Heredity (2007) 98, 392–400

www.nature.com/hdy

ORIGINAL ARTICLE

Fluctuating asymmetry and genetic variability in the roe deer (*Capreolus capreolus*): a test of the developmental stability hypothesis in mammals using neutral molecular markers

FE Zachos[1], GB Hartl[1] and F Suchentrunk[2]
[1]*Zoological Institute, Christian-Albrechts-University, Kiel, Germany and* [2]*Research Institute of Wildlife Ecology, University of Veterinary Medicine, Vienna, Austria*

Fluctuating asymmetry (FA), used as an indicator of developmental stability, has long been hypothesized to be negatively correlated with genetic variability as a consequence of more variable organisms being better suited to buffer developmental pathways against environmental stress. However, it is still a matter of debate if this is due to metabolic properties of enzymes encoded by certain key loci or rather to overall genomic heterozygosity. Previous analyses suggest that there might be a general difference between homeo- and poikilotherms in that only the latter tend to exhibit the negative correlation predicted by theory. In the present study, we addressed these questions by analysing roe deer (*Capreolus capreolus*) from five German populations with regard to FA in metric and non-metric skull and mandible traits as well as variability at eight microsatellite loci. Genetic variability was quantified by heterozygosity and mean d^2 parameters, and although the latter did not show any relationship with FA, we found for the first time a statistically significant negative correlation of microsatellite heterozygosity and non-metric FA among populations. Because microsatellites are non-coding markers, this may be interpreted as evidence for the role of overall genomic heterozygosity in determining developmental stability. To test if the threshold character of non-metric traits is responsible for the metric vs non-metric difference we also carried out calculations where we treated our metric traits as threshold values. This, however, did not yield significant correlations between FA and genetic variability either.
Heredity (2007) **98,** 392–400; doi:10.1038/sj.hdy.6800954; published online 21 March 2007

Keywords: fluctuating asymmetry; developmental stability; microsatellites; roe deer; homeotherms; poikilotherms

Introduction

Fluctuating asymmetry (FA), defined as minor non-directional deviations from bilateral symmetry in morphological characters (Van Valen, 1962), is widely regarded as an index of developmental stability, that is, an organism's ability to precisely express genetically determined developmental pathways despite environmental disturbances. Although there is still a general controversy about the value of FA as an estimator of an organism's fitness (e.g., Rasmuson, 2002), a long-held theory, going back at least to Lerner (1954), connects developmental stability with genetic variability. More heterozygous individuals, the argument goes, more efficiently stabilize their development by buffering against environmental insults. This may be owing to the direct influence of certain key loci, to the influence of loci linked to the loci being studied or due to overall genomic heterozygosity (e.g., Mitton, 1978, 1993). Probably the best argument for the key loci hypothesis are the findings of a direct connection between null alleles at the lactate dehydrogenase loci and asymmetry in rainbow trout (Leary *et al*., 1983a, 1993; Messier and Mitton, 1996). As heterozygosity is inversely related to inbreeding, the relationship between FA and genetic variability may also be the result of different levels of inbreeding (see Clarke *et al*., 1986; Møller and Swaddle, 1997) although there are earlier studies that did not find any correlation between FA and inbreeding coefficients (e.g., Clarke *et al*., 1992; Fowler and Whitlock, 1994). To test further the role of genetic variability many studies have been carried out to detect a possible correlation between FA and genetic variability. Most of these studies measured variability as heterozygosity at allozyme loci (for a review see Møller and Swaddle, 1997). If, however, it is overall heterozygosity rather than specific key loci, which causes developmental stability, variability at neutral non-coding loci should also be correlated with the level of FA. So far, there have only been few studies based on non-coding markers such as microsatellites but first results do not corroborate the hypothesis of a role of overall genomic heterozygosity (Kruuk *et al*., 2003; Borrell *et al*., 2004).

Another important issue which is still being debated is if there is a general difference between homeotherms and poikilotherms with respect to the relationship of FA and genetic variability. It is noteworthy that most of the

Correspondence: Dr FE Zachos, Zoological Institute, Christian-Albrechts-University, Olshausenstrasse 40, 24118 Kiel, Germany.
E-mail: fzachos@zoologie.uni-kiel.de
Received 5 July 2006; revised 2 February 2007; accepted 16 February 2007; published online 21 March 2007

studies yielding a negative correlation between developmental instability (as measured by FA) and genetic variability were carried out on poikilotherms using non-metric traits (e.g., Vrijenhoek and Lerman, 1982; Leary *et al*., 1983b; but see, e.g., Hosken *et al*. (2000) and Réale and Roff (2003) for no such correlation with metric traits), whereas for homeotherms (for which FA is mostly quantified by metric traits, see Discussion section) no such relationship was found in the majority of cases (e.g., Hartl *et al*., 1991; Sert *et al*., 2005; for a review see Novak *et al*., 1993). In line with this, a meta-analysis conducted by Vøllestad *et al*. (1999) yielded the tendency of a positive association between FA and heterozygosity in homeotherms and the tendency of a negative association for poikilotherms whereas altogether there was 'only a weak association between heterozygosity and FA' (Vøllestad *et al*., 1999, page no. 215).

In the present study we analyse five German populations of European roe deer (*Capreolus capreolus*) with respect to genetic variability at selectively neutral microsatellite loci and FA in metric and non-metric skull and mandible traits. We can thus address some of the aforementioned questions on the relationship between developmental stability and genetic variability, in particular (1) if variability of neutral molecular markers also shows associations with FA (which would favour the hypothesis of the role of overall genomic heterozygosity); (2) if there is a difference between metric and non-metric traits in studies on homeotherms and (3) if the hierarchical level of analysis (individual, population) is of relevance. The latter point is important because both Karvonen *et al*. (2003) and Hartl *et al*. (1995) found a relationship on the population level only although neither study yielded a corresponding result for the level of the individual.

