UNIVERSITY OF CINCINNATI

January 27, 2003

I, Elizabeth M. Stillabower,

hereby submit this as part of the requirements for the degree of:

Master of Science

in the department of Biological Sciences

It is entitled The relationship between genotype, fluctuating asymmetry, & mating success in natural populations of Drosophila immigrans.

Approved by:

Jodie R. Shaw

Steven L. Courtright

The relationship between genotype, fluctuating asymmetry, and mating success in natural populations of *Drosophila immigrans*

A thesis submitted to the

Division of Research and Advanced Studies
of the University of Cincinnati

in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE

in the Department of Biological Sciences
of the College of Arts and Sciences

2003

by

Elizabeth M. Stillabower

B.A., Hanover College, 1998

Committee Chair: Dr. Michal Polak

Abstract

Fluctuating asymmetry, or random deviations from perfect bilateral symmetry, is a widely accepted measure of developmental instability. It has been proposed in the literature that genetically-based differences in asymmetry may confer a disadvantage in competition for mates. The present study evaluates this hypothesis using *Drosophila immigrans*. I first examined the magnitude of heritability in sternopleural bristles of this species using an isofemale line approach. Lines were cultured in replicate bottles, so that environmental and genetic sources of variation could be distinguished. Two forms of fluctuating asymmetry exhibited significant genetic variation among lines, and one form showed significant heritability. I examined the relationship between asymmetry and several components of mating success: copulation rate, mating latency, copulation duration, and male fertility. Copulation rate and fertility were not related to developmental instability. However, both mating latency and copulation duration were significantly longer in high-developmental instability males. I tested whether these differences were the result of differences in asymmetry ("direct involvement hypothesis") or differences in underlying vigor ("vigor hypothesis"). I compared mating success between high- and low- developmental instability males with similar asymmetry phenotypes. Mating latency and duration were both significantly longer in high-developmental instability males than in low-developmental instability males, despite the fact that asymmetry was held constant, thus supporting the vigor hypothesis. These results suggest that genetically-based differences in developmental instability can be a significant determinant of male sexual performance, and that developmental instability in natural populations of *Drosophila immigrans* has significant evolutionary potential.

Acknowledgements

Throughout my two-and-a-half years in graduate school many people have provided me with support and guidance. The first of these is my advisor, Dr. Michal Polak, who has helped me to understand the concepts presented in this thesis more clearly, as well as providing the initial ideas for this project, suggesting improvements along the way, and assisting in the completion of portions of some experiments. My other committee members, Drs. Jodi Shann and Iain Cartwright, also offered suggestions from different viewpoints, which I believe helped to improve the quality of this paper and my project. My labmates, Jerry Hinn, Ben Heath, and Chris Nasrallah, spent many hours observing flies during mating experiments.

I would also like to thank my parents, Don and Pat Quinn, who have always encouraged me to do my best. Finally, I need to thank my husband, Scott Stillabower, who has been incredibly supportive and understanding, and has always encouraged me to do what makes me happy.

Table of Contents

List of Abbreviations

FA = Fluctuating Asymmetry

NFA = Numerical Fluctuating Asymmetry

PFA = Positional Fluctuating Asymmetry

DA = Directional Asymmetry

AS = Antisymmetry

DI = Developmental Instability

Introduction

The ability of an organism to develop properly when challenged by perturbation of either environmental or genetic origin has important consequences for its overall fitness (Gibson and Wagner 2000). This is because deviations from "proper" or "ideal" development are assumed to incur physiological costs and fitness disadvantages, so selection is thought to operate to maintain developmental fidelity. While departure from ideal development is not always easy to identify, departures from symmetry in morphological traits may be helpful in this regard because perfect symmetry is generally accepted as the optimal state of bilaterally symmetrical characters (Mather 1953).

Random deviations from perfect symmetry, referred to as fluctuating asymmetry (FA), are a widely accepted measure of developmental instability (DI). Developmental instability is defined as small, random perturbations to developmental processes (Palmer 1994), and may include, for example, random differences in physiological processes within cells, and rates of cell division on either side of the body (Palmer 1994, 1996). Developmental stability ("buffering capacity," VanValen 1962), on the other hand, refers to processes that reduce the effects of developmental perturbation, thus resulting in more buffered development and a consequent reduction in asymmetry (Palmer 1994). One possible example of a developmental stability mechanism involves feedback systems regulating enzyme activity among cells (Palmer 1994).

In order to identify asymmetry as "true FA," a sample of several individuals must be used. For each individual in the sample, the size of one side of a character is subtracted from the size of the other side of the character, and the mean and distribution of these values [i.e., signed right-minus-left (R-L) values] are then examined. Fluctuating asymmetry is

characterized by a normal distribution with a mean of zero (Mather 1953, Palmer 1994, 1996) (Figure 1). Other kinds of asymmetry are distinguished from FA by the shape of the distribution of R-L values (VanValen 1962). Antisymmetry (AS) refers to a pattern of these values where the variation is distributed about a mean of zero, but the frequency distribution departs from normality in the direction of bimodality (Figure 1). On the other hand, directional asymmetry (DA) refers to a pattern of R-L values where the variation is normally distributed but where the mean is significantly different from zero (Figure 1). Whether or not these forms of asymmetry can reflect DI is unclear and even controversial (Palmer 1994). A consensus opinion is that directional asymmetry and antisymmetry usually have a non-random genetic component, and so are unlikely to yield reliable measures of DI (Palmer and Strobeck 1992; Polak 2003).

An important unresolved question in the field of developmental instability is the relative contributions of genotypic versus environmental effects on asymmetry variation in natural populations (Nijhout and Davidowitz 2003). Understanding the relative role of genotype in causing FA differences between individuals is crucial for evaluating a body of popular theory linking FA and sexual selection (Møller and Pomiankowski 1993; Tomkins and Simmons 2003). The importance of genetic and environmental contributions to FA, and important components of sexual selection theory are described in the following sections.

Genotypic and environmental factors influencing FA

The relative contributions of genotype and environment in determining FA have been debated. Some studies suggest that FA is heritable (VanDongen *et al.* 1999, M øller and Thornhill 1997, Polak and Starmer 2001), while many others have failed to detect a

significant heritability (Chapman and Goulson 2000, and see Fuller and Houle 2003). For example, a study of winter moths (*Operophtera brumata*) (VanDongen et al. 1999) did detect a significant average heritability value for three pairs of tibia using a full-sib analysis. Full-sib designs, however, confound dominance effects with additive genetic variation (Fuller and Houle 2003). Møller and Thornhill's (1997) meta-analysis determined that heritability estimates of FA across multiple studies were indeed significant. However, this analysis has received criticism for including studies that did not test for the presence of DA or AS, did not account for measurement error (see Palmer and Strobeck 1997), or selected for DA (see Whitlock and Fowler 1997). In addition, meta-analysis may not be appropriate for heritability estimates because these analyses combine data across many different species and traits, which presumably are controlled by different genetic systems (Markow and Clarke 1997). A parent-offspring regression performed by Chapman and Goulson (2000) detected no heritable component to wing length FA in *Musca domestica*. It is possible, however, that asymmetry in wing length is strongly selected against, as it directly relates to an individual's ability to fly and, presumably, to escape predators (Cadee 2000). It is thought that such strong stabilizing selections would deplete heritable genetic variation underlying FA in traits such as wing length (Møller and Pomiankowski 1993).

Studies of the effects of environmental factors on FA have been widespread with respect to the factor examined. These factors have ranged from chemicals (Ostbye *et al.* 1997, Graham *et al.* 1993, Polak *et al.* 2002) and temperature (Chapman and Goulson 2000) to maternal parasitism (Polak 1997). Some of these studies have detected an effect of environmental factors on FA. For example, in perch (*Perca fluviatilis*), both mandibular pore number FA and a meristic asymmetry index increased with increasing water acidity (Ostbye

et al. 1997). In addition, *D. melanogaster* exposed to lead and benzene during development exhibited increased sternopleural bristle FA (Graham *et al.* 1993). Conversely, exposure of *D. melanogaster* larvae to varying concentrations of arsenic resulted in reduced FA among flies raised in the highest concentrations, but this effect was attributed to high mortality among the least developmentally stable individuals (Polak *et al.* 2002). In house flies (*Musca domestica*), culture temperature appears to be related to wing length FA (Chapman and Goulson 2000). Barn swallows (*Hirundo rustica*) raised in experimentally enlarged broods had significantly higher tarsus FA than did those raised in experimentally reduced broods (Cadee 2000). However, barn swallows did not differ in tail or wing FA regardless of brood size (Cadee 2000). Finally, ectoparasitism of female *Drosophila nigrospiracula* resulted in increased positional FA of sternopleural bristles in male offspring (Polak 1997). Positional FA (PFA) reflects the placement of the bristles on different areas of the sternopleuron, rather than simply the number of bristles present. Maternal age, however, was not significantly related to FA of sternopleural bristles, wing length, or wing area in *D. melanogaster* (Wakefield *et al.* 1994).

Fluctuating Asymmetry and Sexual Selection

Sexual selection is variation in reproductive success that results from competition for mates (Darwin 1871). This variation may result from two forms of competition. First, males may compete among each other to gain access to reproductive females, thus giving rise to intrasexual selection. The outcome of male-male competition will often depend on differences in male phenotypic condition, body size and vigor (Anderson 1994).

A second form of sexual selection, intersexual selection, occurs when females prefer to mate with males expressing certain phenotypic features over males that do not express these features. Thus, the outcome of female choice will often depend on differences among males in the degree of exaggeration of secondary sexual traits (e.g., tail size, body coloration, courtship behavior). By choosing to mate with males with the most extravagant secondary sexual traits, females may acquire direct benefits (e.g., valuable food items, territories etc.) for themselves or their offspring. In contrast, female choice may evolve as a function of indirect benefits, in the form of "good genes" which enhance the attractiveness of their sons or the viability of male and female offspring (Andersson 1994). According to the "good genes" hypothesis, the size of male secondary sexual traits correlates positively with intrinsic, heritable quality.

Anders Pape Møller in the early 1990s was the first to propose a general hypothesis for a role of FA in sexual selection (Møller 1990, 1992; Møller and Pomiankowski 1993). Much of this theory is predicated on the idea that FA reflects genotype-specific "quality" of individuals, defined as an organism's ability to survive and to reproduce in a given environment, or as stated by Leung and Forbes (1997), as "the probability of fitness." In the context of female choice, one important hypothesis is that FA in secondary sexual traits reveals information to the female about a potential mate's general health and parenting ability. According to "good-genes" models of sexual selection, mate choice in favor of low FA mates is favored by selection because of the superior quality inherited by offspring produced from such matings. Thus, this hypothesis also predicts that trait size and FA will be negatively correlated because individual quality increases with trait size. Some studies have demonstrated that females are more likely to choose the most symmetrical males as

mates (Møller 1992, Morris 1998, Uetz and Smith 1999), although other studies have failed to find such an association (Radesater and Halldorsdottir 1993, Tomkins and Simmons 1998, Goulson *et al.* 1999).

