Directional and disruptive selection in populations structured by class and continuous ontogeny under incomplete plasticity

Arthur Weyna[1], Charles Mullon[1], Laurent Lehmann[1]

[1] Department of Ecology and Evolution, University of Lausanne, 1015 Lausanne, Switzerland

Corresponding author: arthur.weyna@unil.ch

Abstract

Many organisms undergo ontogeny, whereby individuals change in state (e.g. in size, morphology, or condition) as they age. Understanding the evolution of traits influencing ontogeny is challenging because their fitness effects unfold across an individual's lifetime and may differ between classes such as sexes. Here, we analyse selection on non-plastic traits (e.g., fixed resource allocation strategies) that determine the development of dynamical states throughout life (e.g., body size), with consequences for fecundity and survival in age- and class-structured populations. Using invasion analysis, we derive expressions for directional and quadratic selection that decompose into age- and class-specific components. This allows us to identify convergence stable trait values, assess whether they are uninvadable or potentially experience evolutionary branching, and pinpoint the age and class pathways through which correlational and disruptive selection act. Applying our results to a model of growth under size-mediated sexual selection, we show how selection on growth rates distributes across ages and sexes, and how its relative strength depends on genetic correlations, individual ploidy and life-history. We also show how sex-specific developmental trade-offs and constraints can generate disruptive selection on male growth and favour the evolution of alternative male life histories. More broadly, our results highlight how adaptation is mediated by the interaction of development and demography, and provide tools to investigate how conflicts across ages and classes may influence senescence, sexual dimorphism, and the diversification of ontogenetic strategies.

Author Summary

How does natural selection shape traits that influence how organisms develop through their lifetime? This question is difficult because the effects of such traits unfold dynamically as individuals change in states such as size or condition, which in turn influence survival and reproduction throughout life. We formulate a mathematical model that follows these causal links and characterises the gradual evolution of non-plastic traits affecting development. Our results apply to populations with different classes, such as males and females, that follow distinct developmental and demographic paths. Our model reveals when selection is stabilizing, holding populations fixed for an optimal trait value, and when it is disruptive, favouring trait diversification and the emergence of alternative life-history strategies. As an illustration, we analyse the evolution of sex-specific growth under size-based sexual selection. We show how the evolution of male and female growth rates depends on genetics, such as individual ploidy or genetic correlation between traits, but also on life history, which determines the relative strength of selection acting on trait expression at

different ages. We further show how developmental trade-offs and demographic differences can generate disruptive selection on male growth, leading to the coexistence of fast-growing short-lived males and slow-growing long-lived ones. More broadly, our results link development, demography, and adaptation, and show how their interplay shapes biological diversity both within and between populations.

1 Introduction

Many traits expressed by an organism throughout its life influence not only its survival and reproduction, but also its ontogeny. The fitness effects of such traits can be immediate, when their expression acts on current reproduction, as well as delayed, when their expression acts on future reproduction by modulating current maintenance and growth. Understanding how selection acts on traits affecting ontogeny is thus difficult because their fitness effects typically unfold over an individual's entire lifetime. Additionally, these fitness effects often depend on the traits expressed by others and vary among classes of individuals, such as individuals of different sexes or ploidies, in different environments, or in different conditions. For instance, growth allocation can lead to different outcomes in males and females of the same population due to sex-specific developmental trajectories and sexual selection (e.g. more individuals of a given size can increase competition and sex-specific size variation can influence mate choice). Such interactions between ontogeny, class structure and social effects make it challenging to understand how selection shapes traits affecting ontogeny.

The inheritance and expression of traits affecting ontogeny can be conceptualised in different ways, with important implications for the operation of natural selection. Consider for example the growth allocation schedule of females and males in some vertebrate species. At one extreme, this schedule could be a completely plastic trait, with allocation varying independently across age and sex. This perspective, often adopted in evolutionary life-history models (e.g. León 1976; Perrin & Sibly 1993; Houston et al. 2024), allows trait expression to be adjusted optimally to each context, albeit within preset physiological constraints (e.g. physical boundaries in trait space, available resources). Such complete plasticity, however, implies the existence of a number of genetic loci large enough that expression can be controlled to match each possible context, allowing mutations to generate age- and/or class-specific changes. It also implies that selection must be strong enough, on these numerous loci, to perfectly fine-tune trait expression. This contrasts with the fact that plasticity is often incomplete owing to genetic constraints (i.e. due to trait expression being controlled by a limited number of loci), leading to similar trait values being expressed across contexts. Under such constraints, selection acts on a shared genetic architecture, potentially generating antagonistic pressures across age (Austad & Hoffman 2018) or classes (Poissant et al. 2010). In humans for instance, some mutations in BRCA genes are associated with increased mortality in older individuals (via higher rates of breast and ovarian cancer), but also with increased fertility in young individuals (via earlier age at paternity in men and lower miscarriage rates in women; Kwiatkowski et al. 2015).

Mathematical models have been useful for analysing selection on traits affecting ontogeny under complete

and incomplete plasticity. These models typically consider gradual trait evolution, whereby under a constant influx of mutations with small effects on traits, a population evolves under directional selection towards a local evolutionary equilibrium. Under complete age-wise plasticity, directional selection on age-specific trait expression depends on its effect on current, as well as future reproduction, via changes in development or survival (e.g., Perrin & Sibly 1993; Day & Taylor 2000; Metz et al. 2016; Avila et al. 2021; González-Forero 2024). These models show that by treating ontogeny as a dynamical system at the individual level, it becomes possible to track the future fitness consequences of age-specific trait expression, such that each instant of expression can be optimized. More recent studies have applied the same future-accounting methods to cases where plasticity is incomplete (Avila et al. 2021; Lehmann 2022). When a shared genetic architecture forces the same trait value to be expressed at different ages or times, directional selection can be understood as a weighted sum of fitness effects over all instants where a trait value is expressed (Avila et al. 2021; Lehmann 2022; Avila & Mullon 2023). The weights in that sum then reflect the discount in selection efficiency that applies to later trait expression, due to the thinning of cohorts and to the reduction in remaining reproduction opportunities (Hamilton 1966). This body of theory provides a way to analyse the outcome of directional selection on traits that affect ontogeny under genetic constraints, while also accounting for age-specific demography and social interactions.

Once a population has converged towards a local evolutionary equilibrium under directional selection (a "convergence stable trait" value, Eshel 1983; Christiansen 1991), it may either remain genetically monomorphic due to stabilising selection or become polymorphic owing to disruptive selection (e.g. Christiansen 1991; Eshel et al. 1997; Geritz et al. 1998). However, disruptive selection on traits that affect ontogeny remains understudied (Avila & Mullon 2023). Nevertheless, empirical evidence suggests selection plays a role in shaping life-history polymorphisms. In the Atlantic salmon for instance, young males can adopt alternative ontogenetic trajectories that favour either short-term mating or delayed reproduction after migration and growth (Lepais et al. 2017). This polymorphism is in part explained by genetic variation in developmental thresholds, but current theory offers no general method to assess whether this variation is favoured by disruptive selection. A second limitation is that most current analyses of selection on traits affecting ontogeny consider populations structured by age only. Yet many natural populations are also structured in various discrete classes (e.g. sexes, ploidies or local environments) in which the consequences of trait expression differ. Polymorphisms in such populations can occur within a single class (Lepais et al. 2017) or span multiple classes, sometimes with overlapping genetic control (Braendle et al. 2005). It is unknown to what extent genetic constraints, by coupling selection pressures acting in different classes, may affect life history evolution and diversification.

Here, we fill these gaps by extending current theory to analyse directional and disruptive selection on a

finite number of quantitative traits expressed constantly throughout life history (i.e. incomplete plasticity) and that affect ontogeny in class-structured populations. We consider selection on traits expressed by individuals belonging to different classes with specific development, ploidy, demography, and social context. We derive expressions for both directional and disruptive selection on these traits, that we partition into class- and age-wise components, and provide numerical tools that facilitate the calculation of these expressions. Additionally, we use these results to study a model of size-mediated sexual selection under a sex-specific survival-growth trade-off and frequency-dependent competition. In this application, we find that under strong sexual selection, disruptive selection on male growth rates leads to the emergence of multiple male life histories: some grow rapidly to large size, thereby obtaining mating opportunities with highly fecund females but suffering higher mortality; others remains small, specialise in less fecund mates, and survive longer. Beyond this specific example, our results allow for analyses of directional and disruptive selection under a wide range of evolutionary contexts including ecological and social feedbacks, and thus opens the door to a sharper understanding of the selective forces that shape life history and its variation within populations.

2 The model

2.1 Biological assumptions

We consider a large age-structured population in continuous time, under constant global abiotic environmental conditions. Each individual in the population belongs to one of a finite number n_{c} of classes, such as sex, mating type, or habitat state. The class of an individual is determined at birth and does not change during its lifetime. We leave unspecified at this point the mode of reproduction and ploidy of each class, but assume that they are constant properties of the population. To implement class-specific ontogeny, we associate to a focal individual of class j and age a a vector $\boldsymbol{x}_j(a) = (x_{j,1}(a), ..., x_{j,n_{\mathrm{x}}}(a))$ of n_{x} class-specific internal state variables, where $x_{j,m}(a)$ is a scalar giving the value of the focal's m-th internal state at age a (see tab. 1 for a summary of the symbols used throughout), and where n_{x} is assumed to be independent of age. Internal states can represent any relevant biological quantitative feature that changes with age, such as the mass or size of a given organ, or the probability of occurrence of a physiological event (e.g. moulting or menopause).

We assume that the rates of change $\mathrm{d}\boldsymbol{x}_j(a)/\mathrm{d}a$ in internal states in a focal individual of age a and class j depend deterministically on its internal states and on a vector $\boldsymbol{z}_j = (z_{j,1}, ..., z_{j,n_{\mathrm{z}}})$ of n_{z} class-specific evolving quantitative traits, where $z_{j,l}$ is a scalar giving the value of the lth trait expressed in class j

Symbol	Meaning
Parameters	
n_{c}	Number of classes
n_{z}	Number of evolving traits in each class
n_{x}	Number of internal states
γ_{ij}	Expected number of alleles transmitted to offspring of class i by a parent of class j
c_{ij}	Proportion of individuals of class i within the offspring of individuals of class j
Variables	
$u_{j,l}$, $v_{j,l}$	Allelic value for trait l in class j encoded by mutant and resident alleles, respectively
$\boldsymbol{u}_j$, $\boldsymbol{v}_j$	Vectors of mutant and resident allelic values in class j, respectively
$\boldsymbol{u}$, $\boldsymbol{v}$	Vectors collecting mutant and resident allelic values across all classes ($\boldsymbol{v}$ is also used as a shorthand for resident traits as $\boldsymbol{z}(\boldsymbol{v},\boldsymbol{v}) = \boldsymbol{v}$)
$z_{j,l}(u_{j,l}, v_{j,l})$	Value of trait l in a mutant of class j
$\boldsymbol{z}_j(\boldsymbol{u}_j, \boldsymbol{v}_j)$	Vector of trait values in a mutant of class j
$\boldsymbol{z}(\boldsymbol{u}, \boldsymbol{v})$	Vector collecting mutant trait values across all classes
Functions	
$l_j(a)$, $l_j^\circ(a)$	Survival probability (survivorship) to age a of a mutant and resident individual of class j, respectively
$\boldsymbol{x}_j(a)$, $\boldsymbol{x}_j^\circ(a)$	Internal state value at age a of a mutant and resident individual of class j, respectively
$f_{ij}(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$	Fecundity rate: rate of production of offspring of class i by a mutant of class j at age a
$f_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$	Total fecundity rate: total rate of production of offspring by a mutant of class j at age a
$\mu_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$	Mortality rate in a mutant of class j and age a
$\boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$	Vector of rates of change in internal states at age a in a mutant of class j
$\mathbf{R}(\boldsymbol{u}, \boldsymbol{v})$	Mutant next-generation matrix
$R_0(\boldsymbol{u}, \boldsymbol{v})$	Mutant basic reproductive number
$\boldsymbol{\nu}^\circ, \nu_i^\circ$	Leading left-eigenvector of $\mathbf{R}^\circ = \mathbf{R}(\boldsymbol{v}, \boldsymbol{v})$ and its element i, i.e. resident allelic reproductive values at birth of class i
$\boldsymbol{q}^\circ, q_j^\circ$	Leading right-eigenvector of $\mathbf{R}(\boldsymbol{v}, \boldsymbol{v})$ and its element j, i.e. the resident allelic frequency at birth of class i
$R_{ij}(\boldsymbol{u}_j, \boldsymbol{v})$	Expected lifetime fecundity via class-i offspring of a mutant of class j
$R_j(\boldsymbol{u}_j, \boldsymbol{v})$	Total expected lifetime fecundity of a mutant of class j
$H_{ij}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j(a), \boldsymbol{v})$	Class-oriented Hamiltonian
$\lambda_{ij}^{\mathrm{l}}(a)$, $\boldsymbol{\lambda}_{ij}^{\mathrm{x}}(a)$	Class-oriented costates associated to survivorship and internal states
$H_j(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j(a), \boldsymbol{v})$	Class-specific Hamiltonian
$\lambda_j^{\mathrm{l}}(a)$, $\boldsymbol{\lambda}_j^{\mathrm{x}}(a)$	Class-specific costates associated to survivorship and internal states, respectively
$\mathbf{s}_j(\boldsymbol{v})$	Selection gradient specific to class j
$s_{j,l}(\boldsymbol{v}), \hat{s}_{j,l}(a, \boldsymbol{v})$	Total and age-specific directional selection on trait (j, l)
$\mathbf{J}(\boldsymbol{v}), \mathbf{J}_{ij}(\boldsymbol{v})$	Jacobian matrix and its block (i, j)
$\mathbf{H}(\boldsymbol{v}), \mathbf{H}_{ij}(\boldsymbol{v})$	Hessian matrix and its block (i, j)
$h_{ij,kl}(\boldsymbol{v}), \hat{h}_{ij,kl}(a, \boldsymbol{v})$	Total and age-specific quadratic selection on traits (i, k) and (j, l)

Tab. 1. List of symbols used in the main text.

(called trait (j, l) hereafter). Trait values are assumed to be constant properties of individuals throughout ontogeny (e.g. constant allocation proportions), fully determined by their genetic value (no stochastic environmental effects on trait expression) under additive genetic action (no dominance). All traits are assumed to be encoded by the same locus under a continuum-of-alleles model without mutational bias (e.g. Kimura 1965; Bürger 2000).

Besides their effect on development, the internal states and traits of an individual of class j are assumed to also influence its vital rates, that is its mortality rate μ_j, and the rate f_{ij} at which it produces offspring of class i. Hereafter, we will refer to those quantities conditional on both a class j parent and a class i offspring as *class-oriented* quantities (e.g. f_{ij} is the class-oriented fecundity rate). An individual's developmental rates and vital rates are further assumed to depend deterministically on a set of environmental variables, that are common to all individuals in the population but may affect individuals differently depending on their internal states. These external environmental variables can be demographic (e.g. population density) or ecological (e.g. available resources, pathogen density), and their dynamics can thus depend on the population.

Taken together, these assumptions let us derive the evolutionary dynamics of the evolving traits in populations structured by class and ontogeny. Our objective is to establish the first and second order conditions necessary to classify the singular points of gradual evolutionary dynamics (Geritz et al. 1998). To this aim, it is sufficient to carry out an invasion analysis (e.g. Charlesworth 1973; Michod 1979; Eshel 1984; Vincent & Brown 1984; Geritz et al. 1998; Avila & Mullon 2023), that is to focus on the invasion process of a rare mutant allele inducing a trait change in a population fixed for some resident allele (i.e. monomorphic for some resident trait).

2.2 The mutant basic reproductive number

To account for the effect of mutant alleles on mutant trait expression under arbitrary ploidy and reproductive systems, we denote by $z_{j,l}(u_{j,l}, v_{j,l})$ the value of trait (j, l) in a mutant individual (i.e. an individual carrying at least one mutant allele), where $u_{j,l}$ is the mutant allelic value specific to trait (j, l) and $v_{j,l}$ is the corresponding resident allelic value. The allelic value of an allele is defined here as the trait value induced by that allele in an individual carrying only copies of that allele. This construction allows to cover simple systems such as, for instance, haploidy (where mutants carry a single mutant allele, $z_{j,l}(u_{j,l}, v_{j,l}) = u_{j,l}$), diploidy under full genetic mixing and dosage compensation (i.e. where all mutant individuals are heterozygotes with trait $z_{j,l}(u_{j,l}, v_{j,l}) = \frac{1}{2}(u_{j,l} + v_{j,l})$) or any combination of both across classes (e.g. haplodiploidy). This notation for trait $z_{j,l}$ also accounts for more complicated

cases, such as systems with polyploid classes, or where mutants can carry more than one mutant allele (e.g. diploid mutants produced via selfing; see for instance Lesaffre & Billiard 2020). With this, we let $\boldsymbol{z}(\boldsymbol{u},\boldsymbol{v}) = (\boldsymbol{z}_1(\boldsymbol{u}_1,\boldsymbol{v}_1),...,\boldsymbol{z}_{n_c}(\boldsymbol{u}_{n_c},\boldsymbol{v}_{n_c}))$ stand for the vector collecting mutant trait values in each class, where $\boldsymbol{z}_j(\boldsymbol{u}_j,\boldsymbol{v}_j) = (z_{j,1}(u_{j,1},v_{j,1}),...,z_{j,n_z}(u_{j,n_z},v_{j,n_z}))$ is the vector of traits expressed by a mutant individual of class j. Therein, $\boldsymbol{u} = (\boldsymbol{u}_1,...,\boldsymbol{u}_{n_c})$, where $\boldsymbol{u}_j = (u_{j,1},...,u_{j,n_z})$ stands for the vector of mutant allelic values specific to class j, and likewise $\boldsymbol{v} = (\boldsymbol{v}_1,...,\boldsymbol{v}_{n_c})$, where $\boldsymbol{v}_j = (v_{j,1},...,v_{j,n_z})$ stands for the vector of resident allelic values specific to class j. Under the assumptions of additive gene action and dosage compensation, individuals carrying only resident alleles have traits equal to the resident allelic value. Thus, resident trait (j,l) can conveniently be written as $z^{\circ}_{j,l} = z_{j,l}(v_{j,l},v_{j,l}) = v_{j,l}$ (here and hereafter, we use the circle $\circ$ superscript to denote evaluation at resident trait values, i.e. $u_{j,l} = v_{j,l}$), such that $\boldsymbol{z}^{\circ} = \boldsymbol{v}$.

Having defined trait values in terms of allelic values, we now turn to the conditions of spread of a mutant allele. It follows from our demographic assumptions, and from standard results on multi-type age-dependent branching processes (Mode 1969), that the fate of a lineage of mutant allelic copies encoding allelic values $\boldsymbol{u}$, and descending from a single copy introduced into a resident population monomorphic for allelic values $\boldsymbol{v}$ at equilibrium (i.e. where environmental processes have reached a fixed attractor; e.g. Michod 1979; Metz et al. 1992; Lion 2018), can be predicted from its basic reproductive number $R_0(\boldsymbol{u},\boldsymbol{v})$. This is the expected number of descendant mutant copies produced over its whole lifetime by an individual carrying the mutant allele, averaged over all possible individual contexts in which the mutant allele can be found in. Specifically, a mutant lineage starting from a single copy goes extinct with certainty when $R_0(\boldsymbol{u},\boldsymbol{v}) \leq 1$, but has a positive probability of not going extinct when $R_0(\boldsymbol{u},\boldsymbol{v}) > 1$.

The basic reproductive number is obtained as the leading eigenvalue of the mutant next-generation matrix $\mathbf{R}(\boldsymbol{u},\boldsymbol{v})$, whose ijth element is the expected number of descendant copies that a mutant allele found in an individual of class j produces over its lifetime, via descendant individuals of class i, when the population is monomorphic for $\boldsymbol{v}$ (Mode 1969; Metz & Leimar 2011; Inaba 2017; see appendix A.1 for details). This element can be written as the product $\gamma_{ij}R_{ij}(\boldsymbol{u}_j,\boldsymbol{v})$, where $R_{ij}(\boldsymbol{u}_j,\boldsymbol{v})$ is the expected lifetime number of offspring of class i produced by a focal mutant individual of class j (i.e. individual class-oriented lifetime fecundity), and where γ_{ij} gives the expected number of mutant allelic copies transmitted to one offspring of class i by a parent of class j (i.e. converting *individual* fecundity into *allelic* fecundity and assumed to be constant). In a well-mixed diploid population with sexual reproduction for instance, $\gamma_{ij} = 1/2$ for any i and j as diploid mutants carry one mutant allele that is transmitted to half their offspring.

We express individual class-oriented lifetime fecundity as the integral

$$R_{ij}(\boldsymbol{u}_j, \boldsymbol{v}) = \int_0^\infty l_j(a) f_{ij}(\boldsymbol{z}_j(\boldsymbol{u}_j, \boldsymbol{v}_j), \boldsymbol{x}_j(a), \boldsymbol{v}) \, \mathrm{d}a, \tag{1}$$

over all possible ages of the product of the probability $l_j(a)$ to survive to age a (i.e. survivorship) with $f_{ij}(\boldsymbol{z}_j(\boldsymbol{u}_j, \boldsymbol{v}_j), \boldsymbol{x}_j(a), \boldsymbol{v})$, the rate at which a mutant individual of age a and class j produces offspring of class i, which depends on the individual's state $\boldsymbol{x}_j(a)$. Survivorship and state values, in turn, constitute the unique solution to the system of autonomous ordinary differential equations (ODEs):

$$\begin{aligned} \frac{\mathrm{d}l_j(a)}{\mathrm{d}a} &= -\mu_j(\boldsymbol{z}_j(\boldsymbol{u}_j, \boldsymbol{v}_j), \boldsymbol{x}_j(a), \boldsymbol{v}) \; l_j(a) \quad \text{with i.c.} \quad l_j(0) = 1, \\ \frac{\mathrm{d}\boldsymbol{x}_j(a)}{\mathrm{d}a} &= \boldsymbol{g}_j(\boldsymbol{z}_j(\boldsymbol{u}_j, \boldsymbol{v}_j), \boldsymbol{x}_j(a), \boldsymbol{v}) \quad \text{with i.c.} \quad \boldsymbol{x}_j(0) = \boldsymbol{x}_{j,0}, \end{aligned} \tag{2}$$

where $\mu_j(\boldsymbol{z}_j(\boldsymbol{u}_j, \boldsymbol{v}_j), \boldsymbol{x}_j(a), \boldsymbol{v})$ and $\boldsymbol{g}_j(\boldsymbol{z}_j(\boldsymbol{u}_j, \boldsymbol{v}_j), \boldsymbol{x}_j(a), \boldsymbol{v})$ are, respectively, the mortality rate and vector of developmental rates of a mutant of class j (with $\boldsymbol{x}_{j,0}$ being a fixed vector of initial values for the internal states).

The fact that the fecundity, mortality, and developmental rates in eqs. (1) and (2) do not depend explicitly on age a reflects the modelling assumption that all environmental and physiological effects on individuals are mediated by their internal state variables. The biological rationale behind this assumption is that organisms typically respond not to time per se, but to internal cues (e.g. energy reserves, physiological state, accumulated damage) or to external cues (e.g. temperature, photoperiod) that may correlate with age or other aspects of condition. Under this rationale, the dependence of eqs. (1)-(2) on the resident traits $\boldsymbol{v}$ summarizes all interactions that a focal mutant has with its environment, when this environment is at an equilibrium (which is assumed throughout and entails no global environmental fluctuations). This is because all demographic and/or ecological state variables are then fully determined by the resident population, and thus by resident traits. In other words, environmental processes, as complex as they may be, do not need to be modelled explicitly for a general characterization of selection in our model. Within specific applications however, the environment does need to be modelled, such that the effect of $\boldsymbol{v}$ on a mutant can be measured (see for instance De Roos 1997 for ecological feedbacks, or Lehmann 2022 for behavioral feedbacks. See also the example model of section 4 where simple density-dependent fecundity is implemented).

