

Evolution in a Changing Environment: A Case Study with Great Tit Fledging Mass

Dany Garant,^{1,*} Loeske E. B. Kruuk,^{2,†} Robin H. McCleery,^{1,‡} and Ben C. Sheldon^{1,§}

1. Edward Grey Institute, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom;

2. Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Submitted January 14, 2004; Accepted July 2, 2004;

Electronically published September 22, 2004

ABSTRACT: Heritable phenotypic traits under significant and consistent directional selection often fail to show the expected evolutionary response. A potential explanation for this contradiction is that because environmental conditions change constantly, environmental change can mask an evolutionary response to selection. We combined an “animal model” analysis with 36 years of data from a long-term study of great tits (*Parus major*) to explore selection on and evolution of a morphological trait: body mass at fledging. We found significant heritability of this trait, but despite consistent positive directional selection on both the phenotypic and the additive genetic component of body mass, the population mean phenotypic value declined rather than increased over time. However, the mean breeding value for body mass at fledging increased over time, presumably in response to selection. We show that the divergence between the response to selection observed at the levels of genotype and phenotype can be explained by a change in environmental conditions over time, that is, related both to increased spring temperature before breeding and elevated population density. Our results support the suggestion that measuring phenotypes may not always give a reliable impression of evolutionary trajectories and that understanding patterns of phenotypic evolution in nature requires an understanding of how the environment has itself changed.

Keywords: natural selection, additive genetic variance, breeding values, fledging mass, environmental deviation, *Parus major*.

A major aim of evolutionary biology is to understand how the combination of genetic variation and natural selection produces evolutionary change and thus adaptation. If selection acts consistently on a heritable trait, the mean value of the trait is expected to change over time (Falconer and MacKay 1996; Lynch and Walsh 1998). However, several studies of natural populations have documented both directional selection and additive genetic variation for a given trait but have found no evidence of the expected microevolutionary change (sometimes even a change in the opposite direction) in the population mean phenotype over time (for a review, see Merilä et al. 2001b). These findings stand in contrast with the otherwise overwhelming evidence of natural selection and adaptation documented in the wild (for reviews, see Endler 1986; Kingsolver et al. 2001; Kinnison and Hendry 2001).

An implicit assumption of any study using long-term data sets to assess microevolutionary change over time is that the environment in which individuals are measured remains constant over the study period. However, Cooke et al. (1990) first suggested that if the environment changes at the same time as the presumed evolutionary changes (e.g., through negative density-dependent processes), then it is possible that a genetic response to selection on a given trait might be masked by opposing changes in the environmental component of that trait. Cooke et al. (1990) argued that this effect might be responsible for the lack of evolutionary response observed for clutch size in lesser snow geese (*Anser caerulescens*) despite positive directional selection and significant heritability. They suggested that if a given population were restricted to a limited area, then environmental factors such as negative density-dependent processes would tend to have a negative impact on the trait under investigation despite the fact that the mean breeding value (or additive genetic merit) of the trait were increasing. For instance, selection on a trait implies that some individuals with particular values of the trait have higher reproductive success and thus produce more offspring, which will tend to resemble their parents. These offspring will subsequently occupy mostly the same habitat and will be subject to the same form of selection as their

* E-mail: dany.garant@zoology.oxford.ac.uk.

† E-mail: loeske.kruuk@ed.ac.uk.

‡ E-mail: robin.mccleery@zoology.oxford.ac.uk.

§ E-mail: ben.sheldon@zoology.oxford.ac.uk.

parents were, given constant selection pressures. Then, as the population density rises, the portion of resources to which each individual will have access should decrease. This decline in the environmental component of the phenotype may reduce, cancel, or even reverse any tendency for the phenotype to increase because of an increase in the population's mean breeding value.

Despite the appeal of this "environmental deterioration" hypothesis to explain the lack of an apparent response to directional selection on heritable traits, no direct test of it has ever been conducted in a wild population. Several potential examples have been identified from long-term studies of bird populations. For example, in populations of snow geese *A. caerulescens* and barnacle geese *Branta leucopsis*, a decline in population mean body size despite opposing directional selection was suggested to be caused by increased competition for food, caused in turn by increasing density over the study period (Cooch et al. 1991; Larsson et al. 1998). Merilä et al. (2001a) suggested another case of environmental degradation opposing selection in a population of the collared flycatcher *Ficedula albicollis*, where they separated genetic and phenotypic change. In this case, a decline in residual fledging mass, despite selection in the opposite direction, was linked to indirect evidence that food for growing nestlings had been reduced by climatic change. However, to understand why a change in the environmental component of variation in a trait might have occurred, detailed data on potential processes such as population dynamics and climate underlying environmental variation are necessary.

In addition to the environmental deterioration model of Cooke et al. (1990), there are several other potential explanations for a mismatch between selection and evolutionary response in wild populations. First, selection may be associated with the nonheritable environmental component of the phenotype, in which case no evolutionary response at the phenotypic level is expected (Price et al. 1988; Alatalo et al. 1990; Rausher 1992; see Kruuk et al. 2002 for an example). Specifically, environmental factors that affect the development of individuals may have independent effects both on the expression of particular traits and on fitness. Under these circumstances, trait values will be phenotypically correlated with fitness, but there will be no genetic basis for the relationship, leading to an overestimate of the expected response to selection in the case of heritable traits (Rausher 1992; Scheiner et al. 2002; Stinchcombe et al. 2002; Kruuk et al. 2003). One way to overcome this problem is to estimate the force of selection acting directly on the additive genetic component of a given trait, quantified by individual breeding values for that trait (Rausher 1992). Second, selection may vary in time and/or in space, potentially reducing any phenotypic response (see Grant and Grant 2002 for an example). The

presence of negative genetic correlations between different components of fitness is a third mechanism that would restrict an evolutionary response (Roff 1997; see Sheldon et al. 2003 for an example), or the trait itself could be associated in an opposite manner with fitness at later stages. Further discussion of these and associated mechanisms can be found in Merilä et al. (2001b) and Kruuk (2004).

In this study, we provide an extensive analysis of patterns of selection on, and the genetic basis of, fledging mass in a population of great tits (*Parus major*). Analyses of heritability using traditional techniques have suggested that fledging mass should be heritable in great tits (van Noordwijk et al. 1988; Gosler and Harper 2000). Moreover, heavier body mass is associated with higher juvenile survival (Perrins 1965; Garnett 1981; Tinbergen and Boerlijst 1990; Both et al. 1999) and is highly correlated with mass at recruitment for both sexes in this species (Perrins and McCleery 2001). However, as we show here, this character has not shown the expected increase over time in our study population. Through a detailed analysis of the mechanisms of inheritance and of natural selection in a long-term study of a wild population, we aimed to explain this mismatch between observation and expectation using the framework above. We show ultimately that the divergence between the response to selection observed at the levels of genotype and phenotype can be explained by a change in environmental conditions over time, which underscores the importance of incorporating the effects of environmental variation when assessing evolutionary dynamics in the wild.

Material and Methods

Study Species and Data Collection

We used data obtained from the long-term study of the great tit population in Wytham Woods, Oxford, United Kingdom (see Perrins 1965, 1979 for more details). All breeding attempts are monitored from the date of egg laying until all nestlings had fledged. At 15 days old, nestlings were weighed (to nearest 0.1 g) and marked with individually numbered aluminium rings. At the same time, their parents were captured and their identity was checked (see Perrins 1979; Perrins and McCleery 2001 for more detailed information), and thus pedigrees for the quantitative genetic analysis were constructed on the assumption that the pair of birds attending a nest were the genetic parents of the nestlings in the nest. Extra-pair paternity in this population has been estimated to be 14% using two allozyme loci (Blakey 1994), although estimates using more reliable methods for other great tit populations suggest a lower figure (7.3%, the combined estimate from studies using DNA fingerprinting; for a review, see Griffith

et al. 2002). This suggests the presence of errors in the pedigree in father-offspring relationships that will tend to bias heritability and additive variance estimates downward. Unmarked birds that were first caught as breeding adults were defined as being immigrants to Wytham. Here we used data from 1965 to 2000 because nest box locations were kept constant from 1964 and because we were able to correct for the effect of laying date, clutch size, and egg weight for these years (all have significant effects on fledgling mass; results of general linear model including all three terms: lay date, $b = -0.095 \pm 0.009$, $t = 10.40$, $df = 38,020$, $P < .001$; clutch size, $b = -0.659 \pm 0.049$, $t = 13.48$, $df = 38,020$, $P < .001$; average egg weight of clutch, $b = 10.793 \pm 0.580$, $t = 18.61$, $df = 38,020$, $P < .001$). In all the analyses that follow, we use residual mass from a general linear model including lay date, clutch size, and egg weight as factors. Analyses using raw fledgling mass are unchanged as regards interpretation, but we present data using residuals as it controls for potentially confounding factors that are also under selection. Second clutches and repeat clutches laid after failure of the first clutch were removed from the data set (as in van Noordwijk et al. 1995). In total, data were available for 4,943 breeding attempts involving 38,024 chicks. There was a total of 59,386 individuals included in the pedigree with at least one lineage link; the maximum length of a lineage was 30 generations (mean \pm SE = 7.0 ± 1.3).

