Positive selection of senescence through increased evolvability: ageing is not a by-product of evolution.

Roget T.[1], Jolivet P.[2], Méléard S.[3] and Rera M.[4]

Affiliations:

[1] Institut Montpelliérain Alexander Grothendieck (IMAG), Université de Montpellier, France

[2] IRIT-APO - Algorithmes Parallèles et Optimisation - Institut de recherche en informatique de Toulouse, France

[3] Institut Universitaire de France et École Polytechnique, CNRS, Institut polytechnique de Paris, 91 128 Palaiseau, France

[4] Centre de Recherche Interdisciplinaire, INSERM UMR U1284, 75004 Paris, France

corresponding author: michael.rera@cri-paris.org

Abstract

The possibility of ageing being directly selected through evolution has been discussed for the past hundred years. As ageing is occurring, by definition, only late in life - i.e. after the organismal development is finalized -, many think that it cannot be actively selected for as a process. In addition, by decreasing an individual's fitness, it is thought unlikely to be selected for. In order to explain the observation of its broad presence in the realm of life, numerous theories have been proposed in the past 75 years, in agreement with this view.

Here, building upon a simple life-history trait model that we recently introduced and that summarizes the life of an organism to its two core abilities - reproduce and thrive -, we discuss the possibility of ageing being selected for through evolution.

Our model suggests that senescence can be positively selected through evolution thanks to the higher evolvability it confers to organisms, not through a given mechanism but through a function "ageing", limiting organismal maintenance and ability to reproduce. It provides an elegant explanation for the apparent tradeoff between longevity and fertility that led to the disposable soma theory without requiring an energy tradeoff while confirming the substrate for mutation accumulation and antagonistic pleiotropy theories. In addition, it predicts that the Lansing effect should be present in organisms showing rapid post-reproductive senescence. This formal and numerical modeling of ageing evolution also provides new hints to test the validity of existing theories.

Introduction

Ageing can be defined as the effect of elapsed time on an organism, it is also often referred to as 'senescence'. It manifests itself in a broad range of age-related patterns of mortality depending on the studied organism, from negligible senescence to post-reproductive death through progressive or brutal age-dependent mortality increase (Jones et al., 2014). Numerous explanatory theories of ageing have been proposed in the past century (reviewed in Kirkwood and Holliday, 1979). Whether they are focusing on the molecular aspects leading to the observed ageing phenomena or on the evolution of the process, one fundamental question remains regarding the selectability of ageing through evolution.

Soon after Charles Darwin published his theory of evolution, August Weismann proposed evolutionary arguments to explain ageing (Weismann, 1882). "His initial idea was that there exists a specific death-mechanism designed by natural selection to eliminate the old, and therefore worn-out, members of a population" (Gavrilov and Gavrilova, 2002). Since then, and probably partially due to Weismann's later changes of mind on his early theory, it is mostly accepted that "ageing is not adaptive since it reduces reproductive potential" (Kirkwood and Holliday, 1979) and

hence, fitness. Two other arguments are usually used against the selection of ageing. First, that this would have required a group selection effect stronger than selection at the individual level, which is very seldom the case (Smith, 1976). Second, the mortality rate is so high in early life that little senility is observed in the wild (reviewed in (Nussey et al., 2013)). There would thus be little opportunity for a removal mechanism to evolve (reviewed in (Johnson et al., 2019)). Since the 1950s onwards, evolutionary theories are mainly building upon the idea that ageing is a byproduct of natural selection (Fabian, 2011). A first example is Peter Medawar's theory of mutation accumulation for which ageing is caused by the progressive accumulation of deleterious mutations with effects that manifest only late in life (Medawar, 1952). Williams' antagonistic pleiotropy theory goes a little further than Medawar's, by inferring the existence of genes and mutations with antagonistic effects: beneficial at an early age, they would not be selected against for negative effects manifesting later in life. Finally, the "disposable soma" theory of ageing proposed by Thomas Kirkwood, is based on the idea that individuals have a limited amount of energy to be split between reproductive functions and (non-reproductive) maintenance of the organism, the "soma". According to Kirkwood's theory, increasing an organism's longevity would thus be associated with a decrease in growth and reproduction rates, delaying death (Kirkwood, 1977). Later, evolutionary conserved genes involved in both the regulation of longevity and organismal growth were discovered in the model organism *C. elegans* (Kenyon et al., 1993), later shown to be conserved in flies (Clancy et al., 2001) , mice (Bluher et al., 2003) and humans (van Heemst et al., 2005). Thus, genetic modulators for longevity exist and manifest themselves through evolutionarily conserved physiological mechanisms.

Nevertheless, although "the evolutionary theories of aging are closely related to the genetics of aging because biological evolution is possible only for heritable manifestations of aging"(Gavrilov and Gavrilova, 2002) and we now have countless examples of evolutionarily conserved genes playing a role in ageing across species (Partridge and Gems, 2002), the subject is still a matter of vivid debate (Kowald and Kirkwood, 2016). In their 2016 review, Axel Kowald and Thomas Kirkwood state that the "idea that aging is a programmed trait that is beneficial for the species [...] is now generally accepted to be wrong". If ageing cannot be positively selected for through evolution, can it be, at least partially, programmed? This question raised no less vivid debates than the previous one in the past century (Austad, 2004; Bredesen, 2004a, 2004b; Gavrilov and Gavrilova, 2002; Kirkwood and Melov, 2011; Longo et al., 2005; Skulachev, 2011). Two good examples of these active debates can be found in (Blagosklonny, 2013) and (Longo et al., 2005), these two works showing dramatically opposed views. The first, affirming that ageing cannot be programmed, the second that it can and will occur as it brings a "kind of population-level selection" that can be explained by kin selection. One of the final arguments given by Kowald and Kirkwood is that if ageing were to be programmed, "it would be possible experimentally to identify the responsible genes and inhibit or block their action".

Programmed or not, ageing in unicellular organisms is associated with mechanisms that discriminate new components from older ones as individuals replicate (Henderson et al., 2014; Lai et al., 2002, p. 2; Nyström, 2007; Sinclair and Guarente, 1997; Steiner, 2021). In multicellular organisms, the Lansing effect is a good candidate for such a mechanism. It is the effect through which the "progeny of old parents do not live as long as those of young parents" in rotifers (Lansing, 1954, 1947). More recently, it has been shown that older drosophila females and to some extent males tend to produce shorter lived offspring (Priest et al., 2002), zebra finch males give birth to offspring with shorter telomere lengths and reduced lifespans (Noguera et al., 2018) and finally in humans, "Older father's children have lower evolutionary fitness across four centuries and in four populations" (Arslan et al., 2017). Despite the absence of consensus on the underlying mechanisms (Monaghan et al., 2020), the near-ubiquity of the Lansing effect is important for our understanding of the selective forces shaping the evolution of life histories. In a recent article (Méléard et al., 2019), we introduced an asexual and haploid age-structured population model implementing a strong Lansing effect, showing that this strong transgenerational effect of ageing could be maintained. Here, by extending

this model to any system able to reproduce and maintain its homeostasis, we show that 1) ageing - i.e. non-infinite reproduction and homeostasis maintenance - is an adaptive force of evolution, 2) such a system evolves towards a configuration where fertility exceeds homeostasis capabilities, 3) propitious for pro-senescence mechanisms to appear - i.e. Lansing effect - within a few dozen generations. 4) Individuals carrying such a transgenerational effect of ageing have a non-null probability to be selected for when competing with individuals deprived of it. In addition, this model 5) allows to explain longevity-fertility tradeoffs mathematically without the need for energy investment strategies, while 6) suggesting that the selection of ageing relies on the maximization of a meta-characteristic named evolvability, in the form of fitness gradient. Finally, the landscape of this fitness gradient shows how mutation accumulation can accompany the evolution of ageing.

Results

A generalized bd model

The model and its population dynamics is similar to those described in (Méléard et al., 2019). Briefly, the model describes an asexual and haploid population structured by a life-history trait that is defined by a pair of parameters - or genes - (x_b, x_d) where x_b is defining the duration of fertility and x_d the age at which mortality becomes non-null. Here we generalized the model to any intensities of birth and death denoted (i_b, i_d) as well as to populations without Lansing effect (Figure 1, see also Annex 1). The selective pressure is enforced by a logistic competition c mimicking a maximum carrying capacity of the environment, thus no explicit adaptive value is given to any particular trait and, for each reproduction event, a mutation (h) of probability p can affect both the genes x_b and x_d, independently, following a Gaussian distribution centered on the parental trait. In Figure 1, the different cases are explored, depending on the respective values of x_b and x_d. Individuals in the Figure 1b-c configuration (for $x_b \leq x_d$) will always give progeny with a genotype (x_b, x_d) $\mp$ (h_b | h_d). The case of individuals carrying a genotype with $x_d < x_b$ (Figure 1a) is subtler, depending on the parental age a and whether the parent carries the possibility for a Lansing effect or not (Figure 1d-f). If $a < x_d$ or if the parent does not carry a Lansing effect, the genotype of the progeny will be as previously described. But if $a > x_d$, and if the parent carries the Lansing effect, then the progeny inherits a dramatically reduced x_d (here x_d is set to 0), mimicking a strong Lansing effect.

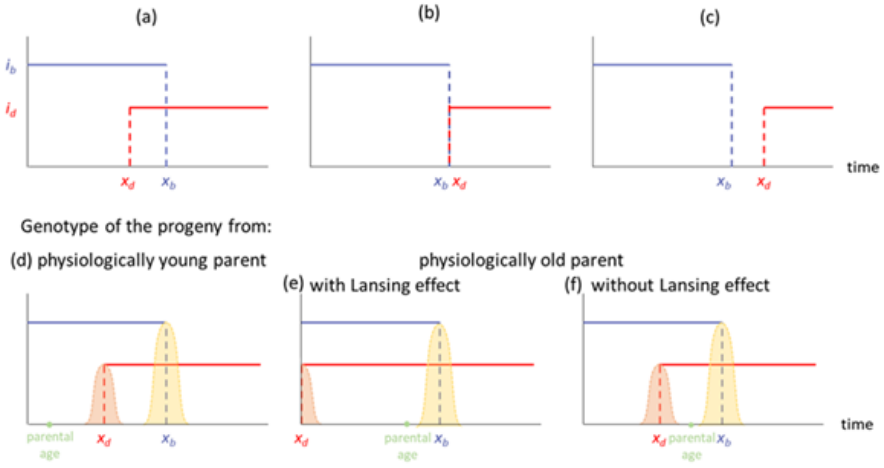


Figure 1: Three typical configurations of the model with $i_b > i_d$ and their effect on progeny's genotypes as a function of parental age. (upper panel) Each haploid individual is defined by a parameter x_b defining its fertility period of intensity i_b and a parameter x_d defining the time during which it will maintain itself, with an intensity i_d. These parameters can be positive or null. (a) 'Too young to die' : it corresponds to configurations satisfying $x_d < x_b$. (b) 'Now useless': it corresponds to configurations where $x_b = x_d$. (c) 'Menopause': it corresponds to configurations where $x_d > x_b$. (lower panel) Each individual may randomly produce a progeny during its fertility period [0; x_b]. (d) In the case of physiologically young parents ($a < x_d$), the progeny's genotype is that of its parent $\mp$ a Gaussian kernel of mutation centered on the parental gene. In the case of the reproduction event occurring after x_d, for configuration (a) above, two cases are observed, (e) if the organism carries a Lansing effect ability, the x_d of its progeny will be strongly decreased. (f) In the absence of the Lansing effect, the default rule applies.

In our previous work (Méléard et al., 2019), we formally and numerically showed the long-time evolution of the model to converge towards (x_b - x_d) = 0 in the case of individuals carrying a Lansing effect. To test whether this convergence of (x_b - x_d) still occurred without this strong transgenerational effect of ageing, we implemented a new version of the model devoid of Lansing effect and simulated its evolution for a viable - i.e. allowing the production of at least one progeny - trait (x_b = 1.2, x_d = 1.6). Surprisingly, we still observe a convergence of (x_b - x_d). The dynamics of the trait (x_b, x_d) is described by the canonical equation of adaptive dynamics depending on the Malthusian parameter and its gradient (Annexe 1). The latter can be interpreted as the age-specific strength of selection in the sense of Hamilton. We observe that the evolution speed of x_b and x_d decreases with time as the previous - less general - form of the model did (Méléard et al., 2019). This allows us to recover the classical age-related decrease in the strength of selection (Hamilton, 1966; J.b.s Haldane, 1941; Medawar, 1952).

Simulations of the generalized bd model presented here shows that the x_b - x_d distance, i.e. the time separating the end of fertility from the increasing risk of death, converges - for any initial trait - towards a positive constant. Thus, it seems that the long term evolution of such a system is a configuration similar to Figure 2a ($x_d < x_b$). The formal analysis of the generalized bd model confirmed that the long-time limit of the traits (x_b - x_d) is the positive constant defined by the formula in Figure 2b (mathematical analysis Annexe 4), reached after a few dozen simulated generations (Figure 2c).

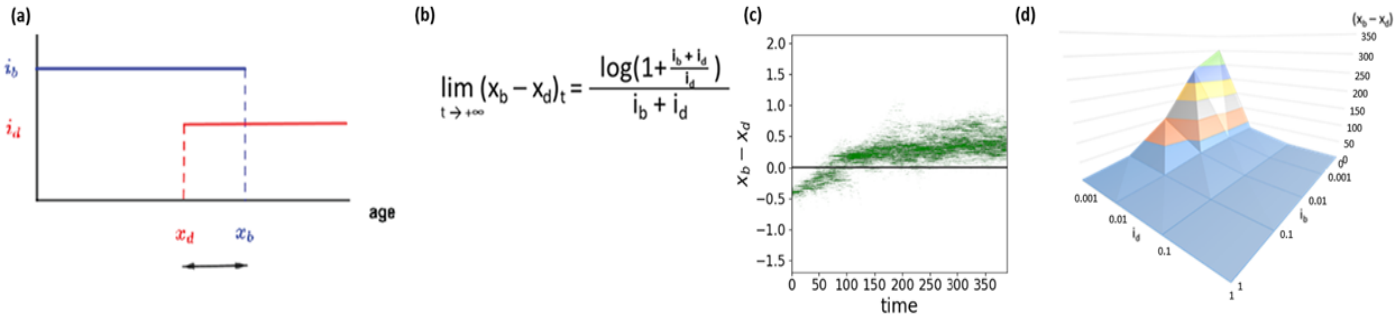


Figure 2: The bd model shows a convergence of x_b - x_d towards a positive value. Dynamics of the individual-based model shows a convergence of x_b - x_d towards a positive constant value in the absence of the Lansing effect. (a) The generalized b-d model shows a convergence of (x_b - x_d) for any ib and id towards a positive value given by **(b) (cf. Annexe 4.3, figure 2). (c)** Simulation of 1000 individuals with initial trait (x_b = 1.2, x_d = 1.6) of intensities i_b = i_d = 1, a competition c = 0.0009 and a mutation kernel (p = 0.1, σ = 0.05) show that the two parameters co-evolve and maintain x_b - x_d ≳ 0 . **(d)** Landscape of solutions (x_b - x_d) as a function of i_b and i_d (colors separate ranges of 50 units on the z-axis).

