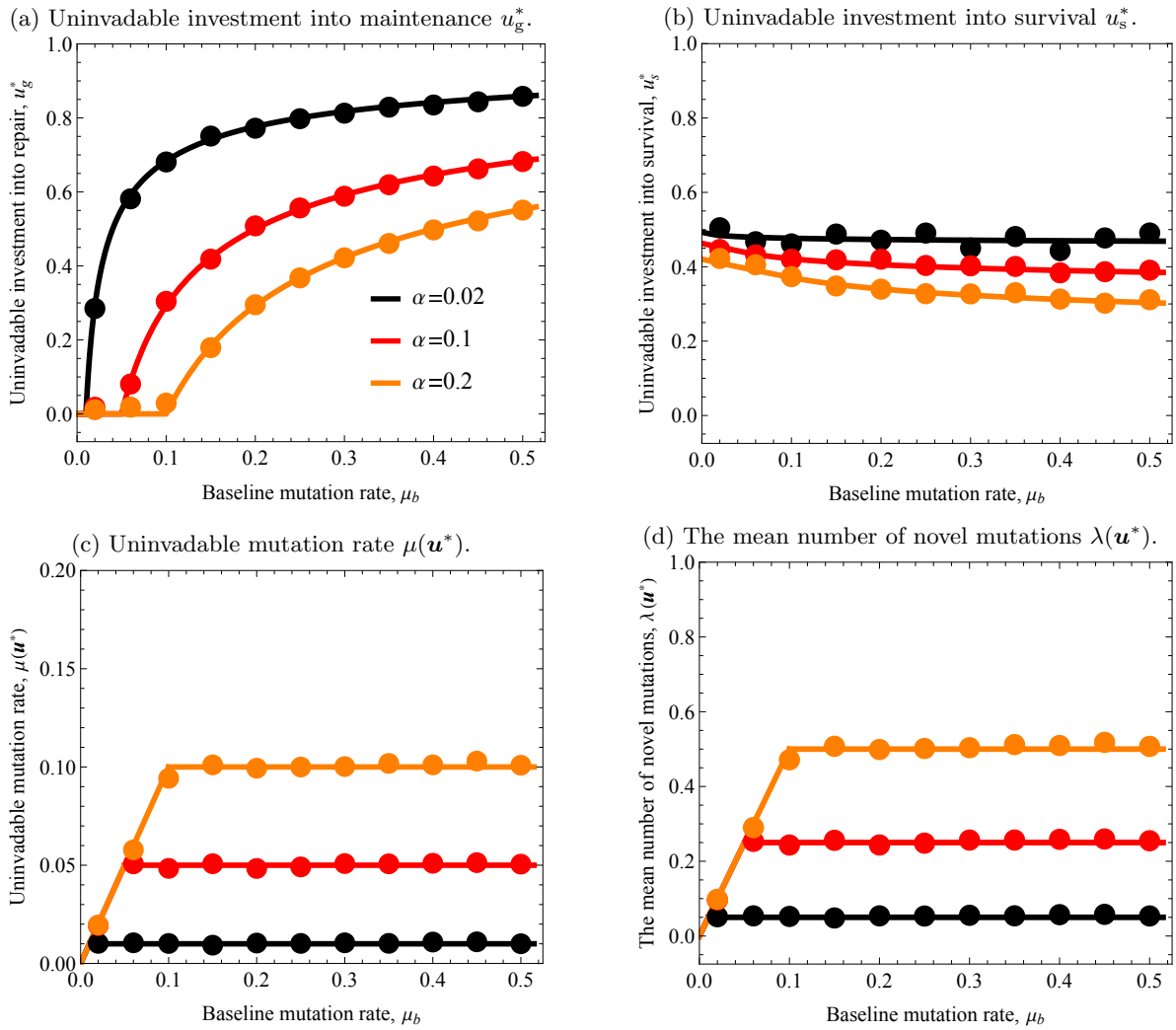

Figure 2: Predictions from the analytical model (solid lines) and from individual-based simulations of a finite population (circles) for the uninvadable life-history strategies $\boldsymbol{u}^* = (u_g^*, u_s^*)$ (panel a and b) and the mean number of novel mutations $\lambda(\boldsymbol{u}^*)$ (panel c) as functions of baseline mutation rate μ_b. The solution for the individual-based simulations is a time-averaged mean values, measured over 7500 generations while starting the simulation at the analytically predicted equilibrium (see Appendix C for details about the simulations and Table C-4 for the time-average standard deviations from the mean; see S.I. for the simulation code). The different colours represent different values of the efficiency α of reproduction and survival (smaller values of α correspond to more efficient returns from investment into vital rates). Parameter values: $f_b = 5$, $\alpha_\mu = 2$, $s_b = 0.5$, $\sigma = \sigma_f = \sigma_s = 0.2$; for simulations: $N = 7500$, $f_b = 5$, $s_b = 0.5$, the mutations in the life-history locus follow a Normal distribution with zero mean and a standard deviation of 0.1.

