Integrative and Comparative Biology

Integrative and Comparative Biology, pp. 1–12
doi:10.1093/icb/icz059

Society for Integrative and Comparative Biology

SYMPOSIUM

Which Line to Follow? The Utility of Different Line-Fitting Methods to Capture the Mechanism of Morphological Scaling

Alexander W. Shingleton[1]

Department of Biological Sciences, University of Illinois at Chicago, Chicago, IL 60607, USA

From the symposium "Allometry, Scaling and Ontogeny of Form" presented at the annual meeting of the Society for Integrative and Comparative Biology, January 3–7, 2019 at Tampa, Florida.

[1]E-mail: ashingle@uic.edu

Synopsis Bivariate morphological scaling relationships describe how the sizes of two traits co-vary among adults in a population. In as much as body shape is reflected by the relative size of various traits within the body, morphological scaling relationships capture how body shape varies with size, and therefore have been used widely as descriptors of morphological variation within and among species. Despite their extensive use, there is continuing discussion over which line-fitting method should be used to describe linear morphological scaling relationships. Here I argue that the "best" line-fitting method is the one that most accurately captures the proximate developmental mechanisms that generate scaling relationships. Using mathematical modeling, I show that the "best" line-fitting method depends on the pattern of variation among individuals in the developmental mechanisms that regulate trait size. For *Drosophila* traits, this pattern of variation indicates that major axis regression is the best line-fitting method. For morphological traits in other animals, however, other line-fitting methods may be more accurate. I provide a simple web-based application for researchers to explore how different line-fitting methods perform on their own morphological data.

Introduction

The morphological scaling relationship between body and trait size among adults in a species—also called static allometry—broadly captures the relative size of structures in the body and therefore characterizes species shape. Correspondingly, scaling relationships have been central to the study of morphology for well over 100 years. Early studies focused on the developmental mechanisms that underlie morphological scaling (Huxley 1924, 1932), but for the last 60 years research has largely concentrated on variation in scaling within and among species and the selective pressures that generate this variation. Morphological scaling relationships among adults in a population are typically linear on a log–log scale, and so can be modeled using the allometric equation $\log y = \log b + \alpha \log x$, where x and y are trait sizes measured in the same dimension, $\log b$ is the intercept and α is the allometric coefficient (Huxley and Tessier 1936). $\log b$ broadly captures the size of y relative to x, while α captures how relative trait size changes with overall size. When α is 1, a condition called isometry, traits scale proportionally with each other, and the size of y relative to x is maintained across body sizes. In contrast, when α is greater than or less than 1, called hyperallometry and hypoallometry, respectively, trait y becomes disproportionally larger ($\alpha > 1$) or smaller ($\alpha < 1$) relative to trait x with an increase in body size. In as much as $\log b$ and α capture key aspects of body proportion, considerable effort has been invested in describing variation in $\log b$ and α among pairs of traits within a species and among species for a pair of traits.

Central to the effort to study variation in morphological scaling is the ability to fit linear relationships to morphological measurements and extract scaling parameters. This effort has been complicated by the existence of various line-fitting methods, which, when applied to the same data, can generate different slopes and intercepts. There has, unsurprisingly, been much debate over which method is "correct," although no consensus has been reached (Madansky 1959; McArdle 1988; Kuhry and Marcus 1977;

Warton et al. 2006; Smith 2009; Taskinen and Warton 2011; Carroll and Ruppert 2012; Hansen and Bartoszek 2012; Pelabon et al. 2013; Kilmer and Rodriguez 2017). Much of this debate has concentrated on the statistical nuances of line fitting, and in particular the assumptions different methods make about the error with which morphological measurements are made (Kilmer and Rodriguez 2017). What is often absent from the discussion is consideration of the biological phenomenon the scaling parameter is trying to capture, and the efficiency with which each method achieves this. Here, I use a well-supported developmental model of how trait size is regulated to explore which line fitting method best captures the developmental process that control the slope of morphological scaling relationships. My goal is not to provide evidence that one method of line fitting is superior to another. Instead, my purpose is to facilitate exploration of alternative line fitting approaches so that researchers can match their statistics to the biological processes that they are attempting to model.

The developmental basis for morphological scaling

Morphological scaling relationships arise because there is variation in body size among individuals in a population and covariation in the size of their morphological traits (Shingleton et al. 2007). This size variation can be generated through variation in environmental factors (producing *environmental* scaling relationships) or variation in genetic factors (producing *genetic* scaling relationships). Regardless of the mechanism that generates size variation, covariation among pairs of traits arises because these traits are exposed to systemic variation in the same environmental or genetic size-regulatory factors. These factors may act directly and autonomously on growing traits—for example temperature—or indirectly and systemically—for example via growth hormones. Intuitively, it is the extent to which a change in a systemic factor generates a change in the size of each trait that determines their size covariation and consequently the slope of their scaling relationship. For example, for two traits x and y, if both traits share the same sensitivity to a size-regulatory factor, they will scale isometrically to one another as size varies with that factor (Fig. 1A, B, D). In contrast, if x is very sensitive to changes in a size-regulatory factor but y is not, then as size varies in response to that factor, x will vary more than y and the slope of their scaling relationship (x versus y) will be flat (Fig. 1B, C, E). Conversely, if y is more sensitive than x, then the slope of their scaling relationship will be steep. From this perspective, the slope of a morphological scaling relationship reflects the relative sensitivity of two traits to common size-regulatory factors (Shingleton et al. 2007).

Empirically estimated scaling relationships are a property of a population. Trait and body size are measured for a group of individuals and fit with a line that describes how trait size changes, on average, with body size among these individuals. Consequently, scaling relationships describe the average relative sensitivity of the two traits to size-regulating factors among all individuals in the group. What is lost with this approach, however, is the relative sensitivity of the two traits for each individual in a population. It is therefore useful to distinguish between an *individual-level scaling relationship*, which is the theoretical scaling relationship between adult traits for a single individual across a range of potential body sizes, and the *population-level scaling relationship*, which is the observed scaling relationship between traits among individual across a range of realized body sizes (Fig. 2; Dreyer et al. 2016; O' Brien et al. 2017).

Problematically, individual scaling relationships are typically unobservable, since it is not possible for the same individual to express more than one final, adult size—the one exception is where multiple individuals with the same genotype are each exposed to different levels of the same size-regulatory factor. In a genetically heterogeneous population, therefore, each individual will occupy a single point on its otherwise cryptic individual scaling relationship (Fig. 2). The observed population-level scaling relationship is that fitted among the observed phenotypes for many individuals, each a point on a cryptic individual scaling relationship. It is the variation among individuals in the relative sensitivity of traits to regulatory factors, manifest as intra-population variation in the individual-level scaling relationships, that is the ultimate target of selection on population-level scaling relationships. The pattern of this variation is thought to affect profoundly the response to selection (Dreyer et al. 2016). In as much as researchers are interested in the mechanisms that generate covariation among traits and how these mechanisms evolve, the observed population-level scaling relationship is therefore only an indirect measurement of the actual relationships of interest: the underlying cryptic individual scaling relationships.