Surprisingly, the limit value of the trait is not affected by x_b or x_d values - the fertility and organismal maintenance durations per se - but only by their respective intensities i_b and i_d. These intensities can be interpreted as the instant mortality risk i_d and the probability to give a progeny i_b. Interestingly, the long-time limit values for any i_b and i_d shows a significantly stronger sensitivity to the increasing mortality risk i_d than to reproduction by almost two orders of magnitude (Figure 2d). In addition, for extremely low values of i_b and i_d - i.e. below 0.01 - the apparent time correlation of the fertility and organismal integrity maintenance period is almost nonexistent since (x_b - x_d) is large. A biological manifestation of this would be a loss of organismal maintenance occurring long before the exhaustion of reproductive capacity. Such an organism would be thus characterized as having no significant fertility decrease during the ageing process. On the other end, for individuals showing either a high instant mortality risk or a high probability to give a progeny, the (x_b - x_d) trait is close to 0, meaning that fertility and organismal integrity maintenance are visibly correlated. Note that this mathematical study concerns individuals for which the mean number of descendants per individual is large enough, allowing us to define a viability set of traits (x_b, x_d) (see Annexe 2.3). Because of these mathematical properties, a tradeoff emerges between i_b, i_d, x_b and x_d. Let's consider an organism - for both the

Lansing and non-Lansing cases - with a low reproductive intensity $i_b = 0.01$ and $i_d = 1$. For it to be viable, the product $i_b * x_b$ has to be strictly superior to 1, hence here $x_b \geqslant 100$ (see Annexe 2.3). In this example, the long-time limit of the trait ($x_b - x_d$) is equal to ln(2), thus $x_b \cong x_d$. With the same reasoning, considering an organism significantly more fertile (with $i_b = 1$, $i_d = 1$), ($x_b - x_d$) long-time evolution lower limit is $1/\sqrt{3}$. This model thus allows an elegant explanation for the apparent negative correlation previously described between longevity and fertility (see Annexe 2.3 - examples).

On the selection of Lansing effect

In our model, whatever the initial trait (x_b, x_d) in the viability set, evolution leads to a configuration of the trait such that the risk of mortality starts to increase before the fertility period is exhausted. Similar to biochemical reactions involved in a given pathway being evolutionarily optimized through tunneled reactions and gated electron transfers, we hypothesize here that such a configuration, caused by simple mathematical constraints, creates the conditions for the apparition, selection and maintenance of a molecular mechanism coupling x_b and x_d. Such a coupling mechanism could thus be the Lansing effect, the only described age-related decline in progeny's quality that seems to affect numerous iteroparous species (Lansing, 1947; Monaghan et al., 2020).

We assessed the likelihood of an organism carrying such a non-genetic pro-senescence mechanism to survive when in competition with a population devoid of such a mechanism. To do so, we considered a population divided into two sub-populations: one made of individuals subject to the Lansing effect and the other made up of individuals not subject to it. We assume, as before, that each individual is under the same competitive pressure. The two initial sub-populations have the same Darwinian fitness approximated by their Malthusian parameter (cf. Annexe 2, supplementary figure 2). Their traits are thus $(1.5; 1.3)_{Lansing}$ and $(1.5; 0.83)_{non\text{-}Lansing}$. In order to simplify the analysis, both the birth and death intensities are as follows: $i_b = i_d$ (the model is nevertheless generalized to any (i_b; i_d), see Annexe 5.1) and evolution was simulated for discrete pairs of mutation rate (**p**) and competition (**c**) parameters. Three indexes were calculated for each set of simulation: Table 1 a) the ratio of Lansing and non-Lansing populations that collapsed (a "-" indicates that all survived), Table 1 b) the ratio of total number of progenies produced during the simulation by each population and Table 1 c) the relative proportion of the Lansing population at the end of the simulation. Our 1200 simulations each of 2.10^5 birth-death events summarized in Table 1 show that the Lansing populations survive at least as well as non-Lansing ones (Table 1a) and show a significantly increased success for a moderate competition ($c = 9.10^{-4}$) and low (in our simulations) mutation rate ($p = 0.1$). With these parameters, Lansing populations show almost half the risk of disappearance of non-Lansing ones (Table 1a), producing nearly thrice as many descendants as non-Lansing populations (Table 1b) for up to a 20% faster growing population (Table 1c). Hence, although the Lansing effect leads to the production of a significant proportion of progeny with an extremely low fitness ($x_d = 0$), it decreases the risk of population collapse for organisms carrying it and seems to allow a slightly better growth of the population, whatever the magnitude of the Lansing effect (Sup. Figure 1).

Mutation probability

Competition	a) Lansing/non-Lansing collapsed population				b) Lansing/non-Lansing number of individuals				c) Lansing/total population size			
	0	0.1	0.5	1	0	0.1	0.5	1	0	0.1	0.5	1
9.10^{-5}	-	-	-	-	1.30	1.38	1.39	1.35	0.57	0.64	0.62	0.59
9.10^{-4}	1.02	0.62	0.56	0.66	1.70	2.84	3.58	3.20	0.49	0.62	0.60	0.55
9.10^{-3}	1.00	1.05	1.13	1.03	1.05	1.31	1.56	1.84	-	0.43	0.44	0.49

Table 1: Populations with Lansing effect are favorably selected under logistic competition when the mutation rate is non-zero. *p* is the mutation rate and *c* the intensity of the logistic competition. For each couple (*p*, *c*), 100 independent simulations were run with 500 individuals per population at t_0 of which traits are $(1.5; 1.3)_{Lansing}$ and $(1.5; 0.83)_{non\text{-}Lansing}$ so that their respective Malthusian parameters are equal. Each simulation corresponds to 2.10^5 events of birth or death. Table **(a)** shows the ratio of Lansing and non-Lansing populations (out of 100 simulations in each case) that did collapse by the end of the simulation. For the lowest competition, none of the populations collapsed within the timeframe of simulations (-). For an intermediate value of competition,

approximately half less Lansing population disappear relative to non-Lansing ones. Table **(b)** shows the ratio of the number of individuals generated between Lansing and non-Lansing populations**.** On average, Lansing populations generate approximately twice as many individuals as non-Lansing ones. **(c)** On average, Lansing populations grow 20% more than the non-Lansing. Values highlighted in green are discussed further below.

The Lansing effect increases fitness in ageing populations by increasing their evolvability.

In order to understand the evolutionary success of a characteristic apparently decreasing an organism's fitness, we focused our attention on their Malthusian parameter - the genotypic rate of increase (Hairston et al., 1970) - in each population through time. Here we present the results for an intermediate set of **c** and **p** - highlighted in green Table 1 - that we identified as associated with the highest success rate of Lansing bearing populations. First, we observe that, on average, Lansing populations (blue) grow while non-Lansing ones (red) have a decreasing size (Figure 3a - blue and red curves represent deciles 1, 5 and 9). Nevertheless, in the simulations where both populations coexist all along, the higher fitness of the Lansing population is marginal, with populations growing 20% more than the non-Lansing population (Figure 3b). This higher success rate seems to be carried by a faster and broader exploration of the Malthusian parameter space in the Lansing population (Figure 3c). This maximization of the Malthusian parameter is not associated with any significant difference in the lifespan (time of death - time of birth) distributions of either population (Figure 3d). Indeed, although carrying the same mutation rate **p** and being subject to the same competition **c**, the distribution of the progeny from non-Lansing populations is essentially of the parental trait in the first 5 generations, while Lansing progenies (not affected by the Lansing effect - we excluded progeny with $x_d = 0$ for the comparison) explore a broader part of the trait space (Figure 3e). This significantly higher success of the Lansing-bearing population is observed although their low fitness progeny ($x_d = 0$) represents up to 10% of the population for a significant amount of time (Figure 3f). This leads to the Lansing populations to reach the equilibrium trait faster than the non-Lansing ones (Figure 3g). Thus, the relatively higher success of Lansing bearing populations seems to be associated with a higher genotypic diversity leading to a broader range of fitness on which natural selection can apply, namely the evolvability.

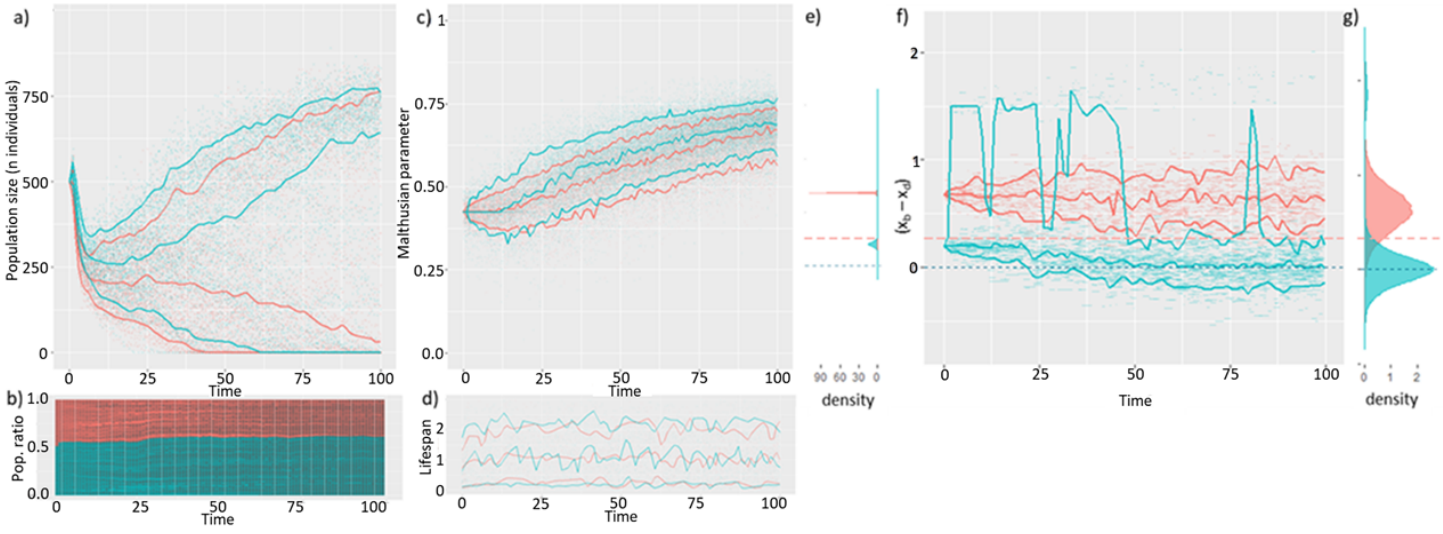


Figure 3: The Lansing effect maximizes populational survival by increasing its evolvability. 100 independent simulations were run with a competition intensity of 9.10^{-4} and a mutation rate p = 0.1 on a mixed population made of 500 non-Lansing individuals and 500 individuals subjected to such effect. At t_0, all individuals are of age 0. Here, we plotted a subset of the 100.10^6 plus individuals generated during the

simulations. Each individual is represented by a segment between its time of birth and its time of death. In each graph, blue and red curves represent deciles 1, 5 and 9 of the distribution at any time for each population type. **(a)** The higher success rate of Lansing bearing populations does not seem to be associated with a significantly faster population growth but with a lower risk of collapse. **(b)** For cohabitating populations, the Lansing bearing population (blue) is overgrowing by only 10% the non-Lansing one (red). **(c)** This higher success rate is associated with a faster and broader exploration of the Malthusian parameter - surrogate for fitness - space in Lansing bearing populations **(d)** that is not associated with significant changes in the lifespan distribution **(e)** but a faster increase in genotypic variability within the [0; 10] time interval. **(f)** This occurs although progeny from physiologically old parents can represent up to 10% of the Lansing bearing population and leads to it reaching the theoretical optimum within the timeframe of simulation **(g)** with the exception of Lansing progenies.

The relative success of Lansing-bearing populations with randomly distributed traits

We have proposed here a simple model showing the mathematical basis driving the evolutive pressure connecting organismal maintenance with reproductive mechanisms. Nevertheless, the numerical exploration of our model's behaviour has been limited so far to initial conditions where the competing populations were of equal Malthusian parameter. The low number of generations involved suggest that the conditions for the development, selection and maintenance of mechanisms of ageing (Lemoine, 2021) would have occurred early on, in a population of mixed individuals. As such, we decided to test the evolution of the trait (x_b - x_d) in Lansing and non-Lansing bearing individuals of traits uniformly distributed on [-10; +10] (Figure 4 - left panel). We chose to plot one (Figure 4 - central panel) of the hundred simulations we made, that is representative of the general results. Simulations show, over 110.10^6+ individuals, an early counter selection of extreme trait values, typically (x_b - x_d) > 4. Interestingly, the whole space of (x_b - x_d) trait is not explored evenly and the positive part of it represents approximately ⅔ of the individuals although the branched evolution process can lead a line on both the positive ('Too young to die' - Figure 1a) and negative ('Menopause' - Figure 1c) sides of the trait space. Both the Lansing and non-Lansing bearing populations manage to co-exist until the end of the simulation, each reaching a distribution centered on their respective theoretical solutions (Figure 4 - right panel), 0 for the Lansing (Méléard et al., 2019) and ln(3)/2 for the non-Lansing. In this context where the initial condition does not restrict the competition to populations of identical Malthusian parameters, the Lansing bearing population is significantly less successful than the non-Lansing one, representing only one third of the final population size. As such, the evolution of a mixed population of individuals with a trait (x_b - x_d) initially uniformly distributed on [-10; +10], bearing or not a strong inter-generational effect, will lead to a mixed solution of individuals carrying a trait that converged towards the theoretical solution such as $x_d \lesssim x_b$ thus allowing the maximization of fertility without cluttering the environment with individuals not producing descendants, very similar to Weismann's first intuition (Weismann, 1882). Nevertheless, this interpretation seems somehow finalist and does not discriminate between the relative success of each of our population types. We explore next, how ageing affects the evolvability of individuals bearing it, leading us to propose another interpretation of the gain of fitness it brings..