Quantitative Genetic Analysis

Variance Components and Breeding Value Estimation. Heritability of fledgling mass and individual breeding values were estimated through a mixed model restricted maximum likelihood (REML) estimation procedure using the software packages VCE4 (Variance Component Estimation, version 4; Neumaier and Groeneveld 1998) and PEST (Prediction and ESTimation; Groeneveld et al. 1992). We thus used the pedigree information to fit an individual "animal model" (Lynch and Walsh 1998), which partitions phenotypic variance in a quantitative character into its additive genetic and other fixed and random components such as common environment (Meyer 1989). Specifically, year of birth was included as a random effect, and the sector of the wood (one of the nine sections of Wytham Wood with different habitat characteristics; see Minot and Perrins 1986) was included as a fixed effect to account for temporal and spatial heterogeneity in environmental effects on the phenotype. The brood identity (combination of parental identities) was also fitted as a random effect to account for common-environment effects specific to the individual brood as well as any nongenetic maternal effects (Kruuk 2004). Partitioning the variance by including brood identity as a random effect just ensures that

significance tests are conducted against the proper residual variance and not inflated by between-brood variance. The narrow-sense heritability (h^2) was estimated as the ratio of the additive genetic variance (V_A) to the total phenotypic variance (V_P): $h^2 = V_A/V_P$. Best linear unbiased predictors of individual estimated breeding values (twice the expected effect of the genes that an individual passes on to its offspring or the total additive effect of its genes on that trait; Falconer and MacKay 1996) were quantified from pedigree information using REML estimates of variance components with the software package PEST. An individual's predicted breeding value is given as a linear sum of its own phenotype (corrected for other effects and scaled by the heritability) plus contributions from the breeding values of all of its relatives scaled by the relatedness of each relative to the focal animal (its relationship matrix). As such, two nonsurvivors in a nest will have exactly the same breeding value only if they have exactly the same phenotype (fledgling mass). Changes in estimated breeding values (EBVs) in different generations will reflect changes in additive genetic effects resulting from selection.

Genetic Covariances. We used a multivariate animal model to calculate the genetic correlation of body mass measured among temporal "sections" of the study to assess any potential genotype by environment interaction that could generate the appearance of a genotypic change and alter the expression of genetic variance in different environments. Specifically, a genetic correlation was estimated among the traits expressed in temporal segments of equal length in years (1965–1982 vs. 1983–2000). Here we defined traits measured in the first part of the study period versus those measured in the last part of the study as different "subtraits." Genetic correlations between these subtraits that are significantly less than unity indicate a genotype-environment interaction (Falconer and MacKay 1996), and this could be taken as evidence of a genetic constraint on evolutionary change as well as potentially confounding tests for genetic change assessed by trends in breeding values as above.

The genetic correlation between the sexes for body mass was also estimated among recruited individuals (the first point at which the sex was known for the majority of individuals). A strong genetic correlation between the sexes might constrain any evolutionary response, particularly if selection acted differently on the two sexes (see Merilä et al. 1997, 1998; Verboven and Visser 1998). In this case, the analysis was restricted to individuals that recruited to the breeding population as adults; year was included as a random effect, and sector and sex (known for individuals recaptured as adults) were included as fixed effects.

Standard errors for heritabilities and genetic correlations were computed by VCE4, and standard errors for variance

components were obtained from VCE3 (see also Neumaier and Groeneveld 1998).

Selection Analyses

Estimates of viability selection on phenotypic and estimated breeding values of fledging mass were based on recapture data under the assumption that nestlings not returning to the study area in subsequent years had died (survival = local survival). A proportion of juveniles disperse to breed elsewhere, but because we were interested in explaining the dynamics of a population, those individuals that disperse outside are effectively dead. Standardized selection differentials on phenotypic values were compared with those on standardized EBVs as a test of whether selection was associated with individual genotypes (see Scheiner et al. 2002; Stinchcombe et al. 2002; Kruuk et al. 2003). To test for temporal heterogeneity in selection, analyses were performed for each of the study years separately as well as for the data pooled over years. In each of the analyses, relative fitness (survivorship divided by mean survival rate) was calculated and fledging mass was standardized (0 mean, unit variance) within the data set used in the given analysis. Standardized directional (*i*) and quadratic (*j*) selection differentials (Endler 1986) were estimated using linear and second-order polynomial regressions, respectively, for relative survival on both standardized (0 mean, unit variance) phenotypic and breeding values of body mass. Statistical significance of the selection differentials was estimated with logistic regression (see Fairbairn and Preziosi 1996). Here, we used a logistic regression model with fixed (phenotype or genotypic values of body mass) and random effects that allowed us to take into account the nonindependence of fledglings within nest boxes by including brood identity as a random effect in the model (using a generalized linear mixed model [GLMM] with binomial errors; implemented in Genstat, ver. 6.2 [VSN 2002]). The significance of body mass as a predictor of survival (0 = died, 1 = survived) was assessed from its Wald statistic, which is distributed as a $\chi^2_{(1)}$ (VSN 2002).

We explored trends in selection differentials in cross-sectional analyses using linear regressions of annual means (for any mass-related measurements) weighted by the number of fledglings in a given year.

We also quantified the genetic correlation between body mass and adult fitness components using multivariate analyses, as in "Genetic Covariances," to explore selection on fledging mass via adult fitness components. Specifically, genetic correlations were estimated from multivariate analyses of body mass, life span (LSP; number of years from birth to last time caught) of recruited individuals, and lifetime reproductive success (LRS; number of offspring

who recruited into the breeding population) of recruits, where year was included as a random effect and sector as a fixed effect. Associations between adult LSP or adult LRS and EBVs for body mass were also tested using GLMMs with Poisson error structure in Genstat (VSN 2002). This analysis was performed to test for potential constraints acting through antagonistic effects on components of fitness in later, adult stages. Brood identity and year were included as random effects to account for repeated measures, and sector was included as a fixed effect in the model. The significance of the fixed effect of body mass breeding value as a predictor of LSP or LRS was assessed by its Wald statistic ($df = 1$).

Effect of Selection on Variance Components

We quantified the effect of viability selection on components of phenotypic variance in fledging mass by comparing the quantitative genetic estimates from two different sets of data. The first was preselection, using data on all nestlings in the database (with restrictions detailed above, $N = 38,024$), which estimated the different causal components of variance before most of the selection had taken place. The second was postselection, using a restricted data set consisting only of phenotypic measurements from those individuals that recruited to the breeding population (recruitment = local recruitment). Because the animal model estimates the components of variance in a base population (Lynch and Walsh 1998), any difference in the estimates from the two approaches will reflect the difference between the constitution of the total population and that of the subset that survived to become breeding adults. It therefore provides an estimate of the average effect (across all years) on the variance components of selection (for examples, see Merilä et al. 2001c; Kruuk et al. 2001) but allows for the fact that as a result of continuous immigration into the study area, different generations are distributed at different times throughout the pedigree. Variance due to environmental effects was defined as the sum of brood, year, and residual components of the phenotypic variance, and the environmental deviations calculated (see next section) represent the sum of these effects at the individual level.

Calculation of the Environmental Deviation

We calculated an "environmental deviation" for each individual by regressing phenotypic fledging mass on estimated breeding value. Thus, the estimation of each individual environmental deviation is made from the residual of each individual phenotype on its breeding value. The mean of all individual environmental deviations (residuals) can then be averaged for each year, and the

pattern of these means can then be compared over time or in relation to measurable environmental variables such as population density and temperature (see next section). For example, when the annual environmental deviation is negative in a given year or with respect to temperature or density, it is because the individual phenotypic values of fledging mass, on average, are smaller than expected, based on individual breeding values for that year (i.e., fledglings had a smaller body mass than expected based on their genotypes; see Kruuk et al. 2002; Sheldon et al. 2003 for other examples of analyses involving the environmental deviation).

Density- and Temperature-Dependent Processes

Overall breeding density was established as the total number of both great tit and blue tit (*Parus caeruleus*) pairs breeding in the wood in each year because both species compete for the same nest sites and for food (see Minot 1981; Minot and Perrins 1986). Temperature data were obtained from the School of Geography, Radcliffe Meteorological Station in Oxford, about 5 km east of Wytham Wood. We used warmth sum (defined as the sum of daily maximum temperature from March 1 to April 25; see McCleery and Perrins 1998) as a measure of spring temperature. Density and temperature are not necessarily independent because temperature could potentially affect the density of the population. We thus used a path analysis (Mitchell 1992; Shipley 2000) and the standardized annual mean values of each parameter over the 36-year period to evaluate the effects of temperature and density on the mean environmental deviation (see "Calculation of the Environmental Deviation") of fledging mass. The analysis was run using the SEPATH module in Statistica package, which allows a maximum likelihood estimation of the discrepancy function (the function that is minimized to yield parameter estimates; Statsoft 1999).

