

# Domestication and fitness in the wild: A multivariate view

Jarle Tufto\*

Wed Jul 12 23:29:49 2017

## Abstract

Domesticated species continually escaping and interbreeding with wild relatives impose a migration load on wild populations. As domesticated stocks become increasingly different as a result of artificial and natural selection in captivity, fitness of escapees in the wild is expected to decline, reducing the effective rate of migration into wild populations. Recent theory suggest that this may alleviate and eventually eliminate the resulting migration load. I develop a multivariate model of trait and wild fitness evolution resulting from the joint effects of artificial and natural selection in the captive environment. Initially, the evolutionary trajectory is dominated by the effects of artificial selection causing a fast initial decline in fitness of escapees in the wild. In later phases, through the counteracting effects of correlational multivariate natural selection in captivity, the mean phenotype is pushed in directions of weak stabilizing selection, allowing a sustained response in the trait subject to artificial selection. Provided that there is some alignment between the adaptive landscapes in the wild and in captivity, these phases are associated with slower rates of decline in wild fitness of the domesticated stock, suggesting that detrimental effects on wild populations are likely to remain a concern in the foreseeable future.

## Introduction

Domesticated crop (Ellstrand et al., 1999), terrestrial animal (Kidd et al., 2009; Verardi et al., 2006) and aquatic fish species (Naylor et al., 2005) often

---

\*Centre for Biodiversity Dynamics/Department of Mathematical Sciences, Norwegian University of Science and Technology, 7491 Trondheim, Norway. E-mail: jarle.tufto@ntnu.no

interbreed with populations of wild relatives from which they originate. This has led to concerns over possible detrimental effects on wild populations (Hindar et al., 1991). Fitness of escapees relative to fitness of their wild relatives is a critical parameter influencing the migration load in the recipient population (Lorenzen, 2005; Tufto, 2010; Baskett and Waples, 2013), the expected change in population size (Ronce and Kirkpatrick, 2001; Tufto, 2001), as well as gene flow of neutral markers useful in monitoring introgression (Huisman and Tufto, 2013; Karlsson et al., 2014, 2016).

Domesticated species typically undergo continual artificial selection for traits of economic interest. A sustained response to such selection over many generations is often observed in artificial selection experiments and in animal and plant breeding programs, with little or no slowdown in the response per generation (Hill and Caballero, 1992). This is consistent with the view that most traits are polygenic and that large phenotypic changes on the basis of standing genetic variation can be achieved through only small changes in underlying allele frequencies (Barton and Keightley, 2002). A sustained response also suggests that counteracting natural selection in the captive environment must be weak, and perhaps negligible as assumed in some models (e.g. Wei et al., 1996; Tufto, 2010). Weak natural selection, however, only follows if all evolutionary changes occur in the trait under direct artificial selection as assumed by Zeng and Hill (1986) (no genetic correlations and no correlational selection).

In the wild, natural selection is, under quite general conditions, expected to move the mean phenotype towards an optimum maximizing mean population fitness (Lande, 1982). Any change in the mean phenotype away from a peak in this adaptive landscape is therefore expected to be accompanied by a reduction in mean relative fitness. A Gaussian fitness function of the form  $w(z) = \exp\{-\frac{1}{2}s(z - \theta)^2\}$  introduced by Haldane (1954) is an often used model of stabilizing natural selection and can be viewed as a parsimonious second-order approximation of the true fitness function, appropriate when the deviation of the phenotype  $z$  from the optimum  $\theta$  is moderate. For a domesticated stock showing a sustained response to artificial selection, extrapolation based on a univariate Gaussian model would suggest that their fitness in the wild would follow a bell-shaped curve, tending to zero in a predictable way as the total response to artificial selection in  $z$  accumulates. In conservation genetic settings with accidental gene flow from the domesticated stock into populations of wild relatives, this would imply that detrimental genetic effects of interbreeding are at their largest for intermediate levels of genetic differentiation, and that detrimental effects might become resolved once the species undergoing domestication becomes sufficiently different from its wild relatives (Huisman and Tufto, 2012, their Fig. 8), suggesting a “make

them different” management strategy sensu Baskett and Waples (2013).

Here I argue that this view is likely to be too simplistic. Despite the fact that a sustained response to artificial selection is often seen experimentally, I argue that natural selection in the captive environment also can be strong. Using a multivariate model similar to Zeng (1988) involving additional unobserved latent traits, I model the evolutionary trajectory resulting from the joint effects of artificial selection and natural correlational, stabilizing selection in the captive environment. I assume some alignment of the adaptive landscapes in captivity and in the wild (shared eigenvectors) but different optima. Observable evolutionary changes are shown to involve fast and slow phases over different characteristic time scales associated with the strength of stabilizing natural selection acting on a set of decoupled, independently evolving linear transformations of the original traits. In particular, fitness of individuals from the domesticated stock escaping to the wild is predicted to exhibit a fast initial decline followed by slower subsequent phases towards a limiting positive plateau, even under sustained response to artificial selection. Additionally, the model provides new insights into the effects of supportive captive breeding programs corresponding to the special case of no intentional artificial selection in captivity.

The model is in some ways similar to Tufto (2010), who modelled migration load resulting from immigration of domesticated escapees and stabilizing, correlational selection in the wild of the same form as assumed here. Although predicting this migration load partially motivates the current work, migration-selection balance and the resulting migration load are not considered explicitly here. Instead the aim is to focus in detail on multivariate evolutionary changes occurring in captivity during the domestication process, and the associated implications for fitness of domesticated escapees in the wild. Provided that their fitness reaches a positive, non-negligible plateau, general findings in Tufto (2000, 2010) (see Discussion) then suggest that this will incur a considerable migration load in the wild.

Separate from these issues but important from a more theoretical point of view, the model is validated against and provides novel interpretations of patterns often observed in artificial selection experiments.

## Model and Results

### MODEL RATIONALE

Artificial selection is usually employed at some life stage that follows considerable prior mortality. This together with possible differences in fertility

provides ample opportunity for natural selection to act. Zeng and Hill (1986) considered the selection limit resulting from a conflict between artificial truncation selection and natural stabilizing selection, using a univariate, single trait model. As emphasized by Lande and Arnold (1983); Blows (2007); Arnold et al. (2008); Kirkpatrick (2009), however, adaptation is inherently a multivariate process. Patterns of long-term response to artificial selection seen experimentally also support a multivariate view (see Discussion).

