PATTERNS OF THRESHOLD EVOLUTION IN POLYPHENIC INSECTS UNDER DIFFERENT DEVELOPMENTAL MODELS

Joseph L. Tomkins[1,2] and Armin P. Moczek[3,4]

[1]*The Centre for Evolutionary Biology, School of Animal Biology, University of Western Australia, Crawley 6009, Western Australia, Australia*

[2]*E-mail: jtomkins@cyllene.uwa.edu.au*

[3]*Department of Biology, Indiana University, 915 E. Third Street, Myers Hall 150, Bloomington IN 47405-7107*

[4]*E-mail: armin@indiana.edu*

Two hypotheses address the evolution of polyphenic traits in insects. Under the developmental reprogramming model, individuals exceeding a threshold follow a different developmental pathway from individuals below the threshold. This decoupling is thought to free selection to independently hone alternative morphologies, increasing phenotypic plasticity and morphological diversity. Under the alternative model, extreme positive allometry explains the existence of alternative phenotypes and divergent phenotypes are developmentally coupled by a continuous reaction norm, such that selection on either morph acts on both. We test the hypothesis that continuous reaction norm polyphenisms, evolve through changes in the allometric parameters of even the smallest males with minimal trait expression, whereas threshold polyphenisms evolve independent of the allometric parameters of individuals below the threshold. We compare two polyphenic species; the dung beetle *Onthophagus taurus,* whose allometry has been modeled both as a threshold polyphenism and a continuous reaction norm and the earwig *Forficula auricularia*, whose allometry is best modeled with a discontinuous threshold. We find that across populations of both species, variation in forceps or horn allometry in minor males are correlated to the population's threshold. These findings suggest that regardless of developmental mode, alternative morphs do not evolve independently of one another.

KEY WORDS: Alternative tactics, dimorphism, earwig, horned beetles, *Onthophagus*, plasticity, polyphenism, sexual selection.

Understanding threshold evolution is important for the interpretation of phenotypic plasticity and the evolution of alternative life histories (Tomkins and Hazel 2007). In particular, thresholds have been considered important to the origins of diversity, where alternative phenotypes are thought to reflect selection acting independently on either side of a developmental schism, thereby producing divergent phenotypic optima (West-Eberhard 1986, 1989, 2003). Threshold models for the evolution of alternative reproductive morphologies in insects, such as those in dimorphic beetles, have concentrated on a developmental model in which a cue, such as body size, leads to a reprogramming event that radically alters the developmental trajectory of the larva (Fig. 1A;

Emlen and Nijhout 1999; Nijhout 1999; Emlen 2000; Emlen and Nijhout 2000, 2001; Moczek and Nijhout 2002; Nijhout 2003; Emlen and Allen 2004; Hartfelder and Emlen 2005). The reprogramming event in these models is of fundamental importance to the notion that selection can act independently on the phenotypes on either side of the threshold (West-Eberhard 1986, 1989; Emlen and Nijhout 2000; West-Eberhard 2003).

Recently, Tomkins et al. (2005) used Nijhout and Wheeler's (1996) models of polyphenic development in insects to suggest that some patterns of dimorphism that were thought to be caused by developmental reprogramming could be readily explained by a reaction norm of extreme positive allometry. In other words, that
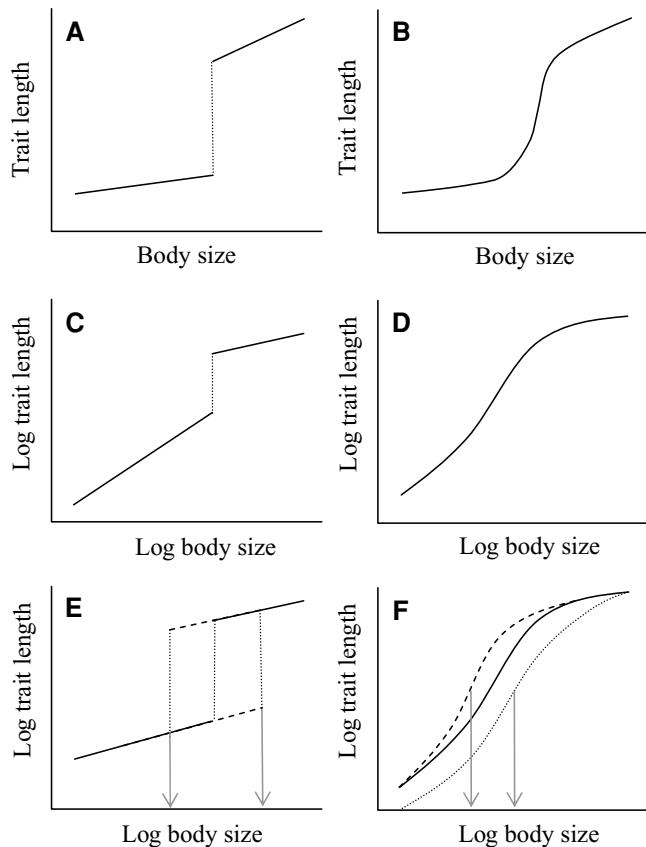
Figure 1. Patterns of dimorphism and changes in threshold switchpoint or reaction norm under two alternative developmental models. Both the reprogramming (A) and the positive allometry model (B) give rise to similar patterns when they are presented on a linear scale. On the log scale the pattern under the reprogramming model (C) retains the vertical displacement of the linear sections whereas the positive allometry model (D) produces a single continuous reaction norm (see Fig. 2). We hypothesis that changes in switchpoint (gray arrows) under the reprogramming model (E) take place because of a shift in the critical size at reprogramming but leave allometry unchanged, whereas under the positive allometry model (F) changes in the body size at which trait expression appears to suddenly increase can only result from changes in either the slope (dashed line) or the intercept (dotted line) of the trait allometry (Tomkins et al. 2005).

continuous phenotypic modulation, could account for patterns attributed wholly to developmental conversion (Smith-Gill 1983). For example, in two species of dimorphic dung beetle, *Onthophagus taurus* and *O. binodis*, the dimorphic allometry of the male horns loses the step function believed to be characteristic of reprogramming following the transformation to the appropriate log scale (Figs. 1B,D,F, and 2B,D; Tomkins et al. (2005). In contrast when similar transformations were made to forceps lengths and body size in the dimorphic earwig *Forficula auricularia*, the step function characteristic of reprogramming remains (Figs. 1A,C,E, and 2A,C; Tomkins et al. 2005).

