Selection of lamb size and early pregnancy in Soay sheep (*Ovis aries*)

Maria João Janeiro^{1,2}, Jonathan M. Henshaw^{3,4}, Josephine M. Pemberton⁵, Jill G. Pilkington¹, and Michael B. Morrissey^{1,*}

¹School of Biology, University of St Andrews, St Andrews, KY16 9TH, UK

²CESAM, Department of Biology, University of Aveiro, Aveiro, 3810-193, Portugal

³Department of Biological Sciences, University of Idaho, Moscow, 3051, Idaho, USA

⁴Institute of Biology I, University of Freiburg, Hauptstraße 1, 79104 Freiburg i. Br., Germany

⁵Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh, EH9 3FL, UK

*corresponding author, email: michael.morrissey@st-andrews.ac.uk

September 16, 2020

Abstract

The paradox of stasis – the unexpectedly slow evolution of heritable traits under direct selection – has been widely documented in the last few decades. This paradox is often particularly acute for body size, which is often heritable and where positive associations of size and fitness are frequently identified, but constraints to the evolution of larger body sizes are often not obvious. Here, we identify a trade-off between survival and size-dependent reproduction in Soay sheep (*Ovis aries*), contributes to selection against large body size. Using recently developed theory on non-linear developmental systems, then decompose total selection of ewe lamb mass along different causal paths to fitness. Larger lambs are more likely to become pregnant, which has a large viability cost. After controlling for this pathway, however, the association between lamb mass and subsequent lifetime fitness is positive. Thus this trade-off does not fully explain stasis of size in this population, but it does substantially reduce the strength of positive directional selection of size that would otherwise occur. While selection currently favours reduced probability of early pregnancy, largely irrespective of body size, it is likely that the occurrence of early pregnancy could result from adaptation to conditions during a recent period during which population density was much lower.

Introduction

The *paradox of stasis* (Sogard, 1997) and the pattern of *bigger-is-better* with respect to selection on body size have received great attention in the last few decades from both macro- and microevolutionary biologists. Positive directional selection on size appears to be widespread (e.g. Kingsolver &

Pfennig, 2004) and genetic variability is generally sufficient for rapid evolution (e.g. Postma, 2014), yet both fossil records (Estes & Arnold, 2007; Hunt, 2007) and extant taxa (Larsson et al., 1998; Knouft & Page, 2003; Ozgul et al., 2009) generally show very little change in body size over time. Both theoretical and methodological explanations have been identified for this mismatch (Bradshaw, 1991; Merilä et al., 2001; Hansen & Houle, 2004). One such mechanism, which is expected to be common in nature, is the existence of opposing selective forces (i.e. *trade-offs*), such as those between natural and sexual selection (e.g. Roff, 2000; Roff & Fairbairn, 2007).

Fecundity and sexual selection are major evolutionary forces selecting for larger body size in both females and males (Fairbairn, 1997; Fairbairn et al., 2007). These may be counterbalanced by viability selection against bigger individuals (Blanckenhorn, 2000), leading to a trade-off between sexual (or fecundity) selection and viability selection (Roff, 2000; Roff & Fairbairn, 2007). The benefits and costs associated with larger body size provide the mechanisms underlying such trade-offs and have been extensively discussed (Shine, 1988; Honek, 1993; Anderson, 1995; Sogard, 1997; Blanckenhorn, 2000; Sokolovska et al., 2000). This interplay of costs and benefits is naturally not limited to adult body size. In particular, Blanckenhorn (2000) emphasizes the distinction between the costs of *becoming* as opposed to *being* large, and elaborates on how forces such as predation, parasitism and starvation play different roles in juveniles and adults.

Trade-offs are most readily identified through a deep understanding of a species' biology. Once a trade-off's mechanistic basis is hypothesised, multivariate statistical methods can be used to characterise trade-offs in evolutionary quantitative genetic terms. Ultimately, trade-offs result when a character has both positive and negative effects on fitness, acting through different mechanisms. Path analysis, developed by Wright (1921, 1934) to deal "with a system of interrelated variables" (Wright, 1960), is a natural approach for quantifying trade-offs, as it provides the means to disentangle the various mechanisms by which body size can affect fitness. Path analysis underlies the concept of the *extended* selection gradient, formally defined by Morrissey (2014) as the total causal effects of traits on fitness, i.e. taking into account the whole system of causal associations among traits. Extended selection gradients are therefore akin to the concept of *selection for*, defined by Sober (1986) in opposition to that of *selection of* (the total association of traits with fitness). The initial linear formulation of path analysis and extended selection gradients was subsequently extended to non-linear development systems (Morrissey, 2015). Such non-linear systems are typical in studies of selection, which often include variables (e.g. survival probability or offspring number) that are non-linear functions of underlying traits. Recently, Henshaw et al. (2020) provided a means of decomposing the total causal effect of a trait on fitness (i.e. the extended selection gradient) into

components associated with multiple non-linear paths.

The Soay sheep (*Ovis aries*) on St Kilda (Clutton-Brock & Pemberton, 2004b) is an example of a system in which the paradox of stasis occurs in relation to body size. Positive directional selection (Milner et al., 2004; Morrissey et al., 2012a; Hunter et al., 2018) and reasonably high heritability (Milner et al., 2000; Wilson et al., 2007; Bérénos et al., 2014) have been reported for body size in this system, alongside stasis (Ozgul et al., 2009). Faster growth may be harmful in juvenile females due to an association between size and early fecundity, combined with a viability cost of reproduction. A non-negligible percentage of ewes get pregnant as lambs and those ewes are more likely to die during their first year of life (Clutton-Brock et al., 2004). We hypothesise that early pregnancy is size-dependent, with larger ewes being more likely to get pregnant and, therefore, to die. A mechanism regulating body size emerges directly from this hypothesis – genetic variation in body mass could be maintained by a trade-off between fitness components in female lambs: being larger is associated with higher rates of survival and total offspring production over a lifetime (Clutton-Brock et al., 1992, 1996; Milner et al., 1999), but being pregnant as a lamb would be associated with lower first-year survival.

In this study, we first model the relationship lamb body mass and early pregnancy, as well as their effects on first-year survival and lifetime offspring production, showing that: (1) early pregnancy is size-dependent, (2) there is a positive additive genetic correlation between lamb body mass and early pregnancy, (3) there is viability selection against early pregnancy, and (4) lifetime rearing success (LReS; lambs reared to independence) in surviving ewes is size-dependent, but very similar in ewes that did and did not become pregnant in their first year. We subsequently draw on these models, in combination with recently developed theory for the analysis and decomposition of selection in non-linear systems (Morrissey, 2015; Henshaw et al., 2020) to implement a formal quantitative genetic analysis of natural selection. We estimate overall selection on body mass to be positive, and we quantify how much stronger it would be in the absence of lamb pregnancy. We also show that about half the selection on lamb body mass occurs later in life, suggesting that this trade-off also regulates body size in adults through across-age correlation in body size.

Study system

Data collection

The population of Soay sheep inhabiting the Village Bay on St Kilda in the Outer Hebrides (57°49'N 08°34'W) has been the subject of intensive, individual-based study since 1985 (Clutton-Brock &

Pemberton, 2004a). More than 95% of the individuals living in the study area have been marked with plastic ear tags shortly after birth to enable identification throughout their lifetimes, at which time maternal identity and twin status is ascertained from field observations. Regardless of whether a ewe's lamb is captured (e.g., in cases of neonatal mortality), reproductive status is extensively recorded, allowing assessment of annual and lifetime fitness of ewes. Regular censuses and mortality searches provide death dates for the majority of the sheep. Each year in August a large portion of the study population is captured and phenotyped for multiple traits, including body mass. We focus on phenotypic and life history data on ewe lambs from 1991 to 2015, corresponding to 3457 ewes, as there is no systematic record of pregnancy status from post-mortems of ewes that died during the winter before 1991.

The pedigree of this population has been constructed through a combination of observational field data and molecular markers for maternal links, and using molecular markers only for paternal links (Johnston et al., 2013; Bérénos et al., 2014). 315 polymorphic and unlinked SNP markers were used in molecular parentage assignments (for 4371 individuals) with 100% confidence in the R package MasterBayes (Hadfield et al., 2006). Polymorphic microsatellite markers were also used when SNP genotypes were not available for either lamb or candidate fathers. In those cases, for a total of 222 lambs, 14-18 polymorphic microsatellite markers were used in assignments with confidence >95% in MasterBayes (Morrissey et al., 2012b). The resulting pedigree has a maximum depth of 10 generations and consists of 6740 individuals, of which 6336 are nonfounders (i.e. have one or two known parents).

Summary statistics of mass and early pregnancy

Early pregnancy, here defined as a pregnancy occurring during the first year of life, occurs at a rate of 37 per year in Soay sheep (914 records documented from 1991 to 2015). The frequency of lamb pregnancy varies greatly among years, but in the raw data no apparent temporal trend is found (Fig. 1a). A major driver of these pregnancies seems to be body size, as the probability of pregnancy increases significantly with body mass (Fig. 1b). A lamb weighing around 10 kg has a very low chance of becoming pregnant, whereas a lamb weighing around 15 kg is more likely to become pregnant than not (Fig. 1c). It is well known that body mass in Soay sheep depends on population size, on average decreasing in years with higher density (Clutton-Brock & Pemberton, 2004b; Ozgul et al., 2009). Conditional on population density, there is a temporal trend of decreasing body mass in Soay sheep, including in lambs (Ozgul et al., 2009), and also in ewe lambs (Fig. 1d, upper panel). This trend is apparently concomitant with a decrease in the rate of early pregnancy over the years, when density

is again accounted for (Fig. 1d, lower panel). Substantial costs and few benefits would explain this trend in early pregnancy, as, for example, ewe lamb descendants having significant lower survival when compared to offspring born to older ewes (Fig. 1e).

Models of size, early pregnancy, survival and fitness

In this section, we investigate associations among body size, early pregnancy, survival, and parental success. The purpose of these models is to dissect all of the individual associations among key variables that each contribute to some aspect of overall selection on size and early pregnancy. Ultimately, these models will parameterise the functions relating variables in a path analysis, allowing to disentangling selection *of* and *for* body mass and early pregnancy. All models were fit with MCMCglmm (Hadfield, 2010), using inverse gamma distributions for residual variances (of non-binary traits), and inverse Wishart distributions for the (co)variance components associated with random effects as prior distributions. Also, parameter expansion was implemented to increase convergence efficiency (Gelman, 2006; Hadfield, 2010).