We hypothesized that the path model shown in figure 1A would provide an explanation of annual variation in mean environmental deviation. The direct effect of density in the current year was included as a path of direct causation because there is evidence that this parameter restricts food and territory availability through intra- and interspecific competition (Minot 1981). Direct effects of temperature in the previous year ($T-1$) and in the current year (T) were included because temperature has a clear impact on reproductive timing and on the amount of the main food resources available for parents to feed their offspring (Perrins 1991; Buse et al. 1999). The indirect effect of temperature on density at $T-1$ and T was also included because a warmer year is potentially beneficial for survival (Cresswell and McCleery 2003). The only path of direct causation excluded was the direct influence of

density in the previous year ($T-1$) on the environmental deviation. Including this factor as a direct determinant of environmental deviation yielded a model with a higher Akaike Information Criterion (Burnham and Anderson 1998) when testing for the best subset determining environmental deviation (established from the All Subsets Regression procedure for linear models implemented in Genstat) and also in the path model. It was, however, kept in the model as an indirect determinant of environmental deviation because of its indirect effect on density in the following year (fig. 1A).

Results

Heritability and Variance Components

The means and standard deviations of uncorrected fledging mass for pre- and postselection individuals are given in table 1. We found a significant heritability of body mass (0.239 ± 0.015 ; table 1). The largest component of the phenotypic variance in the preselection sample was that attributable to brood identity (common environment), which accounted for 41.0% of the variance (table 1). Selection significantly reduced both additive (-23.5%) and environmental (-39.2% ; $V_E = \text{year} + \text{brood} + \text{residual}$) variance. Consequently, heritability did not differ in the postselection and preselection samples ($t = 1.47$, $df = 41,714$, $P = .143$).

Selection Differentials

Across the whole study period, there was evidence of strong directional selection on body mass, with a mean standardized selection differential of 0.209 ± 0.016 ($t = 12.82$, $df = 35$, $P < .001$; see table 2). Thirty-five out of 36 standardized selection differentials on fledging mass for a given year were positive, with 17 of them being significantly >0 ($P < .05$). Natural selection on EBVs was also significantly directional, with a mean selection differential across the whole study period of 0.141 ± 0.016 ($t = 8.58$, $df = 35$, $P < .001$; see table 2); there were 33 positive and 11 significant standardized selection differentials. Overall directional selection differentials calculated with nonstandardized phenotypic and breeding values (Stinchcombe et al. 2002) were also significant (phenotype = 0.196 ± 0.017 , breeding value = 0.491 ± 0.059 , both $P < .001$). The relationships between the linear selection differentials on phenotype and on breeding value in each year were strongly positive ($r = 0.88$, $P < .001$). There was also evidence for weak stabilizing selection on the phenotypic values in the estimate combined over all years, but the quadratic terms were statistically significant in only two out of 36 years (table 2).

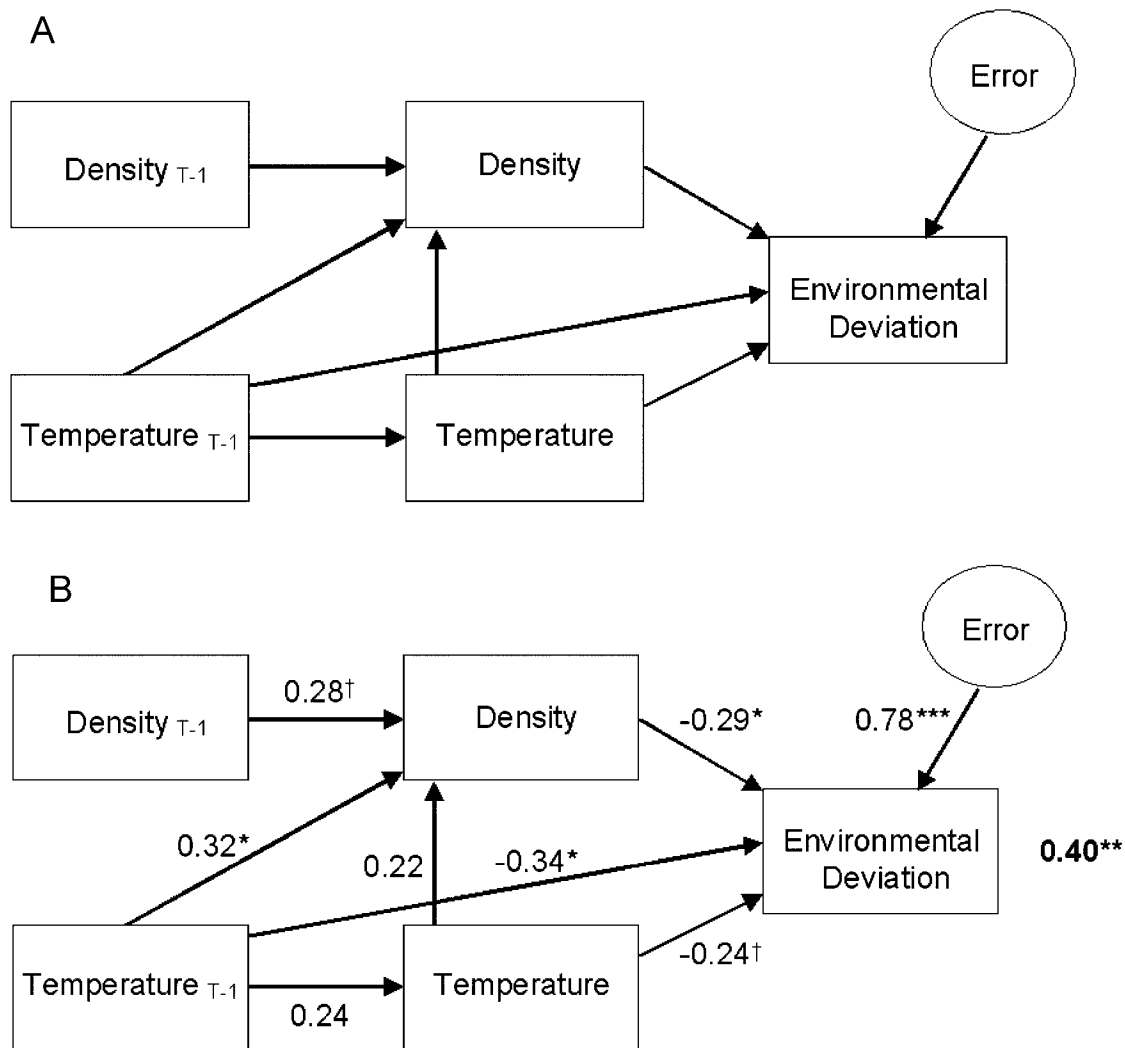


Figure 1: Hypothesized (A) and estimated (B) relationships between temperature, breeding density, and environmental deviation (difference between genotypic and phenotypic components) of fledging mass in great tits. Standardized partial regression coefficients obtained from the path analysis are shown. $T-1$ = previous year. $\dagger = P < .10$, $* = P < .05$, $** = P < .01$, $*** = P < .01$. Value in boldface indicates the squared multiple correlation coefficient of the model.

There was no evidence of temporal fluctuations in the direction of linear selection because none of the selection differentials for either phenotype or genotype was significantly negative (table 2). There also was no suggestion of any consistent change in the strength of selection on phenotype in relation to time ($b = 0.002 \pm 0.003$, $t = 0.68$, $df = 34$, $P = .502$). Selection differentials, however, were greater in years in which the mean body mass was lower than average ($b = -0.224 \pm 0.092$, $t = 2.45$, $df = 34$, $P = .020$), and they tended to be greater when the proportion of individuals surviving to recruit was smaller ($b = -1.127 \pm 0.598$, $t = 1.88$, $df = 34$, $P = .068$). Increased breeding density was associated with increasing

selection differentials ($b = 0.0005 \pm 0.0002$, $t = 2.03$, $df = 34$, $P = .050$) and also had a negative effect on fledging mass ($b = -0.0011 \pm 0.0004$, $t = -2.62$, $df = 34$, $P = .013$). However, there was no association between selection and density after taking account of differences in mass; including density and fledging mass in a multiple linear regression of selection differentials showed that the relationships between mean fledging mass and the strength of selection remained significant ($b = -0.213 \pm 0.102$, $t = -2.10$, $df = 33$, $P = .044$) while the effect of density was no longer significant ($b = 0.0001 \pm 0.0003$, $t = 0.26$, $df = 33$, $P = .796$). Temperature had no significant effect on selection differentials (results not shown).

