Evol Ecol
DOI 10.1007/s10682-006-9106-z

ORIGINAL PAPER

Quantitative genetics of growth and cryptic evolution of body size in an island population

A. J. Wilson · J. M. Pemberton · J. G. Pilkington · T. H. Clutton-Brock · D. W. Coltman · L. E. B. Kruuk

Received: 25 April 2006 / Accepted: 2 August 2006

Abstract While evolution occurs when selection acts on a heritable trait, empirical studies of natural systems have frequently reported phenotypic stasis under these conditions. We performed quantitative genetic analyses of weight and hindleg length in a free-living population of Soay sheep (*Ovis aries*) to test whether genetic constraints can explain previously reported stasis in body size despite evidence for strong positive directional selection. Genetic, maternal and environmental covariance structures were estimated across ontogeny using random regression animal models. Heritability increased with age for weight and hindleg length, though both measures of size were highly heritable across ontogeny. Genetic correlations among ages were generally strong and uniformly positive, and the covariance structures were also highly integrated across ontogeny. Consequently, we found no constraint to the evolution of larger size itself. Rather we expect size at all ages to increase in response to positive selection acting at any age. Consistent with expectation, predicted breeding values for age-specific size traits have increased over a twenty-year period, while maternal performance for offspring size has declined. Re-examination of the phenotypic data confirmed that sheep are not getting larger, but also showed that there are significant negative trends in size at all ages. The genetic evolution is therefore cryptic, with the response to selection presumably being masked at the phenotypic level by a plastic response to changing environmental conditions. Density-dependence, coupled with systematically increasing population size, may contribute to declining body size but is insufficient to completely explain it. Our results demonstrate that an increased understanding of the genetic basis of

A. J. Wilson (✉) · J. M. Pemberton · J. G. Pilkington · L. E. B. Kruuk
Institute of Evolutionary Biology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JT, UK
e-mail: Alastair.Wilson@ed.ac.uk

T. H. Clutton-Brock
Department of Zoology, University of Cambridge, Downing Street, Cambridge, UK

D. W. Coltman
Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada T6G2E9

quantitative traits, and of how plasticity and microevolution can occur simultaneously, is necessary for developing predictive models of phenotypic change in nature.

Keywords heritability · Ovis aries · ontogeny · cryptic evolution · growth

Introduction

Large size frequently confers fitness advantages in the form of increased survival or fecundity in animals. Consequently, while not ubiquitous (Gaillard et al. 2000), positive directional selection on size traits has been commonly reported (e.g., Guinness et al. 1978; Sogard 1997; Merilä et al. 2001a) giving rise to interest in the question of what keeps organisms small (Blanckenhorn 2000). In this study we examine size in a free-living ungulate population in which animals are not getting bigger despite such directional selection. We test two alternate hypotheses that might explain this observation; firstly, that genetic constraints prevent the evolution of body size; and, secondly, that genetic evolution of larger size is occurring but that these changes are masked at the phenotypic level by changes in environmental conditions.

We focus on body size in a free-living population of Soay sheep, *Ovis aries*, resident on the Scottish island of Hirta in the St. Kilda archipelago. While Soay sheep are small relative to other breeds of domestic sheep, size is known to be under selection in this population. Although it is difficult to absolutely preclude the possibility that large size could impose a fitness cost that is yet to be measured, prior studies of this system have consistently provided evidence for positive directional selection on body size. For example, increased birth weight is associated with higher juvenile viability and lifetime fitness (Clutton-Brock et al. 1992; Wilson et al. 2005c), and adult size traits are positively correlated with both survival and reproductive success (Coltman et al. 1999; Milner et al. 1999). Evolution is expected to occur when selection acts on a trait that has a heritable component of phenotypic variation (Falconer and Mackay 1996), and significant heritabilities for size traits (e.g. weight, leg length) have also been estimated in this population (Milner et al. 2000). While theory therefore predicts the evolution of larger size, previous analysis has found no evidence that sheep are getting bigger (Milner et al. 2000).

Phenotypic stasis in traits under selection has been frequently reported in natural populations (Merilä et al. 2001b; Kruuk et al. 2002). One possible explanation for this phenomenon is that stasis reflects genetic constraints arising either from a lack of heritable variation, or from genetic correlations between traits under selection (Arnold 1992). The latter will occur if genetic correlations are negative between traits under similar directions of selection, or if positively correlated traits are subject to antagonistic selection regimes (e.g., Wilson et al. 2003). To date, analyses of genetic correlations between body size and other phenotypic traits have provided little evidence for this type of constraint in natural systems (e.g., Coltman et al. 2005). However, while important traits may have been overlooked in studies to date, it has also been argued that genetic constraints will not always be apparent from consideration of h^2 and r_G alone (Pease and Bull 1988), and that multivariate statistical methods (e.g., eigenvector analysis of genetic variance–covariance matrices) should additionally be used for their detection (Blows and Hoffman 2005).

Significant heritabilities (h^2) for body size have been shown in a number of wild vertebrate systems (e.g., Kruuk et al. 2001; Pakkasmaa et al. 2003), though most studies to date have been performed in a univariate framework. Specifically, they have either focused on size as defined at a single point in ontogeny, or have used analytical methods that implicitly assume constancy of h^2 with age. However, if size is viewed as a series of age-specific traits, then genetic and environmental components of phenotypic variance (and hence h^2) are frequently found to vary over ontogeny (Cheverud et al. 1983a; Réale et al. 1999). Failing to account for this may result in misleading expectations of evolutionary change, particularly if selection acts differently at different ages. Furthermore, genetic covariance structures among age-specific size traits might impose now-later trade-offs (Wilson et al. 2003), or more generally limit the potential for growth trajectories to evolve in certain directions (Kirkpatrick and Lofsvold 1992).

It is therefore expected that multivariate approaches should provide more flexible (and biologically realistic) models of genetic architecture over ontogeny. However, application to natural systems has been limited by statistical considerations since individuals are frequently not sampled at all measurement ages (due to both mortality and field-sampling limitations). This is particularly problematic for determining genetic covariances among ages since accurate and precise estimation generally requires large sample sizes (Lynch and Walsh 1998). Consequently, comparatively few studies have estimated genetic covariance structures for trait ontogenies outside the laboratory, except over limited ontogenetic periods where repeated sampling is relatively feasible (e.g. analyses of nestling growth in passerines; Björklund 1997; Badyaev and Martin 2000).

Nevertheless, the statistical problems associated with incomplete sampling can be reduced through recently developed random regression animal models. If growth is viewed as an infinite-dimensional trait, then each individual's ontogenetic trajectory is defined by a potentially infinite series of size traits along a temporal axis corresponding to age (Kirkpatrick et al. 1990). Within this framework it is possible to model an individual's genetic merit, or ''breeding value'' as a covariance function of age (Kirkpatrick et al. 1990; Meyer 1998; Schaeffer 2004), a technique that allows estimation of age-specific heritabilities and genetic correlations between ages. Importantly, in comparison to more traditional multiple-trait analyses, random regression allows more efficient use of the incomplete data sets typically obtained in natural populations (Wilson et al. 2005b).

Here we apply the random regression approach to quantitative genetic analyses of size and growth in Soay sheep. We estimate the additive genetic variance–covariance matrix for age-specific size traits in order to test for genetic constraints in a multivariate framework. Specifically we test whether genetic constraints can explain the lack of a systematic increase in Soay sheep body size. Furthermore, since our hypothesis of genetic constraint is based on previously reported phenotypic stasis (Milner et al. 2000), we re-examine phenotypic trends (or lack thereof) using an extended data set, and explicitly test for genetic responses to selection using predicted breeding values (PBV). The predicted breeding value is a measure of the additive genetic merit of an individual for the trait in question (Lynch and Walsh 1998), and can be thought of as an estimate of the individual's phenotype corrected, albeit imperfectly (Postma 2006), for environmental effects. In the absence of any genetic constraint it is possible that a genetic response to selection has occurred but is masked at the phenotypic level by environmental effects on body size.

Consideration of predicted breeding values allows us to explicitly test this alternate hypothesis for phenotypic stasis.

