doi:10.1111/evo.12984

Cryptic individual scaling relationships and the evolution of morphological scaling

Austin P. Dreyer,[1] Omid Saleh Ziabari,[2] Eli M. Swanson,[3] Akshita Chawla,[4] W. Anthony Frankino,[5] and Alexander W. Shingleton[2,6]

[1]*Department of Integrative Biology, Michigan State University, East Lansing, Michigan 48824*

[2]*Department of Biology, Lake Forest College, Lake Forest, Illinois 60045*

[3]*College of Biological Sciences, University of Minnesota, St Paul, Minnesota 55108*

[4]*Department of Statistics and Probability, Michigan State University, East Lansing, Michigan 48824*

[5]*Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204*

[6]*E-mail: shingleton@lakeforest.edu*

Received March 18, 2016
Accepted May 27, 2016

Morphological scaling relationships between organ and body size—also known as allometries—describe the shape of a species, and the evolution of such scaling relationships is central to the generation of morphological diversity. Despite extensive modeling and empirical tests, however, the modes of selection that generate changes in scaling remain largely unknown. Here, we mathematically model the evolution of the group-level scaling as an emergent property of individual-level variation in the developmental mechanisms that regulate trait and body size. We show that these mechanisms generate a "cryptic individual scaling relationship" unique to each genotype in a population, which determines body and trait size expressed by each individual, depending on developmental nutrition. We find that populations may have identical population-level allometries but very different underlying patterns of cryptic individual scaling relationships. Consequently, two populations with apparently the same morphological scaling relationship may respond very differently to the same form of selection. By focusing on the developmental mechanisms that regulate trait size and the patterns of cryptic individual scaling relationships they produce, our approach reveals the forms of selection that should be most effective in altering morphological scaling, and directs researcher attention on the actual, hitherto overlooked, targets of selection.

KEY WORDS: **Allometry, body size, body proportion, morphological scaling, mathematical modelling, nutrition, shape.**

Background

Morphological scaling–-the relationship between trait size and body size among adults in a population–-is central to the expression and evolution of animal form. As growth varies in response to environmental factors such as temperature or access to nutrition, the developmental mechanisms that regulate scaling maintain proper size covariation among traits, and retain structural and functional cohesion across a range of adult body sizes (Bertalanffy 1964; Gould 1966). Because morphological scaling describes the relationship between trait size and overall body size, changes in scaling result in changes in body shape. Unsurprisingly, therefore, scaling relationships can vary dramatically among species or even within them, the latter most obvious in species exhibiting high sexual dimorphism (Fairbairn 1997). Consequently, the evolution of morphological scaling is the major mechanism by which all morphological diversification occurs (Thompson 1917; Stern and Emlen 1999).

Morphological scaling relationships fit to individuals of a population (called a *static allometry*) are typically linear on a log–log scale. Consequently, they can be described using the allometric equation: $\log y = \log b + \alpha \log x$, where y and x are the size of morphological traits, typically organ (y) and body (x) size, and α is the allometric coefficient (Huxley and Tessier 1936). For many traits $\alpha \approx 1$, a condition called isometry, where the size of y relative

EVOLUTION
INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

to x is constant across all trait sizes. When α is $>$ or $<$ 1, called hyper- and hypoallometry, respectively, the size of y relative to x changes with overall size. Hyperallometric traits ($\alpha > 1$) are disproportionally large in large individuals and provide some of the most striking examples of morphology, such as the oversized claw of male fiddler crabs (Huxley 1932), the massive antlers of the extinct Irish elk (Gould 1974), and the elongated eye-stalks of stalk-eyed flies (Wilkinson 1997). Hypoallometric traits ($\alpha < 1$) are disproportionally smaller in large individuals and while less dramatic, are still important biologically. Examples include brain size in humans (Koh et al. 2005), and male genital size in most insects and possibly mammals (Eberhard 2009). Change in log b, the intercept of the allometry, reflects a proportional shift in the size of y relative to x across the full range of trait sizes within a population.

Although the central importance of morphological scaling in generating morphological diversity has been recognized widely for some time (Thompson 1917; Huxley 1932; Gould 1966), an understanding of how selection alters scaling relationships remains elusive. On the one hand, the slopes and intercept of morphological scaling relationships show extensive evolutionary variation, even among closely related taxa (Gould 1966; Wilkinson and Dodson 1997; Stern and Emlen 1999; Voje et al. 2014). However, direct tests of the evolvability of scaling relationship parameters have generated equivocal results. While the intercept of morphological scaling relationships appear to respond rapidly to artificial selection (Frankino et al. 2005; Egset et al. 2012; Bolstad et al. 2015), the same is not true for the slope. There are only three studies that have directly targeted the slope of a morphological scaling relationship using artificial selection. The response was absent in one (Egset et al. 2012) and weak in another (Stillwell et al. 2016). In the third, the derived scaling did not hold across the full range of body sizes displayed by a population experiencing variation in access to nutrition (Bolstad et al. 2015), as natural scaling patterns must. In all experiments, the response to selection varied considerably among replicate lines and from generation to generation. Such an erratic response is consistent with the general hypothesis that the slopes of morphological scaling relationships are developmentally constrained and evolve slowly at a macroevolutionary time scale (Huxley 1932; Gould 1966; Voje and Hansen 2013; Pelabon et al. 2014; Voje et al. 2014).

In contrast, however, artificial selection on absolute or relative trait size has produced a correlated change in the slope of the morphological scaling relationship in several species (Robertson 1962; Weber 1990; Wilkinson 1997; Tobler and Nijhout 2010). Moreover, developmental-genetic manipulations of the developmental mechanisms that regulate the relationship between organ and body size easily change the slope of trait-body size scaling (Tang et al. 2011; Shingleton and Tang 2012). The incongruity between the effects of direct selection on the slope of morphological scaling relationships, versus the effects of indirect selection or developmental manipulation, demonstrates that our understanding of how selection alters morphological scaling is inadequate.

Here, we use a fundamentally new approach to determine how different modes of selection generate change in morphological scaling. We model the population-level scaling relationship as a consequence of variation in the individual-level developmental processes that regulate trait and body growth. These processes can be summarized as a few growth parameters that describe how trait and body size covary for an individual genotype across nutritional conditions, producing what we call the "cryptic individual scaling relationship." We use the term "cryptic" because the individual scaling relationship is not observed. Rather, each individual realizes only a single phenotypic point on its unique individual scaling relationship; the location of this point is determined by access to nutrition during growth. The population-level scaling relationship is therefore fit to a population of such single realized points, each of which rests on its own cryptic individual scaling relationship. Our model explicitly captures the range and distribution of cryptic individual scaling relationships in a population, thereby linking the population-level static allometry with the individual-level variation in developmental mechanisms that generate it. By applying an evolutionary algorithm to our model, we are able to predict the selection response of the proximate developmental mechanisms that ultimately determine the slope and intercept of a population's static allometry in a manner that has not been possible with more traditional approaches.

Model and Results

THE MODEL

Implicit to the concept of scaling is that there is variation in body size that is accompanied by covariation in the size of other traits. Perhaps the major factor that influences body size in most animals is access to nutrition during development (Demment and Van Soest 1985; Bateson et al. 2004; Koyama et al. 2013; Nijhout et al. 2014), and for simplicity we model the evolution of scaling relationships that are a consequence of nutritional variation, that is *nutritional scaling relationships* or nutritional static allometries (Shingleton et al. 2007). The developmental mechanisms that regulate nutritional scaling relationships are best understood in holometabolous insects, and our model is therefore derived from an earlier one that describes body and organ growth in these taxa (Shingleton et al. 2008).

In insects, growth of the body and organs is an approximately exponential process (Bakker 1959; Martin 1982; Bryant and Levinson 1985; Nijhout et al. 2006) that can be modeled using an exponential equation:

$$x_t = ae^{rt}, \quad (1)$$

where x_t is the size of a trait at time t, r is the exponential growth rate, and a is the initial trait size. In holometabolous insects, nutrition affects trait size primarily by influencing growth rate, which can therefore be divided into two components: intrinsic (or morphogenetic) growth and nutrition-dependent growth (Truman 2006). Developmentally, nutrient-dependent growth is mediated via the activity of the insulin-signaling pathway, which comprises a number of positive and negative growth regulators that work together to activate cell growth and proliferation when nutrition is high and suppress it when nutrition is low (Teleman 2010). Importantly, traits can differ in their sensitivities to changes in insulin signaling, producing variation among traits in their growth response to variation in access to nutrition during development (Shingleton et al. 2005; Shingleton and Tang 2012). The growth of a trait can therefore be modeled as:

$$x_t = ae^{(Sk+i)t}, \tag{2}$$

where S is the systemic level of insulin-signaling an individual generates in response to developmental nutrition and affects all tissues, k is the level of a trait's insulin-sensitivity, and i is the morphogenetic growth rate, which we assume is organ-autonomous (Truman 2006). S can be either positive or negative, such that low nutrition reduces growth rate below the intrinsic rate while high nutrition elevates it above the intrinsic rate. We can therefore model final trait size as:

$$T = ae^{(Sk+i)d}, \tag{3}$$

where d is the duration of trait growth (Shingleton et al. 2008). This can be log-transformed as:

$$\log T = \log a + \log(Sk + i)\, d. \tag{4}$$

Within an individual, two traits for the same individual may have different insulin-sensitivities and different intrinsic growth rates (Tang et al. 2011) but will share the same level of systemic insulin signaling and duration of growth (Shingleton et al. 2008):

$$\log T_1 = \log a_1 + (Sk_1 + i_1)\, d, \tag{5}$$

$$\log T_2 = \log a_2 + (Sk_2 + i_2)\, d, \tag{6}$$

where T_1 is the size of one trait, for example body size, and T_2 is the size of another trait, for example organ size. We can calculate the allometric relationship between $\log T_1$ and $\log T_2$ across a range of nutritional levels for a particular genotype by solving equation (5) for S and inserting into equation (6).

