Highlights

A mathematical framework for evo-devo dynamics

Mauricio González-Forero

- We formulate a framework integrating evolutionary and developmental dynamics.
- We derive equations describing the evolutionary dynamics of traits considering their developmental process.
- This yields a description of the evo-devo process in terms of closed-form formulas that are simple and insightful, including for genetic covariance matrices.

A mathematical framework for evo-devo dynamics

Mauricio González-Forero[a,*]

[a] *School of Biology, University of St Andrews, Dyers Brae, St Andrews, KY16 9TH, Fife, UK*

Abstract

Natural selection acts on phenotypes constructed over development, which raises the question of how development affects evolution. Classic evolutionary theory indicates that development affects evolution by modulating the genetic covariation upon which selection acts, thus affecting genetic constraints. However, whether genetic constraints are relative, thus diverting adaptation from the direction of steepest fitness ascent, or absolute, thus blocking adaptation in certain directions, remains uncertain. This limits understanding of long-term evolution of developmentally constructed phenotypes. Here we formulate a general tractable mathematical framework that integrates age progression, explicit development (i.e., the construction of the phenotype across life subject to developmental constraints), and evolutionary dynamics, thus describing the evolutionary developmental (evo-devo) dynamics. The framework yields simple equations that can be arranged in a layered structure that we call the evo-devo process, whereby five core elementary components generate all equations including those mechanistically describing genetic covariation and the evo-devo dynamics. The framework recovers evolutionary dynamic equations in gradient form and describes the evolution of genetic covariation from the evolution of genotype, phenotype, environment, and mutational covariation. This shows that genotypic and phenotypic evolution must be followed simultaneously to yield a dynamically sufficient description of long-term phenotypic evolution in gradient form, such that evolution described as the climbing of a fitness landscape occurs in "geno-phenotype" space. Genetic constraints in geno-phenotype space are necessarily absolute because the phenotype is related to the genotype by development. Thus, the long-term evolutionary dynamics of developed phenotypes is strongly non-standard: (1) evolutionary equilibria are either absent or infinite in number and depend on genetic covariation and hence on development; (2) developmental constraints determine the admissible evolutionary path and hence which evolutionary equilibria are admissible; and (3) evolutionary outcomes occur at admissible evolutionary equilibria, which do not generally occur at fitness landscape peaks in geno-phenotype space, but at peaks in the admissible evolutionary path where "total genotypic selection" vanishes if exogenous plastic response vanishes and mutational variation exists in all directions of genotype space. Hence, selection and development jointly define the evolutionary outcomes if absolute mutational constraints and exogenous plastic response are absent, rather than the outcomes being defined only by selection. Moreover, our framework provides formulas for the sensitivities of a recurrence and an alternative method to dynamic optimization (i.e., dynamic programming or optimal control) to identify evolutionary outcomes in models with developmentally dynamic traits. These results show that development has major evolutionary effects.

Keywords:
evolutionary dynamics, developmental constraints, genetic constraints, life-history theory, matrix population models, adaptive dynamics

1. Introduction

Development may be defined as the process that constructs the phenotype over life (Barresi and Gilbert, 2020). In particular, development includes "the process by which genotypes are transformed into phenotypes" (Wolf et al., 2001). As natural selection screens phenotypes produced by development, a fundamental evolutionary problem concerns how development affects evolution. Interest in this problem is long-standing (Baldwin 1896, Waddington 1959 p. 399, and Gould and Lewontin 1979) and has steadily increased in recent decades. It has been proposed that developmental constraints (Gould and Lewontin, 1979; Maynard Smith et al., 1985; Brakefield, 2006; Klingenberg, 2010), causal feedbacks over development occurring among genes, the organism, and environment (Lewontin, 1983; Rice, 2011; Hansen, 2013; Laland et al., 2015), and various development-mediated factors (Laland et al., 2014, 2015), namely plasticity (Pigliucci, 2001; West-Eberhard, 2003), niche construction (Odling-Smee et al., 1996, 2003), extra-genetic inheritance (Baldwin, 1896; Cavalli-Sforza and Feldman, 1981; Boyd and Richerson, 1985; Jablonka and Lamb, 2014; Bonduriansky and Day, 2018), and developmental bias (Arthur, 2004; Uller et al., 2018), may all have important evolutionary roles. Understanding how development — including these elements acting individually and together — affects the evolutionary process remains an outstanding challenge (Baldwin, 1896; Waddington, 1959; Müller, 2007; Pigliucci, 2007; Laland et al., 2014, 2015; Galis et al., 2018).

Classic evolutionary theory indicates that development affects evolution by modulating the genetic covariation upon which selection acts. This can be seen as follows. In quantitative genetics, an individual's i-th trait value x_i is written as $x_i = \bar{x}_i + \sum_j \alpha_{ij}(y_j - \bar{y}_j) + e_i$, where the overbar denotes population average, y_j is the individual's gene content at the j-th locus, α_{ij} is the partial regression coefficient of the i-th trait deviation from the average on the deviation from the average of the j-th locus content, and e_i is the residual (Fisher, 1918; Crow and Kimura, 1970; Falconer and Mackay, 1996; Lynch and Walsh, 1998; Walsh and Lynch, 2018). The quantity α_{ij} is Fisher's additive effect of allelic substitution (his α; see Eq. I of Fisher 1918 and p. 72 of Lynch and Walsh 1998) and is a description of some of the linear effects of development, specifically of how genotypic change is transformed into phenotypic change. In matrix notation, the vector of an individual's trait values is $\mathbf{x} = \bar{\mathbf{x}} + \boldsymbol{\alpha}(\mathbf{y} - \bar{\mathbf{y}}) + \mathbf{e}$, where the matrix $\boldsymbol{\alpha}$ corresponds to what Wagner (1984) calls the developmental matrix (his $\mathbf{B}$). The breeding value of the multivariate phenotype $\mathbf{x}$ is defined as $\mathbf{a_x} \equiv \bar{\mathbf{x}} + \boldsymbol{\alpha}(\mathbf{y} - \bar{\mathbf{y}})$, which does not consider the residual that includes non-linear effects of genes on phenotype. Breeding value thus depends on development via the developmental matrix $\boldsymbol{\alpha}$. The Lande (1979)

*Corresponding author:
Email address: mgf3@st-andrews.ac.uk (Mauricio González-Forero)

equation describes the evolutionary change due to selection in the mean multivariate phenotype $\bar{\mathbf{x}}$ as $\Delta\bar{\mathbf{x}} = \mathbf{G}\partial \ln \bar{W}/\partial\bar{\mathbf{x}}$, where the additive genetic covariance matrix is $\mathbf{G} \equiv \mathrm{cov}[\mathbf{a}_{\mathbf{x}}, \mathbf{a}_{\mathbf{x}}] = \boldsymbol{\alpha}\mathrm{cov}[\mathbf{y}, \mathbf{y}]\boldsymbol{\alpha}^{\mathsf{T}}$ (e.g., Wagner 1984), mean absolute fitness is $\bar{W}$, and the selection gradient is $\partial \ln \bar{W}/\partial\bar{\mathbf{x}}$, which points in the direction of steepest increase in mean fitness (here and throughout we use matrix calculus notation described in Appendix A). An important feature of the Lande equation is that it is in gradient form, so the equation shows that, within the assumptions made, phenotypic evolution by natural selection proceeds as the climbing of a fitness landscape. Moreover, the Lande equation shows that additive genetic covariation, described by $\mathbf{G}$, may divert evolutionary change from the direction of steepest fitness ascent, and may prevent evolutionary change in some directions if genetic variation in those directions is absent (in which case $\mathbf{G}$ is singular). Since additive genetic covariation depends on development via the developmental matrix $\boldsymbol{\alpha}$, the Lande equation shows that development affects evolution by modulating genetic covariation via $\boldsymbol{\alpha}$ (Charlesworth et al., 1982; Cheverud, 1984; Maynard Smith et al., 1985).

However, the Lande equation might provide a limited description of how development affects evolution for two reasons. First, the Lande equation does not explicitly consider phenotype construction, that is, development, but rather a regression-based description of development in terms of $\boldsymbol{\alpha}$. This limits the possibility of translating a mechanistic understanding of development, for instance, expressed as developmentally dynamic equations, into a mechanistic understanding of genetic covariation. Second, the Lande equation describes short-term evolution by assuming negligible allele frequency change (Walsh and Lynch, 2018, pp. 504 and 879). More specifically, the Lande equation describes the evolution of mean traits $\bar{\mathbf{x}}$ but not of mean gene content $\bar{\mathbf{y}}$, that is, it does not describe change in allele frequency but it depends on allele frequency; for instance, since $\boldsymbol{\alpha}$ is a matrix of regression coefficients calculated for the current population, $\boldsymbol{\alpha}$ depends on the current state of the population including allele frequency $\bar{\mathbf{y}}$. The Lande equation considers negligible genetic evolution by postulating Fisher's (1918) infinitesimal model, whereby each trait is assumed to be controlled by an arbitrarily large number of loci such that allele frequency change per locus per generation is negligible (Bulmer, 1971, 1980; Turelli and Barton, 1994; Barton et al., 2017; Hill, 2017). Yet, over the long-term there is genetic evolution that is mapped to phenotypic evolution via development, but by assuming negligible genetic evolution the effect of this mapping might not be revealed by the Lande equation.

There is a lack of equations describing the long-term phenotypic evolutionary dynamics in gradient form while explicitly considering the developmental dynamics of phenotype construction and how development translates into genetic covariation. Various research lines have enabled the mathematical modelling of either long-term phenotypic evolutionary dynamics, age progression, development, or how development translates into genetic covariation, without including all these elements at the same time. Both the classic Lande equation (Lande, 1979) and the classic canonical equation of adaptive dynamics (Dieckmann and Law, 1996) describe the evolutionary dynamics of a multivariate trait in gradient form without an explicit account of development, by considering no explicit age progression nor development. Extensions to these equations have considered explicit age progression by implementing age structure, which allows individuals of different ages to coexist and to have age-specific survival and fertility rates (Lande, 1982; Charlesworth, 1993, 1994; Durinx et al., 2008). An important feature of age-structured models is that the forces

of selection decline with age due to demography, in particular due to mortality and fewer remaining reproductive events as age advances (Medawar, 1952; Hamilton, 1966; Caswell, 1978; Caswell and Shyu, 2017). Although age progression is often taken as describing development, age structure does not explicitly describe development defined as phenotype construction.

Another research line in life-history theory has extended age-structured models to explicitly consider phenotype construction (Gadgil and Bossert, 1970; Taylor et al., 1974; León, 1976; Schaffer, 1983; Houston et al., 1988; Roff, 1992; Houston and McNamara, 1999; Sydsæter et al., 2008). This line has considered models with two types of age-specific traits: genotypic traits called control variables, which are under direct genetic control, and developed traits called state variables, which are constructed according to developmentally dynamic equations called dynamic constraints. This explicit consideration of development in an evolutionary context has mostly assumed that the population is at an evolutionary equilibrium. Thus, this approach identifies evolutionarily stable (or uninvadable) controls and associated states using techniques from dynamic optimization such as optimal control and dynamic programming (Gadgil and Bossert, 1970; Taylor et al., 1974; León, 1976; Schaffer, 1983; Houston et al., 1988; Roff, 1992; Houston and McNamara, 1999). While the assumption of evolutionary equilibrium yields great insight, it does not address the evolutionary dynamics which would provide a richer understanding. Moreover, the relationship between developmental dynamics and genetic covariation is not made evident by this approach.

Yet another research line has made it possible to mathematically model the evolutionary dynamics of developmentally constructed traits, but this has remained prohibitively challenging. A first step in this research line has been to consider function-valued or infinite-dimensional traits, which are genotypic traits indexed by a continuous variable (e.g., age) rather than a discrete variable as in the classic Lande equation. Thus, the evolutionary dynamics of univariate function-valued traits (e.g., body size across continuous age) has been described in gradient form by the Lande equation for function-valued traits (Kirkpatrick and Heckman, 1989) and the canonical equation for function-valued traits (Dieckmann et al., 2006). Although function-valued traits may depend on age, they are not constructed according to developmentally dynamic equations, so the consideration of the evolutionary dynamics of function-valued traits alone does not model the evolutionary dynamics of developmentally constructed traits. To our knowledge, Parvinen et al. (2013) were the first to mathematically model the evolutionary dynamics of developmentally constructed traits (the first to do it analytically; there have also been numerical models, for instance, integrating mathematical modeling of the developmental dynamics and individual-based modeling of the evolutionary dynamics; Salazar-Ciudad and Marín-Riera, 2013 and Watson et al., 2013). Parvinen et al. (2013) did so by considering the evolutionary dynamics of a univariate function-valued trait (control variable) that modulates the developmental construction of a multivariate developed trait (state variables; they refer to these as process-mediated models). However, the analysis of these models poses substantial technical challenges, by requiring calculation of functional derivatives and the solution of (partial) differential equations at evolutionary equilibrium with terminal age conditions in addition to the equations describing the developmental dynamics with initial age conditions (a two-point boundary value problem) (Dieckmann et al., 2006; Parvinen et al., 2013; Metz et al., 2016; Avila et al., 2021). Moreover, these models have yielded evolutionary

dynamic equations in gradient form for genotypic traits, but not for developed traits (Dieckmann et al., 2006), so they have left unanswered the question of how the evolution of traits constructed via developmental dynamics proceeds in the fitness landscape. Additionally, these models have not provided a link between developmental dynamics and genetic covariation (Metz 2011; Dieckmann et al. 2006 discuss a link between constraints and genetic covariation in controls, not states; see Supplementary Information (SI) section S1.1 for further details).

Finally, a separate research line in quantitative genetics has analyzed how implicit development translates into genetic covariation. This line has considered models where a set of traits are functions of underlying traits such as gene content or environmental variables (Wagner, 1984, 1989; Hansen and Wagner, 2001; Rice, 2002; Martin, 2014; Morrissey, 2014, 2015). This dependence of traits on other traits is used by this research line to describe development and the genotype-phenotype map. However, this research line considers short-term evolution, no explicit age progression, and no explicit development (i.e., no developmentally dynamic phenotype construction). Although this line has provided an understanding of how implicit development translates into genetic covariation, it might miss evolutionary effects of the mapping of genotype to phenotype provided by development because of its assumption of negligible genetic evolution. Furthermore, this line has not provided equations describing the long-term evolution in gradient form of traits constructed according to explicit developmental dynamics.

Here we formulate a tractable mathematical framework to model the evolutionary dynamics of developmentally constructed traits. The framework provides closed-form equations to model the evolutionary dynamics of genotypic traits and the concomitant developmental dynamics of developed traits subject to developmental constraints. Moreover, the framework provides equations describing the long-term evolutionary dynamics in gradient form for developmentally constructed traits and equations that translate explicit development into genetic covariation. The framework is based on adaptive dynamics assumptions (Dieckmann and Law, 1996; Metz et al., 1996; Champagnat, 2006; Durinx et al., 2008). We consider general developmental dynamics that allow the phenotype to be "predisposed" to develop in certain ways, thus allowing for developmental bias (Arthur, 2004; Uller et al., 2018). We allow development to depend on the environment, which allows for a mechanistic description of plasticity (Pigliucci, 2001; West-Eberhard, 2003). We also allow development to depend on social interactions, which allows for a mechanistic description of extra-genetic inheritance (Boyd and Richerson, 1985; Jablonka and Lamb, 2014; Bonduriansky and Day, 2018) and indirect genetic effects (Moore et al., 1997). Social development entails that a mutant's phenotype may change as the mutant genotype spreads, which complicates evolutionary invasion analysis. In turn, we allow the environment faced by each individual to depend on the traits of the individual and of social partners, thus allowing for individual and social niche construction although we do not consider ecological inheritance (Odling-Smee et al., 1996, 2003). We also let the environment depend on processes that are exogenous to the evolving population, such as eutrophication or climate change caused by members of other species, thus allowing for exogenous environmental change. To facilitate analysis, we let population dynamics occur over a short time scale, whereas environmental and evolutionary dynamics occur over a long time scale. Crucially, we measure age in discrete time, which simplifies the mathematics yielding closed-form formulas for otherwise implicitly defined quantities. Our methods use concepts from optimal control

(Sydsæter et al., 2008) and integrate tools from adaptive dynamics (Dieckmann and Law, 1996) and matrix population models (Caswell, 2001; Otto and Day, 2007). In particular, we conceptualize the genotype as "control" variables and the phenotype as "state" variables whose developmental dynamics are modulated by controls. While we use concepts from optimal control, we do not use optimal control itself. Instead, we derive a method to model the evolutionary dynamics of controls, which yields an alternative method to optimal control that can be used to obtain optimal controls in a broad class of evolutionary models with dynamic constraints. Our use of optimal control concepts is thus useful to see how our results relate to optimal control, which has wide-ranging theory and applications.

We obtain three sets of main results. First, we derive several closed-form formulas for the total selection gradient of genotypic traits $\bar{\mathbf{y}}$ (i.e., of control variables) that affect the development of the phenotype $\bar{\mathbf{x}}$ (i.e., of state variables), formulas that can be easily computed with elementary operations. Second, we derive equations in gradient form describing the evolutionary dynamics of developed traits $\bar{\mathbf{x}}$ and of the niche-constructed environment. These equations depend on a mechanistic counterpart of the developmental matrix $\boldsymbol{\alpha}$, for which we obtain developmentally explicit and evolutionarily dynamic formulas for a broad class of models. Such formulas are in terms of closed-form formulas that we derive for the sensitivity of a system of recurrence equations, which are of use beyond evolutionary or biological applications. The obtained equation describing the evolutionary dynamics of the developed traits $\bar{\mathbf{x}}$ is generally dynamically insufficient because it depends on the genotypic traits $\bar{\mathbf{y}}$, whose evolutionary dynamics are not described by the equation, a problem similarly present in the classic Lande equation. Third, we obtain synthetic equations in gradient form simultaneously describing the evolutionary dynamics of genotypic, developed, and environmental traits. These equations are in gradient form and are dynamically sufficient in that they include as many evolutionarily dynamic equations as evolutionarily dynamic variables, which enables one to describe the long-term evolution of developed multivariate phenotypes as the climbing of a fitness landscape. Such equations describe the evolutionary dynamics of the constraining matrix analogous to $\mathbf{G}$ as an emergent property, where genotypic traits $\bar{\mathbf{y}}$ play an analogous role to that of allele frequency under quantitative genetics assumptions while linkage disequilibrium is not an issue as we assume clonal reproduction. The obtained dynamically sufficient gradient system depends on a constraining matrix that is always singular, which is mathematically trivial, but biologically crucial. The reason is that this singularity entails that genetic variation is necessarily absent in certain directions such that adaptive evolution at best converges to outcomes defined by both selection and development rather than by selection alone. Consequently, we find that development plays a major evolutionary role.

2. Problem statement

We begin by describing the mathematical problem we address. We consider a finite age-structured population with deterministic density-dependent population dynamics with age measured in discrete time. Each individual is described by three types of traits that we call genotypic, phenotypic (or developed), and environmental, all of which can vary with age and can evolve. We let all traits take continuous values, which allows us to take derivatives. Genotypic

traits are defined by being directly specified by the genotype. Phenotypic traits are defined by being constructed over life subject to a developmental constraint: for instance, a phenotypic trait may be body size subject to the influence of genes, developmental history, environment, social interactions, and developmental processes constructing the body. Environmental traits are defined as describing the local environment of the individual subject to an environmental constraint: for instance, an environmental trait may be ambient temperature, which the individual may adjust behaviorally such as by roosting in the shade. We assume that reproduction transmits genotypic traits clonally, but developed and environmental traits need not be transmitted clonally due to social interactions. Given clonal reproduction of genotypic traits, we do not need to specify the genetic architecture (e.g., ploidy, number of loci, or linkage) and it may depend on the particular model. We assume that the genotypic traits are *developmentally independent*, whereby genotypic traits are entirely specified by the individual's genotype and do not depend on other traits expressed over development: in particular, this means that a given genotypic trait at a given age can only be modified by mutation, but does not depend on other genotypic traits, the phenotype, or the environment. Developmental independence corresponds to the notion of "open-loop" control of optimal control theory (Sydsæter et al., 2008). Genotypic traits may still be *mutationally* correlated, whereby genotypic traits may tend to mutate together or separately. We assume that environmental traits are mutually independent, which facilitates derivations. We obtain dynamically sufficient equations in gradient form for the evolution of the phenotype by aggregating the various types of traits. We give names to such aggregates for ease of reference. We call the aggregate of the genotype and phenotype the geno-phenotype. We call the aggregate of the genotype, phenotype, and environment the geno-envo-phenotype.

The above terminology departs from standard terminology in adaptive dynamics as follows. In adaptive dynamics, our genotypic traits are referred to as the phenotype and our phenotypic traits as state variables. We depart from this terminology to follow the biologically common notion that the phenotype is constructed over development. In turn, adaptive dynamics terminology defines the environment as any quantity outside the individual, and thus refers to the global environment. In contrast, by environment we refer to the local environment of the individual. This allows us to model niche construction as the local environment of a mutant individual may differ from that of a resident.

We use the following notation (Table 1). Each individual can live from age 1 to age $N_\mathrm{a} \in \{2, 3, \ldots\}$. Each individual has a number N_g of genotypic traits at each age. A mutant's genotypic trait $i \in \{1, \ldots, N_\mathrm{g}\}$ at age $a \in \{1, \ldots, N_\mathrm{a}\}$ is $y_{ia} \in \mathbb{R}$. For instance, y_{ia} may be the value of a life-history trait i at age a assumed to be directly under genetic control (i.e., a control variable in life-history models; Gadgil and Bossert, 1970; Taylor et al., 1974; León, 1976; Schaffer, 1983). Given our assumption of developmental independence of genotypic traits, the genotypic trait value y_{ia} for all $i \in \{1, \ldots, N_\mathrm{g}\}$ and all $a \in \{1, \ldots, N_\mathrm{a}\}$ of a given individual is exclusively controlled by her genotype but mutations can tend to change the value of y_{ia} and y_{kj} simultaneously for $k \neq i$ and $j \neq a$. Additionally, each individual has a number N_p of developed traits, that is, of phenotypes at each age. A mutant's phenotype $i \in \{1, \ldots, N_\mathrm{p}\}$ at age $a \in \{1, \ldots, N_\mathrm{a}\}$ is $x_{ia} \in \mathbb{R}$. Moreover, each individual has a number N_e of environmental traits that describe her local environment at each age. A mutant's environmental trait $i \in \{1, \ldots, N_\mathrm{e}\}$ at age $a \in \{1, \ldots, N_\mathrm{a}\}$ is $\epsilon_{ia} \in \mathbb{R}$. We do not consider the developmental or evolutionary change of the number of traits (i.e., of N_g, N_p, or N_e), but our framework allows for the

modeling of the developmental or evolutionary origin of novel traits (e.g., the origin of a sixth digit where there was five previously in development or evolution; Chan et al., 1995; Litingtung et al., 2002; Müller, 2010) by implementing a suitable codification (e.g., letting x_{ia} mean sixth-digit length, being zero in a previous age or evolutionary time).

We use the following notation for collections of these quantities. A mutant's i-th genotypic trait across all ages is denoted by the column vector $\mathbf{y}_i = (y_{i1}; \ldots; y_{iN_\mathrm{a}}) \in \mathbb{R}^{N_\mathrm{a}\times 1}$, where the semicolon indicates a line break, that is, $\mathbf{y}_i = (y_{i1}, \ldots, y_{iN_\mathrm{a}})^\intercal$; we write $\mathbb{R}^{N_\mathrm{a}\times 1}$ rather than simply $\mathbb{R}^{N_\mathrm{a}}$ as it will be important to clearly distinguish column from row vectors. A mutant's i-th phenotype across all ages is denoted by the column vector $\mathbf{x}_i = (x_{i1}; \ldots; x_{iN_\mathrm{a}}) \in \mathbb{R}^{N_\mathrm{a}\times 1}$. A mutant's i-th environmental trait across all ages is denoted by the column vector $\boldsymbol{\epsilon}_i = (\epsilon_{i1}; \ldots; \epsilon_{iN_\mathrm{a}}) \in \mathbb{R}^{N_\mathrm{a}\times 1}$. A mutant's genotype across all genotypic traits and all ages is denoted by the block column vector $\mathbf{y} = (\mathbf{y}_1; \ldots; \mathbf{y}_{N_\mathrm{g}}) \in \mathbb{R}^{N_\mathrm{a}N_\mathrm{g}\times 1}$. A mutant's phenotype across all developed traits and all ages is denoted by the block column vector $\mathbf{x} = (\mathbf{x}_1; \ldots; \mathbf{x}_{N_\mathrm{p}}) \in \mathbb{R}^{N_\mathrm{a}N_\mathrm{p}\times 1}$. A mutant's environment across all environmental traits and all ages is denoted by the block column vector $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1; \ldots; \boldsymbol{\epsilon}_{N_\mathrm{p}}) \in \mathbb{R}^{N_\mathrm{a}N_\mathrm{e}\times 1}$. To simultaneously refer to the genotype and phenotype, we denote the geno-phenotype of the mutant individual at age a as $\mathbf{z}_a = (\mathbf{x}_a; \mathbf{y}_a) \in \mathbb{R}^{(N_\mathrm{p}+N_\mathrm{g})\times 1}$, and the geno-phenotype of a mutant across all ages as $\mathbf{z} = (\mathbf{x}; \mathbf{y}) \in \mathbb{R}^{N_\mathrm{a}(N_\mathrm{p}+N_\mathrm{g})\times 1}$. Moreover, to simultaneously refer to the genotype, phenotype, and environment, we denote the geno-envo-phenotype of a mutant individual at age a as $\mathbf{m}_a = (\mathbf{z}_a; \boldsymbol{\epsilon}_a) \in \mathbb{R}^{(N_\mathrm{p}+N_\mathrm{g}+N_\mathrm{e})\times 1}$, and the geno-envo-phenotype of the mutant across all ages as $\mathbf{m} = (\mathbf{z}; \boldsymbol{\epsilon}) \in \mathbb{R}^{N_\mathrm{a}(N_\mathrm{p}+N_\mathrm{g}+N_\mathrm{e})\times 1}$. We denote resident values analogously with an overbar (e.g., $\bar{\mathbf{z}}$ is the resident geno-phenotype).

The developmental process that constructs the phenotype is as follows, with causal dependencies described in Fig. 1. We assume that an individual's multivariate phenotype at a given age is a function of the genotypic, phenotypic, and environmental traits that the individual had at the immediately previous age as well as of the social interactions experienced at that age. Thus, we assume that a mutant's multivariate phenotype at age $a + 1$ is given by the developmental constraint

$$\mathbf{x}_{a+1} = \mathbf{g}_a(\mathbf{m}_a, \bar{\mathbf{z}}) \tag{1}$$

for all $a \in \{1, \ldots, N_\mathrm{a} - 1\}$ with initial condition $\mathbf{x}_1 = \bar{\mathbf{x}}_1$. The function

$$\mathbf{g}_a(\mathbf{m}_a, \bar{\mathbf{z}}) = (g_{1a}(\mathbf{m}_a, \bar{\mathbf{z}}); \ldots; g_{N_\mathrm{p}a}(\mathbf{m}_a, \bar{\mathbf{z}}))$$

is the developmental map at age a, which we assume is a continuously differentiable function of the individual's geno-envo-phenotype at that age and of the geno-phenotype of the individual's social partners who can be of any age; thus, an individual's development directly depends on the individual's local environment but not directly on the local environment of social partners. For simplicity, we assume that the phenotype $\mathbf{x}_1 = \bar{\mathbf{x}}_1$ at the initial age is constant and does not evolve. This assumption corresponds to the common assumption in life-history models that state variables at the initial age are given (Gadgil and Bossert, 1970; Taylor et al., 1974; León, 1976; Schaffer, 1983; Sydsæter et al., 2008). The term developmental function can be traced back to Gimelfarb (1982) through Wagner (1984).