Tomkins et al. (2005) suggested that the sigmoid shape of the dimorphic allometries in species like *O. taurus* may arise because of extremely steep allometric exponents across all but the largest males. Hence, rather than there being a reprogramming event at a particular body size the sigmoid nature of the scaling relationship of horn length on pronotum width (viewed on the linear scale, Fig. 1B) may arise because positive allometry causes an upwardly sweeping curve, followed by some limitation on horn growth, that produces the asymptote (Tomkins et al. 2005). This developmental model of polyphenic expression, which we refer to as the positive allometry hypothesis is not new, being presented by Nijhout and Wheeler (1996), but does come as a challenge to the view that the polyphenic patterns found in many dung beetles, and in *O. taurus* in particular, are examples of reprogramming (Emlen 2000; Emlen and Allen 2004; Emlen et al. 2005; Hartfelder and Emlen 2005). Instead, the suggestion is that developmental reprogramming is not required to generate substantial phenotypic plasticity and dimorphic patterns of trait expression; something that has received attention in caste polymorphism in ants (Wilson 1953; Wheeler 1991) but seems less widely recognized in species with alternative reproductive tactics. Although the phenotypes produced by the two developmental models represent convergent evolution on what researchers would perceive as a threshold trait, there are important differences that arise from the different developmental modes. For example, where a polyphenism arises through positive allometry, alternative phenotypes are not developmentally dissociated from one another, and may be unable to respond to selection independently. Having said that, the idea of a developmental dissociation between morphs in species with reprogramming has not received much empirical attention either, and it may be that even where reprogramming does occur there might remain a level of developmental congruence between morphs.

Comparing the positive allometry model with the reprogramming model introduces several semantic challenges. Both models give rise to a norm of reaction linking variation in environmental cues to variation in phenotypic expression. Furthermore, both models have the capacity to generate polyphenisms, that is the expression of alternative morphs, although the degree of discreteness of morphs is likely to depend at least in part on criteria used for separation. However, the models differ dramatically with respect to the notion of a threshold. Although the definition of a threshold, and by extension threshold evolution, is unambiguous in the context of a sigmoid allometry thought to have arisen through reprogramming, a proper counterpart does not exist for the positive allometry model. In a mathematical sense positive allometries do not have thresholds, and the evolution of thresholds can thus not be modeled using a positive allometry model in a strict sense. Instead, in the positive allometry model, a "threshold" is perceived because the appreciable growth of structures
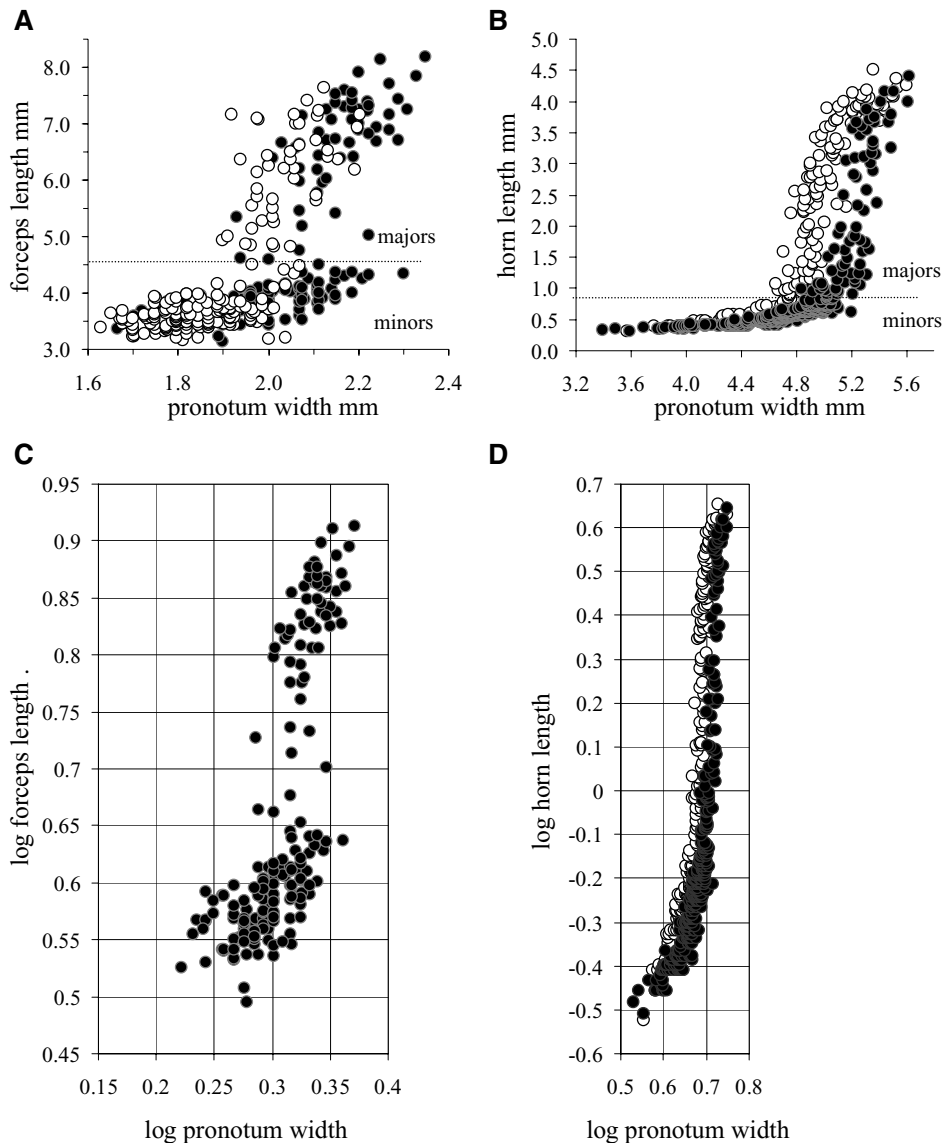
Figure 2. (A) Sigmoid relationship between forceps length and pronotum width in two field collected populations of *F. auricularia* that have diverged in threshold switchpoint; ○, Knoxes Reef, one of the Farne Islands (UK) and ●, Bass Rock in the Firth of Forth (UK). (B) Sigmoid relationship between horn length and pronotum width in two populations of *Onthophagus taurus* reared in a common garden experiment that have diverged in the position of the reaction norm; ○, North Carolina and ●, Western Australia. In (A) and (B), the horizontal line shows the divisions between the horn size classes of the males. (C) Log–log plot where $\Delta X = \Delta Y$ for the *F. auricularia* from Bass Rock only, showing how the step function characteristic of reprogramming remains following log–log transformation. (D) Log–log plot where $\Delta X = \Delta Y$ for horn length on pronotum width for the *O. taurus* data showing the continuous nature of the reaction norm.

