

1 Decomposing phenotypic skew and its effects on the
2 predicted response to strong selection

3 Joel L. Pick^{1,2*}, Hannah E. Lemon¹, Caroline E. Thomson¹ & Jarrod D. Hadfield¹

4 ¹Institute of Evolutionary Biology, University of Edinburgh, Edinburgh,
5 United Kingdom

6 ²Centre of Biodiversity Dynamics, Norwegian University of Science and
7 Technology, Trondheim, Norway

8 * Corresponding Author: joel.l.pick@gmail.com

9 **The major frameworks for predicting evolutionary change assume that a phenotype's**
10 **underlying genetic and environmental components are normally distributed. However,**
11 **the predictions of these frameworks may no longer hold if distributions are skewed.**
12 **Despite this, phenotypic skew has never been decomposed, meaning the fundamental**
13 **assumptions of quantitative genetics remain untested. Here, we demonstrate that**
14 **the substantial phenotypic skew in the body size of juvenile blue tits (*Cyanistes***
15 **caeruleus) is driven by environmental factors. Although skew had little impact on**
16 **our predictions of selection response in this case, our results highlight the impact**
17 **of skew on the estimation of inheritance and selection. Specifically, the non-linear**
18 **parent-offspring regressions induced by skew, alongside selective disappearance, can**
19 **strongly bias estimates of heritability. The ubiquity of skew and strong directional**
20 **selection on juvenile body size implies that heritability is commonly overestimated,**
21 **which may in part explain the discrepancy between predicted and observed trait**
22 **evolution.**

23 Quantitative genetics describes how traits respond to selection in terms of selection and
24 inheritance. Typically we use two equations to describe this, the breeder's equation (¹
25 Chapter 12) and Lande's gradient equation (² Eq 7). The breeder's equation gives the
26 predicted response to selection as the heritability (h^2) multiplied by the selection differential
27 (S), whereas Lande's gradient equation describes the response to selection as the additive
28 genetic variance (V_A) of the trait multiplied by the selection gradient (β). Although these
29 frameworks are generally thought to be interchangeable, they only converge when phenotypes
30 (and their genetic and environmental components) are normally distributed or fitness functions
31 (the relationship between a trait and fitness) are linear (³ Chapter 29). Given that fitness
32 functions are highly unlikely to be linear in practice^{4;5}, any deviation from normality can
33 lead to problems with the application of these equations. Consequently, normality is seen
34 as a fundamental assumption in quantitative genetics⁶⁻⁸, yet to our knowledge has not been
35 directly tested, despite the major consequences it has for how traits are predicted to respond
36 to selection⁹⁻¹⁷.

37 The most natural interpretation of heritability in the context of the breeder's equation is the
38 slope of a *linear* parent-offspring (PO) regression^{12-14;18;19}, whilst S (the covariance between
39 a trait and fitness) describes the *linear* relationship between a phenotype and fitness. The
40 accuracy of the breeder's equation relies heavily on the linearity of both of these functions -
41 if both are non-linear, the residuals from the linear functions may be correlated, creating a
42 'spurious response to selection'¹⁴. The linearity of the parent-offspring relationship breaks
43 down when the amount of skew (asymmetry) differs between genetic and environmental
44 components^{20;21}, with genetic and environmental skew causing curvature in opposite directions
45 (Figure 1). Formally, a distribution is skewed when it has a non-zero third central moment.

46 Whilst the gradient equation is robust to environmental skew, it doesn't correctly describe the
47 response to selection in the presence of genetic skew if the fitness function is non-linear (¹¹ Eq
48 42). Environmental skew, through its contribution to phenotypic skew, can, however, impact
49 the estimation of β when it is approximated using Lande-Arnold regression^{5;17;22}.

50 Although extensions to these two equations have been derived that allow for the non-linearity
51 of the PO-regression¹² and the non-normality of genetic values¹¹, the majority of the work
52 in this area remains theoretical. Non-linearity in PO-regressions has been demonstrated in

53 the lab^{12;23-27} and ad-hoc methods have been used to test for skew at the genetic level^{28;29}.
54 Nevertheless, to our knowledge, no study has 1) relaxed the normality assumptions when
55 making statistical inferences to examine the origin and extent of skew at different levels, and
56 2) explored how observed patterns of natural selection interact with skew to determine how
57 well these two equations predict selection response in the wild.

58 Juvenile body size is under strong, persistent, directional selection across taxa³⁰, yet is known
59 to show little response to this selection³¹. We show that juvenile body size is highly negatively
60 skewed (long tail of small individuals) across bird species, but the origin of this skew is
61 unknown. To determine this, we developed statistical methods to decompose the phenotypic
62 distribution into a set of skew-t distributions, and predict the shape of PO-regression based
63 on the estimated skew. We applied these methods to data from a long-term cross-fostering
64 experiment of a wild bird population. By estimating survival selection acting on juvenile body
65 size, we tested the robustness of the predicted response to selection from the breeder's and
66 gradient equations.

67 **Results**

68 **Prevalence of Phenotypic Skew**

69 Across 27 species of birds, tarsus length (a common measure of structural size) was substantially
70 negatively skewed (long tail of small individuals) in juveniles (coefficient of skew: -1.054
71 [-1.394, -0.686], pMCMC<0.001), but not adults (-0.302 [-0.641, 0.052], pMCMC=0.086),
72 with tarsus length being significantly more skewed in juveniles than adults (difference = -0.752
73 [-1.124, -0.366], pMCMC<0.001; Figure 2).

74 **Decomposing Phenotypic Skew**

75 Using data on four juvenile body size traits (tarsus length, head-bill length, mass and wing
76 length), measured on 15 day old chicks from a long-term cross-fostering experiment on a wild
77 population of blue tits, we decomposed phenotypic skew into genetic, between- and within-nest
78 environmental components. We used a mixed model approach with skew-t distributed random
79 effects which allowed the extent and direction of skew to vary between these levels. There
80 was considerable phenotypic skew in all four traits, with the coefficient of skew ranging from
81 -0.51 to -1.60 (Figure 3). There was little evidence of genetic skew in any trait (Figure 3,
82 Tables S5, S8, S11 and S12 and further discussion in supplementary methods). Phenotypic
83 skew was instead driven by considerable environmental skew at both between- and within-nest
84 levels, with the relative magnitude of this skew varying between traits (Figure 3, Tables S6,
85 S9, S12 and S15).

86 Given the environmental origin of the negative phenotypic skew, we would expect a convex
87 PO-regression for all traits²⁰ (Figure 1c). By deriving a method to compute this non-linear
88 PO-regression (Equation 1), we can show that for all traits the slope in the lower tail of the
89 distributions is close to zero, but becomes steeper with increasing body size (Figure 3).

90 Selection on Juvenile Body Size

91 To quantify selection acting on body size, we estimated the linear and quadratic effects of body
92 size on survival from both day 15 to fledging and fledging to local recruitment in a bivariate
93 probit event-history model. As expected, all traits showed significant positive linear effects of
94 body size on survival at both stages, with survival increasing at larger body sizes (Figure 4,
95 Tables S16-19). Interestingly, all quadratic effects of juvenile size on survival between day 15
96 and fledging were positive, with these effects being suggestive and significant for mass and
97 wing length, respectively (Figure 4, Tables S16-19), indicating an accelerating effect of size
98 on offspring survival. In contrast, negative quadratic effects were typical for survival from
99 fledging to recruitment although this effect was only suggestive in the case of tarsus length
100 (Figure 4, Tables S16-19). The fitness functions over both events were generally concave
101 (Figure 4), which would indicate stabilising selection, but the hypothesis that the optimal
102 trait value lay outside of the observed phenotypic range for any trait could not be rejected
103 (proportion of iterations with an internal optimum: tarsus 0.853; head-bill 0.543; mass 0.757;
104 wing 0.017).

105 Using these fitness functions, we were able to estimate selection gradients (β) for each trait by
106 taking the partial derivative of the individual relative fitness function with respect to the trait
107 and averaging it over the trait's distribution. However, β is more frequently approximated
108 using a Lande-Arnold regression of fitness on a trait²² and phenotypic skew can bias this
109 approximation when the fitness function is not linear or quadratic (as is the case for survival
110 functions)²². To test this, we calculated the expected estimates of β that would be obtained
111 from the Lande-Arnold approach without (β_1) and with (β_2) a quadratic term fitted^{22;32;33},
112 over the posterior distribution of the survival models (Equations 10 and 11). Figure 4 shows
113 that generally there is little meaningful difference between estimates, with the exception of
114 wing length, where there is suggestive evidence that β_1 would underestimate β by nearly 30%
115 (β_1/β : 0.711 [0.532, 0.915], pMCMC=0.012).

116 Predicted Response to Selection

117 In the absence of genetic skew, the correct response to selection is given by Lande's gradient
118 equation ($V_A\beta$), which for these traits gives: tarsus: 0.085mm [0.034, 0.127]; head-bill:
119 0.069mm [0.037, 0.102]; mass: 0.094g [0.052, 0.139]; wing: 0.175mm [0.077, 0.280]. The
120 breeder's equation is equal to the gradient equation when the Lande-Arnold regression without
121 the quadratic term gives good estimates of the selection gradient, irrespective of whether
122 the PO-regression is linear or not (i.e if $\beta_1 = \beta$ then $h^2S = V_A\beta$;³ Chapter 29). Given
123 the similarity between β and β_1 for tarsus, head-bill and mass, the breeder's equation will
124 therefore give accurate predictions of the selection response for these traits. However, it
125 underestimates the response to selection in wing length by nearly 30%, as the proportional
126 change in the predicted response to selection is equal to β_1/β (shown above).

127 Selection Bias and Heritability Estimation

128 The heritability in the breeder's equation is the heritability *before* selection (h_b^2) which can be
129 interpreted as the slope of the PO-regression averaged over all individuals irrespective of their
130 fitness. However, direct estimates of the PO-regression can only be obtained from individuals

131 that survive to become parents and so to some extent measure the heritability *after* selection
 132 (h_a^2 ; note the terms heritability before and after selection are used in a broader sense than
 133 in¹⁴, and capture a different bias; see³ p171 for a clear explanation of Heywood's usage).
 134 Since larger individuals are more likely to survive, and the PO-regression is steeper for these
 135 individuals, direct estimates of the PO-regression are likely to be upwardly biased estimates
 136 of heritability. To demonstrate this, we obtained direct estimates of the PO-regression from
 137 the 182 individuals (118 male and 64 female) that were measured as chicks and survived to
 138 produce offspring that were also measured. Although the estimated linear regression (blue line
 139 in Figure 5) is similar to the predicted non-linear PO-regression (red line in Figure 5) for the
 140 large surviving individuals (the linear and non-linear regressions fit the data equally well for all
 141 traits; tarsus $p = 0.195$, head-bill $p = 0.087$, mass $p = 0.060$ and wing $p = 0.052$), the two
 142 diverge substantially at small body sizes (Figure 5). In order to directly compare h_a^2 and h_b^2 ,
 143 we used the parameters of the quantitative genetic and survival models described above to
 144 calculate h_a^2 as the linear PO-regression weighted by the fitness of the parents (Equation 16)
 145 and h_b^2 as V_A/V_P . For tarsus, head-bill and mass, h_a^2 was substantially and significantly higher
 146 than h_b^2 , with a proportional increase in h_a^2 of over 60% for head-bill and mass (h_a^2/h_b^2 : tarsus
 147 1.223 [1.137, 1.333], pMCMC=0.002; head-bill 1.664 [1.421, 1.951], pMCMC<0.001; mass
 148 1.645 [1.325, 2.046], pMCMC<0.001; wing 1.584 [0.373, 2.551], pMCMC=0.372).