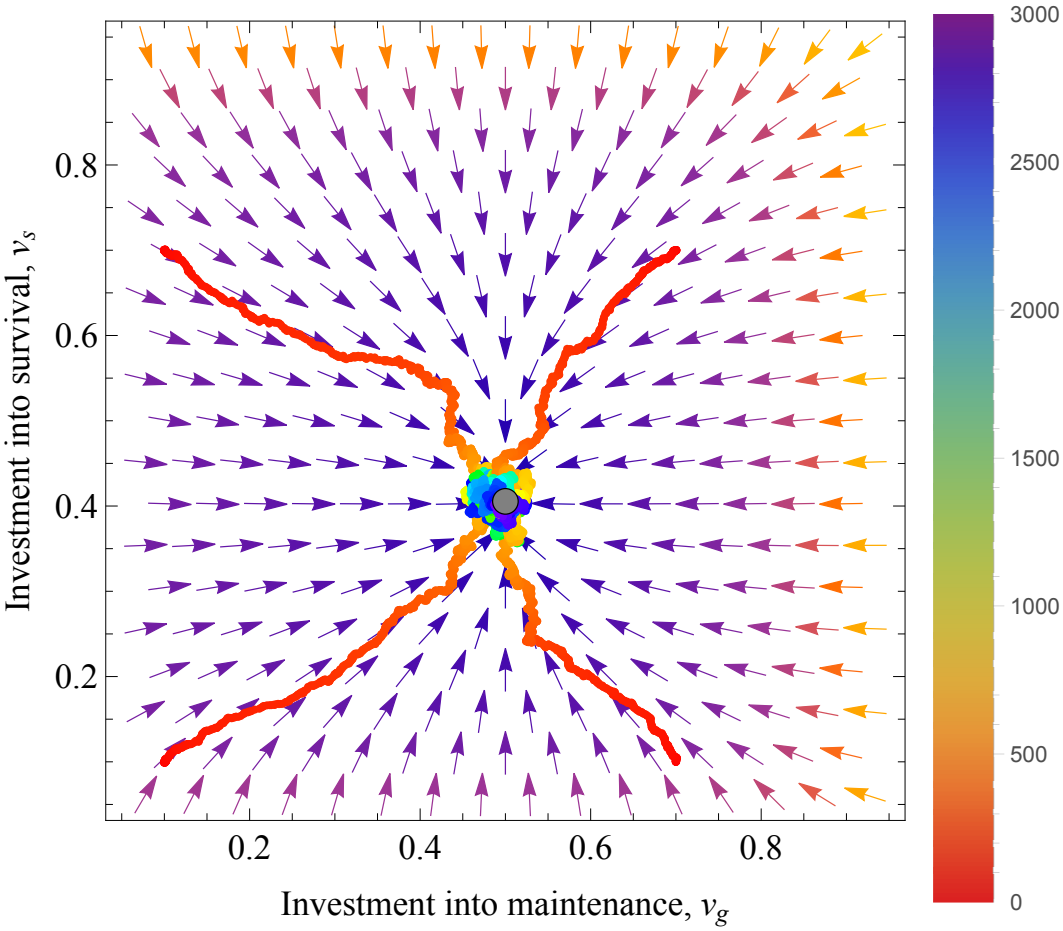

Figure 3: Evolutionary convergence towards the uninvadable life-history strategy $\boldsymbol{u}^* = (u_g^*, u_s^*) \approx (0.50, 0.40)$ (grey circle). The arrows give the analytic direction of selection at any population state (eqs. 8 and 9) and the colourful jagged lines represent the evolution of population average trait values over evolutionary time in simulations (from initial time, up to 3000 generations). Simulation were started from four different initial conditions: (i) $v_g = 0.1$, $v_s = 0.1$, (ii) $v_g = 0.1$, $v_s = 0.7$, (iii) $v_g = 0.7$, $v_s = 0.1$, and (iv) $v_g = 0.7$, $v_s = 0.7$. The colour of jagged lines indicates the number of generations since the start of the simulation (the color bar on the right-hand-side indicates the the number of generations). The simulations indicate that the population converges close to the uninvadable strategy within 3000 generations. Parameter values: $f_b = 5$, $\alpha_\mu = 2$, $s_b = 0.5$, $\sigma = \sigma_f = \sigma_s = 0.2$; for simulations: $N = 6000$, $f_b = 5$, $s_b = 0.5$, the mutations in the life-history locus follow a Normal distribution with zero mean and a standard deviation of 0.05.

Table 3: List of key symbols of "Coevolution of age at maturity and germline maintenance model".

