Cryptic selection forces and dynamic heritability in generalized phenotypic evolution

William Gilpin [a],[*], Marcus W. Feldman [b]

[a] *Department of Applied Physics, Stanford University, Stanford, CA, United States*
[b] *Department of Biology, Stanford University, Stanford, CA, United States*

ARTICLE INFO

ABSTRACT

Individuals with different phenotypes can have widely-varying responses to natural selection, yet many classical approaches to evolutionary dynamics emphasize only how a population's average phenotype increases in fitness over time. However, recent experimental results have produced examples of populations that have multiple fitness peaks, or that experience frequency-dependence that affects the direction and strength of selection on certain individuals. Here, we extend classical fitness gradient formulations of natural selection in order to describe the dynamics of a phenotype distribution in terms of its moments—such as the mean, variance, and skewness. The number of governing equations in our model can be adjusted in order to capture different degrees of detail about the population. We compare our simplified model to direct Wright–Fisher simulations of evolution in several canonical fitness landscapes, and we find that our model provides a low-dimensional description of complex dynamics not typically explained by classical theory, such as cryptic selection forces due to selection on trait ranges, time-variation of the heritability, and nonlinear responses to stabilizing or disruptive selection due to asymmetric trait distributions. In addition to providing a framework for extending general understanding of common qualitative concepts in phenotypic evolution – such as fitness gradients, selection pressures, and heritability – our approach has practical importance for studying evolution in contexts in which genetic analysis is infeasible.

1. Introduction

The effects of evolutionary forces may be apparent in natural populations even when their underlying genetic consequences are not known. The size of river guppies increases when their natural predators are depleted (Reznick et al., 1997); the beaks of Darwin's finches grow larger after droughts (Grant and Grant, 1995); and mammals grow smaller in response to climate change (Gingerich, 2003). These and other natural and experimental studies demonstrate that rapid selection can produce noticeable changes in specific traits, underscoring the importance of considering phenotypic models of natural selection. These models are particularly relevant to studies in the field or of the fossil record where genetic analysis is unavailable or infeasible (Gingerich, 1976).

A widely-used framework for such theories is the fitness landscape (Wright, 1932), an abstract function that describes the collective survival or reproductive benefits conferred by a given phenotype: an evolving population typically approaches a locally or globally maximum value in this space, subject to constraints on its rate of adaptation. But the underlying dynamics of this process may depend strongly on the context (Mustonen and Lässig, 2009), and molecular techniques have only recently begun to shed light on the individual steps of adaptation and the intermediate phenotypes they produce (Poelwijk et al., 2007; Hartl, 2014). Moreover, macroscale analyses have produced examples of non-monotone fitness functions with elaborate, multipeaked topographies subject to strong frequency dependence (Martin and Wainwright, 2013), selection that acts on trait ranges in addition to average values (Shen et al., 2008), and selection forces with nonlinear effects on metric traits (Miles, 2004). These studies indicate the need for simple, analytical models that can provide heuristic insight into the complex dynamics of phenotypic adaptation. Of particular importance are cases in which traditional experimental metrics, such as the realized heritability or selection response, have complex time-dependence over long timescales (Björklund et al., 2013; Steppan et al., 2002), rendering derived experimental quantities such as selection coefficients insufficient as large-scale descriptors of population dynamics.

Many classical approaches to phenotypic evolutionary theory readily describe the evolution of a bell-shaped and fixed-width trait distribution within a given fitness landscape, resulting in the widely-known result that the rate of adaptation of the mean trait is directly proportional to the gradient of the mean fitness (Arnold,

* Corresponding author.
E-mail address: wgilpin@stanford.edu (W. Gilpin).

1992; Lande, 1976; Abrams and Matsuda, 1997; Cortez and Weitz, 2014). This depiction of evolution as a fitness gradient-climbing problem is particularly intuitive, and it mirrors earlier work on the adaptive landscape of individual genes (Wright, 1932). However, many phenotypes fail to satisfy these conditions (Karlin, 1987; Shen et al., 2008; Bolnick et al., 2011), as a result of which many experimental and theoretical studies have noted strong limitations on the applicability of fitness gradient dynamics (Waxman and Gavrilets, 2005; Lynch and Walsh, 1998; Barton and Polechová, 2005; Turelli and Barton, 1994). Efforts to establish more general rules for the evolution of an arbitrary trait distribution typically reformulate the underlying mathematics in terms of agent-based rules or a stochastic transmission kernel (Nowak and Sigmund, 2004; Sato and Kaneko, 2007; Sasaki and Ellner, 1995; Prügel-Bennett, 1997); however, these alternative formulations are difficult to compare directly with classical fitness gradient dynamics, and they typically introduce new assumptions regarding the underlying genetic processes or functional forms of the trait distribution or fitness.

Here, we develop a general model of phenotypic evolution that seeks to reduce functional constraints on the fitness or trait by using a moment series expansion of the fitness landscape. Our approach has its origin in classical approaches that describe evolution of the mean trait as a fitness gradient-climbing problem, but we add additional dynamical equations for the variance, skewness, kurtosis, and finer-scale statistical features of the trait distribution. By adding or removing dynamical equations that describe various moments of the trait distribution, the level of detail our model captures about the evolutionary dynamics may be tuned. Importantly, our model reduces to a classical fitness gradient model when only the mean trait is allowed to vary. Our model explicitly relates the topology of the fitness landscape to the timescales (and thus dynamical relevance) of various phenomena through a series of coupling constants, which we compute analytically for several canonical fitness landscapes in a series of demonstrative examples. Using these examples, we show how even a simple generalization of classical fitness gradient dynamics can lead to a series of surprising effects in the evolutionary dynamics, including cryptic forces of selection that cause changes in the trait distribution even when the local fitness gradient is zero, as well as suppression of disruptive selection due to asymmetry and skewness in the trait distribution. We also show that our model can be re-formulated in terms of the coupled dynamics of the narrow-sense heritability and the mean fitness, and we show how these two quantities may jointly evolve under various conditions.

2. Model

Our model is based on a series expansion of the trait distribution in terms of its moments (such as the mean trait, trait variance, and trait skewness). These quantities correspond to population-level statistical observables for a natural population, such as average height and height variation, which have dynamics that may be recorded across time in lieu of the full trait distribution. Under our approach, each moment has a separate dynamical equation, which is coupled to the remaining moments. For example, variance in the range of traits (such as height variation among a subgroup) may affect how the mean height changes over time; likewise, asymmetry (skew) in the height distribution can affect the dynamics of the trait variance. The strength and magnitude of the coupling of each moment's dynamics to others depend on the specific fitness landscape, and can be summarized in terms of a set of "coupling coefficients" for the landscape. Our basic approach is shown schematically in Fig. 1, and below we describe how we derive dynamical equations for the mean trait $\dot{\bar{z}}$, trait range $\dot{\sigma}$, and mean fitness $\dot{\bar{w}}$. We also describe below the assumptions of our approach, as well as the relationship between our approach and series-based approaches developed by others.

2.1. Trait mean dynamics in an arbitrary phenotype distribution

We follow the classical approach to deriving phenotypic evolution originally developed by Lande (Lande, 1976). We start with the breeder's equation, which relates the dynamics of the trait mean $\bar{z}$ to its change after a period of selection

$$\dot{\bar{z}} = \frac{V}{\sigma^2}(\bar{z}_w - \bar{z}), \tag{1}$$

where the mean trait value $\bar{z}$, mean fitness $\overline{W}$, and mean trait after selection $\bar{z}_w$ are defined in terms of statistical averages over the entire population,

$$\bar{z} = \int zp(z)dz, \qquad \bar{z}_w = \int zp(z)\frac{W(z)}{\overline{W}}dz,$$
$$\overline{W} = \int W(z)p(z)dz. \tag{2}$$

Here $p(z)$ is the frequency of the trait z in the population, $W(z)$ is an individual's fitness as a function of its trait. These integrals, as well as all integrals hereafter, are assumed to be taken over the full trait range $z \in (-\infty, \infty)$. The prefactor V/σ^2 in Eq. (1) is equivalent to the narrow-sense heritability h^2 of the trait, with V representing the additive genetic variance. Eq. (1) incorporates the primary assumptions regarding the underlying genetics; namely, that there is no direct gene-environment interaction, and that there is a linear regression between the selection differential and the mean trait value over short timescales (Lande, 1979; Bulmer, 1971) (although the slope of this regression, h^2, need not be constant). In many previous studies in which Eq. (1) appears, the phenotypic variance σ^2 is also assumed to be fixed, either due to logarithmic ranges in the values of metric traits that suppress the magnitude of fluctuations in trait variance, or due to the assumption of a fixed trait distribution shape that becomes a Gaussian distribution after an appropriate nonlinear transformation (Lande, 1979, 1976). We will relax this requirement here, in order to generalize the phenotypic dynamics equations originally derived by Lande (Lande, 1976).

