How development affects evolution

Mauricio González-Forero*[1] and Andy Gardner[1]

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

Abstract

Natural selection acts on developmentally constructed phenotypes. Ever since the modern synthesis, many researchers have called for integration of development into evolution, but there has been a lack of general mathematical frameworks explicitly integrating the two. This has likely inhibited understanding of the evolutionary effects of development. Here we use a new mathematical framework that integrates development into evolution to analyse how development affects evolution. We show that, whilst selection pushes genetic and phenotypic evolution up the fitness landscape, development determines the admissible evolutionary pathway, such that evolutionary outcomes occur at path peaks rather than landscape peaks. Changes in development can generate path peaks, triggering adaptive radiations, even on constant, single-peak landscapes. Phenotypic plasticity, niche construction, extra-genetic inheritance, and developmental bias alter the evolutionary path and hence the outcome. Thus, extra-genetic inheritance can have permanent evolutionary effects by changing development and so the evolutionary path, even if extra-genetically acquired elements are not transmitted to future generations. Selective development, whereby phenotype construction points in the adaptive direction, may induce adaptive or maladaptive evolution depending on the developmental constraints. Moreover, developmental propagation of phenotypic effects over age enables the evolution of negative senescence. Overall, we find that development plays a major evolutionary role.

Introduction

Development constructs the phenotype across the lifespan. Natural selection must then act upon what development constructs, raising the question of how development affects evolution. Addressing this question involves an integrated consideration of development and evolution. Accordingly, efforts to address it have invoked the need for an analogous synthesis to that achieved between natural selection and genetics in the modern synthesis of the early twentieth century (West-Eberhard, 2003; Pigliucci, 2007; Pigliucci and Müller, 2010; Laland *et al.*, 2014, 2015). Key early steps in the modern synthesis involved the formulation of general mathematical frameworks that integrated natural selection and genetics, namely, population and quantitative genetics. Similarly, general mathematical frameworks integrating development into evolution are likely needed to advance our understanding of the evolutionary effects of development. Large bodies of empirical and theoretical research, as well as entire scientific journals and scientific societies have been increasingly devoted to establish the effects of development on evolution from different angles (Schlosser and Wagner, 2004; Wagner, 2005; Müller, 2007; Carroll, 2008). However, despite steady growth in research efforts on this topic, general mathematical frameworks synthesizing development and evolution have remained lacking.

One of the available frameworks, namely quantitative genetics, provides some of the still fundamental understanding of how development affects evolution. Specifically, a basic principle of adaptation first partly sketched for genetic evolution (Wright, 1932, 1937; Edwards, 2000; Walsh and Lynch, 2018) and later established for phenotypic evolution (Lande, 1979; Dieckmann and Law, 1996) is that, under certain assumptions, evolution by natural selection can be seen as the climbing of a fitness landscape over phenotype space, where selection pushes evolution in the direction of steepest ascent on the landscape and genetic covariation may divert evolution in a less steep direction. Development has thus long been understood to affect evolution by shaping the genetic covariation upon which selection acts, where this covariation is quantified by the **G** matrix (Lande, 1979; Charlesworth *et al.*, 1982; Cheverud, 1984; Maynard Smith *et al.*, 1985; Klingenberg, 2010). This perspective gives development two fundamental possibilities. First, if development yields genetic variation in all directions of phenotype space (i.e., if **G** is non-singular), then development may affect the direction of evolution but evolutionary outcomes still occur at peaks of the fitness landscape. Second, if development yields no genetic variation in some directions of phenotype space (i.e., if **G** is singular), then development not only affects the direction of evolution but may also affect where the evolutionary outcomes are in the fitness landscape, possibly away from fitness landscape peaks even if there is a single fitness peak (Houle, 1991; Kirkpatrick and Lofsvold, 1992; Altenberg, 1995). The first possibility gives development a relatively minor role in evolution, as in such case development is less able to shape evolutionary outcomes. Instead, the second possibility might give development a relatively major evolutionary role, as development is more able to shape evolutionary outcomes.

Progress on establishing which of the two possibilities is the case has been difficult. Despite vigorous research effort, the answer has remained uncertain, in part due to the difficulties of empirically establishing whether **G**

*Corresponding author: mgf3@st-andrews.ac.uk

is singular (Kirkpatrick, 2009). More fundamentally, as will become apparent below, a key reason for this uncertainty may have been that equations describing phenotypic evolutionary change have until recently not been derived from an explicit account of development, but often from considerations of genetic covariation defined in terms of linear regression of trait values on gene content (Fisher, 1918; Lande, 1979). Since regression describes relationships regardless of their underlying mechanisms, this approach to genetic covariation may have limited a mechanistic understanding of the nature of genetic covariation arising from development. In principle, an explicit integration of development into evolution to yield a mechanistic understanding of genetic covariation would provide insights into the nature of the **G**-matrix.

The task of an explicit integration of development into evolution to yield a general mathematical framework would involve many complexities, which has likely contributed to the persistent lack of such attempts. Calls for explicit integration of development into evolution have recently highlighted several developmental factors — including phenotypic plasticity (West-Eberhard, 2003), niche construction (Odling-Smee *et al.*, 1996), extragenetic inheritance (Baldwin, 1896), and developmental bias (Arthur, 2004) — as possibly having important, yet unrecognised, evolutionary consequences (Laland *et al.*, 2014, 2015). Calls for synthesis between developmental and evolutionary biology have also asked for consideration of the mechanistic basis of phenotype construction, non-linear genotype-phenotype maps, gene-gene and gene-environment interactions, non-normal distributions, far-from-equilibrium evolutionary dynamics, dynamic fitness landscapes, evolution and the nature of the **G**-matrix, evolvability and epigenetics, and a variety of other complexities (Pigliucci and Schlichting, 1997; Pigliucci, 2007). Mathematical frameworks and specific mathematical models integrating some of these complexities have become available in recent decades (e.g., Dieckmann and Law, 1996; Caswell, 2001; Hansen and Wagner, 2001; Day and Bonduriansky, 2011; Mullon and Lehmann, 2018; Lande, 2019; Chantepie and Chevin, 2020; Engen and Sæther, 2021). Additionally, theoretical research has often used individual-based simulations integrating some these complexities (e.g., Salazar-Ciudad and Marín-Riera, 2013; Watson *et al.*, 2013; Jones *et al.*, 2014a; Miloco and Salazar-Ciudad, 2022), although we note it is of particular interest to obtain mathematical frameworks with the aim of further deepening insights. Despite the progress made, a synthetic mathematical framework unifying these complexities in a general and tractable manner had remained unavailable until recently.

A new mathematical framework (González-Forero and Gardner, 2021) integrates explicit development into evolution while incorporating the elements listed in the previous paragraph by building upon many advances in evolutionary modelling over the last decades (Dieckmann and Law, 1996; Caswell, 2001) (and references in González-Forero and Gardner 2021). Specifically, this framework integrates conceptual and mathematical advances from adaptive dynamics (Dieckmann and Law, 1996), matrix population models (Caswell, 2001), and optimal control theory as used in life history theory (e.g., Schaffer, 1983; Sydsæter *et al.*, 2008). The framework yields formulas that relate mechanistic descriptions of development to genetic covariation and to plastic change separate from selection. Here we use this framework to analyse how development affects evolution, which sharpens the principle of adaptation as the climbing of a fitness landscape and provides numerous insights into a wide array of long-standing questions.

Framework overview

The framework is based on standard assumptions of adaptive dynamics (Dieckmann and Law, 1996). It considers a resident (i.e., wild-type) population where individuals can be of different ages (i.e., it is age-structured), the genetic system can be haploid or diploid, reproduction is clonal for simplicity (so linkage disequilibrium is not relevant as there is no recombination), and individuals can interact socially but only with non-relatives for simplicity. Each individual has three types of age-specific traits. First, individuals have age-specific traits that are directly controlled by genes, which borrowing terminology from optimal control theory, we call control variables. As these traits are as close to genetic control as possible, we also refer to them as gene expression. Each of these traits thus measures the level of expression of a given gene product at a given age. Such gene expression level can be controlled by one or more loci given our assumption of clonal reproduction. Second, individuals have age-specific developed traits that are constructed over life (e.g., body size) depending on the abiotic and social environment experienced across life. Borrowing terminology from optimal control theory, we call developed traits state variables. As developed traits closely approach the vernacular meaning of phenotype, we refer to them as the phenotype (note gene expression is a phenotype in the classic sense used in adaptive dynamics, but since it is not a developed trait, we do not call it phenotype here). Third, individuals have age-specific environmental traits that describe the individual's local environment, which can be modified by the individual, social partners, or sources exogenous to the population. Once the resident population achieves carrying capacity, rare mutant individuals arise in the population having a marginally different level of gene expression drawn from a symmetric distribution around the resident gene expression; thus, mutation in gene expression is assumed to be rare, weak, and unbiased. Population dynamics is deterministic so the only source of stochasticity is mutation in gene expression. This means that realised variation in genetically controlled traits is solely due to mutation at the current evolutionary time step (whereas expected evolutionary change, which is our interest, depends on the expected variation over the distribution of possible mutant gene expression). Invasion of mutants establishes a new resident, thus yielding evolutionary change.

The framework uses the following notation. The gene expression level across all possible gene products for a

mutant individual of age $a \in \{1,\dots,N_\mathrm{a}\}$ is described by the vector $\mathbf{y}_a = (y_{1a},\dots,y_{N_\mathrm{c}a})^\top$ of N_c possible gene products (control variables). The phenotype of a mutant individual of age a is described by the vector $\mathbf{x}_a = (x_{1a},\dots,x_{N_\mathrm{s}a})^\top$ of N_s developed traits (state variables). The environment of a mutant individual of age a is described by the vector $\boldsymbol{\epsilon}_a = (\epsilon_{1a},\dots,\epsilon_{N_\mathrm{e}a})^\top$ of N_e mutually independent variables describing the individual's environment (e.g., ambient temperature, which the individual may adjust, such as by roosting in the shade). A mutant individual's phenotype, gene expression, and environment across all ages are described respectively by the block-column vectors $\mathbf{x} = (\mathbf{x}_1;\dots;\mathbf{x}_{N_\mathrm{a}})$, $\mathbf{y} = (\mathbf{y}_1;\dots;\mathbf{y}_{N_\mathrm{a}})$, and $\boldsymbol{\epsilon} = (\boldsymbol{\epsilon}_1;\dots;\boldsymbol{\epsilon}_{N_\mathrm{a}})$ (the semicolon indicates a linebreak). As explained below, the framework finds that two aggregates of traits have special importance. First, we aggregate gene expression and phenotype into a single vector called the "geno-phenotype" (see also Feldman and Zhivotovsky 1992 where the term phenogenotype is used). Thus, the geno-phenotype of a mutant individual of age a is described by $\mathbf{z}_a = (\mathbf{x}_a;\mathbf{y}_a)$. The mutant geno-phenotype across all ages is $\mathbf{z} = (\mathbf{x};\mathbf{y})$. Second, we similarly aggregate gene expression, phenotype, and environment into a single vector called the "geno-envo-phenotype". The geno-envo-phenotype of a mutant individual of age a is described by the block-column vector $\mathbf{m}_a = (\mathbf{z}_a;\boldsymbol{\epsilon}_a)$. The mutant geno-envo-phenotype across all ages is $\mathbf{m} = (\mathbf{z};\boldsymbol{\epsilon})$. Resident variables are denoted analogously with an overbar (e.g., $\bar{\mathbf{z}}$).