Materials and Methods

Data and pedigree structure

The Village Bay population of Soay sheep (*Ovis aries*) is resident on the Scottish island of Hirta, in the St. Kilda archipelago in the North Atlantic (57(49(N, 08(34(W). This population has been subject to long-term study, with morphological, life history and genetic data collected on individually tagged sheep throughout their lives since 1985. We used two different measures of body size, weight and hindleg length, that are phenotypically and genetically correlated (Milner et al. 2000; Coltman et al. 2001). Body weight is measured in lambs (within a few days of birth), and both weight and hindleg length are recorded during annual round ups of older animals conducted each August, during which over 50% of the population is captured. More extensive details of the project and field methodology are presented elsewhere (see Clutton-Brock and Pemberton 2004).

Soay sheep are sexually dimorphic, and early size traits are also influenced by natal litter size (twins being smaller than singletons; Clutton-Brock et al. 1992). We removed these known effects prior to estimation of genetic parameters, by fitting linear models of size at each measurement age in which sex and natal litter size were fitted as explanatory variables. Residuals from these models were then used as the measures of ''corrected weight'' (WT_{AGE}) and ''corrected hindleg length'' (HL_{AGE}) at each age. For weight at birth (i.e. AGE = 0) the linear model included the additional explanatory term of capture age (in days) as many individuals are not captured for several days after birth and lambs grow at an average of 0.135 kg/day. Note that there is no available data on leg length at birth.

In order to avoid sample size issues that arise from a lack of records on older animals, we restricted our attention to size traits expressed from birth (AGE = 0 - months) until August in the fifth year of life (AGE = 64 months). Although Soay sheep can sometimes live for 10 years or more, 80% of weight records and 85% of hindleg records relate to this period of ontogeny. Thus we defined corrected size traits at ages 0 months (i.e., at birth which is typically in April) and at 4, 16, 28, 40, 52 and 64 months (i.e. successive August weights up to 5 years old). Analyses of weight are based on 6,871 measurements made on 3,749 distinct individuals between 1985 and 2005. The number of phenotypically informative individuals is 3,054, 1,670, 760, 515, 327, 297 and 248 for WT_0, WT_4, WT_{16}, WT_{28}, WT_{40}, WT_{52} and WT_{64} respectively. Analyses of hindleg length are based on 3,453 measurements made on 1,985 distinct individuals, with the number of phenotypically informative individuals being 1,536, 671, 464, 291, 272 and 219 for HL_4, HL_{16}, HL_{28}, HL_{40}, HL_{52} and HL_{64} respectively.

Quantitative genetic analyses also require pedigree information which has been determined through field observations of maternity and molecular paternity assignment. Paternity assignment was performed using microsatellite data and the maximum-likelihood method CERVUS (Marshall et al. 1998). Paternities were assigned for individuals born before and after 1997 using overlapping panels of 14 and 18 microsatellite loci respectively, and accepted if assigned at a pedigree-wide

confidence level of 80% with a maximum of one locus showing an allelic incompatibility between offspring and putative sire. Complete details of microsatellite protocols and paternity assignment methods are presented elsewhere (Overall et al. 2005). The pedigree structure determined in this way contains 6342 individual records with 3,541 maternal links, and 1,615 paternal links (from 807 distinct dams and 495 distinct sires respectively), and has a maximum depth of 10 generations.

Quantitative genetic analysis

Phenotypic variance for weight and hindleg length was partitioned into genetic and environmental components using animal models (Kruuk 2004). Two modelling approaches were employed for each measure of size: firstly, we estimated variance components under a conventional repeated measures animal model in which all variance components are assumed to be constant across ontogeny; and secondly, we used random regression to fit a series of less constrained animal models in which variance components can vary across ontogeny. Fixed effects were included in all animal models to account for increasing average size with age (i.e., growth) and phenotypic differences caused by variation in environmental conditions among different years of the study. Thus age (in months) and birth year were fitted as factors, and the interaction between these terms was also included. Birth year was fitted as 30 level factor and its inclusion as a fixed effect also accounts for any temporal trend in mean phenotype, reducing the possibility of spurious trends in breeding values resulting from environmentally-induced phenotypic trends (Postma 2006). In all cases parameter estimates were solved for using restricted maximum likelihood (REML) implemented in the program ASReml (VSN International Ltd 2002). The two modelling approaches are presented more fully below. Parallel analyses were performed for corrected weight and hindleg length but for simplicity we show the former only.

Model 1: Repeated measures animal model

Under the repeated measures animal model the corrected weight of individual i with mother j is given as:

$$\mathrm{WT}_{i\mathrm{AGE}} \sim (\mathrm{AGE} + \mathrm{BIRTHYEAR} + \mathrm{AGE} * \mathrm{BIRTHYEAR})_{i\mathrm{AGE}} + a_i + m_j + pe_i + e_i$$

where a_i is the breeding value of individual i which has a population mean of zero and variance of $\sigma^2{}_{\mathrm{A}}$ (the additive genetic variance). Estimating $\sigma^2{}_{\mathrm{A}}$ is possible because the variance–covariance matrix of additive genetic effects is expected to equal to $\mathbf{A}\sigma^2{}_{\mathrm{A}}$ where $\mathbf{A}$ is the additive numerator relationship matrix containing the individual elements $A_{ij} = 2\Theta_{ij}$, and Θ_{ij} is the coefficient of coancestry between individuals i and j obtained from the pedigree structure. Permanent environment effect pe_i, maternal performance m_j, and residual error e_i terms also contribute to the phenotype. These terms were assumed to be normally distributed with means of zero and variance–covariance matrices of $\mathbf{I}\sigma^2{}_{\mathrm{PE}}$, $\mathbf{I}\sigma^2{}_{\mathrm{M}}$ and $\mathbf{I}\sigma^2{}_{\mathrm{R}}$, where $\sigma^2{}_{\mathrm{PE}}$, $\sigma^2{}_{\mathrm{M}}$ and $\sigma^2{}_{\mathrm{E}}$ are the permanent environment, maternal, and residual (temporary environmental) variance components respectively, and $\mathbf{I}$ is an identity matrix with order equal to the number of records as appropriate. The phenotypic variance of weight ($\sigma^2{}_{\mathrm{P}}$) was

determined as the sum of its estimated components, and the direct heritability (h^2), permanent environment effect (pe^2), maternal effect (m^2) and ratio of residual variance (r^2) were then calculated as the ratio of the relevant variance component to σ^2_{P}.

Model 2: Random regression animal model

To test for ontogenetic variation in quantitative genetic parameters and estimate genetic correlations between age-specific weights, additive and maternal effects on the phenotype of individual i were then modelled by regressing on orthogonal (Legendre) polynomials of standardized age ($\mathrm{AGE_{SD}}$), defined as age in months standardized to the interval $-1 \leq \mathrm{AGE_{SD}} \leq 1$ (Kirkpatrick et al. 1990; Meyer 1998). Consequently, at the individual level, the corrected weight ($\mathrm{WT_{AGE}}$) of individual i with mother j at any age is given as:

$$\begin{aligned}\mathrm{WT}_{i\mathrm{AGE}} \sim\ & (\mathrm{AGE} + \mathrm{BIRTHYEAR} + \mathrm{AGE} * \mathrm{BIRTHYEAR})_{i\mathrm{AGE}} \\ & + f_1(a_i, n_1, \mathrm{AGE_{SD}}) + f_2(m_j, n_2, \mathrm{AGE_{SD}}) + e_{i\mathrm{AGE}}\end{aligned}$$

where $f_1(a_i,n_1,\ \mathrm{AGE_{SD}})$ is the random regression function, on orthogonal polynomials of $\mathrm{AGE_{SD}}$ with order n_1, of additive genetic merit values of individuals; and $f_2(m_j,n_2,\ \mathrm{AGE_{SD}})$ is a random regression function with order n_2 of maternal performance values of individuals on $\mathrm{AGE_{SD}}$; and $e_{i\mathrm{AGE}}$ is the age-specific residual error for individual i. The latter term was modelled using a 7×7 unstructured matrix to permit a multivariate error structure, with $e_{i\mathrm{AGE}}$ separately estimated at values corresponding to 0, 4, 16, 28, 40, 52 and 64 months. The unstructured matrix allows residual errors to be correlated across ages within individuals (removing the requirement for including a permanent environment effect).