Symbols for "Coevolution of age at maturity and germline maintenance model".	
u_{g}, v_{g}	Proportional allocation of resources to germline maintenance of mutant and resident individual, respectively.
u_{m}, v_{m}	Proportional allocation of resources to germline maintenance of mutant and resident individual, respectively.
$N(\boldsymbol{v})$	Total population size; endogeneously determined and thus depends on the resident trait $\boldsymbol{v}$.
$x(t)$	Body size of a mutant individual at age t.
$x_{\mathrm{m}}(\boldsymbol{u})$	Body size of a mutant individual at maturity.
$B(x_{\mathrm{m}}(\boldsymbol{u}))$	Surplus energy rate, i.e., rate of energy available to be allocated to life-history functions; we assume that the surplus energy scales as the power with size, i.e. energy available to mature individuals is $B(x_{\mathrm{m}}(\boldsymbol{u})) = a x_{\mathrm{m}}(\boldsymbol{u})^c$.
μ_{f}	Rate at which germline mutations appear in an offspring when giving birth (fixed parameter, $\mu_{\mathrm{f}} = 0$ in simulations).
$\mu_{\mathrm{s}}(u_{\mathrm{g}})$	Rate at which germline mutations appear in a mutant individual over time (independent of age).
μ_{b}	Baseline mutation rate at which germline mutations appear; mutation rate, when no resources are allocated into germline maintenance.
d_{b}	Baseline death rate.

