Available online at www.sciencedirect.com
ScienceDirect
Zoology 109 (2006) 318–330
ZOOLOGY
www.elsevier.de/zool

Sexual size dimorphism in a *Drosophila* clade, the *D. obscura* group

Raymond B. Huey[a,*], Brigitte Moreteau[b], Jean-Claude Moreteau[c], Patricia Gibert[d], George W. Gilchrist[e], Anthony R. Ives[f], Theodore Garland Jr.[g], Jean R. David[b]

[a]*Department of Biology, Box 351800, University of Washington, Seattle, WA 98195-1800, USA*
[b]*Laboratoire Evolution, Génomes et Spéciation, CNRS, 91198 Gif-sur-Yvette Cedex, France*
[c]*Equipe de Démoécologie, Laboratoire Biodiversité et Fonctionnement des Ecosystèmes (LBFE), Université de Metz – UFR Sciences Fondamentales et Appliquées, Campus Bridoux, 57070 Metz, France*
[d]*Laboratoire de Biométrie et Biologie Evolutive (UMR 5558), CNRS, University Lyon 1, 43 bd 11 nov, 69622 Villeurbanne Cedex, France*
[e]*Department of Biology, Box 8795, College of William & Mary, Williamsburg, VA 23187-8795, USA*
[f]*Department of Zoology, University of Wisconsin-Madison, Madison, WI 53706, USA*
[g]*Department of Biology, University of California, Riverside, CA 92521, USA*

Received 28 October 2005; received in revised form 4 April 2006; accepted 10 April 2006

Abstract

The *Drosophila obscura* clade consists of about 41 species, of which 20 were used for analyses of wing and thorax length. Our primary goal was to investigate the magnitude of sexual size dimorphism (SSD) of these traits within this clade and to test Rensch's Rule [when females are larger than males, SSD (e.g., female/male ratio) should decrease with body size]. Our secondary goal was methodological and involved evaluating for these flies alternative measures of SSD (female/male ratio, female/male absolute difference, female/male relative difference), developing a bootstrap method to estimate the magnitude of intraspecific variation in SSD, and applying a new method of estimating allometric relationships that is phylogenetically based and incorporates error variance in both traits. All indices of SSD were strongly correlated for both size traits. Nevertheless, female/male ratio is the best index here: it is easily interpretable and essentially independent of size. For both traits, SSD (F/M) varied interspecifically, showed a strong phylogenetic signal, but did not differ for the main phylogenetic subgroups or correlate with latitude. Factors underlying variation in SSD in this clade are elusive and might include genetic drift. SSD (wing) tended to decrease with increasing size, as predicted by Rensch's Rule, though not consistently so. SSD (thorax) was unrelated to size. However, analysis of published data for thorax length of *Drosophila* spp. ($N = 42$) with a larger size range showed that SSD decreased significantly with increasing size (consistent with Rensch's Rule), suggesting our ability to detect SSD-size relations in the *D. obscura* data may be limited by low statistical power.

Keywords: Comparative method; Female/male ratio; Rensch's rule; Thorax length; Wing length

Introduction

Patterns of sexual size dimorphism (SSD) have long intrigued ecologists as well as evolutionary and

*Corresponding author.
E-mail address: hueyrb@u.washington.edu (R.B. Huey).

0944-2006/$ - see front matter © 2006 Elsevier GmbH. All rights reserved.
doi:10.1016/j.zool.2006.04.003

behavioral biologists. The causes of SSD are complex. They include such proximate factors as innate sex differences in energy intake, expenditure, and allocation, some of which are influenced by lower-level physiology and endocrinology (e.g., Cox et al., 2006). Many of these factors can in turn be influenced by ecological and behavioral differences between the sexes. Moreover, natural and sexual selection often will act differently on the sexes, which may lead to evolutionary divergence in male and female size within a species. Whatever their proximate and ultimate origins, differences in size between females and males can have important impacts on their life history, social behavior, physiology and energetics, and competitive relations (Schoener, 1967; Shine, 1978; Andersson, 1994; Fairbairn, 1997; Cox et al., 2003).

The primary purpose of this paper is to examine SSD in members of the *obscura* group of *Drosophila* as part of an ongoing series of investigations into their comparative biology. Previous studies have examined aspects of cold adaptation (Moreteau et al., 1997; Gibert and Huey, 2001; Gibert et al., 2001) and morphometrical evolution (Moreteau et al., 2003). This clade is attractive for comparative studies for several reasons. About 41 species are known (Ashburner et al., 2005), and most can be cultured in the laboratory, which is crucial for comparative studies that attempt to make evolutionary inferences (Garland and Adolph, 1991). They are largely holarctic in distribution, but some have colonized the afrotropical and neotropical regions. Thus they live in a variety of climates and ecological contexts. Their evolutionary relationships are increasingly well understood (O'Grady, 1999; Renard, 2000; see also references in Moreteau et al., 2003), which allows the application of modern phylogenetically based statistical methods (Garland et al., 2005). Moreover, the magnitude of SSD varies strikingly among species (see below) and thus offers opportunities to search for evolutionary patterns in the magnitude of SSD. Also, we use this opportunity to explore key methodological issues involving allometric scaling. Specifically, we evaluate the reliability of various indices of SSD (e.g., Fairbairn, 1997; Smith and Cheverud, 2002), test for phylogenetic signal in SSD (Blomberg et al., 2003), develop a new bootstrap method for estimating the magnitude of intraspecific variation in SSD, and apply a recently developed (phylogenetic) method to estimate scaling coefficients when both variables have error (Garland et al., 2004; Ives et al., in press).

The magnitude of SSD often varies in a predictable way with body size, and this phenomenon has been called Rensch's Rule (Rensch, 1960; Fairbairn, 1997). For species in which males are larger than females (''male-biased''), SSD typically increases with increasing body size; conversely, in species in which females are the larger sex (''female-biased''), SSD typically decreases with increasing body size (Fairbairn, 1997). In *Drosophila* females are larger than males: thus SSD should decrease with increasing body size if Rensch's Rule holds. We test this expectation for the *D. obscura* group flies. In addition, we analyze published data on thorax length of males and females for 42 species of *Drosophila* from several species groups (Pitnick et al., 1995). Relative to our *D. obscura*-group data set, the Pitnick et al.'s (1995) data set offers a larger sample size, broader phylogenetic coverage, and has a larger range of thorax lengths (~2.6-fold greater on the log scale): the latter two factors should in principle enhance power to detect deviations from isometry of male vs. female body size. However, such increased power may come at a cost: the broader phylogenetic range may obscure patterns in the data (e.g., see discussion and references in Garland et al., 2005). Finally, we look for correlations with latitude because body size in *Drosophila* often varies with latitude (e.g., Karan et al., 2000; Moreteau et al., 2003) and thus SSD might co-vary as well. The causal explanation for any such allometric or geographic patterns – and any others involving SSD – will be obscure as little is known about the natural history and reproductive biology of these flies.

Material and methods

Drosophila strains and morphometrical measurements

Laboratory strains of 20 species (one strain per species) in the *D. obscura* clade were available for study. The size data analyzed here were originally published in Moreteau et al. (2003), which reports their geographic source and latitude, rearing conditions and population sizes, and time in captivity (if known). We assume here that SSD has not evolved significantly in laboratory culture, but of course it might have shifted by drift or by adaptation to the laboratory (Matos et al., 2002; David et al., 2006).

Groups of 10 adult pairs were used as parents. Oviposition took place on a high-nutrient, killed-yeast food (David and Clavel, 1965), which reduces crowding effects (Karan et al., 1999). Larval density was not precisely controlled but was approximately 100 per vial. [In *D. melanogaster*, larval densities up to 300 per vial do not affect adult size traits (Karan et al., 1999).] For each species, at least three different vials were used, originating from different parental groups. All experiments (oviposition and development) were done at 21 °C under a L:D 16:8 photoperiod.