We first fitted a model without additive or maternal effects such that all phenotypic variance is allocated to the residual structure (Model 2.0 in Table 1). The resultant residual matrix is therefore a description of the phenotypic variance–covariance surface for WT where $-1 < \mathrm{AGE_{SD}} < 1$ (i.e., ages 0–64 months). Subsequently, a forward selection procedure was used to compare a series of successively more complex random regression models that differed in the order of polynomial function used to fit additive and maternal effects (see Table 1). Values of n_1 from 0 (a_i as constant) to 3 (a_i as a cubic function of A_{SD}), and n_2 from 0 (m_j as a constant) to 1 (m_j as a linear function of AGE_{SD}) were used. Increasing values of *n1* and *n2* result in an increase in the number of (co)variance components estimated. Consequently, model selection was performed using likelihood-ratio tests to compare pairs of models (Meyer 1992), with the number of degrees of freedom determined as the difference in the number of (co)variance parameters estimated between models.

Following selection of appropriate values of n_1 and n_2, the variance–covariance matrices of random regression parameters obtained for the additive genetic effect (matrix **Q** with dimensions $(n_1 + 1) \times (\ (n_1 + 1)$ was used to derive age-specific genetic parameters. Specifically the additive genetic variance–covariance matrix, **G**, for $\mathrm{WT_{AGE}}$ (for ages 0–64 months) was obtained as **G = Z Q Z',** where **Z** is the vector of orthogonal polynomials evaluated at values of standardized age that correspond to 0, 4, 16, 28, 40, 52 and 64 months (and **Z'** is the transpose of **Z**). An

Table 1 Random regression model selection for weight and hindleg length showing: order of polynomials used for additive and maternal effects, number of (co)variance parameters estimated, and associated log-likelihoods scores (LnLK) for each model

Model	Polynomial order		WEIGHT		HINDLEG	
	Additive n_1	Maternal n_2	Parameters	LnLK	Parameters	LnLK
2.0	NF	NF	28	–4644.54	21	–8258.19
2.1	0	NF	29	–4608.14	22	–8221.23
2.2	0	0	30	–4556.34	**23**	**–8215.58***
2.3	0	1	32	–4555.03	25	–8215.17
2.4	1	NF	31	–4590.24	24	–8221.12
2.5	1	0	32	–4538.09	25	–8215.28
2.6	1	1	34	–4536.72	27	–8213.46
2.7	2	NF	34	–4580.87	27	–8218.57
2.8	2	0	35	–4528.42	28	–8212.84
2.9	2	1	37	–4527.26	30	–8209.44
2.10	3	NF	38	–4575.3	31	–8216.92
2.11	3	0	**39**	**–4522.55***	32	–8211.11
2.12	3	1	41	–4521.17	34	-

NF indicates that an effect was not fitted. Note that convergence of Model 2.12 was not achieved for hindleg length

*Statistically best model

analogous procedure was used to obtain the maternal genetic variance–covariance matrix **M,** and approximate standard errors for the elements of **G** and **M** matrices were determined (according to Fischer et al. 2004). The multivariate residual error structure obtained by solving the random regression model represents the environmental variance–covariance matrix **R**.

At each age of measurement, the phenotypic variance was then determined as the sum of age-specific variance components, and the ratios h^2, m^2 and r^2 were calculated. We also calculated phenotypic (CV_P), additive (CV_A), maternal genetic (CV_M) and residual (CV_R) coefficients of variation [where the coefficient of variation is found as $100 \times (\text{variance}^{0.5})/\text{sample mean}$]. By comparison to the raw variances, coefficients of variation are expected to be less sensitive to scale effects arising from an increasing mean weight with age (Houle 1992). Finally, between each pair of ages, covariance components were rescaled to give genetic (r_G) and maternal (r_M) correlations, and the variance–covariance matrices **G** and **M** were also subjected to eigenvector decomposition in order to summarise the major patterns of variation for individual growth trajectories (following e.g., Cheverud et al. 1983b).

Determination of phenotypic and genetic temporal trends

To test for temporal trends at the phenotypic level, linear regressions on birth year were performed for annual means of each age-specific corrected weight (WT_{AGE}) and hindleg length (HL_{AGE}) with the degrees of freedom determined from the number of years represented in the phenotypic data. Trends at the genetic level were similarly tested for using annual means of the predicted breeding values at each age ($\overline{PBV_{AGE}}$) as the dependent variable, and the number of degrees of freedom determined from the number of birth years represented in the full pedigree structure, which includes animals some born prior to 1985 (N = 30 years).

For each individual, PBV was determined from the animal model as the best linear unbiased predictor (BLUPs) of the breeding value (a_i). The repeated measures animal model yields a single PBV per individual (since a_i is necessarily constant over ontogeny). In contrast, under the random regression model $a_{i\mathrm{AGE}}$, the breeding value of individual i at a given age, is specified by a polynomial function of $\mathrm{AGE_{SD}}$ with order n_1. Thus BLUPs of $n_1 + 1$ coefficients were obtained for each individual, and by evaluating the polynomial we obtained PBVs for each animal at ages 0, 4, 16, 28, 40, 52 and 64 months. Equivalent analyses were then performed to test for temporal trends in maternal performance for offspring size. Thus, mean predicted maternal performance ($\overline{\mathrm{PMV}}$) was regressed on birth year with individual PMVs determined from the BLUPs of the $n_2 + 1$ coefficients in the polynomial function for m_{jA}, the maternal performance at age A.

Results

Genetic and environmental covariance structures for weight

Animal model analyses supported the presence of additive and maternal effects on corrected body weight. Under the repeated measures approach (Model 1), variance components (± standard errors) were estimated as $\sigma^2_{\mathrm{A}} = 0.363 \pm 0.077$, $\sigma^2_{\mathrm{M}} = 0.080 \pm 0.038$, $\sigma^2_{\mathrm{PE}} = 0.764 \pm 0.083$ and $\sigma^2_{\mathrm{R}} = 2.552 \pm 0.059$. The ratios of additive and maternal variance to σ^2_{P} were calculated as $h^2 = 0.096 \pm 0.020$ and $m^2 = 0.021 \pm 0.010$ respectively, while most of the phenotypic variance was attributed to permanent and temporary environmental effects ($pe^2 = 0.203 \pm 0.022$, $r^2 = 0.679 \pm 0.015$). Note that while the maternal effect accounted for only 2% of the phenotypic variance, comparison to a reduced model showed this to be statistically significant ($\chi^2_1 = 3.37$, $P = 0.009$).

The random regression models also demonstrated additive and maternal effects on weight. Furthermore, using this approach provided important evidence that variance components and associated quantitative genetic parameters change across ontogeny, a phenomenon an evolutionary importance that cannot be detected using the conventional repeated measures model (Model 1). Based on log-likelihood tests, Model 2.11 was selected as the best model (Table 1), performing significantly better than models 2.0–2.10, while Model 2.12 was not a significantly better fit ($\chi^2_2 = 1.38$, $P = 0.251$). Thus age-specific genetic parameters were estimated using Model 2.11 in which the breeding value was modelled with a third order polynomial regression on standardized age ($n_1 = 3$), and the maternal performance was modelled using a zero order polynomial ($n_2 = 0$).

