Allometry constrains the evolution of sexual dimorphism in *Drosophila* across 33 million years of divergence

Jacqueline L. Sztepanacz[1,2,3] [iD] and David Houle[2] [iD]

[1]*Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON M5S 3B2, Canada*

[2]*Department of Biology, Florida State University, Tallahassee, Florida 32306*

[3]*E-mail: jsztepanacz@gmail.com*

Sexual dimorphism is widely viewed as adaptive, reflecting the evolution of males and females toward divergent fitness optima. Its evolution, however, may often be constrained by the shared genetic architecture of the sexes, and by allometry. Here, we investigated the evolution of sexual size dimorphism, shape dimorphism, and their allometric relationship, in the wings of 82 taxa in the family Drosophilidae that have been diverging for at least 33 million years. Shape dimorphism among species was remarkably similar, with males characterized by longer, thinner wings than females. There was, however, quantitative variation among species in both size and shape dimorphism, with evidence that they have adapted to different evolutionary optima in different clades on timescales of about 10 million years. Within species, shape dimorphism was predicted by size, and among species, there was a strong relationship between size dimorphism and shape dimorphism. Allometry constrained the evolution of shape dimorphism for the two most variable traits we studied, but dimorphism was evolutionary labile in other traits. The keys for disentangling alternative explanations for dimorphism evolution are studies of natural and sexual selection, together with a deeper understanding of how microevolutionary parameters of evolvability relate to macroevolutionary patterns of divergence.

Phenotypic difference between males and females, sexual dimorphism, is one of the most striking forms of phenotypic variation observed in nature. Sexual dimorphism is widespread among both traits and species, and its evolution has garnered substantial research interest. Males and females share the majority of their genomes as each sex passes alleles to both sons and daughters, which results in cross-sex trait correlations (Lande 1980). However, the sexes may also experience selection toward sex-specific fitness optima (Cox and Calsbeek 2009). Shared genetic architecture can limit trait divergence between the sexes, potentially causing both male and female phenotypes to be suboptimal. This is termed intralocus sexual conflict, and it results in a decrease in population mean fitness (Bonduriansky and Chenoweth 2009; Cox and Calsbeek 2009). There are a number of studies that have demonstrated ongoing antralocus sexual conflict in populations (e.g., Foerster et al. 2007; Collet et al. 2016; Wolak et al. 2018).

The evolution of sexual dimorphism, in general, is thought to be adaptive by enabling each sex to reach, or become close to, its phenotypic optimum (Bonduriansky and Chenoweth 2009). Many authors have also emphasized that sexual dimorphism and sexual conflict are not necessarily related (Bonduriansky and Chenoweth 2009; Cox and Calsbeek 2009; Ingleby et al. 2015). Sexual dimorphism can arise without intralocus sexual conflict as a result of differences in sex-specific genetic variation and cross-sex covariation (Lande 1980; Connallon and Clark 2014; Cheng and Houle 2020). In contrast, sexually homologous traits can, in principle, be subject to persistent antagonistic selection without ever evolving dimorphism.

In this article, we investigate the evolution of sexual dimorphism in wings of 82 taxa of flies in the family Drosophilidae. Drosophilid wings exemplify the paradox of evolutionary stasis, with high inferred evolvability yet low rates of evolution.
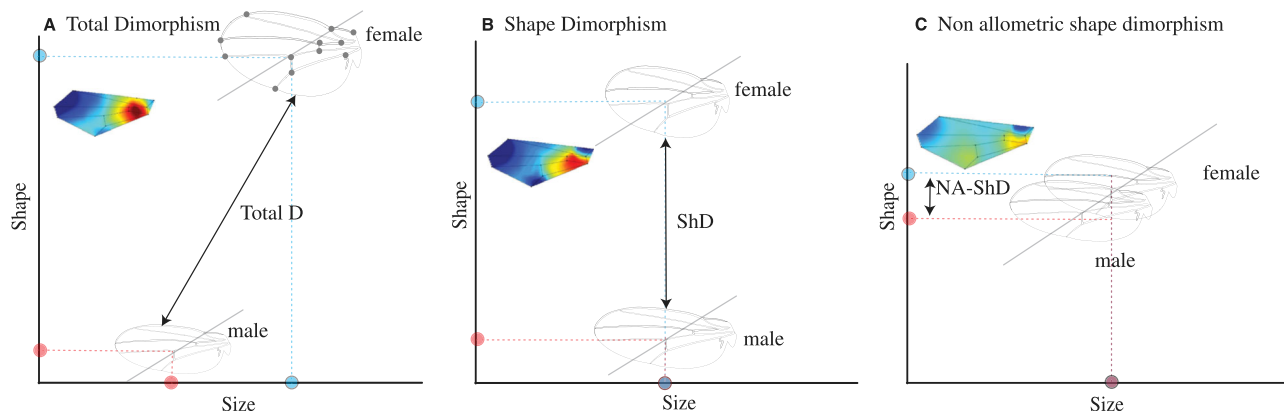
Figure 1. The components of sexual dimorphism in *Drosophila melanogaster* wings, represented in two dimensions. The gray points on the female wing shown in panel A are the 12 vein intersections that we use as landmarks. Red and blue dashed lines and points help to map differences to the *x* and *y* axes for males and females, respectively. The gray regression lines going through male and female wings represent the allometric relationship between size and shape. Males and females are assumed to have the same slope. The distance between male and female wings, shown by the black arrows, is the magnitude of dimorphism. (A) Male and female wings differ in both size and shape, along the *x* and *y* axes. The distance between wings is the total dimorphism (Total Dim). The deformation represented by this panel corresponds to the Total Dim column in Figure 3. (B) Male and female wings have been scaled to the same average centroid size using Procrustes alignment, and now only differ in shape. The distance between wings is shape dimorphism. These shape differences are caused by both allometric and nonallometric effects. The deformation represented by this panel corresponds to the ShD column in Figure 3. (C) Male and female wing shapes have been regressed on size to remove the shape variation caused by allometric scaling. The shape differences remaining in these data are not caused by allometry. The distance between wings is the magnitude of nonallometric shape dimorphism. The deformation represented by this panel corresponds to the NA-ShD column in Figure 3.

Wing-shape has been shown to have genetic variation in all directions of phenotype space in at least two species (Mezey and Houle 2005; Houle and Meyer 2015; Sztepanacz and Blows 2015), and wings can respond to artificial selection in the lab by several genetic standard deviations in a few generations (Weber 1990; Pélabon et al. 2010; Bolstad et al. 2015). Across the Drosophilid phylogeny, however, wing shape evolution is remarkably slow. On average, the rate of wing shape evolution is only 0.7 genetic standard deviations per million years (Houle et al. 2017). So far, we have no satisfactory explanation to reconcile these micro- and macroevolutionary observations. One possibility, however, is that constraints on the evolution of sexual dimorphism may contribute to the broader evolutionary stasis of wings.

In a recent article, we showed that multivariate cross-sex genetic covariances (**B**) can significantly bias the predicted responses to selection in *Drosophila melanogaster* wing shape (Sztepanacz and Houle 2019). Manipulative experiments have also shown this to be the case. When selection on *D. melanogaster* wings was limited to male flies, wings evolved to be more male like in both sexes and females experienced a fitness cost, suggesting that each sex constrains the others phenotype (Abbott et al. 2010). This is consistent with the negative covariance that has been observed between male and female fitness in *D. melanogaster* (Chippindale et al. 2001). Two highlights of our study on cross-sex covariances were that **B** constrained the predicted response to antagonistic selection in (1) the direc-

tion of extant sexual dimorphism in *D. melanogaster*, and (2) the direction of most variation in sexual dimorphism among 74 taxa of Drosophilids (Sztepanacz and Houle 2019). Cross-sex genetic covariances clearly have the potential to constrain the evolution of sexual dimorphism among species, but whether dimorphism evolution is actually constrained across macroevolutionary timescales is unknown.

Most studies of sexual dimorphism evolution have focused on size dimorphism, with a particular interest in the relationship between the evolution of size and size dimorphism, or intraspecific allometry (e.g., studies of Rensch's Rule: Abouheif and Fairbairn 1997; De Lisle and Rowe 2013). There have been many fewer studies that focus on the evolution of shape dimorphism and whether allometry is an important predictor of shape dimorphism, with mixed results regarding the importance of allometry (O'Higgins et al. 1990; Bruner et al. 2005; Schwarzkopf 2005; Bondurianksy 2006). The high-dimensional nature of shape data, however, makes studies of shape dimorphism complicated. In this article, we use geometric morphometrics to study shape dimorphism in the space defined by 12 vein intersections that we use as landmarks on the drosophilid wing (shown in Fig. 1A), focusing on size and three components of dimorphic shape variation.