The developmental constraint (1) allows for a wide range of models. Eq. (1) is a mathematical, deterministic

Table 1: **Notation summary**

Symbol	Meaning
$\mathbf{x}$	Phenotype (developed traits)
$\mathbf{y}$	Genotype (genotypic traits)
$\mathbf{z}$	Geno-phenotype (genotype and phenotype)
$\boldsymbol{\epsilon}$	Environment
$\mathbf{m}$	Geno-envo-phenotype (genotype, phenotype, and environment)
N_{a}	Number of ages
N_{p}	Number of developed traits
N_{g}	Number of genotypic traits
N_{e}	Number of environmental traits
$\mathbf{g}$	Developmental map
$\mathbf{h}$	Environmental map
$\mathbf{n}$	Population density
f	Fertility
p	Survival probability
ℓ	Survivorship
w	Fitness
λ	Invasion fitness
$\mathbf{u}$	Stable age distribution
$\mathbf{v}$	Reproductive value
ϕ	Force of selection on fertility
π	Force of selection on survival
t	Ecological time
τ	Evolutionary time
θ	Socio-devo time
T	Generation time
$\check{\mathbf{x}}$	Resident phenotype in the context of mutant
$\hat{\mathbf{z}}$	Unperturbed geno-phenotype
$\boldsymbol{\zeta}, \boldsymbol{\xi}$	Arbitrary vectors
$\mathbf{b}_{\zeta}$	Mechanistic breeding value of $\boldsymbol{\zeta}$
$\mathbf{b}_{\zeta}^{\mathrm{s}}$	Stabilized mechanistic breeding value of $\boldsymbol{\zeta}$
$\mathbf{H}_{\zeta}$	Mechanistic additive genetic covariance matrix of $\boldsymbol{\zeta}$
$\mathbf{L}_{\zeta}$	Mechanistic additive socio-genetic cross-covariance matrix of $\boldsymbol{\zeta}$
$\frac{\partial \boldsymbol{\zeta}^{\mathsf{T}}}{\partial \boldsymbol{\xi}}$	Direct effects of $\boldsymbol{\xi}$ on $\boldsymbol{\zeta}$
$\frac{\delta \boldsymbol{\zeta}^{\mathsf{T}}}{\delta \boldsymbol{\xi}}$	Total immediate effects of $\boldsymbol{\xi}$ on $\boldsymbol{\zeta}$
$\frac{\mathrm{d}\boldsymbol{\zeta}^{\mathsf{T}}}{\mathrm{d}\boldsymbol{\xi}}$	Total effects of $\boldsymbol{\xi}$ on $\boldsymbol{\zeta}$
$\frac{\mathrm{s}\boldsymbol{\zeta}^{\mathsf{T}}}{\mathrm{s}\boldsymbol{\xi}}$	Stabilized effects of $\boldsymbol{\xi}$ on $\boldsymbol{\zeta}$

description of Waddington's (1957) "epigenetic landscape". Eq. (1) is a constraint in that the phenotype $\mathbf{x}_{a+1}$ cannot take any value but only those that satisfy the equality (e.g., an individual's body size today cannot take any value but depends on her body size, gene expression, and environment since yesterday). The developmental map in Eq. (1) is an

extension of the notions of genotype-phenotype map (often a function from genotype to phenotype, without explicit developmental dynamics) and reaction norm (often a function from environment to phenotype, also without explicit developmental dynamics), as well as of early mathematical descriptions of development in an evolutionary context (Alberch et al., 1979). The developmental constraint (1) can describe gene regulatory networks (Alon, 2020), for instance, letting x_{ia} be the expression level of gene i; learning in deep neural networks (Russell and Norvig, 2021), for instance, letting $\boldsymbol{\epsilon}_a$ describe the network's input at age a and $\mathbf{x}_a$ describe the network's weights at that age; and reaction-diffusion models of morphology (Murray, 2003), for instance, letting x_{ia} be a morphogen's level at the i-th spatial location at age a (SI section S1.2). The developmental map in Eq. (1) may be non-linear and can change over development (e.g., from $g_{ia} = \sin x_{ia}$ to $g_{ij} = x_{ij}^{\beta}$ for $a < j$ and some parameter β, for instance, due to metamorphosis) and over evolution (e.g., from a sine to a power function if $g_{ia} = [y_{ja} \sin x_{ia} + (1 - y_{ja}) x_{ia}^{\beta}]$ as genotypic trait y_{ja} evolves from 0 to 1). The dependence of the mutant phenotype on the phenotype of social partners in (1) allows one to implement Jablonka and Lamb's (2014) notion that extra-genetic inheritance transmits the phenotype rather than the genotype (see their p. 108), such that the mutant phenotype can be a possibly altered copy of social partners' phenotype. Simpler forms of the developmental constraint (1) are standard in life-history models (Gadgil and Bossert, 1970; Taylor et al., 1974; León, 1976; Schaffer, 1983; Sydsæter et al., 2008) and physiologically structured models of population dynamics (de Roos, 1997, Eq. 7).

We describe the local environment as follows. We assume that an individual's local environment at a given age is a function of the genotypic traits, phenotype, and social interactions of the individual at that age, and of processes that are not caused by the population considered. Thus, we assume that a mutant's environment at age a is given by the environmental constraint

$$\boldsymbol{\epsilon}_a = \mathbf{h}_a(\mathbf{z}_a, \bar{\mathbf{z}}, \tau) \tag{2}$$

for all $a \in \{1, \ldots, N_{\mathrm{a}}\}$. The function

$$\mathbf{h}_a(\mathbf{z}_a, \bar{\mathbf{z}}, \tau) = (h_{1a}(\mathbf{z}_a, \bar{\mathbf{z}}, \tau); \ldots; h_{N_{\mathrm{e}}a}(\mathbf{z}_a, \bar{\mathbf{z}}, \tau))$$

is the environmental map at age a, which can change over development and evolution. We assume that the environmental map is a continuously differentiable function of the individual's geno-phenotype at that age (e.g., the individual's behavior at a given age may expose it to a particular environment at that age), the geno-phenotype of the individual's social partners who can be of any age (e.g., through social niche construction), and evolutionary time τ due to slow exogenous environmental change (so the exogenous process changing the environment in Fig. 1 acts as forcing, as τ appears explicitly in Eq. 2). We assume slow exogenous environmental change to enable the resident population to reach carrying capacity to be able to use relatively simple techniques of evolutionary invasion analysis to derive selection gradients. The environmental constraint (2) may also be non-linear and can change over development (i.e., over a) and over evolution (as the genotype or phenotype evolves or exogenously as evolutionary time advances).

The environmental constraint (2) is a minimalist description of the environment of a specific kind (akin to "feed-

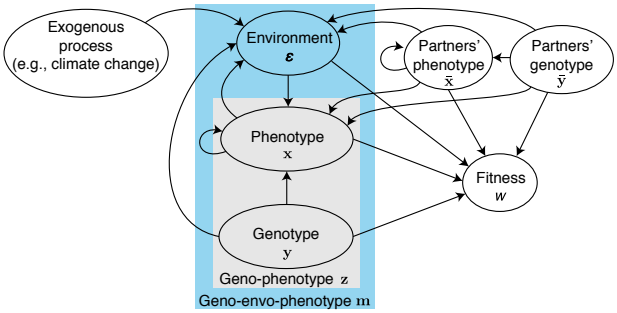

Figure 1: Causal diagram among the framework's components. Variables have age-specific values that are not shown for clarity. The phenotype $\mathbf{x}$ is constructed by a developmental process. Each arrow indicates the direct effect of a variable on another one. A mutant's genotypic traits may directly affect the phenotype (with the slope quantifying developmental bias from genotype), environment (niche construction by genotype), and fitness (direct selection on genotype). A mutant's phenotype at a given age may directly affect her phenotype at an immediately subsequent age (quantifying developmental bias from the phenotype), thus the direct feedback loop from phenotype to itself. A mutant's phenotype may also directly affect her environment (niche construction by the phenotype) and fitness (direct selection on the phenotype). A mutant's environment may directly affect the phenotype (plasticity) and fitness (environmental sensitivity of selection). The social partners' genotype may directly affect their own phenotype (quantifying developmental bias from genotype), the mutant's phenotype (indirect genetic effects from genotypes), and the mutant's fitness (social selection on genotype). The social partners' phenotype at a given age may directly affect their own phenotype at an immediately subsequent age (quantifying developmental bias from phenotypes), thus the direct feedback loop. The social partners' phenotype at a given age may also directly affect the mutant's phenotype (quantifying indirect genetic effects from the phenotype), the mutant's environment (social niche construction), and the mutant's fitness (social selection on the phenotype). The environment may also be directly influenced by exogenous processes. We assume that the genotype is developmentally independent (i.e., controls $\mathbf{y}$ are open-loop), which means that there is no arrow towards the genotype.

back functions" used in physiologically structured models to describe the influence of individuals on the environment; de Roos, 1997). We use the minimalist environmental constraint (2) as a first approximation to shorten derivations; our derivations illustrate how one could obtain equations with more complex developmental and environmental constraints. With the minimalist environmental constraint (2), the environmental traits are mutually independent in that changing one environmental trait at one age does not *directly* change any other environmental trait at any age (i.e., $\partial\epsilon_{kj}/\partial\epsilon_{ia} = 0$ if $i \neq k$ or $a \neq j$). We say that development is social if $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}}|_{\mathbf{y}=\bar{\mathbf{y}}} \neq \mathbf{0}$.

Our aim is to obtain closed-form equations describing the evolutionary dynamics of the resident phenotype $\bar{\mathbf{x}}$ subject to the developmental constraint (1) and the environmental constraint (2). The evolutionary dynamics of the phenotype $\bar{\mathbf{x}}$ emerge as an outgrowth of the evolutionary dynamics of the genotype $\bar{\mathbf{y}}$ and environment $\bar{\boldsymbol{\epsilon}}$. In the SI section S2.5, we show that under our assumptions the evolutionary dynamics of the resident genotype $\bar{\mathbf{y}}$ are given by the canonical equation of adaptive dynamics (Dieckmann and Law, 1996):

$$\frac{\Delta\bar{\mathbf{y}}}{\Delta\tau} \approx \iota\mathbf{H}_{\mathbf{y}} \left.\frac{\mathrm{d}\lambda}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}, \tag{3}$$

where $\lambda = \lambda(\mathbf{m}, \bar{\mathbf{m}})$ is invasion fitness, $\iota = \frac{1}{2}\mu(\bar{\mathbf{m}})\bar{n}^*(\bar{\mathbf{m}})$ is a non-negative scalar measuring mutational input in terms of the mutation rate $\mu(\bar{\mathbf{m}})$ and the carrying capacity $\bar{n}^*(\bar{\mathbf{m}})$, and $\mathbf{H}_{\mathbf{y}} = \mathrm{cov}[\mathbf{y}, \mathbf{y}]$ is the mutational covariance matrix (of

genotypic traits). The selection gradient in Eq. (3) involves total derivatives so we call it the *total* selection gradient of the genotype, which measures the effects of genotypic traits $\mathbf{y}$ on invasion fitness λ across all the paths in Fig. 1 (λ has the same derivatives with respect to mutant trait values as fitness w, defined below). Total selection gradients differ from Lande's selection gradient in that the latter is defined in terms of partial derivatives and so measures only the direct effects of traits on fitness (e.g., $\partial w/\partial \mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}}$ measures the effect of $\mathbf{y}$ on w only across the path directly connecting the two in Fig. 1). We will be concerned with describing the evolutionary dynamics to first-order of approximation, so we will treat the approximation in Eq. (3) as an equality although we keep the approximation symbols throughout to distinguish what is and what is not an approximation.

The arrangement above describes the evolutionary developmental (evo-devo) dynamics: the evolutionary dynamics of the resident genotype are given by the canonical equation (3), while the concomitant developmental dynamics of the phenotype are given by the developmental (1) and environmental (2) constraints evaluated at resident trait values. To complete the description of the evo-devo dynamics, we obtain closed-form expressions for the total selection gradient of the genotype. Moreover, to determine whether the evolution of the resident developed phenotype $\bar{\mathbf{x}}$ can be understood as the climbing of a fitness landscape, we derive equations in gradient form describing the evolutionary dynamics of the resident phenotype $\bar{\mathbf{x}}$, environment $\bar{\boldsymbol{\epsilon}}$, geno-phenotype $\bar{\mathbf{z}}$, and geno-envo-phenotype $\bar{\mathbf{m}}$. Before doing so, we first give an overview of the model underpinning the arrangement above, which describes a complication introduced by social development, how we handle it, and a fitness function that has the same gradient as invasion fitness in age-structured populations. We then use these descriptions to write our results.

3. Model overview

Here we give an overview of the model. We describe it formally in the SI section S2.

3.1. Set up

We base our framework on standard assumptions of adaptive dynamics, particularly following Dieckmann and Law (1996), but social development introduces a non-standard complication. We separate time scales, so developmental and population dynamics occur over a short discrete ecological time scale t and evolutionary dynamics occur over a long discrete evolutionary time scale τ. Although the population is finite, in a departure from Dieckmann and Law (1996), we let the population dynamics be deterministic rather than stochastic for simplicity, so there is no genetic drift. Thus, the only source of stochasticity in our framework is mutation. We assume that mutation is rare, weak, and unbiased. Weak mutation means that the variance of mutant genotypic traits around resident genotypic traits is marginally small (i.e., $0 < \mathrm{E}[\|\mathbf{y}-\bar{\mathbf{y}}\|^2] = \mathrm{tr}(\mathrm{cov}[\mathbf{y},\mathbf{y}]) = \sum_{i=1}^{N_\mathrm{g}}\sum_{a=1}^{N_\mathrm{a}} \mathrm{E}[(y_{ia}-\bar{y}_{ia})^2] \ll 1$). Weak mutation (Gillespie, 1983; Walsh and Lynch, 2018, p. 1003) is also called δ-weak selection (Wild and Traulsen, 2007). Unbiased mutation means that mutant genotypic traits are symmetrically distributed around the resident genotypic traits (i.e., the mutational distribution $M(\mathbf{y}-\bar{\mathbf{y}})$ is even). Unbiased mutation in genotypic traits allows for bias in the distribution

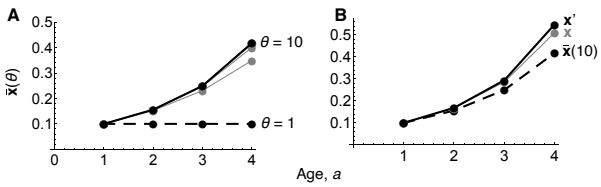

Figure 2: A difficulty introduced by social development. (A) Illustration of socio-devo dynamics converging to a quasi socio-devo stable equilibrium (solid, black line). The dashed line is a socio-devo initial resident phenotype $\bar{x}_a(\theta) = 0.1$ for all $a \in \{1, \ldots, 4\}$, $\bar{x}_1 = 0.1$, and socio-devo time $\theta = 1$. The gray line immediately above is a phenotype developed in the context of such resident, where $\bar{x}_{a+1}(2) = g_a(\bar{\mathbf{z}}_a, \bar{x}_{a+1}) = \bar{x}_a(2) + \bar{y}_a\{\bar{x}_a(2) + q[\bar{x}_{a+1}(1)]^2\}$, with $\bar{y}_a = 0.5$ for all $a \in \{1, \ldots, 4\}$ and $q = 0.5$. Setting this phenotype $\bar{\mathbf{x}}(2)$ as resident and iterating up to $\theta = 10$ yields the remaining gray lines, with iteration 10 given by the black line, where $\bar{x}_{a+1}(10) = \bar{x}_a(10) + \bar{y}_a\{\bar{x}_a(10) + q[\bar{x}_{a+1}(9)]^2\}$ and $\bar{\mathbf{x}}(10) \approx \bar{\mathbf{x}}(9)$ is approximately a socio-devo stable equilibrium, which breeds true. (B) Introducing in the context of such resident $\bar{\mathbf{x}}(10)$ (dashed line) a mutant genotype $\mathbf{y}$ yields the mutant phenotype $\mathbf{x}$ (gray line), where $x_{a+1} = x_a + y_a\{x_a + q[\bar{x}_{a+1}(10)]^2\}$ and $y_a = 0.6$ for all $a \in \{1, \ldots, 4\}$. Such mutant does not breed true: a mutant $\mathbf{x}'$ (solid black line) with the same genotype developed in the context of mutant $\mathbf{x}$ has a different phenotype, where $x'_{a+1} = x'_a + y_a\{x'_a + q[x_{a+1}]^2\}$. One can use socio-devo dynamics (A) to find for such mutant genotype $\mathbf{y}$ a phenotype that breeds true under social development.

of mutant phenotypes (i.e., the distribution of $\mathbf{x} - \bar{\mathbf{x}}$ is not necessarily even). Thus, we do not make the isotropy assumption of Fisher's (1930) geometric model (Orr, 2005), although isotropy may arise for mechanistic breeding values (defined below) with large $N_\mathrm{a}N_\mathrm{g}$ and additional assumptions (e.g., high pleiotropy and high developmental integration) from the central limit theorem (Martin, 2014). We assume that a monomorphic resident population having geno-envo-phenotype $\bar{\mathbf{m}}$ undergoes density-dependent population dynamics that bring it to carrying capacity. At this carrying capacity, rare mutant individuals arise which have a marginally different genotype $\mathbf{y}$ and that develop their phenotype in the context of the resident. If the mutant genotype increases in frequency, it increasingly faces mutant rather than resident individuals. Thus, with social development, the mutant phenotype may change as the mutant genotype spreads, which complicates invasion analysis.

3.2. A complication introduced by social development

With social development, the phenotype an individual develops depends on the traits of her social partners. This introduces a complication to standard evolutionary invasion analysis, for two reasons. First, the phenotype of a mutant genotype may change as the mutant genotype spreads and is more exposed to the mutant's traits via social interactions, making the mutant phenotype frequency dependent. Thus, the phenotype developed by a rare mutant genotype in the context of a resident phenotype may be different from the phenotype developed by the same mutant genotype in the context of itself once the mutant genotype has approached fixation. Second, because of social development, a recently fixed mutant may not breed true, that is, her descendants may have a different phenotype from her own despite clonal reproduction of the genotype and despite the mutant genotype being fixed (Fig. 2; see also Kobayashi et al. 2015, Eq. 14 in their Appendix). Yet, to apply standard invasion analysis techniques, the phenotype of the fixed genotype must breed true, so that the phenotype of a mutant genotype developed in the context of individuals with the mutant genotype have the same phenotype.

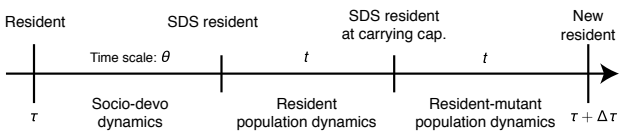

Figure 3: Phases of an evolutionary time step. Evolutionary time is τ. SDS means socio-devo stable. The socio-devo dynamics phase is added to the standard separation of time scales in adaptive dynamics, which only consider the other two phases. The socio-devo dynamics phase is only needed if development is social (i.e., if the developmental map $\mathbf{g}_a$ depends directly or indirectly on social partners' geno-phenotype for some age a).

Thus, to carry out invasion analysis, we proceed as follows. Ideally, one should follow explicitly the change in mutant phenotype as the mutant genotype increases in frequency and achieves fixation, and up to a point where the fixed mutant phenotype breeds true. Yet, to simplify the analysis, we separate the dynamics of phenotype convergence and the population dynamics. We thus introduce an additional phase to the standard separation of time scales in adaptive dynamics so that phenotypic convergence occurs first and then resident population dynamics follow. Such additional phase does not describe a biological process but is a mathematical technique to facilitate mathematical treatment (akin to using best-response dynamics to find Nash equilibria). However, this phase might still be biologically justified under somewhat broad conditions. In particular, Aoki et al. (2012, their Appendix A) show that such additional phase is justified in their model of social learning evolution if mutants are rare and social learning dynamics happen faster than allele frequency change; they also show that this additional phase is justified for their particular model if selection is δ-weak. As a first approximation, here we do not formally justify the separation of phenotype convergence and resident population dynamics for our model and simply assume it for simplicity.

3.3. Phases of an evolutionary time step

To handle the above complication introduced by social development, we partition a unit of evolutionary time in three phases: socio-developmental (socio-devo) dynamics, resident population dynamics, and resident-mutant population dynamics (Fig. 3).

At the start of the socio-devo dynamics phase of a given evolutionary time τ, the population consists of individuals all having the same resident genotype, phenotype, and environment. A new individual arises which has identical genotype as the resident, but develops a phenotype that may be different from that of the original resident due to social development. This developed phenotype, its genotype, and its environment are set as the new resident. This process is repeated until convergence of the geno-envo-phenotype to what we term a "socio-devo stable" (SDS) resident equilibrium or until divergence. These socio-devo dynamics are formally described by Eq. S2.1.1 and illustrated in Fig. 2A. If development is not social, the resident is trivially SDS so the socio-devo dynamics phase is unnecessary. If an SDS resident is achieved, the population moves to the next phase; if an SDS resident is not achieved, the analysis stops. We thus study only the evolutionary dynamics of SDS resident geno-envo-phenotypes. More specifically, we say a geno-envo-phenotype $\bar{\mathbf{m}} = (\bar{\mathbf{x}}; \bar{\mathbf{y}}; \bar{\boldsymbol{\epsilon}})$ is a socio-devo equilibrium if and only if $\bar{\mathbf{x}}$ is produced by development

when the individual has such genotype $\bar{\mathbf{y}}$ and everyone else in the population has that same genotype, phenotype, and environment (Eq. S2). A socio-devo equilibrium $\bar{\mathbf{m}} = (\bar{\mathbf{x}}; \bar{\mathbf{y}}; \bar{\boldsymbol{\epsilon}})$ is locally stable (i.e., SDS) if and only if a marginally small deviation in the initial phenotype $\bar{\mathbf{x}}(1)$ from the socio-devo equilibrium keeping the same genotype leads the socio-devo dynamics (Eq. S1) to the same equilibrium. A socio-devo equilibrium $\bar{\mathbf{m}}$ is locally stable if all the eigenvalues of the matrix

$$\left.\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\bar{\mathbf{x}}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}$$

have absolute value (or modulus) strictly less than one. For instance, this is always the case if social interactions are only among peers (i.e., individuals of the same age) so the mutant phenotype at a given age depends only on the phenotype of immediately younger social partners (in which case the above matrix is block lower triangular so all its eigenvalues are zero; Eq. S5.6.9). We assume that there is a unique SDS geno-envo-phenotype for a given developmental map at every evolutionary time τ.

If an SDS resident is achieved in the socio-devo dynamics phase, the population moves to the resident population dynamics phase. Because the resident is SDS, an individual with resident genotype developing in the context of the resident geno-phenotype is guaranteed to develop the resident phenotype. Thus, we may proceed with the standard invasion analysis. Hence, in this phase of SDS resident population dynamics, the SDS resident undergoes density dependent population dynamics, which we assume asymptotically converges to a carrying capacity.

Once the SDS resident has achieved carrying capacity, the population moves to the resident-mutant population dynamics phase. At the start of this phase, a random mutant genotype $\mathbf{y}$ marginally different from the resident genotype $\bar{\mathbf{y}}$ arises in a vanishingly small number of mutants. We assume that the mutant becomes either lost or fixed in the population (Geritz et al., 2002; Geritz, 2005; Priklopil and Lehmann, 2020), establishing a new resident geno-envo-phenotype.

Repeating this evolutionary time step generates long term evolutionary dynamics of an SDS geno-envo-phenotype.

3.4. Fitness in age structured populations

To compute the total selection gradient of the genotype, we now write a fitness function that is more tractable than invasion fitness and that has the same first-order derivatives with respect to mutant trait values as invasion fitness for age-structured populations. To do this, we first write a mutant's survival probability and fertility at each age. At the resident population dynamics equilibrium, a rare mutant's fertility at age a is

$$f_a = f_a(\mathbf{m}_a, \bar{\mathbf{m}}) \tag{4a}$$

and the mutant's survival probability from age a to $a + 1$ is

$$p_a = p_a(\mathbf{m}_a, \bar{\mathbf{m}}). \tag{4b}$$

The first argument $\mathbf{m}_a$ in Eqs. (4) is the direct dependence of the mutant's fertility and survival at a given age on her own geno-envo-phenotype at that age. The second argument $\bar{\mathbf{m}}$ in Eqs. (4) is the direct dependence on social

partners' geno-envo-phenotype at any age (thus, fertility and survival may directly depend on the environment of social partners, specifically, as it may affect the carrying capacity, and fertility and survival are density dependent).

In the SI section S2.3, we show that the gradients of invasion fitness λ with respect to mutant trait values are equal to (not an approximation of) the corresponding gradients of the relative fitness w of a mutant individual per unit of generation time (Eq. S2.3.9), defined as

$$w = \sum_{j=1}^{N_\mathrm{a}} w_j, \tag{5a}$$

where a mutant's relative fitness at age j is

$$w_j = \frac{1}{T}\left(\phi_j f_j + \pi_j p_j\right), \tag{5b}$$

and generation time is

$$T = \sum_{j=1}^{N_\mathrm{a}} j\ell_j^\circ f_j^\circ \tag{6}$$

(Charlesworth 1994, Eq. 1.47c; Bulmer 1994, Eq. 25, Ch. 25; Bienvenu and Legendre 2015, Eqs. 5 and 12). The superscript $\circ$ denotes evaluation at $\mathbf{y} = \bar{\mathbf{y}}$ (so at $\mathbf{m} = \bar{\mathbf{m}}$ as the resident is a socio-devo equilibrium). The quantity $\ell_j = \prod_{k=1}^{j-1} p_k$ is the survivorship of mutants from age 1 to age j, and ℓ_j° is that of neutral mutants. Thus, the generation time appearing in w is for a neutral mutant, or resident, rather than the mutant. This can be intuitively understood as the mutant having to invade over a time scale determined by the resident rather than the mutant. We denote the force of selection on fertility at age j (Hamilton 1966 and Caswell 1978, his Eqs. 11 and 12) as

$$\phi_j(\bar{\mathbf{m}}) = \ell_j^\circ \tag{7a}$$

and the force of selection on survival at age j (Baudisch 2005, her Eq. 5a) as

$$\pi_j(\bar{\mathbf{m}}) = \frac{1}{p_j^\circ}\sum_{k=j+1}^{N_\mathrm{a}} \ell_k^\circ f_k^\circ, \tag{7b}$$

which are independent of mutant trait values because they are evaluated at the resident trait values. It is easily checked that ϕ_j and π_j decrease with j (respectively, if $p_j^\circ < 1$ and $f_{j+1}^\circ > 0$ provided that $p_j^\circ \approx p_{j+1}^\circ$).

In the SI section S2.4.2, we show that the gradients of invasion fitness λ with respect to mutant trait values are also equal to $1/T$ times the corresponding gradients of a rare mutant's expected lifetime reproductive success R_0 (Eq. S2.4.11) (Bulmer, 1994; Caswell, 2009). This occurs because of our assumption that mutants arise when residents are at carrying capacity (Mylius and Diekmann, 1995). For our life cycle, a mutant's expected lifetime reproductive success is

$$R_0 = \sum_{j=1}^{N_\mathrm{a}} \ell_j f_j, \tag{8}$$

(Caswell, 2001). From the equality of the gradients of invasion fitness and fitness, it follows that invasion fitness λ for age-structured populations is approximately equal to the mutant's relative fitness w to first-order of approximation around resident genotypic traits, that is, $\lambda \approx w$ (Eq. S2.3.11). Similarly, from the equality of the gradients of invasion fitness and a mutant's lifetime reproductive success per generation time, it follows that invasion fitness λ for age-structured populations is given by $\lambda \approx 1 + (R_0 - 1)/T$ to first-order of approximation around resident genotypic traits (Eq. S2.3.13). Taking derivatives of w with respect to mutant trait values is generally simpler than for λ or R_0, so we present most results below in terms of w.

4. Summary of main results

We now give an overview of the three sets of main results. We provide a comprehensive presentation of ancillary results and further analysis in section 5 and SI section S3. The derivations of all these results are in SI section S5 and involve repeated use of the chain rule due to the recurrence and feedbacks involved in the developmental constraint (1).

4.1. Total selection gradient of the genotype

The total selection gradient of the genotype is

$$\begin{aligned}\left.\frac{\mathrm{d}\lambda}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &= \left.\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\ &= \left.\left(\frac{\mathrm{d}\mathbf{x}^{\mathsf{T}}}{\mathrm{d}\mathbf{y}}\frac{\partial w}{\partial \mathbf{x}} + \frac{\partial w}{\partial \mathbf{y}} + \frac{\mathrm{d}\boldsymbol{\epsilon}^{\mathsf{T}}}{\mathrm{d}\mathbf{y}}\frac{\partial w}{\partial \boldsymbol{\epsilon}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\ &\in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{g}}\times 1}.\end{aligned} \tag{9}$$

The chain rule in Eq. (9) is reversed to its standard presentation because the gradient in the canonical equation (3) is a column vector. Yet, the order in Eq. (9) gives a natural biological interpretation: a perturbation in the genotype first affects the phenotype and then the affected phenotype affects fitness, and similarly for the effect of genotypic change on the environment. The total selection gradient of the genotype depends on the block matrix of *total effects of a mutant's genotype on her phenotype*, given by

$$\left.\frac{\mathrm{d}\mathbf{x}^{\mathsf{T}}}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\delta\mathbf{x}^{\mathsf{T}}}{\delta\mathbf{y}}\frac{\mathrm{d}\mathbf{x}^{\mathsf{T}}}{\mathrm{d}\mathbf{x}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{g}}\times N_{\mathrm{a}}N_{\mathrm{p}}}. \tag{10}$$

This matrix is a mechanistic counterpart of Fisher's (1918) additive effects of allelic substitution and of Wagner's (1984) developmental matrix. This matrix also gives the sensitivity of a recurrence of the form (1) to perturbations in parameters $\mathbf{y}$ at any time a. Moreover, this matrix can be interpreted as measuring the total developmental bias of the phenotype from the genotype (Uller et al., 2018).

The block matrix of *total effects of a mutant's phenotype on her phenotype* is

$$\left.\frac{\mathrm{d}\mathbf{x}^{\mathsf{T}}}{\mathrm{d}\mathbf{x}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(2\mathbf{I} - \frac{\delta\mathbf{x}^{\mathsf{T}}}{\delta\mathbf{x}}\right)^{-1}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{p}}\times N_{\mathrm{a}}N_{\mathrm{p}}}. \tag{11}$$

This matrix describes *developmental feedback*, is always invertible (SI section S5.1, Eq. S5.1.15), and gives the sensitivity of a recurrence of the form (1) to perturbations in state variables $\mathbf{x}$ at any time a. Moreover, this matrix can be interpreted as a lifetime collection of total immediate effects of the phenotype on itself. Partitioning this matrix in blocks of size $N_\mathrm{p} \times N_\mathrm{p}$, the aj-th block entry gives the total effects of the phenotype at age a on the phenotype at age j, that is,

$$\frac{\mathrm{d}\mathbf{x}_j^\intercal}{\mathrm{d}\mathbf{x}_a} = \begin{cases} \overset{\curvearrowright}{\prod_{k=a}^{j-1}} \frac{\delta \mathbf{x}_{k+1}^\intercal}{\delta \mathbf{x}_k} = \frac{\delta \mathbf{x}_{a+1}^\intercal}{\delta \mathbf{x}_a} \cdots \frac{\delta \mathbf{x}_j^\intercal}{\delta \mathbf{x}_{j-1}} & \text{for } j > a \\ \mathbf{I} & \text{for } j = a \\ \mathbf{0} & \text{for } j < a. \end{cases} \tag{12}$$

Since matrix multiplication is not commutative, the $\curvearrowright$ denotes right multiplication. Eq. (11) has the same form of an equation for total effects used in path analysis (Greene 1977, p. 380; see also Morrissey 2014, Eq. 2), particularly if $(\delta \mathbf{x}^\intercal / \delta \mathbf{x} - \mathbf{I})|_{\mathbf{y}=\bar{\mathbf{y}}}$ is interpreted as a matrix listing the path coefficients of "direct" effects of the phenotype on itself (if environmental traits are not explicitly considered in the analysis).