Covariates are included in the models such that fitness and selection are obtained for average conditions of such variables (e.g. maternal age at parity). For clarity, simplified versions of model descriptions, excluding these variables, are presented in the main text, but complete information is provided in Appendix B. Two exceptions are made: twin status (singleton or twin) and population density. In the former case, the binary nature of the variable demands its relationship to the variables included in the path analysis to be explicitly modelled. As for the latter, given the dynamics of the Soay sheep population in St Kilda, understanding how fitness and selection vary with population density is key to understanding selection of body size and pregnancy. All interaction terms are omitted in equations shown in the main text, but are reported in Appendix B.

We define parental success as the number of offspring successfully reared to November 1st and distinguish between those offspring successfully reared as a result of early pregnancy (annual rearing success, AReS) and subsequently (lifetime rearing success, LReS). While these measures of reproductive fitness differ from the traditional definitions of breeding success (offspring produced over a lifetime) and reproductive success (offspring successfully raised up until the age of 1; Clutton-Brock 1988), they delineate fitness associated with a ewe's ability to gestate and provision a lamb from subsequent components of lamb survival which may be less directly dependent on the mother.

Lamb body size

A linear regression of the form

$$m_{ijk} = \alpha + \alpha_t I_{t_{ijk}} + \beta_d f(d_{ijk}, 2) + \beta \mathbf{X}_{ijk} + a_i + u_{cj} + u_{mk} + \epsilon_{ijk} \quad (1)$$

was adopted to model August body mass of lamb i , born in year j to mother k . α and α_t are the intercept and the twinning-specific contrast to the intercept, β_d are the linear (slope) and quadratic terms associated to a second order polynomial function on population density, d , and β is a vector with the coefficients for the remaining fixed effects, $\mathbf{X}$ (see Appendix B). Random intercepts to estimate among-mother (u_{mk}) and among-cohort (u_{cj}) variation were also included. Finally, a_i is the breeding value of lamb i and ϵ are the residuals of the model. Parameter estimates are provided in Appendix B, Table B1.

Early pregnancy

The probability of early pregnancy was estimated using a binomial regression with a logistic link function,

$$\ln \left(\frac{\mathbb{E}[p_{ijk}]}{1 - \mathbb{E}[p_{ijk}]} \right) = \alpha + \alpha_t I_{t_{ijk}} + \beta_m f(m_{ijk}, 3) + \beta_d f(d_{ijk}, 2) + \beta \mathbf{X}_{ijk} + u_{cj} + u_{mk} + \epsilon_{ijk}, \quad (2)$$

where β_m corresponds to a vector with coefficients for the linear, quadratic, and cubic terms for August body mass, m . The remaining quantities have analogous meanings as in Equation (1). More generally, although corresponding to different quantities, we use the same Greek letters for equivalent coefficients in different models throughout the manuscript. Finally, as the overdispersion variance is unobservable for binary data, the variance in ϵ was set to one. Parameter estimates are provided in Appendix B, Table B2.

Bivariate association of pregnancy and size

An important step to disentangling the selection *of* and *for* body mass and early pregnancy is to quantify the genetic association between these variables, a signature of the causal relationships between them (Rausher, 1992; Morrissey et al., 2010). To investigate whether such an association exists, in the form of additive genetic covariance, we used an animal model corresponding to a multi-response generalised linear mixed model (Hadfield, 2010) for body mass, m , and the probability of

lamb pregnancy, p . Although the probability of lamb pregnancy is well described by a binomial distribution, a latent scale variable, p_l , can be defined such that $p_l = \ln\left(\frac{p}{1-p}\right)$. In that case, both m and p_l are assumed to be drawn from a multivariate normal distribution, with a mean vector that includes both mean mass, μ_m , and mean pregnancy probability in the logit scale, μ_{p_l} , and a variance-covariance matrix $\mathbf{P}_l$,

$$\begin{bmatrix} m \\ p_l \end{bmatrix} \sim \mathcal{MVN}\left(\begin{bmatrix} \mu_m \\ \mu_{p_l} \end{bmatrix}, \mathbf{P}_l\right). \quad (3)$$

Both μ_m and μ_{p_l} depend on twin status (singleton or twin), population density, and maternal age at parity, whereas the former also depends on birth and measurement dates. The corresponding parameters and their estimates are listed in Appendix B. Note that the subscript l denotes *latent scale*, which here applies to early pregnancy only, since August lamb body mass was modelled on its natural scale. The genetic and environmental contributions to $\mathbf{P}_l$ were partitioned by including random effects on breeding values (animal model, Henderson 1975). In fact, a $\mathbf{G}_l$ matrix and an $\mathbf{E}_l$ matrix can be defined such that

$$\mathbf{P}_l = \begin{bmatrix} \sigma_m^2 & \sigma_{m,p_l} \\ \sigma_{m,p_l} & \sigma_{p_l}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{m_a}^2 & \sigma_{m,p_{l_a}} \\ \sigma_{m,p_{l_a}} & \sigma_{p_{l_a}}^2 \end{bmatrix} + \begin{bmatrix} \sigma_{m_e}^2 & \sigma_{m,p_{l_e}} \\ \sigma_{m,p_{l_e}} & \sigma_{p_{l_e}}^2 \end{bmatrix}, \quad (4)$$

where subscripts a and e denote the additive genetic and the environmental contribution to the overall (co)variances, respectively. For more information on (co)variance partition, see Falconer (1981). To simplify the notation, we do not use any extra subscript to denote phenotypic (co)variances (see matrix $\mathbf{P}_l$ in Equation 4). Besides residual variances and covariances, the matrix $\mathbf{E}_l$ also includes variances and covariances associated with cohort and maternal random effects. Since the overdispersion variance of a generalised linear mixed model (GLMM) is unobservable for binary data, the residual variance for p_l was set to one.

Taking into account the additive genetic and the phenotypic variances in lamb body mass, 0.78 (95% CrI 0.20; 1.27) and 4.29 (95% CrI 3.84; 5.23), respectively, the heritability of this trait is estimated to be 0.14 (95% CrI 0.04; 0.29), very similar to the estimate reported by Bérénos et al. (2014) for both sexes (0.12, SE 0.036). Up until very recently, for binomial variables, heritabilities were typically obtained by approximation, as described in Nakagawa & Schielzeth (2010). Following this approach, we obtain an estimate of 0.15 95% CrI 0.10; 0.70) for the heritability of early pregnancy. There is a strong positive phenotypic correlation between lamb body mass and early pregnancy (0.51, 95% CrI 0.40; 0.66), the genetic contribution to that correlation being very important (0.64, 95% CrI 0.24; 0.90).

210

211 In 2016, Villemereuil et al. derived exact expressions to convert the latent scale parameters
212 estimated in GLMMs into scales on which traits are expressed and selected. We are particularly
213 interested in obtaining parameters in the probability scale (probability of early pregnancy), rather
214 than in the logarithm of the odds of being pregnant as a lamb. Villemereuil et al. (2016) refer both
215 to *expected* and *data* scales, the difference between them being the random noise that distinguishes
216 a model from the data itself. Here, we adopt the expected scale. Following their work, the expected
217 values for body mass and the probability of early pregnancy is given by

$$\bar{z} = \begin{bmatrix} \bar{m} \\ \bar{p} \end{bmatrix} = \int g^{-1}(\mathbf{l}) f_{\mathcal{MVN}}(\mathbf{l}, \boldsymbol{\mu}, \mathbf{P}_l) d\mathbf{l}, \quad (5)$$

218 where g^{-1} corresponds to the inverse link functions used to model body mass and the probability
219 of early pregnancy. As the adopted link functions for these variables were the identity and the logit
220 functions, respectively, their inverses are the identity and the logistic functions, the latter correspond-
221 ing to $\frac{\exp(p_l)}{\exp(p_l)+1}$. $\mathbf{l}$ are the latent values for lamb body mass and early pregnancy. $f_{\mathcal{MVN}}(\mathbf{l}, \boldsymbol{\mu}, \mathbf{P}_l)$ is
222 the probability density of a multivariate normal distribution with vector mean $\boldsymbol{\mu}$ and variance $\mathbf{P}_l$.
223 Likewise, the phenotypic variance-covariance matrix on the expected scale is given by

$$\mathbf{P} = \int (g^{-1}(\mathbf{l}) - \bar{z})^2 f_{\mathcal{MVN}}(\mathbf{l}, \boldsymbol{\mu}, \mathbf{P}_l) d\mathbf{l}. \quad (6)$$

224 The $\mathbf{G}$ matrix on the expected scale,

$$\mathbf{G} = \Phi \mathbf{G}_l \Phi^T, \quad (7)$$

225 can also be derived (Villemereuil et al., 2016), where Φ is the average derivative of the expected values
226 with respect to the latent values,

$$\Phi = \int dv g^{-1} \mathbf{l} f_{\mathcal{MVN}}(\mathbf{l}, \boldsymbol{\mu}, \mathbf{P}_l) d\mathbf{l}. \quad (8)$$

227 Details on how derivatives and integrals were solved are provided in Appendix C. The heritability
228 on the expected scale is obtained directly through its definition, as the proportion of the additive
229 genetic variance relative to the phenotypic variance, using the derived values. The results obtained
230 using the above expressions can also be found in Table B3. The heritability of early pregnancy on
231 the expected scale (probability) is 0.40 (95% CrI 0.11; 0.63), higher than the approximation shown
232 above (but note that this approximation is derived for the data scale), and corroborating the genetic
233 basis of this trait. Using the derived values on the expected scale, we also obtained the conditional

genetic variance of lamb body mass. This measure allows us to evaluate the proportion of the additive genetic variance in lamb body mass that is independent of early pregnancy. Following Hansen et al. (2003), this quantity, $\sigma_{m|p_a}^2$, is given by

$$\sigma_{m|p_a}^2 = \sigma_{m_a}^2 - \frac{\sigma_{m,p_a}^2}{\sigma_{p_a}^2}, \quad (9)$$

which in this particular case equals 0.40 (95% CrI 0.06; 0.85). The analogous metric for early pregnancy corresponds to 0.03 (95% CrI 0.00; 0.05). Since $\sigma_{m_a}^2$ and $\sigma_{p_a}^2$ were estimated to be, respectively, 0.78 (95% CrI 0.20; 1.27) and 0.04 (95% CrI 0.00; 0.08), the additive genetic variance in lamb body mass and early pregnancy that are independent from one another corresponds to 59% (95% CrI 29%; 100%). Substantial proportions of the genetic variances in lamb body mass and early pregnancy are not independent, and therefore *selection for* each one of these traits necessarily results in *selection of* the other.