A multivariate approach is also suggested by theoretical considerations of the evolutionary origin of allometric relationships. Such relationships in the form of simple power laws can be found between body size and many physiological, morphological and life history traits (Schmidt-Nielsen, 1984; West and Brown, 2005). Theory suggests that such scaling laws and, more generally, phenotypic integration, constitute a form of universal optimization of biological organisms and that this optimization ultimately is a result of natural selection (West and Brown, 2005; Armbruster et al., 2014). This implies that the fitness function representing natural selection on such traits must take a specific form. On logarithmic scales, power laws translate to simple linear relationships. In terms of log-transformed traits, selection optimizing the organism according to a given power law can then be envisioned as stabilizing selection perpendicular to a ridge along the allometric linear relationship of trait combinations yielding optimal fitness, or as multivariate stabilizing selection with weak selection along the ridge (Fig. 1a). This evolutionary view for the origin of allometric scaling laws is supported by some studies, for example Frankino et al. (2005) who, after inflating phenotypic variance through artificial selection, demonstrated strong stabilizing selection on the ratio between body size and forewing area in the butterfly *Bicyclus ananana*. Univariate analyses of stabilizing selection, on the other hand, often fail to find evidence for strong stabilizing selection (Kingsolver and Hoekstra, 2001), possibly because only some component of overall fitness is measured. Another possible explanation is the omission of phenotypically correlated traits, which would tend to bias estimates of stabilizing selection on the trait under study downwards towards the strength of (weak) stabilizing selection acting along the allometric slope, in particular if the phenotypic covariance matrix  $\mathbf{P}$ -matrix of measured and unmeasured traits aligns with the adaptive landscape. Such alignment is expected from theory on environmental canalization (Zhang and Hill, 2005) and on the evolution of the  $\mathbf{G}$ -matrix (see discussion).

Below, I build on the model by Zeng (1988) and model natural stabilizing selection both in captivity and in the wild using multivariate Gaussian fitness functions, which again can be seen as a parsimonious second-order approximations of the true fitness surface around the optimum. In the multivariate

case this model involves  $n$  parameters describing strength of stabilizing selection on each of the  $n$  traits, plus  $n(n - 1)/2$  parameters describing correlational selection on each pair of traits. Although parameter values may be different in captivity versus in the wild, it seems reasonable to assume that some properties are shared across both environments, for example, as a result of similar functional constraints and optimization principles underlying allometric patterns applying in both environments. Shared allometry would suggest that the fitness functions describing natural selection in the captive environment should have approximately the same shape and orientation as in the wild (shared eigenvectors), perhaps with some reduction in the overall strength of selection in the domestic environment. This is consistent with idea that adaptive landscapes in different environments align, which through mutation-selection balance, translates to the pattern of variation in  $\mathbf{G}$  often seen empirically. For example, Arnold et al. (2008) found  $\mathbf{G}$ -matrix eigenvectors to be similar in about 75% out of 66 studies. I assume that differences between the captive and wild environments (e.g. the lack of predators or differences in food abundance) lead to shifts in the tradeoffs generating the various allometric relationships. These shifts translate to displacements in the stabilizing selection optima in the captive relative to the wild environment (as in the univariate model of Ford, 2002). For breeding lines originating from the wild, a displaced optimum in captivity would initially generate directional natural selection and hence an asymmetric response, thus being consistent with patterns often seen experimentally (Falconer and Mackay, 1996, p. 212).

## EVOLUTION IN CAPTIVITY

The model assumptions and its behavior are illustrated by the numerical example displayed in Fig. 1. In the following,  $\mathbf{G}$  (shaded grey ellipse in Fig. 1a) refers to the genetic covariance matrix among newborns at the beginning of the life cycle. The initial transient dynamics of  $\mathbf{G}$  during the first generations of domestication, resulting from linkage disequilibrium generated by truncating and stabilizing selection (Bulmer, 1971; Villanueva and Kennedy, 1990), is ignored. The mean phenotype vector  $\bar{\mathbf{Z}}_t = (\bar{Z}_{1,t}, \bar{Z}_{2,t}, \dots, \bar{Z}_{n,t})^\top$  is initially at the optimum in the wild such that  $\bar{\mathbf{Z}}_0 = \mathbf{0}$ .

Natural stabilizing Gaussian selection during domestication acts on the vector of phenotypic values, but unlike Zeng (1988), I analyse the effects of selection by working directly with the distribution of the vector of breeding values  $\mathbf{Z}_t$ . Assuming that environmental effects are independent of  $\mathbf{Z}_t$ , normally distributed and with zero mean, it follows that mean relative fitness

