American Genetic Association

Journal of Heredity, 2019, 383–395
doi:10.1093/jhered/esz018
THE 2017 AGA KEY DISTINGUISHED LECTURE

OXFORD

THE 2017 AGA KEY DISTINGUISHED LECTURE

Estimation of Genetic Variance in Fitness, and Inference of Adaptation, When Fitness Follows a Log-Normal Distribution

Timothée Bonnet, Michael B. Morrissey, and Loeske E. B. Kruuk

From the Division of Ecology and Evolution, Research School of Biology, The Australian National University, 46 Sullivans Creek Road, Canberra, ACT 2600, Australia (Bonnet and Kruuk); and School of Biology, University of St Andrews, St Andrews Fife KY16 9TH, UK (Morrissey).

Address correspondence to Timothée Bonnet at the address above, or e-mail: timotheebonnetc@gmail.com.

Received November 14, 2018; First decision January 23, 2019; Accepted April 7, 2019.

Corresponding Editor: Anne Bronikowski

Loeske Kruuk is a professor in the Research School of Biology at the Australian National University. She did a PhD in population genetics at the University of Edinburgh, followed by a postdoc at the University of Cambridge, and then held a Royal Society University Research Fellowship at the University of Edinburgh and an Australian Research Council Future Fellowship at the Australian National University. Her research focuses on evolutionary ecology and quantitative genetics of natural populations, with the aim of understanding how both evolutionary processes and environmental conditions shape diversity in natural populations. She has worked on a range of species, largely but not entirely using long-term studies of wild vertebrate populations with individual-level field data. A key current interest is to estimate the genetic variance of fitness in wild populations.

Abstract

Additive genetic variance in relative fitness ($\sigma^2_A(w)$) is arguably the most important evolutionary parameter in a population because, by Fisher's fundamental theorem of natural selection (FTNS; Fisher RA. 1930. The genetical theory of natural selection. 1st ed. Oxford: Clarendon Press), it represents the rate of adaptive evolution. However, to date, there are few estimates of $\sigma^2_A(w)$ in natural populations. Moreover, most of the available estimates rely on Gaussian assumptions inappropriate for fitness data, with unclear consequences. "Generalized linear animal models" (GLAMs) tend to be more appropriate for fitness data, but they estimate parameters on a transformed ("latent") scale that is not directly interpretable for inferences on the data scale. Here we exploit the latest theoretical developments to clarify how best to estimate quantitative genetic

parameters for fitness. Specifically, we use computer simulations to confirm a recently developed analog of the FTNS in the case when expected fitness follows a log-normal distribution. In this situation, the additive genetic variance in *absolute* fitness on the latent log-scale ($\sigma_A^2(l)$) equals ($\sigma_A^2(w)$) on the data scale, which is the rate of adaptation *within* a generation. However, due to inheritance distortion, the change in mean relative fitness *between* generations exceeds $\sigma_A^2(l)$ and equals $(\exp(\sigma_A^2(l)) - 1)$. We illustrate why the heritability of fitness is generally low and is not a good measure of the rate of adaptation. Finally, we explore how well the relevant parameters can be estimated by animal models, comparing Gaussian models with Poisson GLAMs. Our results illustrate 1) the correspondence between quantitative genetics and population dynamics encapsulated in the FTNS and its log-normal-analog and 2) the appropriate interpretation of GLAM parameter estimates.

Subject area: Quantitative genetics and Mendelian inheritance
Key words: animal model, fundamental theorem of natural selection, GLMM, heritability, quantitative genetics

The concept of "fitness" has been central to evolutionary biology since the beginnings of the field (Wagner 2010). An increase in mean fitness in a population, for example due to adaptive evolution, can increase a population's growth rate and thus reduce the risks of population extinction in the face of environmental changes (Saccheri and Hanski 2006; Hendry et al. 2018). The rate of genetic evolution in mean fitness (i.e., the change due to a response to selection) therefore provides an integrative measure of adaptation. By Fisher's fundamental theorem of natural selection (FTNS; Fisher 1930), this change in mean relative fitness due to selection equals the additive genetic variance in relative fitness in a population, making the additive genetic variance in relative fitness arguably the most important evolutionary parameter for a population. However, its empirical estimation is exceptionally difficult, and as a result, estimates to date are still relatively scarce (Hendry et al. 2018). One of the most challenging aspects in the estimation of additive genetic variance in fitness is the statistical issue that individual fitness typically follows a non-Gaussian distribution and hence should be analyzed using more complex "generalized" linear models (Bolker et al. 2009). These models come with inherent difficulties of interpretation of parameters on different scales: of particular relevance to any application of Fisher's FTNS is the fact that genetic effects that are additive on the ("latent") scale of estimation are no longer solely additive when transformed back to the scale of the observed data: additive genetic variance on the latent scale, therefore, maps to both additive and nonadditive genetic variance on the original data scale. Usefully, however, recent theoretical developments have provided ways of dealing with these issues, in particular via an analog of Fisher's FTNS that predicts rates of change in mean fitness when fitness follows a log-normal distribution (Morrissey and Bonnet 2019). Here, we present an overview of the issues and use simulated fitness data to ask: how can we predict the rate of evolutionary adaptation when fitness is log-normal?

Before dealing with the complexities of the alternative scales, we first outline some general issues relevant to the estimation of genetic variance in fitness. Many readers will be most familiar with heritability as a measure of genetic variance and as a determinant of the rate of trait evolution, and the narrow-sense heritability of fitness has been a parameter of interest in empirical studies over several decades (Gustafsson 1986; Jones 1987; Kruuk et al. 2000; Merilä and Sheldon 2000; Reid et al. 2011). However, heritability is not a perfect measure of the ability of a trait to evolve (Hansen et al. 2011) and in the case of fitness, a more useful measure of the rate of genetic evolution is the additive genetic variance. To see why this is the case, consider the "breeder's equation" (Lush 1937) prediction for a response to selection applied to relative fitness (w) (see also Walsh and Lynch 2018, chapter 6). Under some assumptions (Heywood 2005; Morrissey et al. 2010), the breeder's equation predicts $R = h^2S$, where R is the response to selection, that is, the expected change in mean relative fitness between 2 successive generations (keeping the same reference to compute relative fitness), S is the selection differential on relative fitness, and h^2 is the narrow-sense heritability, that is, the proportion of additive genetic variance ($\sigma_A^2(w)$) relative to the phenotypic variance in relative fitness ($\sigma_P^2(w)$). Assuming that there is no covariation between genetic and environmental sources of phenotypic variation, $\sigma_P^2(w) = \sigma_A^2(w) + \sigma_E^2(w)$, where $\sigma_E^2(w)$ is the environmental variance in relative fitness.

The heritability of fitness is thus $h^2 = \sigma_A^2(w)/(\sigma_A^2(w) + \sigma_E^2(w))$, so for a given $\sigma_A^2(w)$, increasing $\sigma_E^2(w)$ decreases h^2. This result suggests that an increasing $\sigma_E^2(w)$ should decrease R. However, in the case of fitness, increasing $\sigma_E^2(w)$ also means increasing S. This is because a selection differential is defined as the phenotypic covariance (σ_P) between a trait and relative fitness (Price 1970):

$$\begin{aligned} S &= \sigma_p(w, Z), \text{ where } Z \text{ is any trait under selection} \\ &= \sigma_p(w, w), \text{ because } Z = w \\ &= \sigma_p^2(w), \text{ by definition of variance and covariance} \\ &= \sigma_E^2(w) + \sigma_A^2(w), \text{ assuming no covariation between genetic and environmental deviations} \end{aligned}$$

Therefore,

$$\begin{aligned} R &= h^2 S \\ &= \frac{\sigma_A^2(w)}{\sigma_A^2(w) + \sigma_E^2(w)} \times (\sigma_A^2(w) + \sigma_E^2(w)) \\ &= \sigma_A^2(w) \end{aligned}$$

When $\sigma_E^2(w)$ increases, the increase in selection exactly compensates the decrease in heritability. In other words, the more phenotypic variation in fitness (i.e., opportunity for selection) there is, the more selection can give amplitude to the heritable part of fitness, although selection captures the heritable part less reliably.

Selection on fitness always causes the breeding values for relative fitness to increase by an amount of $\sigma_A^2(w)$ *within* a generation. This is Fisher's FTNS (Fisher 1930), as also recovered above using the breeder's equation (and sometimes written equivalently as the change in mean absolute fitness, $\bar{W}$, by $\Delta\bar{W} = \sigma_A^2(W)/\bar{W}$). In the absence of changes in the environment, the increase in mean breeding values is perfectly transmitted to the next generation and

increases the population growth rate by $\sigma_A^2(w)$. Note that here the "environment" is actually defined in part by genetic properties of the population such as allele frequencies, and indirect genetic effects due to interactions with conspecifics, so in practice the relationship between Fisher's FTNS and population growth rate relies on large assumptions (Price 1972a; Ewens 1989).

Despite these assumptions, for empiricists, $\sigma_A^2(w)$ can give a useful integrative measure of evolution through natural selection (Shaw and Shaw 2014). In other words, it can be used as a proxy of how fast a population is currently evolving, assuming evolutionary forces other than selection are negligible (Morrissey et al. 2010). It can also potentially be used as an indication of whether a population can persist in the face of environmental change, and hence as a link between evolutionary and population dynamics (Asmussen 1983; Gomulkiewicz et al. 2010; McPeek 2017, chapter 3). Finally, $\sigma_A^2(w)$ indicates whether there is the potential for any other trait to evolve, as if $\sigma_A^2(w) = 0$, no trait can evolve by direct adaptation in the current environment (Heywood 2005; Morrissey et al. 2010), although see Bijma 2010; Fisher and McAdam 2019) for indirect adaptation). Empirical knowledge of $\sigma_A^2(w)$ could, therefore, be central to both fundamental research and applications in agronomy and wildlife management. However, this potential importance is at odds with the scarcity of estimates of $\sigma_A^2(w)$ in natural populations.

