A Matter of Measurements: Challenges and Approaches in the Comparative Analysis of Static Allometries

Armin P. Moczek[*]

Department of Biology and Indiana Molecular Biology Institute, Indiana University, Bloomington, Indiana 47405

Submitted March 18, 2005; Accepted September 7, 2005; Electronically published March 13, 2006

ABSTRACT: Comparisons of static allometries are frequently used to gain insights into patterns and processes underlying morphological and developmental evolution. A study by J. L. Tomkins and co-workers, recently published in the *American Naturalist*, examined complex nonlinear allometries in three insect species in which males are dimorphic in the expression of secondary sexual traits. Employing a novel approach to analyzing male allometries in these organisms, the authors were able to show that developmental reprogramming of trait primordia is not necessary to explain allometric scaling in two of the species examined, contrary to several previous studies on the same species. Instead, male dimorphisms could be explained by simple exponential growth, an important result that carries with it major evolutionary and developmental implications. Using this study as an example, I highlight some of the methodological challenges involved in analyzing and comparing static allometries and in inferring the developmental processes that underlie them. I end by discussing how correct application of hypothesis testing, on one side, and basic anatomy and developmental biology, on the other, should guide how morphology is measured.

*Keywords:* alternative phenotypes, *Onthophagus*, plasticity, polyphenism, static allometry, reprogramming.

The evolution and diversification of scaling relationships, or allometries, is a classic theme in evolutionary biology that continues to receive major attention by evolutionary biologists (Huxley 1932; Thompson [1942] 1992; Emlen and Nijhout 2000). In particular, the comparative analysis of static allometries, that is, the relative sizes of the same body parts in individuals that belong to different populations, species, or higher-order taxa, continues to generate a wealth of insights into the allometric diversity that exists

* E-mail: armin@indiana.edu.

in nature and has helped formulate hypotheses regarding the evolutionary and developmental mechanisms that shape this diversity (Wheeler 1991; Nijhout and Wheeler 1996). More recently, insights into the genetic and developmental underpinnings of relative growth are providing opportunities to further evaluate and refine these hypotheses (Abouheif and Wray 2002; Nijhout and Grunert 2002; Frankino et al. 2005; Miura 2005). In this comment, I examine a study recently published in the *American Naturalist* by Tomkins, Kotiaho, and LeBas (2005) that analyzed static allometries to explore the developmental basis of dimorphic growth in male insects. This examination is not meant to question the main findings and conclusions of Tomkins et al.; instead I am using their study as an example to highlight some of the common challenges faced by comparative allometric studies in general and to point out some simple approaches that can help to overcome these challenges.

Methodological Challenges to the Comparative Analysis of Static Allometries: A Case Study