Table 1: Components of phenotypic variance in fledgling mass (g) in preselection (all fledglings) and postselection (only survivors) samples

| Source | Preselection | | Postselection | | Change (%) |
|---------------------|---------------|---------|---------------|---------|------------|
| | Var (SE) | Var (%) | Var (SE) | Var (%) | |
| Additive | .255 (.017)** | 23.90 | .195 (.019)** | 28.30 | −23.5* |
| Year | .080 (.015)** | 7.50 | .043 (.012)** | 6.24 | −46.3* |
| Brood | .437 (.009)** | 40.96 | .210 (.012)** | 30.48 | −52.0** |
| Residual | .295 (.009)** | 27.64 | .241 (.013)** | 34.98 | −18.3** |
| V_p | 1.067 | | .689 | | −35.4 |
| h^2 (SE) | .239 (.015)** | | .283 (.026)** | | +18.4 |
| N (nests) | 4,943 | | 2,375 | | |
| N (nestlings) | 38,024 | | 3,692 | | |
| Mean body mass (SD) | 18.65 (1.46) | | 18.99 (1.18) | | +1.8** |

Note: Var = absolute; Var % = relative contributions of components of variance; Change = (after − before)/before \times 100.

* Significance of the variance components and of the change in sources of total phenotypic variance between pre- and postselection values is based on a t -test; $P < .05$.

** Significance of the variance components and of the change in sources of total phenotypic variance between pre- and postselection values is based on a t -test; $P < .001$.

Phenotypic and Genotypic Response over Time

Despite the positive directional selection on phenotype demonstrated above, we found a significant decrease in the annual mean phenotype over time (fig. 2; $b = -0.009 \pm 0.004$, $t = -2.24$, $df = 34$, $P = .032$; uncorrected mean fledgling mass from 1965 to 2000: $b = -0.013 \pm 0.006$, $t = -2.26$, $df = 34$, $P = .030$). Results of a linear mixed model (LMM) of individual values, controlling for parental pair identity and nonindependence of observations from the same year, confirmed this relationship ($b = -0.010 \pm 0.004$, $t = 2.47$, $df = 38,024$, $P = .013$). Standard quantitative genetic theory predicts that directional selection should induce a corresponding change in the trait of magnitude $R = h^2 i \sigma_p$ in one generation, where h^2 is the heritability, i is the standardized selection differential, and σ_p is the standard deviation of the phenotype (assuming that the trait is the direct target of selection and that all other genetically correlated traits are not subject to any direct selection; Cameron 1997). Thus, the expected phenotypic change given the mean values calculated here is $+0.05$ ($= 0.239 \times 0.209 \times 1.00$ SD) phenotypic standard deviations per generation, giving an expected change of $+0.027$ g year $^{-1}$ (given that one generation = 1.85 year, as estimated from Charlesworth 1994). The estimated rate of change (-0.009 ± 0.004 g year $^{-1}$) thus differs significantly from the predicted rate of $+0.027$ ($t = 9.00$, $df = 34$, $P < .001$; comparisons using uncorrected phenotypes produce the same results). Hence, the predicted change in mean phenotype is not occurring.

In contrast to the phenotypic trend, we found a highly significant increase in the mean EBVs over the study period

($b = 0.002 \pm 0.0003$, $t = 7.01$, $df = 34$, $P < .001$; LMM: $b = 0.002 \pm 0.0003$, $t = 5.51$, $df = 38,024$, $P < .001$; fig. 3). The observed increase in breeding value over time is unlikely to be due to immigration because the EBV of immigrants ($N = 3,349$) showed no tendency to change over time ($b = -0.0001 \pm 0.0003$, $t = -0.43$, $df = 34$, $P = .669$). Further, the mean EBV of immigrants in each year was lower than that of residents (mean \pm SE: immigrants = 0.013 ± 0.003 , residents = 0.103 ± 0.009 ; paired t -test, $t = 9.48$, $df = 35$, $P < .001$), and the proportion of immigrants did not change over time ($b = -0.0004 \pm 0.002$, $t = -0.24$, $df = 34$, $P = .812$).

Environmental Variance Reduction: Intra-brood Component

Comparison of pre- and postselection samples showed that the component of variance subject to the largest reduction in variance was the one corresponding to brood identity (table 1). We therefore further assessed how this component of variance was affected by ranking each individual that survived until day 15 within a nest according to its mass, and we estimated the probability of survival to recruitment according to each individual's rank within the nest. We found that there was a significant effect of individual rank on survival probability (GLMM with year and brood identity as random effect and sector as a fixed effect: $\chi^2_{(13)} = 34.59$, $P < .001$). Higher-ranked individuals had a higher probability of survival (Spearman rank correlation of predicted survival probabilities on rank within nest [limited to the 10 first-ranked chicks per nest] obtained from GLMM: $r_s = -0.867$, $P = .001$; fig. 4), and thus selection was acting

Table 2: Standardized directional (*i*) and nonlinear (*j*) selection differentials on phenotype and breeding values of fledgling mass for survival to recruitment

| Year | N survivors | N total | Phenotype | | Breeding value | |
|------|----------------|------------|----------------|-----------------|----------------|----------------|
| | | | <i>i</i> ± SE | <i>j</i> ± SE | <i>i</i> ± SE | <i>j</i> ± SE |
| 1965 | 31 | 611 | .095 ± .175 | −.246 ± .188 | .012 ± .175 | .020 ± .202 |
| 1966 | 17 | 580 | .405 ± .239 | −.242 ± .294 | .380 ± .239 | −.138 ± .276 |
| 1967 | 33 | 481 | .060 ± .168 | −.228 ± .198 | .102 ± .168 | −.178 ± .202 |
| 1968 | 31 | 294 | .019 ± .171 | −.264 ± .178 | .141 ± .171 | −.100 ± .154 |
| 1969 | 18 | 296 | −.208 ± .229 | .070 ± .280 | −.291 ± .229 | .000 ± .302 |
| 1970 | 24 | 137 | .164 ± .187 | −.058 ± .292 | .013 ± .187 | .080 ± .290 |
| 1971 | 35 | 539 | .361 ± .163* | .058 ± .200 | .388 ± .163* | .266 ± .190 |
| 1972 | 45 | 526 | .040 ± .143 | −.158 ± .144 | .076 ± .143 | −.224 ± .132** |
| 1973 | 31 | 769 | .361 ± .176* | .014 ± .224 | .238 ± .176 | −.034 ± .224 |
| 1974 | 50 | 508 | .161 ± .134 | −.054 ± .132 | .226 ± .134 | .092 ± .150 |
| 1975 | 18 | 511 | .536 ± .231* | .220 ± .252 | .438 ± .231* | .306 ± .248 |
| 1976 | 92 | 527 | .184 ± .095 | .010 ± .112 | .159 ± .095 | .286 ± .130 |
| 1977 | 68 | 839 | .228 ± .116* | .036 ± .126 | .232 ± .116* | .022 ± .122 |
| 1978 | 76 | 1,107 | .071 ± .111 | −.060 ± .120 | .064 ± .111 | .270 ± .138 |
| 1979 | 128 | 979 | .056 ± .083 | −.018 ± .096 | .0002 ± .0825 | .020 ± .108 |
| 1980 | 181 | 1,412 | .008 ± .070 | −.092 ± .064 | .064 ± .070 | .040 ± .072 |
| 1981 | 83 | 1,463 | .222 ± .107* | −.112 ± .114 | −.021 ± .107 | .090 ± .130 |
| 1982 | 211 | 1,184 | .244 ± .062*** | −.046 ± .68 | .093 ± .062 | .130 ± .076 |
| 1983 | 139 | 2,091 | .066 ± .082 | −.130 ± .96 | .019 ± .082 | −.068 ± .096 |
| 1984 | 57 | 1,115 | .296 ± .129* | .112 ± .142 | .327 ± .129* | .102 ± .144 |
| 1985 | 127 | 1,166 | .288 ± .083*** | .46 ± .92 | .233 ± .084** | .022 ± .094 |
| 1986 | 143 | 1,167 | .034 ± .078 | −.024 ± .082 | .068 ± .078 | .098 ± .088 |
| 1987 | 234 | 1,862 | .149 ± .061* | −.066 ± .066 | .112 ± .061 | −.050 ± .074 |
| 1988 | 73 | 1,215 | .506 ± .113*** | .088 ± .150 | .359 ± .113** | .088 ± .134 |
| 1989 | 144 | 1,233 | .353 ± .078*** | −.036 ± .092 | .250 ± .078** | .096 ± .100 |
| 1990 | 58 | 1,108 | .187 ± .128 | −.190 ± .156 | .213 ± .091 | .130 ± .180 |
| 1991 | 86 | 581 | .332 ± .099** | −.016 ± .116 | .348 ± .099** | .044 ± .114 |
| 1992 | 175 | 1,283 | .136 ± .070 | −.118 ± .074 | .097 ± .070 | −.048 ± .086 |
| 1993 | 225 | 1,934 | .109 ± .063 | −.052 ± .070 | .007 ± .063 | .0108 ± .076 |
| 1994 | 136 | 1,624 | .331 ± .082*** | −.058 ± .096 | .204 ± .082* | .222 ± .102 |
| 1995 | 245 | 1,663 | .271 ± .059*** | −.004 ± .062 | .249 ± .059*** | .160 ± .070 |
| 1996 | 149 | 1,472 | .033 ± .078 | −.198 ± .088* | .018 ± .078 | −.116 ± .088 |
| 1997 | 179 | 1,923 | .155 ± .071* | −.048 ± .082 | .093 ± .071 | −.032 ± .084 |
| 1998 | 105 | 1,598 | .280 ± .094** | .016 ± .110 | .177 ± .094 | .196 ± .112 |
| 1999 | 61 | 1,059 | .438 ± .124*** | .100 ± .164 | .385 ± .124** | .286 ± .146 |
| 2000 | 184 | 1,167 | .071 ± .068 | −.126 ± .062* | −.0004 ± .0677 | −.062 ± .066 |
| All | 3,692 | 38,024 | .209 ± .016*** | −.054 ± .018*** | .141 ± .016*** | .046 ± .018 |