Constancy of genetic variance V must be decided empirically for a given system under study; in general, V may change due to drift or direct gene-environment interactions (Bulmer, 1971; Turelli, 1988; Steppan et al., 2002). However, in the model of phenotypic evolution presented here, as in similar approaches based on the dynamics of single continuous traits (Bulmer et al., 1980; Falconer, 1981; Jones et al., 2003), V may be assumed to vary much more slowly than the phenotypic variance σ^2 provided that there is linkage equilibrium, weak selection, or a near-equilibrium trait distribution. Additionally, for cases in which V changes quickly, such as during periods of strong directional selection or from transient effects such as the accumulation of linkage disequilibrium, rapid compensatory effects (for example, recombination) can allow V to stabilize at a well-defined average value over long timescales (Björklund et al., 2013; Reznick et al., 1997). Accordingly, low or stationary genetic variance has been reported in certain experimental systems that may satisfy these conditions (Garant et al., 2008).

Here, we use a standardized trait distribution $p((z - \bar{z})/\sigma)$ and also parametrize the fitness as $W(z - c\bar{z})$, where c is a positive constant dependent the system being studied. Any arbitrary fitness landscape may be expressed in this form via a Taylor expansion around the mean trait; however for some fitness landscapes this series cannot be truncated, and so below we focus on the commonly-studied cases of continuous, metric traits for which truncation is permissible (Bulmer et al., 1980; Falconer, 1981). We note that the two cases $c = 0$ and $c = 1$ are the most common in previous models: $c = 0$ corresponds to a trait that directly affects survival independently of other individuals in the population (examples include metabolic efficiency, camouflage coloration, or
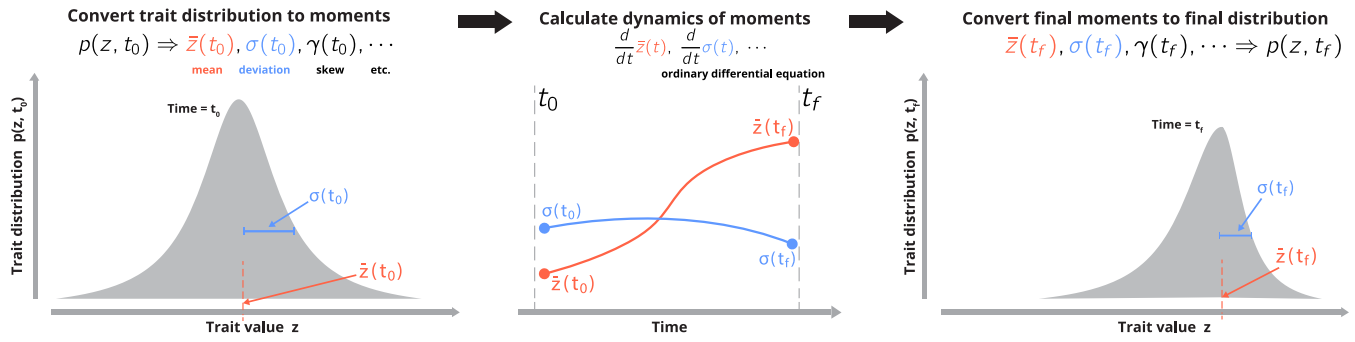
Fig. 1. Overview of modeling approach. (Left) An initial distribution of traits in a population is parametrized into an arbitrary number of moments (the mean trait, trait range, skew, kurtosis, etc.), which, collectively, describe the full population at the initial time t_0. The modeled univariate trait is abstract, but can be thought of as a metric feature that affects individual fitness (e.g. height, fin length, walking speed). (Middle) Using the breeder's equation, a known fitness landscape, and a series expansion of the trait distribution, a set of ordinary differential equations is derived that describes the time-evolution of each moment of the trait distribution, with the moment values in the initial trait distribution acting as initial conditions. (Right) The solutions of these differential equations at some later time may then be used to reconstruct an estimate of the full trait distribution at that time.

immune system competency), whereas $c = 1$ corresponds to a trait that affects fitness of a given individual only relative to others in the population (examples include secondary sex characteristics or running speed relative to a herd). Using this form of the fitness function, taking the derivative of both sides of Eq. (2) with respect to $\bar{z}$ yields a simple relationship between an individual's fitness landscape and the population mean fitness,

$$\frac{d\overline{W}}{d\bar{z}} = (1 - c) \int \frac{dp(z)}{d\bar{z}} W(z) dz. \tag{3}$$

See Supplementary Appendix B for full derivation. We note that if $c = 1$, then $\frac{d\overline{W}}{d\bar{z}} = 0$; thus when the fitness landscape depends only on the difference between each individual trait and the trait mean, the mean fitness itself cannot depend on the mean trait.

Next, we assume that the trait distribution before selection, $p(z)$, is separable into a Gaussian part $\mathcal{N}[\bar{z}, \sigma](z)$ and a dimensionless non-Gaussian component f, which without loss of generality we express in terms of the dimensionless standardized coordinate $\tilde{z} = (z - \bar{z})/\sigma$,

$$p(z) = \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}} \right) f(z) = \mathcal{N}[\bar{z}, \sigma](z) f\left(\frac{z - \bar{z}}{\sigma} \right). \tag{4}$$

We note that $f = 1$ corresponds to a purely Gaussian trait distribution, in which case our theory recreates standard evolutionary dynamics (Lande, 1976; Abrams and Matsuda, 1997; Cortez and Weitz, 2014; Slatkin, 1970).

Inserting Eqs. (3) and (4) into Eq. (1), we arrive at a dynamical equation for the mean trait of a non-Gaussian trait distribution,

$$\dot{\bar{z}} = \frac{V}{\overline{W}} \left(\frac{1}{1 - c} \frac{d\overline{W}}{d\bar{z}} \right) + \frac{V}{\sigma \overline{W}} \int \mathcal{N}[\bar{z}, \sigma](z) f'(\tilde{z}) W(z) dz. \tag{5}$$

where each term arises from the individual terms in Eq. (3) followed by integration by parts, during which a surface term vanishes due to compactness of $p(z)$ (see Supplementary Appendix B). Asymptotic analysis confirms that the first term vanishes when $c = 1$ due to a first-order zero in $\frac{d\overline{W}}{d\bar{z}}$ at $c = 1$ (see Supplementary Appendix C). The first term in Eq. (5) corresponds to a classical fitness gradient dynamics model, and represents the complete dynamics if the trait distribution is Gaussian ($f' = 0$) (Cortez and Weitz, 2014; Gilpin and Feldman, 2017). The second term determines how the higher-order moments of the trait distribution affect the evolutionary dynamics. Importantly, this new term depends explicitly on the values of higher-order moments at each timestep. Therefore, Eq. (5) can only be used to determine the dynamics if additional differential equations are specified for the trait variance, skewness, etc., which makes full determination of

the dynamics a "moment closure" problem because the dynamics of the nth moment will, in general, depend on the $(n + 1)^{th}$ moment (Whittle, 1957; Smerlak and Youssef, 2017). However, Eq. (5) can still be used to find the mean trait equilibrium and its stability as functions of the other moments.

We next expand the trait distribution in Eq. (4) as a Gram–Charlier A series,

$$p(z) = \mathcal{N}[\bar{z}, \sigma](z) \left(1 + \sum_{n=3}^{\infty} c_n He_n \left(\frac{z - \bar{z}}{\sigma} \right) \right), \tag{6}$$

where the expansion coefficients c_n are uniquely determined by the moments of the trait distribution (see Supplementary Appendix A). We use a Gram–Charlier series, as opposed to other series, in order to formulate our model as a perturbation from classical evolutionary dynamics for which $c_n = 0$. Because the trait distribution is a probability distribution, we expect that the Gram–Charlier series will have favorable truncation properties in its coefficients c_n, as well as greater numerical stability, relative to a traditional power series. Additionally, unlike other expansions of probability distributions (such as the Edgeworth series), the Gram–Charlier series allows the action of individual cumulants on the dynamics to be isolated.