Tomkins et al. (2005) analyzed a set of static allometries to examine a specific hypothesis considering the developmental basis of nonlinear allometries: the notion of developmental reprogramming of growth primordia during the development of male dimorphisms in insects (Nijhout and Wheeler 1996; Emlen and Nijhout 2000). In particular, Tomkins et al. examined static allometries of male forceps length in a species of earwig and male horn length in two species of dung beetle. All three species had been the focus of allometric studies in the past (Cook 1987; Simmons and Tomkins 1996; Hunt and Simmons 1997; Tomkins 1999), and one of the dung beetle species, *Onthophagus taurus*, has become somewhat of a model system for studies on the evolution and development of extreme male dimorphisms in insects (Moczek 1998; Emlen and Nijhout 1999, 2000, 2001; Moczek and Emlen 1999; Nijhout and Emlen 1998; Simmons et al. 1999; Hunt and Simmons 2000, 2002; Moczek and Nijhout 2002, 2003). Several of these earlier studies generally concluded that severe reprogramming of growth parameters that govern horn de-
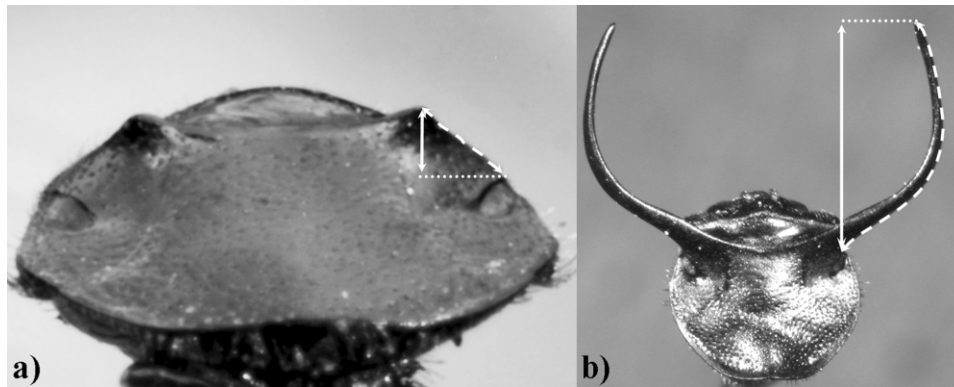
Figure 1: Differences in horn height (*solid line*) and horn length (*dashed line*) measurements of horn size in male *Onthophagus taurus*. Differences are illustrated for a small, minor male (*a*) and a large, major male (*b*).

velopment occurs around a critical threshold body size in male *O. taurus*. Reprogramming of growth parameters was thought to allow males that exceed this size threshold to develop greatly enlarged horns, whereas males that do not exceed this size threshold develop only rudimentary horns. One of the surprising findings of Tomkins et al. was that a simple log transformation of body size and horn length measurements was sufficient to linearize male horn allometries, which was consistent with the hypothesis that simple exponential growth of horn primordia may be sufficient to generate male dimorphism in these organisms. This contradicted the reprogramming hypothesis, which would predict the existence of discontinuous scaling even in log-transformed data. The only discontinuity detected by Tomkins et al. occurred at the very largest body sizes. There the increase of male horn lengths with body size seemed to slow down, and Tomkins et al. interpreted this finding as a possible reflection of local resource exhaustion after exponential growth. The notion that exponential growth of horn primordia occurred even at small body sizes received further support from comparing the slopes of log-transformed horn allometries of small males with those of other, nondimorphic traits, such as length of wings, tibiae, or femurs. The surprising result was that even at small body sizes, male horn length increased with a much higher slope than did the size of other traits. Combined, these results suggested that male horn dimorphism might not be the product of body size–dependent reprogramming of horn growth but instead might simply be the outcome of exponential growth, possibly constrained by resource limitations in the very largest males.

This scenario represents a previously overlooked and possibly very important alternative to the reprogramming hypothesis for the development of highly discontinuous, nonlinear allometries. Specifically, it contradicts the notion

that horned and hornless male morphs constitute alternative phenotypes that are developmentally decoupled from one another. Such developmental decoupling, or semi-independence, of alternative phenotypes represents a crucial component in the argument that conditional alternative phenotypes may represent an important phase in the origin of evolutionary novelties (West-Eberhard 2003) or speciation (Pfennig and Murphy 2000, 2002). Developmental decoupling is thought to permit alternative phenotypes to originate, adapt, and diversify alongside, and without having to abandon, established ancestral phenotypes. *Onthophagus* beetles have been cited frequently as a possible example of such a scenario because they exhibit extreme intra- and interspecific diversity of form as well as phenomenal species richness (reviewed in Moczek 2006). Understanding the mechanisms and phenotypic consequences of developmental decoupling, for example, via reprogramming or lack thereof, is therefore crucial for evaluating the role of alternative phenotypes in the origin of phenotypic diversity in these and other organisms. The results of Tomkins et al. suggest that reprogramming and developmental decoupling do not occur during the development of horned and hornless male morphs, which stands in sharp contrast to previous studies on the same species (e.g., Emlen and Nijhout 1999, 2001; Emlen 2000; Moczek and Nijhout 2002). Below, however, I show that in this particular case the novelty of the findings may simply arise because of a common problem in evolutionary morphology: different studies use different ways to measure the same trait in different populations.