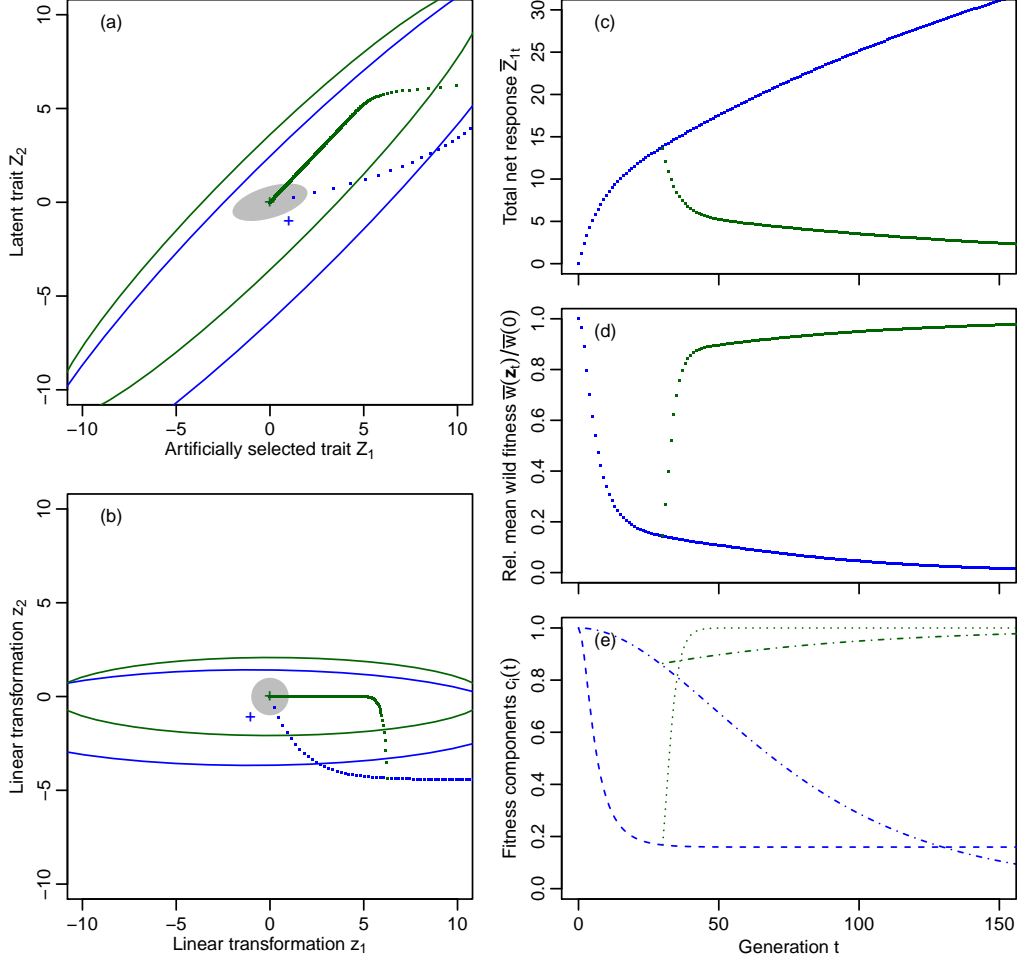


Figure 1: Trait trajectories and evolution of relative mean population fitness for a bivariate version of the model. (a) The genetic covariance matrix  $\mathbf{G}$  (represented by the grey shaded 75%-probability ellipse), matrices  $\mathbf{S}$  and  $\alpha\mathbf{S}$  describing Gaussian fitness functions in captivity and in the wild (displayed by blue and green fitness isocontours), selective optima in captivity and in the wild (blue and green crosses), all in terms of the original traits  $Z_1$  and  $Z_2$ . The blue dots represents the evolutionary trajectory of  $(\bar{Z}_{1,t}, \bar{Z}_{2,t})$  during domestication and the green dots the trajectory after relaxation of artificial selection in the 30th generation. (b) The same quantities but in terms of the linearly transformed decoupled traits  $z_1$  and  $z_2$ . (c) The cumulative response over time to artificial selection  $\bar{Z}_{1,t}$  (blue dots) and regression of  $\bar{Z}_{1,t}$  when artificial selection is relaxed (green dots). (d) The decline over time in relative mean fitness of escapees in the wild during domestication (blue dots) and fitness recovery after relaxation of artificial selection (green dots) equal to the product between the curves in (e) representing components associated with the decoupled traits  $z_1$  (dashed dotted curves) and  $z_2$  (dashed curves). Parameter values used are listed in Appendix A.

of individuals with breeding value  $\mathbf{Z}$  is

$$w(\mathbf{Z}) = \exp \left\{ -\frac{1}{2}(\mathbf{Z} - \mathbf{\Delta})^\top \mathbf{S}(\mathbf{Z} - \mathbf{\Delta}) \right\}, \quad (1a)$$

where the matrix  $\mathbf{S}$  describes stabilizing, correlational selection on the breeding values  $\mathbf{Z}$  (blue ellipse in Fig. 1a). It follows that  $\mathbf{S}^{-1} = \mathbf{S}_P^{-1} + \mathbf{E}$  where  $\mathbf{S}_P$  describes selection on the phenotype and  $\mathbf{E}$  is the covariance matrix of the environmental effects. The vector  $\mathbf{\Delta}$  is the displacement of the optimum in the captive relative to the wild environment (blue cross in Fig. 1a). Because the fitness function and the distribution of breeding values are both Gaussian, the distribution of breeding values after selection is also Gaussian with mean vector

$$(\mathbf{S} + \mathbf{G}^{-1})^{-1}(\mathbf{S}\mathbf{\Delta} + \mathbf{G}^{-1}\bar{\mathbf{Z}}_t). \quad (1b)$$

This can be interpreted as a simple weighted average of the optimum and the mean prior to stabilizing selection in the case of selection on a single trait. The covariance matrix changes to  $(\mathbf{G}^{-1} + \mathbf{S})^{-1} \approx \mathbf{G}$ , that is, it remains approximately unchanged if selection is weak.

Natural selection is followed by artificial selection acting on the first element of the phenotype vector. Genetic correlations, not to be confused with correlational selection, creates an indirect response to artificial selection in other genetically correlated components of  $\mathbf{Z}$ , such that artificial selection overall changes  $\bar{\mathbf{Z}}_t$  by an amount  $\mathbf{A} = (A_1, A_2, \dots, A_n)^\top$  each generation. Given the assumption that  $\mathbf{G}$  remains approximately unchanged by prior weak stabilizing Gaussian selection, the indirect correlated response to artificial selection in  $Z_2, Z_3, \dots, Z_n$  on  $Z_1$  is

$$A_i = \frac{G_{i1}}{G_{11}} A_1 \quad (1c)$$

for  $i = 2, 3, \dots, n$  (Falconer and Mackay, 1996, p. 317). The recursion equation for the overall change in  $\bar{\mathbf{Z}}_t$  from generation  $t$  to generation  $t + 1$  (generating the blue trajectory in Fig. 1a), is then

$$\bar{\mathbf{Z}}_{t+1} = (\mathbf{S} + \mathbf{G}^{-1})^{-1}(\mathbf{S}\mathbf{\Delta} + \mathbf{G}^{-1}\bar{\mathbf{Z}}_t) + \mathbf{A}. \quad (1d)$$

The analysis is simplified by instead working with a linear transformation of genetically uncorrelated, independently selected, decoupled traits  $\mathbf{z} = (z_1, z_2, \dots, z_n)^\top$ . This transformation, illustrated in Fig. 1a versus 1b, is obtained through a rotation (based on the eigenvectors of  $\mathbf{G}$ ), rescaling (based on the eigenvalues of  $\mathbf{G}$ ), and a second rotation based on the eigenvectors of the selection matrix after the two first transformations. This changes (1d) to

$$\bar{\mathbf{z}}_{t+1} = (\mathbf{s} + \mathbf{I})^{-1}(\mathbf{s}\boldsymbol{\delta} + \bar{\mathbf{z}}_t) + \mathbf{a}, \quad (2a)$$

where  $\mathbf{s}$  is a diagonal matrix describing the strength of independent stabilizing selection on each element of  $\mathbf{z}_t$ ,  $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_n)^\top$  is the displacement of the optimum in this new coordinate system, and  $\mathbf{a} = (a_1, a_2, \dots, a_n)^\top$  the direct and indirect response to artificial selection (see Appendix A for details). A resulting evolutionary trajectory is shown in Fig. 1b. The length of the vector  $\mathbf{a}$ ,  $\|\mathbf{a}\|$ , can be shown to be equal to the per generation response to artificial selection measured in genetic standard deviations in the trait (or linear combination of traits) subject to artificial truncating selection. Thus, a reasonable value for  $\|\mathbf{a}\|$ , consistent with the response seen in the Norwegian farmed salmon breeding program (B. Gjerde, pers. comm.) might be somewhere around 0.5, which corresponds to selecting 38% of the population beyond a truncation point of 0.3 phenotypic standard deviation if assuming a heritability of 0.5 (Falconer and Mackay, 1996, eqs. 11.3–5).