Inserting the rightmost parenthetical term of Eq. (6) into Eq. (5) as f, we use the property of Hermite polynomials $He_n' = n He_{n-1}$ and the change of variables $\tilde{z} = (\bar{z} - z)/\sigma$ to find,

$$\dot{\bar{z}} = \frac{V}{\overline{W}} \left(\frac{1}{1 - c} \frac{d\overline{W}}{d\bar{z}} \right) + \frac{V}{\sigma \overline{W}} \sum_{n=3}^{\infty} c_n n \mathcal{W}_{n-1}, \tag{7}$$

where the first-order term corresponds to classical phenotypic evolutionary theory. In the higher-order terms, the singly-indexed family of integrals $\mathcal{W}_n$ is given by

$$\mathcal{W}_n \equiv \int \mathcal{N}[0, 1](\tilde{z}) He_n(\tilde{z}) W(\sigma \tilde{z} + \bar{z}) d\tilde{z}. \tag{8}$$

This equation represents a projection of the fitness landscape onto a basis of Hermite polynomials, with finer-scale features in the fitness landscape being represented by larger values of n in the series. However, if the fitness landscape is sufficiently smooth, there always exists some n above which the sequence $\mathcal{W}_n$ continuously decreases, suggesting that a finite set of $\mathcal{W}_n$ is sufficient to describe many simple fitness landscapes.

We refer to the series $\mathcal{W}_n$ as the "coupling coefficients" because the form of Eq. (7) suggests that the values of $\mathcal{W}_n$ represent the degree of coupling between the fitness landscape and progressively larger cumulants of the trait distribution. Larger-scale features of the fitness landscape affect the dynamics over longer timescales,

and thus appear in lower-order $\mathcal{W}_n$; conversely, higher-order coupling coefficients provide increasingly precise information about the landscape, and small-scale dynamical changes within it. Thus $\mathcal{W}_n$ serve a similar function to the individual selection differentials described in previously-developed models of non-Gaussian breeding value distributions using multilocus genetic theory (Turelli and Barton, 1994; Bürger, 1991). Each $\mathcal{W}_n$ may, in principal, depend on lower-order cumulants of the trait distribution; however this dependence reduces to a scaling factor (see Appendices for further discussion).

Importantly, because the coupling coefficients $\mathcal{W}_n$ depend only on lower-order cumulants of the fitness landscape, the series of dynamical equations may be closed and then computed. The integro-differential dynamics of an arbitrary trait distribution and fitness landscape can be expressed as a series of coupled ordinary differential equations for the time-varying moments (appearing in the individual c_n of the Gram–Charlier series), with the fitness landscape appearing only through the coupling terms $\mathcal{W}_n$. For most fitness landscapes each individual $\mathcal{W}_n$ is straightforward to compute analytically (using generating functions); however, for more complex fitness landscapes they may be computed numerically due to the orthogonality properties of Hermite integrals (Babusci et al., 2012). In either case, the dependence of $\mathcal{W}_n$ on lower moments (such as $\bar{z}$ and σ) appears only as a shift or scaling of the integrand, and so the dependence of the analytical or numerical solution on these moments has predictable properties due to shifting properties of Hermite polynomials (see Supplementary Appendix H). For this reason, a set of $\mathcal{W}_n$ may be "pre-computed" for a given fitness landscape and then inserted as explicit terms into the dynamical equation (7). The dynamics of the trait distribution subject to this fitness landscape can then be determined without the need to perform additional time-dependent integrals over the fitness landscape—as lower moments evolve, the known solutions can be shifted accordingly. If the fitness landscape were to change in time, any change arising from a process independent of the trait distribution (e.g. rainfall variation due to large scale weather patterns, or depletion of predators due to overfishing) would allow the time variation of each $\mathcal{W}_n$ to be solved separately from the trait distribution dynamics. The resulting set of $\mathcal{W}_n(t)$ may then be inserted into Eq. (7) to yield the trait dynamics.

2.2. Trait variance dynamics

The method used above to derive the dynamics of the trait mean may be employed to derive corresponding dynamical equations for any moment or cumulant of the trait distribution. Here, we determine the dynamical equation for the trait standard deviation (and thus variance) using a similar method to that above.

As with the trait mean, we assume that the trait variance $M \equiv \sigma^2$ has linear heritability with dynamics specified by

$$\frac{d}{dt}(\sigma^2) = h_\sigma^2(\sigma_w^2 - \sigma^2), \tag{9}$$

where h_σ^2 is the *variance heritability*, or the degree to which the phenotypic variance in one generation influences the phenotypic variance of the next generation. σ_w^2 is the trait variance after selection, which is defined in a manner analogous to $\bar{z}_w$. If the (unmodeled) recombination and mutation processes are sufficiently smooth, such a linear relationship represents the lowest-order term in a series expansion about the case in which the trait variance varies slowly, due to the summation properties of cumulants of random variables (Rattray and Shapiro, 2001; Turelli and Barton, 1990). This relation follows earlier work on the infinitesimal model (Bulmer, 1971), in which a given trait is continuous due to an effectively infinite number of individual loci contributing to it. Below, we further discuss the limitations of a linear heritability relation.

If the trait distribution has the separable form Eq. (4), then it can be shown (see Supplementary Appendix D) that Eq. (9) simplifies to a dynamical equation of the form,

$$\dot{\sigma} = \frac{U}{\overline{W}} \left(\frac{d\overline{W}}{d\sigma} + \frac{1}{\sigma} \int \tilde{z} \, \mathcal{N}[0, 1](\tilde{z}) W(\sigma\tilde{z} + \bar{z}) f' \, d\tilde{z} \right),$$

where $U \equiv (\sigma \, h_\sigma)^2/2$. Inserting the Gram–Charlier series (Eq. (6)) for $p(z)$ and exploiting the properties of the Hermite polynomials results in a final expression,

$$\dot{\sigma} = \frac{U}{\overline{W}} \frac{d\overline{W}}{d\sigma} + \frac{U}{\sigma \overline{W}} \sum_{n=3}^{\infty} c_n \, n\Big((n-1)\mathcal{W}_{n-2} + \mathcal{W}_n\Big). \tag{10}$$

As with the mean trait dynamics equation (7), the variance dynamics (and thus the width of the trait distribution) simplifies to a gradient of the mean fitness, plus an infinite summation over $\mathcal{W}_n$ that takes into account increasingly fine-scale moments of the trait distribution. Using this construction technique, the dynamics of any arbitrary moment or cumulant of the trait distribution may be expressed as a dynamical equation with a similar form.

2.3. Mean fitness dynamics

Based on the forms of Eqs. (7) and (10), the dynamics of the mean fitness are given by

$$\dot{\overline{W}} = \frac{d\overline{W}}{d\bar{z}}\dot{\bar{z}} + \frac{d\overline{W}}{d\sigma}\dot{\sigma} + \sum_{i=3}^{\infty} \frac{d\overline{W}}{d\kappa_i}\kappa_i, \tag{11}$$

where κ_i are the cumulants of $p(z)$. Because all derivatives of $\overline{W}$ may themselves be expressed in terms of the cumulants of the trait distribution $p(z)$, the dynamics of $\overline{W}$ do not involve an additional dependent variable in the dynamical system. The form of $\overline{W}$ suggests that, even in the absence of additive genetic variance ($V = 0$) or in the case of a fixed mean trait $\dot{\bar{z}} = 0$, the mean fitness may continue to change due to the contribution of higher-order cumulants of the trait distribution to the dynamics. These higher-order effects are absent in the standard breeder's equation and will be discussed further below.

2.4. Gradient dynamics with leading-order corrections

Here, we simplify Eqs. (7) and (10) to account for only the leading-order effects of non-Gaussian moments in the trait distribution, resulting in a closed form for the dynamical equations.

The general framework used above for deriving $\dot{\bar{z}}$ and $\dot{\sigma}$ may, in principle, be used to derive dynamical equations for an arbitrary number of moments of the trait distribution. In these cases, a Gram–Charlier series with a non-Gaussian leading-kernel may be preferable (Berberan-Santos, 2007). However, in the following sections we focus primarily on the dynamics of the first and second moments of the trait distribution ($\dot{\bar{z}}$ and $\dot{\sigma}$ derived above) because in many standard fitness models, selection acts directly on the mean and width of the trait distribution, but not necessarily the skewness and higher moments. As a result, the terms in the series $\mathcal{W}_n$ decrease quickly in magnitude, causing the dynamics of $\bar{z}$ and σ to be nearly uncoupled from the dynamics of higher moments. This is equivalent to assuming that higher-order cumulants of the trait distribution affect the dynamics solely as fixed parameter values in the series terms in Eqs. (7) and (10). This restriction implies that higher moments of the trait distribution have zero effective heritability, and that natural selection, together with reproduction, mutation, and recombination, causes these moments to vary much more slowly than the mean and variance (Prügel-Bennett and Shapiro, 1997; Smerlak and Youssef, 2017). However, if higher moments vary continuously due to natural selection, equilibria in

$\bar{z}$ and σ^2 found under these assumptions would remain the same because the forms of their dynamical equations remain constant, although their stability may be further constrained.

