

---

## Individual fluctuating asymmetry in pied flycatchers (*Ficedula hypoleuca*) persists across moults, but is not heritable and not related to fitness

Leif Christian Stige,\* Tore Slagsvold and Leif Asbjørn Vøllestad

*Department of Biology, University of Oslo, Oslo, Norway*

---

### ABSTRACT

**Questions:** What is the relationship between fluctuating asymmetry (small, random departures from perfect morphological symmetry) and individual fitness? What is the heritability of fluctuating asymmetry? At which developmental stage is plumage fluctuating asymmetry determined?

**Hypotheses:** Fluctuating asymmetry is negatively correlated with fitness. Fluctuating asymmetry is heritable. Individual signed left–right asymmetries persist across moults.

**Organism:** Pied flycatchers.

**Field site:** Mixed forest area supplied with nest boxes. Oslo, Norway.

**Methods:** We monitored breeding or breeding attempts of 300 individually marked birds over a 4-year period. We measured fluctuating asymmetry of up to five plumage traits for the adults each year, as well as for 448 nestlings.

**Results:** Neither composite nor single trait fluctuating asymmetry was associated with individual condition, attractiveness, reproductive success or survival. Estimated heritabilities of signed and unsigned fluctuating asymmetry were low and non-significant. This implies that fluctuating asymmetry was a poor predictor of genetic quality or phenotypic condition. Among-individual differences in fluctuating asymmetry therefore had little biological significance. Individual signed asymmetries persisted from the nestling stage to adulthood, and across at least two moults for adults. Plumage fluctuating asymmetry therefore reflected developmental noise acting on the embryological development of the feather follicles, rather than on the actual growth of the feathers.

*Keywords:* developmental stability, fitness, fluctuating asymmetry, heritability, ontogeny.

### INTRODUCTION

Fluctuating asymmetry refers to small, random departures from perfect morphological symmetry, and is typically measured as differences in sizes of structures between the right and left sides of bilaterally symmetric organisms (Van Valen, 1962). Fluctuating asymmetry is

---

\* Address all correspondence to Leif Christian Stige, Biogéosciences, UMR-CNRS 5561, Université de Bourgogne, 6 Boulevard Gabriel, 21000 Dijon, France. e-mail: leif.stige@u-bourgogne.fr  
Consult the copyright statement on the inside front cover for non-commercial copying policies.

used as a measure of developmental instability, which reflects the ability of an organism to consistently produce a targeted phenotype in a given environment (Zakharov, 1992). Developmental instability can be affected by both genetic and environmental factors, and organisms with superior genotypes that have experienced low environmental stress are thought to have high developmental precision, and thus low fluctuating asymmetry (Møller and Swaddle, 1997). Fluctuating asymmetry has therefore been used as a proxy measure of condition or fitness (e.g. González-Guzmán and Mehlman, 2001; Carbonell *et al.*, 2003), and is potentially important in an evolutionary context (Møller and Swaddle, 1997), for example in connection with sexual selection (reviewed by Møller and Cuervo, 2003; Tomkins and Simmons, 2003).

The relationship of fluctuating asymmetry with fitness (Clarke, 1995, 1998; Palmer, 2000) as well as with other factors [e.g. stress (Bjorksten *et al.*, 2000), sexual selection (Simmons *et al.*, 1999; Palmer, 2000)] is disputed. Empirical results are heterogeneous and equivocal: fluctuating asymmetry–fitness relations are generally weak, and appear to depend on the study system as well as the type of trait (Leung and Forbes, 1996). Using meta-analysis, Leung and Forbes (1996) found that fluctuating asymmetry accounted on average for 6.7% of the variation in fitness components; Møller (1999), also using meta-analysis, estimated that fluctuating asymmetry accounted for 2.1%, 6.0% and 12.3%, respectively, of the variation in growth, fecundity and survival. There are, however, reasons to consider these estimates with some caution. First, they may be inflated by publication bias; Palmer (2000) found that among the studies included in the first meta-analysis, large sample sizes were associated with low effect sizes ( $r^2 = 0.16$ ). Møller (1999) reported a similar correlation for fluctuating asymmetry–survival studies ( $r^2 = 0.18$ ), but not for fluctuating asymmetry–fecundity studies. While this provides indirect evidence of publication bias, alternative explanations exist (see Møller, 1999). There is also indirect evidence of publication bias in studies of the role of fluctuating asymmetry in sexual selection (Simmons *et al.*, 1999; Palmer, 2000; Tomkins and Simmons, 2003), but a direct comparison between 68 published and 11 unpublished studies disclosed no difference in average effect size (Møller and Cuervo, 2003). Second, methodological robustness is crucial in studies of fluctuating asymmetry (Palmer and Strobeck, 1986, 1992, 2003; Merilä and Björklund, 1995) and the estimation of fluctuating asymmetry in some of the early studies may have been confounded by measurement error, the effects of other types of symmetry (directional asymmetry or anti-symmetry) or scaling effects (Simmons *et al.*, 1999; Palmer, 2000). Third, as pointed out by Clarke (1998), comparisons of fluctuating asymmetry with fitness are often made at the population level, and it is not possible to determine whether the relationship holds at the individual level, which is obviously important (e.g. in connection with sexual selection). More methodologically sound studies are required to determine whether low fluctuating asymmetry is generally associated with high individual fitness.

The interpretation of fluctuating asymmetry as an indicator of genetic quality is also challenged by the fact that its heritability is on average estimated to be less than 10% (reviewed by Fuller and Houle, 2003). Although the heritability of the underlying developmental instability is potentially much higher (Gangestad and Thornhill, 1999, 2003), low heritability of fluctuating asymmetry suggests that among-individual differences in fluctuating asymmetry predominantly reflect random or environmental sources of variation. On the other hand, low heritability of fluctuating asymmetry is an advantage if stress effects are to be inferred from such asymmetry, and heritability estimates have little relevance for comparisons at the population level.

One of the main reasons for the controversy surrounding fluctuating asymmetry is our limited knowledge of the processes that generate and cause among-individual variation in

such asymmetry (Klingenberg, 2003). One method to gain some insight into the basis of fluctuating asymmetry is to investigate its ontogeny (see Kellner and Alford [2003] and Møller and van Dongen [2003], and references therein). In particular, time-series of left–right values of single individuals provide information on the time scale on which asymmetries originate and persist. Persistence of individual asymmetries across time suggests that fluctuating asymmetry is fixed early in development (as observed in brachyuran crabs, *Hemigrapsus nudus*; Chippindale and Palmer, 1993), while fluctuations of asymmetries suggest that fluctuating asymmetry mainly reflects the recent growth history (as found in domestic fowl, *Gallus gallus domesticus*; Kellner and Alford, 2003). The first case would suggest that fluctuating asymmetry is determined by genotype and/or environmental conditions early in life, whereas the second is consistent with fluctuating asymmetry as an indicator of present-state condition.

The pied flycatcher, *Ficedula hypoleuca* Pallas (Passeriformes, Muscicapidae), is a frequently used model species (Lundberg and Alatalo, 1992), well suited for studies of fluctuating asymmetry. Pied flycatchers are long-distance migrants, breeding in northern Europe in spring and early summer and wintering in tropical Africa. Males are facultatively polygynous and new pair bonds are formed each year (Lundberg and Alatalo, 1992). This leads to large variation in male reproductive success and to the possibility of repeated assessments of male attractiveness. Males also show great variation in plumage colour, a trait that is important in mate choice (Sætre *et al.*, 1994). Breeding occurs naturally in holes in trees, but nest boxes are preferred. Supplying nest boxes thus makes it possible to monitor breeding success of nearly all individuals in a population. The relatively high breeding site fidelity of adults (but lower for females than males; Lundberg and Alatalo, 1992; Dale *et al.*, 2002) enhances survival estimation and more precise estimation of individual lifetime reproductive success. It also provides the opportunity for measuring individual asymmetries across subsequent moults to investigate the ontogeny of fluctuating asymmetry (FA). Finally, pied flycatchers are robust to handling, which allows detailed morphometric measurements at various stages of breeding, and close monitoring of each breeding attempt. It has been shown that heavy metal pollution stress causes increased tarsus length FA, but not tarsus width FA or primary length FA in pied flycatcher nestlings (Eeva *et al.*, 2000). However, to our knowledge, no studies have tested for general associations between fluctuating asymmetry and fitness or investigated other aspects of developmental instability in pied flycatchers.

In this study, we investigate the relationship between fluctuating asymmetry and individual fitness in a population of pied flycatchers. The fitness components estimated include condition, male attractiveness, reproductive success and survival. A negative relation between fluctuating asymmetry and fitness is predicted. The heritability of fluctuating asymmetry is estimated by offspring–parent regression and by full sibling analysis. Repeated measurements of individuals recaptured as nestlings and as adults, and as adults during subsequent years, provide information on the ontogeny of fluctuating asymmetry. These data are used to determine whether asymmetries in plumage characters of single birds are persistent across subsequent moults.

## METHODS

### Study area and materials

The study was carried out during four subsequent breeding seasons (2000–2003) in a mixed deciduous–coniferous forest area in Sørkedalen near Oslo, south-eastern Norway (59°59'N;

10°38'E). The pied flycatchers were captured using traps at the entrances of nest boxes, or for a few males by mist netting. Males were captured and measured soon after arrival and were given a numbered metal ring and unique combinations of colour rings to permit identification of individuals in the field. In addition, males were given a red pen mark on the throat, chest or under-tail to ease identification from a distance. The mark disappeared at the latest during the autumn moult, but was usually visible for no longer than 2–4 weeks. Females were generally captured, ring-marked and measured during the incubation period. All pied flycatchers observed in the study area were captured, except less than ten apparently transmittal males (generally observed only once) and three females that disappeared around the egg-laying period.

The birds bred in nest boxes placed approximately 1.5 m above the ground, which permitted monitoring of egg laying, hatching and fledging. Arrival date was estimated by the date of first observation. The pairing date of males was estimated by the date of the first visit of a female into his nest box. Females pair soon after arrival (Lundberg and Alatalo, 1992), and we did not attempt to differentiate pairing date and arrival date for females. Onset of egg laying, number of eggs, number of hatchlings and number of fledglings were recorded for breeding pairs. Nestlings were weighed and measured for asymmetry 13 days after hatching, at which time the increase in body mass has reached an asymptote (Lundberg and Alatalo, 1992).

In 2000–2002 all observed individuals were monitored, while in 2003 we only recorded which of the previously marked individuals returned. In 2000 the numbers of hatchlings and fledglings could not be assessed in 12 nests due to nest manipulation for a different project. In total, 11 nests (2000,  $n = 8$ ; 2001,  $n = 2$ ; 2002,  $n = 1$ ) were depredated after hatching, probably by mustelids or cats. This was assumed to have occurred randomly, with little or no relation to any characteristics of the birds, so numbers of fledglings ( $= 0$ ) for these nests were not included in the analyses.

### Morphological measurements

The plumage colour of males was scored on a scale from black-and-white ( $= 1$ ) to brownish and female-like ( $= 7$ ) (Drost, 1936). Half-scores were used for intermediate types. The highest part of the forehead patch of males was measured to the nearest 0.1 mm. Body mass was measured to the nearest 0.1 g with a Pesola spring balance. Since body mass increases predictably during a day (Dale *et al.*, 2002), observed body mass was standardized to 09.00 h (males:  $0.043 \text{ g} \cdot \text{h}^{-1}$ ; cf. Dale *et al.*, 2002; females: regression slope  $0.079 \pm 0.022 \text{ g} \cdot \text{h}^{-1}$ , current data, co-factor year). For females, only body mass recordings from the incubation period were used. For males, the first body mass recording after arrival was used. The body mass of a few males that were not captured until late in the season was adjusted upwards (maximally 0.3 g), based on regression of weight loss during the nestling feeding period for males captured two or more times the same year. Tarsus length was measured with bent toes to the nearest 0.1 mm with a calliper. The length of the first primary and wing length were measured to the nearest 0.5 mm using a flattened and straightened wing. Age (one-year or older) was assessed from morphological characteristics (Karlsson *et al.*, 1986). For birds captured two or more times in the same year, average values of the measurements of morphological traits (other than body mass, see above) were used.

