

Selection on age-specific survival: Constant versus fluctuating environment

Stefano Giaimo

Department of Evolutionary Theory, Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306 Plön, Germany

ARTICLE INFO

Article history:

Received 21 January 2022
Available online 17 May 2022

Keywords:

Aging
Elasticity
Environmental stochasticity
Hamilton
Reproductive value

ABSTRACT

According to a classic result in evolutionary biodemography, selection on age-specific survival invariably declines with reproductive age. The result assumes proportional changes in survival and a constant environment. Here, we look at selection on age-specific survival when changes are still proportional but the environment fluctuates. We find that selection may or may not decline with reproductive age depending on how exactly survival is proportionally altered by mutations. However, interpreted in neutral terms, the mathematics behind the classic result capture a general property that the genetics of populations with age structure possess both in a constant and in a fluctuating environment.

© 2022 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Populations often face changing environmental conditions. As adaptation to one condition may be suboptimal for a different one, variability in the environment can profoundly impact the ecology and evolution of populations (Morris and Doak, 2002; Lande et al., 2003). From a theoretical perspective, inclusion of fluctuating environments in one's model often is neither trivial nor inconsequential. Steiner (2020) has recently recalled two main examples where the work of Tuljapurkar on environmental stochasticity in demography (Tuljapurkar, 1982a,b; Tuljapurkar and Orzack, 1980) has both exposed the risk of not properly accounting for variability in the environment and prompted the reconsideration of results based on constant environmental conditions. One example is about the estimation of the long-run growth of a population that goes through a stochastically determined sequence of environments. Tuljapurkar (1982a) showed how, when environments are independent, one can approximate the long-run growth rate as the difference between the deterministic growth the population would have in the average environment and a nonzero term capturing the covariance structure of traits in the population. This exposes the flaw inherent in using growth in the average environment alone to estimate growth under environmental stochasticity.

The other example reported by Steiner (2020), and the one of interest for the present work, is about how selection acts on traits that have specific ages of expression. In a classic work, Hamilton (1966) computed selection gradients on a number of life history traits for a population that lives in a constant environment. These

gradients display a declining pattern over reproductive ages. Accordingly, any trait change of a given magnitude that is limited to a single age would have a smaller fitness effect the later the age under consideration. With fluctuating environments, however, fitness also depends on the variability in life history traits. This variability is not accounted for in Hamilton's results. Therefore, when the environment changes with time, there is no obvious reason that selection should operate progressively less forcefully with age as it does when the environment is constant (Steiner, 2020).

However, there seems to be no specific analysis in the literature about how Hamilton's study of age-specific selection translates, or fails to do so, to the case of stochastically fluctuating environments. To provide such analysis is the aim of the present work. Attention is here limited to selection on age-specific survival. Hamilton (1966) also studied selection on mutations with other sorts of effects, e.g., protracted over several ages, with influence on the timing of other mutations or on fecundity. But his result about age-declining selection on age-specific survival is traditionally seen as his most notable result (Kirkwood and Holliday, 1979; Partridge and Barton, 1993; Charlesworth, 1993; Baudisch, 2005; Caswell, 2007; Rose et al., 2007; Flatt and Schmidt, 2009; Hughes, 2010; Wachter et al., 2013; Flatt and Partridge, 2018; Maklakov and Chapman, 2019). Together with Hamilton's gradient on age-specific fecundity, his gradient on age-specific survival is often used as a basis for the more general claim that the force of selection always declines with reproductive age.

This is how we shall proceed. Initially, we review demographic notions and the computation of selection gradients for class-structured populations in a constant environments. Then, we recall Hamilton's classic result and we elaborate upon its connection with elasticity analysis of population growth in a constant

E-mail address: giaimo@evolbio.mpg.de.

environment. Subsequently, we review demographic notions and the computation of selection gradients for class-structured populations in a stochastically fluctuating environment. Building on the contributions of Tuljapurkar et al. (1990, 2003) to the elasticity analysis of population growth in variable environments, we then look at three possible ways in which Hamilton’s classic result could be translated to changing demography induced by environmental stochasticity. In one case, we obtain an exact generalization of Hamilton’s result, while in the other two cases we find counterexamples to a steady decline of selection on survival with reproductive age. Finally, we connect Hamilton’s result with the general framework of genetic lineages in class-structured populations. We argue that this connection helps understand both his original result for a constant environment and the one case in which the result generalizes to a fluctuating environment.

2. Selection on survival in a constant environment

2.1. Demographic model

Throughout we adopt the framework of matrix population models, which is standard in the study of the ecology and evolution of class-structured populations under both constant and stochastic demography (Caswell, 2001). In this framework, classes may be based on any relevant individual property, e.g. age, size, physiological state or developmental phase, that can be used to subdivide individuals in the population into separate classes. Class abundances in the population are the components of a vector $\mathbf{x}(t)$, with $x_j(t)$ the abundance in class j at t . The total population size at t is $X(t)$. With a constant environment and no density or frequency dependent dynamics, there always are $a_{i,j}$ individuals in class i at $t+1$ per individual in class j at t . The matrix $\mathbf{A} = [a_{i,j}]$ projects population abundances forward in time: $\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t)$. In the long run, under assumptions about \mathbf{A} , which are here assumed to hold throughout, the population grows at each time step by a factor equal to the dominant eigenvalue $\lambda > 0$ of \mathbf{A} and reaches demographic stability: the class distribution is stable and proportional to the right dominant eigenvector \mathbf{w} of \mathbf{A} . The left dominant eigenvector \mathbf{v} is the reproductive value vector, its components measure the relative contribution of individuals in each class to the population future composition. The components of both these eigenvectors are positive. They are normalized so that

$$\begin{aligned} \mathbf{e}^T \mathbf{w} &= \sum_j w_j = 1, \\ \mathbf{v}^T \mathbf{w} &= \sum_j v_j w_j = 1, \end{aligned} \tag{1}$$

where \mathbf{e} is a vector of 1s and T indicates vector transposition.

2.2. Selection gradients

Before recalling Hamilton’s result, we briefly review the theory of selection in class-structured populations that live a constant environment. Assume that the matrix \mathbf{A} captures the demographic dynamics of a stable resident population. We consider selection in this population combining the approach of Taylor and Frank (1996) and that of Engen et al. (2011), which are here more easily connected with perturbation results for matrix models that will be useful later on. For more general approaches to selection in class-structured populations also including stochastic demography, see Rousset (2004), Rousset and Ronce (2004), Engen and Sæther (2014), Lehmann and Rousset (2014), Lion (2018) and Priklopil and Lehmann (2021).

We start by noting that individuals in a class-structured population are not all equal. An individual has different survival

and reproductive prospects, and therefore different influences on the population future composition, depending on the class this individual currently is in. The demographic contribution of an individual in class j at t to the population at $t+1$ is $\sum_i a_{i,j}$. This contribution, which includes offspring as well as the individual itself if it survives (possibly while transitioning to some other class), will generally be different from the contribution of an individual in some other class $k \neq j$. Moreover, individuals contributed by a single individual over one time step also differ among them, as they can be in different classes at $t+1$. As an extreme case, Ellner (2018) asks us to consider those offspring that die before their first census: they go undetected by the demographic model of Section 2.1 and never count towards the population composition. How should we then treat offspring that survive to first census yet are so small, compared to other offspring, that have few chances of making it to reproductive maturity? The standard way of dealing with disparity among individuals due to class structure is to weight them by their reproductive values (Fisher, 1930; Taylor, 1990; Cochran and Ellner, 1992; Engen et al., 2009; Barton and Etheridge, 2011; Ellner, 2018; Lion, 2018; Grafen, 2020), because reproductive values precisely capture future contributions to the population. In practice, instead of working with the census number $x_j(t)$ of individuals in class j at t , one works with their reproductive value $v_j x_j(t)$. The total reproductive value in the population at t is $V(t) = \sum_j v_j x_j(t)$. At demographic stability, population size and total reproductive value coincide since $V(t) = \sum_j v_j x_j(t) = X(t) \sum_j v_j w_j = X(t)$, where we have used the normalization in Eq. (1). Therefore, at demographic stability both $V(t)$ and $X(t)$ change by a factor λ per time step – note that the total reproductive value $V(t)$, differently from the census size $N(t)$, changes exactly by this factor even when the class distribution is not stable (Fisher, 1930; Caswell, 2001).

We can then define the fitness f_j of a resident individual currently in class j as its contribution to the total reproductive value in the population at the next time point relative to its current reproductive value,

$$f_j = \sum_i \frac{v_i}{v_j} a_{i,j}, \tag{2}$$

(Taylor and Frank, 1996; Engen et al., 2011). In this way, the fitness contribution of the fraction of reproductive-valued weighted individuals in class j at demographic stability is

$$v_j w_j f_j = w_j \sum_i v_i a_{i,j} = v_j w_j \lambda, \tag{3}$$

where we have used the left eigenvector equation for reproductive values. Mean fitness in the population is

$$f = \sum_j v_j w_j f_j = \lambda, \tag{4}$$

where we have used Eq. (3) and the normalization in Eq. (1). Note that mean resident fitness coincides, as it should, with stable geometric growth of the resident population.

Take then a quantitative phenotypic trait, like telomere length or aggressiveness level. The resident trait value is z . The entries of \mathbf{A} are supposed to depend on this trait. Imagine that a mutation emerges in the population so that mutants have trait value $z + \delta$ with $\delta \ll 1$. For $\delta = 0$, resident and mutant are identical. So long as the mutant subpopulation is small, it experiences the set E of conditions imposed by the resident population and mutant demographic dynamics are captured by the matrix $\mathbf{A}(\delta, E)$ with dominant eigenvalue λ^* (Metz et al., 2008). This matrix coincides with the resident projection matrix when $\delta = 0$. Mean mutant fitness is

$$f^* = \sum_j v_j w_j f_j^*(a_{i,j}(\delta, E)), \tag{5}$$

where reproductive values and class distribution are regarded as constants at resident (neutral) values. Whether or not the mutation invades depends on the sign of the quantity

$$S(z) = \left(\frac{1}{f^*} \frac{\partial f^*}{\partial \delta} \right) \Big|_{\delta=0} = \frac{1}{f} \sum_j v_j w_j \frac{\partial f_j^*}{\partial \delta} \Big|_{\delta=0} = \frac{1}{f} \sum_{i,j} v_i w_j \frac{\partial a_{i,j}(\delta, E)}{\partial \delta} \Big|_{\delta=0}, \tag{6}$$

which is the derivative of mutant fitness with respect to mutant deviation when this approaches zero. As [Van Cleve \(2020\)](#) noted, an expression equivalent to Eq. (6) and based on an eigenvalue perturbation result originally introduced to population biology by [Caswell \(1978\)](#) is

$$S(z) = \left(\frac{1}{\lambda^*} \frac{\partial \lambda^*}{\partial \delta} \right) \Big|_{\delta=0} = \frac{\partial \ln \lambda^*}{\partial \delta} \Big|_{\delta=0} = \frac{1}{\lambda} \mathbf{v}^T \frac{\partial \mathbf{A}(\delta, E)}{\partial \delta} \Big|_{\delta=0} \mathbf{w}. \tag{7}$$

The quantity $S(z)$ is the selection gradient on the trait z ([Taylor and Frank 1996](#), Eq. 29; [Rousset 2004](#), Eq. 11.16; [Ronce and Promislow 2010](#), Eq. 2.1). If $S(z) > 0$, the mutation invades. If $S(z) < 0$, the mutation disappears. Observe that division by $f (= \lambda)$ is sometimes omitted from the expression for $S(z)$. However, this scaling is needed when considering fitness on the natural log scale, as originally assumed by [Hamilton \(1966\)](#). Here, we retain his assumption throughout.