The block matrix of *total immediate effects of $\boldsymbol{\zeta}$ on a mutant's phenotype* is

$$\left.\frac{\delta \mathbf{x}^\intercal}{\delta \boldsymbol{\zeta}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\partial \mathbf{x}^\intercal}{\partial \boldsymbol{\zeta}} + \frac{\partial \boldsymbol{\epsilon}^\intercal}{\partial \boldsymbol{\zeta}} \frac{\partial \mathbf{x}^\intercal}{\partial \boldsymbol{\epsilon}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \tag{13}$$

for $\boldsymbol{\zeta} \in \{\mathbf{x}, \mathbf{y}, \bar{\mathbf{x}}, \bar{\mathbf{y}}\}$. This matrix depends on direct niche construction ($\partial \boldsymbol{\epsilon}^\intercal / \partial \boldsymbol{\zeta}$) and direct plasticity ($\partial \mathbf{x}^\intercal / \partial \boldsymbol{\epsilon}$). Thus, the developmental feedback of the phenotype (Eq. 11) depends on direct developmental bias from the phenotype ($\partial \mathbf{x}^\intercal / \partial \mathbf{x}$), direct niche-construction by the phenotype ($\partial \boldsymbol{\epsilon}^\intercal / \partial \mathbf{x}$), and direct plasticity of the phenotype ($\partial \mathbf{x}^\intercal / \partial \boldsymbol{\epsilon}$).

The total selection gradient of the genotype also depends on the block matrix of *total effects of a mutant's genotype on her environment*

$$\left.\frac{\mathrm{d}\boldsymbol{\epsilon}^\intercal}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\mathrm{d}\mathbf{x}^\intercal}{\mathrm{d}\mathbf{y}} \frac{\partial \boldsymbol{\epsilon}^\intercal}{\partial \mathbf{x}} + \frac{\partial \boldsymbol{\epsilon}^\intercal}{\partial \mathbf{y}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{a} N_\mathrm{g} \times N_\mathrm{a} N_\mathrm{e}}. \tag{14}$$

This matrix quantifies the total niche construction by the genotype and depends on direct niche construction by the phenotype and the genotype.

Additionally, the *direct selection gradient of the phenotype* is

$$\left.\frac{\partial w}{\partial \mathbf{x}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\partial w_1}{\partial \mathbf{x}_1}; \cdots ; \frac{\partial w_{N_\mathrm{a}}}{\partial \mathbf{x}_{N_\mathrm{a}}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{a} N_\mathrm{p} \times 1}, \tag{15}$$

where the direct selection gradient of the phenotype at age $a \in \{1, \ldots, N_\mathrm{a}\}$ is

$$\left.\frac{\partial w_a}{\partial \mathbf{x}_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \frac{1}{T}\left(\phi_a \left.\frac{\partial f_a}{\partial \mathbf{x}_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} + \pi_a \left.\frac{\partial p_a}{\partial \mathbf{x}_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}}\right) \in \mathbb{R}^{N_\mathrm{p} \times 1}, \tag{16}$$

and the direct effect on fertility of the mutant phenotype at age $a \in \{1, \ldots, N_\mathrm{a}\}$ is

$$\left.\frac{\partial f_a}{\partial \mathbf{x}_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\partial f_a}{\partial x_{1a}}; \cdots ; \frac{\partial f_a}{\partial x_{N_\mathrm{p} a}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{p} \times 1}. \tag{17}$$

The other direct selection gradients and direct effects on fertility or survival are defined analogously.

In turn, the block matrix of *direct effects of a mutant's phenotype on her phenotype* is

$$\left.\frac{\partial \mathbf{x}^{\mathsf{T}}}{\partial \mathbf{x}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\begin{pmatrix} \mathbf{I} & \frac{\partial \mathbf{x}_2^{\mathsf{T}}}{\partial \mathbf{x}_1} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I} & \frac{\partial \mathbf{x}_{N_\mathrm{a}}^{\mathsf{T}}}{\partial \mathbf{x}_{N_\mathrm{a}-1}} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{I} \end{pmatrix}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{a}N_\mathrm{p}\times N_\mathrm{a}N_\mathrm{p}}, \tag{18}$$

where the matrix of *direct effects of a mutant's phenotype at age a on her phenotype at age a* + 1 is

$$\left.\frac{\partial \mathbf{x}_{a+1}^{\mathsf{T}}}{\partial \mathbf{x}_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\begin{pmatrix} \frac{\partial x_{1,a+1}}{\partial x_{1a}} & \cdots & \frac{\partial x_{N_\mathrm{p},a+1}}{\partial x_{1a}} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_{1,a+1}}{\partial x_{N_\mathrm{p}a}} & \cdots & \frac{\partial x_{N_\mathrm{p},a+1}}{\partial x_{N_\mathrm{p}a}} \end{pmatrix}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{p}\times N_\mathrm{p}}. \tag{19}$$

Given the interpretations of $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\mathbf{y}$ and $\mathrm{d}\boldsymbol{\epsilon}^{\mathsf{T}}/\mathrm{d}\mathbf{y}$, the total selection gradient of the genotype (9) can then be interpreted as measuring total (directional) genotypic selection in a fitness landscape modified by: (1) the interaction of total developmental bias of the phenotype from the genotype and directional selection on the phenotype, and (2) the interaction of total niche construction by the genotype and environmental sensitivity of selection $\partial w/\partial \boldsymbol{\epsilon}$ (Chevin et al., 2010). In a standard quantitative genetics framework, the total selection gradient of the genotype would correspond to Lande's (1979) selection gradient of the genotype if phenotypic and environmental traits are not explicitly included in the analysis.

To build a minimal evo-devo dynamics model using our approach, one first needs expressions for fertility f_a, survival p_a, and development $\mathbf{g}_a$. Then, one computes the partial derivatives (17), (19), and analogous partial derivatives, and finally feeds these partial derivatives to Eqs. (9)-(11), (13)-(16), and (18) to compute the evo-evo dynamics using Eqs. (1)-(3).

4.2. Evolutionary dynamics of the phenotype in gradient form

The evo-devo dynamics above allow for the modeling of the evolutionary dynamics of the phenotype under explicit development, but do not provide an interpretation of the evolutionary dynamics of the phenotype as the climbing of a fitness landscape. To obtain such an interpretation, we obtain equations in gradient form for the evolutionary dynamics of the phenotype.

As a first step, temporarily assume that the following four conditions hold: (I) development is non-social ($\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), and there is (II) no exogenous plastic response of the phenotype ($[(\mathrm{d}\mathbf{x}/\mathrm{d}\boldsymbol{\epsilon}^{\mathsf{T}})(\partial\bar{\boldsymbol{\epsilon}}/\partial\tau)]\,|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), (III) no total immediate selection on the genotype ($\delta w/\delta\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), and (IV) no niche-constructed effects of the phenotype on fitness ($[(\partial\boldsymbol{\epsilon}^{\mathsf{T}}/\partial\mathbf{x})(\partial w/\partial\boldsymbol{\epsilon})]\,|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$). Then, in the limit as $\Delta\tau \to 0$, the evolutionary dynamics of the

resident phenotype satisfies

$$\frac{\mathrm{d}\bar{\mathbf{x}}}{\mathrm{d}\tau} \approx \iota \mathbf{H}_{\mathbf{x}} \left.\frac{\partial w}{\partial \mathbf{x}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \tag{20}$$

This is a mechanistic version of the Lande equation for the phenotype and is in gradient form, where the *mechanistic additive genetic covariance matrix of the phenotype* $\mathbf{x} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{p}}\times 1}$ is

$$\mathbf{H}_{\mathbf{x}} \equiv \mathrm{cov}[\mathbf{b}_{\mathbf{x}}, \mathbf{b}_{\mathbf{x}}] = \left.\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{y}^{\mathsf{T}}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\mathbf{x}^{\mathsf{T}}}{\mathrm{d}\mathbf{y}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{p}}\times N_{\mathrm{a}}N_{\mathrm{p}}} \tag{21}$$

(H for heredity), which guarantees that the developmental constraint (1) is met at all times given the formulas for the total effects of the genotype on phenotype provided in section 4.1. The matrix $\mathbf{H}_{\mathbf{x}}$ describes genetic covariation in the phenotype as the covariation of *mechanistic* breeding value $\mathbf{b}_{\mathbf{x}}$. Mechanistic breeding value is a mechanistic counterpart of breeding value, defined not in terms of regression coefficients but in terms of total derivatives and so has different properties explained in section 5.1.

Eq. (20) describes phenotypic evolution in gradient form but is dynamically insufficient since it depends on the resident genotype $\bar{\mathbf{y}}$ but does not describe its evolutionary dynamics. Indeed, the only reason that the resident phenotype $\bar{\mathbf{x}}$ evolves under the assumptions of Eq. (20) is that the resident genotype $\bar{\mathbf{y}}$ evolves, but the evolution of the resident genotype is not described by such equation. Hence, the mechanistic Lande equation (20) is not sufficient to describe phenotypic evolution as the climbing of a fitness landscape.

4.3. Evolutionary dynamics of the geno-envo-phenotype in gradient form

To describe phenotypic evolution as the climbing of a fitness landscape, we obtain dynamically sufficient equations in gradient form for phenotypic evolution. Dropping assumptions (I-IV) above, one such an equation describes the evolutionary dynamics of the resident geno-envo-phenotype as

$$\frac{\mathrm{d}\bar{\mathbf{m}}}{\mathrm{d}\tau} \approx \left.\left(\iota \mathbf{L}_{\mathbf{m}}\frac{\partial w}{\partial \mathbf{m}} + \frac{\mathrm{s}\mathbf{m}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}}\frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}}, \tag{22}$$

which is an extended mechanistic Lande equation. This equation describes the evolution of the geno-envo-phenotype as the climbing of a fitness landscape and is dynamically sufficient because it describes the evolution of all the variables involved, including the resident genotype $\bar{\mathbf{y}}$. This equation depends on the *mechanistic additive socio-genetic cross-covariance matrix of the geno-envo-phenotype*

$$\begin{aligned}\mathbf{L}_{\mathbf{m}} &\equiv \mathrm{cov}[\mathbf{b}^{\mathrm{s}}_{\mathbf{m}}, \mathbf{b}_{\mathbf{m}}] \\ &= \left.\left(\frac{\mathrm{s}\mathbf{m}}{\mathrm{s}\mathbf{y}^{\mathsf{T}}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\mathbf{m}^{\mathsf{T}}}{\mathrm{d}\mathbf{y}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}(N_{\mathrm{p}}+N_{\mathrm{g}}+N_{\mathrm{e}})\times N_{\mathrm{a}}(N_{\mathrm{p}}+N_{\mathrm{g}}+N_{\mathrm{e}})}\end{aligned} \tag{23}$$

(L for legacy). We say that the matrix $\mathbf{L}_{\mathbf{m}}$ describes the "socio-genetic" covariation of the geno-envo-phenotype, that is, the covariation between the stabilized mechanistic breeding value $\mathbf{b}^{\mathrm{s}}_{\mathbf{m}}$ and mechanistic breeding value $\mathbf{b}_{\mathbf{m}}$. Whereas mechanistic breeding value considers the total effect of mutations on the traits within the individual, stabilized mechanistic breeding value considers the stabilized effect of mutations on the traits across the population, after

social development has stabilized. The matrix $\mathbf{L_m}$ may be asymmetric and its main diagonal entries may be negative (unlike variances) due to social development. Moreover, the matrix $\mathbf{L_m}$ is always singular because $\mathrm{d}\mathbf{m}^\intercal/\mathrm{d}\mathbf{y}$ has fewer rows than columns, regardless of whether development is social. This means that there are always directions in geno-envo-phenotypic space in which there cannot be socio-genetic covariation, so there are always absolute socio-genetic constraints on adaptation of the geno-envo-phenotype (see SI section S2.2 for a definition of absolute mutational or genetic constraints). In particular, this implies that evolution does not generally stop at peaks of the fitness landscape where the direct selection gradient is zero.

Although the extended mechanistic Lande equation (22) is useful to interpret evolution as the climbing of a fitness landscape, in practice it may often be more useful to compute the evo-devo dynamics using Eqs. (1)-(3) since the extended mechanistic Lande equation may still require computing the evo-devo dynamics, particularly when development is social.

In the remainder of this section, we give the formulas for additional quantities involved in the extended mechanistic Lande equation (22). These formulas show that stabilized effects depend on social feedback (Eq. 27), joint total developmental bias (Eq. 26), and joint direct niche construction (Eqs. 29 and 32). Analysis of the obtained equations for genetic covariation and evolutionary dynamics is given in section 5. A reader interested in seeing an illustration of the method may jump to section 6.

The matrix $\mathbf{L_m}$ depends on the block matrix of *total effects of a mutant's genotype on her geno-envo-phenotype*

$$\left.\frac{\mathrm{d}\mathbf{m}^\intercal}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\begin{pmatrix} \frac{\mathrm{d}\mathbf{x}^\intercal}{\mathrm{d}\mathbf{y}} & \mathbf{I} & \frac{\mathrm{d}\boldsymbol{\epsilon}^\intercal}{\mathrm{d}\mathbf{y}} \end{pmatrix}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{a}N_\mathrm{g}\times N_\mathrm{a}(N_\mathrm{p}+N_\mathrm{g}+N_\mathrm{e})}, \tag{24}$$

measuring total developmental bias of the geno-envo-phenotype from the genotype, which is singular because it has fewer rows than columns. The matrix $\mathbf{L_m}$ also depends on the transpose of the matrix of *stabilized effects of a focal individual's genotype on the geno-envo-phenotype*

$$\left.\frac{\mathrm{s}\mathbf{m}}{\mathrm{s}\mathbf{y}^\intercal}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\mathbf{y}^\intercal}; \mathbf{I}; \frac{\mathrm{s}\boldsymbol{\epsilon}}{\mathrm{s}\mathbf{y}^\intercal}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{a}(N_\mathrm{p}+N_\mathrm{g}+N_\mathrm{e})\times N_\mathrm{a}N_\mathrm{g}}. \tag{25}$$

This matrix gives the total effects of the genotype on the geno-envo-phenotype after the effect of the mutation has propagated through the population through social development and such propagation has stabilized. This matrix can be interpreted as measuring stabilized developmental bias of the geno-envo-phenotype from the genotype.

Eq. (25) depends on the transpose of the matrix of *stabilized effects of a focal individual's phenotype or genotype on her phenotype*

$$\left.\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\mathbf{y}^\intercal}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left[\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\bar{\mathbf{x}}^\intercal}\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{y}^\intercal} + \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\bar{\mathbf{y}}^\intercal}\right)\right]\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_\mathrm{a}N_\mathrm{p}\times N_\mathrm{a}N_\mathrm{g}}. \tag{26}$$

This matrix can be interpreted as measuring stabilized developmental bias of the phenotype from the genotype. It depends on the effect of the genotype of the focal individual and of social partners on the phenotype, and on the feedback of such altered phenotype through the population (social feedback described by $\mathrm{s}\mathbf{x}/\mathrm{s}\bar{\mathbf{x}}^\intercal|_{\mathbf{y}=\bar{\mathbf{y}}}$). We say that stabilized developmental bias is "joint" as it includes both the effects of the focal individual and of social partners

(i.e., $\mathrm{d}\mathbf{x}/\mathrm{d}\boldsymbol{\zeta}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$ for $\boldsymbol{\zeta} \in \{\mathbf{y}, \bar{\mathbf{y}}\}$). If development is not social (i.e., $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), then the stabilized developmental bias matrix of the phenotype from the genotype ($\mathrm{s}\mathbf{x}/\mathrm{s}\mathbf{y}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$) reduces to the corresponding total developmental bias matrix ($\mathrm{d}\mathbf{x}/\mathrm{d}\mathbf{y}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$).

The matrix in Eq. 26 depends on *social feedback*, given by the transpose of the matrix of *stabilized effects of social partners' phenotypes on a focal individual's phenotype*

$$\left.\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\bar{\mathbf{x}}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left(\mathbf{I} - \left.\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\bar{\mathbf{x}}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}\right)^{-1} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{p}}\times N_{\mathrm{a}}N_{\mathrm{p}}}. \tag{27}$$

This matrix is invertible by our assumption that all the eigenvalues of $\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{x}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$ have absolute value strictly less than one, to guarantee that the resident is socio-devo stable. The matrix $\mathrm{s}\mathbf{x}/\mathrm{s}\bar{\mathbf{x}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$ can be interpreted as the total effects of social partners' phenotypes on a focal individual's phenotype after socio-devo stabilization; or vice versa, of a focal individual's phenotype on social partners' phenotypes. This matrix corresponds to an analogous matrix found in the indirect genetic effects literature (Moore et al., 1997, Eq. 19b and subsequent text). If development is not social from the phenotype (i.e., $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{x}}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), then the matrix $\mathrm{s}\mathbf{x}/\mathrm{s}\bar{\mathbf{x}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$ is the identity matrix. This is the only stabilized-effect matrix that does not reduce to the corresponding total-effect matrix when development is not social.

The transpose of the block matrix of *total effects of social partners' phenotype on a mutant's phenotype* is

$$\left.\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\bar{\mathbf{x}}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{x}^{\mathsf{T}}}\frac{\delta\mathbf{x}}{\delta\bar{\mathbf{x}}^{\mathsf{T}}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \tag{28}$$

This matrix can be interpreted as measuring total social developmental bias of the phenotype from phenotype, as well as the total effects on the phenotype of extra-genetic inheritance. It is also a mechanistic version of the matrix of interaction coefficients in the indirect genetic effects literature (i.e., $\boldsymbol{\Psi}$ in Eq. 17 of Moore et al. 1997, which is defined as a matrix of regression coefficients). Moreover, the matrix $\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{x}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$ can be interpreted as involving a developmentally immediate pulse of phenotype change caused by a change in social partners' traits followed by the triggered developmental feedback of the mutant's phenotype.

Eq. (25) also depends on the transpose of the matrix of *stabilized effects of a focal individual's genotype on the environment*

$$\left.\frac{\mathrm{s}\boldsymbol{\epsilon}}{\mathrm{s}\mathbf{y}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left[\left(\frac{\partial\boldsymbol{\epsilon}}{\partial\mathbf{x}^{\mathsf{T}}} + \frac{\partial\boldsymbol{\epsilon}}{\partial\bar{\mathbf{x}}^{\mathsf{T}}}\right)\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\mathbf{y}^{\mathsf{T}}} + \frac{\partial\boldsymbol{\epsilon}}{\partial\mathbf{y}^{\mathsf{T}}} + \frac{\partial\boldsymbol{\epsilon}}{\partial\bar{\mathbf{y}}^{\mathsf{T}}}\right]\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{e}}\times N_{\mathrm{a}}N_{\mathrm{g}}}. \tag{29}$$

This matrix can be interpreted as measuring stabilized niche construction by the genotype. This matrix is formed by stabilized developmental bias of the geno-phenotype from genotype followed by joint direct niche construction by the geno-phenotype (Layer 5, Eq. S6a). If development is not social (i.e., $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), then stabilized niche construction by the genotype ($\mathrm{s}\boldsymbol{\epsilon}/\mathrm{s}\mathbf{y}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$) reduces to total niche construction by the genotype ($\mathrm{d}\boldsymbol{\epsilon}/\mathrm{d}\mathbf{y}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$).

The extended mechanistic Lande equation (22) also depends on the transpose of the matrix of *stabilized effects of a focal individual's environment on the geno-envo-phenotype*

$$\left.\frac{\mathrm{s}\mathbf{m}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}}; \mathbf{0}; \frac{\mathrm{s}\boldsymbol{\epsilon}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}(N_{\mathrm{p}}+N_{\mathrm{g}}+N_{\mathrm{e}})\times N_{\mathrm{a}}N_{\mathrm{e}}}, \tag{30}$$

measuring stabilized plasticity of the geno-envo-phenotype.

The transpose of the matrix of *stabilized effects of a focal individual's environment on the phenotype* is

$$\left.\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\bar{\mathbf{x}}^{\mathsf{T}}}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{x}^{\mathsf{T}}}\frac{\partial\mathbf{x}}{\partial\boldsymbol{\epsilon}^{\mathsf{T}}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{p}}\times N_{\mathrm{a}}N_{\mathrm{e}}}. \tag{31}$$

This matrix measures the stabilized plasticity of the phenotype, where the environment first causes direct plasticity in a focal individual, and then there is developmental and social feedback. In contrast to Eq. (26), stabilized plasticity does not depend on the joint effects of the environment. If development is not social (i.e., $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), then stabilized plasticity reduces to total plasticity.

Stabilized environmental feedback is given by the transpose of the matrix of *stabilized effects of a focal individual's environment on the environment*

$$\left.\frac{\mathrm{s}\boldsymbol{\epsilon}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left[\left(\frac{\partial\boldsymbol{\epsilon}}{\partial\mathbf{x}^{\mathsf{T}}} + \frac{\partial\boldsymbol{\epsilon}}{\partial\bar{\mathbf{x}}^{\mathsf{T}}}\right)\frac{\mathrm{s}\mathbf{x}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}} + \mathbf{I}\right]\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{e}}\times N_{\mathrm{a}}N_{\mathrm{e}}}. \tag{32}$$

This matrix depends on stabilized plasticity of the phenotype, joint direct niche construction by the phenotype, and direct mutual environmental dependence. Such direct mutual environmental dependence is here the identity matrix because of our assumption that environmental traits are mutually independent.

5. Top layers of the evo-devo process

The equations listed in section 4 are part of a broader set of equations that we derive to analyze the evolutionary dynamics under explicit development. We term "evo-devo process" this set of equations. The evo-devo process can be arranged in a layered structure, where each layer is formed by components in layers below (Fig. 4). This layered structure helps see how complex interactions between variables involved in genetic covariation are formed by building blocks describing the direct interaction between variables. In this section 5, we present and analyze the top two layers of the evo-devo process, namely, the layer of genetic covariation and the layer of evolutionary dynamics (Fig. 4A-D). In SI section S3, we describe and analyze the bottom layers, starting from the layer of elementary components up to the layer of stabilized effects (Fig. 4E-I). Sections 5, S3, and S5 include multiple ancillary results. For instance, SI section S3.4 includes closed-form equations for the total selection gradients of the genotype, phenotype, environment, geno-phenotype, and geno-envo-phenotype, and alternative arrangements thereof.

5.1. Layer 6: genetic covariation

In this section, we describe and analyze the layer of genetic covariation of the evo-devo process. We first define mechanistic breeding value under our adaptive dynamics assumptions, which allows us to define mechanistic additive genetic covariance matrices under our assumptions. Then, we define (socio-devo) stabilized mechanistic breeding value, which we use to define mechanistic additive socio-genetic cross-covariance matrices. The notions of stabilized mechanistic breeding values and mechanistic socio-genetic cross-covariance generalize the corresponding notions of mechanistic breeding value and mechanistic additive genetic covariance to consider the effects of social development.

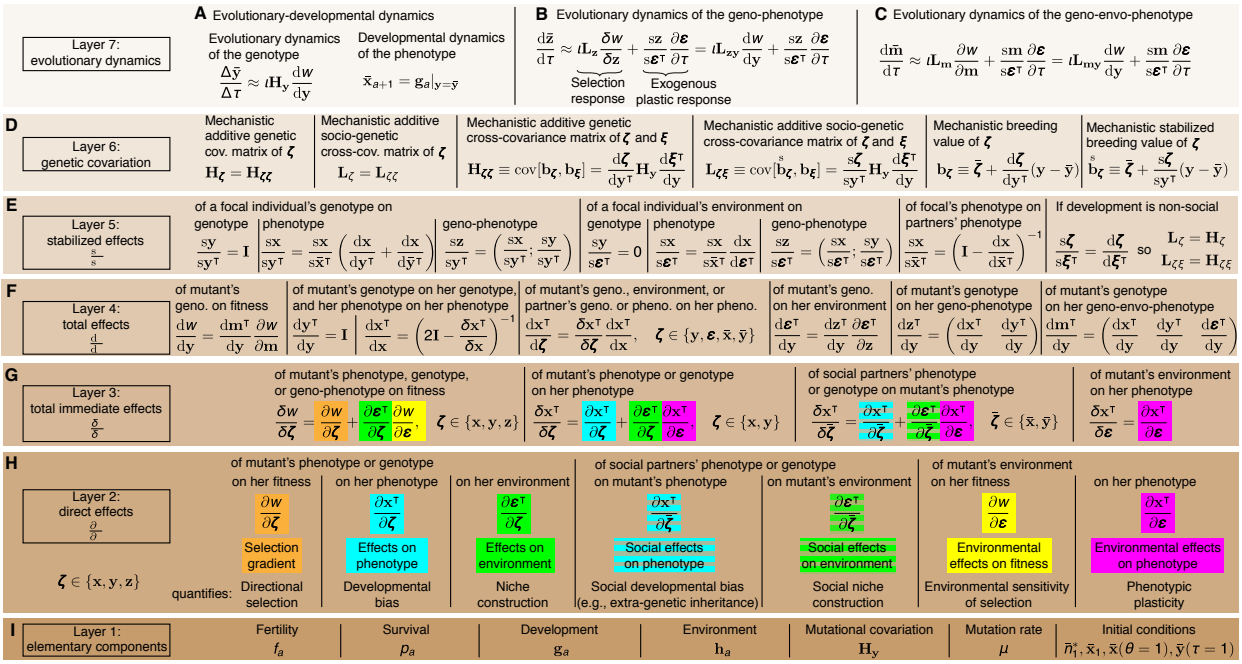

Figure 4: The evo-devo process and its layered structure. Here we summarize the equations composing the evo-devo process arranged in a layered structure. Each layer is formed by components in layers below. Layer 7 describes the evolutionary dynamics as (A) evo-devo dynamics, which in the limit as $\Delta\tau \to 0$ implies (B) the evolutionary dynamics of the geno-phenotype, and (C) the evolutionary dynamics of the geno-envo-phenotype. (D) Layer 6 describes genetic covariation. (E) Layer 5 describes stabilized effects (total derivatives over life after socio-devo stabilization, denoted by s/s). (F) Layer 4 describes total effects (total derivatives over life before socio-devo stabilization, denoted by d/d). (G) Layer 3 describes total immediate effects (total derivatives at the current age, denoted by δ/δ). (H) Layer 2 describes direct effects (partial derivatives, denoted by ∂/∂). (I) Layer 1 comprises the elementary components of the evo-devo process that generate all layers above. All derivatives are evaluated at $\mathbf{y} = \bar{\mathbf{y}}$. See SI section S3.2 for the equations of direct-effect matrices, which have structure due to age structure. See Fig. 1 and Table 1 for the meaning of symbols.

We follow the standard definition of breeding value to define its mechanistic analogue under our assumptions. The breeding value of a trait is defined under quantitative genetics assumptions as the best linear estimate of the trait from gene content (Lynch and Walsh, 1998; Walsh and Lynch, 2018). As stated above, under quantitative genetics assumptions, the i-th trait value x_i is written as $x_i = \bar{x}_i + \sum_j \alpha_{ij}(y_j - \bar{y}_j) + e_i$, where the overbar denotes population average, y_j is the j-th predictor (gene content in j-th locus), α_{ij} is the partial least-square regression coefficient of $x_i - \bar{x}_i$ vs $y_j - \bar{y}_j$, and e_i is the residual; the breeding value of x_i is $a_i \equiv \bar{x}_i + \sum_j \alpha_{ij}(y_j - \bar{y}_j)$. Accordingly, we define the mechanistic breeding value $\mathbf{b}_\zeta$ of a vector $\boldsymbol{\zeta}$ as its first-order estimate with respect to genotypic traits $\mathbf{y}$ around the resident genotypic traits $\bar{\mathbf{y}}$:

$$\begin{aligned}\mathbf{b}_\zeta &\equiv \zeta|_{\mathbf{y}=\bar{\mathbf{y}}} + \left.\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^\intercal}\right|_{\mathbf{y}=\bar{\mathbf{y}}} (\mathbf{y} - \bar{\mathbf{y}}) \\ &= \bar{\boldsymbol{\zeta}} + \left.\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^\intercal}\right|_{\mathbf{y}=\bar{\mathbf{y}}} (\mathbf{y} - \bar{\mathbf{y}}). \qquad \text{(Layer 6, Eq. 1)}\end{aligned}$$

The key difference of this definition with that of breeding value is that rather than using regression coefficients, this

definition uses the total derivatives of $\boldsymbol{\zeta}$ with respect to the genotype, $\mathrm{d}\boldsymbol{\zeta}/\mathrm{d}\mathbf{y}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}}$, which are a mechanistic analogue of Fisher's additive effect of allelic substitution (his α; see Eq. I of Fisher 1918 and p. 72 of Lynch and Walsh 1998).

There are material differences between breeding value and its mechanistic counterpart, as made evident with heritability. Because breeding value under quantitative genetics uses linear regression via least squares, breeding value a_i is guaranteed to be uncorrelated with the residual e_i. This guarantees that heritability is between zero and one. Indeed, the (narrow sense) heritability of trait x_i is defined as $h^2 = \mathrm{var}[a_i]/\mathrm{var}[x_i]$, where using $x_i = a_i + e_i$ we have $\mathrm{var}[x_i] = \mathrm{var}[a_i] + \mathrm{var}[e_i] + 2\mathrm{cov}[a_i, e_i]$. The latter covariance is zero due to least squares, and so $h^2 \in [0, 1]$. In contrast, mechanistic breeding values may be correlated with residuals. Indeed, in our framework we have that the phenotype $x_{ia} = b_{ia} + O(\|\mathbf{y} - \bar{\mathbf{y}}\|^2)$, but mechanistic breeding value b_{ia} is not computed via least squares, so b_{ia} and the error $O(\|\mathbf{y} - \bar{\mathbf{y}}\|^2)$ may covary, positively or negatively. Hence, the classic quantitative genetics partition of phenotypic variance into genetic and "environmental" (i.e., residual) variance does not hold with mechanistic breeding value, as there may be mechanistic genetic and "environmental" covariance. Consequently, since the covariance between two random variables is bounded from below by the negative of the product of their standard deviations, mechanistic heritability defined as the ratio between the variance of mechanistic breeding value and phenotypic variance cannot be negative but it may be greater than one.