In summary, our model decomposes class-oriented lifetime reproductive success in terms of three embedded phenotypic layers. The highest layer is composed of vital rates (i.e. fecundity and mortality rates), which depend on individual traits, current states and the environment. The second layer is that of the

states, which summarize the current behavioural, physiological or morphological states of individuals resulting from past trait expression and past interactions with their environment (since the rates of change in states depend on traits and on the environment). The third layer is that of evolving traits, which depend only on allelic values and summarize how these values are aggregated and expressed by individuals. In turn, allelic values are encoded by single alleles, and the set of alleles carried by individuals constitute their genotype. Because we concentrate on invasion analyses, we need not consider more than two alleles at a time, a mutant and a resident allele.

2.3 Local analysis and evolutionary dynamics

2.3.1 Directional and quadratic selection gradients

We now assume weak mutation effects and write the mutant's allelic value as a small perturbation $u_{j,l} = v_{j,l} + \epsilon\eta_{j,l}$ of the resident's allelic value, where $\eta_{j,l}$ is a trait-specific random mutation effect and $\epsilon \ll 1$ is a small parameter (assuming that mutations are such that $u_{j,l}$ always lies within the domain of $z_{j,l}$, for any trait l in any class j). The basic reproductive number of a mutant allele can then be Taylor expanded around $\epsilon = 0$ as

$$R_0(\boldsymbol{u}, \boldsymbol{v}) = 1 + \sum_{j=1}^{n_\mathrm{c}} \boldsymbol{\eta}_j \cdot \mathbf{s}_j(\boldsymbol{v})\epsilon + \frac{1}{2}\sum_{i=1}^{n_\mathrm{c}}\sum_{j=1}^{n_\mathrm{c}} \boldsymbol{\eta}_i^\intercal \mathbf{H}_{ij}(\boldsymbol{v})\boldsymbol{\eta}_j\epsilon^2 + O(\epsilon^3), \tag{3a}$$

where we used the fact that in a resident population at demographic equilibrium, $R_0(\boldsymbol{v}, \boldsymbol{v}) = 1$, and $O(\epsilon^3)$ is a remainder of order ϵ^3. Here, $\boldsymbol{\eta}_j = (\eta_j, ..., \eta_{j,n_\mathrm{z}})$ is the effect of the mutant in class j; $\mathbf{s}_j(\boldsymbol{v})$ is a $n_\mathrm{z} \times 1$ vector of directional selection coefficients, with element l given by

$$s_{j,l}(\boldsymbol{v}) = \left.\frac{\partial R_0(\boldsymbol{u}, \boldsymbol{v})}{\partial u_{j,l}}\right|_{\boldsymbol{u}=\boldsymbol{v}} = \left.\frac{\partial R_0(\boldsymbol{u}, \boldsymbol{v})}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}} \times \left.\frac{\partial z_{j,l}(u_{j,l}, v_{j,l})}{\partial u_{j,l}}\right|_{\boldsymbol{u}=\boldsymbol{v}}; \tag{3b}$$

$\mathbf{H}_{ij}(\boldsymbol{v})$ is a $n_\mathrm{z} \times n_\mathrm{z}$ matrix of quadratic selection coefficients with (k, l)-element

$$h_{ij,kl}(\boldsymbol{v}) = \left.\frac{\partial^2 R_0(\boldsymbol{u}, \boldsymbol{v})}{\partial u_{i,k}\partial u_{j,l}}\right|_{\boldsymbol{u}=\boldsymbol{v}} = \left.\frac{\partial^2 R_0(\boldsymbol{u}, \boldsymbol{v})}{\partial z_{i,k}\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}} \times \left.\frac{\partial z_{i,k}(u_{i,k}, v_{i,k})}{\partial u_{i,k}}\right|_{\boldsymbol{u}=\boldsymbol{v}} \times \left.\frac{\partial z_{j,l}(u_{j,l}, v_{j,l})}{\partial u_{j,l}}\right|_{\boldsymbol{u}=\boldsymbol{v}}, \tag{3c}$$

where we used the facts that $R_0(\boldsymbol{u}, \boldsymbol{v})$ implicitly depends on $\boldsymbol{z}(\boldsymbol{u}, \boldsymbol{v})$ and that $\boldsymbol{z}^\circ = \boldsymbol{v}$. Finally, '$\cdot$' denotes the dot product, and '$\intercal$' denotes transposition in eq. (3a).

The quantity $s_{j,l}(\boldsymbol{v})$ (eq. 3b) is the effect of a marginal change in the mutant allelic value $u_{j,l}$ on the mutant's reproductive number. It can be decomposed as the product of two components (right hand

side of eq. 3b): the marginal effect of $z_{j,l}$ on $R_0(\boldsymbol{u}, \boldsymbol{v})$, and the effect of a change in allelic value on trait expression. Because we assume additive gene action, this second term is a constant independent of trait values for a given class j: it gives the relative genetic contribution of a mutant allele to the individual's trait value (e.g., one-half under diploidy or one under haploidy). If $s_{j,l}(\boldsymbol{v}) > 0$, selection favours an increase in that trait; if $s_{j,l}(\boldsymbol{v}) < 0$, it favours a decrease (see tab. 2 for a summary of the types of selection considered throughout).

Selection	**Measure**	**Effect**
Directional	$s_{j,l}(\boldsymbol{v})$	Favours an increase ($s_{j,l}(\boldsymbol{v}) > 0$) or decrease ($s_{j,l}(\boldsymbol{v}) < 0$) in the genetic value for trait (j, l). These values are collected in the selection gradient vector $\mathbf{s}_j(\boldsymbol{v})$, whose l-th element is $s_{j,l}(\boldsymbol{v})$ and gives the directional selection in class j.
Quadratic	$h_{ij,kl}(\boldsymbol{v}^*)$	Favours an increase ($h_{ij,kl}(\boldsymbol{v}^*) > 0$) or decrease ($h_{ij,kl}(\boldsymbol{v}^*) < 0$) in the (co)variance between the genetic values for traits (i, k) and (j, l) at a singular strategy $\boldsymbol{v}^*$. These values are collected in the Hessian blocks $\mathbf{H}_{ij}(\boldsymbol{v}^*)$, whose (k, l)-element is $h_{ij,kl}(\boldsymbol{v}^*)$. The full Hessian matrix $\mathbf{H}(\boldsymbol{v}^*)$ consists of all such blocks.
Correlational	$h_{ij,kl}(\boldsymbol{v}^*)$, with $i \neq j$ or $k \neq l$	Favours an increase ($h_{ij,kl}(\boldsymbol{v}^*) > 0$) or decrease ($h_{ij,kl}(\boldsymbol{v}^*) < 0$) in the genetic covariance between different traits traits (i, k) and (j, l). This captures correlational selection across classes (if $i \neq j$) or between different traits within a class (if $i = j$ and $k \neq l$).
Stabilising	$h_{jj,ll}(\boldsymbol{v}^*) < 0$	Favours a decrease in the genetic variance of trait (j, l).
Disruptive	$h_{jj,ll}(\boldsymbol{v}^*) > 0$	Favours an increase in the genetic variance of trait (j, l).
Net stabilising	$\mathrm{eig}\,(\mathbf{H}(\boldsymbol{v}^*)) < 0$	Disfavours any mutant away from $\boldsymbol{v}^*$ when all traits jointly evolve – $\mathrm{eig}\,(\mathbf{H}(\boldsymbol{v}^*))$ refers to the leading eigenvalue of $\mathbf{H}(\boldsymbol{v}^*)$.
Net disruptive	$\mathrm{eig}\,(\mathbf{H}(\boldsymbol{v}^*)) > 0$	Favours divergence away from $\boldsymbol{v}^*$ when all traits evolve jointly. Divergence occurs along the eigenvector associated with the leading eigenvalue of $\mathbf{H}(\boldsymbol{v}^*)$.

Tab. 2. Summary of selection types. Note that a singular strategy $\boldsymbol{v}^*$ is defined here as a vector of trait values for which directional selection vanishes (see eq. 5 for independent traits, and eq. 21 for genetically correlated traits).

Meanwhile, the quantity $h_{ij,kl}(\boldsymbol{v})$ (eq. 3c) is proportional to the multiplicative effects of joint changes in $z_{i,k}$ and $z_{j,l}$ on the mutant's reproductive number, beyond their individual effects. When $i \neq j$ or $k \neq l$, it quantifies how a change in the first trait modulates the effect of a change in the second. In other words, $h_{ij,kl}(\boldsymbol{v})$ describes correlational selection on traits (i, k) and (j, l): positive values indicate that $R_0(\boldsymbol{u}, \boldsymbol{v})$ increases faster with one trait when the other is larger – meaning that quadratic selective effects favours similar changes in the two traits – whereas negative values indicate that $R_0(\boldsymbol{u}, \boldsymbol{v})$ increases faster with one trait when the other is smaller, favouring opposite changes. For $i = j$ and $k = l$, a negative value of $h_{ii,kk}(\boldsymbol{v})$ indicates that $R_0(\boldsymbol{u}, \boldsymbol{v})$ decelerates with increasing deviation from the resident trait along a single trait axis (i, k), while a positive value of $h_{ii,kk}(\boldsymbol{v})$ indicates that $R_0(\boldsymbol{u}, \boldsymbol{v})$ accelerates along that axis – meaning that quadratic selection favours trait values close or away from the current resident value,

respectively.

2.3.2 Mutation limited evolutionary dynamics

Directional and quadratic selection coefficients inform gradual trait evolution when mutations are rare and have weak effects on the traits. Specifically, when mutations have small effects (i.e., $\epsilon \ll 1$), the 'invasion implies substitution' principle for class-structured populations (Priklopil & Lehmann 2020) states that a mutant will increase (respectively, decrease) in frequency and eventually replace the resident (be lost) whenever $\sum_{j=1}^{n_c} \boldsymbol{\eta}_j \cdot \mathbf{s}_j(\boldsymbol{v})$ is positive (respectively, negative). This implies the existence of a mutation-limited regime in which evolution proceeds gradually, as a sequence of trait substitutions along the direction favoured by selection. This sequence halts only when the population reaches a phenotypic point $\boldsymbol{v}$ where

$$\sum_{j=1}^{n_c} \boldsymbol{\eta}_j \cdot \mathbf{s}_j(\boldsymbol{v}) \leq 0 \tag{4}$$

for all mutations $\boldsymbol{\eta} = (\boldsymbol{\eta}_1, ..., \boldsymbol{\eta}_{n_c})$ in a close neighbourhood of $\boldsymbol{v}$. When the trait space is bounded and $\boldsymbol{v}$ is at the boundary, new mutations can only move traits inward or along the boundary. Condition (4) can then be satisfied if directional selection opposes all admissible mutations (i.e. or equivalently if selection favours trait values outside of the authorized range, such that corresponding elements of $\boldsymbol{\eta}_j$ and $\mathbf{s}_j(\boldsymbol{v})$ always have opposing sign). In what follows, we focus on cases where condition (4) is met in the interior of trait space.

When mutation effects on trait values are not exactly equal across traits – that is when the distribution of mutation effect is not confined to a lower-dimensional subspace of trait space (e.g. $\boldsymbol{\eta}$ is sampled from a multi-normal distribution over $n_c \times n_z$ with full-rank covariance matrix) – there always exists some mutation that affects one trait without changing the others. Therefore, on the long term and unless mutation effects are fully correlated across traits, traits can be thought of as being unconstrained genetically. In this case, condition (4) is satisfied in the interior of trait space if, for some trait value vector $\boldsymbol{v}^*$,

$$s_{j,l}(\boldsymbol{v}^*) = 0 \quad \text{for each trait } (j, l). \tag{5}$$

Any $\boldsymbol{v}^*$ satisfying eq. (5) will be called a singular trait value and constitutes a candidate evolutionary equilibrium. But will evolution converge to such a singular trait?

A sufficient condition for $\boldsymbol{v}^*$ to act as a local attractor of the evolutionary dynamics – i.e. for the population to evolve toward $\boldsymbol{v}^*$ via a trait substitution sequence – is that the Jacobian matrix $\mathbf{J}(\boldsymbol{v}^*)$ is negative definite (Leimar 2009, p. 189). This matrix is a $n_c \times n_c$ block matrix whose block (i, j) is a

matrix $\mathbf{J}_{ij}(\boldsymbol{v}^*)$ with element (k,l) given by

$$j_{ij,kl} = \left.\frac{\partial s_{i,k}(\boldsymbol{v})}{\partial v_{j,l}}\right|_{\boldsymbol{v}=\boldsymbol{v}^*}. \tag{6}$$

The matrix $\mathbf{J}(\boldsymbol{v}^*)$ is negative definite if the leading eigenvalue of its symmetric part, $(\mathbf{J}(\boldsymbol{v}^*)+\mathbf{J}^\intercal(\boldsymbol{v}^*))/2$, is negative. In that case, $\boldsymbol{v}^*$ is said to be strongly convergence stable (*sensu* Leimar 2009). The term 'strongly' reflects that convergence occurs from all nearby trait values, regardless of the mutational covariance structure. In contrast, when $\mathbf{J}(\boldsymbol{v}^*)$ is not negative semi-definite, there is some mutational covariance structure for which $\boldsymbol{v}^*$ is a repellor of the evolutionary dynamics (Leimar 2009, p. 198).

By eq. (3a), if the population has reached a singular trait value $\boldsymbol{v}^*$, then a sufficient condition for local uninvadability, i.e. for any nearby mutant to be selected against and thus for selection to be stabilising at $\boldsymbol{v}^*$, is that

$$\sum_{i=1}^{n_c}\sum_{j=1}^{n_c}\boldsymbol{\eta}_i^\intercal\mathbf{H}_{ij}(\boldsymbol{v}^*)\boldsymbol{\eta}_j < 0 \quad \text{for all mutations } \boldsymbol{\eta}=(\boldsymbol{\eta}_1,...,\boldsymbol{\eta}_{n_c}). \tag{7}$$

Eq. (7) holds if and only if the block matrix $\mathbf{H}(\boldsymbol{v}^*)$, whose block (i,j) is $\mathbf{H}_{ij}(\boldsymbol{v}^*)$, is negative definite (i.e. $\boldsymbol{\eta}^\intercal\mathbf{H}(\boldsymbol{v}^*)\boldsymbol{\eta}<0$ for all $\boldsymbol{\eta}$). Since $\mathbf{H}(\boldsymbol{v}^*)$ is a symmetric real matrix, it is negative definite if and only if all its eigenvalues are negative.

When $\mathbf{H}(\boldsymbol{v}^*)$ has at least one positive eigenvalue (i.e. $\mathbf{H}(\boldsymbol{v}^*)$ is not negative semi-definite), there exists a set of mutation directions – specifically those aligned with the leading eigenvector of $\mathbf{H}(\boldsymbol{v}^*)$ – that are favoured on both sides of the singular strategy $\boldsymbol{v}^*$. In other words, selection is disruptive at $\boldsymbol{v}^*$: it favours divergence away from the resident trait value. A sufficient condition for $\mathbf{H}(\boldsymbol{v}^*)$ to not be negative semi-definite is that any diagonal block $\mathbf{H}_{jj}(\boldsymbol{v}^*)$ is not negative semi-definite (since diagonal blocks are symmetric matrices, Kierzkowski & Smoktunowicz 2011, Theorem 2.2, p. 1062). In turn, a diagonal block $\mathbf{H}_{jj}(\boldsymbol{v}^*)$ is not negative semi-definite if any of its diagonal elements is positive, i.e., if $h_{jj,kk}(\boldsymbol{v}^*)>0$ for some trait (j,k). At a singular strategy, the quantity $h_{jj,kk}(\boldsymbol{v}^*)$ captures whether selection on trait (j,k) is locally stabilising (if negative) or disruptive (if positive) when that trait evolves in isolation from all others. More generally, when $\mathbf{H}(\boldsymbol{v}^*)$ is not negative semi-definite and $\boldsymbol{v}^*$ is strongly convergence stable, selection may lead to evolutionary branching, whereby two alleles with slightly different trait values near $\boldsymbol{v}^*$ are both maintained and subsequently diverge as they accumulate further mutations (the exact conditions for this to happen when multiple traits co-evolve remains an open question, Geritz et al. 2016).

Eqs. (3a)–(7) show that under the assumption of rare mutations with small phenotypic effects, key features of the gradual long-term evolution of traits can be predicted from the computation of the selection gradients and the Jacobian and Hessian matrices. Eqs. (5)–(7), however, assume that mutation effects

on trait values are not exactly equal across traits, thereby excluding cases where genetic constraints force traits to be expressed identically across classes. We will later visit such genetically constrained trait expression, which can in fact be analysed using the same results as above, with only minimal adjustments (see appendix A.6 for details). In the following, we give explicit expressions for the directional and quadratic selection coefficients.

3 Analysis of selection

3.1 Directional selection

3.1.1 Age-specific coefficient of directional selection

We show in appendix A.2 that the coefficient of directional selection on trait (j,l) can be written as

$$s_{j,l}(\boldsymbol{v}) = \left[\int_0^\infty \hat{s}_{j,l}(a,\boldsymbol{v})\mathrm{d}a\right] \times \left.\frac{\partial z_{j,l}(u_{j,l},v_{j,l})}{\partial u_{j,l}}\right|_{\boldsymbol{u}=\boldsymbol{v}}, \tag{8}$$

where

$$\hat{s}_{j,l}(a,\boldsymbol{v}) = \sum_{i=1}^{n_c} \nu_i^\circ \gamma_{ij} \left.\frac{\partial H_{ij}(\boldsymbol{z}_j, l_j^\circ(a), \boldsymbol{x}_j^\circ(a), \boldsymbol{v})}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}} q_j^\circ \tag{9}$$

is an age a specific coefficient of directional selection. This coefficient involves three quantities. (i) The leading right-eigenvector $\boldsymbol{q}^\circ$ of the resident next-generation matrix $\mathbf{R}^\circ = \mathbf{R}(\boldsymbol{v},\boldsymbol{v})$, such that q_j° gives the frequency of alleles carried by class-j individuals among resident newborns. (ii) The leading left-eigenvector $\boldsymbol{\nu}^\circ$ of $\mathbf{R}^\circ$, such that ν_i° is the reproductive value of an allele found in a resident newborn of class i (i.e. the expected number of copies this allele will produce over the lifetime of its carrier). We refer to $\boldsymbol{q}^\circ$ and $\boldsymbol{\nu}^\circ$ as allelic class frequencies and reproductive values at birth, respectively. Finally, (iii) the marginal effect $\partial H_{ij}/\partial z_{j,l}$ of variation in trait expression on the so-called Hamiltonian H_{ij}, which we detail below in eq. (10). For now simply note that this derivative measures how a change in trait (j,l) affects the production of class-i offspring by a class-j individual: directly, through a change in fecundity at age a; and indirectly through changes in survival and development at that same age (i.e. influencing future reproduction).

With these definitions in mind, $\hat{s}_{j,l}(a,\boldsymbol{v})$ in eq. (9) can be read from right to left as follows. First, q_j° gives the probability that a mutant allele occurs in a newborn of class j, so that a variation in trait (j,l) will be expressed in such an individual and affect its life history across all ages a. Second, $\partial H_{ij}/\partial z_{j,l}$

quantifies how such age-specific variation affects the production of class-i offspring, accounting for both current and future contributions, i.e. at age a and later. Third, γ_{ij} converts offspring into transmitted mutant allele copies, and finally ν_i° weights each such descendant copy by its reproductive value in class i. Summing over all offspring classes i gives the total fitness consequences of variation in trait (j, l) via its effects on fecundity, survival and development at age a. Directional selection on trait (j, l) is finally obtained by integrating these age-specific coefficients over all ages (eq. 8).

3.1.2 The Hamiltonian: partitioning age-specific fitness effects

Quantifying the fitness effects at a specific age relies on the Hamiltonian:

$$\begin{aligned} H_{ij}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j(a), \boldsymbol{v}) = l_j(a) f_{ij}(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v}) - \lambda^{\mathrm{l}}_{ij}(a)\mu_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v}) l_j(a) \\ + \boldsymbol{\lambda}^{\mathrm{x}}_{ij}(a) \cdot \boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v}), \end{aligned} \tag{10}$$

where $\lambda^{\mathrm{l}}_{ij}(a)$ and $\boldsymbol{\lambda}^{\mathrm{x}}_{ij}(a) = (\lambda^{\mathrm{x}}_{ij,1}(a), ..., \lambda^{\mathrm{x}}_{ij,n_{\mathrm{x}}}(a))$ are so-called costate variables that are given by the solution to the system of ODEs:

$$\begin{aligned} \frac{\mathrm{d}\lambda^{\mathrm{l}}_{ij}(a)}{\mathrm{d}a} &= -\left.\frac{\partial H_{ij}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j^\circ(a), \boldsymbol{v})}{\partial l_j(a)}\right|_{\boldsymbol{z}=\boldsymbol{v}} \quad \text{with i.c.} \quad \lambda^{\mathrm{l}}_{ij}(0) = R^\circ_{ij}, \\ \frac{\mathrm{d}\lambda^{\mathrm{x}}_{ij,m}(a)}{\mathrm{d}a} &= -\left.\frac{\partial H_{ij,m}(\boldsymbol{z}_j, l_j^\circ(a), \boldsymbol{x}_j(a), \boldsymbol{v})}{\partial x_{j,m}(a)}\right|_{\boldsymbol{z}=\boldsymbol{v}} \quad \text{with f.c.} \quad \lim_{a\to\infty} \lambda^{\mathrm{x}}_{ij,m}(a) = 0. \end{aligned} \tag{11}$$

The costates $\lambda^{\mathrm{l}}_{ij}(a)$ and $\lambda^{\mathrm{x}}_{ij,m}(a)$ respectively measure how marginal changes in survival and in internal state m at age a affect the expected remaining reproductive output of a class-j individual via offspring of class i. In particular $\lambda^{\mathrm{l}}_{ij}(a)$ is a class-oriented equivalent of Fisher's definition of the reproductive value of an individual of age a (see box 1). The boundary conditions in eq. (11) are obtained as follows. At birth, and provided fecundity is finite, we have $\lambda^{\mathrm{l}}_{ij}(0) = R^\circ_{ij}$; namely, the resident expected lifetime class-oriented fecundity (e.g. in a haploid population without class structure $\lambda^{\mathrm{l}}_{ij}(0) = 1$ owing to demographic consistency; see appendix A.2.3). Biologically, this reflects that death at birth would eliminate the entire reproductive potential of the individual. By contrast, the development costates $\lambda^{\mathrm{x}}_{ij,m}(a)$ vanish as $a \to \infty$, under the assumption that traits do not affect internal states at birth but do affect their finite terminal values (see appendix A.2.3). Intuitively, this is because at very old ages the expected remaining reproductive output approaches zero, such that marginal changes in state values no longer influence fitness.

