Biol. Rev. (2002), **77**, pp. 27–38 © *Cambridge Philosophical Society*
DOI: 10.1017/S1464793101005796 *Printed in the United Kingdom*

Fluctuating asymmetry as an indicator of fitness: can we bridge the gap between studies?

LUC LENS[1,2]*, STEFAN VAN DONGEN[1], SALIT KARK[3] and ERIK MATTHYSEN[1]

[1] *Laboratory of Animal Ecology, Dept. of Biology, University of Antwerp (UIA), Universiteitsplein* 1, *B*-2610 *Wilrijk, Belgium;*
[2] *Dept. of Ornithology, National Museums of Kenya, PO Box* 40658, *Nairobi, Kenya;*
[3] *Dept. of Biological Sciences, Stanford University, Stanford* 94305-5020, *CA, USA*

(*Received* 19 *February* 2001; *revised* 30 *July* 2001; *accepted* 2 *August* 2001)

ABSTRACT

There is growing evidence from both experimental and non-experimental studies that fluctuating asymmetry does not consistently index stress or fitness. The widely held – yet poorly substantiated – belief that fluctuating asymmetry can act as a universal measure of developmental stability and predictor of stress-mediated changes in fitness, therefore staggers. Yet attempts to understand why the reported relationships between fluctuating asymmetry, stress and fitness are so heterogeneous – i.e. whether the associations are truly weak or non-existent or whether they become confounded during different stages of the analytical pathways – remain surprisingly scarce. Hence, we attempt to disentangle these causes, by reviewing the various statistical and conceptual factors that are suspected to confound potential relationships between fluctuating asymmetry, stress and fitness. Two main categories of factors are discerned: those associated with the estimation of developmental stability through fluctuating asymmetry, and those associated with the effects of genotype and environment on developmental stability. Next, we describe a series of statistical tools that have recently been developed to help reduce this noise. We argue that the current lack of a theoretical framework that predicts if and when relationships with developmental stability can be expected, urges for further theoretical and empirical research, such as on the genetic architecture of developmental stability in stressed populations. If the underlying developmental mechanisms are better understood, statistical patterns of asymmetry variation may become a biologically meaningful tool.

Key words: fluctuating asymmetry (FA), developmental stability, stress, inbreeding, fitness.

CONTENTS

* Corresponding author: Luc Lens, University of Antwerp, Dept. of Biology, Laboratory of Animal Ecology, Universiteitsplein 1, B-2610 Wilrijk, Belgium. Tel: +32 3 820 22 61; Fax: +32 3 820 22 71.
E-mail: llens@uia.ua.ac.be

I. INTRODUCTION

The left and right sides of bilaterally symmetrical traits often show minor differences in their development. Because corresponding body sides presumably share the same genome and (in a homogeneous environment) experience similar external effects, differences in their development cannot be explained by genetic or environmental effects (Reeve, 1960). Rather, the observed asymmetries are believed to reflect the inability of individuals to buffer their development against small, random perturbations of cellular processes ('developmental noise'; Palmer, 1994) and, hence, accurately to develop their expected phenotype given their genotype and the environment (Palmer & Strobeck, 1986). Indeed, even if an organism possesses the most ideally balanced and structured genome for the environment it occupies and that environment is constant during development, there will still be some aspect of randomness in its development (Møller & Swaddle, 1997). The exact processes underlying developmental noise are poorly understood. However, suggested directions refer to perturbation at the molecular level or random variation in rates of physiological processes among cells, both which may affect cell-cell communication and rates of cell growth, division, or elongation (Palmer, 1994; McAdams & Arkin, 1999; Fiering, Whitelaw & Martin, 2000). Since perturbations resulting from stochastic, cellular processes mostly act locally (McAdams & Arkin, 1999), their effects will accumulate on the left and right sides of developing individuals separately. Consequently, they may give rise to left-right asymmetries in development.

Based on the fact that the magnitude of left-right asymmetries tends to decrease as bilateral traits develop (e.g. Swaddle & Witter, 1997), however, it is assumed that homeostatic mechanisms have evolved that buffer the disruptive effects of developmental noise during ontogeny ('developmental stability'; Palmer, 1994). At least six non-mutually exclusive hypotheses have been proposed to describe the development of bilateral asymmetry, some of which have been confirmed with empirical data (Swaddle & Witter, 1997; Tomkins, 1999; but see Aparicio, 1998). Yet, irrespective of the underlying mechanisms involved, non-directional differences in bilateral development can be considered as the outcome of two opposing forces: those tending to disrupt precise development (developmental noise), and those tending to stabilise it (developmental stability) (Van Valen, 1962; Palmer, 1994; Klingenberg & Nijhout, 1999).

Individual- and population levels of bilateral asymmetry have been shown to relate positively to a wide range of abiotic, biotic, and genetic stresses, although the strength of the association varies considerably between taxa, traits, and/or types of stress (e.g. Leary & Allendorf 1989; Kieser & Groeneveld, 1991; Palmer & Strobeck, 1992; Parsons, 1990, 1992; Clarke, 1993*a*, 1995*a*; Polak & Trivers, 1994; Lens *et al.*, 1999, 2000). Environmental or genetic stresses may affect bilateral development in two ways: by increasing the amount of developmental noise (i.e. increased incidence of random perturbations that tend to disrupt the symmetric development of left-right sides), and/or by decreasing the level of developmental stability (i.e. reduced buffering capacity or robustness of the developmental system) (Klingenberg & Nijhout, 1999; Klingenberg, 2001). As both the origin of developmental noise and the basis of developmental stability remain obscure (Palmer, 1996), it is presently unclear whether the same or different types of stress affect either pathway (but see Rutherford & Lindquist, 1998 for an example of extrinsic stress presumably affecting developmental stability rather than noise). Since developmental stability and noise cannot be observed independently but jointly result in a given level of asymmetry, their variation components cannot easily be separated either (see Section III).

The degree of developmental stability of individuals and populations is most often estimated by their level of fluctuating asymmetry (FA; Ludwig, 1932). FA refers to a pattern of bilateral variation in a sample of individuals, where the mean of right minus left values of a trait is zero and the variation is normally distributed about that mean (Palmer,

1994). The appeal of FA to studies of the relationships between developmental stability and stress originates from conceptual as well as methodological considerations. First, FA is one of the few morphological attributes for which the norm, i.e. perfect symmetry, is known (Palmer, 1996). Second, because increased energy expenditure in stressful conditions compromises the maintenance of developmental stability (Koehn & Bayne, 1989; Sommer, 1996), FA is believed to be a more sensitive stress estimator than the traditional use of fitness measures (e.g. survival; Clarke & McKenzie, 1992). This potential to predict future, stress-mediated changes in fitness has promoted the use of FA as a conservation tool (Clarke, 1995*a*; see also Cairns, McCornick & Niederlehner, 1993). Third, although the estimation of variability – such as population FA – is known to be subject to large sampling variation (e.g. Palmer, 1996; Whitlock, 1996), adequate sample sizes for the study of asymmetry can fairly easily be obtained (see Van Dongen, 1999 for a study of the effect of sample size on the accuracy and power in FA studies). This is particularly true in comparison with studies on survival or reproduction, which often require (multiple) recaptures or the fulfilment of stringent model assumptions. Fourth, measurements of FA generally do not require expensive equipment, and can be conducted non-destructively. As a result, FA has become a popular tool to estimate the quality and health of individuals and populations (reviews by Markow, 1995; Leung & Forbes, 1996; Møller, 1997; Clarke, 1998*b*; Møller & Thornhill, 1998).