First-year survival

We investigated whether early pregnancy has fitness costs by evaluating first-year survival in ewe lambs as a function of pregnancy and body mass. First-year survival was defined using the first day of May as the cut-off date to make sure that dying during labour was considered in the first annual cycle of the ewes (89% of births are before 1 May). Ewes with known first-year survival status and that survived until the rut were included in the analysis ($n_2 = 947$). A binomial regression with logit link function of the form

$$\ln \left(\frac{\mathbb{E}[s_{ijk}]}{1 - \mathbb{E}[s_{ijk}]} \right) = \alpha + \alpha_t I_{t_{ijk}} + \alpha_p I_{p_{ijk}} + \beta_m m_{ijk} + \beta_d f(d_{ijk}, 2) + \beta X_{ijk} + u_{c_j} + u_{m_k} + \epsilon_{ijk}, \quad (10)$$

was adopted to model the probability of lamb i , born in year j to mother k , survives its first year of life. α and α_p are the intercept and the pregnancy-specific contrast to the intercept, β_m is the slope for lamb body mass, and β is a vector with the coefficients for the remaining fixed effects. Although not pictured in Equation (10), a pregnancy-specific contrast to the slope for mass is also estimated (Appendix B). As first-year survival, and survival in general, is highly dependent on cohort effects due to variability in winter conditions and food availability (Clutton-Brock et al., 2004), we included cohorts as random effects (u_{c_j}), as well as maternal environmental effects (u_{m_k}). Variance in residuals (ϵ_{ijk}) was set to one, as overdispersion is unobservable in binomial mixed models.

We first establish that first-year survival is size-dependent by adopting a similar model to the one in Equation (10), but excluding the pregnancy-specific contrasts - bigger ewes have higher chance of survival ($p < 0.001$, Fig. 2a, Tab. B4a in App. B). Adopting the full model in Equation (10) shows that for a given body mass, pregnant lambs are significantly less likely to survive their first annual cycle when compared to the ewes that do not get pregnant ($p < 0.001$, Fig. 2c, Tab. B4b in App. B). Lamb pregnancy has a substantial survival cost - the probability of survival for ewe lambs of average body mass that did not get pregnant is 50% (95% CrI 37%; 65%), an estimate that drops to 26% (95% CrI 14%; 39%) in pregnant lambs. Pregnant lambs of all sizes are more likely to die than non-pregnant ewe lambs (Fig. 2c). Although merely a numerical result, it is interesting to notice that mean body mass is lower among *all* lambs that got pregnant (14.00, 95% CrI 13.65; 14.48) than it is in the sub-group of lambs that got pregnant and survived at least a year (14.59, 95% CrI 14.11; 15.03), suggesting that if a ewe is to be large enough to get pregnant, then it is better to be as large as possible. Investigating the simultaneous effect of lamb body mass, early pregnancy and population density on first-year survival exposes an important interaction effect between the former traits and the latter. Although large population size is, in general, associated with lower survival, it is evident that pregnant ewe lambs are more susceptible to variation in population density than lambs that do not get pregnant. First-year survival in pregnant ewes is very low (below 10%) across most possible values of population density, unless body mass is relatively very high, whereas such low yearly survival rates are much less likely in non-pregnant ewes (App. A, Fig. A1a,b).

First-year rearing success (AReS)

Having established that pregnant ewe lambs are more likely to die when compared to same-sized non-pregnant ewe lambs, it is important to understand the fate of those pregnancies. From the 914 documented early pregnancies, only 97 resulted in offspring surviving until the winter (AReS), and only 36, less than 5%, resulted in recruitments to the population, i.e. offspring that survived until their first April 1st. We estimated the AReS of ewe lamb i fitting a binomial regression with a logit link function of the form

$$\ln \left(\frac{\mathbb{E}[\text{AReS}_i]}{1 - \mathbb{E}[\text{AReS}_i]} \right) = \alpha + \beta_m m_i + \beta_d f(d_{ijk}, 2) + \beta \mathbf{X}_i + \epsilon_i, \quad (11)$$

where α is the model intercept, β_m is a slope for lamb body mass, and β is a vector with the coefficients associated to the covariates in $\mathbf{X}$, which includes capture and birth dates, as well a linear and a quadratic terms for population density. The variance in the model residuals, ϵ_i , was set to one. Due to sample size limitations, we did not consider any random effects.

Only 2.2% (95% CrI 0.57%; 6.17%) of the offspring born to ewes getting pregnant as average-sized lambs in average-density years survive until the winter (App. B, Tab. B5). It is interesting to note that the slope for body mass is positive and significantly different from zero, showing that larger ewe lambs are more likely to have surviving offspring, and suggesting that, as for first-year survival, if a ewe lamb is to get pregnant, than the larger the better. Overall, AReS not only increases with ewe body mass, but also significantly decreases as population size increases (App. A, Fig. A1c).

Subsequent lifetime rearing success (LReS)

We also investigated the effect of lamb body mass and early pregnancy on LReS. As the LReS of ewes dying during their first year of life is zero, analyses presented here only include ewes that survived their first winter. The LReS of ewe i , born in year j to mother k was modelled with a Poisson mixed model with a log link function of the form

$$\ln(\mathbb{E}[\text{LReS}_{ijk}]) = \mu + \alpha_t I_{t_{ijk}} + \mu_p I_{p_{ijk}} + \beta_m m_{ijk} + \beta_d f(d_{ijk}, 2) + \beta \mathbf{X}_{ijk} + u_{c_j} + u_{m_k} + \epsilon_{ijk}, \quad (12)$$

where μ is the model intercept and μ_t and μ_p are the twin status and early pregnancy specific contrasts to the intercept, respectively. β_m is the slope for lamb August mass, β_d the linear (slope) and quadratic terms for density, and β a vector with the slopes of the remaining fixed effects, $\mathbf{X}$. Although not shown, a pregnancy-specific contrast to the slope for mass is also estimated (Appendix B). $\mathbf{u}_c$ and $\mathbf{u}_m$ are random effects, assumed to be drawn from independent normal distributions, for cohort and maternal identity. Finally, ϵ_{ijk} is the residual of individual i . To study LReS, the data were further constrained to only include ewes from cohorts that were completely phenotyped for this trait, which was achieved by excluding all ewes born after 2006 ($n_3 = 550$). This procedure resulted in including very few LReS records associated with ewes that were still alive by the cut-off year (2015). While these surviving ewes' lifetime fitness is underestimated, they have reared most the offspring they will rear in their lives. Excluding these records would result in stronger bias because ewes that lived longer had potentially higher rearing success.

We applied the model in Equation (12) without the coefficients associated with early pregnancy and established that LReS is dependent on lamb size (Fig. 2b, Tab. B6a in App. B). Rearing success is an increasing function of lamb August body mass, with bigger animals being more likely to rear more offspring. Results from the full model suggest that rearing success is not determined by early pregnancy status (Fig. 2d, Tab. B6b), although it is unclear whether under particular conditions of

population density this trait could have an effect on LReS that is independent from its correlation with lamb body mass. Neither the contrast to the intercept (0.37; 95% CrI -0.06; 0.82) nor the interaction terms with body mass (0.02; 95% CrI -0.05; 0.07) and population density (3.4×10^{-3} ; 95% CrI -9×10^{-3} ; 7.5×10^{-3}) differ from zero, however most of the posterior density of the pregnancy-related coefficients supports positive values. Such numeric differences result in distinct expected numbers of offspring successfully reared in ewes that were and were not pregnant as lambs. In averaged-sized individuals living under average environmental conditions, 2.48 (95% CrI 1.57; 4.00) offspring are expected to be reared by ewes that survived their first year of life, which is broken down to 3.36 (95% CrI 1.87; 5.54) and 2.30 (95% CrI 1.40; 3.46) in ewes that got and did not get pregnant as lambs, respectively (Fig. 2b). As no strong statistical support is found for these differences, nor is any biological justification known for such an effect of early pregnancy, LReS will be assumed to be independent from this trait in further analyses. The simultaneous effect of body mass and population density on LReS, independent of early pregnancy, is presented in Appendix A (Fig. A1d), showing not only that LReS is highest in ewes born at low population densities, but also that the effect of lamb body mass on LReS is stronger in ewes that were also born in years of low population density.

Twinning

The parameterisation of the path model built in this study requires knowledge of the effects of twin status on size, pregnancy, survival, and rearing success, which are estimated in Equations (1), (2), (10), and (12). Furthermore, given the binary nature of the variable, a model of the probability of being a twin is required so that the quantities such as mean phenotype or selection gradients can be average over this probability. For this purpose, a binomial mixed model was fitted according to

$$\ln \left(\frac{\mathbb{E}[t_{ijk}]}{1 - \mathbb{E}[t_{ijk}]} \right) = \alpha + \beta_{\mathbf{a}} f(d_{ijk}, 2) + \beta \mathbf{X}_{ijk} + u_{cj} + u_{mk} + \epsilon_{ijk} \quad (13)$$

where $\mathbb{E}[t_{ijk}]$ are individual probabilities of twinning, and other terms are defined analogously to other models. Detailed information on the estimated quantities is provided in Appendix B, Table B7.

Selection analysis

The analyses shown so far establish that (1) early pregnancy is size-dependent, (2) there is a fairly strong positive additive genetic correlation between lamb body mass and early pregnancy, (3) there is selection against early pregnancy through viability selection, (4) the ultimate fitness benefit of early pregnancy is very small, and (5) subsequent LReS in surviving ewes is size-dependent, but very similar in ewes that got and did not get pregnant. In this section, we use the information

presented so far in a formal selection analysis to investigate *selection for* and *selection of* lamb size in Soay sheep ewes. In particular, we use a path analysis involving all pertinent traits (Fig. 3), allowing the estimation of an *extended selection gradient* (Morrissey, 2014) for lamb body mass. Using the path rules for developmental systems set by Wright (1934) and expanded to non-linear variables by Morrissey (2015) one can use the twin model approach developed by Henshaw et al. (2020) to decompose the components of the effect of lamb body mass on fitness and that arise via early pregnancy and otherwise. The path diagram shown in Figure 3 includes twinning, t , mass, m , pregnancy, p , survival, s , first-year rearing success, AReS, subsequent rearing success, LReS, and fitness, W , and can be represented by a vector-valued function of the following form

$$\mathbf{z}_{ijk} = \begin{bmatrix} t \\ m \\ p \\ s \\ \text{AReS} \\ \text{LReS} \\ W \end{bmatrix}_{ijk} = \mathbf{f}(\mathbf{l})_{ijk} = \begin{bmatrix} g^{-1}(\alpha + \beta_d f(d_{ijk}, 2) + \beta_{\mathbf{X}} \mathbf{X}_{ijk} + u_{cj} + u_{mk} + \epsilon_{ijk}) (\text{Eqn. 13}) \\ \alpha + \alpha_t I_{t_{ijk}} + \beta_d f(d_{ijk}, 2) + \beta_{\mathbf{X}} \mathbf{X}_{ijk} + a_i + u_{cj} + u_{mk} + \epsilon_{ijk} (\text{Eqn. 1}) \\ g^{-1}(\alpha + \alpha_t I_{t_{ijk}} + \beta_m f(m_{ijk}, 3) + \beta_d f(d_{ijk}, 2) + \beta_{\mathbf{X}} \mathbf{X}_{ijk} + u_{cj} + u_{mk} + \epsilon_{ijk}) (\text{Eqn. 2}) \\ g^{-1}(\alpha + \alpha_t I_{t_{ijk}} + \alpha_p I_{p_{ijk}} + \beta_m m_{ijk} + \beta_d f(d_{ijk}, 2) + \beta_{\mathbf{X}} \mathbf{X}_{ijk} + u_{cj} + u_{mk} + \epsilon_{ijk}) (\text{Eqn. 10}) \\ g^{-1}(\alpha + \beta_m m_{ijk} + \beta_d f(d_{ijk}, 2) + \beta_{\mathbf{X}} \mathbf{X}_{ijk} + \epsilon_{ijk}) (\text{Eqn. 11}) \\ e^{\alpha + \alpha_t I_{t_{ijk}} + \beta_m m_{ijk} + \beta_d f(d_{ijk}, 2) + \beta_{\mathbf{X}} \mathbf{X}_{ijk} + u_{cj} + u_{mk} + \epsilon_{ijk}} (\text{Eqn. 12}) \\ s_{ijk} (\text{AReS}_{ijk} + \text{LReS}_{ijk}) \end{bmatrix}. \quad (14)$$