To incorporate development, we describe an individual's phenotype at a given age as a function of all her traits at the immediately preceding age and of the social environment experienced at that age. Thus, the phenotype of a mutant individual at age $a+1$ satisfies the developmental constraint

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

where the developmental map $\mathbf{g}_a(\mathbf{m}_a,\bar{\mathbf{z}}) = (g_{1a}(\mathbf{m}_a,\bar{\mathbf{z}}),\dots,g_{N_\mathrm{s}a}(\mathbf{m}_a,\bar{\mathbf{z}}))^\top$ depends on the mutant's gene expression, phenotype, and environment at age a and on the gene expression and phenotype of social partners of any age. For simplicity, we assume that the developmentally initial phenotype $\mathbf{x}_1$ is constant and that gene expression $\mathbf{y}$ is entirely determined by genes (i.e., in the terminology of optimal control theory, we assume open-loop control; Sydsæter *et al.* 2008). The developmental map describes the developmental constraints on the phenotype $\mathbf{x}$ which cannot take arbitrary values but only those satisfying Eq. (1). The developmental map can evolve and take any (differentiable) form as gene expression, phenotype, and environment evolve.

An individual's environment may depend on her own gene expression and phenotype, on the gene expression and phenotype of social partners, and on exogenous processes. The dependence of the environment on the gene expression and phenotype of herself and social partners can describe niche construction by her or her social partners. The dependence of the environment on exogenous processes can describe, for instance, eutrophication or climate change caused by members of other species. For simplicity, we assume that exogenous environmental change is slow to allow the population to achieve the carrying capacity. Thus, a mutant's environment at age a satisfies the environmental constraint

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

where the environmental map $\mathbf{h}_a(\mathbf{z}_a,\bar{\mathbf{z}},\tau) = (h_{1a}(\mathbf{z}_a,\bar{\mathbf{z}},\tau),\dots,h_{N_\mathrm{e}a}(\mathbf{z}_a,\bar{\mathbf{z}},\tau))^\top$ depends on the mutant's gene expression and phenotype at that age (e.g., from behavioural choice or environmental modification), on her social partners' gene expression and phenotypes at any age (e.g., from social niche construction), and on the evolutionary time τ due to exogenous environmental change. Consequently, the environment $\boldsymbol{\epsilon}$ can only take values allowed by the environmental map $\mathbf{h}$ describing the environmental constraints. In turn, a mutant's fertility $f_a(\mathbf{m}_a,\bar{\mathbf{m}})$ and survival probability $p_a(\mathbf{m}_a,\bar{\mathbf{m}})$ at age a depend on her gene expression, phenotype, and environment at that age, and on the gene expression, phenotype, and environment of her social partners of any age.

The developmental constraint in Eq. (1) incorporates: developmental bias, in that the phenotype $\mathbf{x}$ may be predisposed to develop in certain ways; phenotypic plasticity, in that the same "genotype" $\mathbf{y}$ can generate a different phenotype $\mathbf{x}$ under a different environment $\boldsymbol{\epsilon}$; adaptive phenotypic plasticity, for example, via somatic selection or reinforcement learning (e.g., if g_{ia} is proportional to the gradient $\partial g_{ja}/\partial x_{ia}$ of a material payoff g_{ja} earned at age a that increases fitness); niche construction, in that environmental variables depend on developmentally constructed traits of the individual and social partners; and extra-genetic inheritance, in that an individual's phenotype at age $a+1$ can be an identical or modified copy of non-relatives' phenotypes. Moreover, the developmental constraint in Eq. (1) can evolve as explained above, be non-linear, and mechanistically describe development, gene-gene interaction, and gene-environment interaction. For illustration, we allow for the possibility that the developmental map $\mathbf{g}_a$ depends on the selection gradient $\partial w/\partial \mathbf{z}_a$ (defined below), in which case we say that development is selective.

Results

Recovery of classic but insufficient results

We investigate how development as described by Eq. (1) affects evolution by using equations that yield a well-defined description of the long-term evolution of a developed phenotype $\bar{\mathbf{x}}$ as the climbing of a fitness landscape. To obtain such equations, we first recover the classic Lande (1979) equation and explain why it is generally insufficient to describe long-term evolution of developed phenotypes.

Assume for now that there is (i) no niche construction ($\partial\boldsymbol{\epsilon}^\top/\partial\mathbf{z} = \mathbf{0}$), (ii) no social development ($\mathrm{d}\mathbf{x}^\top/\mathrm{d}\bar{\mathbf{z}} = \mathbf{0}$), (iii) no exogenous environmental change ($\partial\boldsymbol{\epsilon}/\partial\tau = \mathbf{0}$), and (iv) no direct effects of gene expression on fitness ($\partial w/\partial\mathbf{y} = \mathbf{0}$) (for an explanation of our use of matrix calculus notation, see Supplementary Information section 1). Then,

González-Forero and Gardner (2021) show that the (expected) evolutionary dynamics of the phenotype satisfy

$$\frac{\mathrm{d}\bar{\mathbf{x}}}{\mathrm{d}\tau} = \kappa \mathbf{G}_{\mathbf{x}} \frac{\partial w}{\partial \mathbf{x}}, \tag{3}$$

provided that the developmental constraint in Eq. (1) is satisfied at all times (derivatives are evaluated at $\mathbf{y} = \bar{\mathbf{y}}$ throughout; we use an equal sign in Eq. (3) but it is strictly a first-order approximation). The right-hand side of Eq. (3) has three components. First, κ is a non-negative scalar proportional to the mutation rate and the carrying capacity. Second, $\partial w/\partial \mathbf{x}$ corresponds to Lande's selection gradient, which measures direct directional selection on the phenotype: it is a vector that points in the direction of steepest increase in fitness in phenotype space. Because of age-structure, a mutant's relative fitness is

$$w = \sum_{a=1}^{N_{\mathrm{a}}} w_a = \frac{1}{T} \sum_{a=1}^{N_{\mathrm{a}}} (\phi_a f_a + \pi_a p_a), \tag{4}$$

where T is generation time, and ϕ_a and π_a are the forces (Hamilton, 1966) of selection on fertility and survival at age a (given in Eqs. S4 in the Supplementary Information). The quantities T, ϕ_a, and π_a depend on resident but not mutant values. Third, Eq. (3) depends on the additive genetic covariance matrix of the phenotype

$$\mathbf{G}_{\mathbf{x}} = \mathrm{cov}[\mathbf{a}_{\mathbf{x}}, \mathbf{a}_{\mathbf{x}}] = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{y}^{\mathsf{T}}} \mathbf{G}_{\mathbf{y}} \frac{\mathrm{d}\mathbf{x}^{\mathsf{T}}}{\mathrm{d}\mathbf{y}}, \tag{5}$$

where the breeding value of $\mathbf{x}$ is defined as

$$\mathbf{a}_{\mathbf{x}} = \bar{\mathbf{x}} + \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}\mathbf{y}^{\mathsf{T}}} (\mathbf{y} - \bar{\mathbf{y}}) \tag{6}$$

and $\mathbf{G}_{\mathbf{y}} = \mathrm{cov}[\mathbf{y}, \mathbf{y}]$ is the mutational covariance matrix. Thus, the matrix $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\mathbf{y}$ corresponds to both Fisher's (1918) additive effects of allelic substitution and Wagner's (1984) developmental matrix.

Eq. (3) recovers the classic Lande (1979) equation, apart from some differences due to our adaptive dynamics assumptions but these differences do not affect the points we make. Indeed, the expression for relative fitness in Eq. (4) recovers that of Lande (1982) for quantitative genetics under age structure. The expression for the additive genetic covariance matrix of the phenotype in Eq. (5) has the same form of previous expressions obtained under quantitative genetics assumptions (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). While in quantitative genetics $\mathbf{G}_{\mathbf{y}}$ would describe the realised standing variation in gene content, here $\mathbf{G}_{\mathbf{y}}$ describes the expected variation in gene expression due only to mutation in the current evolutionary time and so we call it the mutational covariance matrix. This different meaning of $\mathbf{G}_{\mathbf{y}}$ in the two frameworks does not mean that $\mathbf{G}_{\mathbf{y}}$ in our framework induces absolute constraints that are absent under quantitative genetics assumptions: our definition of $\mathbf{G}_{\mathbf{y}}$ allows for mutational variation in all directions of gene expression space, or only in some restricted directions (i.e., $\mathbf{G}_{\mathbf{y}}$ may be full rank, or less than full rank), as in quantitative genetics. Moreover, the definition of breeding value in Eq. (6) is the first-order estimate of the phenotype with respect to gene expression as predictor, which corresponds to the quantitative genetics notion of breeding value as the best linear estimate of the phenotype with respect to gene content as predictor (Falconer and Mackay, 1996; Lynch and Walsh, 1998). Hence, gene expression plays here an analogous role to that of gene content in a quantitative genetics framework in that it emerges as the relevant first-order predictor of the phenotype to describe inheritance ($\mathbf{y}$ is a random variable with distribution $M(\mathbf{y}, \bar{\mathbf{y}})$ centered around $\bar{\mathbf{y}}$ with vanishingly small, symmetric variance; using this distribution, the narrow-sense heritability of a single developed trait x can be computed in our framework with the traditional formula $\mathrm{var}[a_x]/\mathrm{var}[x]$).

While Eq. (7) recovers the Lande (1979) equation, it has been derived from an explicit account of development so we have formulas to relate development in Eq. (1) to breeding value and so to genetic covariation $\mathbf{G}_{\mathbf{x}}$ that guarantee that the developmental constraint in Eq. (1) is satisfied at all times. Crucially, these formulas show that $\mathbf{G}_{\mathbf{x}} = \mathbf{G}_{\mathbf{x}}(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ is generally a function of resident gene expression $\bar{\mathbf{y}}$, in particular, if development is non-linear, in which case the developmental matrix $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\mathbf{y}$ depends on $\bar{\mathbf{y}}$. Specifically, gene-gene interaction (products between y's in the developmental map), gene-phenotype interaction (products between y's and x's), and gene-environment interaction (products between y's and ϵ's) cause resident gene expression $\bar{\mathbf{y}}$ to remain in $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\mathbf{y}$. Thus, in general, the right-hand side of Eq. (3) depends on resident gene expression $\bar{\mathbf{y}}$ if the mutation rate, carrying capacity, mutational covariation $\mathbf{G}_{\mathbf{y}}$, the developmental matrix $\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\mathbf{y}$, or direct phenotypic selection $\partial w/\partial \mathbf{x}$ depend on resident gene expression $\bar{\mathbf{y}}$. Consequently, Eq. (3) is under-determined as it generally depends on resident gene expression but does not describe the evolutionary dynamics of resident gene expression. Hence, Eq. (3) cannot be used alone to describe the evolutionary dynamics of the resident phenotype in general.