Note: Significance of fledgling mass or breeding value on survival (0 or 1) was tested with a generalized linear mixed model (binomial error) including brood identity as random effect.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

to remove the smallest individuals within each nest. Furthermore, we established that the mean environmental deviation of individuals that were not recaptured as adults was significantly lower (i.e., they had a smaller body mass than expected based on their breeding value) than individuals that were known to have survived; we tested with a linear mixed model including year and brood identity as random effect and sector as a fixed effect (av-

erage difference between nonsurvivors and survivors: -0.009 ± 0.004 ; $\chi^2_{(1)} = 5.56$, $P = .018$). Thus, selection was also removing the individuals with the largest negative environmental deviations. This pattern is unlikely to be caused by nonrandom emigration from Wytham because the probability that a bird hatched in Wytham dispersed to the surrounding area is known to be independent of its mass as a nestling (Verhulst et al. 1997).

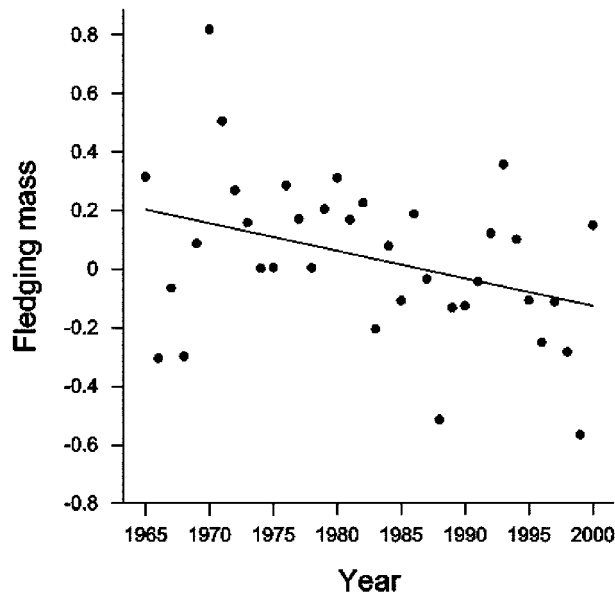


Figure 2: Change in mean fledging mass over time in the great tit population from 1965 to 2000. Mean residual values are weighted by the number of fledglings present in each year. The phenotypic decrease over time is significant ($P = .032$).

Genetic Correlations

GLMM analyses revealed that both LSP and LRS were positively associated with breeding values for fledging mass when including all measured nestlings (LSP: $\chi^2_{(1)} = 54.91$, $P < .001$; LRS: $\chi^2_{(1)} = 33.25$, $P < .001$). The genetic correlations (\pm SE) between fledging mass and LSP and between mass and LRS among survivors were 0.050 ± 0.212 and -0.047 ± 0.135 , respectively, and thus neither of these correlations was significantly different from 0. The genetic correlation for fledging mass between the sexes was equal to 1.000 ± 0.001 ($P < .001$, thus significantly different from 0), implying no differences between the sexes in the genetic architecture for this character. The genetic correlation between temporal segments with equal numbers of years was also equal to 1.0 ($P < .001$ again, significantly different from 0) implying that there is little likelihood that a genotype-by-environment interaction has generated the observed increase in breeding value over time.

Density- and Temperature-Dependent Processes

Population breeding density (great tits: mean \pm SD = 233 ± 73 ; blue tits: 334 ± 63 ; combined: 567 ± 103 pairs of birds) increased from 1965 to 2000 (great tits: $b = 4.54 \pm 0.89$, $t = 5.08$, $df = 34$, $P < .001$; blue tits: $b = 2.45 \pm 0.93$, $t = 2.62$, $df = 34$, $P = .013$; combined:

$b = 6.99 \pm 1.18$, $t = 5.92$, $df = 34$, $P < .001$; fig. 5A). Temperature also increased over the same period ($b = 2.506 \pm 0.758$, $t = 3.30$, $df = 34$, $P = .002$; fig. 5B). Results of the path analysis (see fig. 1B) showed a significant direct negative effect of breeding density (fig. 6A) and of temperature the previous year (fig. 6B) and a marginally nonsignificant negative effect of temperature in the current year on the mean environmental deviation for body mass in a given year. The indirect effect of a warmer temperature in the previous year was to increase breeding density the next year, and this indirect effect thus resulted in a negative influence on the environmental deviation (standardized indirect effect: $0.32 \times -0.29 = -0.09$; see fig. 1B). The indirect effect of density in the previous year on density in the current year also had a resulting negative influence on environmental deviation ($0.28 \times -0.29 = -0.08$). The overall squared multiple correlation of the path model was significant ($r^2 = 0.40 \pm 0.12$, $P = .002$).

Discussion

We used a long-term data set to try to understand the evolutionary trajectory of fledging mass in a wild great tit population. We found that mean fledging mass had declined over a 36-year period despite positive directional selection and a significant heritable component. Hence, the situation seems similar to that observed in the case of

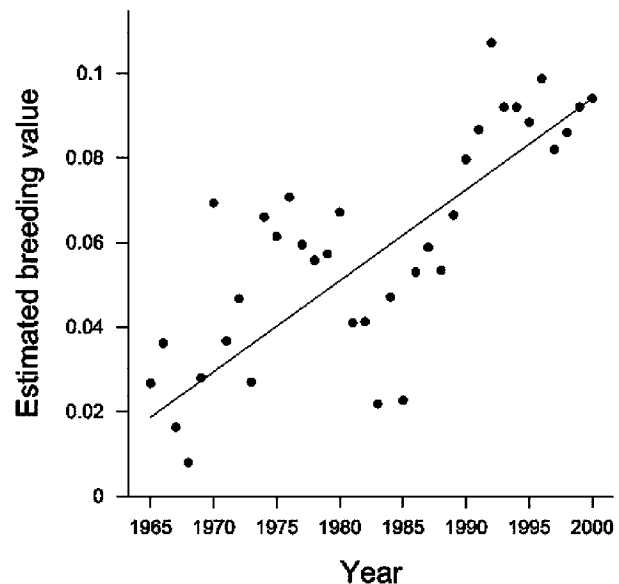


Figure 3: Change of mean estimated breeding value of fledging mass over time in the great tit population from 1965 to 2000. Values are weighted by the number of observations in each year. The genotypic increase over time is highly significant ($P < .001$).