Two size-related traits were measured on each species: total wing length (from the thoracic articulation to the wing tip) and total thorax length (from the neck to the

end of the scutellum). Both measures were taken (using an ocular micrometer) from a left-side, lateral view. Micrometer units were transformed into mm × 100. We measured 25 males and 25 females for each species. We also analyzed published data on average thorax length of males and females of 42 species of *Drosophila* (Pitnick et al., 1995).

Quantifying sexual size dimorphism (SSD)

SSD of species can be quantified in a variety of ways (Fairbairn, 1997; Smith and Cheverud, 2002). We consider three commonly used indices: ratio of female/male size (RA) computed from mean values for males and for females, absolute difference between mean female and mean male size (AD), and relative difference between female and male size, divided by the mean size of the sexes (RD). For the present study, two "size" traits were measured, wing length and thorax length, as described above.

Within-species descriptive statistics for each of the three indices of SSD were estimated by bootstrapping so that we can estimate the standard error of the mean of each SSD index in an unbiased manner. We randomly paired 25 females and 25 males (sampled with replacement) within each species, computed each of the measures of SSD for each pair, and recorded the mean for all 25 pairs. We then computed the bootstrapped mean, standard error, and 95% confidence limits for each trait from 1000 re-samplings. All re-sampling and other statistics were done in R, version 1.8.0 (Ihaka and Gentleman, 1996).

Using a novel approach, we estimated among-species variation in SSD by bootstrapping both the coefficient of variation (CV = standard deviation/mean) and the intra-class correlation coefficient (ICC) from 1000 re-samplings of the entire data set. [CV measures the amount of interspecific variation relative to the overall mean, and ICC estimates the proportion of the total variance that is attributable to differences among species. A high ICC indicates that most of the variation in the sample lies among (rather than within) species.] We took a bootstrapped sample of 25 male and female pairs for each of the 20 species and then computed (using analysis of variance) the within- and among-species variance components for both traits. From these we computed the CV and the ICC. The mean and 95% confidence intervals of these statistics were computed from 1000 bootstrap re-samplings for the size data and the SDD measures (Tables 1 and 2).

Estimating patterns of intraspecific scaling of SSD is not possible unless isofemale lines or other sibships are available (David et al., 2003). Such isofemale line data were not available here.

Phylogenetically based statistical analyses

We performed several statistical analyses that require a phylogeny with a topology of relationships and branch lengths (Garland et al., 2005). Classical taxonomy has

Table 1. Bootstrapped means ± 1 SE for wing and thorax lengths of males and females in 20 species of the *D. obscura* clade

Species	Wing (mm × 100)		Thorax (mm × 100)	
	Female	Male	Female	Male
D. affinis	259.6 ± 1.08	226.5 ± 0.98	105.8 ± 0.44	92.7 ± 0.49
D. ambigua	322.4 ± 1.23	280.8 ± 1.07	127.1 ± 0.53	111.3 ± 0.39
D. athabasca	302.5 ± 1.13	266.4 ± 1.20	115.8 ± 0.46	99.0 ± 0.50
D. azteca	293.5 ± 0.92	251.0 ± 1.17	113.9 ± 0.36	100.9 ± 0.42
D. bifasciata	328.1 ± 1.04	291.0 ± 1.00	128.3 ± 0.52	115.0 ± 0.51
D. bogotana	304.9 ± 0.79	267.4 ± 1.53	115.3 ± 0.43	105.5 ± 0.41
D. guanche	286.2 ± 1.33	277.2 ± 1.47	117.8 ± 0.71	114.6 ± 0.65
D. helvetica	276.8 ± 1.06	233.8 ± 0.91	113.4 ± 0.48	95.5 ± 0.42
D. imaii	273.8 ± 1.19	238.9 ± 1.35	114.3 ± 0.37	101.4 ± 0.73
D. madeirensis	317.9 ± 0.93	283.4 ± 0.86	130.6 ± 0.47	113.8 ± 0.53
D. microlabis	271.5 ± 1.34	242.3 ± 0.96	108.9 ± 0.51	98.7 ± 0.42
D. miranda	343.4 ± 1.14	308.7 ± 1.08	130.0 ± 0.48	115.3 ± 0.40
D. narragansett	263.8 ± 0.99	226.7 ± 1.00	105.0 ± 0.43	87.8 ± 0.48
D. obscura	304.3 ± 1.07	277.4 ± 1.38	121.2 ± 0.50	110.3 ± 0.64
D. persimilis	289.1 ± 1.00	262.9 ± 0.70	115.6 ± 0.37	103.0 ± 0.34
D. pseudoobscura	297.0 ± 0.98	267.7 ± 0.83	116.2 ± 0.46	104.4 ± 0.38
D. subobscura	290.2 ± 0.89	262.0 ± 1.07	118.5 ± 0.38	104.4 ± 0.49
D. subsilvestris	303.9 ± 1.42	262.7 ± 0.96	123.0 ± 0.51	106.0 ± 0.55
D. tolteca	274.1 ± 0.92	224.6 ± 1.13	108.4 ± 0.37	89.4 ± 0.56
D. tristis	319.6 ± 1.50	285.2 ± 1.01	130.7 ± 0.78	115.0 ± 0.59

Table 2. Bootstrapped estimates for the among-species intra-class correlation (ICC) and for coefficients of variation (CV) for female and male body dimensions and for three estimates of sexual size dimorphism: female/male ratio (RA), female–male difference (AD), and female–male relative difference (RD)

	Wing		Thorax	
	Mean ± SE	95% CI	Mean ± SE	95% CI
ICC				
Female	0.94 ± 0.004	(0.936, 0.950)	0.92 ± 0.005	(0.908, 0.927)
Male	0.95 ± 0.003	(0.942, 0.955)	0.92 ± 0.006	(0.908, 0.934)
RA	0.58 ± 0.027	(0.529, 0.632)	0.56 ± 0.029	(0.497, 0.611)
AD	0.53 ± 0.030	(0.470, 0.586)	0.52 ± 0.032	(0.449, 0.578)
RD	0.59 ± 0.029	(0.531, 0.639)	0.56 ± 0.027	(0.504, 0.617)
CV				
Female	7.64 ± 0.087	(7.469, 7.799)	6.88 ± 0.097	(6.686, 7.062)
Male	9.04 ± 0.098	(8.852, 9.229)	8.29 ± 0.114	(8.075, 8.522)
RA	3.39 ± 0.147	(3.099, 3.683)	3.57 ± 0.164	(3.250, 3.872)
AD	24.46 ± 1.254	(21.999, 26.849)	26.22 ± 1.442	(23.448, 28.981)
RD	27.33 ± 1.192	(24.853, 29.605)	28.62 ± 1.400	(25.999, 31.471)

long recognized two subgroups within the *D. obscura* group, namely *obscura* in the Old World and *affinis* in the New World. Numerous publications using molecular data (allozymes or DNA sequences) have confirmed this subdivision (see Moreteau et al., 2003). Within each subgroup, taxonomists have defined lower-level subdivisions, called ''complexes''. In a previous paper (Moreteau et al., 2003), we recognized five such complexes. The reality of one of them, the *microlabis* complex, which comprises afrotropical species, now seems dubious. Recent molecular investigations (O'Grady, 1999; Renard, 2000) suggest that *D. microlabis* (as well as another related African species) should in fact be included in the *D. subobscura* complex. Consequently, we here use only four complexes (*pseudoobscura* and *affinis*, which are nested within the *affinis* subgroup; and *obscura* and *subobscura*, within the *obscura* subgroup). The presumed phylogeny for our 20 species is shown in Fig. 1. As relative branch lengths were not available for this phylogeny, we set them to Pagel's (1992) arbitrary values in the PDTREE program of Garland et al. (1993, 1999), as shown in Fig. 1. For all traits considered, these branch lengths passed the diagnostic test described by Garland et al. (1992). For a sensitivity analysis, we also performed analyses with constant branch lengths (all segments set equal to unity). These also passed the diagnostic test, but mean squared errors calculated with the PHYSIG.M program (Blomberg et al., 2003) were larger for all traits and other results were similar, so only results with Pagel's branch lengths are reported here.