149 Estimates of h_b^2 will only be accurate if they do not suffer from the same selection bias
 150 present in PO-regression. Our experimental cross-fostering design means that the majority
 151 of information used to estimate V_A in our analysis comes from the comparison of siblings
 152 (569 nests have chicks from at least 2 clutches), rather than parents and offspring (182
 153 parent-offspring comparisons). Sibling comparisons are made before selection, and so should
 154 not suffer from the same selection bias as parent-offspring comparisons. However, many
 155 wild bird pedigrees rely largely on information from parent-offspring relationships to estimate
 156 genetic effects - without partial cross-fostering and using social pedigrees (no within-nest
 157 variation in relatedness), sibling comparisons provide little information on genetic effects
 158 because they are confounded with common environment (nest) effects. As both PO-regression
 159 and the animal model assume that the relationship between offspring and parental phenotypes
 160 is linear, animal models relying mainly on the information from parent-offspring comparisons
 161 may also be biased. To test this, we simulated data using the parameters from our quantitative
 162 genetic and selection models for mass, assuming social and genetic monogamy, with and
 163 without skew and with and without partial cross-fostering. As expected, environmental skew
 164 caused h^2 estimated from PO-regressions to be consistently and substantially upwardly biased
 165 by a similar amount as we observed in our data, regardless of cross-fostering (estimated/simulated:
 166 no cross-fostering 1.609; cross-fostering 1.616). Without cross-fostering (information mainly
 167 from parent-offspring comparisons), estimates of V_A , and so heritability, from animal models
 168 were upwardly biased, although less than in the PO-regressions (estimated/simulated: V_A
 169 1.226, h^2 1.228), whereas cross-fostering (information mainly from sibling comparisons) led
 170 to the correct estimation of V_A and h^2 (estimated/simulated: 1.012 and 1.015 respectively;
 171 Table 1).

Discussion

A common assumption in quantitative genetics is that phenotypes, and their underlying genetic and environmental components, are normally distributed. Here we demonstrate that this assumption is commonly violated, and in four morphological traits the observed negative phenotypic skew is driven by environmental, rather than genetic, skew. There was strong directional viability selection acting on all four traits, with non-linear fitness functions. Under these conditions the breeder's equation may give inaccurate predictions for the response to selection, but Lande's gradient equation - which only assumes genetic values are normally distributed - is expected to be accurate¹¹. However, this assumes that the methods used to obtain estimates of β and V_A are robust to deviations from normality. Here we empirically demonstrate that common methods used to estimate both metrics can produce biased estimates in the presence of environmental skew.

Perhaps the most striking result is the apparent absence of genetic skew. Theory shows that directional selection can generate genetic skew, but the direction of the skew differs between models. Under the infinitesimal (Gaussian descendants³⁴) model (assumed in our analyses), directional selection can drive a Gaussian distribution of breeding values to be skewed in the direction of selection through the build up of linkage disequilibrium^{11;35;36}. However, stabilising selection may mitigate this (¹¹ Eq 46) and the breeding value distribution quickly returns to normality if selection ceases (the skew quarters each generation for unlinked loci;³⁵ p149). Finite allele models also generate genetic skew through changes in allele frequency. Under the rare-alleles model, directional selection after a long period of stabilising selection generates skew in the direction of selection^{10;11} but sustained long term directional selection (with new mutations, on average, having effects in the opposite direction) is expected to drive skew in the opposite direction to selection^{37;38}. Given juvenile body size appears to be under sustained positive directional selection³⁰ and gene knockout studies in mice show that loss-of-function mutations reduce size more often than increase it³⁹, we would predict negative genetic skew in our system. However, these models predict that the amount of skew generated through selection should be small, consistent with our finding of no or negligible genetic skew. Other processes, such as few loci, alleles of large effect, extreme allele frequencies or substantial non-additive gene action, particularly directional dominance, could generate greater levels of skew^{21;37;40;41}. This seems unlikely for body size, which appears to be highly polygenic^{42;43}, although the finding that inbred individuals are on average smaller does suggest some directional dominance⁴⁴⁻⁴⁷ which would also generate skew in the opposite direction to selection. Two other studies have looked at the distribution of breeding values (indirectly through estimating the skew of breeding values estimated in a Gaussian model) and while one also found little evidence of skew²⁹, the other found skew in the opposite direction to selection²⁸. Lack of genetic skew would also be a consequence of selection acting on an environmentally correlated trait, rather than acting directly on size^{48;49} (discussed further below). More widespread assessments of the prevalence of genetic skew are needed to assess the generality of these results.

Environmental skew has received little attention from theoreticians, with most studies assuming that environmental effects are normally distributed^{11;12;14}. There are, however, several biological processes that are known to induce environmental skew. As far as we are aware, these processes are all predicted to generate negative environmental skew, which fits with our

216 general observation of negative skew in juvenile body size across species (Figure 2). For
217 example, asymmetric competition, when larger individuals have a disproportionate negative
218 competitive effect on others, can drive negative skew^{50–53}. Blue tits have moderate levels of
219 hatching asynchrony (hatching spread is approximately 2 days; see⁵⁴ for distribution across
220 bird species) which is expected to generate asymmetries in competitive ability⁵⁵ and therefore
221 skew at the within-nest level. However, the dominant source of phenotypic skew is at the
222 between-nest level (contribution to phenotypic skew relates to standardised skew and variance)
223 and so if asymmetric competition was the main driver of phenotypic skew, it would require
224 parental ability to be driven by asymmetric adult competition, perhaps through differences
225 in condition and/or territory quality. An alternative explanation is that (some) chicks have
226 yet to reach their asymptotic size by the time of measurement and so variation in their size
227 at this time is driven by variation in growth rate and asymptotic size. If variation in growth
228 rate is largely at the between-nest level and variation in the asymptote is largely genetic, as
229 has been suggested in great tits⁵⁶, then the non-linearity of growth functions could result in
230 skew that is primarily environmental in origin (see⁵⁷ for a related result). This skew would be
231 expected to disappear further into development as all chicks reach their asymptotic size, but
232 due to the strong selective disappearance of small chicks this may not necessarily manifest
233 itself (see below).

234 The strong, negative environmental skew led the PO-regression in all traits to be convex. This
235 occurs because the long tail of small individuals are primarily small because of environmental
236 factors and so resemble their parents less than larger individuals. Most discussions of the
237 linearity of the PO-regression focus on how, in combination with a non-linear fitness function,
238 a non-linear PO-regression leads the breeder's equation to be inaccurate, through generating a
239 covariance between the residuals from a linear fitness function and the linear PO-regression^{3;14}
240 (see also Figure S18). This 'spurious response to selection'¹⁴ will be largest when the
241 non-linear fitness function and the PO-regression either have the same non-linear shape (e.g.
242 both concave) causing a positive covariance between residuals, leading the breeder's equation
243 to under-estimate the response to selection or opposite shapes (e.g. one concave and one
244 convex), creating negative covariance between residuals and so over-estimation of selection
245 response. Skew generates quite predictable and simple non-linearity in the PO-regression
246 (Figure 1), and so generally accelerating or decelerating fitness functions will be more likely
247 to generate a spurious response to selection, as is seen with wing length (Figure S18).

248 We additionally show that the selective disappearance of small individuals alongside a non-linear
249 PO-regression leads to h^2 estimates that are biased towards the slope of the surviving large
250 individuals. This selection bias is particularly striking in estimates from PO-regression (approx
251 65% increase in h^2 for mass and head-bill length; Figure 5) but importantly also occurs
252 in animal models applied to pedigrees where information about the genetic variance comes
253 primarily from parent-offspring comparisons (e.g. typical bird pedigrees without cross-fostering),
254 although to a lesser degree (23% increase in animal models compared to a 61% increase in
255 PO-regression; Table 1). This bias occurs because both PO-regression and the animal model
256 assume that the relationship between offspring and parental phenotypes is linear, and so
257 assumes the missing parent-offspring comparisons would follow the same slope. It is worth
258 noting that we simulated closed populations and so a higher relatedness structure than in
259 most wild bird populations, which are characterised by low recruitment and high immigration.
260 Thus, our simulations likely underestimated the possible bias in animal models. We also

261 demonstrated that cross fostering eliminated this bias in animal models. This occurs because
262 cross-fostering shifts the majority of the information for estimating V_A from parent-offspring
263 comparisons, to sibling comparisons, and sibling comparisons are made before selection whilst
264 parent-offspring comparisons are made after.

265 Previous work in this system has shown that selection differentially eliminates negative environmental,
266 but not genetic, deviations for mass over the course of development⁵⁸. This was interpreted
267 as mass being an environmentally correlated target of selection rather than the true target
268 (i.e. no causal relationship between size and survival)⁴⁸. However, incorporating skew into
269 our models challenges this interpretation as, under our model, size is the true target of
270 selection. As the long tail of small individuals are small for environmental reasons, the selective
271 disappearance of these individuals drives the observed decrease in environmental variance and
272 skew though ontogeny. Given the selective disappearance previously observed was prior to the
273 measurements analysed here⁵⁸ it seems likely that the environmental skew we observe is an
274 underestimation of the true skew, meaning we are likely underestimating the true non-linearity
275 of the PO-regression. Multivariate methods would account for this selective disappearance⁵⁹,
276 however, these proved too complex to implement in this instance.

277 Given the consistent negative environmental skew we see across the four traits, and the
278 conserved nature of negative phenotypic skew in juvenile (but not adult) size across bird
279 species, we believe a concave PO-regression for juvenile size traits might be a general finding.
280 As found here, juvenile body size is also generally under strong viability selection across
281 taxa³⁰. Together, this suggests that previous heritability estimates of juvenile size are likely
282 to have been systematically over-estimated, especially as a large proportion are based on
283 PO-regressions⁶⁰. Indeed, tarsus length heritability estimates from PO-regressions have been
284 shown to be consistently larger than those from animal models⁶⁰. Juvenile size is a hallmark
285 trait of evolutionary stasis, whereby traits that should respond to selection in the wild appear
286 not to. Although these results do not fully explain this stasis, they do show that the predicted
287 response to selection may be being substantially overestimated in traits with non-Gaussian
288 phenotypic distributions.

289 Lande-Arnold regression is by far the most common method for estimating β ^{5;32;61} and is
290 known to be unbiased in the presence of phenotypic skew only if the fitness function is linear
291 or quadratic *and* this quadratic term is modelled²². Although the estimated survival functions
292 deviated from a quadratic for all traits, estimates of β were close to those that would have
293 been obtained under Lande-Arnold regression including the quadratic term (β_2) for all traits,
294 and without the quadratic term (β_1) for three traits. The near equivalence of these different
295 estimates seems at odds with the conclusions of Bonamour *et al.*¹⁷, who demonstrate that
296 selection gradients approximated with Lande-Arnold regression are biased in the presence of
297 phenotypic skew. However, Bonamour *et al.* only modelled the linear term in the Lande-Arnold
298 regression (β_1) whilst assuming a quadratic fitness function - had the quadratic term also been
299 included, the linear term in the Lande-Arnold regression (β_2) would have been unbiased (²²,³
300 Chapter 29), in correspondence with our wing length results (β_1 underestimated β , but β_2
301 did not). However, there is no reason to believe including a quadratic term in a Lande-Arnold
302 regression will generally result in a good approximation of β . Indeed, Morrissey & Sakrejda⁵
303 compared β with that approximated from a quadratic Lande-Arnold regression and found quite
304 large proportional differences (approx. 30%), although small differences in absolute terms.

305 We therefore urge caution in assuming that our results are a general statement about the
306 accuracy of Lande-Arnold regression under non-normality.