Under Model 2.11, the additive genetic variance σ^2_{A} increases with age (Table 2, Fig. 1), while the residual (environmental) variance also rises from 0 to 4 months before showing relative stability. Maternal variance is constrained to be constant by the choice of $n_2 = 0$. However, since variance is expected to increase over ontogeny as a scale-dependent consequence of increasing phenotypic mean, the coefficients of variation are more appropriate for comparing among measurement ages. The coefficient of phenotypic variation ($\mathrm{CV_P}$) shows a general pattern of decline from age 0 to 64 months, caused by reduction in maternal and residual environmental variation (measured by $\mathrm{CV_M}$ and $\mathrm{CV_R}$ respectively; Table 2). It is these changes, rather than increases in $\mathrm{CV_A}$, that cause heritability of weight to increase from

Table 2 Estimated phenotypic means, variance components, coefficients of variation and ratios to phenotypic variance for age-specific size traits. Estimates are based on Model 2.11 for weight and Model 2.2 for hindleg, and approximate standard errors are shown in parentheses where available

	Age	Mean	σ^2_A	σ^2_M	σ^2_R	CV_P	CV_A	CV_M	CV_R	h^2	m^2	r^2
Weight (kg)	0	2.13	0.025 (1.263)	0.101 (0.013)	0.231 (0.011)	28.0	7.36	14.94	22.6	0.069	0.284	0.647
	4	13.3	0.284 (0.566)	0.101 (0.013)	5.59 (0.211)	18.4	4.00	2.39	17.8	0.047	0.017	0.936
	16	19.5	2.07 (0.113)	0.101 (0.013)	6.43 (0.531)	15.1	7.38	1.63	13.0	0.240	0.012	0.748
	28	23.9	2.99 (0.321)	0.101 (0.013)	5.66 (0.592)	12.4	7.24	1.33	9.96	0.342	0.012	0.647
	40	24.8	3.02 (0.534)	0.101 (0.013)	4.34 (0.595)	11.0	7.00	1.28	8.39	0.405	0.014	0.581
	52	25.5	3.14 (0.719)	0.101 (0.013)	5.51 (0.707)	11.6	6.94	1.25	9.20	0.359	0.012	0.630
	64	25.7	4.30 (1.202)	0.101 (0.013)	4.39 (0.916)	11.6	8.08	1.24	8.16	0.489	0.012	0.499
Hindleg (mm)	4	160	28.4 (5.14)	6.83 (2.44)	65.3 (3.86)	6.28	3.34	1.64	5.06	0.283 (0.045)	0.068 (0.024)	0.649 (0.045)
	16	177	28.4 (5.14)	6.83 (2.44)	41.1 (3.69)	4.93	3.01	1.48	3.62	0.372 (0.057)	0.090 (0.030)	0.538 (0.056)
	28	182	28.4 (5.14)	6.83 (2.44)	66.3 (5.77)	5.54	2.93	1.44	4.47	0.280 (0.046)	0.067 (0.024)	0.653 (0.046)
	40	181	28.4 (5.14)	6.83 (2.44)	30.3 (3.96)	4.46	2.94	1.44	3.03	0.434 (0.065)	0.104 (0.036)	0.462 (0.062)
	52	182	28.4 (5.14)	6.83 (2.44)	29.7 (4.08)	4.43	2.93	1.44	2.99	0.437 (0.066)	0.105 (0.036)	0.457 (0.063)
	64	181	28.4 (5.14)	6.83 (2.44)	28.7 (4.11)	4.42	2.94	1.44	2.96	0.444 (0.066)	0.107 (0.037)	0.449 (0.069)

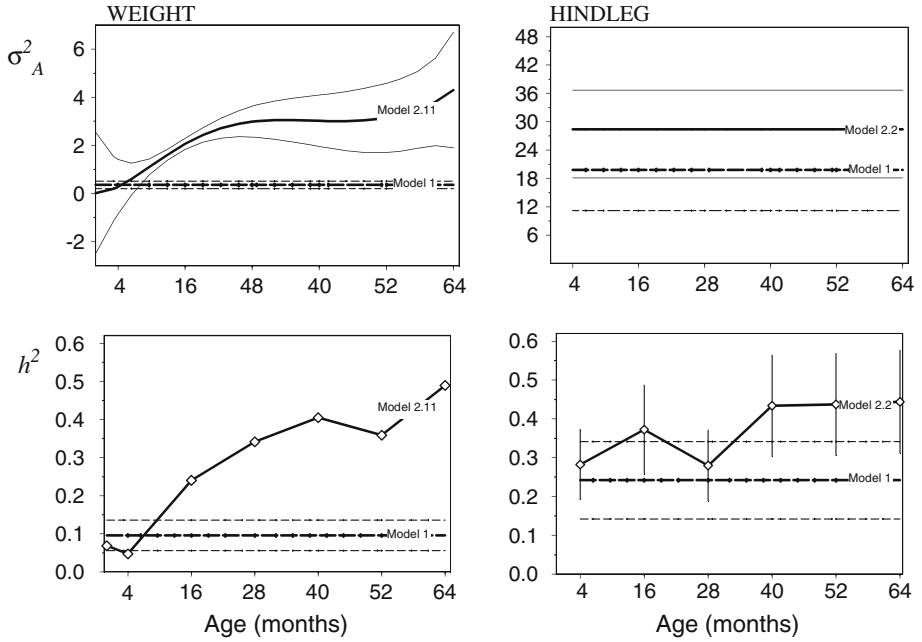

Fig. 1 Estimated σ^2_A (upper panels) and h^2 (lower panels) under Model 1 (dashed line), and best fitting models for weight (Models 2.11) and hindleg (Model 2.2) (solid lines). 95% confidence intervals as thinner lines where available

$h^2 = 0.069$ at birth to $h^2 = 0.489$ for WT_{64} (Fig. 1). As a proportion of phenotypic variance, the maternal effect m^2 declines from 0.284 to 0.012 over the same period (Table 2).

Additive genetic covariances and correlations were positive between all age-specific traits (Table 3). Residual (environmental) and total phenotypic covariances between ages were also uniformly positive (results not shown), while the maternal correlation is constrained to $r_M = +1$ between any pair of ages under Model 2.11. Among traits expressed from age 4 months, genetic correlations were effectively equal to +1 (Table 3). Note that the model 2.11 was unconstrained such that some estimates of r_G lie just outside the strictly permissible parameter space of $-1 \le r_G \le +1$). In contrast,

Table 3 Estimated additive genetic covariances (below diagonal) and correlations (above diagonal) between age-specific weights estimated under Model 2.11. Approximate standard errors are shown in parentheses where available

	Age (months)						
	0	4	16	28	40	52	64
0	-	0.327	0.202	0.259	0.332	0.344	0.227
4	0.027 (0.375)	-	0.992	0.998	0.994	0.981	0.99
16	0.045 (0.387)	0.759 (0.264)	-	0.998	0.986	0.975	1.009
28	0.07 (0.629)	0.919 (0.505)	2.48 (0.389)	-	0.995	0.991	1.025
40	0.091 (0.706)	0.921 (0.582)	2.465 (0.478)	2.992 (0.494)	-	1.001	1.027
52	0.095 (0.744)	0.926 (0.551)	2.485 (0.465)	3.037 (0.497)	3.087 (0.611)	-	1.009
64	0.074 (0.913)	1.094 (0.66)	3.006 (0.513)	3.673 (0.528)	3.703 (0.622)	3.709 (0.786)	-

Table 4 Loading coefficients for the first two eigenvectors of the **G** matrix of age-specific weight (WT_{AGE}) as estimated under Model 2.11

Eigenvector	% variation explained	Loadings on age-specific weight Age						
		0	4	16	28	40	52	64
1	99.95	–0.008	–0.131	–0.359	–0.435	–0.438	–0.442	–0.526
2	0.04	–0.102	–0.077	+0.126	+0.362	+0.423	+0.104	–0.804

genetic correlations between birth weight (WT_0) and traits expressed later were substantially lower (ranging from 0.202 to 0.343; Table 3).

Decomposition of the estimated genetic variance–covariance **G** for age-specific weights showed that over 99.9 % of the variation in the data was explained by the first eigenvector (Table 4). For age-specific weight traits, the loading coefficients associated with this vector were uniform in sign with a general trend of increasing magnitude with age consistent with the age-related trend in σ^2_A (Table 4). The variation described by this eigenvector therefore corresponds to additive effects that are highly integrated across ontogeny, with alleles having a uniformly positive or negative influence on weight at all ages. Note since all estimated elements of matrix **M** are equal, no eigenvector decomposition was performed.