If every member of a population had the same individual scaling relationship, and body and trait size varied only in response to variation in systemic factors, be it environmental or genetic, then the

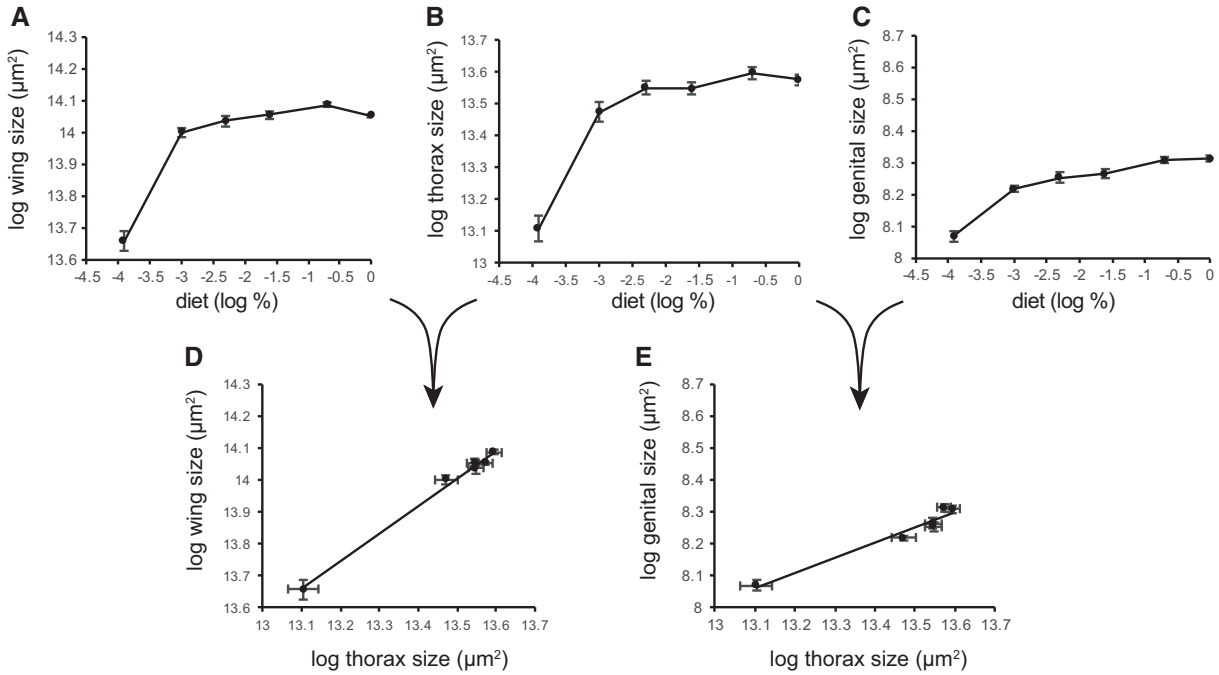

Fig. 1 The slope of morphological scaling relationships reflects relative trait plasticity. Traits vary in the relative sensitivity to changes in developmental nutrition. In *Drosophila melanogaster*, some traits, for example the wing (**A**) and thorax (**B**), are more sensitive to changes in developmental nutrition than other, for example the male genitalia (**C**). Consequently, the slope of the morphological scaling relationship where both traits are more nutritionally sensitive (**D**) is steeper than the slope where one trait is less nutritionally sensitive (**E**). Data from Shingleton et al. (2009).

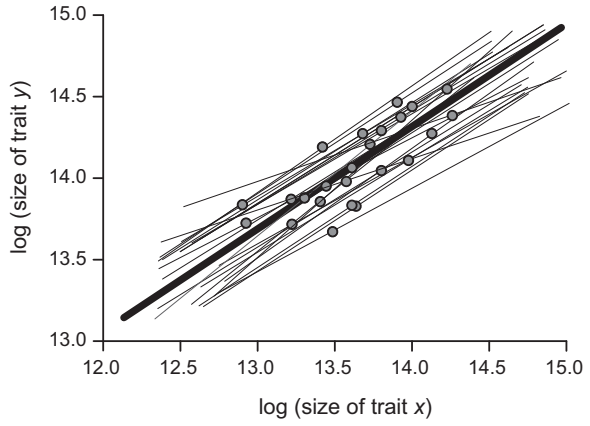

Fig. 2 The relationship between individual and population scaling relationships. Each individual genotype will express a scaling relationship across a range of environmental conditions (e.g., a nutritional gradient) (thin black lines). Because each individual's genotype is (typically) only exposed to a single developmental environment, it will sit at a single point on its cryptic individual scaling relationship. The observed population scaling relationship (thick black line) is the relationship between these individual points in the population.

population scaling relationship would be identical to all individual scaling relationships. However, in the real world, there is always scatter around a population-level scaling relationship, which has the potential to weaken the link between it and the underlying individual-level scaling relationship(s). Different line-fitting methods use different approaches to deal with this scatter. Thus, it is important to understand the cause of scatter if we are to make informed about decision about which line-fitting method to use.

There are three factors that generate scatter around the population-level scaling relationship (Fig. 3). First, variation among individuals in the relative sensitivity of their traits to a systemic size-regulatory factor will generate variation in the slope of their individual scaling relationships and hence generate scatter around the observed population scaling relationship (Fig. 3A, B). Second, genetic variation in trait-autonomous size regulators will generate variation in the intercept of their individual scaling relationships, uncorrelated with variation in other traits (Fig. 3C). Third, developmental instability, small-scale environmental heterogeneity that impacts individual traits, and measurement error will generate genotype-specific variation and also add to scatter around the population scaling relationship (Fig. 3D).