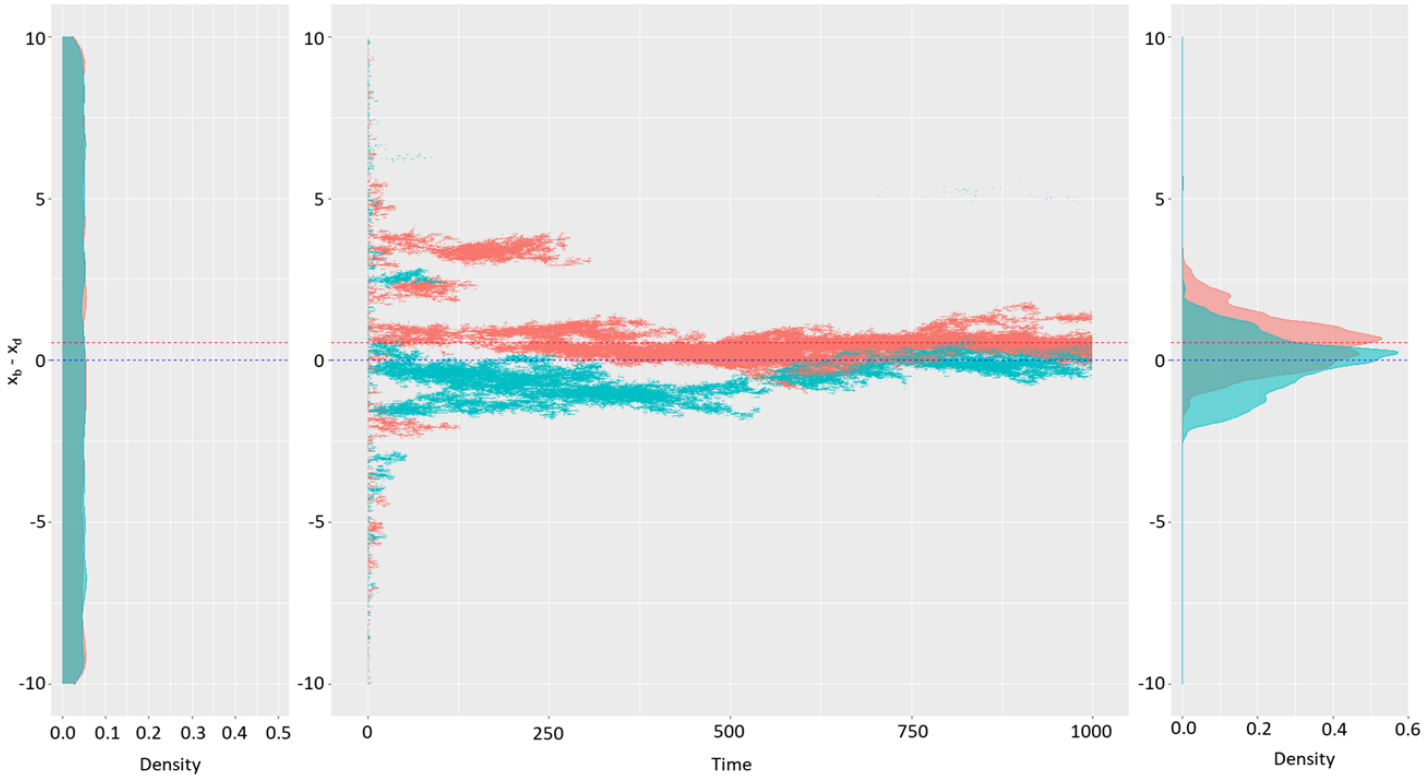


Figure 4: Mixed populations lead to (x_b - x_d) theoretical limit in a limited time and cohabitation of Lansing and non-Lansing populations. Starting with an homogenous population of 5000 Lansing bearing and 5000 non-Lansing individuals with traits uniformly distributed from -10 to +10 (left panel), we ran 100 independent simulations on time in [0; 1000]. (center panel) Plotting the trait (x_b - x_d) as a function of time for one simulation shows a rapid elimination of extreme traits and branching evolution. (right panel) The final distribution of traits in each population type is centered on the theoretical convergence limit for each. $N_{total} \cong 110.10^6$ individuals, $c = 9.10^{-4}$, $p = 0.1$

The fitness gradient, a mediator of evolvability

In order to understand the origin of this relative success of individuals carrying the ability to transmit an ageing information to the next generation, we focused our attention on the differential landscape of the Malthusian parameters as a function of the trait (x_b; x_d) for both Lansing and non-Lansing populations. We built this landscape numerically using the Newton method implemented in Annex 2. First of all, it is interesting to notice that we have derived from the equations that the maximum rate of increase for Malthusian parameters is $1/i_d$ with a maximum fitness value capped by i_b (Annex 2). Consistent with our previous characterization of the Trait Substitution Sequence in populations with Lansing effect (Méléard et al., 2019), Lansing individuals have a symmetrical fitness landscape (Figure 5, blue lines) centered on the diagonal $x_b = x_d$ (Figure 5, green diagonal). Along the latter, we can directly observe what is responsible for the "selection shadow". As (x_b; x_d) increases, a mutation of same magnitude has less and less effects on the fitness, thus allowing the accumulation of mutations (Figure 5, blue arrows). The case of non-Lansing individuals is asymmetrical, the rupture of symmetry occurring on the $x_b = x_d$ diagonal. For $x_d > x_b$ (Figure 5, upper diagonal), fitness isoclines fully overlap thus showing an equal response of both Lansing and non-Lansing fitness to mutations. In addition, as expected, the fitness of Lansing individuals is equal to that of non-Lansing ones for a given trait. On the lower part of the graph, corresponding to $x_d < x_b$, non-Lansing fitness isoclines separate from that of Lansing individuals, making the fitness of non-Lansing individuals higher to that of Lansing ones for a given trait. Nevertheless, the fitness gradient is significantly stronger for Lansing individuals as represented on Figure 5 by

the yellow arrow and associated yellow area. For an individual of trait (x_b = 2.45; x_d = 1.05) a mutation making a non-Lansing individual 0.1 in fitness (isocline 0.7 to isocline 0.8) will make a Lansing individual increase its own by 0.42 (isocline 0.1 to above isocline 0.5). With a 4-fold difference, the Lansing population produces 4 times as many individuals as the non-Lansing ones for a given mutation probability. But this reasoning can be extended to any trait (x_b; x_d) with or without Lansing effect. Organisms ageing rapidly - i.e. with low x_b and x_d - will see their fitness significantly more affected by a given mutation *h* than individuals with slower ageing affected by the same mutation. As such, ageing favors the emergence of genetic variants, it increases an organism's evolvability.

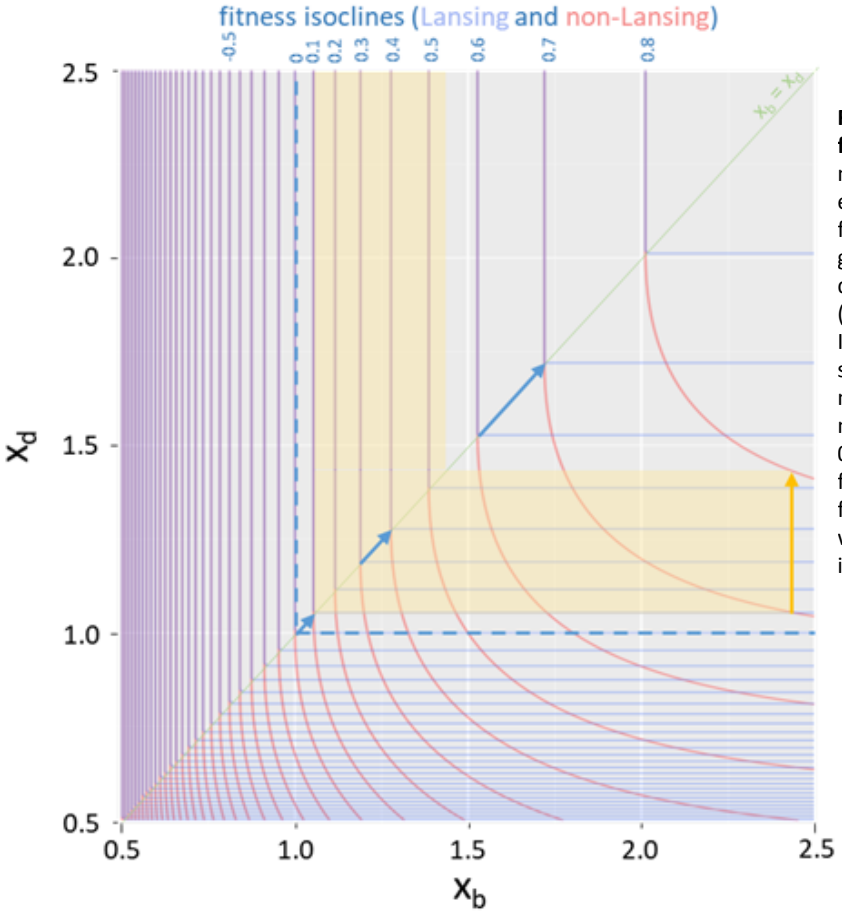


Figure 5: The Lansing effect is associated with an increased fitness gradient. We were able to derive Lansing and non-Lansing Malthusian parameters from the model's equations (see Annexe 1-2.3 and 1-5) and plot them as a function of the trait (x_b; x_d). The diagonal $x_b = x_d$ is drawn in light green. The corresponding isoclines are overlapping above the diagonal but significantly differ below, with non-Lansing fitness (red lines) being higher than that of Lansing's (light blue lines). In addition, the distance between two consecutive isoclines is significantly more important in the lower part of the graph for non-Lansing than Lansing bearing populations. As such, a mutation leading a non-Lansing individual's fitness going from 0.7 to 0.8 (yellow arrow) corresponds to a Lansing individual's fitness going from 0.1 to 0.52. Finally, Hamilton's decreasing force of selection with age can be observed along the diagonal with a growing distance between two consecutive fitness isoclines as x_b and x_d continue increasing.

Discussion

Ageing affects a broad range of organisms with various flavours (Jones et al., 2014). Yet, its characteristics show evolutionarily conserved so-called hallmarks of ageing (Lopez-Otin et al., 2013), the most evolutionarily ancient one being the loss of protein stability (Lemoine, 2021). For almost a century and a half, researchers interested in understanding the evolutive role of ageing have been strongly debating why ageing even exists. Although early theories (Weismann, 1882) proposed a population-based adaptive role for it, the 20th century seems to have seen the debate getting settled after seminal works showing the time-dependent decline of selective pressure (Hamilton, 1966; Medawar, 1952) allowed for the existence of theories (Kirkwood, 1977; Medawar, 1952; Williams, 1957) proposing that ageing is nothing more but a by-product of evolution.

The model we presented here allows us to propose an alternative theory where ageing necessarily emerges for any system showing the two minimal properties of life (Trifonov, 2011), namely a) reproduction with variation (x_b) and b)

organismal maintenance (x_d). We formally show that an haploid and asexual organism with these two properties will rapidly evolve, within a few dozen generations, towards a solution such as (x_b - x_d) is strictly positive. More importantly, the time separating both parameters is independent from their absolute values and only depends on the rate of each, respectively i_b for x_b and i_d for x_d. This property allows us to explain the observations that fed the disposable soma theory of ageing, based on an apparent trade-off existing between the fertility of an organism and its lifespan. Indeed, the lower limit condition for the viability of an individual in our model is $x_b * i_b > 1$. As such, an organism with a low fertility ($i_b << 1$) will require a long fertility time ($x_b >> 1$) to be viable. The formally shown properties of our model imply that the duration for maintenance of homeostasis (x_d) will rapidly evolve towards (x_b - x_d) equal to a positive constant, so $x_d >> 1$. On the other hand, a highly fertile organism will evolve towards its minimum viable condition requiring only a small x_d. The apparent trade-off between fertility and longevity is thus solely a consequence of $x_b * i_b > 1$ and $\lim_{+\infty}(x_b - x_d)_t$. By constraining x_b and x_d to converge, evolution also creates the conditions favoring the apparition of a phase of life in which an individual's fertility drops while its risk of dying becomes non-zero. This time-coupling of the two characteristics would thus facilitate the selection of any molecular mechanism functionally coupling the two properties (Echave, 2021) and, on the contrary to what was suggested in (Stearns, 1989), we observe that two genes with no common genetic basis can be co-selected although not linked by any direct tradeoff. By testing such a coupling mechanism, materialized as a strong Lansing effect, we showed that not only is this apparently fitness-decreasing mechanism able to, at least slightly, increase fitness, but it actually significantly increases the probability of a population bearing it to thrive when in competition with a population of equal Malthusian parameter where the effect is absent. We observed numerically that this slight increase in fitness mediated by the Lansing effect materializes itself by an increase in the genetic variability produced within the population. Hence, we propose that active mechanisms of ageing are selected for during evolution through their ability to increase an organism's evolvability. The concept of evolvability comes from the EvoDevo community. It is “an abstract, robust, dispositional property of populations, which captures the joint causal influence of their internal features upon the outcomes of evolution” (Brown, 2014). In other terms, it is “the capacity to generate heritable selectable phenotypic variation” (Kirschner and Gerhart, 1998). It is an interesting concept as it allows for a character that has no direct effect on fitness - or even a negative one (Maynard Smith, 1971) - to be under strong selection simply for its ability to birth the genetic-phenotypic variation that is the support of evolution. Furthermore, such a mechanism, triggered when age > x_d, would be of great advantage in a constantly varying environment. Indeed, when environmental conditions become less permissive, x_d might be affected and individuals pushed to enter the [x_d; x_b] space earlier, thus increasing the evolvability of the population. By constraining the long-term evolution of the trait (x_b - x_d) without any a priori on the underlying mechanism, our model predicts a high evolutionary conservation of the function “ageing” - ensuring that an individual has a limited time to propagate its genes - but not necessarily to a conservation of the underlying mechanisms. This would lead to a layering of mechanisms through evolution as described recently in (Lemoine, 2021). Hamilton’s result (Hamilton, 1966) that we previously derived from the mathematical analysis of our model (Méléard et al., 2019) confirm such evolving organisms to be sensitive to Medawar’s mutation accumulation (Medawar, 1952) and Williams’ antagonistic pleiotropy (Williams, 1957).