In order to illustrate potential applications and to provide closed-form results, in what follows we truncate at fourth order the infinite series c_n appearing in Eqs. (7) and (10). This fourth-order closure is used in order to isolate the effects of asymmetry (c_3) and heavy-tailedness (c_4) on the dynamics of the trait distribution: neither effect can be described by the original breeder's equation due to its implicit assumption of a Gaussian trait distribution. In principle, additional $\mathcal{W}_n$ can be added to account for other effects; however, the contribution of these higher-order terms to the dynamics of lower-order moments is bounded due to the scaling properties of the Gram–Charlier series ($c_n \sim 1/\sigma^n$). Similar cumulant closure relations appear in moment-based models of ecological dynamics (Whittle, 1957; Matis and Kiffe, 1999; Krishnarajah et al., 2005). In previous work by other authors (Turelli and Barton, 1994; Bürger, 1991), the series terms in a discrete-time cumulant dynamical model were computed explicitly for the case of truncation selection, for which the short-time dynamics primarily depend on the leading moments. Additionally, in some discrete-time models of multilocus selection, the cumulant dynamical equations intrinsically contain a finite number of terms and cross terms (Barton and Turelli, 1991; Neher and Shraiman, 2011).

Together, these assumptions result in a simplified set of dynamical equations,

$$\dot{\bar{z}} = \frac{V}{W}\left(\frac{1}{1-c}\frac{d\overline{W}}{d\bar{z}} + \frac{1}{2\sigma}\left(\gamma\mathcal{W}_2 + \frac{1}{3}(k-3)\mathcal{W}_3\right)\right) \qquad (12)$$

$$\dot{\sigma} = \frac{U}{W}\left(\frac{d\overline{W}}{d\sigma} + \frac{1}{2\sigma}\left(\gamma(\mathcal{W}_2 + \mathcal{W}_0) + \frac{1}{3}(k-3)(\mathcal{W}_3 + 2\mathcal{W}_1)\right)\right), \qquad (13)$$

where γ and k are, respectively, the skewness and kurtosis of the trait distribution. While γ may take any value, the kurtosis is mathematically bounded from below by $k \geq \gamma^2 + 1$. Together, Eqs. (12) and (13) may be considered a first-order "correction" to the classical fitness gradient dynamics equation, and they account for the leading-order effects of non-Gaussian features of the trait distribution. When the fitness landscape is centered ($c = 1$), the first terms vanish from both of these equations. In this case, if the trait distribution itself is purely Gaussian ($\gamma = 0, k = 3$), then the remaining terms vanish from the right hand sides of both equations and the trait distribution evolves under classical phenotypic evolutionary theory. However, if the trait distribution is non-Gaussian but $c = 1$, then $\bar{z}$ and σ will vary entirely due to the non-Gaussian components of the trait distribution.

3. Assumptions and related models

Our model is comparable to moment-series approaches previously used to study natural selection in phenotypic and genotypic systems under various assumptions (Neher and Shraiman, 2011; Barton and Turelli, 1987; Bürger, 1991; Zeng, 1987; Turelli and Barton, 1994; Smerlak and Youssef, 2017; Bürger, 1993); we review these approaches and in greater detail in Supplementary Appendix K. As in previous models, we assume that the full dynamics of the trait distribution may be approximated by a finite series of ordinary differential equations, thus reducing a complex partial differential equation problem to a lower-dimensional moment evolution problem. Mathematically, the set of techniques upon which we base our analysis parallels those found in models of genetic processes under the infinitesimal model, in which a given continuous trait is

assumed to depend on an arbitrary number of alleles—in particular, the use of a Gram–Charlier series as a starting point for cumulant iteration equations was pioneered in genetics by Zeng (1987), as well as by Turelli and Barton (1994). Additionally, we note that several related works have focused on the distribution of fitness values (Good and Desai, 2013; Neher and Shraiman, 2011), including recent work producing the intriguing result that many fitness distributions asymptotically approach a fixed class of distributions (Smerlak and Youssef, 2017).

Our assumption of a linear heritability for higher cumulants, Eq. (9), represents the primary assumption of our model regarding the underlying mechanisms of genetic inheritance in our system; it thus introduces the primary limitations of this purely phenotypic approach because it does not include an explicit inheritance mechanism. For extremely strong selection (leading to large changes in the trait distribution within one generation), our model may fail due to both the continuous time assumption and the presence of higher-order terms in Eq. (9). The form of these terms depends on the underlying genetic process, and their general form has previously been found using multilocus theory (Turelli and Barton, 1994; Barton and Turelli, 1991). These subleading terms affect the dynamics over long timescales, and may alter the stability criteria of equilibria.

Our continuous time phenotypic equations may be compared to previous work on cumulant dynamics that treat selection as a discrete-time repeated sampling process weighted by the fitness (Rogers and Prügel-Bennett, 2000). Thus we test our findings below against a simulated Wright–Fisher process, and we find general agreement (see Supplementary Appendix M for details of this numerical work). We thus emphasize that our model is best applied to the study of short-term phenotypic evolutionary trends when genetic assays are unavailable or infeasible; however, over longer timescales in which the additive genetic variance parameter V varies, we expect high-order effects in heritability to manifest. These limitations are consistent with classical usages of "gradient dynamics" models (which our model generalizes), which have found particular utility for the study of coupled ecological and evolutionary processes (Cortez and Weitz, 2014; Lande, 1976; Abrams and Matsuda, 1997; Gilpin and Feldman, 2017).

4. Results

4.1. Cryptic forces of selection arise from non-Gaussian trait distributions

In classical phenotypic evolution, the first term in Eq. (12) is associated with the "force of selection" on the trait mean (Lande and Arnold, 1983). However, the remaining terms in Eqs. (12) and (13) show that the trait distribution can change even when this term is zero, allowing the trait distribution to evolve in the absence of any apparent force of selection. These cryptic selection forces can have significant effects on the overall dynamics of natural selection.

As a demonstration of this effect, Fig. 2 shows the behavior of the system Eq. (12), Eq. (13) relative to a null model in which the trait distribution is always Gaussian (in which case all terms containing $\mathcal{W}_n$ vanish in the dynamical equations). The figure illustrates two separate cryptic forces of selection: the excess selection force on the mean trait (left plots) and the excess selection force on the trait variance (right plots). For the figure, we use a simple fitness landscape consisting of exponential directional selection (Lande, 1981; Turelli and Barton, 1990; Balagam et al., 2011),

$$W = W_0 e^{sz}. \qquad (14)$$

Such a landscape represents limiting case in several common contexts, including selection on metric traits (which frequently have
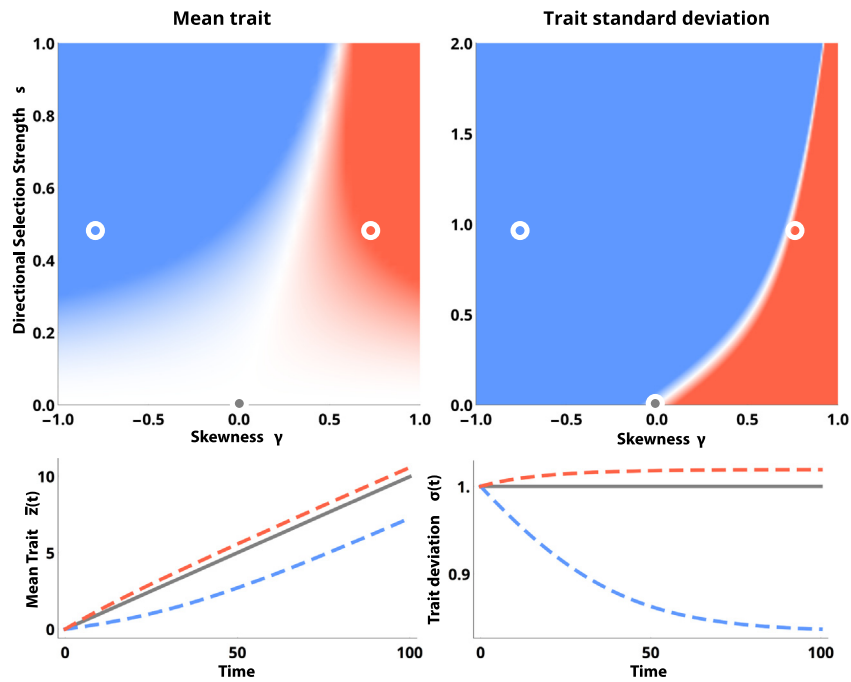
Fig. 2. Cryptic forces of selection under directional selection. Top: The excess force of selection due to the non-Gaussian form of the trait distribution, for the mean trait dynamics (left) and trait standard deviation dynamics (right). Colored shading indicates the relative direction and magnitude of the cryptic terms in Eqs. (12) and (13), with red (blue) indicating cryptic forces that accelerate (retard) the growth of each moment. Shading runs from −1 (deepest blue) to 1 (darkest red). Lower plots represent example dynamics of the mean and standard deviation for representative points on each color plot. White circles on upper plots indicate parameter values (γ, s) used for the trajectories in the lower plots: the gray trace indicates classical fitness gradient dynamics with an unperturbed Gaussian trait distribution, while red and blue traces indicate cases in which higher trait moments speed or hinder the evolutionary dynamics, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