Figure 1 illustrates how we can decompose the total dimorphism in male and female wings into size-free shape differences, allometric shape differences, and nonallometric shape differences. In each panel, the magnitude of dimorphism is the

Euclidian distance between male and female wings, represented by the black line. Panel A illustrates the observed dimorphism, where male and female wings differ in both size and shape, which we call total shape dimorphism (Total D). Size-independent shape data are obtained by scaling all wings to the same average centroid size, with the result shown in Panel B. We refer to size-independent shape dimorphism in this space as shape dimorphism, ShD. Although shape data are independent of average size, part of the variation in shape data may be due to allometric scaling, which is not accounted for by simply scaling wings to the same average centroid size. To partition ShD into its allometrically and nonallometrically determined components, we use a regression of biological wing size (the square root of the area of the wing) on shape to obtain residual shape data that are independent of the allometric scaling relationship between size and shape. This component of shape dimorphism is shown in Panel C, where male and female wing shapes are oriented along the slope of the regression of size on shape. Allometric shape dimorphism (A-ShD) is the difference between ShD in Panel B and NA-ShD in Panel C.

Using these components of dimorphism, we characterize the evolution of sexual size and shape dimorphism of Drosophilid wings among 82 taxa that have been diverging for at least 33 million years. We start by describing the patterns of size, shape, and nonallometric shape dimorphism among taxa. We then use a phylogenetic comparative approach to identify how sexual dimorphism evolves, and whether it is likely to experience strong evolutionary constraints. Finally, we test the specific hypotheses that (1) variation in shape dimorphism among taxa is a consequence of size dimorphism, or static allometry, and (2) that the allometric relationship between size dimorphism and shape dimorphism observed within taxa evolves across the phylogeny.

Methods
DATA
The data used in these analyses are a subset of the species "R" dataset of 21,138 wings from 117 taxa that was analyzed in Houle et al. (2017). The complete methods used to obtain wing measurements can be found in Houle et al. (2017). Briefly, the x and y coordinates of 12 vein intersections (Fig. 1A) on the left wing of each fly were characterized using a semiautomated system that uses the program Wings 3.72 (Houle et al. 2003) to place landmarks. The data were geometrically aligned within each taxon using Procrustes least squares superimposition, and outliers were identified using the robust Minimum Volume Ellipsoid algorithm. Wings that were more than four robust standard deviations from the mean were excluded. We then pruned these data to the 112 taxa from the family Drosophilidae, and then to those taxa that

had a minimum of 80 measurements for each sex. The resulting dataset is composed of 18,145 wings from 82 taxa, with an average of 110 female and 110 male flies per taxa. Taxon identities and sample sizes are shown in Table S1.

We estimated wing area as the area of a polygon defined by the landmarks that fall along the outer edge of the wing (landmarks 1, 2, 3, 4, 5, 6, and 12), using the surveyor's formula. The square root of wing area was our measure of wing size. This measure of wing size is preferable to centroid size, because it is not affected by the placement of inner landmarks on the wing. We analyzed variation in wings using two different scalings, referred to as size-shape and shape. Size-shape data include the total variation in both size and shape of the wings, as represented in Figure 1A. Shape data describe the shape variation remaining after average size is removed, as represented in Figure 1B.

Size-shape data are obtained by multiplying centroid size (in mm) by the Procrustes-aligned x- and y-coordinates of each landmark. Therefore, size-shape data are in units of millimeters and describe the total variation in size and shape of the wings simultaneously. These data are composed of 24 traits (12 x- and 12 y-coordinates); however, 3 degrees of freedom are used to rotate and center the wings on a common origin, so there are only 21 variable dimensions in size-shape data. We refer to sexual dimorphism in size-shape data as total dimorphism (Total D).

Shape data consist of the Procrustes-aligned landmarks. One additional degree of freedom is used to remove average centroid size, so there are 20 variable dimensions in shape data, and its units are in centroid size. We projected shape data into the 20-dimensional subspace where variation remains after the Procrustes alignment using the 20 phenotypically and genetically variable principal components of wing shape estimated in an independent dataset of *D. melanogaster* (Sztepanacz and Houle 2019). Therefore, our shape data are composed of 20 traits called PC1 to PC20 that are directly comparable among species and sexes. We refer to sexual dimorphism in these data as shape dimorphism (ShD; Fig. 1B).

SEXUAL DIMORPHISM IN WING SIZE AND SHAPE
To determine whether wing size varied among species and the sexes, we tested the effects of sex, species, and their interaction on the natural logarithm of wing size with an ANOVA. Significance was determined using Type III sums of squares. To determine which of the species were size dimorphic, we used an ANOVA to test the effect of sex on wing size separately for each of the 82 taxa.

To determine whether wing shape varied among species and the sexes, we used a MANOVA to test the effects of sex, wing size, species, and all two- and three-way interactions of these variables on our 20 wing shape traits. There was a significant species-by-sex-by-size interaction (see Results) indicating that

shape varied among species and that size predicted shape variation. To determine which species were dimorphic in wing shape (ShD), we used an ANOVA to test the effect of sex on wing shape separately for each of the 82 taxa.

ALLOMETRIC SEXUAL SHAPE DIMORPHISM

To describe the effect of size on shape dimorphism (ShD) within species (i.e., static allometry), we used two approaches that focused on the magnitude and orientation of allometric (A-ShD) versus nonallometric (NA-ShD) shape dimorphism. The difference between these two measures of dimorphism is represented in the contrast between Figures 1B and 1C, which show shape dimorphism (ShD) and nonallometric shape dimorphism (NA-ShD), respectively. ShD is quantified as the Euclidian distance between male and female wing shape phenotypes in each species (Fig. 1B). The Euclidian distance is a linear approximation to Procrustes tangent distance, and when shape differences are small, as in our data, it produces almost identical results. Next, we quantified the ShD attributable to static allometry by using the residuals from a regression of the 20 shape traits on the natural log of the square root of wing area, for each species separately. The Euclidian distance between male and female wing shape phenotypes within each species using these data was the magnitude of NA-ShD (Fig. 1C). Finally, the difference between ShD and NA-ShD was the magnitude of A-ShD.

To identify differences in orientation of ShD and NA-ShD, we calculated the angle between the multivariate trait combination that varied the most among species in ShD and the multivariate trait combination that varied the most among species in NA-ShD. If these two vectors point in the same direction, the pattern of shape dimorphism that arises as a consequence of allometry is similar to the pattern of shape dimorphism that arises from other factors, such as sex differences in selection. If these vectors point in different directions, allometry causes different patterns of shape dimorphism than the other causes of dimorphism. We estimated these multivariate trait combinations using two different multifactor discriminant models. Model 1 estimated ShD, and model 2 estimated NA-ShD. Model 1 was a MANOVA with the 20 wing shape traits as response variables, and three predictors: a main effect of species, a main effect of sex, and a species-by-sex interaction. Model 2 fit the same MANOVA but added the natural log of the square root of wing area as a covariate. From these MANOVAs, we extracted the sums-of-squares and cross product hypothesis matrices for the species-by-sex interaction ($\mathbf{H}$) and for the residual ($\mathbf{E}$). $\mathbf{H}_{max}$, the first eigenvector of the hypothesis matrix, is the trait combination that varies the most in sexual dimorphism among species. The angle between $\mathbf{H}_{max}$ from model (1) and from model (2) indicates the similarity in pattern of ShD and NA-ShD for the most variable sexually dimorphic trait combination.

$\mathbf{E}^{-1}\mathbf{H}_{max}$ is the first eigenvector of $\mathbf{E}^{-1}\mathbf{H}$, whose eigenvalues form the test statistic (Wilks' lambda) used to test whether there is a significant species-by-sex interaction overall. Therefore, $\mathbf{E}^{-1}\mathbf{H}_{max}$ is the statistically supported multivariate trait combination with the most variation in sexual dimorphism among species. The angle between $\mathbf{E}^{-1}\mathbf{H}_{max}$ from model (1) and from model (2) indicates the similarity in pattern of ShD and NA-ShD for the statistically supported most variable trait combination.

EVOLUTION OF SIZE AND SHAPE DIMORPHISM
Single optimum models

We used the Drosophilid phylogeny of van der Linde and Houle (2008) to conduct comparative analyses, modeling the evolution of size, size dimorphism, and shape dimorphism in our data. The unresolved phylogenetic relationship between some of the taxa in our data set, and short or zero branch lengths, caused convergence problems for our models. Therefore, we removed the data from five taxa with the smallest sample sizes within polytomies (*Scaptodrosophila lebanonensis casteeli*, *Scaptodrosophila galloi*, *Drosophila stonei*, *Drosophila pseudoobscura bogotana*, and *Drosophila sulfurigaster*). The evolution of these traits in the remaining 78 taxa were first modeled as Ornstein-Uhlenbeck (OU) processes, in which there was a single primary optimum that the population would approach given unlimited time. The OU process includes a deterministic tendency to adapt toward that optimum, and a stochastic component that represents, among other things, the effects of drift and selection on correlated traits (Hansen 1997). The OU process is represented by the stochastic differential equation:

$$dy = -\alpha(y - \theta(x))dt + \sigma dB,$$

where dy is the change in the trait over time interval dt, α measures how strongly the trait is attracted toward the primary optimum, θ is the optimum state, dB is a white noise process, and σ is the standard deviation of the white noise. Brownian motion is a special case of an OU model in which the stochastic component overwhelms any deterministic process of adaptation toward an optimum. By employing an OU approach, our analyses are able to detect evolutionary regimes ranging from strong stabilizing selection on dimorphism to random drift (i.e., Brownian motion).