The absolute magnitude of $S(z)$ is evolutionarily relevant too, as it is usually interpreted as the present force of selection on the trait z , e.g. [Hamilton \(1966\)](#), [Baudisch \(2005\)](#), [Ronce and Promislow \(2010\)](#), [Caswell and Shyu \(2017\)](#), [Gaillard and Lemaître \(2017\)](#), [Sun et al. \(2020\)](#), [Giaino and Traulsen \(2022\)](#). This parallels a traditional interpretation of selection gradients for populations without explicit class structure, e.g. [Arnold and Wade \(1984\)](#), [Meszéná G. Kisdi et al. \(2001\)](#)

2.3. Demographic model with classification by age

To get to the result by [Hamilton \(1966\)](#), we must consider the special case of age classes. For this case, the projection matrix \mathbf{A} takes the form of the Leslie matrix, which we denote by \mathbf{L} . The probability that an individual of age j at t survives to be observed in age class $j+1$ at $t+1$ is p_j , while m_j is the fecundity of this individual. Reproduction starts at age α and it does not cease until some later age ω , which also is the maximum attainable age. Age-specific fecundity weighted by newborn survival populates the first row of \mathbf{L} . Remaining survival populates the Leslie matrix subdiagonal. Zeros are everywhere else. Hence, \mathbf{L} takes the form

$$\mathbf{L} = \begin{pmatrix} p_0 m_1 & \dots & \dots & p_0 m_\omega \\ p_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ 0 & 0 & p_{\omega-1} & 0 \end{pmatrix}, \tag{8}$$

([Charlesworth, 1994](#)). At demographic stability, the age distribution is stable and the Euler–Lotka equation,

$$\sum_{i=1}^{\omega} \left(\prod_{k=0}^{i-1} p_k \right) m_i \lambda^{-i} = 1, \tag{9}$$

relates the resident life history, i.e. the time-homogeneous schedule of age-specific survival (p_j) and fecundity (m_j), with its demographically stable geometric growth λ .

2.4. Hamilton’s result

We can now recall the original result by [Hamilton \(1966\)](#). He computed the selection gradient on age-specific survival assuming mutations of small proportional effects. Up to a linear

approximation in the mutant deviation, introducing a proportional increase in survival p_j at age j is equivalent to introducing an additive increase in $\ln p_j$. Using the above formalism, the trait of Hamilton’s interest then was $z = \ln p_j$. He assumed no density or frequency dependent dynamics in the population, so that there is no need to specify the set E of conditions imposed by the resident because different strategies do not interfere with each other. Inserting a factor $(1+\delta)$ next to p_j in Eq. (9) and substituting resident growth with mutant growth, one obtains the Euler–Lotka equation for the initial mutant population. Implicitly differentiating $\ln \lambda^*(\delta)$ with respect to δ in this equation and evaluating at $\delta = 0$, [Hamilton \(1966\)](#) found that

$$S(\ln p_j) = \sum_{i=j+1}^{\omega} \frac{\left(\prod_{k=0}^{i-1} p_k \right) m_i \lambda^{-i}}{T} > 0, \tag{10}$$

where

$$T = \sum_{i=1}^{\omega} i \left(\prod_{k=0}^{i-1} p_k \right) m_i \lambda^{-i} \tag{11}$$

is a measure of the generation time in the resident population ([Charlesworth, 1994](#)). As Hamilton noted, this selection gradient has a general age pattern. As we compute the gradient for each successive age j , we drop one term at a time. But $S(\ln p_j)$ is a sum of nonnegative terms, which are zero at pre-reproductive ages ($m_i = 0$ with $i < \alpha$), while they are positive at reproductive ages ($m_i > 0$ with $\alpha \leq i \leq \omega$). Therefore,

$$S(\ln p_0) = S(\ln p_1) = \dots = S(\ln p_{\alpha-1}) > S(\ln p_\alpha) > \dots > S(\ln p_{\omega-1}). \tag{12}$$

In words, fitness is more sensitive to a proportional change in survival at a reproductive age than to a change of the same proportion at any later age. The change has the same fitness effect at all pre-reproductive ages. This age pattern is general in the sense that it is entirely independent of survival and fecundity schedules of the resident population.

2.5. Hamilton’s result as an elasticity

Before introducing fluctuating environments, we recall a known connection between Hamilton’s result and eigenvalue perturbation methods for matrix population models. These methods are useful for our purposes because their extension for environmental stochasticity are needed when we try to translate Hamilton’s result from the case of a constant environment to the case of a variable environment.

The slope of the dominant eigenvalue λ of a projection matrix \mathbf{A} as a function of some parameter θ upon which the matrix depends is

$$\frac{\partial \lambda}{\partial \theta} = \mathbf{v}^T \frac{\partial \mathbf{A}}{\partial \theta} \mathbf{w}, \tag{13}$$

where the derivative is evaluated at θ ’s present value, see [Caswell \(1978\)](#).

Elasticity analysis studies the gradient, in a log–log plot, of λ as a function of θ when this parameter is assumed non-negative ([de Kroon et al., 1986, 2000; Caswell, 2001](#)). The elasticity of λ to θ is

$$e_\lambda(\theta) = \frac{\theta}{\lambda} \frac{\partial \lambda}{\partial \theta} = \begin{cases} \frac{\partial \ln \lambda}{\partial \ln \theta} & \theta > 0 \\ 0 & \theta = 0 \end{cases} \tag{14}$$

From Eqs. (13)–(14), the elasticity of λ to $\theta = a_{i,j}$ is

$$e_\lambda(a_{i,j}) = \frac{a_{i,j} v_i w_j}{\lambda}, \tag{15}$$

see Caswell (2001). Using this expression and the eigenvector equations $\lambda v_j = \sum_i v_i a_{i,j}$ and $\lambda w_j = \sum_i a_{j,i} w_i$, van Groenendael et al. (1994) were the first to observe a fundamental property of the elasticities of λ for any demographic classification:

$$\sum_i e_\lambda(a_{i,j}) = \frac{w_j}{\lambda} \sum_i a_{i,j} v_i = w_j v_j = \frac{v_j}{\lambda} \sum_i a_{j,i} w_i = \sum_i e_\lambda(a_{j,i}), \tag{16}$$

the sum of elasticities of λ to entries in column j equals the sum of elasticities of λ to entries in row j .

Hamilton's selection gradient on survival in Eq. (10) is known to be formally equivalent to the elasticity of λ to survival, e.g. (Caswell, 2001; Jones, 2009),

$$e_\lambda(p_j) = S(\ln p_j) \quad j = 0, 1, \dots, \omega - 1. \tag{17}$$

Here, we expand upon this observation and derive a result that, as we will see in Section 5, has an exact counterpart under environmental stochasticity. It seems that it has not been acknowledged so far that the age-pattern of Hamilton's gradient is a direct consequence of the general balance property of elasticities in Eq. (16). To show this, recall that $e_\lambda(a_{i,j})$ are nonnegative quantities because λ is an increasing function of the matrix entries. In particular, $e_\lambda(a_{i,j}) = 0$ only when the (i, j) -entry is zero, as Eq. (15) makes evident. Let us then look at row $j+1$ and column $j+1$ of the Leslie matrix,

$$\mathbf{L} = \begin{pmatrix} p_0 m_1 & \dots & \dots & p_0 m_{j+1} & \dots & \dots & \dots & p_0 m_\omega \\ p_1 & 0 & \dots & 0 & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & \vdots & \dots & \dots & \dots & 0 \\ 0 & \dots & p_j & 0 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & p_{j+1} & \ddots & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & p_{j+2} & \ddots & \dots & 0 \\ 0 & \dots & \dots & \vdots & \dots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \dots & \dots & p_{\omega-1} & 0 \end{pmatrix}, \tag{18}$$

which are here highlighted in gray. Applying the result in Eq. (16) to the Leslie matrix, we have that

$$e_\lambda(p_j) = e_\lambda(p_0 m_{j+1}) + e_\lambda(p_{j+1}) \tag{19}$$

The elasticity $e_\lambda(p_0 m_{j+1})$ is positive when $j+1$ is a reproductive age, i.e. when $m_{j+1} > 0$ so that $p_0 m_{j+1} > 0$, while this elasticity vanishes at juveniles ages, i.e. when $m_{j+1} = 0$. Therefore, Eq. (19) implies that $e_\lambda(p_j) = e_\lambda(p_{j+1})$ when $j+1 < \alpha$, where α is the first reproductive age, while $e_\lambda(p_j) > e_\lambda(p_{j+1})$ when $j+1 \geq \alpha$. Applying then Eq. (16) to the first row and first column of a Leslie matrix, $\sum_{i=1}^\omega e_\lambda(p_0 m_i) = e_\lambda(p_0 m_1) + e_\lambda(p_1)$. If $\alpha > 1$, then $e_\lambda(p_0 m_1) = 0$ and $\sum_{i \geq 1}^\omega e_\lambda(p_0 m_i) = e_\lambda(p_1)$. If $\alpha = 1$, then $e_\lambda(p_0 m_1) > 0$ and $\sum_{i \geq 1}^\omega e_\lambda(p_0 m_i) > e_\lambda(p_1)$. Putting everything together,

$$\sum_{i=1}^\omega e_\lambda(p_0 m_i) = e_\lambda(p_1) = \dots = e_\lambda(p_{\alpha-1}) > e_\lambda(p_\alpha) > \dots > e_\lambda(p_{\omega-1}). \tag{20}$$

Moreover,

$$\begin{aligned} \sum_{i=1}^\omega e_\lambda(p_0 m_i) &= \sum_{i=\alpha}^\omega \frac{\partial \ln \lambda}{\partial \ln(p_0 m_i)} = \sum_{i=\alpha}^\omega p_0 \frac{\partial(p_0 m_i)}{\partial p_0} \frac{\partial \ln \lambda}{\partial(p_0 m_i)} \\ &= p_0 \frac{\partial \ln \lambda}{\partial p_0} = e_\lambda(p_0) \end{aligned} \tag{21}$$

Therefore, from this and Eq. (17), we have that Eq. (20) corresponds to Eq. (12). In other words, we have just obtained the age-pattern of Hamilton's selection gradient on age-specific survival. Hamilton derived this pattern by reasoning on the algebra of the analytic expression of his gradient. The present derivation shows that such expression, i.e. a specific version of Eq. (6) (or Eq. (15)) for the age-classified case, is not needed to get the pattern. Structural properties of the Leslie matrix, i.e. its location of zeros and nonnegative entries, and the balance properties of elasticities for general matrix models of populations suffice.

3. Environmental stochasticity and selection

3.1. Demographic model

We now introduce stochastic fluctuations in the environment. We do so by using the model introduced by Tuljapurkar (1990). Suppose that the environment the resident population lives in can be in m possible states. A stochastic process governs the sequence of environmental states. In different environmental states, the population exhibits different vital rates so that to each environmental state there corresponds a projection matrix. Demographic projection then takes the form

$$\mathbf{x}(t+1) = \mathbf{A}(t)\mathbf{x}(t), \tag{22}$$

where the matrix $\mathbf{A}(t) \in \{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m\}$ contains the vital rates induced by the environment the population experiences at time t . In practice, if at t the environment is in state l , then $\mathbf{A}(t) = \mathbf{A}_l$. When we consider the age-classified case, the matrix in Eq. (22) is a Leslie matrix with fixed α (age at first reproduction) and ω (maximum age) for all t . Define

$$\lambda(t) = \frac{X(t+1)}{X(t)}, \tag{23}$$

the time-specific growth rate, which is the factor by which the population size X changes between t and $t+1$. Mild assumptions, here assumed to hold throughout, about the stochastic process that drives the sequence of environments and about the set of projection matrices corresponding to the environments (Cohen, 1977a,b; Tuljapurkar, 1990) ensure that the time average of the logarithm of age-specific growth rates over any realization of the demographic process eventually converges to the so-called stochastic growth rate $\ln \lambda_s$, which is given by

$$\ln \lambda_s = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \ln \lambda(t), \tag{24}$$

and is independent of the initial class distribution.

Tuljapurkar (1990) also introduced time-specific analogs of the class distribution vector and the reproductive value vector,

$$\mathbf{w}(t+1) = \frac{\mathbf{A}(t)\mathbf{w}(t)}{\lambda(t)}, \tag{25a}$$

$$\mathbf{v}^\top(t) = \frac{\mathbf{v}^\top(t+1)\mathbf{A}(t)}{\lambda(t)}, \tag{25b}$$

which are obtained from a realization of the demographic process generated by iterating Eq. (22) from $t = 0$ up to $t = \tau - 1$ and where $\mathbf{w}(0)$ and $\mathbf{v}(\tau)$ are arbitrary nonnegative vectors scaled so that the components of $\mathbf{w}(0)$ add up to 1 and $\mathbf{v}^\top(\tau)\mathbf{w}(\tau) = 1$.