Fluctuating asymmetry measurements were made with a calliper to the nearest 0.1 mm. To increase precision, all traits were measured twice, but only the average value of each side was recorded and used in the later analyses. In 2000 eight FA traits were measured. Four of

these traits (measurements on the inside of tail feathers [cf. traits 3 and 4; Table 1] and width and height of male forehead patch) were excluded, since fluctuating asymmetry was not significant relative to measurement error. A ninth trait was introduced in 2001. Consequently, five FA traits were included in the final analyses (Table 1). For nestlings, only FA traits 1 and 2 could be measured, as the tail feathers were not developed enough. In 2000 FA measurements were made for the two largest nestlings in each nest (i.e. the ones most likely to return a later year), and in 2001 all nestlings were measured. In 2002 nestling FA was not measured. Moulting stage was noted when measurements were made late in the season, which confirmed that the autumn moult had not yet started for any of the measured feathers. Measurements of visibly worn or damaged feathers were not used. Additionally, on several tail feathers, the white patch extended to the feather tip, or there was no patch, or the patch edge was diffuse and poorly defined; in all these cases, measurements could not be done. These criteria excluded more than half of the tail measurements (cf. Table 2).

Independent replicate measurements of adults were obtained by recapturing and measuring most males in 2000 and 2001 (and eight in 2002) a second time later in the season. Replicate measurements of nestlings were obtained in 2001 by measuring one randomly chosen nestling from each nest twice. Independence between measurements was assured by performing the second measurement of a nestling after all nestlings in the nest had been measured once without knowing its identity (the ring number was read after measuring).

Fluctuating asymmetry was calculated as the absolute value of the difference between the left (L) and right (R) sides ( $FA = |L - R|$ ) for traits 1–3 for adult birds. For traits 4 and 5, and for traits 1 and 2 for nestlings, fluctuating asymmetry was calculated from log-transformed values ( $FA = |\ln(L + 1) - \ln(R + 1)|$ ), since fluctuating asymmetry calculated from untransformed values correlated positively with trait size. Following this transformation, the fluctuating asymmetry of no traits correlated significantly with trait size (Pearson's correlation coefficient between fluctuating asymmetry and trait size:  $r = 0.01, 0.07, 0.12, 0.08, 0.15$  for traits 1–5, respectively; d.f.  $\geq 110, P > 0.05$ ).

### Observer bias

Altogether, six observers were involved in morphological measuring (2000: T.S., H.P.R., B.R.B., P.A.; 2001: L.C.S., T.S.T., T.S.: six birds; 2002: L.C.S.). After a preliminary analysis of observer error and observer bias, all morphological measurements on adult birds by two

**Table 1.** FA traits: each character is measured on the right (R) and left (L) side of the bird, and fluctuating asymmetry is calculated as the difference between sides

- 
1. *Primaries 1, 2.* Difference in length between the first and second primaries (i.e. flight feathers). The bird was held in the left hand with its belly up and the wings resting as undisturbed as possible.
  2. *Primaries 2, 3.* Difference in length between the second and third primaries. The bird was held as for the previous measure.
  3. *Tail patch 1.* The distance from the tip of the outermost tail feather to the distal end of a white patch found on the outside of the shaft on the same feather. The tail was seen from the side of the bird's back.
  4. *Tail patch 2.* Similar to the previous measure, but on the second outermost tail feather.
  5. *Tail 1, 2.* Difference in length between the outermost and second outermost tail feathers. The tail was seen from the side of the bird's belly. Only 2001 and 2002.
-

observers (B.R.B. and P.A.), on nestlings by one observer (P.A.) and of FA trait 5 by one observer (T.S.T.) were excluded from further analysis. Since nearly all of these birds had also been measured by one of the other observers, this had a negligible influence on the final sample size.

Systematic differences between observers were corrected for if the differences were significant ( $P < 0.05$ ) in paired *t*-tests of birds measured by two observers in one year. In 2000 measurements by H.P.R. were adjusted as follows: forehead patch  $-0.44$  mm, wing length  $-0.86$  mm, length of first primary  $-0.27$  mm, tarsus length  $+0.86$  mm; and measurements by T.S. of signed asymmetry ( $L - R$ ) of FA trait 1 were adjusted by  $+0.22$  mm (based on 27–34 paired measurements). In 2001 wing length measurements by T.S.T. were adjusted by  $-0.66$  mm and measurements of unsigned asymmetry ( $|L - R|$ ) of FA trait 1 by T.S.T. were multiplied by 0.72 (based on 26–29 paired measurements on the same birds by T.S.T. and L.C.S.). Measurements in 2001 and 2002 were adjusted to the 2000 level if there was a significant year effect in an analysis of variance (ANOVA) on measurements on all birds, with sex, age and year as explanatory variables (measurements on the six birds by T.S. in 2001 excluded). Consequently, wing lengths were adjusted by  $+0.63$  mm and  $+0.27$  mm, the lengths of first primaries by  $+0.91$  mm and  $+0.26$  mm, and tarsus lengths by  $+0.09$  mm and  $-0.23$  mm for 2001 and 2002, respectively. Mean values of signed or unsigned FA did not differ between years ( $P > 0.05$ , ANOVA, residual d.f. = 110–375).

### Measurement error and asymmetry distribution

The measurement error for fluctuating asymmetry, estimated as the absolute value of the difference between replicate estimates of signed FA by the same or different observers, was not correlated with trait size ( $r^2 < 0.02$ ,  $P > 0.05$ ,  $n = 20$ –120) and did not differ significantly between years ( $P > 0.05$ , ANOVA on the same data).

To estimate the magnitude of fluctuating asymmetry in relation to measurement error and directional asymmetry, a mixed model approach was used (cf. van Dongen *et al.*, 1999). All raw FA trait measurements of each trait were analysed by restricted maximum-likelihood (REML) mixed effects models (Pinheiro and Bates, 2000). The measured trait size is considered a function of side (fixed) + individual (random) + side nested in individual (random) + measurement error (random). Fluctuating asymmetry is represented by the random side effect. The standard deviation of this variance component is analogous to FA10a of Palmer and Strobeck (2003), and multiplying it by  $0.798 \cdot \sqrt{2} = 1.13$  gives the expected mean  $|L - R|$  value, corrected for effects of directional asymmetry and measurement error. The significance of skewness ( $g_1$ ) was calculated according to Sokal and Rohlf (1995, p. 138), and that of kurtosis according to Palmer and Strobeck (2003, table 17.5, 'equation 6', given at a 1% and a 5% level). The analyses showed that although measurement error was high, variation due to fluctuating asymmetry was significant for all traits (Table 2, one-tailed likelihood ratio tests; the tests were one-tailed since variances cannot be negative: van Dongen *et al.*, 1999). Directional asymmetry (mean  $L - R$  different from zero) was significant for trait 2 (*F*-test), but of low magnitude relative to fluctuating asymmetry (Table 2). To correct for directional asymmetry, FA for trait 2 was calculated as the absolute value of the deviation from the population mean  $L - R$  value. There was significant skewness for trait 1 and leptokurtosis for all traits (Table 2). Visual inspection of the data suggested that the departures from normality were caused by long tails of the  $L - R$  distributions. To determine whether the final results were influenced by the inclusion of traits exhibiting directional asymmetry or

**Table 2.** Distributional characteristics of FA traits

Trait <sup>a</sup>	<i>n</i> <sub>2000</sub>	<i>n</i> <sub>2001</sub>	<i>n</i> <sub>2002</sub>	<i>n</i> <sub>REP</sub>	DA	FA	ME	Skewness	Kurtosis
<b>Adults</b>									
1	125	134	140	106	0.005	0.313***	0.271	0.35**	1.50**
2	123	136	142	91	0.087***	0.220***	0.298	-0.10	2.50**
3	36	54	25	37	0.085	0.662**	0.908	0.29	4.40**
4	49	75	50	42	0.000	0.052***	0.044	0.06	2.60**
5	0	67	69	20	0.025	0.133***	0.089	0.32	5.30**
<b>Nestlings</b>									
1	78	370	0	66	0.001	0.016***	0.009	-0.14	0.71**
2	78	370	0	66	0.028***	0.040***	0.046	-0.32**	2.80**

*Note:* Directional asymmetry (DA), fluctuating asymmetry (FA) and measurement error (ME) were estimated from mixed-effects REML models. DA is the expected  $L - R$  value (mm) for traits 1–3 in adults; for other traits it is the expected value of  $\ln(L + 1) - \ln(R + 1)$ . FA and ME are estimated standard deviations. *n* denotes the number of individuals measured each year and the total number of individuals measured twice in any one year. <sup>a</sup> See Table 1. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

skewness, analyses were also performed with either trait 1 or 2 excluded. The results were qualitatively the same, and analyses based on all traits are presented.

### Final FA measures

To avoid undue leverage from the largest FA-values in the further analyses, they were rank-transformed to a scale from 0 to 1. The transformation consisted in replacing each FA-value with the quantile of its rank among all individuals' FA-values of the given trait (pooled across years and sexes) ( $FA1_{YR} - FA5_{YR}$ ). The same procedure was applied to signed asymmetry values ( $sFA1_{YR} - sFA5_{YR}$ , scale = -0.5, +0.5). The subscript 'YR' signifies that a FA index is year specific. In addition, mean fluctuating asymmetry across years was calculated for each trait ( $FA1_{ALL} - FA5_{ALL}$ ). Composite fluctuating asymmetry was calculated as mean rank-transformed FA across traits for each year ( $cFA_{YR}$ ) and for all years ( $cFA_{ALL}$ ). The composite index is similar to the CFA3 of Leung *et al.* (2000), except that ranks are scaled and averaged and not summed. Analyses were also performed with  $cFA$  calculated from untransformed FA-values, similar to the CFA2 of Leung *et al.* (2000), but averaging instead of summing FA-values. The results (not shown) were qualitatively the same as those presented.

### Year-specific and overall morphological indices (subscripts <sub>YR</sub> and <sub>ALL</sub> respectively)

Both for fluctuating asymmetry and for the other morphological variables, year-specific and overall indices were calculated. As overall indices of SIZE, COND, PLUM and FOREHEAD (see Table 3), the individual random effect estimates from generalized linear mixed models (GLMMs) accounting for year and age (see below) were used. The overall indices are most likely to reflect permanent characteristics of individuals (e.g. genetic effects), while the year-specific indices also reflect within-individual variation (e.g. condition-dependent effects).

**Table 3.** Explanation of fitness parameters and other variables used in fitness models**Fitness parameters**

1. **SIZE:** Score on the first axis in principal components analysis (PCA) on body weight, tarsus length and wing length; sex-specific scores.
2. **COND:** Condition. Residuals of regression of body weight on tarsus length; sex-specific scores.
3. **PLUM:** Male plumage colour (Drost, 1936; from 1 = brightest to 7 = female-like). Females prefer brightly coloured males (Sætre *et al.*, 1994).
4. **FOREHEAD:** Height of white patch on forehead of males (mm). Females in some populations, though not the Norwegian one, prefer males with large forehead patches (see Dale *et al.*, 1999, and references therein).
5. **ARRIVAL:** Date of first observation (females: arrival = pairing date). Individuals that arrive earlier can choose better nest-sites, find higher-quality mates and can start egg laying earlier (Lundberg and Alatalo, 1992). May 1st = 1.
6. **P.PAIRED:** Whether or not a male was paired (0/1). In the study population, 7–10% of the males each year remained unpaired throughout the breeding season.
7. **P.POLYG:** Whether or not a male was polygynous, if paired (0/1). In the study population, 15–17% of all males each year were bigamous and 0–1.5% were trigamous.
8. **PAIRING:** Date of pairing (covariate: arrival date; males only). The most attractive males are likely to attract a mate first. May 1st = 1.
9. **LAYING:** Date of onset of egg laying (covariate: pairing date). Females differ in their capacity to lay early, which affects fitness since earlier broods have higher breeding success, probably due to declining food supplies (Lundberg and Alatalo, 1992). May 1st = 1.
10. **CLUTCH-SIZE:** Number of eggs laid (polygynous males: primary nest only).<sup>1</sup>
11. **HATCH.SUCC:** Hatching success (= number of hatchlings/number of eggs, arcsine transformed and scaled to maximum = 1; bigamous males: primary nest only).<sup>2</sup>
12. **FLEDG.SUCC:** Fledging success (= number of fledglings/number of hatchlings, arcsine transformed and scaled to maximum = 1; polygynous males: primary nest only).
13. **HATCH.N:** Total number of hatchlings (polygynous males: all nests).
14. **FLEDG.N:** Total number of fledglings (polygynous males: all nests).
15. **FLEDG.WT:** Average body mass of fledglings (at day 13 after hatching) (g).
16. **RECRUITS:** Number of offspring recaptured in a later year (0–3).