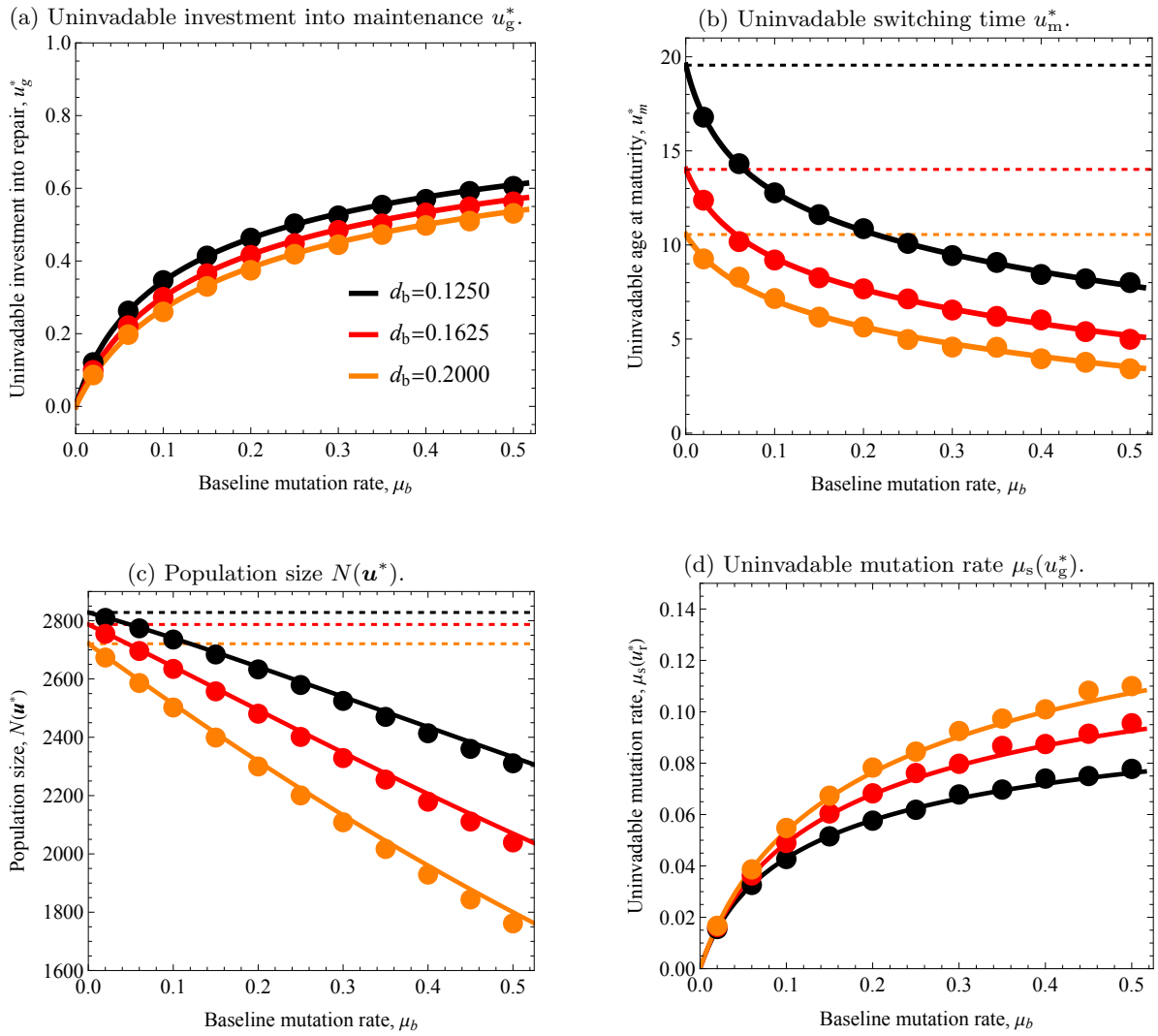

Figure 4: Predictions from the analytical model (solid lines) and from individual-based simulations (circles obtained as averages) for uninvadable life-history strategies $\boldsymbol{u}^* = (u_m^*, u_g^*)$ (panel a and b), population size $N(\boldsymbol{u}^*)$ (panel c) and mutation rate $\mu_s(u_g^*)$ as functions of baseline mutation rate μ_b for different values of baseline mortality d_b ($d_b = 0.1250$ - black, $d_b = 0.1625$ - red, $d_b = 0.2$ - orange). The dashed lines represent the "classical life-history" prediction (i.e. when $\mu(u_g^*) \to 0$ and $u_g^* \to 0$), where the colours of the dashed represent the different values for baseline death rate d_b parameter and match the values of solid lines ($d_b = 0.1250$ - black, $d_b = 0.1625$ - red, $d_b = 0.2$ - orange). The solution for the individual-based simulations are obtained as time-averaged mean values measured over 3000 "generations" while starting the simulation at analytically predicted equilibrium for the trait values and population size (Table C-5 for the time-average standard deviations from the mean; see S.I. section 2.3 for the code and for more details). The different colours represent different values of baseline mortality rate d_b. Parameter values: $\sigma = \sigma_b = \sigma_d = 0.2$, $x_0 = 1$, $a = 0.9$, $c = 0.75$, $\gamma = 0.00035$, $\beta = 1$, $\mu_f = 0$, the mutations in the life-history locus follows a Normal distribution with zero mean and a standard deviation of 0.07.