Materials and methods

Specimens studied

The study is based on 105 roe deer specimens from five different populations in Schleswig–Holstein, northern Germany (Figure 1). Foehr (FO, $n=18$) and Fehmarn (FM, $n=26$) are island populations, whereas Nordfriesland (NF, $n=22$), Schleswig (SL, $n=16$) and Rantzau (RA, $n=23$) are mainland stocks. The two island populations were artificially founded in the 1930s, but repeated subsequent introductions mitigated the founder effect in both populations (see Zachos *et al*., 2006 for details). Owing to overhunting the complete North-German roe deer population went through a bottleneck in the middle of the nineteenth century, but there have been numerous introductions since. Thus, Zachos *et al*. (2006) found high variability of the Schleswig-Holstein roe deer at microsatellite loci and the mitochondrial control region (although not at allozyme loci). The animals sampled were adults only with a minimum age of 2 years as determined qualitatively from dentition, skull sutures, antler structure and, in some cases, recognition of individual specimens.

Microsatellite data

Microsatellites are short tandem-repetitive DNA units with repeat lengths of 1–6 base pairs. Mutation rates are very high leading to large allele numbers with the various alleles differing in repeat unit number and hence in length. The 105 roe deer were genotyped for eight polymorphic microsatellite loci: OarFCB304, RT1, RT7, ILSTS008, ILSTS058, NVHRT16, NVHRT21 and NVHRT24. The genotypic data were the same as used by Zachos *et al.* (2006, see this article for references on the loci used).

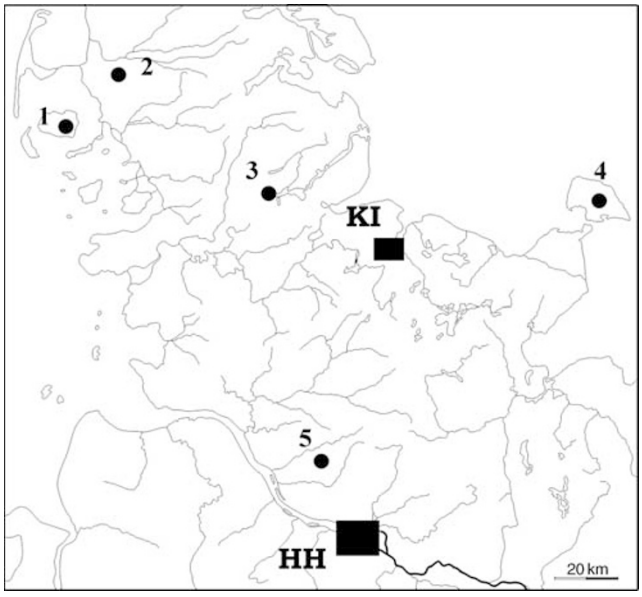

Figure 1 Map of northern Germany showing the geographical location of the populations studied (taken from Zachos *et al*., 2006). FO, Foehr; NF, Nordfriesland; SL, Schleswig; FM, Fehmarn; RA, Rantzau; KI, Kiel; HH, Hamburg.

Genetic variability of the individuals was measured as percentage of heterozygous loci (individual heterozygosity H) and as mean d^2 parameters. The latter is calculated as the sum of the squared differences of the two alleles at a locus divided by the number of loci analysed (Coulson *et al*., 1998). In addition to the common mean d^2 parameter we also used the outbreeding-mean d^2 values (mean d^2_{out}; Coulson *et al*., 1999) which are calculated as explained above except that homozygous loci are neglected. The rationale of calculating these parameters is that the length difference between the two alleles at a locus in an individual contains information on the genetic distance between the two gametes that gave rise to this individual and thus may serve as an indicator of genetic variability (Pemberton *et al*., 1999). Although mean d^2 parameters have recently been criticized and shown to have generally less power in detecting genotype-fitness correlations than heterozygosity (Tsitrone *et al*., 2001), we included them in our calculations to produce data directly comparable to previous analyses on genetic variability and FA (Borrell *et al*., 2004; Kruuk *et al*., 2003).

Population values for H and mean d^2 parameters were calculated as the mean variability values over all individuals of the population. Expected heterozygosity (H_E) was calculated as the mean over all eight loci for each population using the Arlequin software (Schneider *et al*., 2000).

In all analyses (genetic and morphological), Bonferroni procedures were used with a nominal α of 0.05 to correct the critical values for multiple tests (Rice, 1989).

Deviation from normal distribution and homogeneity of variance of H and mean d^2 were tested for each population separately by means of Kolmogorov–Smirnov and Levene tests. As none of the tests was significant a two-way ANOVA was conducted for each of the three variability parameters to test for differences between the two sexes and the five populations. Correlation between H and the two mean d^2 parameters was tested for with a Pearson test. All three tests showed positive correlations, especially the two mean d^2 parameters ($r = 0.863$, $P < 0.0005$). In order not to obtain redundant results, only H and mean d^2_{out}, whose correlation with H was much weaker than the one of mean d^2 and not significant after Bonferroni correction ($r = 0.176$, $P = 0.037$), were used in tests of correlation with FA.

FA in metric traits

For assessing FA in metric traits 15 bilateral skull and mandible measurements were taken on the left and right side of each specimen (Figure 2). Measurements were taken exclusively by one of the authors (FEZ) to avoid possible inter-observer variability (Lee, 1990). The bilateral traits were measured twice (and then averaged) with digital callipers to the nearest 0.01 mm or, if exceeding 15 cm in length, with common callipers to the nearest 0.05 mm. As right-left differences in general are often small and (just like FA!) normally distributed around a mean of 0, estimation of measurement error is indispensable (Merilä and Björklund, 1995). Its relative contribution was determined as follows (cf. Hartl *et al*., 1995). Each measurement was repeated the three times on each side in 20 individuals. On the basis of resulting data set of 120 measurements for each trait (two sides × three repeated measurements × 20 individuals) two-way ANOVAs with individual and side as fixed factors were carried out. Measurement error was considered insignificant if the sum of the variance due to side and of the variance owing to side/individual interaction was at least twice as high as the residual variance. Using this criterion, none of the FA estimates was affected by measurement error.