suddenly becomes easily detectable to the observer at a particular body size, and "threshold evolution" proceeds via changes in the intercept or slope of the positive allometry. To still facilitate comparisons between classic threshold models, which have traditionally been applied to polyphenic organisms such as earwigs or horned beetles, and the positive allometry model we will use the term "threshold" when comparing both models, and the term "threshold switchpoint" for comparisons that apply solely to the reprogramming model.

Tomkins et al. (2005) hypothesized that under the positive allometry model, the evolution of dimorphic reaction norms would occur either through a shift in the intercept of trait size on body size, and/or a change in the slope itself (Figs. 1F and 3), because both parameters independently affect the body size at which exponential growth (characterized by the steep upward curve) is detectable on a linear scale (Fig. 3). The shifts in slope or intercept that cause reaction norm evolution should be detectable in minor males because the positive allometry hypothesis is based
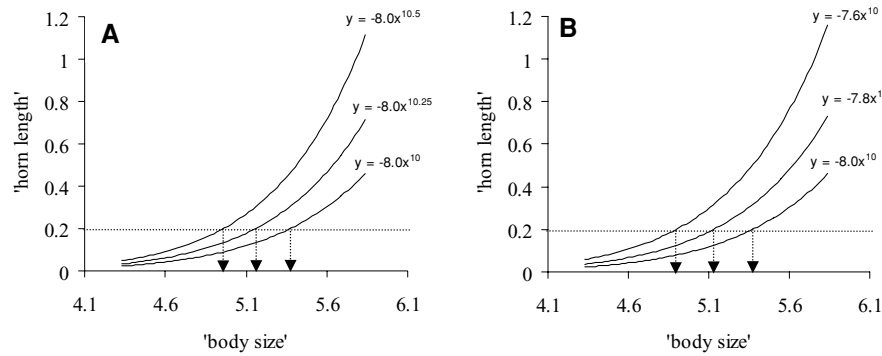
Figure 3. Hypothetical relationships between trait length and body size showing how changes either in the slope (A) or intercept (B) of a power curve have similar effects on the body size that trait expression appears to suddenly increase. Switchpoint is defined here as the body size where "horn length" is greater than 0.2, this is similar to cases in which horns were considered rudimentary below a critical value.

on the notion that the reaction norm is continuous. A change in allometric slope or intercept among minor males that correlates with variation in "threshold" is therefore predicted under the positive allometry hypothesis and detecting such patterns would reveal the mechanism of evolutionary changes in morph expression in those species in which positive allometry generates a polyphenism.

Under the reprogramming model, developmental decoupling ought to mean that selection for larger or smaller traits on one side of the developmental schism or threshold should be dissociated from what happens on the other side (West-Eberhard 1989, 2003). Hence, a priori this hypothesis predicts that evolutionary changes in threshold are more likely to be independent of any correlated changes in intercept and slope (Fig. 1E). Here we investigate patterns of allometry and threshold evolution in populations of two species believed to employ different modes of polyphenic development: the dung beetle *O. taurus* and the European earwig *F. auricularia*. We find that horn allometry in *O. taurus* has diverged in minor males in a manner that is very tightly related to the population's threshold, consistent with the mechanism of threshold evolution proposed by Tomkins et al. (2005). Contrary to the reprogramming model proposed for *F. auricularia*, however, we find that forceps allometry in *F. auricularia* also correlates with variation in threshold switchpoint. Although the pattern in *F. auricularia* is weaker and different from *O. taurus*, it nevertheless raises intriguing questions about polyphenic diversity and threshold evolution.

Materials and Methods
FIELD COLLECTED AND MUSEUM SAMPLES
Switchpoint variation in the forceps length threshold of *F. auricularia* has been described before (Tomkins 1999; Tomkins and Brown 2004). This variation occurs across island populations in the Firth of Forth in Scotland and the North Sea off the coast of Northumberland in England. Samples from each population were taken by trapping or collecting by hand between August and

September of 2001 and 2002 (Tomkins and Brown 2004). Further populations, hand collected from Rathlin Island in Northern Ireland (August 2003), from Krakow, Poland (September 2001) from Pinnaroo, South Australia (November 2002), and from Pemberton Western Australia (June 2008) were also included. These were measured and analyzed using the same methodology used for the other samples.