307 Quantitative genetics uses two main frameworks to predict how traits will respond to selection.
308 Here we demonstrate how both of these frameworks are affected by skew at the environmental
309 and genetic levels. Genetic skew can lead both the breeder's equation and Lande's gradient
310 equation to be inaccurate. Although little or no genetic skew has been found in the few
311 studies that have tried to quantify it, it remains unknown to what extent this is a generality,
312 and will be highly dependent on the genetic architecture of specific traits. In the absence of
313 genetic skew, the gradient equation presents an accurate prediction of selection response¹¹,
314 although environmental skew provides challenges to the accurate estimation of both β and V_A .
315 Whilst the breeder's equation may provide a more intuitive way of thinking about selection
316 response, the extensions to this framework that allow for non-linearity¹² are complex and
317 computationally expensive. We therefore recommend a focus on the gradient equation (and
318 its extensions¹¹) in wild systems, where fitness functions are highly likely to be non-linear and
319 trait distributions are commonly skewed.

320 Methods

321 This study was preregistered (see <https://osf.io/7qyp4/>). We have highlighted in the following
322 sections where our methods deviate from those planned.

323 Meta-analysis of Skew

324 We collected raw data on juvenile and adult tarsus length from several sources: we used a
325 mailing list to request data, we searched the dryad repository for 'tarsus', we emailed groups
326 with known long-term avian datasets that were not represented in these sources and included
327 any tarsus length data that we otherwise encountered. When datasets from different studies of
328 the same population overlapped in time, we use the largest single dataset available. Datasets
329 were taken from ^{43;62-99}.

Sample standardised skew was estimated from raw data z as

$$\frac{\frac{1}{n} \sum_{i=1}^n (z_i - \hat{\mu})^3}{\left[\frac{1}{n} \sum_{i=1}^n (z_i - \hat{\mu})^2 \right]^{3/2}} \frac{\sqrt{n(n-1)}}{n-2}$$

with sampling variance as

$$\frac{6n(n-1)}{(n-2)(n+1)(n+3)}$$

330 where n is sample size and $\hat{\mu}$ the estimate of the trait mean¹⁰⁰.

331 Using these data, we ran a random-effect meta-analytic model in MCMCglmm with age
332 (juvenile or adult) as a fixed factor and random effects of species and study. Models were run
333 for 65000 iterations, with a burnin of 15000 and a thinning intervals of 50. The priors for the
334 random-effect variances were scaled (by 100) $F_{1,1}$ and the prior for the residual variance was
335 inverse-gamma with a shape and scale of 0.001. The fixed effects had a diffuse normal prior
336 (mean=0, variance= 10^{10}).

337 Study population

338 We used data from a nest-box population of blue tits (*Cyanistes caeruleus*), on the Dalmeny
339 estate, Edinburgh, United Kingdom, collected from 2011 to 2018, with 253 nest-boxes over
340 two sites. Detailed methods are described in ^{58;101}. Briefly, all nests were visited regularly until
341 the discovery of the first egg, and then daily for egg cross-fostering, when eggs were weighed.
342 From 2011-2013 and 2016-2018 a partial egg cross-fostering design was used to enable additive
343 genetic and nest-of-rearing effects on offspring size to be separated⁵⁸. In 2014-2015 a mixture
344 of full and partial cross-fostering was used as part of a separate experiment. Full details of
345 cross-fostering can be found in¹⁰². After egg laying was complete, nests were left undisturbed
346 for 11 days and then checked daily for hatching. At hatching (day 0), all chicks were uniquely
347 marked (within a nest). The chicks had blood samples taken at day 3 and were given a unique
348 metal ring at day 9. At day 15, chick's tarsus, wing and head-bill lengths were measured and
349 they were weighed. For the morphometric measurements, one chick from each nest was
350 measured twice in order to account for measurement error⁵⁸. From day 10, adults were

351 caught at the nest in order to identify them; blood samples and morphometric measurements
352 were taken and the birds were uniquely ringed. At the end of the season we checked all
353 nests and recorded any dead chicks left in the nest. From this we could infer which chicks
354 fledged. Chicks were considered recruited if they were recaptured as breeders in subsequent
355 years. Permission to monitor, catch and ring the birds was given by Scottish Natural Heritage
356 and the British Trust for Ornithology and permission to take blood samples was granted by
357 the UK Government's Home Office. All permission and licenses were granted to JDH.

358 Social parentage was assigned through catching parents at the nest. When no female was
359 caught, the social female was assigned a dummy mother identity. When no male was caught,
360 the social father was assigned as the genetic sire with the largest proportion of paternity in
361 a nest, either a male caught at a different nest that year, or an unsampled male assigned a
362 dummy identity.

363 For the assignment of genetic parentage and chick sex, genotypes were obtained using blood
364 and tissue samples from adults and chicks. Genotyping and pedigree reconstruction largely
365 followed protocols outlined in⁵⁸ and¹⁰¹. However, adults not caught in the focal year but
366 that were known to be alive (because they were caught in subsequent years and were aged 2
367 years or over) were allowed to be parents of chicks in the focal year. The distance between
368 the nest-of-origin of the chicks and the nest at which these candidate parents were caught in
369 the subsequent year was fitted as a covariate. Mothers were allowed to be polygamous when
370 (half) sib-ships were assigned to chicks with unknown fathers (see Supplementary Methods).
371 When assigning chick sex, we used morphological sexing of recruits over molecular sexing from
372 chicks (sexing didn't match for 5 chicks).

373 For our analysis we included data on chick size measured at day 15 post-hatching, collected on
374 this project from 2011-2018, and additionally chick recruitment data from 2019 and 2020. We
375 included all nests for which hatching date was known. Although similar morphological data
376 was collected in 2010, we excluded all records from this year as egg size was not measured.
377 Egg size was used to account for nest-of-origin effects in our models (see below). We also
378 excluded data from an additional two nests where egg size was not measured, from chicks for
379 which molecular sexing was not successful (n=20 chicks) and where we did not have one of
380 the day 15 measurements (n=11 chicks). In total, we had records of 5123 day 15 chicks in
381 715 nests, with 642 chicks repeatedly measured.

382 **Statistical analysis**

383 All models were run in a Bayesian framework. From all models posterior means and 95%
384 credible intervals are presented. A p-value for the fixed effects and covariances in these
385 models was approximated (pMCMC) as two times the smaller number of iterations where the
386 parameter value is either less than zero or greater than zero¹⁰³. We use a threshold of 0.005
387 to refer to results as significant and those between 0.05 and 0.005 as suggestive¹⁰⁴.

388 **Decomposing phenotypic skew using hierarchical models**

389 We modelled the four traits (tarsus length, head-bill length, mass and wing length) measured
390 at day 15 using linear mixed effects models with sex (2 level factor), year (8 level factor), time
391 of day (continuous - hours from midnight) and egg size (continuous) as fixed. Additive genetic

392 and nest-of-rearing effects were modelled as random. Because we have repeated measurements
393 of tarsus, wing and head-bill lengths, we additionally modelled measurement error effects in
394 these traits, by including bird identity effects, which are equivalent to the residuals in a model
395 without repeat measures, and the residuals are measurement error effects⁵⁸. In contrast to
396 past analyses^{58;101}, we do not model nest-of-origin effects but rather include egg size as a
397 covariate to account for these effects (see⁵⁸ and Supplementary materials). As estimating
398 skew-t distributed random effects (see below) is parameter heavy, including a covariate rather
399 than a random effect is preferable, especially as nest-of-origin effects are very small for these
400 traits^{58;101}.

401 Skew due to the fixed effects was obtained by multiplying the fixed effect design matrix by the
402 fixed effects and estimating the parameters for the skew-t distribution of the resulting variable.
403 These were used when calculating the non-linear parent offspring regression and when plotting
404 the sample skew. This method assumes that the joint distribution of the covariates is equal to
405 the empirical distribution we observe. In combination with a diffuse prior on the fixed effects,
406 this assumption probably leads to a small inflation in the estimated (absolute) skew. Time
407 of day was excluded from this estimate as any skew induced by this is due to our sampling
408 design rather than being biologically relevant.

409 In order to estimate skew in the random effects, we fitted random effects with skew-t
410 distributions. The residuals for the repeat measured traits were treated as Gaussian as these
411 represent measurement errors. As with the normal distribution, the skew-t distribution^{105–108}
412 has a location ξ and scale ω parameter, but also parameters δ and ν which modify the skew and
413 tailness, respectively. The distribution converges on a normal distribution when $\delta = 0$ and ν
414 approaches infinity. As δ moves away from 0 and ν decreases the (absolute) skew in a variable
415 increases, with the sign of δ signifying the direction of the skew. The skew-t distribution is
416 unbounded and readily allows for considerable amounts of positive and negative skew. The
417 reasons for the use of this distribution are further discussed in the supplementary materials.
418 Our approach to modelling the additive genetic effects is to extend standard quantitative
419 genetic models by allowing the base population breeding values to have a skew-t distribution,
420 with normally distributed Mendelian sampling deviations in the descendants (with variance
421 $\omega^2(1 - \bar{F})/2$ where \bar{F} is the average inbreeding coefficient of the individual's parents). This
422 assumes that inheritance occurs under the Gaussian descendants infinitesimal model^{34;109}; i.e.
423 the Mendelian sampling deviations are normally distributed within families, and any genetic
424 skew results from selection. In practice, however, the Mendelian sampling deviations are largely
425 confounded with residual effects in our data because there are few parent-offspring comparisons
426 (due to high migration and low recruitment) and so inferences are probably quite robust to
427 any violation of the Gaussian descendants assumption. Initially we tried to fit this model in an
428 animal model framework, but due to poor mixing we chose to approximate the model using
429 a dam-sire model. This model discards information about the Mendelian-sampling deviations
430 and subsumes them in the residual effects which then come from a mixture distribution¹¹⁰.
431 Given there is little information in our data about the Mendelian-sampling deviations the
432 dam-sire and animal models are expected to give almost identical answers (see Supplementary
433 Materials). Although this method allows us to directly estimate skew in breeding values, when
434 the environmental residuals are skew-t, as assumed here, the mixture distribution does not
435 have standard form. Here, we approximate the mixture distribution as skew-t and although
436 we cannot derive the full distribution of the environmental residuals we are able to obtain their

437 variance and skew. These models provided little evidence for genetic skew in any trait and so
 438 we reverted to an animal model with normally distributed breeding values - the animal model
 439 approach having the advantage that the environmental residual skew can then be directly
 440 estimated. The dam and sire effects were modelled in a multi-membership model where the
 441 two sets of effects were constrained to having the same skew-t distribution.

442 Initially, we intended to model chick mass over ontogeny in a multivariate framework (see
 443 preregistration), as in previous studies of this population^{58;101}. However, implementing the
 444 required multivariate skew-t models proved too challenging. Since there is strong directional
 445 selection on chick body mass throughout ontogeny^{58;101}, our estimates of skew at day 15
 446 are likely underestimates as the univariate analysis used will fail to account for selective
 447 disappearance prior to day 15^{58;101}. We also planned to have a global box-cox parameter
 448 in case there was a single transformation that would make everything linear and additive.
 449 However, given the problems we had with implementing more complex models, we chose not
 450 to include this additional complexity.

451 It should also be noted that estimates from these skew-t models seem to be more sensitive
 452 to unmodelled heteroskedasticity than standard Gaussian mixed effects models, even when
 453 skew exists, and this can lead to biased fixed effect and variance estimates. This led us
 454 to fit a reduced set of fixed effects compared with previous analyses^{58;101} and outlined in
 455 our pre-registration (see Supplementary materials). To partly address this issue we also ran
 456 equivalent Gaussian models for all skew-t models, and present the results in the Supplementary
 457 materials. There were small differences between models but the results remain qualitatively
 458 the same (see SM; Figure S17, Tables S4-15).