To test for the occurrence of directional asymmetry we used sign-tests (right vs left measure). Antisymmetry was examined using Kolmogorov–Smirnov tests of frequency distributions of right-left differences compared to an expected normal distribution (Palmer and Strobeck, 1986).

For single traits the following FA index was used:

$$|R - L|/[(R + L)/2],$$

where R and L are the measurements on the right and the left body side, respectively. The denominator corrects for trait size in order not to weight traits with respect to their absolute length. This FA index has often been used before and, thus, our results are directly comparable to a variety of previous studies (e.g., Hartl *et al*., 1995; Suchentrunk *et al*., 1998; Kruuk *et al*., 2003; Sert *et al*., 2005). Pairwise correlation of single FA indices was tested by means of Spearman rank tests.

An overall FA index that combines all metric traits of each individual (FA_M) was calculated as the arithmetic mean of all scorable FA values. Owing to damage during preparation of the skulls not all traits could be measured in all individuals. We only included individuals with at least eight scorable FA values (102 out of 105 specimens).

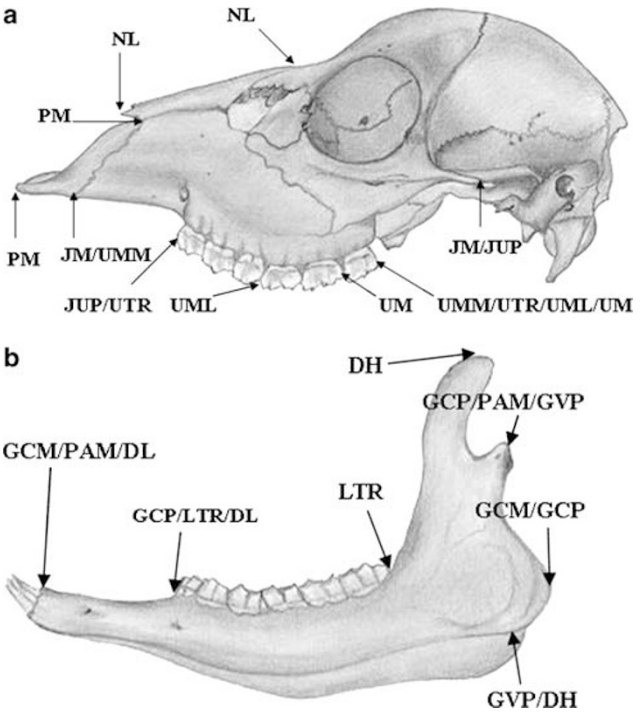

Figure 2 Bilateral metric skull (**a**) and mandible (**b**) measurements examined in the roe deer. JUP, jugal – upper second premolar; JM, jugal – tip of maxillary; UMM, upper third molar – tip of maxillary; PM, premaxillary length; NL, nasal length; UTR, upper tooth row length; UML, upper molar row length; UM, length of upper third molar; GCM, Gonion caudale – tip of mandible; GCP, Gonion caudale – lower second premolar; PAM, Processus articularis – tip of mandible; LTR, lower tooth row length; DL, diastema length; GVP, Gonion ventrale – Processus articularis; DH, dental height. The two points used for taking each measurement are indicated by arrows.

The influence of sex and population on metric overall FA was tested with a one-way ANOVA (sex) and, owing to variance heterogeneity among populations (Levene test), a Kruskal–Wallis test (population).

To test if the potential difference between metric and non-metric traits is caused by the threshold character of non-metric traits we also carried out tests where we treated our metric data as threshold traits and then conducted calculations as for the non-metric traits. We set a threshold by defining – for each single metric trait – only the 10 (in a second approach: 20) individuals with the highest FA value as asymmetric and the rest as symmetric. The reason for the arbitrarily chosen 10 or 20 specimens is that we did not want our samples to become too small, and any choice of a threshold of this kind inevitably has to be arbitrary. We then calculated an individual asymmetry index as the proportion of asymmetric traits and carried out correlation analyses as described for the non-metric traits.

FA in non-metric traits

A total of 18 non-metric skull and mandible characters were used for assessing asymmetry in non-metric traits (Table 1), many of which have been used in FA studies on roe deer before (Markowski and Markowska, 1988; Markowski, 1993). For each trait only the presence or absence of

Table 1 Non-metric bilateral skull and mandible characters used in the present study

Code	*Anatomical designation of the character*
NM1	Number of additional internal Foramina hypoglossi
NM2	Internal foramen condylare
NM3	Number of Foramina supraoccipitalia
NM4	Number of foramina on the Facies temporalis of the squamosal
NM5	Additional foramen at the foramen ovale
NM6	Additional foramen at the foramen orbitorotundum
NM7	Additional foramen at the foramen opticum
NM8	Double foramen supraorbitale
NM9	Additional foramen at the posterior palatal foramen
NM10	Additional foramen infraorbitale
NM11	foramen zygomaticum anterior
NM12	Additional foramen zygomaticum anterior
NM13	foramen infralacrimale
NM14	Upper foramen lacrimale
NM15	Additional upper foramen mentale
NM16	Additional proximal lower foramen mentale
NM17	Posterior foramen mentale
NM18	Additional posterior foramen mentale

Except for NM1, NM3 and NM4 all characters were scored as either present or absent.

symmetry (i.e., the occurrence of equal or different character states on the right and left side) was scored.