The denominator $\lambda(t)$ on the right hand side of the expressions in Eq. (25) takes care of keeping this normalization throughout so that

$$\begin{aligned} \mathbf{e}^\top \mathbf{w}(t) &= \sum_j w_j(t) = 1, \\ \mathbf{v}^\top(t) \mathbf{w}(t) &= \sum_j v_j(t) w_j(t) = 1, \end{aligned} \tag{26}$$

for $0 \leq t \leq \tau$. In a constant environment, i.e. $\mathbf{A}(t) = \mathbf{A}$ for all t , the vectors in Eq. (25) converge to the right and left dominant eigenvectors of \mathbf{A} .

3.2. Selection gradients

The stochastic growth rate $\ln \lambda_s$ is the relevant measure of fitness when the environment fluctuates (Tuljapurkar, 1982b; Metz et al., 1992; Rand et al., 1994; Ferriere and Gatto, 1995), i.e. a rare mutant that experiences the set E of conditions imposed by the resident and has stochastic growth rate greater than the resident invades. As we did for the case of a constant environment, we briefly review here the theory of selection under environmental stochasticity. We will need this theory in our attempt to extend Hamilton’s result.

To compute the selection gradient on a trait z upon which the projection matrices $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_m$ depend, we proceed analogously to Section 2.2 (but see references therein for a more general approach). In particular, we retain reproductive-value weighting of individuals. Define the fitness at t of a resident individual in class j as

$$f_j(t) = \sum_i \frac{v_i(t+1)}{v_j(t)} a_{ij}(t), \tag{27}$$

which is the contribution of the individual to the total reproductive value at $t + 1$ relative to the current reproductive value of this individual. Thus, averaging over the relative abundances in the classes at t while weighting individuals by their reproductive values, mean fitness in the resident population at t is

$$\begin{aligned} f(t) &= \sum_j v_j(t) w_j(t) f_j(t) \\ &= \sum_j w_j(t) \sum_i v_i(t+1) a_{ij}(t) \\ &= \sum_j v_j(t) w_j(t) \lambda(t) \\ &= \lambda(t), \end{aligned} \tag{28}$$

where we have used Eqs. (25b)–(26) to simplify. Eqs. (24) and (28) imply that the long-run average of time-specific mean resident fitness on the log scale equals the stochastic growth rate of the resident population.

Consider now a small subpopulation of mutants with trait value $z + \delta$ that experience the conditions set by the residents. The projection matrices for the mutant subpopulation are $\mathbf{A}_1(\delta, E), \mathbf{A}_2(\delta, E), \dots, \mathbf{A}_m(\delta, E)$, which coincide with the resident matrices when $\delta = 0$. The mutant has stochastic growth rate $\ln \lambda_s^*$. Mean mutant fitness at t is

$$f^*(t) = \sum_j v_j(t) w_j(t) f_j^*(a_{ij}(\delta, E, t)), \tag{29}$$

where reproductive values and class distribution are taken at resident (neutral) values. The derivative of $\ln f^*(t)$ with respect

to δ at $\delta = 0$ is

$$\begin{aligned} S(z, t) &= \left(\frac{\partial \ln f^*(t)}{\partial \delta} \right) \Big|_{\delta=0} \\ &= \left(\frac{1}{f^*(t)} \frac{\partial f^*(t)}{\partial \delta} \right) \Big|_{\delta=0} \\ &= \frac{1}{f(t)} \sum_j v_j(t) w_j(t) \frac{\partial f_j^*(t)}{\partial \delta} \Big|_{\delta=0} \\ &= \frac{1}{f(t)} \sum_{i,j} v_i(t+1) w_j(t) \frac{\partial a_{ij}(\delta, E, t)}{\partial \delta} \Big|_{\delta=0}. \end{aligned} \tag{30}$$

However, the mutant effect may be apparent at any time point and we are interested in the expected geometric growth of the mutant on the log scale. Hence, the selection gradient of our interest, i.e. the counterpart of Eqs. (6)–(7) under environmental stochasticity, is

$$\begin{aligned} S(z) &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} S(z, t) \\ &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \frac{\mathbf{v}^\top(t+1) \frac{\partial \mathbf{A}(\delta, E, t)}{\partial \delta} \Big|_{\delta=0} \mathbf{w}(t)}{f(t)}, \end{aligned} \tag{31}$$

assuming this limit exists and where we have used matrix notation to make this expression more compact. Although the approach to $S(z)$ here has been heuristic, Eq. (31) corresponds to prior, established results by Tuljapurkar (1990) and Caswell (2005) about perturbations of the stochastic growth rate. This reassures us about the existence of the limit in Eq. (31) and tells us that we can also express this equation as

$$S(z) = \left(\frac{\partial \ln \lambda_s^*}{\partial \delta} \right) \Big|_{\delta=0}. \tag{32}$$

Looking for the cumulative change in a trait under selection in the special case of a periodic environment, Lion (2018, Eq. 23) derived a result analogous to Eq. (31). This result by Lion appears to be consistent with the perturbation analysis by Caswell (2005) of the stochastic growth rate in the presence of periodicity in the environment.

Note that the expression for $S(z)$ in Eq. (31) does not require the variability between entries of the resident projection matrices to be small. In particular, this expression does not rely on the small-noise approximation, see Section 6, to the stochastic growth rate (Tuljapurkar, 1982a; Tuljapurkar et al., 2003).

4. Proportional changes under a fluctuating environment

As we saw in Section 2.5, Hamilton’s result connects with the elasticity of λ to the entries of the projection matrix. This is because Hamilton studied the selection gradient on age-specific survival by assuming mutations with proportional effect and by equating fitness with $\ln \lambda$. Age-specific survival is an entry of the Leslie matrix and the elasticity of the dominant eigenvalue of \mathbf{L} to any entry of this matrix is a proportional sensitivity.

In the attempt of translating Hamilton’s result to stochastic environments, it is then natural to look at results in the elasticity analysis of λ_s . In particular, we recall here the contribution of Tuljapurkar et al. (2003) to this analysis. Consider any class structure. In a constant environment, introducing a proportional change solely in the (i, j) -entry of a matrix model is an unambiguous move: the resident value is the fixed quantity a_{ij} , while the mutant value is $a_{ij}(1 + \delta)$, which also is a fixed quantity. With stochastically fluctuating environments, the number $a_{ij}(t)$ of individuals in class i at $t + 1$ per individual in class j at

t in the resident population is a random variable which takes values $(\mathbf{A}_1)_{i,j}, (\mathbf{A}_2)_{i,j}, \dots, (\mathbf{A}_m)_{i,j}$, where $(\mathbf{A}_k)_{i,j}$ is the (i, j) -entry of \mathbf{A}_k . The probability to take one or other value depends on the stochastic process governing the sequence of environments; i.e. if at t the environment is in state l , then $a_{i,j}(t) = (\mathbf{A}_l)_{i,j}$. The random variable $(\mathbf{A}_1(\delta, E))_{i,j}, (\mathbf{A}_2(\delta, E))_{i,j}, \dots, (\mathbf{A}_m(\delta, E))_{i,j}$ describing the same quantity for the mutant subpopulation may differ in a proportional fashion from the resident in more than one way.

Assume, as in Hamilton (1966), that population dynamics are density and frequency independent so that there is no need to specify a set E of conditions imposed by the resident. Tuljapurkar et al. (2003) distinguished three main ways in which a new mutation in this population can induce a proportional change that only affects the (i, j) -entry of the projection matrices:

1. Mutant values are

$$(\mathbf{A}_k(\delta, E))_{i,j} = (\mathbf{A}_k)_{i,j} + \delta(\mathbf{A}_k)_{i,j}, \quad k = 1, 2, \dots, m. \quad (33)$$

In this way, all values of the random variable are proportionally perturbed and the mutant deviates in $a_{i,j}(t)$ of the same proportion with respect to the resident in all m environments, i.e. at all t . Since a small proportional increase in a positive quantity is equivalent, up to a linear approximation, to a small additive increase of this quantity on the natural log scale, the trait of interest here is $z = \ln a_{i,j}(t)$. Using Eqs. (28), (31) and (33), we have that the relevant selection gradient is

$$S(\ln a_{i,j}(t)) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \frac{v_i(t+1)a_{i,j}(t)w_j(t)}{\lambda(t)}, \quad (34)$$

which is formally equivalent to the so-called stochastic elasticity of λ_s to $a_{i,j}(t)$,

$$e_{\lambda_s}(a_{i,j}) = \frac{\partial \ln \lambda_s}{\partial \ln a_{i,j}} = S(\ln a_{i,j}(t)), \quad (35)$$

see Caswell (2001, Eq. 14.98)

2. Suppose that the mutation induces a proportional change in the standard deviation $\sigma_{i,j}$ of $a_{i,j}(t)$ while keeping the mean $\bar{a}_{i,j} = \mathbb{E}(a_{i,j}(t))$ unchanged, where \mathbb{E} is the expectation operator. The trait under selection is $z = \ln \sigma_{i,j}$. Recall that, for any random variable X and constant a , $\text{Var}(X + \delta(X + a)) = \text{Var}(X) + \delta^2 \text{Var}(X) + 2\delta \text{Cov}(X, X) = (1 + \delta)^2 \text{Var}(X)$. Hence, the random variable $X + \delta(X + a)$ has standard deviation $(1 + \delta)\sigma(X)$, where $\sigma(X)$ is the standard deviation of X . Moreover, if we set $a = -\mathbb{E}(X)$, we have that X and $X + \delta(X - \mathbb{E}(X))$ share the same mean, since $\mathbb{E}(X + \delta(X - \mathbb{E}(X))) = \mathbb{E}(X) + \delta(\mathbb{E}(X) - \mathbb{E}(X)) = \mathbb{E}(X)$. In the light of these considerations, mutant values are

$$(\mathbf{A}_k(\delta, E))_{i,j} = (\mathbf{A}_k)_{i,j} + \delta((\mathbf{A}_k)_{i,j} - \bar{a}_{i,j}), \quad k = 1, 2, \dots, m. \quad (36)$$

Using Eqs. (28), (31) and (36), the relevant selection gradient is

$$S(\ln \sigma_{i,j}) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \frac{v_i(t+1)(a_{i,j}(t) - \bar{a}_{i,j})w_j(t)}{\lambda(t)}, \quad (37)$$

This gradient is formally equivalent to the elasticity of λ_s to $\sigma_{i,j}$:

$$e_{\lambda_s}(\sigma_{i,j}) = \frac{\partial \ln \lambda_s}{\partial \ln \sigma_{i,j}} = S(\ln \sigma_{i,j}), \quad (38)$$

see Tuljapurkar et al. (2003, Eqs. 6, 9-10).

3. Finally, the mutation may induce a proportional change in the mean of $a_{i,j}(t)$ while keeping its standard deviation

unchanged so that $z = \ln \bar{a}_{i,j}$. Since for any random variable $\mathbb{E}(X + \delta \mathbb{E}(X)) = (1 + \delta)\mathbb{E}(X)$ and the standard deviation is invariant to location parameters, mutant values in this case are

$$(\mathbf{A}_k(\delta, E))_{i,j} = (\mathbf{A}_k)_{i,j} + \delta \bar{a}_{i,j}, \quad k = 1, 2, \dots, m. \quad (39)$$

Using Eqs. (28), (31) and (39), the relevant selection gradient is

$$S(\ln \bar{a}_{i,j}) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \frac{v_i(t+1)\bar{a}_{i,j}w_j(t)}{\lambda(t)}, \quad (40)$$

which is formally equivalent to the elasticity of λ_s to $\bar{a}_{i,j}$:

$$e_{\lambda_s}(\bar{a}_{i,j}) = \frac{\partial \ln \lambda_s}{\partial \ln \bar{a}_{i,j}} = S(\ln \bar{a}_{i,j}), \quad (41)$$

see Tuljapurkar et al. (2003, Eqs. 6, 7-8).