Yet, besides staunch support, the use of FA as an indicator of stress and fitness has also generated strong scepticism, especially in the literature dealing with sexual trait selection (e.g. Palmer, 2000). In a recent review of 21 experimental papers examining the relationship between FA and environmental stress, Bjorksten, David, Pomiankowski & Fowler (2000*b*) found that a third of the studies showed a consistent increase in FA with stress, one-third showed a trait- or stress-specific increase, and a further third discovered no effect. Likewise, various other studies reported inconsistencies in the relationships of FA with inbreeding (Leary, Allendorf & Knudsen, 1984; Clarke, 1993*b*; Fowler & Whitlock, 1994; Völlestad, Hindar & Møller, 1999) and with various components of fitness (Clarke, 1995*a*, *b*, 1998*a*; Markow, 1995; Leung & Forbes, 1996, 1997; Dufour & Weatherhead, 1998).

Despite this vivid debate between advocates and critics of the use of FA as a bio-indicator, attempts to understand the factors that may cause the observed discordance in reported relationships remain surprisingly scarce. On the one side, it has been suggested that particular groups of traits and/or organisms might provide stronger FA-stress or FA-fitness relationships than others, e.g. sexually selected versus non-sexually selected traits (Møller, 1993; Manning & Chamberlain, 1994; Hill 1995), performance versus non-performance traits (Eggert & Sakaluk, 1994; Palmer, 1994), or in poikilotherms versus homeotherms (Wooten & Smith, 1986, Novak *et al.*, 1993). Yet, neither trait type nor thermy proved predictive of the presence or strength of these relationships in a subsequent meta-analysis (Leung & Forbes, 1996; but see Møller & Thornhill, 1998). On the other side, it has been argued that heterogeneity in the reported relationships might be due to statistical problems inherent to the estimation of developmental stability. For instance, measurement error is known to weaken the relationships with FA, by biasing FA estimation and reducing statistical power (i.e. reducing effective sample sizes) (Palmer, 1994). As sample sizes are combined across studies in meta-analyses, this might explain why the latter continue to yield significant FA-stress and FA-fitness relations (e.g. Leung & Forbes, 1996), even though the primary literature appears equivocal. Besides, low repeatability of individual FA estimates causes a downward bias in between-trait correlations in FA (see section II below), and hence, may cause underestimation of the 'true' between-trait correlations in developmental stability (Whitlock, 1996; Van Dongen, 1998*a*). After statistically correcting for the presumed weak correlation between FA and developmental stability, Van Dongen & Lens (2000) showed a substantial increase in the strength of between-trait correlations in stability as compared to the original studies using non-transformed data.

Hence, while relationships between FA, developmental stability, stress, and fitness are generally not believed to be spurious (Leung & Forbes, 1996; but see Bjorksten *et al.*, 2000*a*, *b* for a different opinion), FA does not consistently index stress or fitness, either because relationships are truly weak or non-existent, or because they become confounded during different stages of the analytical pathways. This paper aims to disentangle some of these causal factors, (i) by reviewing statistical and conceptual factors that are currently known, or suspected, to confound potential relationships between FA, stress and fitness, and (ii) by describing a suite of newly developed statistical tools that have proved to enhance the accuracy of developmental stability estimation.

II. FACTORS THAT POTENTIALLY BIAS THE ESTIMATION OF INDIVIDUAL DEVELOPMENTAL STABILITY

(1) Heterogeneity in the relationship between FA and developmental stability

Because each trait develops only once under exactly the same environmental conditions, a sample of two data points (i.e. left- and right-side measures for the computation of FA) is used to estimate a variance (developmental stability). Given that, in general, variances are notoriously difficult to estimate (Palmer 1994; Van Dongen, 1999), variance components estimated from just two data points can be expected to be unreliable. This is reflected by the low, overall repeatability of developmental stability estimates obtained through the study of FA (Whitlock, 1996; Houle, 1997). For instance, based on the kurtosis of distributions of the signed FA, Gangestad & Thornhill (1999) estimated that, on average, a single trait's FA typically owes only approximately 7% of its variance to underlying individual differences in developmental stability. Consequently, due to sampling variability, individuals with the same underlying level of stability may express different levels of FA (Whitlock, 1998).

Apart from sampling error, the random nature of the developmental processes that generate FA may further confound the estimation of developmental stability by FA, i.e. developmentally unstable individuals may still display low FA if they are 'lucky' enough to experience low levels of developmental noise (Table 1; see also Palmer, 1994; Leung & Forbes, 1997). Overall, the relationship between FA and the presumed underlying stability can be expected to be more reliable under high than under low levels of environmental stress. This is because under stressful conditions, 'low-quality' individuals face increasing difficulties in allocating sufficient energy to maintain high levels of developmental stability, hence become 'unmasked' by their higher levels of FA (see below for a discussion of stress effects on the expression of other types of bilateral asymmetry). Differences in FA between individuals might thus reflect differences in the extent or outcome of developmental noise, rather than stability (Table 1).

(2) Bias in the estimation of individual FA

Levels of FA are usually very subtle, typically in the order of 1% or less of the size of the trait (e.g. Møller & Swaddle, 1997). Because the degree of asymmetry is so small, and some traits cannot be measured with high accuracy, measurement error can be expected to account for a large fraction of the between-sides variance. In particular, measurement error has been shown to cause an upward bias in the variance if not appropriately corrected for, both at the individual- and the population level (Palmer & Strobeck, 1986; Merilä & Björklund, 1995; Van Dongen, 1999, 2000). Consequently, because the ratio of signal (FA) to noise (measurement error) is low, the power of detecting relationships with FA is low as well.