Although this simple model helps us to see ageing as an evolutionarily adaptive force for various ageing types, it is still a toy model. We are now developing more complex versions of it, notably to assess the interactions existing between i_b, i_d, x_b and x_d, but also to extend it to joint evolution of maturation, sex, ploidy and varying environmental conditions and ageing.

Materials and Methods

See Annex 2 for code, packages and the software used.

Authors Contribution

RT wrote the Python code presented in Annexe 2 and developed the mathematical analysis presented in Annexe 1, JP translated the Python code into C and ran the broad traits simulations. MS verified the mathematical analysis from Annexe 1, developed the analysis presented in Annexe 2 and wrote the manuscript. RM designed the study, designed the figures and wrote the manuscript.

Acknowledgments

We would like to thank Dr. Sarah Kaakai for her help in transposing our initial simulation Python codes into the IBMPopSim framework, Dr. Allon Weiner for a few hours of “naive” discussion that helped explore the interpretations of this model’s impact on our perception of ageing Dr. André Klarsfeld for his numerous useful comments on the manuscript and Dr. Bastian Greshake Tsovaras for helping with setting up the mybinder.org implementations of the simulations’ codes. This work was granted access to the GENCI-sponsored HPC resources of TGCC@CEA under allocation A0090607519. This work has been supported by the Chair “Modélisation Mathématique et Biodiversité” of Veolia Environnement-Ecole Polytechnique-Museum National d’Histoire Naturelle-Fondation X.

References

Austad SN. 2004. Rebuttal to Bredesen: 'The non-existent aging program: how does it work?' *Aging Cell* **3**:253–254. doi:10.1111/j.1474-9728.2004.00119.x

Blagosklonny MV. 2013. Aging is not programmed: Genetic pseudo-program is a shadow of developmental growth. *Cell Cycle* **12**:3736–3742. doi:10.4161/cc.27188

Bluher M, Kahn BB, Kahn CR. 2003. Extended longevity in mice lacking the insulin receptor in adipose tissue. *Science* **299**:572–4. doi:10.1126/science.1078223

Bredesen DE. 2004a. The non-existent aging program: how does it work? *Aging Cell* **3**:255–259. doi:10.1111/j.1474-9728.2004.00121.x

Bredesen DE. 2004b. Rebuttal to Austad: 'Is aging programmed? *Aging Cell* **3**:261–262. doi:10.1111/j.1474-9728.2004.00120.x

Brown RL. 2014. What Evolvability Really Is. *Br J Philos Sci* **65**:549–572. doi:10.1093/bjps/axt014

Clancy DJ, Gems D, Harshman LG, Oldham S, Stocker H, Hafen E, Leevers SJ, Partridge L. 2001. Extension of life-span by loss of CHICO, a Drosophila insulin receptor substrate protein. *Science* **292**:104–6. doi:10.1126/science.1057991

Echave J. 2021. Evolutionary coupling range varies widely among enzymes depending on selection pressure. *Biophys J* **120**:4320–4324. doi:10.1016/j.bpj.2021.08.042

Fabian D. 2011. The Evolution of Aging. *Evol Ageing*. https://www.nature.com/scitable/knowledge/library/the-evolution-of-aging-23651151/

Gavrilov LA, Gavrilova NS. 2002. Evolutionary theories of aging and longevity. *ScientificWorldJournal* **2**:339–56. doi:10.1100/tsw.2002.96

Hairston NG, Tinkle DW, Wilbur HM. 1970. Natural Selection and the Parameters of Population Growth. *J Wildl Manag* **34**:681–690. doi:10.2307/3799132

Hamilton WD. 1966. The moulding of senescence by natural selection. *J Theor Biol* **12**:12–45. doi:10.1016/0022-5193(66)90184-6

Henderson KA, Hughes AL, Gottschling DE. 2014. Mother-daughter asymmetry of pH underlies aging and rejuvenation in yeast. *eLife* **3**:e03504. doi:10.7554/eLife.03504

J.b.s Haldane. 1941. New Paths In Genetics.

Johnson AA, Shokhirev MN, Shoshitaishvili B. 2019. Revamping the evolutionary theories of aging. *Ageing Res Rev* **55**:100947. doi:10.1016/j.arr.2019.100947

Jones OR, Scheuerlein A, Salguero-Gómez R, Camarda CG, Schaible R, Casper BB, Dahlgren JP, Ehrlén J, García MB, Menges ES, Quintana-Ascencio PF, Caswell H, Baudisch A, Vaupel JW. 2014. Diversity of ageing across the tree of life. *Nature* **505**:169–173.

Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang R. 1993. A C. elegans mutant that lives twice as long as wild type. *Nature* **366**:461–4. doi:10.1038/366461a0

Kirkwood TB. 1977. Evolution of ageing. *Nature* **270**:301–4.

Kirkwood TB, Melov S. 2011. On the programmed/non-programmed nature of ageing within the life history. *Curr Biol* **21**:R701-7. doi:10.1016/j.cub.2011.07.020

Kirkwood TBL, Holliday R. 1979. The evolution of ageing and longevity. *Proc R Soc Lond B* **205**:531–546. doi:10.1098/rspb.1979.0083

Kirschner M, Gerhart J. 1998. Evolvability. *Proc Natl Acad Sci U S A* **95**:8420–8427.

Kowald A, Kirkwood TBL. 2016. Can aging be programmed? A critical literature review. *Aging Cell* **15**:986–998. doi:10.1111/acel.12510

Lai C-Y, Jaruga E, Borghouts C, Jazwinski SM. 2002. A mutation in the ATP2 gene abrogates the age asymmetry between mother and daughter cells of the yeast Saccharomyces cerevisiae. *Genetics*

162:73–87. doi:10.1093/genetics/162.1.73

Lansing AI. 1954. A Nongenic Factor in the Longevity of Rotifers. *Ann N Y Acad Sci* **57**:455–464. doi:10.1111/j.1749-6632.1954.tb36418.x

Lansing AI. 1947. A Transmissible, Cumulative, and Reversible Factor in Aging. *J Gerontol* **2**:228–239. doi:10.1093/geronj/2.3.228

Lemoine M. 2021. The Evolution of the Hallmarks of Aging. *Front Genet* **0**. doi:10.3389/fgene.2021.693071

Longo VD, Mitteldorf J, Skulachev VP. 2005. Programmed and altruistic ageing. *Nat Rev Genet* **6**:866–872. doi:10.1038/nrg1706

Lopez-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. *Cell* **153**:1194–217. doi:10.1016/j.cell.2013.05.039

Maynard Smith J. 1971. What use is sex? *J Theor Biol* **30**:319–335. doi:10.1016/0022-5193(71)90058-0

Medawar PB. 1952. An unsolved problem of biology. *Med J Aust* **1**:854–5.

Méléard S, Rera M, Roget T. 2019. A birth–death model of ageing: from individual-based dynamics to evolutive differential inclusions. *J Math Biol* **79**:901–939. doi:10.1007/s00285-019-01382-z

Monaghan P, Maklakov AA, Metcalfe NB. 2020. Intergenerational Transfer of Ageing: Parental Age and Offspring Lifespan. *Trends Ecol Evol* **35**:927–937. doi:10.1016/j.tree.2020.07.005

Nussey DH, Froy H, Lemaitre JF, Gaillard JM, Austad SN. 2013. Senescence in natural populations of animals: widespread evidence and its implications for bio-gerontology. *Ageing Res Rev* **12**:214–25. doi:10.1016/j.arr.2012.07.004

Nyström T. 2007. A Bacterial Kind of Aging. *PLOS Genet* **3**:e224. doi:10.1371/journal.pgen.0030224

Partridge L, Gems D. 2002. Mechanisms of ageing: public or private? *Nat Rev Genet* **3**:165–75. doi:10.1038/nrg753

Priest NK, Mackowiak B, Promislow DEL. 2002. The role of parental age effects on the evolution of aging. *Evol Int J Org Evol* **56**:927–935. doi:10.1111/j.0014-3820.2002.tb01405.x

Sinclair DA, Guarente L. 1997. Extrachromosomal rDNA Circles— A Cause of Aging in Yeast. *Cell* **91**:1033–1042. doi:10.1016/S0092-8674(00)80493-6

Skulachev VP. 2011. Aging as a particular case of phenoptosis, the programmed death of an organism (A response to Kirkwood and Melov “On the programmed/non-programmed nature of ageing within the life history”). *Aging* **3**:1120–1123. doi:10.18632/aging.100403

Smith JM. 1976. What Determines the Rate of Evolution? *Am Nat* **110**:331–338.

stearns SC. 1989. Trade-offs in life-history evolution. *Funct Ecol* **3**:259–268.

Steiner UK. 2021. Senescence in Bacteria and Its Underlying Mechanisms. *Front Cell Dev Biol* **9**:1461. doi:10.3389/fcell.2021.668915

Trifonov EN. 2011. Vocabulary of Definitions of Life Suggests a Definition. *J Biomol Struct Dyn* **29**:259–266. doi:10.1080/073911011010524992

van Heemst D, Beekman M, Mooijaart SP, Heijmans BT, Brandt BW, Zwaan BJ, Slagboom PE, Westendorp RG. 2005. Reduced insulin/IGF-1 signalling and human longevity. *Aging Cell* **4**:79–85. doi:10.1111/j.1474-9728.2005.00148.x

Weismann A. 1882. Ueber die Dauer des Lebens; ein Vortrag. Jena: G. Fischer.

Williams GC. 1957. Pleiotropy, Natural Selection, and the Evolution of Senescence. *Evolution* **11**:398–411. doi:10.2307/2406060

Annexe 1: mathematical proofs

Supplementary - Positive selection of senescence through increased evolvability: ageing is not a by-product of evolution.

Roget T., Jolivet P., Méléard S. and Rera M.

March 7, 2022

1 The mathematical individual-based bd model

We model an haploid and asexual population of individuals with evolving life-histories by a stochastic individual-based model, similar to the one introduced in [8] and a particular case of [3]. Each individual is characterized by its age and by a life-history trait $x = (x_b, x_d) \in \mathbb{R}_+^2$ that describes for each individual the age x_b at the end of reproduction and the age x_d when mortality becomes positive. The trait can change through time, by mutations occuring continuously in time.

More precisely, the Markovian dynamics of the population process is defined as follows. The individuals reproduce and die independently. An individual with trait (x_b, x_d) reproduces at rate i_b as long as it is younger than x_b. Further, he cannot die as long as it is younger than x_d and has a natural death rate i_d after age x_d.

The life-history of an individual with trait $x = (x_b, x_d)$ is described by the couple of birth and death functions (B_x, D_x^c) defined on $\mathbb{R}_+$ by

$$\forall a \in \mathbb{R}_+, \qquad B_x(a) = i_b \mathbb{1}_{a \leqslant x_b}, \quad D_x^c(a) = i_d \mathbb{1}_{a > x_d} + cN. \tag{1}$$

Here, the individual age a is the physical age, N the (varying) population size and $c > 0$ the competition pressure exerted by an individual on another one. The death rate will be extended to

$$D_x^c(0) = +\infty, \text{ for } x_b < 0 \text{ or } x_d > 0,$$

meaning that an individual appearing by mutation will be able to survive only if the two components if its trait are non negative.

Note that the date of birth and lifespan of an individual are stochastic and the law of the lifespan on an individual with trait x born at time τ is given by $f_x(s) = D_x^c(s) \exp(-\int_\tau^{\tau+s} D_x^c(a)da)$.

We also take into account genetic mutations which create phenotypic variation, and which added to competition between individuals, will lead to natural selection.

At each reproduction event, a mutation appears instantaneously on each trait x_b and x_d independently with probability $p \in]0, 1[$. If the trait x_b mutates (resp. if x_d mutates),

the trait of the newborn is $x_b + h_b$ (resp. $x_d + h_d$). The mutation effect h_b (resp. h_d) is distributed following a centered Gaussian law with variance σ^2. This Gaussian law is denoted by $k(h)dh$.

Note that a similar model has been defined in [8], including a Lansing effect on the reproductive lineage of "old" individuals.

2 The Malthusian parameter

2.1 The demographic parameters

We now introduce the classical demographic parameters for age-structured (without competition) population, where all individuals have the same trait $x \in \mathbb{R}_+^2$ (cf. [1]). We are looking for a triplet $(\lambda(x), N_x, \phi_x)$ where $\lambda(x) \in \mathbb{R}$ is the Malthusian parameter, $N_x(a), a \in \mathbb{R}_+$ the stable age distribution and $\phi_x(a), a \in \mathbb{R}_+$ the reproductive value. They describe the asymptotic growth of the population dynamics and measure the fitness of life-histories: $\lambda(x)$ is the growth rate of the population at its demographic equilibrium, N_x the age distribution of the population and $\phi_x(a)$ is the probability that an individual with trait x has a newborn after age a. It is known (cf. [1]), that $(\lambda(x), N_x, \phi_x)$ is solution of the direct and dual eigenvalue problems:

$$\begin{cases} -\partial_a N_x(a) - D_x(a)N_x(a) = \lambda(x)N_x(a) \\ N_x(0) = \int_0^{+\infty} B_x(\alpha)N_x(\alpha)d\alpha, \quad N_x(0) = 1, \end{cases} \tag{2}$$

$$\begin{cases} \partial_a \phi_x(a) - D_x(a)\phi_x(a) + B_x(a)\phi_x(0) = \lambda(x)\phi_x(a) \\ \phi_x(0) = 1, \end{cases} \tag{3}$$

where $B_x(a) = i_b \mathbb{1}_{\{a \leqslant x_b\}}$ and $D_x(a) = i_d \mathbb{1}_{\{a > x_d\}}$.