The Hamiltonian eq. (10) quantifies how a class-j individual, at age a, contributes to its lifetime production of class-i offspring through three distinct components: (1) current offspring production through

Consider the remaining number of offspring of class i that a mutant individual of class j and age a is expected to produce until its death,

$$\hat{R}_{ij}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j(a), \boldsymbol{v}, a) = \int_a^\infty l_j(\tau) f_{ij}(\boldsymbol{z}_j, \boldsymbol{x}_j(\tau), \boldsymbol{v})\, \mathrm{d}\tau, \tag{I-A}$$

such that $\hat{R}_{ij}(\boldsymbol{z}_j, l_j(0), \boldsymbol{x}_j(0), \boldsymbol{v}, 0) = R_{ij}(\boldsymbol{u}_j, \boldsymbol{v})$ (see eq. 1). Using results from Avila et al. (2021) (their appendix B.1), we obtain that the co-states $\lambda^{\mathrm{l}}_{ij}(a)$ and $\boldsymbol{\lambda}^{\mathrm{x}}_{ij}(a)$ satisfy

$$\lambda^{\mathrm{l}}_{ij}(a) = \left.\frac{\partial \hat{R}_{ij}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}^\circ_j(a), \boldsymbol{v}, a)}{\partial l_j(a)}\right|_{\boldsymbol{z}=\boldsymbol{v}} \quad \text{and} \quad \boldsymbol{\lambda}^{\mathrm{x}}_{ij}(a) = \left.\frac{\partial \hat{R}_{ij}(\boldsymbol{z}_j, l^\circ_j(a), \boldsymbol{x}_j(a), \boldsymbol{v}, a)}{\partial \boldsymbol{x}_j(a)}\right|_{\boldsymbol{z}=\boldsymbol{v}}. \tag{I-B}$$

Costates can therefore be interpreted as the change in the expected remaining class-oriented reproductive output of an individual of class j and age a, resulting from a change in its current survivorship or internal states values. The costate associated to survivorship, in particular, can be expressed as

$$\lambda^{\mathrm{l}}_{ij}(a) = \frac{1}{l^\circ_j(a)} \hat{R}_{ij}(\boldsymbol{v}_j, l^\circ_j(a), \boldsymbol{x}^\circ_j(a), \boldsymbol{v}, a). \tag{I-C}$$

We see from eq. (I-C) that $\lambda^{\mathrm{l}}_{ij}(a)$ is the expected remaining class-oriented reproductive output of a resident individual, conditional on this individual having survived to age a. Thus, $\lambda^{\mathrm{l}}_{ij}(a)$ is the class-oriented equivalent of Fisher's reproductive value (Fisher 1930; Goodman 1982; Charlesworth 1994). Note that, since according to eq. (I-C) we have $\lambda^{\mathrm{l}}_{ij}(0) = \hat{R}_{ij}(\boldsymbol{v}_j, l^\circ_j(0), \boldsymbol{x}^\circ_j(0), \boldsymbol{v}, 0) = R_{ij}(\boldsymbol{v}_j, \boldsymbol{v})$, we obtain from the definition of the left-eigenvector $\boldsymbol{\nu}^\circ$ of $\mathbf{R}^\circ$ (eq. A-3) that

$$\nu^\circ_j = \sum_{i=1}^{n_{\mathrm{c}}} \nu^\circ_i \gamma_{ij} \lambda^{\mathrm{l}}_{ij}(0). \tag{I-D}$$

This connects costates to the elements ν°_j of the leading left-eigenvector of the next-generation matrix (i.e. allelic reproductive values at birth). While costates are age-specific measures of future reproductive output at the individual level, reproductive values in $\boldsymbol{\nu}$ are lifetime measures at the allelic level that take into account the value of each type of descendants (i.e. ν°_i in eq. I-D).

Box 1: Costates and reproductive values in age- and class-structured populations.

fecundity f_{ij}; (2) future reproductive output depending on mortality rate μ_j at that age, weighted by $\lambda^{\mathrm{l}}_{ij}(a)$; and (3) future reproductive output depending on development rate $\boldsymbol{g}_j$ at that age, weighted by $\boldsymbol{\lambda}^{\mathrm{x}}_{ij}(a)$. This decomposition reflects the fact that at any given age, an individual can influence its eventual production of class-i offspring not only through its immediate fecundity, but also through its current survival and developmental rates, which determine its future reproduction. The Hamiltonian can thus be interpreted as an age-specific measure of reproductive success—one that, when integrated over age (as in eq. 8), allows us to quantify the total effect of a trait change on an individual's lifetime reproductive success.

An important practical point is that in the coefficient of directional selection (eq. 9) the derivative of the Hamiltonian is evaluated at resident trait values. As a result, when computing the Hamiltonian and solving the ODEs for the costates (eq. 11), we only need to consider resident survivorship and state

dynamics. Concretely, it is sufficient to solve

$$\begin{aligned}\frac{\mathrm{d}l_j^\circ(a)}{\mathrm{d}a} &= -\mu_j(\boldsymbol{v}_j, \boldsymbol{x}_j^\circ(a), \boldsymbol{v})\; l_j^\circ(a) \quad \text{with i.c.} \quad l_j^\circ(0) = 1,\\ \frac{\mathrm{d}\boldsymbol{x}_j^\circ(a)}{\mathrm{d}a} &= \boldsymbol{g}_j(\boldsymbol{v}_j, \boldsymbol{x}_j^\circ(a), \boldsymbol{v}) \quad \text{with i.c.} \quad \boldsymbol{x}_j^\circ(0) = \boldsymbol{x}_{j,0}.\end{aligned} \tag{12}$$

rather than the mutant system in eq. (2). This can greatly simplify the analysis as we will illustrate later in our example.

3.1.3 Life history trade-offs in structured population

The age-specific coefficient of directional selection (eq. 9) is expressed in terms of the effect of a trait change on the Hamiltonian, which in turn separates the relevant fitness components at each age. To gain further insights, one can substitute eq. (10) into eq. (9), giving

$$\begin{aligned}\hat{s}_{j,l}(a, \boldsymbol{v}) = \sum_{i=1}^{n_\mathrm{c}} \nu_i^\circ \gamma_{ij} \Bigg(& l_j^\circ(a) \left.\frac{\partial f_{ij}(\boldsymbol{z}_j, \boldsymbol{x}_j^\circ(a), \boldsymbol{v})}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}} \\ & - \lambda_{ij}^{\mathrm{l}}(a) l_j^\circ(a) \left.\frac{\partial \mu_j(\boldsymbol{z}_j, \boldsymbol{x}_j^\circ(a), \boldsymbol{v})}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}} + \boldsymbol{\lambda}_{ij}^{\mathrm{x}}(a) \cdot \left.\frac{\partial \boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j^\circ(a), \boldsymbol{v})}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}} \Bigg) q_j^\circ.\end{aligned} \tag{13}$$

Eq. (13) highlights the fundamental trade-off between current and future reproduction, while accounting for age, class structure, ontogeny and ploidy. This is made explicit by showing that directional selection boils down to a weighted effect of a trait change on the three core life-history components: fecundity, mortality, and development. The weights q_j°, ν_i° and γ_{ij} capture the relative importance of different parent and offspring classes, $l_j^\circ(a)$ weights age-specific effects by the probability of surviving to age a, and the costates $\lambda_{ij}^{\mathrm{l}}(a)$ and $\boldsymbol{\lambda}_{ij}^{\mathrm{x}}(a)$ measure how a trait change at age a influences future reproduction through survival and development. Taken together, these weights quantify how selection balances all the potential first-order effects of trait expression. They make it possible to distinguish three types of trade-offs: (i) within-age trade-offs, between the contributions to fecundity, survival or development at a given age (the term within parentheses in eq. 13); (ii) among-class trade-offs, reflecting the relative value of producing offspring of different classes (arising from the summation over classes in eq. 13); and (iii) among-age trade-offs, arising when eq. (13) is integrated over ages in eq. (8), and which depend on how survivorship $l_j^\circ(a)$ declines with age and on how the costates $\lambda_{ij}^{\mathrm{l}}(a)$ and $\boldsymbol{\lambda}_{ij}^{\mathrm{x}}(a)$ vary with age—either decreasing or increasing depending on developmental effects and life history.

3.1.4 Connections with previous expressions for directional selection

Eq. (13) connects to several results from the literature. First, in the absence of ontogeny, it reduces to the standard selection gradient in class-structured populations (Taylor 1990; Taylor & Frank 1996; Caswell 2000; Rousset 2004). Second, in the absence of class structure, it reduces to the selection gradient on constant traits (controls) affecting life history (Avila et al. 2021; Lehmann 2022; Avila & Mullon 2023) and extend it to arbitrary ploidy. Third, the term within parenthesis in eq. (13) is conceptually similar to expressions for the fitness effects of age-specific trait expression under complete plasticity (usually derived in the absence of class structure, e.g., León 1976; Goodman 1982; Schaffer 1983; Iwasa & Roughgarden 1984; Perrin & Sibly 1993; Parvinen et al. 2013; Metz et al. 2016; Houston et al. 2024, but see Avila et al. 2019).

Finally, when mutations act only on mortality at age a, eq. (13) reduces to Hamilton's first indicator of the strength of selection (Hamilton 1966, eq. 9, p. 17; see also Wensink et al. 2017), with selection being proportional to $l_j^\circ(a)\lambda_{jj}^1(a) = \int_a^\infty l_j^\circ(\tau)\, f_{jj}(\boldsymbol{v}_j, \boldsymbol{x}_j^\circ(\tau), \boldsymbol{v})\, \mathrm{d}\tau$, i.e. to the expected remaining reproductive output (eq. I-C); while for mutations acting only on fecundity at age a, eq. (13) reduces to Hamilton's second indicator i.e. survivorship to age a, $l_j^\circ(a)$. Eq. (13), however, shows that selection generally acts not only through mortality and fecundity, but through development as well, taking into account the specificities of those processes within each class. A relevant implication of including development, in particular, is that it can significantly alter among-age trade-offs between life-history components. Indeed, developmental effects can be strongly non-monotonic with respect to age (e.g. during metamorphosis), and as a result the costates associated with internal states do not necessarily decline with age. Consequently, the age-specific strength of selection also need not decrease monotonically with age (Vaupel et al. 2004; Monaghan et al. 2008), contrary to a common interpretation of Hamilton's seminal results (Hamilton 1966). We illustrate this in our example model (Section 4).

3.2 Quadratic selection

Once the population expresses a singular trait value $\boldsymbol{v}^*$ where directional selection vanishes (i.e. such that eq. 5 holds), the nature of selection is given by the class-oriented Hessian matrices $\mathbf{H}_{ij}(\boldsymbol{v}^*)$ (eq. 3c). We show in appendix A.3 that the (k,l)-element of these matrices can be expressed as

$$h_{ij,kl}(\boldsymbol{v}^*) = \left[\int_0^\infty \hat{h}_{ij,kl}(a,\boldsymbol{v}^*)\mathrm{d}a\right] \times \left.\frac{\partial z_{i,k}(u_{i,k}, v_{i,k})}{\partial u_{i,k}}\right|_{\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{v}^*} \times \left.\frac{\partial z_{j,l}(u_{j,l}, v_{j,l})}{\partial u_{j,l}}\right|_{\boldsymbol{u}=\boldsymbol{v}=\boldsymbol{v}^*}, \tag{14}$$

where $\hat{h}_{ij,kl}(a, \boldsymbol{v}^*)$ is an age-specific coefficient of quadratic selection, measuring the effects of a joint change in traits (i, k) and (j, l) on fitness through their influence on life-history components at age a (i.e. fecundity, survival, and development), that can be decomposed as

$$\hat{h}_{ij,kl}(a, \boldsymbol{v}^*) = \begin{cases} \hat{h}^{\mathrm{zz}}_{jj,kl}(a, \boldsymbol{v}^*) + \hat{h}^{\mathrm{xx}}_{jj,kl}(a, \boldsymbol{v}^*) + \hat{h}^{\mathrm{zx}}_{jj,kl}(a, \boldsymbol{v}^*) & \\ \qquad + \hat{h}^{\mathrm{zl}}_{jj,kl}(a, \boldsymbol{v}^*) + \hat{h}^{\mathrm{lx}}_{jj,kl}(a, \boldsymbol{v}^*) + \hat{h}^{\mathrm{zq}}_{jj,kl}(a, \boldsymbol{v}^*) & \text{if } i = j \\ \hat{h}^{\mathrm{zq}}_{ij,kl}(a, \boldsymbol{v}^*) & \text{if } i \neq j. \end{cases} \tag{15}$$

Each term captures a relevant age-specific effect that can generate correlational selection among traits (i, k) and (j, l) (or disruptive selection on a trait when $i = j$ and $k = l$) that we detail below.

3.2.1 Age-specific direct trait effects

The first element in the right-hand side of eq. (15) with $i = j$ is given by

$$\hat{h}^{\mathrm{zz}}_{jj,kl}(a, \boldsymbol{v}^*) = \sum_{m=1}^{n_{\mathrm{c}}} \nu^{\circ}_m \gamma_{mj} \left. \frac{\partial^2 H_{mj}(\boldsymbol{z}_j, l^{\circ}_j(a), \boldsymbol{x}^{\circ}_j(a), \boldsymbol{v})}{\partial z_{j,k} \partial z_{j,l}} \right|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} q^{\circ}_j. \tag{16a}$$

Eq. (16a) describes the direct joint effects of traits (j, k) and (j, l) on the Hamiltonian H_{mj} (eq. 10), summed over all offspring classes m with appropriate weights. Biologically, the derivative in eq. (16a) quantifies whether simultaneous changes in the two traits, expressed in the same individual of class j, increase or decrease its production of class-m offspring. As with directional selection (eq. 13), these fitness effects can manifest through fecundity at age a or through survival and development at that same age, which influence future reproduction.

Eq. (16a) can thus be viewed as the direct extension of the classical measure of correlational selection (i.e. described in the absence of any population structure, Phillips & Arnold 1989). In that setting, correlational selection is due to multiplicative effects of traits on reproduction only. By contrast, correlational selection here can also arise because traits have multiplicative effects on survival rates (i.e. on $\mu_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$) or on development rates (i.e. on $\boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$) at age a. For instance, if two traits interact to determine the growth rate – say one trait sets foraging effort, which determines resource acquisition, and the other governs how much of these resources are allocated to growth – and if larger size later increases reproduction, eq. (16a) indicates that selection will favour a positive association between these two traits.

These different paths, which become apparent by substituting eq. (10) into eq. (16a) (though we refrain from doing so explicitly for succinctness), show that population structure and ontogeny create additional

opportunities for correlational selection, and thus for the emergence of polymorphism via disruptive selection. As we will now see, there are in fact many other paths for correlational selection generated by ontogeny and class structure.

3.2.2 Age-specific indirect effects

The next element in the right-hand side of eq. (15), which for $i = j$ is given by

$$\hat{h}^{\mathrm{xx}}_{jj,kl}(a, \boldsymbol{v}^*) = \sum_{m=1}^{n_{\mathrm{c}}} \nu^{\circ}_m \gamma_{mj} \left[\sum_{k'=1}^{n_{\mathrm{x}}} \sum_{l'=1}^{n_{\mathrm{x}}} \left(\frac{\partial^2 H_{mj}(\boldsymbol{z}_j, l^{\circ}_j(a), \boldsymbol{x}_j(a), \boldsymbol{v})}{\partial x_{j,k'}(a) \partial x_{j,l'}(a)} \frac{\partial x_{j,k'}(a)}{\partial z_{j,k}} \frac{\partial x_{j,l'}(a)}{\partial z_{j,l}} \right) \bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \right] q^{\circ}_j, \quad (16b)$$

describes the effect of joint change in all pairs of internal states at age a ($x_{j,k'}(a)$ and $x_{j,l'}(a)$), that are induced by changes in traits (j, k) and (j, l). Put differently, this term captures how simultaneous changes in two traits alter state variables at age a in ways that synergistically influence the different components of life history—fecundity, survival, or development—where fecundity affects reproduction at that age, while survival and development determine future reproduction (as per eq. 10). As an example where eq. (16b) may be relevant, consider a scenario where one trait regulates allocation to body growth while another controls allocation to sexual maturation. These respectively determine the internal state variables 'size' and 'maturity' at a given age. When particular combinations of these internal states yield higher fitness than intermediate ones (e.g. small but early-maturing individuals, or large but late-maturing ones, as observed in Atlantic salmon, Lepais et al. 2017), eq. (16b) will be positive, indicating that selection favours an association between the two traits that induce these internal state combinations.

Correlational selection may also be due to the third element in the right-hand side eq. (15), which for $i = j$ reads

$$\begin{aligned} \hat{h}^{\mathrm{zx}}_{jj,kl}(a, \boldsymbol{v}^*) = \sum_{m=1}^{n_{\mathrm{c}}} \nu^{\circ}_m \gamma_{mj} \Bigg[& \sum_{k'=1}^{n_{\mathrm{x}}} \left(\frac{\partial^2 H_{mj}(\boldsymbol{z}_j, l^{\circ}_j(a), \boldsymbol{x}_j(a), \boldsymbol{v})}{\partial z_{j,k} \partial x_{j,k'}(a)} \frac{\partial x_{j,k'}(a)}{\partial z_{j,l}} \right) \bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \\ & + \sum_{l'=1}^{n_{\mathrm{x}}} \left(\frac{\partial^2 H_{mj}(\boldsymbol{z}_j, l^{\circ}_j(a), \boldsymbol{x}_j(a), \boldsymbol{v})}{\partial z_{j,l} \partial x_{j,l'}(a)} \frac{\partial x_{j,l'}(a)}{\partial z_{j,k}} \right) \bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \Bigg] q^{\circ}_j, \end{aligned} \quad (16c)$$

and which describes the interaction effects between traits and internal states, i.e. it captures situations where the state-mediated influence of one trait on fitness contributions at age a depends on another trait. More specifically, eq. (16c) shows that correlational selection associates two traits whenever some internal state varies with the value of one trait, while the effect of that internal state on the Hamiltonian depends on the other trait.

For example, suppose the first trait determines the growth and thus morphology of some sexual traits (e.g. genital morphology) and the second determines the preference for mates with a compatible morphology. In this case, correlational selection arises because the benefit of expressing a given reproductive morphology is conditional on simultaneously expressing a preference for mates with the matching morphology. Conversely, the benefit of the preference trait depends on having the compatible morphology. Eq. (16c) captures this interdependence. Similar mechanisms apply whenever behaviours or other traits have state-dependent effects, such as risk-taking behaviours that depend on body condition, or condition-dependent costs of sexual traits as in handicap models.

Another interaction with traits that can contribute to correlational selection is given by the fourth element in the right-hand side of eq. (15), which for $i = j$ is

$$\begin{aligned}\hat{h}^{\mathrm{zl}}_{jj,kl}(a,\boldsymbol{v}^*) = \sum_{m=1}^{n_{\mathrm{c}}} \nu^{\circ}_m \gamma_{mj} \Bigg[& \left(\frac{\partial^2 H_{mj}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}^{\circ}_j(a), \boldsymbol{v})}{\partial z_{j,k} \partial l_j(a)} \frac{\partial l_j(a)}{\partial z_{j,l}} \right)\Bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \\ & + \left(\frac{\partial^2 H_{mj}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}^{\circ}_j(a), \boldsymbol{v})}{\partial z_{j,l} \partial l_j(a)} \frac{\partial l_j(a)}{\partial z_{j,k}} \right)\Bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \Bigg] q^{\circ}_j .\end{aligned} \tag{16d}$$

This term describes the interaction effects between traits and survivorship to age a. In biological terms, correlational selection arises because of this term whenever one trait modifies survivorship to age a, while another affects fecundity or mortality at that age (but not growth, since growth rates do not depend on survival, see eq. 10).

As an illustration, consider a trait that improves survival to later ages, and another trait that enhances reproductive success specifically at those later ages (e.g. investment in traits expressed after maturity, such as secondary sexual displays). In this case, the benefit of the survival trait is conditional on the reproductive trait, and, conversely, the benefit of the reproductive trait depends on surviving long enough to express it. Such an interaction is captured by eq. (16d), with correlational selection favouring an association between longevity and late-life reproductive investment, and disruptive selection potentially arising if intermediate strategies perform poorly compared to the extremes.

Interaction effects that are entirely mediated by life history can also generate correlational selection. This occurs when trait changes influence both internal states at a certain age and survivorship to that same age, as long as the Hamiltonian depends multiplicatively on these two components. This case is captured

by the fifth element in the right-hand side of eq. (15) that for $i = j$ reads

$$\hat{h}^{\mathrm{lx}}_{jj,kl}(a,\boldsymbol{v}^*) = \sum_{m=1}^{n_{\mathrm{c}}} \nu^{\circ}_m \gamma_{mj} \left[\sum_{k'=1}^{n_{\mathrm{x}}} \left(\frac{\partial l_j(a)}{\partial z_{j,k}} \frac{\partial x_{j,k'}(a)}{\partial z_{j,l}} + \frac{\partial l_j(a)}{\partial z_{j,l}} \frac{\partial x_{j,k'}(a)}{\partial z_{j,k}} \right)\Bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \right.$$
$$\left. \times \frac{\partial^2 H_{mn}(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j(a), \boldsymbol{v})}{\partial l_j(a) \partial x_{j,k'}(a)}\Bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \right] q^{\circ}_j. \tag{16e}$$

Eq. (16e) shows that correlational selection arises whenever an internal state at age a affects the fecundity or mortality rates (since both rates are multiplied by survivorship in eq. 10), provided traits influence both survivorship and the internal state.

For example, one trait may enhance survival through investment in immune function while another governs allocation to sexual development. The benefit of improved immunity is conditional on surviving long enough to express the sexual state, while the benefit of sexual development depends on immunity ensuring survival. Eq. (16e) would therefore be positive in this case, indicating that correlational selection favours a positive associations between these two traits through their joint effects on survivorship and development.

The final element in the right-hand side of eq. (15) is

$$\hat{h}^{\mathrm{zq}}_{ij,kl}(a,\boldsymbol{v}^*) = \sum_{m=1}^{n_{\mathrm{c}}} \nu^{\circ}_m \left[\gamma_{mi} \left(\frac{\partial H_{mi}(\boldsymbol{z}_i, l^{\circ}_i(a), \boldsymbol{x}^{\circ}_i(a), \boldsymbol{v})}{\partial z_{i,k}} \frac{\partial q_i(\boldsymbol{z}, \boldsymbol{v})}{\partial z_{j,l}} \right)\Bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \right.$$
$$\left. + \gamma_{mj} \left(\frac{\partial H_{mj}(\boldsymbol{z}_j, l^{\circ}_j(a), \boldsymbol{x}^{\circ}_j(a), \boldsymbol{v})}{\partial z_{j,l}} \frac{\partial q_j(\boldsymbol{z}, \boldsymbol{v})}{\partial z_{i,k}} \right)\Bigg|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \right], \tag{16f}$$

which consists of the product between the effect of one trait on the Hamiltonians for a given parental class and the effect of the other trait on the probability that mutant alleles are found in newborns of that parental class (i.e. on mutant allelic frequencies at birth $\boldsymbol{q}(\boldsymbol{z}, \boldsymbol{v})$). In essence, this means that correlational selection will favour an association between two traits when one increases the likelihood of being in a certain class, while the other increases reproductive success in that class. For example, one trait may affect the development of forms that are particularly suited to reproduction as a female but not as a male, while another biases the sex ratio at birth so that a larger fraction of offspring are female. In this case, eq. (16f) would be positive, indicating that correlational selection favours an association between these two traits. This is because such an association leads genotypes that produce better female forms to be more often expressed in females, while those producing better male forms are more often expressed in males.

In contrast to the previous components (eqs. 16a–16e), this term $\hat{h}^{\mathrm{zq}}_{ij,kl}$ contributes to correlational selec-

tion even when $i \neq j$ (see eq. 15). The reason is that a change in a trait expressed in one class j can modify the allelic class frequencies at birth in another class $q_j(\boldsymbol{z}, \boldsymbol{v})$ with $i \neq j$. For example, if sex ratio is under paternal control, then a male-expressed trait that biases the proportion of daughters directly affects the allelic frequency of the female class at birth.

Taken together, the different parts of eq. (15) show that correlational selection can arise not only through direct interactions among traits (eq. 16a), but also indirectly via their effects on internal states, survival, or class representation at birth (eqs. 16b–16f), thereby encompassing all pathways by which trait associations influence fitness in class-structured populations with ontogeny.

3.2.3 Trait effects on internal states, survivorship and class structure

The indirect pathways through which correlational selection acts (eqs. 16b–16f) ultimately depend on how traits affect internal states, survivorship, and class structure. Here we summarise how these effects, or perturbations, can be computed (see appendix A.4 for details).