logarithmic ranges) (Karlin, 1987; Lynch and Walsh, 1998), evolution of biochemical reactions subject to microscale energetic constraints (Tsimring et al., 1996; Haldane et al., 2014), and cases in which fitness scales with mutation count (Tripathi et al., 2012; Saakian et al., 2014). The parameter s in Eq. (14) describes the relative strength of selection on the trait z. For this fitness landscape, the individual coupling coefficients are observed to obey the general relationship $\mathcal{W}_n \sim s^n$ (Supplementary Appendix G); thus the broadest features of the fitness landscape (lowest order $\mathcal{W}_n$) contribute the most to directional selection.

The upper portion of the figure describes the relative magnitude of the excess force of selection as a function of the trait skewness γ and the selection strength s; for simplicity, the kurtosis k is held fixed at its theoretical minimum $k = \gamma^2 + 1$. At the origin, the trait distribution is Gaussian and so the cryptic selection force is zero; however, as either the skewness γ or the selection strength s increase, the relative contributions of the non-Gaussian terms in Eq. (12), Eq. (13) increase. Red regions in Fig. 2 correspond to cases in which the cryptic selection force is positive (and thus assists directional selection), whereas blue regions correspond to cases in which the non-Gaussian contributions retard directional selection. The plot suggests that positive skewness (corresponding to a trait distribution with a long tail of large trait values) generally speeds the evolutionary dynamics of both the trait mean and trait variance, primarily due to the added contributions of extremal individuals. The opposite is true for negatively-skewed populations in which most individual trait values exceed the trait mean, due to outlier individuals producing offspring with lower fitness.

In the lower portions of the figure, trajectories of $\bar{z}(t)$ and $\sigma(t)$ are shown for a numerically-constructed non-Gaussian trait distribution (Supplementary Appendix L) using (γ, s) values marked by open circles in the blue and red regions of the upper plots; for comparison, a trajectory consisting of the "null" case of a Gaussian trait distribution is shown in gray. The dynamics of each

distribution relative to the Gaussian case proceed as predicted by the magnitude of the cryptic force, with the primary advantage/detriment due to skewness occurring initially before the rate of evolution eventually stabilizes. Notably, directional selection causes a continued increase in the mean trait ($\bar{z} \rightarrow \infty$ as $t \rightarrow \infty$), but the trait distribution width σ stabilizes to a constant value that is proportional to the skewness. Thus the effect of non-Gaussian features in the trait distribution may manifest experimentally as a constant variance that is larger or smaller than that predicted under classical evolutionary theory.

In Supplementary Appendix M, we compare these results to Wright–Fisher simulations of phenotypic evolution in a population initialized to the same starting values of the mean, variance, skewness, and kurtosis as was used in these equations, and we find general agreement. In the simulations, as in real populations, the values of the skewness and kurtosis drift over time due to accumulated sampling errors (this corresponds to a timescale over which the "fixed cumulant" assumption above is no longer valid). However we observe that trait means and variances tend to drift monotonically under exponential directional selection, and so even over long timescales the calculated direction of the cryptic force of selection correctly predicts the dynamics relative to the Gaussian distribution.

4.2. Transient evolutionary responses to stabilizing and disruptive selection

In addition to generating qualitatively different dynamics, non-Gaussian features of the trait distribution may affect the long-term duration and direction of natural selection. In order to illustrate this effect in a fitness landscape with a variable number of local maxima, we next consider a general fitness landscape described

by an arbitrary polynomial,

$$W(z) = W_0 + \sum_{m=1}^{\infty} \alpha_m (z - c\bar{z})^m, \qquad (15)$$

where the relative magnitudes of the various coefficients α_m determine the number of local maxima and minima of the landscape. In general, if the largest non-zero α_m is positive, then for large values the fitness landscape looks like a "U" ($\lim_{z \to \infty} W(|z|) = \infty$), and so the value of the mean trait $\bar{z}$ eventually diverges: natural selection proceeds indefinitely in the system as the mean trait continuously increases. Sub-leading terms in the polynomial equation (15) induce short term transients that may affect the dynamics only temporarily depending on the initial conditions. Globally, however, natural selection will proceed continuously in a U-shaped landscape, as has been reported experimentally (Kupriyanova, 2014). Conversely, if the largest non-zero α_m is negative, then the fitness landscape looks like a hill at large z and so the mean trait and trait variance will always eventually equilibrate at an intermediate stable solution—in which case natural selection proceeds transiently until this solution is reached (Cortez and Ellner, 2010).

We can investigate the manner in which skewness and kurtosis affect the timescale of natural selection by studying the specific case of a quartic fitness landscape under transient natural selection ($\alpha_m = 0$ for $m > 4$; $\alpha_4 < 0$). This landscape can have either one or two local maxima depending on the relative magnitudes of α_2 and α_4 in Eq. (15). Thus for a continuous, unbounded trait, a quartic fitness landscape represents the simplest landscape that can model both stabilizing (one maxima) or disruptive (two maxima) selection (Gilpin and Feldman, 2017).

For each value of the skewness γ and kurtosis k, the equilibrium solutions of Eqs. (12) and (13) are independent of the initial conditions and can be found analytically along with the Jacobian matrix describing their local stability. Fig. 3 shows plots of the maximum eigenvalue of this Jacobian matrix associated with the intermediate phenotype, parametrized by the skewness γ and kurtosis k (the kurtosis lower bound $k > \gamma^2 + 1$ is indicated by a solid gray line). For cases in which the dynamical equations yield multiple solutions, the intermediate equilibrium phenotype is defined as the solution of Eqs. (12) and (13) with $\bar{z}$ closest to 0. In the blue regions on the plots, this intermediate phenotype is stable and so the mean trait approaches this point. In the red regions of the plots, the intermediate phenotype is unstable and so the mean trait instead approaches another, extremal equilibrium point. The relative darkness of the plot colors represents the relative speed of the evolutionary dynamics: darker blue indicates that the intermediate phenotype is achieved relatively quickly, while darker red indicates that extremal phenotypes are reached quickly.

The point $\gamma = 0, k = 3$ on each plot corresponds to the default case of a purely Gaussian trait distribution with a time-varying mean and standard deviation. For stabilizing selection this point resides within a blue region, consistent which the intuitive result that a Gaussian trait distribution evolves towards the location of the global maximum of the fitness landscape, and that the range of trait values decreases over time as more and more individuals in the population approach this optimum ($\sigma(t) \to 0$) (Tsimring et al., 1996). Indeed, for stabilizing selection, non-Gaussian features of the trait distribution barely affect the rate of the dynamics, consistent with experimental results suggesting that traits under stabilizing selection universally attain intermediate optima (Hoffmann and Merilä, 1999). Additionally, this suggests that if γ and k were themselves dynamical variables that changed either due to selection or mating effects, the mean fitness would nonetheless always reach a finite value dictated by the maximum of the fitness distribution.

In a disruptive landscape, the intermediate phenotype is disfavored for almost all trait distributions near the classical case of a Gaussian distribution at $\gamma = 0, k = 3$ (red regions in Fig. 3). This is expected because the maxima of the disruptive landscape occur away from the intermediate phenotype at $z = 0$, and so the trait moves towards these dispersed values in the absence of additional destabilizing dynamics. However, when the trait distribution is strongly non-Gaussian, the intermediate phenotype near $z = 0$ regains stability, an effect that is particularly pronounced when the trait distribution is strongly asymmetric and flat (large γ, $k \approx \gamma^2 + 1$). High skewness and kurtosis correspond to the trait distribution containing a relatively large fraction of individuals with extremal phenotypes, which represent a large enough trait range that the distance between the two symmetric equilibria in the disrupting fitness landscape becomes relatively insignificant (the locations of these off-center equilibria depend on α_2/α_4 in Eq. (15)). As a result, the overall average phenotype returns to the center due to the leading-order effect of the negative α_4 term in Eq. (15).