Threshold variation across a number of populations of *O. taurus* have been described previously in the literature (Moczek and Nijhout 2003, 2004). These represent samples from museums that were field collected specimens from the species native range in Europe and North Africa, France, Spain, Morocco, Tunisia, Turkey, Iran, and Greece (Moczek and Nijhout 2003) and from four populations in the Eastern US (Knoxville, Durham, Statesville, and Sylvania), Western Australia (Narrikup, Harvey, Pinjarra, and Witchcliffe), and Eastern Australia (Canberra, Cootamundra, Tallengatta, and Tumut) where they are introduced (Moczek and Nijhout 2004).

COMMON GARDEN EXPERIMENT
Western Australian (WA) and North Carolinean (NC) *O. taurus* were collected in the field and maintained in the laboratory as described in Moczek et al. (2002) and Moczek and Nijhout (2003). Approximately 2000 beetles collected from one pasture near Busselton (WA) and 1500 beetles collected from pastures in Durham County (NC) were used to found two laboratory colonies used for common garden rearing. Both laboratory colonies were kept in the same insectary at Duke University at 26°C and 60% relative humidity under a 16:8 light:dark cycle. Beetles were bred as described in Moczek et al. (2002). To minimize inbreeding, individual adult beetles were allowed to reproduce only once and were then removed from the colony. Different generations were kept in separate containers. Over 1000 individuals were reared each generation for each population. Both laboratory colonies received the same treatment and breeding setup. Here we examine individuals

from the second filial generation that resulted from this breeding design.

MORPHOLOGICAL MEASUREMENTS

Pronotum width and left and right forceps length were measured for male *F. auricularia.* Measurements were made under a dissecting microscope using an eye-piece graticule, or using Scion image, image-analysis software. Forceps length measurements were the straight-line distances from the base of the forceps to the tip, pronotum was measured at its widest point.

All individuals of *O. taurus* were measured using a standard two-dimensional image analysis setup, at the Duke University Morphometrics Laboratory (for details see Moczek and Emlen 1999). We used the width of the pronotum as an estimate for body size and measured horn length by following the outermost edge of the horn from the upper edge of the eye cavity to the tip of the horn (as described in Moczek 2006b). This way of measuring horn length is one of several employed by different laboratories, and Tomkins et al. (2006) and Moczek (2006b) showed recently that the use of different land marks and measurement techniques can affect the slope in allometric analyses. The method employed here generally underestimates allometries of small males compared to other measures (Tomkins et al. 2006). This underestimation leads to an increased divergence between the slope of horn length in minor males and the slope of horn length in preasymptotic major males. Even so, because we use identical methods to compare different populations, and the beetles were measured before we were aware of between-laboratory differences in measurement techniques—or indeed this hypothesis, we consider the analysis presented here as an objective contrast of allometric differences between these populations. A comparison between the slopes derived by using different landmarks and the pitfalls therein, are discussed fully in Tomkins et al. (2006) and Moczek (2006b).

COMPARING ALLOMETRIES AND INTERCEPTS

Testing the hypothesis that population variation in morph expression has arisen through changes in the allometry of horn length, we concentrated only on minor male earwigs and beetles. These males are the significant part of the population for the discrimination between the competing hypotheses, because the positive allometry hypothesis predicts a strong correlation between the allometric slope and intercept of horn or forceps length and population variation in morph expression, whereas the developmental reprogramming model does not (Fig. 1). We used the model of Kotiaho and Tomkins (2001) and R code for the model supplied by Ken Wilson and Rob Knell to assign males to either the major or minor morph in *F. auricularia* (Tomkins and Brown 2004), the average across the populations was for minor males to have forceps less than 4.51 mm (e.g., Fig. 1A). Switchpoints in *F. auricularia* were estimated by using Eberhard and Gutiérrez's (1991) model

and the same SPlus Code. The discrimination of morphs in *O. taurus* is more difficult because of the continuous nature of the reaction norm but to avoid including any males showing the beginnings of major type of horns we used only males with a horn length of less than 0.75 mm, (Fig. 2B). This provides a reliable estimate of the allometry in minor males; the correlation between body size and horn size is strong (mean and median $r^2 = 0.5$) and only one regression (Statesville) was nonsignificant due to low sample size ($N = 7$). The removal of this population increases the significance of the relationships reported below, but we have left it in the analysis because, significant or not, it estimates the allometric parameters in that population. The horn length switchpoints for *O. taurus* were taken from Moczek (2003) and Moczek and Nijhout (2003), which were calculated using the modified formula of a sigmoid curve, the details of which can be sourced from the original papers.

We used the Smatr program (Falster et al. 2006) to estimate the reduced major axis (RMA) slopes and intercepts of log forceps or horn length on log pronotum width for the males. Smatr also estimates 95% confidence intervals (CI) around the slope and can be used to test both slopes and intercepts against each other in a manner equivalent to ANCOVA.