Our definition of mechanistic breeding value recovers Fisher's (1918) infinitesimal model under certain conditions, but we do not assume the infinitesimal model. According to Fisher's (1918) infinitesimal model, the normalized breeding value excess is normally distributed as the number of loci approaches infinity. Using Layer 6, Eq. 1, we have that the mechanistic breeding value excess for the i-th entry of $\mathbf{b}_{\boldsymbol{\zeta}}$ is

$$b_{\zeta_i} - \bar{\zeta}_i = \sum_{k=1}^{N_\mathrm{g}} \sum_{a=1}^{N_\mathrm{a}} \left.\frac{\mathrm{d}\zeta_i}{\mathrm{d}y_{ka}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} (y_{ka} - \bar{y}_{ka}).$$

Let us denote the mutational variance for the k-th genotypic trait at age a by

$$\sigma_{ka}^2 = \mathrm{E}[(y_{ka} - \bar{y}_{ka})^2],$$

and let us denote the total mutational variance by

$$s_{N_\mathrm{g}N_\mathrm{a}}^2 = \sum_{k=1}^{N_\mathrm{g}} \sum_{a=1}^{N_\mathrm{a}} \sigma_{ka}^2.$$

If the y_{ka} are mutually independent and Lyapunov's condition is satisfied, from the Lyapunov central limit theorem we have that, as either the number of genotypic traits N_g or the number of ages N_a tends to infinity (e.g., by reducing the age bin size), the normalized mechanistic breeding value excess

$$\frac{1}{s_{N_\mathrm{g}N_\mathrm{a}}}(b_{\zeta_i} - \bar{\zeta}_i)$$

is normally distributed with mean zero and variance 1. Thus, this limit yields Fisher's (1918) infinitesimal model, although we do not assume such limit. Our framework thus recovers the infinitesimal model as a particular case,

when either N_{g} or N_{a} approaches infinity (provided that the y_{ka} are mutually independent and Lyapunov's condition holds).

From our definition of mechanistic breeding value, we have that the mechanistic breeding value of the genotype is the genotype itself. Indeed, the mechanistic breeding value of the genotype $\mathbf{y}$ is

$$\mathbf{b}_{\mathbf{y}} = \bar{\mathbf{y}} + \left.\frac{\mathrm{d}\mathbf{y}}{\mathrm{d}\mathbf{y}^{\intercal}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} (\mathbf{y} - \bar{\mathbf{y}}) = \bar{\mathbf{y}} + \mathbf{y} - \bar{\mathbf{y}} = \mathbf{y},$$

since $\mathrm{d}\mathbf{y}/\mathrm{d}\mathbf{y}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{I}$ because, by assumption, the genotype does not have developmental constraints and is developmentally independent (Layer 4, Eq. S5). In turn, from Layer 6, Eq. 1, the expected mechanistic breeding value of vector $\boldsymbol{\zeta}$ is

$$\bar{\mathbf{b}}_{\zeta} \equiv \mathrm{E}[\mathbf{b}_{\zeta}] = \bar{\boldsymbol{\zeta}}.$$

We now define mechanistic additive genetic covariance matrices under our assumptions. The additive genetic variance of a trait is defined under quantitative genetics assumptions as the variance of its breeding value, which is extended to the multivariate case so the additive genetic covariance matrix of a trait vector is the covariance matrix of the traits' breeding values (Lynch and Walsh, 1998; Walsh and Lynch, 2018). Accordingly, we define the *mechanistic additive genetic covariance matrix* of a vector $\boldsymbol{\zeta} \in \mathbb{R}^{m\times 1}$ as the covariance matrix of its mechanistic breeding value:

$$\begin{aligned}
\mathbf{H}_{\zeta} &\equiv \mathrm{cov}[\mathbf{b}_{\zeta}, \mathbf{b}_{\zeta}] \\
&= \mathrm{E}[(\mathbf{b}_{\zeta} - \bar{\mathbf{b}}_{\zeta})(\mathbf{b}_{\zeta} - \bar{\mathbf{b}}_{\zeta})^{\intercal}] = \mathrm{E}[(\mathbf{b}_{\zeta} - \bar{\boldsymbol{\zeta}})(\mathbf{b}_{\zeta} - \bar{\boldsymbol{\zeta}})^{\intercal}] \\
&= \mathrm{E}\left[\left(\left.\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^{\intercal}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} (\mathbf{y} - \bar{\mathbf{y}})\right)\left(\left.\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^{\intercal}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} (\mathbf{y} - \bar{\mathbf{y}})\right)^{\intercal}\right] \\
&= \mathrm{E}\left[\left.\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^{\intercal}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} (\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^{\intercal} \left.\frac{\mathrm{d}\boldsymbol{\zeta}^{\intercal}}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}\right] \\
&= \left.\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^{\intercal}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \mathrm{E}\left[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^{\intercal}\right] \left.\frac{\mathrm{d}\boldsymbol{\zeta}^{\intercal}}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\
&= \left.\left(\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^{\intercal}} \mathbf{H}_{\mathbf{y}} \frac{\mathrm{d}\boldsymbol{\zeta}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{m\times m}, \qquad \text{(Layer 6, Eq. 2)}
\end{aligned}$$

where the fourth line follows from the property of the transpose of a product (i.e., $(\mathbf{AB})^{\intercal} = \mathbf{B}^{\intercal}\mathbf{A}^{\intercal}$) and the last line follows since the mechanistic additive genetic covariance matrix of the genotype $\mathbf{y}$ is

$$\mathbf{H}_{\mathbf{y}} \equiv \mathrm{cov}[\mathbf{b}_{\mathbf{y}}, \mathbf{b}_{\mathbf{y}}] = \mathrm{cov}[\mathbf{y}, \mathbf{y}] \in \mathbb{R}^{N_{\mathrm{a}}N_{\mathrm{g}}\times N_{\mathrm{a}}N_{\mathrm{g}}}.$$

Layer 6, Eq. 2 has the same form of previous expressions for the additive genetic covariance matrix under quantitative genetics assumptions, although using least-square regression coefficients in place of the derivatives if the classic partitioning of phenotypic variance is to hold (see Eq. II of Fisher 1918, Eq. + of Wagner 1984, Eq. 3.5b of Barton and Turelli 1987, and Eq. 4.23b of Lynch and Walsh 1998; see also Eq. 22a of Lande 1980, Eq. 3 of Wagner 1989, and Eq. 9 of Charlesworth 1990). We denote the matrix $\mathbf{H}$ rather than $\mathbf{G}$ to note that the two are different, particularly as the former is based on mechanistic breeding value. Note $\mathbf{H}_{\zeta}$ is symmetric.

In some cases, Layer 6, Eq. 2 allows one to immediately determine whether a mechanistic additive genetic covariance matrix is singular, which means there are directions in the matrix's space in which there is no genetic variation. Indeed, a matrix with fewer rows than columns is always singular (Horn and Johnson, 2013, section 0.5 second line), and if the product $\mathbf{AB}$ is well-defined and $\mathbf{B}$ is singular, then $\mathbf{AB}$ is singular (this is easily checked to hold). Hence, from Layer 6, Eq. 2 it follows that $\mathbf{H}_{\zeta}$ is necessarily singular if $\mathrm{d}\boldsymbol{\zeta}^{\intercal}/\mathrm{d}\mathbf{y}$ has fewer rows than columns, that is, if $\mathbf{y}$ has fewer entries than $\boldsymbol{\zeta}$. Since $\mathbf{y}$ has $N_{\mathrm{a}}N_{\mathrm{g}}$ entries and $\boldsymbol{\zeta}$ has m entries, then $\mathbf{H}_{\zeta}$ is singular if $N_{\mathrm{a}}N_{\mathrm{g}} < m$. Moreover, Layer 6, Eq. 2 allows one to immediately identify bounds for the "degrees of freedom" of genetic covariation, that is, for the rank of $\mathbf{H}_{\zeta}$. Indeed, for a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$, we have that the rank of $\mathbf{A}$ is at most the smallest value of m and n, that is, $\mathrm{rank}(\mathbf{A}) \leq \min\{m, n\}$ (Horn and Johnson, 2013, section 0.4.5 (a)). Furthermore, from the Frobenius inequality (Horn and Johnson, 2013, section 0.4.5 (e)), for a well-defined product $\mathbf{AB}$, we have that $\mathrm{rank}(\mathbf{AB}) \leq \mathrm{rank}(\mathbf{B})$. Therefore, for $\boldsymbol{\zeta} \in \mathbb{R}^{m\times 1}$, we have that

$$\mathrm{rank}(\mathbf{H}_{\zeta}) \leq \min\{N_{\mathrm{a}}N_{\mathrm{g}}, m\}. \qquad \text{(Layer 6, Eq. 3)}$$

Intuitively, this states that the degrees of freedom of genetic covariation are at most given by the lifetime number of genotypic traits (i.e., $N_{\mathrm{a}}N_{\mathrm{g}}$). So if there are more entries in $\boldsymbol{\zeta}$ than there are lifetime genotypic traits, then there are fewer degrees of freedom of genetic covariation than traits. This point is mathematically trivial but biologically crucial because the evolutionary dynamic equations in gradient form that are generally dynamically sufficient involve an $\mathbf{H}_{\zeta}$ whose $\boldsymbol{\zeta}$ necessarily has more entries than $\mathbf{y}$. These points on the singularity and rank of $\mathbf{H}_{\zeta}$ also hold under quantitative genetics assumptions, where the same structure (Layer 6, Eq. 2) holds, except that $\mathbf{H}_{\mathbf{y}}$ does not refer to mutational variation but to standing variation in allele frequency and total effects are measured with regression coefficients. Considering standing variation in $\mathbf{H}_{\mathbf{y}}$ and regression coefficients does not affect the points made in this paragraph.

Consider the following slight generalization of the mechanistic additive genetic covariance matrix. We define the mechanistic additive genetic cross-covariance matrix between a vector $\boldsymbol{\zeta} \in \mathbb{R}^{m\times 1}$ and a vector $\boldsymbol{\xi} \in \mathbb{R}^{n\times 1}$ as the cross-covariance matrix of their mechanistic breeding value:

$$\mathbf{H}_{\zeta\xi} \equiv \mathrm{cov}[\mathbf{b}_{\zeta}, \mathbf{b}_{\xi}] = \left(\frac{\mathrm{d}\boldsymbol{\zeta}}{\mathrm{d}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\boldsymbol{\xi}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{m\times n}. \qquad \text{(Layer 6, Eq. 4)}$$

Thus, $\mathbf{H}_{\zeta\zeta} = \mathbf{H}_{\zeta}$. Note $\mathbf{H}_{\zeta\xi}$ may be rectangular, and if square, asymmetric. Again, from Layer 6, Eq. 4 it follows that $\mathbf{H}_{\zeta\xi}$ is necessarily singular if there are fewer entries in $\mathbf{y}$ than in $\boldsymbol{\xi}$ (i.e., if $N_{\mathrm{a}}N_{\mathrm{g}} < n$). Also, for $\boldsymbol{\xi} \in \mathbb{R}^{n\times 1}$, have that

$$\mathrm{rank}(\mathbf{H}_{\zeta\xi}) \leq \min\{N_{\mathrm{a}}N_{\mathrm{g}}, n\}.$$

In words, the degrees of freedom of genetic cross-covariation are at most given by the lifetime number of genotypic traits.

The mechanistic additive genetic covariance matrix of the phenotype takes the following form. Evaluating Layer

6, Eq. 2 at $\boldsymbol{\zeta} = \mathbf{x}$, the mechanistic additive genetic covariance matrix of the phenotype $\mathbf{x} \in \mathbb{R}^{N_a N_p \times 1}$ is

$$\mathbf{H}_{\mathbf{x}} = \left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{y}^{\intercal}} \mathbf{H}_{\mathbf{y}} \frac{\mathrm{d}\mathbf{x}^{\intercal}}{\mathrm{d}\mathbf{y}} \right)\Bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_a N_p \times N_a N_p}, \qquad \text{(Layer 6, Eq. 5)}$$

which is singular because the developmental matrix $\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}}$ is singular since the developmentally initial phenotype is not affected by the genotype and the developmentally final genotypic traits do not affect the phenotype (SI section S5.2, Eq. S5.2.16). This singularity is immaterial for our purposes because a dynamical system consisting only of evolutionary dynamic equations for the phenotype thus having an associated $\mathbf{H}_{\mathbf{x}}$-matrix is underdetermined in general because the system has fewer dynamic equations (i.e., the number of entries in $\mathbf{x}$) than dynamic variables (i.e., the number of entries in $(\mathbf{x}; \mathbf{y}; \boldsymbol{\epsilon})$). Indeed, the evolutionary dynamics of the phenotype generally depends on the resident genotype, in particular, because the developmental matrix depends on the resident genotype (Layer 4, Eq. S2), for instance, due to non-linearities in the developmental map involving products between genotypic traits, or between genotypic traits and phenotypes, or between genotypic traits and environmental traits, that is, gene-gene interaction, gene-phenotype interaction, and gene-environment interaction, respectively. Thus, evolutionary dynamic equations of the phenotype alone generally have either zero or an infinite number of solutions for any given initial condition and are thus dynamically insufficient. To have a determined system in gradient form that is dynamically sufficient in general, we follow the evolutionary dynamics of both the phenotype and the genotype, that is, of the geno-phenotype, which depends on $\mathbf{H}_{\mathbf{z}}$ rather than $\mathbf{H}_{\mathbf{x}}$ alone.

The mechanistic additive genetic covariance matrix of the geno-phenotype takes the following form. Evaluating Layer 6, Eq. 2 at $\boldsymbol{\zeta} = \mathbf{z}$, the mechanistic additive genetic covariance matrix of the geno-phenotype $\mathbf{z} \in \mathbb{R}^{N_a(N_p+N_g)\times 1}$ is

$$\mathbf{H}_{\mathbf{z}} = \left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{y}^{\intercal}} \mathbf{H}_{\mathbf{y}} \frac{\mathrm{d}\mathbf{z}^{\intercal}}{\mathrm{d}\mathbf{y}} \right)\Bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_a(N_p+N_g)\times N_a(N_p+N_g)}. \qquad \text{(Layer 6, Eq. 6)}$$

This matrix is necessarily singular because the geno-phenotype $\mathbf{z}$ includes the genotype $\mathbf{y}$ so $\mathrm{d}\mathbf{z}^{\intercal}/\mathrm{d}\mathbf{y}$ has fewer rows than columns (Layer 4, Eq. S7). Intuitively, Layer 6, Eq. 6 has this form because the phenotype is related to the genotype by the developmental constraint (1). From Layer 6, Eq. 3, the rank of $\mathbf{H}_{\mathbf{z}}$ has an upper bound given by the number of genotypic traits across life (i.e., $N_a N_g$), so $\mathbf{H}_{\mathbf{z}}$ has at least $N_a N_p$ eigenvalues that are exactly zero. Thus, $\mathbf{H}_{\mathbf{z}}$ is singular if there is at least one trait that is developmentally constructed according to the developmental constraint (1) (i.e., if $N_p > 0$). This is a mathematically trivial singularity, but it is biologically key because it is $\mathbf{H}_{\mathbf{z}}$ rather than $\mathbf{H}_{\mathbf{x}}$ that occurs in a generally dynamically sufficient evolutionary system in gradient form (provided the environment is constant; if the environment is not constant, the relevant matrix is $\mathbf{H}_{\mathbf{m}}$ which is also always singular if there is at least one phenotype or one environmental trait).

Another way to see the singularity of $\mathbf{H}_{\mathbf{z}}$ is the following. From Layer 6, Eq. 6, we can write the mechanistic additive genetic covariance matrix of the geno-phenotype as

$$\mathbf{H}_{\mathbf{z}} = \begin{pmatrix} \mathbf{H}_{\mathbf{zx}} & \mathbf{H}_{\mathbf{zy}} \end{pmatrix},$$

where the mechanistic additive genetic cross-covariance matrix between $\mathbf{z}$ and $\mathbf{x}$ is

$$\mathbf{H}_{\mathbf{zx}} = \left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\mathbf{x}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}(N_{\mathrm{p}}+N_{\mathrm{g}})\times N_{\mathrm{a}}N_{\mathrm{p}}},$$

and the mechanistic additive genetic cross-covariance matrix between $\mathbf{z}$ and $\mathbf{y}$ is

$$\mathbf{H}_{\mathbf{zy}} = \left(\frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\mathbf{y}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{N_{\mathrm{a}}(N_{\mathrm{p}}+N_{\mathrm{g}})\times N_{\mathrm{a}}N_{\mathrm{g}}}.$$

Thus, using Layer 4, Eq. S5, we have that

$$\mathbf{H}_{\mathbf{zx}} = \mathbf{H}_{\mathbf{zy}}\frac{\mathrm{d}\mathbf{x}^{\intercal}}{\mathrm{d}\mathbf{y}}\bigg|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Layer 6, Eq. 7)}$$

That is, some columns of $\mathbf{H}_{\mathbf{z}}$ (i.e., those in $\mathbf{H}_{\mathbf{zx}}$) are linear combinations of other columns of $\mathbf{H}_{\mathbf{z}}$ (i.e., those in $\mathbf{H}_{\mathbf{zy}}$). Hence, $\mathbf{H}_{\mathbf{z}}$ is singular.

The mechanistic additive genetic covariance matrix of the geno-phenotype is singular because the geno-phenotype includes the genotype ("gene content"). The singularity arises because the mechanistic breeding value of the phenotype is a linear combination of the mechanistic breeding value of the genotype by definition of mechanistic breeding value, regardless of whether the phenotype is a linear function of the genotype and regardless of the number of phenotypic or genotypic traits. In quantitative genetics terms, the $\mathbf{G}$-matrix is a function of allele frequencies (which corresponds to our $\bar{\mathbf{y}}$), so a generally dynamically sufficient Lande system would require that allele frequencies are part of the dynamic variables considered; consequently, if the geno-phenotypic vector $\bar{\mathbf{z}}$ includes allele frequencies $\bar{\mathbf{y}}$, then the associated $\mathbf{G}$ is necessarily singular since by definition, breeding value under quantitative genetics assumptions is a linear combination of gene content.

The definition of mechanistic breeding value implies that if there is only one phenotype and one genotypic trait, with a single age each, then there is a perfect correlation between their mechanistic breeding values (i.e., their correlation coefficient is 1). This also holds under quantitative genetics assumptions (Via and Lande, 1985), in which case the breeding value a of a trait x is a linear combination of a single predictor y, so the breeding value a and predictor y are perfectly correlated (i.e., $\mathrm{cov}[a, y]/\sqrt{\mathrm{var}[a]\mathrm{var}[y]} = \mathrm{cov}[\alpha y, y]/\sqrt{\mathrm{var}[\alpha y]\mathrm{var}[y]} = (\alpha/\alpha)\mathrm{cov}[y, y]/\sqrt{\mathrm{var}[y]\mathrm{var}[y]} = 1$). The perfect correlation between a single breeding value and a single predictor arises because, by definition, breeding value excludes the residual e. Note this does not mean that the phenotype and genotype are linearly related: it is (mechanistic) breeding values and the genotype that are linearly related by definition of (mechanistic) breeding value (Layer 6, Eq. 1).

A standard approach to remove the singularity of an additive genetic covariance matrix is to remove some traits from the analysis (Lande, 1979). To remove the singularity of $\mathbf{H}_{\mathbf{z}}$ we would need to remove at least either all phenotypic traits or all genotypic traits from the analysis. However, removing all phenotypic traits from the analysis prevents analyzing phenotypic evolution as the climbing of a fitness landscape whereas removing all genotypic traits from the analysis renders the analysis dynamically insufficient in general because the evolutionary dynamics of some

variables is not described. Thus, in general, to analyze a dynamically sufficient description of phenotypic evolution as the climbing of a fitness landscape, we must keep the singularity of $\mathbf{H}_{\mathbf{z}}$.

We now extend the notion of mechanistic breeding value to consider social development. As stated above, we denote a matrix of stabilized effects as $\mathrm{s}\boldsymbol{\zeta}^{\intercal}/\mathrm{s}\boldsymbol{\xi}$, which gives the total effects of $\boldsymbol{\xi}$ on $\boldsymbol{\zeta}^{\intercal}$ over the individual's development and after the effects of the perturbation have stabilized in the population. We define the stabilized mechanistic breeding value of a vector $\boldsymbol{\zeta}$ as:

$$\begin{aligned}\mathbf{b}^{\mathrm{s}}_{\zeta} &\equiv \boldsymbol{\zeta}|_{\mathbf{y}=\bar{\mathbf{y}}} + \left.\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\mathbf{y}^{\intercal}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}(\mathbf{y}-\bar{\mathbf{y}}) \\ &= \bar{\boldsymbol{\zeta}} + \left.\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\mathbf{y}^{\intercal}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}(\mathbf{y}-\bar{\mathbf{y}}). \qquad \text{(Layer 6, Eq. 8)}\end{aligned}$$

The stabilized-effect matrix $\mathrm{s}\boldsymbol{\xi}/\mathrm{s}\mathbf{y}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}}$ equals the total-effect matrix $\mathrm{d}\boldsymbol{\xi}/\mathrm{d}\mathbf{y}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}}$ if development is non-social (SI section S3.5). Thus, if development is non-social, the stabilized mechanistic breeding value $\mathbf{b}^{\mathrm{s}}_{\zeta}$ equals the mechanistic breeding value $\mathbf{b}_{\zeta}$. Also, note that $\mathrm{E}[\mathbf{b}^{\mathrm{s}}_{\zeta}] = \bar{\boldsymbol{\zeta}}$.

We use this definition to extend the notion of mechanistic additive genetic covariance matrix to include the effects of social development as follows. We define the *mechanistic additive socio-genetic cross-covariance matrix of* $\boldsymbol{\zeta} \in \mathbb{R}^{m\times 1}$ as

$$\begin{aligned}\mathbf{L}_{\zeta} &\equiv \mathrm{cov}[\mathbf{b}^{\mathrm{s}}_{\zeta}, \mathbf{b}_{\zeta}] \\ &= \left.\left(\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\boldsymbol{\zeta}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{m\times m}. \qquad \text{(Layer 6, Eq. 9)}\end{aligned}$$

Note $\mathbf{L}_{\zeta}$ may be asymmetric and its main diagonal entries may be negative (unlike variances). If development is non-social, $\mathbf{L}_{\zeta}$ equals $\mathbf{H}_{\zeta}$. As before, $\mathbf{L}_{\zeta}$ is singular if $\boldsymbol{\zeta}$ has fewer entries than $\mathbf{y}$. Also, for $\boldsymbol{\zeta} \in \mathbb{R}^{m\times 1}$, have that

$$\mathrm{rank}(\mathbf{L}_{\zeta}) \leq \min\{N_{\mathrm{a}}N_{\mathrm{g}}, m\}.$$

That is, the degrees of freedom of socio-genetic covariation are at most also given by the lifetime number of genotypic traits.

As before, we generalize this notion and define the *mechanistic additive socio-genetic cross-covariance matrix between* $\boldsymbol{\zeta} \in \mathbb{R}^{m\times 1}$ *and* $\boldsymbol{\xi} \in \mathbb{R}^{n\times 1}$ as

$$\begin{aligned}\mathbf{L}_{\zeta\xi} &\equiv \mathrm{cov}[\mathbf{b}^{\mathrm{s}}_{\zeta}, \mathbf{b}_{\xi}] \\ &= \left.\left(\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\boldsymbol{\xi}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{m\times n}. \qquad \text{(Layer 6, Eq. 10)}\end{aligned}$$

Again, if development is non-social, $\mathbf{L}_{\zeta\xi}$ equals $\mathbf{H}_{\zeta\xi}$. Note $\mathbf{L}_{\zeta\xi}$ may be rectangular and, if square, asymmetric. Also, $\mathbf{L}_{\zeta\xi}$ is singular if $\boldsymbol{\xi}$ has fewer entries than $\mathbf{y}$. For $\boldsymbol{\xi} \in \mathbb{R}^{n\times 1}$, have that

$$\mathrm{rank}(\mathbf{L}_{\zeta\xi}) \leq \min\{N_{\mathrm{a}}N_{\mathrm{g}}, n\}.$$

That is, the degrees of freedom of socio-genetic cross-covariation are at most still given by the lifetime number of genotypic traits.

In particular, some $\mathbf{L}_{\zeta\xi}$ matrices are singular or not as follows. The mechanistic additive socio-genetic cross-covariance matrix between $\boldsymbol{\zeta}$ and the geno-phenotype $\mathbf{z}$

$$\mathbf{L}_{\zeta\mathbf{z}} = \left(\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\mathbf{z}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{m\times N_{\mathrm{a}}(N_{\mathrm{p}}+N_{\mathrm{g}})} \qquad \text{(Layer 6, Eq. 11)}$$

is singular if there is at least one phenotype (i.e., if $N_{\mathrm{p}} > 0$). Thus, $\mathbf{L}_{\zeta\mathbf{z}}$ has at least $N_{\mathrm{a}}N_{\mathrm{p}}$ eigenvalues that are exactly zero. Also, the mechanistic additive socio-genetic cross-covariance matrix between $\boldsymbol{\zeta}$ and the geno-envo-phenotype $\mathbf{m}$

$$\mathbf{L}_{\zeta\mathbf{m}} = \left(\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\frac{\mathrm{d}\mathbf{m}^{\intercal}}{\mathrm{d}\mathbf{y}}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{m\times(1+N_{\mathrm{a}})(N_{\mathrm{p}}+N_{\mathrm{g}}+N_{\mathrm{e}})} \qquad \text{(Layer 6, Eq. 12)}$$

is singular if there is at least one phenotype or one environmental trait (i.e., if $N_{\mathrm{p}} > 0$ or $N_{\mathrm{e}} > 0$). Thus, $\mathbf{L}_{\zeta\mathbf{m}}$ has at least $N_{\mathrm{a}}(N_{\mathrm{p}} + N_{\mathrm{e}})$ eigenvalues that are exactly zero. In important contrast, the mechanistic additive socio-genetic cross-covariance matrix between a vector $\boldsymbol{\zeta} \in \{\mathbf{y}, \mathbf{z}, \mathbf{m}\}$ and the genotype $\mathbf{y}$

$$\mathbf{L}_{\zeta\mathbf{y}} = \left(\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\mathbf{y}^{\intercal}}\mathbf{H}_{\mathbf{y}}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}} \in \mathbb{R}^{m\times N_{\mathrm{a}}N_{\mathrm{g}}} \qquad \text{(Layer 6, Eq. 13)}$$

is non-singular if $\mathbf{H}_{\mathbf{y}}$ is non-singular because the genotype is developmentally independent (SI sections S5.7 and S5.9). The $\mathbf{L}$-matrices share various properties with similar generalizations of the $\mathbf{G}$-matrix arising in the indirect genetic effects literature (Kirkpatrick and Lande, 1989; Moore et al., 1997; Townley and Ezard, 2013).

5.2. Layer 7: evolutionary dynamics

Now we describe the top layer of the evo-devo process, that of the evolutionary dynamics. This layer contains equations describing the evolutionary dynamics under explicit developmental and environmental constraints. In the SI sections S2.5 and S5.6-S5.9, we show that, in the limit as $\Delta\tau \to 0$, the evolutionary dynamics of the resident phenotype, genotype, geno-phenotype, environment, and geno-envo-phenotype (i.e., for $\boldsymbol{\zeta} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\epsilon}, \mathbf{m}\}$) are given by

$$\frac{\mathrm{d}\bar{\boldsymbol{\zeta}}}{\mathrm{d}\tau} \approx \left(\iota\mathbf{L}_{\zeta\mathbf{m}}\frac{\partial w}{\partial \mathbf{m}} + \frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\boldsymbol{\epsilon}^{\intercal}}\frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}}, \qquad \text{(Layer 7, Eq. 1a)}$$

which must satisfy both the developmental constraint

$$\bar{\mathbf{x}}_{a+1} = \mathbf{g}_a(\bar{\mathbf{m}}_a, \bar{\mathbf{z}}) \text{ for all } a \in \{1, \ldots, N_{\mathrm{a}} - 1\} \text{ with fixed } \bar{\mathbf{x}}_1, \qquad \text{(Layer 7, Eq. 1b)}$$

and the environmental constraint

$$\bar{\boldsymbol{\epsilon}}_a = \mathbf{h}_a(\bar{\mathbf{z}}_a, \bar{\mathbf{z}}, \tau) \text{ for all } a \in \{1, \ldots, N_{\mathrm{a}}\}. \qquad \text{(Layer 7, Eq. 1c)}$$

If $\boldsymbol{\zeta} = \mathbf{z}$ in Layer 7, Eq. 1a, then the equations in Layers 2-6 guarantee that the developmental constraint is satisfied for all $\tau > \tau_1$ given that it is satisfied at the initial evolutionary time τ_1. If $\boldsymbol{\zeta} = \mathbf{m}$ in Layer 7, Eq. 1a, then the equations in Layers 2-6 guarantee that both the developmental and environmental constraints are satisfied for all $\tau > \tau_1$ given that they are satisfied at the initial evolutionary time τ_1. Both the developmental and environmental constraints can evolve as the genotype, phenotype, and environment evolve and such constraints can involve any family of curves as long as they are continuously differentiable.