In Supplementary Appendix M, we compare the dynamics predicted by local stability analysis of Eqs. (12) and (13) with the dynamics of a population subject to Wright–Fisher dynamics with starting cumulants matching those used here. We find that the short-time dynamics of the Wright–Fisher process are directly analogous to those predicted by the local analysis, particularly because additional effects (such as variation in the genetic variance V, and fluctuations in the values of higher cumulants due to sampling drift) do not manifest over the infinitesimally-short timescales which local stability analysis applies.

Significant skewness and flatness in the trait distribution can obscure the effects of disruptive selection and potentially serve as a mechanism to preserve or enhance phenotypic variance, even when selection itself acts on the trait variance—potentially implicating higher trait moments in speciation and ecological phenomena that depend on the phenotypic variance (Bolnick et al., 2011). Although the dynamical equations would become substantially more complex if the skewness and kurtosis also responded to selection, results from other cumulant dynamical systems (Rattray and Shapiro, 2001; Turelli and Barton, 1990) suggest that genetic or mating processes that preserve a given high order moment would result in similar "freezing" of the dynamics of lower-order cumulants. Thus, even if natural selection alters arbitrarily high moments of the trait distribution, if mutation or mating serve as a "source" that constrains an arbitrary moment of the trait distribution, then lower-order moments such as the phenotypic variance would also be prevented from reaching zero in certain regions of parameter space. This process represents a generalization of the concept of a "mutation-selection balance", a concept typically invoked in order to justify holding the phenotypic variance fixed during selection (Smerlak and Youssef, 2017; Neher and Shraiman, 2011; Lande, 1976). However, as with the traditional breeder's equation, an additional equation specifically incorporating the underlying genetics would be required in order to justify this process biologically (Barton and Turelli, 1987).

4.3. Time-variation of the heritability

Experimental studies of phenotypic evolution in artificial selection regimes generally quantify genetic effects using the narrow sense heritability, $h(t)^2 \equiv V/\sigma(t)^2$, defined as the ratio of additive genetic variance V to overall phenotypic variance σ^2. In large populations, h^2 changes slowly enough that it can be estimated without the need for explicit identification of a trait's genetic origin. However, recent experimental results have suggested that h^2 may change appreciably during short periods of strong selection, especially when the underlying genetics (and thus V) exhibit complex dynamics (Lynch and Walsh, 1998). In particular, rapid changes in the phenotypic variance σ^2 may underlie this phenomenon over short timescales (Bolnick et al., 2011; Hoffmann and Merilä, 1999),
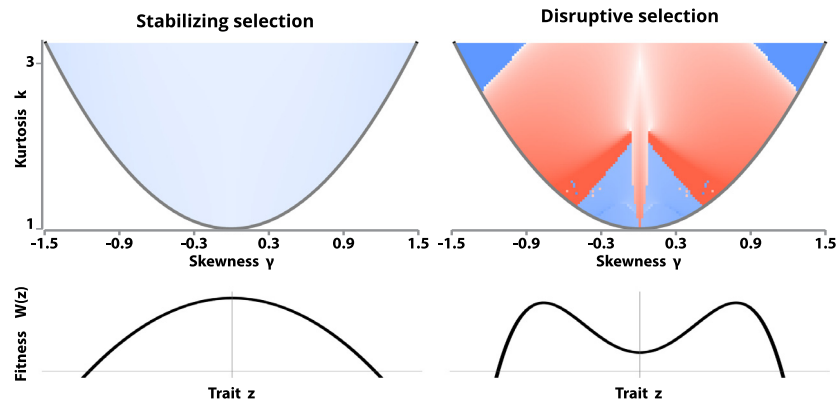
Fig. 3. Evolutionary dynamics under stabilizing or disruptive selection. Colored shading represents the maximum eigenvalue associated with the Jacobian matrix of $\dot{\bar{z}}, \dot{\sigma}$ on a logarithmic scale, evaluated at the equilibrium nearest to $\bar{z} = 0$ (the intermediate phenotype) and parametrized by the skewness and kurtosis of the trait distribution (all other parameters are held constant). Negative maximum eigenvalues (blue regions) represent dynamics that eventually converge to the intermediate phenotype, positive maximum eigenvalues (red regions) represent dynamics that approach an extremal phenotype, and the intensity of shading indicates the instantaneous rate of the dynamics. The solid gray line indicates the lowest mathematically-valid value for the kurtosis, $k > \gamma^2 + 1$. Beneath each figure is a diagram of the fitness landscape $W(z)$ that produced it. For this figure, $c = 1$, $V = 1$, $U = 10$ in Eqs. (12) and (13). Shading ranges from eigenvalue values of -2 (dark blue) to 2 (dark red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

even when an insufficient number of generations have elapsed for V to change appreciably.

We note that by including changes in higher moments than the mean, our framework naturally describes changes in the observed heritability arising from the dynamics of trait variation. In order to study this effect, we re-parametrize our model in terms of quantities more readily measured experimentally: our two dynamical variables $\bar{z}(t)$ and $\sigma(t)$ may be replaced by the equivalent conjugate variables of heritability $h^2(t)$ and mean fitness $\overline{W}(t)$ (see Supplementary Appendix F). Because the exact underlying trait values or distributions may not be accessible in certain experimental contexts, these variables are more descriptive of macroscopic population trends in population-level assays, such as long-term bacterial evolution experiments (Good et al., 2017).

In Fig. 4, we show the evolutionary dynamics of the heritability under a linear directional fitness landscape,

$$W(z) = \alpha_1(z - c\bar{z}), \tag{16}$$

for which the coupling coefficients are computed in Supplementary Appendix J. Because any arbitrary, smooth fitness landscape may be approximated as linear in the neighborhood of the mean trait, this parametrization illustrates a general relationship between the heritability and mean fitness dynamics over short timescales. Setting $\alpha_1 > 0$ results in positive directional selection ($\dot{\overline{W}} > 0$, $\dot{\bar{z}} > 0$) because regions of a fitness landscape with positive slope tend to drive a population towards a higher mean fitness. The solutions to Eq. (12), Eq. (13) corresponding to each timepoint are shown in Fig. 4A, where they are parametrized in terms of the heritability as a function of mean fitness (which itself increases in time). In the figure, different traces correspond to various values of the skewness γ and kurtosis k.

In general, a Gaussian trait distribution ($\gamma = 0, k = 3$ in Eq. (12) and Eq. (13)) always produces the default case of constant heritability ($\dot{h}^2 = 0$, gray dashed line in Fig. 4A). An evolving population with positive skewness (blue traces) exhibits heritability that eventually decreases in time, primarily due to a high fraction of outlier individuals in the high-fitness tail of the trait distribution. These individuals produce higher-fitness offspring quickly enough that the overall trait range increases in time, leading to a corresponding decrease in the overall heritability—which is consistent with several field studies showing that increasing mean fitness also increases trait variation and lowers the observed heritability (Kruuk et al., 2000; Hoffmann and Merilä, 1999). This growth of the variance due to tail effects is consistent with prior analytic

solutions for the case of directional selection (Bürger, 1991; Bürger and Lande, 1994); however, we note that we hold the skewness and higher moments fixed in these calculations. Intriguingly, we find that high values of kurtosis retard this process at short timescales, producing some scenarios where the heritability appears to increase transiently, before eventually relaxing non-monotonically to zero over longer timescales (topmost blue trace). Such nonmonotonicity in the heritability under directional selection would be an observable experimental signature of non-Gaussian dynamics.

Conversely, a population with negative skewness (red traces in Fig. 4A) has a long tail of individuals with comparatively low fitness; these individuals serve to counteract the tendency of directional selection to increase trait ranges, and thus cause the trait variation to decrease in time—leading to an accompanying increase in the trait heritability. Because this effect is driven by the lower-fitness tail of the distribution, it may partly explain experimental results that have reported a disproportionately large contribution of rare phenotypes to the heritability of certain deleterious traits (Mancuso et al., 2016; Schork et al., 2009). The differential effect of left and right skewness is further apparent in the trait distributions at three representative timepoints in the dynamics shown in Fig. 4B, where the apparent width of the trait distribution varies non-monotonically and causes continuous changes in the heritability observed in the upper portion of the figure. This substantial variation in the distribution's width at various points in the dynamics may confound efforts to study experimentally the evolution of ecological niche width (Bolnick et al., 2010), as non-Gaussian features in the trait distribution may cause transient contraction and expansion of the observed trait range, even in the absence of competition.