459 These models were run using Stan (version 2.21.0)¹¹¹ using the cmdstanr package (Stan
 460 Development Team, 2019) in R (version 4). Four chains were run for each model with
 461 a warmup of 4000 iterations and 6000 iterations post-warmup, with the exception of the
 462 dam-sire wing length model which was run with a warmup of 5000 iterations and 10000
 463 iterations post-warmup. Convergence of individual chains was visually assessed, as well as
 464 ensuring that the Gelman–Rubin diagnostic (R-hat) across chains was less than 1.1¹¹². We
 465 used diffuse normal priors for fixed effects (mean=0 and standard deviation=100), half-Cauchy
 466 priors (mean=0 and standard deviation=10) for standard deviations and uniform priors from
 467 -1 to 1 for δ and 4 to 40 on ν . The choice of priors is discussed further in the Supplementary
 468 materials.

469 **Non-Linear Parent-Offspring Regression**

470 The PO-regression function is defined as $E[z_o|z]$ where z_o is the phenotype of offspring from a
 471 parent with phenotype z . Assuming random mating and environmental values in the offspring
 472 (e_o) are independent of parental phenotypes this becomes $\frac{1}{2}E[g|z] + \frac{1}{2}E[g] + E[e_o]$ under
 473 the Gaussian descendants assumption, where g is breeding value. Have θ_g be the parameters
 474 of the breeding value distribution and θ_e the parameters of the environmental distribution.
 475 Then,

$$E[g|z] = \frac{\int (z-e)p(z-e|\theta_g)p(e|\theta_e)de}{\int p(z-e|\theta_g)p(e|\theta_e)de} \quad (1)$$

476 The integrals have to be evaluated numerically, which is time consuming, and so the regression
 477 function was evaluated at the posterior mean of the parameters from the skew-t animal models
 478 to give $E[z_o|z]$ for each trait (Figure 5). Also, note that in the presence of pre-breeding survival
 479 selection, the term $\frac{1}{2}E[g]$ in the intercept of the regression function should be replaced by
 480 $\frac{1}{2}(E[g] + \Delta g)$ where Δg is the change in mean breeding value due to selection such that
 481 $E[g] + \Delta g$ is the expected breeding value of the other parent¹⁴.

482 Selection on chick body mass

483 Given that we were not able to model chick body mass in a multivariate framework, we did
 484 not model survival throughout ontogeny as originally planned (see preregistration), but rather
 485 modelled survival from day 15 to fledging and fledging to recruitment. We modelled this as
 486 an event history in a probit regression (binomial error distribution and probit link function)
 487 including a quadratic effect of chick size at day 15 on both events, allowing us to model the
 488 stabilising component of selection. These models accounted for measurement error in tarsus,
 489 head-bill and wing lengths, using the repeated measurements of these traits. Originally we
 490 planned to correct our measurements for time of day effects (see preregistration). However,
 491 these effects proved to be very small and for most traits non-significant (see Supplementary
 492 Results). We therefore decided not to add this extra complexity into our models.

493 Sex, day of hatching within the nest, year, clutch size, male presence, nest hatch date were
 494 also included as fixed effects. All fixed effects were allowed to differ between the two events.
 495 Finally we modelled the 2x2 covariance matrix of nest-of-rearing effects. This model was
 496 run using Stan. Four chains were run for each model with 5000 iterations and a warmup of
 497 2500 iterations with a thinning interval of 10. Convergence of chains was assessed as above.
 498 Diffuse priors for fixed effects (mean=0 and standard deviation=100), half-Cauchy priors for
 499 all standard deviations (mean=0 and standard deviation=10) and LKJ priors on correlations
 500 with shape=2¹¹³ were used.

501 The Individual Relative Fitness Function

502 Partitioning the linear predictors for each survival event (1: day 15 to fledging, 2: fledging to
 503 recruitment) into a part due to the trait and a part due to remaining terms (denoted η), and
 504 assuming that the distribution of $\eta^{(1)}$ and $\eta^{(2)}$ are bivariate normal conditional on the trait z ,
 505 then the absolute fitness function has the form:

$$W(z) = F_{MVN}(s|\Sigma) \quad (2)$$

506 where F_{MVN} is the multivariate normal cumulative density function in which the first argument
 507 is the quantile to be evaluated and the second argument is the (co)variance of the variates
 508 (the means are zero and are therefore not given). For event i

$$s^{(i)} = E[\eta^{(i)}] + \frac{COV(\eta^{(i)}, z)}{\mu_2}(z - \mu) + \beta^{(i)}z + \frac{1}{2}\gamma^{(i)}z^2 \quad (3)$$

509 where $\beta^{(i)}$ and $\frac{1}{2}\gamma^{(i)}$ are the linear and quadratic effect of the trait on event i , μ is the trait
 510 mean and μ_i the i^{th} central moment of the phenotypic distribution.

$$\Sigma^{(i,j)} = COV(\eta^{(i)}, \eta^{(j)}) - \frac{COV(\eta^{(i)}, z)COV(\eta^{(j)}, z)}{\mu_2} + COV(u^{(i)}, u^{(j)}) + \delta^{(i,j)} \quad (4)$$

511 where $u^{(i)}$ are the nest effects for event i and $\delta^{(i,j)} = 1$ when $i = j$ and represents the residual
 512 variance.

513 The partial derivative of $W(z)$ with respect to z is given by

$$\begin{aligned} \frac{\partial W(z)}{\partial z} = & f_N(s^{(1|2)} | \Sigma^{(1|2)}) \left(\frac{COV(\eta^{(1)}, z)^2}{\mu_2} + \beta^{(1)} + \gamma^{(1)}z - \frac{\Sigma^{(1,2)}}{\Sigma^{(2)}} \left(\frac{COV(\eta^{(2)}, z)^2}{\mu_2} + \beta^{(2)} + \gamma^{(2)}z \right) \right) \\ & F_N(s^{(2)} | \Sigma^{(2)}) + f_N(s^{(2)} | \Sigma^{(2)}) \left(\frac{COV(\eta^{(2)}, z)^2}{\mu_2} + \beta^{(2)} + \gamma^{(2)}z \right) F_N(s^{(1|2)} | \Sigma^{(1|2)}) \end{aligned} \quad (5)$$

514 where f_N and F_N are the density and cumulative density functions for a centred normal
 515 distribution, and

$$s^{(1|2)} = s^{(1)} - \frac{\Sigma^{(1,2)}}{\Sigma^{(2)}} s^{(2)} \quad \Sigma^{(1|2)} = \Sigma^{(1)} - \frac{(\Sigma^{(1,2)})^2}{\Sigma^{(2)}} \quad (6)$$

516 Solving Equation 5 to find the stationary point(s), and therefore the optimal trait value, is
 517 difficult. Instead we evaluated the derivative of Equation 5 at the minimum and maximum
 518 observed trait value and assessed whether the derivative at the minimum is positive and
 519 negative at the maximum. This condition implies an optimal trait value within the range of
 520 observed trait values.

521 Selection Gradients

522 The Lande-Arnold method²² for estimating the selection gradient is only robust to phenotypic
 523 skew if the fitness function is quadratic and both the mean-centered trait value and its square
 524 are fitted in the regression^{3;22}. We therefore computed three selection gradients. Using the
 525 notation in³³, we calculated our best estimate of it¹¹⁴,

$$\beta = E \left[\frac{\partial w(z)}{\partial z} \right] = \int \frac{\partial w(z)}{\partial z} p(z) dz \approx \frac{1}{n} \sum_{i=1}^n \frac{\partial w(z)}{\partial z} \Big|_{z_i} \quad (7)$$

526 where $p(z)$ is the probability density function for z , $w(z)$ is the relative fitness function
 527 obtained by dividing $W(z)$ by mean fitness ($E[W] = \int W(z)p(z)dz$) and z_i are the observed
 528 trait values. Put simply, we calculated the mean partial derivative of individual fitness function
 529 (from Equation 5) across our observed phenotypic distributions, divided by mean fitness.

530 The linear selection differential is defined as

$$S = \int zw(z)p(z)dz - \mu \approx \frac{1}{n} \sum_{i=1}^n z_i w(z_i) - \hat{\mu} \quad (8)$$

531 and the quadratic selection differential as

$$C = \int (z - \mu)^2 p(z) w(z) dz - \mu_2 \approx \frac{1}{n} \sum_{i=1}^n (z_i - \hat{\mu})^2 w(z_i) - \hat{\mu}_2 \quad (9)$$

532 From these we can calculate the expected linear regression coefficient from the Lande-Arnold
533 method when only the linear term was fitted:

$$\hat{\beta}_1 = \frac{\hat{S}}{\hat{\mu}_2} \quad (10)$$

534 and the linear regression coefficient from the Lande-Arnold method when both the linear and
535 quadratic term are fitted (Eq. 29.28a from³):

$$\hat{\beta}_2 = \frac{(\hat{\mu}_4 - \hat{\mu}_2^2)\hat{S} - \hat{\mu}_3\hat{C}}{\hat{\mu}_2(\hat{\mu}_4 - \hat{\mu}_2^2) - \hat{\mu}_3^2} \quad (11)$$

536 Selection cannot operate on between-sex differences in trait values (the average fitness of
537 the two sexes is constrained to be equal) and we assume that selection does not operate on
538 between-year differences in trait values (which might occur if juvenile size impacts on adult
539 survival). We therefore estimated each β as the average of each sex by year combination
540 (Figure 4 e-h), calculated across the posterior distribution of the survival model.

541 **Response to Selection**

542 The extension of Lande's gradient equation to a non-normal distribution of genetic effects is
543 (combining Equations 26 and 42 from¹¹):

$$\Delta\mu = \sum_{j=1}^{\infty} K^{j+1}(g) \frac{1}{j!} \int \frac{\partial^j w(z)}{\partial z^j} p(z) dz \quad (12)$$

544 where $K^j(x)$ denotes the j^{th} cumulant of x , which up to the third cumulant (skew) is

$$\Delta\mu = V_A E \left[\frac{\partial w(z)}{\partial z} \right] + \frac{S_A}{2} E \left[\frac{\partial^2 w(z)}{\partial z^2} \right] \quad (13)$$

545 where S_A is the skew in the additive genetic effects. When the distribution of additive genetic
546 values is normal and/or the fitness function is linear, Equation 12 reduces to Lande's gradient
547 equation

$$\Delta\mu = V_A E \left[\frac{\partial w(z)}{\partial z} \right] = V_A \beta \quad (14)$$

548 since all cumulants > 2 of the genetic distribution are zero.

549 Heritability

550 We compared how well our inferred non-linear PO-regression (Equation 1) performed at
551 predicting offspring phenotype compared to linear single-parent mid-offspring regression. Using
552 the 182 individuals (118 male and 64 female) that were measured as chicks at day 15 and
553 survived to produce offspring that were also measured at day 15, we fitted a weighted (by
554 family size) regression with our inferred non-linear PO-regression fitted as an offset. We then
555 compared the fit of this model to an identical model but where the raw parental phenotype
556 was also fitted as a covariate with a free parameter.

557 We then compared estimates of the heritability before and after selection (h_b^2 and h_a^2 , respectively).
558 The heritability can be defined as the regression coefficient of a linear mid-PO-regression, and
559 can be calculated before selection

$$h_b^2 = 2 \frac{COV(z_o, z)}{\mu_2} = \frac{V_A}{V_P} \quad (15)$$

560 or after selection

$$h_a^2 = 2 \frac{E[w(z)z_o z] - E[w(z)z_o]E[w(z)z]}{E[w(z)z^2] - E[w(z)z]^2} \quad (16)$$

561 The posterior distribution of h_b^2 was evaluated directly, but the i^{th} posterior sample of h_a^2
562 was obtained by simulating 10^4 values of z and z_o using the parameters sampled at the
563 i^{th} iteration of the trait model, calculating expected fitness for each sampled z using the
564 parameters sampled at the i^{th} iteration of the fitness model, and then evaluating the relevant
565 expectations.