First, we show in appendix A.4.1 that the effect of a trait l on internal states at age a in class j is obtained as

$$\left.\frac{\partial \boldsymbol{x}_j(a)}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} = \int_0^a \boldsymbol{\Psi}_j(a,\tau) \left.\frac{\partial \boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j^\circ(\tau), \boldsymbol{v})}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \mathrm{d}\tau, \tag{17}$$

where $\boldsymbol{\Psi}_j$ is the so-called fundamental matrix collecting the cumulative effects of states on their own rates of change and satisfying

$$\frac{\mathrm{d}\boldsymbol{\Psi}_j(a,\tau)}{\mathrm{d}a} = \left.\frac{\partial \boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})}{\partial \boldsymbol{x}_j(a)}\right|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \boldsymbol{\Psi}_j(a,\tau) \text{ with i.c. } \boldsymbol{\Psi}_j(\tau,\tau) = \mathbf{I}, \tag{18}$$

where $\mathbf{I}$ is the $(n_\mathrm{x} \times n_\mathrm{x})$ identity matrix. Usefully, in the special case where rates of change in states are constant with respect to time (i.e. $\frac{\mathrm{d}}{\mathrm{d}a}(\partial \boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})/\partial \boldsymbol{x}_j(a)) = 0$), the fundamental matrix has an explicit exponential form that can be plugged directly into eq. (17) (appendix A.4.1).

The effect of trait l on survivorship to age a in class j is given by

$$\left.\frac{\partial l_j(a)}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} = -l_j^\circ(a) \int_0^a \left.\left(\frac{\partial \mu_j(\boldsymbol{z}_j, \boldsymbol{x}_j(\tau), \boldsymbol{v})}{\partial \boldsymbol{x}_j(\tau)} \frac{\partial \boldsymbol{x}_j(\tau)}{\partial z_{j,l}} + \frac{\partial \mu_j(\boldsymbol{z}_j, \boldsymbol{x}_j^\circ(\tau), \boldsymbol{v})}{\partial z_{j,l}}\right)\right|_{\boldsymbol{z}=\boldsymbol{v}=\boldsymbol{v}^*} \mathrm{d}\tau, \tag{19}$$

that is, as the discounted integral of all past (direct and indirect) effects of trait expression on mortality.

Finally, appendix A.4.2 shows that derivatives of mutant allelic class frequencies at birth with respect to

mutant traits at $\boldsymbol{v}^*$ can be obtained as

$$\left.\frac{\partial \boldsymbol{q}(\boldsymbol{z},\boldsymbol{v}^*)}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}^*} = -\left((\mathbf{R}^{\circ\intercal}-\mathbf{I})(\mathbf{R}^{\circ}-\mathbf{I})+\mathbf{1}^{(n_c\times n_c)}\right)^{-1}(\mathbf{R}^{\circ\intercal}-\mathbf{I})\left.\frac{\partial \mathbf{R}(\boldsymbol{z},\boldsymbol{v}^*)}{\partial z_{j,l}}\right|_{\boldsymbol{z}=\boldsymbol{v}^*}\boldsymbol{q}^{\circ}, \tag{20}$$

where $\mathbf{1}^{(n_c\times n_c)}$ is a matrix of ones of dimension $(n_c \times n_c)$. See appendix A.4.2 for the derivatives of mutant allele frequencies when $\boldsymbol{v} \neq \boldsymbol{v}^*$.

Together, the above results provide all the ingredients needed to compute the indirect pathways of correlational selection in class-structured populations with ontogeny. We will return to how to combine these ingredients more precisely in section 3.3, where we summarise our results and highlight potential computational challenges.

3.2.4 Connections with previous expressions for quadratic selection

Eqs. (14)–(16) connect to previous results on quadratic selection in class-structured populations found in the literature. First, in the absence of ontogeny, age-, and class-structure, eq. (14) reduces to the standard multivariate quantitative genetics estimates of quadratic selection, where correlational selection is given by the cross derivative of individual fitness with respect to individual trait values (Lande 1979; Lande & Arnold 1983; Phillips & Arnold 1989). Second, in the absence of ontogeny but with class structure (so that Hamiltonians reduce to fecundity rates), eq. (14) reduces to the sum of eqs. (16a) and (16f), consistent with the results of Ohtsuki et al. (2020) for a single trait (see their eq. 34 without relatedness). Finally, in the absence of class-structure and for a single trait, eqs. (14)–(16) reduce to eq. (3.13) of Avila & Mullon (2023) (which can be recovered by substituting the Hamiltonian into our expressions). The model therefore extends these results not only to class-structured populations but also to arbitrary ploidy and any number of evolving traits. Moreover, whereas Ohtsuki et al. (2020, eq. 35) provided a system of equations for the perturbation of class frequencies in terms of resident frequencies and derivatives of fitness components (that cannot always be solved by direct methods), and whereas Avila & Mullon (2023) left the effects of traits on internal states and survivorship implicit, here we resolved these expressions fully (see section 3.2.3). In particular, eq. (20) shows that the perturbation of class frequencies can be expressed directly in terms of resident frequencies and derivatives of fitness components. This then allows for a complete characterisation of directional, correlational, and disruptive selection in class-structured populations with and without ontogeny, opening the way for the analysis of adaptation under many biologically relevant scenarios.

3.3 Summary of the analysis of selection

Building on the above, we now outline how the different components of our model come together in practice. In particular, we summarise the steps required to evaluate directional and quadratic selection, and to classify evolutionary singular points (e.g. Geritz et al. 1998).

Once eq. (1) and eq. (2) have been specified for a given biological scenario, the analysis proceeds as follows. First, the selection gradient is characterised using eqs. (8) and (13), which requires solving for resident costate dynamics (eq. 11) and state dynamics (eq. 12). This allows identification of singular trait values $\boldsymbol{v}^*$ (by solving eq. 5) and assessment of their convergence stability (via the Jacobian matrix, eq. 6). Second, quadratic selection is characterised using eqs. (14)–(20), which makes it possible to determine whether convergence-stable trait values are uninvadable or candidate evolutionary branching point leading to polymorphism (when the Hessian matrix is negative-definite or not negative semi-definite, respectively). While some calculations simplify when class proportions at birth are constant (i.e. trait independent, see box 2), in most cases internal states and costate dynamics, as well as internal states and survivorship perturbations, cannot be solved analytically. This makes numerical approaches unavoidable.

To this end, we provide a set of Mathematica (Wolfram Research, Inc. 2019) functions that can help perform these computations numerically. These functions only require the specification of mutant developmental and vital rates in terms of traits, survivorship and internal state values (eq. 2 and fecundity rates $f_{ij}(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$). Given a set of parameter values, the workflow is as follows. The selection gradient is first computed on a grid of resident allelic values $\boldsymbol{v}$ by constructing Hamiltonians, solving eqs. (11) and (12) numerically, and evaluating eqs. (8) and (13). This gradient is then interpolated numerically, its zeros identified as singular trait values $\boldsymbol{v}^*$, and the Jacobian matrix approximated numerically at these $\boldsymbol{v}^*$ to identify convergence stable trait values (eq. 6). Finally, the Hessian matrix at $\boldsymbol{v}^*$ is computed using eq. (14), via a separate calculation of each term in eq. (15) (from eqs. 16-20). This procedure accommodates any number of traits, and can exploit computational shortcuts available when class proportions at birth are constant (box 2).

Two caveats are worth keeping in mind. First, our analysis is based on the reproductive number R_0 of a mutant allele, which is sign equivalent around one to the geometric growth rate of the mutant allele, but not strictly equal to it. This means that the coefficients of selection we derive are comparable relative to one another, but in absolute terms they should be scaled by the inverse of generation time (see for instance Metz & Leimar 2011). Therefore, while our analysis is well suited to describe how natural selection acts on life-history trade-offs (e.g. allocation to growth vs. fecundity at different ages), caution

In the general case described in the main text, computing selection (eq. 8 or eq. 14) involves solving eq. (11) for each offspring class i and parental class j (i.e. a total system of size $n_c \times n_c$). We show in appendix A.5.1 that this problem can be reduced to solving only n_c equations (i.e. one per parent class j) under the assumption that individuals of class j allocate a constant proportion $c_{ij} \in [0,1]$ of their reproductive output into individuals of class i, irrespectively of age or state, such that their overall fecundity f_j satisfies $f_{ij}(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v}) = c_{ij} f_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v})$ with $\sum_{i=1}^{n_c} c_{ij} = 1$. Several biological scenarios can be reduced to the case of constant class proportion. For instance, exogenous environmental sex-determination, sex chromosome systems coupled to fair meiosis, or random dispersal at birth between habitats, all fall into this category. Then, the age-specific directional selection pressure on trait (j,l) (i.e. eq. 9) simplifies to

$$\hat{s}_{j,l}(a, \boldsymbol{v}) = \check{\nu}_j^{\circ} \left. \frac{\partial H_j(\boldsymbol{z}_j, l_j^{\circ}(a), \boldsymbol{x}_j^{\circ}(a), \boldsymbol{v})}{\partial z_{j,l}} \right|_{\boldsymbol{z}=\boldsymbol{v}} q_j^{\circ}, \tag{II-E}$$

where

$$\check{\nu}_j^{\circ} = \sum_{i=1}^{n_c} \nu_i^{\circ} \gamma_{ij} c_{ij} \tag{II-F}$$

is the allelic reproductive value at birth of one average offspring of individuals of class j, and where

$$\begin{aligned} H_j(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j(a), \boldsymbol{v}) = {} & l_j(a) f_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v}) - \lambda_j^{\mathrm{l}}(a) \mu_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v}) l_j(a) \\ & + \boldsymbol{\lambda}_j^{\mathrm{x}}(a) \cdot \boldsymbol{g}_j(\boldsymbol{z}_j, \boldsymbol{x}_j(a), \boldsymbol{v}) \end{aligned} \tag{II-G}$$

is the Hamiltonian specific to class j. In equation (II-G), the class-specific costates are solutions to the systems of ODEs,

$$\begin{aligned} \frac{\mathrm{d}\lambda_j^{\mathrm{l}}(a)}{\mathrm{d}a} &= - \left. \frac{\partial H_j(\boldsymbol{z}_j, l_j(a), \boldsymbol{x}_j^{\circ}(a), \boldsymbol{v})}{\partial l_j(a)} \right|_{\boldsymbol{z}=\boldsymbol{v}} \qquad \text{with i.c.} \qquad \lambda_j^{\mathrm{l}}(0) = \sum_{i=1}^{n_c} R_{ij}^{\circ}, \\ \frac{\mathrm{d}\boldsymbol{\lambda}_j^{\mathrm{x}}(a)}{\mathrm{d}a} &= - \left. \frac{\partial H_j(\boldsymbol{z}_j, l_j^{\circ}(a), \boldsymbol{x}_j(a), \boldsymbol{v})}{\partial \boldsymbol{x}_j(a)} \right|_{\boldsymbol{z}=\boldsymbol{v}} \qquad \text{with f.c.} \qquad \lim_{a \to \infty} \boldsymbol{\lambda}_j^{\mathrm{x}}(a) = \boldsymbol{0}. \end{aligned} \tag{II-H}$$

Intuitively, these results mean that under constant class proportions at birth, the optimization of fitness reduces to the optimization of overall lifetime reproductive output.

Box 2: Constant class proportions at birth.

is needed if one's desire is to compare selection to other evolutionary processes such as mutation or drift. This is especially important for drift, since effective population size is also influenced by generation time (Charlesworth 1994).

Second, our analysis has so far assumed that mutation effects on trait values are not exactly equal across traits thereby excluding cases where strong genetic constraints force traits to evolve identically across classes. Such fully correlated traits, however, are expected to play a role especially in systems where class-structure recently evolved and where genetic conflicts have not been resolved. With some adjustments, it turns out that the above results can be applied to such cases. In particular, we can readily consider the case where two traits, say (i,k) and (j,l), are underlain by the same pleiotropic genes such that the traits always mutate in the same way (i.e. with $\eta_{i,k} = \eta_{j,l}$). In this case, it is as though we have a single

trait, such that $u_{i,k} = u_{j,l}$ and $v_{i,k} = v_{j,l}$, and a singular trait value $\boldsymbol{v}^*$ must then satisfy

$$s_{i,k}(\boldsymbol{v}^*) + s_{j,l}(\boldsymbol{v}^*) = 0 \quad \text{while} \quad s_{m,n}(\boldsymbol{v}^*) = 0 \;\; \text{for all } (m,n) \neq (i,k),(j,l) \tag{21}$$

under the constraint that $v^*_{i,k} = v^*_{j,l}$. Eq. (21) shows how genetic constraints can impose additional trade-offs across classes. For instance, plugging eqs. (8)–(9) into eq. (21) reveals that when there is antagonism between classes (i.e. when $s_{i,k}(\boldsymbol{v}) \neq s_{j,l}(\boldsymbol{v})$ with $i \neq j$), a singular trait value will reflect a compromise between selection in both classes, giving more weight to the most common class (i.e. the class j one with the largest q°_j). We detail how such genetic constraints and resulting antagonisms modify the analyses of convergence stability and uninvadability in appendix A.6.

4 Example: ontogenetic sexual niche partitioning

We now use the above results to study the joint evolution of female and male traits that govern the ontogeny of sexually relevant morphology. Our goal is twofold: (i) to illustrate the general approach and numerical procedure on a concrete biological scenario; and (ii) to gain theoretical insight into how sex-specific developmental schedules, by shaping size or structures used in competition and mate acquisition, can feed back on the strength and direction of sexual selection.

4.1 The model

4.1.1 Population and traits

We consider a diploid, dioecious, age-structured population with two classes, females and males, denoted by subscripts f and m respectively (for readability, we use these instead of using 1 and 2 to denote classes here). Sex is determined at birth with a fixed primary sex ratio c (the proportion of males in newborns, see tab. 3 for symbols). Class proportions at birth are thus constants in this model. This lead to several simplifications that are detailed in box 2. In the notation introduced there, we have $c_{\mathrm{mf}} = c_{\mathrm{mm}} = c$ and $c_{\mathrm{ff}} = c_{\mathrm{fm}} = 1 - c$. Age-specific vital rates (survival, development, and offspring production) are allowed to differ between the sexes and to depend on sex-specific traits as we detail below.

We study the co-evolution of female and male traits z_{f} and z_{m} that regulate the allocation of resources into the growth of a sex-specific morphological, physiological, or behavioural module, represented by the internal states $x_{\mathrm{f}}(a)$ and $x_{\mathrm{m}}(a)$ (and as a baseline we present the case without sex or development, where

Symbol	Meaning
Parameters	
c	Primary sex-ratio i.e. proportion of males in offspring at birth
α	Rate at which mass is lost
$\mu_{\mathrm{e},j}$	Extrinsic mortality rate in sex $j \in \{\mathrm{f}, \mathrm{m}\}$
β_j	Survival cost to growth in sex j
κ	Strength of sexual selection (based on size matching with females)
K	Density-dependence parameter (modulating female fecundity)
Variables/functions	
$z_{\mathrm{m}}(u_{\mathrm{m}}, v_{\mathrm{m}}), z_{\mathrm{f}}(u_{\mathrm{f}}, v_{\mathrm{f}}), z_{\mathrm{l}}(u_{\mathrm{l}}, v_{\mathrm{l}})$	Male, female and shared (i.e. under genetic constraints) basal growth rates
$x_j(a)$, $x_j^{\circ}(a)$	Size at age a in a mutant and resident of sex j
N_{T}°	Resident population density
$\chi_j^{\circ}(x_j)$	Probability density of size x_j within residents of sex j
$\tilde{C}(x_{\mathrm{m}}(a), x_{\mathrm{f}})$, $C(x_{\mathrm{m}}(a), x_{\mathrm{f}})$	Absolute and relative competitive ability of a mutant male of size $x_{\mathrm{m}}(a)$ in fertilising the eggs of a female of size x_{f}

Tab. 3. List of symbols used in section 4.

$z_{\mathrm{f}} = z_{\mathrm{m}} = x_{\mathrm{f}}(a) = x_{\mathrm{m}}(a)$ in appendix B.2). Biologically, the most straightforward way to think about this module is as overall body size or body mass, where female fecundity scales with size and male compatibility with females is mediated by size matching. For ease of presentation, this is the interpretation we will adopt throughout. However, the model also applies to specialised reproductive modules such as genital morphology, gamete recognition characters (e.g. fertilisation proteins), or sexually selected phenotypes (e.g. eyestalks, ritualised dances) when female choice correlates with her internal states.

4.1.2 Growth and survival

We assume that the rate of change in the internal state $x_j(a)$ of an individual of sex $j \in \{\mathrm{f}, \mathrm{m}\}$ at age a does not depend on its environment (i.e. no density dependence, no environmental effect in resource availability), and write

$$\frac{\mathrm{d}x_j(a)}{\mathrm{d}a} = g_j(z_j, x_j(a)) = z_j - \alpha x_j(a) \quad \text{with i.c.} \quad x_j(0) = 0, \tag{22}$$

where α is the rate at which size (or mass) is lost due to physiological constraints (e.g. respiration), and $z_j = z_j(u_j, v_j) = (u_j + v_j)/2$ is the trait value of a mutant individual, corresponding to its sex-specific basal growth rate, which can be thought of as the baseline energetic investment into growth. (Since there is a single trait per sex, we simplify notation by writing $z_j = z_{j,1}$, $u_j = u_{j,1}$, and $v_j = v_{j,1}$).

There is a trade-off between allocation to growth and survival such that the sex-specific mortality rate

of a mutant individual of sex j is

$$\mu_j(z_j) = \mu_{\mathrm{e},j} + \beta_j z_j^2, \tag{23}$$

where $\mu_{\mathrm{e},j}$ is a constant extrinsic mortality rate (e.g. due to sex-specific risks) and β_j modulates the cost of growth on survival. The rate of change of survivorship in a focal mutant of sex j and age a is thus

$$\frac{\mathrm{d}l_j(a)}{\mathrm{d}a} = -\mu_j(z_j)\, l_j(a), \tag{24}$$

(from eq. 2).

Solving eqs. (22) and (24) at resident allelic values $\boldsymbol{v}$, the size and survivorship of a resident individual of sex j and age a are

$$x_j^\circ(a) = \frac{1-e^{-\alpha a}}{\alpha}\, v_j \quad \text{and} \quad l_j^\circ(a) = e^{-\mu_j^\circ a}, \tag{25}$$

where $v_j = z_j(v_j, v_j)$ and $\mu_j^\circ = \mu_j(v_j)$. We show in appendix B.1 that the corresponding equilibrium probability density of individuals of size x_j and sex j in the resident population is

$$\chi_j^\circ(x_j) = \frac{\mu_j^\circ}{v_j}\left(\frac{v_j}{v_j - \alpha x_j}\right)^{\frac{\alpha-\mu_j^\circ}{\alpha}}, \tag{26}$$

which is defined for values of x_j between zero and the maximal resident size $\tilde{x}_j^\circ = v_j/\alpha$ (obtained from eq. 25 in the limit of $\alpha \to \infty$ and entailing that $\int_0^{\tilde{x}_j^\circ} \chi_j^\circ(x_j)\mathrm{d}x_j = 1$). Accordingly, the shape of the size density distribution in each sex qualitatively depends on the balance between α and the mortality rate $\mu_j^\circ = \mu_{\mathrm{e},j} + \beta_j v_j^2$. When $\alpha > \mu_j^\circ$, the density $\chi_j^\circ(x_j)$ increases with size x_j, indicating that many individuals survive to reach large sizes. When $\alpha < \mu_j^\circ$, the density decreases with size, indicating that most individuals die before reaching their maximum size. This is illustrated in fig. 1, and plays an important role in evolutionary dynamics.

Fig. 1. Equilibrium distribution of size in the resident population. Each line represents the equilibrium probability density of individuals of size x_j and sex j among residents. With increasing investment into growth, the maximum size increases (compare black, brown and orange lines), but so does the resident mortality rate $\mu_j^\circ = \mu_{\mathrm{e},j} + \beta_j v_j^2$, such that the proportion of smaller-sized individuals also increases. Under a stronger cost β_j of growth on survival, the switch from a distribution with a majority of larger individuals to a population with mostly smaller individuals occurs at smaller values of v_j^* (compare plain and dotted lines). Fixed parameter values: $\alpha = 0.1$ and $\mu_{\mathrm{e},j} = 0.05$.

4.1.3 Female and male fecundity

The remaining model components we need to specify are the female and male fecundity rates, denoted $f_{\mathrm{f}}(x_{\mathrm{f}}(a), \boldsymbol{v})$ and $f_{\mathrm{m}}(x_{\mathrm{m}}(a), \boldsymbol{v})$. We assume that a female produces eggs at a rate $f_{\mathrm{f}}(x_{\mathrm{f}}(a), \boldsymbol{v})$ and that each egg is fertilised (i.e. no mate or sperm limitation). Male fecundity $f_{\mathrm{m}}(x_{\mathrm{m}}(a), \boldsymbol{v})$ then reflects the outcome of competition among males for the fertilisation of the eggs.

For females, we assume that fecundity increases linearly with size $x_{\mathrm{f}}(a)$ and decreases with total population density, for example due to competition for resources. Specifically, we assume:

$$f_{\mathrm{f}}(x_{\mathrm{f}}(a), \boldsymbol{v}) = x_{\mathrm{f}}(a)\left(1 - \frac{N_{\mathrm{T}}^{\circ}}{K}\right), \tag{27}$$

where $K > 0$ controls (inversely) the strength of density dependence (large values of K mean low density-dependent competition), and N_{T}° is the equilibrium total density of individuals in the resident population (sum of male and female densities). We show in appendix B.1 that this is given by

$$N_{\mathrm{T}}^{\circ} = K\left(1 - \frac{1}{(1-c)\int_0^{\infty} l_{\mathrm{f}}^{\circ}(a)x_{\mathrm{f}}^{\circ}(a)\mathrm{d}a}\right), \tag{28}$$

where $(1-c)\int_0^{\infty} l_{\mathrm{f}}^{\circ}(a)x_{\mathrm{f}}^{\circ}(a)\mathrm{d}a$ is the expected lifetime production of female offspring by a single resident female in the absence of density-dependence.

