VOL. 164, NO. 3 THE AMERICAN NATURALIST SEPTEMBER 2004

E-ARTICLE

Components of Variance Underlying Fitness in a Natural Population of the Great Tit *Parus major*

R. H. McCleery,[1,*] R. A. Pettifor,[2] P. Armbruster,[2,3] K. Meyer,[4] B. C. Sheldon,[1] and C. M. Perrins[1]

1. Department of Zoology, Edward Grey Institute, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom;
2. Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, United Kingdom;
3. Department of Biology, University of Vermont, Burlington, Vermont 05405-0086;
4. Animal Genetics and Breeding Unit, University of New England, Armidale, New South Wales 2351, Australia

Submitted February 21, 2004; Accepted March 9, 2004; Electronically published July 23, 2004

ABSTRACT: Traits that are closely associated with fitness tend to have lower heritabilities (h^2) than those that are not. This has been interpreted as evidence that natural selection tends to deplete genetic variation more rapidly for traits more closely associated with fitness (a corollary of Fisher's fundamental theorem), but Price and Schluter (1991) suggested the pattern might be due to higher residual variance in traits more closely related to fitness. The relationship between 10 different traits for females, seven traits for males, and overall fitness (lifetime recruitment) was quantified for great tits (*Parus major*) studied in their natural environment of Wytham Wood, England, using data collected over 39 years. Heritabilities and the coefficients of additive genetic and residual variance (CV_A and CV_R, respectively) were estimated using an "animal model." For both males and females, a trait's correlation (r) with fitness was negatively related to its h^2 but positively related to its CV_R. The CV_A was not related to the trait's correlation with fitness in either sex. This is the third study using directly measured fitness in a wild population to show the important role of residual variation in determining the pattern of lower heritabilities for traits more closely related to fitness.

Keywords: heritability, additive genetic variance, fitness, lifetime reproductive success, *Parus major*, animal model.

Understanding the relative contribution of genetic and environmental factors to the variation in the fitness of individuals is a fundamental goal of evolutionary biology (Lai et al. 1994; Fowler et al. 1997; Kruuk et al. 2000; Merilä and Sheldon 2000). The total phenotypic variance of a trait (V_P) can be expressed as $V_P = V_A + V_D + V_I + V_E$, where V_A is the additive genetic variance, V_D and V_I are the dominance and epistatic variances, respectively (nonadditive sources of genetic variance), and V_E is the environmental variance (Falconer and Mackay 1996; Lynch and Walsh 1998). The narrow-sense heritability is then defined as $h^2 = V_A/V_P$. The narrow-sense heritability (h^2) has been the focus of literally thousands of studies in evolutionary biology (see Falconer and Mackay 1996; Lynch and Walsh 1998; Roff 1997) because, since it is the slope of the relationship between change in phenotypic value from one generation to the next and selection differential, it determines the efficiency of response to natural selection (e.g., Lynch and Walsh 1998, p. 50).

One apparently robust pattern to emerge from studies of heritability is that traits more closely related to fitness (e.g., fecundity, survivorship) tend to have lower heritabilities than traits less intimately associated with fitness (e.g., morphometric, behavioral traits; Gustafsson 1986; Mousseau and Roff 1987; Roff and Mousseau 1987; Houle 1992; Kruuk et al. 2000; Merilä and Sheldon 2000). Initially, this result was interpreted as support for a corollary of Fisher's (1930) fundamental theorem of natural selection (Gustafsson 1986; Mousseau and Roff 1987; Roff and Mousseau 1987), namely, that selection should deplete genetic variation for traits more closely associated with fitness at a more rapid rate relative to those traits less closely associated with fitness. It follows that in populations near an evolutionary equilibrium, traits more closely related to fitness should have lower levels of additive genetic variance. More recently, Price and Schluter (1991) have pointed out that in many cases, although a relationship between a trait's association with fitness and its heritability can be demonstrated, the relationship between additive genetic variance and association with fitness predicted from Fisher's fundamental theorem is not always clear. They suggested that traits more closely associated with fitness might have lower heritabilities because of higher levels of residual (nonadditive genetic,

---

[*] Corresponding author; e-mail: robin.mccleery@zoo.ox.ac.uk.

Am. Nat. 2004. Vol. 164, pp. E62–E72.

Table 1: Analyzed traits

Trait	Comment
1. Clutch size	Number of eggs laid by a female in a completed clutch; where known, all repeat clutches and second broods were excluded
2. Lay date	Calculated as the day on which the first egg of each clutch was laid by a female in its first year of life; where identified, the lay dates of repeat and second broods were excluded
3. Egg mass	Mean mass (in grams) of a sample of eggs in each nest, weighed before the onset of incubation
4. Number fledged	Number of young known to have successfully fledged
5. Dispersal	Distance in meters from the natal nest
6. Nestling mass	Mass (in grams) of each individual on day 15 of its nestling period (day of hatch = day 1)
7. Recruitment	Number of young that were recorded entering the Wytham population as breeders, originating from each individual breeding attempt by birds in their first year
8. Longevity	Age at which each individual was last known to be alive; birds breeding for the first time in 1994 or later were excluded to prevent entering incomplete data for any individual
9. Lifetime fledging success	Total number of young fledged from all breeding attempts by an individual
10. Lifetime reproductive success	Total number of recruits produced by each individual over its lifetime
11. Wing length	The only morphological character (apart from body mass, which depends on time of day and time of season in a complex way) for which we have reliable measurements from adults throughout the data set

maternal, and environmental) variation. Price and Schluter (1991) also pointed out that there was no simple prediction regarding the levels of additive genetic variance for traits more or less closely associated with fitness. Recent empirical studies (Houle 1992; Kruuk et al. 2000; Merilä and Sheldon 2000) have supported the model outlined by Price and Schluter (1991).