Each element of  $\mathbf{z}_t$  in the rescaled model, because all matrices in (2a) are diagonal, evolves independently according to first order linear difference equations. Assuming that the breeding line originates from a population at the wild optimum such that  $\bar{\mathbf{z}}_0 = \mathbf{0}$ , the solutions are

$$\bar{z}_{i,t} = \left( \delta_i + a_i \left( 1 + \frac{1}{s_i} \right) \right) \left( 1 - \frac{1}{(1 + s_i)^t} \right), \quad (2b)$$

for  $i = 1, 2, \dots, n$  (Fig. 1b). Each independently evolving trait  $\bar{z}_{i,t}$  thus evolves towards an equilibrium  $\delta_i + a_i(1 + 1/s_i)$  at which the (direct and indirect) response to artificial selection  $a_i$  is balanced by counteracting stabilizing selection of strength  $s_i$  towards  $\delta_i$ . Importantly, the displacement from the optimum  $\delta_i$  resulting from artificial selection is inversely related to the strength of stabilizing natural selection towards it. Each  $\bar{z}_{i,t}$  evolves exponentially towards the optimum over characteristic times scales  $\tau_i \approx 1/s_i$  (defined by equating  $1/(1 + s_i)^t$  to  $e^{-t/\tau_i}$ ), that is, more slowly in directions of weak selection.

## FITNESS OF ESCAPEES IN THE WILD

Having solved for the evolutionary trajectory resulting from artificial and natural selection in the captive environment, I now consider the corresponding mean fitness of individuals from the domesticated stock when released into the wild environment, relative to mean fitness of a population at the optimum in the wild. This is a critical parameter determining further evolutionary changes in the wild resulting from continual migration and selection, as shown in other work (Tufto, 2000, 2010; Baskett and Waples, 2013; Baskett et al., 2013). I assume a Gaussian fitness function also in the wild with selection matrix  $\mathbf{s}$  differing by a factor of  $\alpha$  from that in captivity (implying



stronger selection in the wild for  $\alpha > 1$ ) (green ellipses in Figs 1a and b). With the optimum at  $\mathbf{z} = \mathbf{0}$ , the mean fitness of an individual with breeding value  $\mathbf{z}$  averaged over the distribution of environmental effects as in (1a) becomes

$$w(\mathbf{z}) = \exp \left\{ -\frac{1}{2} \alpha \mathbf{z}^\top \mathbf{s} \mathbf{z} \right\}. \quad (3a)$$

Averaging over the Gaussian distribution of breeding values with mean vector  $\bar{\mathbf{z}}_t$  and genetic covariance matrix  $\mathbf{I}$  in the rescaled model, mean fitness of escapees in the wild is given by

$$\bar{w}(\bar{\mathbf{z}}_t) = |\alpha \mathbf{s} + \mathbf{I}|^{-1/2} \exp \left\{ -\frac{1}{2} \bar{\mathbf{z}}_t^\top ((\alpha \mathbf{s})^{-1} + \mathbf{I})^{-1} \bar{\mathbf{z}}_t \right\}. \quad (3b)$$

Substituting the solution for the evolutionary trajectory of  $\mathbf{z}_t$  in captivity (2b) into (3b) and using the diagonality of  $\mathbf{s}$ , mean fitness of domesticated individuals in the wild relative to a population at the optimum (blue dotted curve in Fig. 1d) has a factorization  $\bar{w}(\bar{\mathbf{z}}_t)/\bar{w}(\mathbf{0}) = \prod_{i=1}^n c_i(t)$  where the components

$$c_i(t) = \exp \left\{ -\frac{\left( \delta_i + a_i \left( 1 + \frac{1}{s_i} \right) \right)^2 \left( 1 - \left( \frac{1}{1+s_i} \right)^t \right)^2}{2 \left( \frac{1}{\alpha s_i} + 1 \right)} \right\}, \quad (4)$$

$i = 1, 2, \dots, n$  (blue curves in Fig. 1e) represent the separate effects of independent evolution in  $\bar{z}_{1,t}, \bar{z}_{2,t}, \dots, \bar{z}_{n,t}$  on overall wild fitness.

Like each  $\bar{z}_{i,t}$ , these components decline towards limiting values over characteristic time scales  $\tau_i \approx 1/s_i$  inversely proportional to the strength of stabilizing selection on each  $z_i$ . The observable decline in overall relative mean fitness (blue curve in Fig. 1d) in general thus involves several phases occurring over different time scales associated with each direction of independent evolution, as opposed to predictions of univariate models without natural selection in captivity. For the numerical example displayed in Fig. 1, with much weaker selection on  $z_1$  than on  $z_2$ ,  $c_1(t)$  and  $c_2(t)$  decline over characteristic time scales of  $\tau_1 = 200$  and  $\tau_2 = 6.5$  generations, respectively.

The ultimate contribution to loss of fitness associated with evolutionary changes in a particular direction  $z_i$  (the limiting value of  $c_i(t)$  as  $t \rightarrow \infty$ ) depends on the displacement  $\delta_i$  of the optimum, the response to artificial selection  $a_i$ , and the strength of selection  $s_i$  in this direction. Returning to the numerical example in Fig. 1, component  $c_2(t)$  associated with  $z_2$  tends to a limiting value of  $c_2(\infty) = 0.16$  (dashed curve in Fig 1e) resulting from the

displacement of the optimum  $\delta_2$  in this direction and an additional deviation  $a_2(1 + 1/s_2)$  from the displaced optimum caused by artificial selection. Because stabilizing selection on  $z_2$  is strong, these effects are of the same order of magnitude. For the other trait  $z_1$  subject to weak natural stabilizing selection, the contribution to the ultimate fitness loss is, according to the model and somewhat counter intuitively, much larger with  $c_1(t)$  tending to a limiting value  $c_1(\infty) = 3.1 \times 10^{-4}$  (dashed dotted green curve in Fig. 1e). This is caused by the much larger overall response to artificial selection tending to a limiting value of  $\bar{z}_{1,\infty} = 46.6$  genetic standard deviations (outside the range of displayed  $z_1$ -values). Essentially, for small  $s_i$ , the limiting value of the response  $\bar{z}_{i,\infty} \approx a_i/s_i$  and the associated fitness component tends to the limiting value  $c_i(\infty) \approx \exp\{-\frac{1}{2}\alpha s_i \bar{z}_{i,\infty}^2\} = \exp\{-\frac{1}{2}\alpha a_i^2/s_i\}$ , thus producing a larger fitness reduction when stabilizing selection is weak ( $s_i$  small).