Importantly, although Layer 7, Eq. 1a describes the evolutionary dynamics of $\boldsymbol{\zeta}$, such equation is guaranteed to be dynamically sufficient only for certain $\boldsymbol{\zeta}$. Layer 7, Eq. 1a is dynamically sufficient if $\boldsymbol{\zeta}$ is the genotype $\mathbf{y}$, the geno-phenotype $\mathbf{z}$, or the geno-envo-phenotype $\mathbf{m}$, provided that the developmental and environmental constrains are satisfied throughout. In contrast, Layer 7, Eq. 1a is dynamically insufficient if $\boldsymbol{\zeta}$ is the phenotype $\mathbf{x}$ or the environment $\boldsymbol{\epsilon}$, because the evolution of the genotype is not followed but it generally affects the system.

Layer 7, Eq. 1a describes the evolutionary dynamics as consisting of selection response and exogenous plastic response. Layer 7, Eq. 1a contains the term

$$\iota \mathbf{L}_{\zeta\mathbf{m}} \left.\frac{\partial w}{\partial \mathbf{m}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}, \qquad \text{(Layer 7, Eq. 2)}$$

which comprises direct directional selection on the geno-envo-phenotype ($\partial w/\partial \mathbf{m}|_{\mathbf{y}=\bar{\mathbf{y}}}$), socio-genetic cross-covariation between $\boldsymbol{\zeta}$ and the geno-envo-phenotype ($\mathbf{L}_{\zeta\mathbf{m}}$), and mutational input (ι). The term in Layer 7, Eq. 2 is the *selection response* of $\bar{\boldsymbol{\zeta}}$ and is a mechanistic generalization of Lande's (1979) generalization of the univariate breeder's equation (Lush, 1937; Walsh and Lynch, 2018). Additionally, Layer 7, Eq. 1a contains the term

$$\left.\left(\frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\boldsymbol{\epsilon}^{\intercal}} \frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}}, \qquad \text{(Layer 7, Eq. 3)}$$

which comprises the vector of environmental change due to exogenous causes ($\partial \bar{\boldsymbol{\epsilon}}/\partial \tau$) and the matrix of stabilized plasticity ($\mathrm{s}\boldsymbol{\zeta}/\mathrm{s}\boldsymbol{\epsilon}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}}$). The term in Layer 7, Eq. 3 is the *exogenous plastic response* of $\bar{\boldsymbol{\zeta}}$ and is a mechanistic generalization of previous expressions (cf. Eq. A3 of Chevin et al. 2010). Note that the *endogenous* plastic response of $\bar{\boldsymbol{\zeta}}$ (i.e., the plastic response due to endogenous environmental change arising from niche construction) is part of both the selection response and the exogenous plastic response (Layers 2-6).

Selection response is relatively incompletely described by direct directional selection on the geno-envo-phenotype. Indeed, we saw that the matrix $\mathbf{L}_{\zeta\mathbf{m}}$ is always singular if there is at least one phenotype or one environmental trait (Layer 6, Eq. 12). Consequently, the selection response of $\bar{\boldsymbol{\zeta}}$ can invariably vanish with persistent direct directional selection on the geno-envo-phenotype.

Selection response is also relatively incompletely described by total immediate selection on the geno-phenotype. We can rewrite the selection response, so the evolutionary dynamics of $\bar{\boldsymbol{\zeta}}$ for $\boldsymbol{\zeta} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\epsilon}, \mathbf{m}\}$ (Layer 7, Eq. 1a) is equivalently given by

$$\frac{\mathrm{d}\bar{\boldsymbol{\zeta}}}{\mathrm{d}\tau} \approx \left.\left(\iota \mathbf{L}_{\zeta\mathbf{z}} \frac{\delta w}{\delta \mathbf{z}} + \frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\boldsymbol{\epsilon}^{\intercal}} \frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Layer 7, Eq. 4)}$$

This equation now depends on total immediate selection on the geno-phenotype ($\delta w/\delta \mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}}$), which measures total immediate directional selection on the geno-phenotype (or in a quantitative genetics framework, it is Lande's (1979) selection gradient of the allele frequency and phenotype if environmental traits are not explicitly included in the analysis). The total immediate selection gradient of the geno-phenotype can be interpreted as pointing in the direction of steepest ascent on the fitness landscape in geno-phenotype space after the landscape is modified by the interaction of direct niche construction and environmental sensitivity of selection (Layer 3, Eq. S1). We saw that the matrix $\mathbf{L}_{\zeta\mathbf{z}}$ is always singular if there is at least one phenotype (Layer 6, Eq. 11). Consequently, the selection response of $\bar{\boldsymbol{\zeta}}$ can invariably vanish with persistent directional selection on the geno-phenotype after niche construction has modified the geno-phenotype's fitness landscape.

In contrast, selection response is relatively completely described by total genotypic selection. We can further rewrite the selection response, so the evolutionary dynamics of $\bar{\boldsymbol{\zeta}}$ for $\boldsymbol{\zeta} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\epsilon}, \mathbf{m}\}$ (Layer 7, Eq. 1a) is equivalently given by

$$\frac{\mathrm{d}\bar{\boldsymbol{\zeta}}}{\mathrm{d}\tau} \approx \left(\iota \mathbf{L}_{\zeta\mathbf{y}} \frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}} + \frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\boldsymbol{\epsilon}^{\intercal}} \frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\bigg|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Layer 7, Eq. 5)}$$

This equation now depends on total genotypic selection ($\mathrm{d}w/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}}$), which measures total directional selection on the genotype considering downstream developmental effects (or in a quantitative genetics framework, it is Lande's (1979) selection gradient of allele frequency if neither the phenotype nor environmental traits are explicitly included in the analysis). In Eq. 9 we saw that the total selection gradient of the genotype can be interpreted as pointing in the direction of steepest ascent on the fitness landscape in genotype space after the landscape is modified by: (1) the interaction of total developmental bias from the genotype and directional selection on the phenotype and (2) the interaction of total niche construction by the genotype and environmental sensitivity of selection. In contrast to the other arrangements of selection response, in SI sections S5.7 and S5.9 we show that $\mathbf{L}_{\zeta\mathbf{y}}$ is non-singular for all $\boldsymbol{\zeta} \in \{\mathbf{y}, \mathbf{z}, \mathbf{m}\}$ if $\mathbf{H}_{\mathbf{y}}$ is non-singular (i.e., if there is mutational variation in all directions of genotype space); this non-singularity of $\mathbf{L}_{\zeta\mathbf{y}}$ arises because genotypic traits are developmentally independent by assumption. Consequently, the selection response of the genotype, geno-phenotype, or geno-envo-phenotype can only vanish when total genotypic selection vanishes if there is mutational variation in all directions of genotype space.

Let us revisit why following the evolutionary dynamics of the phenotype alone is dynamically insufficient in order to compare such an approach with a dynamically sufficient extension. As stated above, Layer 7, Eq. 1a and its equivalents are generally dynamically insufficient if only the evolutionary dynamics of the phenotype are considered (i.e., if $\boldsymbol{\zeta} = \mathbf{x}$). Let us temporarily assume that the following four conditions hold: (I) development is non-social ($\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\bar{\mathbf{z}}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), and there is (II) no exogenous plastic response of the phenotype ($[(\mathrm{d}\mathbf{x}/\mathrm{d}\boldsymbol{\epsilon}^{\intercal})(\partial\bar{\boldsymbol{\epsilon}}/\partial\tau)]\,|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), (III) no total immediate selection on the genotype ($\delta w/\delta \mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), and (IV) no niche-constructed effects of the phenotype on fitness ($[(\partial\boldsymbol{\epsilon}^{\intercal}/\partial\mathbf{x})(\partial w/\partial\boldsymbol{\epsilon})]\,|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$). Then, the evolutionary dynamics of the phenotype reduces to

$$\frac{\mathrm{d}\bar{\mathbf{x}}}{\mathrm{d}\tau} \approx \iota \mathbf{H}_{\mathbf{x}} \frac{\partial w}{\partial \mathbf{x}}\bigg|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Layer 7, Eq. 6)}$$

This is a mechanistic version of the Lande equation for the phenotype. The mechanistic additive genetic covariance matrix of the phenotype (Layer 6, Eq. 5) in this equation is singular because the developmentally initial phenotype is not affected by the genotype and the developmentally final genotypic traits do not affect the phenotype (so $\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}}$ has rows and columns that are zero; SI section S5.2, Eq. S5.2.16). This singularity might disappear by removing from the analysis the developmentally initial phenotype and developmentally final genotypic traits, provided additional conditions hold. Yet, the key point here is that a system describing the evolutionary dynamics of the phenotype alone is dynamically insufficient because such system depends on the resident genotype whose evolution must also be followed. In particular, setting $\mathrm{d}\bar{\mathbf{x}}/\mathrm{d}\tau = \mathbf{0}$ does not generally imply an evolutionary equilibrium, or evolutionary stasis, but only an evolutionary nullcline in the phenotype, that is, a transient lack of evolutionary change in the phenotype. To guarantee a dynamically sufficient description of the evolutionary dynamics of the phenotype, we simultaneously consider the evolutionary dynamics of the phenotype and genotype, that is, the geno-phenotype.

A dynamically sufficient system can be obtained by describing the dynamics of the geno-phenotype alone if the environment is constant or has no evolutionary effect. Let us now assume that the following three conditions hold: (i) development is non-social ($\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\bar{\mathbf{z}}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), and there is (ii) no exogenous plastic response of the phenotype ($[(\mathrm{d}\mathbf{x}/\mathrm{d}\boldsymbol{\epsilon}^{\intercal})(\partial\bar{\boldsymbol{\epsilon}}/\partial\tau)]\,|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$), and (iii) no niche-constructed effects of the geno-phenotype on fitness ($[(\partial\boldsymbol{\epsilon}^{\intercal}/\partial\mathbf{z})(\partial w/\partial\boldsymbol{\epsilon})]\,|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$). Then, the evolutionary dynamics of the geno-phenotype reduces to

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} \approx \iota\mathbf{H}_{\mathbf{z}}\left.\frac{\partial w}{\partial \mathbf{z}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Layer 7, Eq. 7)}$$

This is an extension of the mechanistic version of the Lande equation to consider the geno-phenotype. The mechanistic additive genetic covariance matrix of the geno-phenotype (Layer 6, Eq. 6) in this equation is singular because the geno-phenotype $\mathbf{z}$ includes the genotype $\mathbf{y}$ (so $\mathrm{d}\mathbf{z}^{\intercal}/\mathrm{d}\mathbf{y}$ has fewer rows than columns; Layer 4, Eq. S7). Hence, the degrees of freedom of genetic covariation in geno-phenotype space are at most given by the number of lifetime genotypic traits, so these degrees of freedom are bounded by genotypic space in a necessarily larger geno-phenotype space. Thus, $\mathbf{H}_{\mathbf{z}}$ is singular if there is at least one trait that is developmentally constructed according to the developmental map (Layer 7, Eq. 1b). The evolutionary dynamics of the geno-phenotype is now fully determined by Layer 7, Eq. 7 provided that i-iii hold and that the developmental (Layer 7, Eq. 1b) and environmental (Layer 7, Eq. 1c) constraints are met, which $\mathbf{H}_{\mathbf{z}}$ guarantees they are if they are met at the initial evolutionary time τ_1. In such case, setting $\mathrm{d}\bar{\mathbf{z}}/\mathrm{d}\tau = \mathbf{0}$ does imply an evolutionary equilibrium to first order of approximation, but this does not imply absence of direct directional selection on the geno-phenotype (i.e., it is possible that $\partial w/\partial\mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}} \neq \mathbf{0}$) since $\mathbf{H}_{\mathbf{z}}$ is always singular. Due to this singularity, if there is any evolutionary equilibrium, there is an infinite number of them. Kirkpatrick and Lofsvold (1992) showed that if $\mathbf{G}$ is singular and constant, then the evolutionary equilibrium that is achieved depends on the initial conditions. Our results extend the relevance of Kirkpatrick and Lofsvold's (1992) observation by showing that $\mathbf{H}_{\mathbf{z}}$ is always singular and remains so as it evolves. Moreover, since both the developmental (Layer 7, Eq. 1b) and environmental (Layer 7, Eq. 1c) constraints must be satisfied throughout the evolutionary process, the developmental and environmental constraints determine the admissible evolutionary trajectory and the admissible evolutionary equi-

libria if mutational variation exists in all directions of genotype space. Therefore, developmental and environmental constraints together with direct directional selection jointly define the evolutionary outcome if mutational variation exists in all directions of genotype space.

Since selection response is relatively completely described by total genotypic selection, further insight can be gained by rearranging the extended mechanistic Lande equation for the geno-phenotype (Layer 7, Eq. 7) in terms of total genotypic selection. Using the rearrangement in Layer 7, Eq. 5 and making the assumptions i-iii in the previous paragraph, the extended mechanistic Lande equation in Layer 7, Eq. 7 becomes

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} \approx \iota \mathbf{H}_{\mathbf{zy}} \left.\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Layer 7, Eq. 8)}$$

If the mutational covariance matrix $\mathbf{H}_{\mathbf{y}}$ is non-singular, then the mechanistic additive genetic cross-covariance matrix between geno-phenotype and genotype $\mathbf{H}_{\mathbf{zy}}$ is non-singular so evolutionary equilibrium ($\mathrm{d}\bar{\mathbf{z}}/\mathrm{d}\tau = \mathbf{0}$) implies absence of total genotypic selection (i.e., $\mathrm{d}w/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$) to first order of approximation. Moreover, dropping assumptions i-iii and to first order, lack of total genotypic selection provides a necessary and sufficient condition for evolutionary equilibria in the absence of exogenous environmental change and of absolute mutational constraints (Layer 7, Eq. 5). Consequently, evolutionary equilibria depend on development and niche construction since total genotypic selection depends on the developmental matrix and on total niche construction by the genotype (Eq. 9). However, since $\mathrm{d}w/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$ has only as many equations as there are lifetime genotypic traits and since not only the genotype but also the phenotype and environmental traits must be determined, then $\mathrm{d}w/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{0}$ provides fewer equations than variables to solve for. Hence, absence of total genotypic selection still implies an infinite number of evolutionary equilibria. Again, only the subset of evolutionary equilibria that satisfy the developmental (Layer 7, Eq. 1b) and environmental (Layer 7, Eq. 1c) constraints are admissible, and so the number of admissible evolutionary equilibria may be finite. Therefore, admissible evolutionary equilibria have a dual dependence on developmental and environmental constraints: first, by the constraints' influence on total genotypic selection and so on evolutionary equilibria; and second, by the constraints' specification of which evolutionary equilibria are admissible.

Because we assume that mutants arise when residents are at carrying capacity, the analogous statements can be made for the evolutionary dynamics of a resident vector in terms of lifetime reproductive success (Eq. 8). Using the relationship between selection gradients in terms of fitness and of expected lifetime reproductive success (Eq. S2.3.12), the evolutionary dynamics of $\bar{\boldsymbol{\zeta}}$ for $\boldsymbol{\zeta} \in \{\mathbf{x}, \mathbf{y}, \mathbf{z}, \boldsymbol{\epsilon}, \mathbf{m}\}$ (Layer 7, Eq. 1a) are equivalently given by

$$\frac{\mathrm{d}\bar{\boldsymbol{\zeta}}}{\mathrm{d}\tau} \approx \left.\left(\iota \frac{1}{T} \mathbf{L}_{\zeta\mathbf{m}} \frac{\partial R_0}{\partial \mathbf{m}} + \frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}} \frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \qquad \text{(Layer 7, Eq. 9a)}$$

$$= \left.\left(\iota \frac{1}{T} \mathbf{L}_{\zeta\mathbf{z}} \frac{\delta R_0}{\delta \mathbf{z}} + \frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}} \frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \qquad \text{(Layer 7, Eq. 9b)}$$

$$= \left.\left(\iota \frac{1}{T} \mathbf{L}_{\zeta\mathbf{y}} \frac{\mathrm{d}R_0}{\mathrm{d}\mathbf{y}} + \frac{\mathrm{s}\boldsymbol{\zeta}}{\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}} \frac{\partial \boldsymbol{\epsilon}}{\partial \tau}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Layer 7, Eq. 9c)}$$

To close, the evolutionary dynamics of the environment can be written in a particular form that is insightful. In SI

section S5.8, we show that the evolutionary dynamics of the environment satisfy

$$\frac{\mathrm{d}\bar{\boldsymbol{\epsilon}}}{\mathrm{d}\tau} = \left(\frac{\partial(\boldsymbol{\epsilon} + \check{\boldsymbol{\epsilon}})}{\partial \mathbf{z}^{\intercal}} \frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} + \frac{\partial \boldsymbol{\epsilon}}{\partial \tau} \right)\Bigg|_{\mathbf{y}=\bar{\mathbf{y}}} . \qquad \text{(Layer 7, Eq. 10)}$$

Thus, the evolutionary change of the environment comprises "joint" direct niche construction and exogenous environmental change.

6. Example: allocation to growth vs reproduction

We now provide an example that illustrates the method and some of the points above. To do this, we use a life-history model rather than a model of morphological development as the former is simpler yet sufficient to illustrate the points. In particular, this example shows that our results above enable direct calculation of the evo-devo dynamics and the evolution of the constraining matrices $\mathbf{H}$ and $\mathbf{L}$ and provide an alternative method to dynamic optimization to identify the evolutionary outcomes under explicit developmental constraints. We first describe the example where development is non-social and then extend the example to make development social.

6.1. Non-social development

We consider the classic life-history problem of modeling the evolution of resource allocation to growth vs reproduction (Gadgil and Bossert, 1970; León, 1976; Schaffer, 1983; Stearns, 1992; Roff, 1992; Kozłowski and Teriokhin, 1999). Let there be one phenotype (or state variable), one genotypic trait (or control variable), and no environmental traits. In particular, let x_a be a mutant's phenotype at age a (e.g., body size or resources available) and $y_a \in [0, 1]$ be the mutant's fraction of resource allocated to phenotype growth at that age. Let the mutant survival probability $p_a = p$ be constant for all $a \in \{1, \ldots, N_\mathrm{a} - 1\}$ with $p_{N_\mathrm{a}} = 0$, so survivorship is $\ell_a = p^{a-1}$ for all $a \in \{1, \ldots, N_\mathrm{a}\}$ with $\ell_{N_\mathrm{a}+1} = 0$. Let the mutant fertility be

$$f_a = d(\bar{\mathbf{z}})(1 - y_a)x_a,$$

where $(1 - y_a)x_a$ is the resource a mutant allocates to reproduction at age a and $d(\bar{\mathbf{z}})$ is a positive density-dependent scalar that brings the resident population size to carrying capacity. Let the developmental constraint be

$$x_{a+1} = g_a(\mathbf{z}_a, \bar{\mathbf{z}}) = x_a + y_a x_a = (1 + y_a)x_a, \qquad \text{(Example, Eq. 1)}$$

with initial condition $x_1 = \bar{x}_1$, where $y_a x_a$ is the resource a mutant allocates to growth at age a. These equations are a simplification of those used in the classic life-history problem of finding the optimal resource allocation to growth vs reproduction in discrete age (Gadgil and Bossert, 1970; León, 1976; Schaffer, 1983; Stearns, 1992; Roff, 1992; Kozłowski and Teriokhin, 1999). In life-history theory, one assumes that at evolutionary equilibrium, a measure of fitness such as lifetime reproductive success is maximized by an optimal control $\mathbf{y}^*$ yielding an optimal pair $(\mathbf{x}^*, \mathbf{y}^*)$ that is obtained with dynamic programming or optimal control theory (Sydsæter et al., 2008). Instead, here we

illustrate how the evolutionary dynamics of $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ can be analyzed with the equations derived in this paper, including identification of an optimal pair $(\mathbf{x}^*, \mathbf{y}^*)$.

We begin by obtaining analytical expressions for the phenotype and the density dependent scalar. Although an analytical expression for the phenotype is possible for this simple model, numerical solution is sufficient for more complex models. Iterating the recurrence given by the developmental constraint (Example, Eq. 1) yields the mutant phenotype at age a

$$x_a = x_1 \prod_{k=1}^{a-1} (1 + y_k). \qquad \text{(Example, Eq. 2)}$$

To find the density-dependent scalar, we note that a resident at carrying capacity satisfies the Euler-Lotka equation $\sum_{a=1}^{N_\mathrm{a}} f_a^\circ \ell_a = 1$ (Eq. S2.4.9), which yields

$$d(\bar{\mathbf{z}}) = \frac{1}{\sum_{a=1}^{N_\mathrm{a}} (1 - \bar{y}_a) \bar{x}_a \ell_a}.$$

We now calculate the elements of the bottom layers that we need to calculate genetic covariation and the evolutionary dynamics. Because there are no environmental traits, total immediate effects equal direct effects (i.e., layer 3 is immaterial). Also, because development is non-social, stabilized effects equal total effects (except for social feedback, which is simply the identity matrix; i.e., layer 5 is immaterial). So, we only need to compute layers 2 and 4, that is, direct and total effects. Using Eq. (5a), the entries of the direct selection gradients are given by

$$\begin{aligned}
\left.\frac{\partial w}{\partial x_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &= \sum_{j=1}^{N_\mathrm{a}} \left.\frac{\partial w_j}{\partial x_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\frac{\partial w_a}{\partial x_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\
&= \frac{1}{T} \left.\left(\phi_a \frac{\partial f_a}{\partial x_a} + \pi_a \frac{\partial p_a}{\partial x_a} \right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \frac{1}{\tilde{T}} \ell_a (1 - \bar{y}_a), \\
\left.\frac{\partial w}{\partial y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &= \sum_{j=1}^{N_\mathrm{a}} \left.\frac{\partial w_j}{\partial y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\frac{\partial w_a}{\partial y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\
&= \frac{1}{T} \left.\left(\phi_a \frac{\partial f_a}{\partial y_a} + \pi_a \frac{\partial p_a}{\partial y_a} \right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} = -\frac{1}{\tilde{T}} \ell_a \bar{x}_a. \qquad \text{(Example, Eq. 3)}
\end{aligned}$$

where the generation time without density dependence is

$$\tilde{T} = \sum_{j=1}^{N_\mathrm{a}} j \ell_j (1 - \bar{y}_j) \bar{x}_j.$$

Thus, there is always direct selection for increased phenotype and against allocation to growth (except at the boundaries where $\bar{y}_a = 1$ or $\bar{x}_a = 0$). The entries of the matrices of direct effects on the phenotype (a: row, j: column) are given by

$$\left.\frac{\partial x_j}{\partial x_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \begin{cases} 1 + \bar{y}_a & \text{if } j = a + 1 \\ 1 & \text{if } j = a \\ 0 & \text{otherwise,} \end{cases}$$

$$\left.\frac{\partial x_j}{\partial y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \begin{cases} \bar{x}_a & \text{if } j = a+1 \\ 0 & \text{otherwise.} \end{cases}$$

Using Eq. (12) and Eq. S5.2.15, the entries of the matrices of total effects on the phenotype are given by

$$\begin{aligned}
\left.\frac{\mathrm{d}x_j}{\mathrm{d}x_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &= \begin{cases} \prod_{k=a}^{j-1} \left.\frac{\partial x_{k+1}}{\partial x_k}\right|_{\mathbf{y}=\bar{\mathbf{y}}} & \text{if } j > a \\ 1 & \text{if } j = a \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \prod_{k=a}^{j-1} (1+\bar{y}_k) & \text{if } j > a \\ 1 & \text{if } j = a \\ 0 & \text{otherwise,} \end{cases} \\
\left.\frac{\mathrm{d}x_j}{\mathrm{d}y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &= \begin{cases} \left.\left(\frac{\partial x_{a+1}}{\partial y_a} \prod_{k=a+1}^{j-1} \frac{\partial x_{k+1}}{\partial x_k}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} & \text{if } j > a+1 \\ \left.\frac{\partial x_{a+1}}{\partial y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} & \text{if } j = a+1 \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \bar{x}_a \prod_{k=a+1}^{j-1} (1+\bar{y}_k) & \text{if } j > a+1 \\ \bar{x}_a & \text{if } j = a+1 \\ 0 & \text{otherwise.} \end{cases} \qquad \text{(Example, Eq. 4)}
\end{aligned}$$

Then, using Layer 4, Eq. S20 and Eq. (9), the entries of the total selection gradients are given by

$$\begin{aligned}
\left.\frac{\mathrm{d}w}{\mathrm{d}x_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &= \left.\left(\frac{\mathrm{d}\mathbf{x}^\intercal}{\mathrm{d}x_a}\frac{\partial w}{\partial \mathbf{x}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \sum_{j=1}^{N_\mathrm{a}} \left.\left(\frac{\mathrm{d}x_j}{\mathrm{d}x_a}\frac{\partial w}{\partial x_j}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\
&= \left.\left(\frac{\mathrm{d}x_a}{\mathrm{d}x_a}\frac{\partial w}{\partial x_a} + \sum_{j=a+1}^{N_\mathrm{a}} \frac{\mathrm{d}x_j}{\mathrm{d}x_a}\frac{\partial w}{\partial x_j}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\
&= \frac{1}{\tilde{T}}\left[\ell_a(1-\bar{y}_a) + \sum_{j=a+1}^{N_\mathrm{a}} \ell_j(1-\bar{y}_j)\prod_{k=a}^{j-1}(1+\bar{y}_k)\right], \\
\left.\frac{\mathrm{d}w}{\mathrm{d}y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &= \left.\left(\frac{\partial w}{\partial y_a} + \frac{\mathrm{d}\mathbf{x}^\intercal}{\mathrm{d}y_a}\frac{\partial w}{\partial \mathbf{x}}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \left.\left(\frac{\partial w}{\partial y_a} + \sum_{j=1}^{N_\mathrm{a}} \frac{\mathrm{d}x_j}{\mathrm{d}y_a}\frac{\partial w}{\partial x_j}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\
&= \left.\left(\frac{\partial w}{\partial y_a} + \frac{\mathrm{d}x_{a+1}}{\mathrm{d}y_a}\frac{\partial w}{\partial x_{a+1}} + \sum_{j=a+2}^{N_\mathrm{a}} \frac{\mathrm{d}x_j}{\mathrm{d}y_a}\frac{\partial w}{\partial x_j}\right)\right|_{\mathbf{y}=\bar{\mathbf{y}}} \\
&= -\frac{1}{\tilde{T}}\ell_a\bar{x}_a + \bar{x}_a\frac{1}{\tilde{T}}\ell_{a+1}(1-\bar{y}_{a+1})
\end{aligned}$$

$$+ \sum_{j=a+2}^{N_\mathrm{a}} \bar{x}_a \prod_{k=a+1}^{j-1} (1+\bar{y}_k) \frac{1}{\tilde{T}} \ell_j (1-\bar{y}_j)$$

$$= \frac{1}{\tilde{T}} \bar{x}_a \left[-\ell_a + \sum_{j=a+1}^{N_\mathrm{a}} \ell_j (1-\bar{y}_j) \prod_{k=a+1}^{j-1} (1+\bar{y}_k) \right], \qquad \text{(Example, Eq. 5)}$$

where we use the empty-product notation such that $\prod_{k=a}^{a-1} F_k = 1$ and the empty-sum notation such that $\sum_{k=a}^{a-1} F_k = 0$ for any F_k. There is thus always total selection for increased phenotype (except at the boundaries), although total selection for allocation to growth may be positive or negative.

Now, using Eqs. (1) and (3), the evo-devo dynamics are given by

$$\begin{aligned} \frac{\Delta \bar{\mathbf{y}}}{\Delta \tau} &\approx \iota \mathbf{H}_\mathbf{y} \left. \frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}} \right|_{\mathbf{y}=\bar{\mathbf{y}}} \\ \bar{x}_{a+1} &= g_a(\bar{\mathbf{z}}_a, \bar{\mathbf{z}}). \end{aligned} \qquad \text{(Example, Eq. 6)}$$

Using Layer 7, Eq. 1a, Layer 7, Eq. 4, and Layer 7, Eq. 5, the evolutionary dynamics of the resident phenotype in the limit as $\Delta\tau \to 0$ are given by

$$\frac{\mathrm{d}\bar{\mathbf{x}}}{\mathrm{d}\tau} \approx \iota \mathbf{H}_\mathbf{xz} \left. \frac{\partial w}{\partial \mathbf{z}} \right|_{\mathbf{y}=\bar{\mathbf{y}}} = \iota \mathbf{H}_\mathbf{xy} \left. \frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}} \right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Example, Eq. 7)}$$

Note these are not equations in Lande's form. In particular, the mechanistic additive genetic-cross covariance matrices involved are not symmetric and the selection gradients are not those of the evolving trait in the left-hand side; Example, Eq. 7 cannot be arranged in Lande's form because the genotypic trait directly affects fitness (i.e., $\partial w/\partial \mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}} \neq \mathbf{0}$; Example, Eq. 3). Importantly, $\mathbf{H}_\mathbf{xz}$ and $\mathbf{H}_\mathbf{xy}$ depend on $\bar{\mathbf{y}}$ because of gene-phenotype interaction in development (i.e., the developmental map involves a product $y_a x_a$ such that the total effect of the genotype on the phenotype depends on the genotype; Example, Eq. 4); consequently, Example, Eq. 7 is dynamically insufficient because the system does not describe the evolution of $\bar{\mathbf{y}}$. In turn, the evolutionary dynamics of the geno-phenotype are given by

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} \approx \iota \mathbf{H}_\mathbf{z} \left. \frac{\partial w}{\partial \mathbf{z}} \right|_{\mathbf{y}=\bar{\mathbf{y}}} = \iota \mathbf{H}_\mathbf{zy} \left. \frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}} \right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Example, Eq. 8)}$$

This system contains dynamic equations for all the evolutionarily dynamic variables, namely both the resident phenotype $\bar{\mathbf{x}}$ and the resident genotype $\bar{\mathbf{y}}$, so it is determined and dynamically sufficient. The first equality in Example, Eq. 8 is in Lande's form, but $\mathbf{H}_\mathbf{z}$ is always singular. In contrast, the matrix $\mathbf{H}_\mathbf{zy}$ in the second equality is non-singular if the mutational covariance matrix $\mathbf{H}_\mathbf{y}$ is non-singular. Thus, the total selection gradient of the genotype provides a relatively complete description of the evolutionary process of the geno-phenotype.