Results
FIELD COLLECTED AND MUSEUM SAMPLES

To determine whether the allometric parameters of minor male *F. auricularia* influenced the position of the switchpoint across populations, population switchpoint was analyzed in a multiple regression with the RMA slope and intercept of the minor male allometries and mean pronotum width. There was a significant effect of pronotum width (whole model, $F_{3,22} = 11.42$, $P = 0.001$; pronotum width, $t = 5.36$, $P < 0.001$) indicating that populations that were larger on the whole had switchpoints at larger body sizes, however there was no effect of switchpoint on either the slope ($t = 1.109$, $P = 0.28$) or the intercept ($t = 0.456$, $P = 0.65$) of the minor male allometry across the populations. Slope and intercept are highly significantly negatively correlated however ($r_{23} = -0.96$, $P < 0.001$), and cause a colinearity problem in the model (variance inflation factor VIF >15). To overcome this we separated the effects of slope and intercept by taking the first and second eigenvectors of the relationship between the two. The first eigenvector summarizes the covariation between slope and intercept. The second (PC2) encapsulates variation in slope controlling for intercept, and intercept controlling for slope because it takes the variation perpendicular to the major axis of the scatter of points. The first principle component (PC1) accounted for 98.34% of the original variation, whereas PC2 accounted for the remaining 1.66% of residual variation. The multiple regression, with switchpoint as a dependent variable and pronotum width

and the covariance in slope and intercept (PC1) and the residual variation between slope and intercept (PC2) as dependents, was significant overall (Whole model, $F_{3,22} = 11.42$, $P = 0.001$); pronotum width was significant as before ($t_{1,22} = 6.32$ $P < 0.001$), there was a significant effect on the switchpoint of the covariation in slope and intercept summarized as PC1 ($t_{1,22} = 2.517$ $P = 0.021$), but there was no effect of the residual variation in slope and intercept ($t_{1,20} = 0.7$ $P = 0.5$). The effect of PC1 is such that where the population's threshold switchpoint is far to the right of the body size distribution (i.e., where major males are rare), there is a shallower slope of forceps length on pronotum width and a greater intercept.

When variation in population switchpoint in *O. taurus* was analyzed in a multiple regression with both pronotum width and RMA slope and intercept of the minor males' horn allometry, both allometric terms were highly significantly negatively related to the "switchpoints" reported by Moczek (2003) and Moczek and Nijhout (2003), but pronotum width was not (whole model, $F_{3,15} = 6.77$, $P = 0.004$, intercept $t = -4.5$, $P = 0.001$, slope $t = 4.49$ $P = 0.001$, pronotum width $t = 1.37$, $P = 0.19$). Similar to *F. auricularia*, however, there is a very strong negative correlation between the slope and the intercept ($r_{19} = -1$, $P = 0.000$); causing a colinearity problem in the regression (VIF < 1150). Despite the strength of the correlation between slope and intercept there was a small amount of residual variance and again we extracted this as the first and second eigenvectors from a principal component analysis. The first principle component summarized the covariation in slope and intercept and accounted for 99.67% of the original variation, whereas the second principle component only summarized 0.024% of the original variation. When the covariance in slope and elevation (PC1) and the residual variance (PC2) and pronotum width were entered in the multiple regression, there was a strong and significant effect of residual variance in slope and intercept (PC2) on population switchpoint (Whole model, $F_{3,15} = 6.773$, $P = 0.004$, residual variation in slope and intercept, $t = 4.65$, $P = 0.001$, Fig. 4A), but no effect of the covariance in slope and intercept ($t = 1.42$, $P = 0.18$) or pronotum width ($t = 1.12$, $P = 0.28$).

O. TAURUS MINOR MALE HORN ALLOMETRY IN THE COMMON GARDEN

For minor male *O. taurus* reared under common garden conditions the slope for the NC population (RMA = 3.00, CI = 2.71–3.31, $N = 108$) was not significantly different from the WA population (RMA = 2.96, CI = 2.79–3.13, $N = 279$; $F_{1,385} = 0.06$, $P = 0.819$; Fig. 4). Analysis of covariance, based on RMA regression, was used to compare the intercept of each slope (based on the common RMA slope of 2.97). The intercept (-2.19) of the NC population was significantly greater (less negative) than the intercept (-2.25; $F_{1,385} = 140$; $P < 0.001$; Fig. 2D) of the WA population.
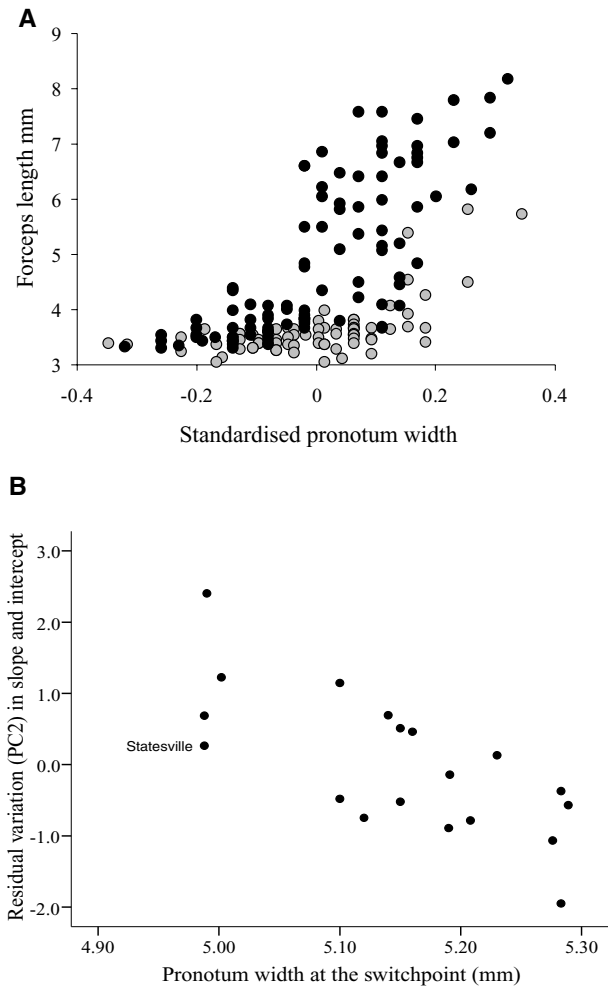
Figure 4. (A) Scatterplot of forceps length on male pronotum width (standardized to have the same mean) for geographically close (<2 km) populations from the islands of Lamb (○) and Craigleith (●) in the Firth of Forth, Scotland. Unlike our hypothesized pattern (Fig. 1E) the slope and intercept of the minor male allometry do change in concert with changes in the position of the threshold. The Craigleith population has a significantly positive allometry in minor male forceps length (RMA = 1.57, 95% CI = 1.28–1.94), whereas in the Lamb population minor male forceps are isometric with pronotum width (RMA = 1.03, 95% CI = 0.82–1.28). (B) Correlation between residual variation (summarized from the relationship between slope and intercept of log horn on log pronotum width) in horn length, and threshold switchpoint variation across populations of *O. taurus*. In *O. taurus* body size at the "switchpoint" was smaller when intercepts were high given the slope, and where slopes were high given the intercept. Lower intercepts for a given slope or low slopes for a given intercept both yield larger values of body size at the switchpoint. In the population from Statesville the correlation between log horn and log pronotum was nonsignificant.