Because of practical constraints, the majority of life-history studies are conducted in a laboratory setting, and this restricts the extent to which the results can be used to understand evolutionary processes in the wild for several reasons. First, patterns of mating in the laboratory may be unnatural. Also, in comparison to a natural environment, novel laboratory environments may change relative levels of genetic and environmental variation and genetic correlations among traits (Service and Rose 1985). Most laboratory studies (and many wild ones) examine fitness correlates rather than overall fitness itself. In the case of laboratory studies, it is often unclear how these fitness correlates relate to lifetime fitness in the wild. Like Merilä and Sheldon (2000) and Kruuk et al. (2000, 2001), we studied a wild population in a natural environment so that we are able to get close to a direct measure of fitness.

We have followed Kruuk et al. (2000) by using analytical models from animal breeding research (Meyer 1988, 1989) to partition the causal components of phenotypic variance for traits underlying fitness, taking advantage of a long-term study started by David Lack in 1947 on the great tit (*Parus major*) in Wytham Wood, Oxford. Because most of the individuals in this population have been uniquely marked and reproductive performance was monitored every year for the past 50 years, we were able to construct a pedigree of >2,100 individual females and 1,800 males, generating >2,750 kin relationships for each sex. In conjunction with the phenotypic measurements of reproductive effort in this population, we are therefore able to estimate the additive genetic and residual (nonadditive genetic + environmental) components of variance affecting these traits, quantify the relationship of each trait to overall fitness (lifetime recruitment), and explicitly test the relationships between the extent of selection experienced by a trait (its association with fitness) and the level of additive genetic and residual variation exhibited by that trait. Our study complements previous efforts in this field. We have more than twice as many years of data as Merilä and Sheldon (2000), amounting to >30 generations, and our study population is larger and less isolated than that of Kruuk et al. (2000).

Methods

Study Population

Great tits *Parus major* have been studied in Wytham Woods (230 ha), near Oxford, over the past 50 years: the data used here were collected between 1960 and 1998. The birds freely use nest boxes in which to breed. The area of the wood used and the number of boxes present increased over the period 1960–1964, by which time most of the area of Wytham wood was included, with 1,020 boxes. It is thought that almost all great tits breeding in Wytham Woods do so in these nest boxes. There are two grounds for believing this. East and Perrins (1988) showed that removal of nest boxes resulted in a substantial drop in great tit numbers compared to a control area where boxes were undisturbed. Second, in one season territorial males were identified by recording song and later matched with nests in nest boxes. No territorial males were left unaccounted for in that season (P. McGregor, unpublished observations). Since 1964, the population size of breeding pairs of great tits has fluctuated between 120 and 415 ($\bar{x} = 237$, $SD = 71$, $n = 38$), with signs of an increase in recent years. Thirty-five percent of breeding females were caught in 1960, but thereafter an average of 80% of all breeding females were caught in any one year (range 42%–93%). Males were not consistently caught until 1964 (16%), averaging 66% thereafter (range 22%–85%).

All nest boxes were checked at weekly intervals during the beginning of each breeding season, and clutch size, lay date, mean egg mass, and number of young fledged were recorded for each occupied box (see trait definitions in table 1). Young were banded between days 7 and 15 of the 20-day nestling period, and all young were weighed on day 15. Parents were captured during the mid- to late-nestling period and banded (where necessary), sexed, and weighed. There is therefore a bias toward recapture of successful adults because birds that bred but did not get as far as the mid-nestling period were not recorded as breeding individuals in that season. Because all young raised in Wytham each year are banded, it is possible to distinguish between immigrant and resident parents. Recruits were defined as those young born in Wytham that entered the study breeding population, generally in the year following their birth. On average, 53% of females and 60% of males breeding in any year within Wytham were birds that were born in the wood, while an unknown number of young emigrated from the wood before breeding. Perrins (1965, 1979, 1998) provides further details of the study, its protocol, and the general ecology and biology of the Wytham great tit.

Data Selection and Traits Analyzed

We analyzed components of variance for the 11 traits shown in table 1. The mean and standard deviation for each trait is presented in table 2. Traits 1–7 were restricted to birds breeding in their first year of life because there are significant differences between first-year and older birds in these measurements (Perrins and McCleery 1985, 1989). Although a mixed model could incorporate age, we felt that it would increase the number of parameters in an already heavily parameterized model and that confining ourselves to first-year performance would not unduly reduce the power of the analysis. Traits 1–3 were used for females only. It is of course possible that female performance in these traits is dependent on characteristics of her mate. For example, good quality males might acquire better territories on which their mates lay larger, heavier clutches earlier in the year. However, estimates of variance in these traits due to the male and female of the pair show that males have very little influence (R. A. Pettifor, B. C. Sheldon, W. J. Browne, J. Rasbash, and R. H. McCleery, unpublished manuscript).

During the course of this study, a number of nests have been experimentally manipulated for one purpose or another (e.g., Perrins and Moss 1975; Pettifor et al. 1988, 2001). With the exception of clutch size, lay date, and egg mass, we have excluded from analysis all data for those individuals whose young were manipulated and all offspring of manipulated nests because these would introduce errors in the pedigree. The mean lifespan of breeders is less than 2 years (female $\bar{x} = 1.80$ years, $SD = 1.13$, $n = 4{,}650$; males $\bar{x} = 1.84$ years, $SD = 1.20$, $n = 3{,}751$), and the maximum recorded lifespan is 9 years. None of the 1992-born cohort was known to be alive in 1998; consequently, in order to calculate longevity (trait 8) and lifetime measures of reproductive success (traits 9 and 10), only data up to 1993 have been included. Similarly, as >80% of resident breeders (i.e., born in Wytham) were first recaptured as breeders in their first or second year of age, analysis of annual recruitment has been restricted to breeding records prior to 1996. Even so, the estimate of longevity (trait 8) is a minimum estimate due to the probability that an animal survived and bred but was not observed breeding. McCleery et al. (1996) combined Cormack-Jolly-Seber recapture estimates with records of occupied nests where the adults were not trapped to deduce that in any season about 85% of birds that had bred previously were seen if they were alive and that 70% of those alive but not seen did attempt breeding but evaded capture because their nests failed before they could be trapped.