Let the entries of the mutational covariance matrix be given by

$$H_{y_a,y_j} = \begin{cases} \gamma \bar{y}_a (1-\bar{y}_a) & \text{if } j = a \\ 0 & \text{otherwise,} \end{cases}$$

where $0 < \gamma \ll 1$ so the assumption of marginally small mutational variance, namely $0 < \mathrm{tr}(\mathbf{H}_\mathbf{y}) \ll 1$, holds. Thus, $\mathbf{H}_\mathbf{y}$ is diagonal and becomes singular only at the boundaries where the resident genotype is zero or one. Then, from

Example, Eq. 6, the evolutionary equilibria of the genotypic trait at a given age and their stability are given by the sign of its corresponding total selection gradient.

Let us now find the evolutionary equilibria and their stability for the genotypic trait. Using Example, Eq. 5, starting from the last age, the total selection on the genotypic trait at this age is

$$\left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_\mathrm{a}}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} \propto -\ell_{N_\mathrm{a}},$$

which is always negative so the stable resident genotypic trait at the last age is

$$\bar{y}^*_{N_\mathrm{a}} = 0. \qquad \text{(Example, Eq. 9a)}$$

That is, no allocation to growth at the last age. Continuing with the second-to-last age, the total selection on the genotypic trait at this age is

$$\begin{aligned}\left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_\mathrm{a}-1}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &\propto -\ell_{N_\mathrm{a}-1} + \sum_{j=N_\mathrm{a}}^{N_\mathrm{a}} \ell_j(1-\bar{y}_j)\prod_{k=N_\mathrm{a}}^{j-1}(1+\bar{y}_k)\\ &= -\ell_{N_\mathrm{a}-1} + \ell_{N_\mathrm{a}}(1-\bar{y}_{N_\mathrm{a}}).\end{aligned}$$

Evaluating at the optimal genotypic trait at the last age (Example, Eq. 9a) and substituting $\ell_a = p^{a-1}$ yields

$$\left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_\mathrm{a}-1}}\right|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} \propto -p^{N_\mathrm{a}-2} + p^{N_\mathrm{a}-1} \propto -1+p,$$

which is negative (assuming $p < 1$) so the stable resident genotypic trait at the second-to-last age is

$$\bar{y}^*_{N_\mathrm{a}-1} = 0. \qquad \text{(Example, Eq. 9b)}$$

Continuing with the third-to-last age, the total selection on the genotypic trait at this age is

$$\begin{aligned}\left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_\mathrm{a}-2}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &\propto -\ell_{N_\mathrm{a}-2} + \sum_{j=N_\mathrm{a}-1}^{N_\mathrm{a}} \ell_j(1-\bar{y}_j)\prod_{k=N_\mathrm{a}-1}^{j-1}(1+\bar{y}_k)\\ &= -\ell_{N_\mathrm{a}-2} + \ell_{N_\mathrm{a}-1}(1-\bar{y}_{N_\mathrm{a}-1})\\ &\quad + \ell_{N_\mathrm{a}}(1-\bar{y}_{N_\mathrm{a}})(1+\bar{y}_{N_\mathrm{a}-1}).\end{aligned}$$

Evaluating at the optimal genotypic trait at the last two ages (Example, Eq. 9a and Example, Eq. 9b) and substituting $\ell_a = p^{a-1}$ yields

$$\left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_\mathrm{a}-2}}\right|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} \propto -p^{N_\mathrm{a}-3} + p^{N_\mathrm{a}-2} + p^{N_\mathrm{a}-1} \propto -1+p+p^2,$$

which is positive if

$$p > p^*_{N_\mathrm{a}-2} = \frac{1}{2}(-1+\sqrt{5}) \approx 0.62,$$

(thus, $p^*_{N_a-2}$ is the reciprocal of the golden ratio). So the stable resident genotypic trait at the third-to-last age is

$$\bar{y}^*_{N_a-2} = \begin{cases} 0 & \text{if } p < p^*_{N_a-2} = \frac{1}{2}(-1+\sqrt{5}) \approx 0.62 \\ 1 & \text{if } p > p^*_{N_a-2} = \frac{1}{2}(-1+\sqrt{5}) \approx 0.62. \end{cases} \tag{Example, Eq. 9c}$$

If $p = p^*_{N_a-2}$, the genotypic trait at such age is selectively neutral, but we ignore this case as without an evolutionary model for p it is biologically unlikely that survival is and remains at such precise value. Hence, there is no allocation to growth at this age for low survival and full allocation for high survival. Continuing with the fourth-to-last age, the total selection on the genotypic trait at this age is

$$\begin{aligned} \left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_a-3}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} &\propto -\ell_{N_a-3} + \sum_{j=N_a-2}^{N_a} \ell_j(1-\bar{y}_j) \prod_{k=N_a-2}^{j-1}(1+\bar{y}_k) \\ &= -\ell_{N_a-3} + \ell_{N_a-2}(1-\bar{y}_{N_a-2}) \\ &\quad + \ell_{N_a-1}(1-\bar{y}_{N_a-1})(1+\bar{y}_{N_a-2}) \\ &\quad + \ell_{N_a}(1-\bar{y}_{N_a})(1+\bar{y}_{N_a-2})(1+\bar{y}_{N_a-1}). \end{aligned}$$

Evaluating at the optimal genotypic trait at the last three ages (Example, Eq. 9a-Example, Eq. 9c) and substituting $\ell_a = p^{a-1}$ yields

$$\begin{aligned} \left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_a-3}}\right|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} &\propto -p^{N_a-4} + p^{N_a-3}(1-y^*_{N_a-2}) + p^{N_a-2}(1+y^*_{N_a-2}) \\ &\quad + p^{N_a-1}(1+y^*_{N_a-2}) \\ &\propto -1 + p(1-y^*_{N_a-2}) + p^2(1+p)(1+y^*_{N_a-2}). \end{aligned}$$

If $p < p^*_{N_a-2}$, this is

$$\left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_a-3}}\right|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} \propto -1 + p + p^2(1+p),$$

which is positive if

$$p > p^*_{N_a-3} \approx 0.54.$$

If $p > p^*_{N_a-2}$, the gradient is

$$\left.\frac{\mathrm{d}w}{\mathrm{d}y_{N_a-3}}\right|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} \propto -1 + 2p^2(1+p),$$

which is positive if

$$p > \tilde{p}^*_{N_a-3} \approx 0.565.$$

Hence, the stable resident genotypic trait at the fourth-to-last age is

$$\bar{y}^*_{N_\mathrm{a}-3} = \begin{cases} 0 & \text{if } p < p^*_{N_\mathrm{a}-3} \approx 0.54 \\ 1 & \text{if } p > p^*_{N_\mathrm{a}-3} \approx 0.54, \end{cases} \qquad \text{(Example, Eq. 9d)}$$

for $p \neq p^*_{N_\mathrm{a}-2} \approx 0.62$. Again, this is no allocation to growth for low survival, although at this earlier age survival can be smaller for allocation to growth to evolve. Numerical solution for the evo-devo dynamics using Example, Eq. 6 is given in Fig. 5. The associated evolution of the $\mathbf{H_z}$ matrix, plotting Layer 6, Eq. 6, is given in Fig. 6. The code used to generate these figures is in the SI.

6.2. *Social development*

Consider a slight modification of the previous example, so that development is social. Let the mutant fertility be

$$f_a = d(\bar{\mathbf{z}})(1 - y_a)(x_a + q\bar{x}_{a+1}),$$

where the available resource is now given by $x_a + q\bar{x}_{a+1}$ for some constant q (positive, negative, or zero). Here the source of social development can be variously interpreted, including that an immediately older resident contributes to (positive q) or takes from (negative q) the resource of the focal individual, or that the focal individual learns from the older resident (positive or negative q depending on whether learning increases or increases the phenotype). Let the developmental constraint be

$$x_{a+1} = g_a(\mathbf{z}_a, \bar{\mathbf{z}}) = x_a + y_a(x_a + q\bar{x}_{a+1}),$$

with initial condition $x_1 = \bar{x}_1$.

To see what the stabilization of social development is, imagine the first individual in the population having a resident genotype developing with such developmental constraint. As it is the first individual in the population, such individual has no social partners so the developed phenotype is

$$\tilde{x}_{a+1} = \tilde{x}_a + \bar{y}_a(\tilde{x}_a + q \times 0),$$

with $\tilde{x}_1 = \bar{x}_1$. Imagine a next resident individual who develops in the context of such initial individual. This next individual develops the phenotype

$$\tilde{\tilde{x}}_{a+1} = \tilde{\tilde{x}}_a + \bar{y}_a(\tilde{\tilde{x}}_a + q\tilde{x}_{a+1}),$$

with $\tilde{\tilde{x}}_1 = \bar{x}_1$, but then $\tilde{\tilde{x}}_{a+1} \neq \tilde{x}_{a+1}$. Iterating this process, the resident phenotype $\bar{\mathbf{x}}$ may converge to a socio-devo equilibrium $\bar{\mathbf{x}}^{**}$, which satisfies

$$\bar{x}^{**}_{a+1} = \bar{x}^{**}_a + \bar{y}_a(\bar{x}^{**}_a + q\bar{x}^{**}_{a+1}).$$

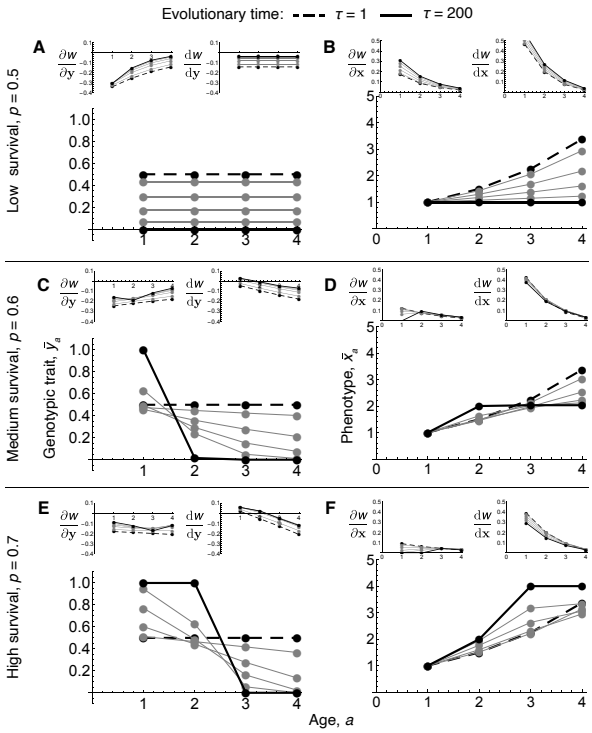

Figure 5: Example. Numerical solution of the evolutionary dynamics of the genotype and associated developmental dynamics of the phenotype. Large plots give the resident genotype (left) or phenotype (right) vs age over evolutionary time for various survival probabilities p. For low survival, (A) the genotypic trait evolves from mid allocation to growth over life to no allocation to growth over life. (B) This leads the phenotype to evolve from indeterminate growth to no growth over life. For medium survival, (C) the genotypic trait evolves from mid allocation to growth over life to full allocation to growth at the first age and no allocation to growth at later ages. (D) This leads the phenotype to evolve from indeterminate growth to determinate growth for one age. For high survival, (E) the genotypic trait evolves from mid allocation to growth over life to full allocation to growth at the first two ages and no allocation to growth at later ages. (F) This leads the phenotype to evolve from indeterminate growth to determinate growth for two ages, thus reaching the largest size of all these scenarios. Small plots give the associated direct and total selection gradients; the costate variable of x is proportional to the total selection gradient of the phenotype at large τ. In all scenarios, there is always negative direct selection for allocation to growth (small left plots in A,C,E) and non-negative direct selection for the phenotype (small left plots in B,D,F); hence, the genotype does not reach a fitness peak in geno-phenotype space, but for medium and high survival the phenotype at early ages reaches a fitness peak. Total selection for growth (small right plots in A,C,E) is positive only for early ages and sufficiently high survival; it remains non-zero at evolutionary equilibrium because $\bar{y}_a$ cannot evolve further as mutational variation vanishes as $\bar{y}_a$ approaches 0 or 1. Total selection for the phenotype (small right plots in B,D,F) is always positive, even at evolutionary equilibrium; the phenotype cannot evolve further despite persistent total selection for it because genetic variation vanishes as mutational variation vanishes. The numerical evolutionary outcomes match the analytical expressions for the genotype (Example, Eq. 9) and associated phenotype (Example, Eq. 2). $x_1 = \bar{x}_1 = 1$. From Eq. S2.3.5a, the carrying capacity is $\bar{n}^* = \bar{n}_1^* \sum_{a=1}^{4} p^{a-1}$. We let $\bar{n}_1^* = 2/(\mu\gamma)$, so $\iota = \gamma^{-1} \sum_{a=1}^{4} p^{a-1}$.

Solving for $\bar{x}_{a+1}^{**}$ yields a recurrence for the resident phenotype at socio-devo equilibrium

$$\bar{x}_{a+1}^{**} = \frac{1 + \bar{y}_a}{1 - q\bar{y}_a} \bar{x}_a^{**} \qquad \text{(Example, Eq. 10)}$$

provided that $1 - q\bar{y}_a \neq 0$.

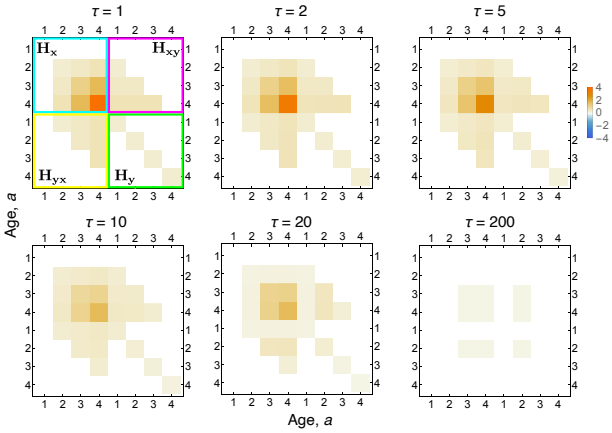

Figure 6: Resulting evolutionary dynamics of the mechanistic additive genetic covariance matrix $\mathbf{H_z}$. The upper-left quadrant (blue) is the mechanistic additive genetic covariance matrix $\mathbf{H_x}$ of the phenotype, that is, of the state variable. For instance, at the initial evolutionary time, the genetic variance for the phenotype is higher at later ages, and the phenotype at age 3 is highly genetically correlated with the phenotype at age 4. As evolutionary time progresses, genetic covariation vanishes as mutational covariation vanishes ($\mathbf{H_y}$ becomes singular) as genotypic traits approach their boundary values. $p = 0.7$. The evolutionary times τ shown correspond to those of Fig. 5.

Iterating Example, Eq. 10 yields the resident phenotype at socio-devo equilibrium

$$\bar{x}_a = \bar{x}_1 \prod_{k=1}^{a-1} \frac{1+\bar{y}_k}{1-q\bar{y}_k}, \qquad \text{(Example, Eq. 11)}$$

where we drop the ** for simplicity. To determine when this socio-devo equilibrium is socio-devo stable, we find the eigenvalues of $\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{x}}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}}$ (Eq. 28) as follows. The entries of the matrix of the direct social effects on the phenotype are given by

$$\left.\frac{\partial x_j}{\partial \bar{x}_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \begin{cases} \bar{y}_{a-1}q & \text{if } j = a \\ 0 & \text{otherwise.} \end{cases}$$

Hence, from Eq. S5.6.8 and Eq. S5.6.9, $\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{x}}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}}$ is lower-triangular, so its eigenvalues are the values in its main diagonal, which are given by $\partial x_a/\partial \bar{x}_a|_{\mathbf{y}=\bar{\mathbf{y}}} = \bar{y}_{a-1}q$. Thus, the eigenvalues of $\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{x}}^{\intercal}|_{\mathbf{y}=\bar{\mathbf{y}}}$ have absolute value strictly less than one if $|q| < 1$, in which case the socio-devo equilibrium in Example, Eq. 11 is socio-devo stable.

Therefore, let $\bar{\mathbf{x}}$ be the SDS resident phenotype given by Example, Eq. 11 with $|q| < 1$. Then, the evo-devo dynamics are given by Example, Eq. 6. Using Layer 7, Eq. 1a, Layer 7, Eq. 4, and Layer 7, Eq. 5, the evolutionary dynamics of the phenotype in the limit as $\Delta\tau \to 0$ are now given by

$$\frac{\mathrm{d}\bar{\mathbf{x}}}{\mathrm{d}\tau} \approx \iota \mathbf{L_{xz}} \left.\frac{\partial w}{\partial \mathbf{z}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \iota \mathbf{L_{xy}} \left.\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Example, Eq. 12)}$$

This system is dynamically insufficient as $\mathbf{L_{xz}}$ and $\mathbf{L_{xy}}$ depend on $\bar{\mathbf{y}}$ because of gene-phenotype interaction in development. In turn, the evolutionary dynamics of the geno-phenotype are given by

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} \approx \iota \mathbf{L_z} \left.\frac{\partial w}{\partial \mathbf{z}}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \iota \mathbf{L_{zy}} \left.\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}}\right|_{\mathbf{y}=\bar{\mathbf{y}}}. \qquad \text{(Example, Eq. 13)}$$

This system is dynamically sufficient as it contains dynamic equations for all evolutionarily dynamic variables, namely both $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$. While $\mathbf{L_z}$ in the first equality is always singular, the matrix $\mathbf{L_{zy}}$ in the second equality is non-singular if the mutational covariance matrix $\mathbf{H_y}$ is non-singular. Thus, the total selection gradient of the genotype still provides a relatively complete description of the evolutionary process of the geno-phenotype.

We can similarly find that the total selection gradient of the genotypic trait at age a is

$$\left.\frac{\mathrm{d}w}{\mathrm{d}y_a}\right|_{\mathbf{y}=\bar{\mathbf{y}}} = \frac{1}{\tilde{T}}\bar{x}_a\frac{1+q}{1-q\bar{y}_a}\left[-\ell_a + \sum_{j=a+1}^{N_\mathrm{a}} \ell_j(1-\bar{y}_j)\prod_{k=a+1}^{j-1}(1+\bar{y}_k)\right],$$

where the generation time without density dependence is now

$$\tilde{T} = \sum_{j=1}^{N_\mathrm{a}} j\ell_j(1-\bar{y}_j)\bar{x}_j\frac{1+q}{1-q\bar{y}_j}.$$

This total selection gradient of the genotypic trait at age a has the same sign as that found in the model for non-social development (Example, Eq. 5). Hence, the stable evolutionary equilibria for the genotype are still given by Example, Eq. 9. Yet, the associated phenotype, given by Example, Eq. 11, may be different due to social development (Fig. 7). That is, social development here does not affect the evolutionary equilibria, as it does not affect the zeros of the total selection gradient of the genotype which gives the zeros of the evolutionary dynamics of the geno-phenotype (Example, Eq. 13). Instead, social development affects here the developmental constraint so it affects the admissible evolutionary equilibria of the phenotype. Numerical solution for the evo-devo dynamics using Example, Eq. 6 is given in Fig. 7. For the q chosen, the phenotype evolves to much larger values due to social feedback than with non-social development although the genotype evolves to the same values. The associated evolution of the $\mathbf{L_z}$ matrix, using Layer 6, Eq. 9, is given in Fig. 8. The code used to generate these figures is in the SI.

7. Discussion

We have addressed the question of how development affects evolution by formulating a mathematical framework that integrates explicit developmental dynamics into evolutionary dynamics. The framework integrates age progression, explicit developmental constraints according to which the phenotype is constructed across life, and evolutionary dynamics. This framework yields a description of the structure of genetic covariation, including the additive effects of allelic substitution $\mathrm{d}\mathbf{x}^\intercal/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}}$, from mechanistic processes. The framework also yields a dynamically sufficient description of the evolution of developed phenotypes in gradient form, such that their long-term evolution can be described as the climbing of a fitness landscape within the assumptions made. This framework provides a tractable method to model the evo-devo dynamics for a broad class of models. We also obtain formulas to compute the sensitivity of the solution of a recurrence (here, the phenotype) to perturbations in the solution or parameters at earlier times (here, ages), which are given by $\mathrm{d}\mathbf{x}^\intercal/\mathrm{d}\boldsymbol{\zeta}$ for $\boldsymbol{\zeta} \in \{\mathbf{x}, \mathbf{y}\}$. Overall, the framework provides a theory of long-term constrained evolutionary dynamics, where the developmental and environmental constraints determine the admissible evolutionary path (Layer 7, Eq. 1).

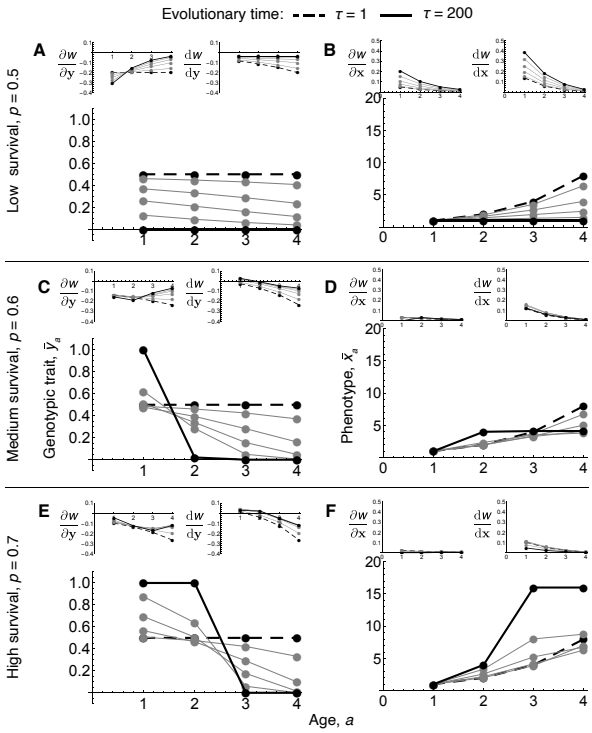

Figure 7: Example with social development. The genotype evolves to the same values as those with non-social development in Fig. 5. However, the phenotype evolves to much larger values due to social development. Total and direct selection for the phenotype (small plots in B,D,F) are smaller than with non-social development, to the point of there being almost no direct selection on the phenotype with high survival indicating that the phenotype is near a fitness peak even for late ages (small left plot in F). Large plots give the resident genotype (left) or phenotype (right) vs age over evolutionary time for various survival probabilities p. Small plots give the associated direct and total selection gradients. The numerical evolutionary dynamics of the genotype match the analytical expressions for the genotype (Example, Eq. 9) and associated phenotype (Example, Eq. 11). ι is as in Fig. 5. $q = 0.5$.

Previous understanding suggested that development affects evolution by inducing genetic covariation and genetic constraints, although the nature of such constraints had remained uncertain. We find that genetic constraints are necessarily absolute in a generally dynamically sufficient description of long-term phenotypic evolution in gradient form. This is because dynamic sufficiency in general requires that not only phenotypic but also genotypic evolution is followed. Because the phenotype is related to the genotype via development, simultaneously describing the evolution of the genotype and phenotype in gradient form entails that the associated constraining matrix ($\mathbf{H_z}$ or $\mathbf{L_z}$) is necessarily singular with a maximum number of degrees of freedom given by the number of lifetime genotypic traits ($N_\mathrm{a}N_\mathrm{g}$). Consequently, genetic covariation is necessarily absent in as many directions of geno-phenotype space as there are lifetime developed traits ($N_\mathrm{a}N_\mathrm{p}$). Since the constraining matrix is singular, direct directional selection is insufficient to identify evolutionary equilibria in contrast to common practice. Instead, total genotypic selection, which depends on development, is sufficient to identify evolutionary equilibria if there are no absolute mutational constraints and no exogenous plastic response. Yet, absence of total genotypic selection is insufficient to identify evolutionary outcomes

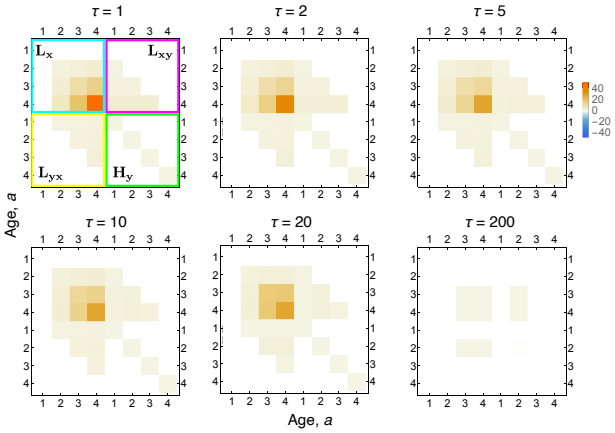

Figure 8: Resulting dynamics of the mechanistic additive socio-genetic cross-covariance matrix $\mathbf{L}_\mathbf{z}$. The structure and dynamics of $\mathbf{L}_\mathbf{z}$ here are similar to those of $\mathbf{H}_\mathbf{z}$ in Fig. 8 but the magnitudes are an order of magnitude larger (compare bar legends). $p = 0.7$, $q = 0.5$. The evolutionary times τ shown correspond to those of Fig. 7.

because the singularity of the constraining matrix associated to direct geno-phenotypic selection entails that if there is any evolutionary equilibrium and no exogenous plastic response, then there is an infinite number of evolutionary equilibria that depend on development. Which of these equilibria is attained also depends on development as it determines the admissible evolutionary trajectory and so the admissible equilibria. The adaptive topography in phenotype space is often assumed to involve a non-singular $\mathbf{G}$-matrix where evolutionary outcomes occur at fitness landscape peaks (i.e., where $\partial w/\partial \mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} = \mathbf{0}$). In contrast, we find that the evolutionary dynamics differ from that representation in that evolutionary outcomes occur at best (i.e., without absolute mutational constraints) at peaks in the admissible evolutionary path determined by development (i.e., where $\mathrm{d}w/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} = \mathbf{0}$), and that such path peaks do not typically occur at landscape peaks (so generally $\partial w/\partial \mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}=\bar{\mathbf{y}}^*} \neq \mathbf{0}$).

The singularity of the constraining matrix ($\mathbf{H}_\mathbf{z}$ or $\mathbf{L}_\mathbf{z}$) is not due to our adaptive dynamics assumptions. Under quantitative genetics assumptions, the additive genetic covariance matrix of phenotype $\mathbf{x}$ is $\mathbf{G}_\mathbf{x} = \boldsymbol{\alpha}_\mathbf{x}\mathrm{cov}[\mathbf{y},\mathbf{y}]\boldsymbol{\alpha}_\mathbf{x}^\mathsf{T}$ as described in the introduction, and here we use the subscripts $\mathbf{x}$ to highlight that this $\boldsymbol{\alpha}$ matrix is for the regression coefficients of the phenotype with respect to gene content. Under quantitative genetics assumptions, the matrix $\mathrm{cov}[\mathbf{y},\mathbf{y}]$ describes the observed covariance in allele frequency due to any source, so it describes standing covariation in allele frequency. Under our adaptive dynamics assumptions, we obtain an $\mathbf{H}_\mathbf{x}$ matrix that has the same form of $\mathbf{G}_\mathbf{x}$, but where $\mathrm{cov}[\mathbf{y},\mathbf{y}]$ describes the covariance in genotypic traits only due to mutation at the current evolutionary time step among the possible mutations, so it describes (expected) mutational covariation. Regardless of whether $\mathrm{cov}[\mathbf{y},\mathbf{y}]$ describes standing covariation in allele frequency or mutational covariation, the additive genetic covariance matrix in geno-phenotype space $\mathbf{G}_\mathbf{z} = \boldsymbol{\alpha}_\mathbf{z}\mathrm{cov}[\mathbf{y},\mathbf{y}]\boldsymbol{\alpha}_\mathbf{z}^\mathsf{T}$ is always singular because the developmental matrix of the geno-phenotype $\boldsymbol{\alpha}_\mathbf{z}^\mathsf{T}$ has fewer rows than columns: that is, the degrees of freedom of $\mathbf{G}_\mathbf{z}$ have an upper bound given by the number of loci (or genetic predictors) while the size of $\mathbf{G}_\mathbf{z}$ is given by the number of loci and of phenotypes. Thus, whether one

considers standing or mutational covariation, the additive genetic covariance matrix of the geno-phenotype is always singular. Eliminating traits from the analysis to render $\mathbf{G_z}$ non-singular as traditionally recommended (Lande, 1979) either renders the gradient system underdetermined and so dynamically insufficient in general (if allele frequency $\bar{\mathbf{y}}$ is removed), or prevents a description of phenotypic evolution as the climbing of a fitness landscape (if the mean phenotype $\bar{\mathbf{x}}$ is removed). The singularity of $\mathbf{H}$ and $\mathbf{L}$ in geno-phenotype space persists despite evolution of the developmental map, regardless of the number of genotypic traits or phenotypes provided there is any phenotype, and in the presence of endogenous or exogenous environmental change. Thus, we find that a dynamically sufficient description of phenotypic evolution in gradient form generally requires a singular constraining matrix. It remains to be seen whether these conclusions hold under other types of developmental constraints, for instance, inequality constraints or stochastic development.