Equation umbers given in parentheses reference the statistical models used to parameterise each component. As before, although corresponding to different quantities, we use the same Greek letters for equivalent coefficients in all models: α represents the intercepts, α_t and α_p the twinning and pregnancy-specific contrasts to the intercept, respectively, β_m and $\beta_{\mathbf{m}}$ the slope and the linear (slope), quadratic and cubic terms for mass, β_d the linear and quadratic terms for population density and, finally, β is a vector containing the effects associated with the remaining covariates. β also includes interaction terms for body mass and population density, interaction terms between early pregnancy status and population density, and between these two and body mass, in all models where the respective predictors are included. u_{cj} and u_{mk} are random effects associated with cohort j and mother k , and ϵ_{ijk} is the residual of individual i . The model for mass also includes breeding values, a , allowing us to segregate the additive genetic variance from other sources of variation. Twinning, pregnancy, survival, and AReS were assumed to follow binomial distributions, mass to follow a Gaussian distribution, and LReS to follow a Poisson distribution with additive overdispersion. Fitness was defined as the sum of AReS and LReS of ewes surviving their first annual cycle, weighted by the probability of first-year survival. Note that although population density is not included in the path diagram of Figure 3, all the results presented in this section were obtained by marginalising over the observed population density in all cohorts. Equation (14) was evaluated for the values of population density of each cohort and the mean value taken as the best estimate of $\bar{z}$.

Selection on mass and the probability of pregnancy

Mean phenotype, $\bar{\mathbf{z}}$, and the phenotypic variance-covariance matrix, $\mathbf{P}$, on the expected scale were obtained using Equations (5) and (2). We estimate $\bar{\mathbf{z}}$ as

$$\bar{\mathbf{z}} = \begin{bmatrix} \bar{t} \\ \bar{m} \\ \bar{p} \\ \bar{s} \\ \text{A}\bar{\text{ReS}} \\ \text{L}\bar{\text{ReS}} \\ \bar{W} \end{bmatrix} = \begin{bmatrix} 0.20 (0.17; 0.25) \\ 12.71 (12.57; 12.86) \\ 0.48 (0.46; 0.49) \\ 0.57 (0.56; 0.58) \\ 0.08 (0.07; 0.08) \\ 3.00 (2.54; 3.52) \\ 1.97 (1.67; 2.32) \end{bmatrix},$$

and $\mathbf{P}$ as shown in Table 2. As variance in breeding values was only estimated for lamb body mass, which was modelled on its natural scale, the additive genetic variance in mass on the expected scale is the one estimated in the model, 0.67 (95% CrI 0.07; 1.14). The fact that the correlations between lamb body mass and first-year survival (0.42, 95% CrI 0.33; 0.51) and between early pregnancy and first-year survival (0.11, 95% CrI 0.03; 0.19) are both positive suggests that the equilibrium between lamb body mass affecting first-year survival positively (direct effect) and negatively (through its effect on early pregnancy) is leaning towards the former (Tab. 2). To investigate the consequences of this trade-off in the selection of lamb body mass, we calculated its *extended* directional selection gradient. For trait z , the extended directional selection gradient is defined as the average derivative of expected fitness with respect to latent value (Morrissey, 2015; Villemereuil et al., 2016),

$$\eta_z = \frac{\partial \bar{W}(\bar{\mathbf{l}})}{\partial l_j} \bar{W}^{-1}. \quad (15)$$

This selection gradient thus represents the ultimate effect of small changes in phenotype (possibly, latent values for phenotypes that are modelled probabilistically, as is pregnancy) on fitness, scaled to relative fitness, averaged over the range of conditions experienced by individuals in the population. The algorithm for implementing the calculation described by equation 15 is detailed in appendix C.

The extended selection gradient for mass is positive, $\eta_m = 0.26$ (95% CrI 0.25; 0.28), whereas for lamb pregnancy the extended selection gradient is negative, $\eta_p = -0.56$ (95% CrI -0.64; -0.47). Importantly, the conflict between viability and fecundity is density-dependent, and is almost nonexistent when population density is very low. In such circumstances, virtually all ewe lambs survive, and therefore early pregnancy does not have a significant cost. As a consequence, η_p is higher (i.e., less strongly favouring low values) when population density is at its lowest and decreases with population

density (Fig. 4, Tab. 3). On the other hand, at high densities, when the probability of first-year survival is lowest, selection favouring larger body size is stronger.

To investigate the effect of pregnancy on the total effect of mass on fitness, we calculated a direct selection gradient, β_m , as opposed to the *extended* directional selection gradient, η_m . The direct selection gradient describes that part of the effect of a focal trait (mass) on fitness that occurs independently of other focal traits (in this case, pregnancy). We calculated the direct selection gradient using the twin model approach described in , which amounts to re-calculating the selection gradient for mass, but using a modified path diagram (as Figure ??), in which causal links from the focal trait to other traits (in Figure ??, this is the effect of mass on pregnancy) are omitted. β_m is thus blind to the positive association that exists between lamb body mass and early pregnancy, and therefore to pregnancy-mediated viability selection on body mass. On average direct and total effects of mass on fitness do not differ much, and β_m is only marginally larger (0.29; 95% CrI 0.27; 0.30). The reason for that lies again in the dependency of these quantities on population density. η_m and β_m are very similar when population density is very low, as viability selection tends to zero and, therefore, there is no cost to early pregnancy. Indeed, we show that η_m and β_m do diverge substantially with increasing population density (Fig. 4, Tab. 3), indicating that the mediation of selection for mass by pregnancy is highly environment-dependent.

The path analysis represented by Figure 3 invokes a definition of fitness that includes the entire life cycle. However, the mechanisms through which selection acts, survival and fecundity, occur at different points of the life cycle. Viability selection is particularly strong in the first annual cycle, whereas most offspring are reared at older ages. To evaluate the strength of the two mechanisms is, therefore, also to understand in which period selection is the strongest. We quantify the contribution of selection on lamb body mass that occurs during the first year of life using a path diagram similar to the one in Figure 3, but defining fitness as first-year survival. Measured this way, η_m is 0.12 (95% CrI 0.11; 0.13), and therefore corresponds to about half of the total lifetime selection on body mass. Selection acting on lamb body mass through viability costs during the first year of life is proportionally very low when population density is lowest (around 30%), is highest at intermediate populations densities and decreases when population density is very high. A slightly different pattern is obtained if β_m , instead of η_m , is used to quantify the strength of selection. In this case, the relative importance of viability selection does not decrease when population density is at its highest. These two trajectories, put together, suggest that at low densities, when both fecundity and survival are very high (Fig. 4), fecundity governs selection on lamb body mass, as mothers will successfully rear several offspring. As population size increases and first-year survival decreases (Fig. 4), viability se-

lection becomes proportionally more important as it occurs before any opportunity for reproduction. Under exceptional high population density, the probability of survival is very low. As most ewes die, fecundity selection becomes slightly more important. That is not observed when the correlation between lamb body mass and early pregnancy is dismissed, because the pregnancy-related viability costs are not considered.

Selection of the probabilistic reaction norm for early pregnancy

Finally, we investigated in more detail the conditions under which early pregnancy could be advantageous, or at least not maladaptive, to Soay sheep ewes. Knowledge of the selection pressures acting on different components of reaction norms of body mass, along with their potential to evolve, could shed light on the persistence of early pregnancy despite its large viability cost. In this context, we assess whether there is genetic variation in, and selection acting on, the slope of the reaction norm for pregnancy across differences in body mass and environment (population size). First, to quantify genetic variance, we fit a random regression animal model of the form

$$\ln \left(\frac{\mathbb{E}[p_{ik}]}{1 - \mathbb{E}[p_{ik}]} \right) = \alpha + \beta_m m_{ik} + \beta_a f(d_{ik}, 2) + u_{mk} + a_{int_i} + a_{sl_i} \times m_{ik} + \epsilon_{ijk}, \quad (16)$$

a version of Equation (2) where variance in breeding values is estimated, associated with both the intercept (a_{int}) and the slope for body mass (a_{sl}). As for assessing selection, we (1) evaluated the fitness of pregnant and non-pregnant ewes as a function of size and population density, thus obtaining *conditional individual fitness functions*, and we (2) calculated extended selection gradients for the intercept and slope for mass across different contexts of mean phenotype and environment. In (1) we were particularly interested in assessing whether the *conditional individual fitness functions* derived for pregnant and non-pregnant ewes would cross, as that would be a clear indication of variation in selection. In both analyses, we evaluated Equation (14) at fixed values of body mass and population density, setting the probability of pregnancy to 0 or 1 in analysis (1).

Fitting the random regression animal model to estimate additive genetic reaction norms of early pregnancy on body mass was numerically very challenging, both due to the binary nature of the response variable and the lack of repeated measures. Nonetheless, we found strong evidence of additive genetic variance in both the intercepts (2.04, 95% CrI 0.13; 9.19, App. B) and slopes (0.19, 95% CrI 0.05; 0.67) of the random regression being significantly different from zero, confirming a genetic basis of the reaction norms. Additionally, a positive additive genetic covariance between intercepts and slopes also seems to exist (0.61, 95% CrI 0.15; 2.35). On the other hand, we found very little evidence

of variation in selection acting on the slope for body mass (Fig. 6). Most conditional individual fitness functions for pregnant and non-pregnant ewes evaluated across a range of population densities never cross, suggesting that the effect of body mass on fitness does not differ qualitatively between pregnant and non-pregnant ewes. Although some of these functions cross, visual inspection suggests that this happens only when functions for pregnant and non-pregnant lambs are very similar (Fig. 6a). Finally, across sensible values of mean phenotype and environmental conditions, estimates of η for the slope of the reaction norm of early pregnancy on August mass are very close to zero (Fig. 6b), providing little evidence for variation in selection for plasticity of pregnancy.