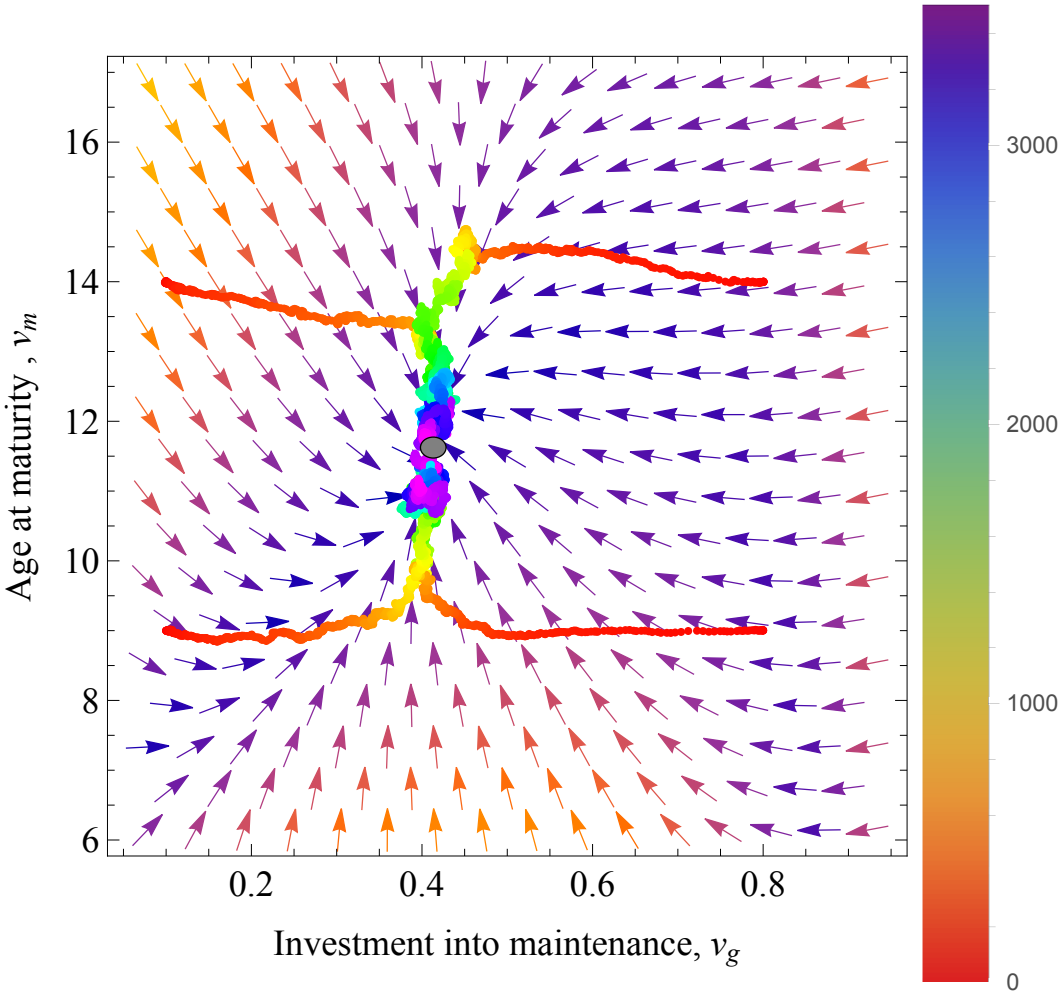

Figure 5: Evolutionary convergence to the uninvadable life-history strategy $\boldsymbol{u}^* = (u_{\mathrm{g}}^*, u_{\mathrm{m}}^*) \approx (0.41, 11.6)$ (grey circle). The arrows give the direction of selection at any resident population state (eqs. 16 and 17) and the colourful jagged lines represent the evolution of the population average trait values over evolutionary time in simulations (from initial time, up to 3500 generations). Simulation were started from four different initial conditions: (i) $v_{\mathrm{g}} = 0.1$, $v_{\mathrm{m}} = 9$, (ii) $v_{\mathrm{g}} = 0.1$, $v_{\mathrm{m}} = 14$, (iii) $v_{\mathrm{g}} = 0.8$, $v_{\mathrm{m}} = 9$, and (iv) $v_{\mathrm{g}} = 0.8$, $v_{\mathrm{s}} = 14$. The colour of jagged lines indicates the number of generations since the start of the simulation (the color bar on the right-hand-side indicates the the number of generations). The simulations indicate that the population converges close to the uninvadable strategy within 3500 generations. Parameter values: $\sigma = \sigma_{\mathrm{b}} = \sigma_{\mathrm{d}} = 0.2$, $x_0 = 1$, $a = 0.9$, $c = 0.75$, $\gamma = 0.00035$, $\beta = 1$, $\mu_{\mathrm{f}} = 0$, $d_{\mathrm{b}} = 0.1250$, the mutations in the life-history locus follow a Normal distribution with zero mean and a standard deviation of 0.07.