In our results, dynamic sufficiency for phenotypic evolution in gradient form requires that the constraining matrix is in geno-phenotype space particularly because of non-linear development. The $\mathbf{H}$-matrix in phenotype space generally depends on the resident genotype via both the mutational covariance matrix and the developmental matrix. The developmental matrix depends on the resident genotype due to non-linear development, particularly gene-gene interaction, gene-phenotype interaction, and gene-environment interaction (see text below Layer 6, Eq. 5). The analogous dependence of $\mathbf{G}$ on allele frequency holds under quantitative genetics assumptions for the same reasons (Turelli, 1988; Service and Rose, 1985). If development is linear (i.e., the developmental map for all phenotypes is a linear function in all its variables at all ages), the developmental matrix no longer depends on the resident genotype (or allele frequency under quantitative genetics assumptions). If in addition the mutational covariance matrix is independent of the resident genotype, then the constraining matrix $\mathbf{H}$ in phenotype space is no longer dependent on the resident genotype. Thus, if one assumes linear development and both mutational covariation and phenotypic selection being independent of the resident genotype (in addition to no social interactions, no exogenous plastic response, no total immediate genotypic selection, and no niche-constructed effects of the phenotype on fitness; Layer 7, Eq. 6), the $\mathbf{H}$ matrix in phenotype space becomes constant and the mechanistic Lande equation (Layer 7, Eq. 6) becomes dynamically sufficient. However, even simple models of explicit development involve non-linearities (e.g., Example, Eq. 1) and mutational covariation depends on the resident genotype whenever the genotype is constrained to take values within a finite range (e.g., between zero and one). Thus, consideration of even slightly realistic models of development seems unlikely to allow for a long-term dynamically sufficient mechanistic Lande equation (i.e., following only phenotypic evolution).

Extensive research efforts have been devoted to determining the relevance of constraints in adaptive evolution (Arnold, 1992; Hine and Blows, 2006; Hansen and Houle, 2008; Jones et al., 2014; Hine et al., 2014; Engen and Sæther, 2021). Empirical research has found that the smallest eigenvalue of $\mathbf{G}$ in phenotype space is often close to zero (Kirkpatrick and Lofsvold, 1992; Hine and Blows, 2006; McGuigan and Blows, 2007). Mezey and Houle (2005) found a non-singular $\mathbf{G}$-matrix for 20 morphological (so, developed) traits in fruit flies. Our results suggest $\mathbf{G}$ singularity would still arise in these studies if enough traits are included to provide a dynamically sufficient description

of long-term phenotypic evolution on an adaptive topography (i.e., if allele frequency were included in the analysis as part of the multivariate "geno-phenotype").

Previous theory has offered limited predictions as to when the **G**-matrix would be singular. These include that incorporating more traits in the analysis renders **G** more likely to be singular as the traits are more likely to be genetically correlated, such as in infinite-dimensional traits (Gomulkiewicz and Kirkpatrick, 1992; Kirkpatrick and Lofsvold, 1992). Suggestions to include gene frequency as part of the trait vector in the classic Lande equation (e.g., Barfield et al., 2011) have been made without noticing that doing so entails that the associated **G**-matrix is necessarily singular. Kirkpatrick and Lofsvold (1992, p. 962 onwards) showed that, assuming that **G** in phenotypic space is singular and constant, then the evolutionary trajectory and equilibria depend on the evolutionarily initial conditions of the phenotype. Such dependence on the evolutionarily initial conditions is sometimes called phylogenetic constraints (Hansen, 1997). Our results extend the relevance of Kirkpatrick and Lofsvold's (1992) analysis by our observation that the constraining matrix is always singular in a dynamically sufficient gradient system for long-term phenotypic evolution, even with few traits and evolution of the constraining matrix. In our framework, the evolutionarily initial conditions of the phenotype are given by the developmental constraint evaluated at the evolutionarily initial genotype and environment. Hence, the evolutionary trajectory, equilibria, and outcomes depend on the developmental constraint, which provides the admissible evolutionary path.

Multiple mathematical models have studied whether **G** is singular. Recently, simulation work studying the effect of pleiotropy on the structure of the **G**-matrix found that the smallest eigenvalue of **G** is very small but positive (Engen and Sæther, 2021, Tables 3 and 5). Our findings indicate that this model and others (e.g., Wagner, 1984; Barton and Turelli, 1987; Wagner, 1989; Wagner and Mezey, 2000; Martin, 2014; Morrissey, 2014, 2015) would recover **G**-singularity by considering the geno-phenotype, so that both allele frequency and phenotype change are part of the gradient system. Other recent simulation work found that a singular **G**-matrix due to few segregating alleles still allows the phenotype to reach its unconstrained optimum if all loci have segregating alleles at some point over the long run, thus allowing for evolutionary change in all directions of phenotype space in the long run (Barton, 2017, Fig. 3). Our results indicate that such a model attains the unconstrained optimum because it assumes that fitness depends on a single phenotype at a single age, that additive effects of allelic substitution are always positive, and that there is no direct genotypic selection and no niche-constructed effects of the genotype on fitness (i.e., there $\partial w/\partial \mathbf{y} = \mathbf{0}$ and $(\mathrm{d}\boldsymbol{\epsilon}^{\mathsf{T}}/\mathrm{d}\mathbf{y})(\partial w/\partial \boldsymbol{\epsilon}) = \mathbf{0}$, so $\mathrm{d}w/\mathrm{d}y_{ia} = \sum_{j=1}^{N_\mathrm{a}} \sum_{k=1}^{N_\mathrm{p}} (\mathrm{d}x_{kj}/\mathrm{d}y_{ia})(\partial w/\partial x_{kj})$, which since fitness depends on a single trait k at a single age j further reduces to $(\mathrm{d}x_{\mathrm{kj}}/\mathrm{d}y_{ia})(\partial w/\partial x_{\mathrm{kj}})$; hence, assuming $\mathrm{d}x_{\mathrm{kj}}/\mathrm{d}y_{ia} > 0$, there we have that $\mathrm{d}w/\mathrm{d}y_{\mathrm{ij}} = 0$ for any locus i implies $\partial w/\partial x_{\mathrm{kj}} = 0$; Layer 4, Eq. S21). Our results show that when at least one of these assumptions does not hold, the unconstrained optimum is not necessarily achieved (as illustrated in Example, Eq. 3 and Fig. 5). In our framework, phenotypic evolution converges at best to constrained fitness optima, which may under certain conditions coincide with unconstrained fitness optima. Convergence to constrained fitness optima under no absolute mutational constraints occurs even with the fewest number of traits allowed in our framework: two, that is, one genotypic trait and one phenotype with one age each (or in a standard quantitative genetics framework, allele frequency at a single

locus and one quantitative trait that is a function of such allele frequency). Such constrained adaptation has important implications for biological understanding (see e.g., Kirkpatrick and Lofsvold, 1992; Gomulkiewicz and Kirkpatrick, 1992). It is also consistent with empirical observations of lack of selection response in the wild despite selection and genetic variation (Merilä et al., 2001; Hansen and Houle, 2004; Pujol et al., 2018), and of relative lack of stabilizing selection across many species (Kingsolver et al., 2001; Kingsolver and Diamond, 2011).

Other modelling work studied the evolutionary relevance of developmental factors without concluding that there are necessarily absolute genetic constraints. Wagner (1984, 1989) constructed and analyzed evolutionary models considering developmental maps, and wrote the **G**-matrix in terms of his developmental matrix to assess its impact on the maintenance of genetic variation. Wagner (1984, 1988, 1989) and Wagner and Mezey (2000) did not simultaneously track the evolution of genotypes and phenotypes, so did not conclude that the associated **G**-matrix is necessarily singular or that the developmental matrix affects evolutionary equilibria. Altenberg (1995) used Wagner's (1984, 1989) developmental matrix to model constrained adaptation (see Altenberg 1995 Fig. 2) without evolutionary dynamics, but found an analytical expression for the constrained optimum phenotype that equals the globally optimum phenotype (i.e., mapping the constrained optimal genotype in his Eq. 10 to phenotype, $A\hat{\mathbf{x}}$ in his notation, yields the globally optimum phenotype). This and other studies (Via and Lande 1985, Houle 1991, his Fig. 2 and Kirkpatrick and Lofsvold 1992, their Fig. 5) have illustrated how constrained evolutionary dynamics would proceed if the **G** matrix is singular, considering it as a possible case rather than a necessary case. Other models have found that epistasis can cause the evolutionary dynamics to take an exponentially long time to reach fitness peaks (Kaznatcheev, 2019). We find that as the constraining matrix in geno-phenotype space has at least as many zero eigenvalues as there are lifetime phenotypes (i.e., $N_{\mathrm{a}}N_{\mathrm{p}}$), even if there were infinite time, the population does not necessarily reach a fitness peak in geno-phenotype space. The population eventually reaches a fitness peak in genotype space if there are no absolute mutational constraints after the landscape is modified by: (1) the interaction of the total effects of the genotype on phenotype and direct phenotypic selection and (2) the total niche-constructed effects of the genotype on fitness.

We find that total genotypic selection provides more information regarding selection response than direct directional selection or other forms of total selection. We show that evolutionary equilibria occur when total genotypic selection vanishes if there are no absolute mutational constraints and no exogenous plastic response. Direct selection or total selection on the phenotype need not vanish at evolutionary equilibria, even if there are no absolute mutational constraints and no exogenous plastic response. As total genotypic selection depends on development rather than exclusively on direct selection, and as development determines the admissible evolutionary trajectory along which developmental and environmental constraints are satisfied, our findings show that development plays a major evolutionary role by sharing responsibility with selection for defining evolutionary equilibria and for determining the admissible evolutionary path.

Total selection gradients correspond to quantities that have received various names. Such gradients correspond to Caswell's (1982, 2001) "total derivative of fitness" (denoted by him as $\mathrm{d}\lambda$), Charlesworth's (1994) "total differential" (of the population's growth rate, denoted by him as $\mathrm{d}r$), van Tienderen's (1995) "integrated sensitivity" (of the popu-

lation's growth rate, denoted by him as IS), and Morrissey's (2014, 2015) "extended selection gradient" (denoted by him as $\boldsymbol{\eta}$). Total selection gradients measure total directional selection, so in our framework they take into account the downstream developmental effects of a trait on fitness. In contrast, Lande's (1979) selection gradients measure direct directional selection, so in our framework's terms they do not consider the developmentally immediate total effects of a trait on fitness nor the downstream developmental effects of a trait on fitness. We obtained compact expressions for total selection gradients as linear transformations of direct selection gradients, arising from the chain rule in matrix calculus notation (Layer 4, Eq. S19), analogously to previous expressions in terms of vital rates that have no developmental constraints (Caswell, 2001, Eq. 9.38).

Our mechanistic approach to total selection recovers the regression approach to total selection of Morrissey (2014). Morrissey (2014) defined the extended selection gradient as $\boldsymbol{\eta} = \boldsymbol{\Phi}\boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is Lande's selection gradient and $\boldsymbol{\Phi}$ is the matrix of total effects of all traits on themselves (computed as regression coefficients between variables related by a path diagram rather than as total derivatives, which entails material differences with our approach as explained above). Morrissey (2014) used an equation for the total-effect matrix $\boldsymbol{\Phi}$ (his Eq. 2) from path analysis (Greene, 1977, p. 380), which has the form of our matrices describing developmental feedback of the phenotype and the geno-phenotype ($\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}}$ and $\mathrm{d}\mathbf{z}^{\intercal}/\mathrm{d}\mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}}$; Eq. (11) and Layer 4, Eq. S8). Thus, interpreting Morrissey's (2014) $\boldsymbol{\Phi}$ as our $\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}}$ (resp. $\mathrm{d}\mathbf{z}^{\intercal}/\mathrm{d}\mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}}$) and $\boldsymbol{\beta}$ as our $\delta w/\delta\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}}$ (resp. $\delta w/\delta\mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}}$) (i.e., Lande's selection gradient of the phenotype or the geno-phenotype if environmental traits are not explicitly included in the analysis), then Layer 4, Eq. S20 (resp. Layer 4, Eq. S23) shows that the extended selection gradient $\boldsymbol{\eta} = \boldsymbol{\Phi}\boldsymbol{\beta}$ corresponds to the total selection gradient of the phenotype $\mathrm{d}w/\mathrm{d}\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}}$ (resp. of the geno-phenotype $\mathrm{d}w/\mathrm{d}\mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}}$). We did not show that $\mathrm{d}\mathbf{m}^{\intercal}/\mathrm{d}\mathbf{m}|_{\mathbf{y}=\bar{\mathbf{y}}}$ has the form of the equation for $\boldsymbol{\Phi}$ provided by Morrissey (2014) (his Eq. 2), but it might indeed hold. If we interpret $\boldsymbol{\Phi}$ as our $\mathrm{d}\mathbf{m}^{\intercal}/\mathrm{d}\mathbf{m}|_{\mathbf{y}=\bar{\mathbf{y}}}$ and $\boldsymbol{\beta}$ as our $\partial w/\partial\mathbf{m}|_{\mathbf{y}=\bar{\mathbf{y}}}$ (i.e., Lande's selection gradient of the geno-envo-phenotype thus explicitly including environmental traits in the analysis), then Layer 4, Eq. S24 shows that the extended selection gradient $\boldsymbol{\eta} = \boldsymbol{\Phi}\boldsymbol{\beta}$ corresponds to the total selection gradient of the geno-envo-phenotype $\mathrm{d}w/\mathrm{d}\mathbf{m}|_{\mathbf{y}=\bar{\mathbf{y}}}$.

Not all total selection gradients provide a relatively complete description of the selection response. We show in the SI sections S5.7 (Eq. S5.7.5) and S5.9 (Eq. S5.9.5) that the selection response of the geno-phenotype or the geno-envo-phenotype can respectively be written in terms of the total selection gradients of the geno-phenotype $\mathrm{d}w/\mathrm{d}\mathbf{z}|_{\mathbf{y}=\bar{\mathbf{y}}}$ or the geno-envo-phenotype $\mathrm{d}w/\mathrm{d}\mathbf{m}|_{\mathbf{y}=\bar{\mathbf{y}}}$, but such total selection gradients are insufficient to predict lack of selection response because they are premultiplied by a singular socio-genetic cross-covariance matrix. Also, the selection response of the phenotype can be written in terms of the total selection gradient of the phenotype $\mathrm{d}w/\mathrm{d}\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}}$, but this expression for the selection response has an additional term involving the total immediate selection gradient of the genotype $\delta w/\delta\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}}$, so the total selection gradient of the phenotype is insufficient to predict lack of selection response (even more considering that following the evolutionary dynamics of the phenotype alone is generally dynamically insufficient). In contrast, we have shown that the total selection gradient of the genotype $\mathrm{d}w/\mathrm{d}\mathbf{y}|_{\mathbf{y}=\bar{\mathbf{y}}}$ predicts lack of selection response if there are no absolute mutational constraints. Thus, out of all total selection gradients considered, only total genotypic selection provides a relatively complete description of the selection response. Morrissey (2015)

considers that the total selection gradient of the genotype (his "inputs") and of the phenotype (his "traits") would be equal, but the last line of Layer 4, Eq. S21 shows that the total selection gradients of the phenotype and genotype are different in general, particularly due to direct genotypic selection and the total effects of genotype on phenotype.

Our results allow for the modeling of evo-devo dynamics in a wide array of settings. First, developmental and environmental constraints (Eqs. 1 and 2) can mechanistically describe development, gene-gene interaction, and gene-environment interaction, while allowing for arbitrary non-linearities and evolution of the developmental map. Several previous approaches have modelled gene-gene interaction, such as by considering multiplicative gene effects, but general analytical frameworks mechanistically linking gene-gene interaction, gene-environment interaction, developmental dynamics, and evolutionary dynamics have previously remained elusive (Rice, 1990; Hansen and Wagner, 2001; Rice, 2002; Hermisson et al., 2003; Carter et al., 2005; Rice, 2011). A historically dominant yet debated view is that gene-gene interaction has minor evolutionary effects (Hansen, 2013; Nelson et al., 2013; Paixão and Barton, 2016; Barton, 2017). Our finding that the constraining matrix ($\mathbf{H}$ or $\mathbf{L}$) is necessarily singular in a long-term dynamically sufficient phenotypic adaptive topography entails that evolutionary equilibria depend on development and consequently on gene-gene and gene-environment interactions. Hence, gene-gene and gene-environment interaction can generally have strong and permanent evolutionary effects in the sense of defining together with selection what the evolutionary equilibria are (e.g., via developmental feedbacks described by $\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}}$) even by altering the constraining matrix alone. This contrasts with a non-singular constraining matrix whereby evolutionary equilibria are pre-determined by selection.

Second, our results allow for the study of long-term evolution of the constraining matrix as an emergent property of the evolution of the genotype, phenotype, and environment (i.e., the geno-envo-phenotype). In contrast, it has been traditional to study short-term evolution of $\mathbf{G}$ by treating it as another dynamic variable under constant allele frequency (Bulmer, 1971; Lande, 1979; Bulmer, 1980; Lande, 1980; Lande and Arnold, 1983; Barton and Turelli, 1987; Turelli, 1988; Gavrilets and Hastings, 1994; Carter et al., 2005; Débarre et al., 2014). Third, our results allow for the study of the effects of developmental bias, biased genetic variation, and modularity (Wagner, 1996; Pavlicev and Hansen, 2011; Pavlicev et al., 2011; Wagner and Zhang, 2011; Pavlicev and Wagner, 2012; Watson et al., 2013). While we have assumed that mutation is unbiased for the genotype, our equations allow for the developmental map to bias the phenotype distribution and the genetic variation of the phenotype. This may lead to modular effects of mutations, whereby altering a genotypic trait at a given age tends to affect some phenotypes but not others.

Fourth, our equations facilitate the study of life-history models with dynamic constraints. Life-history models with dynamic constraints have typically assumed evolutionary equilibrium, so they are analyzed using dynamic optimization techniques such as dynamic programming and optimal control (e.g., León, 1976; Iwasa and Roughgarden, 1984; Houston and McNamara, 1999; González-Forero et al., 2017; Avila et al., 2021). In recent years, mathematically modeling the evolutionary dynamics of life-history models with dynamic constraints, that is, of what we call the evo-devo dynamics, has been made possible with the canonical equation of adaptive dynamics for function-valued traits (Dieckmann et al., 2006; Parvinen et al., 2013; Metz et al., 2016). However, such an approach poses substantial

mathematical challenges by requiring derivation of functional derivatives and solution of associated differential equations for costate variables (Parvinen et al., 2013; Metz et al., 2016; Avila et al., 2021). By using discrete age, we have obtained closed-form equations that facilitate modeling the evo-devo dynamics. By doing so, our framework yields an alternative method to dynamic optimization to analyze a broad class of life-history models with dynamic constraints (Example and González-Forero 2023).

Fifth, our framework allows for the modeling of the evo-devo dynamics of pattern formation by allowing the implementation of reaction-diffusion equations in *discrete space* in the developmental map, once equations are suitably written (e.g., Eq. 6.1 of Turing, 1952; Tomlin and Axelrod, 2007; Deutsch and Dormann, 2017; SI section S1.2). Thus, the framework may allow one to implement and analyze the evo-devo dynamics of existing detailed models of the development of morphology (e.g., Salazar-Ciudad and Jernvall, 2010; Sheth et al., 2012), to the extent that developmental maps can be written in the form of Eq. (1). Sixth, our framework allows for the mechanistic modeling of adaptive plasticity, for instance, by implementing reinforcement learning or supervised learning in the developmental map (Sutton and Barto, 2018; Paenke et al., 2007).

Our approach might also enable modelling developmental innovation, whereby fully formed traits emerge de novo in an individual, such as an extra digit (Goldschmidt, 1940; Gould, 1977; Orr and Coyne, 1992; Orr, 2005; Müller, 2010). The developmental constraint (1) admits that a slight perturbation in the geno-envo-phenotype at an early age yields a large change in the phenotype at a later age, possibly changing it from zero to an appreciable value. This may be used to model developmental innovation, possibly via exploratory processes described by Gerhart and Kirschner (2007) and Kirschner and Gerhart (2010) provided that a mathematical model of such processes satisfies Eq. 1. However, slight perturbations yielding large phenotypic effects might violate our assumption that invasion implies fixation. It has previously been established that invasion implies fixation if mutant *genotypes* $\mathbf{y}$ do not deviate substantially from resident genotypes $\bar{\mathbf{y}}$ (Geritz et al., 2002; Geritz, 2005; Dieckmann et al., 2006; Priklopil and Lehmann, 2020; Avila et al., 2021), which we assume. Application of our framework to model developmental innovation will require explicitly determining whether large deviations in mutant phenotypes in our sense of the word still entail that invasion implies fixation because of small deviations in mutant genotypes.

By allowing development to be social, our framework allows for a mechanistic description of extra-genetic inheritance and indirect genetic effects. Extra-genetic inheritance can be described since the phenotype at a given age can be an identical or modified copy of the geno-phenotype of social partners. Thus, social development allows for the modeling of social learning (Sutton and Barto, 2018; Paenke et al., 2007) and epigenetic inheritance (Jablonka et al., 1992; Slatkin, 2009; Day and Bonduriansky, 2011). However, in our framework extra-genetic inheritance is insufficient to yield phenotypic evolution that is independent of both genetic evolution and exogenous plastic change (e.g., in the framework, there cannot be cultural evolution without genetic evolution or exogenous environmental change). This is seen by setting mutational covariation and exogenous environmental change to zero (i.e., $\mathbf{H}_{\mathbf{y}} = \mathbf{0}$ and $\partial\bar{\boldsymbol{\epsilon}}/\partial\tau = \mathbf{0}$), which eliminates evolutionary change (i.e., $\mathrm{d}\bar{\mathbf{m}}/\mathrm{d}\tau = \mathbf{0}$). The reason seems to be that although there is extra-genetic *inheritance* in our framework, there is no extra-genetic *variation* because both development is

deterministic and we use adaptive dynamics assumptions: without mutation, every SDS resident develops the same phenotype as every other resident. Extensions to consider stochastic development might enable extra-genetic variation and possibly phenotypic evolution that is independent of genetic and exogenously plastic evolution. Also, we have only considered social interactions among non-relatives, so our framework at present only allows for social learning or epigenetic inheritance from non-relatives.

Our framework can mechanistically describe indirect genetic effects via social development because the developed phenotype can be mechanistically influenced by the genotype or phenotype of social partners. Indirect genetic effects mean that a phenotype may be partly or completely caused by genes located in another individual (Moore et al., 1997). Indirect genetic effect approaches model the phenotype considering a linear regression of individual's phenotype on social partner's phenotype (Kirkpatrick and Lande, 1989; Moore et al., 1997; Townley and Ezard, 2013), whereas our approach constructs individual's phenotype from development depending on social partners' genotype and phenotypes. We found that social development generates social feedback (described by $\mathrm{s}\mathbf{x}/\mathrm{s}\bar{\mathbf{x}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$, Eq. 27), which closely though not entirely corresponds to social feedback found in the indirect genetic effects literature (Moore et al., 1997, Eq. 19b and subsequent text). The social feedback we obtain depends on total social developmental bias from the phenotype ($\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{x}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$, Eq. 28); analogously, social feedback in the indirect genetic effects literature depends on the matrix of interaction coefficients ($\boldsymbol{\Psi}$) which contains the regression coefficients of phenotype on social partner's phenotype. Social development leads to a generalization of mechanistic additive genetic covariance matrices $\mathbf{H} = \mathrm{cov}[\mathbf{b}, \mathbf{b}]$ into mechanistic additive socio-genetic cross-covariance matrices $\mathbf{L} = \mathrm{cov}[\mathbf{b}^{\mathrm{s}}, \mathbf{b}]$; similarly, indirect genetic effects involve a generalization of the $\mathbf{G}$-matrix, which includes $\mathbf{C}_{\mathbf{ax}} = \mathrm{cov}[\mathbf{a}, \mathbf{x}]$, namely the cross-covariance matrix between multivariate breeding value and phenotype (Kirkpatrick and Lande, 1989; Moore et al., 1997; Townley and Ezard, 2013).

However, there are differences between our results and those in the indirect genetic effects literature. First, social feedback (in the sense of inverse matrices involving $\boldsymbol{\Psi}$) appears twice in the evolutionary dynamics under indirect genetic effects (see Eqs. 20 and 21 of Moore et al. 1997) while it only appears once in our evolutionary dynamics equations through $\mathrm{s}\mathbf{x}/\mathrm{s}\bar{\mathbf{x}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$ (Layer 6, Eq. 10). This difference may stem from the assumption in the indirect genetic effects literature that social interactions are reciprocal, while we assume that they are asymmetric in the sense that, since mutants are rare, mutant's development depends on residents but resident's development does not depend on mutants (we thank J. W. McGlothlin for pointing this out). Second, our $\mathbf{L}$ matrices make the evolutionary dynamics equations depend on total social developmental bias from the genotype ($\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{y}}^{\mathsf{T}}|_{\mathbf{y}=\bar{\mathbf{y}}}$, Eq. 26) in a non-feedback manner (specifically, not in an inverse matrix) but this type of dependence does not occur in the evolutionary dynamics under indirect genetic effects (Eqs. 20 and 21 of Moore et al. 1997). This difference might stem from the absence of explicit tracking of allele frequency in the indirect genetic effects literature in keeping with the tradition of quantitative genetics, whereas we explicitly track genotypic traits. Third, "social selection" (i.e., $\partial w/\partial \bar{\mathbf{z}}$) plays no role in our results consistently with our assumption of a well-mixed population, but social selection plays an important role in the indirect genetic effects literature even if relatedness is zero (McGlothlin et al., 2010, e.g., setting $r = 0$ in their Eq. 10 still

leaves an effect of social selection on selection response due to "phenotypic" kin selection).

Our framework offers formalizations to the notions of developmental constraints and developmental bias. The two notions have been often interpreted as equivalents (e.g., Brakefield, 2006), or with a distinction such that constraints entail a negative, prohibiting effect while bias entails a positive, directive effect of development on the generation of phenotypic variation (Uller et al., 2018; Salazar-Ciudad, 2021). We defined developmental constraint as the condition that the phenotype at a given age is a function of the individual's condition at their immediately previous age, which both prohibits certain values of the phenotype and has a "directive" effect on the generation of phenotypic variation. We offered quantification of developmental bias in terms of the slope of the phenotype with respect to itself at subsequent ages. No bias would lead to zero slopes thus to identity matrices (e.g., $\partial\mathbf{x}^{\intercal}/\partial\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{I}$ and $\mathrm{d}\mathbf{x}^{\intercal}/\mathrm{d}\mathbf{x}|_{\mathbf{y}=\bar{\mathbf{y}}} = \mathbf{I}$) and deviations from the identity matrix would constitute bias.

Our results clarify the role of several developmental factors previously suggested to be evolutionarily important. We have arranged the evo-devo process in a layered structure, where a given layer is formed by components of layers below (Fig. 4). This layered structure helps see that several developmental factors previously suggested to have important evolutionary effects (Laland et al., 2014) but with little clear connection (Welch, 2017) can be viewed as basic elements of the evolutionary process. Direct-effect matrices (Layer 2) are basic in that they form all the components of the evolutionary dynamics (Layer 7) except mutational covariation and exogenous environmental change. Direct-effect matrices quantify direct (i) directional selection, (ii) developmental bias, (iii) niche construction, (iv) social developmental bias (e.g., extra-genetic inheritance and indirect genetic effects; Moore et al. 1997), (v) social niche construction, (vi) environmental sensitivity of selection (Chevin et al., 2010), and (vii) phenotypic plasticity. These factors variously affect selection and development, thus affecting evolutionary equilibria and the admissible evolutionary trajectory.

Our approach uses discrete rather than continuous age, which substantially simplifies the mathematics. This treatment allows for the derivation of closed-form expressions for what can otherwise be a difficult mathematical challenge if age is continuous (Kirkpatrick and Heckman, 1989; Dieckmann et al., 2006; Parvinen et al., 2013; Metz et al., 2016; Avila et al., 2021). For instance, costate variables are key in dynamic optimization as used in life-history models (Gadgil and Bossert, 1970; León, 1976; Schaffer, 1983; Stearns, 1992; Roff, 1992; Kozłowski and Teriokhin, 1999; Sydsæter et al., 2008), but general closed-form formulas for costate variables were previously unavailable and their calculation often limits the analysis of such models. In SI section S4, we show that our results recover the key elements of Pontryagin's maximum principle, which is a central tool of optimal control theory to analytically solve dynamic optimization problems (Sydsæter et al., 2008). Under the assumption that there are no environmental traits (hence, no exogenous plastic response), in SI section S4, we show that an admissible locally stable evolutionary equilibrium solves a local, dynamic optimization problem of finding a genotype that both "totally" maximises a mutant's lifetime reproductive success R_0 and "directly" maximises the Hamiltonian of Pontryagin's maximum principle. We show that this Hamiltonian depends on costate variables that are proportional to the total selection gradient of the phenotype at evolutionary equilibrium (Eq. S4.3), and that the costate variables satisfy the costate equations of

Pontryagin's maximum principle. Thus, our approach offers an alternative method to optimal control theory to find admissible evolutionary equilibria for the broad class of models considered here. By exploiting the discretization of age, we have obtained various formulas that can be computed directly for the total selection gradient of the phenotype (Layer 4, Eq. S20), so for costate variables, and of their relationship to total genotypic selection (fifth line of Layer 4, Eq. S21), thus facilitating analytic and numerical treatment of life-history models with dynamic constraints. Although discretization of age may induce numerical imprecision relative to continuous age (Kirkpatrick and Heckman, 1989), numerical and empirical treatment of continuous age typically involves discretization at one point or another, with continuous curves often achieved by interpolation (e.g., Kirkpatrick et al., 1990; Rao, 2009). Numerical precision with discrete age may be increased by reducing the age bin size (e.g., to represent months or days rather than years; Caswell, 2001), at a computational cost.