We used the R package slouch (Kopperud et al. 2019) to implement these analyses. We analyzed the species means for female and male wing size ($\ln(\sqrt[2]{\text{wing area}})$; data shown in Fig. 2B), size dimorphism ($\ln(\sqrt[2]{\frac{\text{male area}}{\text{female area}}})$) and shape dimorphism (ShD; data shown in Figs. 2C and 3). Estimation error of the means (and ratios of means) was several orders of magnitude smaller than the means, so we did not account for it in these analyses. We also analyzed ShD in the first seven principal components of wing shape (defined in the DATA subsection),
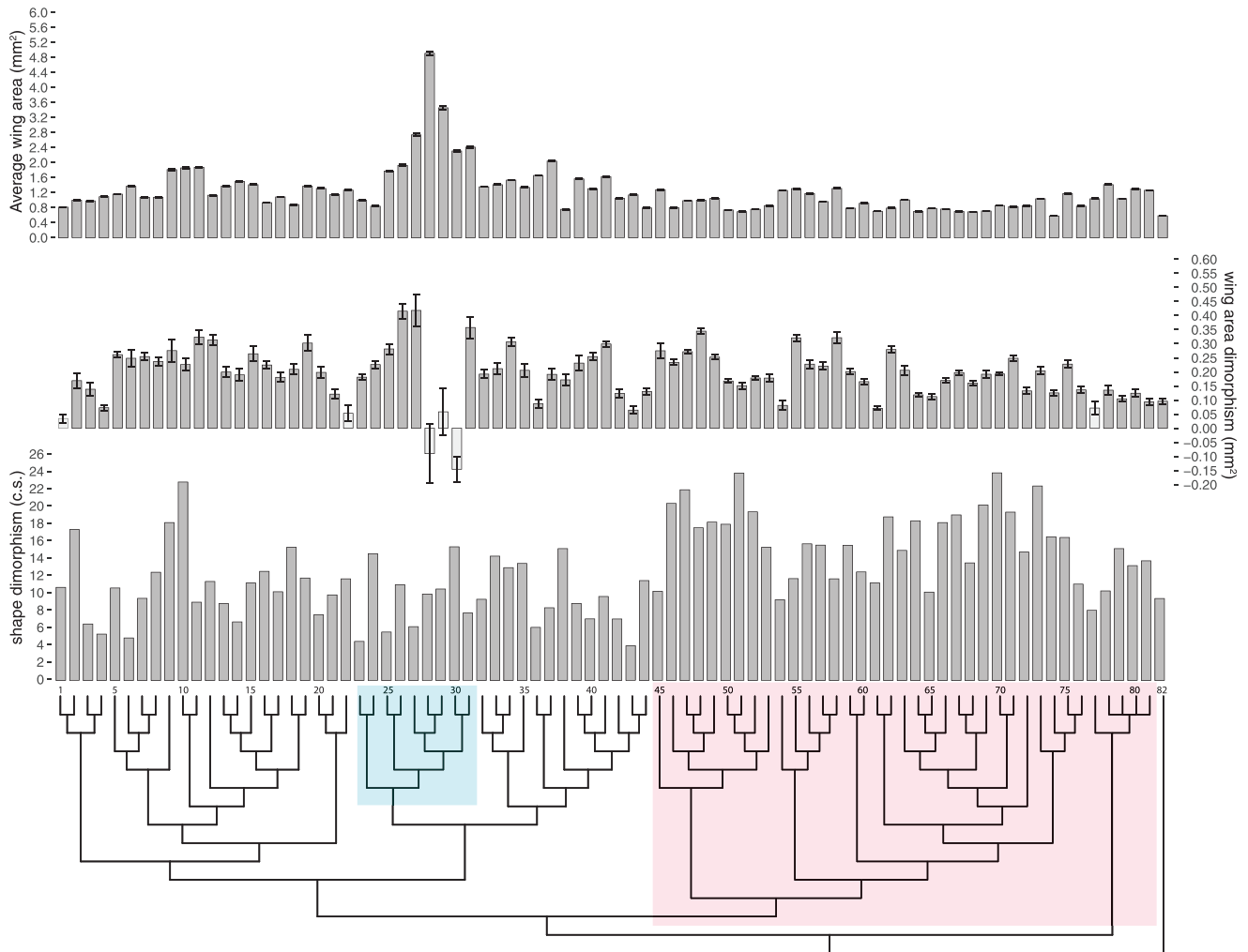
Figure 2. (A) Average wing area across the Drosophilid phylogeny (mm²). The taxon identities are shown in Figure 3. (B) Wing area sexual dimorphism (mm²). Positive values occur when females are larger than males, and negative values when males are larger. The light-colored bars depict the male-female pairs where sex differences in ln($\sqrt[2]{\text{area}}$) were not statistically supported. (C) The magnitude of sexual dimorphism in wing shape (ShD), calculated as the Euclidian distance between male and female wing shapes. ShD is in units of centroid size for each species ($\times$ 1000). Taxa shaded in red correspond to subgenus *Sophophora*, unshaded species to subgenus *Drosophila*, and green-shaded taxa to the Hawaiian *Drosophila*.

as the difference between average male and average female trait value. These seven principal component traits, on average, captured 83% (range: 74–92%) of the variation in total wing shape among the 78 taxa. Estimation error of sexual dimorphism in the principal component traits was larger than in the other traits we analyzed. However, accounting for measurement error in an analysis of the trait with the most error did not have an appreciable effect on the results (see Table S2). Therefore, for computational efficiency, we did not account for measurement error in the analyses presented.

Two specific hypotheses that we wanted to test were (1) whether size dimorphism predicted the evolution of ShD and (2) whether the phenotypic variance in each of the seven principal component traits predicted the evolution of ShD. Therefore, we

also fit size dimorphism and the phenotypic variance in male and female traits as linear predictors in our models. These continuous predictors were assumed to follow a Brownian motion process of evolution (Hansen et al. 2008; Labra et al. 2009).