We performed univariate tests for phylogenetic signal (i.e., the tendency for related taxa to resemble each other) using a recent approach developed for continuous traits such as SSD (Blomberg et al., 2003). This test can be implemented either with phylogenetically independent contrasts (Felsenstein, 1985) or via generalized least-squares models (GLS). When implemented with contrasts, the test compares the variance of the standardized contrasts for the real data in their correct phylogenetic positions with the distribution of variances of contrasts for data sets that have been randomized across the tips of the phylogenetic tree, thus destroying any phylogenetic signal that may be present. If the

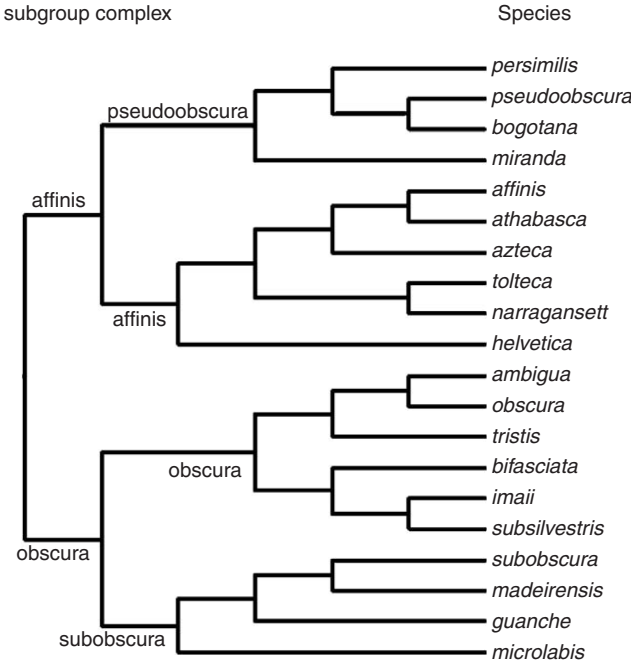

Fig. 1. Phylogenetic hypothesis of species of the *Drosophila obscura* group that are sampled here (see text). Pagel's (1992) arbitrary branch lengths, as used in statistical analyses, are shown.

variance for the data in their correct position is smaller than 5% of the variances for the randomized data sets, then the null hypothesis of no phylogenetic signal is rejected. The equivalent in the GLS implementation uses the mean squared error. We implemented the GLS computations in the Matlab program PHYSIG.M (Blomberg et al., 2003). Computed *P*-values are based on 1000 permutations. As a descriptive statistic for the amount of phylogenetic signal present, we computed the *K* statistic, which takes on a value of 1.00 when a trait shows the amount of signal that would be expected if the trait had evolved via Brownian motion along the specified topology and branch lengths (see Blomberg et al., 2003). A *K* smaller than 1 indicates less signal than expected, whereas a *K* larger than 1 indicates more. Blomberg et al. (2003) provide a survey of *K* values for a wide range of traits in various organisms.

Many different methods have been used to compute how the magnitude of SSD varies with body size (Fairbairn, 1997). For example, some index of SSD (e.g., RA) is often regressed on average female size, on average male size, or on the average of both sexes (see Fairbairn, 1997). As a simple and intuitive test for whether SSD varied significantly with body size, we computed the correlation between female/male size and female size (Smith and Cheverud, 2002), using phylogenetically independent contrasts (Felsenstein, 1985; Garland et al., 1992).

How best to estimate an allometric slope between two traits is very complicated when neither trait is a dependent or an independent variable (as in ordinary least-squares regression analysis) and when both traits are measured with error (e.g., see Rayner, 1985; Garland et al., 1992). Therefore, we performed allometric analyses with new statistical models that can incorporate both phylogenetic information (topology and branch lengths) as well as information on "measurement error" contained in the tip values (estimates of mean values for species). As described and justified elsewhere (Garland et al., 2004; Ives et al., in press), we used bootstrapped estimates of the standard errors of the species mean values (see Table 1) to provide information on measurement error. In the present context, we seek what is often termed the "functional relation" between the traits. The reduced major axis (rma) slope yields unbiased estimates of relations, even with measurement error in one or both traits (Rayner, 1985). Thus, we applied several statistical models using Matlab (programs available on request from A.R.I. or T.G.): rma using a "star" phylogeny (i.e., all species directly descended from a common ancestor) and assuming no measurement error [rma(0)]; rma using a hierarchical phylogeny and no measurement error [rma(*C*)]; rma using a star phylogeny and measurement error [rmaM(0)]; rma using a hierarchical phylogeny and measurement error [rmaM(*C*)]; rma assuming stabilizing selection via an Ornstein–Uhlenbeck (OU) model of character evolution (see Blomberg et al., 2003) [rma(*d*)]; and rma assuming an OU model and the presence of measurement error [rmaM(*d*)]. When $d = 1$, the operational tree is the original phylogeny, while $d = 0$ gives a star. Values between 0 and 1 give a continuum between these two extremes (see Blomberg et al., 2003). The parameter *d* is estimated simultaneously with estimation of the other parameters in the model (i.e., mean of trait 1, variance of trait 1, intercept of the line, slope of the line, variance around the line). Estimation is performed by maximum likelihood (ML), and confidence intervals and/or standard errors are obtained in three ways: (1) using Eq. (15) of Rayner (1985) for the cases without measurement error; (2) using a standard ML asymptotic approximation (which assumes large sample sizes); and (3) via bootstrapping with 2000 samples. The bootstrapped confidence intervals are the most reliable in the present case due to the small sample size (20 species for the *obscura* clade), but the others are presented for comparison. Models can be compared using both their ln likelihoods (negative LLs are given, with smaller [more negative] corresponding to a better fit) and the Akaike Information Criterion (AIC) statistic derived from information theory (smaller values are better). The AIC is calculated as $-2*\mathrm{LL} + 2*p$ (where p = number of parameters in the model).

For the data of Pitnick et al. (1995), we used the phylogenetic tree presented in their Fig. 1, which includes branch lengths in units of estimated divergence times. However, the branch length diagnostic test of Garland et al. (1992) indicated a significant lack of fit for female/male thorax, for ln female thorax, and for both ln and raw male thorax (significant negative correlations between absolute values of standardized contrasts and their standard deviations). However, Pagel's (1992) arbitrary branch lengths indicated no significant lack of fit for any trait; so we used them for analyses, as with analysis of the *obscura* clade data. This also facilitates comparison of the two data sets. As for the *obscura* data set, standard errors associated with the mean values for species were exceedingly small (some reported as 0.00 mm), so we did not attempt to incorporate them in analyses.

To test for an evolutionary correlation between SSD and latitude in the *obscura* clade data, we used bootstrapped mean SSD indices from wing and thorax values (Table 1) and latitudes available in Moreteau et al. (2003). [However, the published latitude for *D. subobscura* (44.6°) is incorrect and was changed to 48.4°]. Statistical correlations were computed for standardized phylogenetically independent contrasts via the PDTREE module of PDAP (Garland et al., 1993, 1999). Differences in mean values between the two subclades were tested by determining whether the basal contrast was unusually large, as described on page 278 of

Garland et al. (1993). Differences in the average (minimum) rate of evolution between the two subclades were tested as described in Garland (1992).

Results

Basic size data

Bootstrapped estimates of wing and thorax length of males and females for each species are given in Table 1. Wing and thorax length of individuals are correlated within species of *Drosophila* in general (David et al., 1994) and also among the 20 species investigated here (average correlation for females within species = 0.70 ± 0.02, for males = 0.71 ± 0.02, see Moreteau et al., 2003). Not surprisingly, bootstrapped estimates of average wing length and average thorax length (Table 1) are also strongly correlated among species, either with standard Pearson correlations ($r = 0.93$ and 0.91, $P < 0.001$) or with phylogenetically independent contrasts ($r = 0.94$ and 0.91, respectively; $P < 0.001$).