Genetic and environmental covariance structures for hindleg length

Additive and maternal effects were also found to be significant for hindleg length. Under Model 1, variance components (± standard errors) were estimated as $\sigma^2_A = 19.8 \pm 4.30$, $\sigma^2_M = 7.41 \pm 2.18$, $\sigma^2_{PE} = 33.7 \pm 3.67$ and $\sigma^2_R = 20.9 \pm 0.806$, with corresponding ratios σ^2_P of; $h^2 = 0.242 \pm 0.050$, $m^2 = 0.091 \pm 0.026$, $pe^2 = 0.412 \pm 0.046$, and $r^2 = 0.256 \pm 0.013$). In contrast to weight, random regression models did not provide evidence for changing levels of additive variance over ontogeny, and the best model was found to be Model 2.2 in which both σ^2_A and σ^2_M are constrained to be constant (Table 2). However, Model 2.2 uses a multivariate error structure, allowing σ^2_R to differ between measurement ages which provided a significantly better fit to the data than the univariate error structure of Model 1 (Likelihood ratio test, $\chi^2_{19} = 528$, $P < 0.001$). Consequently, while σ^2_A is constant, we found a general (though imperfect) trend of increasing heritability of hindleg length from $h^2 = 0.283$ for HL_4 to $h^2 = 0.444$ for HL_{64} (Table 2, Figure 1). This trend is driven by a decline in residual variation as indicated by CV_R (Table 2). Note that under Model 2.2, genetic correlations (r_G) are by definition constrained to +1 between all pairs of measurement ages (as are the maternal correlations). Eigenvector decompositions of **G** and **M** matrix were not performed since estimated matrices contain no variation. Residual and total phenotypic covariances (and correlations) between ages were again positive (results not shown).

Phenotypic and genetic temporal trends for body size

Regressions of mean age specific size traits on birth year were all negative, and significant (or marginally non-significant) in 11 of 13 cases (Table 5). Thus, as measured by both weight and hindleg length, sheep are actually getting smaller at all ages (Table 5, Fig. 2). Rates of phenotypic decline are substantial in some cases with a maximum of –0.221 kg $year^{-1}$ for weight (at 16 months), and –0.635 mm $year^{-1}$ for

Table 5 Temporal trends in mean age–specific weight and hindleg traits estimated from linear models with birth year fitted as an explanatory covariate. Also shown are equivalent models of mean predicted breeding values ($\overline{\text{PBV}}$) and mean predicted maternal performance ($\overline{\text{PMV}}$) as determined under Models 2.11 and 2.2 for weight and hindleg length respectively. Note age-specific predicted breeding values were not obtained for hindleg because the best fit model suggested a constant additive genetic variance with age. Significance of all temporal trends assessed from least-squares regression with residual degrees of freedom as indicated

	Age	Weight			Hindleg		
		Trend (kg/year)	DF	*P*	Trend (mm/year)	DF	*P*
Mean phenotype	0	–0.016 (0.011)	19	0.150	-		-
	4	–0.129 (0.036)	19	0.002	–0.431 (0.054)	16	0.002
	16	–0.221 (0.046)	19	<0.001	–0.412 (0.059)	16	<0.001
	28	–0.113 (0.054)	19	0.051	–0.281 (0.082)	16	0.038
	40	–0.092 (0.066)	19	0.179	–0.470 (0.097)	16	<0.001
	52	–0.180 (0.050)	18	0.002	–0.182 (0.103)	15	0.066
	64	–0.146 (0.049)	17	0.008	–0.505 (0.106)	15	0.002
$\overline{\text{PBV}}$	0	0.001 (2×10^{-3})	29	<0.001	-		-
	4	0.003 (4×10^{-3})	29	<0.001	-		-
	16	0.006 (0.001)	29	<0.001	-		-
	28	0.008 (0.001)	29	<0.001	-		-
	40	0.008 (0.001)	29	<0.001	-		-
	52	0.009 (0.001)	29	<0.001	-		-
	64	0.010 (0.001)	29	<0.001	-		-
	Pooled	-	29	-	0.011 (0.005)	29	0.004
$\overline{\text{PMV}}$	Pooled	–0.005 (0.002)	29	0.012	–0.014 (0.006)	29	0.002

hindleg length (at 16 months). In contrast, at the genetic level, regressions of mean predicted breeding values (as determined by the best fitting models) showed significant increases in genetic merit for size with time (Table 5, Fig. 2). Note that age specific values of $\overline{\text{PBV}_{\text{AGE}}}$ were obtained for weight (Table 5) but not hindleg, since using the zero order polynomial to model the additive covariance function ($n_1 = 0$ in Model 2.2) constrains a_i to be constant over ontogeny. Similarly, since $n_2 = 0$ in both Models 2.11 and 2.2 such age-specific analyses of maternal performance are not applicable. In contrast to breeding values, the average maternal performance for size was actually found to be declining over time. Significant declines in ($\overline{\text{PMV}}$) were determined for weight (–0.005 kg year^{-1}, $P = 0.012$) and hindleg length (–0.014 mm year^{-1}, $P = 0.002$) with maternal performance determined under Models 2.11 and 2.2 respectively.

Note that the above trends in breeding values and maternal performance for weight and hindleg length were also found with analyses conducted under the more conventional repeated measures model (Model 1). For comparison, under Model 1 the temporal trends were: for weight, $\overline{\text{PBV}}$ trend = 0.003 kg year^{-1} ($P < 0.001$), $\overline{\text{PMV}}$ trend = –0.002 kg year^{-1} ($P < 0.001$); for hindleg length, $\overline{\text{PBV}}$ trend = 0.009 mm year^{-1} ($P = 0.046$), $\overline{\text{PMV}}$ trend = –0.020 mm year^{-1} ($P = 0.017$).

Discussion

Our results confirm that body size, as measured by both weight and hindleg length, is heritable across ontogeny in Soay sheep. Furthermore, our estimation of the **G** matrices revealed strong positive genetic correlations among ages, with highly

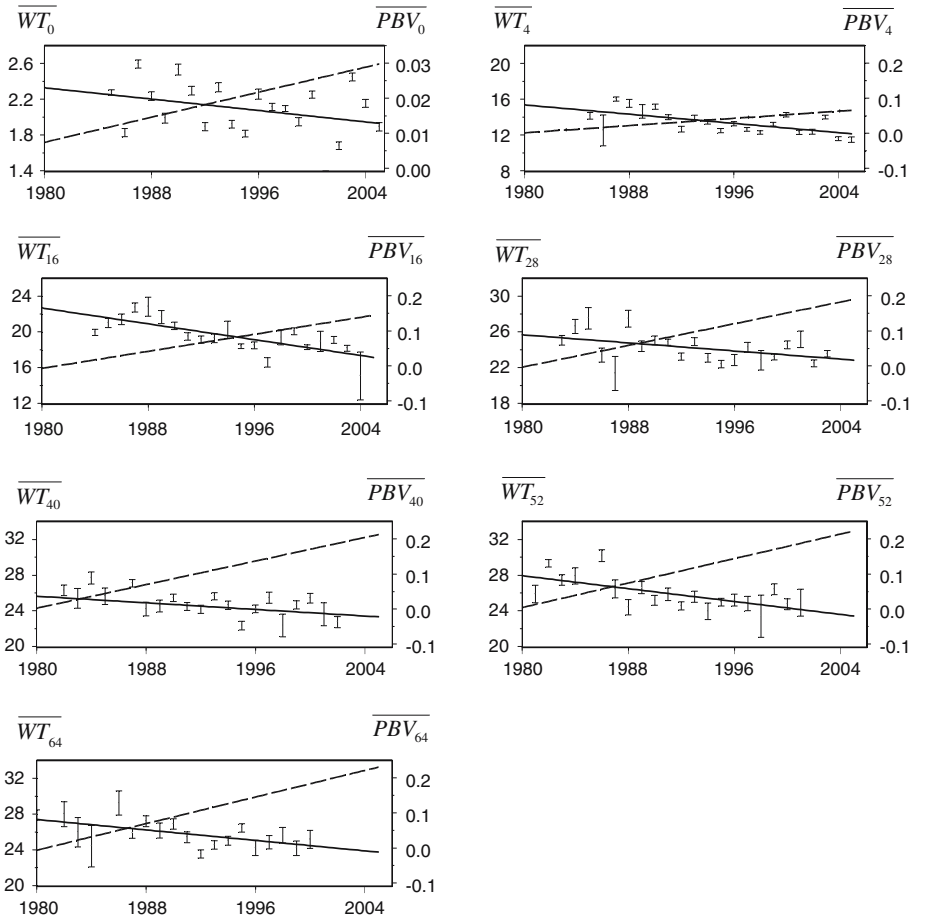

Fig. 2 Temporal trends in phenotype and breeding value for weight. For each age, bars indicate phenotypic mean by birth year (±SE). Lines show best fit from least-squares regression of mean phenotype (solid line), and mean predicted breeding value ($\overline{\mathrm{PBV}}_{\mathrm{AGE}}$; dotted line) on birth year. Note that Y-axis scaling at 0 months differs from other plots for clarity