Proposition 2.1. *For all $x \in \mathbb{R}_+^2$, there exists a unique solution $(\lambda(x), N_x, \phi_x) \in \mathbb{R} \times L^1(\mathbb{R}_+) \times L^\infty(\mathbb{R}_+)$ of (2) and (3). The Malthusian parameter $\lambda(x)$ is the unique solution of the equation:*

$$i_b \int_0^{x_b} e^{-i_d(a-x_d)_+ - \lambda(x)a} da = 1. \tag{4}$$

The stable age distribution N_x and the reproductive value ϕ_x verify

$$N_x(a) = e^{-i_d(a-x_d)_+ - \lambda(x)a}, \quad \phi_x(a) = \frac{i_b \mathbb{1}_{a \leqslant x_b}}{N_x(a)} \int_a^{x_b} N_x(\alpha)d\alpha. \tag{5}$$

Proof. The proof is straightforward by solving the first equations in (2) and (3), and then using the equations satisfied by the boundary conditions. □

Remark 2.2. *The quantities $\lambda(x), N_x, \phi_x$ are the eigenelements (cf. Proposition 2.1) associated with the linear operator that generates the dynamics $v_x(t,a)$ of a non density dependent population with age structure and birth-death rates given by (B_x, D_x). More precisely, $v_x(t,a)$ satisfies the McKendrick Von-Foerster Equation*

$$\begin{cases} \partial_t v_x(t,a) + \partial_a v_x(t,a) = -D_x(a)v_x(t,a), \qquad t \geqslant 0, a \geqslant 0 \\ v_x(t,0) = \int_{\mathbb{R}_+} B_x(\alpha)v_x(t,\alpha). \end{cases}$$

The use of these quantities as an indicator of fitness is justified by the convergence of $e^{-\lambda(x)t}\, v_x(t,a)$ to $(\int_{\mathbb{R}_+} v_x(0,\alpha)\phi_x(\alpha)d\alpha)N_x(a)$ as t tends to infinity (cf. [10] for example).

2.2 Computation and regularity of the Malthusian parameter

The Malthusian parameter $\lambda(x)$ is defined as the unique real number such that

$$i_b \int_0^{x_b} e^{-i_d(a-x_d)_+ - \lambda(x)a} da = 1.$$

Let us introduce

$$U_1 = \{x \in \mathbb{R}_+^2 : x_b < x_d\} \ , \ U_2 = \{x \in \mathbb{R}_+^2 : x_d < x_b\} \ , \ \mathcal{H} = \{x \in \mathbb{R}_+^2 : x_b = x_d\}. \quad (6)$$

For all $x \in U_1 \cup \mathcal{H}$, the Malthusian parameter $\lambda(x)$ satisfies:

$$i_b \int_0^{x_b} e^{-\lambda(x)a} da = 1 = \frac{i_b}{\lambda(x)}(1 - e^{-x_b\lambda(x)}).$$

Then $\lambda(x)$ can be numerically computed by Newton's method applied to the function $K_{x_b}(\lambda) = \frac{1}{\lambda}(1 - e^{-x_b\lambda}) - \frac{1}{i_b}$, since $\lambda(x)$ is solution of $K_{x_b}(\lambda) = 0$, .

In the case where $x \in U_2$, we have

$$i_b \int_0^{x_b} e^{-i_d(a-x_d)_+ - \lambda(x)a} da = i_b \int_0^{x_d} e^{-\lambda(x)a} da + i_b \int_{x_d}^{x_b} e^{-i_d(a-x_d) - \lambda(x)a} da$$

$$= i_b \left\{ \frac{1}{\lambda(x)}(1 - e^{-\lambda(x)x_d}) + \frac{e^{i_d x_d}}{\lambda(x) + i_d}\left(e^{-(\lambda(x)+i_d)x_d} - e^{-(\lambda(x)+i_d)x_b}\right) \right\},$$

which has to be equal to 1. That involves a function

$$H_{(x_b,x_d)}(\lambda) = \frac{1}{\lambda}(1 - e^{-\lambda_d}) + \frac{e^{i_d x_d}}{\lambda + i_d}\left(e^{-(\lambda+i_d)x_d} - e^{-(\lambda+i_d)x_b}\right) - \frac{1}{i_b}.$$

Newton's method still allows to resolve numerically the equation and find $\lambda(x)$.

Let us now prove some regularity properties of the Malthusian parameter. We show that its gradient is a simple function of the stable age distribution, the reproductive value and the mean generation time G defined for all x by

$$G(x) = i_b \int_0^{x_b} a N_x(a) da.$$

Proposition 2.3. *The function $x \in (\mathbb{R}_+^*)^2 \mapsto \lambda(x)$ is of class $\mathcal{C}^1$ and we have:*

$$\forall x \in (\mathbb{R}_+^*)^2, \qquad \nabla\lambda(x) = \frac{1}{G(x)}\left(i_b N_x(x_b), i_d N_x(x_d)\phi_x(x_d)\right).$$

Note that the derivatives are positive, meaning that $x_b \to \lambda(x)$ and $x_d \to \lambda(x)$ are non decreasing.

Proof. Coming back to the definition of λ and using the implicit function theorem, we obtain that λ is differentiable and

$$\begin{aligned}
\forall x \in U_1, \quad \frac{\partial\lambda(x)}{\partial x_b} &= \frac{i_b e^{-\lambda(x)x_b}}{G(x)} = \frac{i_b N_x(x_b)}{G(x)} \quad ; \quad \frac{\partial\lambda(x)}{\partial x_d} = 0 = \frac{N_x(x_d)\phi_x(x_d)}{G(x)}; \\
\forall x \in U_2, \quad \frac{\partial\lambda(x)}{\partial x_b} &= \frac{i_b e^{-i_d(x_b-x_d)} e^{-\lambda(x)x_b}}{G(x)} = \frac{i_b N_x(x_b)}{G(x)} \\
\frac{\partial\lambda(x)}{\partial x_d} &= \frac{i_b i_d e^{i_d x_d} \int_{x_d}^{x_b} e^{-(i_d+\lambda(x))a} da}{G(x)} = \frac{i_b i_d N_x(x_d)\phi_x(x_d)}{G(x)}.
\end{aligned} \quad (7)$$

We deduce that λ has continuous partial derivatives, which concludes the proof. □

2.3 Viability set

The viability set is the set $\mathcal{V} \subset \mathbb{R}_+^2$ of traits $x = (x_b, x_d)$ such that $\lambda(x) > 0$. From Equation (4), $\lambda(x) > 0$ if and only if the mean number $R(x_b, x_d)$ of descendants per individual is larger than one, i.e if and only if we have:

$$R(x_b, x_d) := i_b \int_0^{x_b} e^{-i_d(a-x_d)_+} da > 1. \tag{8}$$

A precise characterization of the set $\mathcal{V}$ is given in Lemma 2.4. In Figure 1, we represent the set $\mathcal{V}$ for $i_b = 1.5$ and $i_d = 2$.

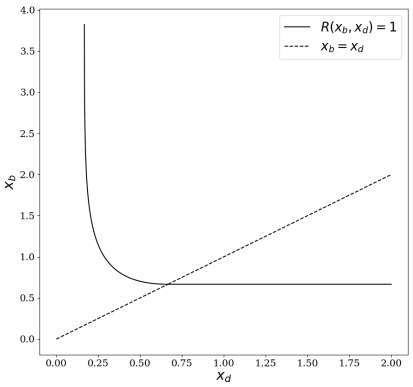


Figure 1: ***The set*** $\mathcal{V} = \{(x_b, x_d) \in \mathbb{R}_+^2; R(x_b, x_d) > 1\}$ ***is the convex set delimited by the black curve with equation*** $R(x_b, x_d) = 1$***.***

Lemma 2.4. [1] *We have:*

$$\begin{aligned} \mathcal{V} &= \{x \in \mathbb{R}_+^2 \ ; \ x_b > x_d - \log(i_d x_d + 1 - (i_b/i_d)) \text{ if } x_b > x_d \text{ and } i_b\, x_b > 1 \text{ if } x_b \leqslant x_d\} \\ &\subset \{(x_b, x_d); i_b\, x_b > 1\}, \end{aligned}$$

and for all $x \in \mathcal{V}$, $\lambda(x) \leqslant i_b$. *Moreover, the map* $x \in \mathcal{V} \mapsto \nabla\lambda(x)$ *is Lipschitz continuous.*

Proof. Let us first note that for any $x \in \mathbb{R}_+^2$, $R(x) \leqslant i_b\, x_b$. We are looking for which $x = (x_b, x_d) \in \mathbb{R}_+^2$, the mean number of descendants $R(x_b, x_d)$ is greater than 1. Recall that $R(x) = i_b \int_0^{x_b} \exp\left(-i_d \int_0^a \mathbb{1}_{\alpha > x_d} d\alpha\right) da$. For $x \in U_1$ (defined in (6)), we have $R(x) = i_b x_b$ and $R(x) > 1$ if and only if $i_b x_b > 1$. For $x \in U_2$, we have $R(x) = i_b x_d + (i_b/i_d) - (i_b/i_d) e^{-i_d(x_b - x_d)}$ and $R(x) > 1$ if and only if $x_b > x_d - \log(i_d x_d + 1 - (i_b/i_d))$. We conclude for the first assertion arguing that the map $\lambda \mapsto i_b \int_0^{x_b} \exp\left(-i_d \int_0^a \mathbb{1}_{\alpha > x_d} i_d \alpha - \lambda a\right) da$ is decreasing. Let us now show that $\lambda(x)$ is upper-bounded by i_b. Assume that there exists $x \in \mathcal{V}$ such that $\lambda(x) > i_b$. Then

$$1 = i_b \int_0^{x_b} e^{-i_d(a-x_d)_+ - \lambda(x)a} da < i_b \int_0^{x_b} e^{-i_b a} da = 1 - e^{-x_b},$$

which is absurd and allows us to conclude. The next claim is shown arguing that the map $x \in \mathcal{V} \mapsto \nabla\lambda(x)$ is differentiable on $U_1 \cup U_2$ and admits bounded partial derivatives. □

[1] Note that notation log will always mean Neperian logarithm

Let us develop different examples:

In the case where $i_b = 0.01$ and $i_d = 1$, we obtain

$$R(x_b, x_d) > 1 \Longleftrightarrow x_d + \left(1 - \left(2.01\right)^{-\frac{1}{1.01}}\right) > 100,$$

which gives essentially that x_d has to be greater than 100.

In the case where $i_b = i_d$, the formula is simpler. We obtain

$$R(x_b, x_d) > 1 \Longleftrightarrow x_d + \frac{1}{i_b}\left(1 - 3^{-\frac{1}{2}}\right) > \frac{1}{i_b}.$$

We deduce

$$R(x_b, x_d) > 1 \Longleftrightarrow x_d > \frac{1}{i_b\sqrt{3}}.$$

If we assume that $i_b = i_d = 1$ then we obtain that

$$R(x_b, x_d) > 1 \Longleftrightarrow x_d > \frac{1}{\sqrt{3}} = 0.577.$$

Let us finally note that if we assume to be in the limit of the canonical equation and then to be in the case when $x_b - x_d = \frac{\log 3}{2i_b}$ (cf. Theorem 4.4), we also obtain a characterization of the viability set using x_b:

$$R(x_b, x_d) > 1 \Longleftrightarrow x_b > \frac{1}{i_b}\left(\frac{1}{\sqrt{3}} + \frac{\log 3}{2}\right).$$

For $i_b = 1$, that gives $x_b > 1.126$

3 Monomorphic equilibrium

Let us come back to the general case with competition, but for a monomorphic population with trait x (and then without mutation). It can be proved (cf. [7] Proposition 2.4) that for a large population, the stochastic process converges in probability to the solution of the following Gurtin-MacCamy partial differential equation (see [4]).

$$\begin{cases} \partial_t n_x(t,a) + \partial_a n_x(t,a) = -\left(D_x(a) + c\int_{\mathbb{R}_+} n_x(t,\alpha)d\alpha\right) n_x(t,a) \\ n_x(t,0) = \int_0^{+\infty} B_x(\alpha) n_x(t,\alpha)d\alpha, \quad (t,a) \in \mathbb{R}_+^2 . \end{cases} \tag{9}$$

This equation describes the density-dependent dynamics of a large population with trait x (without mutation). The trait $x \in \mathbb{R}_+^2$ being given, let us study the positive equilibria of the equation

For $x \in \mathcal{V}$, Equation (9) admits a unique non-trivial solution:

Proposition 3.1. *For all $x \in \mathcal{V}$, there exists a unique globally stable equilibrium $\overline{n}_x$ to Equation (9), i.e a solution of*

$$\begin{cases} -\partial_a \overline{n}_x(a) - \left(D_x(a) + c\int_{\mathbb{R}_+} \overline{n}_x(\alpha)d\alpha\right) \overline{n}_x(a) = 0 \\ \overline{n}_x(0) = \int_0^{+\infty} B_x(\alpha)\overline{n}_x(\alpha)d\alpha, \end{cases} \tag{10}$$

which satisfies $\lambda(x) = c\int_{\mathbb{R}_+} \overline{n}_x(\alpha)d\alpha$.

Note that

$$\overline{n}_x(0) = \frac{\lambda(x)}{c\int_0^{+\infty} N_x(\alpha)d\alpha}.$$

Proof. The existence part of the proof is trivial from (2) and Proposition 2.1 using that $\mathcal{V} = \{x \in \mathbb{R}_+^2 : \lambda(x) > 0\}$. The long-time behavior of the solutions of (9) is studied in [11, Section 5.4]. □

4 Canonical equation

4.1 Invasion fitness

We now compute the invasion fitness function associated with the individual-based model. We use the definition of invasion fitness given in [7]. The invasion fitness $1 - z(y, x)$ of a mutant with trait y in a resident population with trait x is defined as the survival probability of an age-structured branching process with birth rates $B_x(a)$ and death rates $D_x(a) + c\int_{\mathbb{R}_+} \overline{n}_x(a)da$.

Proposition 4.1. *Let* $y \in (\mathbb{R}_+^*)^2$ *and* $x \in \mathcal{V}$, *we have*

$$1 - z(y, x) = \Big[\frac{\lambda(y) - \lambda(x)}{i_b}\Big]_+.$$

Proof. The proof is a direct application of Equation (3.6) in [7]. □

4.2 Trait Substitution Sequence and canonical equation

For this part, we refer principally to [7] where the theory of adaptive dynamics is rigorously developed for general age-structured populations.