**Other variables used as predictor variables in fitness models**

1. **AGE:** 1-year-old (0) or older than 1 year (1).
2. **EXP:** previous breeding experience (0/1). Whether or not captured in the study area in a previous year.
3. **STATUS:** Female breeding status. 0: Paired with monogamous male, or primary female of bigamous male. 1: Secondary or tertiary female of polygynous male. All observed females were paired.
4.  $Y_{2001}$ ,  $Y_{2002}$ : Year effects. Year 2000 is used as reference level.

<sup>1</sup> Two nests where the female disappeared in the laying period excluded. <sup>2</sup> Two nests where the female disappeared in the incubation period excluded.

**Heritability of fluctuating asymmetry**

Estimates of the heritability of fluctuating asymmetry were obtained by mean offspring–mean parent regression corrected for unequal family size and by sibling analysis (Roff, 1997). The offspring–parent regression slope estimates narrow sense heritability ( $h_N^2$ ; additive genetic variance/total phenotypic variance) directly. In the sibling analysis, the variance

among families ( $\sigma_{AF}^2$ ) reflects additive (and partly also non-additive) genetic variance, and heritability is estimated as  $2\sigma_{AF}^2/(\sigma_{AF}^2 + \sigma_{AP}^2)$ , where  $\sigma_{AP}^2$  is variance among progeny. Both types of estimates are inflated by environmental effects shared among siblings (e.g. maternal and paternal effects) and can be viewed as maximum estimates of  $h_N^2$ . Analyses were based on the rank-transformed FA-values (FA1, FA2 and cFA of nestlings, together with FA1<sub>ALL</sub>, FA2<sub>ALL</sub> and cFA<sub>ALL</sub> of adults). Heritability was also estimated from untransformed FA-values, which gave similar estimates (not shown). Variance components in sibling analysis were estimated by REML estimation. Confidence intervals of the parameters from the parent–offspring regression and the sibling analyses were obtained by non-parametric and parametric bootstrapping, respectively (Efron and Tibshirani, 1993).

### Fitness effects

A set of fitness parameters was constructed based on the morphological measurements and the observational data (Table 3). Fluctuating asymmetry–fitness relations were analysed for each sex and each fitness parameter separately. For all parameters except adult survival, a GLMM approach was used. Data from all years were analysed in one model with the individual effect included as a random intercept. The random individual effect thus represents the variation in a given fitness component that is consistent across years for single individuals, and the error term represents the variation among years for each individual. Before testing for FA effects, co-factors were selected for each fitness trait. The selection of co-factors was done by a stepwise selection procedure, starting with the most complex model, and omitting terms with the highest *P*-value one at a time until only terms with *P*-values less than 0.05 remained. Model selection for models with logit or log link functions was done in the opposite direction, starting with the simplest models, since too complex models did not converge with the estimation method used for these. Only interaction terms that were considered biologically meaningful were included in the models. After the selection of co-factors, it was tested for unique effects of fluctuating asymmetry, where all co-factors were included, and for marginal effects, where female status and morphological factors other than fluctuating asymmetry were omitted from the models. Models with identity link functions were fitted by REML estimation (Pinheiro and Bates, 2000). Models with log or logit link functions were fitted by maximum likelihood estimation (Broström, 2003). The effects of fluctuating asymmetry and other fixed effects were tested by *F*-tests (identity link models) or  $\chi^2$ -tests (log or logit link models). The significance of variance terms was tested by one-tailed likelihood-ratio tests (cf. van Dongen *et al.*, 1999). For models with log or logit link, the random individual effect could not be tested due to methodological limitations. In addition, the effects of cFA were tested by non-parametric bootstrapping (resampling of individuals). Standard errors and *P*-values were similar with bootstrap and with parametric methods, with a few exceptions that are noted in the Results.

Adult return rate was analysed by maximum likelihood capture–recapture models using the program MARK (White and Burnham, 1999). Using this method, two sets of parameters are simultaneously estimated: apparent survival probabilities and recapture probabilities. Apparent survival is the probability that an individual remains in the population from one capture occasion (year) to the next. Individuals leave the population either due to mortality or permanent emigration, which are necessarily confounded factors. The recapture parameter estimates the probability of being captured if present in the population or being temporarily absent from it. A logit link function was used. Selection of model structure for

co-factors was done so as to minimize Akaike's information criterion corrected for low sample size ( $AIC_c$ ) (Anderson and Burnham, 1999).

### Power simulation

The power of the tests of fluctuating asymmetry–fitness associations was estimated by computer simulations. Fluctuating asymmetry and fitness data that resembled the real data with respect to sample size and distributions of fluctuating asymmetry and fitness variables were randomly generated and analysed as the real data set. Power was estimated as the proportion of test results significant at a 1% level after 500 iterations. Individual-specific developmental instability values ( $DI_{IND}$ ) that affected both fitness and fluctuating asymmetry were generated from gamma distributions (with expectations = 1). To assess the influence of among-individual variation in developmental instability, simulation regimes with different coefficients of variation of DI ( $CV_{DI}$ ) were simulated ( $CV_{DI} = 23, 45$  or  $100$ ). The generation of FA-values in principle followed the model of Houle (2000). For each individual and trait, a signed FA-value was randomly generated from a normal distribution with expectation zero and a standard deviation given by a trait-specific constant multiplied by  $DI_{IND}$ . Individual FA-values were assumed to be constant across years (in accordance with the results reported below) and were measured with measurement error. Fluctuating asymmetry and measurement error were scaled according to the estimates for the real data (Table 2). Values of the different fitness variables were randomly generated from models described by Table 6. The effects of developmental instability were included by adding or subtracting (depending on the fitness variable) the term:  $k (SD_{DI})^{-1} (1 - DI_{IND}) SD_{FIT}$ , where  $k = 1$  or  $2$ ,  $SD_{DI} = 0.23, 0.45$  or  $1.0$  (depending on  $CV_{DI}$ ), and  $SD_{FIT} = \sqrt{(\sigma_{IND}^2 + \sigma_{RES}^2)}$  from Table 6. Thus the effect of  $DI_{IND}$  on fitness was modelled to be of equal magnitude to the estimated standard deviation of fitness ( $k = 1$ ) or two times this magnitude ( $k = 2$ ). Only fitness variables that could be approached by normal distributions were modelled (i.e. those with identity link; Table 6).

The program R (R Development Core Team, 2003) was used for bootstrapping, power simulations, the fitting of GLMMs with log or logit link functions and for variance estimation in sibling analysis. S-PLUS (Insightful, 2001) was used for all other statistical analyses. In the bootstrapping of cFA versus fitness, 1000 resamples were drawn; otherwise, 5000 resamples were drawn. Non-parametric bootstrapping (Efron and Tibshirani, 1993) was used if not stated otherwise. A 5% critical level is used in statistical tests if not stated otherwise.

## RESULTS

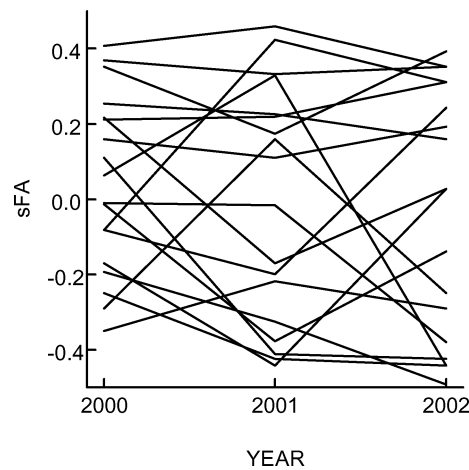
### Ontogeny of fluctuating asymmetry

For individual birds, signed asymmetry (sFA) values were consistent across years, both for nestlings recaptured as adults and between years for adult birds that were caught in two or more subsequent years (Table 4, Fig. 1). Although the confidence intervals are wide, the correlations between measurements 2 years apart appear to be no lower than correlations between measurements 1 year apart. The correlations of unsigned FA-values between years were generally positive, but not significant (Table 4). For trait 1 there were sufficient replicated measurements to quantify the variation in signed asymmetry among individuals compared with the variation among years within individuals (GLMM, identity link,

**Table 4.** Correlation of signed and unsigned asymmetry values (sFA<sub>YR</sub> and FA<sub>YR</sub>) between years for birds caught in two or more subsequent years

Trait	Lapse of 1 year						Lapse of 2 years		
	Nestling–adult			Adult–adult			Adult–adult		
	<i>r</i>	CI	<i>n</i>	<i>r</i>	CI	<i>n</i>	<i>r</i>	CI	<i>n</i>
sFA1 <sub>YR</sub>	0.78	(0.51, 0.91)	16	0.39	(0.18, 0.55)	94	0.71	(0.42, 0.86)	21
sFA2 <sub>YR</sub>	0.31	(−0.24, 0.70)	17	0.51	(0.27, 0.65)	92	0.24	(−0.24, 0.62)	20
sFA3 <sub>YR</sub>				0.29	(−0.08, 0.60)	30	0.61	–	4
sFA4 <sub>YR</sub>				0.32	(−0.03, 0.61)	34	0.28	–	6
sFA5 <sub>YR</sub>				0.44	(0.01, 0.72)	25	–	–	0
FA1 <sub>YR</sub>	0.37	(−0.13, 0.72)	16	0.21	(−0.01, 0.40)	94	0.13	(−0.24, 0.49)	21
FA2 <sub>YR</sub>	0.22	(−0.33, 0.66)	17	0.12	(−0.08, 0.32)	92	0.01	(−0.37, 0.40)	20
FA3 <sub>YR</sub>				0.05	(−0.32, 0.47)	30	−0.71	–	4
FA4 <sub>YR</sub>				0.22	(−0.19, 0.53)	34	0.64	–	6
FA5 <sub>YR</sub>				0.43	(0.02, 0.71)	25	–	–	0

Note: *r* = Pearson’s correlation coefficient, CI = 95% bias-corrected bootstrap confidence interval, *n* = sample size. For one-year comparisons, data for 2000–2001 and 2001–2002 are pooled. For nestlings, only FA traits 1 and 2 were measured.



**Fig. 1.** Signed asymmetry values (sFA1<sub>YR</sub>) for 18 birds in the years 2000, 2001 and 2002.

response: raw signed FA estimates (mm), random effects: Individual, and Year nested in Individual). The within-individual among-year variation (SD = 0.15, 95% CI = 0.05, 0.42) was estimated to be of smaller magnitude than the among-individual variation (SD = 0.43, 95% CI = 0.37, 0.51) and measurement error (SD = 0.39, 95% CI = 0.35, 0.44). The within-individual among-year variance component was not significant ( $\chi^2_{0.5} = 0.96$ ,  $P = 0.16$ , one-tailed likelihood ratio test). In contrast, the among-individual variance component was highly significant ( $\chi^2_{0.5} = 76.8$ ,  $P < 0.0001$ ).