Comparing alternative measures of SSD

Females are invariably larger than males in the *obscura* group (Fig. 2), and the magnitude of sexual size dimorphism is pronounced in these flies (Moreteau et al., 2003). The bootstrapped estimates for the three measures of SSD are presented in Fig. 3. Estimates for wing length are plotted on the left and those for thorax length are on the right. Sexual size *mono*-morphism would be indicated (from top to bottom) by RA of 1, AD of 0, or RD of 0.

All three indices of dimorphism examined here are highly correlated in between-species comparisons (data not shown): in fact, ratios of female to male size, absolute differences, and relative differences all have correlations greater than 0.92 for wing length (all $P < 0.002$) and greater than 0.95 for thorax length (all $P < 0.002$). Further, SSD for wings vs. that for thoraxes are also highly correlated for these three indices: correlations range from 0.60 to 0.69 (all $P < 0.01$).

Coefficients of variation (CV) and intra-class correlations (ICC) for the bootstrapped mean size components as well as for the three measures of sexual size dimorphism are reported in Table 2. CVs are intermediate for the two measured traits ($\bar{X} = 7.96 \pm 0.46$, $n = 4$), relatively small for the F/M ratio ($\bar{X} = 3.48 \pm 0.09$, $n = 2$), but strikingly high for the other two estimates of SSD (24.5–43.3).

The intra-class correlation describes the proportion of the variance among the bootstrap values that is associated with differences in size dimorphism among species. These correlations are near unity for the morphological traits and are around 0.5 for the three reliable measures of SSD for both wing and thorax size (Table 2). Thus, virtually all variation in size itself, but only about half the variation in SSD, is among species rather than within species.

For our data, F/M ratio thus seems to be the most useful index. It is intuitive and simple to interpret (Lovich and Gibbons, 1992), dimension free,

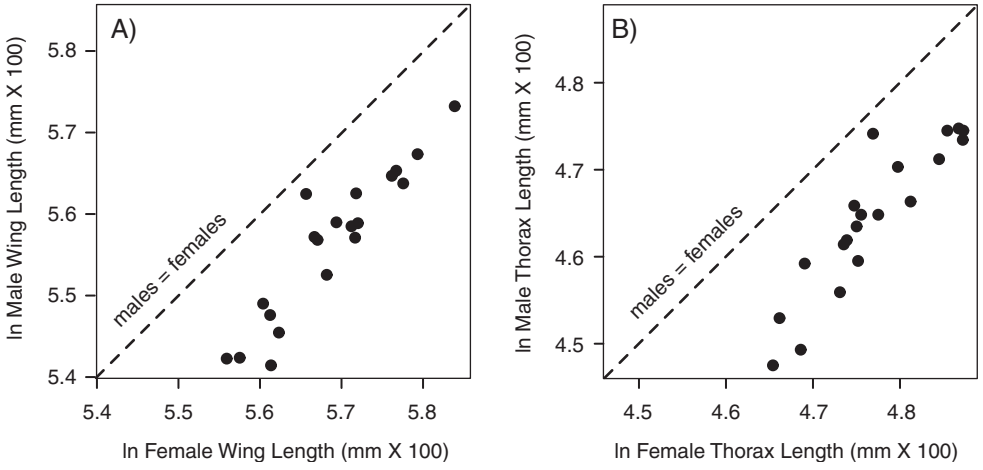

Fig. 2. Relation between average female vs. male wing (A) and thorax (B) length for the 20 species of *Drosophila* in the *obscura* lineage (bootstrapped means from Table 1). The dashed lines indicate a 1:1 ratio and illustrate that females are always larger than males but suggest that the dimorphism appears somewhat reduced for species of larger body size. Nevertheless, whether the null hypothesis of interspecific isometry (slope = 1.00) can be rejected depends on the statistical model used (see Table 4 and text).

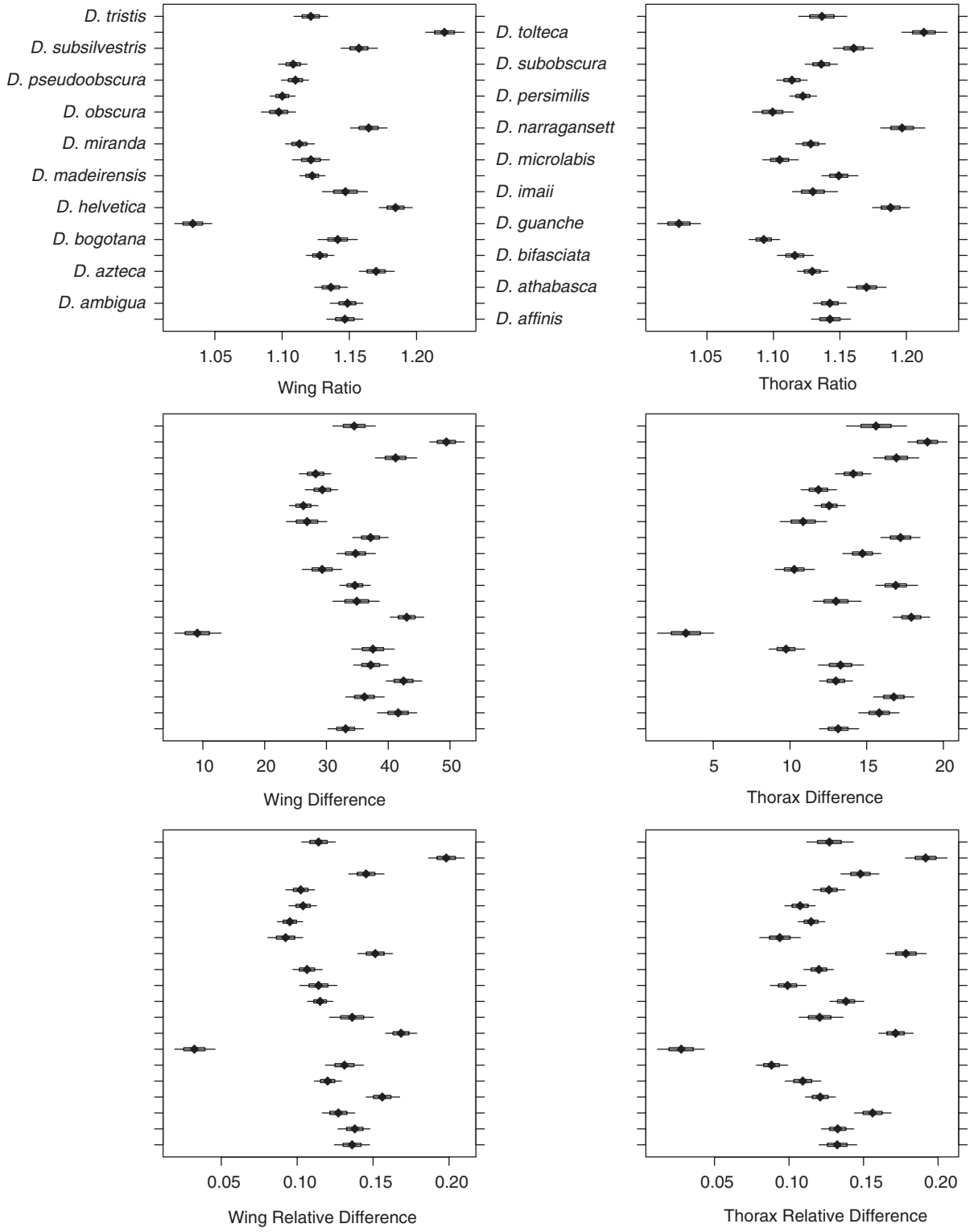

Fig. 3. Bootstrapped values (by species) for wing (left column) and thorax length (right column) for the three measures of SSD. Species names are alternated on the side of the upper left panel. Plotted are the mean (dot), standard errors of mean (box) and 95% confidence intervals (whiskers). Ratio (RA) = female/male; absolute difference (AD) = female–male (mm × 100); relative difference (RD) = difference divided by the mean.

strongly correlated with other measures, and gives good separation of species (high intra-class correlation coefficients, Table 2).