Shifting optimum models

The phylogenetic comparative methods described above were developed to test specific hypotheses about evolution, for example, whether species that occupy different niches adapt toward different evolutionary optima. Aside from the two specific hypotheses we state above, we have no such hypotheses for the evolution of sexual dimorphism in Drosophilid wings. An alternative approach is then to estimate the dynamics of phenotypic optima directly from the data without a priori hypotheses (Ingram and
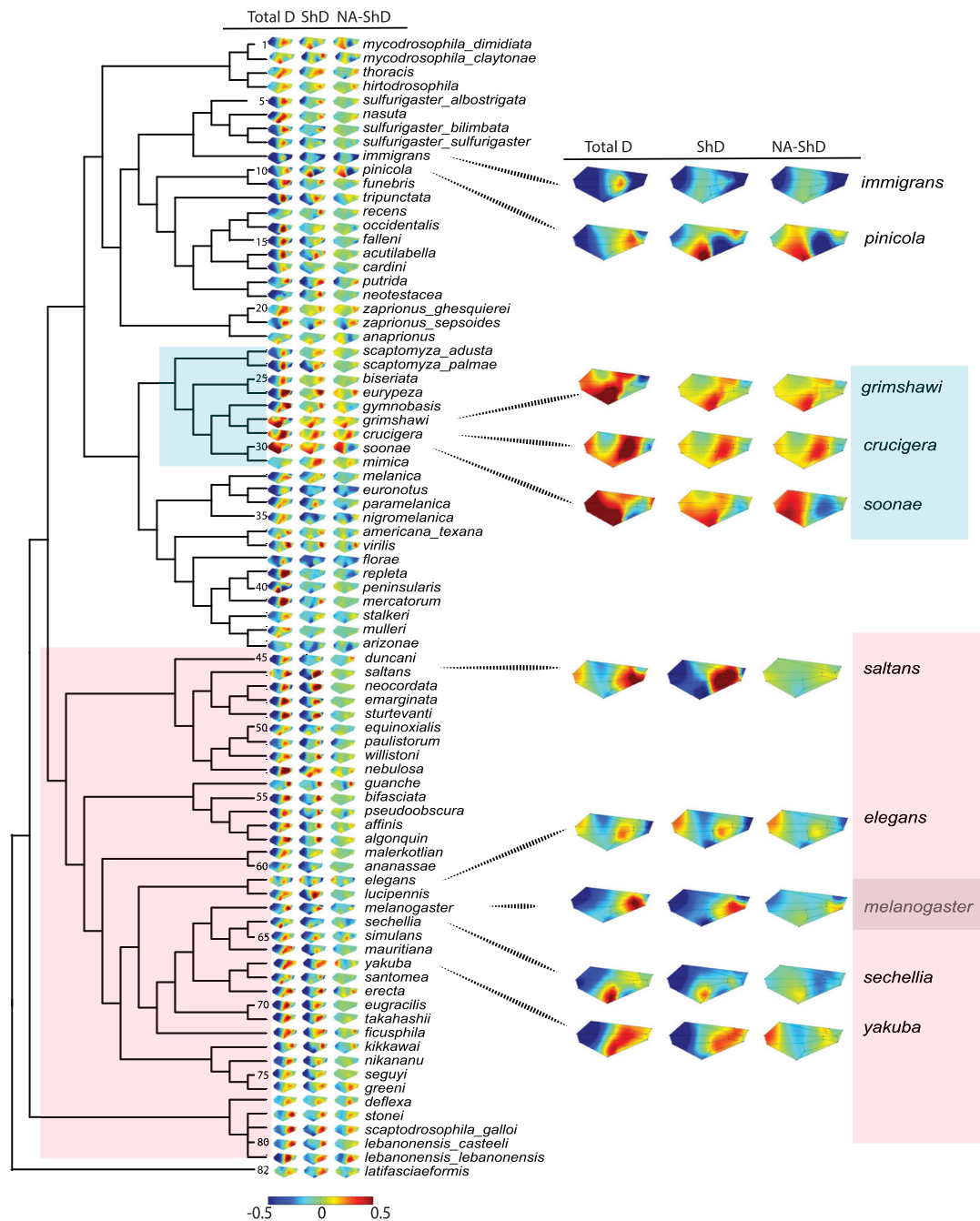
Figure 3. Wing shape dimorphism on the Drosophilid phylogeny. Taxa shaded in red correspond to the subgenus *Sophophora*, unshaded taxa correspond to the paraphyletic subgenus *Drosophila*, and green shaded taxa correspond to the Hawaiian *Drosophila*. Each picture represents the shape deformations of the average female wing compared to the average male wing for each taxon, with warm colors depicting areas of local expansion and cool colors depicting areas of local contraction. Shape deformations are on a log₂ scale where a value of 0 indicates no change, −0.5 is a local halving of the area, and 1 is a local doubling. Deformations are magnified by 5× in all cases. Wings of the dimorphic taxa that are discussed in the text have been expanded for easier viewing. Dimorphism in *D. melanogaster* is also shown for reference. The first column of wings, Total Dim, shows sexual dimorphism for each taxon encompassing sex differences in both size and shape. Wings were scaled by the average centroid size of all 82 taxa prior to plotting the deformations, therefore this column is not comparable in scale to ShD and NA-ShD. The second column, ShD, is sexual dimorphism in wings after scaling the data by centroid size using Procrustes alignment. It represents dimorphism in both allometric and nonallometric shape. The third column, NA-ShD, is the nonallometric component of ShD, or the dimorphism that is not due to a scaling relationship with wing size within taxa. The scales of deformation in ShD and NA-ShD are directly comparable to each other.

Table 1. ANOVA testing the effects of species, sex, and their interaction on the natural logarithm of wing size. Estimates of effect sizes for regressions in each species are given in Table S2.

Effect	d.f.	Type III Sums of squares	*F*	*P*
Species	82	649.55	2483.742	<0.001
Sex	1	31.90	10,125.747	<0.001
Species*Sex	82	6.53	24.975	<0.001

d.f. = degrees of freedom.

Mahler 2013). We used this approach to fit multi-optima OU models to determine whether there is evidence for shifting adaptive optima (θ) of size (data shown in Fig. 2B) and shape (ShD; data shown in Figs. 2B and 4) dimorphism across the phylogeny. Briefly, we used a Bayesian approach that modifies the standard OU model to implement a reversible-jump algorithm that estimates the number, location, and magnitude of shifts in adaptive optima from the data, while jointly sampling the standard OU parameters. For details of model, see Uyeda and Harmon (2014). We fit these models for the same traits as the single optimum models, using the R package bayou. The single optimum analyses (see Results) showed that male and female size evolution can best be described by a Brownian motion process. Therefore, we excluded male and female size from this analysis, and focus on size and shape dimorphism. We specified relatively uninformative priors for our models using a conditional Poisson distribution with a maximum number of shifts in optima equal to half the number of tips in the phylogeny. The priors on the adaptive optima were normally distributed with a mean and variance defined by the mean and variance observed in our data, an equal probability of each branch having a shift, and a maximum of one shift allowed per branch.

The results of the single optimum models showed that size dimorphism was important in predicting shape dimorphism, ShD. To determine whether the relationship between size dimorphism and ShD evolved across the phylogeny, we tested the fit of three models, described in the Results, for the evolution of the relationship between size dimorphism and ShD (Uyeda et al. 2017). For these analyses, we used an informative prior on the relationship between size dimorphism and ShD, specifying a normal distribution around 0.7. This was the correlation between size dimorphism and shape dimorphism among species that we observed in our data (see Fig. 4B).

Results
SEXUAL DIMORPHISM IN WING SIZE AND SHAPE
We found a species-by-sex interaction for wing size (Table 1; Fig. 2), indicating that sexual size dimorphism varied among species. On average, male wings were 0.92 mm smaller than female wings, with statistical support for size dimorphism in 80 out of 82 taxa at $\alpha = 0.05$ (Table S3). In every taxon with statistically supported size dimorphism, female wings were larger than male wings. Variation across the phylogeny in wing size and size dimorphism is shown in Figures 2A and 2B, respectively.

Multivariate wing shape was also sexually dimorphic. A species-by-sex-by-size interaction (Table 2) indicated that shape was sexually dimorphic, that it varied among species, and that it also varied with wing size. To determine which taxa were driving this pattern, we estimated the effect of sex, size, and their interaction, in each taxon separately. Of the 82 taxa we studied, the sex-by-size interaction was significant in 40 of them at $\alpha = 0.05$ (Table S4). Variation in the magnitude of shape dimorphism, ShD ($\times 1000$), among taxa is shown in Figure 2C.

The pattern of shape dimorphism in each taxon is shown in Figure 3. The first column of wings shows total dimorphism in the original 12 *x* and *y* coordinate traits (Total D). For presentation, the wing sizes in this figure were scaled to the average centroid size of all species; however, we are still defining biological size by wing area rather than centroid size. Due to the scaling, Total D deformations in this figure are not on an absolute scale, but are relative to each other. In many cases, the distal portion of the wing tip is contracted in females compared to males and the interior proximal region of the wing is expanded. The second column of wings shows shape dimorphism (ShD) of wings after scaling by centroid size within each species. The magnitude of contraction and expansion cannot be directly compared between Total D and ShD, but the patterns of shape deformations can, and in many cases they remain similar. For example, the pattern of contraction in the wing tip of female flies evident in the Total D column is also common in the ShD column. This similarity is not universal, and we highlight one exception in *Drosophila saltans*, where the pattern of dimorphism is opposite when comparing Total D to ShD. The distal area of female wings is expanded in size compared to males; however, once the wings are scaled by centroid size to get size-independent ShD, the shape of this area of the wing is contracted.

Figure 4. (A) Allometric and nonallometric dimorphism in wing shape that is determined by size. Wing shape dimorphism was defined as the Euclidian distance between male and female wing shapes (×1000). There is a weak relationship between nonallometric shape dimorphism and dimorphism in wing size, whereas wing size strongly predicts allometric dimorphism in wing shape. (B) The proportion of shape dimorphism determined by size dimorphism.

Table 2. MANOVA (Type III) testing the effects of species, sex, the natural logarithm of wing size, and their interactions, on wing shape. Wing shape traits were the scores of scaled landmark coordinates on the 20 variable principal components of wing shape in an independent population of *D. melanogaster*. Landmark coordinates were scaled by dividing each *x* and *y* coordinate by the centroid size for an individual, and multiplying by 1000. The principal components were calculated from measurements on over 17,000 wings from Sztepanacz and Houle (2019).

Effect	Pillai's trace	Approx. F	Num d. f.	Den d.f.	P
Species	4.9490	70.55	1660	356,180	<0.001
Sex	0.0038	3.41	20	17,790	<0.001
ln(size)	0.0062	5.57	20	17,790	<0.001
Species*Sex	0.3286	3.58	1660	356,180	<0.001
Species*ln(size)	0.5136	5.66	1660	356,180	<0.001
Sex*ln(size)	0.0023	2.05	20	17,790	0.0037
Species*Sex*ln(size)	0.1647	1.78	1660	356,180	<0.001

ALLOMETRIC SEXUAL SHAPE DIMORPHISM

The species-by-sex-by-size interaction suggested that at least some part of ShD we observed was due to the relationship between size and shape within species, or static allometry. We, therefore, separated ShD in each species into allometric (A-ShD) and nonallometric (NA-ShD) dimorphism as represented in Figure 1. On average, the magnitude of ShD was 0.013 in units of centroid size (shown in Fig. 2C ×1000), and the magnitude of NA-ShD was 0.007 in units of centroid size. Therefore, 53% of ShD, on average, was due to allometry.