Hamilton's classic result in Eq. (10) assumes proportional changes in age-specific survival. The above distinctions show that this proportionality assumption has at least three possible counterparts under environmental stochasticity. Hence, in the attempt of extending Hamilton's result, we should consider the behavior of three separate quantities: $S(\ln p_j(t))$, $S(\ln \sigma_j)$ and $S(\ln \bar{p}_j)$, where σ_j and \bar{p}_j are, respectively, the standard deviation and the mean in survival at age j .

5. Uniform proportional changes in survival across environments

We first look at $S(\ln p_j(t))$, which is the selection gradient on survival at age j when mutations have proportional effects on survival at this age that are of the same magnitude across all environments. To study this quantity, we recall the result by Claessen (2005) that stochastic elasticities obey the same balance property in Eq. (16) as the elasticities of stable population growth. Using Eq. (25) and (34)–(35), the result by Claessen is

$$\begin{aligned} \sum_i e_{\lambda_s}(a_{i,j}) &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \sum_i \frac{v_i(t+1)a_{i,j}(t)w_j(t)}{\lambda(t)} \\ &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} v_j(t)w_j(t) \\ &= \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=1}^{\tau} \sum_i \frac{v_j(t)a_{j,i}(t-1)w_i(t-1)}{\lambda(t-1)} \\ &= \sum_i e_{\lambda_s}(a_{j,i}), \end{aligned} \quad (42)$$

which parallels Eq. (16). Applying this property to the specific case of age structure and invoking Eq. (35), the same arguments, mutatis mutandis, we put forth to show the equivalence between Eq. (12) and (16) imply that

$$\begin{aligned} S(\ln p_0(t)) &= S(\ln p_1(t)) = \dots \\ &= S(\ln p_{\alpha-1}(t)) > S(\ln p_{\alpha}(t)) > \dots > S(\ln p_{\omega-1}(t)), \end{aligned} \quad (43)$$

a result that appears to be new in the literature. Since $S(\ln p_j(t)) > 0$, there always is positive selection on mutations with uniform proportional effects on survival at age j . We can then conclude that Eq. (43) exactly replicates Hamilton's result under environmental stochasticity: the force of selection on a mutation that proportionally increases age-specific survival of the same magnitude across all environments invariably decreases with reproductive age. As in the case of a constant environment, structural properties of the Leslie matrix together with balance properties of the stochastic elasticities in general matrix models imply

the result. To get to it, no specific version of Eq. (34) for the age-classified case is needed.

6. Model with three age classes and two environments

To understand whether or not the selection gradients $S(\ln \sigma_j)$ and $S(\ln \bar{p}_j)$ follow the same age pattern as Hamilton’s original gradient, we take an approach based on counterexamples. We look for cases where $S(\ln \sigma_j)$ and $S(\ln \bar{p}_j)$ fail to decrease in magnitude with reproductive age. This would be at variance with the age pattern that is distinctive of Hamilton’s original gradient, which steadily declines with reproductive age.

To find these counterexamples, we restrict our attention to a simple demographic model for the resident population. We assume that there are three age classes, all of them reproductive, and two equiprobable environmental states. The Leslie projection matrices for the two states are

$$L_1 = \begin{pmatrix} a & b & c \\ d & 0 & 0 \\ 0 & \bar{p}_2 - \frac{\epsilon}{2} & 0 \end{pmatrix}, \tag{44a}$$

$$L_2 = \begin{pmatrix} a & b & c \\ g & 0 & 0 \\ 0 & \bar{p}_2 + \frac{\epsilon}{2} & 0 \end{pmatrix}. \tag{44b}$$

Average survival at age 1 is $\bar{p}_1 = (d + g)/2$. We use the parameter ϵ to control the variance $(\epsilon/2)^2$ in resident survival at age 2, while keeping the mean \bar{p}_2 fixed.

To obtain analytical results, we make use of the small-noise approximation of Tuljapurkar (1982a) to the stochastic growth rate. This approximation allows one to approximate $\ln \lambda_s$ in Eq. (24) for any class structure provided variability in vital rates is small. When environments are independent and identically distributed, this approximation takes the form

$$\ln \lambda_s \approx \ln \lambda - \frac{1}{2} \sum_{i,j} \sum_{l,k} \frac{\partial \ln \lambda}{\partial \bar{a}_{i,j}} \frac{\partial \ln \lambda}{\partial \bar{a}_{k,l}} \text{Cov}(a_{i,j}, a_{k,l}), \tag{45}$$

where λ is the dominant eigenvalue of the average projection matrix and the derivatives are those of $\ln \lambda$ with respect to the entries $\bar{a}_{i,j}$ of the average projection matrix. It has been suggested that approximations to perturbations of the stochastic growth rate can be obtained by perturbing Eq. (45) (Caswell, 2001; Morris and Doak, 2005; Shyu and Caswell, 2014), when the approximation therein is accurate. This will be the approach here to approximate $S(\ln \sigma_j)$ and $S(\ln \bar{p}_j)$ for the age-classified model in Eq. (44). The approximation in Eq. (45) to the stochastic growth rate of the resident for this model is

$$\ln \lambda_s \approx \ln \lambda - \left(\frac{\partial \ln \lambda}{\partial \bar{p}_1} \right)^2 \frac{(g-d)^2}{8} - \left(\frac{\partial \ln \lambda}{\partial \bar{p}_2} \right)^2 \frac{\epsilon^2}{8} - \frac{\partial \ln \lambda}{\partial \bar{p}_1} \frac{\partial \ln \lambda}{\partial \bar{p}_2} \frac{\epsilon(g-d)}{4}, \tag{46}$$

where λ is the largest real root of Euler–Lotka equation in Eq. (9) parametrized with the average matrix entries. Equivalently, λ is the dominant eigenvalue of $0.5L_1 + 0.5L_2$. For this model, Euler–Lotka equation takes the form

$$a\lambda^{-1} + \bar{p}_1 b \lambda^{-2} + \bar{p}_1 \bar{p}_2 c \lambda^{-3} = 1. \tag{47}$$

The dominant root of this equation can be found in closed form via the cubic formula,

$$\lambda = \frac{A^{\frac{1}{3}}}{3} - \frac{B}{3A^{\frac{1}{3}}} + \frac{a}{3} \tag{48}$$

with

$$A = a^3 + \frac{1}{2} \sqrt{(2a^3 + 9ab\bar{p}_1 + 27\bar{p}_1\bar{p}_2c)^2 + 4B^3}$$

$$+ \frac{9}{2} ab\bar{p}_1 + \frac{27}{2} \bar{p}_1\bar{p}_2c$$

$$B = -a^2 - 3b\bar{p}_1$$

Part of the calculation of λ via the formula in Eq. (48) may involve complex numbers, although the final result is guaranteed to be real (i.e. zero imaginary part) and positive by Perron–Frobenius theorem (Caswell, 2001). Adapting Eqs. (10)–(11) to our model, the derivatives in Eq. (46) are

$$\frac{\partial \ln \lambda}{\partial \bar{p}_1} = \frac{1}{\lambda^2 T} \left(b + \frac{c\bar{p}_2}{\lambda} \right), \tag{49a}$$

$$\frac{\partial \ln \lambda}{\partial \bar{p}_2} = \frac{c\bar{p}_1}{\lambda^3 T}, \tag{49b}$$

where

$$T = a\lambda^{-1} + 2\bar{p}_1 b \lambda^{-2} + 3\bar{p}_1 \bar{p}_2 c \lambda^{-3}, \tag{50}$$

is the generation time of the resident population for the model in Eq. (44).

7. Proportional change in the standard deviation of survival

The gradient $S(\ln \sigma_j)$ quantifies selection on the standard deviation in survival at age j when mutations have proportional effects on this standard deviation, but they leave mean survival \bar{p}_j unchanged. Instead of working directly with an expression for $S(\ln \sigma_j)$ derived from Eq. (37) for the age-classified case, we obtain approximations to $S(\ln \sigma_1)$ and $S(\ln \sigma_2)$ for the model in Eq. (44) by using Eq. (46). We then use these approximations to find a counterexample to the decrease in selection with reproductive age that we may expect if the selection gradient $S(\ln \sigma_j)$ would follow the same age-pattern as Hamilton’s original gradient.

We first consider $S(\ln \sigma_1)$. The only difference between resident and mutant is that the mutant standard deviation in survival at age 1 deviates by a factor $1 + \delta$ from the corresponding resident standard deviation. Looking at variances in survival, the deviation factor is $(1 + \delta)^2$. Thus, the stochastic growth rate of the mutant approximately is

$$\ln \lambda_s^*(\delta) \approx \ln \lambda - \left(\frac{\partial \ln \lambda}{\partial \bar{p}_1} \right)^2 (1 + \delta)^2 \frac{(g-d)^2}{8} - \left(\frac{\partial \ln \lambda}{\partial \bar{p}_2} \right)^2 \frac{\epsilon^2}{8} - \frac{\partial \ln \lambda}{\partial \bar{p}_1} \frac{\partial \ln \lambda}{\partial \bar{p}_2} (1 + \delta) \frac{\epsilon(g-d)}{4}, \tag{51}$$

where we have used the fact that resident and mutant have the same average projection matrix and, therefore, the two matrices have the same dominant eigenvalue. Differentiating Eq. (51) with respect to δ and evaluating at $\delta = 0$, we get

$$S(\ln \sigma_1) \approx - \left(\frac{\partial \ln \lambda}{\partial \bar{p}_1} \right)^2 \frac{(g-d)^2}{4} - \frac{\partial \ln \lambda}{\partial \bar{p}_1} \frac{\partial \ln \lambda}{\partial \bar{p}_2} \frac{\epsilon(g-d)}{4}, \tag{52}$$

which is an approximation to the selection gradient on the standard deviation in survival at age 1 under proportional effects of mutations. With the same strategy, we also get

$$S(\ln \sigma_2) \approx - \left(\frac{\partial \ln \lambda}{\partial \bar{p}_2} \right)^2 \frac{\epsilon^2}{4} - \frac{\partial \ln \lambda}{\partial \bar{p}_1} \frac{\partial \ln \lambda}{\partial \bar{p}_2} \frac{\epsilon(g-d)}{4}, \tag{53}$$

which is an approximation to the selection gradient on the standard deviation in survival at age 2 under proportional effects of mutations. In virtue of Eq. (48), (49)–(50), both approximations are in closed form.

Using these expressions, we report in Fig. 1 a region of the parameter space where, for a range of values of ϵ , the inequality

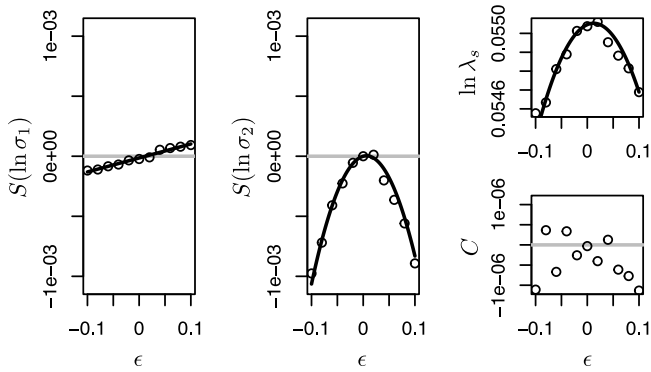


Fig. 1. Selection on standard deviation in survival at two successive reproductive ages. The panel on the left and the one at the center report estimates (dots) of, and analytic approximation (black lines) to, $S(\ln \sigma_1)$ and $S(\ln \sigma_2)$, which are the selection gradients on the standard deviation in survival at age 1 and at age 2, respectively, when mutations act in a proportional fashion. For the underlying matrix population model see main text. The parameter ϵ modulates variability in resident survival at age 2. Values for other model parameters are $a=0.01$, $b=0.05$, $c=2.4$, $d=0.9$, $g=0.88$ and $\bar{p}_2=0.525$. For a range of ϵ values, $|S(\ln \sigma_1)|$ is smaller than $|S(\ln \sigma_2)|$, thereby indicating weaker selection at age 1 than at age 2. For some $\epsilon > 0$, $S(\ln \sigma_1)$ and $S(\ln \sigma_2)$ take opposite sign. Analytic approximations come from Eqs. (52)–(53). Estimates are from simulations of the stochastic demographic process using Eq. (37) by setting $\tau=100000$. Code for these simulations is in the Supporting information. As a control for the goodness of the agreement between approximations and simulations, the upper right panel reports the resident stochastic growth rate $\ln \lambda_s$ as approximated both via Eq. (24) and from simulations. As a control for the power of simulations, the lower right panel reports $C=e_{\lambda_s}(b)+e_{\lambda_s}(p_2)-e_{\lambda_s}(p_1)$, which should be close to zero, see Eq. (42). Where relevant, a horizontal gray line is added to the panels as a visual aid to identify the zero value.