If FA is a reliable indicator of an individual's quality, males of the same species may also be selected to assess each other's fighting ability through symmetry, thereby reducing the potential costs of conflict. Low FA individuals may enjoy a mating advantage simply because they are more vigorous and energetic, hence better able to outcompete sexual rivals either through direct competition, or through an enhanced ability to find and copulate with a receptive female ("scramble competition", Thornhill and Alcock 1983). An important distinguishing feature of this "vigor hypothesis," is that it predicts that male mating success may correlate with FA in *non*-secondary sexual characteristics (i.e., in those characteristics not directly assessed by choosy mates or competitors), as well as in secondary sexual traits. Low FA males in these cases are not more successful because their degree of symmetry is directly assessed (the "direct assessment hypothesis"), but because low FA reflects males' low DI, and hence high intrinsic quality.

Møller (1992) found that female barn swallows choose mates based on tail feather symmetry. Møller (1996) also found that, when given a choice of males, female house flies (*M. domestica*) chose to mate first with males having more symmetric wings and tibia. Males also chose to mate first with females having the most symmetrical wings (Møller 1996). In the field, unmated *M. domestica* had significantly more asymmetric wings than did mated individuals (Møller 1996). Thus, asymmetry may be a *reflection* of metabolic inefficiency and reduced overall vigor (Møller 1996). Alternatively, it is possible that asymmetry in barn swallow tail feathers and house fly wings *causes*

individuals to be less energetically efficient, thus not allowing them to use as much energy in courtship as symmetrical individuals. Polak (1997) found no difference between mated and unmated *Drosophila nigrospiracula* in FA of the sternopleural bristles, a character that may be less likely to directly affect performance.

Overall, there are few studies that have found support for the role of FA in sexual selection (Tomkins and Simmons 2003). Moreover, in cases in which FA does covary with male mating success, the extent to which genetic variation exists for FA is generally unknown; understanding the genetic basis of FA is an important component of the FA-sexual selection hypothesis (see above). Thus, there is a need for studies that test for an effect of FA on mating success, and for a genetic basis to FA simultaneously in the same species.

Research objectives

The present study employs an iso-female line approach to estimate the magnitude of genotypic effects relative to environmental effects underlying FA variation in bristle characteristics in natural populations of *Drosophila immigrans* (Diptera: Drosophilidae). In the event of uncovering significant genetic involvement for FA, I report heritability estimates as well as genetic correlations between FAs and trait size. A second component of the study is to examine the relationship between genotype-specific differences in FA and male mating success. To this end I employ a novel technique, which entails contrasting mating success estimates between high asymmetry males and low asymmetry males chosen from high FA genetic lines and low FA genetic lines, respectively. This method is expected to be a more sensitive assay of the relationship

between FA and fitness correlates compared to methods of measuring mating success of males sampled from a genetically heterogeneous population, hence of an unknown genetic background. I also examine the relationship between genotype-specific FA and number of offspring produced by individual males, a more direct measure of fitness than estimates of mating success based, for example, on counting numbers of copulations. Finally, I examine the relationship between mating success and underlying developmental stability by contrasting mating performance between males with similar FA phenotypes chosen from either high or low FA genetic lines (i. e., high- or low-DI lines). This allowed me to evaluate the relative importance of FA versus underlying DI in affecting male performance.

Methods

Isofemale lines

Isofemale lines were derived from wild-caught female *D. immigrans*. *D. immigrans* is a cosmopolitan species (Patterson 1943), which is relatively abundant and easily attracted to fruit and vegetable baits in the field. Females initiating lines were collected from the field by baiting in June 2000, from two sites in the United States: Independence, KY (IKY: 38.929° N, 84.538° W), and Fayetteville, NY (FNY: 43.030° N, 75.998° W). Wild-caught females were presumed to have mated with multiple males prior to capture. In the laboratory, females were allowed to oviposit into individual 8 fluid-dram shell vials containing powdered Betty Crocker® instant mashed potato flakes, instant *Drosophila* medium (Carolina Biological Supply Co.), distilled water, and 4-5 mg of Fleischmann's active dry yeast. Virgin offspring of each female were harvested and allowed to reach sexual

maturity in single-sex food vials. A single virgin brother-sister pair was randomly selected from each group of offspring and allowed to mate. The offspring of each such pair were used to propagate each isofemale line. Lines were cultured in half-pint bottles containing 12g powdered mashed potato flakes, 3 g *Drosophila* instant medium, 60 ml distilled water, and 10-12 mg yeast. Adults were transferred to a fresh bottle after two days of oviposition in order to retain moderate and similar larval densities across bottles and lines; thus, there were two bottles per line. This method of controlling larval densities across bottles was preferred over manual manipulation because it avoids handling the larvae, which I suspected could be stressful to them and influence adult FAs. Bottles were placed at random within the incubator to avoid any systematic environmental effects, and all flies were cultured and maintained in a Percival incubator with a 12 (25 °C):12 (23 °C) L:D photoperiod. During the first 4 d of eclosion, approximately 20 adult progeny (ten flies of each sex) were randomly selected from each bottle and bristle number and thorax length were determined following procedures described next.

Determination of fluctuating asymmetry

Fluctuating asymmetry was estimated in sternopleural bristle number. Each fly was killed with ether fumes, and under an Olympus SZX12 stereomicroscope, its sternopleural bristles were counted. Anterior and transverse bristles on the sternopleuron were distinguished (see Fig. 1 in Polak 1997). Signed asymmetry was calculated as the value of a given trait on the right side of the body minus its value on the left (R-L). Overall FA was taken as the total number of bristles (i.e., anterior plus transverse bristles) on the right side minus the total number on the left. FA1 and FA2 refer to asymmetry in the anterior and

transverse sternopleural bristles, respectively. Overall FA, FA1, and FA2 are measures of numerical FA (NFA), reflecting the difference in the number of bristles present on the two sides of the sternopleuron. Positional asymmetry (Polak 1997), which reflects the difference between the two sides of the body in the placement of bristles on the sternopleuron, was calculated by two methods:

PFA1 = (number of right anterior bristles / number of right transverse) – (number of left anterior bristles / number of left transverse), and

PFA2= ln (|number of right anterior bristles – number right transverse| +0.5) – ln (|number left anterior – number left transverse| +0.5).

Trait size is defined as the total number of a given bristle type. Tot1 and tot2 are the total number of anterior and transverse strenopleural bristles, respectively. Tot1,2 is the sum of tot1 and tot2, and is the associated measure of trait size for overall FA, PFA1, and PFA2. Thorax length, measured from the anterior end of the thorax to the posterior end of the scutellum using an ocular micrometer of a stereomicroscope, estimated adult body size.

Distributions of signed asymmetry values were examined for the presence of directional asymmetry and antisymmetry. Normality of each distribution was assessed using skewness and kurtosis values. Standard errors of skewness and kurtosis values were calculated according to Palmer (1994). Tests of significance for these moments were calculated following Sokal and Rohlf (1995, pp. 174-175).

Measurement Error

Measurement error is particularly important to address in studies of FA because real differences between sides are often very small, and are easily overwhelmed by measurement

error (Palmer 1994; Merila and Bjorklund 1995). Two replicate counts were conducted on a sample of 28 flies. After the first count, flies were stored frozen for three days, and a second set of counts was made in an order different from the first and which was unknown to the observer. Counts were analyzed using a two-factor ANOVA (Palmer 1994), in which individual and side were entered as factors. In this ANOVA a significant individual*side interaction indicates that non-directional asymmetry (i. e., FA and/or antisymmetry) is significantly larger than measurement error (Palmer 1994). Variation due to measurement error was 0.0926 (mean square $(MS)_{error}$), which was 2.1% of the individual x side MS (4.30). In addition, the individual x side interaction was strongly significant ($F_{26, 54}$=46.44, $p<0.0001$). Thus, variation due to measurement was negligible compared to actual FA.

Genotypic and environmental components of variation

A total of 22 lines (11 per site) were characterized in respect to FA, trait size and thorax length. For each isofemale line, the sample consisted of approximately ten individuals (full sibs) of each sex per bottle, with two bottles per line; the use of replicate bottles permitted environmental effects to be partitioned (Polak and Starmer 2001).

To test for site, line, and bottle effects on FA traits, a nested ANOVA was performed, in which bottle was nested within line, which in turn was nested within site. ANOVAs were performed on unsigned (|R-L|) FA values. Means of unsigned values provide an estimate of the variation in the FA distribution, whereas means of signed FA values should always be zero (Palmer 1996). Bottle and line were entered as random factors, so that F-statistics for site and line effects were computed using line(site) and bottle(line(site)) denominator mean squares, respectively. Because all unsigned FA variables, with the exception of PFA2, were

significantly correlated with trait size (Table 1), trait size was entered as a covariate in all ANOVAs on FA variables.

Because ANOVAs revealed significant line effects (see Results), lines with highest mean FA are referred to as "high-DI" lines, whereas lines with the smallest mean FA values are "low-DI" lines. These lines were used in subsequent tests of male reproductive vigor (see "Effects of Line and FA on Reproductive Traits", below).

To test whether line-specific FA values persisted across generations in the laboratory, I used four lines: Data from four generations were available for two lines, whereas data from six generations were available for the other two lines. An ANOVA was performed for each FA trait, with line and generation entered as factors, and trait size entered as the covariate. Variation due to bottle and site were not considered in this analysis, as these two factors were found to have no effect on FA (see Results). All statistical analyses were performed using the GLM procedure of SAS (1990).

Heritability and genetic correlations among traits

Prior to heritability calculations, variation due to trait size was removed from all FA traits by regressing absolute FA (|R-L|) on trait size. Log10 (Y+3)-transformed residuals from these analyses were then entered into model 3 of the PC-2 version of the mixed-model least-squares and maximum likelihood program (Harvey 1990) to estimate variance components between and within lines. Site and line within site were entered as factors. Heritability (h^2, the intraclass correlation) and its standard error (s.e.) were calculated according to Hoffmann and Parsons (1988), but corrected by deleting the square of the

denominator in their equation for the standard error of h^2 (e.g., Loeschcke *et al.* 1999). This results in the standard error of the intraclass correlation (h^2) being multiplied by: $(2/(1-(h^2/2))$. Genetic correlations between FA traits and trait size were also calculated using model 3 of the PC-2 version of the mixed-model least-squares and maximum likelihood program (Harvey 1990).