566 Simulations

567 To test how different sampling designs and standard estimation procedures (PO-regression
568 and Gaussian animal model) impact estimates of heritability in the presence of skew and
569 selection, we simulated data according to the posterior mean of the parameters from our
570 skew-t quantitative genetic and selection models for mass. A closed population with 1000
571 breeding pairs was simulated over three generations, with 10 measured full-sib offspring per
572 pair. Four scenarios were simulated: either nests were not cross-fostered or they were paired
573 and five offspring reciprocally crossed, and the random effects were either skew t-distributed
574 (with ω , δ and ν parameters set to their posterior means) or they were normally distributed but
575 with matching variance. The probability of a chick recruiting to be a parent was obtained by
576 applying the estimated survival model for chick mass to the simulated phenotype. Each of the
577 four scenarios were simulated 2000 times and for each data set the heritability was estimated
578 directly using PO-regression and as the estimate of the additive genetic variance over the sum
579 of all variances estimated from a Gaussian animal model fitted in ASReml-R¹¹⁵.

580 Data availability

581 All data and code can be found at <https://doi.org/10.5281/zenodo.5342526>.

582 **Acknowledgements**

583 We thank our many field assistants for help with data collection, Shinchi Nakagawa, Anders
584 Moller, Diego Santiago-Alarcon, Roger Jovani, Sergi Sales and Nuria Rodriguez for providing
585 raw data, Eryn McFarlane, Julie Gauzere and Ed Ivimey-Cook for helpful discussions and two
586 anonymous reviewers for their comments on the manuscript. This work was funded by Natural
587 Environment Research Council (NE/P000924/1) and Royal Society Fellowship to JDH, and
588 supported by Lord Rosebery and Dalmeny estate.

589 **Author contributions**

590 JLP and JDH conceived and designed the project. JLP, HEL, CET and JDH generated the
591 data. JLP and JDH analysed the data and wrote the paper. All authors have read and
592 approved the paper.

593 **Competing interests**

594 The authors declare no competing interests

595 **References**

- 596 [1] Lush, J. L. *Animal Breeding Plans* (Iowa State College Press, Ames, Iowa, 1937).
- 597 [2] Lande, R. Natural selection and random genetic drift in phenotypic evolution. *Evolution*
598 **30**, 314–334 (1976).
- 599 [3] Walsh, B. & Lynch, M. *Evolution and Selection of Quantitative Traits* (Oxford
600 University Press, Oxford, UK, 2018).
- 601 [4] Schluter, D. Estimating the Form of Natural Selection on a Quantitative Trait. *Evolution*
602 **42**, 849–861 (1988).
- 603 [5] Morrissey, M. B. & Sakrejda, K. Unification Of Regression-Based Methods For The
604 Analysis Of Natural Selection. *Evolution* **67**, 2094–2100 (2013).
- 605 [6] Falconer, D. & Mackay, T. F. *Introduction to Quantitative Genetics* (Longman, New
606 York, 1996), 4th edn.
- 607 [7] Lynch, M. & Walsh, B. *Genetics and Analysis of Quantitative Traits* (Sinauer
608 Associates, Inc., Sunderland, MA, 1998).
- 609 [8] Roff, D. A. *Evolutionary quantitative genetics* (Springer Science & Business Media,
610 2012).
- 611 [9] Nishida, A. & Abe, T. Non-Linear Heritability and Asymmetrical Selection Responses
612 caused by Skewed Distribution of Breeding Value in Selected Population. *Japanese*
613 *Journal of Zootechnical Science* **51**, 495–500 (1980).

- 614 [10] Barton, N. H. & Turelli, M. Adaptive landscapes, genetic distance and the evolution of
615 quantitative characters. *Genetical Research* **49**, 157–173 (1987).
- 616 [11] Turelli, M. & Barton, N. H. Genetic and statistical analyses of strong selection on
617 polygenic traits: What, me normal? *Genetics* **138**, 913–941 (1994).
- 618 [12] Gimelfarb, A. & Willis, J. H. Linearity versus nonlinearity of offspring-parent regression:
619 An experimental study of *Drosophila melanogaster*. *Genetics* **138**, 343–352 (1994).
- 620 [13] Rice, S. H. *Evolutionary theory: mathematical and conceptual foundations* (Sinauer
621 Associates, 2004).
- 622 [14] Heywood, J. S. An exact form of the breeder's equation for the evolution of a
623 quantitative trait under natural selection. *Evolution* **59**, 2287–2298 (2005).
- 624 [15] Jones, A. G., Bürger, R., Arnold, S. J., Hohenlohe, P. A. & Uyeda, J. C. The effects
625 of stochastic and episodic movement of the optimum on the evolution of the G-matrix
626 and the response of the trait mean to selection. *Journal of Evolutionary Biology* **25**,
627 2210–2231 (2012).
- 628 [16] Urban, M. C., Bürger, R. & Bolnick, D. I. Asymmetric selection and the evolution of
629 extraordinary defences. *Nature Communications* **4** (2013).
- 630 [17] Bonamour, S., Teplitsky, C., Charmantier, A., Crochet, P. A. & Chevin, L. M. Selection
631 on skewed characters and the paradox of stasis. *Evolution* **71**, 2703–2713 (2017).
- 632 [18] Jacquard, A. Heritability: one word, three concepts. *Biometrics* 465–477 (1983).
- 633 [19] Charlesworth, B. The Heritability of Fitness. In Bradbury, J. & Andersson, M. (eds.)
634 *Sexual Selection: Testing the Alternatives*, 21–40 (John Wiley & Sons Limited, 1987).
- 635 [20] Nishida, A. & Abe, T. Distribution of Genetic and Environmental Effects and Linearity
636 of Heritability. *Canadian Journal of Genetics and Cytology* **16**, 3–10 (1974).
- 637 [21] Robertson, A. The non-linearity of offspring-parent regression. In Pollak, E.,
638 Kempthorne, . & Bailey Jr, T. B. (eds.) *Proceedings of the International Conference*
639 *on Quantitative Genetics*, 297–304 (Iowa State University Press, Ames, 1977).
- 640 [22] Lande, R. & Arnold, S. J. . The Measurement of Selection on Correlated Characters.
641 *Evolution* **37**, 1210–1226 (1983).
- 642 [23] Beardsley, J. P., Bratton, R. & Salisbury, G. The Curvilinearity of Heritability of
643 Butterfat Production. *Journal of Dairy Science* **33**, 93–97 (1950).
- 644 [24] Nishida, A. Some Characteristics of Parent-Offspring Regression in Body-Weight of
645 Mus-Musculus at Different Ages. *Canadian Journal of Genetics and Cytology* **14**,
646 293–303 (1972).
- 647 [25] Mäki-Tanila, A. *The Validity of the Heritability Concept in Quantitative Genetics*. Ph.D.
648 thesis, University of Edinburgh (1982).
- 649 [26] Gifford, D. R. & Barker, J. S. The nonlinearity of offspring-parent regression for
650 total sternopleural bristle number of *Drosophila melanogaster*. *Theoretical and Applied*
651 *Genetics* **82**, 217–220 (1991).

- 652 [27] Koerhuis, A. N. Non-normality of egg production distributions in poultry and the
653 effects of outlier elimination and transformation on size and curvilinearity of heritability.
654 *Livestock Production Science* **45**, 69–85 (1996).
- 655 [28] McGuigan, K., Van Homrigh, A. & Blows, M. W. Genetic analysis of female preference
656 functions as function-valued traits. *The American Naturalist* **172**, 194–202 (2008).
- 657 [29] Reid, J. M. *et al.* Immigration counter-acts local micro-evolution of a major fitness
658 component: Migration-selection balance in free-living song sparrows. *Evolution Letters*
659 **5**, 48–60 (2021).
- 660 [30] Rollinson, N. & Rowe, L. Persistent directional selection on body size and a resolution
661 to the paradox of stasis. *Evolution* **69**, 2441–2451 (2015).
- 662 [31] Merilä, J., Sheldon, B. & Kruuk, L. Explaining stasis: microevolutionary studies in
663 natural populations. *Genetica* **112**, 199–222 (2001).
- 664 [32] Kingsolver, J. G. *et al.* The strength of phenotypic selection in natural populations.
665 *The American Naturalist* **157**, 245–261 (2001).
- 666 [33] Geyer, C. J. & Shaw, R. G. Commentary on Lande-Arnold analysis. technical report no.
667 670 (2008).
- 668 [34] Turelli, M. Commentary: Fisher’s infinitesimal model: A story for the ages. *Theoretical*
669 *Population Biology* **118**, 46–49 (2017).
- 670 [35] Bulmer, M. G. *The Mathematical Theory of Quantitative Genetics* (Oxford University
671 Press, Oxford,UK, 1980).
- 672 [36] Turelli, M. & Barton, N. H. Dynamics of polygenic characters under selection.
673 *Theoretical Population Biology* **38**, 1–57 (1990).
- 674 [37] Zeng, Z. B. Genotypic distribution at the limits to natural and artificial selection with
675 mutation. *Theoretical Population Biology* **32**, 90–113 (1987).
- 676 [38] Keightley, P. D. & Hill, W. G. Directional selection and variation in finite populations.
677 *Genetics* **117**, 573–582 (1987).
- 678 [39] Reed, D. R., Lawler, M. P. & Tordoff, M. G. Reduced body weight is a common effect
679 of gene knockout in mice. *BMC Genetics* **9** (2008).
- 680 [40] Fisher, R. A., Immer, F. R. & Tedin, O. The genetical interpretation of statistics of the
681 third degree in the study of quantitative inheritance. *Genetics* **17**, 107–124 (1932).
- 682 [41] Gimelfarb, A. Offspring-parent genotypic regression: how linear is it? *Biometrics* **42**,
683 67–71 (1986).
- 684 [42] Santure, A. W. *et al.* Replicated analysis of the genetic architecture of quantitative
685 traits in two wild great tit populations. *Molecular Ecology* **24**, 6148–6162 (2015). URL
686 <https://onlinelibrary.wiley.com/doi/abs/10.1111/mec.13452>.
- 687 [43] Silva, C. N. S. *et al.* Insights into the genetic architecture of morphological traits in two
688 passerine bird species. *Heredity* **119**, 197–205 (2017). URL <https://www.nature.com/articles/hdy201729>.
- 689