Observed variance in FA among individuals may be due to at least four sources: heterogeneity in the outcome of developmental noise; variance in expression of FA due to the sampling of two sides; variance in measurement error related to the (repeated) measurement of both sides; and between-individual variability in the degree to which their development is buffered against random pertur-

Table 1. *Hypothetical source of heterogeneity in the relationship between fluctuating asymmetry (FA) and developmental stability, in four model individuals.*

[For FA to be a reliable estimator of developmental stability, low levels of stability should be reflected by high levels of FA, and *vice versa*. However, due to the random nature of developmental noise (column 2), developmental unstable individuals may display low FA due to chance (i.e. compare individuals B and D in columns 3 and 4). Consequently, between-individual differences in FA may partly reflect heterogeneity in developmental noise, rather than in developmental stability.]

Individual	Developmental noise	Developmental stability	Observed FA	FA as reliable estimator of developmental stability
A	high	high	low	yes
B	high	low	high[a]	yes
C	low	high	low	yes
D	low	low	low	no

[a] Average expectation, since FA could still be low because of sampling variation (Whitlock, 1996; Houle, 1997).

bations (i.e. true variance in developmental stability).

(3) Admixture of FA with other types of asymmetry

There is growing evidence from theoretical (Graham, Freeman & Emlen, 1993*a*), experimental (McKenzie & Clarke, 1988; Graham, Roe & West, 1993*b*; Henshel *et al*., 1993; Leamy, Doster & Huet-Hudson, 1999) and non-experimental studies (Kark, 1999; Lens & Van Dongen, 2000; Kark *et al*., 2001) that stress may cause transitions from FA to admixtures with other types of bilateral asymmetry, more particularly directional asymmetry (normal distribution with non-zero mean) or anti-symmetry (bimodal distribution with zero mean) (Palmer & Strobeck, 1992). Unlike FA, these asymmetry types are considered inappropriate for the estimation of developmental stability, due to their presumed heritable component (Palmer & Strobeck, 1992; Palmer, Strobeck & Chippindale, 1994; but see Leamy, 1999 for an example of low heritabilities of directional asymmetry). Yet, work in the past decade has challenged this concept on both theoretical and empirical grounds, and suggests a new concept – all three types of asymmetry are part of a continuum rather than separate forms, and all may represent the ability of an individual to develop along a stable pathway (Graham *et al*., 1993*a*, 1998). Whether stress-mediated shifts in asymmetry types reflect real changes in developmental stability, or are due to changes in the expression of genetic variation (see Hoffman & Merilä, 1999), remains unclear. Irrespective of the mechanism involved, however, failure to detect and quantify admixtures of asymmetry types may confound the presumed underlying relationships with stability. For instance, admixture of FA with directional asymmetry skews the distribution of the signed asymmetry, while admixture of FA with anti-symmetry results in both leptokurtic or platykurtic distributions with zero mean (Palmer & Strobeck, 1992; Van Dongen, 1998*b*). These patterns violate the assumptions for translating observed patterns of FA into the presumed underlying developmental stability (Section III below for methodological details).

(4) Absence of between-trait correlation in FA

Evolutionary models of FA generally assume organism-wide asymmetry, i.e. the tendency that an individual which is more asymmetrical for one trait is more asymmetrical for other traits too (Dufour & Weatherhead, 1996; Clarke, 1998*a*). However, organism-wide asymmetry is only rarely detected, even for traits that are developmentally correlated (Clarke, Oldroyd & Hunt, 1992; Leamy, 1993; Clarke, 1998*b*; but see Lens & Van Dongen, 1999). Yet, for FA to be a convenient estimator of stress and fitness, FA levels should be correlated between traits, within individuals. If not, its indicator ability would depend on choosing the 'right' (i.e. most sensitive) trait, given the stress(es) under study (Palmer, 1994; Hill, 1995; Bjorksten *et al*., 2000*b*). Due to the lack of a unified theory of FA, trait sensitivity is not predictable from first principles and can only be known *post hoc*.

The overall failure to detect between-trait correlations at the individual level may have different, non-exclusive causes. First, it may reflect true between-trait differences in the relationships between FA, developmental stability, stress and fitness, i.e. due to (i) Heterogeneity in timing of expression: if the timing during ontogeny when environmental perturbations can cause aberrant phenotypes is trait specific, developmental noise acting randomly on different traits cannot be expected to cause concordance at the individual level. This is independent of whether the genetic basis of developmental stability is genome-wide or character dependent (Clarke, 1998*b*). (ii) Heterogeneity in FA-fitness relationships: direct, mechanical impacts of single-trait FA on fitness (such as of wing asymmetry on flight ability; Swaddle *et al*., 1996; see also Eggert & Sakaluk, 1994; Palmer, 1994) do not automatically imply organism-wide association between developmental stability and fitness (Bennett & Hoffmann, 1998). (iii) Heterogeneity in selection pressure: traits under strong directional selection, e.g. sexually selected traits (Møller, 1993; Manning & Chamberlain, 1994; Watson & Thornhill, 1994; Hill, 1995), may be more susceptible to stress than traits under stabilizing selection due to the loss of developmental canalization (Pomiankowski & Møller, 1995; but see Bjorksten *et al*. 2000*b*).

Second, failure to detect between-trait correlations may be due to methodological problems related to the estimation of individual developmental stability. For example, as a direct consequence of the low repeatability of individual FA estimates (see section (1) above), between-trait correlations in FA are biased downward, and therefore underestimate between-trait correlations in developmental stability (Whitlock, 1996; Van Dongen, 1998*a*). Addition-

ally, FA values computed from meristic traits are believed to be less sensitive indicators of stability than those obtained from metric traits (Swain, 1987), which may lead to further discrepancies.

III. CONCEPTS AND TOOLS DEVELOPED TO IMPROVE THE ESTIMATION OF INDIVIDUAL DEVELOPMENTAL STABILITY

(1) Reducing the estimation bias of individual FA and developmental stability

Early workers in the field of bilateral asymmetry variation already suggested that the confounding effect of measurement error in the computation of FA can be corrected for by repeatedly measuring the left and right trait sides (e.g. Lundström, 1960). Meanwhile, statistical tools have been developed that effectively allow the separation of measurement error from the analysis of left-right asymmetry. One procedure uses two-way mixed analysis of variance (ANOVA) models, with 'measurement' as the dependent variable, 'individual' as the fixed factor, and 'side' and the 'individual × side' interaction as random factors (Palmer & Strobeck, 1986; Palmer, 1994; Swaddle, Witter & Cuthill 1994; Merilä & Björklund, 1995). More recently, Van Dongen, Molenberghs & Matthysen (1999*b*) introduced a mixed regression model with restricted maximum likelihood parameter estimation. While this routine yields identical FA estimates as the ANOVA model, it further allows modeling of heterogeneity in measurement error among samples, to test for non-zero directional asymmetry and for the statistical significance of FA, and to obtain unbiased estimates of individual FA levels. The analysis can be performed with standard software (e.g. Proc Mixed in SAS; Littell *et al*., 1996), and involves five consecutive steps: (i) examining whether the variance due to measurement error is heterogeneously distributed between samples (e.g. populations, years, treatments); (ii) separating FA (variance components of the random side effect) from measurement error (residual variance); (iii) testing for directional asymmetry; (iv) testing for the statistical significance of FA; and (v) computing unbiased FA values per individual as the random slopes of the individual regression lines. Details on the statistical background and use of the mixed regression procedure (including SAS listings) are given in Van Dongen *et al*. (1999*b*).