Effects of Line and FA on Reproductive Traits

I evaluated whether FA in sternopleural bristles at either the level of isofemale lines or individual males was associated with components of male mating success: copulation rate (number of copulations in a two-hour observation period), mating latency (time elapsed from introduction of females to first copulation), copulation duration (time elapsed between onset and termination of copulation), and fertility (number of larvae produced). For males mating more than once, an average value was used for copulation duration. I also tested whether FA per se (the "direct involvement" hypothesis) as opposed to underlying attributes such as vigor (the "underlying vigor" hypothesis) was responsible for any uncovered associations between FA and performance.

Choice of Lines. - Test lines were chosen on the basis of FA2 and PFA2, so that both numerical FA (NFA, reflecting the difference in numbers of bristles on the two sides of the sternopleuron) and positional FA (PFA, reflecting differences in the placement of the bristles on the two sides of the sternopleuron) were represented in the choice of experimental lines (males within lines were selected on the basis of PFA2 values only, see below). Both FA2 and PFA2 varied significantly among lines (see Results). The correlation between FA2 and

PFA2 across individual males was highly significant ($r=0.37$, $p<0.0001$, $n=886$ males). The correlation between the means of these two variables across lines was also significant and positive ($r=0.65$, $p=0.001$, $n=22$ lines); means used here are least square means from nested ANOVAs testing for site, line, and bottle effects on FA traits, and included trait size as a covariate (see "Genotypic and Environmental Components of Variation" in Methods section above).

Lines were ranked in ascending order of FA2 means, then in ascending order of PFA2 means. Ranks were averaged for each line, and the two lines with the lowest average ranks (FNY16 and FNY1) were chosen as "low-DI" lines (lines L1 and L2). Similarly, the two lines with the highest average ranks (IKY24 and FNY13) were chosen as "high-DI" lines (lines H1 and H2). Least-square means (corrected for trait size) for these four test lines are presented in Figures 2 and 3.

Choice of Males. - Males were then chosen from the above lines for the purpose of evaluating whether FA *per se*, as opposed to underlying vigor, was responsible for any associations uncovered between FA and mating success components. My approach was predicated upon assaying performance of males originating from "high-DI" and "low-DI" lines that expressed either similar or different levels of FA.

For one set of experiments (set A), males from each high line exhibiting the 50% highest PFA2 (high FA males) and males from each low line exhibiting the 50% lowest PFA2 (low FA males) were tested in mating trials (see Components of Mating Success, below). One to two days before each trial, bristles were counted in 40-50 males (per line) from which test males were then chosen. Set A consisted of five trials, each performed on a

different day. During each trial, an approximately equal number of high and low FA males were assayed. A total of 208 individuals were tested in this set.

In set B (three trials), males exhibiting signed PFA2 values between -0.25 and +0.25 units were contrasted. Thus, set B contrasted males within a similar range of FA values, but from different genetic lines of origin. The above boundaries were chosen by counting bristles for 120 males (60 from each line) up to two days before the experiment, and choosing individuals whose signed PFA2 values fell between +/- 0.75 standard deviations from the low line mean (mean of signed values, which was therefore near zero). These same boundaries (+/- 0.25 PFA2 units) were used in all subsequent trials of this set. Each trial was performed on a different day, and an approximately similar number of "high-DI" and "low-DI" males were assayed per trial. A total of 178 individuals were assayed across the three trials of set B.

Components of male mating success. - All flies used in mating experiments were between 11 and 14 days old, and males and females used on the same day differed in age by no more than one day. All females were collected from general stocks as virgins and held at densities of ten flies per vial in single-sex vials containing agar medium and 10-12 mg yeast. Females were moved to fresh vials every 1-2 days. Males from "high-DI" and "low-DI" lines were collected within 24 hours of eclosion and held in densities of 8-11 flies per agar vial without live yeast. Males were moved to fresh vials every 2-3 days.

All mating assays began within approximately 30 minutes after incubator lights were turned on. Males were placed individually into clean shell vials without food. After 15 min of acclimation, two virgin females were gently introduced into each vial using an aspirator.

Two females were used as a precaution in case one female was not receptive, for example, because of any subtle injury that was not apparent, or because of unknown causes. Vials were each observed by 2-3 observers for exactly 2 h. During this time, the beginning and ending times of all copulations were recorded. Immediately following each copulation, both females were removed from the vial and replaced with two fresh, virgin females. Females that had been in vials in which copulations occurred were retained in empty vials until the completion of the experiment to discourage oviposition. Females were eventually transferred to an oviposition substrate so that male fertility could be estimated (see Male Fertility, below). Thorax length of all males was measured, using an ocular micrometer of a dissecting microscope, after the mating experiments had ended.

Data analysis. - In a first general test, I evaluated the relationship between FA and performance values across individual males using data pooled across lines. I also examined whether components of mating success differed among isofemale lines (n = 4 lines).

Chi-square tests were used to detect any dependence between line and whether or not a male copulated, as well as between line and whether a male did not mate, mated once, or mated more than once. Because PFA2 was the only FA trait to show a significant isofemale line heritability (see Results), I examined the relationship between this variable and four components of mating success: copulation rate, mating latency, copulation duration, and male fertility. Multiple regression analysis was used, with thorax length included as a covariate.

To discriminate between the "direct involvement" and the "underlying vigor" hypotheses, ANOVAs were used to analyze variation in copulation rate, mating latency and

copulation duration. Data from both sets (A and B) were included in these analyses. Factors included in these analyses were set, trial(set), line and the line*set interaction. Thorax length was entered as a covariate. Because only lines H1 and L1 were assayed in set B, only data from these two lines were used in this analysis, so that the two sets would be directly comparable. Predictions generated by the "direct involvement" and "underlying vigor" hypotheses are presented in Table 3.

Male fertility.- Following mating assays in set A, all females exposed to a given male were placed in a culture vial such that all larvae in a given vial would be the offspring of the one male. Females were moved to fresh culture vials every two days. Three days after the females were removed from a vial, larvae were counted. This process continued until either no larvae were found in two consecutive vials or all females from a particular male had died. To count larvae, a saturated sucrose solution was stirred into the culture medium, and this mixture was poured into a petri dish. Larvae floated to the surface and were counted. Male fertility was defined as the number of larval offspring produced by a given male.

Effects of male starvation.- In all set A trials except for the first one, half of the males from each line were starved for 11-12 hours prior to the start of the experiment. All males in the first trial of set A were starved. Starvation was accomplished by placing males in individual vials containing only a moist tissue. Males that were not starved were held in individual agar-molasses food vials prior to experiments.

Results

Distributions

Descriptive statistics for signed FA traits examined in the heritability experiment are listed in Table 4. When the data from both sites are pooled, FA1 and overall FA means differ significantly from zero. In the case of FA1, the significant variation appears to be due primarily to the IKY sample. The IKY sample also had a non-zero mean for both PFA1 and PFA2, although this significance is lost when the two sites are pooled. While none of the FA distributions exhibited significant skewness when both sites were pooled, the FNY sample had a significantly left-skewed PFA1 distribution, and the IKY sample had a significantly right-skewed PFA2 distribution. Significant leptokurtosis was detected in all PFA2 distributions. Leptokurtic distributions have a concentration of values at the mean and/or in the tails of the distribution, far from the mean, resulting in a more narrow peak than a normal distribution (Zar 1999, p. 69). It has been suggested that leptokurtosis indicates significant genotypic variation (Palmer and Strobeck, 1986; Gangestad and Thornhill, 1999), and has been detected in PFA2 distributions in *Drosophila falleni* (Polak and Starmer 2001). Leptokurtosis was also significant for the IKY sample in both FA2 and PFA1, and in the FNY sample in overall FA.

Means and Heritability

Trait size-corrected least-square FA means for each line within each site are listed in Table 5. A nested ANOVA detected no effect of site or bottle on any of the five FA variables (Table 6). No line effects were evident for FA1, overall FA, or PFA1. However, significant line effects were detected for both FA2 and PFA2.

I tested the null hypothesis of no difference in FAs across generations. After removing line effects, no significant differences were detected among generations ($p>0.20$ for all traits, Table 7), indicating stability of FA differences detected previously, and strengthening the hypothesis of a genetic basis for FA.

Because FA2 and PFA2 showed significant between-line variation, heritability (h^2) values were calculated for these traits. PFA2 was the only trait for which its heritability estimate (0.102 +/- 0.046 SE) was significant (Table 8).

Effects of Line and FA on Reproductive Traits

I evaluated the relationships between FA and components of mating success by pooling data from all four lines in set A and both lines in set B. As starvation was not found to be a significant factor in any analyses performed for set A (copulation rate: $F= 0.47$, $DF= 1$, $p= 0.50$; mating latency: $F= 1.01$, $DF= 1$, $p= 0.32$; duration: $F= 0.53$, $DF= 1$, $p=0.57$; fertility: $F= 1.15$, $DF= 1$, $p= 0.29$), males were not starved in set B, and the effects of male starvation will not be addressed further.

Copulation Rate. – Copulation rate was defined as the number of copulations initiated in a two-hour observation period. In sets A and B combined, 96 (24.9%) of the 386 males tested did not mate, while 290 (75.1%) did mate. One hundred seventy-six males (45.6% of all males) mated only once, 106 (27.5%) mated twice, and 5 (1.3%) mated three times. Differences in numbers of copulations by males in high and low lines are illustrated in Figure 4. Whether or not a male mated was not dependent on line ($x^2=0.1252$, $DF=1$, $0.75>p>0.50$).

Whether a male did not mate, mated once, or mated more than once was also independent of line (x^2=2.2068, DF=2, 0.50>p>0.25).

Mean number of copulations per male was 1.03 (s.e.=0.0572, n=191) for the high lines and 1.10 (s.e.= 0.0569, n=194) for the low lines. An ANOVA testing for effects of body size, trial, set, and line revealed that the difference between lines was not significant (Table 9). Finally, a regression of the $\log_{10}$ of copulation rate on FA (all data from sets A and B) had a nonsignificant slope (slope=-0.0337, t=-1.67, p=0.10). Thus, copulation rate does not appear to differ between high- and low-DI males.

Mating Latency.- Mating latency was defined as the time elapsed between the introduction of females to a male's vial and the beginning of the first copulation. The high-DI lines had a significantly longer average mating latency (22.9 min, s.e.= 1.9 min, n= 145) than did the low-DI lines (15.9 min, s.e.= 1.9 min, n= 144) (Table 9). The high-DI lines showed a 67.5% greater mating latency compared to the low-DI lines. In addition, a multiple regression of mating latency on PFA2 and body size (using all data from sets A and B) revealed a positive and significant slope for mating latency and PFA2 (slope=7.56, t=2.57, p=0.01; Figure 5a), suggesting that males with higher FA take longer to begin mating than males with lower FA. A negative, though non-significant slope was detected for mating latency and body size (slope= -40.25, t=-1.87, p=0.06).