- 690 [44] Becker, P. J., Hegelbach, J., Keller, L. F. & Postma, E. Phenotype-associated
691 inbreeding biases estimates of inbreeding depression in a wild bird population. *Journal*
692 *of Evolutionary Biology* **29**, 35–46 (2016).
- 693 [45] Huisman, J., Kruuk, L. E., Ellisa, P. A., Clutton-Brock, T. & Pemberton, J. M.
694 Inbreeding depression across the lifespan in a wild mammal population. *Proceedings of*
695 *the National Academy of Sciences of the United States of America* **113**, 3585–3590
696 (2016).
- 697 [46] Pemberton, J. M., Ellis, P. E., Pilkington, J. G. & Bérénos, C. Inbreeding depression
698 by environment interactions in a free-living mammal population. *Heredity* **118**, 64–77
699 (2017).
- 700 [47] Hajduk, G. K. *et al.* Inbreeding, inbreeding depression, and infidelity in a cooperatively
701 breeding bird*. *Evolution* **72**, 1500–1514 (2018).
- 702 [48] Rausher, M. D. The measurement of selection on quantitative traits: biases due to
703 environmental covariances between traits and fitness. *Evolution* **46**, 616–626 (1992).
- 704 [49] Alatalo, R. V., Gustafsson, L. & Lundberg, A. Phenotypic Selection on Heritable
705 Size Traits: Environmental Variance and Genetic Response. *The American Naturalist*
706 **135**, 464–471 (1990). URL [https://www.journals.uchicago.edu/doi/10.1086/](https://www.journals.uchicago.edu/doi/10.1086/285056)
707 [285056](https://www.journals.uchicago.edu/doi/10.1086/285056).
- 708 [50] Koyama, H. Intraspecific competition among higher plants. VIII. Frequency distribution
709 of individual plant weight as affected by the interaction between plants. *J Inst Polytech*
710 *Osaka Cy University* **7**, 73–94 (1956).
- 711 [51] Mock, D. W. & Parker, G. A. *The Evolution of Sibling Rivalry* (Oxford University Press,
712 Oxford, UK, 1997).
- 713 [52] Weiner, J. Asymmetric competition in plant populations (1990).
- 714 [53] Bassar, R. D. *et al.* The effects of asymmetric competition on the life history of
715 Trinidadian guppies. *Ecology Letters* **19**, 268–278 (2016).
- 716 [54] Muller, M. & Groothuis, T. G. Within-clutch variation in yolk testosterone as an
717 adaptive maternal effect to modulate avian sibling competition: evidence from a
718 comparative study. *The American Naturalist* **181**, 125–136 (2013).
- 719 [55] Nilsson, J.-A. & Svensson, M. Sibling competition affects nestling growth strategies in
720 marsh tits. *Journal of Animal Ecology* 825–836 (1996).
- 721 [56] Gebhardt-Henrich, S. & Van Noordwijk, A. The genetical ecology of nestling growth
722 in the great tit. environmental influences on the expression of genetic variances during
723 growth. *Functional ecology* 469–476 (1994).
- 724 [57] Gebhardt-Henrich, S. Heritability of growth curve parameters and heritability of final
725 size: a simulation study. *Growth, development, and aging: GDA* **56**, 23–33 (1992).
- 726 [58] Hadfield, J. D., Heap, E. A., Bayer, F., Mittell, E. A. & Crouch, N. M. A. Disentangling
727 genetic and prenatal sources of familial resemblance across ontogeny in a wild passerine.
728 *Evolution* **67**, 2701–13 (2013).

- 729 [59] Hadfield, J. D. Estimating evolutionary parameters when viability selection is operating.
730 *Proceedings of the Royal Society B: Biological Sciences* **275**, 723–734 (2008).
- 731 [60] Postma, E. Four decades of estimating heritabilities in wild vertebrate populations:
732 Improved methods, more data, better estimates? In Charmantier, A., Garant, D. &
733 Kruuk, L. E. B. (eds.) *Quantitative Genetics in the Wild*, 16–33 (Oxford University
734 Press, Oxford, 2014).
- 735 [61] Dingemanse, N. J., Araya-Ajoy, Y. G. & Westneat, D. F. Most published selection
736 gradients are underestimated: Why this is and how to fix it. *Evolution* 1–13 (2021).
- 737 [62] Arct, A., Drobniak, S., Mellinger, S., Gustafsson, L. & Cichon, M. Data from:
738 Parental genetic similarity and offspring performance in blue tits in relation to brood
739 size manipulation (2020). URL <https://doi.org/10.5061/dryad.v6r0758>.
- 740 [63] Bebbington, K. *et al.* Data from: Consequences of sibling rivalry vary across life in a
741 passerine bird (2016). URL <https://doi.org/10.5061/dryad.12np0>.
- 742 [64] Bebbington, K. *et al.* Data from: Telomere length reveals cumulative individual and
743 transgenerational inbreeding effects in a passerine bird (2016). URL <https://doi.org/10.5061/dryad.52fp4>.
- 744 [65] Becker, P. J. J. *et al.* Data from: Mother-offspring and nest mate resemblance but no
745 heritability in early-life telomere length in white-throated dippers (2015). URL <https://doi.org/10.5061/dryad.b2v37>.
- 746 [66] Berzins, L. L., Gilchrist, H. G. & Burness, G. Data from: No assortative mating
747 based on size in black guillemots breeding in the Canadian Arctic (2015). URL <https://doi.org/10.5061/dryad.1bm5t>.
- 748 [67] Caizergues, A. E., Gregoire, A. & Charmantier, A. Data from: Urban versus forest
749 ecotypes are not explained by divergent reproductive selection (2018). URL <https://doi.org/10.5061/dryad.tv45802>.
- 750 [68] Camacho, C., Canal, D. & Potti, J. Data from: Nonrandom dispersal drives phenotypic
751 divergence within a bird population (2014). URL <https://doi.org/10.5061/dryad.h22n9>.
- 752 [69] Class, B. & Brommer, J. Data from: Can dominance genetic variance be ignored in
753 evolutionary quantitative genetic analyses of wild populations? (2020). URL <https://doi.org/10.5061/dryad.zpc866t6d>.
- 754 [70] Cornell, A., Gibson, K. F. & Williams, T. D. Data from: Physiological maturity at a
755 critical life-history transition and flight ability at fledging (2017). URL <https://doi.org/10.5061/dryad.c2n66>.
- 756 [71] Cox, A. R., Robertson, R. J., Lendvai, A. Z., Everitt, K. & Bonier, F. Data from: Rainy
757 springs linked to poor nestling growth in a declining avian aerial insectivore (*Tachycineta*
758 *bicolor*) (2019). URL <https://doi.org/10.5061/dryad.7m41jd8>.
- 759 [72] DeSimone, J. G., Clotfelter, E. D., Black, E. C. & Knutie, S. A. Data from: Avoidance,
760
761
762
763
764
765
766

- 767 tolerance, and resistance to ectoparasites in nestling and adult tree swallows (2017).
768 URL <https://doi.org/10.5061/dryad.9bb60>.
- 769 [73] Dubuc-Messier, G. *et al.* Data from: Gene flow does not prevent personality and
770 morphological differentiation between two blue tit populations (2018). URL <https://doi.org/10.5061/dryad.31tc3s8>.
771
- 772 [74] Grunst, M. L., Raap, T., Grunst, A. S., Pinxten, R. & Eens, M. Data from: Artificial
773 light at night does not affect telomere shortening in a developing free-living songbird:
774 a field experiment (2019). URL <https://doi.org/10.5061/dryad.8216g63>.
- 775 [75] Husby, A., Schielzeth, H., Forstmeier, W., Gustafsson, L. & Qvarnström, A. Data
776 from: Sex chromosome linked genetic variance and the evolution of sexual dimorphism
777 of quantitative traits (2012). URL <https://doi.org/10.5061/dryad.451n7>.
- 778 [76] Ihle, M. *et al.* Data from: Rearing Success Does Not Improve With Apparent Pair
779 Coordination in Offspring Provisioning (2019). URL <https://zenodo.org/record/3459642>.
780
- 781 [77] Jacob, S. *et al.* Data from: Microbiome affects egg carotenoid investment, nestling
782 development and adult oxidative costs of reproduction in Great tits (2015). URL <https://doi.org/10.5061/dryad.9n741>.
783
- 784 [78] Krause, E. T., Krüger, O. & Schielzeth, H. Data from: Long-term effects of early
785 nutrition and environmental matching on developmental and personality traits in zebra
786 finches (2018). URL <https://doi.org/10.5061/dryad.6j700>.
- 787 [79] Krist, M., Janča, M., Edme, A. & Dzuro, R. Data from: Are prenatal maternal resources
788 more important in competitive than in benign postnatal environments? (2016). URL
789 <https://doi.org/10.5061/dryad.823f0>.
- 790 [80] Krist, M., Remeš, V., Uvírová, L., Nádvorník, P. & Bureš, S. Data from: Egg size and
791 offspring performance in the collared flycatcher (*Ficedula albicollis*): a within-clutch
792 approach (2010). URL <https://doi.org/10.5061/dryad.1758>.
- 793 [81] Kvalnes, T. *et al.* Data from: Offspring fitness and the optimal propagule size in a
794 fluctuating environment (2018). URL <https://doi.org/10.5061/dryad.m74c7m9>.
- 795 [82] Kvalnes, T. *et al.* Data from: Reversal of response to artificial selection on body size
796 in a wild passerine (2017). URL <https://doi.org/10.5061/dryad.v50r8>.
- 797 [83] Moiron, M. *et al.* Data from: Functional relations between body mass and risk-taking
798 behavior in wild great tits (2018). URL <https://doi.org/10.5061/dryad.14cn58v>.
- 799 [84] Nishida, Y. & Takagi, M. Data from: Song performance is a condition-dependent
800 dynamic trait honestly indicating the quality of paternal care in the Bull-headed Shrike
801 (2018). URL <https://doi.org/10.5061/dryad.c84f7c4>.
- 802 [85] Nord, A. & Nilsson, J.-A. Data from: Incubation temperature affects growth and
803 energy metabolism in blue tit nestlings (2011). URL <https://doi.org/10.5061/dryad.jb314>.
804

- 805 [86] Pap, P. L. *et al.* Data from: Selection on multiple sexual signals in two Central- and
806 Eastern-European populations of the barn swallow (2019). URL [https://doi.org/
807 10.5061/dryad.64p7k2f](https://doi.org/10.5061/dryad.64p7k2f).
- 808 [87] Perrier, C., Delahaie, B. & Charmantier, A. Data from: Heritability estimates from
809 genome wide relatedness matrices in wild populations: application to a passerine, using
810 a small sample size (2018). URL <https://doi.org/10.5061/dryad.k6r1mk8>.
- 811 [88] Podofillini, S. *et al.* Data from: Benefits of extra food to reproduction depend on
812 maternal condition (2020). URL <https://doi.org/10.5061/dryad.5db0168>.
- 813 [89] Poissant, J., Morrissey, M. B., Gosler, A. G., Slate, J. & Sheldon, B. C. Data from:
814 Multivariate selection and intersexual genetic constraints in a wild bird population
815 (2016). URL <https://doi.org/10.5061/dryad.qt745>.
- 816 [90] Poorboy, D. *et al.* Data from: Experimental cross-fostering of eggs reveals effects of
817 territory quality on reproductive allocation (2018). URL [https://doi.org/10.5061/
818 dryad.h8v8157](https://doi.org/10.5061/dryad.h8v8157).
- 819 [91] Rioux Paquette, S., Pelletier, F., Garant, D. & Bélisle, M. Data from: Severe recent
820 decrease of adult body mass in a declining insectivorous bird population (2014). URL
821 <https://doi.org/10.5061/dryad.67t23>.
- 822 [92] Sakaluk, S. K. *et al.* Data from: Genetic and environmental variation in condition,
823 cutaneous immunity, and haematocrit in house wrens (2014). URL [https://doi.
824 org/10.5061/dryad.jk2m0](https://doi.org/10.5061/dryad.jk2m0).
- 825 [93] Simpson, R. K. & McGraw, K. J. Data from: Multiple signaling in a variable
826 environment: expression of song and color traits as a function of ambient sound and
827 light (2017). URL <https://doi.org/10.5061/dryad.1j81k>.
- 828 [94] Song, Z. *et al.* Data from: Silver spoon effects of hatching order in an asynchronous
829 hatching bird (2018). URL <https://doi.org/10.5061/dryad.184c1dj>.
- 830 [95] Torres, R., Chin, E., Rampton, R. & Williams, T. D. Data from: Are there synergistic
831 or antagonistic effects of multiple maternally-derived egg components (antibodies and
832 testosterone) on offspring phenotype? (2019). URL [https://doi.org/10.5061/
833 dryad.j348s75](https://doi.org/10.5061/dryad.j348s75).
- 834 [96] Vermeulen, A., Müller, W. & Eens, M. Data from: Vitally important – does early
835 innate immunity predict recruitment and adult innate immunity? (2016). URL [https:
836 //doi.org/10.5061/dryad.p0s3g](https://doi.org/10.5061/dryad.p0s3g).
- 837 [97] Weber, B. M. *et al.* Data from: Pre- and post-natal effects of experimentally
838 manipulated maternal corticosterone on growth, stress reactivity, and survival of nestling
839 house wrens (2019). URL <https://doi.org/10.5061/dryad.16049f4>.
- 840 [98] Santiago-Alarcon, D. & Parker, P. G. Sexual Size Dimorphism and Morphological
841 Evidence Supporting the Recognition of two Subspecies in the Galápagos Dove. *The
842 Condor* **109**, 132–141 (2007).
- 843 [99] Santos, E. S. A. & Nakagawa, S. Breeding Biology and Variable Mating System of a