Burt (1995) identified 12 estimates of $\sigma_A^2(w)$ from natural populations, with values ranging from 0 to 0.2. More recently, Hendry et al. (2018) identified 22 studies of 16 species estimating $\sigma_A^2(w)$ in wild plants and animal populations, with values ranging from 0 to 0.85 but mostly below 0.2. These few estimates form, to our knowledge, the entirety of current knowledge about this crucial parameter in natural populations—a paucity that contrasts markedly with the large number of estimates of genetic variance in a range of other traits (e.g., Postma 2014) reports upwards of 16 000 estimates in natural populations). Furthermore, many of the current estimates of $\sigma_A^2(w)$ may be unreliable because of data limitations and statistical problems, which we discuss in detail below. Given the tremendous theoretical importance of genetic variation in fitness, why are there so few reliable estimates of $\sigma_A^2(w)$ in natural populations?

A first explanation for the scarcity of estimates of $\sigma_A^2(w)$ is that researchers do not attempt to estimate it because they expect the value will be close to zero (Shaw and Shaw 2014). This expectation comes from the intuition that populations appear generally adapted to their environments or at least are thought to have lived in their environment long enough to be at evolutionary equilibria (e.g., Jones 1987). Theory shows that at evolutionary equilibrium, mutation, migration, and spatiotemporal variation in selection typically introduce only negligible additive genetic variance in fitness (Charlesworth 1987). However, most populations are probably not at evolutionary equilibrium (Shaw and Shaw 2014), especially not in contemporary times affected by anthropogenic environmental changes (Pelletier and Coltman 2018). Furthermore, the few estimates available suggest that additive genetic variance in fitness can be significant and may be larger than additive genetic variance in other traits when expressed as a coefficient of variation (Merilä and Sheldon 2000)—although this may be difficult to interpret given that right-skewed traits (such as fitness) typically have higher coefficients of variation (Kruuk et al. 2000). Of relevance here is the fact that, in the case of fitness, the (square of the) coefficient of genetic variation in absolute fitness is equivalent to the genetic variance in relative fitness ($\sigma_A^2(w)$).

Second, the estimation of $\sigma_A^2(w)$ requires measures of the fitness of individuals in a population, combined with estimates of their genetic relatedness. The measurement of individual fitness and relatedness used to be very challenging, if not impossible, in natural populations, and still requires intense research efforts (Clutton-Brock and Sheldon 2010; Hendry et al. 2018). Indeed, measuring fitness means tracking the survival and reproduction of most individuals in a population through their lifetimes and their natural environment. As a consequence, many more studies estimate genetic parameters for components of fitness than for total fitness, and therefore are not able to capture the full genetic variance in fitness, which could be either larger or smaller depending on genetic variances and covariances between all components of fitness (Postma 2014; Shaw and Shaw 2014). Fortunately, long-term monitoring of wild populations is becoming more common, and studies are accumulating the necessary data, the quality of which is continually improving with the arrival of new techniques. The accumulation of fitness and relatedness data over several generations thus increases the range of models that can be fitted and the complexity of questions that can be addressed (Clutton-Brock and Sheldon 2010; Kruuk et al. 2014a).

A third problem relates to why the few estimates of $\sigma_A^2(w)$ available to date may be unreliable. Quantitative genetic methods used for the first estimates of $\sigma_A^2(w)$ in natural populations, for instance using parent–offspring regression (POr), are prone to biases due to transgenerational shared environments among relatives (Kruuk 2004) and lack of statistical power (de Villemereuil et al. 2013). "Animal models," a special type of mixed-effect models using pairwise relatedness matrices (Henderson 1950), are more powerful, better able to separate genetic and nongenetic causes of similarity between relatives and, unlike POr, able to directly estimate σ_A^2 for a trait. Animal models have been widely adopted in recent decades for studies of genetic variation in natural populations (Kruuk 2004; Wilson et al. 2010) and continue to be developed to better address questions in natural populations (e.g., Kruuk et al. 2014a).

The 3 problems described above are generally acknowledged and, while collection of suitable empirical data will always be challenging, should not hinder further attempts at estimation of $\sigma_A^2(w)$. By contrast, a fourth problem has received less attention and hence is our focus here: estimates for $\sigma_A^2(w)$ to date may be few and may possibly be unreliable because of difficulties inherent in modeling fitness distributions.

Quantitative genetics most often uses statistic analyses which assume that traits follow a normal distribution (e.g., Lush 1937; Lande 1979). However, in nature, many traits of interest, and in particular fitness, follow highly non-Gaussian distributions often with high skewness (e.g., Walsh and Lynch 2018, chapter 24), and it is generally unclear what the consequences of breaking the assumption of normality are. To date, the problem has primarily been addressed through various approximations that can perform well for some forms of distributions and some parameter values (e.g., for binary data, van Vleck 1972; Roff 2001; de Villemereuil et al. 2013). However, the statistical methods required to estimate $\sigma_A^2(w)$ appropriately have improved dramatically in recent years. Empirical analyses of fitness can be conducted using "generalized linear animal models" (GLAMs): "generalized" in that they allow for error distribution models other than a normal distribution (Bolker et al. 2009) and "animal models," in that one of the random effects is an additive genetic effect with covariance structure determined by a pairwise relatedness matrix (for examples of GLAMs, see Hadfield 2010; Milot et al. 2011; Wilson et al. 2011; de Villemereuil et al. 2013; Mair et al. 2015; see also the Aster method, Geyer et al. 2007).

When fitness is non-Gaussian, GLAMs will afford a more appropriate fit to the data, better model predictions, and more reliable

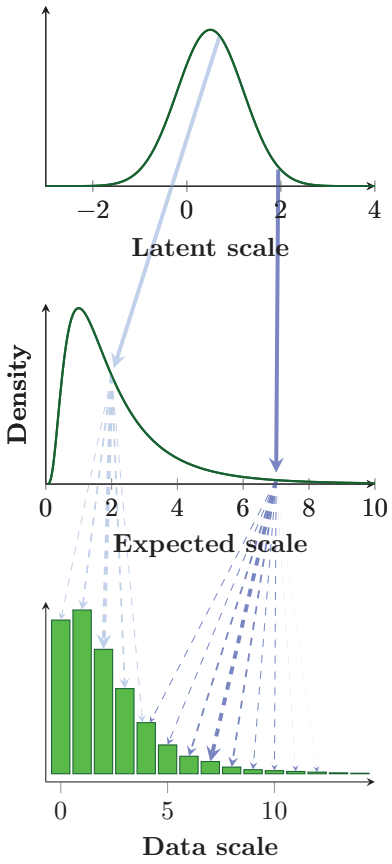

Figure 1. Relationships between the scales of the generalized linear (animal) model using a Poisson trait distribution and a logarithm-link function. *x* axes represent absolute fitness on different scales, and *y* axes represent probability density. Solid arrows represent the deterministic, one-to-one relationship linking the *latent* scale to the *expected* data scale (determined by the link function). The expected scale relates to expected fitness values that can be understood as the average fitness of an individual that would emerge from observing their fitness many times. Expected fitness is on a continuous scale because these averages can take any positive value. However, lifetime absolute fitness is expressed only once for each individual and, as defined here in terms of counts of offspring, can only be an integer value. Therefore, realized fitness on the *data* scale will generally deviate from the expected value, with the deviation being stochastic and following a Poisson distribution of discrete, integer values. Dashed arrows represent the stochastic relationships linking one value on the expected scale to multiple possible realization of values on the data scale, with the thickness of each line representing the relative probability of reaching different data scale values. Figure modified from de Villemereuil et al. (2016).

ascertainment of uncertainty than linear animal models. However, any generalized linear model (GLM) estimates parameters on a transformed scale (the "latent scale") based on a link function of the data, and estimates on this latent scale may not be directly interpretable or applicable to theoretical predictions, as they no longer involve the same units as the original scale on which the data were observed. Furthermore, linear models predict values on a continuous scale, which will differ from the discrete (e.g., integer) values taken by traits such as fitness, introducing an additional step of translation from an "expected" (continuous) to "observed" (discrete) data scale. Following de Villemereuil et al. (2016), we set out the relationship between the latent and the 2 data scales in Figure 1. The Gaussian distribution assumed on the latent scale of a GLAM (or of any GLM) maps onto an "expected" data scale via a link function (such as a logarithm). This "expected scale" relates to the actual observed "data scale" via a stochastic data-generating process, such as a Poisson process, with a single value on the expected scale generating a range of realized values on the data scale. This second step is what sets GLAMs apart from animal models using Gaussian models of simple data-transformations, such as taking the logarithm of fitness plus one (e.g., Kruuk et al. 2000). In any GLM (or data-transformed linear model), the latent-scale estimates obtained may be biologically difficult to interpret (Mitchell-Olds and Shaw 1987; Hendry et al. 2018). However, it is now possible to transform estimates of latent-scale parameters from GLMs, and in particular from GLAMs, back to the required data scale. An important recent step forward has been the development of general methods to back-transform parameters on the latent scales of generalized linear mixed models (including GLAMs) to the data scales (Morrissey 2015; de Villemereuil et al. 2016).

We consider here the case in which expected fitness is log-normal, so that log transformation yields a Gaussian distribution. GLMs of common distributions such as the Poisson or negative binomial typically use a logarithmic link function. Importantly, for a GLAM of *absolute* fitness, back-transformation of values of latent-scale additive genetic variance can correctly estimate the additive genetic variance in *relative* fitness on the data scale (de Villemereuil et al. 2016), which is the rate of increase in relative fitness *within* a generation. However, in a sense, estimates generated from back-transformation of parameters from GLAMs may not be appropriate for prediction of data-scale change in mean fitness *between* generations. This is because the link functions used in GLAMs can introduce nonlinearities of breeding values. so that perfectly additive genetic effects on the latent scale translate into both additive *and* nonadditive genetic effects (epistasis and dominance) on the data scale (Morrissey 2015). The presence of nonadditive genetic effects means that the change within a generation is not accurately transmitted to the next generation. As a result, the inheritance of fitness is distorted, and the increase in mean fitness in the next generation exceeds that due strictly to selection alone (Morrissey 2015).