Small Differences in Methodology Can Go a Long Way

Tomkins et al. (2005) used the same measure of body size as previous studies: the width of the pronotum at its widest

point. They differ, however, in how they measured horn size. Tomkins et al. measured horn height, that is, the vertical distance between the eye and the tip of the horn (J. L. Tomkins, personal communication). This measure always measures the vertical component of horn growth, regardless of width or curvature of the horn (fig. 1). Previous studies, in particular those that ended up favoring the reprogramming hypothesis (e.g., Emlen and Nijhout 1999, 2000, 2001; Emlen 2000; Moczek and Nijhout 2003) measured horn length, that is, the distance between where the eye cavity meets the edge of the head and the tip of the horn. Horn length measurements follow the outer edge of the horn and take changes in width and curvature into account (fig. 1; see also Moczek 1998; Moczek and Emlen 1999). To illustrate the consequences of these superficially subtle differences in measurement technique, I used both techniques to measure the same set of 50 individuals. Two major differences emerge. In small males, horn height measurements generate greater differences in horn sizes between males than do horn length measurements (fig. 2). Simply put, one horn may be twice as high as another, but it is not twice as long. The difference between height and length measurements is a function of the cone shape of small horns (wide at the base, pointy on top) and decreases with increasing horn size or as horn shape begins to approximate that of a tube. The use of height instead of length measures is responsible for the much faster increase in horn size values among small males in the Tomkins et al. study, compared to those in previous studies (fig. 2).

The second major discrepancy arises in large males. As male size increases, horn shape changes from straight to curved (fig. 1). The largest males develop horns that bend nearly 90°. Tomkins et al.'s height measurements, however, measure only the vertical component of horn growth, regardless of shape, whereas horn length measurements follow the outer edge of the horn and thus are able to take some shape changes into account. As a consequence, two horns can have the same height but very different lengths. This difference causes the horn height measurements in the largest males to be consistently smaller than horn length measurements (fig. 2). This reduces the differences in relative horn sizes between small and large males in the Tomkins et al. study, compared to those in previous publications.

Both discrepancies summarized above are evident in both nontransformed and log-transformed data (fig. 2a, 2b). More importantly, they also generate differences in the degree to which they support or contradict the reprogramming and exponential-growth hypotheses. Horn length measurements generate strikingly discontinuous allometries even in the log-transformed data, as predicted by the reprogramming hypothesis, whereas horn height measurements do not

(fig. 2b). The slope of log-transformed horn length allometries among minor males (defined for the purposes of this analysis as males with a pronotum width of <4.7 mm; $n = 30$) is 1.98 (± 0.55), which contradicts the notion of highly exponential horn growth in small males. The slope of log-transformed horn height allometries is 11.04 (± 0.44), which supports highly exponential growth. The difference between the two slope estimates is highly significant (Welch's *T*-test: $t = 12.853$, df $= 79$, $P < .0001$). Clearly, differences in how the size of horns is measured have a substantial effect on the results of the analysis and the conclusions that may be drawn from them.

Hypotheses and Development Should Guide Method Selection

So how should morphological traits like beetle horns be measured? Generally, any linear measure used to estimate the size of a three-dimensional structure is bound to have shortcomings, as it will always be able to depict only one dimension of a multidimensional trait. This can be overcome only through multivariate analyses such as geometric morphometrics, which allows for the capture of a large number of dimensions and the quantification and comparison of shape. This method, however, is also methodologically and computationally intense, causing many researchers to continue to prefer simple, bivariate plots of a single-component measure of a trait of interest against body size. The question then becomes which component should be used to measure a given trait of interest and whether there are general guidelines that can be used in the selection of such a component. Below I use the Tomkins et al. (2005) study and findings as an example to argue that the answers to the above questions should depend, on one side, on the nature of the hypothesis that is being tested and, on the other, on the specific anatomical and developmental underpinnings of the trait in question.