This problem equally applies to the classic Lande equation under quantitative genetics assumptions. Since breeding value under quantitative genetics assumptions depends on the regression coefficients of phenotype on gene content at the current time, such regression coefficients depend on the current state of the population, including allele frequency. Thus, the classic Lande equation depends on allele frequency but it does not describe allele frequency change. The standard approach to address this problem is to assume Fisher's infinitesimal model, whereby each phenotype is controlled by an arbitrarily large number of loci so in the short term allele frequency change per locus is negligible (Barton *et al.*, 2017; Hill, 2017; Walsh and Lynch, 2018). Then, $\mathbf{G}$ is assumed constant or its evolution is described by the Bulmer equation, which considers change in $\mathbf{G}$ due to change in linkage disequilibrium while still assuming negligible allele frequency change (Lande and Arnold, 1983; Turelli, 1988; Barton *et al.*, 2017). Such an approach allows for the classic Lande equation to describe evolution only in the short term, where allele frequency change per locus remains

negligible (Barton *et al.*, 2017). However, long-term evolution involves non-negligible allele frequency change, thus limiting the ability of the classic Lande equation to describe long-term phenotypic evolution.

In turn, the canonical equation of adaptive dynamics (Dieckmann and Law, 1996) describes long-term phenotype evolution, but the canonical equation does not describe long-term evolution of developed traits as an adaptive topography. Indeed, the classic canonical equation of adaptive dynamics describes long-term evolution of genetic traits that do not have developmental constraints (Dieckmann and Law, 1996). Extensions of the canonical equation (Dieckmann *et al.*, 2006) allow for developed traits (Parvinen *et al.*, 2013), but these extensions still describe the evolutionary dynamics in gradient form only for genetic traits, not developed traits. Evolutionary dynamic equations in gradient form for developed traits (i.e., with explicit developmental constraints) such as in Eq. (3) only became available recently (González-Forero and Gardner, 2021). This makes it possible to analyse how the evolution of developed traits proceeds on the fitness landscape.

Development blocks evolutionary change

We have seen that the classic Lande (1979) equation in Eq. (3) is insufficient to describe the evolutionary dynamics of developed phenotypes in general. Fortunately, such equation can be extended to allow for this.

To see this, remove the assumption (iv) above, and still assume that there is (i) no niche construction ($\partial\boldsymbol{\epsilon}^\top/\partial\mathbf{z} = \mathbf{0}$), (ii) no social development ($\mathrm{d}\mathbf{x}^\top/\mathrm{d}\bar{\mathbf{z}} = \mathbf{0}$), and (iii) no exogenous environmental change ($\partial\boldsymbol{\epsilon}/\partial\tau = \mathbf{0}$). Then, González-Forero and Gardner (2021) show that the evolutionary dynamics of the phenotype and gene expression are given by

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} = \kappa\mathbf{G}_\mathbf{z}\frac{\partial w}{\partial\mathbf{z}}, \tag{7}$$

provided that the developmental constraint in Eq. (1) is satisfied at all times. Eq. (7) now depends on the additive genetic covariance matrix of the geno-phenotype

$$\mathbf{G}_\mathbf{z} = \mathrm{cov}[\mathbf{a}_\mathbf{z}, \mathbf{a}_\mathbf{z}] = \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{y}^\top}\mathbf{G}_\mathbf{y}\frac{\mathrm{d}\mathbf{z}^\top}{\mathrm{d}\mathbf{y}}, \tag{8}$$

where the breeding value $\mathbf{a}_\mathbf{z}$ of the geno-phenotype is given by Eq. (6) replacing $\mathbf{x}$ with $\mathbf{z}$.

Eq. (7) is an extended Lande equation that, in contrast to the classic Lande equation, provides a well-defined description of long-term phenotypic evolution. Eq. (7) can do this because it describes the evolutionary dynamics of all the variables involved. In particular, the matrix $\mathbf{G}_\mathbf{z} = \mathbf{G}_\mathbf{z}(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ is still a function of the resident phenotype and gene expression, but now the evolutionary dynamics of both are described by Eq. (7). The right-hand side of Eq. (7) may also be a function of the resident environmental traits, but because of assumptions (i) and (iii), niche construction and exogenous environmental change are absent so the environment remains constant. Hence, the extended Lande equation in Eq. (7) yields a well-defined description of long-term phenotypic evolution as the climbing of a fitness landscape by simultaneously describing genetic and phenotypic evolution since $\mathbf{z} = (\mathbf{x}; \mathbf{y})$ includes both gene expression and phenotype.

Having identified an equation that fully describes the evolutionary dynamics of developed traits as an adaptive topography, it is now relevant to analyse the effect of development on genetic covariation and evolution. A biologically crucial property of Eq. (7) is that it implies that there are always absolute genetic constraints to adaptation. Indeed, to yield a well-defined description of long-term phenotypic evolution as the climbing of a fitness landscape, Eq. (7) simultaneously describes genetic and phenotypic evolution, so the climbing of the fitness landscape is in geno-phenotype space rather than only phenotype space as in the classic Lande equation. Yet, because the geno-phenotype $\mathbf{z}$ contains gene content $\mathbf{y}$, genetic covariation in geno-phenotype is absolutely constrained by the genetic space (i.e., $\mathbf{G}_\mathbf{z}$ is always singular because $\mathrm{d}\mathbf{z}^\top/\mathrm{d}\mathbf{y}$ has fewer rows than columns). Thus, there cannot be genetic variation in as many directions in geno-phenotype space as there are phenotypes across life. Along such directions evolutionary change is blocked. In other words, as developed traits have an indirect genetic basis, the ability of natural selection to shape them is constrained by the ability of genes to shape developed traits. In this sense, development can be seen as blocking evolutionary change in some directions. Therefore, evolutionary stasis ($\mathrm{d}\bar{\mathbf{z}}/\mathrm{d}\tau = \mathbf{0}$) can occur away from landscape peaks, where direct directional selection persists ($\partial w/\partial\mathbf{z} \neq \mathbf{0}$; evolutionary dynamics under absolute genetic constraints for specific scenarios have been previously analysed by Houle 1991, Kirkpatrick and Lofsvold 1992, and Altenberg 1995). The singularity of $\mathbf{G}_\mathbf{z}$ means that the evolutionary dynamics are strongly non-standard; in particular, consideration of direct directional selection alone ($\partial w/\partial\mathbf{z}$) is not sufficient for predicting possible evolutionary outcomes. Although direct directional genetic selection ($\partial w/\partial\mathbf{y}$) is traditionally assumed to be zero in quantitative genetics, life history models with explicit developmental dynamics show this is not necessarily the case (see example in González-Forero and Gardner, 2021, where gene expression modulates resource allocation to fertility and is thus under direct selection in a classic life history model). In any case, the need to use the extended Lande equation (7) persists as absence of direct directional genetic selection does not necessarily eliminate the dependence of the right-hand side of Eq. (3) on $\bar{\mathbf{y}}$ (e.g., if there is gene-gene interaction, so $\mathrm{d}\mathbf{x}^\top/\mathrm{d}\mathbf{y}$ is a function of $\bar{\mathbf{y}}$). So absolute genetic constraints (i.e., the singularity of $\mathbf{G}_\mathbf{z}$) govern the evolutionary dynamics in general even without direct directional genetic selection.

The singularity of $\mathbf{G}_\mathbf{z}$ may help explain common empirical observations of evolutionary stasis in wild populations despite directional selection and genetic variation, termed the "paradox of stasis" (Merilä *et al.*, 2001; Kirkpatrick, 2009; Kingsolver and Diamond, 2011). Because $\mathbf{G}_\mathbf{z}$ is singular, evolutionary stasis can generally occur with persistent directional selection and genetic covariation. We note that there are other explanations for the paradox

of stasis, including measurement error, but such empirical observations are as expected given that genetic constraints on the geno-phenotype are necessarily absolute.

Before moving on, we note that even though plasticity may be present in the extended Lande equation (7), plasticity has no effect under the assumptions made. Indeed, the formulas for $\mathbf{G_z}$ show that, under the framework's assumptions, plasticity (i.e., $\partial\mathbf{z}^\top/\partial\boldsymbol{\epsilon}$) only affects the evolutionary dynamics by interacting with niche construction or exogenous environmental change (see below). Thus, although in Eq. (7) there may be plasticity, it has no effect since by assumptions (i) and (iii) niche construction and exogenous environmental change are absent so the environment remains constant, but plasticity requires environmental change.

Development determines the path

While development can be seen as blocking evolutionary change, development can be more specifically seen as determining the admissible evolutionary path. This is because the developmental constraints in Eq. (1) must be satisfied at all times, thus the developmental constraints provide the admissible path. Moreover, development determines evolutionary equilibria, and which of these are admissible.

To see this, continue to assume that there is (i) no niche construction ($\partial\boldsymbol{\epsilon}^\top/\partial\mathbf{z} = \mathbf{0}$), (ii) no social development ($\mathrm{d}\mathbf{x}^\top/\mathrm{d}\bar{\mathbf{z}} = \mathbf{0}$), and (iii) no exogenous environmental change ($\partial\boldsymbol{\epsilon}/\partial\tau = \mathbf{0}$). Then, the evolutionary dynamics of the phenotype and gene expression are equivalently given by

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} = \kappa\mathbf{G_{zy}}\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}}, \tag{9}$$

again provided that the developmental constraint in Eq. (1) is satisfied at all times. This equation now depends on the total selection gradient of gene expression

$$\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}} = \frac{\mathrm{d}\mathbf{z}^\top}{\mathrm{d}\mathbf{y}}\frac{\partial w}{\partial\mathbf{z}} = \frac{\partial w}{\partial\mathbf{y}} + \frac{\mathrm{d}\mathbf{x}^\top}{\mathrm{d}\mathbf{y}}\frac{\partial w}{\partial\mathbf{x}}, \tag{10}$$

which measures total genetic selection, that is, total directional selection on genes considering the ability of genes to affect the phenotype (in contrast, $\partial w/\partial\mathbf{y}$ measures direct directional selection on genes without considering such ability). The total selection gradient of gene expression is a specific case of Caswell's (1982, 2001) total derivative of fitness, Charlesworth's (1994) total differential of the population's growth rate, van Tienderen's (1995) integrated sensitivity of the population's growth rate, and Morrissey's (2014, 2015) extended selection gradient. Moreover, Eq. (9) depends on the additive genetic cross-covariance matrix between the geno-phenotype and gene expression

$$\mathbf{G_{zy}} = \mathrm{cov}[\mathbf{a_z}, \mathbf{a_y}] = \frac{\mathrm{d}\mathbf{z}}{\mathrm{d}\mathbf{y}^\top}\mathbf{G_y}. \tag{11}$$

Crucially, the matrix $\mathrm{d}\mathbf{z}/\mathrm{d}\mathbf{y}^\top$, which maps change in gene expression to change in phenotype, is non-singular. Thus, Eq. (9) implies that evolutionary stasis occurs when total genetic selection vanishes ($\mathrm{d}w/\mathrm{d}\mathbf{y} = \mathbf{0}$) if there are no absolute mutational constraints (i.e., if $\mathbf{G_y}$ is non-singular). That is, total genetic selection can identify evolutionary equilibria, even though direct directional selection typically cannot. From Eq. (10), we find that total genetic selection depends on the matrix $\mathrm{d}\mathbf{x}^\top/\mathrm{d}\mathbf{y}$, which as previously stated corresponds to Fisher's (1918) additive effects of allelic substitution and to Wagner's (1984) developmental matrix (see Fig. S1 for a description of all the factors affecting total genetic selection). This means that evolutionary equilibria depend on development. Moreover, there are always an infinite number of evolutionary equilibria, because $\mathrm{d}w/\mathrm{d}\mathbf{y} = \mathbf{0}$ provides fewer equations than those in $\mathrm{d}\bar{\mathbf{z}}/\mathrm{d}\tau = \mathbf{0}$. But, since the developmental constraints in Eq. (1) must be satisfied at all times, then the developmental constraints determine not only the evolutionary equilibria, but also the admissible evolutionary trajectory and so the admissible equilibria.