$|S(\ln \sigma_1)| < |S(\ln \sigma_2)|$ holds, which means that the force of selection on the standard deviation on survival is stronger at age 2 than at age 1. The same figure shows that, for some values of $\epsilon > 0$, the gradients $S(\ln \sigma_1)$ and $S(\ln \sigma_2)$ take opposite signs, which indicates that there may be positive selection on variability in survival at one age and positive selection on variability in survival at another age. Therefore, when mutations induce proportional changes, selection limited to the standard deviation in age-specific survival does not necessarily decrease with reproductive age. This selection can even increase with reproductive age and take different directions at different ages. This second finding recapitulates for the specific case of age classes the general finding of Tuljapurkar et al. (2003) for any class structure that the quantity $e_{\lambda_s}(\sigma_{i,j})$ may take different signs for different $a_{i,j}(t)$. These analytic results are supported by estimates based on stochastic simulations.

8. Proportional changes in mean survival

The gradient $S(\ln \bar{p}_j)$ quantifies selection on mean survival at age j when mutations have proportional effects on mean survival, yet they leave the standard deviation in survival σ_j unchanged. As we did in the previous section, instead of working directly with an expression for $S(\ln \bar{p}_j)$ derived from Eq. (40) for the age-classified case, we obtain approximations to $S(\ln \bar{p}_1)$ and $S(\ln \bar{p}_2)$ for the model in Eq. (44) by using Eq. (46). We then use these approximations to find a counterexample to the decrease in selection with reproductive age that we may expect if the selection gradient $S(\ln \bar{p}_j)$ would follow the same age-pattern as Hamilton’s original gradient.

Before obtaining these approximations, however, we recall two properties of the more general quantity $S(\ln \bar{a}_{i,j})$. First, Tuljapurkar et al. (2003) showed that, for any class structure, the

quantity $e_{\lambda_s}(\bar{a}_{i,j})$ in Eq. (41) is always positive when $(\mathbf{A}_k)_{i,j} > 0$ for at least some k . Therefore, the formally equivalent quantity $S(\ln \bar{a}_{i,j})$ always is positive whenever there is some demographic contribution of class j to class i . This indicates that there never is selection against $\bar{a}_{i,j}$. In our case, this means that there never is selection against mean age-specific survival. Second, the expressions in Eqs. (34), (37) and (40) imply that

$$S(\ln a_{i,j}(t)) = S(\ln \bar{a}_{i,j}) + S(\ln \sigma_{i,j}). \tag{54}$$

as shown in Tuljapurkar et al. (2003). Hence, mutations that proportionally increase $a_{i,j}(t)$ in all environments by the same magnitude can be interpreted as having a double effect: proportionally increasing by the same amount both the mean $\bar{a}_{i,j}$ and the standard deviation $\sigma_{i,j}$ of the demographic contribution of class- j individuals to class i . This has an immediate implication for our model when we look for a case where $S(\ln \bar{p}_1) < S(\ln \bar{p}_2)$, i.e. when selection on mean survival is stronger at age 2 than at age 1. Applying Eqs. (42) and (54) to our model, we get $S(\ln \bar{p}_1) = S(\ln p_1(t)) - S(\ln \sigma_1) = S(\ln b) + S(\ln p_2(t)) - S(\ln \sigma_1)$ and $S(\ln \bar{p}_2) = S(\ln p_2(t)) - S(\ln \sigma_2)$. Hence,

$$S(\ln \bar{p}_1) - S(\ln \bar{p}_2) = S(\ln b) - S(\ln \sigma_1) + S(\ln \sigma_2). \tag{55}$$

Having already approximations to $S(\ln \sigma_1)$ and $S(\ln \sigma_2)$, see Eqs. (52)–(53), we only need $S(\ln b) = e_{\lambda_s}(b)$ to evaluate the expression in Eq. (55). Using Eqs. (34)–(35), this elasticity is

$$e_{\lambda_s}(b) = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \frac{v_1(t+1)bw_2(t)}{\lambda(t)}. \tag{56}$$

However, we here approximate this quantity by differentiating Eq. (46) with respect to $\ln b$,

$$S(\ln b) \approx \frac{\partial \ln \lambda}{\partial \ln b} - \frac{(g-d)^2}{4} \frac{\partial \ln \lambda}{\partial \bar{p}_1} \left(\frac{\partial}{\partial \ln b} \frac{\partial \ln \lambda}{\partial \bar{p}_1} \right) - \frac{\epsilon^2}{4} \frac{\partial \ln \lambda}{\partial \bar{p}_2} \left(\frac{\partial}{\partial \ln b} \frac{\partial \ln \lambda}{\partial \bar{p}_2} \right) - \frac{\epsilon(g-d)}{4} \left[\left(\frac{\partial}{\partial \ln b} \frac{\partial \ln \lambda}{\partial \bar{p}_1} \right) \frac{\partial \ln \lambda}{\partial \bar{p}_2} + \frac{\partial \ln \lambda}{\partial \bar{p}_1} \left(\frac{\partial}{\partial \ln b} \frac{\partial \ln \lambda}{\partial \bar{p}_2} \right) \right], \tag{57}$$

where, by implicit differentiation of Eq. (47),

$$\frac{\partial \ln \lambda}{\partial \ln b} = \frac{b\bar{p}_1}{\lambda^2 T}, \tag{58}$$

and, by differentiating Eqs. (49a)–(49b),

$$\frac{\partial}{\partial \ln b} \frac{\partial \ln \lambda}{\partial \bar{p}_1} = \frac{1}{\bar{p}_1} \frac{\partial \ln \lambda}{\partial \ln b} - \frac{3\bar{p}_2 c + 2b\lambda}{\lambda^3 T} \frac{\partial \ln \lambda}{\partial \ln b} - \frac{\partial \ln \lambda}{\partial \bar{p}_1} \frac{\partial \ln T}{\partial \ln b}, \tag{59a}$$

$$\frac{\partial}{\partial \ln b} \frac{\partial \ln \lambda}{\partial \bar{p}_2} = -\frac{\partial \ln \lambda}{\partial \bar{p}_2} \left(3 \frac{\partial \ln \lambda}{\partial \ln b} + \frac{\partial \ln T}{\partial \ln b} \right), \tag{59b}$$

with

$$\frac{\partial \ln T}{\partial \ln b} = \frac{\partial \ln \lambda}{\partial \ln b} \left(2 - \frac{a\lambda^2 + 9\bar{p}_1\bar{p}_2 c + 4\bar{p}_1 b\lambda}{\lambda^3 T} \right). \tag{60}$$

Using Eqs. (48)–(50), (52)–(53) and (57)–(60), we can approximate the right hand side of Eq. (55) in closed form. We use the obtained expression to find regions of the parameter space where selection on mean survival is stronger at age 2 than at age 1 so that $S(\ln \bar{p}_1) < S(\ln \bar{p}_2)$. Fig. 2 shows one such region. This finding is supported by estimates based on stochastic simulations. Hence, under proportional changes, selection limited to mean age-specific survival may increase with reproductive age, differently from Hamilton’s original selection gradient in Eq. (10).

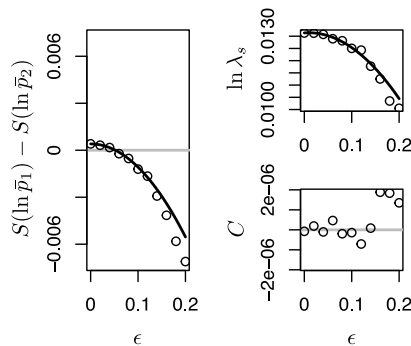


Fig. 2. Selection on mean survival at two successive reproductive ages. The left panel reports estimates (dots) of, and approximation (black lines) to, the difference between the selection gradient $S(\ln \bar{p}_1)$ on mean survival at age 1 and the selection gradient $S(\ln \bar{p}_2)$ on mean survival at age 2 when mutations have proportional effects. For the underlying matrix population model see main text. The parameter ϵ modulates variability in resident survival at age 2. Values for other model parameters are $a=0.08$, $b=0.0016$, $c=3.2$, $d=0.72$, $g=0.704$ and $\bar{p}_2=0.42$. As variability in resident survival at age 2 increases, positive selection on mean survival at age 2 intensifies to eventually become stronger than positive selection on mean survival at age 1 so that the difference $S(\ln \bar{p}_1) - S(\ln \bar{p}_2)$ becomes negative. Analytical approximations are from Eqs. (52), (53) and (55)–(60). Estimates are from simulations of the stochastic demographic process using Eq. (39) by setting $\tau=100000$. Code for these simulations is in the Supporting information. As a control for the goodness of the agreement between approximations and simulations, the upper right panel reports the resident stochastic growth rate $\ln \lambda_s$ as approximated both via Eq. (24) and from simulations. As a control for the power of simulations, the lower right panel reports $C=e_{\lambda_s}(b)+e_{\lambda_s}(p_2)-e_{\lambda_s}(p_1)$, which should be close to zero, see Eq. (42). Where relevant, a horizontal gray line is added to the panels as a visual aid to identify the zero value.

9. Revisiting Hamilton’s result

Under environmental stochasticity, whether or not selection on age-specific survival declines with reproductive age seems to depend on how proportional changes in survival are introduced. But, as we will show in this section, Hamilton’s result still contains a general truth about populations with age structure that holds both when the environment is constant and when it stochastically fluctuates. To get there, we will analyze for the special case of age structure the model of genetic lineages that is behind much of the contemporary, general theory of selection in class-structured populations under both constant and stochastic demography, e.g. Taylor (1990), Taylor and Frank (1996), Rousset (2004), Rousset and Ronce (2004), Lehmann and Rousset (2014), Gardner (2015), Lion (2018), Priklopil and Lehmann (2021).