For males, we assume that fecundity — that is, the fertilisation of eggs — depends on matching with female size (e.g. due to mechanical constraints during mating or female preferences for size-similar partners) and on competition with other males. Specifically, we write the fecundity rate of a focal mutant male of size $x_{\mathrm{m}}(a)$, i.e. the total number of offspring he is expected to sire across all females in the population, as

$$f_{\mathrm{m}}(x_{\mathrm{m}}(a), \boldsymbol{v}) = \int_0^{\tilde{x}_{\mathrm{f}}^{\circ}} (1-\tilde{c}^{\circ})N_{\mathrm{T}}^{\circ}\chi_{\mathrm{f}}^{\circ}(x_{\mathrm{f}})f_{\mathrm{f}}(x_{\mathrm{f}}, \boldsymbol{v})C(x_{\mathrm{m}}(a), x_{\mathrm{f}})\mathrm{d}x_{\mathrm{f}}, \tag{29}$$

where we recall that $\tilde{x}_j^{\circ} = v_j/\alpha$ is the maximum size attainable by residents of sex j, and where

$$\tilde{c}^{\circ} = \frac{c\mu_{\mathrm{f}}^{\circ}}{(1-c)\mu_{\mathrm{m}}^{\circ} + c\mu_{\mathrm{f}}^{\circ}} \tag{30}$$

is the resident secondary sex-ratio (i.e. the proportion of males in the resident population; see appendix B.1.1 for derivation). Accordingly, $(1-\tilde{c}^{\circ})N_{\mathrm{T}}^{\circ}\chi_{\mathrm{f}}^{\circ}(x_{\mathrm{f}})$ in eq. (29) is the total number of females of size x_{f}, and $(1-\tilde{c}^{\circ})N_{\mathrm{T}}^{\circ}\chi_{\mathrm{f}}^{\circ}(x_{\mathrm{f}})f_{\mathrm{f}}(x_{\mathrm{f}}, \boldsymbol{v})$ is the rate at which eggs are produced by these females collectively (i.e. the total fecundity rate of females of size x_{f}). For each of these eggs, we assume that the probability that it

is fertilised by a focal mutant male of size $x_{\mathrm{m}}(a)$ is given by

$$C(x_{\mathrm{m}}(a), x_{\mathrm{f}}) = \frac{\tilde{C}(x_{\mathrm{m}}(a), x_{\mathrm{f}})}{\int_0^{\tilde{x}_{\mathrm{m}}^{\circ}} \tilde{c}^{\circ} N_{\mathrm{T}}^{\circ} \chi_{\mathrm{m}}^{\circ}(x_{\mathrm{m}}) \tilde{C}(x_{\mathrm{m}}, x_{\mathrm{f}}) \mathrm{d}x_{\mathrm{m}}}, \tag{31}$$

where the numerator gives the competitive ability of the focal male to fertilise a female of size x_{f} and the denominator normalises over all resident males. We assume that competitive ability declines with the difference between male and female size according to

$$\tilde{C}(x_{\mathrm{m}}(a), x_{\mathrm{f}}) = \exp\left[-\kappa(x_{\mathrm{m}}(a) - x_{\mathrm{f}})^2\right], \tag{32}$$

where $\kappa \geq 0$ is a parameter controlling the strength of size matching: when $\kappa = 0$, all males are equally competitive regardless of their size (i.e. $\tilde{C} = 1$); as κ increases, fertilisation becomes increasingly size-assortative.

Equation (32) is equivalent to the compatibility function ψ used in earlier models of sexual selection (e.g. Gavrilets 2000; Gavrilets & Waxman 2002), where trait matching determines mating success and can favour polymorphism. However, whereas other models assume fixed trait values, here we apply this mating function to phenotypes that change with development over individuals' lifetimes. From a mechanistic point of view, eq. (32) applies indiscriminately to systems where female directly express preference for similar-sized males (e.g. as in the convict cichlid *Archocentrus nigrofasciatus*; Beeching et al. 2010), where they co-occur more often with similar-sized males (e.g. as in the striped goby cichlid *Eretmodus cyanostictus*; Taborsky et al. 2014), or where mating success depends on size-similarity due to morphological constraints (e.g. as in the Jerusalem cricket *Stenopelmatus sp.*; Weissman et al. 2008). Finally, note that eq. (29) links male and female fecundity such that it satisfies Fisher's condition (i.e. any newborn has one parent of each sex; Houston & McNamara 2006; Fromhage & Jennions 2016; see appendix A.5.2).

4.2 Analysis

4.2.1 Hamiltonians and costates

Since class proportions at birth are constant in this model, we only need class–specific Hamiltonians and costates (see Box 2). Substituting the model ingredients (eqs. 22, 23, 27, and 29) into eq. (II-G) of Box 2

gives

$$H_{\mathrm{f}}(z_{\mathrm{f}}, l_{\mathrm{f}}(a), x_{\mathrm{f}}(a), \boldsymbol{v}) = l_{\mathrm{f}}(a) f_{\mathrm{f}}(x_{\mathrm{f}}(a), \boldsymbol{v}) - \lambda_{\mathrm{f}}^{\mathrm{l}}(a) \left[\mu_{\mathrm{e,f}} + \beta_{\mathrm{f}} z_{\mathrm{f}}^2\right] l_{\mathrm{f}}(a) + \lambda_{\mathrm{f}}^{\mathrm{x}}(a) \left[z_{\mathrm{f}} - \alpha x_{\mathrm{f}}(a)\right]$$
$$H_{\mathrm{m}}(z_{\mathrm{m}}, l_{\mathrm{m}}(a), x_{\mathrm{m}}(a), \boldsymbol{v}) = l_{\mathrm{m}}(a) f_{\mathrm{m}}(x_{\mathrm{m}}(a), \boldsymbol{v}) - \lambda_{\mathrm{m}}^{\mathrm{l}}(a) \left[\mu_{\mathrm{e,m}} + \beta_{\mathrm{m}} z_{\mathrm{m}}^2\right] l_{\mathrm{m}}(a) + \lambda_{\mathrm{m}}^{\mathrm{x}}(a) \left[z_{\mathrm{m}} - \alpha x_{\mathrm{m}}(a)\right], \tag{33}$$

where the sex-specific costates satisfy the system of ODEs:

$$\frac{\mathrm{d}\lambda_{\mathrm{f}}^{\mathrm{x}}(a)}{\mathrm{d}a} = \alpha \lambda_{\mathrm{f}}^{\mathrm{x}}(a) - l_{\mathrm{f}}^{\circ}(a) \left(1 - \frac{N_{\mathrm{T}}^{\circ}}{K}\right), \tag{34a}$$

$$\frac{\mathrm{d}\lambda_{\mathrm{m}}^{\mathrm{x}}(a)}{\mathrm{d}a} = \alpha \lambda_{\mathrm{m}}^{\mathrm{x}}(a) - l_{\mathrm{m}}^{\circ}(a) \left.\frac{\partial f_{\mathrm{m}}(x_{\mathrm{m}}(a), \boldsymbol{v})}{\partial x_{\mathrm{m}}(a)}\right|_{\boldsymbol{z}=\boldsymbol{v}}, \tag{34b}$$

$$\frac{\mathrm{d}\lambda_{\mathrm{f}}^{\mathrm{l}}(a)}{\mathrm{d}a} = \mu_{\mathrm{f}}^{\circ} \lambda_{\mathrm{f}}^{\mathrm{l}}(a) - x_{\mathrm{f}}^{\circ}(a) \left(1 - \frac{N_{\mathrm{T}}^{\circ}}{K}\right), \tag{34c}$$

$$\frac{\mathrm{d}\lambda_{\mathrm{m}}^{\mathrm{l}}(a)}{\mathrm{d}a} = \mu_{\mathrm{m}}^{\circ} \lambda_{\mathrm{m}}^{\mathrm{l}}(a) - f_{\mathrm{m}}(x_{\mathrm{m}}^{\circ}(a), \boldsymbol{v}). \tag{34d}$$

These have boundary conditions $\lim_{a\to\infty} \lambda_j^{\mathrm{x}}(a) = 0$ and $\lambda_j^{\mathrm{l}}(0) = 2$ since in a resident sexual population at demographic equilibrium, each individual has an expected lifetime reproductive output of two offspring (one daughter and one son on average). We solve these ODEs numerically.

4.2.2 Directional selection : balancing future effects of growth across ages

Age-specific coefficients of directional selection are obtained by substituting eq. (33) into eq. (II-E) of Box 2. In addition, since we are considering a diploid population with constant primary sex ratio, the class frequencies and reproductive values are $q_{\mathrm{f}}^{\circ} = (1-c)$, $q_{\mathrm{m}}^{\circ} = c$ and $\check{\nu}_{\mathrm{f}}^{\circ} = \check{\nu}_{\mathrm{m}}^{\circ} = \frac{1}{2}$ respectively (see Box 2 and appendix A.5.2 for derivation). Using these quantities, we have

$$\hat{s}_{\mathrm{f}}(a, \boldsymbol{v}) = \check{\nu}_{\mathrm{f}}^{\circ} \left.\frac{\partial H_{\mathrm{f}}(z_{\mathrm{f}}, l_{\mathrm{f}}^{\circ}(a), x_{\mathrm{f}}^{\circ}(a), \boldsymbol{v})}{\partial z_{\mathrm{f}}}\right|_{\boldsymbol{z}=\boldsymbol{v}} q_{\mathrm{f}}^{\circ} = \frac{1-c}{2} \left(\lambda_{\mathrm{f}}^{\mathrm{x}}(a) - \lambda_{\mathrm{f}}^{\mathrm{l}}(a) l_{\mathrm{f}}^{\circ}(a) 2\beta_{\mathrm{f}} v_{\mathrm{f}}\right), \tag{35a}$$

$$\hat{s}_{\mathrm{m}}(a, \boldsymbol{v}) = \check{\nu}_{\mathrm{m}}^{\circ} \left.\frac{\partial H_{\mathrm{m}}(z_{\mathrm{m}}, l_{\mathrm{m}}^{\circ}(a), x_{\mathrm{m}}^{\circ}(a), \boldsymbol{v})}{\partial z_{\mathrm{m}}}\right|_{\boldsymbol{z}=\boldsymbol{v}} q_{\mathrm{m}}^{\circ} = \frac{c}{2} \left(\lambda_{\mathrm{m}}^{\mathrm{x}}(a) - \lambda_{\mathrm{m}}^{\mathrm{l}}(a) l_{\mathrm{m}}^{\circ}(a) 2\beta_{\mathrm{m}} v_{\mathrm{m}}\right). \tag{35b}$$

Eq. (35) shows that age-specific selection at age a on the basal growth rate balances two antagonistic effects on future reproduction. On the one hand, future fitness effects mediated by growth tend to favour larger z_j especially at a young age (captured by $\lambda_j^{\mathrm{x}}(a)$ in eq. 35 see fig. 21). On the other hand, future fitness effects mediated by survival favour smaller z_j (captured by the term $\lambda_j^{\mathrm{l}}(a) l_j^{\circ}(a)\, 2\beta_j v_j$ in eq. 35). The reason is that higher growth raises mortality and hence lowers the expected remaining reproductive output, $\lambda_j^{\mathrm{l}}(a) l_j^{\circ}(a)$ here, which itself declines with age (fig. 2B). In our model, $\lambda_j^{\mathrm{x}}(a)$ declines more steeply

than $\lambda_j^{\mathrm{l}}(a)l_j^{\circ}(a)$ with age a, so the balance between these two antagonistic effects changes with age: selection favours faster growth at younger ages but slower growth later in life (fig. 2C).

Fig. 2. Age-specific directional selection arises from opposing effects of growth on reproductive value and survival. A. Age-specific benefit of increased size, captured by the costate $\lambda_j^{\mathrm{x}}(a)$, declines with age. This reflects the diminishing marginal returns of size on future reproduction as individuals age, especially when size loss is high (α large). **B.** Expected remaining reproductive output at age a, given by $\lambda_j^{\mathrm{l}}(a)l_j^{\circ}(a)$, also declines with age. However, it is more sensitive to α in females than in males (lighter curves). This is because higher α reduces the maximum attainable size and thus lifetime fecundity in females. These opposing age-specific effects (A vs. B) jointly generate the pattern of directional selection in panel C: early selection favours growth due to future size benefits, while late selection disfavours it due to survival costs. **C.** Age-specific directional selection on growth (from eqs. 35 and 38) at the convergence stable strategies ($\boldsymbol{v}^* = (v_{\mathrm{f}}^*, v_{\mathrm{m}}^*) = (0.518, 0.432)$ for independent traits; $v_{\mathrm{l}}^* = 0.356$ for a shared trait). Each curve integrates to zero (owing to eqs. 5 and 8), but its shape reveals when selection favours increased or decreased allocation to growth. Selection tends to favour faster growth at early ages but reduced allocation later in life. Parameter values: $c = 0.5$, $\mu_{\mathrm{e,f}} = \mu_{\mathrm{e,m}} = 0.1$, $\beta_{\mathrm{f}} = \beta_{\mathrm{m}} = 0.2$, $\kappa = 1$, and $K = 10^4$. Parameter α has value 0.2 in panel C.

Selection eventually favours a compromise between these antagonistic effects. This is formalised by unique convergence stable trait values in females and males $\boldsymbol{v}^* = (v_{\mathrm{f}}^*, v_{\mathrm{m}}^*)$ that satisfy

$$v_j^* = \frac{1}{2\beta_j} \frac{\int_0^\infty \lambda_j^{\mathrm{x}}(a)\mathrm{d}a}{\int_0^\infty \lambda_j^{\mathrm{l}}(a)l_j^{\circ}(a)\mathrm{d}a} \tag{36}$$

for sex j (obtained by plugging eq. 35 into eq. 8). Note that eq. (36) is implicit as both costates and internal states depend on $\boldsymbol{v}^*$. It nevertheless makes clear that the equilibrium growth rate increases when the cumulative benefits of growth (i.e. $\int_0^\infty \lambda_j^{\mathrm{x}}(a)\mathrm{d}a$) outweigh the cumulative benefits of survival (i.e. $\int_0^\infty \lambda_j^{\mathrm{l}}(a)l_j^{\circ}(a)\mathrm{d}a$).

A numerical analysis of eq. (36) together with eq. (34) allows us to investigate how key parameters influence $\boldsymbol{v}^*$ and thus life histories and demography at evolutionary equilibrium. Unsurprisingly, greater growth costs β_j select for slower growth in both females (fig. 3A) and males (fig. S1A), leading to longer lives with delayed reproduction. By contrast, higher extrinsic mortality $\mu_{\mathrm{e},j}$ reduces the expected remaining reproductive output $\lambda_j^{\mathrm{l}}(a)l_j^{\circ}(a)$ and therefore selects for faster growth in sex j (fig. 3B for females, fig. S1B for males), leading to "live-fast–die-young" types of life histories (André & Rousset 2020). In females, faster growth partly compensates for the reduction in average size (and hence fecundity) caused by extrinsic mortality, such that population extinction ($N_{\mathrm{T}}^{\circ} \leq 0$) occurs only at relatively high

values of $\mu_{\mathrm{e,f}}$ (fig. 3B, dark gray area).

Fig. 3. Female mortality and growth parameters influence sex-specific evolution and extinction risk. Each panel shows convergence stable growth rates $\boldsymbol{v}^*$ (solid lines) and corresponding average sizes (dashed lines) for females (red) and males (blue), as a function of the parameter indicated on the x-axis (computed from eq. 36 for trait values and B-99 for average sizes). Trait values under genetic constraint (i.e. identical growth rate in males and females, v_{l}^*) and the resulting female size are shown in green (from eq. 39). **A.** Higher cost of growth to survival in females (β_{f}) selects for slower growth and smaller average size. Males evolve smaller size in response, to match females owing to size-matching sexual selection. **B.** Higher female extrinsic mortality ($\mu_{e,f}$) favours faster growth. Above a threshold, female fecundity becomes insufficient to maintain recruitment and the population goes extinct ($N_{\mathrm{T}}^{\circ} \leq 0$; dark grey area). This extinction threshold is lower when genetic correlations constrain male and female growth to evolve identically (light grey area), because sexual antagonism prevents females from evolving growth rates high enough to offset increased mortality (green). Note that convergence stable growth rates $\boldsymbol{v}^*$ and average sizes are reported beyond the extinction thresholds (transparent lines) for graphical purposes, since the model provides result even when $N_{\mathrm{T}}^{\circ} \leq 0$. **C.** A higher rate of size loss (α) prevents individuals from reaching larger sizes. This favours faster growth in females, but slower growth in males. See main text for explanation. Parameter values (unless varied on x-axis): $c = 0.5$, $\alpha = 0.2$, $\mu_{\mathrm{e,f}} = \mu_{\mathrm{e,m}} = 0.1$, $\beta_{\mathrm{f}} = \beta_{\mathrm{m}} = 0.2$, $\kappa = 1$, and $K = 10^4$.

The parameter α, which determines the rate at which size is lost (eq. 22), has opposite consequences in the two sexes. In females, greater α selects for faster growth (larger v_{f}^*), while in males it selects for slower growth (smaller v_{m}^*; fig. 3C). The reason is that when α is large, it prevents males and females from reaching higher sizes. This is reflected in the size distributions being shifted towards smaller values when α is increased (compare the histograms of fig. 4A and C). In females, this prevents individuals from reaching high fecundity, lowering their remaining reproductive output at any given age (fig. 2B), and thus favouring growth over survival. For males, meanwhile, because greater α reduces average female size and male success depends on matching female size (eq. 32), this favours slower growth as long as $\kappa > 0$.

In fact, the strength of size-based competition κ plays a decisive role in male evolution. When $\kappa = 0$, size brings no reproductive benefit and males are selected to invest nothing in growth ($v_{\mathrm{m}}^* = 0$). When κ is large, size matching is crucial for male fitness and selection favours growth rates that allow males to match female size. Simulations confirm this: κ has no effect on female evolution v_{f}^* (red curves, left vs. right columns in fig. 4), but v_{m}^* evolves to be larger when κ is large (left curves, left vs. right columns in fig. 4), thereby reducing the mismatch between female and male size distributions (compare red and blue histograms in fig. 4).

Fig. 4. Evolutionary branching of male growth generates male life-history polymorphism. Each panel shows results from individual-based simulations (appendix B.3). The top part shows the expressed trait values in females (red) and males (blue): randomly sampled allelic values (dots; 10 copies every 10 time units), population mean (solid lines), and analytically predicted convergence stable trait values (dashed lines). Reported values give: the male-specific component $h_{\mathrm{mm}}(\boldsymbol{v}^*)$ (eq. 14); its size-mediated component $h_{\mathrm{mm}}^{\mathrm{xx}}(\boldsymbol{v}^*) = 0.25 \int_0^\infty \hat{h}_{\mathrm{mm}}^{\mathrm{xx}}(a, \boldsymbol{v}^*)\,\mathrm{d}a$ (see eqs. 37 and A-37, with 0.25 for diploidy); and the leading eigenvalue of $\mathbf{H}(\boldsymbol{v}^*)$. As expected, when this eigenvalue is negative, the population remains monomorphic (panels A–C). When it is positive, evolutionary branching occurs in males (panel D, split in blue lines). The bottom part of each panel shows male and female size distributions at equilibrium as histograms (averaged between time steps 10^5 and $2{\times}10^5$), along with the analytically predicted distribution at $\boldsymbol{v}^*$ (solid lines, eq. 26). The Kullback–Leibler divergence between male and female distributions is also reported. **A–B**: Male and female distributions are similar in shape as both sexes have the same demographic parameters (see left-hand side of panels for parameter values). However, females grow larger on average due to direct selection on size whereas males only track female size via selection for compatibility. **C–D**: Sex-specific demography generates asymmetries in size distributions with a higher frequency of small females. When sexual selection is strong enough, this leads to evolutionary branching (panel D) with one morph specialising on mating with smaller females. The smaller allele is seldom found at a homozygous state and the smaller large size mode is composed of heterozygotes. Other parameters: $c = 0.5$, $\mu_{\mathrm{e,f}} = 0.1$, $K = 10^4$.

4.2.3 Quadratic selection: sexual selection based on trait matching favours life-history polymorphism in males

Once the population has converged to the equilibrium $\boldsymbol{v}^*$ characterized by eq. (36), does it remain monomorphic under stabilising selection or can it become polymorphic due to disruptive selection? Numerical evaluations of eq. (14) show that the Hessian matrix can be positive-definite, and thus selection can be disruptive, but only when

$$\begin{aligned}\hat{h}_{\mathrm{mm}}^{\mathrm{xx}}(a, \boldsymbol{v}^*) = {} & \frac{\kappa(1-e^{-\alpha a})^2}{2(1-c)\alpha^2} l_{\mathrm{m}}^{\circ}(a) \\ & \times \left[\int_0^{\tilde{x}_{\mathrm{f}}^{\circ}} (1-\tilde{c}^{\circ}) N_{\mathrm{T}}^{\circ} \chi_{\mathrm{f}}^{\circ}(x_{\mathrm{f}}) f_{\mathrm{f}}(x_{\mathrm{f}}, \boldsymbol{v}^*) C(x_{\mathrm{m}}^{\circ}(a), x_{\mathrm{f}}) \left(2\kappa(x_{\mathrm{m}}^{\circ}(a) - x_{\mathrm{f}})^2 - 1\right) \mathrm{d}x_{\mathrm{f}}\right]\end{aligned} \tag{37}$$

is positive and large, i.e. when positive multiplicative effects of size on male fitness at age a induced by changes in growth rate are large (eq. 16b, as all other components of eq. 15 were found to be too small to drive disruptive selection). Such multiplicative effects via size suggests that disruptive selection should favour males with either fast or slow growth such that small and large males of the same age coexist. Individual-based simulations confirm this (appendix B.3 for procedure). These show evolutionary branch-

ing leading to the coexistence of two diverged alleles coding for different growth rates, such that three types of males occur: two homozygotes with opposite growth strategies and an intermediate heterozygote (fig. 4D, blue dots). Some males thus grow substantially slower, remaining smaller for most of their lives but being longer-lived. At the population level this produces a multimodal male size distribution and a closer match with the female size distribution (compare panels C and D of fig. 4).

A closer inspection of eq. (37) reveals the drivers of polymorphism. Disruptive selection on male growth is favoured when the number of competitors is large (i.e., when c is high or male survivorship is high), when size matching is strong (large κ), and when individuals can attain large sizes (small α). In addition, the integral in eq. (37) shows that a necessary condition is that the fecundity-weighted average of $(2\kappa(x_{\mathrm{m}}^{\circ}(a) - x_{\mathrm{f}})^2 - 1)$ is positive, which occurs when male and female size distributions diverge. In particular, large positive values of $\hat{h}_{\mathrm{m}m}^{\mathrm{xx}}(a, \boldsymbol{v}^*)$, and thus disruptive selection, occur when males are predominantly large ($\alpha > \mu_{\mathrm{m}}^{\circ}$; see fig. 1) whereas females are predominantly small ($\alpha < \mu_{\mathrm{f}}^{\circ}$). In this case, at the convergence stable trait value v_{m}^* most males are close to the average female size, while only a few match the smaller sizes more common in females (see fig. 4C). Under high κ, this leads to intense competition among larger males and thus to a fitness advantage for rare smaller males. Polymorphism therefore evolves because males specialize: some grow rapidly and die young, specializing into fertilising larger and more fecund females, while others grow slowly and live longer, and mate mostly with smaller females. In our simulations, such disruptive selection can be so strong that it overpowers directional selection before the population has completely converged to the singular strategy, leading to early branching (see fig. 4D). Under parameter values that favour a very wide distribution of female sizes, strong disruptive selection can also lead to the coexistence of more than two male morphs (see fig. S2).

Evolutionary branching in our model hinges upon an advantage for males to specialize in different female sizes. Our results therefore connect to resource-competition and niche-partitioning models, where polymorphism evolves as individuals specialise to escape competition (e.g. Kisdi & Geritz 1999; Pásztor et al. 2016; Schmid et al. 2023), except that here the "resource" consists of potential mating partners. They also build on compatibility-based sexual selection models (Gavrilets 2000; Gavrilets & Waxman 2002), which treat individual states as fixed traits determined at birth. In those models, male diversification arises in response to female genetic variation in compatibility traits generated by sexual conflict. By contrast, in our model females remain genetically monomorphic: variation in female states arises because individuals reproduce continuously while changing size through ontogeny. As a result, the population maintains a standing distribution of female sizes, which creates multiple mating niches; competition for these niches makes male fitness frequency dependent and can favour trait diversification. Another feature necessarily absent from these previous studies is that trait diversification here has knock-on effects on demography,

altering the age and state structure of the population. In our model, the trait under sexual selection in males (investment into growth or sexual structures) also affects mortality via a survival–growth trade-off (eq. 23). Consequently, when selection favours alternative male growth strategies, evolutionary branching in that trait necessarily induces divergence in survival and lifespan. The resulting polymorphism therefore spans both mating strategy and life history, feeding back on population age structure and demography with the coexistence of fast-growing, short-lived males and slow-growing, long-lived males.