We introduce the canonical equation describing the evolution of the trait $x = (x_b, x_d)$ at a mutation time-scale, under the assumptions of adaptive dynamics (large population, rare and small mutation, invasion and fixation principle, as well known since Metz et al. [9], Dieckman-Law [2]). In [7], it is shown that this equation can be obtained as a two-step limit from the individual based model. The first step consists in defining the Trait Substitution process describing the successive advantageous mutant invasions in monomorphic populations at equilibrium. It is obtained as support dynamics of the measure-valued limit of the rescaled population process (at the mutation time-scale), when mutations are rare (but not small). The measure-valued limiting process is rigorously derived from the individual-based model in [7] Section 3. It jumps from a state $\delta_x(dz)\overline{n}_x(a)da$ to a state $\delta_y(dz)\overline{n}_y(a)da$. The trait support process takes values in $\mathcal{V}$ and its dynamics is described as follows.

Definition 4.2. *The Trait Substitution Sequence is the càdlàg process* $(X_t, t \geqslant 0)$ *with values in* $\mathcal{V}$ *whose law is characterized by the infinitesimal generator* L *defined for all bounded and measurable function* $\varphi : \mathcal{V} \to \mathbb{R}$ *by:*

$$L\varphi(x) = \int_{\mathbb{R}^2} (\varphi(x + (h_1, h_2)) - \varphi(x))\Big[\frac{\lambda(x + (h_1, h_2)) - \lambda(x)}{i_b}\Big]_+ \frac{\lambda(x)}{c\int_0^{+\infty} N_x(a)da}\,\mu(dh_1, dh_2),$$

where $\mu(dh) = \frac{\delta_0(dh_2)k(h_1)dh_1 + \delta_0(dh_1)k(h_2)dh_2}{2}$ *and the distribution* k *has been defined in Section 1.*

Note that since by Proposition 2.3, the partial derivatives of λ are positive, then the increment $\lambda(x + (h_1, h_2)) - \lambda(x)$ is non negative if and only if h_1 and h_2 are non negative.

The second step consists in assuming that mutation amplitudes are small and of order ϵ, for $\epsilon > 0$. We then define the rescaled process X^ϵ by $X^\epsilon(t) = \epsilon X(\frac{t}{\epsilon^2})$. Making ε tend to 0 leads to the canonical equation.

Proposition 4.3. *Let* $T > 0$. *Assume that* $X^\epsilon(0)$ *converges to* $x^0 \in \mathcal{V}$ *in probability. Then the sequence of processes* $(X^\epsilon)_\epsilon$ *converges in law in the Skorohod space* $\mathbb{D}([0, T], \mathcal{V})$ *to the solution* $(x(t), t \geqslant 0)$ *of the ordinary differential equation:*

$$\frac{dx}{dt} = \frac{\lambda(x)}{4c\int_0^{+\infty} N_x(\alpha)d\alpha} \frac{\nabla\lambda(x)\sigma^2}{i_b}, \quad x \in \mathcal{V} \subset \mathbb{R}_+^2 \tag{11}$$

Recall that the Malthusian parameter $\lambda(x)$ is defined in (4), the stable age distribution N_x is defined in (5) and $\sigma^2(x)$ denotes the variance of the mutation kernel. Recall that (see Proposition 2.3)

$$\nabla\lambda(x) = \frac{1}{G(x)} \left(i_b N_x(x_b)), i_d N_x(x_d)\phi_x(x_d))\right). \tag{12}$$

It describes the strength of selection at ages x_b and x_d. Hence, this canonical equation allows to interpret the age specific strength of selection at ages x_b and x_d as the evolution speed of the traits x_b and x_d respectively, under the assumptions of adaptive dynamics.

Proof. The proof is classical and can be easily adapted from that of [7, Theorem 4.1]. The canonical equation only charges the set $\mathcal{V} \subset \mathbb{R}_+^2$ (defined in Section 2.1) and writes as follows:

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\nabla_y z(x, x)\frac{\overline{n}_x(0)}{2} \int_{\mathbb{R}_+} h^2 k(h)dh, \quad x \in \mathcal{V}. \tag{13}$$

The set $\mathcal{V}$ is the set of traits that admit a positive stable monomorphic equilibrium $\overline{n}_x$ in a such way that $\overline{n}_x(0)$ equals the birth rate of a mutant (see Proposition 3.1); σ^2 is the variance of the mutations and $1 - z(y, x)$ is the invasion fitness. Computing these parameters gives (11). □

In Figure 2, we present a simulation of a solution of (11). We observe that the traits x_b and x_d increase with time (cf. Figure 2 (a),(b)), with decreasing speed tending to zero. The trait $x_b(t) - x_d(t)$ converges to some positive number (cf. Figure 2 (c)) that we can rigorously compute. That is the aim of the next section.

4.3 Long-time behaviour of the canonical equation

In this section we study the long-time behaviour of the solutions of the canonical equation (11). We prove the following theorem.

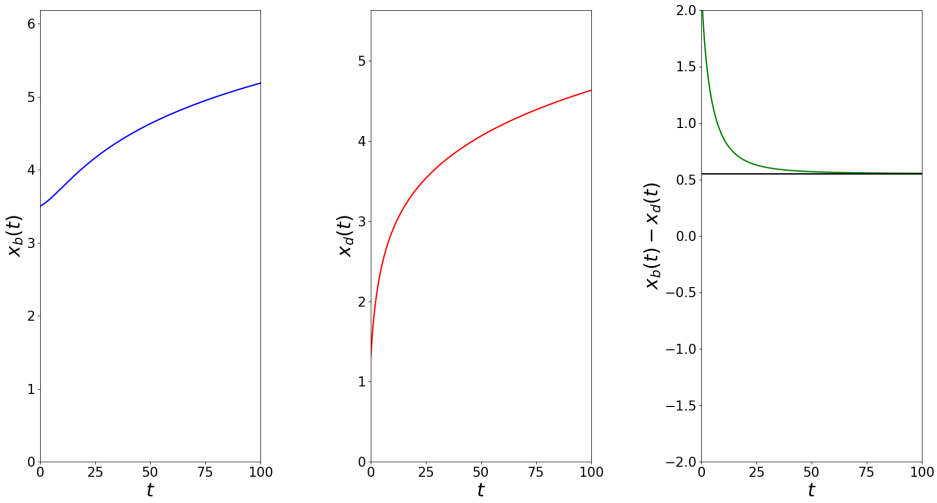


Figure 2: ***Simulation of the canonical equation with*** $x^0 = (3.5, 1.3)$ ***and*** $i_b = i_d = 1$**.** *(a): Dynamics of* x_b*. (b): Dynamics of* x_d*. (c): Dynamics of* $x_b - x_d$*, the black curve has equation* $y = \log(3)/2$*.*

Theorem 4.4. *Let* $x^0 \in \mathcal{V}$ *and let* $(x(t), t \geqslant 0)$ *be the solution of (11) started at* $x^0 \in \mathcal{V}$*. Then we have:*

$$x_b(t) - x_d(t) \underset{t\to+\infty}{\longrightarrow} \frac{\log(1 + \frac{i_b+i_d}{i_d})}{i_b + i_d}.$$

We first prove the following lemma. We always denote $U_1 = \{x \in \mathcal{V} : x_b < x_d\}$, $U_2 = \{x \in \mathcal{V} : x_d < x_b\}$ and $\mathcal{H} = \{x \in \mathcal{V} : x_d = x_b\}$.

Lemma 4.5. *Let* $x^0 \in \mathcal{V}$ *and let* $(x(t), t \geqslant 0)$ *be the solution of (11) started at* $x^0 \in \mathcal{V}$*. Then we have:*

(i) There exists $T > 0$*, such that for all* $t \geqslant T$*,* $x(t) \in U_2 \cup \mathcal{H}$*.*

(ii) There exists $C > 0$ *such that for all* $t \geqslant 0$*,* $|x_b(t) - x_d(t)| < C$*,*

(iii) We have $x_b(t)$ *increases to* $+\infty$*,* $x_d(t)$ *increases to* $+\infty$ *and* $\lambda(x(t)) \to i_b$ *as* $t \to +\infty$*.*

Proof. For all $x \in \mathcal{V}$, let us define:

$$v(x) = \frac{\lambda(x) \int_0^{+\infty} h^2 k(h) dh}{2 i_b c G(x) \int_0^{+\infty} N_x(\alpha) d\alpha}.$$

and we remark that there exist $\underline{v}(x^0), \overline{v}(x^0) > 0$ such that $\underline{v}(x^0) \leqslant v(x) \leqslant \overline{v}(x^0)$.
(i): Let $T := \inf\{t \geqslant 0 : x(t) \in U_2 \cup \mathcal{H}\} \in [0, +\infty]$. We first show that $T < +\infty$. If $x^0 \in U_2 \cup \mathcal{H}$, it is obvious. If $x^0 \in U_1$, assume that $T = +\infty$. Then for all $t \geqslant 0$, $x_d(t) = x_d^0$. Indeed, as soon as $x_b < x_d$, $\phi(x_d) = 0$ and the trait x_d does not move (see (7)). We obtain that

$$\forall t \geqslant 0, \quad \frac{\mathrm{d}(x_b(t) - x_d(t))}{\mathrm{d}t} = \frac{\mathrm{d}x_b(t)}{\mathrm{d}t} \geqslant v_1(x) i_b e^{-\lambda(x) x_b} \geqslant \underline{v}(x^0) i_b e^{-\lambda(x) x_b} > 0,$$

that allows us to obtain the contradiction. So we have $T < +\infty$. We conclude the proof arguing that for all $t \geqslant 0$ such that $x(t) \in \mathcal{H}$, $\mathrm{d}x_d(t)/\mathrm{d}t = 0$ and $\mathrm{d}x_b(t)/\mathrm{d}t > 0$.
(ii): By (i), we assume without loss of generality that $(x(t), t \geqslant 0) \subset U_2 \cup \mathcal{H}$. By Equation (11), we obtain that:

$$\begin{aligned}
\frac{\mathrm{d}(x_b(t) - x_d(t))}{\mathrm{d}t} &= i_b v(x(t)) \left(e^{-i_d(x_b(t)-x_d(t))-\lambda(x)(t)x_b(t)} - i_d \int_{x_d(t)}^{x_b(t)} e^{-i_d(a-x_d(t))-\lambda(x(t))a} da \right) \\
&= i_b v(x(t)) \Bigg(e^{-i_d(x_b(t)-x_d(t))-\lambda(x)(t)x_b(t)} \\
&\qquad - \frac{i_d e^{i_d x_d(t)}}{i_d + \lambda(x(t))} (e^{-(i_d+\lambda(x(t)))x_d(t)} - e^{-(i_d+\lambda(x(t)))x_b(t)}) \Bigg) \\
&= \frac{i_b v(x(t))}{i_d + \lambda(x(t))} \left((2i_d + \lambda(x(t))) e^{-i_d(x_b(t)-x_d(t))-\lambda(x(t))x_b(t)} - i_d e^{-\lambda(x(t))x_d(t)} \right) \\
&= \frac{i_b i_d v(x(t)) e^{-\lambda(x(t))x_d(t)}}{i_d + \lambda(x(t))} \left(\frac{2i_d + \lambda(x(t))}{i_d} e^{-(i_d+\lambda(x(t)))(x_b(t)-x_d(t))} - 1 \right).
\end{aligned} \tag{14}$$

By (14) and using the fact that for $x \in \mathcal{V}$, $0 < \lambda(x) \leqslant i_b$ (cf. Lemma 2.4), we obtain that

$$\frac{\mathrm{d}(x_b(t) - x_d(t))}{\mathrm{d}t} \leqslant \frac{i_b i_d v(x(t)) e^{-\lambda(x(t))x_d(t)}}{i_d + \lambda(x(t))} \left(\frac{2i_d + i_b}{i_d} e^{-i_d(x_b(t)-x_d(t))} - 1 \right).$$

From the previous inequality, we deduce that on the set

$$\left\{ t \geqslant 0 : x_b(t) - x_d(t) > \frac{\log(\frac{2i_d+i_b}{i_d})}{i_d} \right\},$$

the quantity $x_b(t) - x_d(t)$ is decreasing, which allows us to conclude.

(iii): As before and by (i), we assume without loss of generality that $(x(t), t \geqslant 0) \subset U_2 \cup \mathcal{H}$. Using (ii) and since $\lambda(x) \leqslant i_b$ (cf. Lemma 2.4), we obtain that

$$\begin{aligned}
\frac{\mathrm{d}x_b(t)}{\mathrm{d}t} &= i_b v(x(t)) e^{-i_d(x_b(t)-x_d(t))} e^{-\lambda(x(t))x_b(t)} \\
&\geqslant i_b \underline{v}(x^0) e^{-C} e^{-i_b x_b(t)},
\end{aligned}$$

that allows to conclude that $x_b(t)$ increases to $+\infty$ and by (ii) we also have a similar behavior for $x_d(t)$. We now prove that $\lambda(x(t)) \to i_b$ as $t \to +\infty$. Let us recall that for all $t \geqslant 0$, $\lambda(x(t))$ is the unique solution of

$$i_b \int_0^{x_b(t)} e^{-i_d(a-x_d(t))_+ - \lambda(x(t))a} da = 1,$$

that we rewrite

$$i_b \int_0^{x_d(t)} e^{-\lambda(x(t))a} da + i_b \frac{(e^{-\lambda(x(t))x_d(t)} - e^{-i_d(x_b(t)-x_d(t))-\lambda(x(t))x_b(t)})}{i_d + \lambda(x(t))} = 1. \tag{15}$$

The map $t \mapsto \lambda(x(t))$ is clearly increasing (using (7) and the positivity of $x_b'(t)$ and $x_d'(t)$) and bounded by i_b. So there exists $\lambda^* > 0$ such that $\lambda(x(t)) \to \lambda^*$. By taking the limit $t \to +\infty$ in (15) and using the previous part of the proof, we deduce that

$$i_b \int_0^{+\infty} e^{-\lambda^* a} da = 1,$$

and $\lambda^* = i_b$ that concludes the proof. □

We now prove Theorem 4.4.