For differentiating between fluctuating and directional asymmetry we used Wilcoxon signed-rank tests. Pairwise association of single non-metric traits was tested with χ^2 tests.

As an individual index of overall non-metric fluctuating asymmetry (FA_{NM}), we chose the proportion of traits asymmetric in each individual (Suchentrunk, 1993; Hartl *et al*., 1995). Using χ^2 tests, we tested the data of the (arbitrarily chosen) right side for significant differences among the five populations studied to rule out the possibility of a purely numerical correlation of non-metric FA and character distribution, that is, different liabilities of FA of non-metric characters due to different character frequencies in the five populations (e.g., a trait with a very low frequency in one population can only have a low FA in this particular population, whereas the same character might have a higher level of FA in populations with higher frequencies of that trait). Since frequencies of occurrence did not vary significantly across populations in any of the 18 traits, all characters were used in further analyses.

The influence of sex and population on FA_{NM} was tested with a two-way ANOVA after Kolmogorov–Smirnov and Levene tests had shown their normal distribution and variance homogeneity, respectively.

To test whether metric and non-metric FA were correlated, as would be expected if these indices reflect overall developmental stability, Spearman rank tests were carried out on the individual level within and across populations and on the population level across populations. Population-specific FA_M and FA_{NM} were calculated as the arithmetic mean of all individuals from the respective population.t

FA and genetic variability

To test whether some subgroups of metric traits (e.g., mandible traits belonging to the same functional unit) are more prone to be correlated with genetic variability than others we first carried out two Spearman rank tests on each single metric trait (one for correlation with H and one for correlation with mean d^2_{out}) over all individuals.

FA_M and FA_{NM} were tested for correlations with genetic variability on three levels: (1) over all individuals across populations ($n = 102$ for metric and 105 for non-metric traits), (2) over all individuals within populations and (3) among populations using mean values calculated as explained above. For levels (1) and (2) we only used H and mean d^2_{out} as indices of genetic variability, for level (3) we also tested for a relationship between FA and expected heterozygosity (H_E). Tests on the individual level (1, 2) were carried out using Pearson correlations (normal distribution was shown by Kolmogorov–Smirnov tests) for FA_M and Spearman rank tests for FA_{NM}. Tests on the population level (3) were all conducted with Spearman rank tests.

Results

Microsatellite variability

Individual heterozygosity H of the 105 roe deer ranged from 0.125 (one out of eight loci heterozygous) to 1.0 (all loci heterozygous). Neither H nor the two mean d^2 parameters showed significant differences between the sexes or among the populations ($0.347 < P < 0.896$). Population-specific values of H, H_E and mean d^2_{out} are shown in Table 2.

FA in metric traits

Right-left differences of each measurement were normally distributed so that antisymmetry could be ruled out. The sign test for the nasal length data, however, yielded a value significantly different from zero ($P = 0.002$). Owing to this directional asymmetry the trait was not included in further analyses.

The Spearman tests showed significant correlations between three pairs of single-trait FA: JUP/JM, UTR/UML and GCI/GCP. However, since the correlation coefficients were rather small (0.494, 0.395 and 0.459, respectively), we considered the traits to yield sufficiently independent information and, thus, they were all used for further calculations. Individual overall asymmetry (FA_M) showed neither sex nor population dependence. Table 3 summarizes descriptive statistical parameters of the 14 metric traits used in this study.

FA in non-metric traits

No cases of directional asymmetry were found and the χ^2 tests did not yield significant associations between any two traits. FA_{NM} neither depended on sex nor on population. Table 4 summarizes descriptive statistical parameters of the 18 non-metric traits.

Metric and non-metric asymmetry were not correlated on any of the three levels studied (over all individuals pooled, over individuals within populations, among populations).

FA and genetic variability

None of the 14 single metric traits showed any statistically significant correlation with H or mean d^2_{out} ($P \geqslant 0.014$ with a Bonferroni-corrected significance level of 0.0036).

Table 2 Population-specific values of microsatellite variability used for the correlation analyses with FA

Population	H_E	H	*Mean* d^2_{out}
FO	0.74	0.549	23.684
NF	0.78	0.636	29.197
SL	0.76	0.547	28.783
FM	0.76	0.577	23.153
RA	0.79	0.636	26.540

H_E, expected heterozygosity, H, mean individual heterozygosity.
H_E was calculated as the arithmetic mean over all eight loci, H and mean d^2_{out} were calculated as the arithmetic mean of the individual values within each population. For population abbreviations see text.

Table 3 Asymmetry of metric traits measured in the present study

Trait	n	*(R+L)/2±s.e.*	*(R-L)±s.e.*	*p (sign)*	*FA index±s.e.*
JUP	102	93.98±0.41	−0.148±0.100	0.550	0.00851±0.00066
JM	89	128.48±0.57	−0.237±0.139	0.043	0.00718±0.00080
UMM	87	96.43±0.37	0.111±0.104	0.236	0.00658±0.00080
PM	99	47.19±0.36	0.027±0.104	0.688	0.01603±0.00160
UTR	104	58.26±0.21	0.027±0.049	0.844	0.00670±0.00053
UML	103	32.86±0.14	0.097±0.046	0.194	0.01005±0.00106
UM	100	11.24±0.06	−0.018±0.022	0.919	0.01518±0.00133
GCI	98	150.52±0.49	0.250±0.095	0.006	0.00497±0.00041
GCP	92	104.13±0.35	0.100±0.089	0.466	0.00646±0.00054
PAI	84	150.06±0.67	−0.010±0.117	0.743	0.00543±0.00051
LML	88	65.39±0.27	−0.055±0.049	0.749	0.00518±0.00051
DL	95	39.48±0.29	0.057±0.065	1.000	0.01175±0.00112
GVPA	97	57.14±0.32	−0.281±0.088	0.004	0.01063±0.00124
DH	99	84.25±0.38	−0.159±0.102	0.610	0.00862±0.00090

n = sample size, (R+L)/2 = mean of the variable length, s.e. = standard error of mean.
Mean and s.e. of right-left differences (R−L) along with the *P*-values of the sign-tests (Bonferroni-corrected significance level: 0.0033) are given to indicate the absence of directional asymmetry. (R+L)/2 and (R-L) are given in mm. The FA index given (mean over all individuals for each trait) is |R−L|/[(R+L)/2]. The nasal length is not listed due to directional asymmetry (see text).