Alternative male mating morphs with markedly different life history traits are widespread (e.g. sneaker versus dominant males in fish, fighter versus sneaker males in beetles, and multiple morphs in isopods and ruffs). These polymorphisms are generally explained not by compatibility *per se*, but more often by sexual conflict (e.g. coercion versus resistance) or by condition-dependent development, where high- and low-condition males follow different trajectories (Gross 1996; Mank 2023). While our results here apply mostly to the evolution of genetically determined alternative reproductive strategies (Engqvist & Taborsky 2016), our model could readily be adapted to incorporate developmental plasticity and thus help disentangle the relative contributions of sexual selection, sexual conflict, and condition dependence in the evolution of sex-specific development. More generally, it provides a way to better understand the emergence and coexistence of alternative life-history strategies, together with their demographic causes and consequences.

4.2.4 Extensions

We extended the model to explore two cases. First, we investigated the sensitivity of our results to a change in ploidy by considering a haplodiploid system, where males are haploid and females are diploid. This boils down to adjusting the weights in the sums that define selection, using $\gamma_{\mathrm{fm}} = 1$, $\gamma_{\mathrm{mm}} = 0$, $q_{\mathrm{f}}^{\circ} = 2(1-c)/(2-c)$, $q_{\mathrm{m}}^{\circ} = c/(2-c)$, $\check{\nu}_{\mathrm{f}}^{\circ} = \check{\nu}_{\mathrm{m}}^{\circ} = 2c/3$ (see box 2 and appendix A.5.2), and $\partial z_{\mathrm{m}}(u_{\mathrm{m}}, v_{\mathrm{m}})/\partial u_{\mathrm{m}} = 1$ (i.e. assuming complete dosage compensation). Because vital rates are defined as functions of individual phenotype and genetic effects are additive, haplodiploidy does not affect the singular trait values (fig. S3). Male haploidy does however affect quadratic selection by fully exposing mutations in the male trait. By increasing $\partial z_{\mathrm{m}}(u_{\mathrm{m}}, v_{\mathrm{m}})/\partial u_{\mathrm{m}}$, haploidy amplifies the male-specific component $h_{\mathrm{mm}}(\boldsymbol{v}^{*})$ of the Hessian matrix (see eq. 14). As a result, disruptive selection is stronger, especially under low male mortality and high sexual selection (fig. S3D). Evolutionary branching then produces the coexistence of two allelic types where the rarer, smaller type encodes larger male trait values than in the diploid case (compare fig. S3D and fig. 4D).

Second, we considered the effect of genetic constraints across sexes by assuming that males and females share the same growth rate (i.e. see eq. 21). In a diploid population with additive gene action

(i.e. $\partial z_{\mathrm{f}}(u_{\mathrm{f}}, v_{\mathrm{f}})/\partial u_{\mathrm{f}} = \partial z_{\mathrm{m}}(u_{\mathrm{m}}, v_{\mathrm{m}})/\partial u_{\mathrm{m}} = 1/2$), the age-specific directional selection coefficient on the shared trait $z_{\mathrm{l}}(u_{\mathrm{l}}, v_{\mathrm{l}})$ is the sum of the female and male gradients shown in eq. (35):

$$\hat{s}_{\mathrm{l}}(a, v_{\mathrm{l}}) = \hat{s}_{\mathrm{f}}(a, v_{\mathrm{l}}) + \hat{s}_{\mathrm{m}}(a, v_{\mathrm{l}}), \tag{38}$$

resulting in sexual antagonism (i.e. when selection favours different trait values in males and females showing genetic correlations, Bonduriansky & Chenoweth 2009; Cox & Calsbeek 2009; fig. 2C, green line). Solving the integral of eq. (38) for zero, we obtain a unique convergence-stable singular trait

$$v_{\mathrm{l}}^* = \frac{1}{2} \frac{\int_0^\infty \left[(1-c)\lambda_{\mathrm{f}}^{\mathrm{x}}(a) + c\lambda_{\mathrm{m}}^{\mathrm{x}}(a)\right] \mathrm{d}a}{\int_0^\infty \left[(1-c)\lambda_{\mathrm{f}}^{\mathrm{l}}(a) l_{\mathrm{f}}^{\circ}(a)\beta_{\mathrm{f}} + c\lambda_{\mathrm{m}}^{\mathrm{l}}(a) l_{\mathrm{m}}^{\circ}(a)\beta_{\mathrm{m}}\right] \mathrm{d}a}. \tag{39}$$

Eq. (39) captures the lifetime trade-off between growth (numerator) and survival (denominator), averaged across sexes. In spite of this averaging, the equilibrium growth rate does not fall between the female and male optima found in the absence of genetic constraints (i.e. between v_{f}^* and v_{m}^* given by eq. 36) but instead evolves below both (fig. 3). This results from an evolutionary feedback: selection on males favours being smaller than females (owing to the costs of growth), genetic constraints then entail a reduction in female size, which in turn selects males to be even smaller, and so on. As a consequence of this sexually antagonistic feedback, females cannot evolve sufficiently large growth rates in response to high extrinsic mortality, making populations more prone to extinction than in the absence of genetic constraints (compare the green and red lines in fig. 3B).

Once the population has converged to v_{l}^*, whether or not polymorphism evolves depends on the age-specific quadratic selection coefficient $\hat{h}_{\mathrm{ll}}(a, v_{\mathrm{l}}^*) = \hat{h}_{\mathrm{ff}}(a, v_{\mathrm{l}}^*) + \hat{h}_{\mathrm{mm}}(a, v_{\mathrm{l}}^*)$, which is the sum of the female and male components (given by eq. 15; note that the mixed term $\hat{h}_{\mathrm{fm}}(a, v_{\mathrm{l}}^*)$ is always zero in our model, see eq. A-86). By computing $\hat{h}_{\mathrm{ll}}(a, v_{\mathrm{l}}^*)$ numerically, we find that genetic constraints across sexes prevent disruptive selection through two mechanisms. First, stabilizing selection on females is stronger, i.e. $\hat{h}_{\mathrm{ff}}(a, v_{\mathrm{l}}^*)$ is more negative at all ages, than when male and female traits evolve independently (fig. S4, red lines). Sexual antagonism thus prevents disruptive selection from materialising in males because selection in females counteracts it. Second, because the equilibrium growth rate v_{l}^* (eq. 39) is lower owing to sexual antagonism, the variance in female size in the population is not large enough to generate disruptive selection on males, and for rare smaller males to gain an advantage (fig. S5). This effect is especially pronounced under strong sexual selection (see fig. S5D and fig. S4, dashed blue line).

5 Discussion

We derived different quantities for understanding selection on quantitative traits that influence fecundity, mortality, and development rates in age- and class-structured populations. First, the coefficients of directional selection (eqs. 8–13) allow us to determine whether rare mutant alleles inducing small trait changes can invade and fix, and thus to identify convergence stable trait values toward which the population evolves. Second, the coefficients of quadratic selection (eqs. 14–20) indicate whether such trait values are evolutionarily stable or prone to diversification via evolutionary branching, depending on whether nearby allelic types are eliminated by stabilizing selection or favoured by disruptive selection. Whether evolutionary branching occurs or not, these coefficients also identify the correlations among traits that are expected to be shaped by selection.

A key feature of our approach is that both directional and quadratic selection are expressed as a sum of age-specific allelic fitness effects. This provides a biologically meaningful partition of selective forces across ages. As shown in earlier single-class models (Avila et al. 2021; Lehmann 2022), constructing such sums requires quantifying how trait expression at a given age influences both current reproduction (via fecundity at that age) and future reproduction (via changes in development and survival). Delayed effects are tracked using costates, which measure how changes in current internal states and survivorship affect future reproductive output (eq. 11). In class-structured populations, these delayed effects must be characterised not only for each parental class but also with respect to each possible offspring class, requiring $n_c \times n_c$ systems of class-oriented costates (unless there are additional reproductive constraints e.g. constant class proportions at birth; see box 2). This reflects the biological reality that development can influence the production of different types of offspring in different ways, and that these effects can also depend on the parental class (e.g. sexes, ploidies, or habitats). By accounting for this, our results can be used to describe selection in systems where ontogeny interacts with reproductive choices (e.g. maternal effects on sex ratio; Macdonald & Johnson 2008), reproductive mode (e.g. allocation between sexual and vegetative reproduction; Méndez & Obeso 1993), or dispersal (e.g. maternal effects on offspring dispersal; Massot & Clobert 1995).

With selection expressed in terms of age-specific fitness effects, we can consider and better understand the consequences of age-wise antagonistic selection, where different trait values are favoured at different ages but cannot evolve independently due to incomplete plasticity. This type of antagonism is an important aspect of life-history evolution and is often invoked in theories of senescence, because it enables the fixation of alleles that reduce survival later in life while providing advantages early on (Hamilton 1966; Austad & Hoffman 2018). Nevertheless, most continuous-time life-history models neglect this possibility

by assuming complete age-wise plasticity in trait expression (e.g. León 1976; Goodman 1982; Schaffer 1983; Perrin & Sibly 1993; Day & Taylor 2000; Dieckmann et al. 2006; Parvinen et al. 2013; Metz et al. 2016; Houston et al. 2024). By considering non-plastic traits and properly discounting selective effects at later ages (Hamilton 1966), our model makes it possible to analyse the coevolution of ontogeny and senescence in the presence of genetic constraints. In our application for instance, intermediate longevity evolves because higher survival (i.e. via slower growth) is counter selected in early life, showing one possible way how age-wise antagonism may influences the evolution of life history.

By partitioning quadratic selection into age-specific components, our analysis also reveals cases of age-wise antagonistic quadratic selection, where stabilizing and disruptive selection act at different ages (or positive and negative correlational selection). In the sexual selection example, selection is stabilizing on younger males but disruptive on older ones because only the latter benefit from escaping intense competition for larger females (fig. S4). More generally, our decomposition of quadratic selection shows that age-specific effects can be generated indirectly, i.e. through the cumulative influence of trait expression on internal state dynamics, survivorship dynamics, and class-structure (eqs. 16b–16f). The dominant contribution to disruptive selection in the sexual selection example came from effects mediated by internal states, i.e. by size in this case (eq. 37). Because developmental effects accumulate with age, while survivorship gradually declines, the resulting quadratic selection is strongest at intermediate ages here (fig. S4). But in other systems different components may generate disruptive selection and thus follow different patterns of antagonism. Either way, the quadratic selection coefficients presented here can disentangle the mechanisms by which correlational and disruptive selection act through ontogeny and thereby help understand phenomena such as late differentiation, where morphs diverge only after a period of shared development (Ryan & Plague 2004), as well as other life-history polymorphisms where alternative reproductive tactics or morphologies emerge following common juvenile growth trajectories.

Our results also reveal how genetic constraints across classes can generate another form of antagonistic selection, when the same trait is expressed in multiple classes but selection favours different values in each. Such class-wise antagonism, particularly between sexes, has been argued to play an important role in shaping life-history evolution (Bonduriansky et al. 2008). In primates, for example, strong sexual selection for increased size in males appears to cause a correlated increase in female size, owing to the shared genetic architecture of body size across sexes, despite being deleterious for females (Lindenfors 2002). In our example model, sex-wise antagonistic selection prevents the diversification of male life histories, both because stabilizing selection on female growth counteracts disruptive selection on male growth, and because the resulting equilibrium size distribution is less conducive to disruptive selection. In future work, the derived selection coefficients could be used to study more systematically the role of

such genetic constraints in shaping sex-specific ontogeny and alternative reproductive strategies, including cases where the underlying pleiotropic genes are expressed in classes with different ploidies (e.g. when traits are encoded on sex chromosomes). In fact, we have shown in our application that ploidy at the loci encoding a given trait can affect the likelihood to observe evolutionary branching, as well the composition of the population after such event (see fig. S3).

Antagonistic selection across ages and classes is expected to favour the evolution of plasticity, allowing traits to vary with context (Scheiner & Lyman 1991). At first sight, our choice to model traits as fixed scalars may therefore seem restrictive, and even a step back compared to standard life-history models assuming complete plasticity (e.g. León 1976; Perrin & Sibly 1993). However, modelling scalar traits does not mean that plasticity cannot be studied. Any dynamical system can in principle be represented by a finite recurrent neural network (Haykin 1999). By treating the network's weights as evolving scalar traits (the vector $\boldsymbol{v}$ here), it becomes possible to generate reaction norms from constant genetic values in response to environmental or internal cues (Ezoe & Iwasa 1997). In this way, scalar-trait models can be extended to capture the complex mappings between gene expression to life-history strategies, while integrating both external variables and internal state information (e.g. Deshpande & Fronhofer 2025), and these considerations apply to behavioral strategies alike (Lehmann 2022). Understanding how developmental and behavioral systems evolve to generate the required reaction norms, and thus studying how plasticity itself evolves gradually, remains an interesting challenge. By developing evolutionary models of scalar traits, including ontogeny and class structure, our results provide a step toward combining adaptive dynamics with systems approaches to better understand the long-term evolution of biological structures as integrated dynamical systems.

To conclude, we have decomposed directional and quadratic selection into age- and class-specific components, making it possible to identify the many trade-offs and antagonisms that can influence life-history evolution in structured populations. This provides a sharper understanding of adaptation in complex demographic settings and opens the door to a range of problems where context- or class-specific development plays a central role. These include the evolution of sex-specific ontogeny, state-dependent reproductive strategies, class-structured polymorphisms in plants and animals, dispersal strategies that depend on age or condition, and host–parasite interactions shaped by age-structure. More broadly, our approach contributes to linking the dynamics of development, demography, and adaptation, and to understanding how their interactions shape biological diversity.

Acknowledgments

We thank Dr. Ludovic Maisonneuve for constructive feedback.

References

André JB, Rousset F. Does extrinsic mortality accelerate the pace of life? A bare-bones approach. Evol. Hum. Behav. 2020;41:486–92.

Austad SN, Hoffman JM. Is antagonistic pleiotropy ubiquitous in aging biology? Evol. Med. Public Health 2018;2018:287–94.

Avila P, Fromhage L, Lehmann L. Sex-allocation conflict and sexual selection throughout the lifespan of eusocial colonies. Evolution 2019;73:1116–32.

Avila P, Mullon C. Evolutionary game theory and the adaptive dynamics approach: adaptation where individuals interact. Philos. Trans. R. Soc. B Biol. Sci. 2023;378:20210502.

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

Beeching S, Wack C, Ruffner G. Female convict cichlids (Archocentrus nigrofasciatus) prefer to consort with same-sized males. Ethology Ecology & Evolution 2010;July 2004:209–16.

Bonduriansky R, Maklakov A, Zajitschek F, Brooks R. Sexual selection, sexual conflict and the evolution of ageing and life span. Funct. Ecol. 2008;22:443–53.

Bonduriansky R, Chenoweth SF. Intralocus sexual conflict. Trends Ecol. Evol. 2009;24:280–8.

Braendle C, Friebe I, Caillaud MC, Stern DL. Genetic variation for an aphid wing polyphenism is genetically linked to a naturally occurring wing polymorphism. Proc. R. Soc. B Biol. Sci. 2005;272:657–64.

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

Caswell H. Matrix population models. Massachusetts: Sinauer Associates Inc, 2000.

Charlesworth B. Selection in populations with overlapping generations. V. Natural selection and life histories. Am. Nat. 1973;107:303–11.

Charlesworth B. Evolution in age-structured populations. Cambridge: Cambridge University Press, 1994.

Christiansen FB. On conditions for evolutionary stability for a continuously varying character. Am. Nat. 1991;138:37–50.

Cox RM, Calsbeek R. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. Am. Nat. 2009;173:176–87.

Day T, Taylor PD. A generalization of pontryagin's maximum principle for dynamic evolutionary games among relatives. Theor. Popul. Biol. 2000;57:339–56.

De Roos AM. A gentle introduction to physiologically structured population models. In: Structured-Population Models in Marine, Terrestrial, and Freshwater Systems. Massachusetts: Springer US, 1997:119–204.

Deshpande JN, Fronhofer EA. A gene-regulatory network model for density-dependent and sex-biased dispersal evolution during range expansions. bioRxiv 2025.

Dieckmann U, Heino M, Parvinen K. The adaptive dynamics of function-valued traits. J. Theor. Biol. 2006;241:370–89.

Durinx M, Metz JAJ, Meszéna G. Adaptive dynamics for physiologically structured population models. J. Math. Biol. 2008;56:673–742.

Engqvist L, Taborsky M. The evolution of genetic and conditional alternative reproductive tactics. Proc Biol Sci 2016;283:20152945.

Eshel I. On the survival probability of a slightly advantageous mutant gene in a multitype population: A multidimensional branching process model. J. Math. Biol. 1984;19:201–9.

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

Eshel I, Motro U, Sansone E. Continuous stability and evolutionary convergence. J. Theor. Biol. 1997;185:333–43.

Ezoe H, Iwasa Y. Evolution of condition-dependent dispersal: A genetic-algorithm search for the ESS reaction norm. Res. Popul. Ecol. 1997;39:127–37.

Fisher RA. The genetical theory of natural selection. Oxford, England: Clarendon Press, 1930:xiv, 272.

Fox RL, Kapoor MP. Rates of change of eigenvalues and eigenvectors. AIAA J. 1968;6:2426–9.

Fromhage L, Jennions MD. Coevolution of parental investment and sexually selected traits drives sex-role divergence. Nat. Commun. 2016;7:12517.

Gavrilets S. Rapid evolution of reproductive barriers driven by sexual conflict. Nature 2000;403:886–9.

Gavrilets S, Waxman D. Sympatric speciation by sexual conflict. Proc. Natl. Acad. Sci. 2002;99:10533–8.

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

Geritz SAH, Metz JAJ, Rueffler C. Mutual invadability near evolutionarily singular strategies for multivariate traits, with special reference to the strongly convergence stable case. J. Math. Biol. 2016;72:1081–99.

González-Forero M. A mathematical framework for evo-devo dynamics. Theor. Popul. Biol. 2024;155:24–50.

Goodman D. Optimal life histories, optimal notation, and the value of reproductive value. Am. Nat. 1982;119:803–23.

Gross MR. Alternative reproductive strategies and tactics: diversity within sexes. Trends Ecol. Evol. 1996;11:92–8.

Hamilton WD. The moulding of senescence by natural selection. J. Theor. Biol. 1966;12:12–45.

Haykin SS. Neural networks and learning machines. Prentice Hall, 1999.

Houston AI, Fromhage L, McNamara JM. A general framework for modelling trade-offs in adaptive behaviour. Biol. Rev. 2024;99:56–69.

Houston AI, McNamara JM. John Maynard Smith and the importance of consistency in evolutionary game theory. Biol. Philos. 2006;20:933–50.

Inaba H. Age-structured population dynamics in demography and epidemiology. Singapore: Springer, 2017.

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

Kamien MI, Schwartz NL. Dynamic optimization, second edition: the calculus of variations and optimal control in economics and management. Courier Corporation, 2013.

Kierzkowski J, Smoktunowicz A. Block normal matrices and Gershgorin-type discs. Electron. J. Linear Algebra 2011;22:1059–69.

Kimura M. A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proc. Natl. Acad. Sci. 1965;54:731–6.

Kisdi É, Geritz SAH. Adaptive dynamics in allele space: evolution of genetic polymorphism by small mutations in a heterogeneous environment. Evolution 1999;53:993–1008.

Kwiatkowski F, Arbre M, Bidet Y, Laquet C, Uhrhammer N, Bignon YJ. BRCA mutations increase fertility in families at hereditary breast/ovarian cancer risk. PLOS ONE 2015;10:e0127363.

Lande R. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 1979;33:402–16.

Lande R, Arnold SJ. The measurement of selection on correlated characters. Evolution 1983;37:1210–26.

Last G, Penrose M. Lectures on the Poisson process. Cambridge: Cambridge University Press, 2017.

Lehmann L. Hamilton's rule, the evolution of behavior rules and the wizardry of control theory. J. Theor. Biol. 2022;555:111282.

Lehmann L, Mullon C, Akçay E, Van Cleve J. Invasion fitness, inclusive fitness, and reproductive numbers in heterogeneous populations. Evol. Int. J. Org. Evol. 2016;70:1689–702.

Leimar O. Multidimensional convergence stability. Evol. Ecol. Res. 2009;11:191–208.

León JA. Life histories as adaptive strategies. J. Theor. Biol. 1976;60:301–35.

Lepais O, Manicki A, Glise S, Buoro M, Bardonnet A. Genetic architecture of threshold reaction norms for male alternative reproductive tactics in Atlantic salmon (*Salmo salar L.*) Sci. Rep. 2017;7:1–13.

Lesaffre T, Billiard S. The joint evolution of lifespan and self-fertilization. J. Evol. Biol. 2020;33:41–56.

Liberzon D. Calculus of variations and optimal control theory: a concise introduction. Princeton University Press, 2012.

Lindenfors P. Sexually antagonistic selection on primate size. J. Evol. Biol. 2002;15:595–607.

Lion S. Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. Am. Nat. 2018;191:21–44.

Macdonald DW, Johnson PJ. Sex ratio variation and mixed pairs in roe deer: evidence for control of sex allocation? Oecologia 2008;158:361–70.

Mank JE. Sex-specific morphs: the genetics and evolution of intra-sexual variation. Nat. Rev. Genet. 2023;24:44–52.

Massot M, Clobert J. Influence of maternal food availability on offspring dispersal. Behav. Ecol. Sociobiol. 1995;37:413–8.

Méndez M, Obeso JR. Size-dependent reproductive and vegetative allocation in *Arum italicum* (Araceae). Can. J. Bot. 1993;71:309–14.

Metz JAJ, Leimar O. A simple fitness proxy for structured populations with continuous traits, with case studies on the evolution of haplo-diploids and genetic dimorphisms. J. Biol. Dyn. 2011;5:163–90.

Metz JAJ, Nisbet RM, Geritz SAH. How should we define 'fitness' for general ecological scenarios? Trends Ecol. Evol. 1992;7:198–202.

Metz JAJ, Staňková K, Johansson J. The canonical equation of adaptive dynamics for life histories: from fitness-returns to selection gradients and Pontryagin's maximum principle. J. Math. Biol. 2016;72:1125–52.

Michod RE. Evolution of life histories in response to age-specific mortality factors. Am. Nat. 1979;113:531–50.

Mode CJ. Applications of generalized multi-type age-dependent branching processes in population genetics. Bull. Math. Biophys. 1969;31:575–89.

Monaghan P, Charmantier A, Nussey DH, Ricklefs RE. The evolutionary ecology of senescence. Funct. Ecol. 2008;22:371–8.

Nelson RB. Simplified calculation of eigenvector derivatives. AIAA J. 1976;14:1201–5.

Ohtsuki H, Rueffler C, Wakano JY, Parvinen K, Lehmann L. The components of directional and disruptive selection in heterogeneous group-structured populations. J. Theor. Biol. 2020;507:110449.

Parvinen K, Heino M, Dieckmann U. Function-valued adaptive dynamics and optimal control theory. J. Math. Biol. 2013;67:509–33.

Pásztor L, Botta-Dukát Z, Magyar G, Czárán T, Meszéna G. Theory-based ecology: a Darwinian approach. Oxford: Oxford University Press, 2016.

Perrin N, Sibly RM. Dynamic models of energy allocation and investment. Annu. Rev. Ecol. Syst. 1993;24:379–410.

Phillips PC, Arnold SJ. Visualizing multivariate selection. Evolution 1989;43:1209–22.

Poissant J, Wilson AJ, Coltman DW. Sex-specific genetic variance and the evolution of sexual dimorphism: a systematic review of cross-sex genetic correlations. Evolution 2010;64:97–107.

Priklopil T, Lehmann L. Invasion implies substitution in ecological communities with class-structured populations. Theor. Popul. Biol. 2020;134:36–52.

Rao CR, Mitra SK. Generalized inverse of a matrix and its applications. In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Theory of Statistics. Vol. 6.1. University of California Press, 1972:601–21.

Rousset F. Genetic structure and selection in subdivided populations. Princeton University Press, 2004.

Ryan TJ, Plague GR. Hatching asynchrony, survival, and the fitness of alternative adult morphs in *Ambystoma talpoideum*. Oecologia 2004;140:46–51.