Pedigree

Because the majority of the individuals in the Wytham great tit population have been individually banded as chicks over the last 50 years, we were able to construct a pedigree of 2,233 individual females and 2,236 individual males that spans the 39 years included in this study. To be included, females had to be observed breeding in their first year but could be daughters of a female breeding as

Table 2: Means (SD); additive genetic (V_A), environmental ($V_{\text{nest box}}$ and V_{group}), and residual (V_{residual}) variance; narrow-sense heritability (h^2); correlation with total fitness (r_P); and standardized genetic variance (V_w) for the 11 female and eight male traits associated with lifetime fitness considered in this study

	Population mean	Population SD	Number of animals	V_A	$V_{\text{nest box}}$	V_{group}	V_{residual}	h^2	r_P	V_w
Female trait:										
Lifetime recruitment[a]	1.113	1.535	1,795	.004 (.078)	.047 (.049)	.000 (.023)	2.096 (.117)	.002 (.036)	(1)	.0032
Clutch size	8.552	1.686	1,795	.797*** (.157)	.229*** (.063)	.077* (.036)	1.362 (.149)	.334*** (.062)	.170***	.011
Lay date (April days)	28.839	8.654	1,777	5.018** (1.879)	5.072*** (1.015)	2.403*** (.697)	21.540 (2.024)	.159** (.059)	−.121**	.006
Egg mass (g)	1.682	.128	1,540	.006*** (.001)	.000 (.000)	.000 (.000)	.009 (.001)	.396*** (.066)	−.023	.002
Dispersal from natal nest (m)	980.498	654.418	1,758	.157*** (.038)	.000 (.015)	.000 (.006)	.465 (.040)	.253*** (.058)	.023	.000
First-year fledging success	6.451	3.296	1,794	.000 (.444)	.000 (.204)	.385* (.158)	8.772 (.569)	.000 (.051)	.300***	.000
Lifetime fledged	10.587	8.171	1,795	.000 (2.509)	1.821 (1.493)	.229 (.685)	59.815 (3.545)	.000 (.041)	.641***	.000
Longevity	1.62	1.01	1,795	.038 (.043)	.041 (.025)	.027 (.015)	.881 (.057)	.039 (.045)	.526***	.014
First-year recruits	.68	.978	1,795	.010 (.033)	.000 (.019)	.011 (.010)	.812 (.048)	.013 (.041)	.687***	.022
Nestling mass (g)	18.553	1.08	1,595	.547*** (.080)	.024 (.029)	.000 (.014)	.549 (.730)	.489*** (.064)	−.001	.002
Wing length (mm × .1)	727.089	14.178	1,424	122.688*** (15.101)	9.340 (5.302)	11.776** (3.908)	56.722 (12.299)	.650*** (.065)	.026	.0002
Male trait:										
Lifetime recruitment[a]	1.108	1.433	1,631	.031 (.072)	.073 (.051)	.000 (.023)	1.716 (.105)	.017 (.040)	(1)	.025
Dispersal from natal nest (m)	726.665	575.808	1,607	.168*** (.044)	.0198 (.019)	.000 (.007)	.494 (.005)	.247*** (.063)	.047	.000
First-year fledging success	7.046	2.831	1,629	.000 (.388)	.214 (.193)	.161 (.108)	6.892 (.49)	.000 (.055)	.226***	.000
Lifetime fledged	10.701	7.197	1,631	.000 (2.078)	.000 (1.245)	.000 (.625)	48.234 (2.987)	.000 (.043)	.611***	.000
Longevity	1.541	.969	1,631	.000 (.040)	.000 (.024)	.000 (.017)	.888 (.055)	.000 (.045)	.520***	.000
First-year recruits	.692	.954	1,631	.020 (.033)	.027 (.022)	.007 (.010)	.725 (.046)	.026 (.043)	.687***	.042
Nestling mass (g)	19.418	1.197	1,478	.825*** (.110)	.033 (.038)	.000 (.014)	.541 (.093)	.590*** (.067)	.054*	.002
Wing length (mm × .1)	756.321	14.479	1,380	134.668*** (16.351)	7.187 (5.928)	9.710* (4.034)	58.393 (13.186)	.673*** (.066)	.005	.0002

Note: The standard errors associated with the parameter estimates are given in parentheses. For each trait, number of animals is the total number of individuals involved in the analysis. Standardized additive genetic variance is just CV_A^2. Burt (1995, 2000) uses the notation V_W^2 for this quantity. The notation r_P is used to distinguish this correlation from the conventional use of r for fitness.

[a] Lifetime reproductive success.

* $P < .05$.

** $P < .01$.

*** $P < .001$.

an older bird. The resulting pedigree contained 2,233 individuals at its base and included 921 mothers with progeny and 849 fathers with progeny in addition to all other possible kin pairings. The exact number of individuals included in each analysis varied slightly due to missing values for some traits for some individuals. The restricted maximum likelihood analysis we employed (described below) utilizes information from all available relationships in the pedigree and is robust to unbalanced family structure, making this method significantly more powerful than traditional regression-based (i.e., parent-offspring or full sib–half sib) experimental designs (Lynch and Walsh 1998). Estimates of extra-pair paternity in the Wytham great tit population suggest that approximately 14% of the nestlings result from fertilization with a male other than the resident male of the nest pair (Blakey 1994). These offspring of extra-pair matings will slightly decrease the phenotypic covariance among relatives and thus will also slightly bias downward our estimates of both the additive genetic variance (V_A) and the narrow-sense heritability (h^2). The effect will be sex specific as there are essentially no errors in assigning mothers to offspring.