Duration.- Copulation duration was defined as either the amount of time a male spent in copula or, for males mating more than once, the average duration of a copulation. The high lines had a significantly longer average duration (63.2 min, s.e.= 2.2 min, n=141) than did the

low lines (50.1 min, s.e.= 2.2 min, n=141) (Table 9). Multiple regression of all data from sets A and B combined revealed a positive, significant relationship between duration and PFA2 (slope=17.17, t=4.78, p<0.0001; Figure 5b), and a positive, but non-significant relationship between duration and body size (slope= 32.55, t= 1.25, p=0.21). These results suggest that males with higher FA mate longer than males with lower FA.

Fertility.- Fertility was defined as the total number of larvae produced by each male. The two high lines combined produced an average of 214 larvae per male (s.e.= 42.1, n= 30), while the two low lines combined produced an average of 182 larvae per male (s.e.= 49.4, n= 23). This difference is not significant (Table 9). A regression of number of larvae produced on PFA2 was also insignificant (slope=21.72, t=0.33, p=0.74), suggesting that offspring production did not differ between high- and low-DI males in this experiment.

"Direct involvement" versus "underlying vigor" hypotheses

Table 3 presents the predictions generated from each hypothesis. In set A, males at the extremes of the genotypic and phenotypic FA distribution were contrasted, so both hypotheses predict significant differences between lines. In set B, I contrasted males from high- and low-DI genotypic lines while experimentally holding the FA phenotype constant (near zero), so that only the underlying vigor hypothesis predicts differences between lines in reproductive traits. Because lines H2 and L2 were not used in set B, only data from lines H1 and L1 are presented here, in an effort to make as valid comparisons as possible between the two sets. In set B, mean PFA2 for line H1 was 0.13 (s.e.=0.011, n=82), and for line L1 was 0.10 (s.e.=0.011, n=97). An ANOVA testing for effects of line, trial, and line*trial, with tot

1,2 entered as a covariate, confirmed that the difference in PFA2 between lines was not significant (F=3.70, DF=1, p=0.06). As mating latency and copulation duration were the only traits to differ between lines in the pooled data set, only these traits will be addressed here.

Mating Latency.- The high line had a longer mating latency than did the low line in both sets A and B (Figure 6a). In set A, the high line shows a 58.7% greater mating latency than the low line, while in set B the high line has a 30.2% greater mating latency. An ANOVA of data from sets A and B testing for effects of body size, set, trial, and line revealed a significant difference in mating latency between the two lines, and an insignificant line x set interaction (Table 10), suggesting that high-DI males have longer mating latencies than low-DI males, even when their phenotypes are similar. However, Tukey-Kramer post-hoc tests did not detect a significant difference between the two lines in either set A or B separately (set A: p=0.23, n= 93; set B: p=0.52, n= 139). Because considering the two sets separately reduced the sample size, thus the power of the test, the lack of significance in the post-hoc tests may simply be a result of reduced power. Nevertheless, in both sets, latency was greater among males from the high-DI lines.

Duration.- Line H1 showed a 19.4% increase in duration over line L1 in set A, and a 28.1% increase in set B (Figure 6b). An ANOVA testing for effects of body size, set, trial, and line revealed a significant difference between the two lines overall, but an insignificant line x set interaction (Table 10). This suggests that high-DI males have longer copulations than do low-DI males, even when their FA phenotypes are similar. Tukey-Kramer post-hoc tests

revealed that the difference in duration least square means between the high and low lines in set A (differing phenotypes) alone was not significant ($p=0.11$, $n=93$). The difference between the two lines was significant in set B (similar phenotypes) ($p<0.01$, $n=139$).

Discussion

Genetic Variation and Heritability of Fluctuating Asymmetry

The present study detected significant genetic variation in one form of NFA and one form of PFA in sternopleural bristles of *D. immigrans*. Only PFA2 had a significant isofemale line heritability, with a value of 0.102 (s.e.= 0.046, $p=0.02$). However, because the isofemale line approach cannot separate additive genetic effects from other forms of genetic activity, such as dominance or epistasis (Bubily *et al.* 2000), these heritability values are probably an overestimate of the true "narrow sense" heritability (i. e., the ratio of additive genetic variance to total phenotypic variance) (Falconer and Mackay 1996). This heritability estimate should not be interpreted as heritability of DI. As FA is due to both DI and nongenetic sources of variance, heritability of DI is typically higher than heritability of FA (Fuller and Houle 2003). Nevertheless, the combined results of ANOVA and heritability calculations indicate that significant genetic variation exists for sternopleural bristle FA2 and PFA2.

My heritability estimate for PFA2 is within the range of heritability estimates in other studies, though most studies of the heritability of FA have not found it to be significant (Fuller and Houle 2003). PFA2 was examined with an iso-female line approach in *Drosophila falleni* (Polak and Starmer 2001). The magnitude of h^2 for PFA2 in this case was 0.21- approximately twice that reported in the present study. While neither study reported

significant heritability for FA2, the magnitude of h^2 for this form of FA was again roughly double in *D. falleni* (0.024, Polak and Starmer 2001), when compared to *D. immigrans* (Table 8). A full-sib analysis of FA in the tibia of winter moths (*Operophtera brumata*) (Van Dongen et al. 1999), found an average heritability of only 0.05 to 0.07 for FA in three pairs of tibia. While this value was significant, heritabilities for FA in each set of tibia individually ranged from 0 to 0.04 and were not significant. A meta-analysis (Møller and Thornhill 1997) also detected a small (h^2=0.27), but significant value for heritability of FA across multiple characters and species. However, this analysis has been criticized for reasons outlined above (see Introduction), and its results should thus be regarded with some caution. Heritability of FA has also been addressed by Chapman and Goulson (2000), who conducted parent-offspring regressions of wing length FA in house flies (*Musca domestica*). While heritabilities varied from -0.00082 (regression with mean parental FA) to 0.070 (regression with paternal values), none of the regressions were significant. In addition, FA varied significantly due to temperature, with the highest mean FA occurring at the lowest temperature used (15°C), suggesting that FA may have been more affected by environmental factors than by genotype in this study.

In comparison to FA traits, heritabilities reported for other types of traits are often higher. For example, in a meta-analysis of data from a variety of taxa, Mousseau and Roff (1987) reported heritabilities of 0.46 for morphological traits (e. g., sizes of metric characters), 0.26 for life history traits (e. g., fecundity, rates of development, life span), and 0.30 for behavioral traits (e. g., alarm reactions and activity levels). These results follow the same pattern as those reported for *Drosophila* (Roff and Mousseau 1987), in which a heritability of 0.32 was reported for morphological traits, 0.18 for behavioral traits, and only

0.12 for life history traits. These data are interpreted as support for the prediction that traits more closely related to fitness will have lower heritabilities (Mousseau and Roff 1987). It follows, therefore, that PFA2, with a significant iso-female line heritability of 0.10, may also be closely related to fitness. The heritability of FA may also be low due to random variation within individuals or the environment, both of which can obscure true differences in DI among individuals and/or lines.

Genetic Correlations Between Fluctuating Asymmetry and Trait Size

The FA-sexual selection hypothesis predicts that if FA of secondary sexual characters truly reflects quality, then individuals that are able to produce larger traits should also be able to develop the most symmetric traits (Tomkins and Simmons 2003, Møller and Swaddle 1997). Thus, in sexually selected characters, trait size should be negatively correlated with FA. In the present study, the genetic correlation between PFA1 and trait size was negative and significant (Table 2), suggesting that individuals with more sternopleural bristles exhibit more symmetric placement of these bristles. The genetic correlation between PFA2 and trait size was also negative, but was not significant. These results, combined with the significant relationships between PFA2 and both mating latency and copulation duration, suggest that PFA is involved in sexual selection in *D. immigrans*. The genetic correlation between FA1 and trait size was positive and significant (Table 2), suggesting that individuals with more anterior sternopleural bristles exhibit more asymmetry in these bristles. FA2 exhibited a positive, but nonsignificant genetic correlation with trait size. These results suggest that NFA is not correlated with individual quality (if FA indicates quality, the FA-sexual

selection hypothesis predicts that FA and trait size should be negatively correlated, see Introduction).

The results discussed here agree with those discovered by both Scheiner *et al.* (1991) and Polak and Starmer (2001). Scheiner *et al.* (1991) also detected a positive (though nonsignificant) genetic correlation between trait size and FA in the anterior bristles. Polak and Starmer (2001) detected a positive correlation between NFA and trait size and a negative correlation between PFA and trait size, though this study was conducted on *Drosophila falleni.*

Fluctuating Asymmetry, Developmental Instability, and Mating Success

The present study evaluated two hypotheses relating to FA, DI, and mating success. The first ("Direct Involvement" hypothesis) predicts that any differences in mating success between males from high- and low-DI lines are a result of their differing phenotypes (FA). The second hypothesis ("Vigor" hypothesis) predicts that males from low-DI lines will have higher mating success than males from high-DI lines due to increased vigor and mating propensity associated with their genotype. Thus, according to the vigor hypothesis, high- and low-DI males should differ in mating success regardless of their phenotypes.

Copulation Rate.- Copulation rate (number of copulations in a two-hour observation period) does not appear to differ between males from the high- and low-DI lines used in this experiment. Thus, it is possible that high- and low-DI males simply do not differ in copulation rate. Another possibility is that high- and low-DI males do differ in copulation rate, but the divergence in DI between the two lines used was not great enough to result in a

difference in copulation rate. While I defined "high" and "low" FA as 0.75 standard deviations from the mean, another study (Woods *et al.* 2002) defined extreme individuals as being two or more standard deviations from the mean. A third possibility is that the difference between the two lines may only be evident under more stressful conditions (see Leung and Forbes 1997).

Similar to my results, Polak (1997) found no significant differences in either NFA or PFA of sternopleural bristles between male *Drosophila nigrospiracula* collected both in copula and singly in the field. Bourget (2000) also found no relationship between sternopleural bristle FA and male mating success in *D. melanogaster*.

In contrast with my results, Møller (1996) found that asymmetry in wings of unmated male house flies (*Musca domestica*) collected in the field was significantly higher than in mated males. Møller (1996) also discovered that, when given a choice between two males, female house flies (*M. domestica*) chose to mate first with males having more symmetric wings and tibia, which could indicate that symmetric males are chosen more often as mates in the wild. Also contrasting with my results, male tarsus asymmetry in black grouse (*Tetrao tetrix*) has been significantly negatively correlated with both number of copulations and centrality of location on the lek (Rintamaki *et al.* 1997). The authors suggest that males with asymmetric tarsi are less successful at competing for central locations due to decreased vigor, and that their lower copulation rate is due to females mating more frequently with males in the center of the lek. However, it seems that in this case the decreased vigor exhibited by asymmetric males during competition could actually be the result of asymmetry of the tarsi, as opposed to asymmetry being an indication of decreased vigor. That is, perhaps asymmetric tarsi prevent some males from competing as efficiently as others.