Figure 3 shows the patterns of ShD and NA-ShD in columns two and three, respectively. Qualitatively, ShD and NA-ShD appear similar in most cases. To highlight one exception as an example, ShD and NA-ShD are opposite in effect on *Drosophila yakuba* wings. Females have a contracted wing tip and expanded proximal interior when considering ShD, but NA-ShD expands the female wing tip and has limited effects on the interior of the wing. We quantified the average similarity in orientation of ShD and NA-ShD effects for two multivariate trait combinations: the multivariate trait combination that is the most variable in sexual dimorphism ($\mathbf{H}_{max}$) among species, and the statistically supported multivariate trait combination that is most variable in sexual dimorphism ($\mathbf{E}^{-1}\mathbf{H}_{max}$). In these data, $\mathbf{H}_{max}$ and $\mathbf{E}^{-1}\mathbf{H}_{max}$ have a vector correlation of 0.20 to each other. If ShD and NA-ShD have similar phenotypic effects, the angle between their effect vectors would be 0; if they have unrelated effects, the angle would be 90. The angle between ShD and NA-ShD for $\mathbf{H}_{max}$ was 78 degrees, and for $\mathbf{E}^{-1}\mathbf{H}_{max}$ was 34 degrees, suggesting that there is some similarity in the allometric and nonallometric components of shape dimorphism, particularly in the direction in which dimorphism is best supported statistically.

ShD and NA-ShD were both variable among species (Fig. S2; Table S5), as was size dimorphism (Fig. 2B). The relationship between ShD and NA-ShD and size, however, was very different. Shape dimorphism in species that were highly size dimor-

phic was almost all a consequence of allometric scaling, whereas shape dimorphism in species that were not size dimorphic was mostly nonallometric (Fig. 4A). The magnitude of shape dimorphism that was caused by size was strongly correlated with the magnitude of size dimorphism ($\rho = 0.73$, $P < 0.01$) (Fig. 4B).

EVOLUTION OF SIZE AND SHAPE DIMORPHISM

The evolution of male and female wing size was most consistent with a Brownian motion process. In both sexes, the phylogenetic half-life in the OU model exceeded the depth of the phylogeny, indicating a persistent signal of ancestral state, as expected under Brownian motion (Table 3). Males evolved 0.0259 mm per million years, with females evolving at almost the same rate of 0.0255 mm per million years.

The evolution of sexual size dimorphism (square roots of the values shown in Fig. 2B) was more consistent with an OU process, with a phylogenetic half-life of about 13 million years, or 40% of the depth of the tree. Average size significantly predicted size dimorphism, and explained 11% of its variation (Table 3). Bayou models that estimated the number of adaptive optima for size dimorphism provided marginal evidence for three shifts in optimal size dimorphism from the ancestral state (Fig. 5A). The subgenera *Drosophila* (root $\theta = 0.084$; 95% Higher Posterior Density interval (HPD): 0.0063, 0.105) and *Sophophora* (posterior probability = 0.20, $\theta = 0.114$, SE = 0.0007) were estimated to have different optima. Within the *Drosophila* subgenus, the sister species *Drosophila grimshawi* and *Drosophila crucigera* (posterior probability = 0.21, $\theta = -0.055$, SE = 0.003) shared a unique optimum, as did *Drosophila soonae* (posterior probability = 0.25, $\theta = -0.075$, SE = 0.003). *Drosophila grimshawi*, *D. crucigera*, and *D. soonae* are all in the Hawaiian species group, which is the only group where female wings are smaller than male wings (Fig. 2B).

The evolution of total ShD (the Euclidian distance between male and female wing shapes; values shown in Fig. 3C) and ShD

Table 3. Parameter estimates for single-optimum OU models of trait evolution. Female and male wing size is the natural log of the square root of wing area in each sex. Size dimorphism is the natural log of female size divided by male size. Average size is the natural log of the average of male and female sizes. Total shape dimorphism is the Euclidian distance between male and female wing shapes in units of centroid size (×1000). The table shows the maximum likelihood estimates and the 2-unit support intervals in parentheses for the phylogenetic half-life ($t_{1/2}$) in units of million years (the maximum branch length of the tree is 33 million years). The equilibrium stochastic variance, v_y, is in units of trait squared, and the intercept and slope are in units of the trait.

| Trait and predictor	$t_{1/2}$ (support region)	v_y	Intercept (±SE)	Slopes (±SE)	R^2	AICc	ΔAICc
Size							
Female wing size	38.38 (15.15, ∞)	0.04	0.068 (0.054)	...	...	...	
Male wing size	42.42 (16.16, ∞)	0.05	−0.013 (0.059)	...	...	...	
Size dimorphism	12.93 (7.27, 41.21)	0.00	0.082 (0.009)	...	...	−303.00	
average size	8.79 (2.43, 25.15)	0.00	0.089 (0.007)	−0.124 (0.040)	0.11	−308.00	5
ShD							
Total shape dimorphism	9.33 (6.00, 18.00)	24.55	12.355 (0.894)	...	...	460.00	
size dimorphism	6.00 (2.00, 12.00)	18.18	8.28 (1.352)	67.778 (18.248)	0.15	452.00	8
PC1 dimorphism	11.33 (7.33, 24.00)	37.27	6.606 (1.227)	...	...	489.00	
size dimorphism	7.33 (3.33, 16.67)	14.55	−1.713 (1.499)	147.288 (19.357)	0.42	449.00	40
PC2 dimorphism	0.00 (0.00, 6.00)	25.46	0.113 (0.571)	...	...	480.00	
size dimorphism	1.33 (0.00, 6.67)	25.46	1.422 (1.420)	−16.749 (15.814)	0.01	481.00	1
PC3 dimorphism	9.33 (5.33, 20.67)	14.55	3.102 (0.688)	...	...	420.00	
size dimorphism	9.29 (4.85, 21.01)	14.55	3.150 (1.184)	−0.903 (18.378)	0.00	422.00	2
PC4 dimorphism	8.00 (3.33, 18.67)	12.73	−0.849 (0.590)	...	...	412.00	
size dimorphism	8.89 (3.64, 26.26)	12.42	0.733 (1.108)	−29.962 (16.848)	0.04	412.00	0
PC5 dimorphism	10.00 (5.33, 20.00)	9.09	−0.635 (0.566)	...	...	379.00	
size dimorphism	8.48 (4.04, 18.59)	7.27	1.439 (0.869)	−38.259 (12.850)	0.10	373.00	6
PC6 dimorphism	8.67 (5.33, 17.33)	8.18	−1.170 (0.495)	...	...	377.00	
size dimorphism	7.68 (3.23, 14.95)	7.27	0.429 (0.845)	−28.279 (12.316)	0.06	374.00	3
PC7 dimorphism	2.00 (0.00, 5.33)	2.43	−0.558 (0.184)	...	...	297.00	
size dimorphism	0.67 (0.00, 4.67)	2.22	0.465 (0.418)	−12.195 (4.507)	0.09	293.00	5
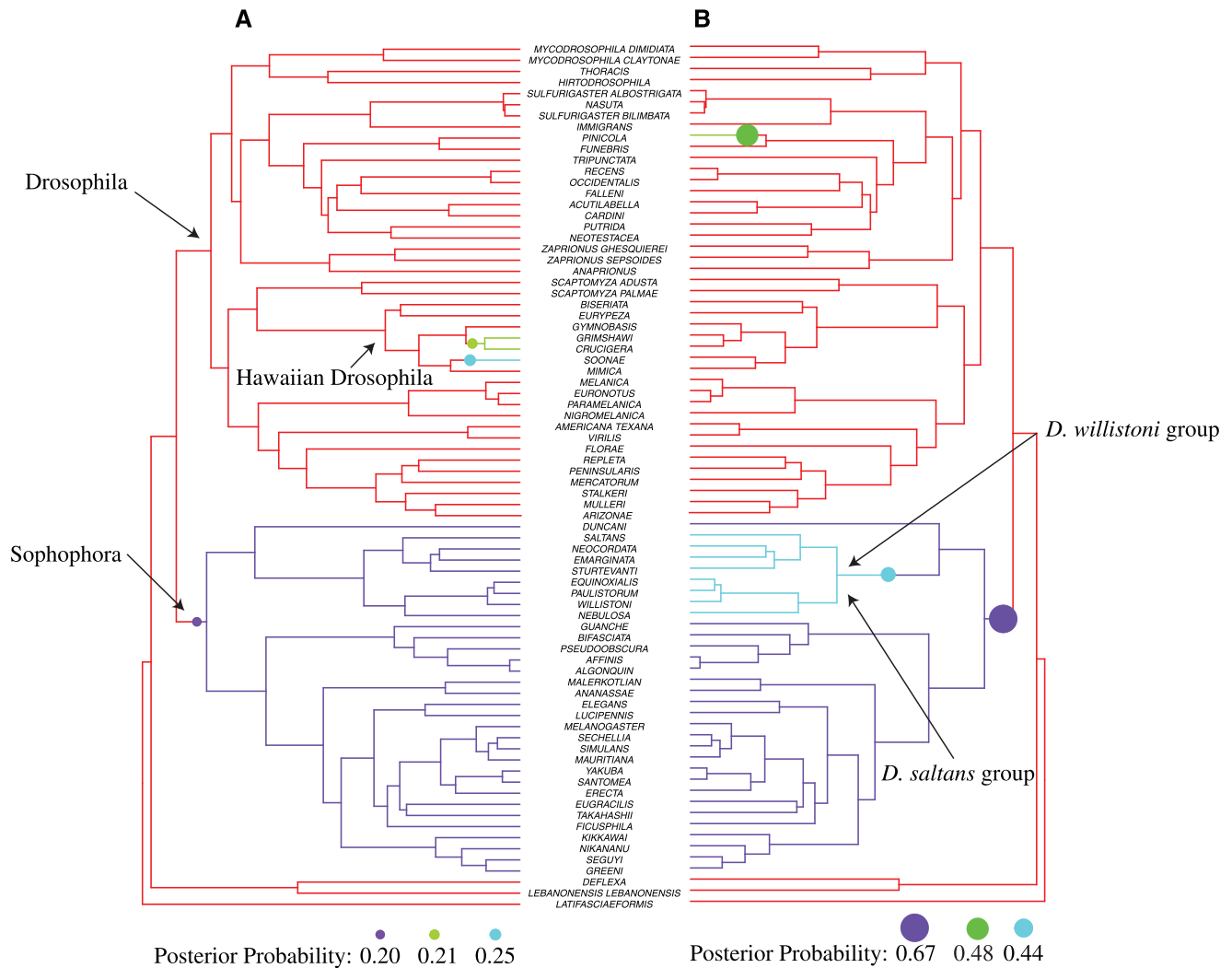
Figure 5. (A) Evolutionary optima for size dimorphism from a Bayesian reversible jump OU model. Red taxa share the estimated optimum at the root, whereas purple, blue, and green taxa share different optimal values of size dimorphism. (B) Evolutionary optima for shape dimorphism. Red taxa share the estimated root optimum, whereas purple, blue, and green taxa share different optimal values of shape dimorphism. In both panels, circles denote where a shift in optimal value occurs, with the area of the circle proportional to the posterior probability of the shift from the Bayesian model. The posterior probabilities are shown below each panel.