Phylogenetic signal

As shown in Table 3, for the *obscura* clade all traits showed statistically significant phylogenetic signal, with the exception of female wing length, when Pagel's (1992) arbitrary branch lengths were used. Similar significant results were found for constant branch lengths (data not shown). Therefore, statistical methods that incorporate phylogenetic relationships are generally to be preferred, at least for univariate analyses.

Allometric scaling of SSD

SSD does not vary significantly with body size for this sample of species of the *obscura* clade, as correlations were invariably weak. For thorax size, the contrasts correlation between F/M thorax and female thorax was 0.273 (2-tailed $P>0.2$), whereas the conventional correlation was −0.225 (2-tailed $P>0.3$). For wing

Table 3. Phylogenetic analysis of data for 20 species in the *D. obscura* clade

Trait	Diagnostic correlation[a]	*P* for phylogenetic signal	*K*	Contrasts correlation with latitude[b]	Conventional correlation with latitude[b]
Latitude	−0.156	0.044	0.683		
Wing, female	−0.303	0.146	0.561	0.284	0.428
Wing, male	−0.002	0.022	0.739	0.389	0.397
Thorax, female	−0.052	0.009	0.777	0.379	0.553*
Thorax, male	0.311	< 0.001	1.167	0.235	0.395
F/M wing	−0.202	0.009	0.753	−0.329	−0.111
F/M thorax	−0.107	0.005	0.824	0.291	0.134

Diagnostic tests for adequacy of Pagel's (1992) arbitrary branch lengths in phylogenetically independent contrasts analyses Garland et al. (1992), tests for phylogenetic signal and *K* statistic describing amount of signal (Blomberg et al., 2003), and correlations of each trait with latitude. Values for first five traits were taken from Moreteau et al. (2003). Values for F/M ratios are from Table 1.

*$P<0.02$.

[a]Critical value for 2-tailed test = ±0.456 for $P = 0.05$ with 1.17 d.f.

[b]Critical value for 2-tailed test = ±0.444 for $P = 0.05$ with 1.18 d.f.

Table 4. Allometric analyses of the relation between ln male trait and ln female trait for the *D. obscura* clade (mean values from Table 1 and shown in Fig. 2)

Model[a]	Slope	Negative ln likelihood	AIC[b]	*d*	95% C.I.[c]	SE ML approx.	SE bootstrap	95% C.I. bootstrap
Wing								
rma(0)	1.1993	−82.3777	−154.7554	0	0.9875, 1.4566	0.0969	0.1075	1.0043, 1.4430
rma(*C*)	1.0379	−80.9409	−151.8818	1	0.8511, 1.2659	0.0854	0.0957	0.8721, 1.2456
rmaM(0)	1.1992	−82.4894	−154.9788	0		0.0962	0.1047	1.0119, 1.4248
rmaM(*C*)	1.0421	−81.2019	−152.4038	1		0.0852	0.0905	0.8799, 1.2226
rma(*d*)	1.1150	−83.1318	−154.2636	0.3361		0.1049	0.0981	0.9332, 1.3208
rmaM(*d*)	1.1169	−83.2450	−154.4900	0.3416		0.1030	0.0993	0.9336, 1.3384
Thorax								
rma(0)	1.2317	−83.1722	−156.3444	0	0.9780, 1.5512	0.1146	0.1229	1.0064, 1.4965
rma(*C*)	0.9954	−86.7441	−163.4882	1	0.7412, 1.3368	0.1115	0.1202	0.7879, 1.2547
rmaM(0)	1.2306	−83.3143	−156.6286	0		0.1138	0.1240	1.0102, 1.4982
rmaM(*C*)	1.0069	−86.8429	−163.6858	1		0.1125	0.1232	0.7850, 1.2731
rma(*d*)	1.0446	−87.0855	−162.1710	0.6877		0.1296	0.1288	0.8185, 1.3339
rmaM(*d*)	1.0536	−87.1832	−162.3664	0.6856		0.1327	0.1321	0.8259, 1.3527

[a]All may be considered variants of "reduced major axis" (e.g., Rayner, 1985). In parentheses, 0 indicates use of star phylogeny, C indicates hierarchical phylogeny as shown in Fig. 1, d indicates phylogeny altered to optimize fit of tree to data under the Ornstein–Uhlenbeck model of character evolution using transformation parameter d (see text). "M" following rma indicates that standard errors (from Table 1) were incorporated into analyses.

[b]A smaller AIC score is better.

[c]Eq. (15) of Rayner (1985).

length, the contrasts correlation was 0.089 ($P>0.5$), and the conventional correlation was −0.266 ($P>0.2$).

The allometric analysis of SSD (M directly on F, Fig. 2) yielded somewhat complicated results (Table 4). For both wing and thorax length, slope estimates were approximately 1.2, which is consistent with Rensch's Rule: thus SSD tends to be reduced at larger body sizes, as is evident in Fig. 2. For both traits, the analysis assuming a star phylogeny yielded bootstrapped 95% confidence intervals that just excluded isometry (slope = 1.00), whether or not information on the standard errors of the species means was incorporated.

In contrast, the null hypothesis of isometry could not be rejected by any of the phylogenetic analyses (Table 4). For wing length, the likelihoods and AICs give us little reason to prefer the phylogenetic analyses over those that assume a star phylogeny. [Although the negative log likelihood for rma(d) is (necessarily) lower than that for rma(0), its AIC is higher due to the added parameter d that must be estimated.] For thorax length, however, both statistics suggest that the phylogenetic models are substantially better. These models suggest that male vs. female sizes do not deviate significantly from isometry.

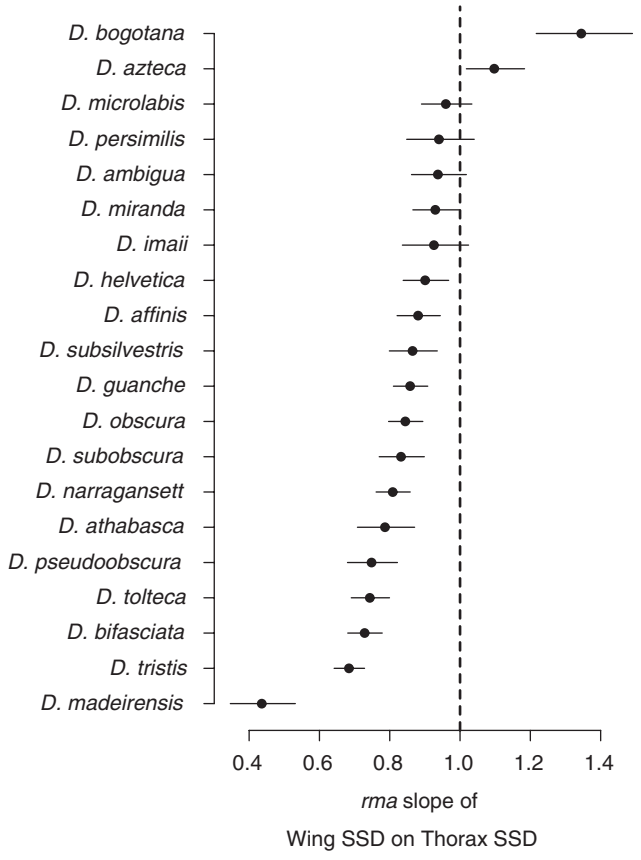

Fig. 4. Intraspecific allometry of wing vs. thorax SSD ratios. The values shown are the slopes and 99% confidence intervals from the regression of bootstrapped estimates of ln-transformed wing on thorax RA. Isometry is indicated by a dashed, vertical line.

Analyses of intraspecific allometry of SSD require males and females from sib groups or isofemale lines (David et al., 2003), which were not available here. However, we were able to examine the pattern of allometric variation between RA of wings vs. RA of thoraxes. We used reduced major axis regression (rma) of the ln-transformed ratios of the bootstrapped means to compute the slopes (Fig. 4, major axis regression yields similar results). The allometric scaling of wings on thoraxes is less than 1 for most species, indicating a relatively larger increase in thorax SSD for each incremental increase in wing SSD.