Mating Latency.- In the present study the high-DI line had a longer mating latency (time elapsed to first copulation after introduction of females to vial) than did the low-DI line both when FA phenotypes differed (set A) and when they were similar (set B). An ANOVA of all data revealed a significantly longer mating latency for the high-DI line, as well as an non-signficant line x set interaction. As a longer mating latency reduces opportunities for future copulations, I interpret this as a disadvantage for the high-DI line, and as support for the vigor hypothesis.

Duration.- High-DI males had significantly longer average copulation durations than did low-DI males both when FA phenotypes differed and when they were similar. Longer copulation durations may be advantageous, as they allow males time to transfer more sperm to females. As *Drosophila* often mate multiply, a male that transfers more sperm to a given female may have a better chance of fertilizing many of her eggs than males that transfer fewer sperm (sperm mixing and sperm loading in Simmons 2001). Alternatively, high-DI males may simply take longer to transfer an ejaculate similar in size to that of low-DI males.

Given that the high- and low-DI males did not produce a significantly different number of offspring, I interpret longer copulations as a disadvantage for high-DI males, and as support for the vigor hypothesis. Because high-DI males take longer to both begin copulation (mating latency) and complete copulation, they should have fewer opportunities to mate on any given day, which may translate to a reduction in lifetime fitness. Alternatively, it is possible that longer copulation durations with each female allow high-DI males to compensate for their longer mating latencies and/or low genetic quality. More reliable

feritilty data (see below), as well as greater knowledge of sperm usage in *Drosophila immigrans*, would aid in evaluating this hypothesis.

Fertility.- I found no significant difference in number of larvae produced between high- and low-DI males. Thus, it is possible that high- and low-DI males do not differ in fertility. However, I am not confident that my fertility data are reliable. In many cases one or more females from a vial died before zero larvae were found in two consecutive vials (see Methods). In order to eliminate the possibility that differences in male fertility were due to female mortality, I eliminated data from males for which any females had died, thus reducing the sample size to 53 males, and reducing the power of the ANOVA. In addition, a correlation between fertility and number of copulations was not significant ($r=-0.018$, $p=0.90$, $n=53$). As more copulations would seem to lead logically to more offspring, this non-significant correlation casts doubt on the validity of the fertility data.

Regardless of the quality of the fertility data in the present study, some other studies have also failed to find a relationship between FA and number of offspring produced. Bourget (2000) found no relationship between sternopleural bristle FA in *D. melanogaster* and a combined measure of fecundity and adult survivorship. Woods *et al.* (2002) also compared high- and low-FA (based on a combined measure of several characters) *D. melanogaster* and obtained similar results. Møller (1997), on the other hand, reports a negative relationship between FA and traits such as clutch size and number of clutches produced in 16 of 17 studies reviewed. However, the only study examined involving *Drosophila* did not test for measurement error, and therefore should be regarded with some caution, as true FA may have been obscured by measurement error. Other studies examined

by Møller (1997) tested humans, plants, birds, and other species, and therefore may not be directly comparable to the present study.

Summary and General Conclusions

The findings of my research can be summarized as follows:

1. A total of twenty-two lines derived from two populations of *D. immigrans* were characterized in terms of the degree of FA in sternopleural bristles. Only one form of FA, namely PFA2 (which reflects the difference in the placement of bristles on the two sides of the sternopleuron), showed significant isofemale line heritability (h^2= 0.102, s. e.= 0.046). This suggests that DI may be heritable, as the heritability of DI is typically higher than that of FA.
2. Two high-DI lines and two low-DI lines were compared with respect to four components of male mating success.
    a) High- and low-DI lines did not differ significantly in average copulation rate. Whether a male did not mate, mated once, or mated more than once was independent of line.
    b) The high-DI lines had a significantly longer average mating latency (22.9 min, s.e.=1.9 min, n=145) than did the low-DI lines (15.9 min, s.e.=1.9 min, n=144). In addition, a multiple regression revealed a positive and significant effect of PFA2 on mating latency, even when body size was included. These results are interpreted as a disadvantage for high-DI males, as they will have fewer opportunities to mate.

c) The high-DI lines had a significantly longer average copulation duration (63.2 min, s.e.=2.2 min, n=141) than did the low lines (50.1 min, s.e.=2.2 min, n=141). In a multiple regression of duration on PFA2 and body size, the relationship between duration and PFA2 was positive and significant. These results are interpreted as a disadvantage for high-DI males, as they will have fewer opportunities to mate.

d) No significant difference was detected in fertility (number of larvae produced) between high- and low-DI lines.

3. Two sets of males from one high- and one low-DI line were compared with respect to two components of mating success. In an attempt to differentiate between FA and underlying vigor as the cause of the differences between the two lines, set A used the highest-FA males from the high-DI line and the lowest –FA males from the low-DI line, while set B used males with similar FA (near zero) from both lines.

a) The high-DI line had a longer average mating latency than did the low-DI line, even when FA phenotypes were similar. Though post-hoc tests did not detect significant differences between the lines when the two sets were considered separately, the overall analysis revealed a significant effect of line and a non-significant set x line interaction. Together, these results favor the vigor hypothesis.

b) The high-DI line had a longer average copulation duration than did the low-DI line in both sets. Post-hoc tests detected a significant difference between the two lines in copulation duration in set B (when FA phenotypes were similar). These results seem to favor the vigor hypothesis.

4. Genetic correlations between PFA and trait size were negative and, in the case of PFA1, significant. Genetic correlations between trait size and both FA1 and FA2 were positive, though only the correlation with FA1 was significant. Based on these results and an FA-sexual selection theory, PFA is likely to be involved in sexual selection, though this is unlikely for NFA.

5. Thus, regardless of their FA phenotypes, low-DI males began copulating sooner after exposure to a female than did high-DI males. Low-DI males also completed copulations more quickly than did high-DI males, and thus may have more opportunities to mate throughout their lifetimes. These findings, together with significant heritability of FA, suggest that DI should be selected against in natural populations of *D. immigrans*.

References

Andersson, M. 1994. Sexual selection. Princeton University Press, Princeton, N.J.

Bourget, D. 2000. Fluctuating asymmetry and fitness in *Drosophila melanogaster*. Journal of Evolutionary Biology 13: 515-521.

Bubily, O. A., V. Loeschcke, and A. G. Imasheva. 2000. Effect of stressful and nonstressful growth temperatures on variation of sternopleural bristle number in *Drosophlia melanogaster*. Evolution 54: 1444-14449.

Cadee, N. 2000. Genetic and environmental effects on morphology and fluctuating asymmetry in nestling barn swallows. Journal of Evolutionary Biology 13: 359-370.

Chapman, J. W. and D. Goulson. 2000. Environmental versus genetic influences on fluctuating asymmetry in the house fly, *Musca domestica*. Biological Journal of the

Linnean Society 70:403-413.

Darwin, C. R. 1871. The Descent of Man and Selection in Relationship to Sex. John Murray, London.

Falconer, D. S. and T. F. Mackay. 1996. Introduction to Quantitative Genetics, 4th edition. Longman, London.

Fuller, R. C. and D. Houle. Inheritance of Developmental Instability. 2003. In: Polak, M. Developmental Stability: Causes and Consequences. Oxford University Press, New York, in press.

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

Gibson, G., and G. Wagner, 2000. Canalization in evolutionary genetics: A stabilizing theory. BioEssays 22:72-380.

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

Harvey, W. R. 1990. User's guide for LSMLMN and MIXMDL PC-2 version. Copyright W. R. Harvey, Department of Dairy Science, Ohio State University, Columbus, OH.

Hoffmann, A. A., and P. A. Parsons. 1988. The analysis of quantitative variation in natural populations with isofemale strains. Genet. Sel. Evol. 20: 87-98.

Leung, B. and M. R. Forbes. 1997. Fluctuating asymmetry in relation to indices of quality and fitness in the damselfly, *Enallagma ebrium* (Hagen). Oecologia 110: 472-477.

Loeschcke, V., Brudgaard, J., and Barker, J.S.F. (1999) Reaction norms across and genetic parameters at different temperatures for thorax and wing size traits in *Drosophila aldrichi* and *D. buzzatii*. Journal of Evolutionary Biology 12:605-623.

Markow, T. A. and G. M. Clarke. 1997. Meta-analysis of the heritability of developmental stability: a giant step backward. Journal of Evolutionary Biology 10: 31-37.

Mather, K. 1953. Genetical control of stability in development. Heredity 7: 297-336.

Merila, J. and M. Bjorklund. 1995. Fluctuating asymmetry and measurement error. Systematic Biology 44(1): 97-101.

Møller, A. P. 1990. Fluctuating asymmetry in male sexual ornaments may reliably reveal male quality. Animal Behaviour 40: 1185-1187.

Møller, A. P. 1992. Female swallow preference for symmetrical male sexual ornaments. Nature (London) 357: 238-240.

---------. 1996. Sexual selection, viability selection, and developmental stability in the domestic fly *Musca domestica*. 1996. Evolution 50: 746-752.

---------. 1997. Developmental stability and fitness: A review. The American Naturalist 149: 916-932.

Møller, A. P., and A. Pomiankowski. 1993. Fluctuating asymmetry and sexual selection. Genetica 89: 267-279.

Møller, A. P., and Swaddle, J. P. 1997. Asymmetry, Developmental Stability, and Evolution. Oxford University Press, Oxford.

Møller, A. P. and R. Thornhill. 1997. A meta-anlaysis of the heritability of developmental stability. Journal of Evolutionary Biology 10: 1-16.

Morris, M. R. 1998. Female preference for trait symmetry in addition to trait size in swordtail fish. Proc R Soc London B 265:907-911.

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

Nijhout, H. F., and Davidowitz, G. 2003. Developmental perspectives on phenotypic variation, canalization, and fluctuating asymmetry. Pp. 3-13, in: Polak, M. (ed.) Developmental Instability: Causes and Consequences. Oxford University Press, New York.

Ostbye, K., S. A. Oxnevad, and L. A. Vollestad. 1997. Developmental stability in perch (*Perca fluviatilis*) in acidic aluminum-rich lakes. Canadian Journal of Zoology 75: 919-928.

Palmer, A. R. 1994. Fluctuating asymmetry analyses: a primer. Pages 335-364 in T. A. Markow, ed. Developmental instability: its origins and evolutionary implications. Kluwer, Norwell.

---------- 1996. Waltzing with asymmetry. Bioscience 46: 518-?.

Palmer, A. R. and C. Strobeck. 1986. Fluctuating asymmetry: measurement, analysis, patterns. Annual Review of Ecology and Systematics 17: 391-421.

---------- 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.

---------- 1997. Fluctuating asymmetry and developmental stability: heritability of observable variation versus heritability of inferred cause. Journal of Evolutionary Biology 10: 51-55.