In addition to measurement error, the estimation of individual developmental stability is subject to large sampling variability (see Section II), irrespective of the number of within-subject repeats (by contrast, the sampling error of developmental stability estimated at the population level decreases with the number of individuals sampled, since each individual represents one data point). The resulting downward bias in the relationship between individual FA and other variables of interest, however, can be statistically corrected for by applying the hypothetical repeatability (R) of individual FA. R is defined as the ratio of the between-individual component of variation in the unsigned FA, divided by the total variance (Whitlock, 1996, 1998, Van Dongen, 1998*a*, Gangestad & Thornhill, 1999). As such, it represents the proportion of variation in individual FA due to between-individual variation in developmental stability. R can be calculated from the distributional characteristics of the signed and/or unsigned asymmetry (Whitlock, 1996, 1998; Van Dongen, 1998*a*), or directly from the parameter estimates of an admixture of FA components with different variances (Van Dongen, Lens & Molenberghs, 1999*a*; see below). Correct use of R allows translation of patterns of FA into the presumed underlying stability. Yet, this implies that no types of asymmetry other than FA are present, whether or not they reflect developmental stability.

The potential impact of heterogeneity in developmental noise (which is usually assumed to be zero) on between-individual variation in fluctuating asymmetry (see Section II) may be statistically explored by introducing different degrees of heterogeneity to a Bayesian model and evaluating the impact on parameter estimation (S. Van Dongen, unpublished). Recently, Van Dongen (2001) and Van Dongen *et al*. (2001*a*) introduced Bayesian modeling routines to the study of developmental stability, and further developments (including the incorporation of between-individual variation in developmental noise) are in progress. Yet, without experimental study on the biological mechanisms of developmental noise and stability (see above), such developments may have limited practical applicability.

(2) Separating FA from directional asymmetry and anti-symmetry

Van Dongen *et al*. (1999*a*) introduced the use of mixture analysis to identify the three types of bilateral asymmetry from the distribution of the signed asymmetry, based on the following two assumptions: (i) FA, directional asymmetry and anti-symmetry can be described by normal distri-

butions (Palmer & Strobeck, 1992); and (ii) between-individual heterogeneity in developmental stability typically results in a blend of normal distributions with different variances and zero mean (Whitlock, 1996). Mixture analysis (Böhning, Schlattmann & Lindsay, 1992) is a general statistical technique by which combinations of normal distributions can be modelled. When applied to the study of asymmetry, it follows three steps: (1) determining the number of components required to approximate the observed distribution of the signed asymmetry; (2) testing whether or not component means differ from zero; and (3) testing how adequately the selected mixture models describe the observed distributions (Van Dongen *et al*., 1999*a*). A Turbo-Pascal program for conducting mixture analysis of bilateral asymmetry is freely available at http://www.uia.ac.be/u/svdongen/index.html.

(3) Analysing multiple-trait FA

When organism-wide asymmetry is expected (but see Section II), analyses that combine information across traits are believed to be more powerful in detecting relationships between developmental stability, stress and fitness than single-trait analyses (Leary & Allendorf, 1989; Watson & Thornhill, 1994). There are three main statistical routines by which multiple-trait FA can be studied.

First, single composite indices of FA (CFAs) can be computed across traits, e.g. by summing or averaging FA values (Clarke & McKenzie, 1992; Whitlock, 1993), or by standardising FA to control for size effects or differences in presumed stability across traits (Palmer & Strobeck, 1986; Thornhill, Gangestad & Conner, 1995; Leung, Forbes & Houle, 2000). The relative validity (type 1 error rates) and power of detecting relationships with developmental stability of different CFAs depend on the level of kurtosis of the FA distributions, degree of heterogeneity of FA distributions between traits, and sample and effect sizes (Leung *et al*., 2000). If there are *a priori* expectations that FA in particular traits may estimate developmental stability more reliably than in others, weighting factors can be included, e.g. by taking into account measurement error as a measure of (un)reliability. However, Leung *et al*. (2000) plead for caution when applying such weighting procedures, as optimal methods have yet to be determined.

Second, multiple-trait FA can be studied through multivariate analysis, e.g. by generating variance-covariance matrices and computing one overall, standardised variance to quantify FA (Zhivotovsky, 1992), by two-way ANOVA analysis (Palmer, 1994) or mixed regression analysis with repeated measure structure (treating traits as repeated measures; e.g. Lens *et al*., 2000), or by Multivariate analysis of variance (MANOVA) (Alados, Escos & Emlen, 1993; Clarke, 1993*b*). These statistical routines generally produce results comparable to those obtained with CFAs. Because the degrees of freedom in multivariate models reflect the number of individuals rather than the number of traits by individuals, pseudoreplication is avoided. In addition, the models offer the advantage of allowing testing for interactions between stresses and traits (see Lens *et al*., 2000 for an example). Yet, they may also require consideration of more complicated issues, such as the structure of variance-covariance matrices (Leung *et al*., 2000).

Third, it can be achieved by structural equation modelling, a statistical routine which tests models of association between latent variables (i.e. not directly observed) and variation in observable traits (Bollen, 1989). When applied to the study of asymmetry, covariances between FA of individual traits (which are used as 'observable' markers of the underlying 'latent' developmental stability) and the outcome variable(s) of interest are interpreted within the context of an explicit model of associations between variables. Based on the covariations between observed traits, modelling procedures lead to (i) parameter estimates for the paths between traits (under a specific model), (ii) overall measures of fit between variances and the model (indicating whether the model is plausible), and (iii) tests of the statistical significance of particular paths (effects) within the model (S. W. Gangestad, K. Bennett & R. Thornhill, unpublished).

IV. CONFOUNDING EFFECTS OF BIOLOGICAL INTERACTIONS

When attempting to compare FA-fitness relationships between studies, it is necessary to consider the potential effects of biological interactions, such as between different stresses or components of fitness, on the expression of these relationships. While statistical routines have been developed that substantially improve the estimation of developmental stability (see Section III), it is often difficult to account for among-population differences in genetic and/or environmental factors in studies of natural populations, especially when robust experimental

designs such as cross-fostering (e.g. Cadée, 2000) or split-brood experiments (e.g. Van Dongen, Sprengers & Löfstedt, 2001*b*) cannot be applied. Better knowledge of the potential effects of biological interactions can be expected to allow more sound interpretation of the observed patterns in FA, and ultimately, of the developmental stability mechanisms underlying these patterns.