Table 4 Asymmetry of non-metric traits measured in the present study

Code	*n*	*Symmetrical*	*Asymmetrical*	P *(Wilcoxon)*
NM1	104	49 (47.1%)	55 (52.9%)	0.371
NM2	105	93 (88.6%)	12 (11.4%)	1.000
NM3	104	65 (62.5%)	39 (37.5%)	0.084
NM4	103	34 (33.0%)	69 (67.0%)	0.213
NM5	105	82 (78.1%)	23 (21.9%)	0.835
NM6	103	79 (76.7%)	24 (23.3%)	0.041
NM7	104	75 (72.1%)	29 (27.9%)	0.853
NM8	104	70 (67.3%)	34 (32.7%)	0.303
NM9	103	72 (69.9%)	31 (30.1%)	0.369
NM10	105	84 (80.0%)	21 (20.0%)	0.827
NM11	105	100 (95.2%)	5 (4.8%)	0.655
NM12	105	77 (73.3%)	28 (26.7%)	0.450
NM13	105	88 (83.8%)	17 (16.2%)	0.029
NM14	105	74 (70.5%)	31 (29.5%)	0.857
NM15	105	78 (74.3%)	27 (25.7%)	0.336
NM16	105	82 (78.1%)	23 (21.9%)	0.297
NM17	105	91 (86.7%)	14 (13.3%)	1.000
NM18	105	97 (92.4%)	8 (7.6%)	0.157

n = sample size, *P* (Wilcoxon) = *P*-values of the Wilcoxon tests for directional asymmetry (Bonferroni-corrected significance level: 0.0028).

The results of the correlation analyses for FA_M and FA_{NM} on all three levels are shown in Tables 5 and 6. For the metric traits, there was merely a nonsignificant (after Bonferroni correction) tendency for FA_M to be weakly negatively correlated with H ($r = -0.236$, $P = 0.008$) across all individuals but for the non-metric traits we found a strongly negative and significant correlation ($r = -0.975$, $P = 0.002$) between non-metric FA and average individual heterozygosity among populations. All other correlations, including those involving H_E, yielded nonsignificant results. This also held for the tests where we treated our metric traits as threshold characters (not shown in further detail).

Discussion

It is generally assumed that functionally important traits are subject to a stronger pressure of balancing selection,

Table 5 Pearson (r_p) and Spearman (r_s) correlations between asymmetry and genetic variability across all individuals and among populations

FA index	*H*	*Mean* d^2_{out}	H_E
Individuals			
FA_M	$r_p = -0.236$, $P = 0.008$	$r_p = 0.052$, $P = 0.303$	—
FA_{NM}	$r_s = 0.064$, $P = 0.258$	$r_s = 0.007$, $P = 0.470$	—
Populations			
FA_M	$r_s = 0.205$, $P = 0.370$	$r_s = -0.400$, $P = 0.252$	$r_s = 0.103$, $P = 0.435$
FA_{NM}	$r_s = -0.975$, $P = 0.002$	$r_s = -0.300$, $P = 0.312$	$r_s = -0.564$, $P = 0.161$

Expected heterozygosity (H_E) was only calculated on the population level. FA for populations was calculated as the mean over all individuals.

Table 6 Pearson (metric FA) and Spearman (non-metric FA) correlations between asymmetry and genetic variability across individuals within populations

Population	*Metric FA/H*	*Metric FA/mean* d^2_{out}	*Non-metric FA/H*	*Non-metric FA/mean* d^2_{out}
FO	−0.286 (0.125)	0.227 (0.183)	0.008 (0.488)	0.042 (0.434)
NF	−0.152 (0.255)	−0.067 (0.387)	−0.138 (0.270)	−0.082 (0.358)
SL	−0.466 (0.035)	0.308 (0.123)	0.061 (0.411)	−0.228 (0.198)
FM	−0.239 (0.131)	0.008 (0.485)	0.282 (0.082)	−0.034 (0.435)
RA	−0.213 (0.164)	0.282 (0.096)	0.206 (0.173)	0.210 (0.168)

Correlation coefficients are given first, *P*-values are in parentheses (none of these is significant after Bonferroni correction).

which, in the case of bilaterally symmetric traits, is supposed to result in lower levels of asymmetry (e.g., Stearns, 1992). The characters in our study did not yield any conspicuous FA patterns; in particular, tooth characters did not consistently show low FA as found in other studies (e.g., Suchentrunk, 1993; Suchentrunk *et al*., 1994). The lack of correlation between single-trait FA demonstrates that FA in a given character is hardly predictive for FA in other characters and thus does not corroborate the hypothesis of the existence of population or individual asymmetry parameters (PAP, IAP, Soulé, 1967; Clarke, 1998), that is, a consistent ranking of individuals or populations for the asymmetry of several characters. The lack of a consistent correlation between metric and non-metric FA is in line with the results of Zima *et al.* (1989) who did not find a 'clear-cut correlation' (p. 54) between metric and non-metric distances between roe deer populations either.