$$\log T_2 = \frac{k_2}{k_1} \log T_1 + \log a_2 - \frac{k_2}{k_1} \log a_1 - \frac{k_2}{k_1} i_1 d + i_2 d. \tag{7}$$

This linear equation describes the genotype-specific cryptic individual scaling relationship (indSR): that is, the scaling relationship that would result if individuals sharing that genotype were reared under a range of nutritional conditions, for example rearing flies from isogenic lines of *Drosophila* under different diets (Shingleton et al. 2009). The allometric coefficient for the relationship between $\log T_1$ and $\log T_2$ is k_2/k_1: the relative insulin-sensitivity of trait 2 versus trait 1. When $k_2 > k_1$, trait 2 will be hyperallometric to trait 1. When $k_2 < k_1$, trait 2 will be hypoallometric to trait 1. It follows that changes in k_2 relative to k_1 will alter the slope of a genotype's individual scaling relationship (Fig. 1A and B). This is supported by developmental-genetic studies in *Drosophila* where wing-autonomous changes in insulin-sensitivity make the wing-body nutritional scaling relationship more hyperallometric (Tang et al. 2011). The observed population level scaling relationship (popSR) will result from fitting an allometry to the phenotypes expressed by a population of genetically unique individuals, each occupying a single point on the their correspondingly unique indSR (Fig. 1C and D).

We used our model to generate a popSR using 1000 genetically unique individuals, each possessing a unique indSR. To reduce the number of model parameters, we assume that there is no variation among individuals in developmental time ($d = 1$) and that initial trait size is the same among individuals ($a = 1$). Note that variation in a and d influences only the intercept but not the slope of the allometry (eq. 7), and so these parameters are unimportant when considering evolution of the allometric coefficient. The size of individual traits becomes:

$$\log T_1 = Sk_1 + i_1, \tag{8}$$

$$\log T_2 = Sk_2 + i_2, \tag{9}$$

and the allometric relationship between T_1 and T_2 becomes:

$$\log T_2 = \frac{k_2}{k_1} \log T_1 - \frac{k_2}{k_1} i_1 + i_2. \tag{10}$$

To generate the static allometry for a population, we assign each individual a genotypic value for k_1, k_2, i_1, and i_2. The value of these parameters is determined by a "diallelic" system with codominance where each value is the sum of two "alleles," a and b. For example, an individual's value for k_1 is the sum of $k_{1.a}$ and $k_{1.b}$. The initial population contains a specified number of alleles, the values of which are sampled from a normal distribution with a known mean (μ) and standard deviation (σ) (e.g., $N(\mu_{k1.a},\sigma^2_{k1.a})$, $N(\mu_{k1.b},\sigma^2_{k1.b})$, etc.). Alleles for the same parameter (k_1, k_2, i_1, and i_2) are sampled from the same normal distribution. Consequently, the μ and σ of the allelic values are half the μ and σ of the genotypic value (e.g., $\mu_{k1.a} = \mu_{k1.b} = \frac{1}{2}\, \mu_{k1}$). The result is a population of individuals, each possessing its own cryptic

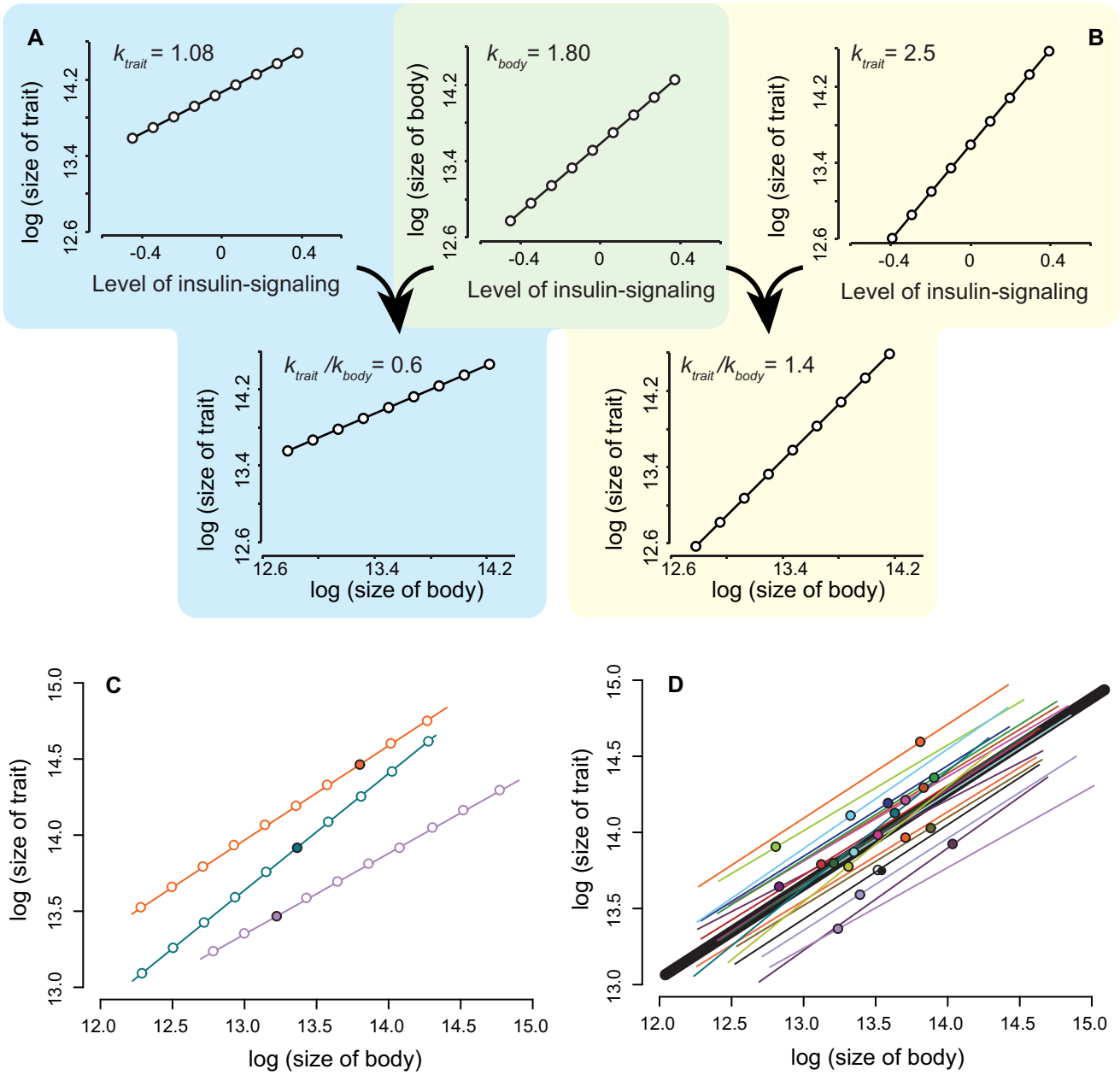

Figure 1. **Cryptic individual scaling relationships (indSRs) and population scaling relationships (popSRs). (A) For a particular genotype, the slope of the nutritional individual scaling relationship (indSR) between trait size and body size will reflect the extent to which the size of each responds to changes in the systemic level of insulin-signaling (*k*). (B) A change in a trait's insulin-sensitivity will alter the slope of the trait's scaling relationship with body size. (C) Variation among individuals in trait insulin-sensitivity and intrinsic growth rates generates variation in the slopes and intercepts of the indSRs (lines), although each genotype will express only a single point on its cryptic indSR (filled circles). (D) The observed scaling relationship for a population (popSR) (filled circles, heavy black line) is comprised of individuals occupying a single point on their cryptic indSRs (fine lines).**

indSR. Finally, each individual is assigned a value of S, the level of nutritionally regulated insulin signaling, again sampled from a normal distribution with known mean (μ_S) and standard deviation (σ_S) ($N(\mu_S,\sigma^2{}_S)$). The parameter S determines the phenotype expressed by each individual along it is cryptic indSR and describes the effect of the nutritional environment on trait size. Fitting a regression to the values for $\log T_1$ against $\log T_2$ for all individuals in the population generates a population scaling relationship that is distributed around the bivariate mean of the two trait distributions (Fig. 1D). The expected bivariate mean and slope of the popSR–whether estimated by ordinary least square (OLS), major axis (MA), or standardized major axis (SMA)–can be calculated

directly from the μ and σ of the parameter values, as described in Supporting Information.