The reprogramming hypothesis was initially developed only for holometabolous insects, and it proposed that discontinuity of adult form arises as a consequence of environment- and hormone-mediated changes in the critical body size at which a developing larva takes the first physiological steps toward metamorphosis (Wheeler 1991; Nijhout and Wheeler 1996). Since then, the notion of developmental reprogramming has been extended to the level of individual body parts, such as imaginal disks, where it can result in the discontinuous expression of traits (such as horns) relative to a continuous expression of body size. The reprogramming hypothesis initially emphasized the role of endocrine factors as reprogramming agents but left open exactly what aspects of development were being reprogrammed in the process. Subsequent studies have since identified a range of developmental mechanisms that
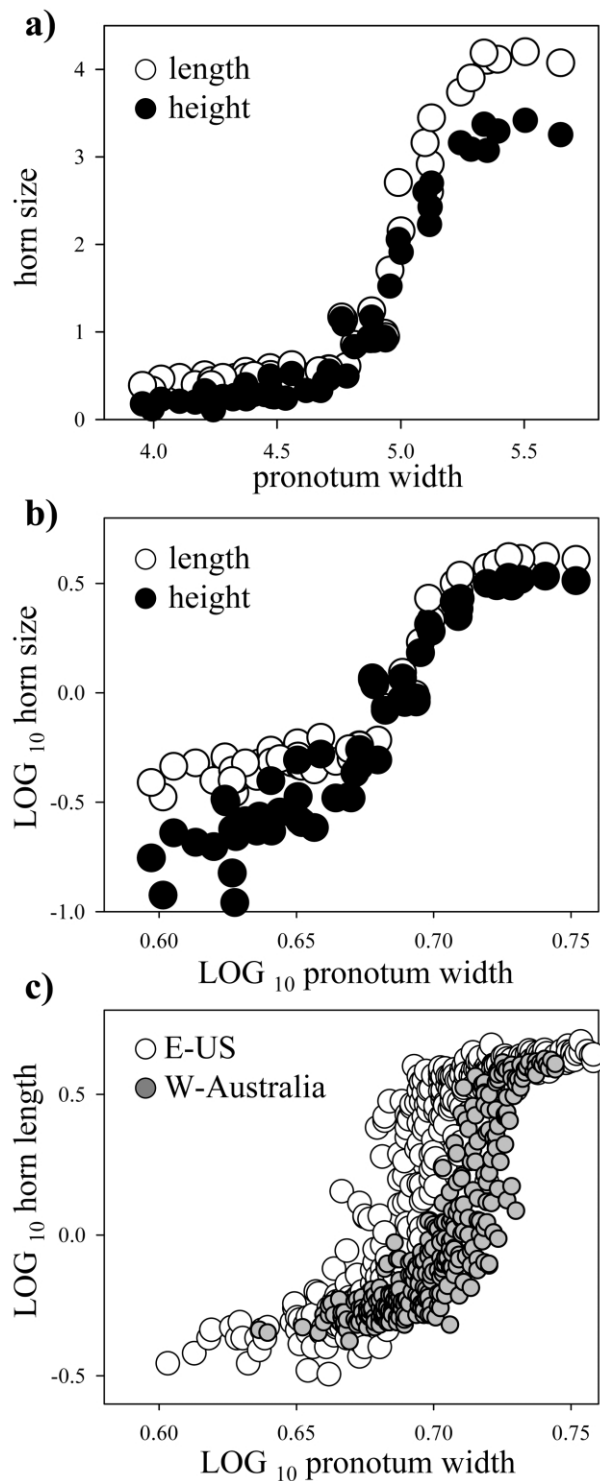
represent targets of reprogramming, such as the induction or inhibition of gene expression (Evans and Wheeler 1999; Abouheif and Wray 2002), differential apoptosis of organ primordia (Sameshima et al. 2004), or up- or downregulation of cell size or rate of cell division (Nijhout and Grunert 2002). An important insight from these studies is also that reprogramming commonly affects some components of a given trait but not others, for example, the length of an appendage but not necessarily its width at the base. The crucial implication of these observations is that in order to test for developmental reprogramming in a given multidimensional trait, one must measure it in a variety of ways, for not all of its components will be affected by a hypothetical reprogramming event. If none of the measurements indicate discontinuity, reprogramming may be rejected. However, relying on a single measure, such as horn height in the Tomkins et al. study, to reject developmental reprogramming during growth is simply premature, especially if earlier studies have come to the opposite conclusion using different measures. All that Tomkins et al. can conclude is that with respect to the vertical component of horn growth, no reprogramming appears to occur. Consequently, while single measures may suffice to support the reprogramming hypothesis, falsifying it with confidence can be achieved only by measuring multiple components of the same trait.