- 844 Population of Introduced Dunnocks (*Prunella modularis*) in New Zealand. *PLOS ONE*
845 **8**, e69329 (2013).
- 846 [100] Joanes, D. N. & Gill, C. A. Comparing measures of sample skewness and kurtosis.
847 *Journal of the Royal Statistical Society: Series D (The Statistician)* **47**, 183–189 (1998).
848 URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9884.00122>.
- 849 [101] Thomson, C. E. *et al.* Selection on parental performance opposes selection for larger
850 body size in a wild population of blue tits. *Evolution* **71**, 716–732 (2017).
- 851 [102] Thomson, C. E. & Hadfield, J. D. No evidence for sibling or parent–offspring
852 coadaptation in a wild population of blue tits, despite high power. *Evolution* **73**, 28–41
853 (2019).
- 854 [103] Baayen, R. H., Davidson, D. J. & Bates, D. M. Mixed-effects modeling with crossed
855 random effects for subjects and items. *Journal of memory and language* **59**, 390–412
856 (2008).
- 857 [104] Benjamin, D. J. *et al.* Redefine statistical significance. *Nature Human Behaviour* **2**,
858 6–10 (2018).
- 859 [105] Branco, M. D. & Dey, D. K. A general class of multivariate skew-elliptical distributions.
860 *Journal of Multivariate Analysis* **79**, 99–113 (2001).
- 861 [106] Azzalini, A. & Capitanio, A. Distributions generated by perturbation of symmetry with
862 emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society.*
863 *Series B: Statistical Methodology* **65**, 367–389 (2003).
- 864 [107] Arellano-Valle, R. B. & Azzalini, A. On the Unification of Families of Skew-normal
865 Distributions. *Scandinavian Journal of Statistics* **33**, 561–574 (2006).
- 866 [108] Azzalini, A. *The Skew-Normal and Related Families*. Institute of Mathematical Statistics
867 Monographs (Cambridge University Press, 2013).
- 868 [109] Barton, N. H., Etheridge, A. M. & Véber, A. The infinitesimal model: Definition,
869 derivation, and implications. *Theoretical Population Biology* **118**, 50–73 (2017).
- 870 [110] Quaas, R. L. & Pollak, E. J. Mixed model methodology for farm and ranch beef cattle
871 testing programs. *journal of Animal Science* **51**, 1277–1287 (1980).
- 872 [111] Carpenter, B. *et al.* Stan: A probabilistic programming language. *Journal of Statistical*
873 *Software* **76** (2017).
- 874 [112] Gelman, A., Rubin, D. B. *et al.* Inference from iterative simulation using multiple
875 sequences. *Statistical science* **7**, 457–472 (1992).
- 876 [113] Lewandowski, D., Kurowicka, D. & Joe, H. Generating random correlation matrices
877 based on vines and extended onion method. *Journal of multivariate analysis* **100**,
878 1989–2001 (2009).
- 879 [114] Janzen, F. J. & Stern, H. S. Logistic regression for empirical studies of multivariate
880 selection. *Evolution* **52**, 1564–1571 (1998).

⁸⁸¹ [115] Butler, D., Cullis, B. R., Gilmour, A. R., Gogel, B. J. & Thompson, R. *ASReml-R*
⁸⁸² *Reference Manual Version 4* (VSN International Ltd, Hemel Hempstead, UK, 2017).

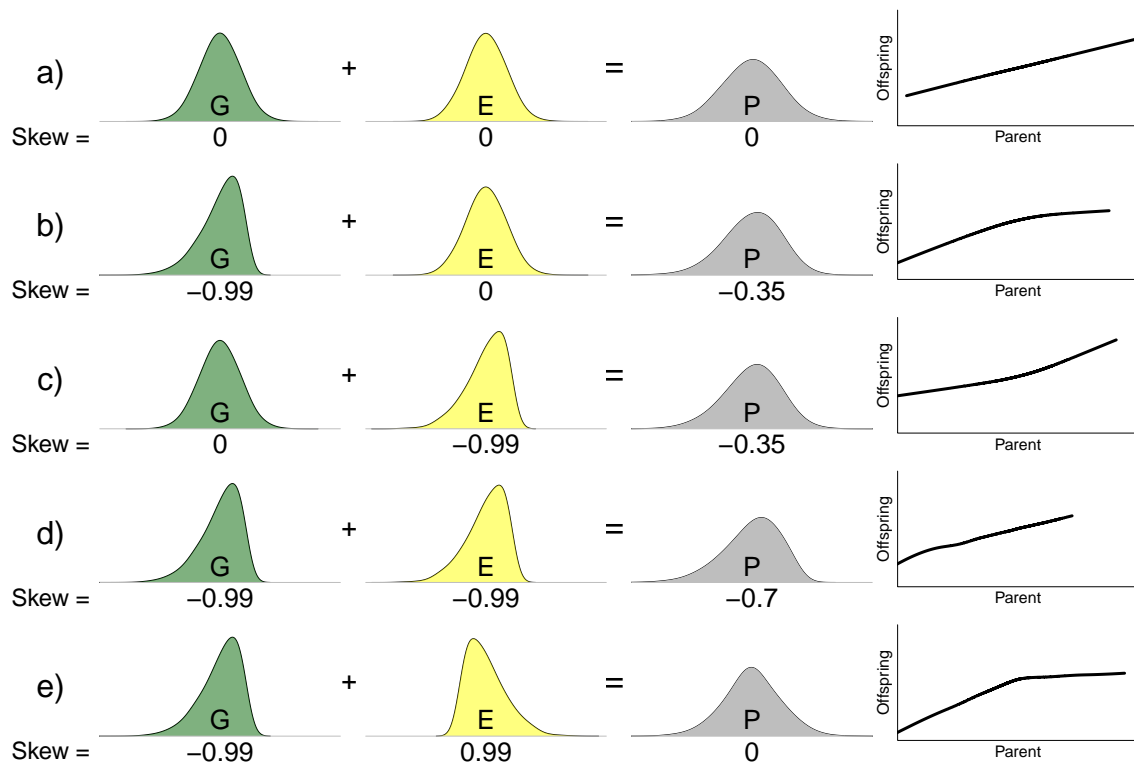


Figure 1: The effects of different distributions of breeding values (G) and environmental values (E) on the distribution of phenotypes (P) and the shape of the PO-regression. When both genetic and environmental values are normally distributed (a), as typically assumed, there is a linear PO-regression. Negative genetic (b) and environmental (c) skew affect the shape of the parent-offspring relationship in opposite directions, whilst inducing the same phenotypic skew. If genetic and environmental distributions are skewed in the same direction (d) their effects on the parent-offspring relationship can cancel each other out, giving a linear parent-offspring relationship, despite considerable phenotypic skew. If genetic and environmental are skewed in opposite directions (e), although they may cancel each other out at the phenotypic level, they induce a highly non-linear parent-offspring relationship. 1-5) are all simulated with a heritability (V_A/V_P) of 0.5.

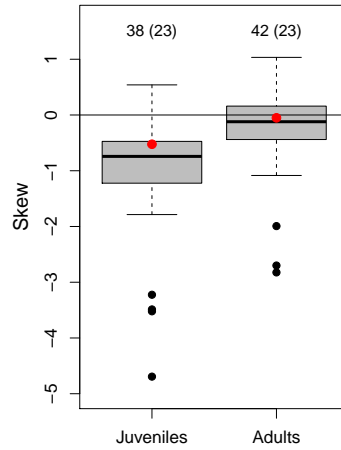


Figure 2: Skew in the distribution of avian tarsus lengths across different species, measured as the coefficient of skew. In the boxplots, the center line shows the median; box limits show upper and lower quartiles; whiskers show 1.5x interquartile range; points show outliers. Numbers above the plots show the number of estimates, and species in parenthesis. The red points show the skew in our blue tit data.

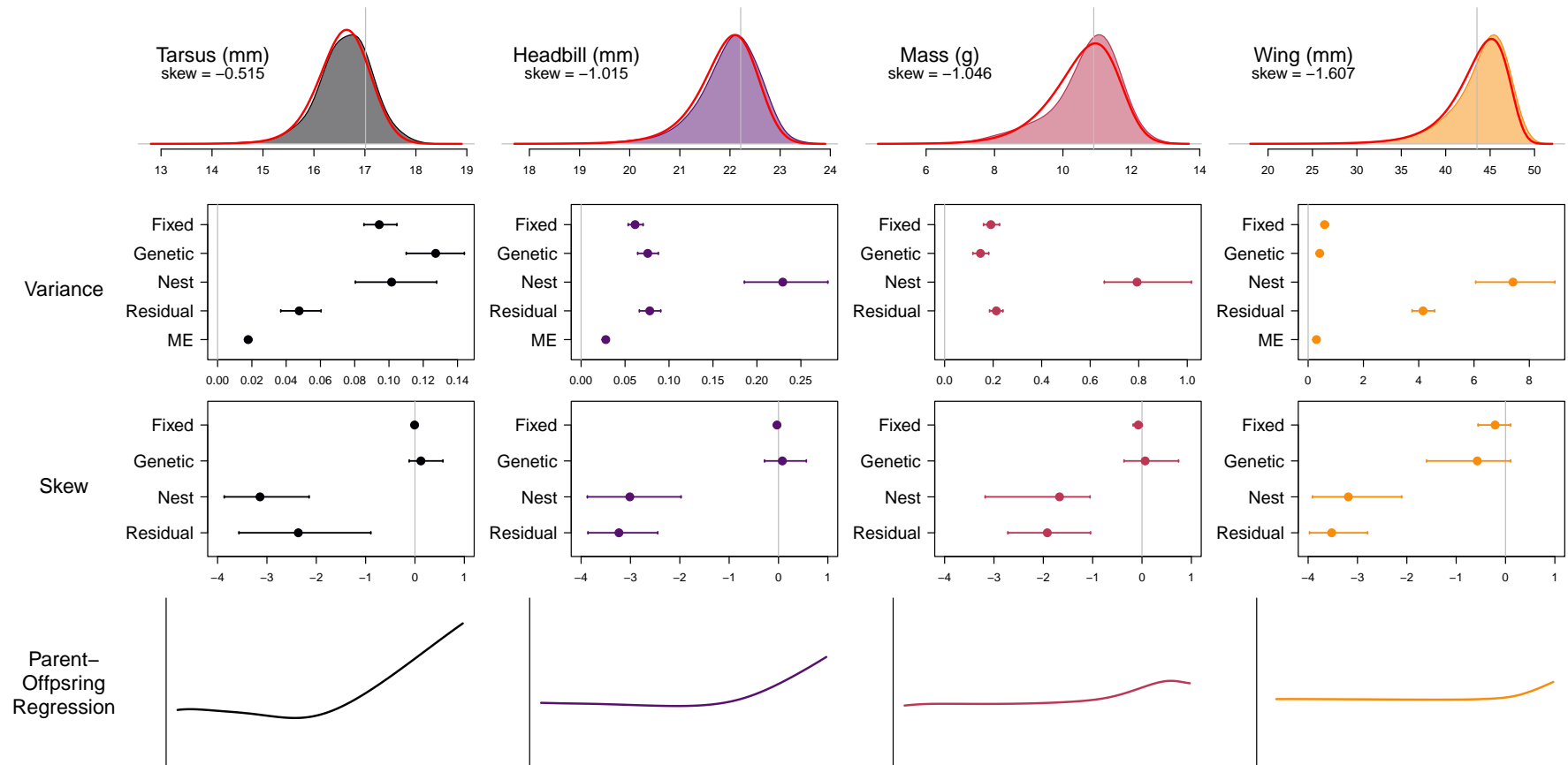


Figure 3: Decomposition of variance and skew in juvenile body size traits in blue tits. Top plots shows the phenotypic distribution of the traits, with the red line showing the distribution predicted from the skew models. The middle rows show the variance and skew (top and bottom, respectively) for each component for all four traits, with all model estimates coming from the skew-t animal model, except the genetic skew which was estimated in the skew-t dam-sire model (see methods). ME stands for measurement error. The bottom row shows the predicted shape of the PO-regression based on the model estimates.