Patterson, J.T. 1943. Studies in the genetics of *Drosophila* III. The Drosophilidae of the Southwest. The University of Texas Publication No. 4313, The University of Texas, Austin.

Polak, M. 1997. Ecotoparasitism in mothers causes higher positional fluctuating asymmetry in their sons: implications for sexual selection. The American Naturalist 149: 955-974.

Polak, M., editor. 2003. Developmental Stability: Causes and Consequences. Oxford University Press, New York.

Polak, M., Opoka, R., and Cartwright, I. 2002. Response of fluctuating asymmetry to arsenic toxicity: support for the developmental selection hypothesis. Environmental Pollution, 118:19-28.

Polak, M. and W. T. Starmer. 2001. The quantitative genetics of fluctuating asymmetry. Evolution 55: 498-511.

Polak, M., L. L. Wolf, and W. T. Starmer. 2001. Function of the mating plug in *Drosophila hibisci* Bock. Behavioral Ecology and Sociobiology 49: 196-205.

Radesater, T. and H. Halldorsdottir. 1993. Fluctuating asymmetry and forceps size in earwigs, *Forficula auricularia*. Animal Behaviour 45: 626-628.

Rintamaki, P. T., R. V. Alatalo, J. Hoglund, and A. Lundberg. 1997. Fluctuating asymmetry and copulation success in lekking black grouse. Animal Behaviour 54: 265-269.

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

Scheiner, S. M., R. L. Caplan, and R. F. Lyman. 1991. The genetics of phenotypic plasticity. III. Genetic correlations and fluctuating asymmetries. Journal of

Evolutionary Biology 4: 51-68.

Simmons, L. W. 2001. Sperm competition and its evolutionary consequences in the insects. Princeton University Press, Princeton.

Thornhill, R. and J. Alcock. 1983. The evolution of insect mating systems. Harvard University Press, Cambridge.

Tomkins, J. L., and L. W. Simmons. 1998. Female choice and manipulations of forceps size and symmetry in the earwig *Forficula auricularia* L. Animal Behaviour 56: 347-356.

Tomkins, J. L., and L. W. Simmons. 2003. Fluctuating asymmetry and sexual selection: paradigm shifts, publication bias, and observer expectation. In: Polak, M. Developmental Stability: Causes and Consequences. Oxford University Press, New York, in press.

Uetz, G. and E. Smith. 1999. Asymmetry in a visual signaling character and sexual selection in a wolf spider. Behav. Ecol. Sociobiol. 45: 87-93.

Van Dongen, S., E. Sprengers, C. Lofstedt, and E. Matthysen. 1999. Heritability of tibia fluctuating asymmetry and developmental instability in the winter moth (*Operophtera brumata* L.) (Lepidoptera, Geometridae). Heredity 82: 535-542.

Van Valen, L. 1962. A study of fluctuating asymmetry. Evolution 16: 125-142.

Wakefield, J., K. Harris, and T. A. Markow. 1994. Parental age and developmental stability in *Drosophila melanogaster*. In: Markow, T. A. Developmental instability: Its origins and evolutionary implications. Kluwer Academic Publishers.

Whitlock, M. C., and K. Fowler. 1997. The instability of studies of instability. Journal of Evolutionary Biology 10: 63-67.

Woods, R. E., C. M. Sgro, M. J. Hercus, and A. A. Hoffmann. 2002. Fluctuating asymmetry, fecundity and development time in *Drosophila*: is there an association under optimal and stress conditions? Journal of Evolutionary Biology 15: 146-157.

Zar, J. H. 1999. Biostatistical Analysis, 4th edition. Prentice-Hall, Inc., New Jersey.

Table 1. Correlation coefficients for unsigned fluctuating asymmetry versus trait size at IKY and FNY sites, separately and pooled.

		FA measure				
Site	N	FA1	FA2	FA	PFA1	PFA2
Cincinnati	440	0.110*	0.129**	0.162***	-0.086	-0.020
Syracuse	446	0.119*	0.144**	0.132**	-0.055	-0.030
Pooled	886	0.114***	0.135***	0.149***	-0.072*	-0.023

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

Table 2. Genetic correlations between FA traits and trait size. A genetic correlation could not be calculated for overall FA and trait size.

Trait	Genetic Correlation
FA1	0.726*
FA2	0.568
PFA1	-0.732**
PFA2	-0.348

* $p<0.05$; ** $p<0.01$

Table 3. Predictions in respect to components of male mating success generated by the "direct involvement" and "underlying vigor" hypotheses. "<" and "=" refer to relative quality of males.

	Predictions	
Hypothesis	Set A (FA different between high-and low-DI lines)	Set B (FA similar between high- and low-DI lines)
Direct involvement	High-DI males < low-DI males	High-DI males = low-DI males
Underlying Vigor	High-DI males < low-DI males	High-DI males < low-DI males

Table 4. Mean signed fluctuating asymmetry, skewness, and kurtosis of sternopleural bristles of *Drosophila immigrans*.

Trait	Site (n)	Mean (Variance)	p[1]	Skewness	Kurtosis
FA1	IKY (440)	-0.2477 (1.991)	0.0003	-0.116	0.037
	FNY(446)	-0.0538 (2.020)	0.42	-0.018	0.299
	Pooled (886)	-0.1501 (2.012)	0.0017	-0.064	0.175
FA2	IKY (440)	0.0545 (3.004)	0.51	-0.198	0.908**
	FNY(446)	-0.1233 (2.369)	0.09	0.141	0.154
	Pooled (886)	-0.0350 (2.689)	0.53	-0.038	0.615
Overall FA	IKY (440)	-0.1932 (4.903)	0.07	0.019	0.291
	FNY(446)	-0.1771 (4.290)	0.07	0.065	0.466*
	Pooled (886)	-0.185 (4.589)	0.01	0.039	0.374
PFA1	IKY (440)	-0.0269 (0.031)	0.001	0.094	0.427*
	FNY(446)	0.0054 (0.028)	0.49	-0.217**	0.054
	Pooled (886)	-0.0106 (0.029)	0.07	-0.065	0.200
PFA2	IKY (440)	0.0829 (0.386)	0.005	0.179*	1.395**
	FNY(446)	-0.0506 (0.369)	0.08	-0.084	1.484**
	Pooled (886)	0.0157 (0.382)	0.45	0.057	1.448**

[1] P evaluating H_o: mean=0

* $p<0.05$; ** $p<0.01$

Table 5. Least-square means and standard errors (SE) for unsigned fluctuating asymmetry measures in *Drosophila immigrans* by site and line within site. Values are taken from an ANOVA with trait size entered as a covariate. Sample size for each line is 40, with the exception of FNY18, in which n=44.

Site/Line	FA1		FA2		Overall FA		PFA1		PFA2	
IKY	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
11	1.014	0.15	1.431	0.17	1.759	0.22	0.140	0.017	0.403	0.066
13	1.229	0.15	1.122	0.17	1.711	0.22	0.129	0.017	0.32	0.065
15	0.894	0.15	1.569	0.17	2.031	0.23	0.138	0.018	0.489	0.068
17	1.109	0.15	1.311	0.17	1.734	0.22	0.146	0.017	0.554	0.065
19	0.986	0.15	1.422	0.18	1.808	0.24	0.158	0.018	0.603	0.071
20	1.097	0.15	1.087	0.17	1.855	0.22	0.125	0.017	0.461	0.065
22	1.227	0.15	1.474	0.18	1.763	0.23	0.137	0.018	0.380	0.068
23	1.024	0.15	1.282	0.17	1.555	0.22	0.140	0.017	0.459	0.065
24	0.982	0.15	1.483	0.17	1.522	0.22	0.160	0.017	0.683	0.066
25	1.046	0.15	1.148	0.17	1.508	0.22	0.146	0.017	0.451	0.066
26	1.293	0.15	1.035	0.19	1.509	0.24	0.107	0.018	0.218	0.070
Pooled	1.082	0.044	1.306	0.051	1.705	0.066	0.139	0.0051	0.458	0.020
FNY										
1	0.894	0.15	1.035	0.17	1.476	0.22	0.112	0.017	0.378	0.065
7	1.056	0.15	1.378	0.17	1.808	0.22	0.137	0.017	0.525	0.065
10	1.410	0.15	1.283	0.17	1.787	0.22	0.153	0.017	0.487	0.065
11	1.178	0.15	1.315	0.17	1.482	0.22	0.168	0.017	0.614	0.066
13	0.798	0.15	1.696	0.17	1.792	0.22	0.151	0.017	0.624	0.066
14	1.453	0.15	1.127	0.17	1.633	0.22	0.142	0.017	0.373	0.066
15	1.018	0.15	1.074	0.17	1.639	0.22	0.096	0.017	0.247	0.065
16	0.819	0.15	0.887	0.17	1.371	0.22	0.099	0.017	0.292	0.065
17	1.050	0.15	1.025	0.17	1.243	0.22	0.137	0.017	0.512	0.066
18	1.192	0.14	1.068	0.16	1.537	0.21	0.129	0.016	0.414	0.061
23	0.930	0.15	1.062	0.17	1.545	0.22	0.115	0.017	0.449	0.066
Pooled	1.072	0.044	1.177	0.050	1.574	0.066	0.131	0.0050	0.447	0.019

Table 6. Results of nested ANOVA for fluctuating asymmetry, testing for site, line, and bottle effects. Trait size was entered as a covariate. N=886 for all traits.

Trait	Source	DF	MS	F
FA1	Site	1	0.019	0.02
	Line(site)	20	1.24	1.64
	Bottle(line(site))	22	0.76	0.88
	Error	841	0.86	
FA2	Site	1	3.66	2.39
	Line(site)	20	1.53	2.88*
	Bottle(line(site))	22	0.53	0.47
	Error	841	1.13	
Overall FA	Site	1	3.80	3.35
	Line(site)	20	1.13	0.72
	Bottle(line(site))	22	1.58	0.82
	Error	841	1.92	
PFA1	Site	1	0.013	0.92
	Line(site)	20	0.014	1.89
	Bottle(line(site))	22	0.0077	0.68
	Error	841	0.011	
PFA2	Site	1	0.025	0.01
	Line(site)	20	0.57	3.70*
	Bottle(line(site))	22	0.15	0.91
	Error	841	0.17	

* $p<0.01$.

Table 7. Results of ANOVA testing for an effect of generation on line- and trait size- corrected FA traits. Analysis includes four lines and up to six generations.

FA trait	F	p
FA1	0.18	0.97
FA2	0.65	0.66
Overall FA	1.18	0.32
PFA1	0.99	0.42
PFA2	1.36	0.24

Table 8. Isofemale line heritability (h^2) and standard error (SE) estimates for fluctuating asymmetry (FA) in the sternopleural bristles of *Drosophila immigrans*. FA data are corrected for trait size. Bottle and site effects were not significant, therefore, data are not corrected for these factors. N=886.