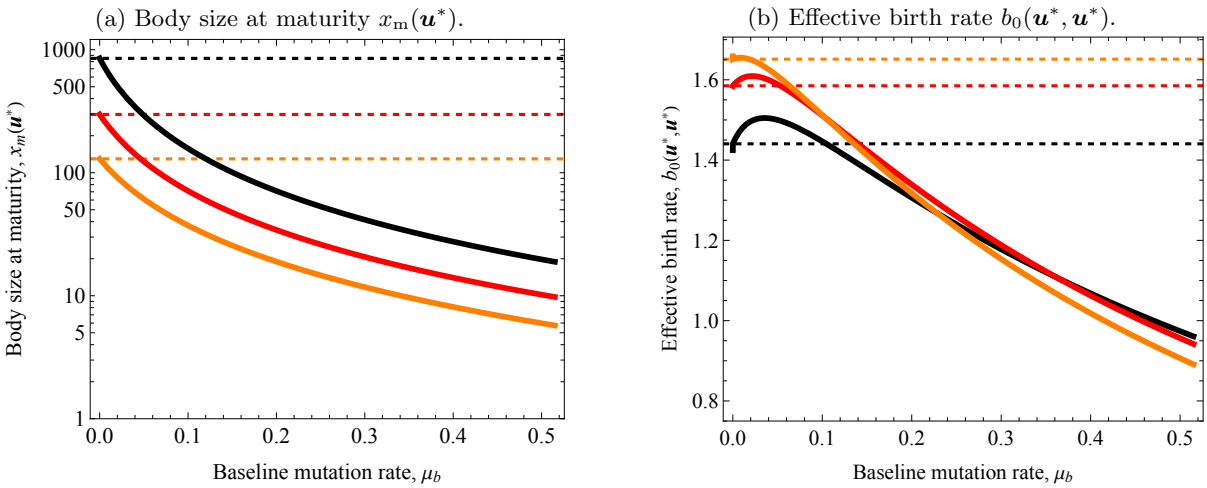

Figure 6: Predictions from the analytical model for the body size at maturity $x_m(\boldsymbol{u}^*)$ and the effective birth rate $b_0(\boldsymbol{u}^*, \boldsymbol{u}^*)$ at the uninvadable population state as a function of baseline mutation rate for different values of mortality rate d_b ($d_b = 0.1250$ - black, $d_b = 0.1625$ - red, $d_b = 0.2$ - black). The dashed lines represent the "classical life-history" prediction (i.e. when $\mu(u_g^*) \to 0$ and $u_g^* \to 0$), where the colours of the dashed represent the different values for d_b parameter and match the values of solid lines ($d_b = 0.1250$ - black, $d_b = 0.1625$ - red, $d_b = 0.2$ - black). Parameter values: $\sigma = \sigma_b = \sigma_d = 0.2$, $x_0 = 1$, $a = 0.9$, $c = 0.75$, $\gamma = 0.00035$, $\beta = 1$, $\mu_f = 0$.

Table C-4: List of standard deviations from the mean (measured over 7500 generations), for the same parameter values as in Figure 2.

Standard deviations from the mean	
$\alpha = 0.02$, black dots in Figure 2	
u_g^*	0.000777063, 0.000249208, 0.0000854773, 0.0000947966, 0.000101641, 0.0000581778, 0.0000428978, 0.0000407067, 0.0000428447, 0.000030804, 0.0000336098
u_s^*	0.000391751, 0.0005146, 0.000661941, 0.000727254, 0.000479774, 0.000537005, 0.000484383, 0.000848656, 0.00056459, 0.000804211, 0.000868807
$\alpha = 0.1$, black dots in Figure 2	
u_g^*	0.0000302984, 0.0000981018, 0.000179639, 0.0000996037, 0.000115564, 0.0000812885, 0.000067068, 0.000065316, 0.0000504356, 0.0000463299, 0.0000454094
u_s^*	0.000179931, 0.000183683, 0.000207313, 0.000220466, 0.000222421, 0.000228642, 0.000146617, 0.000105734, 0.000203143, 0.000130473, 0.000201466