Schaffer WM. The application of optimal control theory to the general life history problem. Am. Nat. 1983;121:418–31.

Scheiner SM, Lyman RF. The genetics of phenotypic plasticity. II. Response to selection. J. Evol. Biol. 1991;4:23–50.

Schmid M, Rueffler C, Lehmann L, Mullon C. Resource variation within and between patches: where exploitation competition, local adaptation, and kin selection meet. Am. Nat. 2023;203:E20–E34.

Süli E, Mayers DF. An introduction to numerical analysis. Cambridge: Cambridge University Press, 2003.

Taborsky B, Guyer L, Demus P. 'Prudent habitat choice': a novel mechanism of size-assortative mating. J Evol Biol 2014;27:1217–28.

Taylor PD, Frank SA. How to make a kin selection model. J. Theor. Biol. 1996;180:27–37.

Taylor PD. Allele-frequency change in a class-structured population. Am. Nat. 1990;135:95–106.

Vaupel JW, Baudisch A, Dölling M, A. Roach D, Gampe J. The case for negative senescence. Theor. Popul. Biol. 2004;65:339–51.

Vincent TL, Brown JS. Stability in an evolutionary game. Theor. Popul. Biol. 1984;26:408–27.

Weber TA. Optimal Control Theory with Applications in Economics. Cambridge, Massachusetts London, England: The MIT Press, 2011.

Weissman DB, Judge KA, Williams SC, Whitman DW, Lee VF. Small-male mating advantage in a species of Jerusalem cricket (Orthoptera: Stenopelmatinae: Stenopelmatus). orth 2008;17:321–32.

Wensink MJ, Caswell H, Baudisch A. The rarity of survival to old age does not drive the evolution of senescence. Evol. Biol. 2017;44:5–10.

Wolfram Research, Inc. Mathematica, Version 12. 2019.

Supplementary figures

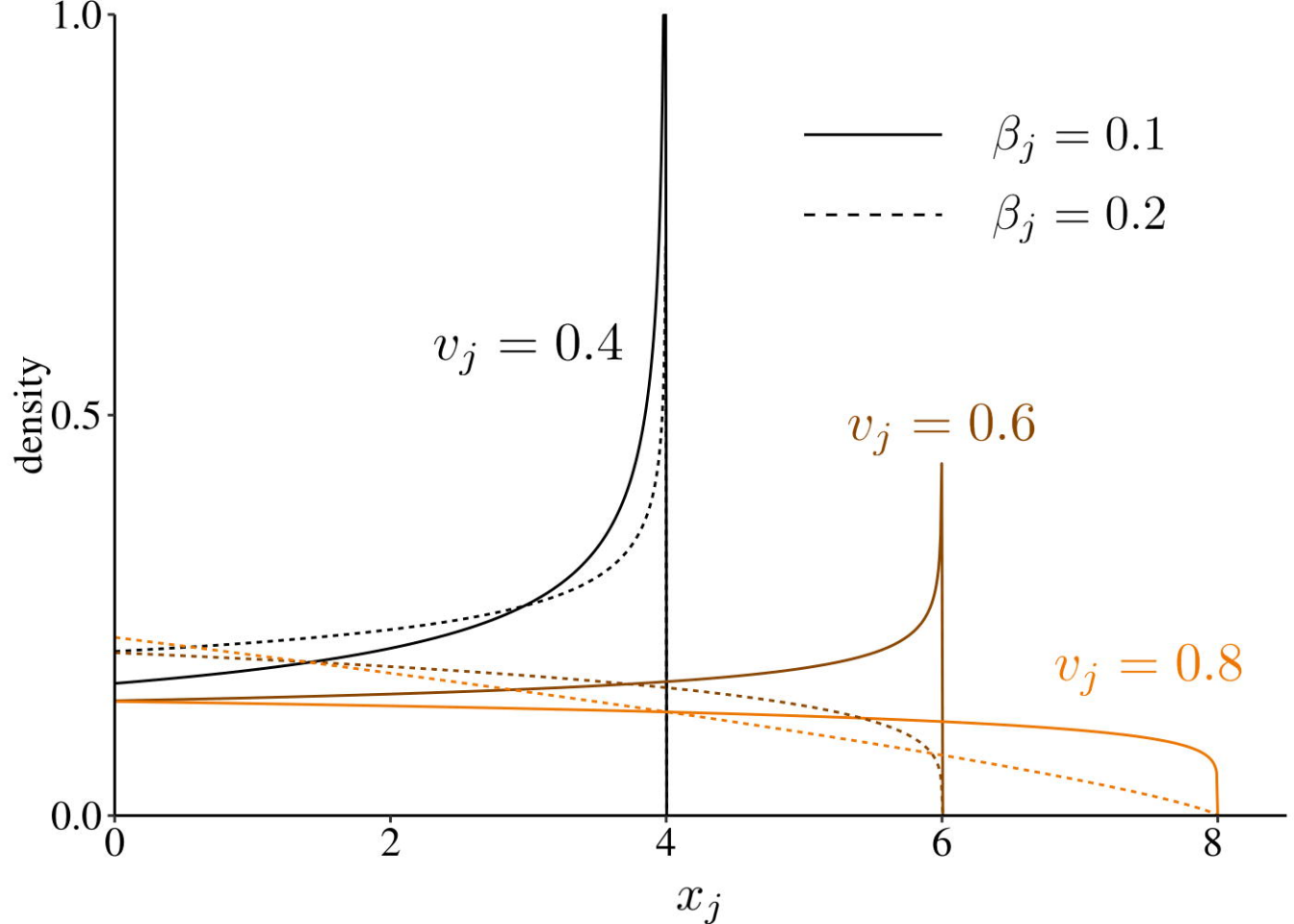
Fig. S1. Male mortality and growth parameters influence male-specific evolution. Each panel shows convergence stable growth rates $\boldsymbol{v}^*$ (solid lines) and corresponding average sizes (dashed lines) for females (red) and males (blue), as a function of the parameter indicated on the x-axis (computed from eq. 36 for trait values and B-99 for average sizes). Trait values under genetic constraint (i.e. identical growth rate in males and females, v_1^*) and the resulting female size are shown in green (from eq. 39). **A**. Higher cost of growth to survival in males (β_{m}) selects for slower growth and smaller average size in males but does not affect females. **B**. Higher male extrinsic mortality ($\mu_{e,m}$) favours faster growth in males but does not affect females. Under genetic constraints, high values of $\mu_{e,m}$ also increase the frequency of females, tipping the balance in favour of even faster growth. Parameter values (unless varied on x-axis): $c = 0.5$, $\alpha = 0.2$, $\mu_{\mathrm{e,f}} = \mu_{\mathrm{e,m}} = 0.1$, $\beta_{\mathrm{f}} = \beta_{\mathrm{m}} = 0.2$, $\kappa = 1$ and $K = 10000$.

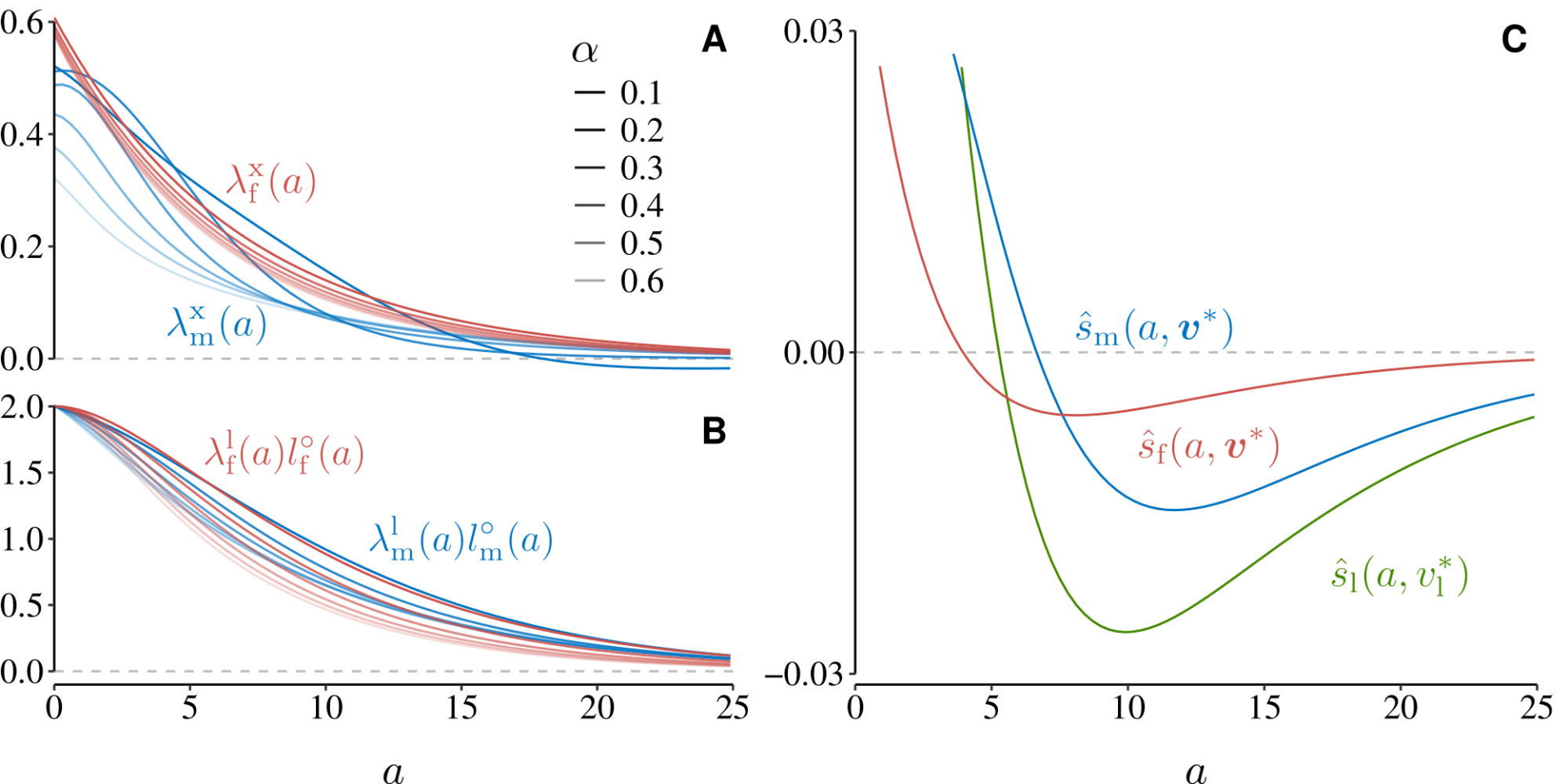
Fig. S2. Multiple branching and the coexistence or more than two male morphs. The figure shows results from individual-based simulations (appendix B.3) with independent growth rates and diploidy. The top part shows the expressed trait values in females (red) and males (blue): randomly sampled allelic values (dots; 20 copies every 10 time units) and population mean (solid lines). The bottom part shows male and female size distributions at equilibrium as histograms (averaged between time steps 10^5 and $2{\times}10^5$). Parameter values: $c = 0.5$, $\alpha = 0.02$, $\mu_{\mathrm{e,f}} = 0.1$, $\mu_{\mathrm{e,m}} = 0.01$, $\beta_{\mathrm{f}} = 0.1$, $\beta_{\mathrm{m}} = 0.01$, $\kappa = 2$ and $K = 10000$.

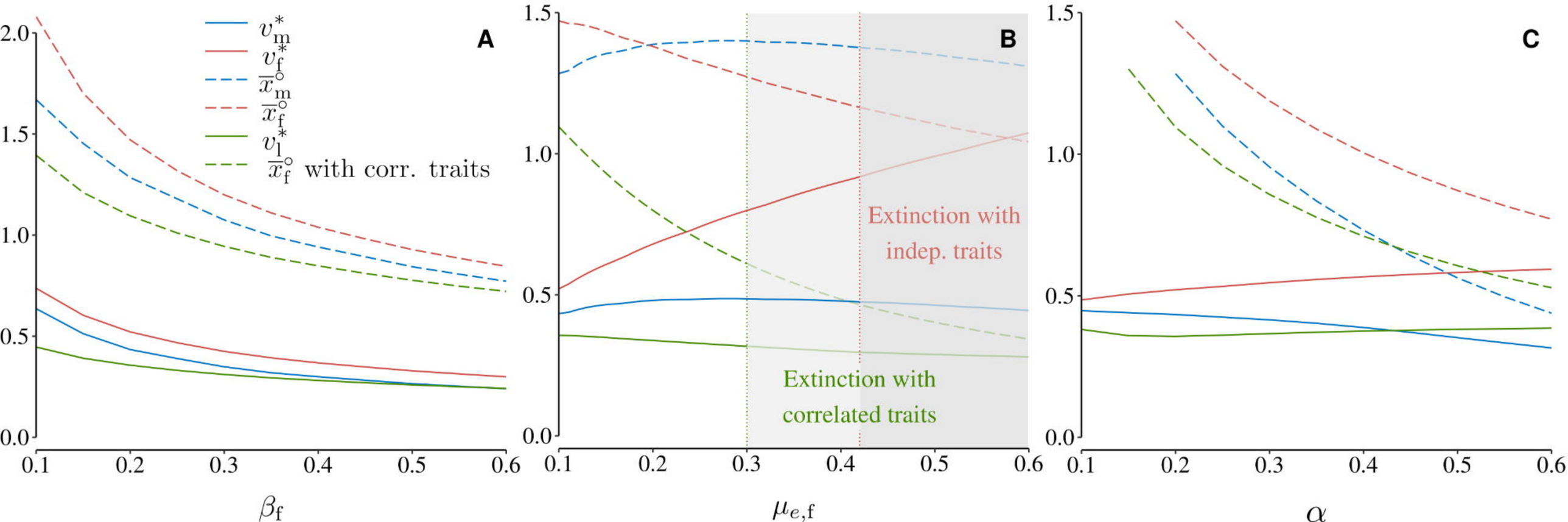
Fig. S3. Evolutionary branching of male growth under haplodiploidy. Each panel shows results from individual-based simulations (appendix B.3). The top part shows the expressed trait values in females (red) and males (blue): randomly sampled allelic values (dots; 10 copies every 10 time units), population mean (solid lines), and analytically predicted convergence stable trait values (dashed lines). Reported values give: the male-specific component $h_{\mathrm{mm}}(\boldsymbol{v}^*)$ (eq. 14); its size-mediated component $h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = 0.25 \int_0^\infty \hat{h}^{\mathrm{xx}}_{\mathrm{mm}}(a, \boldsymbol{v}^*)\,\mathrm{d}a$ (see eqs. 37 and A-37, with 0.25 for diploidy); and the leading eigenvalue of $\mathbf{H}(\boldsymbol{v}^*)$. As expected, when this eigenvalue is negative, the population remains monomorphic (panels A–C). When it is positive, evolutionary branching occurs in males (panel D, split in blue lines). The bottom part of each panel shows male and female size distributions at equilibrium as histograms (averaged between time steps 10^5 and $2{\times}10^5$), along with the analytically predicted distribution at $\boldsymbol{v}^*$ (solid lines, eq. 26). The Kullback–Leibler divergence between male and female distributions is also reported. **A–B**: Male and female distributions are similar in shape as both sexes have the same demographic parameters (see left-hand side of panels for parameter values). However, females grow larger on average due to direct selection on size whereas males only track female size via selection for compatibility. **C–D**: Sex-specific demography generates asymmetries in size distributions with a higher frequency of small females. When sexual selection is strong enough, this leads to evolutionary branching (panel D) with one morph specialising on mating with smaller females. Under haplodiploidy, net disruptive selection is stronger at the singular trait than in the diploid case, and the smaller male morph evolves smaller genetic values (compare values in fig 4D and S3D), but the distribution of male size at the evolutionary equilibrium is similar. Other parameters: $c = 0.5$, $\mu_{\mathrm{e,f}} = 0.1$, $K = 10^4$.

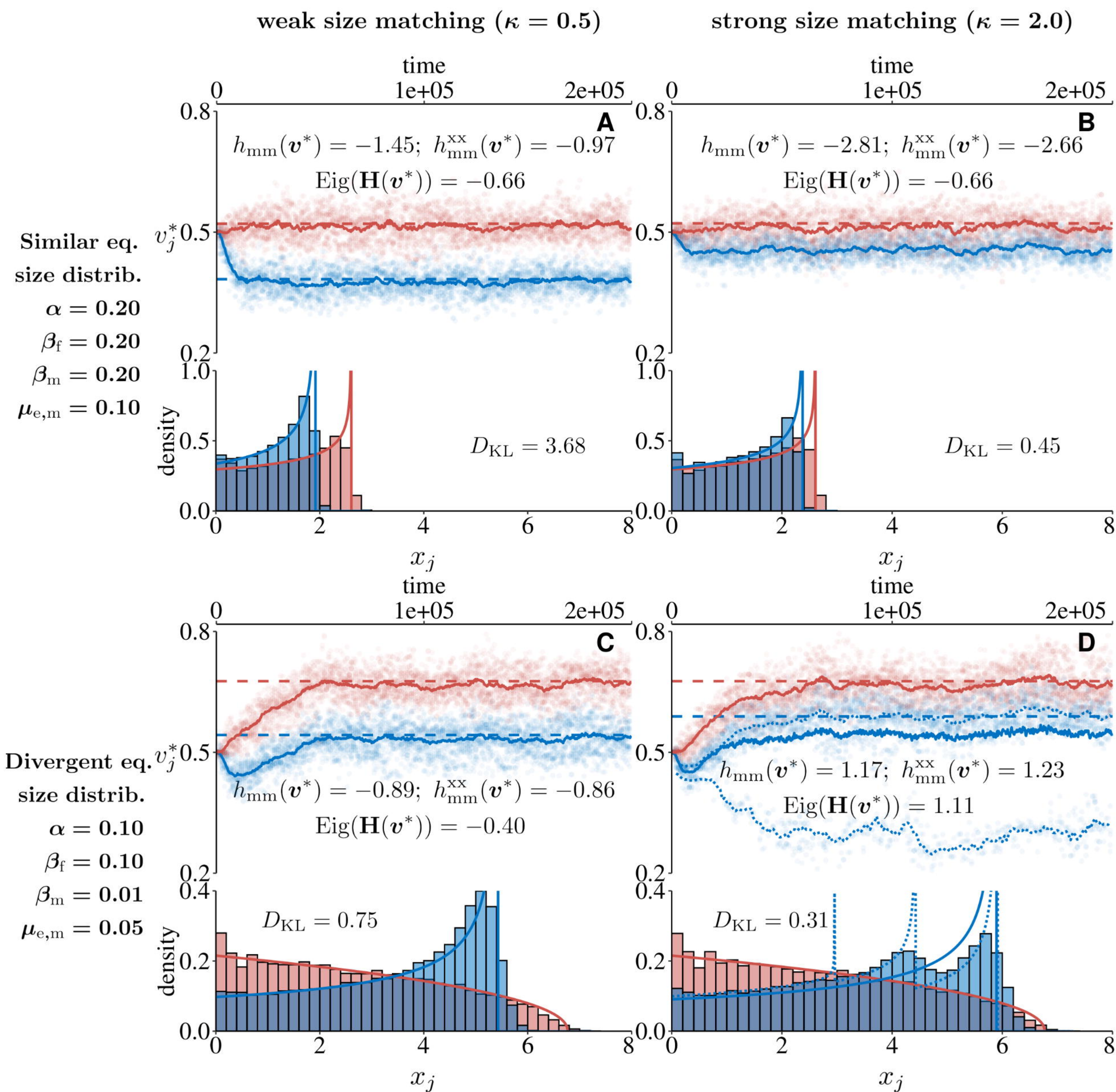
Fig. S4. Age-specific quadratic selection. Age-specific quadratic selection on growth (from eqs. 35a and 35b) at the convergence stable strategies ($\boldsymbol{v}^* = (v^*_{\mathrm{f}}, v^*_{\mathrm{m}}) = (0.676, 0.589)$ for independent traits; $v^*_{\mathrm{l}} = 0.219$ for a shared trait), under parameter values that produce disruptive selection on the male growth rate ($c = 0.5$, $\alpha = 0.1$, $\mu_{\mathrm{e,f}} = 0.1$, $\mu_{\mathrm{e,m}} = 0.05$, $\beta_{\mathrm{f}} = 0.1$, $\beta_{\mathrm{m}} = 0.01$, $\kappa = 2$ and $K = 10^4$, see fig. 4D). Plain lines correspond to the male and female age-specific components of disruptive selection under independent traits (given by eq. 15; recall that here $\hat{h}_{jj} = \hat{h}_{jj,11}$). Dashed lines represent the same components, as well as their sum (green line) under genetic constraints (see eq. A-86).

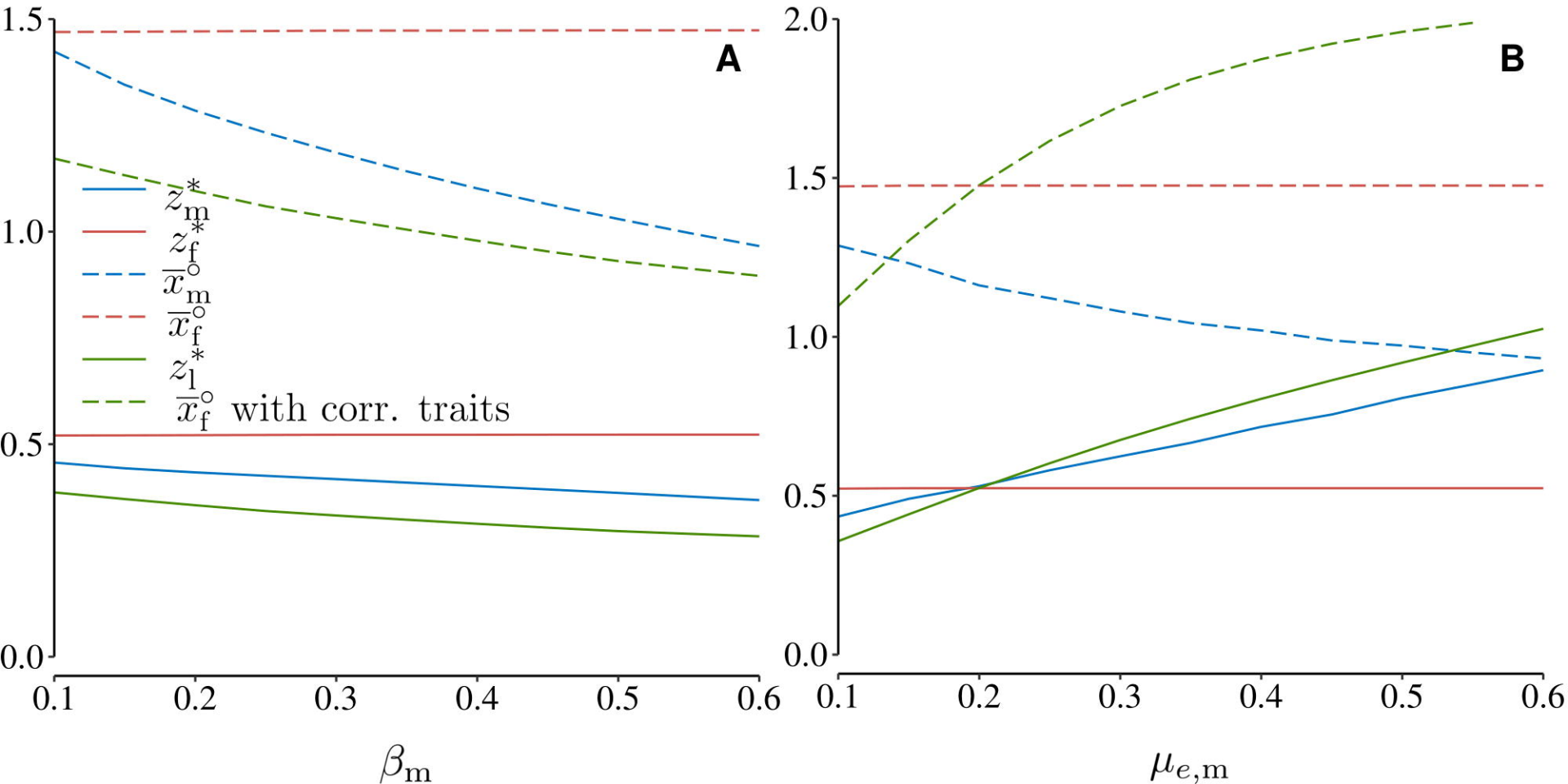
Fig. S5. Genetic constraints between sexes prevent evolutionary branching. Each panel shows results from individual-based simulations (appendix B.3). The top part shows the expressed trait values in females and males (green): randomly sampled allelic values (dots; 10 copies every 10 time units), population mean (solid lines), and analytically predicted convergence stable trait values (dashed lines). Reported values give: the size-mediated component of quadratic selection $h_{\mathrm{ll}}^{\mathrm{xx}}(v_{\mathrm{l}}^{*}) = 0.25 \int_0^\infty \hat{h}_{\mathrm{ll}}^{\mathrm{xx}}(a, \boldsymbol{v}^{*})\mathrm{d}a$ (where $\hat{h}_{\mathrm{ll}}^{\mathrm{xx}}(a, v_{\mathrm{l}}^{*})$ is computed as in eq. A-86 using eq. A-37, with 0.25 for diploidy); and the leading eigenvalue of $\mathbf{H}(v_{\mathrm{l}}^{*})$. Both are always negative showing size-mediated and net stabilising selection on growth, respectively. The bottom part of each panel shows male and female size distributions at equilibrium as histograms (averaged between time steps 10^5 and $2{\times}10^5$), along with the analytically predicted distribution at v_{l}^{*} (solid lines, eq. 26). The Kullback–Leibler divergence between male and female distributions is also reported. **A–B**: Male and female distributions are the same since both sexes have the same growth rate and demographic parameters (see left-hand side of panels for parameter values). **C–D**: Sex-specific demography generates asymmetries in size distributions with a higher frequency of small females. Under genetic constraints, these differences are not sufficient to generate evolutionary branching. Other parameters: $c = 0.5$, $\mu_{\mathrm{e,f}} = 0.1$ and $K = 10000$, and with simulated time steps of 0.2 time units.