Discussion

There is much interest in the evolution and development of scaling relationships in insects, and much of this work concentrates on

static "environmental" allometries, where genotypically diverse individuals develop along an environmental gradient (Nijhout and Wheeler 1996; Nijhout 1999, 2003; Shingleton et al. 2007). Two hypotheses have been proposed to explain evolutionary changes in threshold expression in species like dung beetles and earwigs. On one hand, the developmental reprogramming model predicts that the critical size at which reprogramming occurs can change between populations, thereby altering the position of the threshold (Fig. 1E). On the other, the positive allometry model predicts that what appear to be thresholds are in fact continuous reaction norms that change in position through changes in the slope, and/or the intercept of the allometry (Figs. 1F and 3).

Contrary to our prediction that threshold position would be independent of switchpoint in *F. auricularia*, we found a significant association between the covariance in slope and intercept of the static allometry of forceps length (PC1) in minor males and variation in switchpoints across the populations. We did not however find any association between the residual variation in slope and elevation and threshold position as we did in *O. taurus*. This is despite greater power to do so, both in terms of the number of populations and in terms of the variance available as PC2. The pattern in *F. auricularia* therefore raises the question as to how independent morphs are, even where reprogramming occurs. To our knowledge there are no similar tests of this hypothesis that might elucidate this question.

For *O. taurus*, our data follow the prediction that there is a tight relationship between the allometric parameters and the body size at which horns suddenly become apparent in these animals. Unlike *F. auricularia* the negative covariation in slope and intercept (PC1) in *O. taurus* showed no significant effect on threshold variation, this variation in *O taurus* effectively holds the thresholds within a narrow range of values; for example, the steepest horn allometry (Cootamundra) and the shallowest (Turkey) have the same threshold values because their intercepts are at the opposite extremes of the range. However it is variation around the relationship between slope and intercept (PC2), which results in threshold variation in *O. taurus*. Figure 3A shows how increasing the slope of horn length on pronotum width (with constant intercept) will reduce the body size at which horns become obvious and conversely increasing the intercept, but leaving the slope constant will have the same effect (Fig. 3B). In the common garden experiment we found an example of the latter case: the divergence in the "threshold" between the two populations (Moczek et al. 2002) could be accounted for by changes in the intercept of the horn allometry such that the population with the threshold at a smaller body size (NC) had the greater intercept. Our interpretation of the changed intercept but comparable slope is therefore that the divergence between these populations arose through selection acting on the intercept of the horn length reaction norm. Conversely, field populations of Tallengatta (Aus) and Sylvania (USA) differ

in threshold by 55% (of the range of switchpoints), however the intercept of the minor male allometry is the same in both populations, so that the divergence in these populations is therefore due only to a difference in the slope of the reaction norm.

Divergence in the slope and intercept of static allometries is likely achieved through different physiological processes (Shingleton et al. 2007). Changes in intercept can be brought about by modification of the terminal growth period of appendage primordia (Edgar 2006; Mirth and Riddiford 2007; Shingleton et al. 2007). Changes in the timing of sensitive periods during which target tissues are able to respond to circulating hormone titres, such as juvenile hormone, provide a possible mechanism by which the terminal growth period of appendage primordia can contract or expand, thereby altering the relative sizes of adult structures via changes in the intercept (Shingleton et al. 2007). In contrast, changes in slope may be brought about by evolutionary changes in sensitivity to hormonal signals or growth factors per se. For example, in *Drosophila* organs that are less sensitive to insulin signaling have lower slopes than those that are more sensitive (Shingleton et al. 2005). Despite this separation of physiological process, it is hard to envisage how large changes in slope can occur without concomitant changes in intercept, when traits co-vary with body size. Indeed, across field-collected populations of both *O. taurus* and *F. auricularia* there was a tight negative correlation between the allometric slope and the intercept.

The correlation between slope and intercept in both species shows how these parameters and their underlying physiology evolve together. In *F. auricularia* it is easy to envisage how lower slopes and greater intercepts might yield threshold values further to the right of the distribution (Fig. 4A). In *O. taurus* it is very subtle variation around this tight correlation that reveals how different populations have "solved the problem" of reaction norm evolution. Some recent studies artificially selecting for relative wing size in butterflies have achieved the allometric divergence similar to that seen in the comparison of the WA and NC populations in which the intercept of the allometry changed through selection (Frankino et al. 2005). Emlen's (1996) artificial selection of *O. acuminatus* probably yielded the same pattern since relative horn length diverged, although it was not analyzed as change in intercept. Artificially selecting on relative eye-span in stalk-eyed flies (Wilkinson 1993) affected both the slope and the intercept of the static allometry showing how simply selecting on the ratio of one trait to another can lead to different allometric responses. One explanation for the change in slope in the stalk-eyed fly study was that it arose as a consequence of a constraint, possibly the genetic correlation with female eye-stalk length (Wilkinson 1993). Why some reaction norms in *O. taurus* evolve through intercept more than slope and some vice versa is intriguing.