This problem can now be circumvented using an analog to Fisher's fundamental theorem for log-normal fitness, proposed by Morrissey and Bonnet (2019), which predicts the proportional change in mean fitness (absolute or relative if the same reference is kept across generations) on the data scale, based on the amount of additive genetic variance on a logarithmic latent scale while accounting for inheritance distortion: $\Delta\bar{w} = \exp(\sigma^2_A(l)) - 1$, where l is absolute fitness on the latent scale. Poisson GLAMs using a logarithm link function to model fitness match this theoretical framework because they assume that expected values of fitness follow a log-normal distribution. Thus if $\sigma^2_A(l)$ is the estimate of latent-scale additive genetic variance for fitness from a Poisson GLAM, the rate of increase in mean fitness—or the rate of adaptation—can be predicted as $\exp(\sigma^2_A(l) - 1)$ (Morrissey and Bonnet 2019).

In this article, we use individual-based simulations to verify the FTNS and this recent analog when fitness follows a log-normal distribution. We generate pedigrees and individual-fitness data with known underlying genetic variance and statistical distribution, across several generations. We then start by illustrating key relationships between the population dynamics and the quantitative genetic parameters of log-normal fitness that have not yet been described in the evolutionary literature or may not be well known to evolutionary biologists. In particular, we show how additive genetic variance in absolute fitness on the latent scale $\sigma^2_A(l)$ translates into additive and nonadditive genetic variance in relative fitness on the

data scale, which in turn predicts the rate of increase in mean fitness across generations due to selection. We confirm that the rate of adaptation is predicted by $\exp(\sigma^2_A(l)) - 1$ and approximated by $\sigma^2_A(w) = \sigma^2_A(l)$. In contrast, we also illustrate how the heritability of fitness on the data scale is not a good measure of the rate of adaptation. Second, we compare the performance of animal models of the simulated log-normal fitness data for the estimation of additive genetic variance in relative fitness. Although the heritability of fitness does not figure directly in the mechanics of evolutionary adaptation, it remains an intuitive representation of the amount of genetic variability (in a strictly limited sense, Hansen et al. 2011) that has a wide appeal. We therefore also consider how well heritability of relative fitness is estimated by the different empirical modeling options. Finally, we consider previous empirical estimates of $\sigma^2_A(w)$ in the light of our results and discuss what is known about the rate of adaptation in natural populations.

Methods

We used individual-based simulations to generate fitness data of known properties across several generations. Individual-based simulations let us model individual fitness and population dynamics, and therefore allow us to link the quantitative genetics of fitness to the dynamics of population growth rate. This provides an intuitive link between the adaptive process and population dynamics and also lets us confirm the analog of the fundamental theorem for log-normal fitness (Morrissey and Bonnet 2019).

First, we simulated data using our chosen values for additive genetic and environment variance in fitness, assuming fitness follows a log-normal distribution. Simulations were performed in an enhanced version of the C++ program Volator (Bonnet and Postma 2018). From the simulations, we extracted realized breeding values, environmental values, and hence phenotypes for individual fitness, parent–offspring relationships between individuals (i.e., pedigrees), and population size through time. From these, we could confirm and illustrate the relationship between the fundamental parameters of additive genetic variance, environmental variance, and heritability (we mean "narrow-sense" heritability, here and throughout in this publication) in fitness on different scales, and the rate of increase in mean fitness.

In a second step, we used the individual-fitness data and pedigrees to fit animal models and compare the input simulation parameter values to the parameters estimated by animal models. We did this using either Poisson or Gaussian fitness distributions.

Simulations and Theoretical Expectations

Life-Cycle and Fitness Model

We simulated nonoverlapping generations, with each generation equivalent to 1 time step (or year). A year starts with adult individuals only, males and females. Adults reproduce in a 2-sex panmictic way and then die. Their offspring, born with an even sex ratio, then all recruit as reproducing adults for the next year.

We model a single trait, latent-scale reproductive success (l), which is also the latent-scale fitness in our model (because we do not simulate other fitness components). The trait follows the infinitesimal quantitative genetic model (Fisher 1918; Barton et al. 2017; Walsh and Lynch 2018, chapter 24) with additive effects of genes and the environment on a latent scale. Breeding values are transmitted to the progeny as the mean breeding value of parents plus a Mendelian segregation variance equal to half the simulated additive genetic variances.

For an individual i,

$$l_i = \mu + a_i + e_i,$$

where μ is a population mean for l, a is the latent breeding value, and e is the latent environmental deviation (which we assume is due to the combined action of a large number of plastic responses and developmental noise). a and e follow a multivariate normal distribution with means equal to zero, variances equal to $\sigma^2_A(l)$ and $\sigma^2_E(l)$, and zero covariance between them; note that each of them contribute "overdispersion" to the distribution of fitness, as they ultimately generate variance on the data over and above that due to the random Poisson process. Table 1 summarizes the definitions of the main parameters. We keep $\sigma^2_A(l)$ constant across generations, following the infinitesimal model, which assumes that genetic variation in the trait of interest is due to a large number of genetic loci, so that any response to selection has a negligible effect on allele frequencies and on the total genetic variance (Barton et al. 2017). $\sigma^2_E(l)$ is also constant across generations.

Realized reproductive success on the data scale (W, Figure 1 bottom row) is simulated as an overdispersed Poisson process generated by a Poisson random draw with heterogeneous expectation: $W_i = P(\exp(l_i))$ where $\exp(l_i)$ is the exponential of l_i and the expected reproductive success for individual i on the expected data scale (see Figure 1 for an illustration of the 3 scales).

Fisher's FTNS itself does not make stringent assumptions (Grafen 2018), but assumptions are necessary to directly observe the theorem from population dynamics as we do here. Indeed, Fisher's theorem

Table 1. Definition of the main parameters and the scales on which they are expressed

Symbol	Scale	Meaning
l	Latent	Absolute fitness on the latent scale
$\exp(1)$	Expected	Exponential of l, and absolute fitness on the expected scale
W	Data	Absolute fitness on the data scale
$\bar{W}$	Data	Mean W in a given generation
$\bar{W}_0$	Data	Mean W in the first generation
w	Data	Relative fitness on the data scale ($W/\bar{W}$)
$\sigma^2_A(l)$	Latent	Latent-scale additive genetic variance in fitness
$\sigma^2_E(l)$	Latent	Latent-scale environmental variance in fitness, also the amount of nongenetic overdispersion
$\sigma^2_A(w)$	Data	Additive genetic variance in relative fitness on the data scale
$\sigma^2_G(w)$	Data	Total genetic variance in relative fitness on the data scale, generated by latent-scale additive genetic variance $\sigma^2_A(l)$[a]

The 3 different scales are illustrated in Figure 1.

[a]Note that σ^2_G excludes contributions from any nonadditive genetic variance on the latent scale, as we do not consider these here.

relates to the part of change in mean fitness that is due to selection, but other evolutionary forces could change mean fitness too. On the latent scale, that is, for l, we assume no other evolutionary force than selection and genetic drift are present, so that *on average* Fisher's theorem can be observed directly from changes in mean fitness, without having to isolate the independent role of selection. We consider nonoverlapping generations because overlapping generations can influence the rate of response to selection (which is adaptation when the trait responding to selection is fitness) in a population (Hill 1974). We also, again for simplicity, assume no competition, no frequency or density dependence, no changes in the mean environmental values of phenotypic variation, no gene-by-environment interactions, no mutations, and no immigration. Despite these assumptions, inheritance distortion appears as one more evolutionary force when considering the data scale rather than the latent scale and must be accounted for to predict the change in data-scale mean fitness across generations (see below).

Furthermore, again for simplicity, we consider only genetic variance in fitness in females (following equations presented below). Having determined individual females' reproductive success according to the variation in fitness, paternities are assigned randomly for each offspring. Thus, male reproductive success varies entirely stochastically, but note that males still carry and transmit breeding values for the trait of female fitness.

Simulation Parameters

We started simulations with 100 females and 100 males, and an expected initial population growth rate of $\lambda_0 = 0.96$. Population size was not regulated (i.e., fitness was density-independent), and the expected population growth rate, therefore, tended to increase in the presence of additive genetic variance in fitness. We simulated 6–11 time steps ("years" thereafter), reducing the number of years for scenarios with high additive genetic variance in fitness, because the population growth was very fast and simulations became computationally demanding with large $\sigma^2_A(l)$.

We simulated $\sigma^2_A(l)$ with values {0, 0.01, 0.05, 0.1, 0.3}, and $\sigma^2_E(l)$ with values {0, 0.1, 0.5}. For each of the 15 combinations of $\sigma^2_A(l)$ and $\sigma^2_E(l)$ values, we simulated 100 data sets. For each latent phenotypic variance ($\sigma^2_P(l) = \sigma^2_A(l) + \sigma^2_E(l)$), we altered the latent mean fitness, so that the initially expected data scale mean fitness was always $\bar{W}_0 = 1.92$ (corresponding to $\lambda_0 = 0.96$). Some distributions produced by the combinations of latent means and variances are shown in Supplementary Figure 2.

We compared the realized parameters in the simulated data sets with the original simulated values. The realized variance in simulated breeding values for fitness on the latent Poisson scale matched the pre-set values for all 3 levels of environmental variance (Supplementary Figure 1).

Step 1: Calculating Genetic Variances and Rate of Adaptation on the Data Scale

Conveniently, the additive genetic variance in absolute fitness on the latent scale is exactly equivalent to the additive genetic variance in relative fitness on the data scale ($\sigma^2_A(l) = \sigma^2_A(w)$) in a log-link GLAM (deduced from equation 7 in Morrissey (2015). However, this quantity does not perfectly predict changes in mean fitness on the data scale when fitness is log-normal. This is because the curvature of the exponential transforms some of the additive genetic variance on the latent scale into nonadditive genetic variance on the data scale. To put it simply, on the latent scale, the breeding value of an individual is on average the mean of its parents' breeding values on the latent scale, but this expectation does not survive the exponential-transformation. The phenotype corresponding to an offspring's expected breeding value on the latent scale is larger than the mean of its parents' expected phenotypes. Therefore, mean fitness on the data-scale increases faster than $\sigma^2_A(w)$, specifically at a rate of $\exp(\sigma^2_A(l)) - 1$ (Morrissey and Bonnet 2019). Incidentally, $\exp(\sigma^2_A(l)) - 1$ happens to be the genetic variance for relative fitness ($\sigma^2_G(w)$; on the data scale) that originates from latent additive genetic variance in the absence of nonadditive variance on the latent scale (Morrissey and Bonnet 2019).