Apart from the general nature of the hypothesis that is being examined, the specific anatomical and developmental underpinnings of a trait in question may also help narrow down the most appropriate ways of measuring it, albeit with more limited ability to generalize across traits. In the case of beetle horns, the situation is actually relatively simple. Horns do not contain joints, muscles, or nerves and are instead simple hollow outgrowths bordered on the outside by a single layer of epidermal cells, which in turn is covered with a thick cuticle secreted by these same cells (Emlen and Nijhout 1999; Moczek and Nagy 2005). Horn precursors grow underneath the larval cuticle and first become visible as they expand outward during the molt to the pupal and, subsequently, adult stages. The more horn precursor cells divide before synthesis of the future cuticle, the more horn growth takes place, independent of exactly how the entire structure will ultimately unfold. In this case, horn height measurements can capture only investment into growth that increases the vertical component of horn size but ignore investment into cell growth, division, and cuticular synthesis of those portions

Figure 2: Consequences of different horn size measurements. *a*, Comparison of horn height (*filled circles*) and horn length (*open circles*) measurements of the same 50 individuals. *b*, The same measurements after $\log_{10}$ transformation. Note maintenance of sigmoid, discontinuous scaling in horn length measurements but not in horn height measurements. *c*, Comparison of $\log_{10}$-transformed horn length measurements from populations of *Onthophagus taurus* in the eastern United States (*open circles*) and Western Australia (*gray circles*). Data from Moczek (2003).

of the horn contributing to nonvertical growth, such as curvature. Horn curvature can at times be extreme and is probably functionally important (Moczek and Emlen 2000), yet horn height measurements treat these structures as if they were straight outgrowths. Horn height measurements may also be inappropriate for measuring straight horns of small males, but for a different reason. Small horns resemble short, pointy cones, and as small males increase in body size, horns increase primarily in height but not in width at the base. Like large horns, they consist of a hollow single layer of epidermal cells covered in cuticle. To increase the height of such a cone-shaped horn, the underlying epidermal precursor cells need to multiply, yet doubling the height of such a horn does not require doubling the number of underlying epidermal cells, doubling cuticular synthesis, or, in general, doubling horn growth. This would be true only if the horns of small males were shaped like tubes of constant diameter, which they are not (in contrast to, e.g., the tibia or femur, which do roughly approximate tubes; see below). Thus, if only a single linear measure is to be used to measure horn size of small males (rather than a three-dimensional, multivariate description of horn size), horn length may be a developmentally more sensible means of capturing differences in male horn growth as a function of male body size.