Line fitting methods

All regression analyses assume that there is an underlying bivariate relationship between two

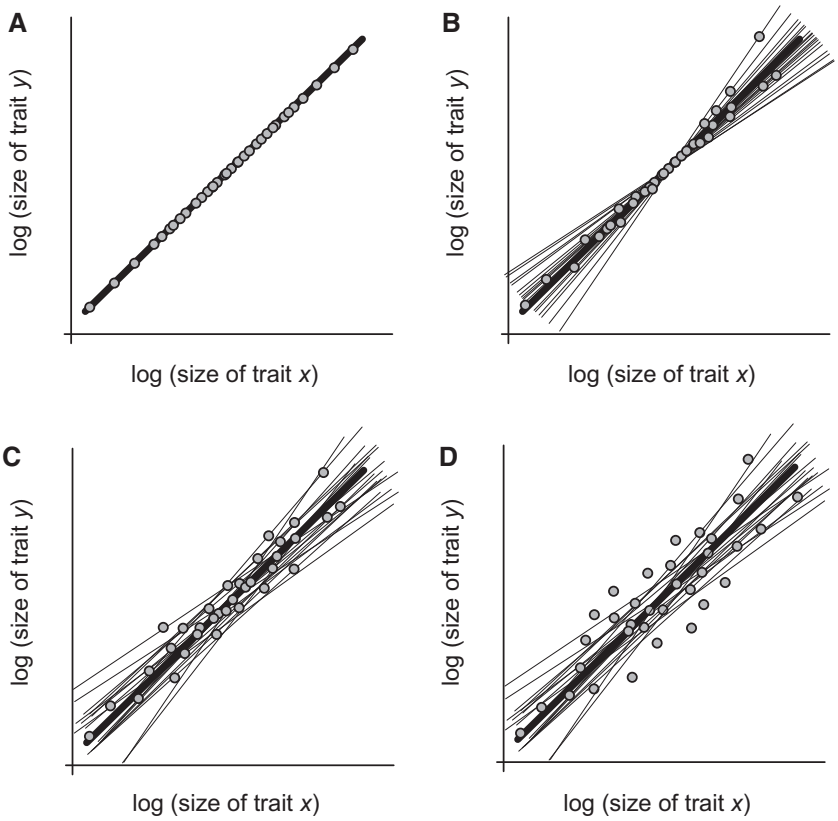

Fig. 3 The processes that generate scatter in population scaling relationships. (**A**) When all individuals in a population have the same individual scaling relationship, each individual (circles) lies along the observed population scaling relationship (thick black line), which is identical to the individual scaling relationship (thin black lines, hidden). (**B**) When there is variation among individuals in the relative sensitivity of their traits to environmental variation, this generates variation in the slope of the individual scaling relationships (thin black lines), and generates individual scatter around the population scaling relationship. (**C**) When there is trait-autonomous genetic variation in trait size, this will add variation to the intercept of the individual scaling relationships, further increasing individual scatter around the population scaling relationship. (**D**) When there is developmental instability, small-scale environmental heterogeneity that impacts individual traits, or error in measuring trait size, this will add additional individual scatter around the population scaling relationship.

measurements (x and y), but that error introduces scatter around this relationship. The regression therefore attempts to fit a line through the data that minimize the residuals; that is, the difference between an observed value and that predicted by the regression line. There are three primary methods used to fit lines to bivariate data: ordinary least squares (OLS), major axis (MA) regression and standardized major axis (SMA) regression (also referred to as reduced major axis [RMA] regression) (Warton et al. 2006; Smith 2009; Hansen and Bartoszek 2012). These three regression methods differ in what they use as residuals, and hence what they are minimizing (McArdle 1988; Warton et al. 2006). OLS regression fits a line to bivariate data such that the vertical distance between the regression line and each point squared and summed across all points is minimized. MA regression fits a line such that the perpendicular distance between the regression line and each point, squared and summed across all points, is minimized. SMA regression fits a line such that the product of the vertical and horizontal distance from the line to each point, summed across all points, is minimized. The SMA is the same as the MA, but the regression line is fitted to data that are standardized so that both variables have the same standard deviation, and then rescaled back to the original axes. Because OLS regression only minimizes the residuals on one axis (the y axis), the slope of the OLS for x against y is not the inverse of the slope of the OLS for y against x. This is not true for MA or SMA regression, where residuals are minimized across both axes, and x and y are functionally interchangeable.

Additional details of the different line-fitting methods, along with how they are calculated, is provided in the Supplementary Material.

Which line fitting method to use?

Much of the discussion about which line-fitting method should be used to estimate scaling relationships centers on the nature of the scatter around the regression line, and in particular, measurement error. Specifically, several authors have rejected the OLS method for fitting scaling relationships because it assumes that the x trait is measured without error, and therefore biases the slope downward (Ricker 1973; McArdle 1988; Ebert and Russell 1994; Green 1999; Bonduriansky 2007). Others have countered this argument by demonstrating that measurement error in x has a marginal effect on OLS slope estimations (Hansen and Bartoszek 2012; Pelabon et al. 2013; Kilmer and Rodriguez 2017). A number of authors have observed, however, that measurement error is not the only, and unlikely to be the most important, factor that causes scatter around a regression line (Warton et al. 2006; Smith 2009; Hansen and Bartoszek 2012). As discussed above, in the case of morphological scaling relationships, scatter is also generated by variation among individuals in the relative sensitivity of traits to systemic size-regulators, and variation in trait-autonomous size-regulators. While previous authors have acknowledged the existence of this scatter, referred to as *biological deviance*, its impact on the performance of different line-fitting methods for static allometry has not been explored.

Here I describe a model of individual and population scaling relationships based on the developmental mechanisms that regulate trait size (Dreyer et al. 2016). I then used this model to explore how assumptions about the nature of biological deviance and measurement error impact the ability of different line-fitting methods to estimate the relative sensitivities of traits to systemic regulators of size. Other mechanisms can generate population scaling relationships—for example, genetic linkage between alleles that independently regulate the size of different traits. However, we are explicitly only interested in developmental processes that regulate trait size systemically and in a coordinated manner across the body. These are likely the primary mechanism generating population scaling relationships and will be the target of selection that generates variation in scaling among populations and species.

A model of individual morphological scaling

The model has been published previously in a paper exploring the selection pressures that drive evolutionary changes in morphological scaling relationships (Dreyer et al. 2016). Briefly, the model assumes that trait growth is exponential, and that growth rate is regulated by two factors: systemic factors, such as circulating growth hormones or temperature, and trait-autonomous factors, such as morphogens gradients within the trait. Consequently, within an individual, trait size can be modeled as:

$$t = ae^{(Sk+i)d}, \tag{1}$$

where t is the trait size, a is the initial size of the trait, S is the level of a systemic growth factor, k is the sensitivity of the trait to the systemic growth factor, i is the organ autonomous growth rate, and d is the duration of growth. Log-transforming the equation generates the linear equation:

$$\log t = T = \log a + (Sk + i)d. \tag{2}$$

For two traits (x and y) in the same body:

$$T_x = \log a_x + (Sk_x + i_x)d. \tag{3}$$
$$T_y = \log a_y + (Sk_y + i_y)d. \tag{4}$$

When S is an environmental factor, Equations (3) and (4) describe the reaction norm of trait size against the environmental variable (e.g., Fig. 1A, B). As body size varies in response to changes in systemic growth factors, the individual scaling relationship between T_x and T_y can be described as:

$$T_y = \frac{k_y}{k_x} T_x + \log a_y - \frac{k_y}{k_x} \log a_x - \frac{k_y}{k_x} i_x d + i_y d. \tag{5}$$

Note that the slope of the individual scaling relationship of *trait x* (x-axis) against *trait y* (*y-axis*) is captured by the relative sensitivity of the two traits to the systemic growth factor, k_y/k_x. This is supported by experimental work in *Drosophila* that demonstrates that changes in the sensitivity of a trait to changes in nutrition alter the slope of its morphological scaling relationship with other traits among genetically identical individuals, when body size varies due to diet (Tang et al. 2011; Shingleton and Tang 2012).

In a genetically heterogeneous population, individuals will vary in k_x, k_y, i_x, i_y, a_x, a_y, and d. For simplicity, I will assume that initial trait size does not vary among individuals in a population. I will also, initially, assume that there is no variation in developmental time, although I will return to this point later in this paper. The size of individual traits therefore becomes:

$$T_x = Sk_x + i_x, \quad (6)$$
$$T_y = Sk_y + i_y, \quad (7)$$

and the individual scaling relationship between T_x and T_y can be described as:

$$T_y = \frac{k_y}{k_x} T_x - \frac{k_y}{k_x} i_x + i_y. \quad (8)$$

A model of population morphological scaling

While Equations (5) and (8) describe the scaling relationship between two traits across a range of S within an individual, each individual in a population is only exposed to a single level of S and thus only occupies a single point on this cryptic individual scaling relationship. The population-level scaling relationship between T_x and T_y estimated from multiple individuals in a population therefore depends on population-level variation in k_x, k_y, i_x, i_y, and S. We can model this using Equations (6) and (7), and by assuming that each individual's value of k_x, k_y, i_x, i_y, and S is sampled from normal distributions such that:

$$X \sim N(\mu_X, \sigma_X^2), \quad (9)$$

where X is the parameter value.

Thus, μ_S and σ_S^2 are the mean level of the systemic growth regulator, and its variance among individuals in a population, respectively; μ_{k_x} and $\sigma_{k_x}^2$ are the mean sensitivity of *trait x* to the systemic growth regulator, and variance in this sensitivity among individuals in a population, respectively, and μ_{i_x} and $\sigma_{i_x}^2$ are the mean trait-autonomous growth rate of *trait x* and variance in the trait-autonomous growth among individuals in a population, respectively. It is important to note that, for simplicity, $\sigma_{i_x}^2$ incorporates any factor that generates non-correlated variation in the size of *trait x*, which includes environmental and genetic factors, as well as developmental instability and measurement error. It is also important to note that σ_S^2 captures variation among individuals in the level of systemic growth regulators, which may be genetic or environmental in origin. When σ_S^2 is solely a consequence of environmental variation (i.e., when all individuals are genetically identical), the population-level scaling relationship is an environmental static allometry (Shingleton et al. 2007). When σ_S^2 is solely a consequence of genetic variation, the population-level scaling relationship is a genetic static allometry (Shingleton et al. 2007). The parameter values and their meaning are summarized in Table 1.

Using the mean and variance of the parameter values, it is possible to calculate the population mean, variance, and covariance of T_x and T_y and from them the expected slope and intercept of the OLS, MA, and SMA regressions for the population scaling relationship (Supplementary Material). These equations are shown in Table 2.

Using the model to assess line fitting methods

As discussed above, we are interested in the developmental processes that regulate trait size systemically and in a coordinated manner across the body and how these processes vary within and between populations and species. Consequently, we are interested in how well, and under what conditions, the different methods of calculating the population scaling relationship capture μ_{k_y}/μ_{k_x}—the relative mean sensitivity of traits to systemic factors that cause covariation in trait size and that generate scaling relationships. To achieve this, the equations for the OLS, MA, and SMA slope can be re-arranged to give a "bias factor" that each slope is multiplied by to give μ_{k_y}/μ_{k_x}.

For the OLS method, unless $\sigma_{i_x}^2$ (variance in trait-autonomous growth rate) and $\sigma_{k_x}^2$ (variance in trait sensitivity to systemic growth regulators) are zero, which is biologically unreasonable, this bias factor is always greater than one. Consequently, the OLS slope always underestimates μ_{k_y}/μ_{k_x}. In contrast, for the MA and SMA method, the bias factor can be greater or less than 1, and so they can both potentially perfectly capture μ_{k_y}/μ_{k_x}, although the conditions under which they do so may be biologically restrictive. Specifically, both α_{SMA} and $\alpha_{MA} = \mu_{k_y}/\mu_{k_x}$ when $\sigma_{k_x}^2 = \sigma_{k_y}^2$, $\sigma_{i_x}^2 = \sigma_{i_y}^2$, and $\mu_{k_x} = \mu_{k_y}$ (see Supplementary Material for mathematical details). That is, both methods will capture the mean individual-level scaling relationship between two traits when the traits scale (on average) isometrically, and when they show the same level of trait-autonomous size variation and variation in sensitivity to systemic growth regulators. Below I examine how reasonable these conditions are.

For many animals, body proportion appears to be largely maintained across a range of body sizes, suggesting that most traits scale near-isometrically and $\mu_{k_x} = \mu_{k_y}$. However, this may not be the case; a recent meta-analysis of 553 static allometries indicated that for traits not obviously subject to sexual selection, the average slope is 0.87 (95% CI: 0.79–0.94) (Voje 2016). This suggests that slight hypoallometry is the most common scaling relationship. One caveat

Table 1 Estimated distribution of parameter values for model of trait growth (Equations 6 and 7) based on morphological measurements of 50 males from each of 40 isogenic *Drosophila* lineages (see Supplementary Material for details)

Parameter	Biological meaning	Thorax	Wing	Palp	Anal plate	Femur	Genital arch
μ_S	Mean level of systemic growth regulator[a]				0		
σ_S	SD of level of systemic growth regulator[b]				0.165		
μ_k	Mean sensitivity to systemic growth regulator	0.842	0.581	0.583	0.431	0.559	0.299
σ_k	SD of sensitivity to systemic growth regulator	0.605	0.818	0.895	0.983	0.889	0.734
μ_i	Mean trait-autonomous growth rate	13.945	13.545	9.106	8.884	12.594	8.026
σ_i	SD of trait-autonomous growth rate	0.109	0.105	0.106	0.086	0.083	0.077

[a]The mean level of systemic growth regulator is set to unity to allow calculation of the other parameters.
[b]This is the standard deviation for the level of *S* for individuals across all traits and lineages.

Table 2 The expected slopes and intercepts of the population static allometry for population of individuals with trait sizes described by Equations (6) and (7)

	OLS	MA	SMA
Slope (α)	$\frac{\sigma_S^2\mu_{k_x}\mu_{k_y}}{\sigma_S^2\sigma_{k_x}^2+\sigma_S^2\mu_{k_x}^2+\sigma_{k_x}^2\mu_S^2+\sigma_{i_x}^2}$	$\frac{\left(\sigma_{T_y}^2-\sigma_{T_x}^2\right)+\sqrt{\left(\sigma_{T_y}^2-\sigma_{T_x}^2\right)^2+4\sigma_{T_x,T_y}^2}}{2\sigma_{T_x,T_y}}$	$\frac{\sigma_{T_y}}{\sigma_{T_x}}$
Intercept (b)		$\mu_{T_2}-\alpha\mu_{T_x}$	
Bias factor[a]	$\frac{\sigma_S^2\mu_{k_x}^2+\sigma_S^2\sigma_{k_x}^2+\sigma_{k_x}^2\mu_S^2+\sigma_{i_x}^2}{\sigma_S^2\mu_{k_x}^2}$	$\frac{2\sigma_S^2\mu_{k_y}^2}{\left(\sigma_{T_y}^2-\sigma_{T_x}^2\right)+\sqrt{\left(\sigma_{T_y}^2-\sigma_{T_x}^2\right)^2+\left(2\sigma_S^2\mu_{k_x}\mu_{k_y}\right)^2}}$	$\frac{\mu_{k_y}\sigma_{T_x}}{\mu_{k_x}\sigma_{T_y}}=\frac{\mu_{k_y}}{\mu_{k_x}}\sqrt{\frac{\left(\sigma_S^2\sigma_{k_x}^2+\sigma_S^2\mu_{k_x}^2+\sigma_{k_x}^2\mu_S^2+\sigma_{i_x}^2\right)}{\left(\sigma_S^2\sigma_{k_y}^2+\sigma_S^2\mu_{k_y}^2+\sigma_{k_y}^2\mu_S^2+\sigma_{i_y}^2\right)}}$

[a]The slope of the OLS, MA, or SMA is multiplied by this factor to give μ_{k_y}/μ_{k_x}. If the bias factor is >1, the slope of the OLS/MA/SMA underestimates μ_{k_y}/μ_{k_x}, while if the bias factor is <1 the slope of the OLS/MA/SMA overestimates μ_{k_y}/μ_{k_x}.

is that many, if not all, of these slopes were calculated using an OLS against body size. This will lower the estimate of the slope whenever there is uncorrelated variation in whatever proxy of body size is used, either due to measurement error or because of variation in trait-autonomous size-regulators, or whenever there is variation in the sensitivity of body size to systemic size-regulators. Thus, it is not clear if slight to moderate hypoallometry is the norm. Further, even if most traits scale isometrically, much of the research on allometry concentrates on traits that most obviously deviate from isometry, for example the hyperallometric secondary sexual characteristics of stalk-eyed flies or horned beetle, or the hypoallometric genitalia of male arthropods. For these traits, $\mu_{k_x} \neq \mu_{k_y}$.

Both the SMA and the MA can, however, capture μ_{k_y}/μ_{k_x}even when $\mu_{k_x} \neq \mu_{k_y}$. For the SMA, it can be seen from Table 2 that as μ_{k_y} and the slope increase, the standard deviation of T_y (σ_{T_y}) will also increase but at a slower rate, increasing the extent to which the SMA underestimates μ_{k_y}/μ_{k_x}. Under these conditions, the SMA slope will only capture μ_{k_y}/μ_{k_x} if the other factors that contribute to σ_{T_y}—that is, $\sigma_{k_y}^2$ and $\sigma_{i_y}^2$—increase also. From a biological perspective, this would imply that the mechanisms that regulate k, the sensitivity of a trait to systemic growth regulators, also regulate the population-level variance of k and of i, the trait-autonomous growth rate. While it is conceivable that the same mechanism could regulate the mean and variance of k (e.g., Emlen et al. 2012) it is difficult to envision how this mechanism could also regulate the variance of i, which by definition acts trait-autonomously. In contrast, the MA can capture μ_{k_y}/μ_{k_x} when $\mu_{k_x} \neq \mu_{k_y}$, if $\sigma_{k_x}^2 = \sigma_{k_y}^2$ and $\sigma_{i_x}^2 = \sigma_{i_y}^2$ (see Supplementary Material for mathematical details). This makes intuitive sense. An assumption of MA regression is that residual variance in x is equal to the residual variance in y. This will be true if traits share the same variation in sensitivity to systemic growth regulators ($\sigma_{i_x}^2 = \sigma_{i_y}^2$), and share the same level of trait-autonomous variation in growth rate ($\sigma_{i_x}^2 = \sigma_{i_y}^2$). While we know the developmental mechanisms that regulate a trait's sensitivity to at least one systemic growth regulator (insulin-like

peptides), and have elucidated many of the developmental mechanisms that regulate organ autonomous growth, there has been no study to directly measure genetic variation in these mechanisms regarding trait size. However, as I discuss below, it is possible to get an idea of what this level of variation is, at least indirectly.

Fitting data to the model: a biological example

The model captures the population scaling relationship based on the pattern of individual scaling relationships in a population. The parameters of the model describe concrete biological processes and so, in principle, it is possible to determine these parameter values for a population. However, these parameters can be difficult to measure. In particular, measuring genetic variation in the sensitivity of organs to changes in systemic growth regulators requires rearing the same genotype across environmental conditions that change systemic growth regulators, for example by varying developmental nutrition. This in turn requires multiple individuals of the same genotype, which is possible only in organisms that reproduce clonally—for example aphids, and many plants and fungi—or organisms that have isogenic lineages through artificial or natural inbreeding—for example *Caenorhabditis elegans* and the mangrove rivulus fish *Kryptolebias marmoratus* (Mesak et al. 2015). While this appears to be a stringent requirement, there are nevertheless published data that can be used to estimate the parameters of the model for a particular species. For example, Dreyer and Shingleton (2011) measured the size of the wing, the femur of the first leg, maxillary palp, posterior lobe of the genital arch, and anal plate of 50 males from each of 40 isogenic *Drosophila* lineages (Supplementary Fig. S1). The scaling relationships among trait sizes within a lineage is the cryptic individual scaling relationship, while the relationship of mean trait sizes among lineages is the population scaling relationship.

The among-lineage variation in the slopes and intercepts of individual scaling relationships can be used to estimate the mean and standard deviation of the sensitivity of different traits to environmentally-regulated growth regulators (μ_k and σ_k), of trait autonomous growth (μ_i and σ_i), and of environmentally-regulated systemic size regulator (μ_S and σ_S). The mathematical details are described in the Supplementary Material, and the R-scripts to conduct the analysis are provided on Dryad. When applied to the Dreyer and Shingleton (2011) data (Table 2), the analysis indicates that the variances in trait-autonomous growth and the variances in trait-sensitivity to changes in systemic size regulators—σ_i and σ_k, respectively—do not vary significantly among *Drosophila* traits (Brown-Forsythe test, $P > 0.5$ for both [Feltz and Miller 1996]). Based on the discussion above, these published morphological data therefore support the application of MA regression to estimate μ_{k_y}/μ_{k_x}, that is the relative mean sensitivity of traits to systemic factors that cause co-variation in trait size and that generate scaling relationships.

To specifically test the hypothesis that the MA regression best captures μ_{k_y}/μ_{k_x} in *Drosophila*, I used a second data set by Bakota et al. (this volume) of wing and body (pupal) size of male and female flies from 87 isogenic lineages (data available on Dryad). Flies were reared across a nutritional gradient to generate extensive variation in body size within each lineage. As for the Dreyer and Shingleton (2011) data, I used these data to calculate the ratio of mean nutritional-sensitivity of wing size relative to the mean nutritional sensitivity of body size, among lineages; that is $\mu_{k_{\text{wing}}}/\mu_{k_{\text{pupal}}}$. I then randomly sampled one individual from each lineage to generate a population scaling relationship of wing size against body size. This was repeated 1000 times to generate a mean and 95% confidence intervals for the slope of the population scaling relationship using each line fitting method. I also generated population static allometries for the mean size of the wing and thorax for each lineage, again using OLS, MA, and SMA regression. For both the "sampled" population scaling relationship and "mean" population scaling relationship, the MA was a better estimator of the $\mu_{k_{\text{wing}}}/\mu_{k_{\text{thorax}}}$ among lineages (Table 3). See Supplementary Material for additional details.

Although both these data sets were generated to explicitly explore the effect of genetic and environmental factors on trait size (co-)variation, there are likely other published data that can be used to estimate the parameters of the model in other organisms. What is required is multi-trait measurements made on multiple individuals, ideally reared across a range of environmental conditions, from multiple isogenic lineages. Data from isogenic lineages used in genome wide association studies should be useful (e.g., Lafuente et al. 2018), although environmental variation may be quite low. Studies that look at trait variation within and between multiple sibling groups may also be useful. While siblings are not genetically identical, the scaling relationship within a sibling group can be used as an individual scaling relationship, and variation in these "sibling scaling

Table 3 Relationship between $\mu_{k_{wing}}/\mu_{k_{pupa}}$ and the slope (with 95% confidence intervals) of the OLS, MA, and SMA regression for the population scaling relationships between wing and pupa size for flies from 87 isogenic *Drosophila melanogaster* lineages

	$\mu_{k_{wing}}/\mu_{k_{pupa}}$	OLS	MA	SMA
"Sampled" population scaling relationship[a]	1.076	0.948 (0.851–1.041)	1.071 (0.963–1.182)	1.062 (0.967–1.160)
"Mean" population scaling relationship[b]		0.840 (0.706–0.975)	1.058 (0.901–1.245)	1.047 (0.921–1.190)

[a]The mean population scaling relationship generated using trait sizes from one individual sampled from each lineage, fitted using OLS, MA, or SMA, and repeated 1000 times.
[b]The population scaling relationship generated using mean trait size from each lineage.

relationships" could be used to estimate the model's parameters.

Exploring the parameter space

While data from *Drosophila* may support the application of MA regression to calculate the slope of allometric relationships, the same need not be true for other traits in other organisms. I therefore developed an interactive interface using *shinyapp* that allows a user to determine which line fitting method (OLS, SMA, MA) best captures the average slope of individual scaling relationships (μ_{k_y}/μ_{k_x}) in a population under different conditions. The application can be accessed at https://shingletonlab.shinyapps.io/linefitting/. Alternatively, the *R*-scripts that run the application are available on Dryad, which a user can download and run locally on their computer.

The application allows the user to explore how well different line-fitting methods fit simulated data across a range of model parameter values. At the bottom of the interface, the user can assign parameter values to the model. The user can then select one of three plots to explore the effect of the parameter values on the utility of the different line-fitting methods. The interface is described in more detail in the Supplementary Material. The application assumes the "best" line fitting method is the one that produces a slope closest to μ_{k_y}/μ_{k_x}. Unsurprisingly, this depends on the parameters used to generate the population scaling relationship. Nevertheless, there are two general trends that are worth highlighting.

First, the model suggests that it is trait-autonomous variation in growth rate (σ_i) and variation in trait sensitivity to systemic growth regulators (σ_k) that have the greatest influence on which line fitting method best captures μ_{k_y}/μ_{k_x} (Fig. 4A). As outlined above, the MA perfectly captures μ_{k_y}/μ_{k_x} when $\sigma_{k_x} = \sigma_{k_y}$ and $\sigma_{i_x} = \sigma_{i_y}$. However, when $\sigma_{i_x} \neq \sigma_{i_y}$ or $\sigma_{k_x} \neq \sigma_{k_y}$, which is the better line fitting method depends on a number of factors. Broadly speaking, as uncorrelated variation in the *x trait* increases (i.e., σ_{i_x} and σ_{k_x}), the OLS, SMA, and MA estimates of μ_{k_y}/μ_{k_x} decrease. However, because the OLS slope always underestimates the true slope while the MA and SMA tend to overestimate the true slope (at least when σ_{i_x} and σ_{k_x} are small), the result is that the OLS becomes a less precise estimate of μ_{k_y}/μ_{k_x}, while the MA and SMA become more precise. This effect is greater when the true slope is hyperallometric ($\mu_{k_y}/\mu_{k_x} > 1$). In contrast, as uncorrelated variation in the *y trait* (i.e., σ_{i_y} and σ_{k_y}) increases, there is no effect of the OLS estimate of μ_{k_y}/μ_{k_x}, while the MA and SMA estimates increase. The results are that the OLS becomes the more precise estimator of the true slope, solely as a consequence of the MA and SMA estimates becoming less precise. This effect is greater when the true slope is hypoallometric ($\mu_{k_y}/\mu_{k_x} < 1$). Collectively, therefore, if the *y trait* scales hypoallometrically to the *x trait* but is expected to show more uncorrelated variation in size (e.g., because it is measured with less precision), then the OLS may best capture the relative sensitivity of the two traits to systemic growth regulators (Table 4). In contrast, if the *y trait* scales hyperallometrically to the *x trait* but is expected to show less uncorrelated variation in size, then the SMA may best capture μ_{k_y}/μ_{k_x} (Table 4). If both traits are expected to show more-or-less the same level of uncorrelated variation, as seen in *Drosophila*, the MA is the best estimator (Table 4).

Second, apart from μ_{k_x} and μ_{k_y}, the means of the other parameter values have comparatively little influence on how well each line fitting method captures μ_{k_y}/μ_{k_x}. Specifically, μ_i has no effect on the regression slope using any line fitting method, which is unsurprising, since μ_i regulates the intercepts of the underlying individual scaling relationships but not the slope. The mean level of the systemic growth factor, μ_s, only substantially affects the MA estimation, primarily when $\mu_s \gg 0 \ll \mu_s$. The parameter μ_s controls the pattern of underlying individual scaling relationships, specifically where in the range of observed trait sizes the individual scaling relationships tend to rotate (Fig. 3B). This pattern has important

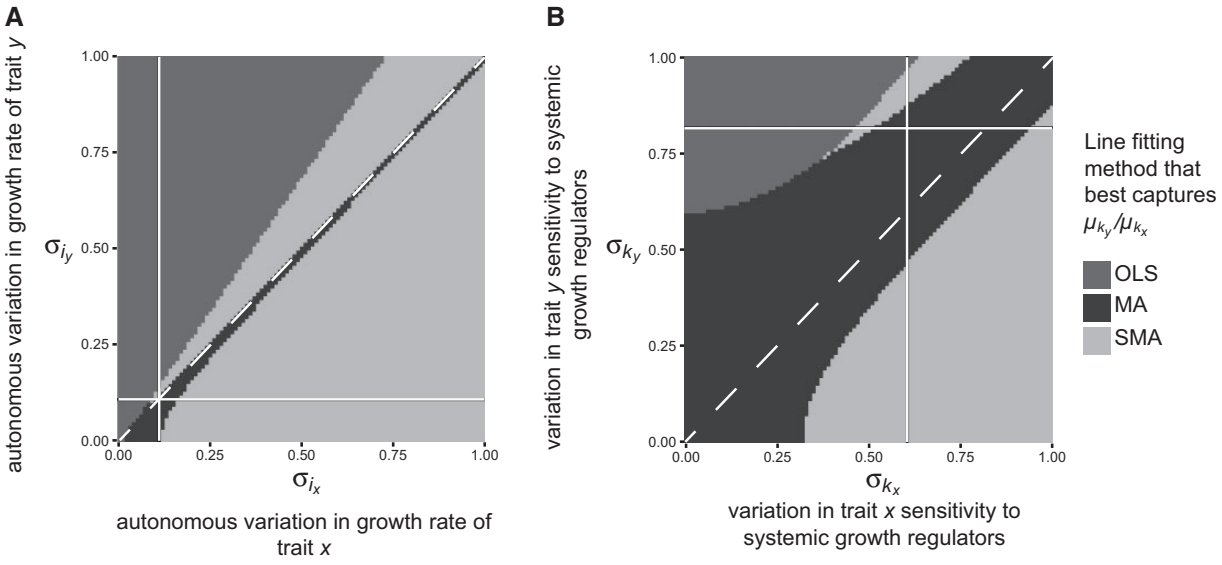

Fig. 4 The influence of different aspects of trait variation on the effectiveness of different line-fitting methods to capture the slope of morphological scaling relationships. At each point, the method (dark gray: OLS; black: MA; light gray: SMA) that generates a slope closest to the slope of the average individual scaling relationship μ_{k_y}/μ_{k_x} for the population is displayed. (**A**) The influence of trait-autonomous variation in growth rate, σ_i, on the effectiveness of different line-fitting methods. (**B**) The influence of variation in trait sensitivity to systemic growth regulators, σ_k, on the effectiveness of different line-fitting methods. All parameter values are for the thorax (*x*-axis) and wing (*y*-axis) in Table 1. Broken white line is $x = y$. Solid white lines are the observed values of σ_k and σ_i for the thorax (*x*-axis) and wing (*y*-axis) in *Drosophila melanogaster*, and where they intersect indicates which regression method best captures μ_{k_y}/μ_{k_x}.

Table 4 Summary of the utility of different line fitting methods for bivariate population scaling relationships, under different model parameters

Putative relationship[a]	Both traits have the same level of uncorrelated variation[b]	Uncorrelated variation greater for *y*-trait (T_y) than *x*-trait (T_x)	Uncorrelated variation greater for *x*-trait (T_x) than *y*-trait (T_y)
Hypoallometric	MA	OLS	SMA
Isometric	MA	SMA	SMA
Hyperallometric	MA	SMA	SMA

[a]Hypoallometric: $\mu_{k_y}/\mu_{k_x} \ll 1$; isometric: $\mu_{k_y}/\mu_{k_x} \approx 1$; hyperallometric: $\mu_{k_y}/\mu_{k_x} \gg 1$.
[b]Uncorrelated variation is generated by variation in trait-autonomous growth rate (σ_i) and/or trait-autonomous variation in sensitivity to a systemic growth factor (σ_k).

implications for how morphological scaling relationships respond to selection, and is considered in more detail in Dreyer et al. (2016).

Conclusions and future directions

The goal of this study was to examine the performance of different line fitting methods in capturing the mechanisms that produce covariation in trait size and generate morphological scaling relationships within populations. The model highlights that the phenotype most morphological researchers measure when studying scaling relationships—the population-level scaling relationship—is an imperfect representation of the relationship they are, in many cases, implicitly most interested in—the individual-level scaling relationship. The observed population-level scaling relationship is not generated by a "true" scaling relationship, with individual scatter around this relationship being a consequence of observation error or stochastic biological processes. Rather, the population-level scaling relationship is the observable part of a population of cryptic individual-level scaling relationships. The key insight is that departure from the population-level scaling relationship is, in part, due to variation among the slopes of individual-level scaling relationships for the group. It is this variation that evolution acts upon to generate changes in allometry. Explicitly incorporating this variation into the model of morphological

scaling not only allows one to better understand how scaling evolves, but also what statistical methods one should use to detect when evolution has occurred. Finally, this study has quantified in *Drosophila* the level of genetic variation in key developmental parameters that regulate morphological scaling, finding no evidence for differences among traits in their variance.

This is certainly not the first study to explore which line-fitting method should be employed to model morphological static allometries. However, most earlier studies did not fully consider the sources of variation that generate scatter around scaling relationships. One important exception is the work of Hansen and Bartoszek (2012) who applied a similar model to explore the interplay between biological and measurement error in evolutionary regressions, including evolutionary scaling relationships (i.e., morphological scaling relationships among species). As with this study, they started with the premise that all line-fitting methods impose bias. However, they concluded that the bias imposed by OLS regression is less severe than that imposed by MA and SMA regression, and therefore favored the OLS method of line-fitting to evolutionary and static allometric regressions. Their model did not, however, incorporate the developmental sources of biological variation that generate scatter in population scaling relationships. Nevertheless, both Hansen and Bartoszek (2012)'s and this study reiterate the importance of considering the sources of variation when applying regression models to biological data.

Which method the reader uses will depend on the purpose of their regression and the levels of variation that generate scatter in their morphological scaling relationship. If the reader is interested in estimating the relative mean sensitivity of traits to systemic size-regulators that generate covariation in size (μ_{k_y}/μ_{k_x}), then our model suggests that the "best" method for modeling morphological scaling relationships is most dependent on the level of uncorrelated variation in trait size, generated by both autonomous variation in trait size (σ_i) including measurement error, and variation in trait-autonomous sensitivity to changes in systemic growth factors (σ_k). If these are expected to be different between the two traits, then the MA regression may not be the best estimate of μ_{k_y}/μ_{k_x}.

Finally, it is important to note that the insight provided by any model is limited by how well the model captures the biological processes it describes. My model has the advantage of being very simple, but assumes that traits have linear "reaction norms" in response to variation in systemic regulators of size (Equations 6 and 7) (here, "reaction norm" includes trait response to systemic factors that may be genetic). For many organismal traits, this will not be true. Nevertheless, even if the trait "reaction norms" are not linear, they can still generate linear individual scaling relationships (Fig. 1), as described in the Supplementary Material. Further, it is possible to include developmental time in the model (Equation 5), also detailed in the Supplementary Material.

Any growth model can be used to explore how variation in underlying growth parameters affect the efficiency of different line fitting methods to capture the values of those parameters. Needless to say, the more complex the developmental model, the more difficult it is to mathematically describe the slopes and intercepts of the population scaling relationship. Even with the most complex models of individual growth, however, it is trivial to generate a simulated population scaling relationship *in silico*, and explore how changes in model parameters affect the slopes and intercept of the population-level scaling relationship when fit using different line-fitting methods. There are likely multiple developmental mechanisms that generate co-variation in trait size among individuals in a population, and a corresponding number of models. As we learn more of these mechanisms, our statistical methods should be adapted to better capture their key characteristics. It is these mechanisms, after all, that are ultimately the target of selection for changes in morphological scaling.

Acknowledgments

I thank Tony Frankino, Sonia Messar, and Austin Wilcox for comments on early versions of this manuscript. I also thank Fred Nijhout and Ken McKenna for inviting me to participate in the *Allometry, Scaling and Ontogeny of Form* symposium at SICB 2019.

Funding

This work was supported by the NSF IOS-1901727.

Supplementary data

Supplementary data are available at *ICB* online.

References

Bonduriansky R. 2007. Sexual selection and allometry: a critical reappraisal of the evidence and ideas. Evolution 61:838–49.

Carroll RJ, Ruppert D. 2012. The use and misuse of orthogonal regression in linear errors-in-variables models. Am Stat 50:1–6.

Dreyer AP, Saleh Ziabari O, Swanson EM, Chawla A, Frankino WA, Shingleton AW. 2016. Cryptic individual scaling relationships and the evolution of morphological scaling. Evolution 70:1703–16.

Dreyer AP, Shingleton AW. 2011. The effect of genetic and environmental variation on genital size in male *Drosophila*: canalized but developmentally unstable. PLoS ONE 6:e28278.

Ebert TA, Russell MP. 1994. Allometry and model II non-linear regression. J Theor Biol 168:367–72.

Emlen DJ, Warren IA, Johns A, Dworkin I, Lavine LC. 2012. A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. Science 337:860–4.

Feltz CJ, Miller GE. 1996. An asymptotic test for the equality of coefficients of variation from k populations. Stat Med 15:646–58.

Green A. 1999. Allometry of genitalia in insects and spiders: one size does not fit all. Evolution 53:1621–4.

Hansen TF, Bartoszek K. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. Syst Biol 61:413–25.

Huxley J. 1924. Constant differential growth-ratios and their significance. Nature 114:895–6.

Huxley JS, Teissier G. 1936. Terminology of relative growth. Nature 137:780–1.

Huxley JS. 1932. Problems of relative growth. London: Methuen & Co. Ltd.

Kilmer JT, Rodriguez RL. 2017. Ordinary least squares regression is indicated for studies of allometry. J Evol Biol 30:4–12.

Kuhry B, Marcus LF. 1977. Bivariate linear models in biometry. Syst Zool 26:201.

Lafuente E, Duneau D, Beldade P. 2018. Genetic basis of thermal plasticity variation in *Drosophila melanogaster* body size. PLoS Genet 14:e1007686.

Madansky A. 1959. The fitting of straight lines when both variables are subject to error. J Am Stat Assoc 54:173.

McArdle BH. 1988. The structural relationship: regression in biology. Can J Zool 66:2329–39.

Mesak F, Tatarenkov A, Avise JC. 2015. Transcriptomics of diapause in an isogenic self-fertilizing vertebrate. BMC Genomics 16:989.

O'Brien DM, Katsuki M, Emlen DJ. 2017. Selection on an extreme weapon in the frog-legged leaf beetle (*Sagra femorata*). Evolution 71:2584–98.

Pelabon C, Bolstad GH, Egset CK, Cheverud JM, Pavlicev M, Rosenqvist G. 2013. On the relationship between ontogenetic and static allometry. Am Nat 181:195–212.

Ricker WE. 1973. Linear regressions in fishery research. J Fish Res Board Can 30:409–34.

Shingleton AW, Estep CM, Driscoll MV, Dworkin I. 2009. Many ways to be small: different environmental regulators of size generate distinct scaling relationships in *Drosophila melanogaster*. Proc R Soc Lond B Biol Sci 276:2625–33.

Shingleton AW, Frankino WA, Flatt T, Nijhout HF, Emlen DJ. 2007. Size and shape: the developmental regulation of static allometry in insects. BioEssays 29:536–48.

Shingleton AW, Tang HY. 2012. Plastic flies—the regulation and evolution of trait variability in *Drosophila*. Fly 6:147–52.

Smith RJ. 2009. Use and misuse of the reduced major axis for line-fitting. Am J Phys Anthropol 140:476–86.

Tang HY, Smith-Caldas MS, Driscoll MV, Salhadar S, Shingleton AW. 2011. FOXO regulates organ-specific phenotypic plasticity in *Drosophila*. PLoS Genet 7:e1002373.

Taskinen S, Warton DI. 2011. Robust estimation and inference for bivariate line-fitting in allometry. Biom J 53:652–72.

Voje KL. 2016. Scaling of morphological characters across trait type, sex, and environment. Am Nat 187:89–98.

Warton DI, Wright IJ, Falster DS, Westoby M. 2006. Bivariate line-fitting methods for allometry. Biol Rev Camb Philos Soc 81:259–91.