Variance Component Analysis

Components of variance and the narrow-sense heritability (h^2) for each trait were estimated by restricted maximum likelihood with DFREML software, version 3.0β (Meyer 1998). We fitted a mixed model of the form

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{b} + \mathbf{Z}\boldsymbol{a} + \boldsymbol{e}, \tag{1}$$

where $\boldsymbol{y}$ is a vector of phenotypic values, $\mathbf{X}$ and $\mathbf{Z}$ are the incidence (design) matrices of fixed and random effects, respectively, $\boldsymbol{a}$ and $\boldsymbol{b}$ are the vectors of fixed and random effects, respectively, and $\boldsymbol{e}$ is the vector of residual errors (Shaw 1987).

Univariate analyses were carried out for each sex separately. Reproductive output is known to vary among habitat types and breeding year in Wytham Wood (Perrins 1979). We therefore defined a "contemporary group" for each breeding year and habitat type and fitted this term as a fixed effect, thus ensuring that only animals subject to the same systematic effects were compared directly with each other. This variable can be thought of as a "separate slopes" parameterization of a year × habitat interaction. Reproductive output also varies between nest boxes within the wood, so we fitted "nest box" as an uncorrelated random effect together with the additive genetic effects (breeding value) of each individual. Note that for traits 8–10 (table 1), which incorporate information across the lifespan of the individual, contemporary group and nest box vary with each breeding attempt, and consequently these were excluded from the model structure. Instead, a cohort grouping (an unique birth-year and habitat classification) and natal nest-box identity were fitted as fixed and random effects, respectively. While we were thus able to estimate environmental influences on traits and control for them, our data are not very suitable for direct estimation of the variance due to maternal effects. The reason for this is that most females breed only once, so that any maternal effect variance is swallowed by the nest box component of variance. Those females that do breed repeatedly often change boxes between attempts, making it difficult to estimate the maternal effect in addition to the environmental effect. Repeated observations of breeding individuals (e.g., Kruuk et al. 2000) are needed to obtain a better estimate of maternal effects. Our inability to estimate the maternal effect variance in this analysis does not influence our conclusions.

For each trait, models were first run using a derivative-free search algorithm (Simplex procedure, Nelder and Mead 1965) with the convergence criterion for acceptance of the log likelihood set at 10^{-15}. Using estimates from this first run as starting values, the model was then rerun using the average information algorithm of Johnson and Thompson (1995). DFREML assumes the residual error terms of fixed and random effects conform to a normal (Gaussian) distribution, but the parameter estimates are robust to departures from normality (Meyer 1985). Nevertheless, analyses were initially run on both log-transformed and untransformed data. We found no significant differences between transformed and untransformed data, and for ease of interpretation, we report results from the untransformed analyses.

We did not attempt to estimate the contribution of nonadditive genetic variance ($V_D + V_I$) to the total phenotypic variance in this population because of insufficient power in the pedigree. In fact, because V_D and V_I will not contribute to the long-term response to selection (Falconer and Mackay 1996; Lynch and Walsh 1998), it is reasonable to combine $V_D + V_I + V_E$ to estimate the residual (nonheritable) variance (V_R) contributing to V_P.

In order to maintain consistency with previous studies when comparing levels of additive genetic and residual variance among traits in our study population, we computed the coefficients of additive genetic variation (CV_A) and coefficients of residual variation (CV_R) for each trait as $CV_A = (V_A)^{1/2}/\overline{X}$ and $CV_R = (V_R)^{1/2}/\overline{X}$, where $\overline{X}$ is the trait mean (Houle 1992; Kruuk et al. 2000; Merilä and Sheldon 2000). Analyses were undertaken using the raw data, data standardized to the yearly median values, and log-transformed values. Results were similar irrespective of the transformation, and we therefore only report on the results from the untransformed data here. We were able to explicitly examine the relationship of each trait in our study to overall fitness by computing the product moment (Pearson) correlation (r_P) between each trait and total lifetime recruitment (fitness). Using this as an index of association, we then correlated the association with fitness (r_P) of each trait and its narrow-sense heritability (h^2), its CV_A, and its CV_R.

While the calculation of coefficients of variation partly controls for the fact that the variance of a trait increases with its mean, it is also sometimes recommended that specific transformations be carried out to stabilize the variance across traits (Lynch and Walsh 1998). We used Kleckowski's transformation (Lynch and Walsh 1998, pp. 301–302) prior to running DFREML but obtained qualitatively similar results to running the analyses on the raw data and then examining the coefficients of variation as described above; only the results from the untransformed analyses are reported here.

Results

Variance Components and Narrow-Sense Heritabilities

The additive genetic variances, residual variances, and narrow-sense heritabilities for all traits considered in this study are presented in table 2. For females, clutch size, lay date, egg mass, nestling mass, dispersal, and wing length (six of 11 traits) had additive genetic variances significantly different from 0. In males, dispersal, nestling mass, and wing length had an additive genetic variance significantly >0. The residual (nonadditive and environmental) variances of all the traits in both males and females were each significantly different from 0 (table 2).

Narrow-sense heritabilities (h^2) ranged from 0.00 to 0.39 for the life-history traits, a typical range (Mousseau and Roff 1987; Roff 1997; Lynch and Walsh 1998), but were much higher (0.65–0.67) for wing length and for nestling mass. Significance tests for h^2 are effectively the same because they depend on V_A. Hence, clutch size, lay date, egg mass, nestling mass, dispersal, and wing length have h^2 values significantly different from 0 in females, while nestling mass, dispersal, and wing length had an h^2 value significantly different from 0 in males. It is notable that the female-only traits all have highly significant heritabilities.