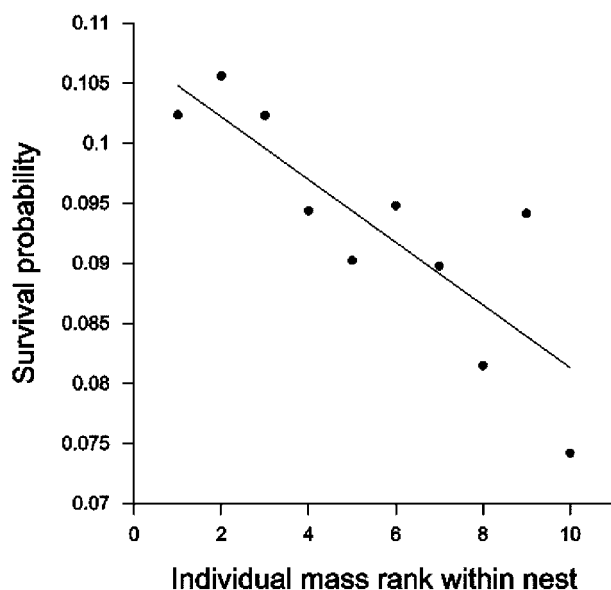


Figure 4: Predicted survival probabilities (obtained from a generalized linear mixed model) depending on rank of an individual within a nest for the 10 first-ranked chicks per nest (Spearman rank correlation: $r_s = -0.867$, $P = .001$).

body size of lesser snow geese (*Anser caerulescens*; Cooch et al. 1991), barnacle geese (*Branta leucopsis*; Larsson et al. 1998), and collared flycatchers (*Ficedula albicollis*; Merilä et al. 2001a). However, estimation of quantitative genetic parameters combined with selection analyses suggests that the additive genetic component of the phenotype has increased over time and has thus responded to selection (as with Merilä et al. 2001a). By combining information about changes in both population breeding density and the early spring temperature over time, we were able to show that the combined action of these two processes can explain a large proportion of the difference between the phenotypic and genotypic responses in this population.

Both population density and the prebreeding spring temperature had increased over the study period, and the combination of both parameters explained a significant proportion of the change in the environmental deviation of the phenotype. The framework suggested by Cooke et al. (1990) therefore proved useful in explaining the departure of the phenotype from the response of the breeding value in wild population. Hence, as argued by Cooke et al. (1990), changes in environmental conditions over time may indeed act to conceal evolution, and as our results suggest, such factors should be taken into account in future studies aiming to predict phenotypic evolutionary responses in the wild. The situation we documented here is also analogous to those often documented in studies of

geographic differentiation called “counter-gradient variation” (Merilä et al. 2001a, 2001b), which is defined as a negative covariance between environmental and genetic influences on a trait that hides spatial genetic divergence among populations (Conover and Schultz 1995; for examples, see Parsons 1997; Laugen et al. 2003). In our study,

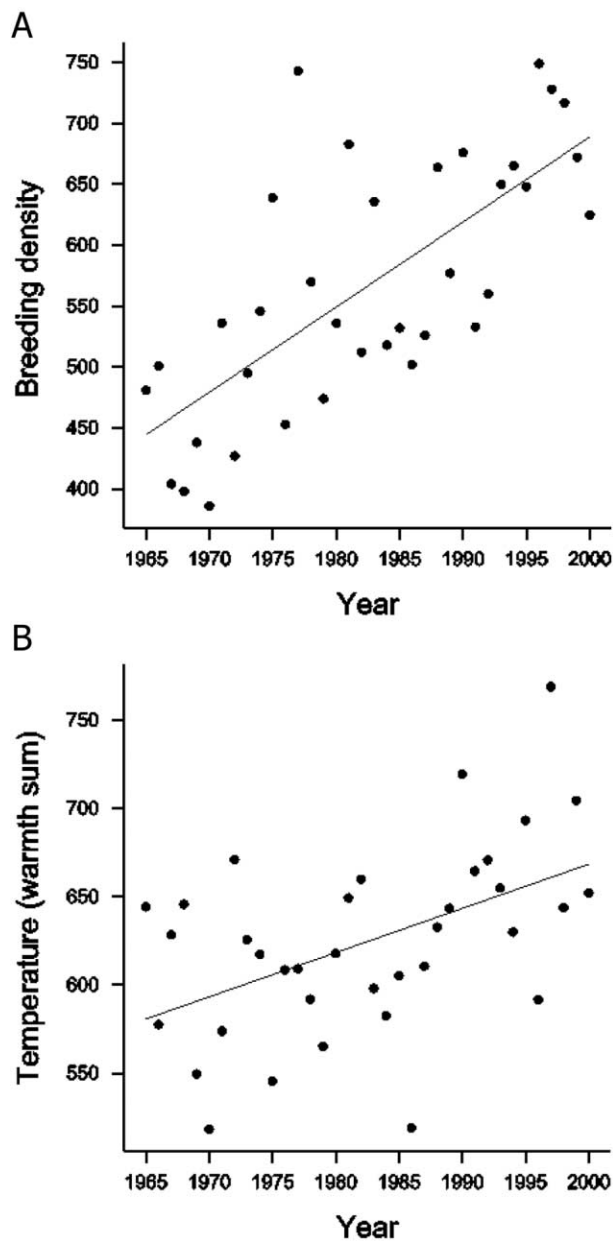


Figure 5: Temporal variation in overall breeding density (total number of pairs of great and blue tits; A) and early spring temperature (B) in Wytham Wood. Both breeding density ($b = 6.99 \pm 1.18$, $t = 5.92$, $df = 34$, $P < .001$) and temperature ($b = 2.506 \pm 0.758$, $t = 3.30$, $df = 34$, $P = .002$) increased significantly over the study period.

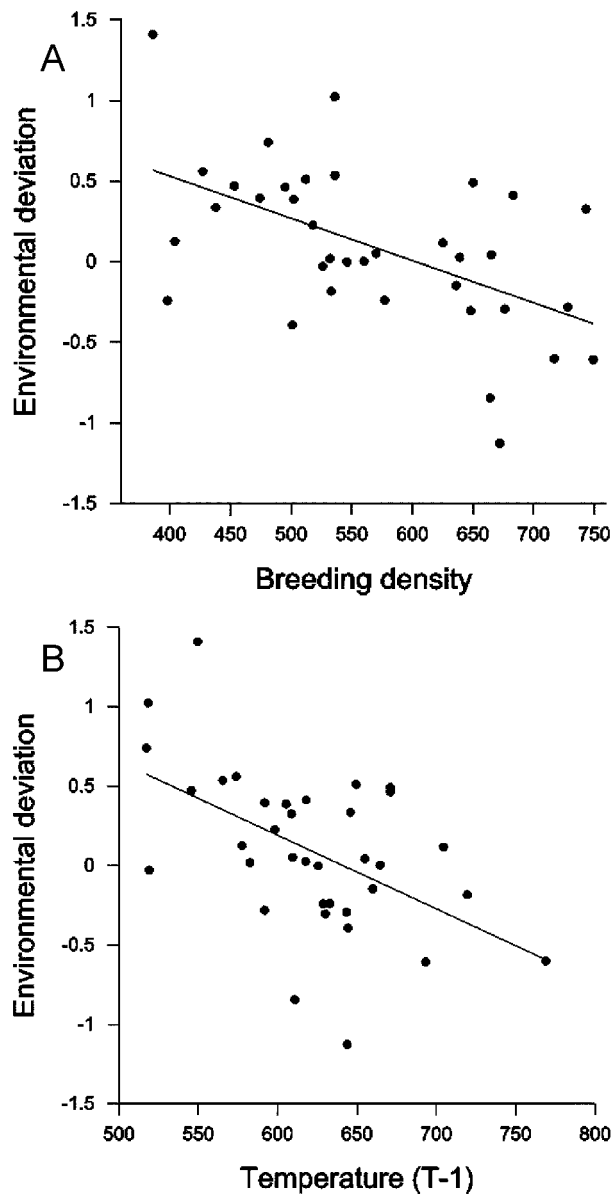


Figure 6: Linear regression of the effects of breeding density (total number of pair of great tits and blue tits; A) and spring temperature in the previous year ($T - 1$; B) on the environmental deviations. Both breeding density ($b = -0.0026 \pm 0.0007$, $t = -3.72$, $df = 34$, $P < .001$) and temperature in the previous year ($b = -0.0046 \pm 0.0013$, $t = -3.52$, $df = 34$, $P = .001$) had a significant negative influence on the environmental component of the phenotype.

changes in environmental conditions increased the selection on a trait and thus augmented the expected micro-evolutionary response but at the same time changed the environmental deviation in the opposite direction, thus reducing the chances of the change being detectable at the phenotypic level.

The negative effect of density on the environmental deviation for fledgling body mass documented here probably has a straightforward explanation because an increase in overall population breeding density will reduce food and territory availability (Minot 1981; Minot and Perrins 1986), which will result in reduced offspring mass. Temperature during the early spring (prebreeding) had both direct and indirect effects on the environmental component of body mass, which is harder to explain. The positive indirect effect of warm temperature on breeding density is most likely to be related to a higher survival of nestlings in warmer years (this study—survival vs. temperature in previous year: $b = 0.0008 \pm 0.0003$, $\chi^2_{(1)} = 5.56$, $P = .018$, results of a GLMM with binomial error also including density, temperature in the current year, and fledging mass as covariates and brood identity as a random effect; see also Cresswell and McCleery 2003). On the other hand, the direct effect of temperature is much more complex because it has potentially different effects in the previous year and in the current year. The direct effect of warmer temperature in the current year might be to reduce fledging mass, potentially through its effect on food availability. Indeed, it has been shown that great tits breed later relative to caterpillars (their main food supply) in years when the temperature is high during the period after laying (Perrins 1991; Buse et al. 1999). The direct effect of temperature in the previous year could be partly related to the fact that it could increase the proportion of smaller individuals (with negative environmental deviation) that survive, which in turn would have smaller offspring the following year. Our data appear to support such a process, as there was a significant effect of temperature in the previous year on environmental deviation of survivors ($b = -0.00412 \pm 0.00129$, $t = -3.20$, $df = 34$, $P = .003$). Thus, temperature in the previous year not only increased survival but also reduced the average size of survivors.