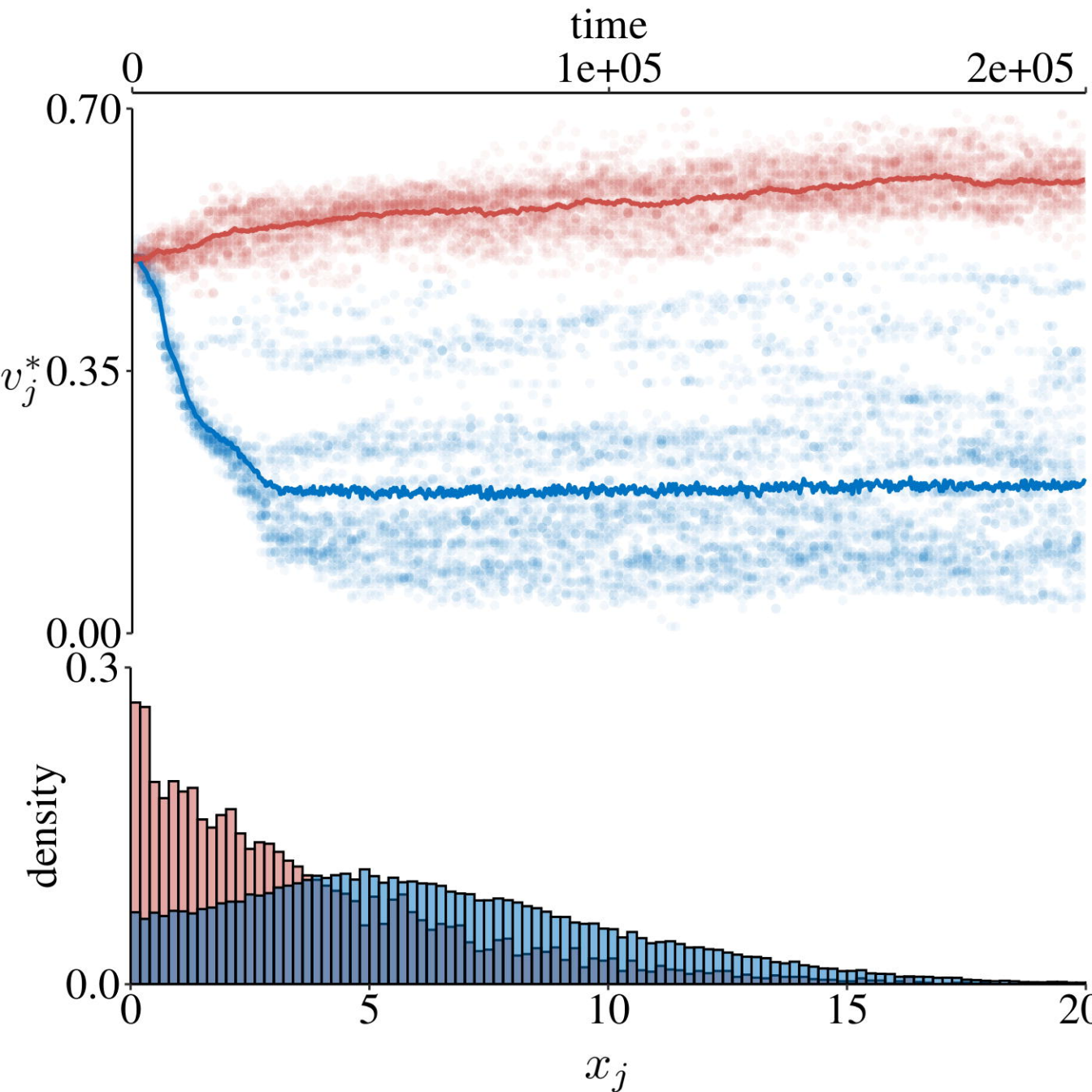

1.0
0.5
0.0
density
0
2
4
6
8
x_j
$\beta_j = 0.1$
$\beta_j = 0.2$
$v_j = 0.4$
$v_j = 0.6$
$v_j = 0.8$

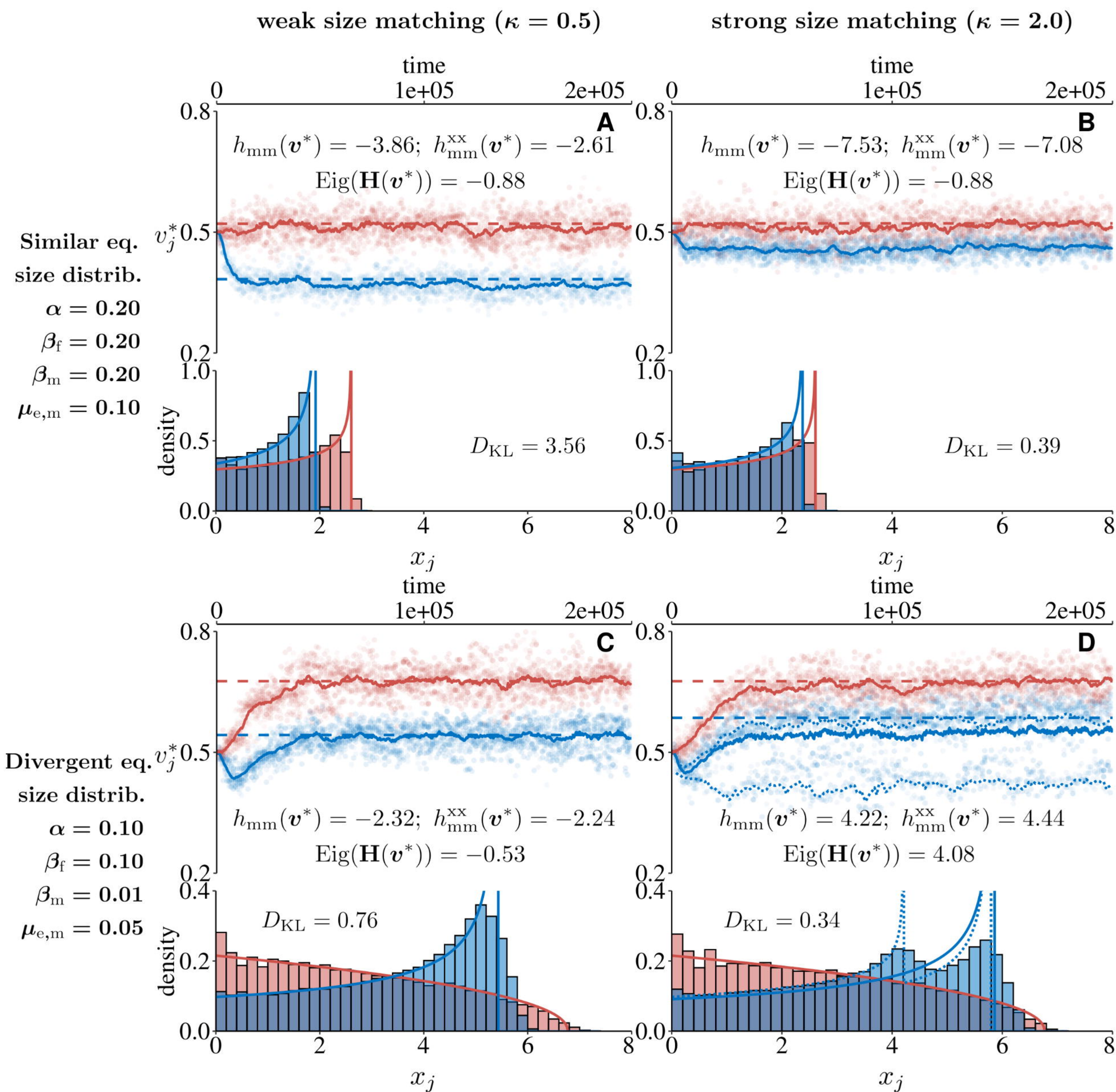

A
B
C
α
0.1
0.2
0.3
0.4
0.5
0.6
$\lambda_{\mathrm{f}}^{\mathrm{x}}(a)$
$\lambda_{\mathrm{m}}^{\mathrm{x}}(a)$
$\lambda_{\mathrm{f}}^{\mathrm{l}}(a)l_{\mathrm{f}}^{\circ}(a)$
$\lambda_{\mathrm{m}}^{\mathrm{l}}(a)l_{\mathrm{m}}^{\circ}(a)$
$\hat{s}_{\mathrm{m}}(a, \boldsymbol{v}^*)$
$\hat{s}_{\mathrm{f}}(a, \boldsymbol{v}^*)$
$\hat{s}_{\mathrm{l}}(a, v_{\mathrm{l}}^*)$
a
a

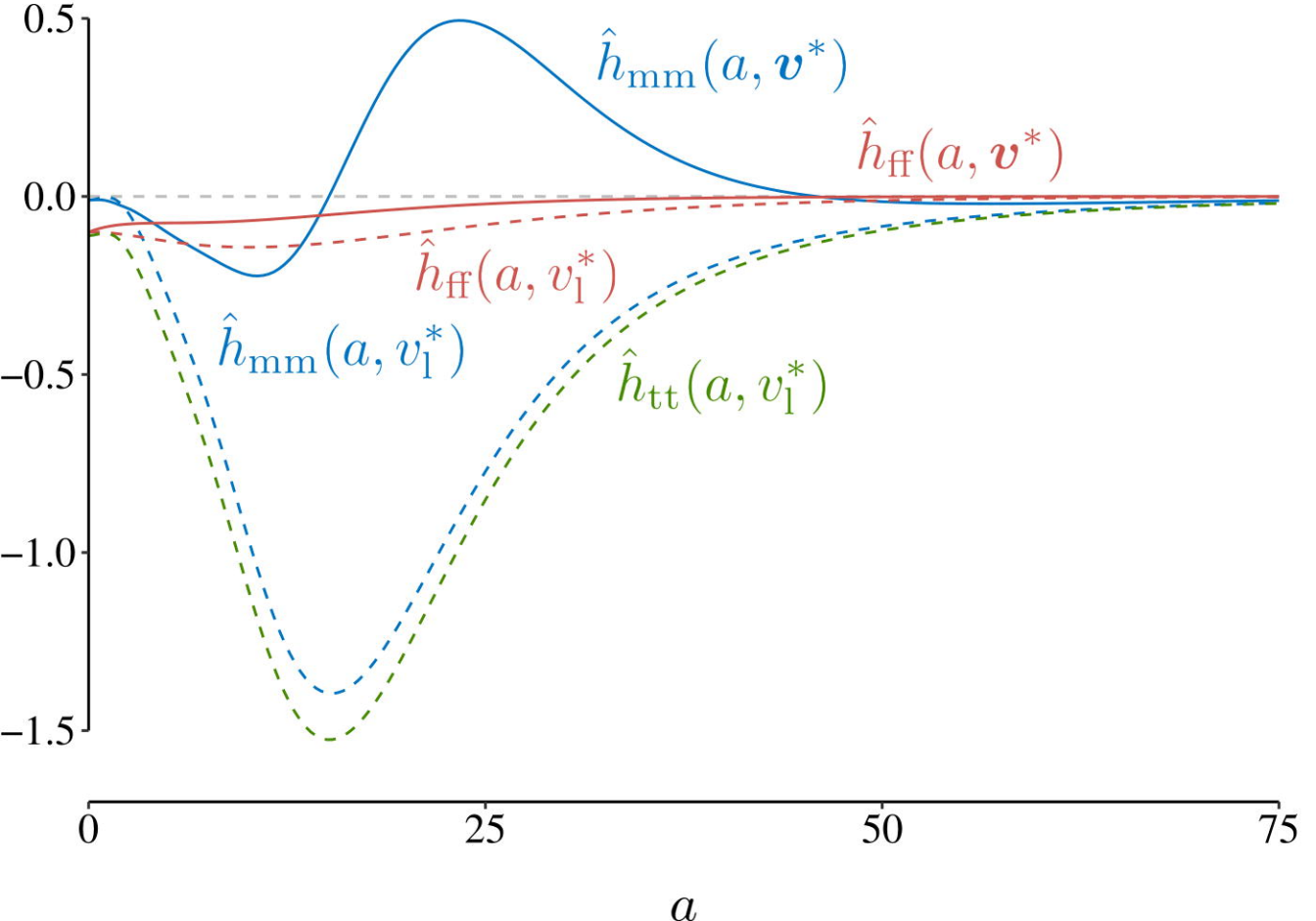

A
B
C
v_m^*
v_f^*
$\overline{x}_m^\circ$
$\overline{x}_f^\circ$
v_l^*
$\overline{x}_f^\circ$ with corr. traits
Extinction with indep. traits
Extinction with correlated traits
β_f
$\mu_{e,f}$
α
0.0
0.5
1.0
1.5
2.0
0.1
0.2
0.3
0.4
0.5
0.6

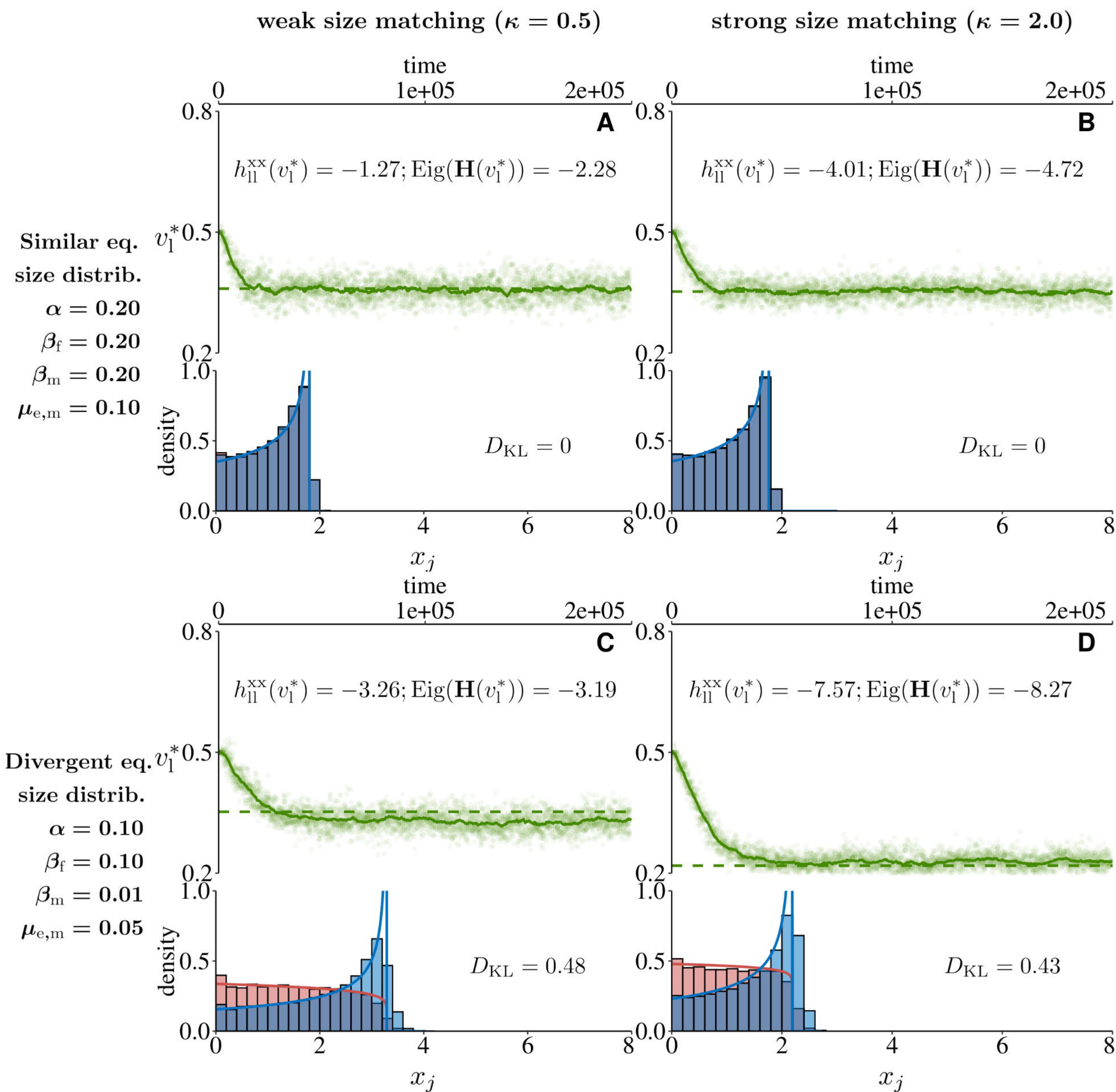

weak size matching ($\kappa = 0.5$)
strong size matching ($\kappa = 2.0$)
time
0
1e+05
2e+05
A
B
C
D
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = -1.45;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = -0.97$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = -0.66$
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = -2.81;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = -2.66$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = -0.66$
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = -0.89;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = -0.86$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = -0.40$
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = 1.17;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = 1.23$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = 1.11$
v_j^*
density
$D_{\mathrm{KL}} = 3.68$
$D_{\mathrm{KL}} = 0.45$
$D_{\mathrm{KL}} = 0.75$
$D_{\mathrm{KL}} = 0.31$
x_j
Similar eq. size distrib.
$\alpha = 0.20$
$\beta_{\mathrm{f}} = 0.20$
$\beta_{\mathrm{m}} = 0.20$
$\mu_{\mathrm{e,m}} = 0.10$
Divergent eq. size distrib.
$\alpha = 0.10$
$\beta_{\mathrm{f}} = 0.10$
$\beta_{\mathrm{m}} = 0.01$
$\mu_{\mathrm{e,m}} = 0.05$


A
B
z_{m}^*
z_{f}^*
$\overline{x}_{\mathrm{m}}^{\circ}$
$\overline{x}_{\mathrm{f}}^{\circ}$
z_{l}^*
$\overline{x}_{\mathrm{f}}^{\circ}$ with corr. traits
1.5
1.0
0.5
0.0
2.0
1.5
1.0
0.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
β_{m}
$\mu_{e,\mathrm{m}}$


time
0
1e+05
2e+05
0.70
0.35
0.00
v_j^*
0.3
0.0
density
0
5
10
15
20
x_j


weak size matching ($\kappa = 0.5$)
strong size matching ($\kappa = 2.0$)
time
0
1e+05
2e+05
A
B
C
D
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = -3.86;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = -2.61$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = -0.88$
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = -7.53;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = -7.08$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = -0.88$
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = -2.32;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = -2.24$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = -0.53$
$h_{\mathrm{mm}}(\boldsymbol{v}^*) = 4.22;\ h^{\mathrm{xx}}_{\mathrm{mm}}(\boldsymbol{v}^*) = 4.44$
$\mathrm{Eig}(\mathbf{H}(\boldsymbol{v}^*)) = 4.08$
v_j^*
density
$D_{\mathrm{KL}} = 3.56$
$D_{\mathrm{KL}} = 0.39$
$D_{\mathrm{KL}} = 0.76$
$D_{\mathrm{KL}} = 0.34$
x_j
Similar eq. size distrib.
$\boldsymbol{\alpha} = \mathbf{0.20}$
$\boldsymbol{\beta}_{\mathrm{f}} = \mathbf{0.20}$
$\boldsymbol{\beta}_{\mathrm{m}} = \mathbf{0.20}$
$\boldsymbol{\mu}_{\mathrm{e,m}} = \mathbf{0.10}$
Divergent eq. size distrib.
$\boldsymbol{\alpha} = \mathbf{0.10}$
$\boldsymbol{\beta}_{\mathrm{f}} = \mathbf{0.10}$
$\boldsymbol{\beta}_{\mathrm{m}} = \mathbf{0.01}$
$\boldsymbol{\mu}_{\mathrm{e,m}} = \mathbf{0.05}$


0.5
0.0
−0.5
−1.0
−1.5
0
25
50
75
a
$\hat{h}_{\mathrm{mm}}(a, \boldsymbol{v}^*)$
$\hat{h}_{\mathrm{ff}}(a, \boldsymbol{v}^*)$
$\hat{h}_{\mathrm{ff}}(a, v_{\mathrm{l}}^*)$
$\hat{h}_{\mathrm{mm}}(a, v_{\mathrm{l}}^*)$
$\hat{h}_{\mathrm{tt}}(a, v_{\mathrm{l}}^*)$


weak size matching ($\kappa = 0.5$)
strong size matching ($\kappa = 2.0$)
time
0
1e+05
2e+05
A
B
C
D
$h_{ll}^{xx}(v_l^*) = -1.27; \mathrm{Eig}(\mathbf{H}(v_l^*)) = -2.28$
$h_{ll}^{xx}(v_l^*) = -4.01; \mathrm{Eig}(\mathbf{H}(v_l^*)) = -4.72$
$h_{ll}^{xx}(v_l^*) = -3.26; \mathrm{Eig}(\mathbf{H}(v_l^*)) = -3.19$
$h_{ll}^{xx}(v_l^*) = -7.57; \mathrm{Eig}(\mathbf{H}(v_l^*)) = -8.27$
v_l^*
0.8
0.5
0.2
Similar eq.
size distrib.
$\alpha = 0.20$
$\beta_f = 0.20$
$\beta_m = 0.20$
$\mu_{e,m} = 0.10$
Divergent eq.
size distrib.
$\alpha = 0.10$
$\beta_f = 0.10$
$\beta_m = 0.01$
$\mu_{e,m} = 0.05$
density
1.0
0.5
0.0
$D_{KL} = 0$
$D_{KL} = 0.48$
$D_{KL} = 0.43$
0
2
4
6
8
x_j