Discussion

We have characterised a trade-off between viability and fecundity selection of body size in female Soay sheep lambs. Larger ewe lambs are more likely to become pregnant during their first annual cycle, and pregnant lambs are more likely to die. As a consequence, natural selection acts on lamb body mass through two distinct pathways: through the direct effect of lamb body mass on fitness, independent of early pregnancy status, and through its effect on early pregnancy. Selection on lamb body mass would, therefore, be stronger (more positive) if not for the relationship between this trait and early pregnancy. As a result, this correlation is one mechanism that constrains the evolution of body size in Soay sheep.

Recent theory on non-linear development systems allows one to decompose selection across its component causal pathways, essentially by using a system of path analyses that allows for non-linear effects and interactions (Morrissey, 2015; Henshaw et al., 2020). We used this theory to disentangle and quantify the direct effect of lamb body mass on fitness and that arising via its effect on early pregnancy. Alternatives to this approach are the use of traditional path analysis (developed for linear traits, e.g. Scheiner et al., 2000; Morrissey, 2014) or multi-response (linear or generalised linear) models that explicitly estimate (additive genetic) correlations among traits (e.g. Bonnet et al., 2017). Our approach specifically harnesses the phenotypic relationship between lamb body mass and early pregnancy, and builds it into a formal model of selection. Such an approach has the potential to be particularly useful when analytical power is a limiting constraint to infer the genetic architecture among traits. If genetic trade-offs are indeed a major explanation for the widespread mismatch between predicted evolutionary change and observed dynamics of body size (Blanckenhorn, 2000; Hansen & Houle, 2004; Kruuk et al., 2008), why are these methods not more frequently adopted? Information available suggests that the most considerable obstacle in dealing with the selection of correlated traits is not purely methodological. Rather, detecting genetic correlations seems to be a

limiting step, with evidence for genetic constraints in the wild being rather scarce (Kruuk et al., 2008). Reasons for this may include insufficient knowledge about the biology and ecology of a species (Pemberton, 2010), but also the presence of confounding effects between genes and environment. Negative genetic covariation between two traits might be obscured at the phenotypic level by environmental covariation among traits (Roff & Fairbairn, 2007; Morrissey et al., 2012c). Put together, evidence so far supports the hypothesis that the paradox of stasis generally has no single and simple solution and that a deep understanding of the species' biology may be critical to picking apart the causes of associations between traits and fitness.

Assessing the impact of early life events on fitness is often difficult because it requires data collection protocols that are not possible in all study systems. Nonetheless, the characterisation of individuals that do not survive is crucial in estimating evolutionary parameters. As an example, failing to identify and to characterise Soay sheep ewes that die due to pregnancy in their first year of life would have led to overestimating the strength of selection acting on August body mass. Essentially, this is because survival data would be due to data *missing not at random*: an invisible fraction of the population would have not been taken into account. Given that viability selection occurring during the first year of life is rather common among animals and plants, this source of bias may be common in studies estimating evolutionary parameters in wild populations (Hadfield, 2008; Nakagawa & Freckleton, 2008). The current study relies heavily on observing phenotypes that would have been expressed of individuals that had died: post-mortems of dead ewe lambs allow us to determine what their fecundity would have been the following spring. These data were instrumental for identifying the density-dependent manner in which pregnancy mediates total selection of body mass. This direct observation of a phenotype that would otherwise be missing is essentially a type of solution to the paradox of stasis, albeit one that may not be widely applicable in other systems.

Soay sheep have existed on an island, Soay, which is adjacent to Hirta, where the long-term study of Soay sheep occurs, since prehistoric times (Clutton-Brock & Pemberton, 2004a). Soay and Hirta have similar population dynamics, and as such, the associations among density, body size, pregnancy, and fitness reported here may be broadly representative of the long-term pattern selection experienced by Soay sheep. However, the current distribution of phenotype in the Soay sheep on Hirta, particularly their propensity to become pregnant in the first year, may be an adaptation to recent selection. Soay sheep were reintroduced to Hirta in 1932, when 107 individuals were moved from Soay (Clutton-Brock & Pemberton, 2004a), following the evacuation of the last human residents of Hirta. During the subsequent decades of much lower population size (Clutton-Brock et al., 2004), selection likely favoured a much faster life history. While results presented here are derived from data

for the range of observed population sizes since 1985, they are strongly indicative of what selection would happen at lower population densities. At lower population densities, survival rates of ewe lambs would be very high, regardless of pregnancy status (Figure A1). Additionally, the survival of offspring of ewes that became pregnant as lambs is extremely strongly density dependent (Table B5), and as such there would be a clear benefit, and little or no survival cost, to early reproduction at smaller population sizes, such as those experienced by Soay sheep on Hirta during a key period in the last century.

Summary

This study provides evidence of a trade-off constraining the evolution of body mass in an unmanaged population of an ungulate species. The Soay sheep population of Hirta has been established as an example of the paradox of stasis, and our results partly elucidate why selection of body mass may not be as strong as it initially appears. We show that selection towards larger body sizes would be stronger if not for the positive genetic correlation between lamb body mass and early pregnancy, and their opposing associations with fitness. Large lamb body mass in females is associated with higher rates of first-year survival and lifetime rearing success, but also with a higher chance of early pregnancy. The latter has a very large viability cost unless population density is low. More generally, this study underscores how a mechanistic understanding of life history patterns allows for a more fine-grained understanding of natural selection, with the potential to explain evolutionary conundrums such as the paradox of stasis.

Acknowledgements

The authors first wish to thank all project members and many volunteers for their contributions to field work on St. Kilda, and to all those who have contributed to keeping the project running over many years, including T. Clutton-Brock, M. Crawley, S. Albon, T. Coulson, L. Kruuk, and D. Nussey. The authors are also grateful to the National Trust for Scotland and Scottish Natural Heritage for permission to work on St. Kilda, and QinetiQ and Euresst for logistics and other support on the island. M. J. Janeiro was supported by a PhD scholarship (SFRH/BD/96078/2013) funded by the Fundação para a Ciência e Tecnologia (FCT). MBM is supported by a University Research Fellowship from the Royal Society (London).

References

- Anderson, M. 1995. *Sexual selection*. Princeton University Press, Princeton, New Jersey.
- Bérénos, C., Ellis, P.A., Pilkington, J.G. & Pemberton, J.M. 2014. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. *Molecular Ecology* **23**: 3434–3451.
- Blanckenhorn, W. 2000. The evolution of body size: what keeps organisms small? *Quarterly Review of Biology* **75**: 385–407.
- Bonnet, T., Wandeler, P., Camenisch, G. & Postma, E. 2017. Bigger is fitter? quantitative genetic decomposition of selection reveals an adaptive evolutionary decline of body mass in a wild rodent population. *PLOS Biology* **15**: 1–21.
- Bradshaw, A.D. 1991. The Croonian Lecture, 1991 Genostasis and the limits to evolution. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **333**: 289–305.
- Clutton-Brock, T., Price, O., Albon, S. & Jewell, P. 1992. Early development and population fluctuations in soay sheep. *Journal of Animal Ecology* **61**: 381–396.
- Clutton-Brock, T.H. 1988. *Reproductive Success*. University of Chicago Press.
- Clutton-Brock, T.H., Grenfell, B.T., Coulson, T., MacColl, A.D.C., Illius, A.W., Forchhammer, M.C., Wilson, K., Lindström, J., Crawley, M.J. & Albon, S.D. 2004. Population dynamics in soay sheep. In: *Soay sheep: dynamics and selection in an island population* (T.H. Clutton-Brock & J.M. Pemberton, eds). Cambridge University Press.
- Clutton-Brock, T.H. & Pemberton, J.M. 2004a. Individuals and populations. In: *Soay sheep: dynamics and selection in an island population* (T.H. Clutton-Brock & J.M. Pemberton, eds). Cambridge University Press.
- Clutton-Brock, T.H. & Pemberton, J.M. 2004b. *Soay Sheep Dynamics and Selection in an Island Population*. Cambridge University Press, Cambridge.
- Clutton-Brock, T.H., Stevenson, I.R., Marrow, P., MacColl, a.D., Houston, a.I. & McNamara, J.M. 1996. Population fluctuations, reproductive costs and life-history tactics in female Soay sheep. *Journal of Animal Ecology* **65**: 675–689.
- Crawley, M.J., Albon, S.D., Bazely, D.R., Milner, J.M., Pilkington, J.G. & Tuke, A.L. 2004. Vegetation and sheep population dynamics. In: *Soay sheep: dynamics and selection in an island population* (T.H. Clutton-Brock & J.M. Pemberton, eds). Cambridge University Press.

- 597 Estes, S. & Arnold, S.J. 2007. Resolving the paradox of stasis: models with stabilizing selection
598 explain evolutionary divergence on all timescales. *The American Naturalist* **169**: 227–244.
- 599 Fairbairn, D.J. 1997. Allometry for sexual size dimorphism: pattern and process in the coevolution
600 of body size in males and females. *Annual Review of Ecology and Systematics* **28**: 659–687.
- 601 Fairbairn, D.J., Blanckenhorn, W.U. & T, S. (eds) 2007. *Sex, Size and Gender Roles: Evolutionary*
602 *Studies of Sexual Size Dimorphism*. Oxford University Press, Oxford.
- 603 Falconer, D. 1981. *Introduction to Quantitative Genetics*, 2nd edn. Longman, New York.
- 604 Feulner, P.G.D., Gratten, J., Kijas, J.W., Visscher, P.M., Pemberton, J.M. & Slate, J. 2013. Intro-
605 gression and the fate of domesticated genes in a wild mammal population. *Molecular Ecology* **22**:
606 4210–4221.
- 607 Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian*
608 *Analysis* **1**: 515–533.
- 609 Hadfield, J.D. 2008. Estimating evolutionary parameters when viability selection is operating. *Pro-*
610 *ceedings of the Royal Society of London B: Biological Sciences* **275**: 723–734.
- 611 Hadfield, J.D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCM-
612 Cglmm R package. *Journal of Statistical Software* **33**: 1–22.
- 613 Hadfield, J.D., Richardson, D.S. & Burke, T. 2006. Towards unbiased parentage assignment: com-
614 bining genetic, behavioural and spatial data in a bayesian framework. *Molecular Ecology* **15**:
615 3715–3730.
- 616 Hansen, T.F., Armbruster, W., Carlson, M. & Pelabon, C. 2003. Evolvability and genetic constraints
617 in Dalechampia Blossoms: genetic correlations and conditional evolvability. *Journal of Experimen-*
618 *tal Zoology* **296B**: 23–39.
- 619 Hansen, T.F. & Houle, D. 2004. Evolvability, stabilizing selection, and the problem of stasis. In:
620 *Phenotypic integration* (M. Pigliucci & K. Preston, eds), pp. 130–150. Oxford University Press,
621 Oxford.
- 622 Henderson, C.R. 1975. Best linear unbiased estimation and prediction under a selection model.
623 *Biometrics* **31**: 423–447.
- 624 Henshaw, J.M., Morrissey, M.B. & Jones, A.G. 2020. Quantifying the causal pathways contributing
625 to natural selection. *Evolution* **in press**.
- 626 Honek, A. 1993. Intraespecific variation in body size and fecundity in insects: a general relationship.
627 *Oikos* **66**: 483–492.