A previous study of this population (Gosler et al. 1995) demonstrated that adult winter mass had shown a response to changes in sparrowhawk (*Accipiter nisus*) abundance, with adults regulating their mass at lower levels following an increase in sparrowhawk abundance between 1965 and 1980. Gosler et al. (1995) were able to show that the majority of this response could be attributed to individual phenotypic plasticity because repeated measures of individuals showed similar trends to the population as a whole. While sparrowhawks are important predators on newly fledged great tits, it is not likely that the change in fledging mass over time can be explained by a change in selection due to sparrowhawk predation. First, there is no evidence for a temporal trend in selection differentials on fledging mass. Second, splitting the analysis into two periods, from

1965 to 1980 (when sparrowhawks recolonized and then increased in population size within the study area) and from 1981 to 2000 (sparrowhawk population roughly stable), provides no evidence that the selective regime differed between the two periods (comparison of annual selection differentials: $t = 1.05$, $df = 34$, $P = .30$). Adriaensen et al. (1998) concluded that the presence of sparrowhawks influenced the form of selection on fledging mass in blue tits, with selection changing from directional to stabilizing as sparrowhawks become more numerous. We found no evidence in support of this conclusion because there was no suggestion that stabilizing selection had become stronger over time (a negative slope for quadratic coefficients is expected if stabilizing selection is becoming stronger over time: $b = 0.001 \pm 0.002$, $t = 0.49$, $df = 34$, $P = .630$).

In their study of a population of collared flycatchers found on Gotland, Sweden, Merilä et al. (2001a) used the synchronization between caterpillar hatching date and the date of bud burst of oak trees on which the caterpillars feed (values used were originally calculated for local conditions found in the Netherlands; see Visser and Holleman 2001) to explain variation in fledgling body condition index. They showed that condition at fledging was higher when synchronization was high (and thus food was likely to have been abundant). Using the same indicator of environmental quality as Merilä et al. (2001a), we found a significant positive relationship with fledging mass (from 1975 and 1999, the years for which synchronization data are available; $b = 0.014 \pm 0.006$, $t = 2.36$, $df = 23$, $P = .027$) and with environmental deviation ($b = 0.029 \pm 0.010$, $t = 2.84$, $df = 23$, $P = .009$). However, synchronization is also highly negatively correlated with increased spring temperatures (our data: $r = -0.69$, $b = -0.096 \pm 0.021$, $t = -4.61$, $df = 23$, $P < .001$; see also Visser and Holleman 2001), implying that the model for path analysis we have defined here includes a large part of this effect. It is, however, interesting to note that environmental data from the Netherlands provided plausible explanations for patterns observed in populations of birds breeding in Sweden and the United Kingdom, implying that a common large-scale environmental process operates across northwest Europe. Large-scale climatic processes such as the North Atlantic Oscillation Index have been shown to affect a range of different traits in animal populations over a wide geographic range (Post and Forchhammer 2002; Sheldon et al. 2003; Garant et al. 2004).

We were able to discount other potential causes for a mismatch between selection and phenotypic response. First, the heritability estimate we obtained for fledging mass was lower than the value obtained in a previous study using parent-offspring regression in the same population ($h^2 = 0.44 \pm 0.05$; Gosler and Harper 2000). Heritability

estimates obtained from animal models are often lower than estimates obtained with parent-offspring regressions (mainly because the common environment variance can be separated from the additive variance; see Kruuk 2004). However, our estimate was still considerable, and thus a lack of genetic variability is unlikely to represent a constraint on the potential evolutionary response to selection. Second, there were no noteworthy fluctuations over time in selection at either the phenotypic or the genotypic level. It is thus doubtful that such a process would hamper the expected evolutionary response of body mass. The situation is thus different from that observed, for example, in a long-term study of two species of Darwin's finches in which temporally fluctuating selection pressures generated unpredictable evolutionary trends for heritable beak size and shape traits (Grant and Grant 2002). Comparisons of immigrants with locally born birds also suggested that there was no evidence for different processes occurring elsewhere. Third, we found no evidence to suggest any constraint imposed by genetic correlations between body mass and fitness components measured later in life (life span and lifetime reproductive success); fledging mass breeding values were positively correlated with LSP and LRS. Opposing selection generated by either negative genetic correlations or reversed associations of fledging mass with fitness traits at a later stage in life are therefore not likely in the current system. Finally, genetic correlations between the sexes and between different segments of the study were not different from unity; both results suggest that sex or time-based genotype-environment interactions are unlikely to influence our conclusions.

It has been suggested previously that directional natural selection could mainly be associated with the environmental component of characters (Price et al. 1988; Alatalo et al. 1990; but see also Kruuk et al. 2001) and that natural selection would thus be removing the individuals with the largest negative environmental deviations from their breeding values. In agreement with this suggestion, we found that larger chicks were more likely to survive, potentially because they were healthier (taking size as a sign of health) and/or because intrabrood competition favors larger individuals (Garnett 1981; although note that there will also be variation in breeding values between chicks within a nest, such that an individual's rank is not entirely due to environmental effects). The result of this was that selection removed individuals with the most negative environmental deviation. However, when we compared the size of the additive genetic (V_A) and environmental ($V_E = \text{year} + \text{nest} + \text{residual}$; see table 1) components of variance before and after selection, we found evidence that both V_A and V_E are targeted by selection because all components were significantly reduced in the postselection sample. As a result, heritability estimates were not different

in our pre- and postselection samples. Finally, we also quantified the force of selection acting on the breeding value of individuals under selection and showed that there was a significant selection differential for breeding values over the study period. Selection associated with environmental deviations is thus not a sufficient explanation for the opposite change in phenotype observed here. It should be acknowledged, however, that our selection estimates rely on local survival and recruitment and are not considering the potential survivors recruiting outside of the study area.

In conclusion, we have demonstrated a microevolutionary response to selection on fledging mass, although it is a response that is masked by a deteriorating environment. In our particular case, there seems no good reason to assume that the selection response and the environmental deterioration were necessarily linked because the environmental change seems most plausibly explained as a large-scale climatic change. In that sense, the processes here differ from the model proposed by Cooke et al. (1990) because their model proposed that the selected population caused its environment to deteriorate. However, our results demonstrate the usefulness of combining quantitative genetic methods and data on population ecology and underline the importance of incorporating the effects of environmental variation when assessing evolution in the wild. Our results also highlight the value of long-term data sets in revealing the effects of environmental change and its interaction with evolutionary processes.

Acknowledgments

We are grateful to the many people who collected data during the long-term tit study in Wytham and to C. Perrins and two anonymous reviewers for comments on the manuscript. D.G. was financially supported by a Natural Sciences and Engineering Research Council of Canada Postdoctoral Research Fellowship and by a Biotechnology and Biological Sciences Research Council grant awarded to B.C.S. and L.E.B.K, who are both Royal Society University Research Fellows.