To gain intuition, consider the following simple example. Let there be one phenotype x and one gene product with expression level y, and let age structure be collapsed such that development occurs instantaneously. Collapsing age structure is a heuristic simplification, but doing so makes $\mathbf{x}$ and $\mathbf{y}$ scalars, which can then be easily visualised. Also, let there be no social interactions, no niche construction, no exogenous environmental change, and no density dependence. By removing social interactions and density dependence, the latter of which is also a heuristic simplification, evolutionary change can be not only abstractly but also visually described as the climbing of a fitness landscape. The extended Lande equation (7) states that selection can be seen as "pushing" the geno-phenotype uphill on the fitness landscape in geno-phenotype space in the direction of steepest ascent, whereas genetic covariation diverts evolutionary change in a possibly less steep direction. If genetic constraints on the traits considered were not absolute (here, if $\mathbf{G_z}$ were non-singular), the population would eventually reach a landscape peak, which is commonly implicitly or explicitly assumed. Yet, as the traits considered must be both the "genotype" and the phenotype arising from the developmental constraint in Eq. (1), genetic constraints are necessarily absolute (i.e., $\mathbf{G_z}$ is singular). Hence, evolution is restricted to an admissible path on the landscape where the developmental constraint is met (Fig. 1a-c; the computer code used to generate all the figures is in the Supplementary Information). Adaptive evolution may thus be understood as the climbing of the fitness landscape along an admissible path determined by development. The evolutionary process eventually reaches a path peak if there are no absolute mutational constraints (Fig. 1a-c). Formally, a path peak is a point in geno-phenotype space where the developmental constraints are met and fitness is locally and totally maximized with respect to gene expression: here "totally" means after substituting in the fitness function for both the developmental and environmental constraints. Selection response vanishes at path peaks which are not necessarily landscape peaks (Fig. 1a-c). The admissible path yields an elevation profile of fitness, namely, the total fitness landscape of gene expression, on which adaptation is constrained to occur, and

can have peaks and valleys created by the developmental map (Fig. 1d-f). Hence, selection pushes gene expression uphill on the total fitness landscape of gene expression, but in this profiled landscape evolutionary change is not blocked provided that there are no absolute mutational constraints. Consequently, without absolute mutational constraints, evolutionary stasis generally occurs at a peak on the total fitness landscape of gene expression, even though this does not generally correspond to a peak on the fitness landscape of the phenotype and gene expression. The $\mathbf{G_z}$-matrix evolves as an emergent property as the resident phenotype and gene expression evolve, and genetic variances and covariances may substantially increase or decrease if development is non-linear (consistent with previous individual-based simulations; Miloco and Salazar-Ciudad, 2022), rather than being approximately constant as is often considered under short-term evolution with an infinite number of loci (Fisher, 1918; Barton *et al.*, 2017) which we do not assume (Fig. 1g-i). Ultimately, development constrains the path of adaptation and thus shapes its outcome (Fig. 1j-l). Overall, development has a permanent dual role in adaptation by determining both evolutionary equilibria (blue lines in Fig. 1j-l) and the admissible evolutionary path (red line in Fig. 1j-l).

That evolution stops at path peaks rather than landscape peaks may help explain abundant empirical data. Empirical estimation of selection has found that direct stabilizing selection is rare relative to direct directional selection across thousands of estimates in more than 50 species (Kingsolver *et al.*, 2001; Kingsolver and Diamond, 2011). The rarity of stabilizing selection has been puzzling under the common assumption that evolutionary outcomes occur at landscape peaks, where stabilizing selection is prevalent (Charlesworth *et al.*, 1982; Walsh and Lynch, 2018). In contrast, the rarity of stabilizing selection is consistent with evolutionary outcomes occurring at path peaks. Indeed, if path peaks occur outside landscape peaks, evolutionary outcomes occur with persistent direct directional selection (Houle, 1991; Kirkpatrick and Lofsvold, 1992; Altenberg, 1995) and relatively weak direct stabilizing selection (because direct stabilizing selection is relatively strong only near landscape peaks). Thus, evolution stopping at path peaks outside landscape peaks may help explain the otherwise puzzling observation of rare stabilizing selection. Again, we note that there are other explanations for the rarity of stabilizing selection, some of which are closely related to ours (Morrissey, 2015).

Development can drive diversification

The evolution of phenotypic diversity is typically explained in terms of selection or drift (Gavrilets and Losos, 2009), but the relevance of development has been less clear. We find that development can drive the evolution of phenotypic diversity, even in a constant, single-peak fitness landscape, where selection provides an evolutionary force while development translates it into diversification. This can happen in two ways. First, if the developmental map changes inducing a shifted path peak, the

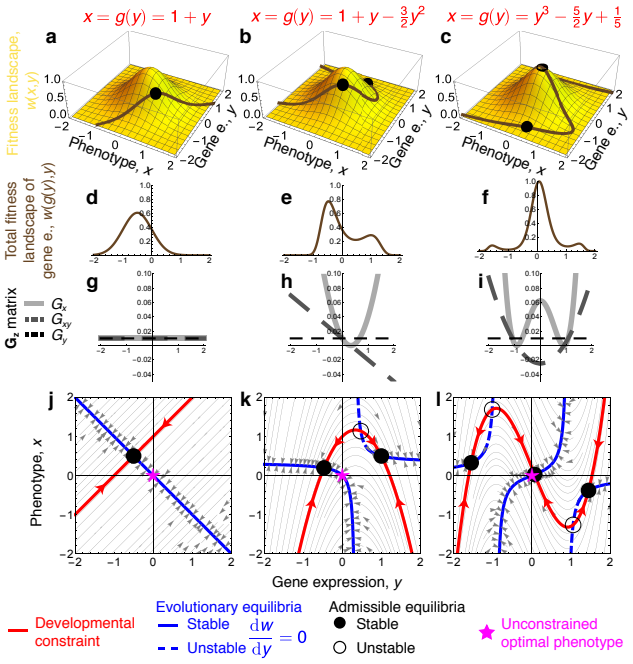

Figure 1: Development determines the evolutionary path. **a-c,** Fitness landscape for a phenotype x and a gene product's expression level y, with three developmental maps. The developmental map determines the path on the landscape as x can only take values along the path. **d-f,** The total fitness landscape of gene expression gives the elevation profile considering developmental and environmental constraints. **g-i,** Additive genetic covariances depend on the developmental map and evolve as gene expression evolves. **j-l,** The evolutionary dynamics occur along the developmental constraint (red; gray arrows are parallel or antiparallel to leading eigenvectors of $\mathbf{G_z}$, called "genetic lines of least resistance"; Schluter, 1996). Evolutionary outcomes (black dots) occur at path peaks and thus depend on development. Evolutionary equilibria (blue) are infinite in number because $\mathbf{G_z}$ is singular and occur when total genetic selection vanishes despite persistent direct selection. The intersection of the developmental map and evolutionary equilibria yields the admissible evolutionary equilibria. Different developmental maps yield different evolutionary trajectories and outcomes with the same single-peak landscape. Developmental bias is quantified by the slope of the red line ($\partial x/\partial y$). Evolutionary change satisfies $d\bar{\mathbf{z}}/d\tau = \mathbf{G_z}\partial w/\partial \mathbf{z}$, where $\mathbf{G_z} = (d\mathbf{z}/dy)G_y(d\mathbf{z}^\intercal/dy) = (dx/dy, 1)^\intercal G_y(dx/dy, 1) = G_y((dx/dy)^2, dx/dy; dx/dy, 1) = (G_x, G_{xy}; G_{xy}, G_y)$, which is singular. Hence, the additive genetic variance of x is $G_y(dx/dy)^2$. The total effects of a mutant's gene expression on her geno-phenotype are $d\mathbf{z}^\intercal/dy = (dx/dy, dy/dy) = (dx/dy, 1)$. Fitness is $w(x, y) = \exp(-x^2 - y^2)$ and mutational variance is $G_y = 0.01$ (so that $0 \leq \mathrm{tr}(\mathbf{G_y}) \ll 1$ as assumed by the framework). As we use our heuristic assumptions of collapsed age structure and no density dependence, we use here a simpler fitness function than that in Eq. (4). Functional forms are chosen for illustration.

population can evolve different phenotypes and gene expression (Fig. 2a-f). Second, the developmental map may change, generating new path peaks, if it is or becomes non-linear, such that the population may evolve different phenotypes and gene expression depending on the initial conditions (Fig. 2g-l). Mathematically, multiple path peaks can arise from non-linear development because it can generate bifurcations altering the number of admissible evolutionary equilibria even though the number of evolutionary equilibria remains the same (i.e., infinite). This is a form of evolutionary branching (Geritz *et al.*, 1998), which may be called evo-devo branching, where the admissible path creates a valley in the total fitness landscape of gene expression so there is disruptive total genetic selection (Fig. 1e; see also Morrissey 2015 where stabilizing total selection is analysed). There may not be a valley in the fitness landscape so no disruptive direct selection on the phenotype or gene expression (Fig. 1b) as evo-devo branching only requires disruptive total genetic selection possibly created by development rather than direct disruptive selection. Thus, evo-devo branching may occur in situations where evolutionary branching driven by direct disruptive selection would not occur. Hence, development can lead to phenotypic diversification, even with a constant single-peak fitness landscape in geno-phenotype or phenotype space. Consequently, adaptive radiations might arise from the evolution of development.