- 628 Hunt, G. 2007. The relative importance of directional change, random walks, and stasis in the
629 evolution of fossil lineages. *Proceedings of the National Academy of Sciences of the United States*
630 *of America* **104**: 18404–8.
- 631 Hunter, D.C., Pemberton, J.M., Pilkington, J.G. & Morrissey, M.B. 2018. Quantification and decom-
632 position of environment-selection relationships. *Evolution* **72**: 851–866.
- 633 Johnston, S.E., Gratten, J., Berenos, C., Pilkington, J.G., Clutton-Brock, T.H., Pemberton, J.M. &
634 Slate, J. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation.
635 *Nature* **502**: 93–5.
- 636 Kingsolver, J.G. & Pfennig, D.W. 2004. Individual-level selection as a cause of Cope’s rule of phyletic
637 size increase. *Evolution* **58**: 1608–1612.
- 638 Knouft, J.H. & Page, L.M. 2003. The evolution of body size in extant groups of North American
639 freshwater fishes: speciation, size distributions, and Cope’s rule. *The American Naturalist* **161**:
640 413–421.
- 641 Kruuk, L.E.B., Slate, J. & Wilson, A.J. 2008. New Answers for Old Questions: The Evolution-
642 ary Quantitative Genetics of Wild Animal Populations. *Annual Review of Ecology Evolution and*
643 *Systematics* **39**: 525–548.
- 644 Larsson, K., van der Jeug, H.P., van der Veen, I.T. & Forslund, P. 1998. Body size declines despite
645 positive directional selection on heritable size traits in a barnacle goose population. *Evolution* **52**:
646 1169–1184.
- 647 Merilä, J., Sheldon, B.C. & Kruuk, L.E.B. 2001. Explaining stasis: Microevolutionary studies in
648 natural populations. *Genetica* **112–113**: 199–222.
- 649 Milner, J.M., Albon, S.D., Illius, A.W., Pemberton, J.M. & Clutton-Brock, T.H. 1999. Repeated
650 selection of morphometric traits in the Soay sheep on St Kilda. *Journal of Animal Ecology* **68**:
651 472–488.
- 652 Milner, J.M., Albon, S.D., Kruuk, L.E.B. & Pemberton, J.M. 2004. Selection on phenotype. In: *Soay*
653 *sheep: dynamics and selection in an island population* (T.H. Clutton-Brock & J.M. Pemberton,
654 eds), pp. 190–216. Cambridge University Press.
- 655 Milner, J.M., Pemberton, J.M., Brotherstone, S. & Albon, S.D. 2000. Estimating variance compo-
656 nents and heritabilities in the wild: a case study using the ‘animal model’ approach. *Journal of*
657 *Evolutionary Biology* **13**: 804–813.
- 658 Morrissey, M.B. 2014. Selection and evolution of causally covarying traits. *Evolution* **68**: 1748–1761.

- 659 Morrissey, M.B. 2015. Evolutionary quantitative genetics of nonlinear developmental systems. *Evo-*
660 *lution* **69**: 2050–2066.
- 661 Morrissey, M.B., Kruuk, L.E.B. & Wilson, A.J. 2010. The danger of applying the breeder’s equation
662 in observational studies of natural populations. *Journal of Evolutionary Biology* **23**: 2277–2288.
- 663 Morrissey, M.B., Parker, D.J., Korsten, P., Pemberton, J.M., Kruuk, L.E.B. & Wilson, A.J. 2012a.
664 The prediction of adaptive evolution: empirical application of the secondary theorem of selection
665 and comparison to the breeder’s equation. *Evolution* **66**: 2399–2410.
- 666 Morrissey, M.B., Parker, D.J., Korsten, P., Pemberton, J.M., Kruuk, L.E.B. & Wilson, A.J. 2012b.
667 The prediction of adaptive evolution: empirical application of the secondary theorem of selection
668 and comparison to the breeder’s equation. *Evolution* **66**: 2399–2410.
- 669 Morrissey, M.B., Walling, C.a., Wilson, A.J., Pemberton, J.M., Clutton-Brock, T.H. & Kruuk, L.E.B.
670 2012c. Genetic analysis of life-history constraint and evolution in a wild ungulate population. *The*
671 *American Naturalist* **179**: E97–114.
- 672 Nakagawa, S. & Freckleton, R.P. 2008. Missing inaction: the dangers of ignoring missing data. *Trends*
673 *in Ecology and Evolution* **23**: 592–596.
- 674 Nakagawa, S. & Schielzeth, H. 2010. Repeatability for Gaussian and non-Gaussian data: a practical
675 guide for biologists. *Biological Reviews of the Cambridge Philosophical Society* **85**: 935–56.
- 676 Ozgul, A., Tuljapurkar, S., Benton, T.G., Pemberton, J.M., Clutton-Brock, T.H. & Coulson, T. 2009.
677 The dynamics of phenotypic change and the shrinking sheep of St. Kilda. *Science* **325**: 464–7.
- 678 Pemberton, J.M. 2010. Evolution of quantitative traits in the wild: mind the ecology. *Philosophical*
679 *Transactions of the Royal Society B: Biological Sciences* **365**: 2431–2438.
- 680 Postma, E. 2014. Four decades of estimating heritabilities in wild vertebrate populations: Improved
681 methods, more data, better estimates? In: *Quantitative Genetics in the Wild*. Oxford (A. Char-
682 mantier, D. Garant & L.E.B. Kruuk, eds), pp. 16–33. Oxford University Press.
- 683 Rausher, M.D. 1992. The measurement of selection on quantitative traits: bi-
684 ases due to environmental covariances between traits and fitness. *Evo-*
685 *lution* **46**: 616–626. doi:10.1111/j.1558-5646.1992.tb02070.x. URL
686 <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1558-5646.1992.tb02070.x>.
- 687 Robert, C.P. & Casella, G. 2005. *Monte Carlo statistical methods*, 2nd edn. Springer texts in statistics.
688 Springer, Berlin. URL <https://cds.cern.ch/record/1187871>.

- 689 Robinson, M.R., Wilson, A.J., Pilkington, J.G., Clutton-Brock, T.H., Pemberton, J.M. & Kruuk,
690 L.E.B. 2009. The impact of environmental heterogeneity on genetic architecture in a wild population
691 of soay sheep. *Genetics* **181**: 1639–1648.
- 692 Roff, D.A. 2000. Trade-offs between growth and reproduction : an analysis of the quantitative genetic
693 evidence. *Journal of Evolutionary Biology* **13**: 434–445.
- 694 Roff, D.A. & Fairbairn, D.J. 2007. The evolution of trade-offs: where are we? *Journal of Evolutionary*
695 *Biology* **20**: 433–47.
- 696 Scheiner, S.M., Mitchell, R.J. & Callahan, H.S. 2000. Using path analysis to measure natural selection.
697 *Journal of Evolutionary Biology* **13**: 423–433.
- 698 Shine, R. 1988. The evolution of large body size in females: a critique of Darwin’s fecundity advantage
699 model. *American Naturalist* **131**: 124–131.
- 700 Sober, E. 1986. Force and disposition in evolutionary theory. In: *Minds, machines and evolution*
701 (C. Hookway, ed), pp. 43–62. Cambridge University Press, Cambridge.
- 702 Sogard, S. 1997. Size-selective mortality in the juvenile stage of teleost fishes: a review. *Bulletin of*
703 *Marine Science* **60**: 1129–1157.
- 704 Sokolovska, N., Rowe, L. & Johansson, F. 2000. Fitness and body size in mature odonates. *Ecological*
705 *Entomology* **25**: 239–248.
- 706 Villemereuil, P.D., Schielzeth, H., Nakagawa, S. & Morrissey, M. 2016. General methods for evolu-
707 tionary quantitative genetic inference from generalised mixed models. *Genetics* **204**: 1281–1294.
- 708 Wilson, A.J., Pemberton, J.M., Pilkington, J.G., Clutton-Brock, T.H., Coltman, D.W. & Kruuk,
709 L.E.B. 2007. Quantitative genetics of growth and cryptic evolution of body size in an island
710 population. *Evolutionary Ecology* **21**: 337–356.
- 711 Wright, S. 1921. Systems of mating. I. The biometric relations between parent and offspring. *Genetics*
712 **6**: 111–123.
- 713 Wright, S. 1934. The method of path coefficients. *The Annals of Mathematical Statistics* **5**: 161–215.
- 714 Wright, S. 1960. Path coefficients and path regressions: alternative or complementary concepts?
715 *Biometrics* **16**: 189–202.