Literature Cited

- Adriaensen, F., A. A. Dhondt, S. Van Dongen, L. Lens, and E. Matthysen. 1998. Stabilizing selection on blue tit fledging mass in the presence of sparrowhawks. *Proceedings of the Royal Society of London B* 265:1011–1016.
- Alatalo, R. V., L. Gustafsson, and A. Lundberg. 1990. Phenotypic selection on heritable size traits: environmental variance and genetic response. *American Naturalist* 135: 464–471.
- Blakey, J. K. 1994. Genetic evidence for extra-pair fertilizations in a monogamous passerine, the great tit, *Parus major*. *Ibis* 136:457–462.
- Both, C., M. E. Visser, and N. Verboven. 1999. Density-dependent recruitment rates in great tits: the importance of being heavier. *Proceedings of the Royal Society of London B* 266:465–469.
- Burnham, K. P., and D. R. Anderson. 1998. *Model selection and inference: a practical information-theoretic approach*. Springer, New York.
- Buse, A., S. J. Dury, R. J. W. Woodburn, C. M. Perrins, and J. E. G. Good. 1999. Effects of elevated temperature on multi-species interactions: the case of pedunculate oak, winter moth and tits. *Functional Ecology* 13(suppl. 1):74–82.
- Cameron, N. D. 1997. *Selection indices and prediction of genetic merit in animal breeding*. CAB, Wallingford.
- Charlesworth, B. 1994. *Evolution in age-structured populations*. Cambridge University Press, Cambridge.
- Conover, D. O., and E. T. Schultz. 1995. Phenotypic similarity and the evolutionary significance of countergradient variation. *Trends in Ecology & Evolution* 10:248–252.
- Cooch, E. G., D. B. Lank, R. F. Rockwell, and F. Cooke. 1991. Long-term decline in body size in a snow goose population: evidence of environmental degradation? *Journal of Animal Ecology* 60:483–496.
- Cooke, F., P. D. Taylor, C. M. Francis, and R. F. Rockwell. 1990. Directional selection and clutch size in birds. *American Naturalist* 136:261–267.
- Cresswell, W., and R. McCleery. 2003. How great tits maintain synchronization of their hatch date with food supply in response to long-term variability in temperature. *Journal of Animal Ecology* 72:356–366.
- Endler, J. A. 1986. *Natural selection in the wild*. Princeton University Press, Princeton, N.J.
- Fairbairn, D. J., and R. F. Preziosi. 1996. Sexual selection and the evolution of sexual size dimorphism in the water strider, *Aquarius remigis*. *Evolution* 50:1549–1559.
- Falconer, D. S., and T. F. C. MacKay. 1996. *Introduction to quantitative genetics*. Longman, Essex.
- Garant, D., B. C. Sheldon, and L. Gustafsson. 2004. Climatic and temporal effects on the expression of secondary sexual characters: genetic and environmental components. *Evolution* 58:634–644.
- Garnett, M. C. 1981. Body size, its heritability and influence on juvenile survival among great tits, *Parus major*. *Ibis* 123:31–41.
- Gosler, A. G., and D. G. C. Harper. 2000. Assessing the heritability of body condition in birds: a challenge exemplified by the great tit *Parus major* L. (Aves). *Biological Journal of the Linnean Society* 71:103–117.
- Gosler, A. G., J. J. D. Greenwood, and C. Perrins. 1995.

- Predation risk and the cost of being fat. *Nature* 377: 621–623.
- Grant, P. R., and B. R. Grant. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296:707–711.
- Griffith, S. C., I. P. F. Owens, and K. A. Thuman. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Molecular Ecology* 11: 2195–2212.
- Groeneveld, E., M. Kovac, T. L. Wang, and R. L. Fernando. 1992. Computing algorithms in a general purpose BLUP package for multivariate prediction and estimation. *Archiv für Tierzucht* 35:399–412.
- Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. *American Naturalist* 157:245–261.
- Kinnison, M. T., and A. P. Hendry. 2001. The pace of modern life. II. From rates of contemporary microevolution to pattern and process. *Genetica* 112–113: 145–164.
- Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the “animal model.” *Philosophical Transactions of the Royal Society of London B* 359:873–890.
- Kruuk, L. E. B., J. Merilä, and B. C. Sheldon. 2001. Phenotypic selection on a heritable size trait revisited. *American Naturalist* 158:557–571.
- Kruuk, L. E. B., J. Slate, J. M. Pemberton, S. Brotherstone, F. E. Guinness, and T. H. Clutton-Brock. 2002. Antler size in red deer: heritability and selection but no evolution. *Evolution* 56:1683–1695.
- Kruuk, L. E. B., J. Merilä, and B. C. Sheldon. 2003. When environmental variation short-circuits natural selection. *Trends in Ecology & Evolution* 18:207–209.
- Larsson, K., H. P. van der Jeugd, I. T. van der Veen, and P. Forslund. 1998. Body size declines despite positive directional selection on heritable size traits in a barnacle goose population. *Evolution* 52:1169–1184.
- Laugen A. T., A. Laurila, K. Räsänen, and J. Merilä. 2003. Latitudinal countergradient variation in the common frog (*Rana temporaria*) development rates: evidence for local adaptation. *Journal of Evolutionary Biology* 16: 996–1005.
- Lynch, M., and B. Walsh. 1998. *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, Mass.
- McCleery, R. H., and C. M. Perrins. 1998. temperature and egg-laying trends. *Nature* 391:30–31.
- Merilä, J., B. C. Sheldon, and H. Ellegren. 1997. Antagonistic natural selection revealed by molecular sex identification of nestling collared flycatchers. *Molecular Ecology* 6:1167–1175.
- Merilä, J., B. C. Sheldon, and H. Ellegren. 1998. Quantitative genetics of sexual size dimorphism in the collared flycatcher, *Ficedula albicollis*. *Evolution* 52:870–876.
- Merilä, J., L. E. B. Kruuk, and B. C. Sheldon. 2001a. Cryptic evolution in a wild bird population. *Nature* 412:76–79.
- Merilä, J., B. C. Sheldon, and L. E. B. Kruuk. 2001b. Explaining stasis: microevolutionary studies in natural populations. *Genetica* 112–113:199–222.
- Merilä, J., L. E. B. Kruuk, and B. C. Sheldon. 2001c. Natural selection on the genetical component of variance in body condition in wild bird population. *Journal of Evolutionary Biology* 13:804–813.
- Meyer, K. 1989. Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. *Genetics Selection Evolution* 23:67–83.
- Minot, E. O. 1981. Effects of interspecific competition for food in breeding blue and great tits. *Journal of Animal Ecology* 50:375–385.
- Minot, E. O., and C. M. Perrins. 1986. Interspecific interference competition: nest sites for blue and great tits. *Journal of Animal Ecology* 55:331–350.
- Mitchell, R. J. 1992. Testing evolutionary and ecological hypotheses using path analysis and structural equation modelling. *Functional Ecology* 6:123–129.
- Neumaier, A., and E. Groeneveld. 1998. Restricted maximum likelihood estimation of covariances in sparse linear models. *Genetics Selection Evolution* 30:3–26.
- Parsons, K. E. 1997. Contrasting patterns of heritable geographic variation in shell morphology and growth potential in the marine gastropod *Bembicium vittatum*: evidence from field experiments. *Evolution* 51:784–796.
- Perrins, C. M. 1965. Population fluctuations and clutch size in the great tit, *Parus major* L. *Journal of Animal Ecology* 34:601–647.
- . 1979. *British tits*. Collins, London.
- . 1991. Tits and their caterpillar food supply. *Ibis* 133(suppl. 1):49–54.
- Perrins, C. M., and R. H. McCleery. 2001. The effect of fledging mass in the lives of great tits *Parus major*. *Ardea* 89(special issue):135–142.
- Post, E., and M. C. Forchhammer. 2002. Synchronization of animal population dynamics by large-scale climate. *Nature* 420:168–171.
- Price, T., M. Kirkpatrick, and S. J. Arnold. 1988. Directional selection and the evolution of breeding date in birds. *Science* 240:798–799.
- Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to environmental covariances between traits and fitness. *Evolution* 46:616–626.

- Roff, D. A. 1997. Evolutionary quantitative genetics. Chapman & Hall, New York.
- Scheiner, S. M., K. Donohue, L. A. Dorn, S. J. Mazer, and L. M. Wolfe. 2002. Reducing environmental bias when measuring natural selection. *Evolution* 56:2156–2167.
- Sheldon, B. C., L. E. B. Kruuk, and J. Merilä. 2003. Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution* 57:406–420.
- Shipley, B. 2000. Cause and correlation in biology: a user's guide to path analysis, structural equations and causal inference. Cambridge University Press, Cambridge.
- Statsoft. 1999. Statistica for Windows (computer program manual). StatSoft, Tulsa, Okla.
- Stinchcombe, J. R., M. T. Rutter, D. S. Burdick, P. Tiffin, M. D. Rausher, and R. Mauricio. 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: theory and practice. *American Naturalist* 160:511–523.
- Tinbergen, J. M., and M. C. Boerlijst. 1990. Nestling weight and survival in individual great tits (*Parus major*). *Journal of Animal Ecology* 59:1113–1127.
- van Noordwijk, A. J., J. H. van Balen, and W. Scharloo. 1988. Heritability of body size in a natural population of the great tit (*Parus major*) and its relation to age and environmental conditions during growth. *Genetical Research* 51:149–162.
- van Noordwijk, A. J., R. H. McCleery, and C. M. Perrins. 1995. Selection for the timing of great tit breeding in relation to caterpillar growth and temperature. *Journal of Animal Ecology* 64:451–458.
- Verboven, N., and M. E. Visser. 1998. Seasonal variation in local recruitment of great tits: the importance of being early. *Oikos* 81:511–524.
- Verhulst, S., C. M. Perrins, and R. Riddington. 1997. Natal dispersal of great tits in a patchy environment. *Ecology* 78:864–872.
- Visser, M. E., and L. J. M. Holleman. 2001. Warmer springs disrupt the synchrony of oak and winter moth phenology. *Proceedings of the Royal Society of London B* 268:289–294.
- VSN. 2002. Genstat. Version 6.2. VSN International, Oxford.

Associate Editor: Trevor Price