PATTERNS OF CRYPTIC INDIVIDUAL SCALING RELATIONSHIPS

The bivariate mean of the popSR is ($\mu_{i1} + \mu_S\ \mu_{k1}$, $\mu_{i2} + \mu_S\ \mu_{k2}$), which, when $\mu_S = 0$, simplifies to (μ_{i1},μ_{i2}) (see Supporting Information). If there is no variation in intrinsic growth rate (σ_{i1} and $\sigma_{i2} = 0$), but variation in trait insulin-sensitivity (σ_{k1} and $\sigma_{k2} > 0$), the model generates a pattern of cryptic indSRs that intersect when $S = 0$, and rotate around (μ_{i1},μ_{i2}) (Fig. 2A). As σ_{i1} and σ_{i2} increase, this point of intersection between indSRs becomes more variable, but the average point of intersection remains ~ (μ_{i1},μ_{i2}) (Fig. 2B). It follows that when $\mu_S = 0$, that is when individuals on average grow at their intrinsic growth rate, the pattern of indSRs across the range of trait sizes observed in the population rotates more-or-less around the bivariate mean of the popSR (Fig. 2B). However, this is not the only possible pattern of cryptic indSRs. It is possible that, because of the nutritional environment experienced by members of a population, individuals grow on average faster or slower than their intrinsic rate and $\mu_S \neq 0$. Under these conditions the indSRs again rotate around ~(μ_{i1},μ_{i2i}), but the bivariate mean of the popSR is now ($\mu_{i1} + \mu_S\ \mu_{k1}$, $\mu_{i2} + \mu_S\ \mu_{k2}$): that is the pattern of indSRs no longer rotates around the bivariate mean of the popSR. When $\mu_S > 0$, the point of indSR rotation lies more towards the origin than the bivariate mean of the popSR (Fig. 2C). Conversely, when $\mu_S < 0$ the point of indSR rotation is further away from the origin than the bivariate mean of popSR (Fig. 2D). For simplicity, as shorthand we reference real-world objects that move in a manner that reflects these three patterns of cryptic indSRs: "seesaw" ($\mu_S = 0$, Fig. 2B), "speedometer" ($\mu_S > 0$, Fig. 2C), and "broomstick" ($\mu_S < 0$, Fig. 2D). Critically, while the underlying pattern of cryptic indSRs may be very different among populations, the covariance of T_1 and T_2 may be the same, producing near-identical phenotype distributions and indistinguishable popSRs with the same slope and intercept (Fig. 2B',C',D').

SELECTION REGIMES AND EVOLUTIONARY ALGORITHM

Using our model (eqs. (8) and (9)), we generated three starting populations of 1000 individuals: one where $\mu_S = 0$ (seesaw), one where $\mu_S > 0$ (speedometer) and one where $\mu_S < 0$ (broomstick). All three populations contained 100 "alleles" each for k_1, k_2, i_1, and i_2 (50 alleles for k_{1a}, 50 alleles for k_{1b}, 50 alleles for k_{2a}, etc.). These alleles were sampled from known normal distributions, as were the values of S. The mean and standard deviations of these normal distributions were chosen such that the resulting popSRs in all three populations were statistically indistinguishable; the relationship between logT_1 and logT_2 had the same slope (OLS, MA, and SMA), and intercept in all three populations (Table 1). Further, the parameter values were chosen so that the popSR resembled the relationship between body size (pupal case area) and wing size in a wild-type outbred population of *Drosophila* (Stillwell et al. 2011).

We subjected each starting population to the forms of selection typically invoked to explain changes in morphological scaling and asked how these influenced: (1) the slope and intercept of the population static allometry; and (2) the mean and standard deviation of the developmental parameters (k_1, k_2, i_1, and i_2) that underlie the population static allometry. The forms of selection were (Fig. 3):

1. Positive (D+) or negative (D−) directional selection on T_2 (Fig. 3A). Here, we selected individuals that had the highest or lowest absolute value of T_2.
2. Positive (P+) or negative (P−) proportional selection on T_2 (Fig. 3B). Here, we selected individuals that had the highest or lowest value of T_2 relative to T_1.
3. Positive (C+) or negative (C−) correlational selection (Fig. 3C). For positive correlational selection we selected individuals that had a large T_1 and disproportionately large T_2 or had a small T_1 and disproportionately small T_2. For negative correlational selection we selected individuals that had a large T_1 and disproportionately small T_2 or had a small T_1 and disproportionately large T_2.
4. Stabilizing (St) selection on T_2 (Fig. 3D). Here, we selected individuals that had an intermediate value of T_2 regardless of their value of T_1.

Specific details of the selection regimes are provided in Supporting Information. These selection regimes all select on some aspect of trait size. Under our model, however, relative insulin-sensitivity k_2/k_1 is the biological factor that controls the slope of the indSRs and accounts for the correlation between T_1 and T_2 in the population, which is in turn captured by the slope of the popSR. We therefore applied a fifth form of selection (K) that targeted an individual's k_2/k_1, that is the slope of its indSR. This selection regime is also detailed in Supporting Information.

To generate "offspring" in the model, selected individuals are randomly designated as "male" or "female," with equal probability. Offspring were generated by randomly sampling one male and one female and randomly selecting one of the two "alleles" for each parameter value from each parent. These parameter values are then assigned to each offspring along with a value for S sampled from $N(\mu_S,\sigma^2_S)$, simulating the effect of a random nutritional environment on trait size. Collectively, these parameter values are used to calculate offspring trait sizes. Each generation, this process was repeated to generate a new population of 1000 individuals. All sampling of male and female parents was done with replacement, so each parent can generate multiple offspring.