Our results partly substantiate a classic explanation of punctuated equilibria in terms of developmental constraints, an explanation that has been dismissed in the light of previous theory. Punctuated equilibria refer to long periods of morphological stasis alternated by periods of rapid change as observed in the fossil record (Eldredge and Gould, 1972; Hunt *et al.*, 2015). The dominant explanation for punctuated equilibria is that fitness landscapes remain constant for long periods so there is stabilizing selection at an unconstrained optimum, and fitness landscapes subsequently undergo change (Charlesworth *et al.*, 1982; Walsh and Lynch, 2018). A classic alternative explanation is that (a) developmental constraints prevent change for long periods and (b) revolutions in the developmental program bring sudden change (Gould, 1980). The ability of constraints to prevent change has been thought to be refuted by the common observation of selection response, sometimes under seeming developmental constraints (Beldade *et al.*, 2002), while revolutions in the developmental program have been refuted on other grounds (Charlesworth *et al.*, 1982; Eldredge *et al.*, 2005; Klingenberg, 2010). Our results substantiate the former, showing that selection response is consistent with developmental constraints preventing change: indeed, even if developmental constraints halt a population at a path peak, selection response is possible since path peaks may change, either by change in the path (e.g., by plasticity, see below) or the landscape (e.g., by artificial selection). Thus, available evidence does not necessarily rule out that developmental constraints can prevent change for long periods in some directions. Punctuated equilibria may then arise by change in the developmental map, rather than only by change of the fitness landscape. In particular, gradual change in the developmental map

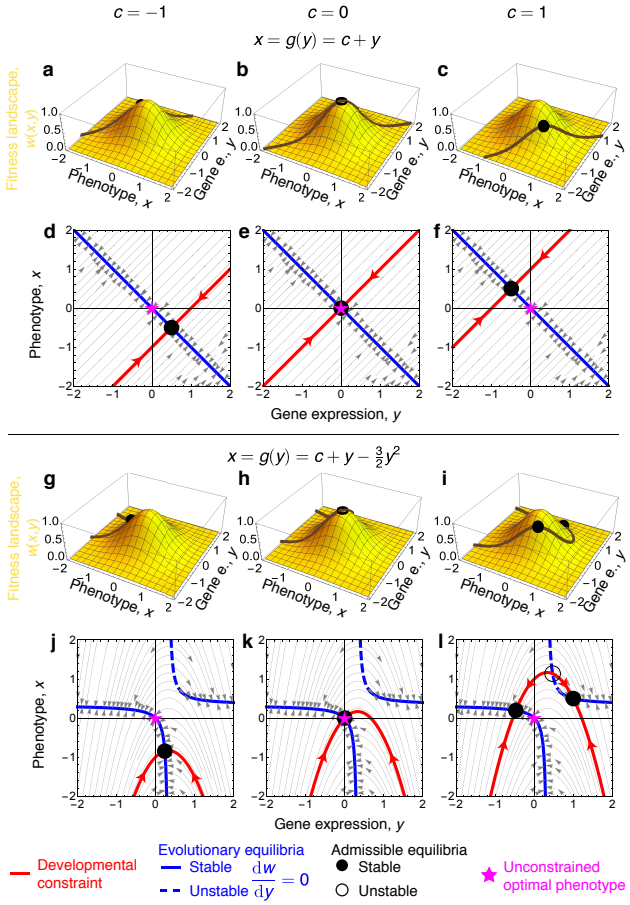

Figure 2: Development-driven diversification. Development may drive phenotypic diversification in two ways. First, by shifting path peaks (**a**-**f**). Second, by creating path peaks (**g**-**l**). The evolutionary dynamics are given as in the legend of Fig. 1. As c changes, development g changes, but total developmental bias $\mathrm{d}x/\mathrm{d}y$ does not change, so the evolutionary equilibria remain the same (blue); however, the admissible evolutionary equilibria do change (open and solid circles) and so does the outcome even though the fitness landscape is constant and single peaked.

may yield sudden path-peak creation events (Fig. 2g-l; see also Fig. 3.4 of Metz 2011). Hence, punctuated equilibria may arise by gradual evolution of the developmental map, without revolutions in the developmental program.

Niche construction reshapes the path

Niche construction — whereby organisms affect their environment — has been suggested to be a developmental factor having major evolutionary effects (Odling-Smee *et al.*, 1996; Laland *et al.*, 2014, 2015). How does niche construction affect evolution? We find that niche construction has a dual effect on adaptation. First, niche construction ($\partial\boldsymbol{\epsilon}^\top/\partial\mathbf{z}$) reshapes the fitness landscape as previously known (Odling-Smee *et al.*, 1996; Laland *et al.*, 2014, 2015). Second, niche construction reshapes the path on the landscape.

To see this, we relax the assumption that niche construction is absent ($\partial\boldsymbol{\epsilon}^\top/\partial\mathbf{z} \neq \mathbf{0}$), whilst maintaining our

assumptions of (ii) no social development ($\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}} = \mathbf{0}$), and (iii) no exogenous environmental change ($\partial\boldsymbol{\epsilon}/\partial\tau = \mathbf{0}$). The evolutionary dynamics of the resident phenotype and gene expression are then given by

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} = \kappa\mathbf{G}_{\mathbf{z}}\frac{\delta w}{\delta \mathbf{z}}, \tag{12}$$

again provided that the developmental constraint in Eq. (1) is met at all times. Eq. (12) depends on the semi-total selection gradient of the phenotype and gene expression

$$\frac{\delta w}{\delta \mathbf{z}} = \frac{\partial w}{\partial \mathbf{z}} + \frac{\partial \boldsymbol{\epsilon}^{\mathsf{T}}}{\partial \mathbf{z}}\frac{\partial w}{\partial \boldsymbol{\epsilon}}, \tag{13}$$

which measures semi-total selection considering environmental constraints but not developmental constraints. So, selection pushes uphill, no longer on the fitness landscape that ignores environmental constraints, but on the reshaped semi-total fitness landscape that considers such constraints (Fig. 3a-e). Eq. (13) shows that such reshaping of the fitness landscape is done by the interaction of niche construction and environmental sensitivity of selection ($\partial w/\partial\boldsymbol{\epsilon}$) (a term coined by Chevin *et al.* 2010 for a more specific notion). The semi-total fitness landscape can have new peaks and valleys relative to the fitness landscape. Additionally, Eq. (12) depends on $\mathbf{G}_{\mathbf{z}}$, which still has the form in Eq. (8) so it is still singular but now depends on the interaction of niche construction and plasticity ($\partial\mathbf{x}^{\mathsf{T}}/\partial\boldsymbol{\epsilon}$). Thus, development determines the path on the semi-total fitness landscape, but now this path is reshaped by the interaction of niche construction and plasticity (Fig. 3f-j). Consequently, niche construction reshapes the fitness landscape by interacting with environmental sensitivity of selection, and reshapes the admissible path on the landscape by interacting with phenotypic plasticity.

Before moving on, we note that Eq. (12) could also be seen as corresponding to the extended Lande equation describing genetic and phenotypic evolution where environmental variables are not explicitly considered in a quantitative genetics analysis. Now, Eq. (12) generally depends on the resident environment $\bar{\boldsymbol{\epsilon}}$, which evolves due to niche construction. Provided that the environmental map is known, the environment is then known so the Eq. (12) still yields a well-defined description of long-term phenotypic evolution despite environmental evolution. Yet, if one is interested in understanding how the evolution of the environment proceeds on the fitness landscape, an equation describing the evolution of the environment in gradient form would be needed. González-Forero and Gardner (2021) derive such an equation, showing it has the form of a further extended Lande equation describing the evolutionary dynamics of the geno-envo-phenotype with an associated additive genetic covariance matrix that is also necessarily singular for the same reasons given above.

Social development alters the outcome

Extra-genetic inheritance, including social learning and epigenetic inheritance, has also been suggested to be a developmental factor having major evolutionary effects (Baldwin, 1896; Laland *et al.*, 2014, 2015). Current understanding is that extra-genetic inheritance has only transient evolutionary effects if elements so inherited are transmitted over only a few generations (e.g., Walsh and Lynch, 2018; Beltran *et al.*, 2020). For instance, the permanent evolutionary effects discussed by Day and Bonduriansky (2011) and Bonduriansky and Day (2018) involve the co-evolution of genetic and extra-genetic elements, which requires transmission of extra-genetic elements over multiple generations. In contrast, we find that extra-genetic inheritance generally has permanent evolutionary effects even without transmission of extra-genetic elements over multiple generations.

Extra-genetic inheritance can be seen as a particular form of social development ($\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}} \neq \mathbf{0}$), where the developed phenotype depends on the phenotype and gene expression of social partners. For instance, acquiring a skill or epimark from a role model or a parent, respectively, involves an interaction between the recipient and donor, and these extra-genetic elements affect the subsequent phenotype of the individual (although recall we assume that social interactions are only with non-relatives for simplicity). Social development can also mechanistically describe both indirect genetic effects, where the genes of social partners affect an individual's phenotype (Moore *et al.*, 1997), and phenotypic effects that need not stem from genes (e.g., if the individual's developed traits depend on social partners' environmentally induced developed traits). How does social development, including extra-genetic inheritance, affect evolution? We find that, since it alters the developmental map, social development alters the evolutionary outcome (i.e., the admissible evolutionary equilibria) by altering the admissible evolutionary path without necessarily altering the evolutionary equilibria themselves. Thus, extra-genetically inherited elements can have permanent evolutionary effects without being transmitted over many generations. What is needed is that the sensitivity to acquire the element extra-genetically is maintained over many generations so that a socially altered development continuous to occur over many generations. To see this, consider the following.

With social development, individuals with the same genotype experiencing the same environment may develop different phenotypes even though development is assumed to be deterministic. This introduces a complication to evolutionary analysis, which we handle with the following notions. We say a gene expression, phenotype, and environment $\mathbf{m}^{**} = (\mathbf{x}^{**};\mathbf{y}^{**};\boldsymbol{\epsilon}^{**})$ is a socio-devo equilibrium if and only if an individual with such gene expression and environment developing in the context of individuals having such gene expression, phenotype, and environment develops their same phenotype. A socio-devo equilibrium is socio-devo stable if and only if socio-devo stabilization dynamics are locally stable to small perturbations in the phenotype (González-Forero and Gardner, 2021). Although social development can yield socio-devo unstable phenotypes, our analyses apply only to socio-devo stable phenotypes (which are guaranteed to be so if all the eigenvalues of $\mathrm{d}\mathbf{x}/\mathrm{d}\bar{\mathbf{x}}^{\mathsf{T}}$ have absolute