By simplifying the mathematics, our approach yields insight that may be otherwise challenging to gain. Life-history models with dynamic constraints generally find that costate variables are non-zero under optimal controls (Gadgil and Bossert, 1970; Taylor et al., 1974; León, 1976; Schaffer, 1983; Houston et al., 1988; Houston and McNamara, 1999; Sydsæter et al., 2008). This means that there is persistent total selection on the phenotype at evolutionary equilibrium. Our findings show that this is to be expected for various reasons including if there are absolute mutational constraints (i.e., active path constraints so controls remain between zero and one, as in the Example), direct genotypic selection, or more state variables than control variables (in which case $\delta\mathbf{x}^{\intercal}/\delta\mathbf{y}$ is singular as it has more rows than columns, even after removing initial states and final controls from the analysis; Eq. S5.2.10) (fifth line of Layer 4, Eq. S21). Thus, zero total genotypic selection at equilibrium may involve persistent total phenotypic selection. Moreover, life-history models with explicit developmental constraints have found that their predictions can be substantially different from those found without explicit developmental constraints. In particular, without developmental constraints, the outcome of parent-offspring conflict over sex allocation has been found to be an intermediate between the outcomes preferred by mother and offspring (Reuter and Keller, 2001), whereas with developmental constraints, the outcome has been found to be that preferred by the mother (Avila et al., 2019). Our results show that changing the particular form of the developmental map may induce substantial changes in predictions by influencing total genotypic selection and the admissible evolutionary equilibria. In other words, the developmental map used alters the evolutionary outcome because it modulates absolute socio-genetic constraints (i.e., the $\mathbf{H}$ or $\mathbf{L}$ matrices in geno-phenotype space).

We have obtained a term that we refer to as exogenous plastic response, which is the plastic response to exogenous environmental change over an evolutionary time step (Layer 7, Eq. 3). An analogous term occurs in previous equations (Eq. A3 of Chevin et al. 2010). Additionally, our framework considers *endogenous* plastic response due to niche construction (i.e., endogenous environmental change), which affects both the selection response and the exogenous plastic response. Exogenous plastic response affects the evolutionary dynamics even though it is not ultimately caused by change in the resident genotype (or in gene frequency), but by exogenous environmental change. In particular, exogenous plastic response allows for a straightforward form of "plasticity-first" evolution (Waddington, 1942, 1961;

West-Eberhard, 2003) as follows. At an evolutionary equilibrium where exogenous plastic response is absent, the introduction of exogenous plastic response generally changes socio-genetic covariation or directional selection at a subsequent evolutionary time, thereby inducing selection response. This constitutes a simple form of plasticity-first evolution, whereby plastic change precedes genetic change, although the plastic change may not be adaptive and the induced genetic change may have a different direction to that of the plastic change.

Empirical estimation of the developmental map may be facilitated by it defining a dynamic equation. Whereas the developmental map defines a dynamic equation to construct the phenotype, the genotype-phenotype map corresponds to the solution of such dynamic equation. It is often impractical or impossible to write the solution of a dynamic equation, even if the dynamic equation can be written in practice. Accordingly, it may often prove impractical to empirically estimate the genotype-phenotype map, whereas it may be more tractable to empirically infer developmental maps. Inference of developmental maps from empirical data can be pursued via the growing number of methods to infer dynamic equations from data (Schmidt and Lipson, 2009; Brunton et al., 2016; Ghadami and Epureanu, 2022; Course and Nair, 2023).

To conclude, we have formulated a framework that synthesizes developmental and evolutionary dynamics yielding a theory of long-term phenotypic evolution on an adaptive topography, which also mechanistically describes the long-term evolution of genetic covariation. This framework shows that development has major evolutionary effects by finding that selection and development jointly define the evolutionary outcomes if mutation is not absolutely constrained and exogenous plastic response is absent, rather than the outcomes being defined only by selection. Our results provide a tool to chart major territory on how development affects evolution.

8. Acknowledgements

I thank Andy Gardner for extensive support throughout this project, by discussing, reading, and commenting on many early drafts, and generously offering interpretation, advice, and funding. I thank K.N. Laland, R. Lande, L.C. Mikula, A.J. Moore, and M.B. Morrissey for comments on previous versions of the manuscript, and A. Krause, J.W. McGlothlin, J.A.J. Metz, L. Milocco, I. Salazar-Ciudad, and D.M. Shuker for discussion. I thank T.F. Hansen, M. Pavlicev, S.J. Schreiber, and three anonymous reviewers for comments that helped to greatly improve the manuscript. I thank M.B. Morrissey for discussion and explanation of his work. This work was funded by an ERC Consolidator Grant to A. Gardner (grant no. 771387), by the School of Biology of the University of St Andrews, and by a John Templeton Foundation grant to K.N. Laland and T. Uller (grant ID 60501). A preprint of the manuscript is available at https://www.biorxiv.org/content/10.1101/2021.05.17.444499.

Appendix A. Matrix calculus notation

Following Caswell (2019), for vectors $\mathbf{a} \in \mathbb{R}^{n\times 1}$ and $\mathbf{b} \in \mathbb{R}^{m\times 1}$, we denote

$$\frac{\partial \mathbf{a}}{\partial \mathbf{b}^\intercal} = \begin{pmatrix} \frac{\partial a_1}{\partial b_1} & \cdots & \frac{\partial a_1}{\partial b_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial a_n}{\partial b_1} & \cdots & \frac{\partial a_n}{\partial b_m} \end{pmatrix} \in \mathbb{R}^{n\times m},$$

so $(\partial \mathbf{a}/\partial \mathbf{b}^\intercal)^\intercal = \partial \mathbf{a}^\intercal/\partial \mathbf{b}$. The same notation applies with total derivatives.

References

Alberch, P., Gould, S.J., Oster, G.F., Wake, D.B., 1979. Size and shape in ontogeny and phylogeny. Paleobiology 5, 296–317.

Alon, U., 2020. An Introduction to System Biology. 2nd ed., Taylor & Francis, Boca Raton, FL, USA.

Altenberg, L., 1995. Genome growth and the evolution of the genotype-phenotype map, in: Banzhaf, W., Eeckman, F.H. (Eds.), Evolution and biocomputation. Springer-Verlag. volume 899 of *Lecture Notes in Computer Science*, pp. 205–259.

Aoki, K., Wakano, J.Y., Lehmann, L., 2012. Evolutionarily stable learning schedules and cumulative culture in discrete generation models. Theor. Popul. Biol. 81, 300–309.

Arnold, S.J., 1992. Constraints on phenotypic evolution. Am. Nat. 140, S85–S107.

Arthur, W., 2004. Biased Embryos and Evolution. Cambridge Univ. Press, Cambridge, UK.

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

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

Baldwin, J.M., 1896. A new factor in evolution. Am. Nat. 30, 441–451.

Barfield, M., Holt, R.D., Gomulkiewicz, R., 2011. Evolution in stage-structured populations. Am. Nat. 177, 397–409.

Barresi, M.J.F., Gilbert, S.F., 2020. Developmental Biology. 12th ed., Oxford Univ. Press, Oxford, UK.

Barton, N.H., 2017. How does epistasis influence the response to selection? Heredity 118, 96–109.

Barton, N.H., Etheridge, A.M., Véber, A., 2017. The infinitesimal model: definition, derivation, and implications. Theor. Popul. Biol. 118, 50–73.

Barton, N.H., Turelli, M., 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. Genet. Res. 49, 157–173.

Baudisch, A., 2005. Hamilton's indicators of the force of selection. Proc. Natl. Acad. Sci. USA 102, 8263–8268.

Bienvenu, F., Legendre, S., 2015. A new approach to the generation time in matrix population models. Am. Nat. 185, 834–843.

Bonduriansky, R., Day, T., 2018. Extended Heredity: a new understanding of inheritance and evolution. Princeton Univ. Press, Princeton, NJ, USA.

Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. The Univ. of Chicago Press, Chicago, IL, USA.

Brakefield, P.M., 2006. Evo-devo and constraints on selection. Trends Ecol. Evol. 21, 362–368.

Brunton, S.L., Proctor, J.L., Kutz, J.N., 2016. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proc. Natl. Acad. Sci. USA 113, 3932–3937.

Bulmer, M., 1994. Theoretical Evolutionary Ecology. Sinauer, Sunderland, MA, USA.

Bulmer, M.G., 1971. The effect of selection on genetic variability. Am. Nat. 105, 201–211.

Bulmer, M.G., 1980. The Mathematical Theory of Quantitative Genetics. Oxford Univ. Press, Oxford, UK.

Carter, A.J.R., Hermisson, J., Hansen, T.F., 2005. The role of epistatic interactions in the response to selection and the evolution of evolvability. Theor. Popul. Biol. 68, 179–196.

Caswell, H., 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. Theor. Popul. Biol. 14, 215–230.

Caswell, H., 1982. Optimal life histories and the age-specific costs of reproduction. J. Theor. Biol. 98, 519–529.

Caswell, H., 2001. Matrix Population Models. 2nd ed., Sinauer, Sunderland, MA, USA.

Caswell, H., 2009. Stage, age and individual stochasticity in demography. Oikos 118, 1763–1782.

Caswell, H., 2019. Sensitivity Analysis: Matrix Methods in Demography and Ecology. Springer Open, Cham, Switzerland.

Caswell, H., Shyu, E., 2017. Senescence, selection gradients and mortality. Cambridge Univ. Press, Cambridge, UK. chapter 4. pp. 56–82.

Cavalli-Sforza, L.L., Feldman, M.W., 1981. Cultural Transmission and Evolution. Princeton Univ. Press, Princeton, NJ, USA.

Champagnat, N., 2006. A microscopic interpretation for adaptive dynamics trait substitution sequence models. Stoch. Process. Their Appl. 116, 1127–1160.

Chan, D.C., Laufer, E., Tabin, C., Leder, P., 1995. Polydactylous limbs in Strong's Luxoid mice result from ectopic polarizing activity. Development 121, 1971–1978.

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

Charlesworth, B., 1993. Natural selection on multivariate traits in age-structured populations. Proc. R. Soc. Lond. B 251, 47–52.

Charlesworth, B., 1994. Evolution in age-structured populations. 2nd ed., Cambridge Univ. Press.

Charlesworth, B., Lande, R., Slatkin, M., 1982. A neo-Darwinian commentary on macroevolution. Evolution 36, 474–498.

Cheverud, J.M., 1984. Quantitative genetics and developmental constraints on evolution by selection. J. Theor. Biol. 110, 155–171.

Chevin, L.M., Lande, R., Mace, G.M., 2010. Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. PLOS Biology 8, 1–8.

Course, K., Nair, P.B., 2023. State estimation of a physical system with unknown governing equations. Nature 622, 261–267.

Crow, J.F., Kimura, M., 1970. An Introduction to Population Genetics Theory. Blackburn Press, Caldwell, NJ, USA.

Day, T., Bonduriansky, R., 2011. A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. Am. Nat. 178, E18–E36.

de Roos, A., 1997. A gentle introduction to physiologically structured population models, in: Tuljapurkar, S. (Ed.), Structured-population models in marine, terrestrial, and freshwater systems. Chapman & Hall, New York, NY, USA, pp. 119–271.

Débarre, F., Nuismer, S.L., Doebeli, M., 2014. Multidimensional (co)evolutionary stability. Am. Nat. 184, 158–171.

Deutsch, A., Dormann, S., 2017. Cellular Automaton Modeling of Biological Pattern Formation. 2nd ed., Birkhäuser, Boston, MA, USA.

Dieckmann, U., Heino, M., Parvinen, K., 2006. The adaptive dynamics of function-valued traits. J. Theor. Biol. 241, 370–389.

Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. J. Math. Biol. 34, 579–612.

Durinx, M., (Hans) Metz, J.A.J., Meszéna, G., 2008. Adaptive dynamics for physiologically structured population models. J. Math. Biol. 56, 673–742.

Engen, S., Sæther, B.E., 2021. Structure of the G-matrix in relation to phenotypic contributions to fitness. Theor. Popul. Biol. 138, 43–56.

Falconer, D.S., Mackay, T.F.C., 1996. Introduction to Quantitative Genetics. 4th ed., Pearson Prentice Hall, Harlow, England.

Fisher, R.A., 1918. XV.—The correlation between relatives on the supposition of Mendelian inheritance. Trans. Roy. Soc. Edinb. 52, 399–433.

Fisher, R.A., 1930. The Genetical Theory of Natural Selection. Oxford Univ. Press, Oxford, UK.

Gadgil, M., Bossert, W.H., 1970. Life historical consequences of natural selection. Am. Nat. 104, 1–24.

Galis, F., Metz, J.A.J., van Alphen, J.J.M., 2018. Development and evolutionary constraints in animals. Annu. Rev. Ecol. Evol. Syst. 49, 499–522.

Gavrilets, S., Hastings, A., 1994. A quantitative-genetic model for selection on developmental noise. Evolution 48, 1478–1486.

Gerhart, J., Kirschner, M., 2007. The theory of facilitated variation. Proc. Natl. Acad. Sci. USA 104, 8582–8589.

Geritz, S.A.H., 2005. Resident-invader dynamics and the coexistence of similar strategies. J. Math. Biol. 50, 67–82.

Geritz, S.A.H., Gyllenberg, M., Jacobs, F.J.A., Parvinen, K., 2002. Invasion dynamics and attractor inheritance. J. Math. Biol. 44, 548–560.

Ghadami, A., Epureanu, B.I., 2022. Data-driven prediction in dynamical systems: recent developments. Phil. Trans. R. Soc. A 380, 20210213.

Gillespie, J.H., 1983. Some properties of finite populations experiencing strong selection and weak mutation. Am. Nat. 121, 691–708.

Gimelfarb, A., 1982. Quantitative character dynamics: Gametic model. Theor. Popul. Biol. 22, 324–366.

Goldschmidt, R., 1940. The Material Basis of Evolution. Yale Univ. Press, New Haven, CT, USA.

Gomulkiewicz, R., Kirkpatrick, M., 1992. Quantitative genetics and the evolution of reaction norms. Evolution 46, 390–411.

González-Forero, M., 2023. Evo-devo dynamics of hominin brain size. Https://www.biorxiv.org/content/10.1101/2023.03.20.533421v2.

González-Forero, M., Faulwasser, T., Lehmann, L., 2017. A model for brain life history evolution. PLOS Comp. Biol. 13, e1005380.

Gould, S.J., 1977. The return of hopeful monsters. Nat. Hist. 86, 30.

Gould, S.J., Lewontin, R.C., 1979. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc. R. Soc. Lond. B 205, 581–598.

Greene, V.L., 1977. An algorithm for total and indirect causal effects. Political Methodology 4, 369–381.

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

Hansen, T.F., 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51, 1341–1351.

Hansen, T.F., 2013. Why epistasis is important for selection and adaptation. Evolution 67, 3501–3511.

Hansen, T.F., Houle, D., 2004. Evolvability, stabilizing selection, and the problem of stasis, in: Pigliucci, M., Preston, K. (Eds.), Phenotypic integration: studying ecology and evolution of complex phenotypes. Oxford Univ. Press. chapter 5.

Hansen, T.F., Houle, D., 2008. Measuring and comparing evolvability and constraint in multivariate characters. J. Evol. Biol. 21, 1201–1219.

Hansen, T.F., Wagner, G.P., 2001. Modeling genetic architecture: a multilinear theory of gene interaction. Theor. Popul. Biol. 59, 61–86.

Hermisson, J., Hansen, T.F., Wagner, G.P., 2003. Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection. Am. Nat. 161, 708–734.

Hill, W.G., 2017. "Conversion" of epistatic into additive genetic variance in finite populations and possible impact on long-term selection response. J. Anim. Breed. Genet. 134, 196–201.

Hine, E., Blows, M.W., 2006. Determining the effective dimensionality of the genetic variance-covariance matrix. Genetics 173, 1135–1144.

Hine, E., McGuigan, K., Blows, M.W., 2014. Evolutionary constraints in high-dimensional traits sets. Am. Nat. 184, 119–131.

Horn, R.A., Johnson, C.R., 2013. Matrix Analysis. 2nd ed., Cambridge Univ. Press, New York, NY, USA.

Houle, D., 1991. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution 45, 630–648.

Houston, A., Clark, C., McNamara, J., Mangel, M., 1988. Dynamic models in behavioural and evolutionary ecology. Nature 332, 29–34.

Houston, A.I., McNamara, J.M., 1999. Models of Adaptive Behaviour: an approach based on state. Cambridge Univ. Press.

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

Jablonka, E., Lachmann, M., Lamb, M.J., 1992. Evidence, mechanisms and models for the inheritance of acquired characters. J. Theor. Biol. 158, 245–268.

Jablonka, E., Lamb, M.J., 2014. Evolution in Four Dimensions. Revised ed., The MIT Press, London, England.

Jones, A.G., Bürger, R., Arnold, S.J., 2014. Epistasis and natural selection shape the mutational architecture of complex traits. Nat. Comm. 5, 3709.

Kaznatcheev, A., 2019. Computational complexity as an ultimate constraint on evolution. Genetics 212, 245–265.

Kingsolver, J.G., Diamond, S.E., 2011. Phenotypic selection in natural populations: What limits directional selection? Am. Nat. 177, 346–357.

Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert, P., Beerli, P., 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157, 245–261.

Kirkpatrick, M., Heckman, N., 1989. A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters. J. Math. Biol. 27, 429–450.

Kirkpatrick, M., Lande, R., 1989. The evolution of maternal characters. Evolution 43, 485–503.

Kirkpatrick, M., Lofsvold, D., 1992. Measuring selection and constraint in the evolution of growth. Evolution 46, 954–971.

Kirkpatrick, M., Lofsvold, D., Bulmer, M., 1990. Analysis of inheritance, selection and evolution of growth trajectories. Genetics 124, 979–993.

Kirschner, M.W., Gerhart, J.C., 2010. Facilitated variation, in: Pigliucci, M., Müller, G.B. (Eds.), Evolution—the extended synthesis. MIT press,

London, UK, pp. 253–280.

Klingenberg, C.P., 2010. Evolution and development of shape: integrating quantitative approaches. Nat. Rev. Genet. 11, 623–635.

Kobayashi, Y., Wakano, J.Y., Ohtsuki, H., 2015. A paradox of cumulative culture. J. Theor. Biol. 379, 79–88.

Kozłowski, J., Teriokhin, A.T., 1999. Allocation of energy between growth and reproduction: the Pontryagin Maximum Principle solution for the case of age- and season-dependent mortality. Evol. Ecol. Res. 1, 423–441.

Laland, K., Uller, T., Feldman, M., Sterelny, K., Müller, G.B., Moczek, A., Jablonka, E., Odling-Smee, J., 2014. Does evolutionary theory need a rethink? Yes, urgently. Nature 514, 161–164.

Laland, K.N., Uller, T., Feldman, M.W., Sterelny, K., Müller, G.B., Moczek, A., Jablonka, E., Odling-Smee, J., 2015. The extended evolutionary synthesis: its structure, assumptions and predictions. Proc. R. Soc. B 282, 20151019.

Lande, R., 1979. Quantitative genetic analysis of multivariate evolution applied to brain: body size allometry. Evolution 34, 402–416.

Lande, R., 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94, 203–215.

Lande, R., 1982. A quantitative genetic theory of life history evolution. Ecology 63, 607–615.

Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. Evolution 37, 1210–1226.

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

Lewontin, R.C., 1983. Gene, organism and environment, in: Bendall, D.S. (Ed.), Evolution from molecules to men. Cambridge Univ. Press, Cambridge, UK. chapter 14, pp. 273–285.

Litingtung, Y., Dahn, R.D., Li, Y., Fallon, J.F., Chiang, C., 2002. *Shh* and *Gli3* are dispensable for limb skeleton formation but regulate digit number and identity. Nature 418, 979–983.

Lush, J.L., 1937. Animal Breeding Plans. Iowa State Univ. Press, Ames, IA, USA.

Lynch, M., Walsh, B., 1998. Genetics and Analysis of Quantitative Traits. Sinauer, Sunderland, MA, USA.

Martin, G., 2014. Fisher's geometrical model emerges as a property of complex integrated phenotypic networks. Genetics 197, 237–255.

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B., Lande, R., Raup, D., Wolpert, L., 1985. Developmental constraints and evolution. Q. Rev. Biol. .

McGlothlin, J.W., Moore, A.J., Wolf, J.B., Brodie III, E.D., 2010. Interacting phenotypes and the evolutionary process. III. Social evolution. Evolution 64, 2558–2574.

McGuigan, K., Blows, M., 2007. The phenotypic and genetic covariance structure of Drosophilid wings. Evolution 61, 902–911.

Medawar, P.B., 1952. An unsolved problem of biology. H. K. Lewis, London, UK.

Merilä, J., Sheldon, B., Kruuk, L., 2001. Explaining stasis: microevolutionary studies in natural populations. Genetica 112, 199–222.

Metz, J., Geritz, S., Meszéna, G., Jacobs, F., van Heerwaarden, J., 1996. Adaptive dynamics, a geometrical study of the consequences of nearly faithful reproduction, in: van Strien, S., Lunel, S.V. (Eds.), Stochastic and spatial structures of dynamical systems. Konink. Nederl. Akad. Wetensch. Verh. Afd. Natuurk. Eerste Reeks, Amsterdam, Netherlands, pp. 183–231.

Metz, J.A.J., 2011. Thoughts on the geometry of meso-evolution: collecting mathematical elements for a postmodern synthesis, in: Chalub, F.A.C.C., Rodrigues, J.F. (Eds.), The Mathematics of Darwin's Legacy. Springer, pp. 193–231.

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

Mezey, J.G., Houle, D., 2005. The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. Evolution 59, 1027–1038.

Moore, A.J., Brodie III, E.D., Wolf, J.B., 1997. Interacting phenotypes and the evolutionary process: I. direct and indirect genetic effects of social interactions. Evolution 51, 1352–1362.

Morrissey, M.B., 2014. Selection and evolution of causally covarying traits. Evolution 68, 1748–1761.

Morrissey, M.B., 2015. Evolutionary quantitative genetics of nonlinear developmental systems. Evolution 69, 2050–2066.

Müller, G.B., 2007. Evo-devo: extending the evolutionary synthesis. Nat. Rev. Genet. 8, 943–949.

Müller, G.B., 2010. Epigenetic innovation, in: Pigliucci, M., Müller, G.B. (Eds.), Evolution—the extended synthesis. MIT press, London, UK, pp. 307–332.

Murray, J.D., 2003. Mathematical Biology II: Spatial Models and Biomedical Applications. 3rd ed., Springer, Berlin, Germany.

Mylius, S.D., Diekmann, O., 1995. On evolutionarily stable life histories, optimization and the need to be specific about density dependence. Oikos 74, 218–224.

Nelson, R.M., Pettersson, M.E., Carlborg, Ö., 2013. A century after Fisher: time for a new paradigm in quantitative genetics. Trends Genet. 29, 669–676.

Odling-Smee, F.J., Laland, K.N., Feldman, M.W., 1996. Niche construction. Am. Nat. 147, 641–648.

Odling-Smee, F.J., Laland, K.N., Feldman, M.W., 2003. Niche Construction: the neglected process in evolution. Princeton Univ. Press, Princeton, NJ, USA.

Orr, H.A., 2005. The genetic theory of adaptation: a brief history. Nat. Rev. Genet. 6, 119–127.

Orr, H.A., Coyne, J.A., 1992. The genetics of adaptation: a reassessment. Am. Nat. 140, 725–742.

Otto, S.P., Day, T., 2007. A Biologist's Guide to Mathematical Models in Ecology and Evolution. Princeton Univ. Press, Woodstock, UK.

Paenke, I., Sendhoff, B., Kawecki, T.J., 2007. Influence of plasticity and learning on evolution under directional selection. Am. Nat. 170, E47–E58.

Paixão, T., Barton, N.H., 2016. The effect of gene interactions on the long-term response to selection. Proc. Natl. Acad. Sci. USA 113, 4422–4427.

Parvinen, K., Heino, M., Dieckmann, U., 2013. Function-valued adaptive dynamics and optimal control theory. J. Math. Biol. 67, 509–533.

Pavlicev, M., Cheverud, J.M., Wagner, G.P., 2011. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. Proc. R. Soc. B 278, 1903–1912.

Pavlicev, M., Hansen, T.F., 2011. Genotype-phenotype maps maximizing evolvability: modularity revisited. Evol. Biol. 38, 371–389.

Pavlicev, M., Wagner, G.P., 2012. A model of developmental evolution: selection, pleiotropy and compensation. Trends. Ecol. Evol. 27, 316–322.

Pigliucci, M., 2001. Phenotypic Plasticity. The Johns Hopkins Univ. Press, London, UK.

Pigliucci, M., 2007. Do we need an extended evolutionary synthesis. Evolution 61, 2743–2749.

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

Pujol, B., Blanchet, S., Charmantier, A., Danchin, E., Facon, B., Marrot, P., Roux, F., Scotti, I., Teplitsky, C., Thomson, C.E., et al., 2018. The missing response to selection in the wild. Trends Ecol. Evol. 33, 337–346.

Rao, A.V., 2009. A survey of numerical methods for optimal control. Adv. Astronaut. Sci. 135, 497–528.

Reuter, M., Keller, L., 2001. Sex ratio conflict and worker production in eusocial hymenoptera. Am. Nat. 158, 166–177.

Rice, S.H., 1990. A geometric model for the evolution of development. J. Theor. Biol. 143, 319–342.

Rice, S.H., 2002. A general population genetic theory for the evolution of developmental interactions. Proc. Natl. Acad. Sci. USA 99, 15518–15523.

Rice, S.H., 2011. The place of development in mathematical evolutionary theory. J. Exp. Zool. 314B, 1–9.

Roff, D.A., 1992. The Evolution of Life Histories. Chapman & Hall, New York, NY, USA.

Russell, S.J., Norvig, P., 2021. Artificial Intelligence: A Modern Approach. 4th, global edition ed., Pearson, Harlow, UK.

Salazar-Ciudad, I., 2021. Why call it developmental bias when it is just development? Biol. Direct 16, 3.

Salazar-Ciudad, I., Jernvall, J., 2010. A computational model of teeth and the developmental origins of morphological variation. Nature 464, 583–586.

Salazar-Ciudad, I., Marín-Riera, M., 2013. Adaptive dynamics under development-based genotype–phenotype maps. Nature 497, 361–364.

Schaffer, W.M., 1983. The application of optimal control theory to the general life history problem. Am. Nat. 121, 418–431.

Schmidt, M., Lipson, H., 2009. Distilling free-form natural laws from experimental data. Science 324, 81–85.

Service, P.M., Rose, M.R., 1985. Genetic covariation among life-history components: the effect of novel environments. Evolution 39, 943–945.

Sheth, R., Marcon, L., Bastida, M.F., Junco, M., Quintana, L., Dahn, R., Kmita, M., Sharpe, J., Ros, M.A., 2012. *Hox* genes regulate digit patterning by controlling the wavelength of a turing-type mechanism. Science 338, 1476–1480.

Slatkin, M., 2009. Epigenetic inheritance and the missing heritability problem. Genetics 182, 845–850.

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

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction. 2nd ed., The MIT Press, Cambridge, Massachusetts, USA.

Sydsæter, K., Hammond, P., Seierstad, A., Strom, A., 2008. Further Mathematics for Economic Analysis. 2nd ed., Prentice Hall.

Taylor, H.M., Gourley, R.S., Lawrence, C.E., Kaplan, R.S., 1974. Natural selection of life history attributes: an analytical approach. Theor. Popul. Biol. 5, 104–122.

Tomlin, C.J., Axelrod, J.D., 2007. Biology by numbers: mathematical modelling in developmental biology. Nat. Rev. Genet. 8, 331–340.

Townley, S., Ezard, T.H.G., 2013. A G matrix analogue to capture the cumulative effects of nongenetic inheritance. J. Evol. Biol. 26, 1234–1243.

Turelli, M., 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. Evolution 42, 1342–1347.

Turelli, M., Barton, N.H., 1994. Genetic and statistical analyses of strong selection on polygenic traits: what, me normal? Genetics 138, 913–941.

Turing, A.M., 1952. The chemical basis of morphogenesis. Phil. Trans. R. Soc. Lond. B 237, 37–72.

Uller, T., Moczek, A.P., Watson, R.A., Brakefield, P.M., Laland, K.N., 2018. Developmental bias and evolution: A regulatory network perspective. Genetics 209, 949–966.

van Tienderen, P.H., 1995. Life cycle trade-offs in matrix population models. Ecology 76, 2482–2489.

Via, S., Lande, R., 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39, 505–522.

Waddington, C.H., 1942. Canalization of development and the inheritance of acquired characters. Nature 150, 563–565.

Waddington, C.H., 1957. The Strategy of the Genes. Allen & Unwin, London, UK.

Waddington, C.H., 1959. Evolutionary adaptation. Perspect. Biol. Med. 2, 379–401.

Waddington, C.H., 1961. Genetic assimilation. Adv. Genet. 10, 257–293.

Wagner, G.P., 1984. On the eigenvalue distribution of genetic and phenotypic dispersion matrices: Evidence for a nonrandom organization of quantitative character variation. J. Math. Biol. 21, 77–95.

Wagner, G.P., 1988. The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. J. Evol. Biol. 1, 45–66.

Wagner, G.P., 1989. Multivariate mutation-selection balance with constrained pleiotropic effects. Genetics 122, 223–234.

Wagner, G.P., 1996. Homologues, natural kinds and the evolution of modularity. Am. Zool. 36, 36–43.

Wagner, G.P., Mezey, J., 2000. Modeling the evolution of genetic architecture: a continuum of alleles models with pairwise $A \times A$ epistasis. J. Theor. Biol. 203, 163–175.

Wagner, G.P., Zhang, J., 2011. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. Nat. Rev. Genet. 12, 204–213.

Walsh, B., Lynch, M., 2018. Evolution and Selection of Quantitative Traits. Oxford Univ. Press, Oxford, UK.

Watson, R.A., Wagner, G.P., Pavlicev, M., Weinreich, D.M., Mills, R., 2013. The evolution of phenotypic correlations and "developmental memory". Evolution 68, 1124–1138.

Welch, J.J., 2017. What's wrong with evolutionary biology? Biol. Philos. 32, 263–279.

West-Eberhard, M.J., 2003. Developmental Plasticity and Evolution. Oxford Univ. Press, Oxford, UK.

Wild, G., Traulsen, A., 2007. The different limits of weak selection and the evolutionary dynamics of finite populations. J. Theor. Biol. 247, 382–390.

Wolf, J.B., Frankino, W.A., Agrawal, A.F., Brodie III, E.D., Moore, A.J., 2001. Developmental interactions and the constituents of quantitative variation. Evolution 55, 232–245.