dimorphism in each of the seven principal components of wing shape were all consistent with an OU processes of evolution. Both ShD and ShD in the trait PC1 were significantly predicted by size dimorphism. Size dimorphism explained 15% of the variation in total ShD, and 42% of the variance in ShD for the trait PC1 (Table 3). Size dimorphism was not a significant predictor of ShD for any of the other shape traits. Adding the phenotypic variance in male and female traits as predictors to our models did not substantially improve the fit of any of the models; phenotypic variance did not significantly predict the evolution of ShD (Table S6).

On average, the phylogenetic half-life of total ShD was 6.2 million years, indicating that the phylogenetic signal is lost relatively quickly. There was strong evidence (posterior prob-

ability of 0.67) for two optima for total ShD, which correspond to the *Drosophila* (root $\theta = 10.68$; 95% HPD: 8.63-12.78) and *Sophophora* ($\theta = 16.48$, SE: 0.026) subgenera (Fig. 5B). There was weaker evidence (posterior probability = 0.48) that *Drosophila pinicola* had a unique optimum of $\theta = 28.83$ (SE: 0.153), and also some evidence (posterior probability = 0.44) that the *Drosophila saltans* and *Drosophila willistoni* species groups differed from the rest of *Sophophora* with $\theta = 19.57$ (SE = 0.042) (Fig. 5B).

As stated above, size dimorphism was an important predictor for the evolution of total ShD and ShD in the trait PC1. To determine whether the relationship between size dimorphism and shape dimorphism evolved across the phylogeny, we tested the fit of three models for the evolution of the relationship between

size dimorphism and shape dimorphism. We compared the fit of models estimating a single slope and intercept for the relationship between size dimorphism and shape dimorphism (model 0), a single slope but multiple intercepts (model 1), and a multiple slopes and intercepts (model 2). The evolution of the relationship between size dimorphism and shape dimorphism among species was best explained by a single slope and intercept for total ShD (Bayes factors: $H_{0 \text{ vs. } 1} = 0.98$, $H_{0 \text{ vs. } 2} = 1.00$, $H_{1 \text{ vs. } 2} = 1.02$), and a single slope and intercept for ShD in the trait PC1 (Bayes factors: $H_{0 \text{ vs. } 1} = 0.98$, $H_{0 \text{ vs. } 2} = 1.00$, $H_{1 \text{ vs. } 2} = 1.01$). Therefore, the allometric relationship between shape dimorphism and size dimorphism is conserved among species.

Discussion

Sexual dimorphism is widely viewed as adaptive, reflecting the evolution of males and females to divergent fitness optima. The evolution of sexual dimorphism, however, may be constrained by cross-sex genetic correlations that can prevent or reduce the responses to divergent selection (Bonduriansky and Chenoweth 2009; Cox and Calsbeek 2009), and by allometry (Bruner et al. 2005; Schwarzkopf 2005; Gidaszewski et al. 2009). Whether dimorphism evolution is actually constrained across macroevolutionary timescales is largely unknown. In this work, we investigated the evolution of sexual size and shape dimorphism in the wings of 82 species of Drosophilids that have been diverging for at least 33 million years. We found that (1) shape dimorphism among species was qualitatively similar in many cases; (2) size and shape dimorphism evolve on moderate timescales of about 10 million years and have adapted to different optimal values in different clades; (3) within species, static allometry partly predicts shape dimorphism; and (4) among species, a single allometric relationship between size dimorphism and shape dimorphism is conserved across the 33 million years of divergence among these species.

Although we can back up each of these declarative statements with the results of specific analyses, it is also important to realize that our results are extraordinarily complex in detail. A glance at the species-specific dimorphisms represented in Figure 3, and particularly the exceptional examples expanded on the right side of the figure, shows a great deal of variation. We can detect this complexity because we have characterized wing size and shape precisely, and in great detail, across a wide diversity of species. Another major result is the very complexity that any simple list of results glosses over. In this high-dimensional system, there are patterns, but important variation around those patterns as well.

Shape dimorphism (Total D and ShD) tends to fall along the proximo-distal axis of the wings, with females typically having a shorter wing that is wider in the middle. There were, however, several exceptions to this pattern, including in the Hawaiian group, and several other species such as *D. elegans*, *D. pinicola*, and *D. immigrans*, among others (Fig. 3). Consistent with these exceptions, we found evidence for quantitative variation among species in both sexual size and shape dimorphism (Table 1; Fig. 3). In a previous study of eight species in the *D. melanogaster* subgroup, Gidaszewski et al. (2009) found no dominant pattern of wing shape dimorphism. The first principal component of among-species shape dimorphism in their data accounted for only 25% of the total variance in dimorphism, and there was no regular configuration of species divergence in shape space for the first three principal components of dimorphism that they studied. They used a phylogenetic permutation test to look for evidence that shape dimorphism was structured phylogenetically, but also found no evidence for this. Our dataset of 82 species, which includes five of the species they studied, provides substantially more power to quantify variation in sexual dimorphism and to quantitatively study how dimorphism evolves on macroevolutionary timescales.

We used a phylogenetic comparative approach to model adaptive evolution of size dimorphism and shape dimorphism, first as an Ornstein-Uhlenbeck (OU) process to a single primary optimum (Hansen 1997). One of the key parameters derived from the OU model is the phylogenetic half-life, or the time it takes to evolve half the distance from an ancestral state to a new optimum. For most of the ShD traits we analyzed, the half-lives were between 7 and 11 million years (Table 3). Two exceptions were evolution in the traits PC2 and PC7, which adapted to their primary optima almost instantaneously. Therefore, the evolution of ShD is not entirely influenced by ancestral state, although a half-life of 11 million years equates to a substantial 33% of the depth of the tree (Table 3). Our multi-optima models suggested that total ShD evolved to three different optima over the 33 million years that these species have been diverging (Fig. 5B). The fact that different optima are supported for the *Drosophila* and *Sophophora* subgenera is perhaps not surprising as there are a large number of species in each of these groups, and therefore high power to detect differences between them. There was also moderate evidence for a distinct evolutionary optimum for the *D. willistoni* and *saltans* species groups. Analyses of each of the seven ShD traits we studied (PC1-PC7) indicate that PC3 and PC5 are driving this result: a unique optimum was estimated for this same group for both of these individual traits (Fig. S1). Ecological factors may be responsible for these shifts in optima. A comparative study of thermal tolerance found that *D. willistoni* and *saltans* had a relatively high critical thermal maximum compared to their relatives (Kellermann et al. 2012), which may reflect adaptation to higher temperatures. Temperature is known to affect wing shape in *Drosophila* (Ray et al. 2016), but how it affects dimorphism is largely unknown. In one study of midges, wing

dimorphism was associated with habitat and flight differences between the sexes (McLachlan 1986), and adaptation to different ecological niches has also been invoked to explain morphological sexual dimorphism in *Anolis* lizards (Butler et al. 2007).