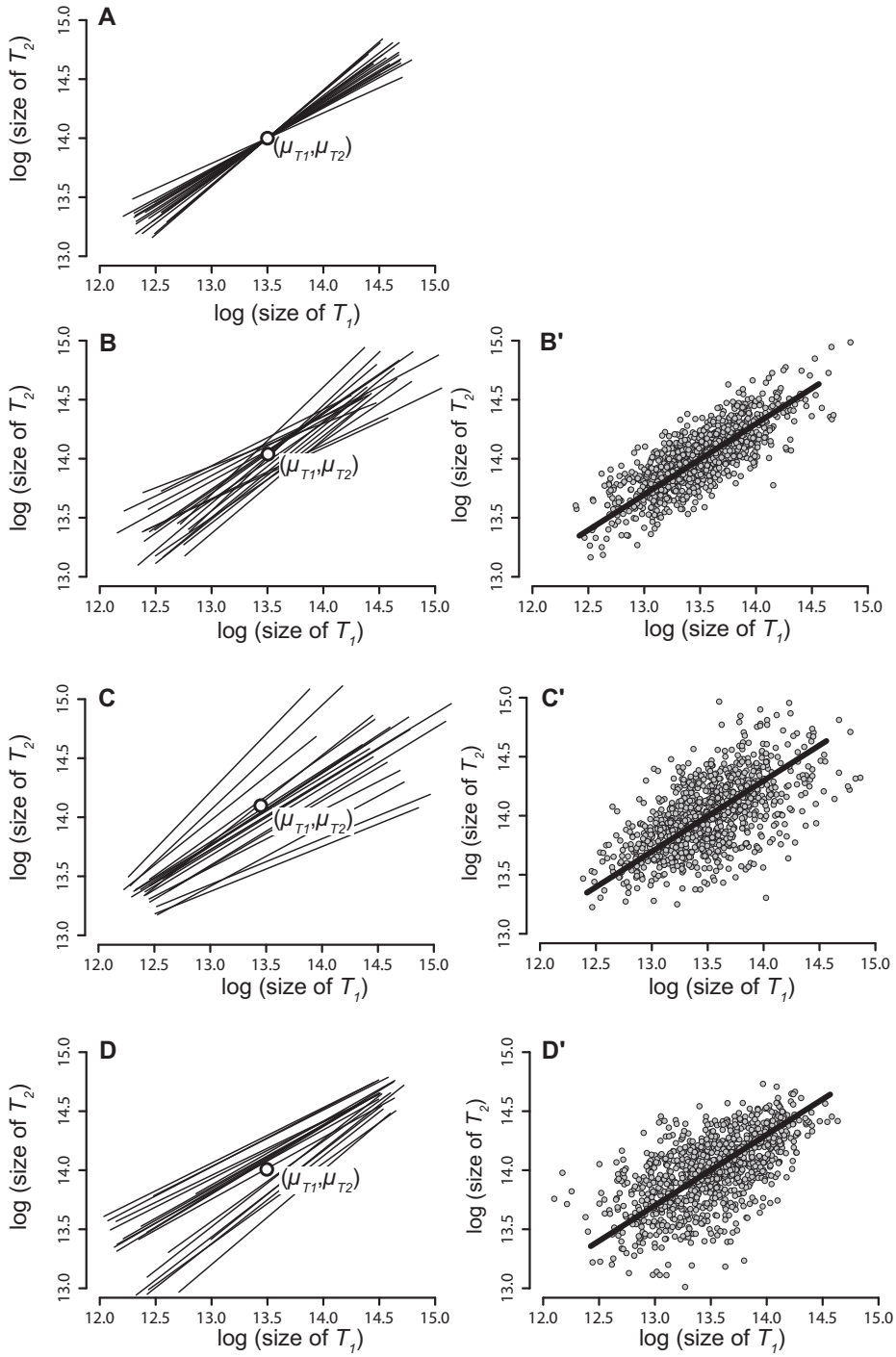

Figure 2. Patterns of cryptic individual scaling relationships (indSRs). (A,A') When $\mu_S = 0$ and there is no variation in trait-autonomous growth rate but variation in trait insulin-sensitivity, the indSRs (A) intersect at (μ_{i1},μ_{i2}), which is the bivariate mean of the popSR (μ_{T1},μ_{T2}). Parameter values are sampled from: $N_{k1}(1.8,0.12^2)$, $N_{k2}(1.08,0.12^2)$, $\mu_{i1} = 13.5$, $\mu_{i2} = 14$. (B) When $\mu_S = 0$ and there is variation in both trait-autonomous growth rate and trait insulin-sensitivity, the indSRs do not intersect at a single point, but more-or-less rotate around the bivariate mean of the popSR (μ_{T1},μ_{T2}) generating a "seesaw" shape. Parameter values are sampled from: $N_S(0,0.2^2)$, $N_{k1}(1.8,0.12^2)$, $N_{k2}(1.08,0.12^2)$, $N_{i1}(13.5,0.13^2)$, $N_{i2}(14,0.13^2)$. (C) When $\mu_S > 0$, the indSRs again more-or-less rotate around (μ_{i1}, μ_{i2}), but this is now closer to the origin than the bivariate mean of the popSR (μ_{T1}, μ_{T2}) generating a "speedometer" shape. Parameter values are sampled from: $N_S(1,0.2^2)$, $N_{k1}(1.8,0.12^2)$, $N_{k2}(1.08,0.12^2)$, $N_{i1}(11.7,0.05^2)$, $N_{i2}(12.92,0.05^2)$. (D) Conversely, when $\mu_S < 0$, the indSRs more-or-less rotate around (μ_{i1}, μ_{i2}), but this is now farther from the origin than the bivariate mean of the popSR (μ_{T1}, μ_{T2}), generating a "broomstick" shape. Parameter values are sampled from: $N_S(-1,0.2^2)$, $N_{k1}(1.8,0.12^2)$, $N_{k2}(1.08,0.12^2)$, $N_{i1}(15.3,0.05^2)$, $N_{i2}(15.08,0.05^2)$. (B',C', D') The popSRs for the three patterns of indSR (B, C, D) all have the same slope and intercept (OLS, MA, and SMA), and the same variance and covariance for T_1 and T_2.

Table 1. Parameter values used for initial populations.

Pattern of cryptic indSRs	Trait	μ_S	σ_S	μ_k	σ_k	μ_i	σ_i	*Mean trait size*	*Slope* $_{OLS}$	*Slope*$_{SMA}$	*Slope*$_{MA}$	μ_{k2}/μ_{k1}
Seesaw	1 (Body)	0	0.2	1.8	0.12	13.5	0.16	13.5	0.529	0.660	0.6	0.6
	2 (Wing	0	0.2	1.08	0.12	14	0.16	14				
Broomstick	1 (Body)	–1	0.2	1.8	0.12	15.3	0.05	13.5	0.529	0.660	0.6	0.6
	2 (Wing	–1	0.2	1.08	0.12	15.08	0.05	14				
Speedometer	1 (Body)	1	0.2	1.8	0.12	11.7	0.05	13.5	0.529	0.660	0.6	0.6
	2 (Wing)	1	0.2	1.08	0.12	12.92	0.05	14				

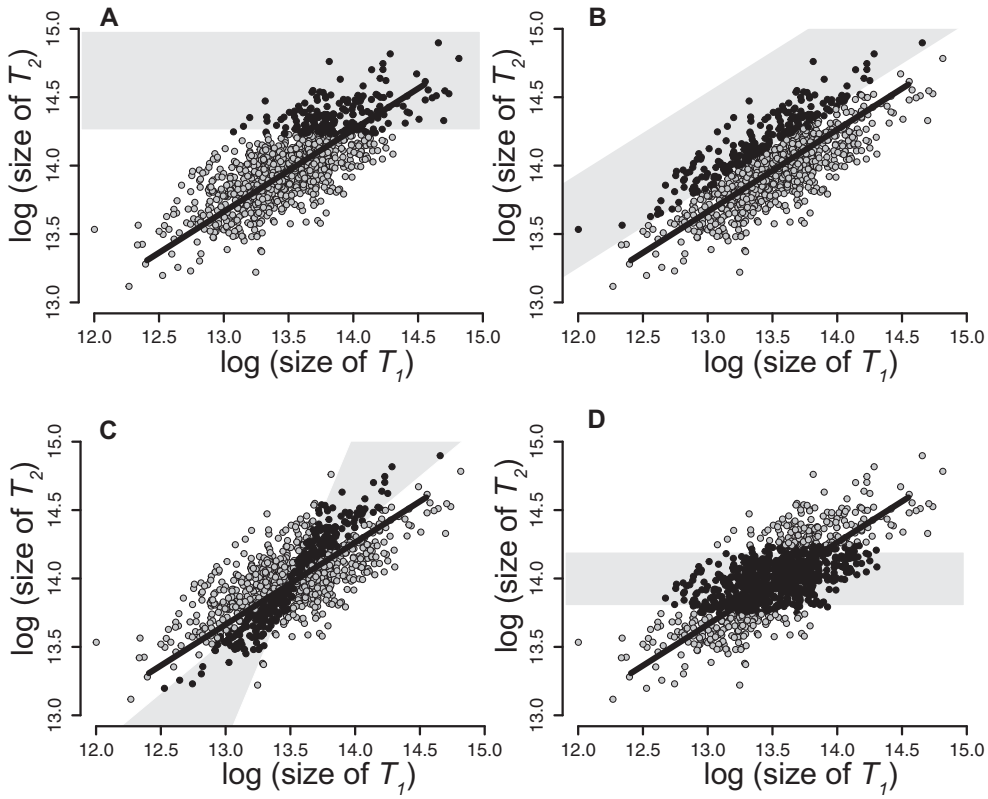

Figure 3. Selection regimes. (A) Directional selection; (B) Proportional selection; (C) Correlational selection; (D) Stabilizing selection. Only positive forms of selection are shown. Details of the methods used to identify selected individuals are detailed in Supporting Information. Black points = selected individuals, gray points = unselected individuals.

For each starting population of 1000 individuals we applied eight selection regimes (D+/-, P+/-, C+/-, St, and K+) for ten generations. Under truncated selection regimes, the intensity of selection at each generation is a function of the proportion of selected individuals (Falconer and Mackay 1996; Lynch and Walsh 1998). Consequently, we selected the same proportion of individuals (50%) at each generation to allow comparisons of the response across generations and among selection regimes. At each generation we recorded: the parameter values of each individual in the population, the selection differential and selection intensity on each parameter, and the slope (α) and intercept of the population scaling relationship using ordinary least squares (OLS), major axis (MA), and standardized major axis (SMA) regression. There has been much debate over which type of regression to use in studies of allometric relationships (e.g., Warton et al. 2006; O'Connor et al. 2007; Al-Wathiqui and Rodríguez 2011; Taskinen and Warton 2011; Hansen and Bartoszek 2012). Our model provides additional insight into this controversy, which we discuss in the Supporting Information. To construct confidence intervals around changes in mean parameter values and changes in slope and intercept of the popSRs, we repeated the analysis 1000 times, initiating each starting population using the same initial parameter distributions (Table 1).

The evolutionary algorithm was applied in R, and performed at Michigan State University's High Performance Computing Facility. The scripts used to run the algorithm are available

at Dryad (http://dx.doi.org/10.5061/dryad.1k17n). All data were subsequently analyzed using R or JMP.

RESPONSE TO SELECTION

Figure 4A–F shows how the slope of the popSR (α_{OLS}, α_{MA}, and α_{SMA}) and the mean slope of the cryptic indSRs ($\mu_{k2/k1}$) in each population class responds to different positive forms of direct, proportional, correlational, and stabilizing selection. These selection regimes have different effects on the slope of a population's static allometry. Further, the response to a particular selection regime depends strongly on the pattern of cryptic individual static allometries in the population. Indeed, the same selection regime can have opposite effects on the allometric coefficient for populations with ostensibly identical popSRs but different patterns of indSRs (e.g., Fig. 4A′-C′ vs. A′′-C′′).

There are several reasons why different selection regimes may vary in their effect on the slope of the popSR. One trivial explanation is that, even though the same proportion of individuals is selected each generation, the regimes may impose different realized selection intensities and therefore vary in their impact on the popSR. To standardize the selection pressure across selection regimes, we plotted the relationship between the MA slope of the popSR against the realized cumulative selection intensity under each selection regime (Fig. 5). These plots indicate that the different selection regimes had more-or-less the same intensity. Differences in selection intensity does not, therefore, explain differences in response to selection among regimes.

A more interesting possibility is that the selection regimes vary in the degree to which they target the developmental mechanisms that regulate the slopes of the individual cryptic scaling relationships that determine the slope of the population level scaling relationship. As discussed above, relative insulin-sensitivity k_2/k_1 is the biological factor that ultimately generates the correlation between the size of T_1 and T_2, the population average of which is captured by the slope of the population static allometry. Unsurprisingly, we found that selecting directly on individual k_2/k_1, (i.e., the slope of the indSR), generates the most rapid response in the slope of the popSR (Fig. 4G, G,' and G"). However, in biological organisms individual k_2/k_1 is phenotypically cryptic. Consequently, selection can only directly target aspects of the size of T_1 and T_2 (e.g., correlational selection on their combination or direct selection on the size of T_2). It follows that selection regimes targeting characters tightly correlated with k_2/k_1 should be more effective at changing the slope of the popSR than should regimes focused on characters weakly correlated with k_2/k_1. To test this, we examined the relationship between k_2/k_1 and the targets of selection in our three starting population classes. In support of our prediction, we found that the tighter the correlation (r^2) between k_2/k_1 and a selected character, the greater the effect of selection on the slope of the popSR (Fig. 6). The character most tightly correlated with an individual's k_2/k_1 varies among the three starting distributions of indSRs, which explains why the same selection regime can have very different effects on the slope of the popSR among population types.