But what about the measurements of wings (width and length), leg segments (tibia length, femur length, tarsus length), head width, and abdominal sternite length used by Tomkins et al. to contrast horn growth to the growth of more typical insect parts? Unlike horns in male *Onthophagus taurus*, legs, wings, heads, and abdominal sclerites do not seem to change their shape much, if at all, with increasing body size. In these cases, a single length or width measurement is more likely to capture the same relative size component across the entire range of male body sizes and thus provides a more adequate means of describing the growth of these structures. In contrast, horn height measurements capture only the vertical component of horn growth, whose contribution to overall horn size changes dramatically as males get larger and horn shape changes.

The above comparisons highlight the importance of selecting the appropriate measurement techniques for morphometric data collection. They also highlight the importance of using identical methods when making comparisons to other studies on the same organism. It is here that Tomkins et al. inadvertently generated a third discrepancy between their results and those of previous studies by measuring individuals derived from a different and morphometrically highly divergent population. Previous studies have relied on individuals collected in the southeastern United States (e.g., Emlen and Nijhout 1999,

2000, 2001), whereas Tomkins et al. used individuals collected in Western Australia. Both U.S. and Australian populations are exotic and are derived from Mediterranean ancestors introduced approximately 40 years ago (reviewed in Moczek 2003). Australian *O. taurus* differ from U.S. populations in that the threshold body size above which males develop large horns has shifted to much larger body sizes (Moczek and Nijhout 2003). At the same time, mean body size of Australian males is smaller than that of U.S. males (Moczek 2003). Combined, these differences cause Australian populations to contain many fewer horned males, but more importantly, they confine the expression of large horns to only a relatively narrow body size range, making it easier for a simple power function to generate a good fit to the data, especially when horn height measurements are used (fig. 2*c*). While by itself such an analysis is, of course, perfectly acceptable, it can be problematic if the results are meant to be contrasted with those of previous studies.

Conclusions

As emphasized above, these arguments are not meant to disqualify the possible significance of the exponential-growth hypothesis, which in fact received strong support from the morphometric study on a second dung beetle species, *Onthophagus binodis*. Instead, my intent was to highlight how at times subtle differences in data collection and comparison can generate strikingly different results with very different implications, which thus need to be interpreted with care. I also hope to have shown that both the nature of the hypothesis that is being investigated and the specific anatomical and developmental underpinnings of the trait in question can provide guidance in deciding how best to quantify and contrast growth and size.

Acknowledgments

Literature Cited

Abouheif, E., and G. A. Wray. 2002. Evolution of the gene network underlying wing polyphenism in ants. Science 297:249–252.

Cook, D. 1987. Sexual selection in dung beetles. I. A multivariate study of morphological variation in two species of *Onthophagus*. Australian Journal of Zoology 35:123–132.

Emlen, D. J. 2000. Integrating development with evolution: a case study with beetle horns. BioScience 50:403–418.

Emlen, D. J., and H. F. Nijhout. 1999. Hormonal control of male horn length dimorphism in the dung beetle *Onthophagus taurus* (Coleoptera: Scarabaeidae). Journal of Insect Physiology 45:45–53.

———. 2000. The development and evolution of exaggerated morphologies in insects. Annual Review of Entomology 45:661–708.

———. 2001. Hormonal control of male horn length dimorphism in *Onthophagus taurus* (Coleoptera: Scarabaeidae): a second critical period of sensitivity to juvenile hormone. Journal of Insect Physiology 47:1045–1054.

Evans, J. D., and D. E. Wheeler. 1999. Differential gene expression between developing queens and workers in the honey bee, *Apis mellifera*. Proceedings of the National Academy of Sciences of the USA 96:5575–5580.

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.

Hunt, J., and L. W. Simmons. 1997. Patterns of fluctuating asymmetry in beetle horns: an experimental examination of the honest signalling hypothesis. Behavioral Ecology and Sociobiology 41:109–114.

———. 2000. Maternal and paternal effects on offspring phenotype in the dung beetle *Onthophagus taurus*. Evolution 54:936–941.