Compared with the study by Markowski (1993), who analysed population differentiation in roe deer based on non-metric traits many of which are identical to traits used in the present study, our FA_{NM} values are considerably higher (0.12–0.18 vs 0.23–0.29). Apart from possible biological reasons (FA might really be higher in the roe deer of the present study) this discrepancy may also be owing to inter-observer variability (Lee, 1990). The claim that non-metric traits are generally superior to metric ones in population studies on mammals owing to their not being correlated with one another and not being sex-dependent (e.g., Berry, 1968; see also Markowski, 1995) has not been corroborated by the results of the present study in that there were no differences between metric and non-metric traits with respect to correlation and sex dependence.

It has long been debated whether FA levels depend upon genetic variability with numerous studies favouring such a relationship and equally numerous studies rejecting it (see Møller and Swaddle, 1997 for a review). The present study did not yield unequivocal results either in that we found a tendency for a (weakly) negative correlation of metric FA and heterozygosity over all individuals (the same nonsignificant trend was found by Alves *et al*., 2001) and a statistically significant strongly negative correlation of non-metric FA and heterozygosity among populations but no further significant relationships. The fact that we did not find any such trend or correlation for genetic variability quantified by mean d^2_{out} parameters is in line with the criticism put forward against these measures (see above). Nevertheless, our study is, to our knowledge, the first one to report a microsatellite-based negative correlation between FA and genetic variability. This is important because it is still an open question if the relationship between developmental stability and genetic variability (if there is any at all) is due to the physiological properties of specific enzymes encoded by certain key loci or owing to overall genomic heterozygosity. Our results based on microsatellites, which are non-coding DNA regions commonly regarded as selectively neutral, favour the latter alternative (against Kruuk *et al*., 2003 and Borrell *et al*., 2004) although linkage with key loci can never totally be ruled out. It has to be borne in mind, however, that all genetic parameters calculated in our study are based on eight loci only, which is probably not enough to be representative of genome-wide heterozygosity (see Chakraborty, 1981; Slate and Pemberton, 2002).

Results showing a negative correlation of FA and genetic variability, as expected *a priori* from the theory, are scarce for homeothermic species. Studies on mammals, according to our knowledge, only yielded the expected negative correlation twice (Hutchison and Cheverud, 1995 on tamarin monkeys and Hartl *et al*., 1995 on brown hares). The study by Hutchison and Cheverud, however, is not directly comparable to similar studies because (1) their analysis comprises two different species and (2) they do not directly measure heterozygosity but rather provide qualitative ranks of genetic variability. Hartl *et al*.(1995) found a negative correlation between non-metric FA and enzyme heterozygosity in the brown hare (*Lepus europaeus*) on the level of the

population but no such relationship for the individuals or metric FA values. They suggest that it might be the choice of characters in FA calculation rather than a principal difference between homeotherms and poikilotherms that accounts for the non-concordant results in these two groups because in studies on poikilotherms measurements of FA are usually based on non-metric traits such as bristle number in insects or fin ray number in fish species, whereas in studies on homeotherms FA is mostly determined through measurement of metric skeletal traits.

Interestingly, the present study yielded results very similar to those arrived at by Hartl *et al.* (1995) in that the only statistically significant correlation between FA and heterozygosity was found among populations and for non-metric traits. Why non-metric traits might be better suited for FA analyses remains unclear but they are, contrary to metric traits, threshold characters: as they do not vary continuously, phenotypic asymmetry (presence vs absence) only occurs when a certain threshold difference between the right and left side is achieved (cf Swain, 1987). If the relationship of FA and genetic variability only holds for very asymmetric individuals asymmetry in non-metric traits may a *priori* be more likely to reveal this relationship. Furthermore, metric traits might be under stronger balancing selection for symmetry since many traits typically measured in FA studies are functionally important (e.g., tooth rows or mandible length). Our approach of treating metric traits as threshold characters, however, did not yield the same results as the non-metric traits. This certainly does not completely refute the threshold explanation of the difference between metric and non-metric traits, but at least we did not find a corroboration for it either. At any rate, the fact that the two negative correlations of FA and genetic variability found in mammals so far were both based on non-metric traits is remarkable given that most studies on mammals were conducted with metric traits. The often claimed difference between homeotherms and poikilotherms (see also the aforementioned meta-analysis by Vøllestad *et al*., 1999) may thus be merely a methodological artefact rather than a biological phenomenon caused by the more homogeneous developmental environment owing to a more constant body temperature in homeotherms. If the constant developmental environment in homeotherms is crucial to FA studies, that is, if developmental noise is relatively more important in homeotherms than it is in poikilotherms because the effects of genetic variability are veiled by the generally lower levels of FA in the former, then mammals might turn out to be particularly well suited to test this hypothesis. The three mammalian subgroups (monotremes, marsupials and placental mammals) differ considerably in their early development with monotremes being oviparous, marsupials having a very short gestation period and giving birth to little-developed young and placental mammals undergoing a long intrauterine phase of development. There is, in other words, an increase in 'developmental homeothermy' from monotremes to placental mammals, and thus, if the hypothesis of a difference between homeo- and poikilotherms holds, we expect a comparative decrease in FA levels from monotremes over marsupials to placental mammals. To date, no such comparative analysis has been conducted.

A further important issue of the present study is the fact that the only clear-cut correlation between FA and genetic variability was found not on the level of the individual but among the populations – another parallel to the results of the study on brown hares by Hartl *et al.* (1995). Similarly, Karvonen *et al.* (2003) found this correlation in greenfinches only when comparing populations – individual heterozygosity (based on allozyme loci) was not related to individual FA. One possible reason for this could be that a limited number of loci used in variability assessments is more likely to uncover differences in general heterozygosity among individuals from different populations than among individuals taken from the same population (Palmer and Strobeck, 1986). Additionally, the fact that measuring FA on the individual level is an attempt to estimate a variance with only two data points, which is likely to result in large sampling errors (Whitlock, 1996), may also be relevant.