The variances due to nest box and to contemporary group are of some interest, though their primary function in this analysis is to account for those components of phenotypic variance that are likely to be due to temporal or known environmental effects. The majority of traits have no significant component of variance attributable to either nest box or contemporary group, though there are some exceptions. The female characters lay date and clutch size have significant components for both. First-year fledging has a V_{group} component for females but not for males. Wing length has a significant V_{group} component for both sexes. These findings thus suggest consistent environmental influences on these characters.

Associations with Fitness

The correlation between each trait considered in this study and total fitness (lifetime recruitment) is given in table 2. Given the sample sizes, it is not surprising that even the lower values for the correlation coefficients in this table differ significantly from 0, though egg mass (for females), natal dispersal, and wing length are not significantly correlated to total fitness. There is a weak ($r_P = -0.054$, $.05 > P > .01$) relationship between total fitness and male nestling mass, but the relationship is not significant for females. The strength of the association of the traits varied, with first-year recruitment, lifetime fledging success, and longevity being the strongest correlates of lifetime fitness in both males and females (table 2). The relationships between the unsigned correlation of each trait with lifetime fitness and h^2, CV_A, and CV_R are plotted in figure 1*A*–1*C* for females and figure 1*D*–1*F* for males, with the statistical analysis summarized in table 3.

Three different correlation coefficients are given for each relationship in table 3. The "linear" correlation is the Pearson correlation of the measure indicated (h^2, CV_A, and CV_R) with r_P. In calculating the significance of this correlation, we used linear regression, and we used model criticism to check the assumptions of a parametric test. In the case of h^2, the plot of the residuals versus the fits indicated that the homogeneity of variance assumption was violated, while for CV_A the residuals were not normally distributed. In both cases, the problem was corrected by taking the square root of the h^2 or CV_A, respectively. A more conservative approach is to use a Spearman's rank correlation in this situation, though the lack of power is a drawback with a very small sample size, as in this case. The results of all three procedures are the same, except that for males the Spearman's correlation between h^2 and the traits' correlations with fitness is not significant. The correlation of fitness with itself was not used in either case, giving 9 degrees of freedom for females and 6 for males.

In both males and females, there was a significant negative relationship between the narrow-sense heritability (h^2) of the traits and their unsigned association with fitness (r_P). No significant relationship was found between the coefficient of additive genetic variance (CV_A) and r_P in females (table 3; fig. 1*B*) or males (table 3; fig. 1*E*). However, for both sexes, the CV_R of a trait and its association with fitness were strongly positively correlated.

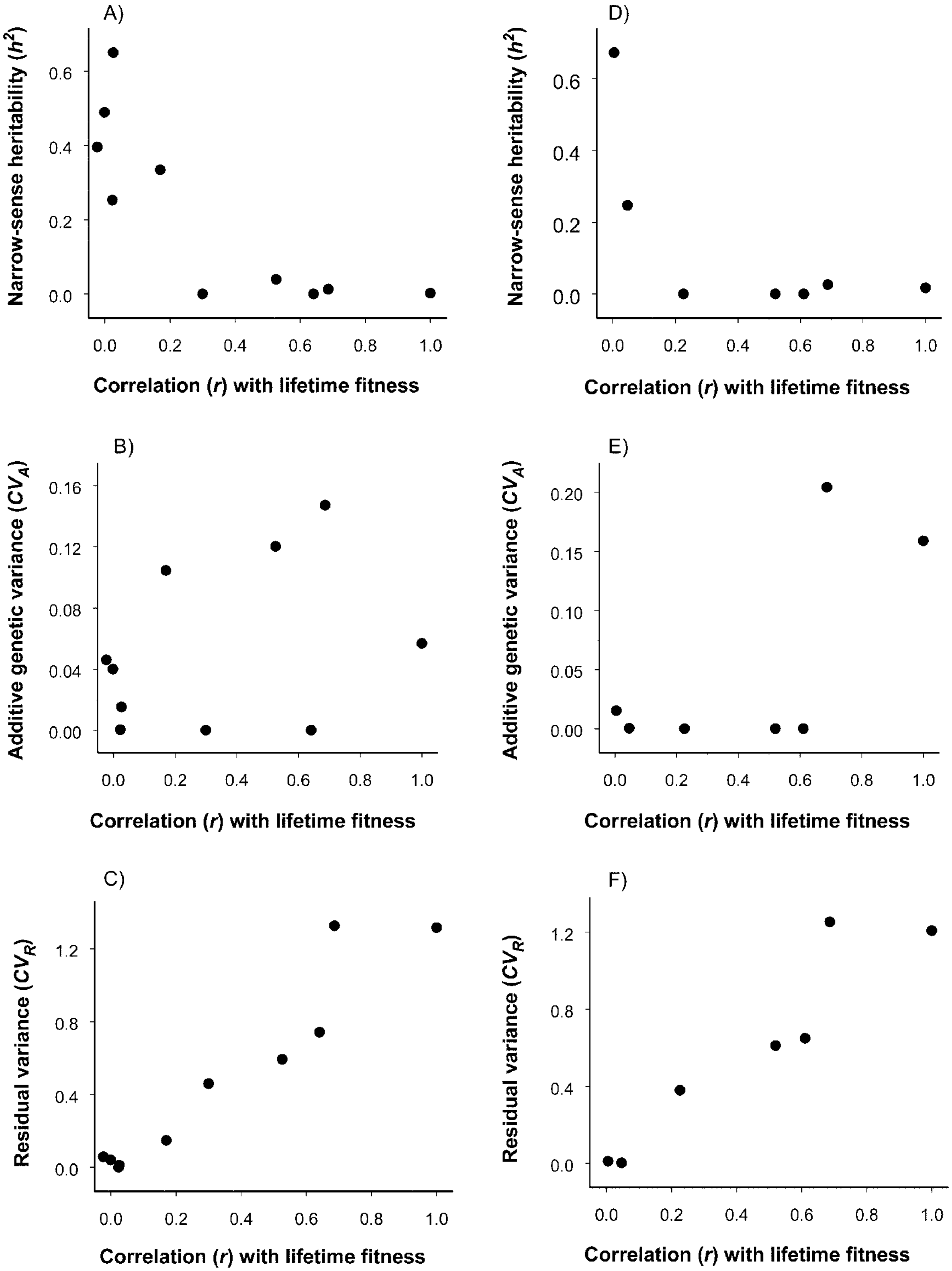

Figure 1: The relationship between a trait's unsigned association (r) with overall fitness (lifetime recruitment) and the narrow-sense heritability (h^2; *A*, *D*), the coefficient of additive genetic variance (CV_A; *B*, *E*), and the coefficient of residual (nonadditive + environmental) variance (CV_R; *C*, *F*). Left-hand plots (*A–C*) are for females; right-hand plots (*D–F*) are for males. As an aid to visual clarity, the symbols have been slightly offset.