9.1. Genetic lineage under a constant environment

Take a matrix model \mathbf{A} that governs the dynamics of a very large, demographically stable population of haploids under neutrality and with any class structure. The probability $q_{i,j}$ that a randomly sampled gene in an individual in class i at some time point descends from a gene in an individual in class j at the previous time point in this population is

$$q_{i,j} = \frac{a_{i,j}w_j}{\lambda w_i}, \tag{61}$$

because, at demographic stability, $a_{i,j}/\lambda$ is the per-time-step demographic contribution of each individual in j to class i relative to population growth and there are constant fractions w_j and w_i of the population in classes j and i , respectively. In the computation, it is irrelevant whether descent of the gene is via survival, i.e. the gene follows its carrier who moves from class j at t to class i at $t+1$, or via reproduction, i.e. a copy of the gene is transmitted

from a parent in class j at t to an offspring that is first observed in the population in class i at $t+1$. The matrix $\mathbf{Q} = [q_{i,j}]$ is the transition matrix of a Markov chain. The “system” the chain describes is a genetic lineage that, in backward time, visits various states, which are the classes in which the population is subdivided (Bodmer and Cavalli-Sforza, 1968; Hill, 1972; Taylor, 1990, 1996; Rousset, 2004; Lehmann, 2014; Bienvenu et al., 2017). At each time point, the genetic lineage is in exactly one class. Only the class where the genetic lineage currently is influences where the genetic lineage will move next. The genetic lineage moves between the classes by residing in genealogically related individuals in the population that move through the classes, e.g. by changing of size in the case of size classes. The genetic lineage resides in one individual up to the individual’s birth, at which point the genetic lineage moves to the parent of this individual. The genetic lineage then resides in the parent individual up to the birth of this, at which point the genetic lineage moves to the grandparent of our initial individual and so on. While in forward time each individual in the population dies, following the genetic lineage in backward time we only encounter rejuvenation and birth events, no deaths, and the genetic lineage persists indefinitely. By standard assumptions (i.e. irreducibility and primitivity) about the matrix model \mathbf{A} , the genetic lineage repeatedly visits all classes without ever getting trapped into one. Let then the vector $\mathbf{u}(t)$ give the current distribution of the lineage in the classes. This vector is updated in backward time via the recursion $\mathbf{u}^\top(t-1) = \mathbf{u}^\top(t)\mathbf{Q}$. The theory of Markov chains tells us that, in the long run, $\mathbf{u}(t)$ approaches the stationary distribution $\boldsymbol{\pi} = (\pi_1, \dots, \pi_n)^\top$, where n is the number of classes, of the chain. The stationary distribution, which gives the long-run fraction of time the genetic lineage spends in each class, is proportional to the left eigenvector of \mathbf{Q} corresponding to the eigenvalue 1. Thus, this distribution has the properties that $\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top \mathbf{Q}$ and $\mathbf{e}^\top \boldsymbol{\pi} = 1$. To find the stationary distribution, one can guess that $\pi_j = w_j v_j$, where $w_j v_j$ is the reproductive value of an individual in class j times the stable population fraction in class j . This quantity is often simply referred to as the reproductive value of class j . Using Eq. (61) to multiply column j of \mathbf{Q} with the vector with components $w_1 v_1, w_2 v_2, \dots, w_n v_n$, we have that

$$\sum_i w_i v_i q_{i,j} = \sum_i w_i v_i \frac{a_{i,j} w_j}{\lambda w_i} = w_j \sum_i v_i \frac{a_{i,j}}{\lambda} = w_j v_j. \tag{62}$$

Hence, setting $\boldsymbol{\pi} = (w_1 v_1, w_2 v_2, \dots, w_n v_n)^\top$ solves both $\boldsymbol{\pi}^\top = \boldsymbol{\pi}^\top \mathbf{Q}$ and $\mathbf{e}^\top \boldsymbol{\pi} = 1$, which reveals that our initial guess was correct: as is well known (Taylor, 1990; Taylor and Frank, 1996; Taylor, 1996; Rousset, 2004; Lehmann, 2014; Bienvenu et al., 2017), the reproductive value of class j corresponds with the long-run fraction of time that the genetic lineage spends in class j .

With this insight, let us go back to the specific case of age classes, i.e. \mathbf{A} is a Leslie matrix. Observe that age classes are peculiar compared to other kinds of demographic classes. For a contrast, adopt for a moment the usual, forward-time perspective on a size-classified population. An individual in this population can grow into one of several distinct classes of size depending on the individual’s current size and on the extent of the individual’s growth over a time step. The individual can persist in a size class, if it fails to grow further, or can even retrogress to a smaller size class because of lack of nutrients to support the current size. Individuals could also place their offspring in different classes, because large individuals may have offspring that are larger than those of small individuals. Similar considerations apply to class structures other than age structure.

With age classes, instead, forward time dictates rigid rules for demographic transitions. An individual is in age class j only after having gone through the exact sequence $1, 2, \dots, j-1$ of age

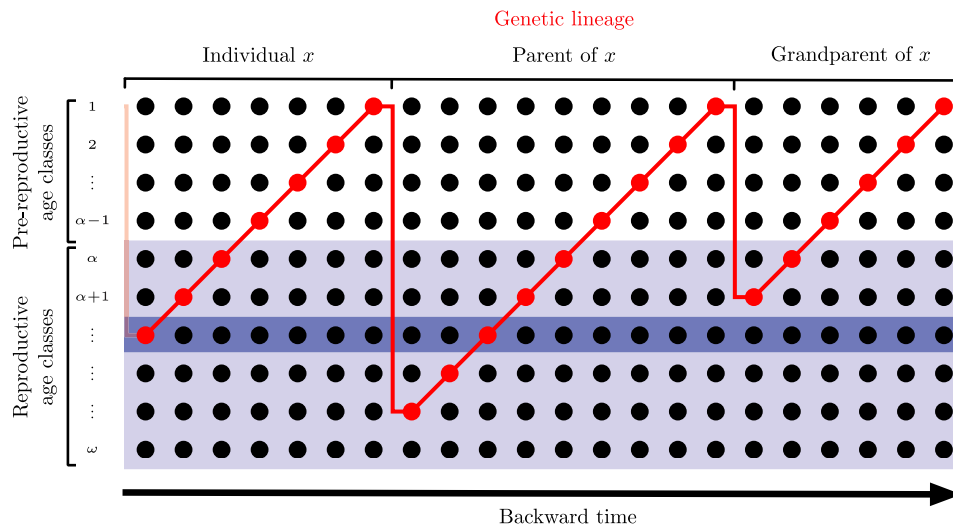


Fig. 3. Path of a genetic lineage in an age-classified population in backward time. Each dot is an age class (row) at a specific time (column). Age classes are distinguished between pre-reproductive (white background) and reproductive (light blue background). Time flows backward from left to right. A genetic lineage (red) follows a trajectory through age classes over time. We distinguish the generational passages of the genetic lineage from offspring to parent. Due to age structure, the genetic lineage always climbs up diagonally from left to right connecting each age class at one time point with the previous age class at the previous time point. Eventually, the genetic lineage hits the first age class. At this point, it “dives” into a reproductive age class and resumes diagonal climbing from there. By “diving”, the genetic lineage moves from one generation to the preceding one. Differently from diagonal climbing, which always is deterministic independently of fluctuations in the environment, the “dive depth” is stochastic: reaching one or another reproductive age class depends on the distribution of parenthood probabilities over ages at that time. Focus on a reproductive age class (dark blue row). The number of times the genetic lineage visits this age class is the number of red dots in this row. After a visit to the focal age class, e.g. at the beginning of the generation of individual x , via diagonal movement the genetic lineage visits all preceding age classes up to the first. The subsequent “dive” into a reproductive age class determines whether the genetic lineage visits the focal age class again in the preceding generation. If the “dive” is to the focal age class or any later age class (i.e. below the dark blue row), then, by diagonal climbing, the genetic lineage will visit the focal age class in this generation. This is what happens, for example, in the generation of the parent of x . If the “dive” is into a reproductive age class preceding the focal (i.e. above the dark blue row), this will not be visited by the genetic lineage in this generation. This is what happens, for example, in the generation of the grandparent of x . The closer a reproductive age class j is to pre-reproductive age classes, the more likely it is to be visited during any one generation. This is because there are more reproductive age classes other than j the genetic lineage can “dive” into at the beginning of the generation (i.e. all age classes subsequent to j) that still ensure a visit to j during that generation. Measuring time in generations, all pre-reproductive age classes and the first reproductive age class receive the same number of visits: the number of generations elapsed. Eventually, the genetic lineage visits each reproductive age class a greater number of times than any later age class.

classes one at a time. If the individual survives through age class j , then this individual is found in age class $j+1$ in the next time step, provided $j < \omega$. The only way an individual can get to the first age class is by being born and surviving to the first birthday. The parent of this individual must be in a reproductive age class ($\geq \alpha$) one time step before the offspring is first observed in the population census.

Fig. 3 illustrates the consequences of all this for the path of a genetic lineage through age classes in backward time. When the genetic lineage moves from offspring to parent, it first finds itself in the reproductive age class the parent was in while having that offspring. Which one reproductive age class exactly is hit depends on the probability distribution of ages at parenthood at that time. The offspring-to-parent transition marks the passage of the genetic lineage from one generation to the previous one. During each generation, the genetic lineage initially visits the parental age class at offspring birth and, then, all preceding age classes up to the first. This has two main implications for the probability that the genetic lineage has of visiting a given age class within any one generation. First, the probability of visiting each pre-reproductive age class and the first reproductive age class (α) is 1, because the genetic lineage will visit for sure all these age classes. Second, the probability of visiting a reproductive age class j equals the probability that the parental age class at offspring birth is j or any later age class, as the genetic lineage will visit all age classes prior to the parental age class at offspring birth. Hence, the earlier the reproductive age class j of interest, the greater the probability that it will be visited within any given generation.

Looking then over many generations backward in time, the genetic lineage visits all pre-reproductive age classes and the first

reproductive age class the same number of times, while each reproductive age class gets visited a smaller number of times than any earlier reproductive age class. Thus, the long-run fraction π_j of time the genetic lineage spends in age class j must obey the inequalities

$$\pi_1 = \pi_2 = \dots = \pi_\alpha > \pi_{\alpha+1} > \dots > \pi_\omega. \tag{63}$$

To the best of our knowledge, that the peculiarity of age structure implies this set of inequalities for the stationary distribution of the Markov chain governed by \mathbf{Q} does not seem to be appreciated in the literature.

We can now connect this back to Hamilton’s result. It is known that one can use the stable age distribution and the reproductive value to express the gradient in Eq. (10) (Goodman, 1971; Kramer et al., 2016; Hitchcock and Gardner, 2020). But see also Hill (1972) and Kobayashi and Yamamura (2007), although these authors do not refer to Hamilton’s work. The structure of the Leslie matrix in Eq. (8) and the right eigenvector equation $\lambda \mathbf{w} = \mathbf{Lw}$ imply that $\lambda w_{j+1} = p_j w_j$, because p_j is the only nonzero entry in row $j + 1$ of the Leslie matrix. Combining this fact with Eqs. (15) and (17),

$$S(\ln p_j) = w_{j+1} v_{j+1}, \tag{64}$$

(Goodman, 1971; Kramer et al., 2016; Hitchcock and Gardner, 2020), where the eigenvectors and their components are those of the Leslie matrix. But the long-run fraction π_{j+1} of time the genetic lineage spends in age class $j + 1$ equals the reproductive value of age class $j+1$, which in turn equals Hamilton’s selection gradient on survival at age j . In Eq. (63), we have then obtained once again the age-pattern (Eq. (12)) of Hamilton’s selection gradient. However, this time we did so by reasoning about the

underlying population genetics. His selection gradient happens to have the same mathematical form as a neutral quantity: the long-run fraction of backward time that a genetic lineage spends in an age class. Age structure imposes a general order relationship to this quantity that the age pattern of the gradient in Eq. (10) reflects.

9.2. Genetic lineage under a fluctuating environment

Does the long-run fraction of time spent by a genetic lineage in the age classes still obey inequalities like those in Eq. (63) when the environment fluctuates? Yes, the scheme in Fig. 3, upon which those inequalities are based, is entirely general. It does not require a stable age distribution or constant demography. The genetic lineage always moves from one age class to the previous one in one step over backward time. Fluctuating environments can only influence the probability distribution of ages at parenthood that, at every passage from one generation to the previous one, governs the move of the genetic lineage from the first age class to a reproductive age class. This distribution depends on schedules of survival and fecundity as well as on the age class abundances that are observed when the genetic lineage moves from one generation to the previous one. Survival, fecundity and age class abundances typically change in response to fluctuations in the environment. And so does the distribution of ages at parenthood, differently from the case of stable demography under a constant environment, where this distribution is constant over time. But our reasoning leading to Eq. (63) does not assume constancy of this distribution. Fluctuating environments notwithstanding, in backward time the genetic lineage eventually visits the age classes 1 through α the same number of times each, while it visits each reproductive age class a smaller number of times than any earlier reproductive age class.