In captive breeding programs, no intentional artificial selection is applied in captivity, which corresponds to  $\mathbf{a} = \mathbf{0}$  in the present model. Apart from this, much of the model rationale arguably still applies. In particular, genetic changes (see e.g. Araki et al., 2007) can still occur through unintentional natural selection in captivity, as a result of a displacement  $\delta$  in the optimum as assumed here. Under this scenario, focusing again on the resulting fitness changes of individuals from the captive stock when released in the wild and the limiting values of fitness components associated with each direction of independent evolution (4), these tend to limiting values of  $c_i(\infty) = \exp\{-\frac{1}{2}\alpha s_i \delta_i^2/(1 + \alpha s_i)\}$ . The reduction in fitness thus increases with the strength  $s_i$  of stabilizing natural selection and the magnitude of the displacement of the optimum, as opposed to the effects of artificial selection. As before, however, the time scale  $\tau_i$  over which such changes occur is inversely proportional to the strength of selection  $s_i$ . Any large reduction in fitness in a captive breeding program is thus likely to occur fast over a small number of generations.

## REGRESSION OF $\bar{Z}_{1,t}$ AND RECOVERY OF WILD FITNESS AFTER RELAXATION OF ARTIFICIAL SELECTION

To validate the model against empirical patterns seen in artificial selection experiments (see Discussion), I derive predictions for the evolutionary trait and fitness trajectories once artificial selection is relaxed. Such predictions are also of interest when considering wild populations completely replaced by escapees, and the recovery of fitness in such populations following the cessation of migration. Similar predictions would apply for the evolutionary recovery of wild populations less heavily impacted by introgression (see e.g. Karlsson et al., 2016; Bolstad et al., 2017), having perhaps reached some in-

intermediate balance between migration and selection (as in Tufto, 2010) prior to the cessation of introgression, the difference being the initial conditions.

Setting  $\mathbf{a} = \mathbf{0}$  in (2a) such that there is no artificial selection after some generation  $t_a$  and solving the resulting difference equation, we find that each component  $\bar{z}_{i,t}, \bar{z}_{2,t}, \dots, \bar{z}_{n,t}$  evolves independently towards the optimum in captivity  $\delta_i$  according to the equations

$$\bar{z}_{i,t} = \left( \frac{1}{1 + s_i} \right)^{t-t_a} \bar{z}_{i,t_a} + \left[ 1 - \left( \frac{1}{1 + s_i} \right)^{t-t_a} \right] \delta_i, \quad (5a)$$

or

$$\bar{z}_{i,t} = \left( \frac{1}{1 + \alpha s_i} \right)^{t-t_a} \bar{z}_{i,t_a}, \quad (5b)$$

$i = 1, 2, \dots, n$ , if we consider a population that is reintroduced and returns to the optimum in the wild (green dotted curves in Fig. 1b). Again, these evolutionary changes occur over the characteristic time scales  $\tau_i \approx 1/s_i$ .

The observable trait under artificial selection  $\bar{Z}_{1,t}$ , however, is a linear combination of  $\bar{z}_{i,t}, \bar{z}_{2,t}, \dots, \bar{z}_{n,t}$ . The regression of  $\bar{Z}_{1,t}$ , once artificial selection is relaxed, will thus involve several phases occurring over the time scales  $\tau_1, \tau_2, \dots, \tau_n$  (green dotted curve in Fig. 1c). The same applies to the recovery of fitness in the wild as seen by substituting (5b) into (3b). For the numerical two-trait example in Fig. 1, fitness is predicted to exhibit a fast, but only partial recovery over a time scale of  $\tau_2 = 6.5$  generations, and a subsequent slower, full recovery over a time scale of  $\tau_1 = 200$  generations (green dotted curves in Fig. 1d and e).

## Discussion

The present model suggests that the decline in wild fitness in organisms undergoing domestication is not easily predictable from current estimates of wild fitness (such as Fleming et al., 2000; McGinnity et al., 2003; Skaala et al., 2012) combined with simple extrapolation arguments. Instead, the model predicts that the loss of fitness caused by the response to artificial selection, may, in part, be alleviated by evolution in other latent traits subject to natural correlational selection in the captive environment, pushing the evolutionary trajectory in directions of weak natural selection, while allowing a sustained response to artificial selection to persist. Provided that there is some alignment between the adaptive landscapes in captivity and in the wild, fitness of escapees in the wild may exhibit a fast initial partial decline followed by a slower decline towards some non-negligible, positive

level. The exact pattern of decline depends on presently unknown information about the overall orientation of the fitness function relative to the direction of overall direct and indirect responses to artificial selection. For example, with sufficiently weak selection in captivity or a larger displacement of the optimum than in Fig. 1, it is possible to obtain model behavior similar to the univariate models of Baskett and Waples (2013) and Huisman and Tufto (2012), with fitness in the wild declining to a level near zero over a small number of generations. However, given the ubiquity of correlational multivariate selection (Walsh and Blows, 2009), model behavior similar to the numerical example in Fig. 1 seems equally possible.

Given the slower than expected rate of decline in fitness of domesticated escapees in the wild, “make them different” strategies considered by Baskett and Waples (2013) and Baskett et al. (2013) may thus not work as intended, a conclusion also reached by these authors on different grounds. Hence, for species such as the Atlantic salmon, under a scenario of continuing artificial selection with continuing accidental release of individuals from the domesticated farmed stock into the wild, detrimental effects of interbreeding with wild relatives (Hindar et al., 1991; Karlsson et al., 2016; Bolstad et al., 2017) may well remain a major concern in the coming decades. Although the resulting migration load has not been considered explicitly here, it seems reasonable to expect conclusions of earlier work to hold up, firstly that the load depends largely on the relative fitness of escapees and only to a small extent on the genetic differentiation of the escapees from the wild population, if keeping the the relative fitness of escapees fixed (Tufto, 2000, Fig. 2a) and secondly, that the load in the wild population at migration-selection balance (unlike evolutionary changes in captivity during domestication) depends only to a small extent on correlational selection (Tufto, 2010, Fig. 4).