integrated ontogenies for both aspects of body size. Thus we found no support for the hypothesis that a response to positive directional selection on size is limited by genetic constraints between traits across ontogeny. Nevertheless, our analyses did reveal important ontogenetic patterns in the quantitative genetic parameters that will influence the evolution of body size. In contrast, our alternate hypothesis, namely that selection is eliciting a genetic response but that this is masked at the phenotypic level by opposing environmental effects, is supported for both weight and hindleg length. In the following discussion we first consider the quantitative genetic architecture of the body size ontogeny before focusing on the evidence for cryptic evolution in this population.

Quantitative genetic architecture of body size

Genetic covariance structures for weight and hindleg length were found to be broadly similar in a qualitative sense, with all analyses providing evidence of additive

and maternal effects. Furthermore, for both measures of size the random regression models showed a trend towards increasing heritability with age. By contrast, the simpler models in which variance components are constrained to be constant provided inferior fits to the data, and also yielded h^2 estimates heavily determined by the early-measured phenotypes (i.e., at 0 and 4 months) that dominate the data set. Increasing h^2 with age was particularly marked for weight where a rapidly diminishing maternal influence on offspring phenotype was also found. Although the declining maternal effect reflects a decrease in absolute levels of maternal variance, the trend in heritability is actually driven by declines in non-genetic (i.e., maternal and environmental) sources of variance rather than by increasing additive variation. Similarly, the increase in h^2 of hindleg length is a consequence of diminishing environmental variation rather than an increase in additive effects.

While qualitatively similar, some quantitative differences are apparent between the genetic architecture of weight and hindleg length. For example, although a lack of hindleg data at birth limits comparison, heritability is noticeably lower for weight than hindleg length in lambs. This is likely because weight is more sensitive to short-term changes in the environment than skeletal traits (e.g. weight loss occurs in response to temporary food shortage while bone resorption does not).

Overall phenotypic variances for both weight and hindleg length were found to decline with increasing age, an observation consistent with the action of directional viability selection which occurs in particular over the first part of ontogeny (i.e. between measurement ages 0 and 16 months; (Overall et al. 2005; Wilson et al. 2005c). However, in addition to viability selection, compensatory growth processes may cause a reduction in variance with age and have commonly been reported in domestic sheep (Wilson and Réale 2006). Although comparative studies are limited, the ontogenetic patterns in quantitative genetic parameters found here are strikingly similar to those reported for live weight in wild bighorn sheep, *Ovis canadensis* (Wilson et al. 2005b), suggesting that they may be quite generalisable.

It should be noted that the viability selection discussed above might have the potential to introduce bias into the estimates of covariance components. While animal models are generally thought to be robust to selection through differential reproductive success (Lynch and Walsh 1998), covariances between ages will depend on age-specific phenotypic distributions that later in ontogeny might represent post-selection distributions. Intuitively, since a continuous covariance function across ontogeny is estimated (i.e., using phenotypic data measured at all ages), random regression models may be less susceptible to bias than a conventional two trait animal model. However, further investigation of these issues is certainly warranted.

A general maternal effect was modelled in our analyses such that the maternal variance may potentially include both environmental and genetic components. Maternal genetic effects arise from allelic differences between individual mothers at loci influencing offspring phenotype, and are themselves a heritable component of phenotypic variance (Wolf et al. 1998). Here, the separation of the maternal variance into genetic and environmental components was not statistically supported for either size trait modelled across ontogeny (results not shown). However maternal genetic effects on birth weight have previously been demonstrated (Wilson et al. 2005a), and failure to partition them here is likely due to statistical considerations arising from the different definition of phenotype (i.e., weight traits across ontogeny rather than at the single stage when maternal genetic effects are most important).

This conclusion is supported by univariate analysis of WT_0 that confirm previous findings using the slightly extended data herein (results not shown).

Since maternal genetic effects are heritable, the ''total heritability'' (incorporating both additive genetic and maternal genetic variance; (Willham 1972) of weight may be somewhat higher than the estimates of direct h^2 reported, particularly for WT_0 where the general maternal effect accounted for 28% of the variance. Nevertheless, maternal genetic effects will not change the general pattern of increasing heritability with age, which would persist even under the extreme (and highly unlikely) scenario of all maternal variance being attributable to maternal genetic effects. All else being equal, selection on size later in life will therefore elicit a more rapid evolutionary response due to higher heritabilities of both weight and hindleg length. However, to offset this prediction, selection is generally stronger on lamb and yearling size than on adult traits (Milner et al. 1999).

Strong positive genetic correlations between age-specific traits (fixed at +1 for hindleg length) are such that that selection at any age should result in positively correlated responses across all ages. A partial caveat to this is that genetic correlations were lower between birth weight and later traits, suggesting genetic influences on WT_0 may differ somewhat from those determining later size and growth. Although the magnitude of correlated responses by later-expressed traits to selection acting on birth weight will thus be reduced accordingly, positive selection should increase the height of the population mean growth curve as a whole.

It is clear from our analyses that the ontogeny of Soay sheep body size is highly genetically integrated (*sensu* Cheverud et al. 1983b). This conclusion is supported by both the estimates of r_G among ages, and also the eigenvector decomposition of the **G** matrix for age-specific traits. For weight, virtually all variation is explained by a single eigenvector corresponding to genetic effects of uniform sign across ontogeny (note that the first eigenvector necessarily explains all variation in **G** for hindleg length and **M** for both size measures as a consequence of model selection). Importantly, this lack of variation associated with additional eigenvectors can be seen as a genetic constraint on the growth curve per se (Kirkpatrick and Lofsvold 1992; Blows and Hoffman 2005). Thus, while our results do not provide evidence of a constraint to the evolution of larger size, they do indicate that evolutionary change will be restricted to an increase (or decrease under negative selection) in the height of the mean growth curve. In contrast, other laboratory and field studies have found that additional eigenvectors can explain substantial genetic variance (e.g., Cheverud et al. 1983b; Ragland and Carter 2004; Wilson et al. 2005b) that may reflect segregating alleles with antagonistic effects at different ages . Under such conditions then the population growth curve might evolve in different directions (e.g., increased size early in life and decreased size later) under an appropriate selection regime.

It should be noted that we restricted our attention to the analysis of size traits only. Though weight and hindleg length, the two aspects of size considered here, are themselves positively genetically correlated (Milner et al. 2000; Coltman et al. 2001), we cannot preclude the possibility of constraints arising from correlations between size and other aspects of phenotype not considered. Furthermore, while accurately estimating the **G** matrix for a wider set of (non-size) phenotypic traits is not trivial in this (or any other) natural system, eigenvector analyses of such a matrix may be useful (Blows and Hoffman 2005).

Cryptic evolution of body size in Soay sheep

Despite selection for increased size and the apparent absence of genetic constraints (discussed above), our analyses of age-specific weight and hindleg lengths confirmed that the sheep are not getting phenotypically larger. Furthermore, in contrast to the previous assertion that there is no temporal trend in weight (Coltman et al. 1999; Milner et al. 1999), we found evidence that sheep are getting significantly lighter at all measurement ages from 0 to 64 months. Similar declines in hindleg length confirm that body size is decreasing in this system. While the difference from previous findings can be largely attributed to the extended data set (including 9 additional years used here), our results are particularly striking in that phenotype is actually changing in the opposite way to that predicted by quantitative genetic theory.