$\alpha = 0.2$, black dots in Figure 2	
u_g^*	0.0000054602, 0.0000182402, 0.0000575673, 0.000186935, 0.000116924, 0.000142281, 0.0000772622, 0.0000947318, 0.0000501496, 0.0000505749, 0.0000701535
u_s^*	0.000127794, 0.000136637, 0.000128277, 0.000117642, 0.000154658, 0.0000981669, 0.0000895615, 0.000126695, 0.000160533, 0.000109858, 0.0000800033

Table C-5: List of standard deviations from the mean (measured over 1500 generations), for the same parameter values as in Figure 4.

Standard deviations from the mean	
$d_\mathrm{b} = 0.1250$ (black dots in Figure 4 panels (a) and (b))	
u_g^*	0.000125415, 0.000103887, 0.0000660474, 0.000065451, 0.0000702922, 0.0000786035, 0.000147816, 0.0000738585, 0.0000553008, 0.0001155, 0.0000481214
u_m^*	0.00256546, 0.000744261, 0.0135737, 0.00215342, 0.0011091, 0.00201475, 0.000621817, 0.00197732, 0.0116396, 0.0176991, 0.0105113
$d_\mathrm{b} = 0.1625$ (red dots in Figure 4 panels (a) and (b))	
u_g^*	0.000103311, 0.0000832005, 0.000120597, 0.0000964527, 0.00010082, 0.0000995489, 0.0000850969, 0.0000694514, 0.0000542216, 0.000100299, 0.0000734388
u_m^*	0.000491412, 0.0050249, 0.00440557, 0.00522081, 0.0133449, 0.000882752, 0.00231452, 0.00266059, 0.00252432, 0.0030314, 0.00192048
$d_\mathrm{b} = 0.2000$ (orange dots in Figure 4 panels (a) and (b))	
u_g^*	0.000153254, 0.000118768, 0.0000812339, 0.00022346, 0.0000685336, 0.000188124, 0.0000844065, 0.0000943718, 0.000145967, 0.000132834, 0.0000880514
u_m^*	0.00237881, 0.000748884, 0.00358715, 0.00309034, 0.00149716, 0.00786109, 0.00192429, 0.0187296, 0.0126454, 0.0081957, 0.00739136