(1) Interactions between fitness components

When testing relationships between developmental stability and fitness, researchers usually measure indirect estimates of fitness, rather than actual fitness (e.g. Møller, 1997). In two case studies on Japanese scorpionflies (*Panorpa japonica*), Thornhill (1992*a*, *b*) showed consistent, inverse relationships between male FA and male-male competitive ability, female preference, and survival. However, there are circumstances where concordance of relationships between FA and different components of fitness cannot be expected. First, components might be traded off with one another. For instance, individuals may sacrifice longevity for increased fecundity (see Ueno, 1994). In such cases, findings largely depend on which fitness estimator is being used. Second, within-cohort selection at early life-stages might reduce variation in developmental stability, hence masking relationships between stability and fitness components at later stages.

(2) Interactions between genotype and environment

Various studies have reported positive relationships between FA and estimates of genetic inbreeding, suggesting that inbreeding reflects 'genetic stress' (Soulé, 1979; Vrijenhoek & Lerman, 1982; Palmer & Strobeck, 1986), and may negatively affect developmental precision (Nilsson, 1994; Palmer, 1996 and references therein). Yet, substantial inconsistency between studies in both the occurrence and magnitude of the association has caused controversy over the generality of these relationships (e.g. Leary *et al.*, 1984; Clarke, 1993*b*; Fowler & Whitlock, 1994; Vøllestad *et al.*, 1999; Hosken, Blanckenhorn & Ward, 2000). At least part of this discordance might be due to genotype-by-environment interactions affecting relationships between developmental stability and environmental or genetic stressors. Proper experiments are needed to clear out such effects. For example, a laboratory study on FA in different genotypes of stalk-eyed flies (*Cyrtodiopsis dalmanni*) across a range of environments that comprised varying degrees of stress, showed significant genotype-by-environment interactions in some traits, but not in others (Bjorksten *et al.*, 2000*a*). Other experiments failed to demonstrate differential responses of genotypes across stress levels (e.g. Van Dongen *et al.*, 2001*b*). Recently, two non-experimental studies on bird populations exposed to different levels of environmental stress provided evidence that the magnitude of the association between stability and inbreeding, itself, may depend on how much environmental stress is experienced during development (Lens *et al.*, 2000; Kark *et al.* 2001).

(3) Heterogeneity in the relationship between quality and fitness

While quality and fitness are often treated as synonyms in the study of developmental stability, Leung & Forbes (1997) illustrated the difference between both concepts by referring to the mediating effect of stress. In a totally benign environment, individual quality can be expected to relate weakly to individual fitness, since the majority of individuals should be able to allocate sufficient energy to the various fitness components. By contrast, in stressed environments, only high-quality individuals might be able to compensate for the increased energy expenditure, while low-quality individuals succumb. Hence, analogous to relationships between FA and developmental stability which may be subject to heterogeneity in the outcome of developmental noise (see Section II), high values of fitness measured under low levels of stress may partly reflect chance effects. Fitness can therefore be expected to be a more reliable indicator of individual quality under stressful conditions and, consequently, relationships between developmental stability and fitness may be difficult to detect under low levels of stress (see Van Dongen & Lens, 2000). In agreement with this hypothesis, Lens, Van Dongen & Matthysen (2001) showed a strong, inverse relationship between FA and survival in a highly stressed natural population of *Turdus helleri*, while the strength of the relationship decreased with increasing habitat quality (i.e. decreasing environmental stress), explaining no variation in the least stressed population.

V. CASE-SPECIFIC SOURCES OF BIAS

Apart from the general factors described under Sections II, III and IV – which are likely to affect most, if not all, studies of developmental

stability – there are other, case-specific sources of noise which may amplify inconsistencies in FA patterns. For instance, heterogeneity in relationships with FA may result from between-study heterogeneity in the statistical power to detect such relationships. As an example, the relationship between heterozygosity, computed over a small number of loci, and that throughout the genome, is assumed to be low in randomly mating populations (Mitton, 1978; Nevo, 1978). If populations strongly differ in effective population sizes, the likelihood of mating between closely related individuals can be expected to differ as well. Due to increased variation in the level of heterozygosity in small populations, coefficients computed over a small number of loci may better reflect genome-wide heterozygosity (Mitton, 1997). In such cases, the closer association between the estimated individual coefficients and the true underlying level of heterozygosity can be expected to increase the statistical power to detect associations with other variables, such as with FA.

Inconsistency between studies may further be increased due to differences in the choice of model species for hypothesis testing. Expectations and predictions are likely to differ between species, and even between populations within a single species, based on their particular life-history patterns. For example, invertebrate species that have the potential to go into diapause, and thus avoid stressful seasons or even years, are expected to show different patterns of FA when compared to populations of birds or other vertebrates that lack this ability (S. Kark, unpublished data). Currently, many studies that examine relationships between asymmetry, stress and fitness fail to validate the presumed negative effect of stress on their study organisms, e.g. by independently testing relationships between the particular stress factor and measures of survival, reproduction or growth. This is, however, important as it might reveal poor choices of stress or fitness traits (e.g. Eggert & Sakaluk, 1994). Hence, future studies should more explicitly explore the biological relevance of stresses or fitness measures that are monitored (Leung & Forbes, 1996).

VI. CONCLUSIONS

(1) Given the growing evidence that relationships between FA, developmental stability, and fitness are stress-dependent, the notion of FA as a universal measure of stability and predictor of stress-mediated changes in fitness, is currently being questioned. We agree with this plea for a critical evaluation of the use of FA.

(2) The current lack of a uniform pattern of relationships with FA, and of insight into how the notions of developmental stability and noise in theories of FA relate to the mechanisms that are the central issue of mainstream developmental biology, does not imply the need for this bio-monitor and potential conservation tool to be abandoned. Rather, we should aim to understand better why in some cases, but not in others, FA reflects stress and fitness, and particularly, what are the factors that cause this discrepancy. This implies further study of the extent to which the observed heterogeneity between studies reflects true differences in developmental stability, or results from statistical, conceptual and/or biological factors that hamper correct interpretation of the observed patterns.

VII. ACKNOWLEDGEMENTS

We thank M. Whitlock, C. P. Klingenberg, T. Backeljau, W. Talloen and two anonymous referees for advice and helpful comments on earlier drafts of this paper. L. L. and S. V. D. hold a postdoctoral fellowship with the Fund for Scientific Research-Flanders (FWO-Vlaanderen).