Evolutionary patterns of SSD

SSD (F/M) clearly varies among species (Fig. 3). SSD was lowest for *D. guanche* (1.03 for both wing and thorax length) and highest for *D. tolteca* (1.22 and 1.21, respectively). SSD for most species were clustered between 1.10 and 1.18 for wing length or 1.09 and 1.20 for thorax length (Fig. 3). Nevertheless, 95% confidence intervals for bootstrap values of SSD were non-overlapping for many species pairs, suggesting significant interspecific heterogeneity. Consequently, it is of interest to search for patterns that might correlate with that heterogeneity.

The *obscura* group is subdivided into 2 subgroups (*affinis* and *obscura*, see Introduction and Fig. 1). Do these two subclades differ in SSD? No, as mean SSD for the *affinis* subgroup was virtually identical to that of the *obscura* subgroup (wing: 1.15 ± 0.011 vs. 1.12 ± 0.111; thorax: 1.15 ± 0.040 vs. 1.12 ± 0.038; $P=0.165$ and 0.315, respectively). A phylogenetic ANOVA (Garland et al., 1993) confirms that the two subgroups do not differ significantly in SSD for either trait, nor for latitudinal range, wing length or thorax length (results not shown). Comparisons of the absolute values of the standardized independent contrasts (Garland, 1992) also indicated that the two subgroups do not differ significantly with respect to average (minimum) rate of evolution for any of the above-mentioned traits (results not shown).

Does size and SSD co-vary with latitude? This might occur, given that body size varies with latitude both within and among species (see references in Moreteau et al., 2003). Source latitudes of these species range widely (1–66°, see Moreteau et al., 2003). A previous non-phylogenetic analysis indicated that average size (PC1 for wing and thorax size) was positively (though weakly) correlated with latitude (combined correlation for both sexes, see Moreteau et al., 2003), as is often the case within species of *Drosophila* (Moreteau et al., 2003). Our results are similar. With phylogenetically independent contrasts, we find positive but not statistically

Table 5. Phylogenetic analysis of data for 42 species of *Drosophila* from Pitnick et al. (1995) using Pagel's (1992) arbitrary branch lengths (see text)

Trait	Diagnostic correlation[a]	*P* for phylogenetic signal	*K*
F/M thorax	−0.277	< 0.001	0.592
Female thorax	0.034	< 0.001	1.017
Male thorax	0.008	< 0.001	1.051
ln Female thorax	0.031	< 0.001	1.069
ln Male thorax	−0.004	< 0.001	1.099

Shown are diagnostic tests for adequacy of branch lengths (Garland et al., 1992), randomization test for presence of phylogenetic signal (Blomberg et al., 2003), and *K* statistic describing amount of signal (Blomberg et al., 2003, see text).

[a]Critical value for 2-tailed test = ±0.308 for $P = 0.05$ with 1.39 d.f.

significant correlations with latitude for both wing and thorax length (Table 3). With conventional statistical analyses, female thorax length shows a significant positive correlation with latitude. SSD (F/M) for thorax or for wing size does not vary significantly with latitude in phylogenetic or in conventional analyses (Table 3).

Analysis of data from Pitnick et al. (1995)

Pitnick et al. (1995) reported thorax lengths of male and female *Drosophila* of 42 species from diverse species groups. The contrasts correlation between F/M thorax and female thorax was −0.351 (2-tailed $P = 0.02$), whereas the conventional correlation was −0.569 (2-tailed $P < 0.001$): both analyses suggest that SSD decreases significantly with body size, consistent with Rensch's Rule. These analyses used Pagel's (1992) arbitrary branch lengths, which passed the diagnostic test of Garland et al. (1992) for all traits (see Table 5). With these branch lengths, all traits showed highly significant phylogenetic signal.

Fig. 5 shows ln male vs. ln female thorax lengths, and clearly illustrates that the difference in size is reduced in large-bodied species. Irrespective of the analytical method employed, the estimates of the slope of the ln–ln relation were always statistically greater than 1.0 (Table 6). The likelihoods and AIC statistics both suggest that the phylogenetic analyses fit the data significantly better than did the star analysis, but estimates of slope are similar for both analyses.

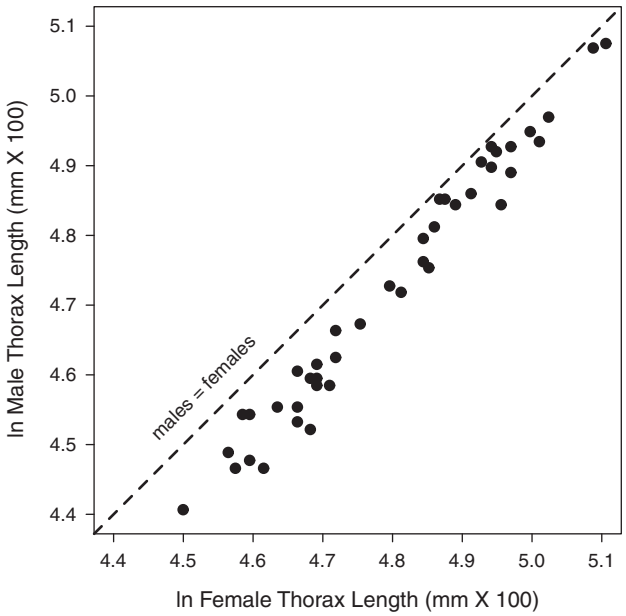

Fig. 5. Relation between average male vs. female thorax length for the 42 species of *Drosophila* reported in Pitnick et al. (1995). The dashed line indicates a 1:1 ratio and illustrates that (1) females are always larger than males but (2) the dimorphism is reduced for species of larger thorax size. Thus, the null hypothesis of interspecific isometry (slope = 1.00) can be rejected for these data (see text and Table 6). Compare with Fig. 2b.

Discussion

Flies in the *obscura* clade of *Drosophila* show marked sexual dimorphism in both wing and thorax length: females in all species are larger than males. Here, we begin by comparing several alternative measures of SSD and then search for evolutionary patterns in SSD in this clade.

As noted above, independent data on thorax size of females and males of three *obscura*-group flies are published in Pitnick et al. (1995). The two sets of measurements differ slightly, at least in part because different measurement techniques were used. Note, however, that RA calculated from their data on thorax length are similar to ours [*D. affinis* 1.17 (Pitnick et al. data) vs. 1.1 (our data), *D. pseudoobscura* 1.08 vs. 1.12, and *D. persimilis* 1.13 vs. 1.12), suggesting that estimates of SSD are reasonably robust and relatively independent of the exact stock, rearing conditions, and measurement technique.

Comparing measures of SSD

We compared three methods of estimating SSD: the ratio of female to male size (RA), the absolute difference between females and males (AD), and the relative difference (RD, scaled by average size). To estimate the within-species variance in these estimates of SSD, we developed a new bootstrapping method in which males

Table 6. Allometric analyses of the relation between ln male trait and ln female trait for the 42 species presented by Pitnick et al. (1995) (mean values as shown in Fig. 5)

Model[a]	Slope	Negative ln likelihood	AIC[b]	*d*	95% C.I.[c]	SE ML approx.	SE bootstrap	95% C.I. bootstrap
rma(0)	1.1466	−144.5648	−279.13	0	1.0845, 1.2123	0.0292	0.0308	1.0882, 1.2067
rma(*C*)	1.1298	−160.1312	−310.26	1	1.0425, 1.2244	0.0409	0.0410	1.0514, 1.2145
rma(*d*)	1.1262	−162.3774	−312.75	0.6703		0.0376	0.0391	1.0446, 1.2044

[a]All models may be considered variants of what is commonly termed the "reduced major axis" (e.g., Rayner, 1985). In parentheses, 0 indicates use of star phylogeny, *C* indicates hierarchical phylogeny as shown in Fig. 1 of Pitnick et al. (1995) but with Pagel's (1992) arbitrary branch lengths, *d* indicates phylogeny altered to optimize fit of tree to data under the Ornstein–Uhlenbeck model of character evolution using transformation parameter *d* (see text). As explained in the text, standard errors reported by Pitnick et al. (1995) were exceedingly small and hence were not incorporated into analyses.