In our simulations, for every data set, we calculated the realized $\sigma^2_G(w)$ across generations as the variance in the individual genotypic values (for an individual i, its data-scale genotypic value for W is $\exp(\mu + a_i + \sigma^2_E(l)/2)$, see Morrissey and Bonnet 2019). We measured realized adaptation using population growth rates. At a given time t with a population size N_t, the population growth rate is $\lambda_t = \frac{N_{t+1}}{N_t}$. The rate of adaptation is calculated as $\frac{\lambda_{t+1}-\lambda_t}{\lambda_t}$. This is equivalent to the relative change in mean individual fitness $\frac{\bar{W}_{t+1}-\bar{W}_t}{\bar{W}_t}$ because only one fitness component is modeled, and generations are nonoverlapping. Therefore, we expect on average $\frac{\lambda_{t+1}-\lambda_t}{\lambda_t} = \exp(\sigma^2_A(l)) - 1$ at any time t. Because in our simulations only females express genetic variation in fitness, we expected the increase in population growth rate to be half that predicted by the equations above.

Step 2: Statistical Estimation From Simulated Data

Animal Models

We used animal models (Henderson 1950; Kruuk 2004) to estimate $\sigma^2_A(l)$, $\sigma^2_E(l)$, and derived parameters from simulated fitness and pedigree data. For each data set, we fitted 1) a linear animal model with a Gaussian error structure and 2) a generalized animal model assuming an overdispersed Poisson distribution and a logarithmic link function, with the response variable of absolute fitness in both cases. Each data set included at least 800 females, across 5–10 generations, for data sets simulated with $\sigma^2_A(l) = 0$. More females were included with increasing values of $\sigma^2_A(l)$ because adaptation caused the population sizes to increase across generations. We used a maximum of 3000 females for each data set (selecting the first 3000 first females born in the population). The analysis involved only female phenotypes (as males do not express phenotypes for the trait, see above), but used the full pedigree to calculate the relatedness matrix.

We fitted models in MCMCglmm, an R package that by default includes an additive overdispersion parameter in Poisson models (Hadfield 2010). We extracted the posterior distributions of the estimates of the intercept, additive genetic variance, and residual (environmental) variance. In the case of the Poisson models, these are direct estimates of the simulation parameters $\sigma^2_A(l)$, $\sigma^2_E(l)$, and μ; as outlined above, the estimate of $\sigma^2_A(l)$ is also an estimator of $\sigma^2_A(w)$ (Morrissey 2015; de Villemereuil et al. 2016). Transformations back to the data scale were performed on the posterior distribution using the R package QGglmm (de Villemereuil et al. 2016), to generate posterior distributions of data scale $\sigma^2_A(w)$ and hence also of data-scale narrow-sense heritability (h^2). In the case of Gaussian models, we divided the estimated additive genetic variance in absolute fitness by the square of the mean fitness in the first generation (using the square of the estimated intercept gave qualitatively identical results) $\sigma^2_A(W)/\bar{W}^2_0$ to obtain an estimator of $\sigma^2_A(w)$. Indeed, animal models

estimate genetic parameters relative to a base population, which in our case exactly matches the first generation. Therefore, the estimate of $\sigma^2_A(W)$ is for the first generation and the conversion to $\sigma^2_A(w)$ must use mean fitness in the first generation. Note however that in our simulations $\sigma^2_A(l)$, and hence $\sigma^2_A(w)$, are constant across generations, which therefore requires $\sigma^2_A(W)$ to increase at the same rate as $\bar{W}^2$. As a consequence, the Gaussian estimate of $\sigma^2_A(W)$ is valid for the first generation only, but we expect the estimate of $\sigma^2_A(w)$ to be valid for any generation.

We used MCMCglmm's default inverse gamma priors with shape and rate parameters both equal to 0.001 (equivalent to a variance and degree of belief equal to 1 and 0.002, respectively). We ran models for 130 000 Markov chain Monte Carlo iterations, with a burn-in of 30 000 and thinning of 100. Visual inspection of a subset of models suggested that a stationary sampling distribution was always reached before the end of the burn-in.

Comparison Between Simulated and Estimated Parameter Values

We extracted the modes and means of the posterior distributions for Gaussian and for Poisson animal models for each simulated data set and plotted them against the simulated parameter values. However, although this approach provides simple visualization, it does not represent estimation performance fully, as it considers only point estimates and no measure of uncertainty. To address this, we therefore also considered coverage and root mean squared error (RMSE) to measure the quality of the estimators of $\sigma^2_A(w)$. For a given set of parameters, coverage was calculated as the proportion of models having a 95% highest probability density credibility interval containing the true value. We did not consider coverage for scenarios simulated with $\sigma^2_A(l) = 0$ because credibility intervals cannot contain exactly zero for variance components (although in practice they might due to arbitrary rounding).

Results

Theoretical Expectations

We start by comparing the observed values of data-scale genetic variance, and also rates of population growth, with the expectations based on simulated levels of latent-scale additive genetic variance. The additive genetic variance in relative fitness on the data scale $\sigma^2_A(w)$, calculated with QGglmm, was, as expected, exactly equal to $\sigma^2_A(l)$. The total genetic variance in relative fitness on the data scale $\sigma^2_G(w)$ calculated from the simulated breeding values matched the theoretical expectation of $\exp(\sigma^2_A(l)) - 1$ (blue lines in Figure 2). At lower values of $\sigma^2_A(l)$, $\sigma^2_G(w)$ was also close to $\sigma^2_A(w)$ (which is equal to $\sigma^2_A(l)$), but it was visibly larger at larger values of $\sigma^2_A(l)$ (red lines in Figure 2). The environmental variance $\sigma^2_E(l)$, which generates overdispersion of the distribution over and above that due to latent genetic differences, had no apparent effect on either the expectation or the spread of the realized values. However, there was considerable spread in the realized values for $\sigma^2_G(w)$ for the largest simulated $\sigma^2_A(l)$ (Figure 2A).

The expectation $\exp(\sigma^2_A(l)) - 1$ (blue lines) was also matched by the observed rate of increase in mean fitness, measured as the proportional change in population growth rates (Figure 2B; note that demographic stochasticity explains the occurrence of negative values). There was a large spread of realized changes in population growth rate, and the spread appeared to increase with environmental variance within each level of simulated $\sigma^2_A(l)$ (Figure 2B). However,

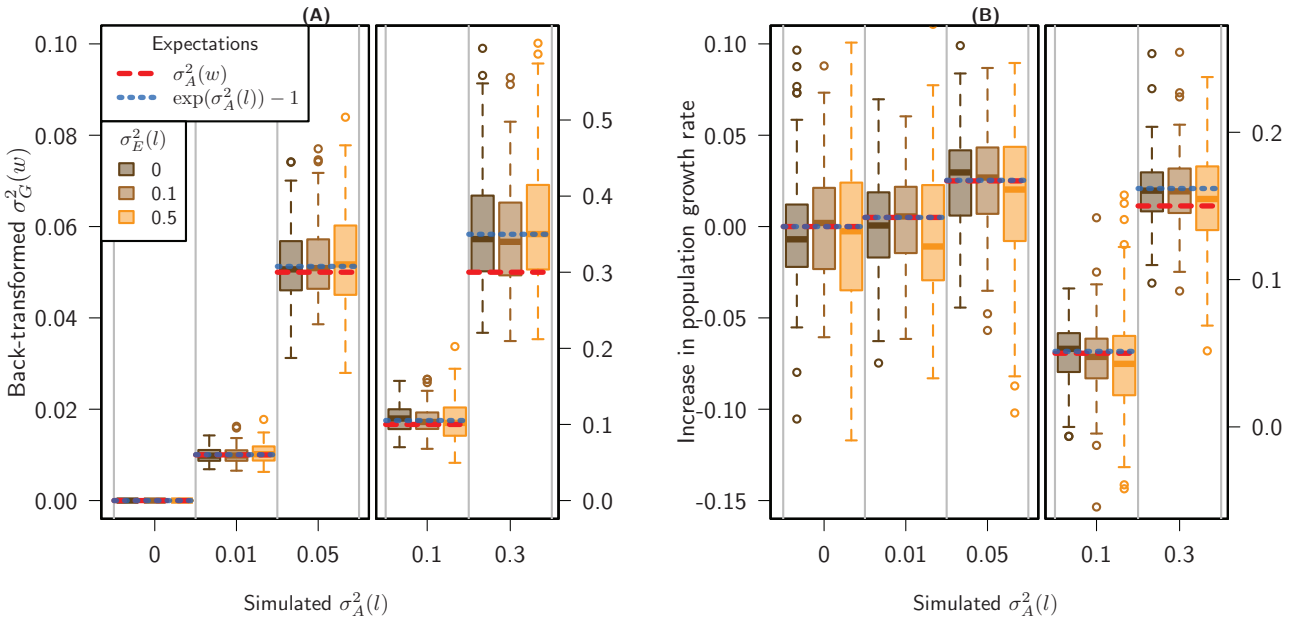

Figure 2. Comparison of realized measures of adaptation in simulations with theoretical expectation: (**A**) realized values of total genetic variance in data-scale relative fitness calculated from simulated breeding values, and (**B**) relative change in population growth rate. The *x* axis shows increasing levels of latent-scale additive genetic variance, $\sigma^2_A(l)$; note that, in both (**A**) and (**B**), the break reflects different *y* axis scales for lower (0, 0.01, and 0.05) versus higher (0.1, 0.3) values of $\sigma^2_A(l)$. The differentially shaded boxes show increasing levels of latent-scale environmental variance, $\sigma^2_E(l)$. Posterior modes were used as point estimates. The box-plots show the percentiles 0.025, 0.25, 0.5, 0.75, and 0.975, and outliers from the posterior modes of the 100 simulated data sets. The red lines indicate the data-scale additive genetic variance in relative fitness: $\sigma^2_A(w) = \sigma^2_A(l)$ in (**A**) or $\sigma^2_A(w)/2 = \sigma^2_A(l)/2$ in (**B**). The blue lines indicate the theoretical expectation for the proportional rate of increase in mean fitness on the data scale: $\exp(\sigma^2_A(l)) - 1$ (for **A**) or $\exp(\sigma^2_A(l)/2) - 1$ (for **B**). Division by the factor of 2 in (**B**) reflects the fact that only females express genetic variance in fitness, so expected rates of adaptation need to be halved.

in general, the simulations support the prediction of the analog of the FTNS, that is, the rate of increase in population growth rate across generations is expected to be $\exp(\sigma^2_A(l)) - 1$.