Our model assumes a linear reaction norm between (log) trait size and the environmental regulator of size (level of insulin-signaling). However, the reaction norm for other environmental regulators of size may be curve-linear (Shingleton et al. 2007; Shingleton et al. 2009). We tested whether the results of our analysis are robust to other types of reaction norms by employing a second nonlinear model to the relationship between trait size T and environmental factor E (see Supporting Information). Although this modified model produces nonlinear individual trait reaction norms, it nevertheless generates linear indSRs (Fig. S1), as observed for nonlinear reaction norms in *Drosophila* (see Fig. 3 in (Shingleton et al. 2007)). We employed our evolutionary algorithm to evolve these populations under different selection regimes as described above. The results of our analysis (Fig. S2) indicate that linear popSRs generated from nonlinear reaction norms respond in the same way to selection as popSRs generated from linear-reaction norms, and that this response also depends on the underlying pattern of indSRs. Thus, our finding that the pattern of indSRs in a population determines how the popSR responds to selection may be a general characteristic of the evolution of environmental static allometries, regardless of the developmental mechanisms generating them.

Discussion

The evolution of morphology has long been associated with the correlation of traits within an organism. In the early 19th century, Cuvier recognized that traits were not independent entities, rather, they were mutually dependent to produce a functioning whole. His "Law of Correlation" was a formalization of his observations, and stated that the interdependency between traits precluded a single trait from changing as the organism could only function as a complete unit (Cuvier 1812). Darwin relaxed the notion of organisms as immutable units, but nevertheless recognized that traits are often correlated with one another and have the potential to constrain morphological diversity. Darwin defined the "correlation of growth" as changes in one trait that modify another trait, due to the tight organization of an organism's growth and development (Darwin 1859). Building on the relationship between traits, Thompson (1917) recognized that changes in the size of one trait relative to another underlies morphological diversity and described this transformation mathematically. Finally, Huxley and Tessier (1936) modeled the size relationship between traits using the allometric equation, and explicitly linked this relationship to the processes that control growth. Thus the modern study of scaling relationships was established.

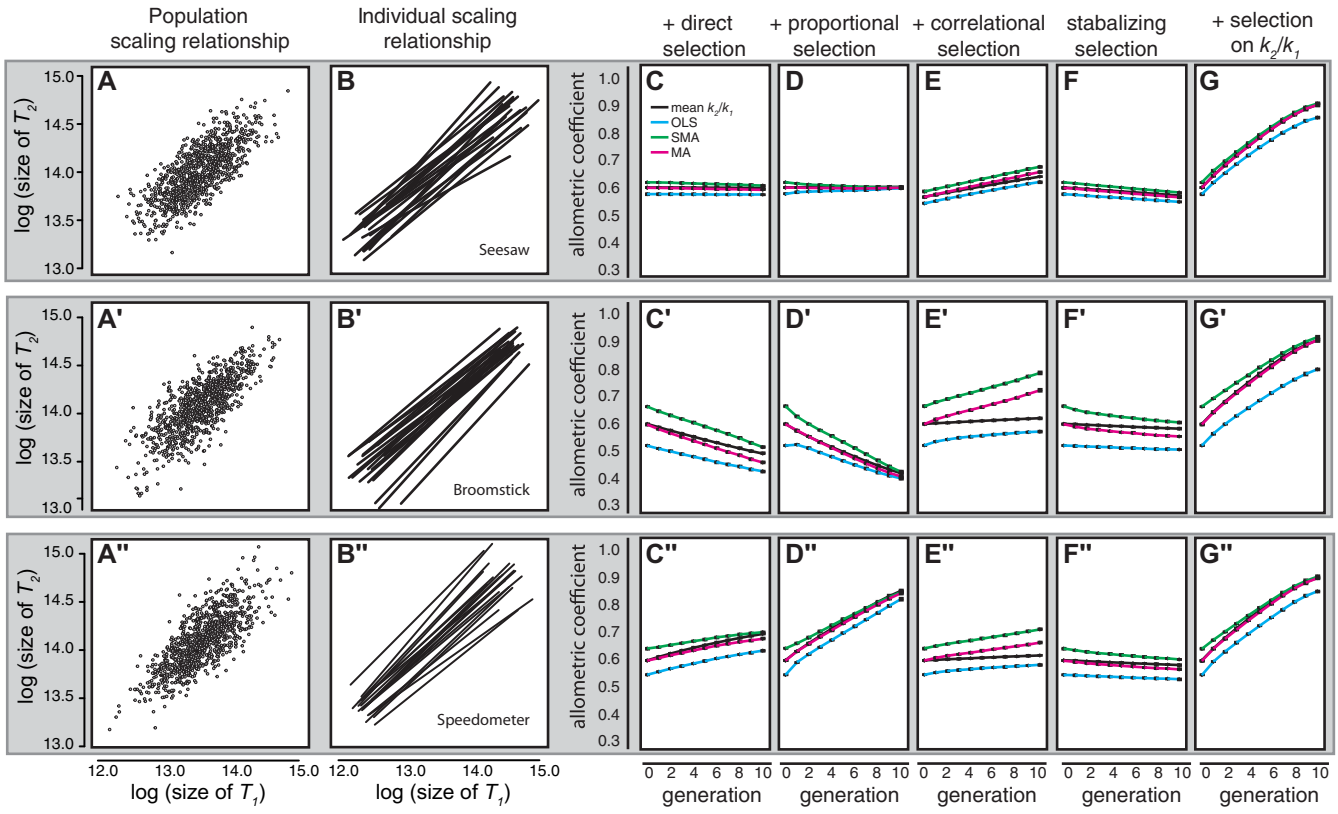

Figure 4. **The response of the allometric coefficient to selection in three different populations. (A, A′, A″) The three starting populations share the same popSR slope (OLS, MA, and SMA), the same bivariate mean and the same variance and covariance for logT_1 and logT_2. (B, B′, B″) The pattern of cryptic indSRs for the three populations are "seesaw," "broomstick," and "speedometer." (C–F″) The response of the popSR allometric coefficient to different forms of selection varies considerably among populations. (G–G″) In all three populations, direct selection on individual k_2/k_1 had the most rapid effect on the slope of the popSR. Points are mean allometric coefficient (OLS, MA, SMA) and $\mu_{k2/k1}$ of 1000 replicate populations across generations. Error bars (partially hidden) are 95% confidence intervals of the mean.**

Despite the centrality of trait correlations and scaling relationships to the evolution of morphology, the modes of selection that alter the intercept, and, particularly, the slope, of morphological scaling relationships remain controversial. This controversy has been confounded by experimental evolution studies that produce equivocal results when selection has ostensibly targeted the slope of a population directly (Egset et al. 2012; Bolstad et al. 2015; Stillwell et al. 2016), but substantial change of the slope when selection targets individual traits such as absolute or relative trait size (Robertson 1962; Wilkinson 1997; Tobler and Nijhout 2010) or when developmental manipulations are applied (Tang et al. 2011). Our model reveals why this may be. Specifically, our model demonstrates that the response of a population-level scaling relationship depends on the pattern of cryptic individual-level scaling relationships in the population. This pattern determines the extent to which a given form of selection targets particular developmental mechanisms that regulate the slope of the individual cryptic scaling relationships, which in turn give rise to the individual phenotypes on which the population level scaling relationship is based. These mechanisms control the relative sensitivity of traits to changes in the environmental factors (nutrition in our model) that generate the scaling relationship (k_2/k_1) and govern which phenotype is expressed along the cryptic individual scaling relationship in a given environment. Importantly, populations may share the same scaling relationship, yet harbor very different underlying patterns of individual-level scaling relationships. Consequently, populations that are ostensibly phenotypically identical may evolve very differently under the same form of selection.

The observation that individual genotypes have their own scaling relationship, although intuitive, is one that has been largely overlooked by previous studies on allometry. In those few studies where it has been considered, assumptions about the pattern of indSRs in the population may have been incorrect. Specifically, studies that attempted to alter the slope of a morphological scaling relationship through correlational selection explicitly or implicitly assumed that individuals in the upper-right and lower-left corners of a bivariate distribution of log(X) on log(Y) have a high indSR slope (Egset et al. 2012; Bolstad et al. 2015; Stillwell et al. 2016). That is, it was assumed that the pattern of cryptic indSRs was a "seesaw" (Fig. 2B). This need not be the case, which may account

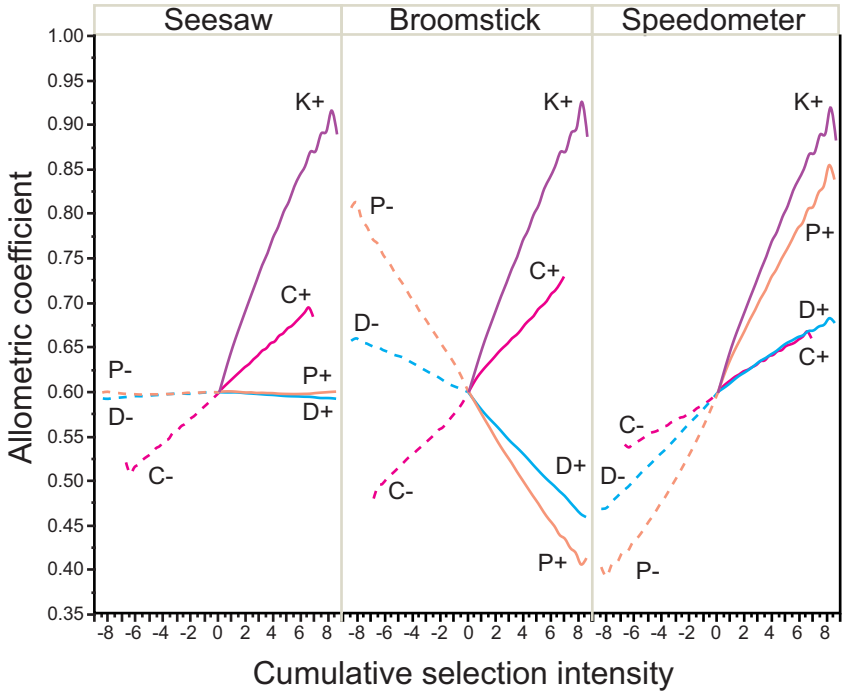

Figure 5. Response of the allometric coefficient to selection, standardized for the cumulative selection intensity. Lines are smooth curve through the data, comprising 1000 replicate populations through 10 generations of selection. The individual popSR allometric coefficients of the populations are shown in Figure 4. Solid lines are positive selection, dashed lines are negative selection, C, correlational selection; P, proportional selection; D, directional selection; and K, selection on k_2/k_1, the slope of the indSR.