Trait	h^2	SE	p (H_0: h^2=0)
FA2	0.011	0.019	0.28
PFA2	0.102	0.046	0.02

Table 9. Results of ANOVA on components of mating success from four lines in set A and two lines in set B. ANOVA tested for effects of trial, line (high or low), and experiment. Thorax length, an estimate of body size, was included as a covariate.

	Source	DF	MS	F
Copulation Rate	Length	1	0.22522	0.39
	Set	1	0.91920	1.78
	Trial (Set)	4	0.51687	0.90
	Line	1	0.40475	0.71
	Line*Set	1	1.0617	1.85
	Error	376	0.57350	
Mating Latency	Length	1	1389.70	2.90
	Set	1	5.156.08	0.03
	Trial (Set)	4	227.68	0.47
	Line	1	3212.94	6.70*
	Line*Set	1	319.64	0.67
	Error	280	479.41	
Duration	Length	1	1258.71	2.01
	Set	1	14415	11.48*
	Trial (Set)	4	1241.50	1.98
	Line	1	10850	17.32***
	Line*Set	1	121.18	0.19
	Error	273	626.41	
Fertility	Length	1	18148	0.36
	Set	1	11010	0.04
	Trial (Set)	2	231977	4.65*
	Line	1	11499	0.23
	Line*Set	1	580.13	0.01
	Error	46	252970	

*p<0.05; **p<0.01; ***p<0.001

Table 10. Results of ANOVA on components of mating success from sets A and B combined. Only lines H1 and L1 are represented. ANOVA tested for effects of trial, line, and set. Thorax length, an estimate of body size, was included as a covariate.

	Source	DF	MS	F
Mating Latency	Length	1	2021.28	4.07*
	Set	1	5.15	0.01
	Trial (Set)	4	489.55	0.99
	Line	1	2624.93	5.28*
	Line*Set	1	173.11	0.35
	Error	228	496.02	
Duration	Length	1	405.19	1.00
	Set	1	1587.89	1.59
	Trial (Set)	4	993.88	2.44
	Line	1	6074.45	14.94**
	Line*Set	1	76.58	0.19
	Error	223	406.70	

*$p<0.05$; **$p<0.001$

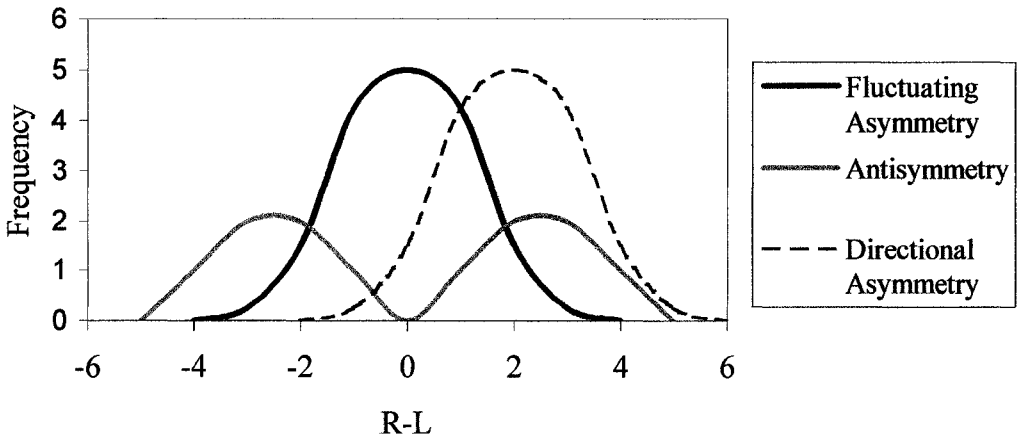

Figure 1. Examples of distributions of fluctuating asymmetry, antisymmetry, and directional asymmetry. Antisymmetry can be more subtle than is shown here, exhibting a platykurtic distribution.

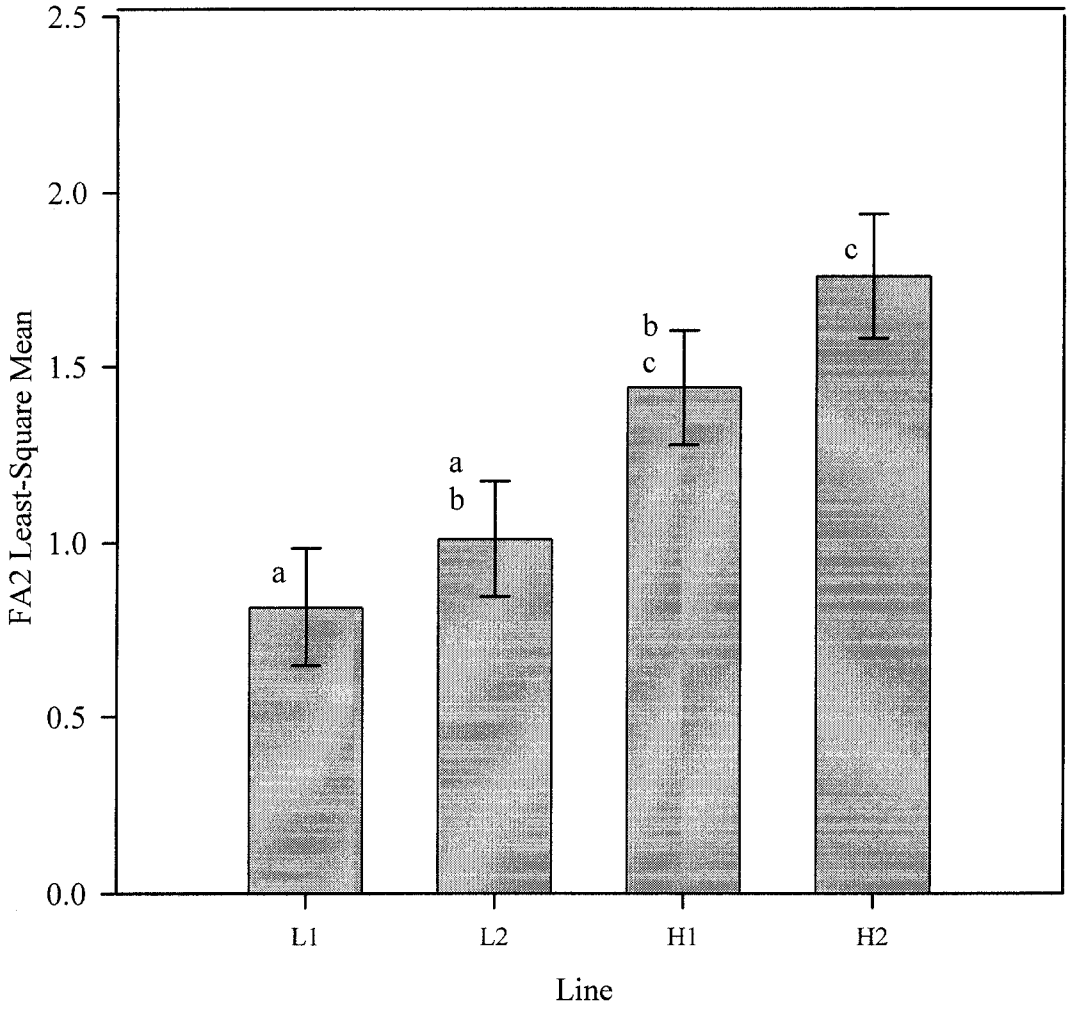

Figure 2. Least-square means of FA2 for high and low FA lines. Least-square means are taken from an ANOVA testing for an effect of line on unsigned FA2 values, with trait size entered as a covariate. Error bars indicate one standard error above and below the mean. Means sharing a letter do not differ significantly from each other.
N (for each mean)=40.

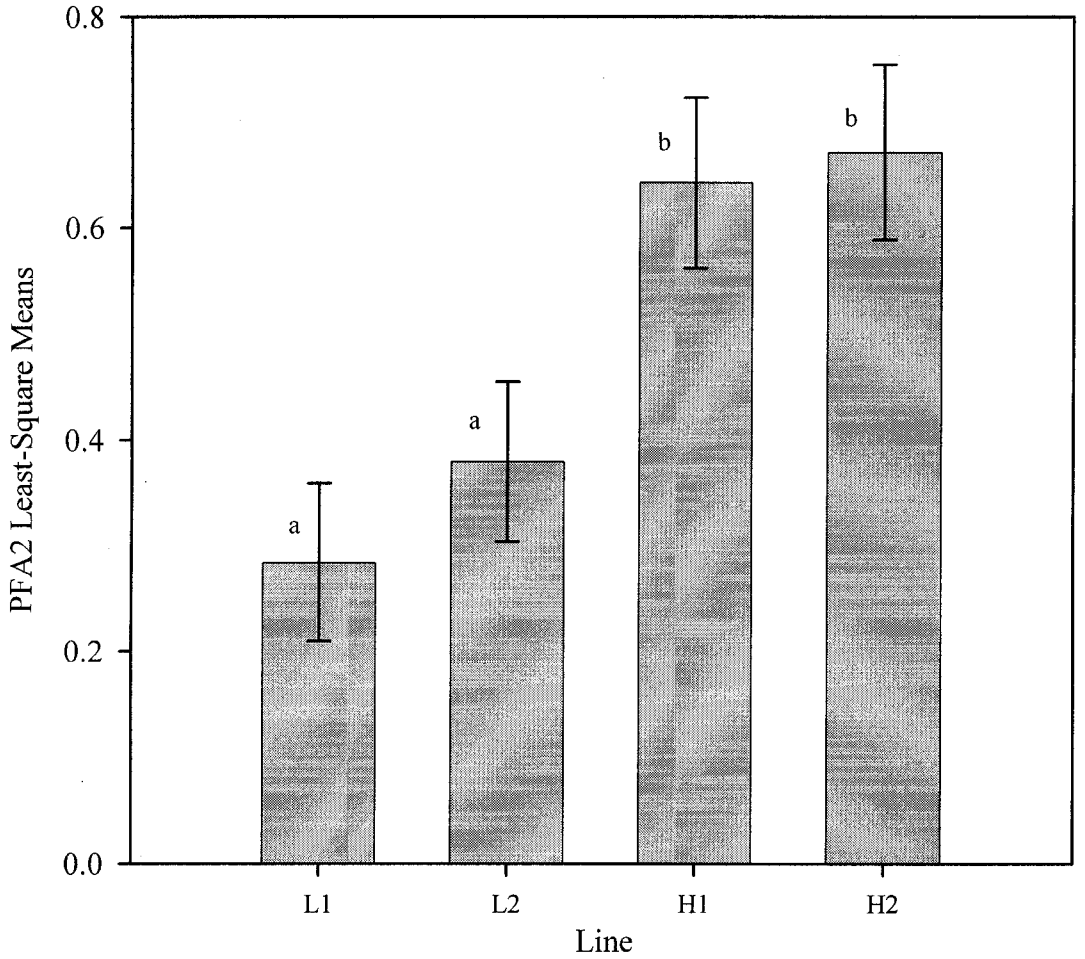

Figure 3. PFA2 least-square means for high and low FA lines. Least-square means are taken from an ANOVA testing for an effect of line on unsigned PFA2 values, with trait size entered as a covariate. Error bars indicate one standard error above and below the mean. Means sharing a letter do not differ significantly from each other.
N (for each mean)=40.