Table 3: Relationship between the heritability, CV_A and CV_R, of each trait in table 2 and its unsigned correlation with lifetime recruitment (r_P in table 2)

	Females			Males		
Measurement	Linear	Square root transform	Spearman	Linear	Square root transform	Spearman
Heritability	−.713 (.02)	−.786 (.007)	−.634 (.05)	−.812 (.026)	−.815 (.025)	−.660 (.107)
CV_A	.338 (.34)	.113 (.755)	.111 (.760)	.429 (.337)	.199 (.699)	−.066 (.888)
CV_R	.907 (<.001)	.899 (<.001)	.722 (.018)	.935 (.002)	.948 (.001)	.864 (.012)

Note: The linear and square root transform columns are Pearson correlations, the latter after square root transforming the measurements for reasons discussed in the text. The Spearman's column is a Spearman rank correlation (ρ). Figures in brackets are the probability for rejecting the null hypothesis of no relationship. Degrees of freedom are 9 for females and 6 for males. The only result affected by the choice of test is the relationship for heritability and correlation with fitness for males.

Standardized Genetic Variance

We also include standardized genetic variance V_w in table 2. It is simply CV_A^2, and it is proposed by Burt (1995, 2000) as a measure of the rate at which fitness will increase in the population across generations. In only one case (lifetime recruitment for males) are the estimated values more than 0.02. The V_w value for fitness (i.e., lifetime recruitment) itself is close to 0 for females, implying that the population is at evolutionary equilibrium for fitness. The value for males is surprisingly high at about 2.5%, though this result should be treated with caution because the calculation involves V_A for males, which does not differ significantly from 0 (0.031, SE 0.072).

Discussion

As in previous studies (Gustafsson 1986; Mousseau and Roff 1987; Roff and Mousseau 1987; Houle 1992; Kruuk et al. 2000; Merilä and Sheldon 2000), we found that traits more closely associated with fitness had lower narrow-sense heritabilites (h^2) than traits more distantly related to fitness (fig. 1). It is particularly interesting to compare our results to the two other analyses of long-term data sets for vertebrates studied in the wild. Kruuk et al. (2000), analyzing data from the Rum red deer (*Cervus elaphus)* long-term study, found a significant negative relationship in both male and female red deer between the h^2 estimates of various traits, including a mixture of reproductive output measures and morphology and their correlation with fitness ($.01 < P < .05$). Merilä and Sheldon (2000), expanding the analyses of Gustafsson (1986) on the Gotland collared flycatcher (*Ficedula albicollis*), obtained a highly significant negative relationship between h^2 and r^2 in male collared flycatchers ($P = .0003$) and a negative but nonsignificant relationship in females ($P = .086$). Thus three independent studies of wild populations now confirm a trend for a negative association between the narrow-sense heritability of traits and their association with fitness. It is worth noting that where heritability is significantly >0, the life-history traits in the great tit most directly associated with fitness all have rather similar h^2 values of around 20%–25% while morphological traits such as nestling mass and especially wing length have much higher h^2 values but much lower correlations with fitness. Inclusion of further morphological traits, had they been available, would have greatly strengthened the negative association between h^2 and r because morphological traits generally tend to be weakly associated with lifetime fitness. For example, heritability estimates of tarsus length and bill length in males and females range from 0.5 to 0.8 (collated in Boag and van Noordwijk 1987), though these estimates may be on the high side because they are all based on parent-offspring correlations without any control for common environments. Kruuk et al. (2001) point out that such estimates may be inflated due to common environmental or maternal effects, and they report a value of 0.35 ± 0.02 for the heritability of tarsus, compared with a value of 0.53 ± 0.07 obtained from parent offspring regression (Merilä and Sheldon 2000). Our estimate for Wytham great tit females is 0.65 ± 0.065, which does not differ from an estimate of 0.68 ± 0.08 based on mother-daughter regression (R. H. McCleery, unpublished analysis).

Our use of lifetime recruitment to the breeding population perhaps needs some justification. In an ecological context, it seems the most natural measure to use, though it is of course a minimum estimate because we have not corrected for the resighting probability of surviving young or tried to estimate emigration. However, theoretical treatments of selection have repeatedly stated that inferences about selection on maternal characters may be incorrect if fitness attributes are erroneously given to the mother, especially when there is genetic covariance between a maternal characteristic (e.g., laying date) and an offspring character that is subject to selection (e.g., fledging mass). Problems of interpretation might also occur when there is environmental variance in the maternal character, as would be the case with clutch size and laying date. (e.g., Lande and Arnold 1983; Cheverud and Moore 1994; Wolf and Wade 2001). Thus in our case, we are assigning an offspring trait, namely, its ability to join the population as a breeder, to its parents by counting their fitness as the number of recruited breeders in their lifetime. We considered that using a direct measure of fecundity such as clutch size or number of fledglings would certainly underestimate the effects we are interested in. For example, mother's laying date is causally related to the probability that her offspring survive to breed in great tits (Brinkhof et al. 1993; Norris 1993; Verhulst et al. 1995), while there are inverse correlations between laying date and clutch size (earlier clutches tend to be larger) and between fledging mass and brood size; heavier fledglings are more likely to survive to breed. Furthermore, at least some of the postfledging survival in this species is due to the activities of the parents, who feed their offspring for a period after they have left the nest. Using total recruitment as a surrogate for fitness means that we are working on the resultant of all these trade-offs by the parents (see also Sheldon et al. 2003; Brommer et al. 2004).

The phenotypic variance in the majority of the traits had no significant component of environmental variation attributable to nest box or to contemporary group; the exceptions are the related female traits lay date and clutch size and wing length in both sexes. As we would expect, nest-box identity accounts for a significant proportion of the variance in clutch size and lay date because some nest boxes are in more favorable locations than others, with more shelter or more food, and females consistently lay more eggs earlier in these. The contemporary group, which is in effect a parameterization of an interaction between year and area of the wood, also accounts for variance in clutch size and lay date. This almost certainly reflects the strong between-year and between-area differences in the timing of the breeding season (Cresswell and McCleery 2003) and consequently in the mean clutch size of the population. The weak relationship with fledging success may also be related to timing but is more likely due to the fact that between 1964 and 1976 there were years when substantial numbers of nests were preyed on by weasels (*Mustela nivalis*; Dunn 1977; McCleery et al. 1996). This effect of nest predation may not be detected in the variance estimates for males because fewer of them were trapped during the relevant period, and they are also less likely to be identified when predation occurs.

Our estimate of the additive genetic variance (and hence narrow-sense heritability) for both total fitness (lifetime recruitment) and lifetime fledging success was close to 0 in both males and females (table 2). Hence the standardized genetic variances, V_W, were also close to 0. However, while the standard errors on these estimates are reasonably small (0.05) from a statistical point of view, the 95% confidence intervals include values that would be of considerable biological significance if correct. For example, the upper bound of V_W for lifetime recruitment would be 0.13 for females and 0.12 for males. As Burt (1995, 2000) points out, V_A (and hence V_W), while theoretically attractive as a measure of natural selection, is extremely hard to determine accurately. It is worth noting that even using the animal model and having far more data than most of the studies cited by Burt (2000), the standard errors of our estimates of h^2 are almost exactly the same. Assuming the random sampling model applies here, this must mean that our data, though a larger sample, are taken from a population with a larger variance, so having more data does not necessarily solve the problem of lack of power in identifying values of V_W that are of biological importance.

We were unable to find a significant relationship between the association of a trait with fitness and its coefficient of additive genetic variation (CV_A) in either females (fig. 1*B*) or males (fig. 1*B*; table 3). These results are similar to Kruuk et al. (2000), who found no relationship between the CV_A of a trait and its association with fitness. Merilä and Sheldon (2000) found a positive relationship, but it was weak for males and nonsignificant for females if lifetime reproductive success (LRS) itself was excluded. It is debatable whether LRS can be included (and we have not done so here) because its correlation with itself is necessarily 1.0. Finally, in agreement with both Kruuk et al. (2000) and Merilä and Sheldon (2000), we found a strong positive relationship between the residual environmental variance affecting a trait and the association of these traits with overall fitness (fig. 1*C*, 1*F*; table 3). These last results thus support the hypothesis of Price and Schluter (1991) by demonstrating that lower narrow-sense heritabilities of traits more closely related to fitness are caused at least in part by higher residual variation. This may not be surprising when one considers that the top three traits in the ranking of associations with fitness, namely lifetime-fledged offspring, longevity, and first-year recruitment, are subject to environmental variation acting over much longer periods than the traits lower down the ranking, which are mainly affected by variation occurring within the breeding season.

Taken together, our results and those of Kruuk et al. (2000) and Merilä and Sheldon (2000) emphasize the important role of residual environmental variance in explaining the widely observed pattern of decreased h^2 for traits closely associated with total lifetime fitness relative to traits less closely associated with fitness. These results appear to be particularly robust given the differences between the three study taxa and because they were derived from wild populations in their natural environment. Kruuk et al. (2000) examined a relatively small, insular island population, while the Wytham great tit population is larger, and on average, 50% of the breeding population each year are immigrants to the wood. The great tit and collared flycatcher's ecology and population structure are similar, except that the flycatchers are migratory and great tits are not. The main difference between our study and that of Merilä and Sheldon (2000) is that they used parent-offspring regression to estimate their genetic parameters (V_A, and by extension CV_A and h^2), while we used an "animal model." As discussed by Merilä and Sheldon (2000), their data set included instances where parents and offspring were reared in the same environment, which may bias their estimates of V_A upward due to parent-offspring environmental covariance. Our analyses, which incorporated an annual area-specific effect (contemporary group as a fixed effect) and a random site effect (individual nest boxes), will be less sensitive to this source of bias in parameter estimation (Meyer 1985). Determining the relative contribution of nonadditive genetic variance (dominance and epistasis) and environmental variance to this residual variance provides an important challenge for future studies.

Acknowledgments

B.C.S. was funded by the Royal Society. We are grateful to anonymous reviewers for comments on the original version of this manuscript.

Literature Cited

Blakey, J. K. 1994. Genetic evidence for extra-pair fertilizations in a monogamous passerine, the great tit *Parus major*. Ibis 136:457–462.

Boag, P. T., and A. J. van Noordwijk. 1987. Quantitative genetics. Pages 45–78 *in* F. Cooke and P. A. Buckley, eds. Avian genetics: a population and ecological approach. Academic Press, London.

Brinkhof, M. W. G., A. J. Cave, F. J. Hage, and S. Verhulst. 1993. Timing of reproduction in the coot *Fulica atra*: evidence for a causal relationship. Journal of Animal Ecology 62:577–587.

Brommer, J., L. Gustafsson, H. Pietiänen, and J. Merilä. 2004. Single-generation estimates of individual fitness as a proxy for long-term genetic contribution. American Naturalist 163:505–517.

Burt, A. 1995. The evolution of fitness. Evolution 49:1–8.

———. 2000. Natural selection in the wild. Trends in Ecology & Evolution 15:306–307.

Cheverud, J. M., and A. J. Moore. 1994. Quantitative genetics and the role of the environment provided by relatives in the evolution of behavior. Pages 67–100 *in* C. R. B. Boake, ed. Quantitative genetic studies of behavior and evolution. University of Chicago Press, Chicago.

Cresswell, W., and R. H. 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.

Dunn, E. 1977 Predation by weasels (*Mustela nivalis*) on breeding tits (*Parus* spp.) in relation to the density of tits and rodents. Journal of Animal Ecology 46:633–652.

East, M. L., and C. M. Perrins. 1988. The effect of nestboxes on breeding populations of birds in broadleaved temperate woodlands. Ibis 130:393–401.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. 4th ed. Longman, Essex.

Fisher, R. A. 1930. The genetical theory of natural selection. Clarendon, Oxford.

Fowler, K., C. Semple, N. H. Barton, and L. Partridge. 1997. Genetic variation for total fitness in *Drosophila melanogaster*. Proceedings of the Royal Society of London B 264:191–199.

Gustafsson, L. 1986. Lifetime reproductive success and heritabilities: empirical support for Fisher's fundamental theorem. American Naturalist 128:761–764.

Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130:195–204.

Johnson, D. L., and R. Thompson. 1995. Restricted maximum likelihood estimation of variance components for univariate animal models using sparse matrix techniques and average information. Journal of Dairy Science 78:449–456.

Kruuk, L. E. B., T. H. Clutton-Brock, J. Slate, J. M. Pemberton, S. Brotherstone, and F. E. Guinness. 2000. Heritability of fitness in a wild mammal population. Proceedings of the National Academy of Sciences of the USA 97:698–703.

Kruuk, L. E. B., J. Merilä, and B. C. Sheldon. 2001. Phenotypic selection on a heritable size trait revisited. American Naturalist 158:557–571.

Lai, C., R. F. Lyman, A. D. Long, C. H. Langley, and T. F. C. Mackay. 1994. Naturally occuring variation in bristle number and DNA polymorphisms at the scabrous locus of *Drosophila melanogaster*. Science 266:1697–1702.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, Mass.

McCleery, R. H., J. Clobert, R. Julliard, and C. M. Perrins. 1996. Nest predation and delayed cost of reproduction in the great tit. Journal of Animal Ecology 65:96–104.

Merilä, J., and B. Sheldon. 2000. Lifetime reproductive success and heritability in nature. American Naturalist 155:301–310.

Meyer, K. 1985. Maximum likelihood estimation of variance components for a multivariate mixed model with equal design matrices. Biometrics 41:153–166.

———. 1988. DFREML: a set of programs to estimate variance components under an individual animal model. Journal of Dairy Science 71(suppl. 2):33–34 (abstract).

———. 1989. Restricted maximum likelihood to estimate variance components for animal models with several random effects using a derivative-free algorithm. Genetics Selection and Evolution 21:317–340.

———. 1998. DFREML, version 3.0β user notes. http://agbu.une.edu.au/~kmeyer/dfreml.html.

Mousseau, T. A., and D. A. Roff. 1987. Natural selection and the heritability of fitness components. Heredity 59: 181–197.

Nelder, J. A., and R. Mead. 1965. A simplex method for function minimization. Computer Journal 7:147–151.

Norris, K. 1993. Seasonal variation in the reproductive success of blue tits: an experimental study. Journal of Animal Ecology 62:287–294.

Perrins, C. M. 1965. Population fluctuations and clutch-size in the great tit (*Parus major*). Journal of Animal Ecology 34:601–647.

———. 1979. British tits. Collins, London.

———. 1998. A long-term study of great tits *Parus major*. Presidential address. Proceedings of the First Meeting of European Ornithologists Union, Bologna, 1997. Biologia y Conservazione della Fauna 102:55–62.

Perrins, C. M, and R. H. McCleery. 1985. The effect of age and pair bond on the breeding success of great tits *Parus major*. Ibis 127:306–315.

———. 1989. Laying dates and clutch size in the great tit. Wilson Bulletin 101:236–253.

Perrins, C. M., and D. Moss. 1975. Reproductive rates in the great tit. Journal of Animal Ecology 44:695–706.

Pettifor, R. A., C. M. Perrins, and R. H. McCleery. 1988. Variation in clutch-size in great tits: evidence for the individual optimisation hypothesis. Nature 336:160–162.

———. 2001. The individual optimisation of fitness: variation in reproductive output, including clutch-size, mean nestling mass and offspring recruitment, in manipulated broods of great tits (*Parus major*). Journal of Animal Ecology 70:1–18.

Price, T., and D. Schluter. 1991. On the low heritability of life-history traits. Evolution 45:853–861.

Roff, D. A. 1997. Evolutionary quantitative genetics. Chapman & Hall, New York.

Roff, D. A., and T. A. Mousseau. 1987. Quantitative genetics and fitness: lessons from *Drosophila*. Heredity 58: 103–118.

Service, P. M., and M. R. Rose. 1985. Genetic covariation among life-history components: the effects of novel environments. Evolution 39:943–945.

Shaw, R. G. 1987. Maximum-likelihood approaches to quantitative genetics of natural populations. Evolution 41:812–816.

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.

Verhulst, S., J. H. van Balen, and J. M. Tinbergen. 1995. Seasonal decline in the reproductive success of the great tit: variation in time or quality? Ecology 76:2392–2403.

Wolf, J. B, and M. J. Wade. 2001. On the assignment of fitness to parents and offspring: whose fitness is it and when does it matter? Journal of Evolutionary Biology 14:347–356.

Associate Editor: Trevor Price