The average unsigned FA for recaptured individuals did not increase or decrease significantly from one year to the next (change in  $FA_{YR}$  2000–2001 =  $-1\%$ , 95% bootstrap confidence interval, CI =  $-12\%$ ,  $+14\%$ ,  $n = 57$ ; 2001–2002 =  $-2\%$ , CI =  $-15\%$ ,  $+13\%$ ,  $n = 51$ ). For individuals that were caught both in 2000 and 2002, mean  $FA_{YR}$  increased 14% (CI =  $-4\%$ ,  $+36\%$ ,  $n = 22$ ). For the population as a whole there was a significant 13% decrease (CI =  $-21\%$ ,  $-4\%$ ,  $n = 285$ ) in  $FA_{YR}$  during the same period (mean  $FA_{YR} = 0.53$ , 0.50, 0.46 for 2000–2002, respectively), possibly caused by increased precision of measurements (although no significant year effect on measurement error was detected). Alternatively, the population decrease in fluctuating asymmetry could be caused by improved environmental conditions or stronger selection against asymmetric individuals, although the lack of significant associations between fluctuating asymmetry and condition or survival (see below) casts doubt on these explanations. The confidence intervals for recaptured birds and for the whole population are tangential, which appears to suggest that estimated FA levels increase with age. However, as all traits increase in size with age ( $t$ -tests between 1-year-olds and older birds,  $t = 2.0$ – $9.2$ , d.f. = 118–383, all  $P < 0.05$ ), and there were positive (but not significant) correlations between trait size and fluctuating asymmetry for all traits, the increase in fluctuating asymmetry with increasing age is confounded by increased trait size. For single FA traits, there were no significant differences between years.

### Heritability of fluctuating asymmetry

The estimated variances in unsigned FA among families were low (Table 5), and not significantly different from zero (one-tailed likelihood ratio tests,  $\chi^2_{0.5} \leq 0.0001$ ,  $P = 0.50$ ). According to the sibling analysis,  $h^2_N$  was lower than 14% (the upper limit of the 95% confidence interval for cFA). The offspring–parent regression gave higher estimates of  $h^2_N$  for FA2 and cFA, but the confidence intervals are wide and include zero for all traits. There was no evidence of heritability of signed asymmetries. The estimated  $h^2_N$  of sFA1 was 0.02 (95% bootstrap CI =  $-0.13$ , 0.15) based on parent–offspring regression and 0.0003 (95% parametric bootstrap CI =  $10^{-5}$ , 0.14) based on sibling analysis. The estimated  $h^2_N$  of sFA2

**Table 5.** Heritability ( $h^2$ ) of unsigned FA

Trait	Offspring–parent regression	Full sibling analysis			$n_{FAMILIES}$	$n_{PROGENY}$
	$h^2$ (CI)	$\sigma_{AF}$ (CI)	$\sigma_{AP}$ (CI)	$h^2$ (CI)		
FA1	0.01 ( $-0.15$ , 0.16)	0.0012 (0.0011, 0.08)	0.29 (0.27, 0.31)	$4 \times 10^{-5}$ ( $3 \times 10^{-5}$ , 0.14)	113	448
FA2	0.12 ( $-0.02$ , 0.26)	0.0014 (0.0010, 0.08)	0.29 (0.27, 0.31)	$5 \times 10^{-5}$ ( $3 \times 10^{-5}$ , 0.15)	113	447
cFA	0.05 ( $-0.11$ , 0.20)	0.0016 ( $8 \times 10^{-4}$ , 0.05)	0.21 (0.19, 0.22)	$1.3 \times 10^{-4}$ ( $3 \times 10^{-5}$ , 0.14)	113	448

*Note:*  $h^2$  is estimated from mean offspring–mean parent regression and from full sibling analysis (REML estimation of variance components).  $\sigma_{AF}$  = estimated SD among families,  $\sigma_{AP}$  = estimated SD among progeny, CI = 95% bootstrap confidence intervals (non-parametric bootstrap for offspring–parent regression, parametric for sibling analysis).

was  $-0.05$  (CI =  $-0.17, 0.08$ ) based on parent–offspring regression and  $0.08$  (CI =  $10^{-4}, 0.24$ ) based on sibling analysis. Heritability was also estimated based on FA measurements of offspring returning as adults, but the sample sizes were too small to make any conclusions (FA1:  $h_N^2 = 0.10$ , CI =  $-0.60, 0.97$ ,  $n = 13$ ; FA2:  $h_N^2 = -0.51$ , CI =  $-1.09, 0.11$ ,  $n = 16$ ).

The hypothetical repeatability,  $R$ , defines ‘the proportion of the variance measured for a trait which is due to real differences among individuals in that trait’ (Whitlock, 1998).  $R$  was estimated to be  $0.17$  (95% bootstrap CI =  $0.07, 0.27$ ) for FA1 and  $0.11$  (CI =  $-0.04, 0.31$ ) for FA2 [calculated according to Whitlock (1998) based on untransformed FA-values]. If the confidence intervals for  $h_N^2$  of FA and those for  $R$  are combined, and the heritability of developmental instability is estimated as  $h_N^2/R$ , the confidence intervals for the heritability of developmental instability include 0 and 1 for both FA1 and FA2.

### Relationship between fluctuating asymmetry and fitness

#### *Morphological variables*

Body size, condition, male plumage score and height of male forehead patch depended on age. In addition, average size and condition varied among years (Table 6). The analyses suggested no strong association between cFA and any of the other morphological variables (Table 7). The overall positive association between  $FA_{YR}$  and plumage score reached significance in the bootstrap test ( $P < 0.05$ ). The significant interaction between year and fluctuating asymmetry suggests that the positive association was most pronounced in 2001. The correlation between  $FA_{YR}$  and plumage score for 2001 alone is, however, not significant (partial correlation correcting for age,  $r_{PART} = 0.09$ , 95% bootstrap CI =  $-0.15, 0.35$ ).

Nestling body mass was not significantly correlated with nestling FA (FA1:  $r = 0.01$ , 95% bootstrap CI =  $-0.10, 0.11$ ,  $n = 390$ ; FA2:  $r = -0.07$ , CI =  $-0.16, 0.04$ ,  $n = 389$ ; cFA:  $r = -0.04$ , CI =  $-0.14, 0.06$ ,  $n = 390$ ).

#### *Arrival date*

Older, larger males with previous breeding experience arrived at the breeding area earlier (Table 6). Females arriving early were older and had lower overall condition. The random individual effect was significant for both sexes, suggesting consistence across years in individual arrival dates.

There was a tendency for the earliest arriving males to be most asymmetric (Table 7). This is opposite to what is expected if fluctuating asymmetry is negatively related to fitness. In analyses of each year independently, effect estimates are negative for all 3 years and of similar or larger magnitude to the global estimate both in 2001 and 2002. Since the observed pattern could be caused by higher measurement error of birds measured earliest in the season, we examined whether there was a decrease in FA estimates for birds measured twice by one observer (L.C.S.) in 2001. No strong trends were found (paired  $t$ -tests for traits 1–5:  $t_{26} = -0.1$ ,  $t_{24} = -0.3$ ,  $t_{10} = 0.8$ ,  $t_{10} = -1.6$ ,  $t_9 = 0.5$ , respectively, all  $P > 0.05$ ). The tests for FA effects on arrival were, however, somewhat influenced by skewness in the distribution of arrival dates, leading to strong leverage from the relatively few late-arriving males. If ranks of arrival dates were used as the response, the  $P$ -values for the unique and marginal effects of  $cFA_{YR}$  were both between 0.01 and 0.05. The effect of cFA on arrival date for females was in the opposite direction to that for males, but it was not significant.

**Table 6.** Co-factor effects on fitness parameters

Response	Linear predictor	$\sigma_{\text{IND}}$	$\sigma_{\text{RES}}$	df <sub>IND</sub>	df <sub>RES</sub>	link
SIZE <sub>M</sub>	$-0.63 \pm 0.13 + 0.37 \pm 0.12 Y_{2001} + 0.45 \pm 0.15 Y_{2002} + 0.50 \pm 0.13 \text{ AGE}$	1.00***	0.59	138	74	identity
SIZE <sub>F</sub>	$0.07 \pm 0.12 - 0.78 \pm 0.12 Y_{2001} - 0.07 \pm 0.14 Y_{2002} + 0.45 \pm 0.13 \text{ AGE}$	0.91***	0.54	178	37	identity
COND <sub>M</sub>	$-0.35 \pm 0.08 + 0.16 \pm 0.08 Y_{2001} + 0.33 \pm 0.09 Y_{2002} + 0.26 \pm 0.08 \text{ AGE}$	0.45***	0.40	129	66	identity
COND <sub>F</sub>	$0.07 \pm 0.10 - 0.65 \pm 0.11 Y_{2001} + 0.37 \pm 0.12 Y_{2002} + 0.22 \pm 0.10 \text{ AGE}$	0.51**	0.57	173	36	identity
PLUM <sub>M</sub>	$4.36 \pm 0.16 - 1.06 \pm 0.17 \text{ AGE}$	1.14***	0.79	132	70	identity
FOREHEAD <sub>M</sub>	$2.90 \pm 0.10 + 0.26 \pm 0.09 \text{ AGE}$	0.85***	0.38	132	67	identity
ARRIVAL <sub>M</sub>	$11.5 \pm 1.0 + 3.9 \pm 1.1 Y_{2001} + 6.4 \pm 1.2 Y_{2002} - 3.9 \pm 1.5 \text{ AGE} - 2.8 \pm 1.4 \text{ EXP} - 1.9 \pm 0.6$ SIZE <sub>ALL</sub>	4.0**	6.2	137	73	identity
ARRIVAL <sub>F</sub>	$13.8 \pm 1.0 + 6.6 \pm 1.1 Y_{2001} + 7.8 \pm 1.1 Y_{2002} - 2.8 \pm 1.0 \text{ AGE} + 3.9 \pm 1.4 \text{ COND}_{\text{ALL}}$	3.5*	5.9	172	37	identity
P.PAIRE <sub>D</sub> <sub>M</sub>	$1.6 \pm 0.3 + 1.3 \pm 0.5 \text{ AGE}$	0.0	1.4	133	69	logit
P.POLY <sub>G</sub> <sub>M</sub>	$-2.9 \pm 0.7 + 1.7 \pm 0.7 \text{ AGE} + 1.0 \pm 0.5 \text{ COND}_{\text{ALL}}$	0.3	1.2	119	64	logit
PAIRING <sub>M</sub>	$6.8 \pm 0.6 + 0.8 \pm 0.0 \text{ ARRIVAL} + 2.9 \pm 0.7 Y_{2001} + 1.7 \pm 0.8 Y_{2002}$	1.4	3.9	120	61	identity
LAYING <sub>F</sub>	$12.0 \pm 0.3 + 0.7 \pm 0.0 \text{ PAIRING} + 2.8 \pm 0.3 Y_{2001} + 1.2 \pm 0.3 Y_{2002}$	0.3	1.9	179	37	identity
CLUTCH-SIZE <sub>M</sub>	$7.6 \pm 0.2 - 0.05 \pm 0.01 \text{ LAYING}$	0.15	0.66	119	62	identity
CLUTCH-SIZE <sub>F</sub>	$7.3 \pm 0.2 - 0.05 \pm 0.01 \text{ LAYING} + 0.37 \pm 0.10 \text{ EXP} + 0.25 \pm 0.13 \text{ COND}_{\text{ALL}}$	0.29	0.56	171	38	identity
HATCH.SUCC <sub>M</sub>	$0.88 \pm 0.02 + 0.05 \pm 0.02 \text{ EXP} + 0.02 \pm 0.01 \text{ PLUM}_{\text{ALL}}$	0.00	0.14	115	61	identity
HATCH.SUCC <sub>F</sub>	$0.91 \pm 0.01$	0.00	0.14	180	37	identity
FLEDG.SUCC <sub>M</sub>	$1.11 \pm 0.10 - 0.06 \pm 0.02 \text{ HATCH.N} + 0.10 \pm 0.04 Y_{2001} + 0.04 \pm 0.04 Y_{2002} + 0.08 \pm 0.04 \text{ AGE}$	0.00	0.20	107	48	identity
FLEDG.SUCC <sub>F</sub>	$1.08 \pm 0.11 - 0.05 \pm 0.02 \text{ HATCH.N} + 0.10 \pm 0.04 Y_{2001} - 0.01 \pm 0.04 Y_{2002} - 0.39 \pm 0.04$ STATUS + 0.07 ± 0.03 AGE	0.12	0.20	163	25	identity
HATCH.N <sub>M</sub>	$4.8 \pm 0.4 + 2.1 \pm 0.4 \text{ AGE}$	1.5*	2.6	133	67	identity
HATCH.N <sub>F</sub>	$5.8 \pm 0.1$	0.46	0.90	180	38	identity
FLEDG.N <sub>M</sub>	$3.4 \pm 0.4 + 0.9 \pm 0.4 Y_{2001} + 0.1 \pm 0.4 Y_{2002} + 2.0 \pm 0.4 \text{ AGE}$	1.0	2.2	122	55	identity
FLEDG.N <sub>F</sub>	$4.4 \pm 0.3 + 0.7 \pm 0.3 Y_{2001} - 0.4 \pm 0.3 Y_{2002} + 0.6 \pm 0.3 \text{ AGE} + 0.3 \pm 0.2 \text{ COND}_{\text{YR}}$	1.1*	1.3	160	26	identity
FLEDG.WT <sub>M</sub>	$14.8 \pm 0.4 - 0.2 \pm 0.1 \text{ HATCH.N} + 0.5 \pm 0.2 Y_{2001} - 0.3 \pm 0.2 Y_{2002} + 0.2 \pm 0.1$ SIZE <sub>ALL</sub> + 0.6 ± 0.3 COND <sub>ALL</sub>	0.31	0.85	101	47	identity
FLEDG.WT <sub>F</sub>	$15.0 \pm 0.5 - 0.3 \pm 0.1 \text{ HATCH.N} + 0.5 \pm 0.2 Y_{2001} - 0.2 \pm 0.2 Y_{2002} - 1.2 \pm 0.2 \text{ STATUS}$	0.84**	0.68	147	25	identity
RECRUITS <sub>M</sub>	$-2.5 \pm 0.5 + 0.9 \pm 0.5 \text{ AGE}$	0.56	0.36	123	59	log
RECRUITS <sub>F</sub>	$-1.7 \pm 0.3 - 1.4 \pm 0.7 \text{ STATUS}$	0.67	0.34	163	31	log