In Supplementary Appendix M, we perform a comparable set of simulations using Wright–Fisher dynamics, and find general agreement in the dynamics, including non-monotonicity in the dynamics at large values of k, as well as a qualitative shift from $h^2 \to 1$ to $h^2 \to 0$ when γ changes sign from negative to positive.

5. Discussion

We have presented a formulation of phenotypic evolution that iteratively relates the dynamics of an arbitrary trait distribution to an arbitrary fitness landscape. This simplified model explains several phenomena observed in numerical simulations of non-Gaussian trait distributions evolving in simple fitness landscapes.
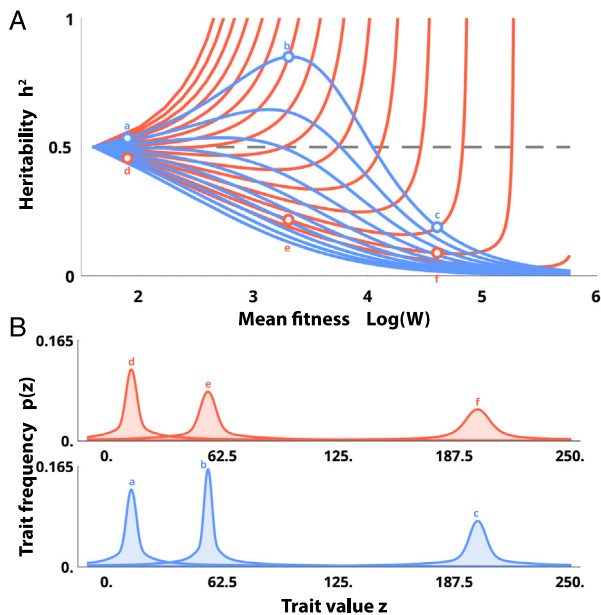
Fig. 4. Time-dependent heritability under directional selection. (A) The heritability h^2 as a function of the mean fitness, which increases continuously over time in a directional fitness landscape. Colors represent cases in which the trait distribution skewness is $\gamma = -0.05$ (red) or $\gamma = 0.05$ (blue). Different traces of the same color correspond to increasing values of the kurtosis k in the range $k = \gamma^2 + 1$ to $k = \gamma^2 + 6$, with slower timescales (lower traces) corresponding to larger values of the kurtosis. The gray dashed line corresponds to evolutionary dynamics with constant heritability, which the dynamical equations recreate when $\gamma = 0, k = 3$. (B) The trait distributions for three representative timepoints (and thus values of $\overline{W}$) marked by open circles in (A). For this figure, $\alpha = V = 1, U = 0.01, c = 0.5$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

We have shown that skewness or asymmetry in the trait distribution can delay directional selection or prevent disruptive selection, and that the heritability – typically assumed to be constant over short timescales – can vary in time with complex, non-monotonic dynamics as the mean fitness changes.

Our results are consistent with empirical and theoretical analyses that formulate phenotypic evolution in terms of the probability distribution of fitness values $p(W)$ rather than in terms of an explicit trait distribution $p(z)$ and auxiliary fitness function $W(z)$ (Good and Desai, 2013). However, the trait-based model here is more readily applied to arbitrarily complex fitness functions (particularly non-invertible and multimodal fitness landscapes) (Neher and Shraiman, 2011), and thus may assist the study of explicitly experimental quantities such as selection coefficients and realized heritability. For fitness distribution models, similar cumulant dynamical equations have been derived under distinct constraints and approximations, and for certain types of experiments (e.g., microbiological lineage assays) these formalisms may be preferable (Smerlak and Youssef, 2017; Neher and Shraiman, 2011; Turelli and Barton, 1990; Good et al., 2017).

Our explicitly phenotypic approach does not involve an explicit underlying genetic model; we assume a linear heritability relation and note that this approximation holds over the relatively short timescales and slow selection regimes observed in our simulations (Bulmer, 1971). A more detailed model would include explicit information about the mechanics of inheritance, and how these mechanics contribute to the breeder's equation and determine the parameters it contains. An important starting point for such work would be models of non-Gaussian evolution based on a multilocus genetic models (Barton and Turelli, 1987; Bürger, 1991), which have shown that the dynamics of the phenotypic moments

may vary appreciably depending on mating effects, transmission effects, and whether selection occurs before or after transmission (Turelli and Barton, 1994).

Other potential improvements include alternative series closure schemes to raw truncation that depend on the type of fitness landscape being evaluated, as well as additional cross terms in the dynamical equations that account for assortative mating effects. Additionally, we have assumed that the additive genetic variance remains fixed during the phenotypic dynamics; however over long timescales natural selection and mating may affect this variance considerably. Coupling genotypes and phenotypes may also require the phenotypic portion of the model to be generalized for multivariate traits by defining a fully time-dependent phenotypic variance–covariance matrix, which may be particularly important for selection experiments in which the traits with the strongest selection responses are unknown a priori (Lande, 1979; Steppan et al., 2002).

Another limitation of our approach comes from the truncation approximations necessary to make the dynamical equations closed-form. Here, we have chosen canonical directional and minimal stabilizing/disruptive fitness landscapes, in order to illustrate the simplest non-trivial dynamics arising from our model and to justify our inclusion of only the skewness and kurtosis (which represent the leading-order departures from Gaussian trait distributions). Additionally, we have retained only two of the dynamical equations in our model: one for the mean trait and one for the trait variance. But for certain types of initial trait distribution or fitness landscape, the leading order Eqs. (12) and (13) may not be sufficiently accurate, and instead the full dynamical equations (7), (10), and their equivalents for higher cumulants may be necessary. The number of necessary dynamical equations, and the number of terms to include in them, depends primarily on the form of the coupling coefficients $\mathcal{W}_n$. A more detailed fitness landscape, such as one derived empirically, may require retention of more terms in the series $\mathcal{W}_n$, because higher-order terms correspond to finer-scale fitness landscape variations which affect dynamics over smaller length and timescales. However, because computing $\mathcal{W}_n$ only requires projecting a given fitness landscape onto a polynomial basis, this step may be performed numerically.

Competing interests

We declare no competing interests.

Appendix A. Supplementary Appendices

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.tpb.2018.11.002.

References

Abrams, P.A., Matsuda, H., 1997. Fitness minimization and dynamic instability as a consequence of predator–prey coevolution. Evol. Ecol. 11 (1), 1–20.

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

Babusci, D., Dattoli, G., Quattromini, M., 2012. On integrals involving Hermite polynomials. Appl. Math. Lett. 25 (8), 1157–1160.

Balagam, R., Singh, V., Sagi, A.R., Dixit, N.M., 2011. Taking multiple infections of cells and recombination into account leads to small within-host effective-population-size estimates of HIV-1. PLoS One 6 (1), e14531.

Barton, N.H., Polechová, J., 2005. The limitations of adaptive dynamics as a model of evolution. J. Evol. Biol. 18 (5), 1186–1190.

Barton, N., Turelli, M., 1987. Adaptive landscapes, genetic distance and the evolution of quantitative characters. Genet. Res. 49 (2), 157–173.

Barton, N., Turelli, M., 1991. Natural and sexual selection on many loci. Genetics 127 (1), 229–255.

Berberan-Santos, M.N., 2007. Expressing a probability density function in terms of another PDF: A generalized Gram-Charlier expansion. J. Math. Chem. 42 (3), 585–594.

Björklund, M., Husby, A., Gustafsson, L., 2013. Rapid and unpredictable changes of the G-matrix in a natural bird population over 25 years. J. Evol. Biol. 26 (1), 1–13.

Bolnick, D.I., Amarasekare, P., Araújo, M.S., Bürger, R., Levine, J.M., Novak, M., Rudolf, V.H., Schreiber, S.J., Urban, M.C., Vasseur, D.A., 2011. Why intraspecific trait variation matters in community ecology. Trends Ecol. Evol. 26 (4), 183–192.

Bolnick, D.I., Ingram, T., Stutz, W.E., Snowberg, L.K., Lau, O.L., Paull, J.S., 2010. Ecological release from interspecific competition leads to decoupled changes in population and individual niche width. Proc. R. Soc. Lond. B Biol. Sci. 277 (1689), 1789–1797.

Bulmer, M., 1971. The effect of selection on genetic variability. Am. Nat. 105 (943), 201–211.

Bulmer, M.G., et al., 1980. The Mathematical Theory of Quantitative Genetics. Clarendon Press.

Bürger, R., 1991. Moments, cumulants, and polygenic dynamics. J. Math. Biol. 30 (2), 199–213.

Bürger, R., 1993. Predictions of the dynamics of a polygenic character under directional selection. J. Theoret. Biol. 162 (4), 487–513.