Although continued migration from artificially selection populations seems likely, model predictions of the evolutionary trajectory without artificial selection (green curves in Fig. 1d) are still of interest, should complete containment of the domesticated stock become a reality some time in the future. While dependent on initial conditions perhaps determined by migration-selection balance as in (Tufto, 2010), the recovery of fitness in such populations (Fig. 1d) is expected to initially exhibit a partial fast phase followed by a slower phase of complete recovery, mirroring the time scales associated with the decline in fitness during domestication. All these predictions of course rely on an idealized model ignoring loss of genetic variance from finite effective population size, in particular the effective size of the breeding nucleus (Tufto and Hindar, 2003), and loss of genetic variance from artificial and stabilizing selection (Barton and Keightley, 2002) and will in reality be blurred by many stochastic phenomena such as the increase in initially rare

alleles of large effects on the traits under selection, fixation of alleles of large effect as well as random allele frequency changes, epistatic effects, linkage disequilibrium, all causing largely unpredictable changes in  $\mathbf{G}$  and in the selection response.

Patterns commonly seen in artificial selection experiments, although exhibiting much more erratic behavior, still appear consistent with predictions of the current model, however, both in terms of the response to artificial selection and in terms of the regression towards the initial mean once artificial selection is relaxed (Fig. 1c). For example, Yoo et al. (1980) observed a decelerating but sustained response to artificial selection on bristle number in *Drosophila*, followed by a fast and then slow regression in the trait after relaxation of artificial selection (his Fig. 1). These patterns were interpreted by Yoo et al. as recessive lethals being involved in the artificial selection response, which would also explain the fast initial rate of decline after relaxation of artificial selection. A second, similar example is the pygmy gene for small size in mice that appeared in the artificial selection experiment of MacArthur (1949). This gene has pleiotropic effects causing sterility in homozygotes. This created counteracting natural selection, a decelerating response and rapid regression back towards the initial mean once artificial selection was relaxed (Falconer and Mackay, 1996, pp. 224-225). The present model provides a complementary and mutually non-exclusive explanation involving correlational, stabilizing selection with other latent traits. A third example of interest is Dobzhansky and Spassky's (1969) experiment on artificial selection for positive and for negative photo- and geotaxis, also in *Drosophila*. As argued by these authors, these traits should be neutral in the lab environment, but regression towards initial means was nevertheless observed after relaxation of artificial selection. Such patterns can be explained by a multivariate model involving a genetically correlated trait  $Z_2$  under stabilizing natural selection in the lab environment, becoming displaced from its optimal value as an indirect effect of artificial selection on  $Z_1$ . Once artificial selection is relaxed, stabilizing selection on  $Z_2$  will cause an indirect response and partial regression of  $Z_1$  and apparent genetic homeostasis. Genetic correlations can also arise temporarily through selective sweeps (see e.g. Rubin et al., 2010) generating linkage disequilibrium with other loci under selection which in turn would generate a similar regression in the focal trait once artificial selection is relaxed. A fourth example is Weber's (1996) experiment selecting for wind tunnel flight speed in *Drosophila*; a sustained response was observed for 100 generations of artificial selection, with practically no loss of fitness between generations 50 and 85. This is consistent with the hypothesis that the overall direction of response to artificial selection had a large component along some direction of weak stabilizing natural

selection or along a ridge in the fitness landscape, with the evolutionary trajectory eventually moving parallel to this ridge after some displacement away from the line of maximum fitness during the first initial generations. Finally, even with no artificial selection as in supportive breeding programs, a large and rapid decline in overall fitness, such as seen in captive bred steelhead trout (*Oncorhynchus mykiss*) (Araki et al., 2007), is another possible model outcome consistent with a large shift in the natural selection optimum in captivity (referred to as unintentional domestication selection by these authors), provided that natural selection in both captivity and in the wild is strong.

Somewhat counter intuitively, the load arising from artificial selection in directions of weak natural selection ultimately becomes larger than the loads associated with directions of stronger natural stabilizing selection. This is because the eventual total response to artificial selection in such directions becomes larger, which more than compensates for the smaller load at any given deviation from the optimum (dashed dotted vs. dashed curve in Fig. 1e). These results of course rely on both the validity of the infinitesimal model over the entire range of phenotypic changes occurring and the validity of the Gaussian approximation of the fitness function for large deviations from the optimum, so it seems clear that these prediction must eventually break down for very weak selection. Analogous theoretical results include the load from sexual selection of constant intensity for increased dimorphism countered by natural Gaussian stabilizing selection (Lande, 1980a), and the evolutionary lag load in models of sustained linear change in the optimal phenotype (Lynch and Lande, 1993; Lande and Shannon, 1996; Chevin et al., 2010), both loads inversely proportional to the strength of stabilizing selection.

The difference between predictions of the multivariate model presented here and those of simpler univariate models depends on the strength of selection on each independent trait in the rescaled model (2a) being different. Although such differences must exist, they may be somewhat smaller than expected. This is indicated by theory and empirical evidence suggesting that  $\mathbf{G}$  evolves towards a balance between the pattern of multivariate selection and the covariance matrix  $\mathbf{M}$  of pleiotropic mutational effects (Lande, 1980b; Arnold et al., 2008). Further theoretical studies suggests that  $\mathbf{M}$  itself ultimately also evolves towards alignment with the adaptive landscape (Jones et al., 2014). Overall, partial alignment would to some degree imply more moderate variation in the strength of selection on each of the rescaled traits  $z_1, z_2, \dots, z_n$ .

Most earlier modelling work relevant in assessing the effects of interbreeding between cultured individuals and wild populations (Tufto, 2000, 2001; Ford, 2002; Huisman and Tufto, 2012; Baskett and Waples, 2013), and specif-

ically interbreeding between farmed salmon escapees with their wild relatives (Hindar et al., 2006), has relied on univariate approaches (but see Castellani et al., 2015). The present model demonstrates the importance of a multivariate approach stressed by earlier authors (Lande and Arnold, 1983; Blows, 2007; Kirkpatrick, 2009) in this conservation genetic context. Tufto (2010) used a similar multivariate approach to model gene flow from domesticated species to wild relatives, focusing on the migration load at migration-selection balance. In that model, introducing more evolutionary degrees of freedom by allowing more traits to evolve led to some alleviation of the migration load at migration-selection balance. This parallels how more evolutionary degrees of freedom in the present model alleviates some of the fitness loss resulting from the response to artificial selection.