The evolution of ShD to different optima suggests that dimorphism evolution in Drosophilids may be more labile than the evolution of shape itself. For example, Houle et al. (2017) estimated low rates of evolution for 20 sex-averaged wing shape traits in a superset of the data we analyzed here. Similarly, Gidaszewski et al. (2009) found no phylogenetic structure of shape dimorphism in their analysis of the *melanogaster* subgroup, but they found a strong phylogenetic signal in the sex-averaged shape data (Klingenberg and Gidaszewski 2010). When studying the evolution of shape or shape dimorphism, a critical consideration is the relationship between size and shape, or allometry. Allometric relationships are often strongly conserved across evolutionary time, leading to a general hypothesis that allometry constrains long-term evolutionary change (Houle et al. 2019). For example, Bolstad et al. (2015) found that the slope of the relationship between size and shape of Drosophilid wings was similar among species and evolved exceedingly slowly over macroevolutionary timescales, suggesting that shape evolution in wings may be dictated by size evolution. There are few studies that address the consequences of allometry for shape dimorphism evolution. In a study of western green lizards, ontogenetic allometry explained much of the variation in head-scale shape dimorphism, whereas in water skinks sexual shape dimorphism of morphology was found between size-monomorphic adults and neonates (Schwarzkopf 2005). Therefore, ontogenetic allometry does not explain all of the variation in shape dimorphism. O'Higgins et al. (1990) found that sexual shape dimorphism in hominid skulls varied among species, and that the most shape-dimorphic species were those whose skulls had the largest sex differences in size. Therefore, much of the variation in shape dimorphism in hominids may be a consequence of static allometry.

Here, we first looked at how size dimorphism evolved across the phylogeny and found its evolution toward an optimum was slow (Table 3). This was not surprising, because we found no evidence that sex-specific size adapted toward a single optimum (Table 3), and low rates of wing size evolution have previously been observed in Drosophilids (Houle et al. 2017). When we added size dimorphism as a linear predictor in our OU models of shape dimorphism evolution, we found that size dimorphism predicted the evolution of total ShD and ShD in the trait PC1. These are the two traits with the most phenotypic variation, including the most allometric variation. Size dimorphism, however, had a marginal effect on the rate of ShD evolution for most traits (Table 3). Our results show that size dimorphism predicts the evolution of shape dimorphism in some multivariate trait combinations, but not all.

If the allometric relationships between size dimorphism and shape dimorphism themselves evolve in different species, we may not see the effect of shape dimorphism on size dimorphism evolution. This was the interpretation offered by Gidaszewski et al. (2009) to explain their observation that the shape features associated with allometry differed among species. We found this was unlikely to be the case in our data. Our evolutionary models that fit a single intercept and slope for the regression of size dimorphism on shape dimorphism fit better than those that allowed for separate intercepts, and those that allowed for separate intercepts and separate slopes. Our results suggest that allometry is a persistent evolutionary constraint for dimorphism evolution in certain directions of phenotype space, but that dimorphism evolution could occur in many other trait combinations that are not constrained by allometry.

In the directions of dimorphism evolution that are not constrained by allometry, genetic architecture may have a larger role in determining dimorphism evolution. High genetic correlations between male and female wing traits have been invoked to explain the remarkable conservatism of sexual shape dimorphism evolution across latitudinal clines that have evolved separately on three continents (Gilchrist et al. 2000). However, over longer evolutionary timescales these constraints may not be relevant. In a previous article, we found that cross-sex genetic covariances of wing shape traits can limit the evolution of sexual shape dimorphism, but that there is genetic variation that would allow slow changes in dimorphism in many aspects of shape (Sztepanacz and Houle 2019). Perhaps the phylogenetic half-lives for ShD evolution of 7–11 million years, that we found for most traits here, reflect the timescale of constraints imposed by **B** on dimorphism evolution. However, understanding how microevolutionary processes lead to macroevolutionary divergence is an ongoing challenge (Uyeda et al. 2011).

Selection will also have a role in the evolution of wing sexual dimorphism. An important component of male courtship is the wing song they produce (e.g., Hoy et al. 1988; Greenspan and Ferveur 2000; Talyn and Dowse 2004; Routtu et al. 2007; Giglio and Dyer 2013). One study that collected flies from copulating pairs found that copulating males had longer wings than the males collected without mates (Taylor and Kekic 1988), suggesting longer wings confer higher male fitness. Artificial selection to increase male wing length has also been shown to confer a mating advantage to males, after controlling for differences in body size (Menezes et al. 2013). Although females may not be subject to sexual selection on wings, functional constraints for flight could have a more important role in this sex (Ray et al. 2016). Larger wings are required to lift the generally heavier bodies of female flies, and wing aspect ratio (length relative to area) has been shown to be correlated with flight performance (Azevedo et al. 1998). Morphometric analyses of wing shape have also shown

that multivariate shape is correlated with *D. suzukii* flight performance at different temperatures (Fraimout et al. 2018). However, we are not aware of any studies that have focused on how wing shape differences between the sexes influence their relative flight performance. Stronger directional sexual selection on male wings paired with stronger stabilizing natural selection on female wings potentially explains our observations of longer winged males and wider winged females. We make this suggestion cautiously, however, as we have limited knowledge of the wing sizes or shapes that selection favors in either sex of Drosophilids.

We have shown that sexual dimorphism in Drosophilid wing shape has a consistent qualitative pattern among species that occupy a variety of ecological niches (Kellermann et al. 2012), and presumably have a variety of wing songs that are favored by sexual selection (Hoikkala et al. 1998; Hoy et al. 1988; Routtu et al. 2007). Despite the qualitative conservatism we observed, we have also shown wing shape dimorphism can and does evolve over moderate timescales to different evolutionary optima, and that some shape dimorphism evolution is a consequence of the allometric relationship between size dimorphism and shape dimorphism. The evolution of allometrically determined shape dimorphism is likely to be conserved across evolutionary time as a consequence of the conserved allometric relationship between size dimorphism and shape dimorphism. Evolution of nonallometric shape dimorphism may be dictated to a larger extent by selection and genetic architecture, and may be more evolutionary labile. Until we know which trait combinations are actually targets of selection, it is difficult to identify the degree to which sexual dimorphism is adaptive or a consequence of constraints. Studies of natural and sexual selection on wings, together with clearer understanding of how microevolutionary parameters of evolvability relate to macroevolutionary patterns of divergence, will be the key for disentangling alternative explanations for the evolution of dimorphism.

LITERATURE CITED

Abbott, J. K., S. Bedhomme, and A. K. Chippindale. 2010. Sexual conflict in wing size and shape in *Drosophila melanogaster*. J. Evol. Biol. 23:1989–1997.

Abouheif, E., and D. J. Fairbairn. 1997. A comparative analysis of allometry for sexual size dimorphism. Am. Nat. 149:540–562.

Azevedo, R. B. R., A. C. James, J. McCabe, and L. Partridge. 1998. Latitudinal variation of wing:thorax size ratio and wing-aspect ratio in *Drosophila melanogaster*. Evolution 52:1353–1362.

Bolstad, G. H., J. A. Cassara, E. Márquez, T. F. Hansen, K. van der Linde, D. Houle, and C. Pélabon. 2015. Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. 112:13284–13289.

Bonduriansky, R. 2006. Convergent evolution of sexual shape dimorphism in Diptera. J. Morphol. 267:602–611.

Bonduriansky, R., and S. F. Chenoweth. 2009. Intralocus sexual conflict. Trends. Ecol. Evol. 24:280–288.

Bruner, E., D. Costantini, A. Fanfani, and G. Dell'Omo. 2005. Morphological variation and sexual dimorphism of the cephalic scales in *Lacerta bilineata*. Acta Zool. 86:245–254.

Butler, M. A., S. A. Sawyer, and J. B. Losos. 2007. Sexual dimorphism and adaptive radiation in Anolis lizards. Nature 447:202–205.

Cheng, C., and D. Houle. 2020. Predicting multivariate responses of sexual dimorphism to direct and indirect selection. The American Naturalist 196(4):391–405.

Chippindale, A. K., J. R. Gibson, and W. R. Rice. 2001. Negative genetic correlation for adult fitness between sexes reveals ontogenetic conflict in *Drosophila*. Proc. Natl. Acad. Sci. USA 98:1671–1675.