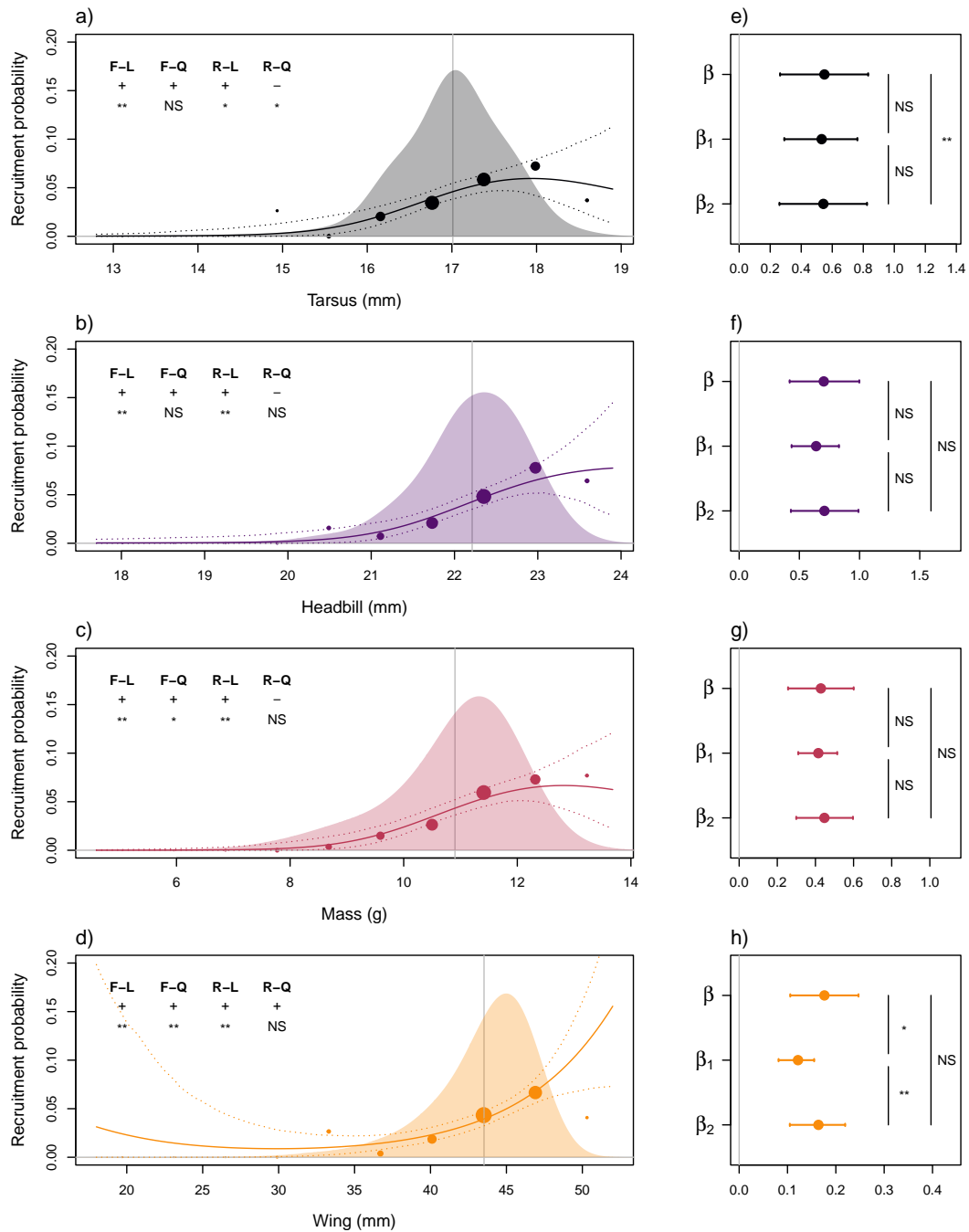


Figure 4: Average (over years and sexes) fitness functions (a-d) and selection gradients (e-h) for tarsus length, head-bill length, mass and wing length, respectively, from day 15 to recruitment. In plots a-d, solid lines show the posterior mean fitness functions, dotted lines show the 95% credible intervals, and points show the average survival of measured individuals from day 15 to recruitment in equally spaced intervals. The size of the points is proportional to the square root of the sample size. The phenotypic distribution of the traits is shown, with the grey vertical line showing the phenotypic mean. The direction and significance of the effect of the trait on fitness is also shown, 'F' and 'R' are survival from day 15 to fledging and from fledging to recruitment respectively, and 'L' and 'Q' and linear and quadratic effects. In plots e-h, β refers to the selection gradient derived from this fitness function, β_1 and β_2 refer to the approximations from the Lande-Arnold regression excluding and including a quadratic term, respectively. In all plots 'NS' indicates $p > 0.05$, '*' indicates $0.05 > p > 0.005$ and '**' $p < 0.005$.

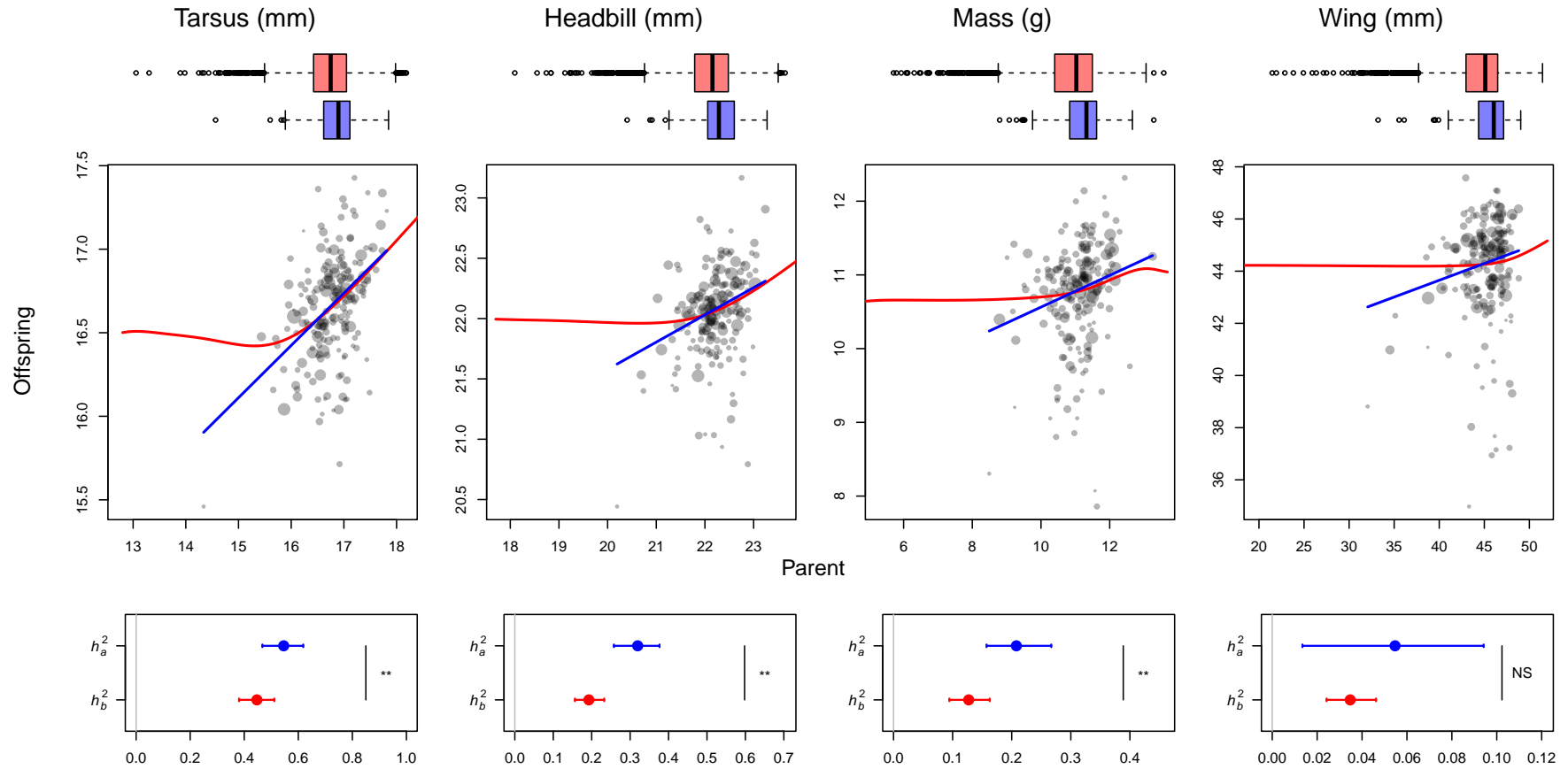


Figure 5: PO-regressions for four body size traits. Top panels show distribution of all chicks (red) and those that survived to recruit (blue), representing the distribution of potential parents before and after selection, respectively. Scatter plots show mid-offspring versus single parental traits. Values are corrected for year, sex and time of day at which they were measured, and the size of the points is proportional to the square root of the family size. The red line is the predicted non-linear PO-regression based on the posterior means of the parameters from the skew-t quantitative genetic model and the blue line is the fit of a weighted (by family size) linear regression to the actual data. These are not corrected for measurement error. Lower panels show the comparison between heritabilities calculated before (h_b^2) and after (h_a^2) selection, calculated across the posterior distribution of the skew-t animal model trait models. In these lower plots all heritabilities account for measurement error. In all plots 'NS' indicates $p > 0.05$, '*' indicates $0.05 > p > 0.005$ and '**' $p < 0.005$.

Table 1: Estimates (mean \pm SE) of heritability and additive genetic variance from PO-regression and Gaussian animal models (AM) across 2000 simulated data sets. Three-generation simulations were set up with either no cross-fostering (N) or with nests paired and half of each nest's offspring reciprocally crossed (X). Phenotypes were simulated according to the model estimated for chick mass exactly (skewed) or as Gaussian with matching variance. The probability of a chick recruiting to be a parent was obtained by applying the estimated survival model for chick mass.

	Simulated	Gaussian		Skewed	
		N	X	N	X
h^2 PO	0.138	0.139 \pm 0.001	0.140 \pm 0.001	0.223 \pm 0.001	0.224 \pm 0.001
h^2 AM	0.138	0.137 \pm 0.001	0.135 \pm 0.000	0.170 \pm 0.001	0.141 \pm 0.000
V_A AM	0.148	0.146 \pm 0.001	0.144 \pm 0.000	0.181 \pm 0.001	0.150 \pm 0.000