Note: For each parameter a GLMM was fitted to the data from all 3 years (random intercept: Individual). Linear predictor: effect estimates ± SE.  $\sigma_{\text{IND}}$  = estimated SD of among-individual variation,  $\sigma_{\text{RES}}$  = estimated SD of within-individual among-year variation. For GLMMs with log or logit link functions, all estimates are shown on log or logit scales. See Table 3 for explanation of variables. \*  $P < 0.05$ , \*\*  $P < 0.01$ , \*\*\*  $P < 0.001$ .

**Table 7.** FA effects on fitness parameters: effect estimates  $\pm$  SE of cFA<sub>ALL</sub> or cFA<sub>YR</sub> added to the basic fitness models (Table 6)

Response	Unique cFA <sub>ALL</sub>	Marginal cFA <sub>ALL</sub>	d.f. <sub>IND</sub>	Unique cFA <sub>YR</sub>	Marginal cFA <sub>YR</sub>	d.f. <sub>RES</sub>	Pred.
SIZE <sub>M</sub>	+0.5 $\pm$ 0.6		127	-0.3 $\pm$ 0.4		66	-
SIZE <sub>F</sub>	-0.4 $\pm$ 0.4		171	-0.3 $\pm$ 0.3		32	-
COND <sub>M</sub>	+0.2 $\pm$ 0.3		125	+0.0 $\pm$ 0.2		64	-
COND <sub>F</sub>	+0.0 $\pm$ 0.3		166	+0.1 $\pm$ 0.2		31	-
PLUM <sub>M</sub>	-0.8 $\pm$ 0.7†		127	+0.8 $\pm$ 0.5‡		68	+
FOREHEAD <sub>M</sub>	+0.2 $\pm$ 0.5		127	+0.2 $\pm$ 0.3		66	-
ARRIVAL <sub>M</sub>	-6.8 $\pm$ 3.7	-7.7 $\pm$ 3.8*	126/127	-8.1 $\pm$ 2.8**	-8.6 $\pm$ 2.8**	65/65	+
ARRIVAL <sub>F</sub>	+3.0 $\pm$ 2.3	+2.6 $\pm$ 2.3	165/171	+2.1 $\pm$ 2.0	+1.9 $\pm$ 2.0	32/32	+
P.PAIRE <sub>M</sub>	+5.0 $\pm$ 2.1*		123	+2.9 $\pm$ 1.8		65	-
P.POLYG <sub>M</sub>	-1.7 $\pm$ 1.5	-1.7 $\pm$ 1.5	112/115	-0.3 $\pm$ 1.3	-0.3 $\pm$ 1.3	60/60	-
PAIRING <sub>M</sub>	-1.7 $\pm$ 2.3		115	-2.1 $\pm$ 1.9		58	+
LAYING <sub>F</sub>	+0.5 $\pm$ 0.6		173	+0.4 $\pm$ 0.6		32	+
CLUTCH-SIZE <sub>M</sub>	+0.2 $\pm$ 0.4		114	+0.2 $\pm$ 0.3		59	-
CLUTCH-SIZE <sub>F</sub>	+0.1 $\pm$ 0.2	+0.1 $\pm$ 0.2	165/173	+0.1 $\pm$ 0.2	+0.1 $\pm$ 0.2	33/33	-
HATCH.SUCC <sub>M</sub>	+0.06 $\pm$ 0.08	+0.03 $\pm$ 0.08	112/113	+0.07 $\pm$ 0.07	+0.05 $\pm$ 0.07	58/58	-
HATCH.SUCC <sub>F</sub>	-0.01 $\pm$ 0.05		173	-0.03 $\pm$ 0.04		34	-
FLEDG.SUCC <sub>M</sub>	+0.03 $\pm$ 0.11		105	+0.02 $\pm$ 0.10		47	-
FLEDG.SUCC <sub>F</sub>	+0.10 $\pm$ 0.08	+0.12 $\pm$ 0.10	158/158	+0.08 $\pm$ 0.07	+0.15 $\pm$ 0.08	22/23	-
HATCH.N <sub>M</sub>	+2.4 $\pm$ 1.5		123	+2.1 $\pm$ 1.2		64	-
HATCH.N <sub>F</sub>	-0.0 $\pm$ 0.3		173	-0.1 $\pm$ 0.3		34	-
FLEDG.N <sub>M</sub>	+1.5 $\pm$ 1.3		115	+1.5 $\pm$ 1.0		54	-
FLEDG.N <sub>F</sub>	+0.5 $\pm$ 0.6	+0.6 $\pm$ 0.6	155/158	+0.5 $\pm$ 0.5	+0.7 $\pm$ 0.5	22/24	-
FLEDG.WT <sub>M</sub>	+0.3 $\pm$ 0.6	+0.4 $\pm$ 0.6	99/101	+0.2 $\pm$ 0.5	+0.3 $\pm$ 0.5	46/46	-
FLEDG.WT <sub>F</sub>	-0.1 $\pm$ 0.4	-0.2 $\pm$ 0.5	143/143	+0.1 $\pm$ 0.3	+0.2 $\pm$ 0.4	22/23	-
RECRUITS <sub>M</sub>	-0.4 $\pm$ 1.3		116	+0.3 $\pm$ 1.0		57	-
RECRUITS <sub>F</sub>	+0.4 $\pm$ 0.8	+0.4 $\pm$ 0.8	158/158	+0.7 $\pm$ 0.7	+0.8 $\pm$ 0.7	28/29	-

Note: cFA<sub>ALL</sub> = individual composite FA averaged across years, cFA<sub>YR</sub> = a year-specific index. *Unique*: Effects of all co-factors are corrected for. *Marginal*: SIZE, COND, PLUM or STATUS are not included among co-factors. Pred. = predicted sign of effect. If not noted otherwise, the interactions between year and fluctuating asymmetry are not significant ( $P > 0.05$ ). See Table 3 for explanation of variables. For models with log or logit link functions (cf. Table 6), estimates are shown on log or logit scales. \* $P < 0.05$ , \*\* $P < 0.01$ .

† Possible interaction with year ( $F_{2,66} = 3.2$ ,  $P = 0.05$ ):

$$\text{PLUM}_M = -0.5 \pm 1.1 \text{ cFA}_{\text{ALL}} + 1.6 \pm 1.3 \text{ cFA}_{\text{ALL}} Y_{2001} - 1.6 \pm 1.4 \text{ cFA}_{\text{ALL}} Y_{2002}$$

‡ Possible interaction with year ( $F_{2,64} = 4.3$ ,  $P = 0.02$ ):

$$\text{PLUM}_M = -0.1 \pm 0.8 \text{ cFA}_{\text{YR}} + 2.7 \pm 1.1 \text{ cFA}_{\text{YR}} Y_{2001} + 0.1 \pm 1.1 \text{ cFA}_{\text{YR}} Y_{2002}$$

### Male pairing success

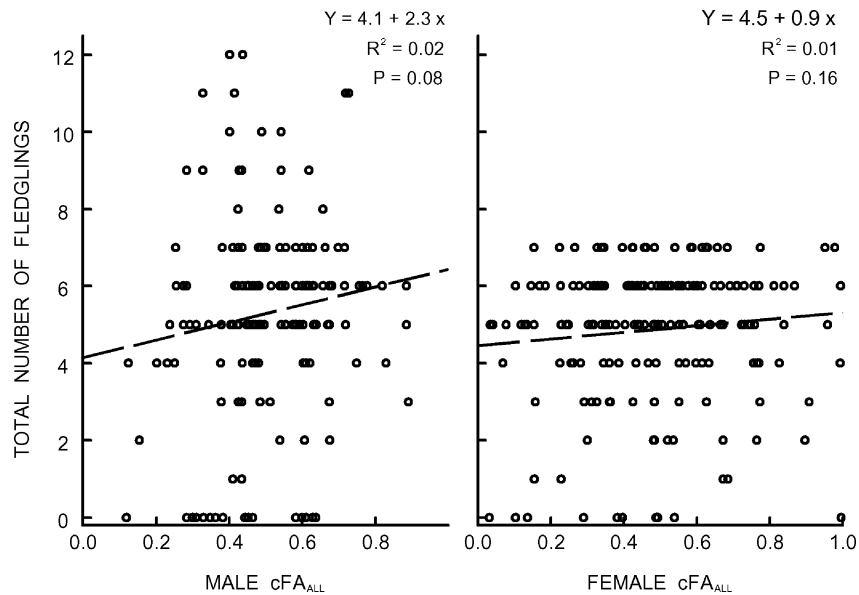
Older males had higher probability of pairing (P.PAIRE<sub>D</sub>) and of becoming polygynous if paired (P.POLYG). The probability of becoming polygynous was also positively associated with male overall condition. A third measure of attractiveness, the time span from arrival to pairing (PAIRING), was not associated with any of the measured male characteristics (Table 6). Associations with cFA were in the opposite direction to that expected for

P. PAIRED and PAIRING and in the expected direction for P.POLYG (Table 7). The only association that was significant was between P.PAIRED and  $cFA_{ALL}$ : more asymmetric males apparently had a better chance of finding a mate.

### Reproductive success

The various parameters contributing to reproductive success (fitness parameters 9–16, Table 3) were associated with several individual characteristics for both sexes (Table 6). Females with previous breeding experience had larger clutch sizes. Clutch size was also positively associated with female overall condition, while female year-specific condition was associated with larger numbers of fledglings. Older females had higher fledging success and a higher number of fledglings in their broods. Secondary or tertiary females of polygynous males had lower fledging success, lower average weights of the nestlings and fewer of their offspring were recruited into the breeding population. Broods of less brightly coloured males and males with previous breeding experience had higher hatching success. Male age was positively associated with total number of hatchlings, fledging success, total number of fledglings and number of offspring recruited into the breeding population. Average weight of fledglings was positively associated with male overall size and male overall condition.