Conclusion

Our study is one of only few to find a negative relationship between FA and genetic variability in homeotherms in general and mammals in particular, although this relationship was only shown for the population level. Ever since the debate about possibly raised asymmetry in the cheetah (Wayne *et al*., 1986; Modi *et al*., 1987; Willig and Owen, 1987; Kieser and Groeneveld, 1991), a species known to be genetically depauperate, this relationship has been uncertain for mammals. The results of the present study, however, suggest, in line with one earlier study (Hartl *et al*., 1995) that the lack of confirmation may be caused by methodological rather than biological factors. Future studies should be carried out bearing in mind that the choice of traits (metric vs non-metric) might bias the results of correlation analyses between FA and genetic variability. The hierarchical level of comparison (individual vs population) may also be important given that in most analyses the number of molecular markers studied is around 10. The question of the genetic factors governing developmental stability is still far from being settled. Although first microsatellite analyses did not produce a negative correlation between FA and variability (Borrell *et al*., 2004), our results show that further research into the relative role of key loci and overall genomic heterozygosity is required.

Acknowledgements

We express our gratitude to all people involved in the collection of samples, in particular the Landesjagdverband Schleswig-Holstein, to K Steffen for the drawings of the roe deer skull and mandible and to H Luttmann and A Oesert for the preparation of the skulls. Two anonymous referees made valuable comments on an earlier version of this manuscript. FE Zachos received financial support from the German National Academic Foundation for the present study.

References

Alves PC, Ferrand N, Suchentrunk F (2001). Developmental stability and protein heterozygosity in a local population of Iberian hares (*Lepus granatensis*). *Mamm Biol* **66**: 238–250.