Our data suggest that some of the hypotheses regarding the developmental biology underlying the male polyphenism in

O. taurus require re-evaluation. Under the reprogramming model, circulating levels of endocrine factors such as JH and possibly ecdysteroids are thought to program the degree to which epithelial tissues would undergo rapid growth during the prepupal stage (Emlen and Nijhout 1999, 2001). Incipient horn tissue of males fated to develop into the minor morph were thought to be programmed to undergo little if any growth, whereas that of males above a critical size threshold would be programmed to undergo disproportionately rapid growth. Differential growth is thought to be achieved via programmed differences in the rate of cell division, the duration of the active growth period, and ultimately reprogramming of genes responsible for coordinating timing and location of outgrowth formation (Emlen and Nijhout 2001; Emlen and Allen 2004). The results of recent studies on gene expression in the horns of developing beetles challenge this perspective by showing that major and minor males share the expression of the same patterning genes thought to be crucial for horn formation, such as the transcription factors *Distal-less, aristaless* (Moczek and Nagy 2005), *dachshund, extradenticle, homothorax* (Moczek 2006a) and the morphogens *wingless* and *decapentaplegic* (B. R. Kesselring and A. P. Moczek, unpubl. ms.). Differences in horn expression can therefore not be explained by the presence or absence of the protein products of these genes in the corresponding tissue regions. Instead, these results suggest that the same genes are activated in similar tissue regions regardless of whether males develop into major or minor morphs, a finding consistent with the positive allometry hypothesis.

Under the positive allometry hypothesis, all males are expected to undergo a similar basic developmental process and it appears that differences between male morphs can arise from a developmental sequence common to all males but terminated by some mechanism able to tightly link magnitude of horn growth to final adult body size across the entire range of male body sizes. A mechanism that is likely to achieve such a size-dependent function is the length of the terminal growth phase of the horn primordia (Tomkins et al. 2005; Shingleton et al. 2007). For example, in *O. taurus* head horn primordia of majors grow over a maximum period of approximately 48 h, whereas the growth of small horns in minor males is completed within the first 24 h of the prepupal stage whereas other appendage primordia (e.g., legs or wings) continue to grow (Moczek and Nagy 2005; Moczek 2006a,c). These observations suggest that simple differences in the duration over which rapid cell proliferation occurs may be sufficient to generate slope variation in *O. taurus* (Nijhout and Wheeler 1996; Moczek and Nagy 2005; Tomkins et al. 2005). If correct, this hypothesis raises the question of how endocrine factors regulate horn development and mediate reaction norm evolution in *Onthophagus* beetles. For example, rather than reprogramming the rate of growth that occurs in a given horn primordium, depending on whether males fall below or above a critical size threshold,

JH may simply regulate the duration over which growth occurs across the entire range of male body sizes (Tomkins et al. 2005). Given the dynamic nature of prepupal horn growth even subtle changes in growth duration are likely sufficient to generate very different ontogenetic outcomes.

Under this developmental model, the evolutionary changes seen in WA and NC populations of *O. taurus*, require evolutionary changes in how JH alters the duration of the prepupal horn growth period for a given adult body size. For example, evolved reductions in JH titres, or a delay in tissue sensitivity to JH would cause males to exhibit a reduced prepupal horn growth period, result in the growth of relatively reduced horn lengths, and cause a shift in the reaction norm of horn expression to larger body sizes. These predictions are consistent with results of endocrine comparisons between WA and NC populations; WA populations require larger body sizes to exhibit the same absolute horn growth, are less sensitive to ectopic JH applications, and exhibit a delayed sensitive period for JH during late larval development compared to their North Carolinian counterparts (Moczek and Nijhout 2002). Furthermore, these differences are observed in minors, as predicted by the positive allometry hypothesis, rather than solely among majors as expected under the reprogramming hypothesis (Emlen 2000; Emlen and Nijhout 2000).

Evolutionary changes in the slopes of *O. taurus* allometries are less readily explicable with existing developmental data, however, promising preliminary findings exist. For example, in *Drosophila* organs that are less sensitive to insulin signaling have lower slopes than those that are more sensitive (Shingleton et al. 2005). Recent microarray experiments in several *Onthophagus* species identified differential expression of insulin pathway genes (e.g., foxo) as well as members of other known growth regulators (T. Kijimoto, D. J. Rose, J. Costello, Z. Tang, A. P. Moczek, and J. Andrews, unpubl. ms.). Whether this differential expression is indicative of altered sensitivities and growth responses is currently being investigated.

Very large changes in trait size can be expected where horn length reflects an underlying positively allometric reaction norm, but changes in shape require the action of sculpting and remodeling genes. Hence despite the explanatory power of the positive allometry hypothesis, size-specific developmental changes of one kind or another are still likely to occur in onthophagine beetles. For example, apart from the seven appendage-patterning genes mentioned above, many other transcripts (T. Kijimoto, D. J. Rose, J. Costello, Z. Tang, A. P. Moczek, and J. Andrews, unpubl. ms.) and proteins (Y. Yerushalmi and A. P. Moczek, unpubl. ms.) are unique to the horns of large-horned males, although their functional significance remains to be examined. Pupal remodeling, too, appears to occur in a discrete, morph-specific manner in at least some species, causing some males to reabsorb or reshape horns far more dramatically than others (Moczek 2007).

Such morph-specific remodeling may also explain a peculiar situation in *O. sloani*, which exhibits a threshold size above which males are hornless and below which males have horns (Emlen et al. 2005), a pattern not otherwise explicable under a simple positive allometry model. The positive allometry model therefore complements, rather than completely replaces, the developmental reprogramming hypothesis. Furthermore, it underscores how apparently similar patterns of morphological variation may be generated by very different developmental mechanisms, with different implications for the evolution of exaggerated structures and alternative phenotypes.