There was no significant association between any of the above parameters and male or female  $cFA_{ALL}$  or  $cFA_{YR}$  based on parametric tests (Table 7). The association between male  $cFA_{ALL}$  and total number of hatchlings was significant by a bootstrap test ( $P < 0.05$ ), but the effect was in the opposite direction to that expected: more asymmetric males tended to produce larger numbers of hatchlings. The best measure of yearly reproductive success is probably the total number of fledglings, which was non-significantly positively associated with  $cFA_{ALL}$  and  $cFA_{YR}$  (Table 7, Fig. 2).



**Fig. 2.** Total number of fledglings in response to  $cFA_{ALL}$  for males and females. Data from all years are pooled. A linear regression line is fitted to each data set. The trends are not significant.

*Return rate*

Of the 448 nestlings that were assessed for fluctuating asymmetry in 2000 or 2001, 23 were recaptured in one of the two following years. The probability of recapture depended positively on nestling body mass (g) at day 13 after hatching (BM) and negatively on hatching date (DAY) (GLMM, logit link; linear predictor =  $-2.5$  (SE = 3.9)  $- 0.10$  (SE = 0.05) DAY + 0.38 (SE = 0.20) BM; estimated SD among families  $\sigma_{AF} = 0.00$ , SD among progeny  $\sigma_{AP} = 0.36$ ). The estimated effect of cFA was small and in the opposite direction to that predicted (unique effect of cFA = +0.02 (SE = 1.16); marginal effect = +0.44 (SE = 1.05), logit scale; d.f.<sub>FAM</sub> = 84 and 120, d.f.<sub>PROG</sub> = 301 and 325, respectively, for unique and marginal effects).

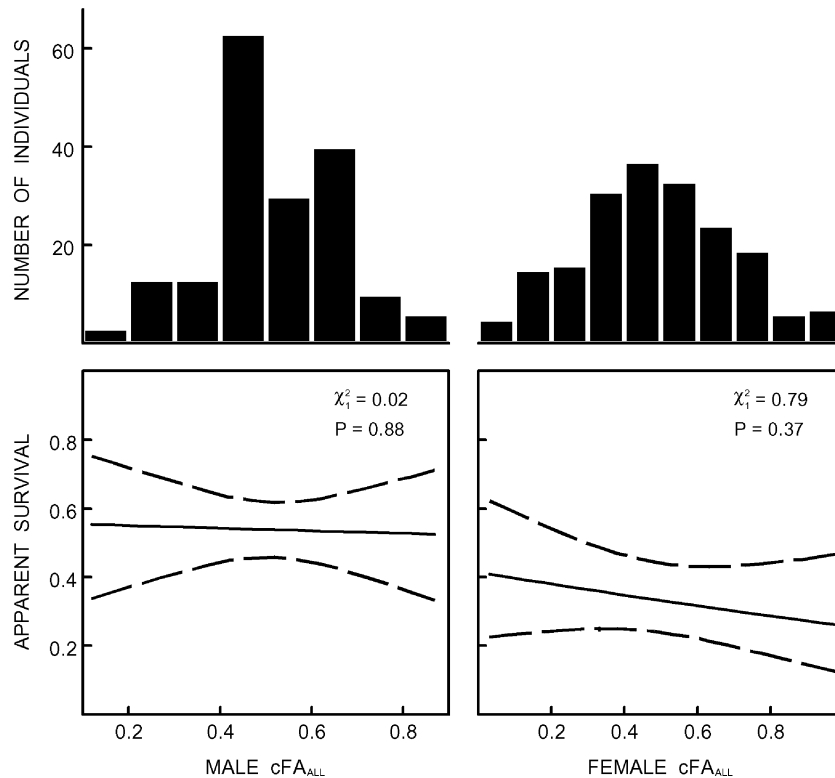
On average, 48% of the captured adult males and 27% of the females that were observed to stay in the study area for more than 3 days in a given year were observed in a later year. None of the males but three of the females present in a given year were observed again after not having been observed the subsequent year (capture history: 101). For males, the recapture probability (conditioned on survival) was fixed to 1 in the model fitting. For females, it was estimated to be 76%. Apparent survival of both sexes increased with body size. For males, the data also suggested a negative age effect. The estimated effects of cFA were of small magnitude, but for females in the predicted direction (Table 8, Fig. 3).

Since the capture–recapture method does not take into consideration age differences of birds when they are caught the first time, we tested for differences in FA of birds caught the first time as 1-year-olds or aged 2 years or older. If older birds had lower FA, it would suggest differential survival (in the absence of age-related changes in FA). This was not observed. Older males had 2% higher cFA<sub>YR</sub> (CI =  $-14\%$ ,  $+17\%$ , total  $n = 97$ ) and 2% lower cFA<sub>ALL</sub> (CI =  $-15\%$ ,  $12\%$ ,  $n = 103$ ) than younger males, and older females had 10% higher cFA<sub>YR</sub> (CI =  $-6\%$ ,  $27\%$ ,  $n = 146$ ) and 5% higher cFA<sub>ALL</sub> (CI =  $-11\%$ ,  $22\%$ ,  $n = 149$ ) than younger females.

**Table 8.** Effects of fluctuating asymmetry and co-factors on adult survival (maximum likelihood capture–recapture estimation)

Response	Estimated co-factor effects				d.f. <sub>IND</sub>
SURV <sub>M</sub>	0.5 ± 0.3 + 0.2 ± 0.3 Y <sub>2001</sub> - 0.6 ± 0.4 Y <sub>2002</sub> - 0.6 ± 0.3 AGE + 0.3 ± 0.2 SIZE <sub>ALL</sub>				130
SURV <sub>F</sub>	-0.1 ± 0.4 - 0.9 ± 0.4 Y <sub>2001</sub> - 1.0 ± 0.4 Y <sub>2002</sub> + 0.3 ± 0.2 SIZE <sub>ALL</sub>				173
	cFA <sub>ALL</sub> Unique	Marginal	cFA <sub>YR</sub> Unique	Marginal	
SURV <sub>M</sub>	-0.2 ± 1.0	+0.1 ± 1.0	+0.2 ± 0.9	+0.4 ± 0.9	119/120/119/120
SURV <sub>F</sub>	-0.7 ± 0.8	-0.8 ± 0.8	-0.4 ± 0.7	-0.5 ± 0.7	167/168/167/168

*Note:* SURV<sub>M</sub> and SURV<sub>F</sub> = apparent survival of males and females. Effect estimates ± SE are given on a logit scale. cFA<sub>ALL</sub> = individual composite FA averaged across years, cFA<sub>YR</sub> = a year-specific index. *Unique:* all co-factors are included. *Marginal:* SIZE is not included among co-factors. *Recapture probabilities:* Males: 1.0 (fixed). Females: 0.76 (estimated).



**Fig. 3.** Estimated survival probability as a function of fluctuating asymmetry for males and females. Bottom panels: Predicted unique effect of  $cFA_{ALL} \pm 2$  SE. Survival is averaged across years, age groups and body sizes. Fluctuating asymmetry has no significant effect on survival. Top panels:  $cFA_{ALL}$  distribution for each sex.

### Single FA traits

Tests similar to those of associations between  $cFA$  and fitness components were conducted for single FA traits. Among 347 tests in which the sample size was at least 10, 5.2% were significant at the 5% level (in the expected direction: 3 tests; in the opposite direction: 15 tests) and 1.2% at the 1% level (in the expected direction: 0 tests; in the opposite direction: 4 tests). Twelve tests (3.5%) suggested significant ( $P < 0.05$ ) interactions between fluctuating asymmetry and year. There was no apparent pattern, suggesting that some traits were more sensitive than others, or any consistency in fluctuating asymmetry–year interactions.

### Power

Results of the power simulation were similar for different fitness variables, and are therefore averaged across variables for ease of presentation. The estimated power depended mainly on the coefficient of variation of developmental instability ( $CV_{DI}$ ). With a  $CV_{DI}$  of 23, the probability of finding significant associations between fluctuating asymmetry and fitness variables was less than 0.4, even if a very strong effect of developmental instability on fitness was assumed (Table 9). This was associated with low correlations between fluctuating

**Table 9.** Power of tests of associations between composite FA (cFA) and normal distributed fitness variables (variables with identity link, Table 6)

DI–fitness effect, $k$ :	$CV_{DI}$					
	23		45		100	
	1	2	1	2	1	2
$r^2$ , fitness–DI	0.43	0.74	0.43	0.74	0.42	0.73
$r^2$ , single-trait FA–DI	0.03	0.03	0.09	0.09	0.27	0.27
$r^2$ , cFA–DI	0.05	0.05	0.13	0.13	0.30	0.30
$r^2$ , cFA–fitness	0.02	0.04	0.06	0.10	0.14	0.23
Power	0.23	0.37	0.69	0.88	0.98	1.00

*Note:* The estimated power is the proportion of test results significant at a 1% level for 500 simulated data sets, averaged across fitness variables. Simulation regimes differed in the coefficient of variation for developmental instability ( $CV_{DI}$ ) and the strength of the effect of DI on fitness ( $k$ , defined in Methods). Individual-specific DI-values were generated from gamma distributions and affected both fitness and the standard deviation of fluctuating asymmetry. The FA and fitness data were randomly generated based on parameters estimated from the real data and analysed as the real data set.

asymmetry and fitness variables ( $r^2 \leq 0.04$ ) and between fluctuating asymmetry and the underlying developmental instability ( $r^2 \leq 0.05$ ). The simulation regime with  $CV_{DI} = 45$  gave fluctuating asymmetry–fitness correlations ( $r^2 = 0.06–0.10$ ) closest to the average effect estimates from the meta-analysis by Møller (1999; growth:  $r^2 = 0.02$ ; fecundity:  $r^2 = 0.12$ ; survival:  $r^2 = 0.06$ ). With a  $CV_{DI}$  of 45 the estimated power was moderate (0.7–0.9). A  $CV_{DI}$  of 100 led to estimated power near 1.0 and correlations between fluctuating asymmetry and fitness somewhat higher than the averages reported in the literature ( $r^2 = 0.14–0.23$ ).

## DISCUSSION

### Ontogeny of fluctuating asymmetry

Individual signed FA-values of primary feather traits were found to persist from the nestling stage to adulthood. The primaries do not moult in the first year of life (Jenni and Winkler, 1994), so this mainly shows that asymmetries persist during the later stages of feather growth and can be measured reliably in both nestlings and adults. More interesting biologically is the finding that individual signed asymmetries persist across years in adults, possibly for 2 years or more. Both primaries and tail feathers of adults moult each year (Jenni and Winkler, 1994). The trends for each of traits 3 and 4 (fluctuating asymmetry of tail patches) are not significant separately, but together they strongly suggest that asymmetries not only in feather lengths but also in colour patterns persist across moults. Analyses on trait 1 (fluctuating asymmetry of primaries 1 and 2) suggest that the individual among-year variation in asymmetry is of small to moderate magnitude. In other words, feather asymmetry is to a large degree fixed early in life, with limited room for later modifications. Such a pattern could potentially suggest heritability of signed asymmetry (Chippindale and Palmer, 1993), but the results of the heritability analyses indicate that heritability of signed FA is low or non-existent. It is more likely that the asymmetry of a given feather is determined

by stochastic incidents during early development that permanently affect the feather follicles on each side.

Persistence of signed FA across moults has also been observed on the tail feathers of adult barn swallows *Hirundo rustica* (Møller, 1996). It has been shown experimentally that the primary feathers of house sparrow *Passer domesticus* develop independently on each side towards an asymptotic size (Aparicio, 1998). The results on pied flycatchers and barn swallows show that the asymptotic size is itself subject to random variation among the two sides of an individual. Developmental instability can thus affect morphogenesis at two levels, the morphogenesis of feather follicles and the morphogenesis of feathers. Different mechanisms may be involved in the two processes. For example, while results from the house sparrow (Aparicio, 1998) show that feedback between sides does not occur in feather morphogenesis, this has no implication for morphogenesis of feather follicles. Our results suggest that disturbances affecting follicle development contribute most to the fluctuating asymmetry of feathers. This implies that the fluctuating asymmetry of bird feathers potentially reflects genetic factors or environmental conditions during embryo development (e.g. maternal effects), and that later environmental effects may be less important. It is, however, possible that more extreme conditions can disturb feather morphogenesis to a larger degree, as indicated by the finding that feathers developed under adverse environmental conditions are less symmetric than feathers on the same birds developed under more benign conditions (Møller, 1996).