Table 1: Coefficients on the latent and expected scales of the multi-response animal model of lamb body mass and early pregnancy in Soay sheep. Lamb mass was modelled on its natural scale, and therefore no distinction between latent and expected scales is necessary. Under the binomial response model for pregnancy, variance components are not meaningful on the expected scale, though means and correlations are. Fixed effects from this model are reported in appendix B.

	latent scale		conversions to the expected scale	
	posterior mode	95% CrI	posterior mode	95% CrI
μ_m	13.82	(12.22; 14.26)		
μ_p	0.00	(-0.85; 0.64)	0.50	(0.39; 0.59)
$\sigma_{m_a}^2$	0.78	(0.20; 1.27)		
$\sigma_{p_a}^2$	0.69	(0.00; 5.86)		
σ_{mp_a}	0.31	(0.00; 1.49)		
ρ_{mp_a}	0.64	(0.24; 0.90)	0.64	(0.24; 0.90)
σ_m^2	4.29	(3.84; 5.23)		
σ_p^2	3.20	(1.99; 10.12)		
σ_{mp}	3.24	(1.49; 3.50)		
ρ_{mp}	0.51	(0.40; 0.66)	0.49	(0.38; 0.64)

Table 2: Phenotypic variance-covariance matrix for the path diagram in Fig. 4, including the probability of twinning (t), August body mass (m), the probability of early pregnancy (p), the probability of survival (s), first year rearing success ($ares$), subsequent lifetime rearing success ($lres$), and fitness, defined as lifetime rearing success (ARes+LReS) in surviving ewes weighed by the probability of survival. The diagonal corresponds to variances, the upper off-diagonal to covariances and the lower off-diagonal to correlations among traits. Values shown correspond to the expected scale.

	t	m	p	s	ares	lres	W
t	0.09 (0.06; 0.11)	-0.25 (-0.34; -0.16)	-0.01 (-0.02; -0.01)	-0.02 (-0.02; -0.01)	0.00 (0.00; 0.00)	-0.05 (-0.13; 0.01)	-0.09 (-0.15; -0.04)
m	-0.43 (-0.49; -0.34)	3.92 (3.54; 4.55)	0.31 (0.27; 0.36)	0.25 (0.20; 0.29)	0.04 (0.03; 0.05)	1.26 (0.86; 2.02)	1.45 (1.17; 2.18)
p	-0.14 (-0.22; -0.08)	0.57 (0.48; 0.64)	0.08 (0.07; 0.09)	0.01 (0.00; 0.02)	0.00 (0.00; 0.00)	0.10 (0.04; 0.18)	0.09 (0.04; 0.15)
s	-0.21 (-0.27; -0.13)	0.42 (0.33; 0.51)	0.11 (0.03; 0.19)	0.09 (0.07; 0.11)	0.00 (0.00; 0.00)	0.08 (0.03; 0.16)	0.31 (0.23; 0.40)
ares	-0.12 (-0.17; -0.06)	0.31 (0.25; 0.36)	0.15 (0.09; 0.21)	0.14 (0.08; 0.19)	0.00 (0.00; 0.01)	0.01 (0.00; 0.04)	0.02 (0.01; 0.04)
lres	-0.05 (-0.11; 0.00)	0.19 (0.11; 0.26)	0.10 (0.03; 0.16)	0.08 (0.02; 0.13)	0.05 (-0.01; 0.13)	11.39 (4.93; 36.23)	8.51 (3.74; 25.96)
W	-0.09 (-0.15; -0.05)	0.27 (0.18; 0.35)	0.10 (0.03; 0.16)	0.34 (0.24; 0.43)	0.11 (0.03; 0.18)	0.87 (0.79; 0.93)	7.55 (3.88; 23.60)

Table 3: Selection gradients of lamb body mass and early pregnancy for three different values of observed population density - minimum, mean and maximum (321, 452, and 671, respectively). Raw (η) and standardised selection gradients, either by the observed standard deviation (η_{sd}) or the observed mean (η_{μ}), are shown. For lamb body mass, selection gradients considering first-year survival as a measure of fitness are also included (η_s).

		d=321	d=452	d=671
mass	η	0.15 (0.14; 0.16)	0.28 (0.27; 0.30)	0.33 (0.30; 0.36)
	η_{sd}	0.35 (0.33; 0.36)	0.66 (0.62; 0.70)	0.76 (0.69; 0.84)
	η_{μ}	1.86 (1.77; 1.96)	3.53 (3.31; 3.76)	4.11 (3.71; 4.50)
	η_s	0.04 (0.04; 0.05)	0.15 (0.13; 0.17)	0.13 (0.11; 0.16)
	$\eta_{s_{sd}}$	0.10 (0.09; 0.11)	0.34 (0.30; 0.38)	0.31 (0.25; 0.37)
	$\eta_{s_{\mu}}$	0.53 (0.46; 0.59)	1.83 (1.62; 2.06)	1.66 (1.37; 1.97)
pregnancy	η	0.01 (0.00; 0.02)	-0.40 (-0.45; -0.34)	-1.79 (-2.08; -1.48)
	η_{sd}	0.00 (0.00; 0.01)	-0.20 (-0.22; -0.17)	-0.89 (-1.04; -0.74)

Figure 1: (a) Number of Soay sheep ewes, ewe lambs and population size from 1991 to 2015 in the Village Bay study area; (b) empirical distribution of body mass in pregnant and non-pregnant ewe lambs; (c) size-dependent probability of early pregnancy. The line in teal corresponds to the prediction of a binomial regression with logit link function (see Appendix B for full specification and parameter estimates), whereas the grey bars correspond to 95% confidence intervals on the binomial probability obtained from the raw data; (d) lamb body mass (upper panel) and probability of early pregnancy (lower panel) across time. The black lines correspond to predictions of a linear and a binomial model (logit link function), respectively, correcting for within-year birth and measurement dates, maternal age and population density (linear and quadratic terms for the latter two covariates). Random effects for cohort and maternal identity were also considered. The grey areas illustrate the variation in lamb body mass and probability of early pregnancy attributed to varying population density in 1 standard deviation from the mean; (e) non-parametric survival curves of Soay sheep according to mother's age at conception. The p value corresponds to a Peto & Peto test for differences among Kaplan-Meier curves.

Figure 2: Probability of first-year survival (a, c) and lifetime rearing success, LReS, (b, d) in Soay sheep ewes as a function of lamb body mass (a, b), and both lamb body mass and early pregnancy status (c, d). Dots correspond to observed data and curves to model predictions. Binomial regressions for first-year survival with logit link function were fit to the data. LReS was estimated in ewes that survived their first year of life and Poisson regressions with exponential link function were fit to the data. In all models, population density and maternal age were included as covariates (linear and quadratic terms were estimated for both), as well as birth and measurement dates (see Tab. B4 and B6 in App. B). Variance among cohorts and maternal identities were also estimated.

Figure 3: Path diagram representing the phenotypic landscape defined in Equation (14). The probability of twinning, *twin*, has an effect on all the other traits, lamb body *mass* affects the probabilities of being pregnant as a lamb, *preg*, surviving the first annual cycle, *surv*, and also first year and subsequent lifetime rearing success, *AReS* and *LReS*; and *preg* affects *surv*. Fitness, *W*, is defined as the sum of *AReS* and *LReS* on surviving ewes weighted by the probability of first-year survival.

Figure 4: The effect of population density on traits depicted in the path diagram in Figure 3, and consequently on selection of lamb body mass and early pregnancy. The upper panel shows variation in mean twinning (t), early pregnancy (p), first-year survival (s), and mean-centred lamb body mass (m). The three lower panels include functions of mean fitness, selection of lamb body mass (η_m and β_m), and selection of early pregnancy (η_p) on population density. All plots include an histogram showing the empirical distribution of population density. Areas in teal and purple correspond to 95% highest posterior density credible intervals.

Figure 5: Proportion of natural selection on lamb body mass occurring through viability selection in Soay sheep ewes as a function of population density. This proportion was obtained both using the extended (η_m) and non-extended (β_m) directional selection gradient of lamb body mass. The histogram in grey shows the empirical distribution of population density.

Figure 6: (a) Conditional individual fitness functions of (mean) August body mass for pregnant and non-pregnant Soay sheep ewes and (b) selection gradients for the intercept (b1) and slope of early pregnancy on body mass (b2) as a function of population density and body mass. Body mass is mean-centred in both (a) and (b). Note that if the effect of body mass on the strength of selection varies between pregnant and non-pregnant ewes the conditional individual fitness functions in (a) are expected to cross.

716 **A Supplementary figure**

Figure A1: Probability of first-year survival in pregnant (a) and non-pregnant (b) ewe lambs, and AReS (c) and LReS (d) in Soay sheep ewes that survived their first annual cycle as a function of lamb body mass and population size. These are predictions based on models in Tables B4, B5 and B6.

717 **B Model summary tables**

718 Lamb body mass in August is dependent on birth and measurement dates. Both are indicators of how
719 long lambs had the opportunity to grow and are particularly relevant because higher growth rates
720 occur during the warmer months (before August), when vegetation is most available (Crawley et al.,
721 2004). As a result, Julian birth and measurement dates were included as covariates when modelling
722 lamb body mass or lamb body mass was used as a predictor. Fixed effects for twinning status,
723 population density and maternal age at parity, including quadratic terms for the latter two, were
724 included in all regressions, as well as the following linear interaction terms: between lamb body mass
725 and population density in models including lamb body mass as predictor, between early pregnancy

status and population density in models including early pregnancy as a predictor, and between these three traits in models including both lamb body mass and early pregnancy status as predictors. Likewise, random effects to estimate among-mother and among-cohort variation were included in all regressions. This fixed and random structure was applied to all statistical models except the one modelling AReS (due to a lack of statistical power).

Table B1: Coefficients of a linear regression of August body mass in Soay ewe lambs. All covariates, except for twin status, were mean-centred. Variances in breeding values, random cohorts and maternal identity were estimated. 95% credible intervals correspond to HPD intervals. See Eq. (1) (main text) for model description.

parameter	posterior mode	95% CrI
intercept	1.60	(1.02; 2.07)
density ($\times 100$)	-0.27	(-0.59; 0.05)
density ² ($\times 100$)	0.00	(-0.01; 0.00)
twin	-2.91	(-3.21; -2.62)
maternal age	0.36	(0.29; 0.40)
maternal age ²	-0.11	(-0.13; -0.10)
birth date	-0.06	(-0.08; -0.05)
measurement date	0.09	(0.04; 0.13)
additive genetic variance:	0.67 (0.07; 1.14)	
variance in cohort effects:	0.71 (0.33; 1.08)	
variance in maternal effects:	0.59 (0.29; 1.21)	
residual variance:	1.81 (1.47; 2.33)	

Table B2: Coefficients of a binomial regression (logit link function) of early pregnancy on body mass in Soay ewe lambs. Other fixed effects include population density, twin status, and maternal age. All covariates, except for twin status, were mean-centred. Variances in random cohorts and maternal identity were also estimated. 95% credible intervals correspond to HPD intervals. See Eq. (2) (main text) for model description.

parameter	posterior mode	95% CrI
intercept	-0.81	(-1.39; -0.12)
body mass	0.69	(0.49; 0.87)
body mass ²	-0.11	(-0.18; -0.07)
body mass ³	0.01	(-0.01; 0.02)
density ($\times 100$)	-0.45	(-0.89; -0.15)
density ² ($\times 100$)	0.00	(0.00; 0.01)
density $\times$ body mass ($\times 100$)	0.11	(0.02; 0.23)
twin	0.76	(0.17; 1.45)
maternal age	-0.03	(-0.13; 0.07)
maternal age ²	0.00	(-0.03; 0.03)
birth date	-0.02	(-0.05; 0.02)
measurement date	-0.07	(-0.12; 0.01)
variance in cohort effects:	0.46 (0.14; 1.27)	
variance in maternal effects:	0.01 (0.00; 1.34)	
residual variance:	set to 1	