Proof of Theorem 4.4. By Lemma 4.5 (i), we assume without loss of generality that $(x(t), t \geqslant 0) \subset U_2 \cup \mathcal{H}$, i.e that for all $t \geqslant 0$, $x_b(t) - x_d(t) \geqslant 0$. We recall that Equality (14) gives:

$$\frac{\mathrm{d}x_b(t) - x_d(t)}{\mathrm{d}t} = \frac{i_b i_d v(x(t)) e^{-\lambda(x(t))x_d(t)}}{i_d + \lambda(x(t))} \left(\frac{2i_d + \lambda(x(t))}{i_d} e^{-(i_d + \lambda(x(t))(x_b(t) - x_d(t))} - 1 \right). \quad (16)$$

We define $f, h : \mathbb{R}_+ \to \mathbb{R}$ by:

$$f(t) = \frac{i_b i_d v(x(t)) e^{-\lambda(x(t))x_d(t)}}{i_d + \lambda(x(t))}$$

and

$$h(t) = \frac{2i_d + \lambda(x(t))}{i_d} e^{-(i_d + \lambda(x(t)))(x_b(t) - x_d(t))} - \frac{2i_d + i_b}{i_d} e^{-(i_d + i_b)(x_b(t) - x_d(t)}.$$

Note that $h(t) \to 0$ as $t \to +\infty$ using Lemma 4.5 (ii). Let us also define $u(t) = x_b(t) - x_d(t)$. So Equation (16) rewrites

$$\frac{\mathrm{d}u(t)}{\mathrm{d}t} = f(t) \left(\frac{2i_d + i_b}{i_d} e^{-(i_d + i_b)u(t)} - 1 + h(t) \right).$$

We deduce that for all $\epsilon > 0$, there exists $t_0 > 0$ such that for all $t \geqslant t_0$,

$$f(t) \left(\frac{2i_d + i_b}{i_d} e^{-2(i_b + i_d)u(t)} - 1 - \epsilon \right) \leqslant \frac{\mathrm{d}u(t)}{\mathrm{d}t} \leqslant f(t) \left(\frac{2i_d + i_b}{i_d} e^{-(i_d + i_b)u(t)} - 1 + \epsilon \right). \quad (17)$$

Let us consider the differential equation

$$\frac{\mathrm{d}w(t)}{\mathrm{d}t} = f(t) \left(\frac{2i_d + i_b}{i_d} e^{-(i_b + i_d)w(t)} - 1 + \epsilon \right).$$

By using the change of variables $s = e^{(i_b + i_d)w}$, we solve the previous equation and we find that there exists a constant $C(x^0)$ such that

$$w(t) = \frac{1}{i_b + i_d} \log \left(\frac{2i_d + i_b}{i_d(1 - \epsilon)} - \frac{C(x^0)}{1 - \epsilon} \exp \left(-(i_b + i_d)(1 - \epsilon) \int_0^t f(s) ds \right) \right).$$

We conclude by proving that the integral above tends to infinity as t tends to infinity. First, the inequality $x_b(t) \geqslant x_d(t)$ implies that

$$f(t) \geqslant \frac{i_b i_d v(x(t)}{i_d + \lambda(x(t))} e^{-\lambda(x(t))x_b(t)}.$$

Moreover, Equation (11) gives that

$$v(x(t))e^{-\lambda(x(t))x_b(t)} = e^{i_d(x_b(t)-x_d(t))}x_b'(t).$$

Since $\lambda(x(t)) \leqslant i_b$, we obtain that

$$f(t) \geqslant \frac{i_b i_d}{i_b + i_d} x_b'(t)$$

and that

$$\int_0^t f(s)ds \geqslant \frac{i_b i_d}{i_b + i_d} x_b(t) \underset{t\to+\infty}{\longrightarrow} +\infty.$$

By (17), we conclude that for all $\epsilon > 0$:

$$\frac{1}{i_b + i_d} \log\left(\frac{2i_d + i_b}{i_d(1+\epsilon)}\right) \leqslant \liminf_{t\to+\infty} u(t) \leqslant \limsup_{t\to+\infty} u(t) \leqslant \frac{1}{i_d + i_b} \log\left(\frac{2i_d + i_b}{i_d(1-\epsilon)}\right)$$

that concludes the proof. □

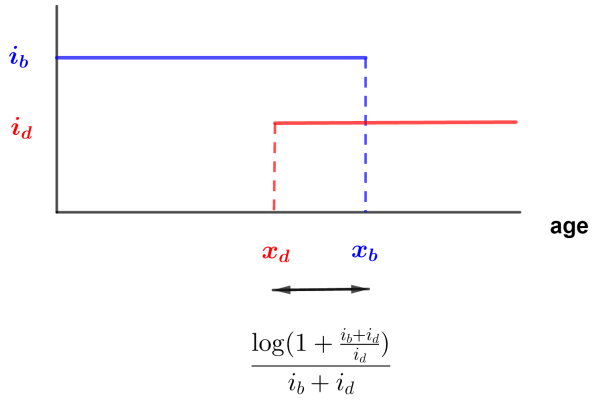


Figure 3: ***Optimal configurations as*** x_b ***and*** x_d ***tend to infinity.***

5 On the selection of Lansing effect

In this section, we ask the question of the apparition of a pro-senescence and non-genetic mechanism similar to the Lansing effect [5, 6]. We recall that the Lansing effect is the effect through which the progeny of old parents do not live as long as those of young parents.

We will show that the Lansing effect can represent a selective advantage, as an accelerator of the evolution.

5.1 The bd model with Lansing effect

The bd-model with Lansing effect is defined by modifying the bd-model that we introduced in Section 1. It was introduced and studied in details in [8] in the case where $i_b = i_d = 1$. The authors show that under the assumptions of the adaptive dynamics theory (large population, rare and small mutations), the evolution of the trait (x_b, x_d) is described by the solutions a differential inclusion which reach the diagonal $\{(x_b, x_d) \in \mathbb{R}_+^2 : x_b = x_d\}$ and then stay on it. The formula given here are generalized to the case where $i_b \neq i_d$.

The model. We assume that an individual which reproduces after age x_d transmits to its descendant a shorter life-expectancy. If an individual with trait $x = (x_b, x_d)$ reproduces at age a, the trait of its descendant is determined by a two-phases mechanism. The first phase is non-genetic and modifies the trait x: if $a < x_d$ we define $\tilde{x} = x$ but if $a > x_d$, $\tilde{x} = (x_b, 0)$. The second phase corresponds to genetic mutations which modify the trait $\tilde{x}$ similarly as in Section 1. Hence, on configurations $\{(x_b, x_d) \in \mathbb{R}_+^2 : x_b < x_d\}$, the dynamics is similar as in the model described in Section 1. Let us note that the population is then composed of two subpopulations, a population with traits $\{(x_b, x_d), x_b > 0, x_d > 0\}$ and a population with traits $\{(x_b, 0), x_b > 0\}$.

Demographic parameters. We now give for the model with Lansing effect, the analogous of the demographic parameters introduced in Section 2. We refer to [8] for the justification. We denote by $\lambda^\ell(x)$ the Malthusian parameter describing the asymptotic growth of the population with Lansing effect. It is solution of

$$i_b \int_0^{x_b \wedge x_d} e^{-\lambda^\ell(x)a} da = 1.$$

Then it can been easily computed by Newton's method (as seen in Section 2) and the set of viability $\mathcal{V}_\ell$ is simple. It is composed of the traits $x = (x_b, x_d)$ such that

$$x_b \wedge x_d > \frac{1}{i_b}.$$

The associated stable age distribution $N_x^\ell(a) = (N_x^{\ell,1}(a), N_x^{\ell,2}(a))$ satisfies

$$N_x^{\ell,1}(a) = i_b\, e^{-i_d(a-x_d)_+ - \lambda^\ell(x)a}, \quad N_x^{\ell,2}(a) = i_b\, F(\lambda^\ell(x)) e^{-(i_d+\lambda^\ell(x))a},$$

where F is some function that we don't detail here (cf. [8, Proposition 3.5]). The functions $N_x^{\ell,1}$ and $N_x^{\ell,2}$ describe the stable age distributions for populations with traits (x_b, x_d) and $(x_b, 0)$ respectively. The generation time $G^\ell(x)$ is given by

$$G^\ell(x) = \int_0^{x_b \wedge x_d} a\, i_b\, e^{-i_d(a-x_d)_+ - \lambda^\ell(x)a} da = i_b \int_0^{x_b \wedge x_d} a e^{-\lambda^\ell(x)a} da. \tag{18}$$

We observe that the Malthusian parameter $\lambda^\ell(x)$ and the mean generation time $G^\ell(x)$ only take into account the individuals reproducing before age $x_b \wedge x_d$.

Evolution of the trait with Lansing effect. Let us now describe the behaviour of the trait.
On the subset $\{x_b < x_d\}$, the Lansing effect doesn't act. So, the dynamics is similar as the one described in the above sections. The trait dynamics is described by the differential equation

$$\frac{\mathrm{d}x_b(t)}{\mathrm{dt}} = \frac{\partial\lambda^\ell(x)}{\partial x_b}\frac{\lambda^\ell(x)}{2i_b\, c\int_{\mathbb{R}_+}(N_x^{\ell,1}+N_x^{\ell,2})(a)da}\sigma^2(x), \quad \frac{\mathrm{d}x_d(t)}{\mathrm{dt}} = 0.$$

Thus, the trait x_b increases while the trait x_d stays constant.
On the subset $\{x_b > x_d\}$, only individuals breeding before the age x_d will have viable offspring. Thus, there is no selective advantage in extending the reproduction phase by increasing x_b, but only in increasing survival by increasing x_d. More precisely, on $\{x_b < x_d\}$, we have:

$$\frac{\mathrm{d}x_d(t)}{\mathrm{dt}} = \frac{\partial\lambda^\ell(x)}{\partial x_d}\frac{\lambda^\ell(x)}{2i_b\, c\int_{\mathbb{R}_+}(N_x^{\ell,1}+N_x^{\ell,2})(a)da}\sigma^2(x), \quad \frac{\mathrm{d}x_b(t)}{\mathrm{dt}} = 0. \tag{19}$$

Indeed, the derivatives of the fitness are given as follows (see [8] Proposition 4.1).

$$\forall x\in U_1, \nabla\lambda^\ell(x) = \left(\frac{i_b e^{-\lambda^\ell(x)x_b}}{G^\ell(x)}, 0\right) \;;\; \forall x\in U_2, \nabla\lambda^\ell(x) = \left(0, \frac{i_b e^{-\lambda^\ell(x)x_d}}{G^\ell(x)}\right).$$

We observe that the trait x_d increases while the trait x_b stays constant. Hence, whatever the initial condition, the trait x reaches in finite time the diagonal $\{x_b = x_d\}$ and then stays on it. On this diagonal the trait can evolve at different speeds (the dynamics is not unique): the global behavior of the trait is described by a differential inclusion (cf. [8, Theorem 4.17]).

5.2 Selection for Lansing effect

Let us first note that for non-Lansing and Lansing populations, as observed in the study of adaptive dynamics, the long time strategy leads to traits x_b and x_d going to infinity, with $x_b - x_d = \log(1+\frac{i_b+i_d}{i_d})/(i_b+i_d)$ in the non-Lansing case and $x_b = x_d$ in the Lansing case (see [8, Theorem 4.17] in that case). It is then easy to deduce that in both cases, the Malthusian parameter, which has been proved to be less than i_b, converges to i_b when t tends to infinity. Therefore the evolution will give the same selective advantage to both populations, making possible the cohabitation of the two populations. In addition, we observe that the partial derivatives of the Malthusian parameters with respect to x_b or x_d (in both cases) are positive, meaning that the convergences are increasing. Let us consider a monotype population with trait $(x_b, x_d)\in U_2$, then by definition, we obtain that

$$\lambda^{n\ell}(x) > \lambda^\ell(x)$$

at time 0. Thus there are periods where the Lansing fitness will increase much more than the non-Lansing one.

In order to assess the relative evolutionary success of non-Lansing/Lansing populations, we consider a population composed of two sub-monomorphic populations with traits respectively $x^\ell = (x_b^\ell, x_d^\ell)$ and $x^{n\ell} = (x_b^{n\ell}, x_d^{n\ell})$, the first one subject to the Lansing effect and the second one which is not affected by this senescence effect, both subjected to the same competitive pressure. The traits have been chosen such that the two sub-populations have the same Darwinian fitness $\lambda^{n\ell}(x^{n\ell}) = \lambda^\ell(x^\ell)$. In each sub-population, the dynamics is described either in Section 1 (without Lansing effect) or in Section 5.1 (with Lansing effect). Let us first note that since $\lambda^{n\ell}(x^{n\ell}) = \lambda^\ell(x^\ell)$ and since by definition,

$$i_b \int_0^{x_d^\ell} e^{-\lambda^\ell(x^\ell)a} da = 1 = i_b \int_0^{x_d^{n\ell}} e^{-\lambda^{n\ell}(x^{n\ell})a} da + i_b \int_{x_d^{n\ell}}^{x_b^{n\ell}} e^{-i_d(a-x_d^{n\ell})} e^{-\lambda^{n\ell}(x^{n\ell})a} da,$$

we deduce immediately that

$$x_b^{n\ell} > x_d^\ell > x_d^{n\ell}.$$

We observe the isoclines of $\lambda^{n\ell}$ and λ^ℓ when they have the same values. Although they are very simple (horizontal or vertical lines) in the Lansing case, and in the region U_1 for the non-Lansing case, they have a more complicated form in the region U_2 for the non-Lansing case (cf. Figure 5 of the main paper).

Let us consider the points $x^{n\ell} \in U_2$ such that $\lambda^{n\ell}(x^{n\ell})$ has a fixed constant value. Using the Implicit Function Theorem, we know the existence of a real-valued smooth function $\varphi_{n\ell}$ such that for all these points, $x_d^{n\ell} = \varphi_{n\ell}(x_b^{n\ell})$. Further,

$$\varphi'_{n\ell}(x_b^{n\ell}) = -\frac{\frac{\partial \lambda^{n\ell}}{\partial x_b^{n\ell}}}{\frac{\partial \lambda^{n\ell}}{\partial x_d^{n\ell}}}(x^{n\ell}).$$

The previous computations showed that the partial derivatives of $\lambda^{n\ell}$ are positive, and then that $\varphi'_{n\ell}(x_b^{n\ell}) < 0$, yielding the function φ_ℓ to be decreasing on U_2. Moreover, the exact computation gives

$$\varphi'_\ell(x_b^{n\ell}) = -\frac{i_b e^{-i_d(x_b^{n\ell} - x_d^{n\ell})} e^{-\lambda(x^{n\ell})x_b^{n\ell}}}{i_b i_d e^{i_d x_d^{n\ell}} \int_{x_d^{n\ell}}^{x_b^{n\ell}} e^{-(i_d + \lambda^{n\ell}(x^{n\ell}))a} da} \geqslant -\frac{1}{i_d(x_b^{n\ell} - x_d^{n\ell})}.$$

The last inequality explains the almost vertical tangent observed when $x^{n\ell}$ is close to the diagonal (see Figure 5 of the main paper).