Bürger, R., Lande, R., 1994. On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. Genetics 138 (3), 901–912.

Cortez, M.H., Ellner, S.P., 2010. Understanding rapid evolution in predator-prey interactions using the theory of fast-slow dynamical systems. Am. Nat. 176 (5), E109–E127.

Cortez, M.H., Weitz, J.S., 2014. Coevolution can reverse predator–prey cycles. Proc. Natl. Acad. Sci. 111 (20), 7486–7491.

Falconer, D.S., 1981. Introduction to Quantitative Genetics, second ed. Pearson Education.

Garant, D., Hadfield, J.D., Kruuk, L.E., Sheldon, B.C., 2008. Stability of genetic variance and covariance for reproductive characters in the face of climate change in a wild bird population. Mol. Ecol. 17 (1), 179–188.

Gilpin, W., Feldman, M.W., 2017. A phase transition induces chaos in a predator-prey ecosystem with a dynamic fitness landscape. PLoS Comput. Biol. 13 (7), e1005644.

Gingerich, P.D., 1976. Paleontology and phylogeny: patterns of evolution at the species level in early Tertiary mammals. Am. J. Sci. 276 (1), 1–28.

Gingerich, P.D., 2003. Mammalian responses to climate change at the Paleocene-Eocene boundary: Polecat Bench record in the northern Bighorn Basin, Wyoming. Special Papers-Geol. Soc. Am. 463–478.

Good, B.H., Desai, M.M., 2013. Fluctuations in fitness distributions and the effects of weak linked selection on sequence evolution. Theor. Popul. Biol. 85, 86–102.

Good, B.H., McDonald, M.J., Barrick, J.E., Lenski, R.E., Desai, M.M., 2017. The dynamics of molecular evolution over 60,000 generations. Nature.

Grant, P.R., Grant, B., 1995. Predicting microevolutionary responses to directional selection on heritable variation. Evolution 49 (2), 241–251.

Haldane, A., Manhart, M., Morozov, A.V., 2014. Biophysical fitness landscapes for transcription factor binding sites. PLoS Comput. Biol. 10 (7), e1003683.

Hartl, D.L., 2014. What can we learn from fitness landscapes? Curr. Opin. Microbiol. 21, 51–57.

Hoffmann, A.A., Merilä, J., 1999. Heritable variation and evolution under favourable and unfavourable conditions. Trends Ecol. Evol. 14 (3), 96–101.

Jones, A.G., Arnold, S.J., Bürger, R., 2003. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. Evolution 57 (8), 1747–1760.

Karlin, S., 1987. Non Gaussian Phenotypic Models of Quantitative Traits. Morrison Institute for Population and Resource Studies at Stanford.

Krishnarajah, I., Cook, A., Marion, G., Gibson, G., 2005. Novel moment closure approximations in stochastic epidemics. Bull. Math. Biol. 67 (4), 855–873.

Kruuk, L.E., Clutton-Brock, T.H., Slate, J., Pemberton, J.M., Brotherstone, S., Guinness, F.E., 2000. Heritability of fitness in a wild mammal population. Proc. Natl. Acad. Sci. 97 (2), 698–703.

Kupriyanova, E.K., 2014. What do egg size distributions in marine invertebrates tell us about validity of fecundity-time models? Mar. Ecol. 35 (2), 249–253.

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

Lande, R., 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33 (1Part2), 402–416.

Lande, R., 1981. Models of speciation by sexual selection on polygenic traits. Proc. Natl. Acad. Sci. 78 (6), 3721–3725.

Lande, R., Arnold, S.J., 1983. The measurement of selection on correlated characters. Evolution 37 (6), 1210–1226.

Lynch, M., Walsh, B., 1998. Genetics and Analysis of Quantitative Traits. Sinauer, URL https://books.google.com/books?id=UhCCQgAACAAJ.

Mancuso, N., Rohland, N., Rand, K.A., Tandon, A., Allen, A., Quinque, D., Mallick, S., Li, H., Stram, A., Sheng, X., et al., 2016. The contribution of rare variation to prostate cancer heritability. Nature Genet 48 (1), 30.

Martin, C.H., Wainwright, P.C., 2013. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. Science 339 (6116), 208–211.

Matis, J.H., Kiffe, T.R., 1999. Effects of immigration on some stochastic logistic models: a cumulant truncation analysis. Theor. Popul. Biol. 56 (2), 139–161.

Miles, D.B., 2004. The race goes to the swift: fitness consequences of variation in sprint performance in juvenile lizards. Evol. Ecol. Res. 6 (1), 63–75.

Mustonen, V., Lässig, M., 2009. From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. Trends Genet. 25 (3), 111–119.

Neher, R.A., Shraiman, B.I., 2011. Statistical genetics and evolution of quantitative traits. Rev. Modern Phys. 83 (4), 1283.

Nowak, M.A., Sigmund, K., 2004. Evolutionary dynamics of biological games. Science 303 (5659), 793–799.

Poelwijk, F.J., Kiviet, D.J., Weinreich, D.M., Tans, S.J., 2007. Empirical fitness landscapes reveal accessible evolutionary paths. Nature 445 (7126), 383.

Prügel-Bennett, A., 1997. Modelling evolving populations. J. Theoret. Biol. 185 (1), 81–95.

Prügel-Bennett, A., Shapiro, J.L., 1997. The dynamics of a genetic algorithm for simple random Ising systems. Physica D 104 (1), 75–114.

Rattray, M., Shapiro, J.L., 2001. Cumulant dynamics of a population under multiplicative selection, mutation, and drift. Theoret. Popul. Biol. 60 (1), 17–31.

Reznick, D.N., Shaw, F.H., Rodd, F.H., Shaw, R.G., 1997. Evaluation of the rate of evolution in natural populations of guppies (Poecilia reticulata). Science 275 (5308), 1934–1937.

Rogers, A., Prügel-Bennett, A., 2000. Evolving populations with overlapping generations. Theor. Popul. Biol. 57 (2), 121–129.

Saakian, D.B., Ghazaryan, M.H., Hu, C.K., 2014. Punctuated equilibrium and shock waves in molecular models of biological evolution. Phys. Rev. E 90 (2), 022712.

Sasaki, A., Ellner, S., 1995. The evolutionarily stable phenotype distribution in a random environment. Evolution 49 (2), 337–350.

Sato, K., Kaneko, K., 2007. Evolution equation of phenotype distribution: general formulation and application to error catastrophe. Phys. Rev. E 75 (6), 061909.

Schork, N.J., Murray, S.S., Frazer, K.A., Topol, E.J., 2009. Common vs. rare allele hypotheses for complex diseases. Curr. Opin. Genet. Dev. 19 (3), 212–219.

Shen, B., Dong, L., Xiao, S., Kowalewski, M., 2008. The Avalon explosion: evolution of Ediacara morphospace. Science 319 (5859), 81–84.

Slatkin, M., 1970. Selection and polygenic characters. Proc. Natl. Acad. Sci. 66 (1), 87–93.

Smerlak, M., Youssef, A., 2017. Limiting fitness distributions in evolutionary dynamics. J. Theoret. Biol. 416, 68–80.

Steppan, S.J., Phillips, P.C., Houle, D., 2002. Comparative quantitative genetics: evolution of the G matrix. Trends Ecol. Evol. 17 (7), 320–327.

Tripathi, K., Balagam, R., Vishnoi, N.K., Dixit, N.M., 2012. Stochastic simulations suggest that HIV-1 survives close to its error threshold. PLoS Comput. Biol. 8 (9), e1002684.

Tsimring, L.S., Levine, H., Kessler, D.A., 1996. RNA virus evolution via a fitness-space model. Phys. Rev. Lett. 76 (23), 4440.

Turelli, M., 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. Evolution 42 (6), 1342–1347.

Turelli, M., Barton, N., 1990. Dynamics of polygenic characters under selection. Theoret. Popul. Biol. 38 (1), 1–57.

Turelli, M., Barton, N., 1994. Genetic and statistical analyses of strong selection on polygenic traits: What, me normal? Genetics 138 (3), 913–941.

Waxman, D., Gavrilets, S., 2005. 20 questions on adaptive dynamics. J. Evol. Biol. 18 (5), 1139–1154.

Whittle, P., 1957. On the use of the normal approximation in the treatment of stochastic processes. J. R. Stat. Soc. Ser. B Stat. Methodol. 268–281.

Wright, S., 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: Proc. 6th Int. Congr. Genet., 1 (1), 356–366.

Zeng, Z.B., 1987. Genotypic distribution at the limits to natural and artificial selection with mutation. Theor. Popul. Biol. 32 (1), 90–113.