In general, the realized values of data-scale heritability did not match the expected rate of adaptation ($\exp(\sigma^2_A(l)) - 1$) well (Supplementary Figure 3). The exception was when heritability was 0 when $\sigma^2_A(l) = 0$, which was the only case where heritability matched the rate of adaptation (which was zero in that case). Otherwise, with any positive value of $\sigma^2_A(l)$, latent-scale heritability was always 1 in the absence of environmental variance ($\sigma^2_E(l) = 0$) and decreased with increasing $\sigma^2_E(l) > 0$. For instance, for $\sigma^2_A(l) = 0.1$, latent-scale heritability can be 1 ($\sigma^2_E(l) = 0$), 0.5 ($\sigma^2_E(l) = 0.1$), or 0.17 ($\sigma^2_E(l) = 0.5$) (Figure 4B). However, the theoretical rate of adaptation ($\exp(\sigma^2_A(l)) - 1$) is independent of $\sigma^2_E(l)$ and, in the previous example, is approximately 0.105 (to be divided by 2 to recover the true rate of increase in mean fitness, given that we consider only females here). The results thus underline the expectation that the data-scale heritability does not provide a decent estimate of the expected rate of adaptation.

Animal Model Estimation

Additive Genetic Variance

Our second set of analyses compared estimates returned from animal models of the simulated data with the original simulated parameter values. As expected from Morrissey (2015), for Poisson models QGglmm converted an amount of additive genetic variance for relative fitness on the data scale exactly equal to the estimate of the additive genetic variance for absolute fitness on the latent scale ($\sigma^2_A(w) = \sigma^2_A(l)$). The back-transformation is mathematically trivial.

Gaussian animal models estimated $\sigma^2_A(w)$ close to the expected values of $\sigma^2_A(l)$ for all simulated values, except $\sigma^2_A(l) = 0.3$ where they overestimated $\sigma^2_A(w)$ by a factor 2–4. Estimates were slightly biased downward for $\sigma^2_A(l) = 0.01$ and 0.1 (Figure 3). Gaussian animal model estimates were dependent on environmental variance, so that the estimate of $\sigma^2_A(w)$ decreased with increasing $\sigma^2_E(l)$ (Figure 3). Poisson models estimated $\sigma^2_A(w)$ close to the expected values of $\sigma^2_A(l)$ for all simulated values, but were slightly biased downward for $\sigma^2_A(l)$ of 0.01 and 0.05. The estimates from Poisson animal models were on average not affected by $\sigma^2_E(l)$, but they became more variable with increasing values of $\sigma^2_E(l)$ (Figure 3). Both Gaussian and Poisson models showed an upward bias for $\sigma^2_A(l) = 0$.

Next, we compared the performance of the different types of models in terms of coverage and uncertainty. RMSE was small and almost independent of the simulated values for Poisson models, except for a slight increase at $\sigma^2_A(l) = 0.05$. In contrast, for Gaussian models, RMSE increased with increasing $\sigma^2_A(l)$ up to a maximum of 21, 3 orders of magnitude above the maximum RMSE of Poisson models (Supplementary Figure 4). For $\sigma^2_A(l) = 0.01$, both Poisson and Gaussian models had mediocre coverage, of about 0.6 (Supplementary Figure 5), although Poisson performed slightly better. For larger simulated values, Poisson models clearly did better than Gaussian models, with their 95% confidence interval (CI) containing the simulated value in about 95% of the data sets, and their coverage was little affected by the amount of nongenetic overdispersion ($\sigma^2_E(l)$). In contrast, the Gaussian models had coverage well below 95%, especially for $\sigma^2_A(l) = 0.3$. In addition, coverage was affected by nongenetic overdispersion (Supplementary Figure 5, especially apparent at $\sigma^2_A(l) = 0.1$) in a counter-intuitive way: with Gaussian models, the estimation apparently improved as the data become noisier, being least bad at $\sigma^2_E(l) = 0.5$). Extrapolating from this trend, one might expect the estimation to get worse again as $\sigma^2_E(l)$ increases beyond the simulated values.

Heritability Estimates

The true narrow-sense heritability on the latent scale is 1 when $\sigma^2_E(l) = 0$ and $\sigma^2_A(l) > 0$, and decreases with increasing values of $\sigma^2_E(l)$. High heritabilities on the latent scale translate to medium

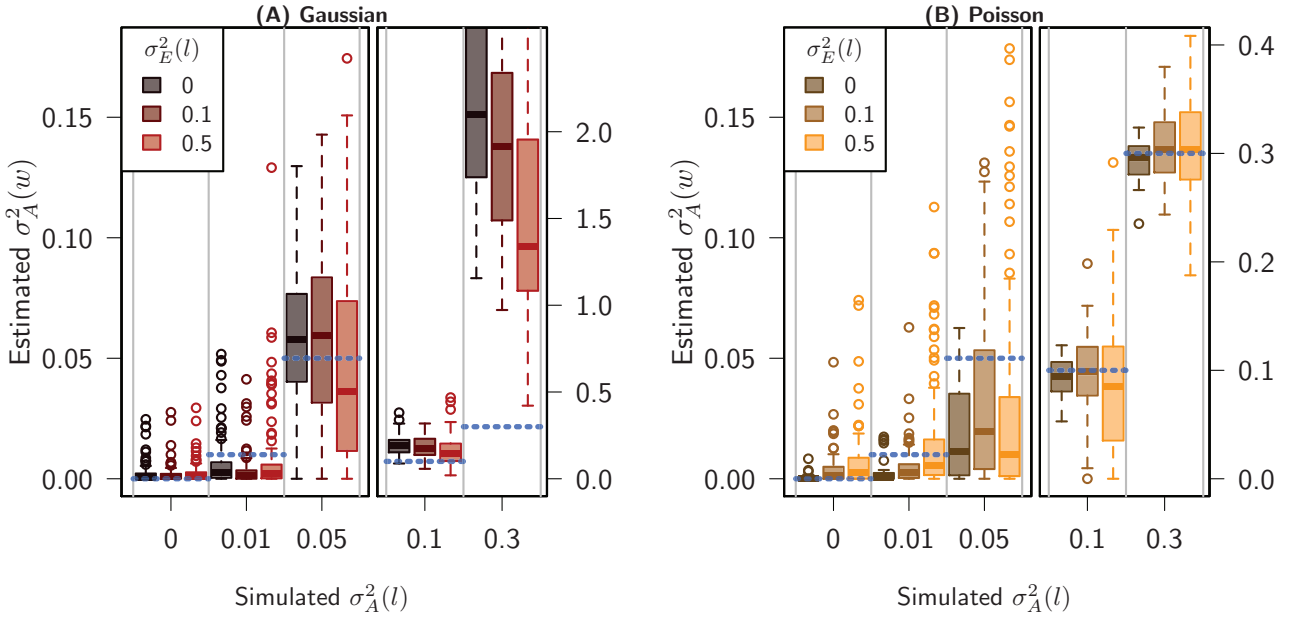

Figure 3. Modes of the posterior distribution of additive genetic variance in data-scale relative fitness $\sigma^2_A(w)$ for different simulated values of latent-scale $\sigma^2_A(l)$ (x axis) and for different simulated values of $\sigma^2_E(l)$ (nongenetic overdispersion; different shades), for (**A**) Gaussian and (**B**) Poisson animal models. Box-plots show the percentiles 0.025, 0.25, 0.5, 0.75, and 0.975, and outliers from the posterior modes of the 100 simulated data sets. The bright blue horizontal lines indicate simulated $\sigma^2_A(l)$, which should equate to the realized $\sigma^2_A(w)$.