In contrast, individual predicted breeding values for weight and hindleg lengths (at all ages) have increased significantly over the time period of the study, providing evidence of a genetic response to positive directional selection on size. Trends in predicted breeding values should generally be interpreted with some caution since PBVs may not be completely independent of the environmental component of an individual's phenotype. In the absence of sufficient data, breeding values may tend to reflect environmentally driven trends in the phenotypic data (Postma 2006). In this case increased confidence in our interpretation stems from two sources. Firstly, we included birth year in the fixed effects of all animal models, thus de-trending the phenotype in models used for breeding value prediction. Secondly, given the potential for environmental bias of PBVs, it has been argued that the safest conclusions will be those based on contrasting patterns of phenotype and PBV, exactly as found here (Postma 2006).

By comparison to changes at the phenotypic level, rates of change in predicted breeding values are modest. In principle, the rate of breeding value increase should be predictable from the genetic covariance structures and the strength of selection acting (Falconer and Mackay 1996). However, several critical assumptions underlying simple evolutionary models (e.g. discrete generations, constant environment) are violated here. For example, St Kilda is characterised by extensive temporal heterogeneity in environmental quality that is known to affect evolutionary parameters (Wilson et al. 2006) as well as population demography (Coulson et al. 2001). Consequently, generating quantitative predictions of phenotypic and genetic change will require more sophisticated approaches to measuring selection (Coulson et al. 2006) and incorporating the effects of environmental heterogeneity on both selection and genetic architecture (Wilson et al. 2006).

Furthermore, while de-trending the phenotypic data increases our confidence that the detected trends in PBVs are real, it also means that the estimated rates of genetic change will be inherently conservative (Postma 2006). Temporal environmental heterogeneity is such that close relatives may often experience more similar conditions complicating the statistical separation of environmental and genetic effects. Thus fitting birth year in the animal models could reduce the signature of genetic change over time, as well as accounting for environmental heterogeneity as intended. For these reasons we focus on the qualitative finding that trends at phenotypic and genetic levels have contrasting signs for both body weight and hind leg length.

Our analyses of predicted breeding values indicate that there are actually genetic responses to this positive selection consistent with the apparent absence of genetic

constraints. However, these responses are cryptic in the sense that they indicate genetic improvement for body size is occurring despite the contrary phenotypic patterns for both weight and hind leg length. Our results parallel those of studies in collared flycatchers, *Ficedula albicollis* (Merilä et al. 2001a) and great tits, *Parus major* (Garant et al. 2004) in which cryptic evolution of fitness-related size traits has also been reported. The current study therefore adds weight to the possibility that this is a relatively common phenomenon (Garant et al. 2004).

At the phenotypic level, declining body size in the Soay sheep is consistent with a wider empirical pattern of dwarfing in large mammals on small islands (the ''Island Rule'' of Foster (1964)). While such island effects are generally viewed as arising from selective forces (Lomolino 2005), here size is declining despite evidence of positive directional selection (Milner et al. 1999) and is not a consequence of additive genetic effects. To some extent declining maternal performance for offspring weight and hindleg length will contribute to the phenotypic trends. For weight the maternal effect is known to include a genetic component (discussed above), such that declining maternal performance may itself reflect microevolution. For example, birth weight is negatively genetically correlated with litter size (Wilson et al. 2005a), which is subject to positive selection through maternal fitness components (Wilson et al. 2005c). While some genetic change may therefore be involved, declining maternal performance for offspring size may also reflect a plastic response to environmental change. More generally, since changes in maternal performance are insufficient to explain the phenotypic declines, observed trends must also include a plastic response at the individual level (Merilä et al. 2001a; Garant et al. 2004).

The current work does not explicitly test possible environmental drivers of phenotypic change. However, population density is known to have a major influence on the biology of this system (Coulson et al. 2001) and, while fluctuating dramatically, the total population of Hirta shows an overall increase from 1985 to 2005 (slope of 31.8 sheep/year, $P = 0.025$). Repeating the linear regressions of age-specific size traits on birth year with density (in the year of phenotypic measurement) included as a covariate suggests that density-dependent effects do contribute to declining body weight (at all ages except 0 months; results not shown). However, significant temporal declines persisted for all weight traits, while density was not a significant explanatory variable for hindleg length. Consequently, while density-dependent effects on size may be occurring, they would appear to be insufficient to explain the observed phenotypic declines in body size. It therefore seems likely that temporal trends in other environmental determinants of growth are also involved. Phenotypic trends in other vertebrate systems have been linked to changes in climate and food abundance [e.g., Réale et al. (2003); Nussey et al. (2005b)], and further investigation of these parameters might prove useful here.

Conclusions

Our quantitative genetic analyses confirmed that body size is heritable in the Soay sheep population. Although age-specific weights and hindleg lengths do differ in heritabilities and hence in potential rates of response to selection, the genetic covariance structures present will cause positively correlated responses across all ages to directional selection. Consequently there is no support for genetic antagonism between early- and later-expressed traits, and no evidence for genetic

constraints on the evolution of increased body size. Given the genetic (co)variance structures estimated, quantitative genetic theory predicts that the entire growth curve should be shifted upwards at all measurement ages for both weight and hindleg length. Increasing breeding values provide evidence of evolutionary change consistent with this prediction, while declining phenotypes are therefore attributable to environmental rather than genetic effects.

More generally, the finding that the directionality of phenotypic change in this system is determined by environmental, rather than genetic, effects has important implications. It is increasingly becoming apparent that selection and heritability can vary dramatically with environmental conditions (Charmantier and Garant 2005; Wilson et al. 2006) as well as with ontogeny as shown here. This, coupled with the frequent presence of overlapping generations, makes quantitative prediction of evolutionary change a difficult task in natural systems that are characterised by environmental heterogeneity. However, even more fundamentally, our results highlight that simple models (e.g. the breeder's equation; Falconer and Mackay 1996) can be limited in their ability to explain even the direction of phenotypic change. For systems characterised by long-term environmental trends, more consideration of the way in which populations, and individuals, respond to changing conditions is therefore needed (Nussey et al. 2005a, c). It is clear that the development of predictive models for phenotypic change will require an increased understanding of how environments are changing, coupled with a view of phenotypic plasticity and microevolution as concurrent, rather than alternative, mechanisms of change.

Acknowledgements We thank the National Trust for Scotland and Scottish Natural Heritage for permission to work on St. Kilda, the Royal Artillery Range (Hebrides) and QinetiQ and Eurest for logistic support. The long-term data collection on St. Kilda has been supported by the Natural Environment Research Council, the Wellcome Trust, the Biotechnology and Biological Sciences Research Council and the Royal Society, through grants to THCB, B.T. Grenfell, M.J. Crawley, T. Coulson, S. Albon, JMP and LEBK. The work described here was funded by a Leverhulme Trust research project grant to LEBK and D.W. Coltman. LEBK is supported by the Royal Society. We also thank the many previous members of the project (including many volunteers) who have collected field data or have contributed genotyping and paternity inference. T. Coulson and two anonymous referees provided useful comments on a previous version of this manuscript.