[b]A smaller AIC score is better.

[c]Eq. (15) of Rayner (1985).

and females were paired randomly (with replacement). SSD estimates of each pair were then calculated. Repeated re-sampling yields empirical distributions of these indices within each species.

All three SSD indices yielded very similar results. We prefer the F/M ratio (RA) as a descriptor of SSD for these species because it has a clear and obvious biological significance: a ratio of 1.13 means that traits are 13% larger in females than those in males (Reeve and Fairbairn, 1999; David et al., 2003).

Ratios are widely used in biology, but are often criticized for two statistical reasons (but see Ranta et al., 1994; Sokal and Rohlf, 1995). First, they often have non-normal distributions, even when the traits themselves are normally distributed. Of course, this can be corrected by suitable transformation of the ratio. In any case, the distribution of SSD ratio in *D. melanogaster* was not significantly different from normal (David et al., 2003). Second, a ratio is expected to be a constant only if the intercept (from a type I regression) is close to zero. The argument is as follows: we assume a linear regression, $F = f(M)$ according to the formula: female $= b(\text{male}) + a$. So the F/M ratio becomes: $b + a/\text{male}$. Thus, the ratio is inversely related to male size, yielding a likely correlation between F/M and M.

How serious is the latter concern? We analyzed this relationship using species means and found intercepts ($a \pm$SE) of 0.89 ± 0.08 and 0.85 ± 0.09 for wing and thorax, respectively. Because the intercepts are both positive but less than 1, we expect a negative correlation between F/M ratio and male value. Empirical data confirmed this expectation, but the coefficients were not significant and very close to zero (−0.14 and −0.15 for wing and thorax, respectively). In other words, the impact of M explains less than 3% of the total variability of the ratio, so that the induced statistical bias can be easily neglected. This exercise validates the use of the F/M ratio as a convenient descriptor of SSD, at least for our data.

Empirical patterns

Interspecific allometry in SSD was examined by regressing male size on female size, as recommended by Fairbairn (1997). We used a generalized form of what is typically termed reduced major axis (rma) regression that can incorporate detailed information on within-species variation ("measurement error") in both conventional and phylogenetic forms. Although results varied with analysis type, we conclude that wing length seems to deviate from isometry in a way consistent with Rensch's Rule (Fairbairn, 1997), although this result is only just barely significant at the 0.05 level. On the other hand, thorax length in the *obscura* clade shows little statistical evidence of deviation from isometry (Table 4 and Fig. 2). However, the larger sample of 42 species (Pitnick et al., 1995) shows a strong deviation from isometry for thorax length (Table 6 and Fig. 5) and is consistent with Rensch's Rule. Both data sets show the same pattern, but differ in statistical significance: this discrepancy may simply relate to the relatively limited power to detect deviations from isometry in the *obscura* analysis, which involves many fewer species and a much narrower range of body sizes (compare Figs. 2 and 5).

One outlier, *D. guanche*, has greatly reduced SSD (Fig. 3). This endemic species from the Canary Islands likely split from a continental ancestor, which now has produced *D. subobscura*, less than one million years ago (Cabrera et al., 1983). Indeed, *D. guanche* and *D. subobscura* hybridize in the laboratory and produce fertile F1 females. The ancestral SSD in this complex is not known, but can be estimated under a Brownian motion assumption of character evolution (Garland et al., 1999). For F/M wing length ratio (using bootstrapped means from Table 1), the hypothetical ancestor at the divergence between *D. guanche* and the ancestor of [*D. subobscura* + *D. madeirensis*] is estimated to have been 1.10 with a 95% confidence interval of 1.06–1.14. For F/M thorax length

ratio, the corresponding values are 1.13 and 1.09–1.18. Thus, the Canary Islands lineage has likely experienced a strong reduction of SSD.

A low SSD is also known in some other drosophilid species, for example, in *Zaprionus indianus* (Karan et al., 2000). Because SSD may evolve quite rapidly, the question is: what are the reasons that favor a low vs. high sexual dimorphism in drosophilids? Darwinian interpretations (Darwin, 1871) consider that SSD may evolve either as a consequence of sexual selection (e.g., male-male competition or female choice) or natural selection (e.g. different adaptations in the two sexes, see also Cox et al., 2003). The striking deviation of *D. guanche* from its relatives would encourage checking these Darwinian alternatives by experiments and related behavioral observations (Llopart et al., 2000; Kopp et al., 2001).

Within-species allometry between SSD on wings and SSD on thoraxes (Fig. 4) is generally characterized by slopes that are less than 1, meaning that SSD of thoraxes is greater than that of wings. Two species (*Drosophila azteca* and particularly *D. bogotana*) have greater dimorphism in the wings than in the thorax. Why these species are different is unclear.

Since diversifying from a common ancestor around 20 million years ago (Powell, 1997), the *D. obscura* clade has produced more than 35 extant species and has colonized different biogeographic areas including the palearctic and nearctic regions, the neotropical region and the afrotropical mountains. The geographic distribution is related to phylogeny and especially the *affinis* subgroup (with its two complexes, *affinis* and *pseudoobscura*) has evolved in the New World where all extant species are found, except the European *D. helvetica*. In contrast, the *D. obscura* subgroup is restricted to the Old World (except for introduced species). Despite a likely long and independent evolution in different regions, the two main subgroups do not differ significantly in SSD.

We also checked whether the magnitude of SSD might change with latitude. In a previous study we showed that average size does increase with latitude, as is usually the case in *Drosophila* (see Moreteau et al., 2003). Nevertheless, SSD was uncorrelated with latitude in our comparisons (phylogenetic or conventional) for either wing or thorax length (Table 5). Cardillo (2002) found a similar lack of clinal variation in SSD in a phylogenetically corrected study of bird species ranging from the equator to the poles. In passing, we note that latitude itself showed significant phylogenetic signal (Table 3), as has been noted in several previous studies (Freckleton et al., 2002; Blomberg et al., 2003; Rezende et al., 2004).

In conclusion, we have examined indices of SSD in the *obscura*-group species of *Drosophila*. These species differ overall in SSD. However, SSD is similar for the two subgroups, even though they occur in largely different geographic areas, and independent of latitude. Rensch's Rule is partially supported: SSD in wing size – but not in thorax size – tends to decrease with body size. However, the Pitnick et al. (1995) data set, which covers a larger size and phylogenetic range of *Drosophila*, does show a significant reduction in SSD with size, clearly consistent with Rensch's Rule. Overall, an understanding of variation in SSD in the *obscura* group will require detailed observations on the natural history of component species.

Acknowledgments

Research was supported by CNRS-NSF agreement 9145 to BM, by NSF INT-09981780, NSF DEB-9981598, and IBN-0416843 to RBH, by NSF DEB0242313 and a Jeffress Memorial Trust grant to GWG, and by NSF DEB-0196384 NSF to TG and ARI. We thank W. Blanckenhorn and an anonymous reviewer for constructive comments.

References

Andersson, M., 1994. Sexual Selection. Princeton University Press, Princeton, NJ.

Ashburner, M., Golic, K.G., Hawley, R.S., 2005. Drosophila, a Laboratory Handbook, second ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Blomberg, S.P., Garland Jr., T., Ives, A.R., 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57, 714–745.

Cabrera, V.M., Gonzalez, A.M., Larruga, J.M., Gullon, A., 1983. Genetic distance and evolutionary relationships in the *Drosophila obscura* group. Evolution 37, 675–689.

Cardillo, M., 2002. The life-history basis of latitudinal diversity gradients: how do species traits vary from the poles to the equator? J. Anim. Ecol. 71, 79–87.