VIII. REFERENCES

Alados, C. L., Escos, J. & Emlen, J. M. (1993). Developmental instability as an indicator of environmental stress in the Pacific lake (*Merluccius productus*). *U.S. National Marine Fisheries Service*: *Fishery Bulletin* **91**, 587–593.

Aparicio, J. M. (1998). Patterns of fluctuating asymmetry in developing primary feathers: a test of the compensational growth hypothesis. *Proceedings of the Royal Society* (*London*) *B* **265**, 2353–2357.

Bennett, D. M. & Hoffmann, A. A. (1998). Effects of size and fluctuating asymmetry on field fitness of the parasitoid *Trichogramma carverae* (Hymenoptera: Trichogrammatidae). *Journal of Animal Ecology* **67**, 580–591.

Böhning, D., Schlattmann, P. & Lindsay, B. (1992). Computer-assisted analysis of mixtures (C. A. MAN): statistical algorithms. *Biometrics* **48**, 283–303.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. John Wiley & Sons, New York.

Bjorksten, T., David, P., Pomiankowski, A. & Fowler, K. (2000*a*). Fluctuating asymmetry of sexual and nonsexual traits in stalk-eyed flies: a poor indicator of developmental stress and genetic quality. *Journal of Evolutionary Biology* **13**, 89–97.

Bjorksten, T. A., Fowler, K. & Pomiankowski, A. (2000*b*). What does sexual trait FA tell us about stress? *Trends in Ecology and Evolution* **15**, 163–166.

Cadée, N. (2000). Genetic and environmental effects on morphology and fluctuating asymmetry in nestling barn swallows. *Journal of Evolutionary Biology* **13**, 359–370.

Cairns, J., McCormick, P. V. & Niederlehner, B. R. (1993). A proposed framework for developing indicators of ecosystem health. *Hydrobiologia* **263**, 1–44.

Clarke, G. M. (1993*a*). Fluctuating asymmetry of invertebrate populations as a biological indicator of environmental quality. *Environmental Pollution* **82**, 207–211.

Clarke, G. M. (1993*b*). Patterns of developmental stability of *Chrysopa perla L.* in response to environmental pollution. *Environmental Entomology* **22**, 1362–1366.

Clarke, G. M. (1993*c*). The genetic basis of developmental stability: I. Relationships between stability, heterozygosity and genomic coadaptation. *Genetica* **89**, 15–23.

Clarke, G. M. (1995*a*). Relationships between developmental stability and fitness: application for conservation biology. *Conservation Biology* **9**, 18–24.

Clarke, G. M. (1995*b*). Relationships between fluctuating asymmetry and fitness: how good is the evidence? *Pacific Conservation Biology* **2**, 146–149.

Clarke, G. M. (1998*a*). Developmental stability and fitness: the evidence is not quite so clear. *American Naturalist* **152**, 762–766.

Clarke, G. M. (1998*b*). The genetic basis of developmental stability. IV. Individual and population asymmetry parameters. *Heredity* **80**, 553–561.

Clarke, G. M. & McKenzie, L. J. (1992). Fluctuating asymmetry as a quality control indicator for insect mass rearing processes. *Journal of Economical Entomology* **85**, 2045–2050.

Clarke, G. M., Oldroyd, B. P. & Hunt, P. (1992). The genetic basis of developmental stability in *Apis mellifera*: heterozygosity vs. genic balance. *Evolution* **46**, 753–762.

Dufour, K. W. & Weatherhead, P. J. (1996). Estimation of organism-wide asymmetry in red-winged blackbirds and its relation to studies of mate selection. *Proceedings of the Royal Society* (*London*) *B* **263**, 769–775.

Dufour, K. W. & Weatherhead, P. J. (1998). Bilateral symmetry as an indicator of male quality in red-winged blackbirds: associations with measures of health, viability, and parental effort. *Behavioral Ecology* **9**, 220–231.

Eggert, A. K. & Sakaluk, S. K. (1994). Fluctuating asymmetry and variation in the size of courtship food gifts in decorated crickets. *American Naturalist* **144**, 708–716.

Fiering, S., Whitelaw, E. & Martin, D. I. K. (2000). To be or not to be active: the stochastic nature of enhancer action. *BioEssays* **22**, 381–387.

Fowler, K. & Whitlock, M. (1994). Fluctuating asymmetry does not increase with moderate inbreeding in *Drosophila melanogaster*. *Heredity* **73**, 373–376.

Gangestad, S. W. & Thornhill, R. (1999). Individual differences in developmental precision and fluctuating asymmetry: a model and its implications. *Journal of Evolutionary Biology* **12**, 402–416.

Graham, J. H., Emlen, J. M., Freeman, D. C., Leamy, L. J. & Kieser, J. A. (1998). Directional asymmetry and the measurement of developmental instability. *Biological Journal of the Linnean Society* **64**, 1–16.

Graham, J. H., Freeman, D. C. & Emlen, J. M. (1993*a*). Antisymmetry, directional asymmetry, and dynamic morphogenesis. *Genetica* **89**, 121–137.

Graham, J. H., Roe, K. E. & West, T. B. (1993*b*). Effects of lead and benzene on the developmental stability of *Drosophila melanogaster*. *Ecotoxicology* **2**, 185–195.

Henshel, D. S., Cheng, K. M., Norstrom, R., Whitehead, P. & Steeves, J. D. (1993). Morphometric and histologic changes in brains of Great Blue Heron hatchlings exposed to PCDDs: preliminary analyses. In: W. G. Landis, J. S. Hughes & M. A. Lewis, eds. *Environmental toxicology and risk assessment*. American Society for Testing and Materials 1179, Philadelphia, pp. 262–277.

Hill, G. E. (1995). Ornamental traits as indicators of environmental health. BioScience **45**, 25–31.

Hosken, D. J., Blanckenhorn, W. U. & Ward, P. I. (2000). Developmental stability in yellow dung flies (Scathophaga stercoraria); fluctuating asymmetry, heterozygosity and environmental stress. *Journal of Evolutionary Biology* **13**, 919–926.

Houle, D. (1997). Comment on "A meta-analysis of the heritability of developmental stability" by Møller and Thornhill. *Journal of Evolutionary Biology* **10**, 17–20.

Kark, S. (1999). Within-population diversity in the distribution range: Partridges as a research model. Ph.D thesis, The Hebrew University of Jerusalem, Jerusalem and http://www.stanford.edu/group/Mooney/salit.

Kark, S., Safriel, U. N., Tabarroni, C. & Randi, E. (2001). Relationship between heterozygosity and asymmetry: a test across the distribution range. *Heredity* **86**, 119–127.

Kieser, J. A. & Groeneveld, H. T. (1991). Fluctuating odontometric asymmetry, morphological variability, and genetic monomorphism in the cheetah *Acinonyx jubatus*. *Evolution* **45**, 1175–1183.