## Acknowledgements

I thank G. H. Bolstad, R. Lande, C. Pelabon, J. Wright, J. Masel and four anonymous reviewers for insightful comments on the manuscript. This work was supported by the Research Council of Norway (QuantEscape, grant 216105).

## Literature Cited

- Araki, H., B. Cooper, and M. S. Blouin, 2007. Genetic effects of captive breeding cause a rapid, cumulative fitness decline in the wild. *Science* 318:100–3.
- Armbruster, W. S., C. Pélabon, G. H. Bolstad, and T. F. Hansen, 2014. Integrated phenotypes: understanding trait covariation in plants and animals. *Phil. Trans. Roy. Soc. B* 369:20130245.
- Arnold, S. J., R. Bürger, P. A. Hohenlohe, B. C. Ajie, and A. G. Jones, 2008. Understanding the evolution and stability of the G-matrix. *Evolution* 62:2451–2461.
- Barton, N. H. and P. D. Keightley, 2002. Understanding quantitative genetic variation. *Nature Rev. Gen.* 3:1–11.
- Baskett, M. L., S. C. Burgess, and R. S. Waples, 2013. Assessing strategies to minimize unintended fitness consequences of aquaculture on wild populations. *Evol. Appl.* 6:1090–1108.

- Baskett, M. L. and R. S. Waples, 2013. Evaluating alternative strategies for minimizing unintended fitness consequences of cultured individuals on wild populations. *Cons. Biol.* 27:83–94.
- Blows, M. W., 2007. A tale of two matrices: multivariate approaches in evolutionary biology. *J. Evol. Biol.* 20:1–8.
- Bolstad, G. H., K. Hindar, G. Robertsen, B. Jonsson, H. Sægvog, O. H. Diserud, P. Fiske, A. J. Jensen, K. Urdal, T. F. Næsje, B. T. Barlaup, B. Florø-Larsen, H. Lo, E. Niemelä, and S. Karlsson, 2017. Gene flow from domesticated escapes alters the life history of wild Atlantic salmon. *Nature Ecology & Evolution* 1:0124.
- Bulmer, M. G., 1971. The effect of selection on genetic variability. *Am. Nat.* 105:201–211.
- Castellani, M., M. Heino, J. Gilbey, H. Araki, T. Svåsand, and K. A. Glover, 2015. IBSEM: An individual-based atlantic salmon population model. *PLoS ONE* 10:e0138444.
- Chevin, L.-M., R. Lande, and G. M. Mace, 2010. Adaptation, plasticity, and extinction in a changing environment: towards a predictive theory. *PLoS biology* 8:e1000357.
- Dobzhansky, T. and B. Spassky, 1969. Artificial and natural selection for two behavioral traits in *Drosophila pseudoobscura*. *P. Nat. Acad. Sci. USA* 62:75–80.
- Ellstrand, N. C., H. C. Prentice, and J. F. Hancock, 1999. Gene flow and introgression from domesticated plants into their wild relatives. *Ann. Rev. Ecol. Syst.* 30:539–563.
- Falconer, D. and T. F. C. Mackay, 1996. Introduction to quantitative genetics. 4 ed. Longman, London.
- Fleming, I. A., K. Hindar, I. B. Mjølnerød, B. Jonsson, T. Balstad, and A. Lamberg, 2000. Lifetime success and interactions of farm salmon invading a native population. *P. Roy. Soc. Lond. B* 267:1517–1523.
- Ford, M. J., 2002. Selection in captivity during supportive breeding may reduce fitness in the wild. *Cons. Biol.* 16:815–825.
- Frankino, W. A., B. J. Zwaan, D. L. Stern, and P. M. Brakefield, 2005. Natural selection and developmental constraints in the evolution of allometries. *Science* 307:718–20.



- Haldane, J. B. S., 1954. The measurement of natural selection. *in* Proc. IX Int. Congr. Genet., Caryologia (Suppl), Pp. 480–487.
- Hill, W. G. and A. Caballero, 1992. Artificial selection experiments. *Ann. Rev. Ecol. Syst.* 23:287–310.
- Hindar, K., I. A. Fleming, P. McGinnity, and O. Diserud, 2006. Genetic and ecological effects of salmon farming on wild salmon: modelling from experimental results. *ICES J. Marine Sci.* 63:1234–1247.
- Hindar, K., N. Ryman, and F. Utter, 1991. Genetic effects of cultured fish on natural fish populations. *Can. J. Fish. Aqua. Sci.* 48:945–957.
- Huisman, J. and J. Tufto, 2012. Comparison of non-Gaussian quantitative genetic models for migration and stabilizing selection. *Evolution* 66:3444–3461.
- , 2013. Modelling wild-domestic interbreeding: How selection on a quantitative trait affects gene flow at a neutral locus. *J. Theor. Biol.* 332:42–51.
- Jones, A. G., R. Bürger, and S. J. Arnold, 2014. Epistasis and natural selection shape the mutational architecture of complex traits. *Nature Commun.* 5:3709.
- Karlsson, S., O. H. Diserud, P. Fiske, and K. Hindar, 2016. Widespread genetic introgression of escaped farmed Atlantic salmon in wild salmon populations. *ICES Journal of Marine Science* 000:fsw121.
- Karlsson, S., O. H. Diserud, T. Moen, and K. Hindar, 2014. A standardized method for quantifying unidirectional genetic introgression. *Ecol. Evol.* 4:3256–3263.
- Kidd, A. G., J. Bowman, D. Lesbarrères, and A. I. Schulte-Hostedde, 2009. Hybridization between escaped domestic and wild American mink (*Neovision vison*). *Molec. Ecol.* 18:1175–1186.
- Kingsolver, J. and H. Hoekstra, 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157:245–261.
- Kirkpatrick, M., 2009. Patterns of quantitative genetic variation in multiple dimensions. *Genetica* 136:271–84.
- Lande, R., 1980a. Sexual Dimorphism, Sexual Selection, and Adaptation in Polygenic Characters. *Evolution* 34:292–305.