### Heritability of fluctuating asymmetry

Both the offspring–parent regression and the full sibling analysis suggested low narrow-sense heritability of fluctuating asymmetry. In the sibling analysis, the among-family variance component reflects additive genetic variance, fractions of dominance and epistatic variance, and variance due to shared environmental effects (Roff, 1997). The low estimated among-family variance thus not only suggests low narrow-sense heritability of nestling fluctuating asymmetry, but also that environmental effects shared among nestlings have little influence on their fluctuating asymmetry. Almost all variation in fluctuating asymmetry was found at the within-nest level. It would therefore be of no purpose to test for associations between offspring fluctuating asymmetry and parent characteristics. Within-nest differences in condition, estimated by nestling body mass, could not explain the variation in fluctuating asymmetry. Nestling FA levels thus appear to be entirely random, with no detectable impact of either genetic or environmental factors.

An objection to the heritability estimation is that the fluctuating asymmetry of nestlings and that of adults were measured at different stages of feather development. Although signed FA-values were strongly correlated between birds captured as nestlings and adults, the moderate correlation between unsigned FA-values (Table 4) adds strength to this objection. The discrepancies reflect a combination of measurement error, wear (although visibly worn feathers were excluded from the analyses) and changes in the magnitude of fluctuating asymmetry in the final stages of feather development. Average values of FA1 and FA2 (in millimetres) were approximately 50% higher in adults than in nestlings (analyses not shown). This suggests that the changes in fluctuating asymmetry from nestlings to adults are not caused by corrections of asymmetries during the final stages of feather development, but rather by additional noise. Unless the sources of developmental noise or the buffering against it change fundamentally during feather development, nestling and adult fluctuating

asymmetry can be assumed to reflect the same components of the underlying developmental instability. The heritability estimates should then be valid. Furthermore, although sample size was too low to draw strong conclusions, the offspring–parent regression based on offspring measured as adults did not suggest a fundamentally different pattern than analyses based on nestlings.

Another potentially confounding factor is extra-pair paternity, which leads to lower average relatedness between ‘social’ fathers and offspring, and also to erroneous estimation of the reproductive success of males. However, the frequency of extra-pair paternity in our study population of pied flycatchers is low (< 10% of nestlings; Lifjeld *et al.*, 1991; Ellegren *et al.*, 1995; Slagsvold *et al.*, 2001), and it did probably not affect any of the results to a large extent.

Low heritability of fluctuating asymmetry is in line with results of most FA studies (Fuller and Houle, 2003). Due to the presumably low correlation between fluctuating asymmetry and developmental instability, the heritability of developmental instability may be considerably higher than that of fluctuating asymmetry (Whitlock, 1996). The hypothetical repeatability,  $R$ , quantifies the amount of variation in fluctuating asymmetry that is due to variation in developmental stability. The estimation of  $R$  relies on a number of assumptions, however, and estimates of  $R$  should be interpreted with some caution (Swaddle, 2003). Due to uncertainty in both the estimation of  $R$  and  $h^2_N$ , it is probably not possible to make inferences about the heritability of developmental instability from the present data alone. This was demonstrated by combining the bootstrap confidence intervals of  $R$  and  $h^2_N$ , which indicated that the heritability of developmental instability could be anywhere from 0 to 1. The low power of single FA studies testing for heritability of developmental instability has also been demonstrated using simulations by Fuller and Houle (2003).

Whether heritability of developmental instability is truly low or fluctuating asymmetry is simply loosely correlated with developmental instability, the low heritability of fluctuating asymmetry does not support the hypothesis that it is a good indicator of genetic quality. Fluctuating asymmetry could, however, potentially be associated with observed individual fitness due to environmental effects affecting fluctuating asymmetry and fitness simultaneously. Fluctuating asymmetry could also play a role in sexual selection, if low fluctuating asymmetry is associated with high phenotypic quality and high parental investment (Swaddle, 2003).

### **The relationship between fluctuating asymmetry and individual fitness**

Almost all relations between fluctuating asymmetry and fitness components were weak and non-significant. Among 82 tests for associations between cFA and fitness components, four were significant at the 5% level, two at the 1% level, and two interactions between cFA and year were significant at a 5% level. Twenty-six of the estimated effects were in the expected direction, while 56 were in the opposite direction to that expected from the hypothesis of a negative relation between fluctuating asymmetry and fitness. There were no apparent systematic differences regarding the response to fluctuating asymmetry between different types of fitness traits, different single-trait FA indices or between the sexes. We have no plausible biological explanation to the finding that more asymmetric males arrived earlier and had higher probabilities of finding a mate. If there had been a functional effect of wing or tail asymmetry in pied flycatchers on flight performance (e.g. Swaddle, 1997; Swaddle and Witter, 1998), or if low fluctuating asymmetry reflected good condition, we would have expected the symmetric individuals to arrive earliest, not latest. The apparent year-dependent

association between fluctuating asymmetry and plumage score could imply that in presumably tough years more asymmetric males became less brightly coloured. However, the other analyses did not suggest similar interactions with year. Taking into consideration the large number of tests, these effects are likely to result from sampling error.

Age was positively associated with attractiveness of males and reproductive success of both males and females. Due to the large reproductive benefit of reaching high age, survival is probably the single factor most closely related to individual fitness. Furthermore, since only high-quality individuals are able to accomplish the demanding journey back and forth to the wintering grounds in tropical Africa, survival is a good indicator of genetic and phenotypic quality. In the survival analyses, permanent emigration is confounded with estimated survival. Natal dispersal is high in pied flycatchers, while breeding dispersal is intermediate in females and low in males (Lundberg and Alatalo, 1992). The present results confirmed this pattern: the return rate was 5% for nestlings, 27% for adult females and 48% for adult males. The survival rates of adult pied flycatchers are assumed to be generally around 40–50% for both sexes, and for populations to remain stable around 20–25% of the fledglings have to enter the breeding population (Lundberg and Alatalo, 1992). If fluctuating asymmetry is correlated with the tendency to disperse, dispersal will bias tests for effects of fluctuating asymmetry on survival. Such a bias could occur if presumably high-quality, symmetric individuals are better able to bear the costs of dispersal, having simultaneously higher survival probabilities and tendencies to disperse. This possibility cannot be tested directly with our data, but the finding that the body sizes of both nestlings and adults were positively associated with return rates suggests that individuals of high quality did not necessarily disperse more. Moreover, it can be argued that high-quality individuals should have a lower not higher tendency to disperse due to higher competitiveness combined with the presumed fitness benefits of breeding in a familiar territory. We therefore do not think dispersal biased the results. There was a non-significant trend in the expected direction for females: more symmetric individuals had higher estimated survival (Fig. 3). While such an effect, if real, would be important in an evolutionary context, much larger sample sizes are needed to demonstrate it conclusively. The lack of any effects for males and nestlings does not support the general existence of survival effects of fluctuating asymmetry, however. The result for nestlings is particularly important because it demonstrates that differential juvenile survival cannot explain why there is no association between fluctuating asymmetry and fitness among adults.

The results therefore fail to demonstrate that fluctuating asymmetry in pied flycatchers is predictably associated with condition, male attractiveness, reproductive success or survival. In the absence of positive results, the power to detect an effect if it exists is important. The power of single studies investigating the relationship between fluctuating asymmetry and individual fitness is generally expected to be low (see review by Lens *et al.*, 2002b). An important reason for the low power is that the relationship between fluctuating asymmetry and fitness is expected predominantly to be indirect, mediated through developmental stability, and fluctuating asymmetry estimates developmental instability with great uncertainty (Palmer, 1994; see also above). Composite FA indices are used to increase the accuracy in the estimation of organism-wide developmental instability (e.g. Leung *et al.*, 2000). In our study, we measured fluctuating asymmetry in five different traits with reasonable accuracy. Precision was also increased by measuring fluctuating asymmetry on the same individuals in several years. Furthermore, the number of individuals in our study is higher than in most studies included in a meta-analysis on the relationship of fluctuating asymmetry with fitness and stress by

Leung and Forbes (1996). The study is also strengthened by the assessment of a range of fitness components over several years. This increased the ability to detect possible trade-offs among fitness components (Clarke, 2003), and the probability of detecting fitness effects that are only evident in stressful conditions (Lens *et al.*, 2002a,b; Hendrickx *et al.*, 2003). We therefore believe that the present study does not have lower statistical power than many previous studies that reported significant associations between fluctuating asymmetry and fitness. The computer simulation suggested reasonably high power ( $\geq 0.7$ ) to detect fluctuating asymmetry–fitness associations if the coefficient of variation of developmental instability ( $CV_{DI}$ ) was 45 or more. The realism of this assumption may be questioned, however. Based on reported kurtosis values of fluctuating asymmetry, Gangestad and Thornhill (2003) estimated mean  $CV_{DI}$  for different species to be in the range 20–25. With a  $CV_{DI}$  of 23, the power was estimated to be less than 0.4 in the present study. However, in our simulation model such a low  $CV_{DI}$  could not possibly lead to correlations between fluctuating asymmetry and fitness as high as that reported in the literature as being typical for fluctuating asymmetry–fecundity correlations ( $r^2 = 0.12$ ; Møller, 1999). In our model, we implicitly assumed that the relatively high reported fluctuating asymmetry–fecundity correlations are due to generally higher  $CV_{DI}$  than suggested by Gangestad and Thornhill (2003). Other explanations are also possible. For example, fluctuating asymmetry may be a more precise measure of developmental instability than assumed here due to a non-linear relation between fluctuating asymmetry and the underlying developmental instability (Klingenberg, 2003), or fecundity may be more directly affected by fluctuating asymmetry through functional effects of asymmetry (Swaddle, 1997; Swaddle and Witter, 1998). We do not expect that such effects, if applicable to the study population, would lead to lower power than estimated. On the other hand, it is also possible that  $CV_{DI}$  is truly low and the adopted model is essentially correct. This would imply that the power to detect indirect associations between fluctuating asymmetry and fitness is low in this as well as in other studies. This option appears to be at odds with the reported general strength of fluctuating asymmetry–fecundity associations.

Given our limited knowledge on the developmental basis of fluctuating asymmetry (Klingenberg, 2003), we do not know whether the lack of significant effects in this and other studies is due to weak relations between fluctuating asymmetry and developmental instability or weak relations between developmental instability and fitness. It is also difficult at present to assess the extent to which differences among studies are based on biological differences or statistical (and possibly publication) artifacts. Several biological factors may affect the strength of fluctuating asymmetry–fitness associations. For example, interactions with environmental conditions have been demonstrated for Taita thrushes (*Turdus helleri*), where tarsus FA is more strongly negatively related to individual survival in more disturbed habitats (Lens *et al.*, 2002a). The choice of traits on which fluctuating asymmetry is measured may be important, although Leung and Forbes (1996) failed to find any difference in strength of fluctuating asymmetry–fitness relations between sexually and naturally selected traits or between functional and non-functional traits. A speculative explanation for the weak fluctuating asymmetry–fitness association in pied flycatchers is that the strong selection resulting from long-distance migration removes low-quality individuals before they have the chance to become measured and monitored by us, but this fails to explain why the wing and tail FA of barn swallows (*Hirundo rustica*) is predictably associated with several fitness components (Møller, 1994). Perhaps the lack of any association between fluctuating asymmetry and fitness in the study population mainly demonstrates the unpredictability

of fluctuating asymmetry–fitness relations, and shows that other morphometric measures (size, condition) are better predictors of individual fitness than fluctuating asymmetry in pied flycatchers.

To summarize, the results suggest that individual fluctuating asymmetries in plumage characters are determined during the development of the feather follicles and persist throughout life. The heritability of both signed and unsigned FA is low or non-existent, suggesting that fluctuating asymmetry is determined by environmental conditions or stochasticity at an early stage. Fluctuating asymmetry is not associated with the condition of nestlings or adults and is weakly or not at all associated with individual fitness. The results therefore provide no support for the use of fluctuating asymmetry as an indicator of genetic quality or phenotypic condition. Variation in fluctuating asymmetry within natural populations of pied flycatchers appears to be of little biological significance.