Collet, J. M., S. Fuentes, J. Hesketh, M. S. Hill, P. Innocenti, E. H. Morrow, K. Fowler, and M. Reuter. 2016. Rapid evolution of the intersexual genetic correlation for fitness in *Drosophila melanogaster*. Evolution 70: 781–795.

Connallon, T., and A. G. Clark. 2014. Evolutionary inevitability of sexual antagonism. Proceedings of the Royal Society B: Biological Sciences 281(1776):2013–2123.

Cox, R. M., and R. Calsbeek. 2009. Sexually antagonistic selection, sexual dimorphism, and the resolution of intralocus sexual conflict. Am. Nat. 173:176–187.

De Lisle, S. P., and L. Rowe. 2013. Correlated evolution of allometry and sexual dimorphism across higher taxa. Am. Nat. 182:630–639.

Foerster, K., T. Coulson, B. C. Sheldon, J. M. Pemberton, T. H. Clutton-Brock, and L. E. B. Kruuk. 2007. Sexually antagonistic genetic variation for fitness in red deer. Nature 447:1107–1110.

Fraimout, A., P. Jacquemart, B. Villarroel, D. J. Aponte, T. Decamps, A. Herrel, R. Cornette, and V. Debat. 2018. Phenotypic plasticity of *Drosophila suzukii* wing to developmental temperature: implications for flight. J. Exp. Biol. 221:jeb166868.

Gidaszewski, N. A., M. Baylac, and C. Klingenberg. 2009. Evolution of sexual dimorphism of wing shape in the *Drosophila melanogaster* subgroup. BMC Evol. Biol. 9:110–111.

Giglio, E. M., and K. A. Dyer. 2013. Divergence of premating behaviors in the closely related species *Drosophila subquinaria* and *D. recens*. Ecol. Evol. 3:365–374.

Gilchrist, A. S., R. B. R. Azevedo, L. Partridge, and P. O'higgins. 2000. Adaptation and constraint in the evolution of *Drosophila melanogaster* wing shape. Evol. Dev. 2:114–124.

Greenspan, R. J., and J. F. Ferveur. 2000. Courtship in *Drosophila*. Annu. Rev. Genet. 34:205–232.

Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51:1341–1351.

Hansen, T. F., J. Pienaar, and S. H. Orzack. 2008. A comparative method for studying adaptation to a randomly evolving environment. Evolution 62:1965–1977.

Hoikkala, A., J. Aspi, and L. Suvanto. 1998. Male courtship song frequency as an indicator of male genetic quality in an insect species, *Drosophila montana*. Proc. Biol. Sci. 265:503–508.

Houle, D., and K. Meyer. 2015. Estimating sampling error of evolutionary statistics based on genetic covariance matrices using maximum likelihood. J. Evol. Biol. 28:1542–1549.

Houle, D., J. Mezey, P. Galpern, and A. Carter. 2003. Automated measurement of *Drosophila* wings. BMC Evol. Biol. 3:25.

Houle, D., G. H. Bolstad, K. van der Linde, and T. F. Hansen. 2017. Mutation predicts 40 million years of fly wing evolution. Nature 548:447–450.

Houle, D., L. T. Jones, R. Fortune, and J. L. Sztepanacz. 2019. Why does allometry evolve so slowly? Integr. Comp. Biol. 59:1429–1440.

Hoy, R. R., A. Hoikkala, and K. Kaneshiro. 1988. Hawaiian courtship songs: evolutionary innovation in communication signals of *Drosophila*. Science 240:217–219.

Ingram, T., and D. L. Mahler. 2013. SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise Akaike Information Criterion. Methods Ecol. Evol. 4:416–425.

Ingleby, F. C., I. Flis, and E. H. Morrow. 2015. Sex-biased gene expression and sexual conflict throughout development. Cold Spring Harb. Perspect. Biol. 7:a017632.

Kellermann, V., J. Overgaard, A. A. Hoffmann, C. Flojgaard, J.-C. Svenning, and V. Loeschcke. 2012. Upper thermal limits of *Drosophila* are linked to species distributions and strongly constrained phylogenetically. Proc. Natl. Acad. Sci. 109:16228–16233.

Klingenberg, C. P., and N. A. Gidaszewski. 2010. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. Syst. Biol. 59:245–261.

Kopperud, B. T., J. Pienaar, K. L. Voje, S. H. Orzack, T. F. Hansen, and M. Grabowski. 2019. Stochastic linear Ornstein–Uhlenbeck comparative hypotheses. R package version 2.

Labra, A., J. Pienaar, and T. F. Hansen. 2009. Evolution of thermal physiology in Liolaemus lizards: adaptation, phylogenetic inertia, and niche tracking. Am. Nat. 174:204–220.

Lande, R. 1980. Sexual dimorphism, sexual selection, and adaptation in polygenic characters. Evolution 34:292–305.

McLachlan, A. J. 1986. Sexual dimorphism in midges: strategies for flight in the rain- pool dweller *Chironomus imicola* (Diptera: Chironomidae). J. Anim. Ecol. 55:261–267.

Menezes, B. F., F. M. Vigoder, A. A. Peixoto, J. Varaldi, and B. C. Bitner-Mathé. 2013. The influence of male wing shape on mating success in *Drosophila melanogaster*. Anim. Behav. 85:1217–1223.

Mezey, J. G., and D. Houle. 2005. The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. Evolution 59:1027–1038.

O'Higgins, P., D. R. Johnson, W. J. Moore, and R. M. Flinn. 1990. The variability of patterns of sexual dimorphism in the hominoid skull. Experientia 46(7):670–672.

Pélabon, C., T. F. Hansen, A. J. R. Carter, and D. Houle. 2010. Evolution of variation and variability under fluctuating, stabilizing, and disruptive selection. Evolution 64:1912–1925.

Ray, R. P., T. Nakata, P. Henningsson, and R. J. Bomphrey. 2016. Enhanced flight performance by genetic manipulation of wing shape in *Drosophila*. Nat. Commun. 7:1–8.

Routtu, J., D. Mazzi, K. van der Linde, P. Mirol, R. K. Butlin, and A. Hoikkala. 2007. The extent of variation in male song, wing and genital characters among allopatric *Drosophila montana* populations. J. Evol. Biol. 20:1591–1601.

Schwarzkopf, L. 2005. Sexual dimorphism in body shape without sexual dimorphism in body size in water skinks (*Eulamprus quoyii*). Herpetologica 61:116–123.

Sztepanacz, J. L., and D. Houle. 2019. Cross-sex genetic covariances limit the evolvability of wing-shape within and among species of *Drosophila*. Evolution 73:1617–1633.

Sztepanacz, J. L., and M. W. Blows. 2015. Dominance genetic variance for traits under directional selection in *Drosophila serrata*. Genetics 200:371–384.

Talyn, B. C., and H. B. Dowse. 2004. The role of courtship song in sexual selection and species recognition by female *Drosophila melanogaster*. Anim. Behav. 68:1165–1180.

Taylor, C. E., and V. Kekic. 1988. Sexual selection in a natural population of *Drosophila melanogaster*. Evolution 42:197–199.

Uyeda, J. C., and L. J. Harmon. 2014. A novel Bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. Syst. Biol. 63:902–918.

Uyeda, J. C., T. F. Hansen, S. J. Arnold, and J. Pienaar. 2011. The million-year wait for macroevolutionary bursts. Proc. Natl. Acad. Sci. USA 108:15908–15913.

Uyeda, J. C., M. W. Pennell, E. T. Miller, R. Maia, and C. R. McClain. 2017. The evolution of energetic scaling across the vertebrate tree of life. Am. Nat. 190:185–199.

van der Linde, K., and D. Houle. 2008. A supertree analysis and literature review of the genus *Drosophila* and closely related genera (Diptera, Drosophilidae). Insect Syst. Evol. 39:241–267.

Weber, K. E. 1990. Selection on wing allometry in *Drosophila melanogaster*. Genetics 126:975–989.

Wolak, M. E., P. Arcese, L. F. Keller, P. Nietlisbach, and J. M. Reid. 2018. Sex-specific additive genetic variances and correlations for fitness in a song sparrow (*Melospiza melodia*) population subject to natural immigration and inbreeding. Evolution 72:2057–2075.

Associate Editor: P. Schmidt
Handling Editor: D. Hall

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supplementary Figure 1: Shifts in optima estimated by bayou for each of the seven principal component traits analyzed.
Supplementary Figure 2: Allometric and Non allometric shape dimorphism among species.
Supplementary Table 1: Sample sizes.
Supplementary Table 2: Parameter estimates for single-optimum OU models of trait evolution in PC1 accounting for measurement error.
Supplementary Table 3: Size dimorphism within species.
Supplementary Table 4: Individual species MANOVA.
Supplementary Table 5: Differences in dimorphism due to allometry.
Supplementary Table 6: Parameter estimates for OU models of trait evolution with phenotypic variance as a predictor.