- , 1980b. The Genetic Covariance Between Characters Maintained by Pleiotropic Mutations. *Genetics* 94:203–215.
- , 1982. A quantitative genetic theory of life history evolution. *Ecology* 63:607–615.
- Lande, R. and S. J. Arnold, 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- Lande, R. and S. Shannon, 1996. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* 50:434–437.
- Lorenzen, K., 2005. Population dynamics and potential of fisheries stock enhancement: practical theory for assessment and policy analysis. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 360:171–189.
- Lynch, M. and R. Lande, 1993. Evolution and extinction in response to environmental change. Pp. 234–250, *in* P. M. Kareiva, J. G. Kingsolver, and R. B. Huey, eds. *Biotic Interactions and Global Change*. John Wiley & Sons.
- MacArthur, J. W., 1949. Selection for Small and Large Body Size in the House Mouse. *Genetics* 34:194–209.
- Mcgininity, P., P. Prodöhl, A. Ferguson, R. Hynes, Niall, N. Baker, D. Cotter, B. O’hea, D. Cooke, G. Rogan, J. Taggart, and T. Cross, 2003. Fitness reduction and potential extinction of wild populations of Atlantic salmon, *Salmo salar*, as a result of interactions with escaped farm salmon. *P. Roy. Soc. Lond. B* 270:2443–2450.
- Naylor, R., K. Hindar, I. A. N. A. Fleming, R. Goldberg, S. Williams, J. Volpe, F. Whoriskey, J. Eagle, D. Kelso, and M. Mangel, 2005. Fugitive Salmon: Assessing the Risks of Escaped Fish from Net-Pen Aquaculture. *BioScience* 55.
- Phillips, P. C., S. J. Arnold, and N. Oct, 1999. Hierarchical Comparison of Genetic Variance-Covariance Matrices . I . Using the Flury Hierarchy 53:1506–1515.
- Ronce, O. and M. Kirkpatrick, 2001. When sources becomes sinks: Migrational meltdown in heterogeneous habitats. *Evolution* 55:1520–1531.

- Rubin, C.-J., M. C. Zody, J. Eriksson, J. R. S. Meadows, E. Sherwood, M. T. Webster, L. Jiang, M. Ingman, T. Sharpe, S. Ka, F. Hallbook, F. Besnier, O. Carlborg, B. Bed’hom, M. Tixier-Boichard, P. Jensen, P. Siegel, K. Lindblad-Toh, and L. Andersson, 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464:587–591. 10.1038/nature08832.
- Schmidt-Nielsen, K., 1984. *Scaling: why is animal size so important?* Cambridge University Press, Cambridge.
- Skaala, Ø., K. A. Glover, B. T. Barlaup, T. Svåsand, F. Besnier, M. M. Hansen, R. Borgstrøm, and I. Fleming, 2012. Performance of farmed, hybrid, and wild Atlantic salmon (*Salmo salar*) families in a natural river environment. *Can. J. Fish. Aqua. Sci.* 69:1994–2006.
- Tufto, J., 2000. Quantitative genetic models for the balance between migration and stabilizing selection. *Genet. Res.* 76:285–293.
- , 2001. Effects of releasing maladapted Individuals: A demographic-evolutionary model. *Am. Nat.* 158.
- , 2010. Gene flow from domesticated species to wild relatives: Migration load in a model of multivariate selection. *Evolution* 64:180–192.
- Tufto, J. and K. Hindar, 2003. Effective size in management and conservation of subdivided populations. *J. Theor. Biol.* 222:273–281.
- Verardi, A., V. Lucchini, and E. Randi, 2006. Detecting introgressive hybridization between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. *Molec. Ecol.* 15:2845–2855.
- Villanueva, B. and B. W. Kennedy, 1990. Effect of selection on genetic parameters of correlated traits. *Theor. Appl. Genet.* 80:746–752.
- Walsh, B. and M. W. Blows, 2009. Abundant Genetic Variation + Strong Selection = Multivariate Genetic Constraints: A Geometric View of Adaptation. *Annual Review of Ecology, Evolution, and Systematics* 40:41–59.
- Weber, K., 1996. Large genetic change at small fitness cost in large populations of *Drosophila melanogaster* selected for wind tunnel flight: Rethinking fitness surfaces. *Genetics* 144:205–213.
- Wei, M., A. Caballero, and W. G. Hill, 1996. Selection Response in Finite Populations. *Genetics* 144:1961–1974.

- West, G. B. and J. H. Brown, 2005. The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. *J. Exper. Biol.* 208:1575–92.
- Yoo, B., F. Nicholas, and K. Rathie, 1980. Long-term selection for a quantitative character in large replicate populations of *Drosophila melanogaster*. *Theor. Appl. Genet.* 57:1–17.
- Zeng, Z., 1988. Long-Term Correlated Response, Interpopulation Covariation, and Interspecific Allometry. *Evolution* 42:363–374.
- Zeng, Z. and W. G. Hill, 1986. The selection limit due to the conflict between truncation and stabilizing selection with mutation. *Genetics* 114:1313–1328.
- Zhang, X.-S. and W. G. Hill, 2005. Evolution of the environmental component of the phenotypic variance: stabilizing selection in changing environments and the cost of homogeneity. *Evolution* 59:1237–1244.

## Appendix A

Let  $\mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$  be the eigendecomposition of the genetic covariance matrix  $\mathbf{G}$ . Transforming  $\mathbf{Z}$  to  $\mathbf{Z}' = \mathbf{\Lambda}^{-1/2}\mathbf{U}^{-1}\mathbf{Z}$  transforms the genetic covariance matrix to the identity matrix  $\mathbf{I}$  and the selection matrix describing selection on  $\mathbf{Z}'$  to  $\mathbf{S}' = \mathbf{\Lambda}^{1/2}\mathbf{U}^\top\mathbf{S}\mathbf{U}\mathbf{\Lambda}^{1/2}$ . Letting  $\mathbf{V}$  be the matrix of eigenvectors in the eigendecomposition of  $\mathbf{S}'^{-1}$ , a subsequent transformation  $\mathbf{z} = \mathbf{V}^{-1}\mathbf{Z}'$  changes (1d) to (2a). It follows that  $\mathbf{s} = \mathbf{V}^\top\mathbf{S}'\mathbf{V}$ ,  $\mathbf{a} = \mathbf{V}^{-1}\mathbf{\Lambda}^{-1/2}\mathbf{U}^{-1}\mathbf{A}$  and  $\boldsymbol{\delta} = \mathbf{V}^{-1}\mathbf{\Lambda}^{-1/2}\mathbf{U}^{-1}\boldsymbol{\Delta}$ .

Parameter values used in Fig. 1 are  $\mathbf{G} = \begin{bmatrix} 4 & 1 \\ 1 & 1 \end{bmatrix}$ ,  $\mathbf{S} = \begin{bmatrix} 0.0513 & -0.0487 \\ -0.0487 & 0.0513 \end{bmatrix}$ ,  $\alpha = 1.5$ ,  $A_1 = 1$  and  $\boldsymbol{\Delta} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ .