Ann. Zool. Fennici 40: 269–280 ISSN 0003-455X
Helsinki 18 June 2003

Heterotic effects on fitness in a wild bird population

Juha Merilä[1], Ben C. Sheldon[2,3] & Simon C. Griffith[3]

[1)] *Ecological Genetics Research Unit, Department of Ecology and Systematics, P.O. Box 65, FIN-00014 University of Helsinki, Finland (e-mail: juha.merila@helsinki.fi)*
[2)] *Department of Animal Ecology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 d, SE-752 36 Uppsala, Sweden*
[3)] *Edward Grey Institute, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK*

Received 3 Sep. 2002, revised version received 7 Oct. 2002, accepted 8 Oct. 2002

Merilä, J., Sheldon, B. C. & Griffith, S. C. 2003: Heterotic effects on fitness in a wild bird population. — *Ann. Zool. Fennici* 40: 269–280.

The role of genetic variation in determining fitness in natural populations has been enigmatic for decades. Both theoretical and empirical work suggest that additive genetic contributions to fitness variation are small, whereas observations of inbreeding depression suggest that non-additive genetic contributions to fitness can sometimes be large. We analysed associations between genetic variability at a small number of microsatellite loci and fitness in a natural population of the collared flycatcher *Ficedula albicollis*, and related these to estimates of additive genetic contributions to the same traits. We found strong single-locus associations between an intra-locus measure of variability (d^2) and variables related to fitness. These heterotic effects were sex-specific, being found only in males, and variation at this locus explained as much as 11% of the variance in male lifetime reproductive success. The size of the heterotic effect on a trait depended on how closely that trait was related to fitness, and there was a negative relationship between the size of the heterotic effect and the proportion of a trait's variance due to additive genetic variance. One possibility is that the heterosis results from introgression due to hybridization with the closely related pied flycatcher *F. hypoleuca*. Our results provide evidence that genetic contributions to fitness can be important in outbred populations.

Introduction

Traits closely associated with fitness are known to exhibit lower heritabilities than traits weakly associated with fitness (Gustafsson 1986, Mousseau & Roff 1987, Kruuk *et al*. 2000, Merilä & Sheldon 2000). This has led to the view that most fitness variation in natural populations has little to do with genetic differences among individuals. However, observations of inbreeding depression (Keller *et al*. 1994, Falconer & Mackay 1996, Lacy *et al*. 1996, Lynch & Walsh 1998), and heterozygote superiority (Allendorf & Leary 1986, Mitton 1994, Britten 1996, Bierne *et al*. 1998, Slate *et al*. 2000) suggest that, under some conditions, genetic differences may contribute importantly to fitness variation. The low heritability of traits closely related to fitness seems to result from elevated residual variance (V_R), rather than from low additive genetic variance (V_A,

where total phenotypic variance, $V_P = V_A + V_R$), since coefficients of additive genetic variance may be quite large for traits closely related to fitness (Houle 1992, Kruuk *et al.* 2000, Merilä & Sheldon 1999, 2000). Residual variance can in theory be further partitioned into terms due to environmentally caused variation, genotype by environment interaction and non-additive genetic variance (dominance and epistasis). However, due to the difficulty of partitioning phenotypic variance into different genetic components, the relative importance of additive and non-additive genetic contributions to fitness in natural populations is largely unknown (Merilä & Sheldon 1999). A study by Crnokrak and Roff (1995), which found that dominance variance accounted for an average of 30% of all phenotypic variance in life history traits of wild outbred species suggests the potential importance of dominance variance. Equally, the fact that fitness-related traits may be influenced by many loci suggests that epistasis (interactions between alleles at different loci) could be an important source of genetic variation for fitness (Lynch & Walsh 1998, Wolf *et al.* 2000).

In this paper we explore associations between a microsatellite-specific measure (d^2) of genetic variation and a number of different life history and morphological traits in a natural population of collared flycatchers (*Ficedula albicollis*). Our interest is, therefore, in asking whether there are associations between genotypes and phenotypes related to fitness. Recent studies of free-ranging bird and mammal populations indicate that mean d^2 can sometimes be a sensitive method for revealing associations between degree of outbreeding and fitness (Coulson *et al.* 1998, 1999, Coltman *et al.* 1998, Amos *et al.* 2001, Hansson *et al.* 2001, Rossiter *et al.* 2001, Höglund *et al.* 2002, though *see* Hedrick *et al.* 2001, Rowe & Beebee 2001, Slate & Pemberton 2002). d^2 is the squared difference in the number of repeat units of the two alleles at a microsatellite locus (Goldstein *et al.* 1995), and its use as a measure of inbreeding is based on the assumption that microsatellites evolve by mutations occurring in a step-wise fashion. If this assumption is correct, or approximately so, the difference in repeat number between two alleles will be related to the time since they shared a common ancestor (Goldstein *et al.* 1995). We first provide evidence to suggest that genotypic variation at a single locus microsatellite maker is strongly linked to individual fitness differences. We then demonstrate that there is a positive relationship between the strength of this relationship and the proportion of variance explained in lifetime reproductive success across different traits, suggesting that while traits closely associated with fitness tend to have low heritabilities, non-additive genetic effects can be substantial.

Materials and methods

Study species and data collection

The collared flycatcher is a small cavity-nesting insectivorous migratory passerine bird native to Europe. Due to its ready acceptance of nest boxes as breeding sites, and high natal and breeding philopatry (Pärt & Gustafsson 1989) on the southern part of the island of Gotland (57°10´N, 18°20´E, Sweden), it has become a model organism for studies in which estimates of individual fitness are required (Gustafsson & Sutherland 1988, Gustafsson & Pärt 1990, Merilä & Sheldon 2000). Data on variation in external morphology and breeding performance have been collected for a large number of individuals (almost 9000) breeding in this population from 1980 to date; data in the present paper were collected up to 1999. For most breeding attempts clutch initiation date (= laying date), clutch size and number of fledged young were recorded, and the majority of adults were measured for tarsus, beak, wing, tail, and first primary length as described in Merilä and Gustafsson (1993). In males the size of the white forehead patch (height and width: Gustafsson *et al.* 1995) was also measured. As many individuals were measured in several years, we calculated the mean value for each trait for each individual, correcting for year and age-associated variation using ANOVAs. Individuals were identified by numbered aluminium rings applied when they were first captured (in many cases as nestlings before fledging). Blood samples, used

for subsequent genetic analyses, were collected between 1993 and 1998 when parents were caught feeding nestlings in nest-boxes, and stored in a variety of buffers and at a number of different temperatures for up to five years before analysis.

In order to determine individual lifespan, lifetime fledging success (henceforth LFS) and lifetime reproductive success (henceforth LRS), we searched our data base covering the years 1980–1999 for records of all adults and their offspring included in this study. Lifespan was defined as the maximum age (to the nearest year) that a given individual had attained when it was last captured, LFS as the total number of offspring fledged from all breeding attempts over an individual's lifetime and LRS as the total number of offspring recruited to the breeding population over an individual's lifetime. In the present paper we obtained genotype data for adults captured between 1993 and 1998, but we restricted our analyses involving these three lifetime measures of fitness to birds sampled in 1993–1996. This was done because a relatively large proportion of individuals sampled in 1997 and 1998 were still alive in 1999. A small proportion of individuals (n = 7 males, 2 females) in the sample of adults from 1993–1996 were known to be still alive in 1999; these were excluded when calculating the three fitness components. For analyses of LRS we also excluded individuals alive in 1998 (n = 18 males, 7 females) as data on recruitment of these individuals' offspring were not complete (some offspring do not recruit until 2 years of age). Inclusion of these individuals decreases the effect of the genetic markers found to be correlated with fitness; this is to be expected given mutual correlations with lifespan. For analyses of relationships between microsatellite variability and morphological variation we used all data at our disposal, as these measures should be unaffected by whether an individual is currently alive or not. Due to the high degree of breeding and natal philopatry exhibited by this population (*see* above), as well as the high capture effort over the years 1993–1999, the estimates of lifetime measures of fitness are probably rather accurate. For example, the mean LRS for males in our sample was 1.68, and for females 1.43, both quite close to the expected value of 2.0 in a stable population.

Estimation of variance components

Heritabilities (and hence variance components) were estimated using (mean) male offspring–male parent, and (mean) female offspring–female regressions in which the slope of the regression estimates half of the heritability (Falconer & Mackay 1996). These estimates were corrected for assortative mating which was, in any case, weak for the traits considered here (Merilä & Sheldon 2000). Most traits also exhibited marked age and cohort-specific variation, so individual trait values were corrected for age and cohort specific variation (i.e. year of birth) using two-factor ANOVAs before estimating the heritabilities. Heritability estimates for morphological traits in this population have been shown to be unconfounded by common environment effects (Gustafsson & Merilä 1994), but part of the resemblance in life history traits may result from environmental correlations between parents and offspring (*see* Merilä *et al*. 2001, Sheldon *et al*. 2003 for examples). Further details of the analytical methods used to obtain heritability estimates and estimates of lifetime reproductive success are given elsewhere (Merilä & Sheldon 2000). Heritability (h^2) estimates the proportion of phenotypic variance due to additive genetic effects (i.e. $h^2 = V_A/V_P$), but as it is calculated as a ratio, low heritability can result from either low V_A or high V_P. Consequently, heritability does not reveal very much about the amount of additive genetic variance for a trait. Houle (1992) has suggested that a more appropriate metric for comparing the size of genetic variance components can be obtained by scaling the variance by the mean; the coefficient of additive genetic variance (CV_A) is thus calculated as: $CV_A = 100[V_A/x]^{0.5}$, where V_A is twice the single-parent offspring covariance (i.e. the additive genetic variance) obtained from a single parent–offspring regression and x is the mean of the parents used to estimate V_A. A similarly standardised measure of the residual variance (CV_R), where $V_R = V_P - V_A$, can be calculated as: $CV_R = 100[(V_P - V_A)/x]^{0.5}$.

Microsatellite genotype data

Genotypes at three previously described microsatellite loci (*FhU2*, *FhU3*, *FhU4*; Primmer *et al*. 1996) were obtained for 123–191 males and 109–156 adult females that bred in the study area between 1993–1998. *FhU2* and *FhU3* are dinucleotide repeats, whereas *FhU4* is a tetranucleotide repeat; all three markers show relatively typical degrees of variation for avian microsatellite loci (Table 1). Allele size determination and general procedures used in genotyping are described in Primmer *et al*. (1996). Most genotypes were obtained in the course of investigations of parentage (Sheldon & Ellegren 1998, 1999, Sheldon *et al*. 1999); an additional 88 adults (39 males, 49 females) were genotyped at *FhU4* after we had uncovered relationships between variability at this locus and fitness measures. d^2 values were calculated as the squared difference in allele size within each locus after dividing by the size of the repeat unit. Mean heterozygosities and mean squared distances (d^2) for the three microsatellite loci are given in Table 1. Observed heterozygoties at the three loci do not differ from Hardy-Weinberg expectation (Sheldon & Ellegren 1996).

Analysis

Associations between d^2 values and different life history and morphological traits were investigated using two different approaches. First, we calculated mean d^2 over the three loci for each individual and used mean d^2 as the independent variable in linear regressions of trait values on mean d^2. Second, we performed separate analyses for each of the loci to investigate their effects on traits of interest. Given the non-normal error-structure of these models, the statistical significance testing of regression coefficients was done using a randomisation approach (5000 permutations of data) using *Resampling Stats* package (Simon 1997). The results are robust to the exclusion of homozygotes at each of the loci, which were few in number in any case: thus, the patterns reported here do not reflect a difference between homozygotes and heterozygotes at the loci in question. As we conducted multiple tests for the relationship between d^2 and phenotypic traits at each locus, we corrected for the inflated type I error rate using a sequential Bonferroni approach (Rice 1989), applied at the family-of-tests level (i.e. within each group of ten tests relating d^2 at a particular locus to phenotypic variation). To investigate relationships between trait-specific d^2 values, heritability and CV_R, all comparisons where sample sizes were low, we used a randomization approach (5000 permutations) to obtain P values for correlation coefficients. An alternative assumption-free approach is to calculate rank correlation coefficients between these variables: our results are unchanged in terms of statistical significance and interpretation if this procedure is followed (data not shown). In order to investigate the interrelationship between d^2 and life-history traits contributing to lifetime reproductive success we constructed path-analytic models. This technique

Table 1. Mean observed heterozygosity (H_o) and mean squared allele difference (d^2) for three microsatellite loci in male (m) and female (f) collared flycatchers.

Locus	*N* alleles	Allele size range (bp)†	Sex	H_o	Mean d^2 ± SE
FhU2	16	138–184	m	0.902	61.62 ± 7.43
	16	142–186	f	0.862	63.70 ± 8.47
FhU3	7	151–177	m	0.656	6.56 ± 1.74
	7	151–177	f	0.807	8.78 ± 1.90
FhU4	12	170–201	m	0.848	33.37 ± 3.50
	12	166–198	f	0.878	26.49 ± 3.00
All loci			m		34.58 ± 3.09
			f		34.44 ± 3.36

†There were no significant differences in allele frequencies among the sexes (*FhU2*: $\chi^2_{18} = 20.85$, $P = 0.29$; *FhU3*: $\chi^2_7 = 10.83$, $P = 0.15$; *FhU4*: $\chi^2_{15} = 23.22$, $P = 0.08$).

is held to be particularly sensitive to non-normal error structure, but residuals from models which included lifespan were normally distributed (Shapiro-Wilk $W \geq 0.96$, $P > 0.05$).

Hence, correlations between d^2 and phenotypic traits were analysed separately for the two sexes. Other reasons for separating the sexes in this way were that some traits were measured in only one sex, and a previous study had found sex-differences in heterosis revealed by d^2 (Coulson *et al.* 1999).

Results

Descriptive results

Correlations between d^2 values at individual loci within individuals were small and not statistically significant ($-0.070 \leq r \leq 0.075$, $n \geq 218$, $P > 0.26$; both sexes pooled). There were no differences in mean d^2 at individual loci when comparing the sexes (*FhU2*: $t_{230} = 0.19$, $P = 0.85$; *FhU3*: $t_{232} = 0.86$, $p = 0.39$; *FhU4*: $t_{344} = 1.49$, $P = 0.14$; mean d^2 across loci: $t_{217} = 0.03$, $P = 0.98$), nor did the variance in d^2 differ between the sexes for *FhU2*, *FhU3* or mean d^2 across loci (Levene's test: $F \leq 1.33$, $P \geq 0.25$). However, the variance of d^2 at *FhU4* in males was significantly higher than in females (Levene's test: $F_{1,345} = 6.84$, $P = 0.009$).

Within locus comparisons — univariate tests

There was no evidence for an association between mean d^2 and trait values in females, but evidence for weak positive correlations between mean d^2 and fitness components in males, none of which were significant after sequential Bonferroni correction (Table 2). Analysing each locus separately revealed that all of the effects found on fitness components for mean d^2 in males were due to highly significant positive correlations between d^2 at *FhU4* and the three fitness components (lifespan, LFS and LRS, respectively; Table 2 and Fig. 1). In addition, there was a negative

Table 2. Relationships between d^2 at three microsatellite loci and phenotypic traits in male and female collared flycatchers. Values are regression coefficients from linear regressions ($\times 10^3$). Values in boldface are those significant at $P < 0.05$ after test-family-wide sequential Bonferroni correction.

Trait	Mean d^2	*FhU2*	*FhU3*	*FhU4*
Females				
First primary length	1.2	–2.7*	12.8	0.8
Tarsus length	–0.5	–0.2	–4.0	–3.0*
Bill length	1.2	–0.4	–1.7	–0.6
Tail length	–4.6	–0.2*	1.0	5.0
Wing Length	–3.1	4.3	1.5*	4.3
Laying Date	–0.5	0.6	–10.8	–4.7
Clutch Size	1.9	0.7	0.4	2.2
Lifespan	1.6	1.2	–2.0	–6.8*
LFS	4.9	3.5	29.6	–0.8
LRS	–4.7	0.5	–3.2	–1.6
Males				
First primary length	–12.2*	**–7.4****	14.0	0.3
Tarsus length	–1.5	0.5	–5.4	**–1.8****
Bill length	3.5	–1.9*	–2.9	0.2
Tail length	6.8	–2.3	23.0	3.3
Wing Length	–9.1	–4.3*	12.2	–4.8
Forehead patch height	–4.5	1.0	3.8	–0.5
Forehead patch width	6.5	1.4	–1.4	–1.5*
Lifespan	20.0**	1.5	–16.8	**7.1****
LFS	54.8	–6.0	103*	**34.7****
LRS	19.9*	–1.0	–2.4	**10.4*****

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

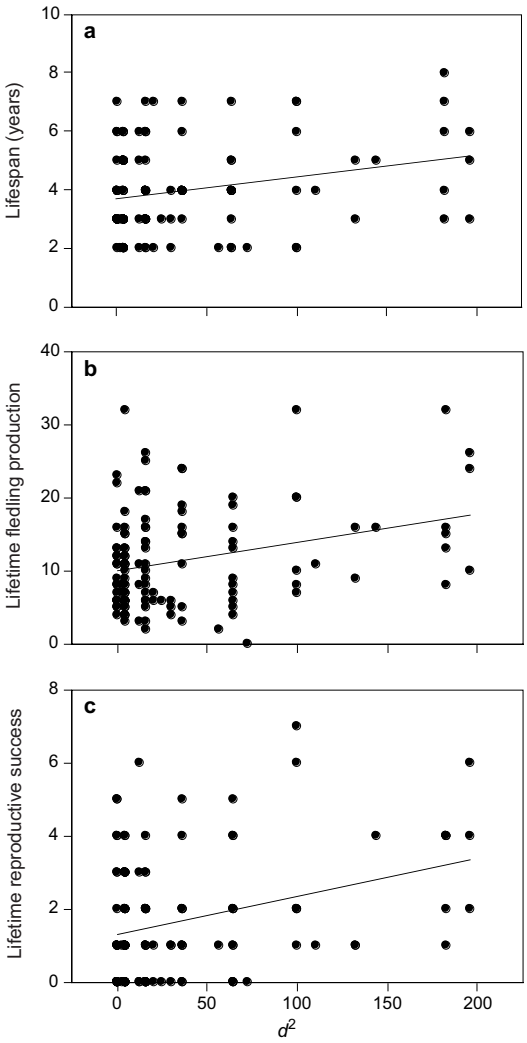

Fig. 1. Relationships between (**a**) lifespan, (**b**) lifetime fledling production, and (**c**) lifetime reproductive success and d^2 in locus *Fhu4* in male collared flycatchers. Each point represents one individual (some points hidden due to overlap). All regressions are significant (*see* Tables 2 and 3 for statistical tests), and there is no evidence for curvlinearity in the relationships (quadratic terms in all cases: $t \leq 0.62$, $P \geq 0.54$).

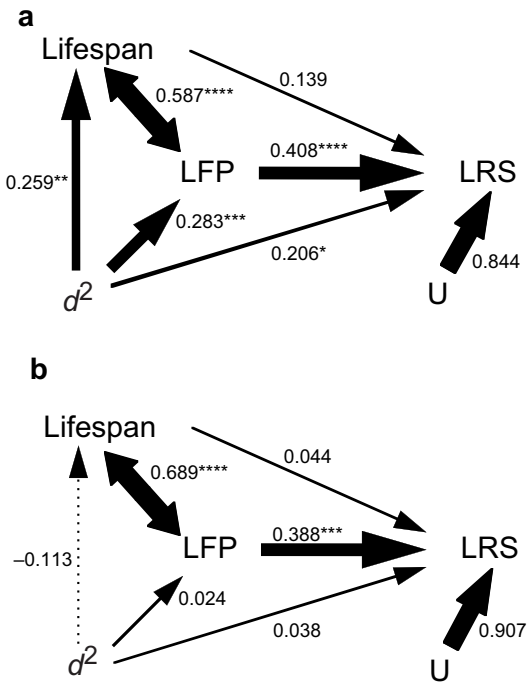

Fig. 2. Path diagrams summarizing relationships among different fitness components and d^2 at *FhU4* in (**a**) male and (**b**) female collared flycatchers. Path coefficients (standardised partial regression coefficients) are given beside the arrows (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$), the thickness of which corresponds to the size of the path coefficient; dashed lines indicate negative path coefficients. Single-headed arrows indicate causal relationships. Path *U* represents unexplained variation, calculated as $(1 - r^2_{\text{total}})^{0.5}$.

correlation between d^2 at *FhU4* and tarsus length in males and a non-significant (after Bonferroni correction) negative correlation between d^2 at *FhU4* and male forehead patch width (Table 2). Five significant correlations between phenotypic traits and single locus d^2 values were found in females, but none of these remained significant after Bonferroni correction (Table 2).

Within locus comparisons — multivariate tests

Both LFS and LRS are strongly influenced by lifespan (within this sample of males: lifespan *vs.* LFS: $r_{124} = 0.66$, $P < 0.0001$; lifespan *vs.* LRS: $r_{115} = 0.40$, $P < 0.0001$). As a consequence, the correlations with d^2 could be driven by a strong correlation with only one of these variables. To explore this possibility, we constructed path analytic models (Fig. 2) with these three variables and d^2 to determine the inter-relationships among these traits. These models revealed that d^2 at *FhU4* affects fitness (LRS) via three independent pathways: (i) via increased lifespan, (ii) via increased LFS, irrespective of lifespan (partial $F_{1,111} = 4.02$, $P = 0.048$) and (iii) via increased recruitment irrespective of LFS (Fig.

2a). Thus, the relationships between d^2 and the three fitness components are not driven simply by a strong correlation with one of them alone: variation at this locus has independent additive effects on the three separate traits. In marked contrast to our findings for males, there were no associations between d^2 values and any of the traits in females, and thus no evidence that d^2 at *FhU4* contributed independently to variation in the three fitness measures (Fig. 2b). The sex difference in relationship between d^2 and proportion of variance explained in LRS is statistically significant for LRS ($F_{1,242} = 5.86$, $P = 0.016$) and Lifespan ($F_{1,248} = 8.53$, $P = 0.004$), and nearly so for LFS ($F_{1,244} = 3.43$, $P = 0.065$). The strength of the relationship that we uncovered between d^2 and LRS, the best measure of fitness that we have for this population in males, is surprisingly strong, explaining 11% of the variance in LRS. In comparison, LFS, the best phenotypic predictor of LRS explains 44% of the variance in LRS.

Among trait comparisons

In males, our analyses revealed striking patterns as regards the importance of d^2 at *FhU4* for explaining variance in a trait and that trait's contribution to fitness, and also between d^2 and genetic variance components. There was a positive relationship between the proportion of variance in LRS that a trait explained, and the proportion of that trait's variance that was explained by d^2 at *FhU4* (Fig. 3a: $r = 0.88$, $P = 0.0008$). Thus d^2 at this locus was more influential the more closely a trait was related to overall fitness. As expected, because of the negative relationship between a trait's relation-

Table 3. Correlation between life history and morphological traits and d^2 values at the locus *FhU4* together with details of their heritabilities (h^2) and the proportion of variance in fitness (LRS) that each trait explains in male and female collared flycatchers. r = Pearson product moment correlation coefficient, b (± s.e.) = the slope ($\times 10^3$) of regression of given trait on d^2, and r^2 = the coefficient of determination from this regression. r^2_{LRS} is the proportion variance in LRS explained by each trait. Data for heritabilities and r^2_{LRS} from Merilä and Sheldon (2000). LFP = lifetime fledgling production; LRS = lifetime reproductive success. *P* values associated with Pearson correlation coefficients were determined by randomisation.

Trait	r	r^2	b ± s.e.	N	h^2	n_{h2}	r^2_{LRS}
Males							
Tarsus length	–0.176**	0.031	–1.76 ± 0.84	138	0.47	782	0.0004
Bill length	0.014	0.000	0.17 ± 1.15	110	0.47	688	0.0000
Tail length	0.109	0.012	3.25 ± 2.79	115	0.35	688	0.0001
First primary length	0.009	0.000	0.27 ± 2.91	111	0.57	667	0.0009
Wing length	–0.014	0.000	–4.80 ± 2.98	141	0.29	795	0.0066
Forehead patch height	–0.025	0.001	–0.45 ± 1.57	141	0.45	765	0.0003
Forehead patch width	–0.094*	0.009	–1.52 ± 1.36	141	0.41	766	0.0014
Lifespan	0.239**	0.057	7.13 ± 2.51	135	0.15	826	0.1511
LFP	0.253**	0.064	34.71 ± 11.58	133	0.02	782	0.2410
LRS	0.324***	0.105	10.40 ± 2.80	125	0.07	652	1.0000
Females							
Tarsus length	–0.101	0.010	–3.01 ± 2.67	126	0.53	866	0.0016
Bill length	–0.047	0.002	–0.59 ± 1.25	99	0.41	768	0.0011
Tail length	0.110	0.012	5.03 ± 4.59	101	0.44	773	0.0062
First primary length	0.017	0.000	0.84 ± 5.08	97	0.59	760	0.0000
Wing length	0.057	0.003	4.34 ± 6.77	126	0.47	885	0.0108
Clutch size	0.123	0.015	2.16 ± 1.57	126	0.35	927	0.0058
Laying date	–0.041	0.002	–4.72 ± 10.29	126	0.41	917	0.0114
Lifespan	–0.155*	0.024	–6.77 ± 3.90	124	0.00	828	0.1843
LFP	–0.005	0.000	–0.84 ± 15.86	122	0.01	872	0.2799
LRS	–0.039	0.002	–1.60 ± 3.74	120	0.21	719	1.0000

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

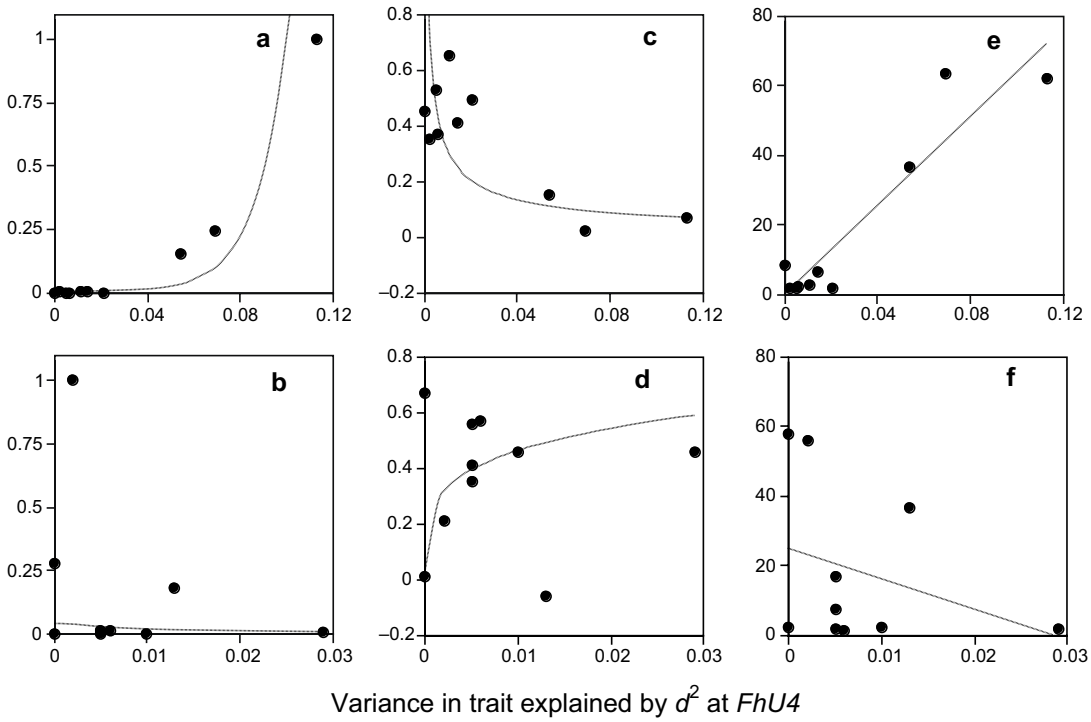

Fig. 3. Relationships between the proportion of variance in a trait explained by d^2 at *FhU4*, and quantitative genetic and fitness parameters. — **a** and **b**: Relationship between proportion of variance in lifetime reproductive success (LRS) explained by trait, and the proportion of variance in trait explained by d^2 at *FhU4* in males and females, respectively. Lines show best-fit exponential curves for ease of interpretation; — **c** and **d**: Relationship between heritability of a trait and proportion of variance in trait explained by d^2 at *FhU4* in males and females, respectively. Lines show best-fit power curve; — **e** and **f**: Relationship between coefficient of residual variance for a trait (CV_R) and proportion of variance in trait explained by d^2 at *FhU4* in males and females respectively. Lines show best-fit linear functions for ease of interpretation.

ship to fitness and its heritability in this population (Gustafsson 1986, Merilä & Sheldon 2000), the relationship between heritability and the proportion of variance explained by d^2 at *FhU4* was negative, (Fig. 3c: $r = -0.83$, $p = 0.006$). Hence, there was a negative relationship between the proportion of a trait's variance due to additive genetic sources and that explained by d^2 at *FhU4*. Low heritability of traits closely related to fitness generally occurs because these traits have very high residual variance (V_R), relative to additive genetic variance (V_A; Merilä & Sheldon 1999). There was a highly significant positive relationship between a standardised measure of the residual variance (CV_R) and the proportion of variance in a trait explained by d^2 at *FhU4* (Fig. 3e: $r = 0.91$, $P = 0.001$). Quite different patterns were found among females, where there was no significant relationship between the amount of variance explained by d^2 at *FhU4* and either a trait's relationship to LRS, or its heritability, or the coefficient of residual variance for that trait (Fig. 3b: $r = -0.17$, $P = 0.42$; Fig. 2d: $r = -0.36$, $P = 0.15$, and Fig. 3f: $r = -0.06$, $P = 0.48$ respectively).

Discussion

We found that an individual-specific measure of within-locus allelic diversity at one microsatellite locus was strongly related to variation in fitness components in male collared flycatchers, with no such effect apparent in females. Males with alleles that differed in size by a large amount at this locus had longer lifespans, fledged more young in each year of their lives, and a higher

proportion of these young recruited to the breeding population. As the measure that we used, d^2, is determined by the difference between the two alleles at a locus, it will not be directly transmitted to offspring. Thus, as a component of genetic variance, the effect that we detect seems likely to be at least partially non-additive. In agreement with this, we found a strong negative correlation between the effect of d^2 on a trait and that trait's heritability, and a strong positive correlation between the effect of d^2 on a trait and that trait's coefficient of residual variance (a term which includes non-additive genetic effects). These analyses tentatively suggest that, in male collared flycatchers, the relative importance of additive and non-additive genetic variation in determining variation in a trait are inversely related, and also that non-additive genetic variance contributes disproportionately to traits that are closely related to fitness. A similar conclusion can be drawn from the positive relationship found between the magnitude of inbreeding depression for a trait, and that trait's partial correlation with fitness in this population (Kruuk *et al*. 2002). In that study it was found that close inbreeding reduced fitness by 94% relative to non-inbred birds, suggesting a substantial genetic load in this population. An alternative possibility, which we cannot discount at present, is that recent hybridization with pied flycatchers *F. hypoleuca* (*see* below) has introduced additive genetic variance for fitness traits to the collared flycatcher population.

Genetic contributions to individual variation in fitness in wild outbred populations are generally believed to small, as most fitness related traits are known to have low heritability (Merilä & Sheldon 1999). For example, the most important determinant of lifetime reproductive success (LRS) in a wide range of vertebrates is lifespan, because individuals surviving through many breeding attempts will produce more offspring than those breeding fewer times (Clutton-Brock 1988, Newton 1989). Over a long period of time, it would seem reasonable that stochastic events (climate, variation in food supply, predation) should play a major role in determining an individual's lifespan, and hence that the great majority of variation in LRS would be due to environmental sources. Our findings suggest that this is not always true: despite screening a very small number of microsatellite loci, we were able to derive a metric which had independent effects on both lifespan and lifetime reproductive success. Together with other recent evidence from natural bird populations (Keller *et al*. 1994) and other taxa (e.g. Ross *et al*. 1996, Keller & Ross 1999, Slate *et al*. 2000) these findings challenge the view that the great majority of variation in fitness in nature arises from non-genetic sources. Although stochastic events may indeed be important as a source of selection in natural populations, their stochasticity does not rule out a role for genetic variation in determining which individuals survive those events.

What explains these findings? In the present study, intrinsic heterozygote advantage is unlikely to explain the observed association between fitness components and d^2 values, since the exclusion of homozygotes did not change our conclusions. Two other explanations for heterosis have been proposed: associative overdominance and genotypic association. Associative overdominance refers to an apparent heterozygote advantage due to associations between marker loci and other selected loci, while genotypic association occurs when non-random mating or inbreeding creates associations between marker genotypes and fitness (*see* Houle 1989, Savolainen & Hedrick 1995). Of these, we favour the latter explanation. Although the effects observed at the *FhU4* locus could be understood in terms of close association between this locus and some other functionally important locus or loci, this would require very high linkage disequilibrium between the two loci. It seems improbable that in selecting only three microsatellite loci we would by chance have selected one with high linkage to a locus with such a large effect on fitness, given the low density yet relatively large absolute number (> 30 000) of microsatellites in avian genomes (Primmer *et al*. 1997).

The collared flycatcher population on Gotland displays typical degrees of genetic variation in microsatellite, allozyme and mtDNA markers (Gelter *et al*. 1989, Tegelström & Gelter 1991) and quantitative traits (Merilä & Gustafsson 1993, Gustafsson & Merilä 1994, Merilä & Sheldon 2000), and close inbreeding is quite rare (only 1% of mating events: Kruuk *et al*.

2002). These observations might suggest little opportunity for genotypic association in this population. However, an additional opportunity for genotypic association arises due to hybridization between collared flycatchers and their close congener the pied flycatcher. The two species hybridize relatively frequently on the island of Gotland (2.6% of pairings are mixed: Alatalo *et al*. 1990, Veen *et al*. 2001), and while hybrid females are usually sterile, the majority of first generation hybrid males are fertile (Gelter *et al*. 1992, Veen *et al*. 2001). There is thus the potential for introgression to occur due to backcrosses between F1 hybrid males and female collared flycatchers (1.8% of pairings in this population: Veen *et al*. 2001). Sætre *et al*. (2003) demonstrate that introgression from pied to collared flycatchers has occurred in this population, though restricted to the autosomes, by assaying variation at single nucleotide polymorphisms that are species-specific in allopatric populations.

The heterosis found is not, however, a straightforward case of F1 'hybrid vigour'. None of the individuals used in the analyses were F1 hybrids (based on phenotypic characters and on genotypes at a microsatellite marker with fixed species-specific alleles: *see* Sætre *et al*. 2001 for further details). Nor, as far as we are aware, were any the offspring of F1 hybrid males backcrossed to female collared flycatchers (F1 hybrid females are sterile). For 37 of the males for which relationships between d^2 at *FhU4* and fitness were investigated, we knew, from pedigree data, the species identity of both of the parents. In all cases, both parents were phenotypically pure collared flycatchers. Although this sample represents only 30% of the total sample of males, the patterns as regards associations between d^2 and phenotypic traits were strikingly similar in this sample to those in the entire sample ($r_S = 0.94$, $n = 10$, $P < 0.0001$). It therefore appears that any hybridization events leading to the introgression causing the heterosis must have occurred before the grandparental generation.

One striking finding was that the heterosis that we detected occurred only in males. Effects of inbreeding on fitness have repeatedly been shown to be environmentally dependent such that the effects tend to be more pronounced when poor conditions prevail (Pray *et al*. 1994, Hauser & Loeschcke 1996, Lynch & Walsh 1998), although no evidence for annual variation in the magnitude of inbreeding depression was found in this population (Kruuk *et al*. 2002). In this light, one could view the two sexes as representing two 'environments' in which the same genes are expressed. From this perspective, the finding that heterosis was sex-dependent may not be that surprising. The two sexes in this species differ widely in their ecology, and also in the way in which natural selection acts on morphological traits (Merilä *et al*. 1997). Because sexual selection acts more strongly on males than on females (Sheldon & Ellegren 1999, Clutton-Brock 1988), one might therefore view males as an 'environment' in which increased stress causes greater expression of heterosis. A recent study of red deer *Cervus elaphus* has also demonstrated sex-dependent heterosis, derived from d^2 values (Coulson *et al*. 1999 — *see* also Rossiter *et al*. 2001), and two other studies suggest that inbreeding depression may be amplified in males relative to females (Slate *et al*. 2000, Meagher *et al*. 2000). An alternative explanation for the sex-dependent heterosis is that heterotic effects in females are counteracted by reduced fertility (and hence LRS) or viability owing to the effects of Haldane's rule, according to which the heterogametic sex in hybrid crosses (females in birds) is the sex that suffers most reduced fertility or viability (Orr 1997). In accordance with this suggestion, female lifespan was weakly negatively related to d^2 at *FhU4* (Table 2). The finding that the variance in d^2 at *FhU4* was higher for males than for females (with a trend in the same direction for the means: Table 1) is also consistent with the sex-biased heterosis seen, since this will result in males with large value of d^2 surviving for longer than females with large values of d^2.

In conclusion, our results substantiate the view that while traits closely related with fitness may have low heritabilities, this does not imply that genetic influences on fitness must be small. In our data, there appears to be negative relationship between the trait's heritability and the degree to which heterosis affects that trait. Whether variation in fitness-related traits showing low heritabilities is generally associated with loci showing strong heterosis remains a challenge to be verified in future studies. Likewise, although

more direct estimates of non-genetic influences on fitness are required, our results suggest that additive and non-additive genetic influences on fitness may be inversely correlated.

The data on microsatellite variation that we analysed in this paper were collected for another purpose, and for that reason a small number of loci were employed. In that sense, it might be thought remarkable that such a strong genetic signal can be detected in our data, and indeed the results might represent a serendipitous choice of a marker closely linked to a QTL with major effects on fitness in males. On the other hand, whether such effects are common or rare in natural populations is an empirical question, and one that can partly be addressed by using data on allele frequencies that already exist.

Acknowledgements

We thank Marie Björklund, Petri Clusius, Pete Davidson, Christer Hemborg, Robert Przybyło and Anna Qvarnström for help. Thanks are also due to Lars Gustafsson for introducing us to the collared flycatcher project, and to Laurent Keller for comments on a previous version of the manuscript. This work was supported by the Swedish Natural Science Research Council (grants to J.M. and B.C.S) and Finnish Academy (J.M). B.C.S. is a Royal Society University Research Fellow, and S.G. a NERC Research Fellow.

References

Alatalo, R. V., Eriksson, D., Gustafsson, L. & Lundberg, A. 1990: Hybridization between pied and collared flycatchers — sexual selection and speciation theory. — *J. Evol. Biol.* 3: 375–389.

Allendorf, F. W. & Leary, L. 1986: Heterozygosity and fitness in natural populations of animals. — In: Soulé, M. (ed.), *Conservation biology: the science of scaricity and diversity*: 57–76. Sinauer Associates, Sunderland, MA.

Amos, W., Wilmer, J. W., Fullard, K., Burg, T. M., Croxall, J. P., Bloch, D. & Coulson, T. N. 2001: The influence of parental relatedness on reproductive success. — *Proc. R. Soc. Lond. B* 268: 2021–2027.

Bierne, N., Launey, S., Naciri-Graven, Y. & Bonhomme, F. 1998: Early effect of inbreeding as revealed by microsatellite analyses on *Ostrea edulis* larvae. — *Genetics* 148: 1893–1906.

Britten, H. B. 1996: Meta-analyses of the association between multilocus heterozygosity and fitness. — *Evolution* 50: 2158–2164.

Clutton-Brock, T. H. 1988. *Reproductive success*. — Chicago Univ. Press, Chicago.

Crnokrak, P. & Roff, D. A. 1995: Dominance variance: associations with selection and fitness. — *Heredity* 75: 530–540.

Coltman, D. W., Bowen, W. & Wright, J. M. 1998: Birth weight and neonatal survival of harbour seal pups are positively correlated with genetic variation measured by microsatellites. — *Proc. R. Soc. Lond. B* 265: 803–809.

Coulson, T. N., Pemberton, J. M., Albon, S. D., Beaumont, M., Marshall, T. C., Slate, J., Guiness, F. E. & Clutton-Brock, T. H. 1998: Microsatellites reveal heterosis in red deer. — *Proc. R. Soc. Lond. B* 265: 489–495.

Coulson, T. N., Albon, S. D., Slate, J. & Pemberton, J. M. 1999: Microsatellite loci reveal sex-dependent responses to inbreeding and outbreeding in red deer calves. — *Evolution* 53: 1951–1960.

Falconer, D. S. & Mackay, T. 1996: *Introduction to quantitative genetics*. — Longman, London.

Gelter, H. P., Tegelström, H. & Ståhl, G. 1989: Allozyme similarity between pied and collared flycatchers (Aves: *Ficedula hypoleuca* and *F. albicollis*). — *Hereditas* 111: 65–72.

Gelter, H. P., Tegelström, H. & Gustafsson, L. 1992: Evidence from hatching success and DNA fingerprinting for the fertility of hybrid pied × collared flycatchers *Ficedula hypoleuca* × *albicollis*. — *Ibis* 134: 62–68.

Goldstein, D. B., Linares, A. R., Cavalli-Sforza, R. R. & Feldman, M. W. 1995: An evaluation of genetic distances for use with microsatellite loci. — *Genetics* 139: 463–471.

Gustafsson, L. 1986: Lifetime reproductive success and heritability: empirical support for Fisher's fundamental theorem. — *Am. Nat.* 128: 761–764.

Gustafsson, L. & Merilä, J. 1994: Foster parent experiment reveals no genotype-environment correlation in the external morphology of *Ficedula albicollis*, the collared flycatcher. — *Heredity* 73: 124–129.

Gustafsson, L. & Pärt, T. 1990: Acceleration of senescence in the collared flycatcher *Ficedula albicollis* by reproductive costs. — *Nature* 347: 279–281.

Gustafsson, L. & Sutherland, W. J. 1988: The costs of reproduction in the collared flycatcher *Ficedula albicollis*. — *Nature* 335: 813–815.

Gustafsson, L., Qvarnström, A. & Sheldon, B. C. 1995: Trade-offs between life history traits and a secondary sexual character in male collared flycatchers. — *Nature* 375: 311–313.

Hansson, B., Bensch, S., Hasselquist, D. & Akesson, M. 2001: Microsatellite diversity predicts recruitment of sibling great reed warblers. — *Proc. R. Soc. Lond. B* 268: 1287–1291.

Hauser, T. P. & Loeschcke, V. 1996: Drought stress and inbreeding depression in *Lychnis flos-cuculi* (Caryophyllaceae). — *Evolution* 50: 1119–1126.

Hedrick, P., Fredrickson, R. & Ellegren, H. 2001: Evaluation of d^2, a microsatellite measure of inbreeding and outbreeding in wolves of known pedigree. — *Evolution* 55: 1256–1260.

Höglund, J., Piertney, S. B., Alatalo, R. V., Lindell, J., Lundberg, A. & Rintamäki P. 2002: Inbreeding depression and fitness in black grouse. — *Proc. R. Soc. Lond. B* 269: 711–715.

Houle, D. 1989. Allozyme-associated heterosis in *Drosophila melanogaster*. — *Genetics* 123: 789–801.

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

Keller, L. & Ross, K. G. 1999: Major gene effects on phenotype and fitness: the relative roles of *Pgm-3* and *Gp-9* in introduced populations of the fire ant *Solenopsis invicta*. — *J. Evol. Biol.* 12: 672–680.

Keller, L. F., Arcese, P., Smith, J. M. N., Hochachka, W. M. & Stearns, S. C. 1994: Selection against inbred song sparrows during a natural population bottleneck. — *Nature* 372: 356–357.

Kruuk, L. E. B., Clutton-Brock, T. H., Slate, J., Pemberton, J. M., Brotherstone, S. & Guinness, F. E. 2000: Heritability of fitness in a wild mammal population. — *Proc. Natl. Acad. Sci. USA* 97: 698–703.

Kruuk, L. E. B., Sheldon, B. C. & Merilä, J. 2002: Severe inbreeding depression in collared flycatchers (*Ficedula albicollis*). — *Proc. R. Soc. Lond. B* 269: 1581–1589.

Lacy, R. C., Alaks, G. & Walsh, A. 1996: Hierarchical analysis of inbreeding depression in *Peromyscus polionotus*. — *Evolution* 50: 187–220.

Lynch, M. & Walsh. B. 1998: *Genetics and analysis of quantitative traits.* — Sinauer, Sunderland, MA.

Meagher, S., Penn, D. J. & Potts, W. K. 2000: Male-male competition magnifies inbreeding depression in wild house mice. — *Proc. Natl. Acad. Sci. USA* 97: 3324-3329.

Merilä, J. & Gustafsson, L. 1993: Inheritance of size and shape in a natural population of collared flycatchers, *Ficedula albicollis*. — *J. Evol. Biol.* 6: 375–398.

Merilä, J. & Sheldon, B. C. 1999: Genetic architecture of fitness and non-fitness traits — empirical patterns and development of ideas. — *Heredity* 83: 103–109.

Merilä, J. & Sheldon, B. C. 2000: Lifetime reproductive success and heritability in nature. *Am. Nat.* 155: 301–310.

Merilä, J., Sheldon, B. C. & Ellegren, H. 1997: Antagonistic natural selection revealed by molecular sex identification of nestling collared flycatchers. — *Mol. Ecol.* 6: 1167–1175.

Merilä, J., Sheldon, B. C. & Kruuk, L. E. B. 2001: Explaining stasis: microevolutionary studies in natural populations. — *Genetica* 112: 199–222.

Mitton, J. B. 1994: Molecular approaches to population biology. — *Annu. Rev. Ecol. Syst.* 25: 45–69.

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

Newton, I. 1989: *Lifetime reproduction in birds*. — Academic Press, London.

Orr, H. A. 1997: Haldane's rule. — *Annu. Rev. Ecol. Syst.* 28: 195–218.

Pärt, T. & Gustafsson, L. 1989: Breeding dispersal in the collared flycatcher (*Ficedula albicollis*): possible causes and reproductive consequences. — *J. Anim. Ecol.* 58: 307–322.

Pray, L., Schwartz, J. M., Goodnight, C. J. & Stevens, L. 1994: Environmental dependency of inbreeding depression — implications for conservation biology. — *Cons. Biol.* 8: 562–568.

Primmer, C. R., Møller, A. P. & Ellegren, H. 1996: New microsatellites from the pied flycatcher *Ficedula hypoleuca* and the swallow *Hirundo rustica* genomes. — *Hereditas* 124: 281–283.

Primmer, C. R., Raudsepp, T., Chowdhary, B., Møller, A. P. & Ellegren, H. 1997: Low frequency of microsatellites in the avian genome. — *Genome Res.* 7: 471–482.

Rice, W. R. 1989: Analysing tables of statistical tests. — *Evolution* 43: 223–225.

Ross, K. G., Vargo, E. L. & Keller, L. 1996: Simple genetic basis for important social traits in the fire ant *Solenopsis invicta*. — *Evolution* 50: 2387–2399.

Rossiter, S. J., Jones, G., Ransome, R. D. & Barratt, E. M. 2001: Outbreeding increases offspring survival in wild greater horseshoe bats (*Rhinolophus ferrumequinum)*. — *Proc. R. Soc. Lond. B* 268: 1055–1061.

Rowe, G. & Beebee, T. J. C. 2001: Fitness and microsatellite diversity estimates are not correlated in two outbred anuran populations. — *Heredity* 87: 558–565.

Sætre, G.-P., Borge, T., Lindell, J., Moum, T., Primmer, C. R., Sheldon, B. C., Haavie, J., Johnsen, A. & Ellegren, H. 2001: Speciation, introgressive hybridization and non-linear rate of molecular evolution in flycatchers. — *Mol. Ecol.* 10: 737–749.

Sætre, G.-P., Borge, T., Lindroos, K., Haavie, J., Primmer, C. R., Sheldon, B. C. & Syvänen, A.-C. 2003. Sex-linked speciation in *Ficedula* flycatchers. — *Proc. R. Soc. Lond. B.* 270: 53–59.

Savolainen, O. & Hedrick, P. 1995: Heterozygosity and fitness: no association in Scots pine. — *Genetics* 140: 755–766.

Sheldon, B. C. & Ellegren, H. 1996: Offspring sex and paternity in the collared flycatcher. — *Proc. R. Soc. Lond. B* 263: 1017–1021.

Sheldon, B. C. & Ellegren, H. 1998: Paternal effort related to experimentally manipulated paternity of male collared flycatchers. — *Proc. R. Soc. Lond. B* 265: 1737–1742.

Sheldon, B. C. & Ellegren, H. 1999: Sexual selection resulting from extrapair paternity in collared flycatchers. — *Anim. Behav.* 57: 285–298.

Sheldon, B. C., Davidson, P. & Lindgren, G. 1999: Mate replacement in experimentally widowed collared flycatchers: determinants and outcomes. — *Beh. Ecol. Sociobiol.* 46: 141–148.

Sheldon, B. C., Kruuk, L. E. B. & Merilä, J. 2003: Natural selection and inheritanc e of breeding time and clutch size in the collared flycatcher. — *Evolution.* [In press].

Simon, J. K. 1997: *Resampling: the new statistics*, 2nd ed. — Resampling Stats Inc, Arlington, VA.

Slate, J., Kruuk, L. E. B, Marshall, T. C., Pemberton, J. M. & Clutton-Brock, T. H. 2000: Inbreeding depression influences lifetime breeding success in a wild population of red deer (*Cervus elaphus*). — *Proc. R. Soc. Lond. B* 267: 1657–1662.

Slate, J. & Pemberton, J. M. 2002: Comparing molecular measures for detecting inbreeding depression. — *J. Evol. Biol.* 15: 20–31.

Tegelström, H. & Gelter, H. P. 1991: Haldane's rule and sex biased gene flow between two hybridizing flycatcher species (*Ficedula albicollis* and *Ficedula hypoleuca*, Aves: Muscicapidae). — *Evolution* 44: 2012–2021.

Veen, T., Borge, T., Griffith, S. C., Sætre, G.-P., Bures, S., Gustafsson, L. & Sheldon, B. C. 2001: Hybridization and adaptive mate choice in flycatchers. — *Nature* 411: 45–50.

Wolf, J. B., Brodie, E. D. & Wade, M. J. 2000: *Epistasis and the evolutionary process*. — Oxford Univ. Press, Oxford.