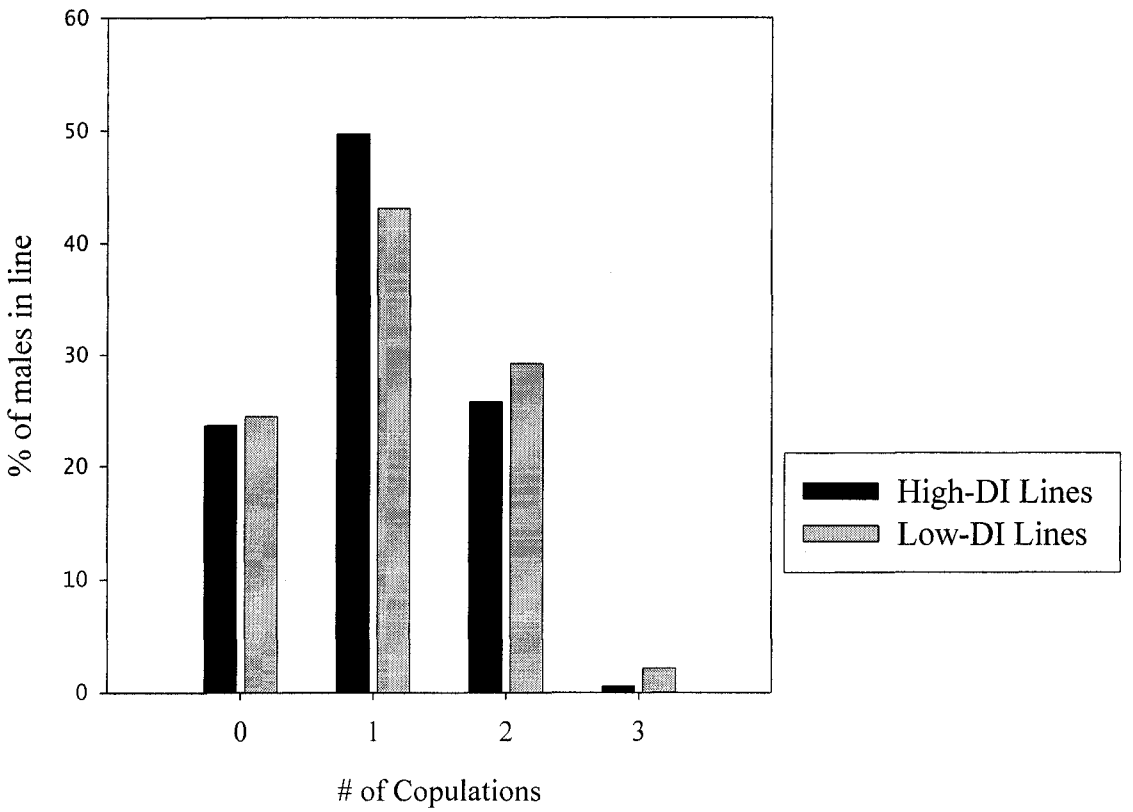

Figure 4. Number of copulations by males in two high- and two low-DI lines in sets A and B combined. N=386.

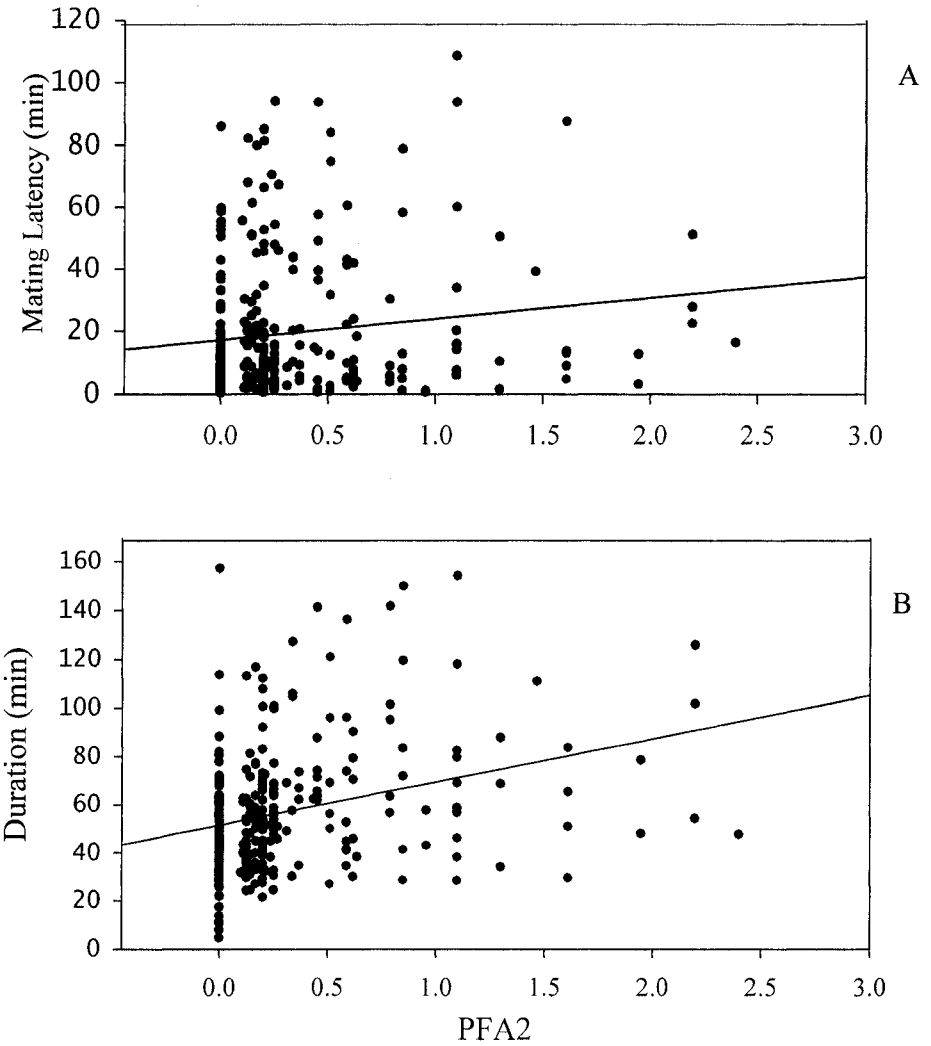

Figure 5. Mating latency (a) and copulation duration (b) plotted against PFA2 values. Data from two high- and two low-DI lines are included. Slopes of both lines are significant (mating speed: slope=7.56, t=2.57, p=0.01; duration: slope=17.91, t=5.05, p<0.001).

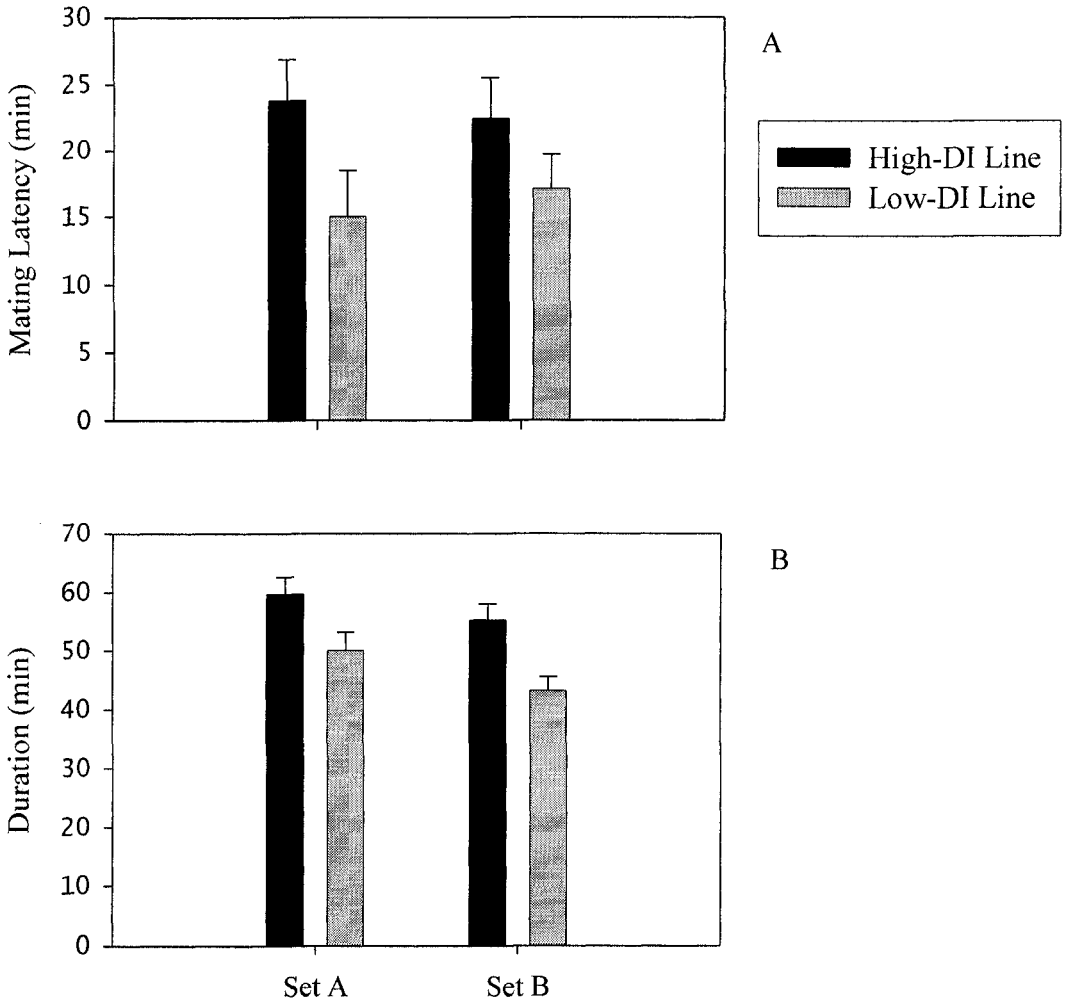

Figure 6. Average mating latency (a), and copulation duration (b) for males from lines H1 and L1 in sets A and B. Values are least-square means from an ANOVA testing for set, trial, and line effects, with thorax length entered as a covariate. Error bars indicate one standard error above the mean.