Klingenberg, C. P. (2001). A developmental perspective on developmental instability: theory, models and mechanisms. In: M. Pòlak, ed. *Developmental instability: causes and consequences*. Oxford University Press, Oxford, In Press.

Klingenberg, C. P. & Nijhout, H. F. (1999). Genetics of fluctuating asymmetry: a developmental model of developmental instability. *Evolution* **53**, 358–375.

Koehn, R. K. & Bayne, B. L. (1989). Towards a physiological and genetic understanding of the energetics of the stress response. *Biological Journal of the Linnean Society* **37**, 157–171.

Leamy, L. (1993). Morphological integration of fluctuating asymmetry in the mouse mandible. *Genetica* **89**, 139–153.

Leamy, L. (1999). Heritability of directional and fluctuating asymmetry for mandibular characters in random-bred mice. *Journal of Evolutionary Biology* **12**, 146–155.

Leamy, L., Doster, M. J. & Huet-Hudson, Y. M. (1999). Effects of methoxychlor on directional and fluctuating asymmetry of mandible characters in mice. *Ecotoxicology* **8**, 63–71.

Leary, R. F. & Allendorf, F. W. (1989). Fluctuating asymmetry as an indicator of stress: implications for conservation biology. *Trends in Ecology and Evolution* **4**, 214–217.

Leary, R. F., Allendorf, F. W. & Knudsen, K. L. (1984). Superior developmental stability of heterozygosity in enzyme loci in salmonid fishes. *American Naturalist* **124**, 540–551.

Lens, L. & van Dongen, S. (1999). Evidence for organism-wide asymmetry in five bird species of a fragmented afrotropical forest. *Proceedings of the Royal Society* (*London*) *B* **266**, 1055–1060.

Lens, L. & van Dongen, S. (2000). Fluctuating and directional asymmetry in natural bird populations exposed to different levels of habitat disturbance, as revealed by mixture analysis. *Ecology Letters* **3**, 516–522.

Lens, L., van Dongen, S., Galbusera, P., Schenck, T., Matthysen, E. & van De Casteele, T. (2000). Developmental instability and inbreeding in natural bird populations exposed to different levels of habitat disturbance. *Journal of Evolutionary Biology* **13**, 889–896.

Lens, L., van Dongen, S. & Matthysen, E. (2001). Fluctuating asymmetry as early warning system in the critically endangered Taita thrush. *Conservation Biology*, In Press.

Lens, L., van Dongen, S., Wilder, C. M., Brooks, T. M. & Matthysen, E. (1999). Fluctuating asymmetry increases with habitat disturbance in seven bird species of a fragmented afrotropical forest. *Proceedings of the Royal Society* (*London*) *B* **266**, 1241–1246.

Leung, B. & Forbes, M. R. (1996). Fluctuating asymmetry in relation to stress and fitness: effect of trait type as revealed by meta-analysis. *Ecoscience* **3**, 400–413.

Leung, B. & Forbes, M. R. (1997). Modelling fluctuating asymmetry in relation to stress and fitness. *Oikos* **78**, 397–405.

Leung, B., Forbes, M. R. & Houle, D. (2000). Fluctuating asymmetry as a bioindicator of stress: comparing efficacy of analyses involving multiple traits. *American Naturalist* **155**, 101–115.

Littell, R. C., Milliken, G. A., Stroup, W. W. & Wolfinger, R. D. (1996). *SAS system for mixed models*. SAS institute Inc., Cary, North Carolina.

Ludwig, W. (1932). *Das Rechts-Links Problem im Tierreich und beim Menschen*. Springer, Berlin.

Lundström, A. (1960). Asymmetries in the number and size of the teeth and their aethological significance. *Transactions of the European Orthodontists Society* **36**, 167–185.

Manning, J. T. & Chamberlain, A. T. (1994). Fluctuating asymmetry in gorilla canines: a sensitive indicator of environmental stress. *Proceedings of the Royal Society* (*London*)*B* **255**, 189–193.

Markow, T. A. (1995). Evolutionary ecology and developmental instability. *Annual Review of Entomology* **40**, 105–120.

McAdams, H. H. & Arkin, A. (1999). It's a noisy business! Genetic regulation at the nanomolecular scale. *Trends in Genetics* **15**, 65–69.

McKenzie, J. A. & Clarke, G. M. (1988). Diazanon resistance, fluctuating asymmetry and fitness in the Australian sheep blowfly, *Lucilia cuprina*. *Genetics* **120**, 213–220.

Merilä, J. & Björklund, M. (1995). Fluctuating asymmetry and measurement error. *Systematic Biology* **44**, 97–101.

Mitton, J. B. (1978). Relationship between heterozygosity for enzyme loci and variation of morphological characters in natural populations. *Nature* **273**, 661–662.

Mitton, J. B. (1997). *Selection in natural populations*. Oxford University Press, Oxford.

Møller, A. P. (1993). Developmental stability, sexual selection, and speciation. *Journal of Evolution Biology* **6**, 493–509.

Møller, A. P. (1997). Developmental stability and fitness: a review. *American Naturalist* **149**, 916–932.

Møller, A. P. & Swaddle, J. P. (1997). *Asymmetry, Developmental Stability, and Evolution*. Oxford: Oxford University Press.

Møller, A. P. & Thornhill, R. (1998). Bilateral symmetry and sexual selection: a meta-analysis. *American Naturalist* **151**, 174–192.

Nevo, E. (1978). Genetic variation in natural populations: patterns and theory. *Theoretical Population Biology* **13**, 121–177.

Nilsson, J. A. (1994). Energetic stress and the degree of fluctuating asymmetry: implications for a long-lasting, honest signal. *Evolutionary Ecology* **8**, 248–255.

Novak, J. M., Rhodes, O. E., Smith, M. H. Jr. & Chesser, R. K. (1993). Morphological asymmetry in mammals: Genetics and homeostasis reconsidered. *Acta Theriologica* **38**, 7–18.

Palmer, A. R. (1994). Fluctuating asymmetry analyses: A primer. In: *Developmental stability: its Origins and Evolutionary Implications* (ed. T. A. Markow), pp. 335–364. Kluwer Academic Publishers, Dordrecht.

Palmer, A. R. (1996). Walzing with asymmetry. Is fluctuating asymmetry a powerful new tool for biologists or just an alluring new dance step? *BioScience* **46**, 518–532.

Palmer, A. R. (2000). Detecting publication bias in meta-analyses: A case study of fluctuating asymmetry and sexual selection. *American Naturalist* **154**, 220–233.

Palmer, A. R. & Strobeck, C. (1986). Fluctuating asymmetry: measurement, analysis and patterns. *Annual Review of Ecology and Systematics* **17**, 391–421.