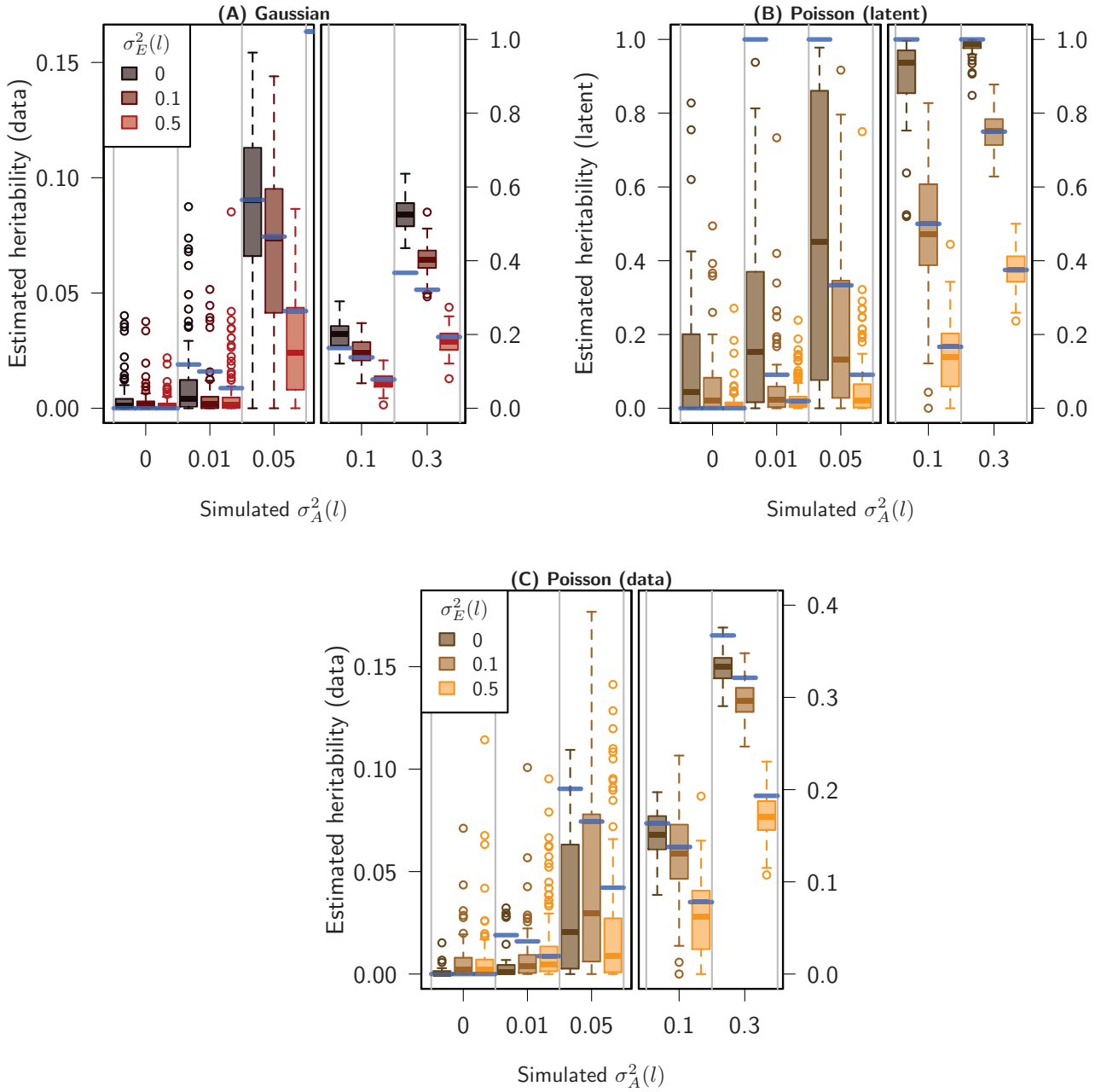

Figure 4. Narrow-sense heritability of fitness estimated from animal models (distribution of posterior modes, reddish box-plots), compared with theoretical expectation (blue lines). (**A**) Data-scale heritability estimated from a Gaussian animal model; (**B**) latent-scale heritability estimated from a Poisson animal model; and (**C**) data-scale heritability estimated from a Poisson animal model. Heritabilities are shown for different values of simulated additive genetic variance in latent fitness ($\sigma^2_A(l)$, *x* axis) and environmental variance ($\sigma^2_E(l)$, tints of red or orange). Expectations for heritabilities on the data scale were computed from QGglmm using the simulated values for latent variance components.

values on the data scale (maximum 0.37 in our simulations, Figure 4) due to the stochasticity added by the Poisson process. Poisson GLAMs gave correct estimates of latent heritability for the higher values of $\sigma^2_A(l)$, but were less reliable for the lower values (Figure 4B). On the data scale, Poisson animal models returned accurate estimates of the data-scale heritability or slightly underestimated it (Figure 4C). The Gaussian animal models gave estimates close to the expected values in all cases, except for scenarios with high $\sigma^2_A(l)$ and low $\sigma^2_E(l)$ (Figure 4A).

Discussion

In this article, we have used simulations to test the relationship between levels of genetic variance in fitness and rates of population adaptation and growth when individual fitness follows a log-normal distribution. Our results confirmed the new theoretical result (Morrissey and Bonnet 2019) that for log-normal fitness, the proportional increase in mean (absolute or relative to the previous generation) fitness on the data scale is predicted by the

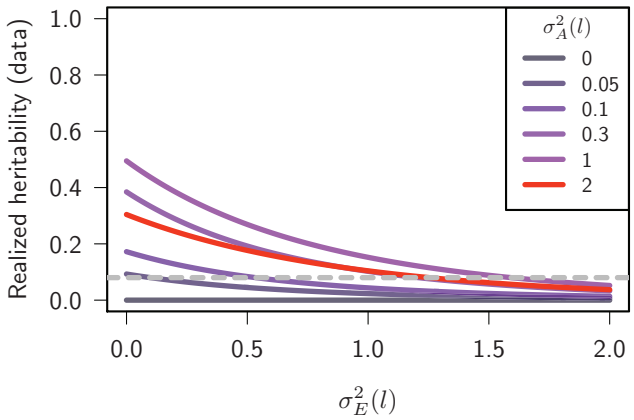

Figure 5. Realized heritability on the data scale against latent environmental variances ($\sigma^2_E(l)$) for several values of latent additive genetic variance ($\sigma^2_A(l)$). The theoretical rate of adaptation is a function of $\sigma^2_A(l)$ only, whereas heritability depends on both $\sigma^2_A(l)$ and $\sigma^2_E(l)$. Data-scale heritabilities were computed by QGglmm, assuming a latent intercept of log(2). $\sigma^2_A(l)$ values above 0.3 are biologically unlikely, because they correspond to extremely fast rates of adaptation, but illustrate how it is theoretically possible for heritability to decrease as $\sigma^2_A(l)$ increases while holding $\sigma^2_E(l)$ constant. This is possible because increased latent genetic variance increases the mean on the expected data scale, which then generates increased Poisson variance on the data scale. The gray dashed line highlights that a data-scale heritability of 0.08 could correspond to any value of $\sigma^2_A(l)$ between 0.05 and 2.

exponential of the additive genetic variance of latent fitness, minus one ($\exp(\sigma^2_A(l)) - 1$). The expression $\exp(\sigma^2_A(l)) - 1$ captures both the pure effect of adaptation, that is, Fisher's FTNS, and the inheritance distortion due to the nonlinear genotype–phenotype map here (for a more detailed explanation, see Morrissey and Bonnet (2019). For small values of latent additive genetic variance, the increase in fitness is well approximated by $\sigma^2_A(w)$ (or equivalently $\sigma^2_A(l)$), but the conversion to the exact expectation is trivial ($\exp(\sigma^2_A(l)) - 1$) and should be used. In addition, the analysis of our simulations has illustrated the fact that narrow-sense heritability, on either scale, is not a useful measure of adaptation. We confirmed that animal models assuming an overdispersed Poisson distribution for fitness estimate $\sigma^2_A(w)$ well. With the same data, Gaussian models gave reasonable point estimates for lower values of $\sigma^2_A(w)$, but estimation was poor for larger values.

What Parameter Is Relevant to Predict Adaptation?

Our results illustrate which parameters are or are not relevant to assessing the rate of adaptation in a population. It has been argued that the highly stochastic nature of variation in individual fitness means that 1) natural selection could rarely manifest itself in small wild populations, 2) the heritability of fitness would generally be small, and therefore that 3) adaptive evolution should be limited and its study obscured by environmental variation (Orzack et al. 2011; Steiner and Tuljapurkar 2012; Hamel et al. 2018). The first conclusion is incorrect, as well-specified models have identified natural selection in numerous small populations (Kingsolver et al. 2001; Kingsolver and Diamond 2011; Bonnet and Postma 2016; Authier et al. 2017). In contrast, the second point is supported by data (Gustafsson 1986; Mousseau and Roff 1987; Kruuk et al. 2000; Teplitsky et al. 2009; Postma 2014; McFarlane et al. 2014) and by our simulations. However, the low heritability of fitness is typically driven by high levels of environmental variance (Price and Schluter 1991). In the case of a Poisson distribution, even a perfect heritability of 1 for values of the expected scale (see Figure 1) does not translate into a large heritability on the observed scale of the data (Figure 5) because the Poisson process adds stochasticity on top of the expected values (de Villemereuil et al. 2016). To reillustrate this point, we computed theoretical data-scale heritabilities using QGglmm and its relationship with adaptation (Figure 5). Note that it is theoretically possible for heritability to decrease as $\sigma^2_A(l)$ increases (for a given value of $\sigma^2_E(l)$). This is possible because increased latent genetic variance increases the mean on the expected data scale, which then generates increased Poisson variance on the data scale. Thus, somewhat counter-intuitively, data-scale heritability for $\sigma^2_A(l) = 2$ will be lower than for $\sigma^2_A(l) = 1$ (Figure 5), although the inversion occurs only for extremely high and unrealistic variance components. Thus, both these theoretical arguments and the simulation results above both illustrate how heritability of fitness is not a good measure of adaptive evolution. As a result, the third point (lack of adaptation) above cannot be deduced from the second point (low heritability of fitness). Rather than the heritability, it is the additive genetic variance in fitness that controls adaptive evolution under the predictions of Fisher's FTNS (Fisher 1930; Queller 2017; Morrissey and Bonnet 2019).

In the specific case of log-normal fitness, the expected rate of increase in fitness is $\exp(\sigma^2_A(l)) - 1$ (Morrissey and Bonnet 2019), which combines strict adaptation with the effects of inheritance distortion due to the nonlinearity of breeding values. This result does not invalidate Fisher's FTNS in the case of log-normal fitness. The FTNS is still appropriate to measure the increase in mean fitness within a generation, or the pure effect of adaptation between generations, as $\sigma^2_A(w)$. However, with log-normal fitness, adaptation comes with an inevitable supplementary increase in mean fitness due to inheritance distortion. Conceptually this is an opposite analog of the deterioration of genetic effects (which Fisher confusingly termed "environmental" deterioration, although it is in part the result of change in allele frequencies and indirect genetic effects, rather than the abiotic environment) where the phenotypic direct effect of genetic change is impeded by the change in the genetic "environment" across generations (Price 1972b). Here the direct effect of genetic change is enhanced by change in the interactive effects of other genes: we can speculate Fisher may have called it "environmental amelioration." This supplementary increase is not strictly speaking an effect of genetic adaptation but is intimately related to it and must be accounted for an accurate prediction of change in mean fitness. However, for small values of latent additive genetic variance ($\sigma^2_A(l) << 1$), the supplementary increase will not be substantial: $\sigma^2_A(l)$ is a reasonable approximation of the rate of increase in mean fitness because $\exp(\sigma^2_A(l) - 1)$ tends toward $\sigma^2_A(l)$ as $\sigma^2_A(l)$ tends toward zero (e.g., the difference is below 5% of $\exp(\sigma^2_A(l) - 1)$ when $\sigma^2_A(l) < 0.1$).