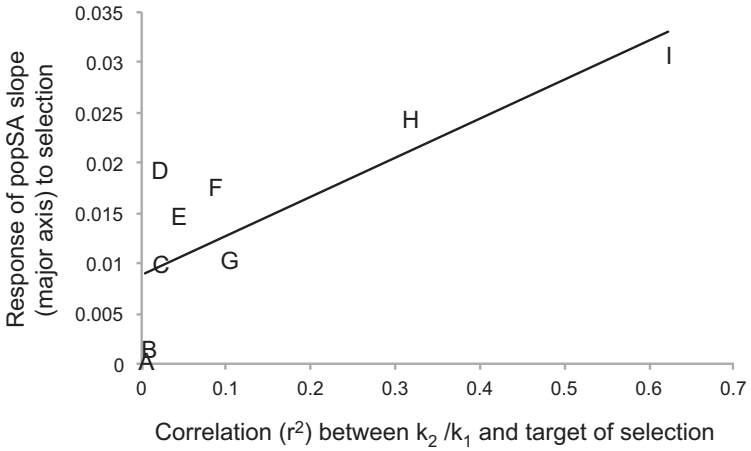

Figure 6. The greater the correlation (r^2) between k_2/k_1 and the target of selection in a population, the greater the response of that population's SA to selection. r^2 is calculated from the relationship between individual k_2/k_1 and the individual values of the targets of selection for a population of 1000 individuals. The response to selection is calculated as the slope of the standardized selection responses shown in Figure 5. Line is linear regression (OLS, $P = 0.01$). Letters indicate population and target of selection: A, seesaw, proportional; B, seesaw, directional; C, broomstick, correlational; D, speedometer, correlational; E, seesaw, correlational; F, broomstick, directional; G, speedometer, directional; H, broomstick, proportional; I, speedometer, proportional.

for the absent or erratic response to selection in these studies. Indeed, artificial selection experiments that have applied directional or proportional selection to single traits have generally generated the most consistent changes in allometric slope (e.g., Robertson 1962; Wilkinson 1997; Tobler and Nijhout 2010). This suggests that the pattern of cryptic indSRs, at least for the focal traits in these studies, was not the "seesaw" but rather was the "speedometer." In general, absent knowledge of the indSR pattern, prediction of the expected response of a population to a particular form of selection on scaling is difficult. Our model therefore provides a

compelling argument for measuring the pattern of indSRs in a population before exploring how that population responds to selection on the allometric slope. To generate indSRs it is necessary to rear genetically identical individuals across an environmental gradient. For organisms where it is easy to generate isogenic lines (e.g., *Drosophila*, Stillwell et al. 2011), or for clonal organisms (e.g., aphids, Stern et al. 1996), this is straightforward. However, for sexually reproducing nonmodel organisms, it may be necessary to use members of sibships as a proxy for genotype to estimate indSRs.

The observation that the distribution of cryptic individual scaling relationships can profoundly affect the response to selection raises the question: What causes a population to exhibit the "seesaw," "speedometer," or "broomstick" pattern of indSRs in the first place? Experimental manipulation of genes in the insulin signaling pathway in *Drosophila* suggest some candidate mechanisms. Variation in expression of the insulin receptor (*Inr*) in the wings of developing flies produces the "speedometer" pattern of wing-body indSRs (Fig. 2A, Shingleton and Tang 2012) whereas variation in expression of the forkhead transcription factor (*FOXO*) produces the "broomstick" (Fig. 2B, Shingleton and Tang 2012). These findings fit well with our model. *Inr* is a positive growth regulator and isogenic lineages that differ in expression of *Inr* in the wings vary in the extent to which nutritionally induced increases in *Inr* activity can raise wing growth rate above the intrinsic rate. In the context of our model, S therefore becomes the level of *Inr* activity. Because *Inr* is a positive growth regulator, S can only be positive and $\mu_S > 0$. Consequently, the pattern of indSRs among genotypes that vary in their sensitivity to this positive growth regulator is the speedometer. In contrast, FOXO is a negative growth regulator and the same arguments explain why variation among genotypes in wing-specific *FOXO* expression generates the broomstick pattern of indSRs. More generally, the pattern of inSRs in a population will reflect the extent to which genotypes vary in their sensitivity to positive versus negative environmentally regulated growth factors: in a population, variation in sensitivity to positive growth factors will produce the speedometer pattern of indSRs whereas variation in sensitivity to negative growth factors will generate the broomstick pattern of indSRs.

There has, historically, been much debate over the biological meaning of the parameters of Huxley and Tessier's allometric equation, that is the slope and intercept of linear morphological scaling relationships plotted on a log–log scale (Huxley 1950; White and Gould 1965; Gayon 2000). Our model undoubtedly simplifies the biological processes that underlie variation in trait size and that generate morphological scaling relationships (Nijhout 2011; Nijhout and German 2012). Nevertheless, the model makes explicit the types of developmental mechanisms that regulate Huxley and Tessier's parameters—insulin-sensitivity and organ-autonomous intrinsic growth rates for the slope and intercept, respectively—hypotheses that are supported by an increasing number of developmental studies (Cheng et al. 2011; Tang et al. 2011; Emlen et al. 2012; Refki et al. 2014; Refki and Khila 2015). The model also addresses the problem of delineating between changes in the slope of the scaling relationship and changes in the intercept (Egset et al. 2012; Bolstad et al. 2015; Stillwell et al. 2016). The model suggests that shifts in the intercept independent of the slope can occur through changes in a trait's intrinsic growth rate (i in eq. 10). The malleability of allometric intercepts is supported by numerous artificial selection experiments (Wilkinson 1997; Emlen 1996; Frankino et al. 2005; Frankino et al. 2007; Egset et al. 2012) and developmental manipulations of the insulin signaling pathway (Shingleton et al. 2005; Tang et al. 2011). In contrast, the model suggests that changing the slope independently of the intercept is much more problematic. This is because the mechanisms that affect the slope of a scaling relationship, insulin-sensitivity (k in eq. 10), pleiotropically affect the intercept, as shown in equation (10). This means that for selection to alter the slope of an allometry without altering the intercept, the effect of any change in insulin-sensitivity (k) on intercept needs to be countered by selection on intrinsic growth rate (i). This may explain why the response to correlational selection on the slope alone has been erratic (Egset et al. 2012; Bolstad et al. 2015; Stillwell et al. 2016): selection must target both mechanisms at the same time, and they must interact in a specific, body-size dependent, manner.

The model allows us to predict the pattern of indSRs and the genetic targets of selection in a population, given the observed response of scaling to a known form of selection. For example, the pleiotropic effect of insulin-sensitivity on slope and intercept likely reveals why directional or proportional selection can lead to the evolution of exaggerated hyperallometric traits often subject to sexual selection sexual selection (Green 1992; Emlen & Nijhout 2000; Koric-Brown et al. 2006; but see Bonduriansky, 2007). Exaggerated traits are those that exhibit an extreme increase in size relative to the ancestral state, such as the horns of male rhinocerous beetles (Enrodi 1985), the eyestalks of male stalk-eyed flies (Wilkinson 1997) and the antlers of male deer (Gould 1974). In many cases, such traits exhibit hyperallometric scaling, making them sensitive indicators of individual condition and hence informative when selecting a mate (Emlen and Nijhout 2000; Kodric-Brown et al. 2006). Exaggerated hyperallometric traits are thought to be more insulin-sensitive than traits that scale isometrically with body size (Shingleton et al. 2008; Emlen et al. 2012). This means that the same developmental mechanism—increased insulin-sensitivity (k in our model)—both exaggerates traits and causes them to be hyperallometric. Consequently, sexual selection for trait exaggeration can result in an indirect increase in signal reliability via the evolution of hyperallometry, creating positive feedback for greater trait exaggeration and increased hyperallometry (Emlen

et al. 2012). However, morphological diversity is replete with examples of enlarged traits that are not hyperallometric, and so it is unclear why sexual selection should target insulin-sensitivity specifically (Shingleton and Frankino 2013). Our model predicts that exaggerated traits have the speedometer pattern of indSRs, as directional or proportional selection for increased trait size from this indSR distribution causes rapid, indirect evolution of hyperallometry (Fig. 4). Further, the model predicts that proportional or directional selection regimes that generate trait hyperallometry will target the mechanisms that increase a trait's sensitivity to positive growth regulators, perhaps by increasing expression of *Inr*.

Conversely, if we know the genetic targets of selection we can use our model to predict the form of selection that generated a change in scaling. For example, hypoallometry of the male genitalia in *Drosophila* is a consequence of an organ-autonomous reduction in expression of *FOXO*, a mechanism that generates the broomstick pattern of indSRs (Shingleton and Tang 2012). Such patterns generate hypoallometry most efficiently through positive proportional or directional selection on trait size (Fig. 4). Consequently, the model predicts that genital hypoallometry is a consequence of positive directional or proportional selection on genital size. There is some evidence that this is the case: hypoallometric genital traits are under positive directional selection in the water strider *Aquarius remigis* (Bertin and Fairbairn 2007). Importantly, this prediction contrasts with existing hypotheses that invoke stabilizing selection on genital size to explain genital hypoallometry (Eberhard et al. 1998; Eberhard 2009), a form of selection that generates a weak evolutionary response in our model (Fig. 4).

Our model considers linear static allometries where size variation results from a single environmental factor. The model can easily be modified to consider alternative patterns of trait growth that may generate nonlinear morphological scaling relationships (Nijhout 2011), and to incorporate multiple environmental factors, where each trait has a unique sensitivity to each factor. Further, while our model addresses the evolution of scaling relationships generated through environmental variation (called environmental static allometries), it can be readily adapted to consider the evolution of scaling relationships generated through genetic variation (called genetic static allometries) (Shingleton et al. 2007), by having the value of S inherited rather than sampled anew each generation. Because genetic static allometries reflect pleiotropic relationships among traits, such a modification would allow the model to be applied to the evolution of pleiotropy. There is an extensive literature on the evolution of pleiotropy (see Pavlicev and Cheverud 2015 for review), an important component of which is epistatic pleiotropy, where the pleiotropic effects of variation at one locus on multiple traits varies with genetic background (Wolf 2005). Epistatic pleiotropy can cause variation in the genetic correlations among traits across genotypes, and thus generate a population of cryptic genetic static allometries analogous to the population of cryptic environmental static allometries described by our model (Pavlicev et al. 2008). Variation in environmental static allometries occurs when different traits vary in their response to the gene-by-environment interactions that generate variation in phenotypic plasticity (Shingleton et al. 2007). Similarly, variation in genetic static allometries occurs when different traits vary in their response to the gene-by-gene interactions that generate epistasis (Pavlicev et al. 2008), a phenomenon called differential epistasis (Cheverud 2001; Cheverud et al. 2004). Several studies have identified loci that influence the pleiotropy between traits and underlie differential epistasis, referred to as *relationship QTLs* (rQTLs) (Cheverud et al. 2004; Pavlicev et al. 2008), and the evolution of these rQTLs has been modelled mathematically (Pavlicev et al. 2011; Watson et al. 2014). Although these rQTL models have been applied to different evolutionary questions than our model, there are interesting and important parallels among all these models that warrant future consideration.

Despite its simplicity, the basic model developed here clearly demonstrates the importance of considering the developmental origins of the trait variation subject to selection. Traditional approaches to modeling the evolution of allometries that focus on patterns of trait and body size covariation (Bonduriansky and Day 2003; Pelabon et al. 2013) would predict that the three populations shown in Figure 2B′, C′ or D′ would have the same response to a given form and intensity of selection. This is because such approaches do not consider the actual target of selection—the mechanisms underlying trait and body size expression that produce the genotype-specific indSR—but instead consider only the observed phenotype expressed by each individual. Thus, these approaches could not predict, or explain, the responses to selection observed in Figure 4.

In conclusion, our model provides a fresh approach to understanding the evolution of morphological scaling relationships, by explicitly addressing how selection will act on the developmental parameters known to regulate growth and thereby determine trait allometries. The key insight is that each genotype encodes a cryptic individual scaling relationship (indSR), and variation in the developmental parameters among individuals generates specific patterns of cryptic indSR in a population. Although only part of the indSR is exposed to selection, the pattern of these indSRs in a population determines how the population-level scaling relationship will respond to selection. Thus, populations that differ in their distribution of indSRs will respond quite differently to the same pattern of selection, even if the populations are phenotypically indistinguishable. By explicitly linking patterns of selection acting on a population with the developmental targets of that selection among individuals, our model allows one to predict how morphological scaling relationships evolve, and to explain how

patterns of scaling relationships common among taxa, such as hyperallometric sexually selected traits, came to be. The model therefore provides a new and rich theoretical framework for future empirical research on the evolution of morphology.

ACKNOWLEDGMENTS

We thank C. Mirth and V. Callier for helpful comments while preparing this manuscript, and C. Anderson for assistance with coding. This work was funded by National Science Foundation grants IOS-0919855, IOS-0845847, and IOS-1557638 (to A.W.S.), IOS-0920720 and IOS-1558098 (to W.A.F.) and National Science Foundation Cooperative Agreement No. DBI-0939454 (to A.P.D.).

DATA ARCHIVING

The doi for our data is 10.5061/dryad.1k17n.

LITERATURE CITED

Al-Wathiqui, N., and R. L. Rodríguez. 2011. Allometric slopes not underestimated by ordinary least squares regression: a case study with enchenopatreehoppers (Hemiptera: Membracidae). Annal. Entomol. Soc. Am. 104:562–566.

Bakker, K. 1959. Feeding period, growth, and pupation in larvae of *Drosophila melanogaster*. Entomol. Exp. Appl. 2:171–186.

Bateson, P., D. Barker, T. Clutton-Brock, D. Deb, B. D'Udine, R. A. Foley, P. Gluckman, K. Godfrey, T. Kirkwood, M. M. Lahr, et al. 2004. Developmental plasticity and human health. Nature 430:419–421.

Bertalanffy, L. V. 1964. Basic concepts in quantitative biology of metabolism. Helgoländer Wiss. Meeresunters 9:5–37.

Bertin, A., and D. J. Fairbairn. 2007. The form of sexual selection on male genitalia cannot be inferred from within-population variance and allometry—a case study in *Aquarius remigis*. Evolution 61:825–837.

Bolstad, G. H., J. A. Cassara, E. Márquez, T. F. Hansen, K. van der Linde, D. HOULE, and C. Pelabon. 2015. Complex constraints on allometry revealed by artificial selection on the wing of *Drosophila melanogaster*. Proc. Natl. Acad. Sci. 112:201505357–201513289.

Bonduriansky, R. 2007. Sexual selection and allometry: a critical reappraisal of the evidence and ideas. Evolution 61:838–849.

Bonduriansky, R., and T. Day. 2003. The evolution of static allometry in sexually selected traits. Evolution 57:2450–2458.

Bryant, P. J., and P. Levinson. 1985. Intrinsic growth control in the imaginal primordia of Drosophila, and the autonomous action of a lethal mutation causing overgrowth. Dev. Biol. 107:355–363.

Cheng, L. Y., A. P. Bailey, S. J. Leevers, T. J. Ragan, P. C. Driscoll, and A. P. Gould. 2011. Anaplastic lymphoma kinase spares organ growth during nutrient restriction in Drosophila. Cell 146:435–447.

Cheverud, J. 2001. The genetic architecture of pleiotropic relations and differential epistasis. *In* G. P. Wagner, ed. The character concept in evolutionary biology. Academic Press, San Diego.

Cheverud, J. M., T. H. Ehrich, T. T. Vaughn, S. F. Koreishi, R. B. Linsey, and L. S. Pletscher. 2004. Pleiotropic effects on mandibular morphology II: differential epistasis and genetic variation in morphological integration. J. Exp. Zool. B Mol. Dev. Evol. 302B:424–435.

Cuvier, G., *Recherches sur les ossemens fossiles de quadrupèdes. Tome 1*. 1812: Deterville, Paris

Demment, M. W., and P. J. Van Soest. 1985. A nutritional explanation for body-size patterns of ruminant and nonruminant herbivores. Am. Nat. 125:641–672.

Eberhard, W. G. 2009. Static allometry and animal genitalia. Evolution 63:48–66.

Eberhard, W. G., B. A. Huber, R. L. Rodriguez, R. D. Briceno, I. Salas, and V. Rodriquez. 1998. One size fits all? Relationships between the size and degree of variation in genitalia and other body parts in twenty species of insects and spiders. Evolution 52:415–431.

Egset, C., T. Hansen, A. Le Rouzic, G. H. Bolstad, G. Rosenqvist, and C. Pelabon. 2012. Artificial selection on allometry: change in elevation but not slope. J. Evol. Biol. 25:938–948.

Emlen, D. J. 1996. Artificial selection on horn length-body size allometry in the horned beetle *Onthophagus acuminatus* (Coleoptera: Scarabaeidae). Evolution 50:1219–1230.

Emlen, D. J., and H. F. Nijhout. 2000. The development and evolution of exaggerated morphologies in insects. Ann. Rev. Entomol. 45:661–708.

Emlen, D. J., I. A. Warren, A. Johns, I. Dworkin, and L. C. Lavine. 2012. A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. Science 337:860–864.

Enrodi, S. 1985. The dynastinae of the world. Dr W Junk, Boston.

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

Falconer, D. S., and T. F. Mackay. 1996. Introduction to quantitave genetics. The Ronald Press Company, New York.

Frankino, W. A., D. L. Stern, B. J. Zwaan, and P. M. Brakefield. 2007. Internal and external constraints in the evolution of morphological allometries in a butterfly. Evolution 61:2958–2970.