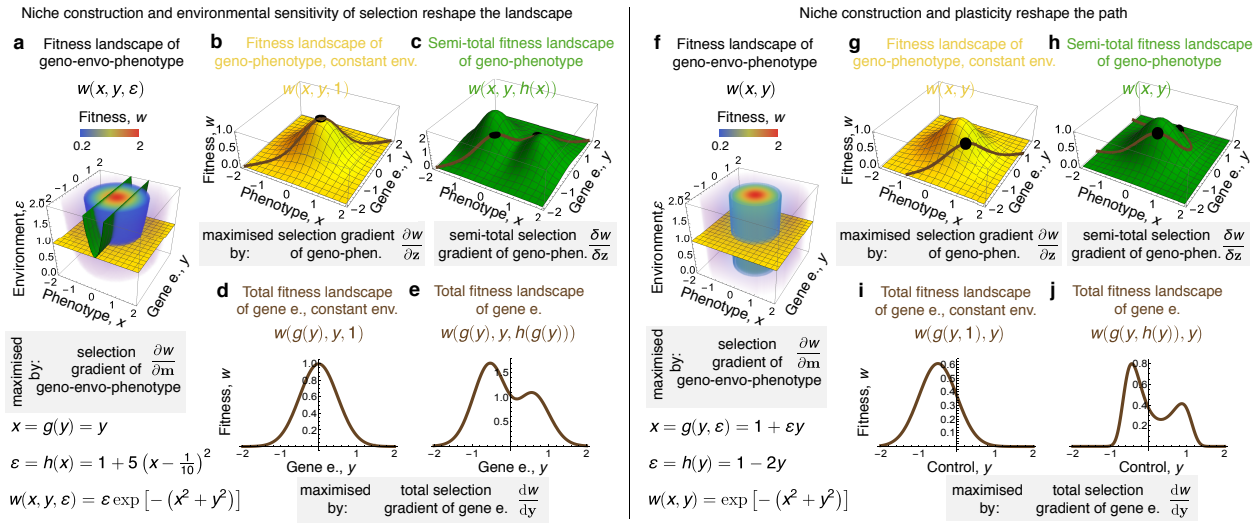

Figure 3: Niche construction reshapes the fitness landscape and the evolutionary path. **a-e,** There is niche construction by the phenotype x which non-linearly affects the environmental variable ϵ, which in turn increases fitness. **a,** Fitness landscape vs. phenotype, gene expression, and environment. Slices for constant environment (yellow) and for the environmental constraint (green) are respectively shown in **b** and **c**. The interaction of niche construction ($\partial\epsilon/\partial x$) and environmental sensitivity of selection ($\partial w/\partial\epsilon$) reshapes the landscape (from **b** to **c**) by affecting semi-total selection $\delta w/\delta\mathbf{z}$ but it does not affect genetic covariation $\mathbf{G_z}$. **f-j,** There is niche construction by gene expression y which linearly affects the environmental variable ϵ, which in turn increases the phenotype. The interaction of niche construction ($\partial\epsilon/\partial y$) and plasticity ($\partial x/\partial\epsilon$) reshapes the path (from **g** to **h**). This interaction affects genetic covariation $\mathbf{G_z}$ but does not affect semi-total selection $\delta w/\delta\mathbf{z}$. In all panels, evolutionary change satisfies $\mathrm{d}\bar{\mathbf{z}}/\mathrm{d}\tau = \mathbf{G_z}\delta w/\delta\mathbf{z}$, where $\delta w/\delta\mathbf{z} = \partial w/\partial\mathbf{z} + (\partial\epsilon/\partial\mathbf{z})(\partial w/\partial\epsilon)$ and $\mathbf{G_z}$ is given in Fig. 1. In **a-e,** $\mathrm{d}x/\mathrm{d}y = \partial x/\partial y$, whereas in **f-j,** $\mathrm{d}x/\mathrm{d}y = \partial x/\partial y + (\partial\epsilon/\partial y)(\partial x/\partial\epsilon)$.

value strictly less than one, which we assume).

Thus, for socio-devo stable phenotypes, their evolutionary dynamics are described by the following equations. Allowing for niche construction ($\partial\boldsymbol{\epsilon}^{\mathsf{T}}/\partial\mathbf{z} \neq \mathbf{0}$) and social development ($\mathrm{d}\mathbf{x}^{\mathsf{T}}/\mathrm{d}\bar{\mathbf{z}} \neq \mathbf{0}$), but assuming (iii) that there is no exogenous environmental change ($\partial\boldsymbol{\epsilon}/\partial\tau = \mathbf{0}$), then the evolutionary dynamics of the resident phenotype and gene expression are given by

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} = \kappa\mathbf{H_z}\frac{\delta w}{\delta\mathbf{z}} = \kappa\mathbf{H_{zy}}\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}}, \tag{14}$$

again provided that the developmental constraint in Eq. (1) is met at all times. This equation now depends on the additive socio-genetic cross-covariance matrix of the geno-phenotype

$$\mathbf{H_z} = \mathrm{cov}[\mathbf{b_z}, \mathbf{a_z}] = \frac{\mathrm{s}\mathbf{z}}{\mathrm{s}\mathbf{y}^{\mathsf{T}}}\mathbf{G_y}\frac{\mathrm{d}\mathbf{z}^{\mathsf{T}}}{\mathrm{d}\mathbf{y}}, \tag{15}$$

where the stabilized breeding value of $\mathbf{z}$ is defined as

$$\mathbf{b_z} = \bar{\mathbf{z}} + \frac{\mathrm{s}\mathbf{z}}{\mathrm{s}\mathbf{y}^{\mathsf{T}}}(\mathbf{y} - \bar{\mathbf{y}}). \tag{16}$$

The matrix $\mathrm{s}\mathbf{z}/\mathrm{s}\mathbf{y}^{\mathsf{T}}$ is the stabilized total derivative of $\mathbf{z}$ with respect to $\mathbf{y}^{\mathsf{T}}$, which measures the total effects of $\mathbf{y}^{\mathsf{T}}$ on $\mathbf{z}$ after the effects of social development have stabilized in the population; $\mathrm{s}\mathbf{z}/\mathrm{s}\mathbf{y}^{\mathsf{T}}$ reduces to $\mathrm{d}\mathbf{z}/\mathrm{d}\mathbf{y}^{\mathsf{T}}$ if development is not social. Thus, the $\mathbf{H_z}$-matrix generalizes $\mathbf{G_z}$ by including the effects of social development ($\mathbf{H_z}$ is analogous to another generalization of the $\mathbf{G}$-matrix used in the indirect genetic effects literature; Moore *et al.*, 1997). As for $\mathbf{G_z}$, the matrix $\mathbf{H_z}$ is always singular because the geno-phenotype contains gene content (i.e., because $\mathrm{d}\mathbf{z}^{\mathsf{T}}/\mathrm{d}\mathbf{y}$ has fewer rows than columns). Hence, also under social development, evolutionary stasis can generally occur with persistent semi-total selection on the phenotype and gene expression. In turn, the second equality in Eq. (14) depends on the additive socio-genetic cross-covariance matrix between the geno-phenotype and gene expression

$$\mathbf{H_{zy}} = \mathrm{cov}[\mathbf{b_z}, \mathbf{a_y}] = \frac{\mathrm{s}\mathbf{z}}{\mathrm{s}\mathbf{y}^{\mathsf{T}}}\mathbf{G_y}, \tag{17}$$

which generalizes the corresponding matrix $\mathbf{G_{zy}}$ we had in Eq. (11) to consider social development. Crucially and similarly to $\mathbf{G_{zy}}$, this matrix $\mathbf{H_{zy}}$ is non-singular if $\mathbf{G_y}$ is non-singular. Thus, evolutionary equilibria still occur when total genetic selection vanishes ($\mathrm{d}w/\mathrm{d}\mathbf{y} = \mathbf{0}$) if there are no absolute mutational constraints.

Having these equations we can now see that social development, including extra-genetic inheritance, may alter both the evolutionary equilibria (by altering $\mathrm{d}w/\mathrm{d}\mathbf{y}$) and the admissible path (by altering $\mathbf{H_{zy}}$; Fig. 4). Note that $\bar{y}$ in Fig. 4**a**,**b** can be interpreted as being an extra-genetically acquired element that modifies development, but such acquired effect is not transmitted to future generations as the individual transmits the y it has. Yet, such extra-genetically acquired element still alters the evolutionary outcome relative to that without extra-genetic inheritance (Fig. 1**a**,**d**,**g**,**j**). Alternatively, social development may not affect the evolutionary equilibria if total genetic selection $\mathrm{d}w/\mathrm{d}\mathbf{y}$ is independent of social partners' phenotype $\bar{\mathbf{z}}$, but social development may still affect the ad-

missible evolutionary path given by the developmental constraint in Eq. (1) (by altering only $\mathbf{H_{zy}}$; Fig. 5). Similarly, $\bar{y}$ in Fig. 5**a**,**b** can also be interpreted as being an extra-genetically acquired element that modifies development, but it is not transmitted to future generations and still alters the evolutionary outcome relative to that without extra-genetic inheritance (Fig. 1**b**,**e**,**h**,**k**). Indeed, in both cases, social development (e.g., extra-genetic inheritance) affects the evolutionary path and so the admissible evolutionary equilibria and hence the evolutionary outcome. Mathematically, this results because $\mathbf{H_z}$ is singular, so changing the path generally changes the outcome. Thus, because $\mathbf{H_z}$ is singular, extra-genetic inheritance can have permanent evolutionary effects even if extra-genetically acquired elements are not transmitted to other individuals.

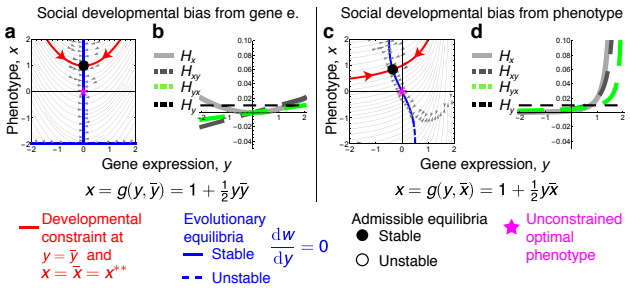

Figure 4: Social development alters the evolutionary outcome. **a-b,** There is social development, particularly social developmental bias from gene expression ($\partial x/\partial \bar{y} \neq 0$). Here social development introduces a non-linearity relative to the developmental map of Fig. 1**a**,**d**,**g**,**j**, which changes the evolutionary equilibria (blue), admissible path (red), and outcome (solid dot). **c-d,** As in **a-b,** but there is social developmental bias from the phenotype ($\partial x/\partial \bar{x} \neq 0$). The developmental map is that of **a-b** replacing $\bar{y}$ with $\bar{x}$. This also changes the evolutionary equilibria, admissible path, and outcome. In all panels, evolutionary change satisfies $\mathrm{d}\bar{\mathbf{z}}/\mathrm{d}\tau = \mathbf{H_z}\partial w/\partial \mathbf{z}$, where $w(x,y) = \exp(-x^2 - y^2)$ and $\mathbf{H_z} = (\mathrm{s}\mathbf{z}/\mathrm{s}y)G_y(\mathrm{d}\mathbf{z}^\top/\mathrm{d}y) = (H_{xx}, H_{xy}; H_{yx}, H_{yy})$, with $\mathrm{s}\mathbf{z}/\mathrm{s}y = (\mathrm{s}x/\mathrm{s}y; 1)$. For **a-b**, $\mathrm{s}x/\mathrm{s}y = \mathrm{d}x/\mathrm{d}y + \mathrm{d}x/\mathrm{d}\bar{y}$ (from Layer 5, Eq. 2a in González-Forero and Gardner 2021). Similarly, for **c-d**, $\mathrm{s}x/\mathrm{s}y = (1 - \partial x/\partial \bar{x})^{-1}\mathrm{d}x/\mathrm{d}y$, with $\mathrm{d}\mathbf{z}^\top/\mathrm{d}y$ and $\mathrm{s}x/\mathrm{s}y$ evaluated at the socio-devo equilibrium x^{**} that solves $x^{**} = g(y, x^{**})$. Since development is social, a fitness landscape visualization as in Fig. 1a is not possible. In **c**, the stream plot (gray arrows) is only locally accurate around the developmental constraint where the socio-devo equilibrium holds. The 1/2 in the developmental constraint is used to meet the assumption $|\mathrm{d}x/\mathrm{d}\bar{x}| < 1$ but does not affect the points made.

Before moving on, we note that when social development alters only the admissible path, it does so because of social developmental bias ($\partial \mathbf{x}^\top/\partial \bar{\mathbf{z}}$; e.g., extra-genetic inheritance and indirect-genetic effects), and/or the interaction of social niche construction ($\partial \boldsymbol{\epsilon}^\top/\partial \bar{\mathbf{z}}$) and plasticity ($\partial \mathbf{x}^\top/\partial \boldsymbol{\epsilon}$) (from Layer 3, Eq. 3 of González-Forero and Gardner 2021).