Implications for Empirical Studies

According to the literature search in Hendry et al. (2018), to date, there are 22 studies estimating $\sigma^2_A(w)$ from good measures of lifetime fitness in wild plant and animal populations. Estimates range from 0 to 0.85, with 89% below 0.2, and 77% below 0.1. Most of these estimates were produced by Gaussian models (animal models, regressions, or structural equation models), while lifetime fitness measures are generally unlikely to have followed Gaussian distributions (although some studies used Gaussian models after log transformation of fitness data; e.g., Kruuk et al. 2000). The reliability of these Gaussian models using log-transformed data remains to be explored, but could

be analyzed using the framework proposed in de Villemereuil et al. (2016) and de Villemereuil (2018). Hendry et al.'s review did identify 4 studies (Milot et al. 2011; McFarlane et al. 2014, 2015; Wolak et al. 2018) estimating additive genetic variance in absolute fitness or its heritability using Poisson models. However, only one of those reported the data-scale variance in absolute fitness ($\sigma_A^2(W)$) (Wolak et al. 2018) and none treated their estimates of $\sigma_A^2(l)$ as estimates of the data-scale additive genetic variance in relative fitness ($\sigma_A^2(w)$), presumably because it was not yet well known that $\sigma_A^2(l) = \sigma_A^2(w)$.

We found that Gaussian models have suboptimal estimation properties for $\sigma_A^2(w)$ when fitness follows an overdispersed Poisson distribution. Thus, Gaussian estimates can be biased, in complicated ways that depends on the level of overdispersion ($\sigma_A^2(w)$ and $\sigma_E^2(w)$) in the data (Figure 3); CIs generally do not contain the parameter value for large simulated $\sigma_A^2(w)$ (Supplementary Figure 5); and estimates come with considerable variance, as evidenced by the high RMSE (Supplementary Figure 4). Estimation problems probably arise because of the skewness of the fitness data distribution and of the breeding values. As we simulated fitness data with a constant mean, increasing overdispersion leads to an increase in the proportion of zeroes and of high values (Supplementary Figure 2), which results in increasingly deviation from the bell shape of a Gaussian distribution. However, these problems were substantial only for "large" values of $\sigma_A^2(w)$ (and hence "large" skewness in breeding values) and may not meaningfully hamper biological interpretation when considering "small" $\sigma_A^2(w)$ values. In our study, Gaussian models were qualitatively correct, that is, they detected the presence of genetic variance when the simulated $\sigma_A^2(w) > 0$ and did not estimate large genetic variance when none was simulated: the largest point estimate was 0.029 and 95% of point estimates were below 0.01 (the next lowest simulated value). Besides, estimation was reasonable for small parameter values (Supplementary Figures 4 and 5). What should be considered "large" or "small" $\sigma_A^2(w)$ then?

Any threshold between "small" and "large" will be arbitrary. Nevertheless, a value of $\sigma_A^2(w) > 0.1$ is clearly large because it suggests an increase in mean fitness of 10% due to adaptation within a generation. If such increase was sustained for a few generations and not counterbalanced by changes in the environment (in the very broad sense of Fisher's "environmental" deterioration), it would lead to an explosive population increase. Published estimates, although potentially not perfectly reliable, suggest that most values are below 0.1 and values of $\sigma_A^2(w) > 0.2$ are rare (Hendry et al. 2018). Therefore, our results suggest that cautious optimism is reasonable when considering past and future estimates based on Gaussian models of untransformed fitness measures. Gaussian models appear fairly reliable for values that are small and likely to actually exist (for instance $\sigma_A^2(w) < 0.1$). Nevertheless, the more accurate and precise Poisson models using a log-link should be preferred whenever data are assumed to follow a Poisson or overdispersed Poisson distribution.

Indeed, we have illustrated here how, when fitness can be described with a Poisson distribution, all the tools are now available to fit quantitative genetic models and extract meaningful estimates of components of variance (de Villemereuil et al. 2016; Morrissey and Bonnet 2019). The interpretation of variance components from GLAMs (and other generalized linear mixed models) generally requires back-transformations involving integrations. These back-transformations are now supported by R packages (de Villemereuil et al. 2016), but these may still be troublesome or intimidating. However, in the case of Poisson animal models, the latent estimate of additive genetic variance is exactly the additive genetic variance for relative fitness ($\sigma_A^2(w) = \sigma_A^2(l)$), meaning that, conveniently, back-transformation is unnecessary and the latent-scale model estimates are directly meaningful. In addition, the expected increase in mean fitness due to adaptation and to inheritance distortion is simply $\exp(\sigma_A^2(l)) - 1$ and does not require complicated back-transformations either. We hope that the relative simplicity of these relationships will encourage researchers to adopt the methods more widely.

Nevertheless, our results also illustrate some practical difficulties with animal model estimation. Our Poisson models matched the simulation data-generating process and could have been expected to have ideal statistical properties here. However, the Poisson models did not perform ideally. First, for simulated $\sigma_A^2(l) = 0$ estimates of $\sigma_A^2(\omega)$ were biased upward, as is always the case for variance components in any mixed models (see for instance Bolker et al. 2009). This upward bias is obvious only for $\sigma_A^2(l) = 0$ but must necessarily exist for small non-null values (although it may be masked by an opposite bias). Second, for non-null small simulated $\sigma_A^2(l)$, estimates of $\sigma_A^2(w)$ were slightly biased downward (Figure 3B), showed a small peak in RMSE (Supplementary Figure 4) and had mediocre coverage (Supplementary Figure 5). These suboptimal properties could suggest an influence of the prior distribution on estimation. The prior used here is very dense close to zero and could potentially pull down the posterior distribution for small parameter values (Gelman 2006). Nevertheless, preliminary re-analyses of the same data sets with restricted maximum likelihood showed the same bias (data not shown), suggesting the prior is not to blame here. Instead, bias could originate from a deeper boundary effect, that is, the tendency for the likelihood distribution to stick close to the value of 0 when the true value of a variance parameter is small and the sample size is limited. These issues may be particularly relevant to studies of the genetic basis of variation in fitness; as small values of genetic variance in fitness are expected, models are likely to frequently encounter the boundary area where both positive and negative biases may appear. This is a topic that requires further detailed investigation.

Limitations and Future Directions

We used an overdispersed Poisson model to simulate multigenerational fitness data and to fit animal models to them. Poisson animal models performed well, not surprisingly because they were ideally specified for our simulated data (Burnham and Anderson 2002, p. 158). In nature, real fitness distributions can follow Poisson distributions or overdispersed Poisson distributions (as assumed in our work; e.g., Kruuk et al. 2014b), but may also often deviate from them (e.g., Reid et al. 2011). In particular, fitness distributions may be zero-inflated compared to an overdispersed Poisson distribution. Standard Poisson animal models may still be able to fit zero-inflated data reasonably well (e.g., Reid et al. 2011; Kruuk et al. 2014b), but it is unclear how well they would estimate quantitative genetic parameters in that case. Thus, as a further extension, zero-inflated overdispersed Poisson animal models could be used to fit even more realistic fitness distributions. However, zero-inflated, or hurdle, models are mixtures of 2 distributions and output 2 sets of parameter estimates on 2 different scales (Atkins et al. 2013). Future work will be needed to focus on clarifying how to convert estimates from the 2 parts of zero-inflated models into a single biologically meaningful estimate of genetic variance or into predictions of changes in mean fitness.

Supplementary Material

Supplementary results are available at *Journal of Heredity* online.

Funding

This work was supported in part by a University Research Fellowship from the Royal Society (London) to M.M.

Acknowledgments

We thank Anne Bronikowski for the invitation to submit to the symposium issue on the Evolutionary Quantitative Genetics of Wild Populations; Pierre de Villemereuil, François Rousset, Simon Evans, and Erik Postma for insightful discussions; and anonymous referees for useful comments. We also thank the Australian National University Bioinformatics Consultancy for access to their Bioinformatics Development Cluster.

References

Asmussen MA. 1983. Density-dependent selection incorporating intraspecific competition. II. A diploid model. *Genetics*. 103:335–350.

Atkins DC, Baldwin SA, Zheng C, Gallop RJ, Neighbors C. 2013. A tutorial on count regression and zero-altered count models for longitudinal substance use data. *Psychol Addict Behav*. 27:166–177.

Authier M, Aubry LM, Cam E. 2017. Wolf in sheep's clothing: Model misspecification undermines tests of the neutral theory for life histories. *Ecol Evol*. 7:3348–3361.

Barton NH, Etheridge AM, Véber A. 2017. The infinitesimal model: definition, derivation, and implications. *Theor Popul Biol*. 118:50–73.

Bijma P. 2010. Fisher's fundamental theorem of inclusive fitness and the change in fitness due to natural selection when conspecifics interact. *J Evol Biol*. 23:194–206.

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MH, White JS. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol*. 24:127–135.

Bonnet T, Postma E. 2016. Successful by chance? The power of mixed models and neutral simulations for the detection of individual fixed heterogeneity in fitness components. *Am Nat*. 187:60–74.

Bonnet T, Postma E. 2018. Fluctuating selection and its (elusive) evolutionary consequences in a wild rodent population. *J Evol Biol*. 31:572–586.

Burnham K, Anderson D. 2002. *Model selection and multimodel inference. A practical information-theoretic approach*. 2nd ed. New York: Springer.