#### ACKNOWLEDGEMENTS

The study was financed by the Norwegian Research Council. We thank Hans Petter Rømme, Trude I.T. Starholm, Bodhild Brunsvik Rekkedal, Pål Abrahamsen and Per Kristian Slagsvold for help in the field. We also thank Helene M. Lampe for information on recaptured birds in 2003 and Dorothee Ehrlich and an anonymous reviewer for helpful and constructive comments on the manuscript.

#### REFERENCES

- Anderson, D.R. and Burnham, K.P. 1999. Understanding information criteria for selection among capture–recapture or ring recovery models. *Bird Study*, **46** (suppl.): S14–S21.
- Aparicio, J.M. 1998. Patterns of fluctuating asymmetry in developing primary feathers: a test of the compensational growth hypothesis. *Proc. R. Soc. Lond. B*, **265**: 2353–2357.
- Bjorksten, T.A., Fowler, K. and Pomiankowski, A. 2000. What does sexual trait FA tell us about stress? *Trends Ecol. Evol.*, **15**: 163–166.
- Broström, G. 2003. Generalized linear models with random intercepts. Department of Statistics, Umeå University, Sweden (available at: <http://www.stat.umu.se/forskning/glmML.pdf>).
- Carbonell, R., Pérez-Tris, J. and Tellería, J.L. 2003. Effects of habitat heterogeneity and local adaptation on the body condition of a forest passerine at the edge of its distributional range. *Biol. J. Linn. Soc.*, **78**: 479–488.
- Chippindale, A.K. and Palmer, A.R. 1993. Persistence of subtle departures from symmetry over multiple molts in individual brachyuran crabs: relevance to developmental stability. *Genetica*, **89**: 185–199.
- Clarke, G.M. 1995. Relationships between fluctuating asymmetry and fitness: how good is the evidence? *Pac. Conserv. Biol.*, **2**: 146–149.
- Clarke, G.M. 1998. Developmental stability and fitness: the evidence is not quite so clear. *Am. Nat.*, **152**: 762–766.
- Clarke, G.M. 2003. Developmental stability–fitness relationships in animals: some theoretical considerations. In *Developmental Instability: Causes and Consequences* (M. Polak, ed.), pp. 187–195. Oxford: Oxford University Press.
- Dale, S., Slagsvold, T., Lampe, H.M. and Sætre, G.-P. 1999. Population divergence in sexual ornaments: the white forehead patch of Norwegian pied flycatchers is small and unsexy. *Evolution*, **53**: 1235–1246.
- Dale, S., Slagsvold, T., Lampe, H.M. and Lifjeld, J.T. 2002. Age-related changes in morphological characters in the pied flycatcher: patterns of delayed maturation and senescence. *Avian Sci.*, **2**: 153–166.

- Drost, R. 1936. Über das Brutkleid männlicher Trauenfliegenfänger, *Muscicapa hypoleuca*. *Vogelzug*, **6**: 179–186.
- Eeva, T., Tanhuanpää, S., Råbergh, C. *et al.* 2000. Biomarkers and fluctuating asymmetry as indicators of pollution-induced stress in two hole-nesting passerines. *Funct. Ecol.*, **14**: 235–243.
- Efron, B. and Tibshirani, R.J. 1993. *An Introduction to the Bootstrap*. New York: Chapman & Hall.
- Ellegren, H., Lifjeld, J.T., Slagsvold, T. and Primmer, C.R. 1995. Handicapped males and extrapair paternity in pied flycatchers: a study using microsatellite markers. *Mol. Ecol.*, **4**: 739–744.
- Fuller, R.C. and Houle, D. 2003. Inheritance of developmental instability. In *Developmental Instability: Causes and Consequences* (M. Polak, ed.), pp. 157–183. Oxford: Oxford University Press.
- Gangestad, S.W. and Thornhill, R. 1999. Individual differences in developmental precision and fluctuating asymmetry: a model and its implications. *J. Evol. Biol.*, **12**: 402–416.
- Gangestad, S.W. and Thornhill, R. 2003. Fluctuating asymmetry, developmental instability, and fitness: toward model-based interpretation. In *Developmental Instability: Causes and Consequences* (M. Polak, ed.), pp. 62–80. Oxford: Oxford University Press.
- González-Guzmán, L.I. and Mehlman, D.W. 2001. Developmental stability across the breeding distribution of the scissor-tailed flycatcher (*Tyrannus forficatus*). *Ecol. Lett.*, **4**: 444–452.
- Hendrickx, F., Maelfait, J.-P. and Lens, L. 2003. Relationships between fluctuating asymmetry and fitness within and between stressed and unstressed populations of the wolf spider *Pirata piraticus*. *J. Evol. Biol.*, **16**: 1270–1279.
- Houle, D. 2000. A simple model of the relationship between asymmetry and developmental stability. *J. Evol. Biol.*, **13**: 720–730.
- Insightful. 2001. *S-PLUS 6 for Windows Guide to Statistics*, Vol. 1. Seattle, WA: Insightful Corporation.
- Jenni, L. and Winkler, R. 1994. *Moult and Ageing of European Passerines*. London: Academic Press.
- Karlsson, L., Perrson, K. and Wallinder, G. 1986. Ageing and sexing in pied flycatchers, *Ficedula hypoleuca*. *Vår Fågelvärld*, **45**: 131–146.
- Kellner, J.R. and Alford, R.A. 2003. The ontogeny of fluctuating asymmetry. *Am. Nat.*, **161**: 931–947.
- Klingenberg, C.P. 2003. A developmental perspective on developmental instability: theory, models, and mechanisms. In *Developmental Instability: Causes and Consequences* (M. Polak, ed.), pp. 14–34. Oxford: Oxford University Press.
- Lens, L., van Dongen, S. and Matthysen, E. 2002a. Fluctuating asymmetry as an early warning system in the critically endangered Taita Thrush. *Conserv. Biol.*, **16**: 479–487.
- Lens, L., van Dongen, S., Kark, S. and Matthysen, E. 2002b. Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies? *Biol. Rev.*, **77**: 27–38.
- Leung, B. and Forbes, M.R. 1996. Fluctuating asymmetry in relation to stress and fitness: effects of trait type as revealed by meta-analysis. *Ecoscience*, **3**: 400–413.
- Leung, B., Forbes, M.R. and Houle, D. 2000. Fluctuating asymmetry as a bioindicator of stress: comparing efficacy of analyses involving multiple traits. *Am. Nat.*, **155**: 101–115.
- Lifjeld, J.T., Slagsvold, T. and Lampe, H.M. 1991. Low frequency of extrapair paternity in pied flycatchers revealed by DNA fingerprinting. *Behav. Ecol. Sociobiol.*, **29**: 95–101.
- Lundberg, A. and Alatalo, R.V. 1992. *The Pied Flycatcher*. London: T & AD Poyser.
- Merilä, J. and Björklund, M. 1995. Fluctuating asymmetry and measurement error. *Syst. Biol.*, **44**: 97–101.
- Møller, A.P. 1994. Sexual selection in the barn swallow (*Hirundo rustica*). IV. Patterns of fluctuating asymmetry and selection against asymmetry. *Evolution*, **48**: 658–670.
- Møller, A.P. 1996. Development of fluctuating asymmetry in tail feathers of the barn swallow *Hirundo rustica*. *J. Evol. Biol.*, **9**: 677–694.
- Møller, A.P. 1999. Asymmetry as a predictor of growth, fecundity and survival. *Ecol. Lett.*, **2**: 149–156.
- Møller, A.P. and Cuervo, J.J. 2003. Asymmetry, size, and sexual selection: factors affecting hetero-

- genity in relationships between asymmetry and sexual selection. In *Developmental Instability: Causes and Consequences* (M. Polak, ed.), pp. 262–275. Oxford: Oxford University Press.
- Møller, A.P. and Swaddle, J.P. 1997. *Asymmetry, Developmental Stability, and Evolution*. Oxford: Oxford University Press.
- Møller, A.P. and van Dongen, S. 2003. Ontogeny of asymmetry and compensational growth in elm *Ulmus glabra* leaves under different environmental conditions. *Int. J. Plant Sci.*, **164**: 519–526.
- Palmer, A.R. 1994. Fluctuating asymmetry: a primer. In *Developmental Instability: Its Origins and Evolutionary Implications* (T.A. Markow, ed.), pp. 335–364. Dordrecht: Kluwer Academic.
- Palmer, A.R. 2000. Quasireplication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annu. Rev. Ecol. Syst.*, **31**: 441–480.
- Palmer, A.R. and Strobeck, C. 1986. Fluctuating asymmetry: measurement, analysis, patterns. *Annu. Rev. Ecol. Syst.*, **17**: 391–421.
- Palmer, A.R. and Strobeck, C. 1992. Fluctuating asymmetry as a measure of development stability: implications of non-normal distributions and power of statistical tests. *Acta Zool. Fenn.*, **191**: 57–72.
- Palmer, A.R. and Strobeck, C. 2003. Fluctuating asymmetry analyses revisited. In *Developmental Instability: Causes and Consequences* (M. Polak, ed.), pp. 279–319. Oxford: Oxford University Press.
- Pinheiro, J.C. and Bates, D.M. 2000. *Mixed-Effects Models in S and S-PLUS*. New York: Springer.
- R Development Core Team. 2003. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing (available at: <http://www.R-project.org>).
- Roff, D.A. 1997. *Evolutionary Quantitative Genetics*. New York: Chapman & Hall.
- Simmons, L.W., Tomkins, J.L., Kotiaho, J.S. and Hunt, J. 1999. Fluctuating paradigm. *Proc. R. Soc. Lond. B*, **266**: 593–595.
- Slagsvold, T., Johnsen, A., Lampe, H.M. and Lifjeld, J.T. 2001. Do female pied flycatchers seek extrapair copulations with familiar males? A test of the incomplete knowledge hypothesis. *Behav. Ecol.*, **4**: 412–418.
- Sokal, R.R. and Rohlf, F.J. 1995. *Biometry: The Principles and Practice of Statistics in Biological Research*. New York: Freeman.
- Swaddle, J.P. 1997. Within-individual changes in developmental stability affect flight performance. *Behav. Ecol.*, **8**: 601–604.
- Swaddle, J.P. 2003. Fluctuating asymmetry, animal behavior, and evolution. *Adv. Stud. Behav.*, **32**: 169–205.
- Swaddle, J.P. and Witter, M.S. 1998. Cluttered habitats reduce wing asymmetry and increase flight performance in European starlings. *Behav. Ecol. Sociobiol.*, **42**: 281–287.
- Sætre, G.-P., Dale, S. and Slagsvold, T. 1994. Female pied flycatchers prefer brightly coloured males. *Anim. Behav.*, **48**: 1407–1416.
- Tomkins, J.L. and Simmons, L.W. 2003. Fluctuating asymmetry and sexual selection: paradigm shifts, publication bias, and observer expectation. In *Developmental Instability: Causes and Consequences* (M. Polak, ed.), pp. 231–261. Oxford: Oxford University Press.
- van Dongen, S., Molenberghs, G. and Matthysen, E. 1999. The statistical analysis of fluctuating asymmetry: REML estimation of a mixed regression model. *J. Evol. Biol.*, **12**: 94–102.
- Van Valen, L. 1962. A study of fluctuating asymmetry. *Evolution*, **16**: 125–142.
- White, G.C. and Burnham, K.P. 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study*, **46** (suppl.): 120–138.
- Whitlock, M. 1996. The heritability of fluctuating asymmetry and the genetic control of developmental stability. *Proc. R. Soc. Lond. B*, **263**: 849–854.
- Whitlock, M. 1998. The repeatability of fluctuating asymmetry: a revision and extension. *Proc. R. Soc. Lond. B*, **265**: 1429–1431.
- Zakharov, V.M. 1992. Population phenogenetics: analysis of developmental stability in natural populations. *Acta Zool. Fenn.*, **191**: 7–30.