Berry RJ (1968). The biology of non-metrical variation in mice and men. In: Brothwell DR (ed). *The Skeletal Biology of Earlier Human Populations*. Pergamon Press: London. pp 103–133.
Borrell YJ, Pineda H, McCarthy I, Vázquez E, Sánchez JA, Lizana GB (2004). Correlations between fitness and heterozygosity at allozyme and microsatellite loci in the Atlantic salmon, *Salmo salar* L. *Heredity* **92**: 585–593.
Chakraborty R (1981). The distribution of the number of heterozygous loci in an individual in natural populations. *Genetics* **98**: 461–466.
Clarke GM (1998). The genetic basis of developmental stability IV. Individual and population asymmetry parameters. *Heredity* **80**: 553–561.
Clarke GM, Brand GW, Whitten MJ (1986). Fluctuating asymmetry: a technique for measuring developmental stress caused by inbreeding. *J Aust Biol Sci* **39**: 145–153.
Clarke GM, Oldroyd BP, Hunt P (1992). The genetic basis of developmental stability in *Apis mellifera*: heterozygosity versus genic balance. *Evolution* **46**: 753–762.
Coulson T, Albon S, Slate J, Pemberton J (1999). Microsatellite loci reveal sex-dependent responses to inbreeding and outbreeding in red deer calves. *Evolution* **53**: 1951–1960.
Coulson TN, Pemberton JM, Albon SD, Beaumont M, Marshall TC, Slate J *et al* (1998). Microsatellites measure inbreeding depression and heterosis in red deer. *Proc R Soc Ser B* **265**: 489–495.
Fowler K, Whitlock MC (1994). Fluctuating asymmetry does not increase with moderate inbreeding in *Drosophila melanogaster*. *Heredity* **73**: 373–376.
Hartl GB, Lang G, Klein F, Willing R (1991). Relationships between allozymes, heterozygosity and morphological characters in red deer (*Cervus elaphus*), and the influence of selective hunting on allele frequency distributions. *Heredity* **66**: 343–350.
Hartl GB, Suchentrunk F, Willing R, Petznek R (1995). Allozyme heterozygosity and fluctuating asymmetry in the brown hare (*Lepus europaeus*): a test of the developmental homeostasis hypothesis. *Philos T Roy Soc B* **350**: 313–323.
Hosken DJ, Blanckenhorn WU, Ward PI (2000). Developmental stability in yellow dung flies (*Scathophaga stercoraria*): fluctuating asymmetry, heterozygosity and environmental stress. *J Evol Biol* **13**: 919–926.
Hutchison DW, Cheverud JM (1995). Fluctuating asymmetry in tamarin (*Saguinus*) cranial morphology: intra- and interspecific comparisons between taxa with varying levels of genetic heterozygosity. *J Hered* **86**: 280–288.
Karvonen E, Merilä J, Rintamäki PT, van Dongen S (2003). Geography of fluctuating asymmetry in the greenfinch, *Carduelis chloris*. *Oikos* **100**: 507–516.
Kieser JA, Groeneveld HT (1991). Fluctuating odontometric asymmetry, morphological variability, and genetic monomorphism in the cheetah *Acinonyx jubatus*. *Evolution* **45**: 1175–1183.
Kruuk LEB, Slate J, Pemberton JM, Clutton-Brock TH (2003). Fluctuating asymmetry in a secondary sexual trait: no associations with individual fitness, environmental stress or inbreeding, and no heritability. *J Evol Biol* **16**: 101–113.
Leary RF, Allendorf FW, Knudsen KL (1983b). Developmental stability and enzyme heterozygosity in rainbow trout. *Nature* **301**: 71–72.
Leary RF, Allendorf FW, Knudsen KL (1993). Null alleles at two lactate dehydrogenase loci in rainbow trout are associated with decreased developmental stability. *Genetica* **89**: 3–13.
Leary RF, Knudsen KL, Allendorf FW (1983a). Developmental instability of heterozygotes for a null allele at an LDH locus in rainbow trout. *Isozyme Bulletin* **16**: 76.
Lee JC (1990). Sources of extraneous variation in the study of meristic characters: the effect of size and of inter-observer variability. *Syst Zool* **39**: 31–39.
Lerner IM (1954). *Genetic Homeostasis*. Oliver and Boyd: Edinburgh.
Markowski J (1993). Fluctuating asymmetry as an indicator for differentiation among roe deer Capreolus capreolus populations. *Acta Theriol* **38** (Suppl 2): 19–31.
Markowski J (1995). Non-metric traits: remarks on sex dependence, age dependence, and on intercorrelations among characters. *Acta Theriol* (Suppl 3): 65–74.
Markowski J, Markowska M (1988). Non-metrical Variation in Three Populations of Roe Deer. *Acta Theriol* **33**: 519–536.
Merilä J, Björklund M (1995). Fluctuating Asymmetry and Measurement Error. *Syst Biol* **44**: 97–101.
Messier S, Mitton JB (1996). Heterozygosity at the malate dehydrogenase locus and developmental homeostasis in *Apis mellifera*. *Heredity* **76**: 616–622.
Mitton JB (1978). Relationship between heterozygosity for enzyme loci and variation of morphological characters in natural populations. *Nature* **273**: 661–662.
Mitton JB (1993). Enzyme heterozygosity, metabolism, and developmental stability. *Genetica* **89**: 47–65.
Modi WS, Wayne RK, O'Brien SJ (1987). Analysis of fluctuating asymmetry in cheetahs. *Evolution* **41**: 227–228.
Møller AP, Swaddle JP (1997). *Asymmetry, Developmental Stability, and Evolution*. Oxford University Press: Oxford, New York.
Novak JM, Rhodes Jr OE, Smith MH, Chesser RK (1993). Morphological asymmetry in mammals: genetics and homeostasis re-considered. *Acta Theriol* **38** (Suppl 2): 7–18.
Palmer AR, Strobeck C (1986). Fluctuating asymmetry: measurement, analysis, patterns. *Annu Rev Ecol Syst* **17**: 391–421.
Pemberton JM, Coltman DW, Coulson TN, Slate J (1999). Using microsatellites to measure the fitness consequences of inbreeding and outbreeding. In: Goldstein DB, Schlötterer C (eds). *Microsatellites. Evolution and Applications*. Oxford University Press: Oxford, New York. pp 151–164.
Rasmuson M (2002). Fluctuating asymmetry – indicator of what? *Hereditas* **136**: 177–183.
Réale D, Roff DA (2003). Inbreeding, developmental stability, and canalization in the sand cricket *Gryllus firmus*. *Evolution* **57**: 597–605.
Rice WR (1989). Analyzing tables of statistical tests. *Evolution* **43**: 223–225.
Schneider S, Roessli D, Excoffier L (2000). *Arlequin ver 2.000: A Software for Population Genetics Data Analysis*. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
Sert H, Suchentrunk F, Ludescher B, Hackländer K (2005). Developmental stability and canalization of limb bones of brown hares *Lepus europaeus* with varying levels of heterozygosity. *Acta Theriol* **50**: 213–226.
Slate J, Pemberton JM (2002). Comparing molecular measures for detecting inbreeding depression. *J Evol Biol* **15**: 20–31.
Soulé M (1967). Phenetics of natural populations. II. Asymmetry and evolution in a lizard. *Am Nat* **101**: 141–160.
Stearns SC (1992). *The Evolution of Life Histories*. Oxford University Press: Oxford, New York.
Suchentrunk F (1993). Variability of minor tooth traits and allozymic diversity in brown hare *Lepus europaeus* populations. *Acta Theriol* **38** (Suppl 2): 59–69.
Suchentrunk F, Hartl GB, Flux JEC, Parkes J, Haiden A, Tapper S (1998). Allozyme heterozygosity and fluctuating asymmetry in brown hares *Lepus europaeus* introduced to New Zealand: Developmental homeostasis in populations with a bottleneck history. *Acta Theriol* (Suppl 5): 35–52.
Suchentrunk F, Willing R, Hartl GB (1994). Non-metrical polymorphism of the first lower premolar (P_3) in Austrian brown hares (*Lepus europaeus*): a study on regional differentiation. *J Zool* **232**: 79–91.
Swain DP (1987). A problem with the use of meristic characters to estimate developmental stability. *Am Nat* **129**: 761–768.
Tsitrone A, Rousset F, David P (2001). Heterosis, marker mutational processes and population inbreeding history. *Genetics* **159**: 1845–1859.
Van Valen L (1962). A study of fluctuating asymmetry. *Evolution* **16**: 125–142.

Vøllestad LA, Hindar K, Møller AP (1999). A meta-analysis of fluctuating asymmetry in relation to heterozygosity. *Heredity* **83**: 206–218.

Vrijenhoek RCm, Lerman S (1982). Heterozygosity and developmental stability under sexual and asexual breeding systems. *Evolution* **36**: 768–776.

Wayne RK, Modi WS, O'Brien SJ (1986). Morphological variability and asymmetry in the cheetah (*Acinonyx jubatus*), a genetically uniform species. *Evolution* **40**: 78–85.

Willig MR, Owen RD (1987). Fluctuating asymmetry in the cheetah: methodological and interpretive concerns. *Evolution* **41**: 225–227.

Whitlock M (1996). The heritability of fluctuating asymmetry and the genetic control of developmental stability. *Proc R Soc Lond B* **263**: 849–854.

Zachos FE, Hmwe SS, Hartl GB (2006). Biochemical and DNA markers yield strikingly different results regarding variability and differentiation of roe deer (*Capreolus capreolus*, Artiodactyla: Cervidae) populations from northern Germany. *J Zool Syst Evol Res* **44**: 167–174.

Zima J, Libosvarsky J, Bauerova Z, Koubek P, Zejda J (1989). Comparison of metric and non-metric morphological distances between four populations of roe-deer (*Capreolus capreolus*). *Folia Zool* **38**: 45–58.