Table B3: Coefficients of the bivariate model of lamb body mass and early pregnancy in Soay sheep ewes. See Eq. (3) and (4) (main text) for model description.

parameter		posterior mode	95% CrI
fixed effects			
body mass	intercept	13.821	(13.215; 14.261)
	density	-0.002	(-0.005; 0.001)
	density ²	0.000	(0.000; 0.000)
	twinning	-2.972	(-3.245; -2.652)
	maternal age	0.391	(0.330; 0.436)
	maternal age ²	-0.108	(-0.125; -0.095)
pregnancy	intercept	0.000	(-0.850; 0.636)
	density	-0.005	(-0.009; -0.001)
	density ²	0.000	(0.000; 0.000)
	twinning	1.851	(1.055; 2.273)
	maternal age	-0.140	(-0.268; -0.023)
	maternal age ²	0.043	(-0.003; 0.071)
random effects			
breeding values	body mass	0.785	(0.199; 1.272)
	body mass:pregnancy	0.307	(-0.001; 1.494)
	pregnancy	0.695	(0.001; 5.859)
cohort	body mass	0.596	(0.266; 1.141)
	body mass:pregnancy	0.385	(0.136; 1.163)
	pregnancy	0.932	(0.136; 1.163)
maternal ID	body mass	0.577	(0.349; 1.014)
	body mass:pregnancy	0.002	(-0.159; 0.414)
	pregnancy	0.005	(0.000; 0.824)
residuals	body mass	1.940	(1.387; 2.224)
	body mass:pregnancy	1.059	(0.699; 1.267)
	pregnancy	1.000	(1.000; 1.000)

Table B4: Coefficients of mixed effect binomial regressions with logit link function exploring the association of first-year survival with size and pregnancy in Soay ewes; (a) size-dependent first-year survival, (b) first-year survival as a function of size and pregnancy. Covariates except for twin and pregnancy status were mean-centred. See Eq. (10) (main text) for model description.

parameter	posterior mode	95% CrI
(a)		
intercept	-0.55	(-1.59; 0.38)
body mass	0.50	(0.36; 0.66)
density ($\times 100$)	-1.38	(-1.95; -0.72)
density ² ($\times 100$)	0.01	(0.00; 0.01)
density $\times$ body mass ($\times 100$)	-0.06	(-0.18; 0.05)
twin	0.06	(-0.70; 0.74)
maternal age	-0.02	(-0.12; 0.12)
maternal age ²	0.01	(-0.03; 0.05)
birth date	0.02	(-0.02; 0.05)
measurement date	-0.10	(-0.20; -0.03)
variance in cohort effects: 0.78 (0.78; 4.18)		
variance in maternal effects: 1.82 (0.61; 4.22)		
residual variance: set to 1		
(b)		
intercept	0.14	(-0.95; 1.23)
pregnancy	-2.15	(-2.73; -1.35)
body mass	0.69	(0.47; 0.89)
body mass $\times$ pregnancy	0.19	(-0.18; 0.43)
density ($\times 100$)	-0.89	(-1.75; -0.28)
density ² ($\times 100$)	0.00	(0.00; 0.01)
density $\times$ pregnancy ($\times 100$)	-1.35	(-2.04; -0.70)
density $\times$ body mass ($\times 100$)	0.10	(-0.07; 0.24)
density $\times$ body mass $\times$ pregnancy ($\times 100$)	-0.04	(-0.34; 0.27)
twin	0.19	(-0.74; 0.88)
maternal age	0.04	(-0.12; 0.14)
maternal age ²	0.00	(-0.04; 0.04)
birth date	0.01	(-0.03; 0.05)
measurement date	-0.15	(-0.23; -0.04)
variance in cohort effects: 2.04 (0.98; 5.11)		
variance in maternal effects: 2.23 (0.85; 4.89)		
residual variance: set to 1		

Table B5: Coefficients of a binomial regression (logit link function) of AReS in surviving ewe lambs on lamb body mass, birth and measurement dates, and population density in Soay sheep ewe lambs. All covariates were mean-centred and the 95% credible intervals correspond to HPD intervals. See Eq. (11) (main text) for model description.

Parameter	Posterior mode	95% CrI
intercept	-4.00	(-5.08; -2.94)
body mass	0.64	(0.43; 1.02)
density ($\times 100$)	-1.06	(-2.41; -0.11)
density ² ($\times 100$)	0.00	(0.00; 0.01)
birth date	0.03	(-0.04; 0.10)
measurement date	0.06	(-0.11; 0.16)
residual variance: set to 1		

Table B6: Coefficients of Poisson regressions with logarithm link functions, exploring the association between LReS in ewes surviving their first year and lamb body mass and early pregnancy; (a) size-dependent LReS, and (b) LReS as a function of lamb August body mass and early pregnancy. Covariates except for twin status and early pregnancy status were mean-centred, and 95% credible intervals correspond to HPD intervals. See Eq. (12) (main text) for model description.

parameter	posterior mode	95% CrI
(a)		
intercept	0.64	(0.12; 0.99)
body mass	0.14	(0.04; 0.23)
density ($\times 100$)	-0.15	(-0.49; 0.01)
density ² ($\times 100$)	0.00	(0.00; 0.00)
density $\times$ body mass ($\times 100$)	0.04	(-0.06; 0.09)
twin	0.07	(-0.38; 0.52)
maternal age	0.00	(-0.08; 0.06)
maternal age ²	0.01	(0.00; 0.03)
birth date	0.01	(-0.02; 0.03)
measurement date	0.01	(-0.06; 0.07)
variance in cohort effects:	0.00 (0.00; 0.37)	
variance in maternal effects:	0.38 (0.00; 0.68)	
residual variance:	0.45 (0.20; 0.86)	
(b)		
intercept	0.38	(-0.07; 0.83)
pregnancy	0.37	(-0.06; 0.82)
body mass	0.14	(0.03; 0.27)
pregnancy $\times$ body mass	0.02	(-0.28; 0.09)
density ($\times 100$)	-0.25	(-0.52; 0.06)
density ² ($\times 100$)	0.00	(0.00; 0.00)
density $\times$ pregnancy ($\times 100$)	0.34	(-0.09; 0.75)
density $\times$ body mass ($\times 100$)	0.10	(0.00; 0.21)
density $\times$ body mass $\times$ pregnancy ($\times 100$)	-0.24	(-0.39; -0.06)
twin	0.14	(-0.34; 0.57)
maternal age	0.00	(-0.08; 0.07)
maternal age ²	0.02	(0.00; 0.04)
birth date	0.00	(-0.02; 0.03)
measurement date	0.02	(-0.05; 0.07)
variance in cohort effects:	0.00 (0.00; 0.38)	
variance in maternal effects:	0.30 (0.00; 0.60)	
residual variance:	0.51 (0.26; 0.91)	

Table B7: Coefficients of a binomial regression (logit link function) of the probability of twinning as a function of mean-centred population density and maternal age. 95% credible intervals correspond to HPD intervals. See Eq. (13) (main text) for model description.

parameter	posterior mode	95% CrI
intercept	-2.47	(-3.37; -1.43)
density ($\times 100$)	-0.07	(-0.47; 0.55)
density ² ($\times 100$)	-0.01	(-0.01; 0.00)
maternal age	0.90	(0.69; 1.14)
maternal age ²	-0.15	(-0.20; -0.08)
variance in cohort effects:	0.72 (0.15; 2.59)	
variance in maternal effects:	9.25 (5.02; 14.35)	
residual variance:	set to 1	

Table B8: Coefficients of a random regression animal model to estimate additive genetic reaction norms of early pregnancy on body mass in Soay ewe lambs. A binomial regression with logit link function was fit to the data. All covariates were mean-centred. Variance in random cohorts was also estimated. 95% credible intervals correspond to HPD intervals.

parameter	posterior mode	95% CrI
intercept	-0.88	(-1.63; -0.17)
body mass	0.74	(0.53; 0.98)
density ($\times 100$)	-0.46	(-1.02; -0.08)
density ² ($\times 100$)	0.00	(0.00; 0.01)
variance in breeding values (intercepts):	2.04 (0.13; 9.19)	
variance in breeding values (slopes for mass):	0.19 (0.05; 0.67)	
covariance between breeding values (intercepts and slopes):	0.61 (0.15; 2.35)	
variance in cohort effects:	0.80 (0.19; 2.25)	
residual variance:	set to 1	

C Details on solving derivatives and integrals

Estimating average trait values (Eq. (5), main text) and variances (Eq. (6), main text) on the expected scale implies solving integrals of high dimensionality in the context of the path analysis, and particularly when these equations are applied to the vector-valued function $\mathbf{f}(\mathbf{l})$ in equation (14). In such circumstances, solving integrals using Monte Carlo simulations is particularly useful as the error is independent from the integrand dimensionality, but mainly because other numerical methods are likely to fail. Monte Carlo integration relies on the fact that the integral $\int_L g(\mathbf{l})f(\mathbf{l})d\mathbf{l}$ (or $\int_L g^{-1}(\mathbf{l})f(\mathbf{l})d\mathbf{l}$ if applied to an inverse link function) corresponds to the expectation of g , $\mathbb{E}_f[g(\mathbf{L})]$. As a consequence, a random sample of latent values, L , generated using its density f can be used to approximate the empirical average of g ,

$$\bar{g} \approx \frac{1}{n} \sum_{i=1}^n g(\mathbf{l}_i), \quad (17)$$

where n is the number of draws from L . $\bar{g}$ is also an approximation to $\int_L g(\mathbf{l})f(\mathbf{l})d\mathbf{l}$. This approximation converges to the real value of the integral by the law of large numbers as $n \rightarrow \infty$ (Robert & Casella, 2005). We sampled a large enough number of values from the distribution of the latent traits on which we evaluated $\mathbf{f}(\mathbf{l})$. Since the statistical models were fitted in a Bayesian framework, this procedure was applied to $n_{post} = 2000$ posterior independent samples.

For lower dimensional integrals, and particularly to obtain Φ (average derivative of the expected values with respect to the latent ones) in equation (8) (main text), we used function *cuhre* from R package **R2Cuba** that implements a deterministic algorithm that uses cubature rules of polynomial degree, also allowing for multidimensional integration.

Likewise, the differentiation in the calculation of extended selection gradients (Eq. 15, main text) was accomplished by numerical linear approximation,

$$W'(l) \approx \frac{W(l+h) - W(l)}{h}, \quad (18)$$

where h is some small value. So, the extended selection gradient for mass would be obtained by evaluating W for a certain value of latent mass l and at for $l+h$. This procedure was applied to $n_{MC} = 2000$ values from the distribution of the latent distributions on which we evaluated $\mathbf{f}(\mathbf{l})$ times $n_{post} = 2000$ posterior independent samples.