Frankino, W. A., B. J. Zwaan, D. L. Stern, and P. M. Brakefield. 2005. Natural selection and developmental constraints in the evolution of allometries. Science 307:718–720.

Gayon, J. 2000. History of the concept of allometry. Am. Zool. 40:748–758.

Gould, S. J. 1966. Allometry and size in ontogeny and phylogeny. Biol. Rev. Camb. Phil. Soc. 41:587–640.

———. 1974. The origin and function of "bizarre" structures: antler size and skull size in the Irish elk, *Megaloceros giganteus*. Evolution 28:191–220.

Green, A. J. 1992. Positive allometry is likely with mate choice, competitive display and other functions. Anim. Behav. 43:170–172.

Hansen, T. F., and K. Bartoszek. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. Syst. Biol. 61:413–425.

Huxley, J. S. 1932. Problems of relative growth. Methuen & Co. Ltd., London.

———. 1950. A discussion on the measurement of growth and form; relative growth and form transformation. Proc. R Soc. Lond. B Biol. Sci. 137:465–469.

Huxley, J. S., and G. Tessier. 1936. Terminology of relative growth. Nature 137:780–781.

Kodric-Brown, A., R. M. Sibly, and J. H. Brown. 2006. The allometry of ornaments and weapons. Proc. Natl. Acad. Sci. USA 103:8733–8738.

Koh, I., M. S. Lee, N. J. Lee, K. W. Park, K. H. Kim, H. Kim, and I. J. Rhyu. 2005. Body size effect on brain volume in Korean youth. Neuroreport 16:2029–2032.

Koyama, T., C. C. Mendes, and C. K. Mirth. 2013. Mechanisms regulating nutrition-dependent developmental plassticity through organ-specific effects in insects. Front. Physiol. 4:263.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer, Sunderland, MA.

Martin, P. F. 1982. Direct determination of the growth rate of Drosophila imaginal discs. J. Exp. Zool. 222:97–102.

Nijhout, H. F. 2011. Dependence of morphometric allometries on the growth kinetics of body parts. J. Theoret. Biol. 288:35–43.

Nijhout, H. F., G. Davidowitz, and D. A. Roff. 2006. A quantitative analysis of the mechanism that controls body size in *Manduca sexta*. J. Biol. 5:16.11–16.15.

Nijhout, H. F., and R. Z. German. 2012. Developmental causes of allometry: new models and implications for phenotypic plasticity and evolution. Integr. Comp. Biol. 52:43–52.

Nijhout, H. F., L. M. Riddiford, and C. Mirth, C., Shingleton, A. W., Suzuki, Y. and Callier, V. 2014. The developmental control of size in insects. WIREs Dev Biol, 3:113–134.

O'Connor, M. P., S. J. Agosta, F. Hansen, S. J. Kemp, A. E. Sieg, J. N. McNair, and A. E. Dunham. 2007. Phylogeny, regression, and the allometry of physiological traits. Am. Nat. 170:431–442.

Pavlicev, M., and J. M. Cheverud. 2015. Constraints evolve: context dependency of gene effects allows evolution of pleiotropy. Ann. Rev. Ecol. Evol. Syst. 46:413–434.

Pavlicev, M., J. M. Cheverud, and G. P. Wagner. 2011. Evolution of adaptive phenotypic variation patterns by direct selection for evolvability. Proc. R Soc. B Biol. Sci. 278:1903–1912.

Pavlicev, M., J. P. Kenny-Hunt, E. A. Norgard, C. C. Roseman, J. B. Wolf, and J. M. Cheverud. 2008. Differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. Evolution 62:199–213.

Pelabon, C., G. H. Bolstad, C. K. Egset, M. P. Cheverug, and G. Rosenqvist. 2013. On the relationship between ontogenetic and static allometry. Am. Nat. 181:195–212.

Pelabon, C., C. Firmat, G. H. Bolstad, K. L. Voje, D. Houle, J. Cassara, A. L. Rouzic, and T. F. Hansen. 2014. Evolution of morphological allometry. Ann. NY Acad. Sci. 1320:58–75.

Refki, P. N., D. Armisen, A. J. Crumiere, S. Viala, and A. Khila. 2014. Emergence of tissue sensitivity to Hox protein levels underlies the evolution of an adaptive morphological trait. Dev. Biol. 392:441–453.

Refki, P. N., and A. Khila. 2015. Key patterning genes contribute to leg elongation in water striders. Evodevo 6:14.

Robertson, F. W. 1962. Changing the relative size of the body parts of Drosophila by selection. Genet. Res. 3:169–180.

Shingleton, A. W., J. Das, L. Vinicius, and D. L. Stern. 2005. The temporal requirements for insulin signaling during development in Drosophila. PLoS Biol. 3:e289.

Shingleton, A. W., C. M. Estep, M. V. Driscoll, and I. Dworkin. 2009. Many ways to be small: different environmental regulators of size generate distinct scaling relationships in *Drosophila melanogaster*. Proc. Roy Soc. Lond. B Biol. Sci. 276:2625–2633.

Shingleton, A. W., and W. A. Frankino. 2013. New perspectives on the evolution of exaggerated traits. BioEssays 35:100–107.

Shingleton, A. W., W. A. Frankino, T. Flatt, H. F. Nijhout, and D. J. Emlen. 2007. Size and shape: the developmental regulation of static allometry in insects. BioEssays 29:536–548.

Shingleton, A. W., C. K. Mirth, and P. W. Bates. 2008. Developmental model of static allometry in holometabolous insects. Proc. R Soc. Lond. B Biol. Sci. 275:1875–1885.

Shingleton, A. W., and H. Y. Tang. 2012. Plastic flies: the regulation and evolution of trait variability in Drosophila. Fly 6:147–152.

Stern, D. L., and D. J. Emlen. 1999. The developmental basis for allometry in insects. Development 126:1091–1101.

Stern, D. L., A. Moon, and C. M. Delrio. 1996. Caste allometries in the soldier-producing aphid pseudoregma—Alexanderi (Hormaphididae, Aphidoidea). Insectes Sociaux 43:137–147.

Stillwell, R. C., I. Dworkin, A. W. Shingleton, and W. A. Frankino. 2011. Experimental manipulation of body size to estimate morphological scaling relationships in Drosophila. J. Vis. Exp. e3162.

Stillwell, R. C., A. W. Shingleton, I. Dworkin, and W. A. Frankino. 2016. Tipping the scales: evolution of the allometric slope independent of average trait size. Evolution, 70:433–444.

Tang, H. Y., M. S. Smith-Caldas, M. V. Driscoll, S. Salhadar, and A. W. Shingleton. 2011. FOXO regulates organ-specific phenotypic plasticity in Drosophila. PLoS Genet. 7:e1002373.

Taskinen, S., and D. I. Warton. 2011. Robust estimation and inference for bivariate line-fitting in allometry. Biometr. J. Biometrische Zeitschrift 53:652–672.

Teleman, A. A. 2010. Molecular mechanisms of metabolic regulation by insulin in Drosophila. The Biochem. J. 425:13–26.

Thompson, D. W. 1917. On growth and form. Cambridge Univ. Press, Cambridge.

Tobler, A., and H. F. Nijhout. 2010. Developmental constraints on the evolution of wing-body allometry in *Manduca sexta*. Evol. Dev. 12:592–600.

Truman, J. W. 2006. Juvenile hormone is required to couple imaginal disc formation with nutrition in insects. Science 312:1385–1388.

Voje, K. L., and T. F. Hansen. 2013. Evolution of static allometries: adaptive change in allometric slopes of eye span in stalk-eyed flies. Evolution 67:453–467.

Voje, K. L., T. F. Hansen, C. K. Egset, G. H. Bolstad, and C. Pelabon. 2014. Allometric constraints and the evolution of allometry. Evolution 68:866–885.

Warton, D. I., I. J. Wright, D. S. Falster, and M. Westoby. 2006. Bivariate line-fitting methods for allometry. Biol. Rev. Cambridge Philos. Soc. 81:259–291.

Watson, R. A., G. P. Wagner, M. Pavlicev, D. M. Weinreich, and R. Mills. 2014. The evolution of phenotypic correlations and "developmental memory." Evolution 68:1124–1138.

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

White, J. F., and S. J. Gould. 1965. Interpretation of the coefficient in the allometric equation. dx.doi.org 99:5-18.

Wilkinson, G. S., and G. N. Dodson. 1997. Function and evolution of antlers and eye stalks in fllies. Pp. 310–327 *in* J. Choe, and B. Crespi, eds. The Evolution of Mating Systems in Insects and Arachnids. Cambridge University Press, Cambridge U.K.

Wilkinson, G. S., and G. N. Dodson. 1997. Function and evolution of antlers and eye stalks in flies. Evol. Mating Syst. Insects Arachnids.

Wolf, J. B. 2005. Epistatic pleiotropy and the genetic architecture of covariation within early and late-developing skull trait complexes in mice. Genetics 171:683–694.

Associate Editor: T. Flatt
Handling Editor: M. Servedio

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Parameter values used for initial populations with non-linear reaction norms.

Figure S1. Non-linear reaction norms. The relationship between log wing area and (a) developmental nutrition expressed as proportion of standard food and (b) developmental temperature.

Figure S2. The response of the allometric coefficient to selection in three different populations, where trait size has a non-linear reaction norm between (log) trait size and an environmental regulator of size. (A-A").