Cox, R.M., Skelly, S.L., John-Alder, H.B., 2003. A comparative test of adaptive hypotheses for sexual size dimorphism in lizards. Evolution 57, 1653–1669.

Cox, R.M., Skelly, S.L., John-Alder, H.B., 2006. Testosterone inhibits growth in juvenile male Eastern Fence lizards (*Sceloporus undulatus*): implications for energy allocation and sexual size dimorphism. Physiol. Biochem. Zool. 78, 531–545.

Darwin, C., 1871. The Descent of Man, and Selection in Relation to Sex, vol. 2. John Murray, London.

David, J., Clavel, M.-F., 1965. Interaction entre le génotype et le milieu d'élevage. Conséquences sur les caractéristiques du développement de la Drosophile. Bull. Biol. France Belgique 99, 369–378.

David, J.R., Moreteau, B., Gauthier, J.P., Pétavy, G., Stockel, J., Imasheva, A., 1994. Reaction norms of size characters in relation to growth temperature in *Drosophila melanogaster*: an isofemale line analysis. Genet. Sel. Evol. 26, 229–251.

David, J.R., Gibert, P., Mignon-Grasteau, S., Legout, H., Pétavy, G., Beaumont, C., Moreteau, B., 2003. Genetic variability of sexual size dimorphism in a natural population of *Drosophila melanogaster*: an isofemale line approach. J. Genet. 82, 101–110.

David, J.R., Araripe, L.O., Bitner-Mathé, B.C., Capy, P., Goni, B., Klaczko, L.B., Legout, H., Martins, M.B., Vouidibio, J., Yassin, A., Moreteau, B., 2006. Sexual dimorphism of body size and sternopleural bristle number: a comparison of geographic populations of an invasive cosmopolitan drosophilid. Genetica, in press.

Fairbairn, D., 1997. Allometry for sexual size dimorphism: pattern and process in the coevolution of body size in males and females. Annu. Rev. Ecol. Syst. 28, 659–687.

Felsenstein, J., 1985. Phylogenies and the comparative method. Am. Nat. 125, 1–15.

Freckleton, R.P., Harvey, P.H., Pagel, M., 2002. Phylogenetic analysis and comparative data: a test and review of evidence. Am. Nat. 160, 712–726.

Garland Jr., T., Adolph, S.C., 1991. Physiological differentiation of vertebrate populations. Annu. Rev. Ecol. Syst. 22, 193–228.

Garland Jr., T., 1992. Rate tests for phenotypic evolution using phylogenetically independent contrasts. Am. Nat. 140, 509–519.

Garland Jr., T., Harvey, P.H., Ives, A.R., 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. Syst. Biol. 41, 18–32.

Garland Jr., T., Dickerman, A.W., Janis, C.M., Jones, J.A., 1993. Phylogenetic analysis of covariance by computer simulation. Syst. Biol. 42, 265–292.

Garland Jr., T., Midford, P.E., Ives, A.R., 1999. An introduction to phylogenetically based statistical methods, with a new method for confidence intervals on ancestral values. Am. Zool. 39, 374–388.

Garland Jr., T., Ives, A.R., Midford, P.E., 2004. Within-species variation and measurement error in phylogenetic comparative methods. Integ. Comp. Biol. 44, 556 (abstract).

Garland Jr., T., Bennett, A.F., Rezende, E.L., 2005. Phylogenetic approaches in comparative physiology. J. Exp. Biol. 208, 3015–3035.

Gibert, P., Huey, R.B., 2001. Chill-coma temperature in *Drosophila*: effects of development temperature, latitude, and phylogeny. Physiol. Biochem. Zool. 74, 429–434.

Gibert, P., Moreteau, B., Pla, E., Petavy, G., Karan, D., David, J.R., 2001. Chill-coma tolerance, a major climatic adaptation among *Drosophila* species. Evolution 55, 1063–1068.

Ihaka, R., Gentleman, R., 1996. R: a language for data analysis and graphics. J. Comp. Graph. Stat. 9, 299–314.

Ives, A.R., Midford, P.E., Garland Jr., T., in preparation. Within-species variation and measurement error in phylogenetic comparative methods. Syst. Biol.

Karan, D., Morin, J.P., Gravot, E., Moreteau, B., David, J.R., 1999. Body size reaction norms in *Drosophila melanogaster*: temporal stability and genetic architecture in a natural population. Genet. Selection Evol. 31, 491–508.

Karan, D., Dubey, S., Moreteau, B., Parkash, R., David, J.R., 2000. Geographical clines for quantitative traits in natural populations of a tropical drosophilid: *Zaprionus indianus*. Genetica 108, 91–100.

Kopp, A., Duncan, I., Godt, D., Carroll, S.B., 2001. Genetic control and evolution of sexually dimorphic characters in *Drosophila*. Nature 410, 611.

Llopart, A., Elwyn, S., Coyne, J.A., 2000. Pigmentation and mate choice in *Drosophila*. Nature 408, 553–559.

Lovich, J.E., Gibbons, J.W., 1992. A review of techniques for quantifying sexual size dimorphism. Growth Dev. Aging 56, 181–269.

Matos, M., Avelar, T., Rose, M.R., 2002. Variation in the rate of convergent evolution: adaptation to a laboratory environment in *Drosophila subobscura*. J. Evol. Biol. 15, 673–682.

Moreteau, B., Morin, J.-P., Gibert, P., Petavy, G., Pla, E., David, J.R., 1997. Evolutionary changes of nonlinear reaction norms according to thermal adaptation: a comparison of two *Drosophila* species. C. R. Acad. Sci. Paris 320, 833–841.

Moreteau, B., Gibert, P., Pétavy, G., Moreteau, J.-C., Huey, R.B., David, J.R., 2003. Morphometrical evolution in a *Drosophila* clade: the *Drosophila obscura* group. J. Zool. Syst. Evol. Res. 41, 64–71.

O'Grady, P.M., 1999. Reevaluation of phylogeny in the *Drosophila obscura* species group based on combined analysis of nucleotide sequences. Mol. Phylogenet. Evol. 12, 124–139.

Pagel, M., 1992. A method for the analysis of comparative data. J. Theor. Biol. 156, 431–442.

Pitnick, S., Markow, T.A., Spicer, G.S., 1995. Delayed male maturity is a cost of producing large sperm in *Drosophila*. Proc. Natl. Acad. Sci. USA 92, 10614–10618.

Powell, J.R., 1997. Progress and Prospects in Evolutionary Biology. The *Drosophila* Model. Oxford University Press, New York, NY.

Ranta, E., Laurila, A., Elmberg, J., 1994. Reinventing the wheel: analysis of sexual dimorphism in body size. Oikos 70, 313–321.

Rayner, J.M.V., 1985. Linear relations in biomechanics: the statistics of scaling functions. J. Zool. London A 206, 415–439.

Reeve, J.P., Fairbairn, D.J., 1999. Change in sexual size dimorphism as a correlated response to selection on fecundity. Heredity 83, 697–706.

Renard, E., 2000. Evolution de la famille amylase chez les drosophiles du groupe *obscura*: caractérisation du nouveau gène Amyrel. Université François Rabelais, Tours, France.

Rensch, B., 1960. Evolution Above the Species Level. Columbia University Press, New York.

Rezende, E.L., Bozinovic, F., Garland Jr., T., 2004. Climatic adaptation and the evolution of basal and maximum rates of metabolism in rodents. Evolution 58, 1361–1374.

Schoener, T.W., 1967. The ecological significance of sexual dimorphism in size in the lizard *Anolis conspersus*. Science 155, 474–477.

Shine, R., 1978. Sexual size dimorphism and male combat in snakes. Oecologia 33, 269–278.

Smith, R.J., Cheverud, J.M., 2002. Scaling of sexual dimorphism in body mass: a phylogenetic analysis of Rensch's Rule in primates. Int. J. Primatol. 23, 1095–1135.

Sokal, R.R., Rohlf, F.J., 1995. Biometry, third ed. W.H. Freeman and Company, New York.