Plastic response alters the path

Research on the evolutionary effects of plasticity has increasingly intensified, with a particular focus on whether adaptive plastic change precedes genetic evolution (West-Eberhard, 2003). How does plasticity affect evolution? We find that plasticity also has a dual evolutionary effect. First, we have seen that plasticity ($\partial \mathbf{x}^\top/\partial \boldsymbol{\epsilon}$) alters the evolutionary path by interacting with niche construction (i.e., endogenous environmental change; $\partial \boldsymbol{\epsilon}^\top/\partial \mathbf{z}$). This is because niche construction alters the environment, which through plasticity alters the developed phenotype and so the evolutionary path. Second, plasticity alters the evolutionary path by interacting with exogenous environmental change (e.g., eutrophication or climate change caused by other species; $\partial \boldsymbol{\epsilon}/\partial \tau$). This is similarly because exogenous environmental change alters the environment, which through plasticity alters the developed phenotype and so the evolutionary path. With exogenous environmental change, the evolutionary dynamics comprise selection response and exogenous plastic response (see also Chevin *et al.*, 2010), the latter of which modifies the developmental map and thus the evolutionary path as evolutionary time advances. Exogenous plastic response has some ability to increase adaptation separate from selection response via selective development, although this ability is limited in various ways.

To see this, assume now that there is niche construction ($\partial \boldsymbol{\epsilon}^\top/\partial \mathbf{z} \neq \mathbf{0}$), social development ($\mathrm{d}\mathbf{x}^\top/\mathrm{d}\bar{\mathbf{z}} \neq \mathbf{0}$), and exogenous environmental change ($\partial \boldsymbol{\epsilon}/\partial \tau \neq \mathbf{0}$). Then, the evolutionary dynamics of the resident phenotype and gene expression are given by

$$\frac{\mathrm{d}\bar{\mathbf{z}}}{\mathrm{d}\tau} = \kappa \mathbf{H_z}\frac{\delta w}{\delta \mathbf{z}} + \frac{\mathrm{s}\mathbf{z}}{\mathrm{s}\boldsymbol{\epsilon}^\top}\frac{\partial \boldsymbol{\epsilon}}{\partial \tau} = \kappa \mathbf{H_{zy}}\frac{\mathrm{d}w}{\mathrm{d}\mathbf{y}} + \frac{\mathrm{s}\mathbf{z}}{\mathrm{s}\boldsymbol{\epsilon}^\top}\frac{\partial \boldsymbol{\epsilon}}{\partial \tau}. \tag{18}$$

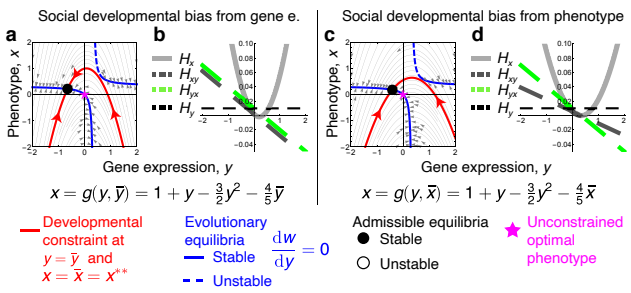

Figure 5: Social development can alter the evolutionary outcome without altering evolutionary equilibria. The plots are as in Fig. 4, except that now the developmental maps are a linear modification of those in Fig. 1**b**,**e**,**h**,**k**. Relative to that figure, social development does not alter total developmental bias $\mathrm{d}x/\mathrm{d}y$ so it does not affect evolutionary equilibria (blue). However, social development affects the admissible path (red) and so the evolutionary outcome (solid dot). Note H_x can be negative in **b** as it gives the covariance between the stabilized breeding value of x and the breeding value of x (in contrast, G_x gives the variance of the breeding value of x, which is non-negative). The 4/5 in the developmental constraint is used to meet the assumption $|\mathrm{d}x/\mathrm{d}\bar{x}| < 1$ but does not affect the points made.

The matrix $\mathrm{s}\mathbf{z}/\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}}$ describes what we call stabilized plasticity of the phenotype and gene expression. Stabilized plasticity includes the effects of socio-devo stabilization and reduces to total plasticity $\mathrm{d}\mathbf{z}/\mathrm{d}\boldsymbol{\epsilon}^{\mathsf{T}}$ if development is non-social. The first term in Eq. (18) is selection response ($\kappa\mathbf{H}_{\mathbf{z}}\delta w/\delta\mathbf{z}$) whereas the second term is exogenous plastic response ($(\mathrm{s}\mathbf{z}/\mathrm{s}\boldsymbol{\epsilon}^{\mathsf{T}})(\partial\boldsymbol{\epsilon}/\partial\tau)$), which makes the developmental map evolve as the environment exogenously changes (Fig. 6). Through exogenous plastic response, the evolutionary effects of development go beyond inducing genetic covariation. Moreover, Eq. (18) shows that exogenous plastic response can induce selection response, which allows for plastic change to preceed genetic change (West-Eberhard, 2003). For example, if the population is at a path peak so there is no selection response ($\mathbf{H}_{\mathbf{z}}\delta w/\delta\mathbf{z} = \mathbf{0}$) and there is exogenous plastic response, this generally changes the path thus inducing future selection response (Fig. 6**a**-**c**). In the idealized case where development is selective in the sense that the developmental map is a function of the selection gradient, stabilized plasticity can depend on the selection gradient. Mechanistically, selective development might occur to some extent with individual or social learning, or with somatic selection (West-Eberhard, 2003). With selective development, exogenous plastic response can induce evolution either towards or away landscape peaks depending on the developmental constraints and initial conditions, by trapping evolution in local peaks of the total fitness landscape of gene expression (Fig. 6). Exogenous plastic response with selective development has a limited ability of inducing evolution towards a landscape peak since plasticity must depend on developmental-past environments, which induces a developmental lag in the plastic response (Supplementary Information section 2). Additionally, exogenous plastic response with selective development may induce evolution away a landscape peak when the rate of environmental change switches sign (Supplementary Information section 2).

Development enables negative senescence

Thus far, we have considered evolutionary effects of development that occur even without considering that development takes time. Indeed, in our illustrative examples we have let age structure be collapsed so development happens instantaneously. We now show that explicit consideration of age progression in development may explain an otherwise puzzling ageing pattern.

Individuals may show senescence as they age, that is, decreasing survival or fertility after the onset of reproduction. Leading hypotheses for the evolution of senescence are based on the fact that the forces of selection decline with age (ϕ_a and π_a in Eq. (4), and Eqs. S8 in the Supplementary Information; Hamilton, 1966). For example, if a mutation has a beneficial effect on survival or fertility early in life and a pleiotropic, deleterious effect of similar magnitude late in life, the mutation would be favoured (Medawar, 1952; Williams, 1957). The universality of declining forces of selection has suggested that senescence should be universal (Hamilton, 1966), at least for organisms with a bodily reset at conception (Lehtonen, 2020).

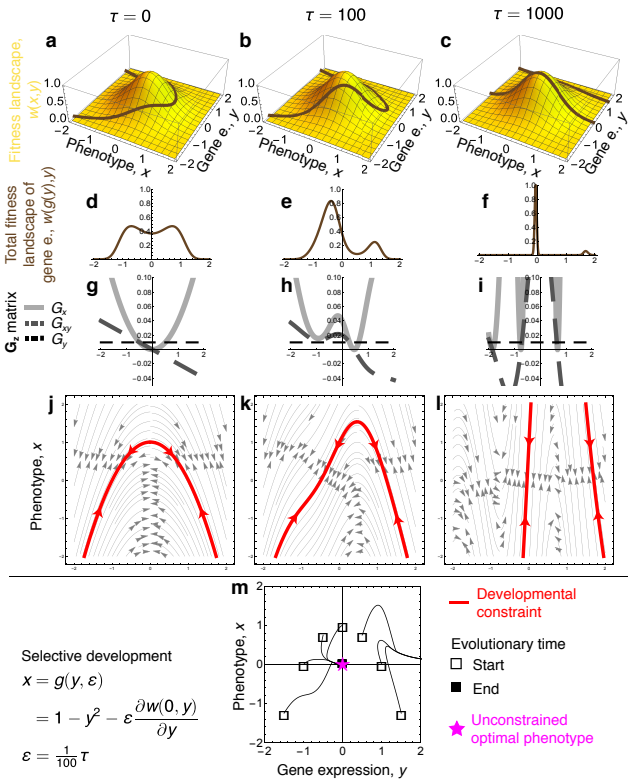

Figure 6: Exogenous plastic response changes the evolutionary path. We let there be exogenous environmental change and plasticity, but no niche construction, so plasticity does not affect $\mathbf{G}_{\mathbf{z}}$ but only the exogenous plastic response. Development is selective by letting plasticity equal the selection gradient at the optimal phenotype, $\partial x/\partial\epsilon = \partial w(0,y)/\partial y$. **a-l,** Exogenous environmental change induces exogenous plastic response which raises part of the path on the landscape. **m,** Resulting evolutionary dynamics showing that selective development induces some evolutionary trajectories to converge to the landscape peak, whereas others do not depending on their initial condition due to the quadratic developmental constraint. Evolutionary change satisfies $\mathrm{d}\bar{\mathbf{z}}/\mathrm{d}\tau = \mathbf{G}_{\mathbf{z}}\partial w/\partial\mathbf{z} + (\mathrm{d}\mathbf{z}/\mathrm{d}\epsilon)(\partial\epsilon/\partial\tau)$, where $\mathrm{d}\mathbf{z}/\mathrm{d}\epsilon = (\mathrm{d}x/\mathrm{d}\epsilon; \mathrm{d}y/\mathrm{d}\epsilon) = (\partial x/\partial\epsilon; 0)$ and $\mathbf{G}_{\mathbf{z}}$ is given in Fig. 1. Fitness is $w(x,y) = \exp(-x^2 - y^2)$.

However, empirical research has increasingly reported organisms with negative senescence (Jones *et al.*, 2014b), that is, where survival or fertility increase with age after the onset of reproduction, and models have suggested that this may stem from indeterminate growth (Vaupel *et al.*, 2004; Lehtonen, 2020).