References

Arnold SJ (1992) Constraints on phenotypic evolution. Am Nat 140:S85–S107

Badyaev AV, Martin TE (2000) Individual variation in growth trajectories: phenotypic and genetic correlations in ontogeny of the house finch (*Carpodacus mexicanus*). J Evol Biol 13:290–301

Björklund M (1997) Variation in growth in the blue tit (*Parus caeruleus*). J Evol Biol 10:139–155

Blanckenhorn WU (2000) The evolution of body size: what keeps organisms small? Q Rev Biol 75:385–407

Blows MW, Hoffman AA (2005) A reassessment of genetic limits to evolutionary change. Ecology 86:1371–1384

Charmantier A, Garant D (2005) Environmental quality and evolutionary potential: lessons from wild populations. Proc R Soc Lond B Bio 272:1415–1425

Cheverud JM, Leamy LJ, Atchley WR, Rutledge JJ (1983a) Quantitative genetics and the evolution of ontogeny. I. Ontogenetic changes in quantitative genetic variance components in randombred mice. Genet Res 42:65–75

Cheverud JM, Rutledge JJ, Atchley WR (1983b) Quantitative genetics of development—genetic correlations among age-specific trait values and the evolution of ontogeny. Evolution 37:895–905

Clutton-Brock TH, Price OF, Albon SD, Jewell PA (1992) Early development and population fluctuations in Soay sheep. J Anim Ecol 61:381–396
Clutton-Brock TH, Pemberton JM (2004) Soay sheep: Dynamics and selection in an island population. Cambridge University Press, Cambridge, 383pp
Coltman DW, Smith JA, Bancroft DR, Pilkington J, MacColl ADC, Clutton-Brock TH, Pemberton JM (1999) Density-dependent variation in lifetime breeding success and natural and sexual selection in Soay rams. Am Nat 154:730–746
Coltman DW, Pilkington J, Kruuk LE, Wilson K, Pemberton JM (2001) Positive genetic correlation between parasite resistance and body size in a free-living ungulate population. Evolution 55:2116–2125
Coltman DW, O'Donoghue P, Hogg JT, Festa-Bianchet M (2005) Selection and genetic (co)variance in bighorn sheep. Evolution 59:1372–1382
Coulson T, Catchpole EA, Albon SD, Morgan BJT, Pemberton JM, Clutton-Brock TH, Crawley MJ, Grenfell BT (2001) Age, sex, density, winter weather, and population crashes in Soay sheep. Science 292:1528–1531
Coulson T, Benton TG, Lundberg P, Dall SRX, Kendall BE, Gaillard J-M (2006) Estimating individual contributions to population growth: evolutionary fitness in ecological time. Proc R Soc Lond B Bio 273:547–556
Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longman, Essex
Fischer TM, Gilmour AR, Van der Werf JHJ (2004) Computing approximate standard errors for genetic parameters derived from random regression models fitted by average information REML. Gen Sel Evol 36:363–369
Foster J (1964) The evolution of mammals on islands. Nature 202:234–235
Gaillard JM, Festa-Bianchet M, Delorme D, Jorgenson J (2000) Body mass and individual fitness in female ungulates: bigger is not always better. Proc R Soc Lond Ser B-Biol Sci 267:471–477
Garant D, Kruuk LEB, McCleery RH, Sheldon BC (2004) Evolution in a changing environment: a case study with great tit fledging mass. Am Nat 164:E115–E129
Guinness FE, Clutton-Brock TH, Albon SD (1978) Factors affecting calf mortality in red deer (*Cervus elaphus*). J Anim Ecol 47:817–832
Houle D (1992) Comparing evolvability and variability of quantitative traits. Genetics 130:195–204
Kirkpatrick M, Lofsvold D, Bulmer M (1990) Analysis of the inheritance, selection and evolution of growth trajectories. Genetics 124:979–993
Kirkpatrick M, Lofsvold D (1992) Measuring selection and constraint in the evolution of growth. Evolution 46:954–971
Kruuk LEB, Merilä J, Sheldon BC (2001) Phenotypic selection on a heritable size trait revisited. Am Nat 158:557–571
Kruuk LEB, Slate J, Pemberton JM, Brotherstone S, Guinness F, Clutton-Brock T (2002) Antler size in red deer: heritability and selection but no evolution. Evolution 56:1683–1695
Kruuk LEB (2004) Estimating genetic parameters in natural populations using the 'animal model'. Philos T Roy Soc B 359:873–890
Lomolino M (2005) Body size evolution in insular vertebrates: generality of the island rule. J Biogeogr 32:1683–1699
Lynch M, Walsh B (1998) Genetics and analysis of quantitative traits. Sinauer Associates, Inc., Sunderland
Marshall TC, Slate J, Kruuk LE, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655
Merilä J, Kruuk LEB, Sheldon BC (2001a) Cryptic evolution in a wild bird population. Nature 412:76–79
Merilä J, Sheldon BC, Kruuk LEB (2001b) Explaining stasis: microevolutionary studies in natural populations. Genetica 112–113:199–222
Meyer K (1992) Variance components due to direct and maternal effects for growth traits of Australian beef cattle. Livest Prod Sci 31:179–204
Meyer K (1998) Estimating covariance functions for longitudinal data using a random regression model. Gen Sel Evol 30:221–240
Milner JM, Albon SD, Illius AW, Pemberton JM, Clutton-Brock TH (1999) Repeated selection of morphometric traits in the Soay sheep on St Kilda. J Anim Ecol 68:472–488
Milner JM, Pemberton JM, Brotherstone S, Albon SD (2000) Estimating variance components and heritabilities in the wild: a case study using the 'animal model' approach. J Evol Biol 13:804–813
Nussey DH, Clutton-Brock TH, Albon SD, Pemberton JM, Kruuk LEB (2005a) Constraints on plastic responses to climate variation in red deer. Biol Lett 1:457–460

Nussey DH, Clutton-Brock TH, Elston DA, Albon SD, Kruuk LEB (2005b) Phenotypic plasticity in a maternal trait in red deer. J Anim Ecol 74:387–396
Nussey DH, Postma E, Gienapp P, Visser ME (2005c) Selection on heritable phenotypic plasticity in a wild bird population. Science 310:304–306
Overall ADJ, Byrne KA, Pilkington J, Pemberton JM (2005) Heterozygosity, inbreeding and neonatal traits in Soay sheep on St Kilda. Mol Ecol 14:3383–3393
Pakkasmaa S, Merilä J, O'Hara RB (2003) Genetic and maternal effect influences on viability of common frog tadpoles under different environmental conditions. Heredity 91:117–124
Pease CM, Bull JJ (1988) A critique of methods for measuring life history trade-offs. J Evol Biol 1:293–303
Postma E (2006) Implications of the difference between true and predicted breeding values for the study of natural selection and micro-evolution. J Evol Biol 19, Doi: 10.1111/j.1420–9101.2005.01007
Ragland GJ, Carter PA (2004) Genetic covariance structure of growth in the salamander *Ambystoma macrodactylum*. Heredity 92:569–578
Réale D, Festa-Bianchet M, Jorgenson JT (1999) Heritability of body mass varies with age and season in wild bighorn sheep. Heredity 83:526–532
Réale D, McAdam AG, Boutin S, Berteaux D (2003) Genetic and plastic responses of a northern mammal to climate change. Pro R Soc Lond B 270:591–596
Schaeffer LR (2004) Application of random regression models in animal breeding. Livest Prod Sci 86:35–45
Sogard SM (1997) Size-selective mortality in the juvenile stage of teleost fishes: a review. B Mar Sci 60:1129–1157
Willham RL (1972) The role of maternal effects in animal breeding: III. Biometrical aspects of maternal effects in animals. J Anim Sci 35:1288–1293
Wilson AJ, Hutchings JA, Ferguson MM (2003) Selective and genetic constraints on the evolution of body size in a stream-dwelling salmonid fish. J Evol Biol 16:584–594
Wilson AJ, Coltman DW, Pemberton JM, Overall ADJ, Byrne KA, Kruuk LEB (2005a) Maternal genetic effects set the potential for evolution in a free-living vertebrate population. J Evol Biol 18:405–414
Wilson AJ, Kruuk LEB, Coltman DW (2005b) Ontogenetic patterns in heritable variation for body size: using random regression models in a wild ungulate population. Am Nat 166:E177–E192
Wilson AJ, Pilkington JG, Pemberton JM, Coltman DW, Overall ADJ, Byrne KA, Kruuk LEB (2005c) Selection on mothers and offspring: whose phenotype is it and does it matter? Evolution 59:451–463
Wilson AJ, Pemberton JM, Pilkington JG, Coltman DW, Mifsud DV, Clutton-Brock TH, Kruuk LEB (2006) Environmental coupling of selection and heritability limits evolution. PLoS Biol 4(7):e216
Wilson AJ, Réale D (2006) Ontogeny of additive and maternal genetic effects: lessons from domestic mammals. Am Nat 167:E23–E38
Wolf JB, Brodie ED, Cheverud JM, Moore AJ, Wade MJ (1998) Evolutionary consequences of indirect genetic effects. Trnds Ecol Evol 13:64–69