The data presented here for both dung beetles and earwigs clearly show that deviations in threshold across populations correlate with allometric variation in minor males, suggesting that for both species and both developmental models, threshold evolution might be achieved by modifying these parameters. The scope for positive allometry to play a role in the extraordinary variation seen in phenotypically plastic traits has been well illustrated in the ants (Wilson 1953; Wheeler 1991). Even so, the notion that dramatic variance in morphology can be explained even where there is developmental congruence between morphs has received less attention than the case of reprogramming where developmental independence has been assumed. Dung beetles in particular have been widely used as an example of the variation that can be generated through reprogramming. In contrast our data show that the type of extraordinary variation in phenotype seen in dung beetles that has traditionally been considered to be the product of the phenotypic expression of divergent and independent selective optima (West-Eberhard 1986, 1989, 2003) appears instead to be constrained to respond to selection elsewhere on the reaction norm. We expected this pattern to occur where there is a continuous positively allometric reaction norm; that we also found it where there is a clear developmental threshold was a surprise. Evidently even reprogrammed morphs are not as independent as was supposed. Whether this is a peculiarity of hemimetabolous development remains to be determined.

ACKNOWLEDGMENTS

LITERATURE CITED

Eberhard, W. G., and E. E. Gutiérrez. 1991. Male dimorphisms in beetles and earwigs and the question of developmental constraints. Evolution 45:18–28.

Edgar, B. A. 2006. How flies get their size: genetics meets physiology. Nat. Rev. Genet. 7:907–916.

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

———. 2000. Integrating development with evolution: a case study with beetle horns. Bioscience 50:403–418.

Emlen, D. J., and C. E. Allen. 2004. Genotype to phenotype: physiological control of trait size and scaling in insects. Integr. Comp. Biol. 43:617–634.

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

———. 2000. The development and evolution of exaggerated morphologies in insects. Annu. Rev. Entomol. 45:611–708.

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

Emlen, D. J., J. Hunt, and L. W. Simmons. 2005. Evolution of male- and sexual-dimorphism in the expression of beetle horns: modularity, lability and developmental constraint. Am. Nat. 166:S42–S68.

Falster, D. S., D. I. Warton, and I. J. Wright. 2006. (S)MATR: Standardised major axis tests and routines. Version 2.0:http:www.bio.mq.edu.au/ecology/SMATR.

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.

Hartfelder, K., and D. J. Emlen. 2005. Endocrine control of insect polyphenism. Pp. 651–703 *in* L. I. Goilbert, K. Iatrou and S. S. Gill, eds. Comprehensive molecular insect Science. Elsevier Ltd, Oxford.

Kotiaho, J. S., and J. L. Tomkins. 2001. The discrimination of alternative male morphologies. Behav. Ecol. 12:553–557.

Mirth, C. K., and L. M. Riddiford. 2007. Size assessment and growth control: how adult size is determined in insects. Bioessays 29:344–355.

———. 2003. The behavioral ecology of threshold evolution in a polyphenic beetle. Behav. Ecol. 14:841–854.

———. 2006a. Integrating micro and macroevolution of development through the study of horned beetles. Heredity 97:168–178.

———. 2006b. A matter of measurements: challenges and approaches in the comparative analysis of static allometries. Am. Nat. 167:606–611.

———. 2006c. Pupal remodeling and the development and evolution of sexual dimorphism in horned beetles. Am. Nat. 168:711–729.

———. 2007. Pupal remodelling and the evolution and development of alternative male morphologies in horned beetles. BMC Evol. Biol. 7:151.

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

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

———. 2003. Rapid evolution of a polyphenic threshold. Evol. Dev. 5:259–268.

———. 2004. Trade-offs during the development of primary and secondary sexual traits in a horned beetle. Am. Nat. 163:184–191.

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

Moczek, A. P., J. Hunt, D. J. Emlen, and L. W. Simmons. 2002. Threshold evolution in exotic populations of a polyphenic beetle. Evol. Ecol. Res. 4:587–601.

Nijhout, H. F. 1999. Control mechanisms of polyphenic development in insects—in polyphenic development, environmental factors alter same aspects of development in an orderly and predictable way. Bioscience 49:181–192.

———. 2003. Development and evolution of adaptive polyphenisms. Evol. Dev. 5:9–18.

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

Shingleton, A. W., J. Das, L. Vinicius, and D. L. Stern. 2005. The temporal requirements for insulin signaling during development in Drosophila. Plos Biol. 3:1607–1617.

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.

Smith-Gill, S. J. 1983. Developmental plasticity—developmental conversion versus phenotypic modulation. Am. Zool. 23:47–55.

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

Tomkins, J. L., and G. S. Brown. 2004. Population density drives the local evolution of a threshold dimorphism. Nature 431:1099–1103.

Tomkins, J. L., and W. Hazel. 2007. The status of the conditional evolutionarily stable strategy. Trends Ecol. Evol. 22:522–528.

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

———. 2006. Major differences in minor allometries: a reply to Moczek. Am. Nat. 167:612–618.

West-Eberhard, M. J. 1986. Alternative adaptations, speciation and phylogeny (A Review). Proc. Natl. Acad. Sci. USA 83:1388–1392.

———. 1989. Phenotypic plasticity and the origins of diversity. Annu. Rev. Ecol. Syst. 20:249–278.

———. 2003. Developmental plasticity and evolution. Oxford Univ. Press, New York.

Wheeler, D. E. 1991. The developmental basis of worker caste polymorphisms in ants. Am. Nat. 138:1218–1238.

Wilkinson, G. S. 1993. Artificial selection alters allometry in the stalk eyed fly *Crytodiopsis dalmanni* (Diptera: Diopsidae). Genet. Res. Cambridge 62:213–222.

Wilson, E. O. 1953. The origin and evolution of polymorphisms in ants. Q. Rev. Biol. 28:136–156.

Associate Editor: D. Pfennig