Palmer, A. R. & Strobeck, C. (1992). Fluctuating asymmetry as a measure of developmental stability: Implications of non-normal distributions and power of statistical tests. *Acta Zoologica Fennica* **191**, 57–72.

Palmer, A. R., Strobeck, C. & Chippindale, A. K. (1994). Bilateral variation and the evolutionary origin of macroscopic asymmetries. In: T. A. Markow, ed. *Developmental Instability: its Origins and Evolutionary Implications*. Kluwer Academic Publishers, Dordrecht, pp. 203–220.

Parsons, P. A. (1990). Fluctuating asymmetry: an epigenic measure of stress. *Biological Review* **65**, 131–154.

Parsons, P. A. (1992). Fluctuating asymmetry: a biological monitor of environmental and genomic stress. *Heredity* **68**, 361–364.

Polak, M. & Trivers, R. (1994). The science of symmetry in biology. *Trends in Ecology and Evolution* **9**, 122–124.

Pomiankowski, A. & Møller, A. P. (1995). A resolution of the lek paradox. *Proceedings of the Royal Society of London Series B* **260**, 21–29.

Reeve, E. C. R. (1960). Some genetic tests on asymmetry of sternopleural chaetae number in *Drosophila*. *Genetical Research* **1**, 151–172.

Rutherford, S. L. & Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature* **396**, 336–342.

Sommer, C. (1996). Ecotoxicology and developmental stability as an in situ monitor of adaptation. *Ambio* **25**, 374–376.

Soulé, M. E. (1979). Heterozygosity and developmental stability: another look. *Evolution* **33**, 396–401.

Swaddle, J. P. & Witter, M. S. (1997). On the ontogeny of developmental stability in a stabilized trait. *Proceedings of the Royal Society* (*London*) *B* **264**, 329–334.

Swaddle, J. P., Witter, M. S. & Cuthill, I. C. (1994). The analysis of fluctuating asymmetry. *Animal Behaviour* **48**, 986–989.

Swaddle, J. P., Witter, M. S., Cuthill, I. C., Budden, A. & McCowen, P. (1996). Plumage condition affects flight performance in common starlings: Implications for developmental homeostasis, abrasion and moult. *Journal of Avian Biology* **27**, 103–111.

Swain, D. P. (1987). A problem with the use of meristic characters to estimate developmental stability. *American Naturalist* **129**, 761–768.

Thornhill, R. (1992*a*). Female preference for the pheromone of males with low fluctuating asymmetry in the Japanese scorpionfly (*Panorpa japonica*: Mecoptera). *Behavioral Ecology* **3**, 277–283.

Thornhill, R. (1992*b*). Fluctuating asymmetry and the mating system of the Japanese scorpionfly *Panorpa japonica*. *Animal Behaviour* **44**, 867–879.

Thornhill, R., Gangestad, S. W. & Comer, R. (1995). Human female orgasm and mate fluctuating asymmetry. *Animal Behaviour* **50**, 1601–1615.

Tomkins, J. L. (1999). The ontogeny of asymmetry in earwig forceps. *Evolution* **53**, 157–163.

Ueno, H. (1994). Fluctuating asymmetry in relation to two fitness components, adult longevity and male mating success in a ladybird beetle, *Harmonia axyridis* (Coleoptera: Coccinellidae). *Ecological Entomology* **19**, 87–88.

van Dongen, S. (1998*a*). How repeatable is the estimation of developmental stability by fluctuating asymmetry? *Proceedings of the Royal Society of London Series B* **265**, 1423–1427.

van Dongen, S. (1998*b*). The distribution of individual fluctuating asymmetry: why are the coefficients of variation so high? *Annales Zoologici Fennici* **35**, 79–85.

van Dongen, S. (1999). Accuracy and power in fluctuating asymmetry studies: effects of sample size and number of within-subject repeats. *Journal of Evolutionary Biology* **12**, 547–550.

van Dongen, S. (2000). Unbiased estimation of individual symmetry. *Journal of Evolutionary Biology* **13**, 107–112.

van Dongen, S. (2001). Modeling developmental instability in relation to individual fitness: a fully Bayesian latent variable model approach. *Journal of Evolutionary Biology* **14**, 552–563.

van Dongen, S. & Lens, L. (2000). The evolutionary potential of developmental instability. *Journal of Evolutionary Biology* **13**, 326–335.

van Dongen, S., Lens, L. & Molenberghs, G. (1999*a*). Mixture analysis of asymmetry: modelling directional asymmetry, antisymmetry and heterogeneity in fluctuating asymmetry. *Ecology Letters* **2**, 387–396.

van Dongen, S., Lens, L. & Molenberghs, G. (2001*a*). Recent developments and shortcomings in the analysis of individual asymmetry: Can Bayesian statistics help us? In: M. Polak, ed. *Developmental instability: causes and consequences*. Oxford University Press, Oxford, In Press.

van Dongen, S., Molenberghs, G. & Matthysen, E. (1999*b*). The statistical analysis of fluctuating asymmetry: REML estimation of a mixed regression model. *Journal of Evolutionary Biology* **12**, 94–102.

van Dongen, S., Sprengers, E. & Lofstedt, C. (2001*b*). Lack of evolutionary potential of developmental instability of front tibia length in the Indian meal moth (*Plodia interpunctella*). *Belgian Journal of Zoology* **131**, 75–82.

van Valen, L. (1962). A study of fluctuating asymmetry. *Evolution* **16**, 125–142.

Völlestad, L.A., Hindar, K. & Möller, A. P. (1999). A meta-analysis of fluctuating asymmetry in relation to heterozygosity. *Heredity* **83**, 206–218.

Vrijenhoek, R. C. & Lerman, S. (1982). Heterozygosity and developmental stability under asexual breeding systems. *Evolution* **36**, 768–776.

Watson, P. J. & Thornhill, R. (1994). Fluctuating asymmetry and sexual selection. *Trends in Ecology & Evolution* **9**, 21–25.

Whitlock, M. (1993). Lack of correlation between heterozygosity and fitness in forked fungus beetles. *Heredity* **70**, 574–581.

Whitlock, M. (1996). The heritability of fluctuating asymmetry and the genetic control of developmental stability. *Proceedings of the Royal Society of London Series B* **263**, 849–854.

Whitlock, M. (1998). The repeatability of fluctuating asymmetry: A revision and extension. *Proceedings of the Royal Society of London Series B* **265**, 1428–1430.

Wooten, M. C. & Smith, M. H. (1986). Fluctuating asymmetry and genetic variability in a natural population of *Mus musculus*. *Journal of Mammalogy* **67**, 725–732.

Zhivotovsky, L. A. (1992). A measure of fluctuating asymmetry for a set of characters. *Acta Zoologica Fennica* **191**, 73–77.