———. 2002. The genetics of maternal care: direct and indirect genetic effects on phenotype in the dung beetle *Onthophagus taurus*. Proceedings of the National Academy of Sciences of the USA 99:6828–6832.

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

Miura, T. 2005. Developmental regulation of caste-specific characters in social-insect polyphenisms. Evolution and Development 7:122–129.

Moczek, A. P. 1998. Horn polyphenism in the beetle *Onthophagus taurus*: larval diet quality and plasticity in parental investment determine adult body size and male horn morphology. Behavioral Ecology 9:636–641.

———. 2003. The behavioral ecology of threshold evolution in a polyphenic beetle. Behavioral Ecology 14:831–854.

———. 2006. Developmental plasticity and the origins of diversity: a case study on horned beetles. *In* T. N. Ananthakrishnan and D. Whitman, eds. Phenotypic plasticity in insects: mechanisms and consequences. Science Publishers, Plymouth (forthcoming).

Moczek, A. P., and D. J. Emlen. 1999. Proximate determination of male horn dimorphism in the beetle *Onthophagus taurus* (Coleoptera: Scarabaeidae). Journal of Evolutionary Biology 12:27–37.

———. 2000. Male horn dimorphism in the scarab beetle *Onthophagus taurus*: do alternative reproductive tactics favor alternative phenotypes? Animal Behaviour 59:459–466.

Moczek A. P., and L. M. Nagy. 2005. Diverse developmental mechanisms contribute to different levels of diversity in horned beetles. Evolution and Development 7:175–185.

Moczek, A. P., and H. F. Nijhout. 2002. Developmental mechanisms of threshold evolution in a polyphenic beetle. Evolution and Development 4:252–264.

———. 2003. Rapid evolution of a polyphenic threshold. Evolution and Development 5:259–268.

Nijhout, H. F., and D. J. Emlen. 1998. Competition among body parts in the development and evolution of insect morphology. Proceedings of the National Academy of Sciences of the USA 95:3685–3689.

Nijhout, H. F., and L. W. Grunert. 2002. Bombyxin is a growth factor for wing imaginal disks in Lepidoptera. Proceedings of the National Academy of Sciences of the USA 99:15446–15450.

Nijhout, H. F., and D. E. Wheeler. 1996. Growth models of complex allometries in holometabolous insects. American Naturalist 148:40–56.

Pfennig, D. W., and P. J. Murphy. 2000. Character displacement in polyphenic tadpoles. Evolution 54:1738–1749.

———. 2002. How fluctuating competition and phenotypic plasticity mediate species divergence. Evolution 56:1217–1228.

Sameshima, S. Y., T. Miura, and T. Matsumoto. 2004. Wing disc development during caste differentiation in the ant *Pheidole megacephala* (Hymenoptera: Formicidae). Evolution and Development 6:336–341.

Simmons, L. W., and J. L. Tomkins. 1996. Sexual selection and the allometry of earwig forceps. Evolutionary Ecology 10:97–104.

Simmons, L. W., J. L. Tomkins, and J. Hunt. 1999. Sperm competition games played by dimorphic male beetles. Proceedings of the Royal Society of London B 266:145–150.

Thompson, D. W. (1942) 1992. On growth and form. Dover, Mineola, NY.

Tomkins, J. L. 1999. Environmental and genetic determinants of the male forceps length dimorphism in the European earwig *Forficula auricularia* L. Behavioral Ecology and Sociobiology 47:1–8.

Tomkins, J. L., J. S. Kotiaho, and N. R. LeBas. 2005. Matters of scale: positive allometry and the evolution of male dimorphisms. American Naturalist 165:389–402.

West-Eberhard, M. J. 2003. Developmental plasticity and evolution. Oxford University Press, New York.

Wheeler, D. E. 1991. Developmental basis of worker caste polymorphism in ants. American Naturalist 138:1218–1238.