Burt A. 1995. The evolution of fitness. *Evolution*. 49:1–8.

Charlesworth B. 1987. The heritability of fitness. In: Bradbury J, Andersson M, editors. *Sexual selection: testing the alternatives*. Chichester: Wiley. p. 21–40.

Clutton-Brock T, Sheldon BC. 2010. Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. *Trends Ecol Evol*. 25:562–573.

de Villemereuil P. 2018. Quantitative genetic methods depending on the nature of the phenotypic trait. *Ann N Y Acad Sci*. 1422:29–47.

de Villemereuil P, Gimenez O, Doligez B. 2013. Comparing parent–offspring regression with frequentist and Bayesian animal models to estimate heritability in wild populations: a simulation study for Gaussian and binary traits. *Methods Ecol Evol*. 4:260–275.

de Villemereuil P, Schielzeth H, Nakagawa S, Morrissey M. 2016. General methods for evolutionary quantitative genetic inference from generalized mixed models. *Genetics*. 204:1281–1294.

Ewens WJ. 1989. An interpretation and proof of the fundamental theorem of natural selection. *Theor Popul Biol*. 36:167–180.

Fisher RA. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans R Soc Edinb*. 52:399–433.

Fisher RA. 1930. *The genetical theory of natural selection*. 1st ed. Oxford: Clarendon Press.

Fisher DN, McAdam AG. 2019. Indirect genetic effects clarify how traits can evolve even when fitness does not. *Evol Lett*. 3:4–14. doi: 10.1002/evl3.98.

Gelman A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal*. 1:515–533.

Geyer CJ, Wagenius S, Shaw RG. 2007. Aster models for life history analysis. *Biometrika*. 94:415–426.

Gomulkiewicz R, Holt RD, Barfield M, Nuismer SL. 2010. Genetics, adaptation, and invasion in harsh environments. *Evol Appl*. 3:97–108.

Grafen A. 2018. The left hand side of the fundamental theorem of natural selection. *J Theor Biol*. 456:175–189.

Gustafsson L. 1986. Lifetime reproductive success and heritability: empirical support for Fisher's fundamental theorem. *Am Nat*. 128:761–764.

Hadfield JD. 2010 . Mcmc methods for multi-response generalized linear mixed models: the MCMCglmm R package. *J Stat Soft*. 33:1–22.

Hamel S, Gaillard JM, Yoccoz NG, Bassar RD, Bouwhuis S, Caswell H, Douhard M, Gangloff EJ, Gimenez O, Lee PC. 2018. General conclusion to the special issue Moving forward on individual heterogeneity. *Oikos*. 127:750–756.

Hansen TF, Pélabon, C, Houle, D. 2011. Heritability is not evolvability. *Evol Biol*. 38:258–277.

Henderson CR. 1950. Estimation of genetic parameters. *Ann Math Stat*. 21:309–310.

Hendry AP, Schoen DJ, Wolak ME, Reid JM. 2018. The contemporary evolution of fitness. *Annu Rev Ecol Evol Syst*. 49:457–476.

Heywood JS. 2005. An exact form of the breeder's equation for the evolution of a quantitative trait under natural selection. *Evolution*. 59:2287–2298.

Hill WG. 1974. Prediction and evaluation of response to selection with overlapping generations. *Anim Prod*. 18:117–139.

Jones JS. 1987. The heritability of fitness: bad news for 'good genes'? *Trends Ecol Evol*. 2:35–38.

Kingsolver JG, Diamond SE. 2011. Phenotypic selection in natural populations: what limits directional selection? *Am Nat*. 177:346–357.

Kingsolver JG, Hoekstra HE, Hoekstra JM, Berrigan D, Vignieri SN, Hill CE, Hoang A, Gibert P, Beerli P. 2001. The strength of phenotypic selection in natural populations. *Am Nat*. 157:245–261.

Kruuk LE. 2004. Estimating genetic parameters in natural populations using the "animal model". *Philos Trans R Soc Lond B Biol Sci*. 359:873–890.

Kruuk LEB, Charmentier A, Garant D. 2014a. The study of quantitative genetics in wild populations. In: Charmentier A, Garant D, Kruuk LEB, editors. *Quantitative genetics in the wild*. 1st ed, Oxford: Oxford University Press. p. 1–15.

Kruuk LEB, Clutton-Brock T, Pemberton JM. 2014b. Case study: quantitative genetics and sexual selection of weaponry in a wild ungulate. In: Charmentier A, Garant D, Kruuk LEB, editors. *Quantitative genetics in the wild*. 1st ed, Oxford: Oxford University Press. p. 160–176.

Kruuk LE, Clutton-Brock TH, Slate J, Pemberton JM, Brotherstone S, Guinness FE. 2000. Heritability of fitness in a wild mammal population. *Proc Natl Acad Sci USA*. 97:698–703.

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

Lush J. 1937. *Animal breeding plans*. Ames (IA): Iowa State College Press.

Mair C, Stear M, Johnson P, Denwood M, Jimenez de Cisneros JP, Stefan T, Matthews L. 2015. A Bayesian generalized random regression model for estimating heritability using overdispersed count data. *Genet Sel Evol*. 47:51.

McFarlane SE, Gorrell JC, Coltman DW, Humphries MM, Boutin S, McAdam AG. 2014. Very low levels of direct additive genetic variance in fitness and fitness components in a red squirrel population. *Ecol Evol*. 4:1729–1738.

McFarlane SE, Gorrell JC, Coltman DW, Humphries MM, Boutin S, McAdam AG. 2015. The nature of nurture in a wild mammal's fitness. *Proc Biol Sci*. 282:20142422.

McPeek MA. 2017. *Evolutionary community ecology*. Princeton (NJ): Princeton University Press.

Merilä J, Sheldon BC. 2000. Lifetime reproductive success and heritability in nature. *Am Nat*. 155:301–310.

Milot E, Mayer FM, Nussey DH, Boisvert M, Pelletier F, Réale D. 2011. Evidence for evolution in response to natural selection in a contemporary human population. *Proc Natl Acad Sci USA*. 108:17040–17045.

Mitchell-Olds T, Shaw RG. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution*. 41:1149–1161.

Morrissey MB. 2015. Evolutionary quantitative genetics of nonlinear developmental systems. *Evolution*. 69:2050–2066.

Morrissey MB, Bonnet T. 2019. Analogues of the fundamental and secondary theorems of selection, assuming a log-normal distribution of expected fitness. *J Hered*. 110:396–402.

Morrissey MB, Kruuk LE, Wilson AJ. 2010. The danger of applying the breeder's equation in observational studies of natural populations. *J Evol Biol*. 23:2277–2288.

Mousseau TA, Roff DA. 1987. Natural selection and the heritability of fitness components. *Heredity (Edinb)*. 59(Pt 2):181–197.

Orzack SH, Steiner UK, Tuljapurkar S, Thompson P. 2011. Static and dynamic expression of life history traits in the Northern Fulmar (*Fulmarus glacialis*). *Oikos*. 120:369–380.

Pelletier F, Coltman DW. 2018. Will human influences on evolutionary dynamics in the wild pervade the Anthropocene? *BMC Biol*. 16:7.

Postma E. 2014. Four decades of estimating heritabilities in wild vertebrate populations: improved methods, more data, better estimates? In: Charmentier A., Garant D, Kruuk LEB, editors. *Quantitative genetics in the wild*. 1st ed. Oxford: Oxford University Press. p. 16–33.

Price GR. 1970. Selection and covariance. *Nature*. 227:520–521.

Price GR. 1972a. Extension of covariance selection mathematics. *Ann Hum Genet*. 35:485–490.

Price GR. 1972b. Fisher's 'fundamental theorem' made clear. *Ann Hum Genet*. 36:129–140.

Price T, Schluter D. 1991. On the low heritability of life-history traits. *Evolution*. 45:853–861.

Queller DC. 2017. Fundamental theorems of evolution. *Am Nat*. 189:345–353.

Reid JM, Arcese P, Sardell RJ, Keller LF. 2011. Additive genetic variance, heritability, and inbreeding depression in male extra-pair reproductive success. *Am Nat*. 177:177–187.

Roff DA. 2001. The threshold model as a general purpose normalizing transformation. *Heredity (Edinb)*. 86:404–411.

Saccheri I, Hanski I. 2006. Natural selection and population dynamics. *Trends Ecol Evol*. 21:341–347.

Shaw RG, Shaw FH. 2014. Quantitative genetic study of the adaptive process. *Heredity (Edinb)*. 112:13–20.

Steiner UK, Tuljapurkar S. 2012. Neutral theory for life histories and individual variability in fitness components. *Proc Natl Acad Sci USA*. 109:4684–4689.

Teplitsky C, Mills JA, Yarrall JW, Merilä J. 2009. Heritability of fitness components in a wild bird population. *Evolution*. 63:716–726.

van Vleck LD. 1972. Estimation of heritability of threshold characters. *J Dairy Sci*. 55:218–225.

Wagner GP. 2010. The measurement theory of fitness. *Evolution*. 64:1358–1376.

Walsh B, Lynch M. 2018. *Evolution and selection of quantitative traits*. Sunderland (MA): Sinauer Associates.

Wilson AJ, Morrissey MB, Adams MJ, Walling CA, Guinness FE, Pemberton JM, Clutton-Brock TH, Kruuk LE. 2011. Indirect genetics effects and evolutionary constraint: an analysis of social dominance in red deer, *Cervus elaphus*. *J Evol Biol*. 24:772–783.

Wilson AJ, Réale D, Clements MN, Morrissey MM, Postma E, Walling CA, Kruuk LE, Nussey DH. 2010. An ecologist's guide to the animal model. *J Anim Ecol*. 79:13–26.

Wolak ME, Arcese P, Keller LF, Nietlisbach P, Reid JM. 2018. Sex-specific additive genetic variances and correlations for fitness in a song sparrow (*Melospiza melodia*) population subject to natural immigration and inbreeding. *Evolution*. 72:2057–2075.