References

[1] B. Charlesworth. *Evolution in age-structured populations*, volume 2. Cambridge University Press Cambridge, 1994.

[2] U. Dieckmann and R. Law. The dynamical theory of coevolution: a derivation from stochastic ecological processes. *Journal of mathematical biology*, 34(5-6):579–612, 1996.

[3] Regis Ferriere and Viet Chi Tran. Stochastic and deterministic models for age-structured populations with genetically variable traits. In *ESAIM: Proceedings*, volume 27, pages 289–310. EDP Sciences, 2009.

[4] M.E. Gurtin and R.C. MacCamy. Non-linear age-dependent population dynamics. *Archive for Rational Mechanics and Analysis*, 54(3):281–300, 1974.

[5] A.I. Lansing. A transmissible, cumulative, and reversible factor in aging. *Journal of Gerontology*, 2(3):228–239, 1947.

[6] A.I. Lansing. A nongenic factor in the longevity of rotifers. *Annals of the New York Academy of Sciences*, 57(1):455–464, 1954.

[7] S. Méléard and V.C. Tran. Trait substitution sequence process and canonical equation for age-structured populations. *Journal of mathematical biology*, 58(6):881–921, 2009.

[8] Sylvie Méléard, Michael Rera, and Tristan Roget. A birth–death model of ageing: from individual-based dynamics to evolutive differential inclusions. *Journal of mathematical biology*, pages 1–39, 2019.

[9] J.A.J. Metz, S.A.H. Geritz, G. Meszéna, F.A.J. Jacobs, and J.S. Van Heerwaarden. Adaptive dynamics: a geometrical study of the consequences of nearly faithful reproduction. *Stochastic and Spatial Structures of Dynamical Systems. Elsevier Science, Burlington, MA*, pages 183–231, 1996.

[10] B. Perthame. *Transport equations in biology*. Springer Science & Business Media, 2006.

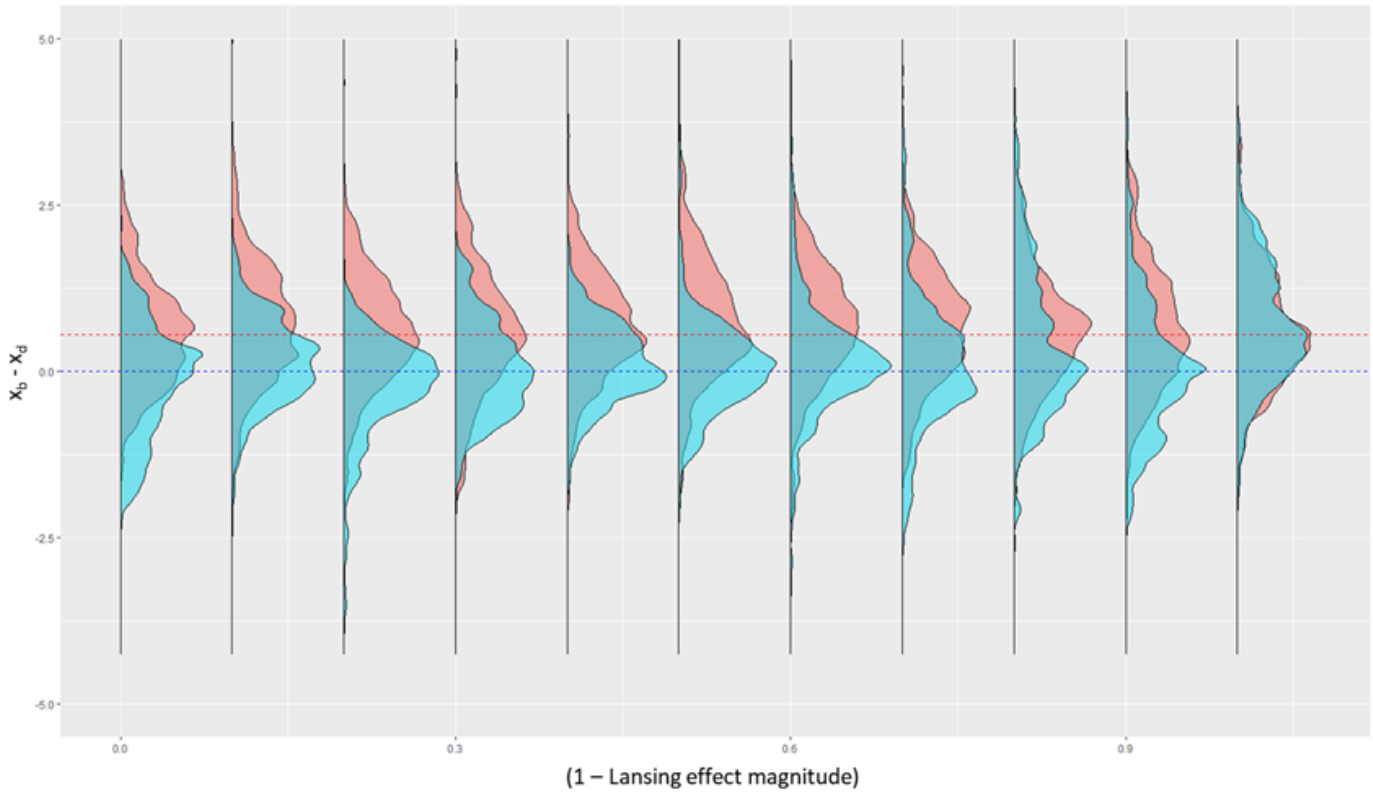
[11] G.F. Webb. *Theory of nonlinear age-dependent population dynamics*. CRC Press, 1985.

Annexe 2: codes for simulations and data visualization

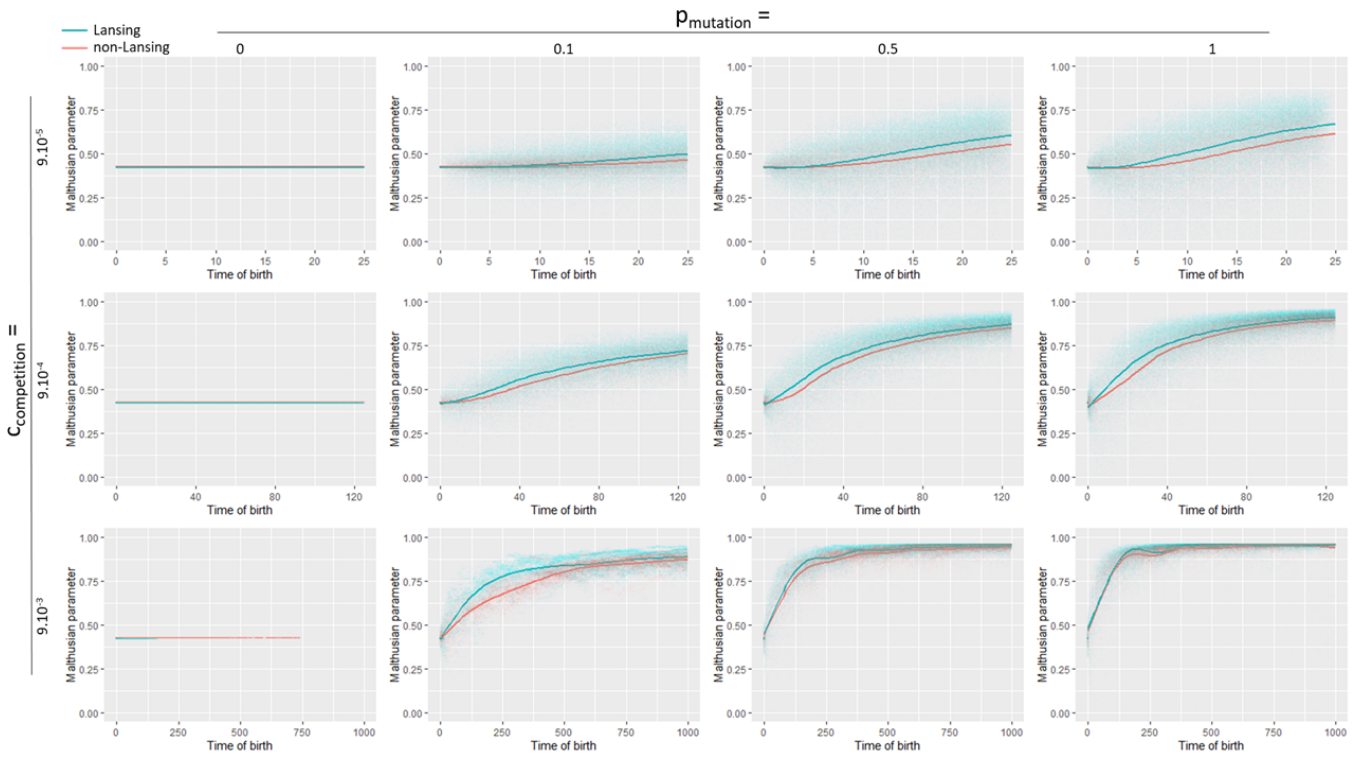
*Package IBMPopSim (R package IBMPopSim v0.3.1):
https://cran.r-project.org/web/packages/IBMPopSim/index.html
*Environment for simulations using IBMPopSim:
https://mybinder.org/v2/gh/MichaelRera/EvoAgeing/HEAD

.Exploring parameters for Lansing populations	Modele_Lansing_evo.ipynb
.Exploring parameters for non-Lansing populations	Modele_nonLansing_evo.ipynb
.Lansing / non-Lansing competition for equal Malthusian parameters	L_nL_compet_eqMalth.ipynb
.Lansing / non-Lansing competition (x_b-x_d) € [-10; 10]	L_nL_compet_heteroPop.ipynb

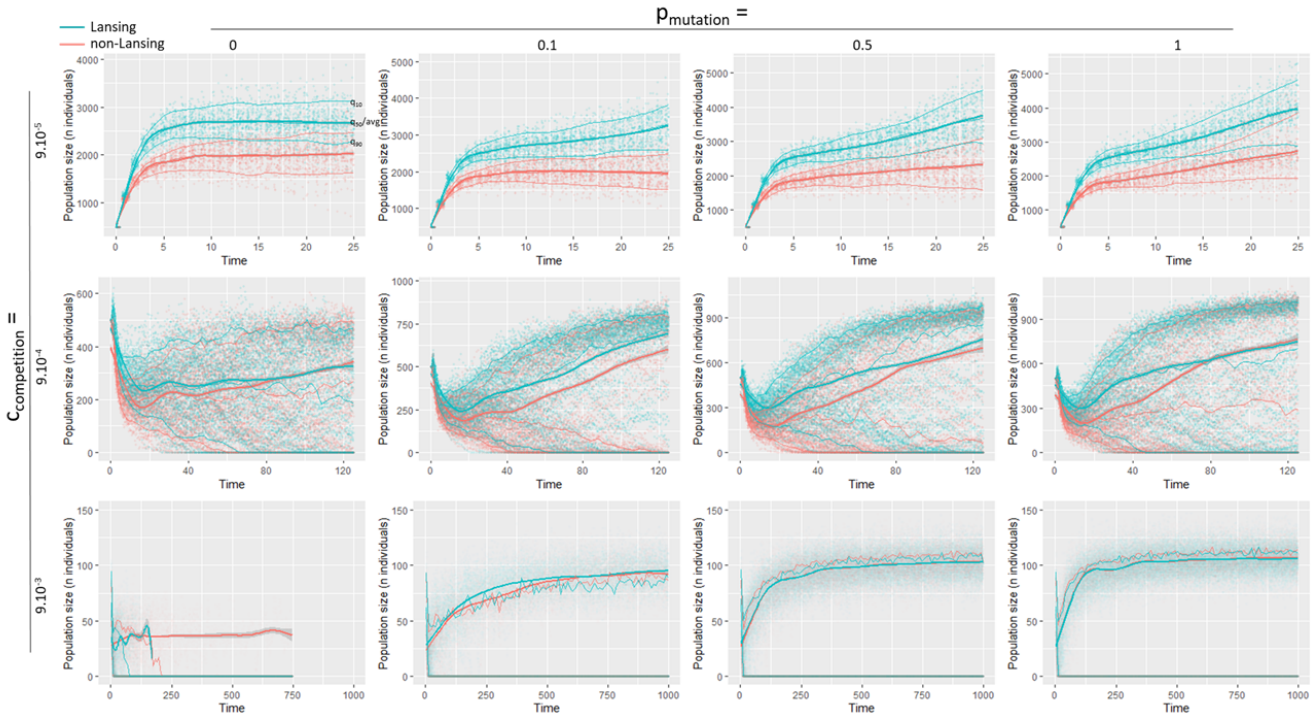
Supplementary figures



Supplementary figure 1: The magnitude of the Lansing effect does not influence the outcome of evolution. 100 independent simulations were run for each Lansing effect magnitude ranging from 0 (no Lansing effect) to 1 (progeny from parents age € [x_d; x_b] have $x_d = 0$), starting with 500 Lansing (1.5; 1.3) and 500 non-Lansing (1.5; 0.83) individuals. We plot here the distribution density of x_b - x_d at the end of the simulation (individuals born in the time interval [990; 1000]), for Lansing populations (blue) and non-Lansing ones (red). Surprisingly, the magnitude of the Lansing effect does not seem to affect the optimal x_b - x_d solution value.



Supplementary figure 2: Evolution of the average Malthusian parameter value in Lansing and non-Lansing populations as a function of time. *p* is the mutation rate and *c* is the logistic competition intensity. Individual values are plotted, the line represents the average value amongst populations. In all conditions with $p > 0$, the Malthusian parameter grows faster and remains slightly higher in the Lansing populations than in the non-Lansing ones.



Supplementary figure 2: Evolution of the Lansing and non-Lansing populations size as a function of time. *p* is the mutation rate and *c* is the logistic competition intensity.