A more formal argument is as follows. Consider any class structure. Under fluctuating environments, the probability in Eq. (61) that a random gene in class i at $t + 1$ comes from class j at t becomes time dependent,

$$q_{i,j}(t) = \frac{a_{i,j}(t)w_j(t)}{\lambda(t)w_i(t+1)}. \tag{65}$$

As shown by Lion (2018), the matrix $\mathbf{Q}(t) = [q_{i,j}(t)]$ updates the vector with j th component equal to $w_j(t)v_j(t)$ in backward time, since

$$\begin{aligned} w_j(t)v_j(t) &= \frac{w_j(t)}{\lambda(t)} \sum_i v_i(t+1)a_{i,j}(t) \frac{w_i(t+1)}{w_i(t+1)} \\ &= \sum_i w_i(t+1)v_i(t+1)q_{i,j}(t), \end{aligned} \tag{66}$$

where we have used Eq. (25). Therefore, the reproductive value $w_j(t)v_j(t)$ contained in class j at t is the probability that, when we look at the population at time t in the past, a genetic lineage currently present in the population is found in class j . Consider now the specific case of age classes. The long-run fraction π_{j+1}^s of time spent in age class $j + 1$ by a genetic lineage under environmental stochasticity is

$$\pi_{j+1}^s = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} v_{j+1}(t+1)w_{j+1}(t+1), \tag{67}$$

i.e. the expected reproductive value of age class $j + 1$, where the limit is assumed to exist. Using Eqs. (25a), (34) and the structure of the Leslie matrix (Eq. (8)), we can rewrite Eq. (67) as

$$\pi_{j+1}^s = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \sum_{t=0}^{\tau-1} \frac{v_{j+1}(t+1)p_j(t)w_j(t)}{\lambda(t)} = S(\ln p_j(t)), \tag{68}$$

where $p_j(t)$ is survival from age class j to age class $j + 1$ at time t . There are two things to note about Eq. (68). First, the existence of the limit in Eq. (34) reassures us about the existence of the limit in Eq. (67). Second, as we have shown in Section 5, the selection gradient $S(\ln p_j(t))$ obeys a specific order relationship with respect to age (Eq. (43)) and, therefore,

$$\pi_1^s = \pi_2^s = \dots = \pi_\alpha^s > \pi_{\alpha+1}^s > \dots > \pi_\omega^s, \tag{69}$$

which means that, under environmental stochasticity, a genetic lineage eventually spends progressively less time in a reproductive age class the later the age under consideration, while it spends the same amount of time in all pre-reproductive age classes and in the first reproductive age class.

10. Results and discussion

Hamilton (1966) found that selection on age-specific survival always gets attenuated with reproductive age. His result assumed a constant environment and proportional effects of mutations. On the basis of Tuljapurkar’s work on environmental stochasticity, Steiner (2020) suggested that Hamilton’s result may not hold in a fluctuating environment. Here, we have shown in detail that this indeed is the case. Proportional changes induced by mutations may alter different aspects of age-specific survival when environmental conditions stochastically vary. Mutations can proportionally change only mean survival, only the standard deviation in survival or both the mean and the standard deviation in survival simultaneously. We have shown analytically that, in the first two cases, selection may increase with reproductive age. When the mean and the standard deviation in survival are simultaneously proportionally altered, instead, we have proven selection on age-specific survival to always follow the age pattern of decline predicted by Hamilton.

Concerning this last case, Tuljapurkar et al. (2003) suggested that the when mutant deviations simultaneously increase the mean and the standard deviation of a trait the corresponding selection gradient may be hard to interpret as capturing a true selective force, since selection can potentially act on the mean and the standard deviation, separately. Adopting the view of these authors, our result that the fitness of an invader with proportionally increased mean and standard deviation in survival declines with reproductive age may not reveal the existence of a selective force on age-specific survival that declines with reproductive age. However, an alternative viewpoint is possible. As Hamilton (1966, p. 35) noted, “[t]o what extent and in exactly what way life schedules will be moulded by natural selection depends on what sort of genetical variation is available”. This suggests that we may shift our focus from how certain selection gradients can be theoretically decomposed to whether actual mutations proportionally alter vital rates by preferentially affecting the mean vital rate, its standard deviation or both. This, however, is an empirical problem that does not seem to be settled.

In this work, we have considered three distinct ways of changing survival in a proportional fashion. What these have in common is that the trait gets modified in some way in all environmental states or habitats. But other ways of introducing proportional changes in a trait are possible. For example, changes may only occur in an environmental state and not in others (Caswell, 2005; Horvitz et al., 2005). Moreover, different aspects of a trait variability can be proportionally modified by mutations, and not only its standard deviation (Morris et al., 2006). In the future, it would be interesting to study in detail the fitness effect of these more subtle changes on age-specific survival. However, the outcome of such study would not alter our finding that, under stochastic demography, selection may or

may not decline with reproductive age depending on how exactly survival is proportionally altered by mutations.

Hamilton (1966) interpreted his result about weakening selection on age-specific survival with age as one piece of evidence that aging is evolutionarily inevitable. Accordingly, a progressive increase in mortality with reproductive age must evolve due to the inherent bias of selection against late ages. It is currently unclear that aging is a universal trait (Jones et al., 2014). Moreover, Baudisch (2005) showed that, under constant demography, Hamilton's result may not hold if changes in age-specific survival are not proportional. The present work shows what may be regarded as a different limitation of Hamilton's result: even keeping his proportionality assumption, the predicted decline of selection on survival with age does not necessarily hold when demography is stochastic. This limitation might even be stronger than the one observed by Baudisch, since fluctuating environments appear to be the norm rather than the exception.

However, as noted above, much depends on the sort of genetic variability that is available to selection. Understanding how survival and, more generally, life history traits change as a result of genetic change, i.e. whether changes are age-specific or not, proportional or additive, is an open empirical problem (Flatt and Schmidt, 2009). Therefore, there is no obvious implication of the present work as to the evolution of aging. A further complication is the fact that mutational changes may also be specific to stages (e.g. size, phase or state) of the life cycle and not to ages (Perrin et al., 1993; McNamara and Houston, 1996; Caswell and Salguero-Gómez, 2013; Steiner et al., 2014). And, in many situations, classification by stage may be biologically more meaningful than classification by age (Caswell, 2001). Due to the non-trivial mapping of life cycle stages into ages, selection gradients for stage-specific traits may reveal an increase of the force of selection with age (Caswell and Salguero-Gómez, 2013; Steiner et al., 2014).

Aside from suggesting what may be another potential limitation to the generality of Hamilton's result, our work also has a *pars construens*. We have shown where the age pattern contained in Hamilton's result originates from. The expression for his selection gradient corresponds to the age-class reproductive value in the resident population (Goodman, 1971; Kramer et al., 2016; Hitchcock and Gardner, 2020), a neutral quantity. Class reproductive values connect with the model of genetic lineages that looks at the long-run fraction of time a gene spends in a demographic class. Using this approach, we have shown that age structure imposes rigid rules on the time spent by a gene in each age class under both constant and stochastic demography. As it turns out, the age pattern of Hamilton's selection gradient is a direct consequence of these rules.

Making explicit the connection between Hamilton's selection gradient on age-specific survival and age-class reproductive value also highlights another aspect concerning the broader interpretation of his result. Hamilton (1966) regarded his selection gradient on survival together with other selection gradients he derived that also decline with age, as evidence that, in general, selection invariably weakens with reproductive age. This idea of his has been widely shared. In particular, later authors have mostly invoked Hamilton's selection gradient on age-specific survival and his gradient on age-specific fecundity, e.g. Charlesworth (1994), Hughes and Reynolds (2005), Rose et al. (2007), Flatt and Partridge (2018), to conclude that selection wanes as age increases. Note that Hamilton's gradient on age-specific fecundity also steadily declines with reproductive age, although under less general conditions than that on survival, but Hamilton's selection gradient on age-specific fecundity presupposes additive, and not proportional, mutational effects (Baudisch, 2005). As briefly reviewed above, the idea that the general decline of selection

with age should be inferred from the age-related decline of some specific selection gradients can be challenged in different ways, e.g. type of mutational effects and appropriateness of age specificity.

A different perspective on the overall force of selection with age is based on the equivalence, which escaped Hamilton's attention (Giaino, 2021), between his selection gradient on age-specific survival and age-class reproductive value. In general, class reproductive values work as weights to the partial derivatives of class-specific fitness with respect to mutant deviations in the computation of selection gradients, as shown by the contemporary theory of selection in class-structured populations (Taylor, 1990; Taylor and Frank, 1996; Rousset, 2004; Rousset and Ronce, 2004; Lehmann and Rousset, 2014; Gardner, 2015; Priklopil and Lehmann, 2021), see Eqs. (6) and (30) in the present work. Consequently, age-class reproductive values are weights to the derivatives of age-specific fitness in selection gradients for age-classified populations (Engen et al., 2011; Hitchcock and Gardner, 2020; Avila et al., 2021). This fact, together with the decline of age-class reproductive value with reproductive age, could serve as a basis for the statement that, in general, selection always tends to weaken with age. The advantage of this perspective is that, to draw a conclusion about overall selection at different ages, no selection gradient on a specific trait, e.g. on age-specific survival and fecundity, is invoked. These gradients require additional assumptions concerning which trait should be chosen and how the trait changes upon mutation, e.g. in an additive or in proportional fashion. Such assumptions are not usually based on first principles (Moorad and Promislow, 2011), but they represent more or less arbitrary modeling choices. Class reproductive values, instead, capture general properties of the genetics of the population and naturally appear in the generalized formula for selection gradients.

The present work shows that, in a fluctuating environment, the expected age-class reproductive value declines with reproductive age. Equating the overall force of age-specific selection with the age-class reproductive value (and not with the age-trajectory of specific selection gradients), some results presented here can be seen as supporting the view that selection invariably declines with age. All in all, it appears that whether one accepts, or rejects, the idea that, in general, the force of selection always weakens with reproductive age has to do with how one conceptualizes this force.

In summary, Hamilton's result identifies a strict age pattern of selection on age-specific survival under a constant environment that may not hold when the environment fluctuates. However, his mathematics still capture a general aspect of the genetics of populations with age structure.

Funding

Funding was provided by the Max Planck Society, Germany.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.tpb.2022.05.001>.