More fundamentally, we find that developmental propagation — whereby an early-acting mutation has a large phenotypic effect late in life — can drive the evolution of negative senescence. For example, if a mutation in gene expression has a deleterious effect on survival or fertility at an early age a, so that $\delta w_a/\delta y_{ia} < 0$, and a pleiotropic, beneficial effect on survival or fertility of similar magnitude at a later age $j > a$, so that $\delta w_j/\delta x_{kj} > 0$ for some phenotype, then semi-total selection is weaker at the later age because of declining selection forces

(i.e., $|\delta w_a/\delta y_{ia}| \geq |\delta w_j/\delta x_{kj}|$). Yet, crucially, such late-beneficial mutation is still favoured if total genetic selection on it is positive, that is, if

$$\frac{\mathrm{d}w}{\mathrm{d}y_{ia}} = \frac{\delta w_a}{\delta y_{ia}} + \frac{\mathrm{d}x_{kj}}{\mathrm{d}y_{ia}}\frac{\delta w_j}{\delta x_{kj}} > 0 \tag{19}$$

(as derived in Supplementary Information section 3), where the effect of the early mutation in gene expression on the late phenotype is

$$\frac{\mathrm{d}x_{kj}}{\mathrm{d}y_{ia}} = \frac{\delta \mathbf{x}_{a+1}^{\mathsf{T}}}{\delta y_{ia}}\left(\prod_{i=a+1}^{j-2}\frac{\delta \mathbf{x}_{i+1}^{\mathsf{T}}}{\delta \mathbf{x}_i}\right)\frac{\delta x_{kj}}{\delta \mathbf{x}_{j-1}} \tag{20}$$

(from Eq. E17 of González-Forero and Gardner 2021). This effect of the early mutation in gene expression on the late phenotype may increase (or decrease) with the phenotype age j because of developmental propagation of phenotypic effects as age advances, a propagation that is described by the term in parentheses in Eq. (20): in particular, indeterminate growth allows for indefinite developmental propagation. So total genetic selection may increase with the age j at which the pleiotropic effect is expressed. Consequently, a mutation with late beneficial effects may be favoured despite early deleterious effects. Thus, developmental propagation may enable selection to shape traits at advanced age, allowing for adaptive evolution of increasing survival or fertility as ageing proceeds despite decreasing selection forces after the onset of reproduction.

Discussion

Evolution by natural selection can be understood as the climbing of a fitness landscape under some assumptions. This was first established for genetic evolution (Wright, 1937). It was later established for phenotypic evolution under the tacit assumption that there is negligible genetic evolution (Lande, 1976, 1979; Barton *et al.*, 2017; Hill, 2017; Walsh and Lynch, 2018). However, long-term phenotypic evolution involves non-negligible genetic evolution. We have shown that understanding long-term phenotypic evolution as the climbing of a fitness landscape requires that genetic and phenotypic evolution are considered simultaneously. Doing so necessarily makes the associated $\mathbf{G}$ (or $\mathbf{H}$) matrix singular because the geno-phenotype contains gene content. Thus, long-term phenotypic evolution necessarily involves absolute genetic constraints, where the evolutionary path is given by the developmental process that associates genes to phenotype.

Consequently, we have shown that development affects evolution by determining the evolutionary path. This sharpens the basic principle of adaptation as the climbing of the fitness landscape (Lande, 1979) (sensu Dieckmann and Law 1996), such that evolutionary outcomes occur at best at peaks on the path rather than on the landscape. Developmental constraints thus stop evolution at path peaks that can happen mid-way on the fitness landscape. Selective development, if feasible, can bring some of the path closer to the landscape peak by means of exogenous plastic response. Our findings substantiate several although not all intuitions previously given in relation to the role of development in evolution (Waddington, 1959; Gould and Lewontin, 1979; West-Eberhard, 2003; Pigliucci, 2007; Laland *et al.*, 2014, 2015). In particular, seemingly disparate development-related factors previously suggested to be evolutionarily important (Waddington, 1959; Gould and Lewontin, 1979; West-Eberhard, 2003; Pigliucci, 2007; Laland *et al.*, 2014, 2015) variously alter the evolutionary path and hence also the outcome. Our analysis offers answers to various major questions, namely the origin of diversity, the punctuated fossil record, the paradox of stasis, the rarity of stabilizing selection, and the origin of negative senescence. Our findings entail that an understanding of development is instrumental for evolutionary understanding: change in development alone changes evolutionary outcomes, equilibria and the admissible path by changing absolute socio-genetic constraints ($\mathbf{H}_{\mathbf{z}}$), even if direct selection ($\partial w/\partial \mathbf{z}$), total selection ($\mathrm{d}w/\mathrm{d}\mathbf{z}$ or $\mathrm{d}w/\mathrm{d}\mathbf{y}$), or the non-social environment ($\boldsymbol{\epsilon}$) remain constant (Fig. 5). These results show that development has major evolutionary consequences.

Acknowledgements

We thank M.W. Feldman, K.N. Laland, R. Lande, L.C. Mikula, A.J. Moore, M.B. Morrissey, and one anonymous reviewer for comments on previous versions of the manuscript, and D.M. Shuker for feedback. This work was funded by an ERC Consolidator Grant to AG (grant no. 771387) and by the School of Biology of the University of St Andrews. AG was also funded by a NERC Independent Research Fellowship (grant no. NE/K009524/1).

References

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

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

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

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

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

Beldade, P., Koops, K. and Brakefield, P.M. (2002). Developmental constraints versus flexibility in morphological evolution. *Nature*, **416**, 844–847.

Beltran, T., Shahrezaei, V., Katju, V. and Sarkies, P. (2020). Epimutations driven by small RNAs arise frequently but most have limited duration in *Caenorhabditis elegans*. *Nat. Ecol. Evol.*, **4**, 1539–1548.

Bonduriansky, R. and Day, T. (2018). *Extended Heredity*. Princeton Univ. Press, Princeton, NJ, USA.

Carroll, S.B. (2008). Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell*, **134**, 25–36.

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*. Sinauer, Sunderland, MA, USA, 2nd edition.

Chantepie, S. and Chevin, L.M. (2020). How does the strength of selection influence genetic correlations? *Evol. Lett.*, **4-6**, 468–478.

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

Charlesworth, B. (1994). *Evolution in age-structured populations*. Cambridge Univ. Press, 2nd edition.

Charlesworth, B., Lande, R. and 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. and Mace, G.M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *PLOS Biology*, **8**(4), 1–8.

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

Dieckmann, U. and Law, R. (1996). The dynamical theory of coevolution: a derivation from stochastic ecological processes. *J. Math. Biol.*, **34**(5), 579–612.

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

Edwards, A.W.F. (2000). Sewall wright's equation $\delta q = (q(1-q)\partial w/\partial q)/2w$. *Theor. Popul. Biol.*, **57**, 67–70.

Eldredge, N. and Gould, S.J. (1972). Punctuated equilibria: an alternative to phyletic gradualism. In J. M. Thomas, editor, *Models in Paleobiology*, pages 82–115. Freeman, Cooper and Company.

Eldredge, N., Thompson, J.N., Brakefield, P.M., Gavrilets, S., Jablonski, D., Jeremy et al (2005). The dynamics of evolutionary stasis. *Paleobiology*, **31**, 133–145.

Engen, S. and 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. and Mackay, T.F.C. (1996). *Introduction to Quantitative Genetics*. Pearson Prentice Hall, Harlow, England, 4th edition.

Feldman, M.W. and Zhivotovsky, L.A. (1992). Gene-culture coevolution: toward a general theory of vertical transmission. *Proc. Natl. Acad. Sci. USA*, **89**, 11935–11938.

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

Gavrilets, S. and Losos, J.B. (2009). Adaptive radiation: Contrasting theory with data. *Science*, **323**, 732–737.

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

González-Forero, M. and Gardner, A. (2021). A mathematical framework for evo-devo dynamics. In review at *Theor. Popul. Biol.* Preprint: https://www.biorxiv.org/content/10.1101/2021.05.17.444499v2.

Gould, S.J. (1980). Is a new and general theory of evolution emerging? *Paleobiology*, **6**, 119–130.

Gould, S.J. and 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.

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

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

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.

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

Hunt, G., Hopkins, M.J. and Lidgard, S. (2015). Simple versus complex models of trait evolution and stasis as a response to environmental change. *Proc. Natl. Acad. Sci. USA*, **112**, 4885–4890.

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

Jones, O.R., Scheuerlein, A., Salguero-Gómez, R., Camarda, C.G., Schaible, R., Casper, B.B. et al (2014b). Diversity of ageing across the tree of life. *Nature*, **505**, 169–173.

Kingsolver, J.G. and 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. et al (2001). The strength of phenotypic selection in natural populations. *Am. Nat.*, **157**, 245–261.

Kirkpatrick, M. (2009). Patterns of quantitative genetic variation in multiple dimensions. *Genetica*, **136**, 271–284.

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

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

Laland, K., Uller, T., Feldman, M., Sterelny, K., Müller, G.B., Moczek, A. et al (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. et al (2015). The extended evolutionary synthesis: its structure, assumptions and predictions. *Proc. R. Soc. B*, **282**, 20151019.

Lande, R. (1976). Natural selection and random genetic drift in phenotypic evolution. *Evolution*, **30**, 314–334.

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. (2019). Developmental integration and evolution of labile plasticity in a complex quantitative character in a multiperiodic environment. *Proc. Natl. Acad. Sci. USA*, **116**, 11361–11369.

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

Lehtonen, J. (2020). Longevity and the drift barrier: bridging the gap between Medawar and Hamilton. *Evol. Lett.*, **4**, 382–393.

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

Maynard Smith, J., Burian, R., Kauffman, S., Alberch, P., Campbell, J., Goodwin, B. et al (1985). Developmental constraints and evolution. *Q. Rev. Biol.*

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

Merilä, J., Sheldon, B. and Kruuk, L. (2001). Explaining stasis: microevolutionary studies in natural populations. *Genetica*, **112**(1), 199–222.

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

Miloco, L. and Salazar-Ciudad, I. (2022). Evolution of the G matrix under nonlinear genotype-phenotype maps. *Am. Nat.*, **199**, 420–435.

Moore, A.J., Brodie III, E.D. and Wolf, J.B. (1997). Interacting phenotypes and the evolutionary process: I. direct and indirect genetic effects of social interactions. *Evolution*, **51**(5), 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.

Mullon, C. and Lehmann, L. (2018). Eco-evolutionary dynamics in metacommunities: Ecological inheritance, helping within species, and harming between species. *Am. Nat.*, **192**, 664–686.

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

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

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

Pigliucci, M. and Müller, G.B., editors (2010). *Evolution: The Extended Synthesis*. The MIT Press, Cambridge, MA, USA.

Pigliucci, M. and Schlichting, C.D. (1997). On the limits of quantitative genetics for the study of phenotypic evolution. *Acta Biotheor.*, **45**, 143–160.

Salazar-Ciudad, I. and 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.

Schlosser, G. and Wagner, G.P., editors (2004). *Modularity in Development and Evolution*. The Univ. of Chicago Press, Chicago, IL, USA.

Schluter, D. (1996). Adaptive radiation along genetic lines of least resistance. *Evolution*, **50**, 1766–1774.

Sydsæter, K., Hammond, P., Seierstad, A. and Strom, A. (2008). *Further Mathematics for Economic Analysis*. Prentice Hall, 2nd edition.

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

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

Vaupel, J.W., Baudisch, A., Dölling, M., Roach, D.A. and Gampe, J. (2004). The case for negative senescence. *Theor. Pop. Biol.*, **65**, 339–351.

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

Wagner, A. (2005). *Robustness and Evolvability in Living Systems*. Princeton Univ. Press, Princeton, NJ, USA.

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. (1989). Multivariate mutation-selection balance with constrained pleiotropic effects. *Genetics*, **122**, 223–234.

Walsh, B. and 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. and Mills, R. (2013). The evolution of phenotypic correlations and "developmental memory". *Evolution*, **68**, 1124–1138.

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

Williams, G.C. (1957). Natural selection, and the evolution of senescence. *Evolution*, **11**, 398–411.

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the Sixth International Congress of Genetics*, **1**, 356–366.

Wright, S. (1937). The distribution of gene frequencies in populations. *Proc. Natl. Acad. Sci. USA*, **23**, 307–320.