References

- Arnold, S.J., Wade, M.J., 1984. On the measurement of natural and sexual selection: Theory. *Evolution* 38, 709–719.
- Avila, P., Priklopil, T., Lehmann, L., 2021. Hamilton's rule, gradual evolution, and the optimal (Feedback) control of phenotypically plastic traits. *J. Theor. Biol.* 526, 110602.
- Barton, N.H., Etheridge, A.M., 2011. The relation between reproductive value and genetic contribution. *Genetics* 188, 953–973.
- Baudisch, A., 2005. Hamilton's indicators of the force of selection. *Proc. Natl. Acad. Sci. USA* 102, 8263–8268.
- Bienvenu, F., Akçay, E., Legendre, S., McCandlis, D., 2017. The genealogical decomposition of a matrix population model with applications to the aggregation of stages. *Theor. Popul. Biol.* 115, 69–80.
- Bodmer, W.F., Cavalli-Sforza, L.L., 1968. A migration matrix model for the study of random genetic drift. *Genetics* 59, 565–592.
- Caswell, H., 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. *Theor. Popul. Biol.* 14, 215–230. [http://dx.doi.org/10.1016/0040-5809\(78\)90025-4](http://dx.doi.org/10.1016/0040-5809(78)90025-4), <http://www.sciencedirect.com/science/article/pii/0040580978900254>.
- Caswell, H., 2001. *Matrix Population Models*, 2nd ed. Sinauer Associates, Sunderland MA.
- Caswell, H., 2005. Sensitivity analysis of the stochastic growth rate: Three extensions. *Austr. New Zealand J. Stat.* 47, 75–85.
- Caswell, H., 2007. Extrinsic mortality and the evolution of senescence. *Trends Ecol. Evol.* 22, 173–174. <http://www.sciencedirect.com/science/article/pii/S0169534707000262>.
- Caswell, H., Salguero-Gómez, R., 2013. Age, stage and senescence in plants. *J. Ecol.* 101, 585–595.
- Caswell, H., Shyu, E., 2017. The evolution of senescence in the tree of life. Cambridge University Press, Cambridge. Chapter Senescence, Selection Gradients, and Mortality, pp. 56–82.
- Charlesworth, B., 1993. Evolutionary mechanisms of senescence. *Genetica* 91, 11–19.
- Charlesworth, B., 1994. *Evolution in age-structured populations*, 2nd ed. Cambridge University Press, Cambridge.
- Claessen, D., 2005. Alternative life-history pathways and the elasticity of stochastic matrix models. *Amer. Nat.* 165, E27–E35. <http://dx.doi.org/10.1086/427091>.
- Cochran, M.E., Ellner, S., 1992. Simple methods for calculating age-based life history parameters for stage-structured populations. *Ecol. Monograph* 62, 345–364. <http://www.jstor.org/stable/2937115>.
- Cohen, J.E., 1977a. Ergodicity of age structure in populations with markovian vital rates, II. general states. *Adv. Appl. Probab.* 9, 18–37.
- Cohen, J.E., 1977b. Ergodicity of age structure in populations with markovian vital rates, III: Finite-state moments and growth rate; An illustration. *Adv. Appl. Probab.* 9, 462–475.
- Ellner, S.P., 2018. Generation time in structured populations. *Amer. Nat.* 192, 105–110. <http://dx.doi.org/10.1086/697539>.
- Engen, S., Lande, R., Saether, B.E., 2011. Evolution of a plastic quantitative trait in an age-structured population in a fluctuating environment. *Evolution* 65, 2893–2906.
- Engen, S., Lande, R., Saether, B.E., Dobson, F.S., 2009. Reproductive value and the stochastic demography of age-structured populations. *Amer. Nat.* 174, 795–804.
- Engen, S., Saether, B.E., 2014. Evolution in fluctuating environments: decomposing selection into additive components of the robertson–price equation. *Evolution* 68, 854–865. <http://dx.doi.org/10.1111/evo.12310>.
- Ferriere, R., Gatto, M., 1995. Lyapunov exponents and the mathematics of invasion in oscillatory or chaotic populations. *Theor. Popul. Biol.* 48, 126–171.
- Fisher, R.A., 1930. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.
- Flatt, T., Partridge, L., 2018. Horizons in the evolution of aging. *BMC Biol.* 16 (93).
- Flatt, T., Schmidt, P.S., 2009. Integrating evolutionary and molecular genetics of aging. *Biochim. Biophys. Acta (BBA) - General Subjects* 1790, 951–962. <http://dx.doi.org/10.1016/j.bbagen.2009.07.010>.
- Gaillard, J.M., Lemaître, J.F., 2017. The williams' legacy: A critical reappraisal of his nine predictions about the evolution of senescence. *Evolution* 71, 2768–2785.
- Gardner, A., 2015. The genetical theory of multilevel selection. *J. Evol. Biol.* 28, 305–319. <http://dx.doi.org/10.1111/jeb.12566>.
- Gaiimo, S., 2021. Medawar and hamilton on the selective forces in the evolution of ageing. *History Philos. Life Sci.* 43 (124).
- Gaiimo, S., Traulsen, A., 2022. The selection force weakens with age because ageing evolves and not vice versa. *Nature Commun.* 13 (686).
- Goodman, L.A., 1971. On the sensitivity of the intrinsic growth rate to changes in the age-specific birth and death rates. *Theor. Popul. Biol.* 2, 339–354. [http://dx.doi.org/10.1016/0040-5809\(71\)90025-6](http://dx.doi.org/10.1016/0040-5809(71)90025-6).
- Grafen, A., 2020. The price equation and reproductive value. *Philos. Trans. R. Soc. B* 375, 20190356.
- van Groenendael, J., de Kroon, H., Kalisz, S., Tuljapurkar, S., 1994. Loop analysis: evaluating life history pathways in population projection matrices. *Ecology* 75, 2410–2415.
- Hamilton, W.D., 1966. The moulding of senescence by natural selection. *J. Theoret. Biol.* 12, 12–45.
- Hill, W.G., 1972. Effective size of populations with overlapping generations. *Theor. Popul. Biol.* 3, 278–289.
- Hitchcock, T.J., Gardner, A., 2020. A gene's-eye view of sexual antagonism. *Proc. R. Soc. B* 287, 20201633.
- Horvitz, C.C., Tuljapurkar, S., Pascarella, J.B., 2005. Plant-animal interactions in random environments: Habitat-stage elasticity, seed predators, and Hurricanes. *Ecology* 86, 3312–3322.
- Hughes, K.A., 2010. Mutation and the evolution of ageing: from biometrics to system genetics. *Philos. Trans. R. Soc. B* 365, 1273–1279.
- Hughes, K.A., Reynolds, R.M., 2005. Evolutionary and mechanistic theories of aging. *Ann. Rev. Entomol.* 50, 421–445.
- Jones, J.H., 2009. The force of selection on the human life cycle. *Evol. Hum. Behav.* 30, 305–314.
- Jones, O.R., Scheuerlein, A., Salguero-Gómez, R., Camarda, C.G., Schaible, R., Casper, B.B., Dahlgren, J.P., Ehrlén, J., García, M.B., Menges, E.S., et al., 2014. Diversity of ageing across the tree of life. *Nature* 505, 169–173.
- Kirkwood, T.B.L., Holliday, R., 1979. The evolution of ageing and longevity. *Proc. R. Soc. Lond. Ser. B* 205, 531–546. <http://www.jstor.org/stable/77444>.
- Kobayashi, Y., Yamamura, N., 2007. How to compute the effective size of spatiotemporally structured populations using separation of time scales. *Theor. Popul. Biol.* 71, 174–181.
- Kramer, B.H., van Doorn, G.S., Weissing, F.J., Pen, I., 2016. Lifespan divergence between social insect castes: challenges and opportunities for evolutionary theories of aging. *Curr. Opin. Insect Sci.* 16, 76–80.
- de Kroon, H., van Groenendael, J., Ehrlén, J., 2000. Elasticities: a review of methods and model limitations. *Ecology* 81, 607–618.
- de Kroon, H., Plaisier, A., van Groenendael, J., Caswell, H., 1986. Elasticity: the relative contribution of demographic parameters to population growth rate. *Ecology* 67, 1427–1431.
- Lande, R., Engen, S., Saether, B., 2003. *Stochastic Population Dynamics in Ecology and Conservation*. Oxford University Press, New York.
- Lehmann, L., 2014. Stochastic demography and the neutral substitution rate in class-structured populations. *Genetics* 197, 351–360. <http://dx.doi.org/10.1534/genetics.114.163345>, <https://www.genetics.org/content/197/1/351>.
- Lehmann, L., Rousset, F., 2014. The genetical theory of social behaviour. *Philos. Trans. R. Soc. B* 369, 20130357.
- Lion, S., 2018. Class structure, demography, and selection: reproductive-value weighting in nonequilibrium, polymorphic populations. *Am. Nat.* 191, 620–637. <http://dx.doi.org/10.1086/696976>.
- Maklakov, A.A., Chapman, T., 2019. Evolution of ageing as a tangle of trade-offs: energy versus function. *Proc. R. Soc. B* 286, 20191604.
- McNamara, J.M., Houston, A.I., 1996. State-dependent life histories. *Nature* 380, 215–221.
- Meszéna G. Kiszdi, U., Geritz, S.A.H., Metz, J.A.J., 2001. Evolutionary optimisation models and matrix games in the unified perspective of adaptive dynamics. *Selection* 2, 193–210.
- Metz, J.A.J., Mylius, S.D., Diekmann, O., 2008. When does evolution optimize? *Evol. Ecol. Res.* 10, 629–654.
- Metz, J., Nisbet, R., Geritz, S., 1992. How should we define 'Fitness' for general ecological scenarios? *Trends Ecol. Evol.* 7, 198–202.
- Moorad, J.A., Promislow, D.E.L., 2011. Evolutionary demography and quantitative genetics: age-specific survival as a threshold trait. *Proc. R. Soc. B* 278, 144–151. <http://dx.doi.org/10.1098/rspb.2010.0992>, <http://rspb.royalsocietypublishing.org/content/278/1702/144.abstract>.
- Morris, W., Doak, D.F., 2002. *Quantitative Conservation Biology*. Sinauer Associates, Sunderland MA.
- Morris, W.F., Doak, D.F., 2005. How general are the determinants of the stochastic population growth rate across nearby sites? *Ecol. Monogr.* 75, 119–137.
- Morris, W.F., Tuljapurkar, S., Haridas, C.V., Menges, E.S., Horvitz, C.C., Pfister, C.A., 2006. Sensitivity of the population growth rate to demographic variability within and between phases of the disturbance cycle. *Ecol. Lett.* 9, 1331–1341.
- Partridge, L., Barton, N.H., 1993. Optimality, mutation and the evolution of ageing. *Nature* 362, 305–311. <http://dx.doi.org/10.1038/362305a0>.
- Perrin, N., Sibly, R., Nichols, N., 1993. Optimal growth strategies when mortality and production rates are size-dependent. *Evol. Ecol.* 7, 576–592.
- Priklopil, T., Lehmann, L., 2021. Metacommunities, fitness and gradual evolution. *Theor. Popul. Biol.* 142, 12–35.
- Rand, D., Wilson, H., McGlade, J., 1994. Dynamics and evolution: Evolutionarily stable attractors invasion exponents and phenotype dynamics. *Philos. Trans. R. Soc. B* 343, 261–283.

- Ronce, O., Promislow, D., 2010. Kin competition, natal dispersal and the moulding of senescence by natural selection. *Proc. R. Soc. B* 277, 3659–3667. <http://dx.doi.org/10.1098/rspb.2010.1095>, <http://rspb.royalsocietypublishing.org/content/277/1700/3659.abstract>.
- Rose, M.R., Rauser, C.L., Benford, G., Matos, M., Mueller, L.D., 2007. Hamilton's forces of selection after forty years. *Evolution* 61, 1265–1276.
- Rousset, F., 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton, NJ.
- Rousset, F., Ronce, O., 2004. Inclusive fitness for traits affecting metapopulation demography. *Theor. Popul. Biol.* 65, 127–141.
- Shyu, E., Caswell, H., 2014. Calculating second derivatives of population growth rates for ecology and evolution. *Methods Ecol. Evol.* 5, 473–482. <http://dx.doi.org/10.1111/2041-210X.12179>.
- Steiner, U.K., 2020. Tuljapurkar and orzack (1980) and tuljapurkar (1982a,b): population dynamics in variable environments. *Theor. Popul. Biol.* 133, 21–22.
- Steiner, U.K., Tuljapurkar, S., Coulson, T., 2014. Generation time, net reproductive rate, and growth in stage-age-structured populations. *Am. Nat.* 183, 771–783.
- Sun, Z., Parvinen, K., Heino, M., Metz, J.A.J., De Roos, A.M., Dieckmann, U., 2020. Evolution of reproduction periods in seasonal environments. *Amer. Nat.* 196, E88–E109.
- Taylor, P.D., 1990. Allele-frequency change in a class-structured population. *Am. Nat.* 135, 95–106.
- Taylor, P.D., 1996. Inclusive fitness arguments in genetic models of behaviour. *J. Math. Biol.* 34, 654–674.
- Taylor, P.D., Frank, S.A., 1996. How to make a kin selection model. *J. Theoret. Biol.* 180, 27–37.
- Tuljapurkar, S., 1982a. Population dynamics in variable environments, II. Correlated environments, sensitivity analysis and dynamics. *Theor. Popul. Biol.* 21, 114–140.
- Tuljapurkar, S., 1982b. Population dynamics in variable environments, III. evolutionary dynamics of R-selection. *Theor. Popul. Biol.* 21, 141–165.
- Tuljapurkar, S., 1990. *Population Dynamics in Variable Environments*. Springer Berlin Heidelberg.
- Tuljapurkar, S., Horvitz, C.C., Pascarella, J.B., 2003. The many growth rates and elasticities of populations in random environments. *Amer. Nat.* 162, 489–502, <http://www.jstor.org/stable/10.1086/378648>.
- Tuljapurkar, S., Orzack, S., 1980. Population dynamics in variable environments I, long-run growth rates and extinction. *Theor. Popul. Biol.* 18, 314–342.
- Van Cleve, J., 2020. Building a synthetic basis for kin selection and evolutionary game theory using population genetics. *Theor. Popul. Biol.* 133, 65–70.
- Wachter, K.W., Evans, S.N., Steinsaltz, D., 2013. The age-specific force of natural selection and biodemographic walls of death. *Proc. Natl. Acad. Sci.* 110, 10141–10146.